MARIE BOUILLÉ

ÉTUDE DE L'ÉVOLUTION MOLÉCULAIRE DU GENRE *PICEA* A. DIETRICH (PINACEAE): PHYLOGÉNIES DE GÈNES ET PHYLOGÉNIES D'ESPÈCES

Thèse présentée
à la Faculté des études supérieures de l'Université Laval
dans le cadre du programme de doctorat en Sciences forestières
pour l'obtention du grade de Philosophiae Doctor (Ph. D.)

DEPARTEMENT DES SCIENCES DU BOIS ET DE LA FORET
FACULTE DE FORESTERIE ET GEOMATIQUE
UNIVERSITE LAVAL
QUEBEC

2007

© Marie Bouillé, 2007

Résumé

La complexité des classifications morphologiques du genre *Picea* et leur manque de cohérence ont justifié le besoin de l'estimation d'une phylogénie moléculaire représentative de l'histoire biogéographique de ce genre conifèrien circumpolaire.

Les essais d'estimation phylogénétique pour trois loci du génome nucléaire ont révélé la présence de polymorphismes intraspécifiques et trans-spécifiques partagés entre trois espèces distinctes du genre *Picea*. La période de divergence entre les trois espèces a été estimée à 13 à 20 millions d'années (mA) et les temps de coalescence des allèles étaient de l'ordre de 10 à 18 mA, ce qui implique que les polymorphismes puissent être partagés depuis l'ancêtre commun aux trois espèces. La fixation des polymorphismes aurait été retardée par de grandes tailles historiques de population, estimées à plus de 100 000 individus. Ces résultats suggèrent que les polymorphismes trans-spécifiques partagés peuvent être fréquents pour les loci nucléaires de plantes et qu'ils peuvent contribuer significativement à la diversité allélique, surtout chez les plantes anémophiles et allogames possédant de grands effectifs de population. Ces polymorphismes appellent à la prudence quant à l'utilisation de gènes nucléaires pour estimer les phylogénies de tels groupes de plantes.

Les efforts d'estimation phylogénétique subséquents se sont donc portés sur les régions des génomes cytoplasmiques haploïdes, dont les séquences sont plus conservées que celles du génome nucléaire. La phylogénie des *Picea* a été estimée suite au séquençage de trois régions du génome chloroplastique, transmis paternellement, ainsi qu'une région du génome mitochondrial, transmis maternellement. Des différences significatives ont été observées entre les phylogénies chloroplastiques, et entre les phylogénies chloroplastiques et la phylogénie mitochondriale. Aucune des phylogénies ne correspondait aux classifications actuelles. La phylogénie mitochondriale était géographiquement plus structurée que les phylogénies chloroplastiques. Sur l'arbre mitochondrial, la plupart des taxons nord-américains formaient un groupe monophylétique. La cohérence géographique

des divers regroupements nord-américains et asiatiques sur cet arbre indique que des phénomènes de spéciation de nature géographique se soient produits. Les incompatibilités entre les phylogénies d'organelles suggéraient l'occurrence de transferts latéraux anciens du génome chloroplastique entre les lignées mitochondriales. Les incompatibilités entre les partitions chloroplastiques suggéraient en outre la présence de recombinaison présumément liée à l'évolution réticulée ancienne. Les marqueurs chloroplastiques seraient donc de moindre valeur phylogénétique et phylogéographique que les marqueurs mitochondriaux pour les genres ayant une transmission paternelle du génome chloroplastique, tel que chez les Pinaceae.

Abstract

Complex morphological classifications for the genus *Picea*, a circumpolar conifer genus, and the lack of coherence among them indicated the need to estimate a molecular phylogeny representative of the biogeographical history of the whole genus.

Preliminary tests for three nuclear loci revealed the presence of intraspecific as well as trans-specific shared polymorphisms among three distantly related species in the genus. The divergence time between species was estimated at 13 to 20 million years (my), and allele coalescence times in the order of 10 to 18 my, which implies the sharing of polymorphisms since common ancestry. Large historical population sizes in excess of 100 000 would have delayed the fixation of polymorphisms. These results suggest that trans-species shared polymorphisms might be frequent at plant nuclear gene loci, contributing significantly to allelic diversity. This trend is more likely in plants characterized by ecological and life-history determinants favoring large population sizes, such as an outcrossing mating system, wind pollination, and a dominant position in ecosystems. These polymorphisms also call for caution in estimating congeneric species phylogenies from nuclear gene sequences in such plant groups.

Subsequent efforts in estimating the biogeographical history of the genus *Picea* were directed at organelle DNA regions, which are more conserved than nuclear DNA at the sequence level. Distinct phylogenies of *Picea* were obtained by sequencing three regions from the paternally inherited chloroplast genome (*trn*K, *rbc*L, and *trn*TLF), and the intron 2 of the maternally transmitted mitochondrial gene *nad1*. Significant differences were found between chloroplastic partitions, and between the cpDNA and mtDNA phylogenies. None of the phylogenies matched current classifications. The mtDNA phylogeny was geographically more structured than the cpDNA phylogenies. Most North American taxa formed a monophyletic group on the mtDNA tree with coherent geographical clustering, indicating geographic speciation by range fragmentation or by dispersal and isolation. Similar patterns, also found among Asian taxa, likely represent a major trend in the

Pinaceae. Incongruences between organelle phylogenies also indicated ancient lateral transfers of the chloroplast genome between mtDNA lineages. Incongruences between cpDNA partitions further suggested heterologous recombination presumably linked to ancient reticulate evolution. These results indicate the reduced value of cpDNA as a phylogenetic and phylogeographical marker in genera with paternal transmission of the chloroplast genome such as in the Pinaceae.

Avant-propos

Les quatre parties distinctes de cet ouvrage se proposent comme suit. Tout d'abord, l'introduction générale présente plusieurs informations qui sont requises à la bonne compréhension du sujet. Puis, les deux chapitres suivants rapportent respectivement les résultats des deux grandes parties distinctes de la recherche. Ces chapitres prennent la forme d'articles scientifiques transcrits tels qu'ils ont été présentés ou publiés dans des revues scientifiques internationales, et pour lesquels l'auteur de cette thèse agit comme auteur principal. Par conséquent, ils sont présentés ici en langue anglaise, dans le format spécifique à chaque revue. Chacun est toutefois précédé d'un résumé court en français. Enfin, la conclusion constitue le quatrième et dernier chapitre, rassembleur des acquis scientifiques de cette recherche, et mentionne certaines perspectives quant au sujet traité.

La réalisation des travaux de cette thèse a été rendue possible grâce à la participation de plusieurs personnes. Je tiens ici à les remercier. De toute première importance, je remercie très sincèrement mon directeur, le Dr Jean Bousquet. Je me dois de souligner son professionnalisme et sa rigueur scientifique exemplaires, appliqués en tous instants. Il s'agit de qualités que je considère indispensables à la recherche scientifique que l'on veut excellente, de qualité, surtout dans un monde parfois trop rapide dans lequel on entend malheureusement souvent des « ça-va-faire-pareil ». D'où l'idée de dépassement, que le Dr Bousquet a su motiver en moi à chaque question, à chaque précision, à chaque discussion et à chaque correction. Je le remercie de sa confiance envers moi.

Je remercie aussi ceux qui ont accepté de faire partie de mon comité d'évaluation: les Drs Louis Bernier, Selvadurai Dayanandan et Maurice Lalonde.

Je remercie également les divers intervenants au laboratoire pour l'aide apportée: Alice Roy, François Larochelle, Nicolas Tremblay, Manuel Lamothe et Sauphie Senneville. J'ai d'ailleurs spécialement apprécié l'aide et l'esprit ingénieux de Monsieur Larochelle qui oeuvre maintenant au Centre de bioinformatique et de biologie computationnelle de l'Université Laval. Monsieur Larochelle est toujours allé de l'avant pour rendre service. Je dois souligner son empressement à régler les problèmes au niveau informatique, logistique ou au laboratoire. Je remercie aussi les divers collaborateurs mentionnés à la section « Remerciements » (*Acknowledgments*) des deux

articles ci-inclus (chapitres 2 et 3). Parmi ceux-ci, je remercie Jérôme Laroche pour certains calculs. Également, merci à Alexis Crête-Lafrenière et à Nicolas Feau qui ont apporté leurs conseils pour certaines questions de statistiques. Merci aussi à Stéphane Larose pour ses conseils appropriés et ses interventions au niveau informatique au Centre de bioinformatique et de biologie computationnelle. Très certainement, pour d'autres personnes qui ont aussi contribué à ces travaux, l'absence de mention ici est bien involontaire. Je les remercie pour toute assistance.

Au niveau administratif et fonctionnel, je remercie chaleureusement Madame Ginette Martel. Madame Martel a toujours réussi à ensoleiller les moments orageux ou simplement ordinaires par de petites conversations intéressantes, à exécuter son travail sans faille et même au-delà des espérances en trouvant des solutions pratiques aux problèmes administratifs, grâce à sa rapidité d'exécution et à sa vivacité d'esprit. Une dame de grande classe que j'apprécie beaucoup.

Je veux souligner la présence de gens agréables que j'ai eu le bonheur de fréquenter à travers divers événements, et avec lesquels j'ai échangé quelques « moments scientifiques » dans le même laboratoire: Juan Pablo, Patrick, Marie-Claude et Betty. Merci. La vie est tellement fascinante lorsque l'on côtoie des personnalités si différentes, toutes belles en soi, qui nous apportent chacune un peu plus de compréhension de la vie avec leurs forces et leurs fragilités propres. Quand l'ambiance est à la cohésion, à l'harmonie, au partage et à l'attitude positive, nous y avons notre force.

J'offre des mercis particuliers à Soeur Estelle Lacoursière, de l'Université du Québec à Trois-Rivières, et à feu Camille Gervais, de l'Herbier Louis-Marie de l'Université Laval, qui m'ont d'abord aiguillée vers le Dr Bousquet. Je leur suis reconnaissante de m'avoir dirigée vers un projet qui a ravivé ma flamme scientifique et qui l'a maintenue jusqu'à ce jour.

Enfin, je remercie le Conseil de la recherche en sciences naturelles et en génie du Canada pour ses bourses et le financement des travaux de la thèse.

Table des matières

Résumé i
Abstract iii
Avant-propos v
Table des matières vii
Liste des tableaux x
Liste des figures xi
Liste des annexes xi

Chapitre 1 **- Introduction générale** 1

1.1- Le genre *Picea*: généralités 3
1.2 - Problématique 3
1.2.1 - Morphologie des *Picea* 3
1.2.2 - Historique de la classification des *Picea* 5
1.2.3 - L'hybridation chez les *Picea* 9
1.2.4 - Histoire biogéographique des *Picea* 12
1.2.4.1 - Les fossiles de *Picea* 12
1.2.4.2 - Hypothèses biogéographiques 15
1.2.4.2.1 - Origine du genre *Picea* 15
1.2.4.2.2 - Dispersion intercontinentale des *Picea* 15
1.2.4.2.3 - Changements topographiques et climatiques globaux 19
1.2.4.3 - La spéciation des *Picea* 21
1.2.4.3.1 - Facteurs et mécanismes de spéciation des *Picea* 21
1.2.4.3.2 - Liens entre les *Picea* eurasiatiques et hypothèses de spéciation géographique 22
1.2.4.3.3 - Liens entre les *Picea* nord-américains et hypothèses de spéciation géographique 26
1.2.5 - Résumé de la problématique 28
1.3 - La phylogénie moléculaire des *Picea* par séquençage de l'ADN 29
1.3.1 - L'avenue moléculaire: généralités 29
1.3.2 - Les études de phylogénie moléculaire chez les conifères 30
1.3.3 - Le choix des régions génomiques 32
1.3.3.1 - L'ADN nucléaire 32
1.3.3.1.1 - Généralités 32
1.3.3.1.2 - L'utilisation des régions nucléaires pour les *Picea* 33
1.3.3.1.3 - La particularité des polymorphismes trans-spécifiques 35
1.3.3.2 - Les génomes d'organelles 39
1.3.3.2.1 - Transmission différentielle des génomes d'organelles: généralités 39

1.3.3.2.2 - Les génomes d'organelles pour la phylogénie des *Picea* 40
1.3.3.2.3 - Le choix et la structure des régions de l'ADN chloroplastique 43
1.3.3.2.4 - Le choix et la structure de la région de l'ADN mitochondrial 45
1.4 - Objectifs de la thèse 48
1.5 - Références 50

Chapitre 2 **- Les polymorphismes trans-spécifiques partagés entre des espèces conifériennes distinctes chez le genre *Picea* (Pinaceae): implications pour le maintien à long terme de la diversité génétique des arbres**.....59

2.1 - Résumé.....60
2.2 - *Abstract* 61
2.3 - *Introduction*.....62
2.4 - *Material and methods* 63
2.4.1 - *Locus and data sampling*.....63
2.4.2 - *Biochemical methods* 64
2.4.3 - *Data analysis* 65
2.5 - *Results* 69
2.5.1 - *Locus* Sb16.....69
2.5.2 - *Locus* Sb29.....74
2.5.3 - *Locus* Sb62.....76
2.5.4 - *Trans-species shared polymorphisms* 78
2.5.5 - *Divergence time between the three* Picea *species and mutation rate at nuclear gene loci* 80
2.6 - *Discussion* 81
2.6.1 - *Diversity at nuclear gene loci within spruce taxa*.....81
2.6.2 - *Allele coalescence and historical population size*.....82
2.6.3 - *Implications for the long-term maintenance of genetic diversity in trees* 85
2.6.4 - *Implications for estimating species phylogenies from nuclear genes* 87
2.7 - *Acknowledgments* 89
2.8 - *References* 89

Chapitre 3 **- Incompatibilité des phylogénies chloroplastiques et mitochondriale chez *Picea*: la spéciation géographique et la réticulation comme mécanismes de l'évolution à long terme chez les genres de Pinaceae**.....94

3.1 - Résumé.....95
3.2 - *Abstract* 96
3.3 - *Introduction* 97
3.4 - *Material and methods* 99
3.4.1 - *Taxa sampling*.....99
3.4.2 - *DNA regions sampling* 99
3.4.3 - *Biochemical methods* 99

3.4.4 - Phylogenetic analyses 101
3.4.5 - Mapping of morphological classifications onto phylogenetic trees 103
3.5 - Results 104
3.5.1 - Sequence variation 104
3.5.2 - Phylogenies based on cpDNA 105
3.5.3 - Phylogeny based on mtDNA 106
3.5.4 - Topological comparisons 112
3.6 - Discussion 113
3.6.1 - cpDNA phylogenies 113
3.6.2 - mtDNA phylogeny 114
3.6.3 - Incongruence with morphological classifications 116
3.6.4 - Evidence for geographic speciation and reticulate evolution 117
3.6.5 - Implications for other Pinaceae and conifers 121
3.7 - Acknowledgments 121
3.8 - Litterature cited 122
3.9 - Appendix 1 127

***Chapitre 4* - Conclusions et perspectives** 129

4.1 - Phylogénie des *Picea* basée sur les gènes nucléaires 129
4.2 - L'histoire biogéographique du genre *Picea* 131
4.3 - Les perspectives 134
4.3.1 - Les études phylogénétiques 134
4.3.2 - Les études de génétique des populations 138
4.3.3 - La cartographie des génomes de *Picea* 139
4.4 - Conclusion 139
4.5 - Références 141

Appendice 143

Liste des tableaux

Tableau 1.1 - Regroupement des taxons de *Picea* dans les divisions principales de quelques classifications morphologiques....8
Tableau 1.2 - Périodes géologiques et âge relatif bornant chacune....13

Table 2.1 - *Parameters derived from the sequence alignment of alleles from three distantly related* Picea *species at three nuclear gene loci*....69
Table 2.2 - *Intraspecific diversity parameters for three nuclear gene loci in three distantly related* Picea *species*....71
Table 2.3 - *Variable sites for the locus* Sb16 *in* Picea *spp*....72
Table 2.4 - *Pairwise divergence per site and net pairwise divergence per site between species for each of three nuclear gene loci*....73
Table 2.5 - *Variable sites for the locus* Sb29 *in* Picea *spp*....75
Table 2.6 - *Variable sites for the locus* Sb62 *in* Picea *spp*....77
Table 2.7 - *Types of polymorphisms detected between* Picea *species at each of three nuclear gene loci*....79
Table 2.8 - *Estimates of historical population size and allele coalescence time from data at three nuclear gene loci in three distantly related* Picea *species*....85

Table 3.1 - *PCR primers used in this study*....100
Table 3.2 - *PCR cycling parameters for gene amplification*....100
Table 3.3 - *Names of taxa analyzed and their subspecific classification according to Liu (1982) and Farjon (1990*....103
Table 3.4 - *Summary of sequence variation and DNA divergence found in the genus* Picea....104
Table 3.5 - *Significant differences involving North American taxa in topology comparisons (Wilcoxon signed-ranks test)*....112

Liste des figures

Figure 1.1 - Répartition générale de toutes les espèces du genre *Picea*.....3
Figure 1.2 - Types de généalogies: monophylie, paraphylie et polyphylie.....36
Figure 1.3 - Organisation des gènes et régions intergéniques pour les loci étudiés.....45

Figure 2.1 - *Estimating the divergence time between* Picea abies, P. glauca, *and* P. mariana.....68
Figure 2.2 - *Phylogeny of spruce alleles at locus* Sb16.....70
Figure 2.3 - *Phylogeny of spruce alleles at locus* Sb29.....76
Figure 2.4 - *Phylogeny of spruce alleles at locus* Sb62.....78

Figure 3.1 - *Position of PCR primers designed in this study for amplification of mitochondrial* nad1 *intron 2*.....100
Figure 3.2 - *Strict consensus trees of* rbcL/trnTLF *partition,* trnK *partition, combined cpDNA data sets, and mtDNA data set*.....107

Liste des appendices

Appendice 1 - Voucher information and GenBank accession numbers.....127
Appendice 2 - Alignements de séquences (programmes informatiques).....143

Chapitre 1

Introduction générale

Les vastes peuplements d'épinettes forment une ressource forestière majeure dans le monde. Au Canada, plantés par centaines de millions lors des reboisements annuels (CCMF 2005), les *Picea* sont considérés comme des plantes polyvalentes et précieuses qui apportent divers avantages environnementaux, économiques et sociaux (Génome Canada 2006).

Une meilleure connaissance fondamentale de cette ressource forestière, à travers son histoire biogéographique et ses particularités génétiques intrinsèques, peut s'avérer utile à plusieurs points de vue. Au niveau pratique, les nouvelles informations phylogéographiques peuvent permettre de mieux comprendre la dispersion des plantes, de même que la spéciation et ses causes sous-jacentes, lesquelles seraient reliées en partie aux changements géologiques et climatiques. Aussi, les informations sur les liens généalogiques découlant des études phylogénétiques devraient permettre de reconnaître plus facilement les espèces génétiquement voisines, afin de mieux comprendre, entre autres, les succès rapportés suite aux essais de croisements interspécifiques chez les *Picea*. Aussi, suivant le degré de proximité phylogénétique et taxonomique nouvellement déterminé entre les espèces, les possibilités de transfert de l'information génomique entre celles-ci seraient mieux définies, plus particulièrement dans le cadre des programmes de génomique et de développement de marqueurs pour la sélection. Par ailleurs, grâce à de nouvelles informations tirées des particularités des génomes des *Picea*, le modèle des *Picea* pourrait être appliqué par extension à d'autres essences coniférienes de mêmes caractéristiques, dans des buts semblables.

Quelles que soient les applications ultimes de l'apport scientifique de la présente étude, il demeure que celle-ci vise à préciser l'histoire évolutive de ce genre au niveau mondial pour permettre d'asseoir sa systématique sur des bases évolutives solides. La

reconnaissance des espèces, la systématique et la taxonomie constituent en effet des instruments fondamentaux en biologie (*e.g.* Davidson 1952, Mayr 1963, Hennig 1966), au même titre que le dictionnaire l'est en linguistique. Naturellement, c'est en s'appuyant sur les connaissances actuelles qu'il sera possible de procéder à de telles inférences pour les *Picea*. Le présent chapitre vise donc à synthétiser l'état des connaissances préalables à l'atteinte des objectifs de cette thèse.

1.1 - Le genre *Picea*: généralités

Le genre conifèrien *Picea*, de la famille des Pinaceae, est représenté par trente à quarante espèces à travers le monde suivant les classifications usuelles (*e.g.* Liu 1982, Farjon 1990). Quoique la majorité de ces espèces croissent en Chine, les *Picea* sont distribués dans tout l'hémisphère nord, soit en Europe, en Asie et en Amérique du Nord (Fig. 1 1). Les épinettes habitent des sites écologiques variés sous des climats différents. En général, dans les régions septentrionales de l'hémisphère nord, elles vivent au niveau des plaines dans les forêts boréales; dans la partie plus méridionale de l'hémisphère, les épinettes se trouvent plutôt en altitude, dans les régions montagneuses ou subalpines (Page et Hollands 1987).

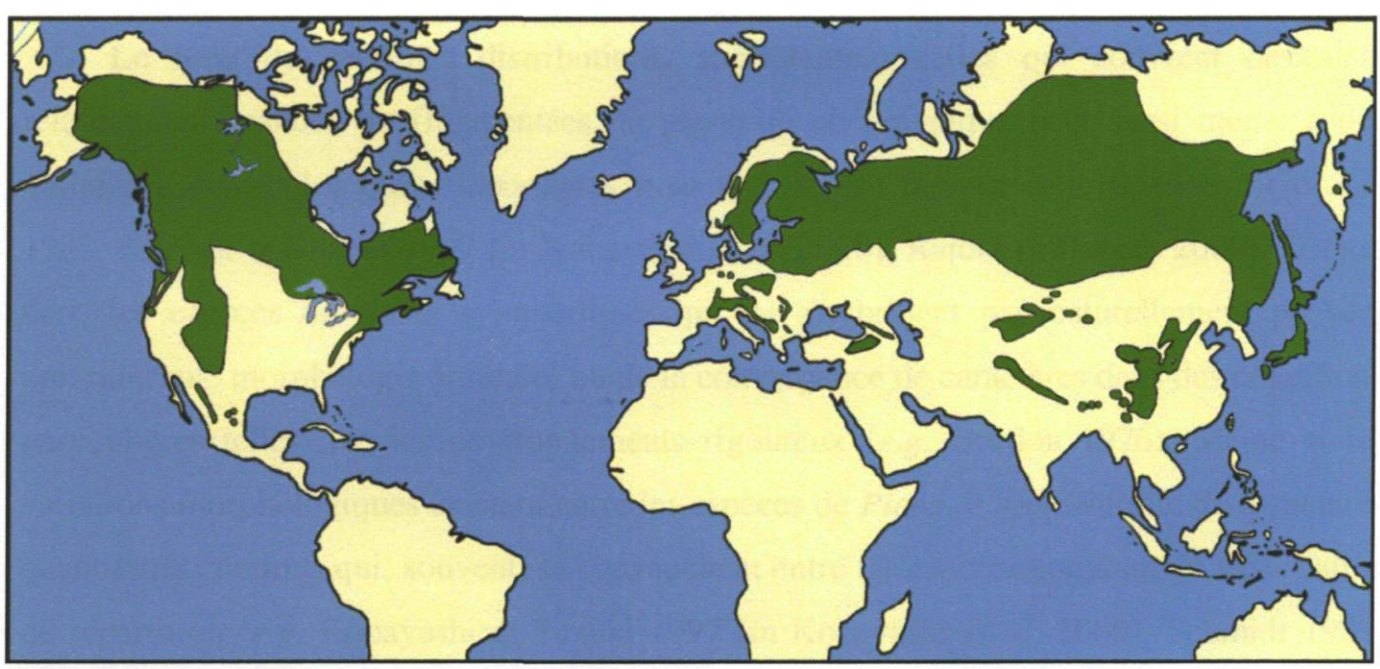

Fig. 1.1 Répartition générale de toutes les espèces du genre *Picea*.

1.2 - Problématique

1.2.1 - Morphologie des *Picea*

Plus de la moitié des espèces de *Picea* actuelles possèdent des aires de distribution allopatriques, ce qui supposerait habituellement la présence de distinctions morphologiques évidentes entre ces espèces. En général, la divergence des

caractères entre deux populations en situation d'allopatrie est liée au processus de spéciation. Pour le genre *Picea*, on devrait donc s'attendre à ce que le grand nombre d'espèces allopatriques présentent à tout le moins quelques caractères morphologiques très distincts. Cependant, aucune coupure morphologique franche ne peut être faite à l'intérieur du genre pour regrouper les espèces en subdivisions nettes (Fowler 1966). Le genre *Picea* ressemble plutôt à un ensemble de quelques groupes d'espèces présentant un gradient morphologique autour du pôle nord (Vidakovic 1991). En effet, les variations morphologiques et anatomiques sont considérées assez uniformes à travers le genre (Wright 1955; Mikkola 1969; Kobayashi *et al.* 2000). Cependant, elles mènent parfois à une forte variance intraspécifique (Liu 1982) en ce qui a trait, par exemple, à la morphologie des cônes, des graines et des ailes des graines (*e.g.* Turna 2004).

Le long de certaines distributions, spécialement celles qui couvrent des aires relativement grandes ou fragmentées, la plasticité phénotypique peut aussi mener à des variations morphologiques importantes pour une espèce donnée (*e.g.* Gordon 1976, Liu 1982, Konishi et Suzuki 1997 [*in* Kobayashi *et al.* 2000], Rajora et Dancik 2000). D'autre part, les espèces de *Picea* sympatriques qui ne s'hybrident pas naturellement peuvent présenter une morphologie similaire ou de la convergence de caractères dans des conditions particulières telles que les environnements rigoureux (*e.g.* Gordon 1976). Même si les variations morphologiques existent entre les espèces de *Picea*, il s'agit surtout de caractères quantitatifs continus qui, souvent, se chevauchent entre les espèces couvrant de vastes aires de répartition (*e.g.* Kobayashi et Suzuki 1997 [*in* Kobayashi *et al.* 2000], Schmidt 1989, Ohsawa 1994). Ainsi, l'utilisation de caractères morphologiques pour la reconnaissance des taxons de *Picea* et pour déduire les lignées évolutives entre eux peut induire en erreur (Kobayashi *et al.* 2000).

1.2.2 - Historique de la classification des *Picea*

D'étude en étude, les nouvelles informations sur les caractères comparatifs utilisés dans les classifications visent à préciser les liens entre les espèces de *Picea* et tentent de justifier l'élévation d'une variété au niveau taxonomique de l'espèce, ou la réduction d'une espèce en sous-espèce ou variété. Aussi, à travers les divers systèmes de classification des *Picea*, jusqu'à 50 taxons de *Picea* différents ont déjà été élevés au rang taxonomique de l'espèce, desquels on en reconnaît aujourd'hui moins d'une quarantaine. La classification des espèces de *Picea* est depuis longtemps problématique (revue dans Liu 1982), principalement à cause de la présence de faibles variations morphologiques interspécifiques (Weng et Jackson 2000; LePage 2001), lesquelles se confinent parfois au niveau de variation que l'on peut retrouver à l'intérieur même de certaines espèces de *Picea* (Kobayashi *et al.* 2000). Wright (1955) voyait même autant de difficulté à identifier les espèces d'épinettes que de discerner les écotypes du frêne blanc! Aussi, la classification de ces espèces peut varier selon l'importance que l'on donne à un caractère ou à un autre. À cela, il faut ajouter le problème de l'inaccessibilité de certaines régions géographiques pour y identifier des spécimens représentatifs de l'étendue de la variabilité morphologique d'une espèce. C'est le cas spécialement en Chine, un pays aux montagnes difficiles d'accès dans lesquelles sont présentes de nombreuses espèces de *Picea* (Farjon 1990).

La restriction à utiliser des variations quantitatives de caractères morphologiques pour distinguer les espèces de *Picea* est étayée par le grand nombre de classifications de ce genre qui ont été élaborées depuis plus d'une centaine d'années. En effet, depuis la fin du 19^{e} siècle sont apparues de nombreuses classifications et considérations taxonomiques différentes pour les *Picea*, dont celles de Willkomm (1887), Mayr (1890), Pardé (1911), Lacassagne (1934), Dallimore et Jackson (1948), Pilger (1960), Debazag (1964), Gaussen (1966), Harrison (1966), Colleau (1968), Bobrov (1970), Schmidt-Vogt (1977), Liu (1982), Krüssmann (1972, 1985), Silba (1984, 1986), Rushforth (1987), Schmidt (1989), Farjon (1990) (voir Liu [1982] et LePage [2001] pour ces références) et LePage (2003a). Des opinions supplémentaires telles que celles de Wright (1955) et Vogellehner (1977 [*in*

Vidacovic 1991]) ont associé les affinités reproductives aux caractéristiques morphologiques et anatomiques et ont tenté de préciser les liens de parenté entre les espèces. Aussi, Vogellehner (1977 [*in* Vidacovic 1991]) s'est-il basé sur les affinités de croisements pour conclure que toutes les espèces chinoises pourraient être regroupées en trois espèces seulement, dont deux seraient étroitement reliées (Vidacovic 1991). Les spécialistes de ces conifères ont proposé, et continuent de proposer, différentes façons d'interpréter les caractères morphologiques, biochimiques, phénologiques et les résultats des essais de croisements (*e.g.* Wright 1955, Gordon 1976 pour quelques espèces américaines, Liu 1982, Farjon 1990, Sigurgeirsson et Szmidt 1993; voir revue dans LePage 2001).

Malgré que certains auteurs soient d'avis qu'il n'y a pas suffisamment de démarcation à travers le genre pour justifier sa division en sections ou en sous-genres (*e.g.* Wright 1955, Page et Hollands 1987, Miller 1989, Vidacovic 1991), la grande majorité des classifications de *Picea* regroupent les espèces en de tels sous-groupes. La plupart des classifications de *Picea* sont basées sur les caractères végétatifs, mais certains auteurs jugent que les caractères reproducteurs offrent une meilleure fiabilité et une plus grande constance parce que ces caractères sont soumis à une pression de sélection plus grande (*e.g.* Farjon 1990, LePage 2001). Par contre, LePage (2001) précisait que l'étendue de la variabilité de certains caractères reproducteurs demeure encore méconnue. Il faut savoir que les caractères végétatifs possèdent également leur lot d'inconnues pour ce qui est de la plus grande partie des espèces de *Picea*. Néanmoins, c'est surtout la combinaison des caractères reproducteurs que Farjon (1990) a employée, en se basant sur la classification de Schmidt (1989 [*in* Farjon 1990]), afin de générer une description du genre *Picea* qui semble aujourd'hui la plus populaire. En plus de nombreuses sous-espèces et variétés, elle fait état de 34 espèces de *Picea* qui sont réparties dans deux sections. Pour le présent ouvrage, c'est ce travail descriptif qui constituera en majeure partie le canevas de dénomination des taxons clés dans le genre *Picea*, pour la plus simple raison qu'il est d'actualité.

Les taxons de *Picea* traités ici sont énumérés au Tableau 1.1. Ils y sont présentés dans les grandes subdivisions de huit classifications différentes, soit celles de Mayr (1890), Lacassagne (1934), Pilger (1960), Gaussen (1966), Bobrov (1970), Liu (1982) et Farjon (1990) (voir Liu [1982] et LePage [2001] pour ces références). La comparaison des grandes subdivisions d'une classification à l'autre permet de réaliser combien se recoupent les différents caractères primordialement considérés pour circonscrire les groupes d'espèces. En effet, les affinités taxonomiques peuvent varier si, par exemple, les caractères reproducteurs sont suivis pour établir les subdivisions de premier niveau (*e.g.* Lacassagne 1934 [*in* LePage 2003a], Liu 1982, Schmidt 1989 [*in* Farjon 1990]). Malgré ces divergences, on peut noter certaines constances d'une classification à l'autre. Quelques espèces semblent souvent classifiées dans les mêmes grandes sections ou les mêmes sous-genres, nonobstant la considération des caractères végétatifs ou reproducteurs, ce qui suggère que les différences sont relativement plus importantes ou plus franches entre ces ensembles d'espèces. Si l'on exclut la classification de Mayr pour laquelle trop peu de taxons peuvent être comparés, on peut remarquer la présence constante des cinq grands regroupements suivants dans plusieurs classifications (voir Tableau 1.1): 1) *P. brachytyla*, *P. breweriana*, *P. omorika* et *P. spinulosa*; 2) *P. glehnii*, *P. koyamai*, *P. maximowiczii*, *P. morrisonicola*, *P. schrenkiana*, *P. smithiana* et *P. wilsonii*; 3) *P. abies*, *P. alcoquiana*, *P. asperata*, *P. glauca*, *P. mariana*, *P. meyeri*, *P. obovata*, *P. retroflexa*, *P. rubens* et *P. torano*; 4) *P. jezoensis* et *P. sitchensis*; 5) *P. engelmannii*, *P. likiangensis*, *P. montigena*, *P. pungens* et - plus récemment inclus - *P. mexicana*. Cependant, les liens de parenté entre ces grands groupes sont variables. Il est à noter que certains taxons tels *P. farreri* et *P. martinezii* n'apparaissent que très récemment dans l'histoire des classifications puisqu'ils ont été découverts respectivement en 1980 et 1988. Sans considérer *P. farreri* ni *P. martinezii*, des lignes sommaires peuvent tout de même être tracées: certains groupes demeurent constants d'une classification à l'autre, mais les taxons *P. chihuahuana*, *P. crassifolia*, *P. koraiensis*, *P. orientalis* et *P. purpurea* restent toujours errants. Par ailleurs, aucun spécimen de *P. neoveitchii*, une espèce asiatique très rare, n'a pu être retracé dans les jardins botaniques et arboretums consultés afin d'en obtenir un échantillon pour le présent travail. Il s'agit de la seule espèce de *Picea* exclue de cet ouvrage.

Tableau 1 1 Regroupement des taxons de *Picea* dans les divisions principales (en caractères gras) de quelques classifications morphologiques. Couleurs ou chiffres: regroupements constants des taxons d'une classification à l'autre. Noter que les groupes constants ne figurent pas dans les mêmes subdivisions (sections ou sous-genres) à l'intérieur des trois dernières classifications (voir Liu [1982] et LePage [2001] pour les références).

Mayr (1890)	Lacassagne (1934)	Pilger (*in* Engler-Prantl 1960)	Gaussen (1966)	Bobrov (1970)	Liu (1982)	Farjon (1990)
Section *Morinda*	**Section *Eupicea***	**Section *Eupicea***	**Section *Eupicea***	**Section *Picea***	**Sous-genre *Picea***	**Section *Picea***
P abies	1 *P abies*	2 *P montigena*	1 *P abies*	1 *P abies*	1 *P abies*	1 *P abies*
P obovata	1 *P obovata*	3 *P koyamai*	1 *P obovata*	1 *P obovata*	1 *P obovata*	1 *P obovata*
P torano	2 *P montigena*	1 *P meyeri*	? *P koraiensis*	? *P koraiensis*	2 *P montigena*	? *P koraiensis*
P mariana	3 *P koyamai*	1 *P asperata*	3 *P koyamai*	3 *P koyamai*	1 *P meyeri*	3 *P koyamai*
P schrenkiana	1 *P meyeri*	1 *P torano*	1 *P meyeri*	1 *P meyeri*	1 *P asperata*	1 *P meyeri*
P smithiana	1 *P asperata*	1 *P retroflexa*	1 *P asperata*	1 *P asperata*	1 *P torano*	1 *P asperata*
P glauca	1 *P torano*	3 *P glehnii*	1 *P torano*	1 *P torano*	1 *P retroflexa*	1 *P torano*
P orientalis	1 *P retroflexa*	3 *P morrisonicola*	1 *P retroflexa*	1 *P retroflexa*	? *P crassifolia*	1 *P retroflexa*
	3 *P glehnii*	1 *P rubens*	3 *P glehnii*	3 *P glehnii*	1 *P rubens*	? *P crassifolia*
Section *Casicta*	3 *P morrisonicola*	1 *P mariana*	3 *P morrisonicola*	3 *P morrisonicola*	1 *P mariana*	3 *P glehnii*
P pungens	1 *P rubens*	3 *P schrenkiana*	1 *P rubens*	1 *P rubens*	? *P purpurea*	3 *P morrisonicola*
P engelmannii	1 *P mariana*	? *P purpurea*	1 *P mariana*	1 *P mariana*	2 *P pungens*	4 *P omorika*
P sitchensis	3 *P schrenkiana*	3 *P smithiana*	3 *P schrenkiana*	3 *P schrenkiana*	2 *P mexicana*	1 *P rubens*
P. jezoensis	3 *P smithiana*	2 *P pungens*	3 *P smithiana*	3 *P smithiana*	2 *P engelmannii*	1 *P mariana*
	2 *P pungens*	2 *P engelmannii*	? *P chihuahuana*	? *P chihuahuana*	5 *P sitchensis*	3 *P schrenkiana*
Section *Omorika*	2 *P engelmannii*	1 *P glauca*	1 *P glauca*	1 *P glauca*	1 *P glauca*	4 *P brachytyla*
P omorika	1 *P glauca*	2 *P likiangensis*	1 *P alcoquiana*	1 *P alcoquiana*	2 *P likiangensis*	4 *P spinulosa*
P glehnii	2 *P likiangensis*	1 *P alcoquiana*	? *P orientalis*	3 *P wilsonii*	1 *P alcoquiana*	3 *P smithiana*
P alcoquiana	1 *P alcoquiana*	3 *P wilsonii*	3 *P wilsonii*	3 *P maximowiczii*	5 *P. jezoensis*	? *P. farreri*
P breweriana	3 *P wilsonii*	3 *P maximowiczii*	3 *P maximowiczii*			? *P chihuahuana*
	3 *P maximowiczii*	? *P crassifolia*		**Section *Casicta***	**Sous-genre *Omorika***	1 *P glauca*
			Section *Casicta*	? *P purpurea*	? *P koraiensis*	1 *P alcoquiana*
	Section *Omorika*	**Section *Omorika***	2 *P montigena*	2 *P pungens*	3 *P koyamai*	? *P orientalis*
	4 *P omorika*	4 *P omorika*	? *P purpurea*	2 *P mexicana*	3 *P glehnii*	3 *P wilsonii*
	? *P crassifolia*	4 *P brachytyla*	2 *P pungens*	2 *P engelmannii*	3 *P morrisonicola*	3 *P maximowiczii*
	? *P purpurea*	4 *P spinulosa*	2 *P engelmannii*	5 *P sitchensis*	4 *P omorika*	4 *P breweriana*
	4 *P brachytyla*	5 *P sitchensis*	2 *P likiangensis*	2 *P likiangensis*	3 *P schrenkiana*	
	4 *P spinulosa*	5 *P. jezoensis*		5 *P. jezoensis*	4 *P brachytyla*	**Section *Casicta***
	5 *P sitchensis*	4 *P breweriana*	**Section *Omorika***		4 *P spinulosa*	2 *P montigena*
	? *P orientalis*		4 *P omorika*	**Section *Omorika***	3 *P smithiana*	? *P purpurea*
	4 *P breweriana*		4 *P brachytyla*	4 *P omorika*	? *P chihuahuana*	2 *P pungens*
			4 *P spinulosa*	4 *P brachytyla*	? *P orientalis*	2 *P mexicana*
			5 *P sitchensis*	4 *P spinulosa*	3 *P wilsonii*	2 *P engelmannii*
			5 *P. jezoensis*	? *P orientalis*	3 *P maximowiczii*	5 *P sitchensis*
			4 *P breweriana*	4 *P breweriana*	4 *P breweriana*	2 *P likiangensis*
						5 *P. jezoensis*

Si la classification traditionnelle utilise les caractères puisés habituellement de la morphologie et de l'anatomie pour créer les groupes précités, c'est que ces caractères sont accessibles, car ils sont directement observables sur les arbres. Toutefois, les résultats des croisements artificiels ont contribué à élucider dans une certaine mesure les liens évolutifs entre les *Picea* ou les groupes de *Picea*, et ils ont tenté de corroborer les résultats découlant des méthodes traditionnelles de classification basées sur la morphologie. Force est d'admettre que, effectivement, d'autres critères d'appariement des espèces s'imposent pour retracer les patrons de spéciation des *Picea* en plus de la classification morphologique. En effet, la classification n'est pas que l'établissement d'une clef d'identification: elle devrait être aussi le reflet de l'évolution naturelle d'un groupe de taxons à travers leurs caractères communs et leurs caractères distincts. Jusqu'ici, les systèmes de classification établis pour les *Picea* ne présentent aucun consensus absolu entre les auteurs quant aux liens de parenté entre les espèces, ni en ce qui a trait au nombre d'espèces (Page et Hollands 1987; LePage 2001). Par conséquent, aucune histoire évolutive ne ressort clairement des observations.

1.2.3 - L'hybridation chez les *Picea*

Le problème de l'établissement des liens de parenté entre les *Picea* tient aussi des phénomènes de l'hybridation et de l'introgression qui viennent brouiller les pistes. D'après le principe de la spéciation suivant l'isolement reproducteur, les taxons qui se croisent facilement devraient posséder des affinités génétiques et des ressemblances plus importantes, ainsi qu'une origine commune plus récente par rapport aux taxons isolés reproductivement (voir les concepts de Dobzhansky [1941], Mayr [1942] et Hennig [1966]). D'ailleurs, quelques couples d'espèces de *Picea* forment des hybrides naturels observables: *P. abies* et *P. obovata* à la jonction de l'Europe et de l'Asie nordiques, *P. mariana* et *P. rubens* dans le nord-est de l'Amérique du Nord, et le trio *P. engelmannii*, *P. glauca* et *P. sitchensis* dans le nord-ouest de l'Amérique du Nord. On pourrait donc s'attendre à ce que ces paires d'espèces soient considérées proches parentes et classifiées comme telles, dans les mêmes subdivisions. C'est le cas des deux premières paires d'espèces seulement. Les membres du dernier trio sont connus pour former de vastes zones d'hybridation naturelle dans le nord-ouest de l'Amérique du Nord (*e.g.* Little 1953 [*in*

Wright 1955], Wright 1955, Rajora et Dancik 2000, Bennuah *et al.* 2004). Il a aussi été démontré que *P. engelmannii* et *P. glauca* devraient plutôt être considérés comme des sous-espèces de *P. glauca* plutôt que des espèces distinctes, justification faite de leur manque d'isolement reproducteur et d'autres caractéristiques (Rajora et Dancik 2000). Cependant, ces trois espèces sont maintenues jusqu'à ce jour dans des sections ou sous-genres différents dans la majorité des classifications traditionnelles (voir Tableau 1.1). D'autre part, des espèces considérablement étudiées comme *P. glauca* et *P. mariana*, qui partagent une immense aire de distribution au Canada d'un océan à l'autre, sont différentes morphologiquement et vraisemblablement isolées reproductivement puisqu'elles ne présentent aucune zone d'hybridation naturelle malgré leur chevauchement d'aires sur de grandes étendues latitudinales et longitudinales. Des essais de croisements artificiels ont démontré que ces deux espèces sont effectivement incompatibles (*e.g.* Mikkola 1969). Puisque leur isolement reproducteur semble complet, elles devraient posséder un ancêtre commun plus lointain dans le temps que les paires d'espèces qui s'hybrident. Paradoxalement, ces deux espèces se retrouvent souvent dans les mêmes subdivisions dans de nombreuses classifications morphologiques, tandis que *P. engelmannii*, *P. sitchensis* et *P. glauca*, qui forment des hybrides, figurent dans des subdivisions différentes (*e.g.* Mayr 1890 et Gaussen 1966 [*in* Liu 1982], Liu 1982, Schmidt 1989 [*in* Farjon 1990]). La classification la plus frappante qui illustre ce paradoxe est celle de Gaussen (1966 [réf. *in* Liu 1982]).

Les résultats positifs de l'hybridation contrôlée entre plusieurs espèces de *Picea* (*e.g.* Wright 1955, Mikkola 1969, Gordon 1976, Gordon 1986), sans malheureusement impliquer le genre en son entier, ont mené à la conclusion que l'hybridation est possible à travers le genre entre de nombreuses espèces qui semblent à première vue évolutivement éloignées. En effet, plusieurs essais de croisements ont très bien réussi même si les taxons traités provenaient de régions géographiques distantes, sur différents continents de surcroît (*e.g.* Mikkola 1969). En ayant tenté de préciser les liens naturels qui persistent entre les espèces de *Picea*, ces études sur l'hybridation interspécifique ont laissé entrevoir que tous les *Picea* semblent assez proches génétiquement les uns des autres, d'un bout à l'autre de la planète. Ceci peut expliquer en partie les difficultés rencontrées lors de la classification des

Picea. En effet, une proximité génétique plus importante sous-tendrait généralement une moins grande divergence de phénotype en milieu contrôlé. Il est donc possible que certaines variations morphologiques considérées pour la classification d'espèces de *Picea* géographiquement très éloignées soient attribuables à la plasticité phénotypique, et que la convergence de certains caractères morphologiques soit plus importante dans ce contexte.

La possibilité d'hybridation entre plusieurs espèces de *Picea* est surprenante compte tenu du grand nombre d'espèces de ce genre. En effet, plus il y a d'événements de divergence qui séparent les taxons, plus les taxons se trouvent éloignés génétiquement suite à l'accumulation des mutations, ce qui diminue graduellement la compatibilité génétique entre les taxons jusqu'à la spéciation complète. De façon générale, on peut donc supposer que la possibilité d'hybridation est d'autant diminuée que la séparation des espèces est lointaine dans le temps (Rieseberg et Wendel 1993). Aussi, l'hybridation naturelle prendrait place plutôt entre des espèces parapatriques ou allopatriques (Wright 1955; Perron et Bousquet 1997) qui ont divergé récemment, comme c'est le cas pour des espèces ayant divergé il y a moins d'un million d'années telles, vraisemblablement, *P. mariana* et son espèce dérivée, *P. rubens* (Perron *et al.* 2000), ou encore *P. glauca* et *P. engelmannii* (*e.g.* Rajora et Dancik 2000), mais aussi entre des espèces qui présentent un contact secondaire récent suite à la dernière glaciation, comme la situation de *P. abies* et *P. obovata* le suppose (Wright 1955).

L'hybridation si répandue chez les *Picea* et leur manque de définition morphologique a mené certains auteurs à croire que le genre est récent (Mikkola 1969) et que les événements de divergence auraient conséquemment été très rapprochés dans le temps. D'autres auteurs considèrent que les *Picea*, tout comme les autres conifères, auraient évolué lentement (Wright 1955; Prager *et al.* 1976), ce qui expliquerait leur uniformité morphologique. Par conséquent, la divergence entre les ancêtres communs aux espèces contemporaines aurait eu lieu à une époque relativement lointaine. En effet, depuis la divergence entre les lignées anciennes menant aux *Picea* et aux *Pinus* contemporains, il y a 120 à 140 millions d'années (Florin 1963; Miller 1989), les plus anciens fossiles d'espèces

de *Picea* aujourd'hui disparues datent d'au moins 45 millions d'années (LePage 2001), et d'autres fossiles découverts, représentatifs de certaines espèces contemporaines distinctes, datent d'au moins 15 à 20 millions d'années (Wolfe 1964).

Donc, compte tenu du grand nombre d'espèces et, surtout, de l'âge apparemment ancien du genre dans son ensemble, l'hybridation interspécifique artificielle répandue à travers le genre apparaît surprenante. Cependant, même dans l'hypothèse où le genre *Picea* est relativement ancien, ses composantes pourraient avoir subi une spéciation très lente par rapport à leur capacité d'envahissement des territoires. Cette spéciation lente serait possiblement liée au fait que les grandes tailles de populations chez *Picea* retarderaient la fixation des nouvelles mutations (Bousquet *et al.* 1992; Bouillé et Bousquet 2005, voir chapitre 2 de la présente thèse).

1.2.4 - Histoire biogéographique des *Picea*

1.2.4.1 - Les fossiles de *Picea*

Des traces anciennes de colonisation des territoires par les *Picea* ont été retrouvées dans l'hémisphère boréal jusque dans l'Arctique (LePage 2001). Les plus anciens fossiles de *Picea* qui ont été retrouvés datent du milieu de l'éocène (voir Tableau 1.2), soit d'environ 45 millions d'années. Il s'agit de *P. heibergii* LePage, *P. nansenii* LePage, *P. palustris* LePage et *P. banksii* Hills et Ogilvie, des fossiles très bien préservés dans l'extrême nord arctique du Canada sur l'île Axel Heiberg (LePage 2001). *P. banksii* est aussi représenté dans le Canada arctique pour la période du miocène (LePage 2001), ce qui lui prête une présence de quelques dizaines de millions d'années dans l'extrême nord de l'Amérique du Nord.

Précédant cette découverte, près d'une quarantaine d'espèces fossiles de *Picea* datant de diverses périodes géologiques ont été mis au jour ailleurs dans le monde (revue dans LePage 2001). Parmi ces espèces fossiles, certaines ont été retrouvées sur des continents aujourd'hui disjoints: *Piceoxylon* datant du miocène dans l'ouest de l'Amérique du Nord et

du pliocène en Allemagne; puis, *P. magna* dans l'ouest de l'Amérique du Nord ainsi que sur la côte orientale russe et au Japon. *P. magna* est celui dont on a retrouvé des traces dans toutes les époques géologiques depuis l'éocène jusqu'au pliocène, soit une présence constante étalée sur plusieurs dizaines de millions d'années, sur des continents différents de surcroît. Ainsi, puisque ces deux espèces fossiles sont représentées sur des continents différents, elles nous informent déjà sur la possibilité d'un mouvement de colonisation intercontinental entre l'Asie et l'Amérique d'une part, et possiblement entre l'Europe et l'Amérique d'autre part.

Tableau 1.2 - Périodes géologiques et âge relatif bornant chacune. mA: millions d'années (adapté de LePage 2003a).

Ère	*Sous-ère*	*Période*	*Époque*		*Âge (mA)*
Cénozoïque	Quaternaire		Holocène		0,01
			Pléistocène		1,64
	Tertiaire	Néogène	Pliocène	Fin	3,4
				Début	5,2
			Miocène	Fin	10,4
				Milieu	16,3
				Début	23,3
		Paléogène	Oligocène	Fin	29,3
				Début	35,4
			Éocène	Fin	38,6
				Milieu	50
				Début	56,5
			Paléocène	Fin	60,5
				Début	65
Mésozoïque		Crétacé			145,6

Parmi les fossiles des *Picea* contemporains, on a retrouvé des indices de la présence de *P. breweriana* en Amérique du Nord datant du miocène près de son aire de répartition actuelle; *P. mariana* au Groenland datant du pliocène/pléistocène; en Asie, *P. jezoensis* et *P. maximowiczii* du pliocène, ainsi que des vestiges de *P. polita* du pliocène et du pléistocène; enfin, en Europe, *P. abies* (syn. *P. excelsa* et possiblement *P. rubra in* LePage

2001), *P. polita* et *P. glehnii* datant tous du pliocène. *P. breweriana* est l'espèce contemporaine considérée la plus ancestrale et la plus divergente en Amérique du Nord d'après des études moléculaires (Ledig *et al.* 2004), son isolement reproducteur des autres espèces de *Picea* (Gordon 1986), et différents attributs morphologiques, anatomiques et génétiques. Sa présence ancienne au miocène synchrone avec cinq autres espèces fossiles incluant *P. magna* et *Piceoxylon* démontre qu'il y avait déjà une variété d'espèces en Amérique du Nord avant la divergence initiale des taxons nord-américains contemporains. L'espèce relique *P. breweriana* est fort possiblement le dernier représentant d'un groupe d'espèces nord-américaines aujourd'hui disparues. Par ailleurs, *P. polita* a une distribution actuelle confinée aux îles japonaises, mais des fossiles de cette espèce ont été retrouvés en Pologne dans un site datant du pliocène. Il s'agit là d'un exemple démontrant que certaines espèces de *Picea* aujourd'hui presque reliques ont pu coloniser de très vastes territoires.

En ce qui concerne la distinction entre les espèces fossiles de *Picea*, elle est reléguée à la même problématique que celle du présent ouvrage concernant l'identification ou la classification d'après les caractères morphologiques. En effet, la comparaison entre les fossiles et les espèces vivantes est déjà difficile à cause du nombre limité de caractères offerts par les fossiles pour effectuer les comparaisons (LePage 2003a). Ce faible nombre de caractères morphologiques pouvant avoir été influencés par la plasticité phénotypique chez les *Picea*, la comparaison entre des fossiles de périodes géologiques différentes, de climats et environnements différents, doit donc être d'autant nuancée. Cependant, les trois espèces fossiles découvertes au même endroit et dans les mêmes couches géologiques au nord de l'Amérique du Nord nous informent de manière plus sûre que la spéciation à l'intérieur du genre avait eu lieu très certainement avant le milieu de l'éocène (LePage 2003a). Ceci rejoint en partie les hypothèses de Miller (1988, 1989), qui avançait que ce genre conifèrien se serait différencié vers la fin du crétacé ou vers le début du paléogène (voir Tableau 1.2). Donc, vu le nombre d'espèces différenciées déjà présentes à l'éocène, l'évolution ou la différenciation des *Picea* se seraient produites assez tôt suite à la divergence du genre (LePage 2003a).

1.2.4.2 - Hypothèses biogéographiques

1.2.4.2.1 - Origine du genre *Picea*

La région d'origine des *Picea* précédant leur diversification hâtive demeure incertaine puisque les opinions divergent. Li (1953) a suggéré que presque tous les conifères tireraient leur origine du bassin Pacifique. Suivant cette hypothèse, les *Picea* trouveraient leur origine en Asie de l'est ou dans l'ouest de l'Amérique. Parallèlement, une étude exhaustive des croisements artificiels en corrélation avec la morphologie, la distribution géographique et la phylogénie des *Picea* a amené Wright (1955) à conclure que l'origine du genre se situerait plus précisément en Asie du sud-est, plus probablement sur le continent asiatique que sur les îles environnantes, à cause de la grande diversité des espèces sur le continent et la présence de *P. koyamai* qu'il voyait comme étant l'espèce la plus primitive. Cette origine aurait été suivie de plusieurs dispersions subséquentes vers l'Amérique. Les hypothèses de Wright (1955) concernant les espèces-liens entre les continents ont été par la suite contredites par Fowler (1966) et Hills et Ogilvie (1970). En effet, Wright (1955) avançait que *P. jezoensis* en Asie faisait le lien entre les taxons asiatiques et les taxons américains, tandis que Fowler (1966) et Hills et Ogilvie (1970) postulaient qu'il s'agissait plutôt de *P. glauca* en Amérique du Nord, ou d'une espèce fossile y étant liée, suivant plutôt une dispersion venant de l'Europe en direction de l'Amérique durant le crétacé (Vidacovic 1991). Plus récemment, en se basant sur une étude moléculaire du génome chloroplastique, Sigurgeirsson et Szmidt (1993) concluaient que l'origine du genre *Picea* se situerait plutôt en Amérique. Cependant, les résultats de la présente thèse remettent en question l'utilité du génome chloroplastique pour vérifier les hypothèses biogéographiques des *Picea* (voir chapitre 3).

1.2.4.2.2 - Dispersion intercontinentale des *Picea*

Pour alimenter ces hypothèses sur les passages intercontinentaux, il y a plusieurs preuves que des liens biogéographiques ont existé depuis au moins le début de l'éocène au nord des océans Atlantique et Pacifique. Au début du tertiaire, il y a moins de 60 millions d'années, des liens intercontinentaux existaient déjà entre l'Asie,

l'Amérique du Nord et l'Europe sous forme de ponts terrestres intermittents qui ont favorisé la dispersion des animaux et des plantes d'un continent à l'autre (LePage 2003a).

Entre l'Europe et l'Amérique, la route DeGeer faisait le lien entre l'Amérique du Nord et la Fennoscandie pendant le paléocène et l'éocène, soit il y a environ 65 à 35 millions d'années (LePage 2003a). Durant la même période, la route *Thulian* constituait aussi un lien entre le Groenland et l'Europe, cependant plus ponctuellement dans le temps ou encore sous forme d'une série d'îles (LePage 2003a). Les preuves géologiques (voir Tiffney 2000) telles que ces routes démontrent qu'il y avait des échanges possibles entre l'Europe et l'Amérique au moins jusqu'à l'éocène, et qu'ensuite, ces ponts terrestres auraient disparu. Cependant, les comparaisons floristiques démontrent que les échanges se seraient perpétués longtemps après l'éocène par ces passages nordiques, même après le déplacement rapide du Groenland vers le nord-ouest suite à sa séparation de l'Europe vers la fin de l'éocène.

Effectivement, pour la période couvrant au moins tout le miocène, divers auteurs ont rapporté des similitudes floristiques significatives entre l'Europe et l'Amérique (Tiffney 2000). Il est aussi possible que des échanges aient eu lieu plus tard que le miocène, lors des différents cycles des dernières périodes glaciaires. Des ponts intercontinentaux pourraient avoir existé durant le pléistocène car les niveaux des mers étaient alors vraisemblablement différents dans les environs du nord de l'Atlantique. À ce sujet, lors de la dernière glaciation, les niveaux océaniques auraient été plus bas d'au moins une centaine de mètres dans les environs de la mer Méditerranée (Rabineau *et al.* 2006). L'expansion importante des populations d'arbres le long du littoral du Portugal il y a environ 200 000 à 300 000 ans (Roucoux *et al.* 2006) démontre qu'il est aussi possible que l'expansion des forêts se soit produite à quelques centaines de kilomètres plus au nord lors des mêmes conditions de niveaux océaniques beaucoup plus bas qu'aujourd'hui. La végétation, et surtout les espèces de conifères pionnières telles les *Picea*, aurait gagné du terrain en colonisant les sols riches nouvellement disponibles. D'ailleurs, une augmentation abrupte de la température de 10^0C au Groenland dans les 100 000 dernières années (Steig 2006) aurait pu permettre à la

végétation de progresser sur ces territoires entre l'Europe et l'Amérique. Pour les 150 000 dernières années, plus d'une augmentation rapide de la température au Groenland ont été inférées, avec des périodes de refroidissement plus lentes (EPICA *community members* 2006). Ces conditions particulières d'oscillations du climat et des niveaux de la mer dans les environs du Groenland auraient offert des possibilités de passages intercontinentaux pour la faune et la flore, et éventuellement pour l'être humain, comme le suppose le rapprochement génétique établi par une variante du « haplogroupe X » du génome mitochondrial entre certains Européens autochtones et des Amérindiens de l'est de l'Amérique du Nord (Brown *et al.* 1998; Reidla *et al.* 2003).

Les fortes similitudes rapprochant évolutivement l'européen *P. omorika* et les nord-américains *P. mariana* et *P. rubens* sont connues depuis longtemps (*e.g.* Wright 1955, Gordon 1976). Cependant, même si la dispersion intercontinentale de *P. mariana*, *P. rubens*, *P. omorika* ou de leur ancêtre commun demeure aussi obscure que celle d'autres groupes de plantes après l'éocène (Tiffney 2000), elle implique probablement la région du Groenland où des fossiles de *P. mariana* datant du pliocène/pléistocène ont été mis au jour. *P. omorika*, qui croît dans les montagnes Tara dans les Balkans, est considéré comme une espèce relique de la période tertiaire (Vidacovic 1991). Ses ancêtres seraient possiblement *P. paleomorika* et *P. omorikoides*, qui étaient dispersés jusque dans le nord de l'Europe et en Asie (Vidacovic 1991). *P. omorikoides* apparaît effectivement parmi les fossiles nord-européens datant de la fin du pléistocène, plus spécifiquement aussi récemment que 57 000 ans (Holyoak 1983). Comme *P. mariana* et *P. rubens* ont vraisemblablement divergé durant la fin du tertiaire, il y a moins d'un million d'années (Perron *et al.* 2000; Jaramillo-Correa *et al.* 2003), et vu le fort rapprochement de ceux-ci avec *P. omorika*, la connexion ancestrale entre les trois n'est probablement pas très lointaine. Cependant, elle demanderait une investigation plus approfondie qui préciserait leur période de divergence et qui déterminerait si, véritablement, les deux taxons ancestraux *P. paleomorika* et *P. omorikoides* sont impliqués dans le lien entre les *Picea* actuels de ce groupe sur les deux continents.

Trois hypothèses peuvent être émises en ce qui concerne la dispersion des plantes telles que les *Picea* entre l'Europe et l'Amérique. Les deux premières hypothèses concernent les échanges floristiques intercontinentaux par le nord de l'Atlantique en rapport avec la dispersion des graines suivant les vents dominants. Si la dispersion des *Picea* s'est effectuée entre les 30° et 60° de latitude nord, alors la direction serait de l'Amérique du Nord vers l'Europe, suivant la prévalence des vents qui étaient à cette époque comme ils le sont aujourd'hui, pour transporter les graines relativement petites des *Picea*. Par contre, si les échanges se sont faits plus haut que 60° nord, il s'agirait d'une dispersion opposée, d'est en ouest, comme la dérive du Groenland le favorise. Pendant le milieu de l'oligocène (29 millions d'années) et jusqu'à la fin du miocène (5 millions d'années), la végétation dominante à l'extrême nord de l'Europe, soit près des paléolatitudes 75 ° à 78° nord, près de Svalbard au nord de la Norvège, était composée de conifères tels les Taxodiaceae et les Pinaceae (Boulter et Manum 1996). Au pliocène, il y a moins de 5 millions d'années, les *Picea* faisaient certainement partie du paysage de l'extrême nord, d'après les vestiges fossiles (Tiffney 2000). Il est donc possible que les *Picea* aient étendu leur aire de répartition entre les 30^{e} et 60^{e} degrés de latitude encore assez récemment, de l'Amérique à l'Europe en passant par le Groenland, comme en témoignent les fossiles de *P. mariana* datant du pliocène/pléistocène; il est aussi plausible que l'événement de dispersion soit survenu plus haut que le 60^{e} degré nord, dans la direction de l'Europe vers l'Amérique, avant que le climat trop froid et la noirceur hivernale persistant de longs mois ne limitent la croissance des arbres. Ces deux options sont pour l'instant équiprobables, et peut-être ont-elles toutes deux participé à la dispersion des *Picea*. Enfin, la troisième hypothèse, peut-être moins probable, établie selon les preuves géologiques, voudrait que les *Picea* aient colonisé les territoires de l'Amérique vers l'Europe, ou l'inverse, en passant par l'Asie (Tiffney 2000), après le retrait des plans d'eau tels le détroit de Turgai à la fin de l'éocène, il y a moins de 38 millions d'années (LePage 2003b). Cette troisième hypothèse biogéographique pourrait être vérifiée à partir de la situation exacte du groupe des trois espèces *P. mariana*, *P. rubens* et *P. omorika* à l'intérieur de la phylogénie moléculaire du genre *Picea* en son entier.

Du côté nord-ouest du continent américain, le corridor de Bering aurait constitué un passage terrestre pour la dispersion de la faune et la flore entre l'Asie et l'Amérique pendant le miocène et ce, jusqu'au pléistocène (LePage 2003a). Sauf durant l'ouverture du détroit de Bering, il y a environ 5 à 7 millions d'années, le passage aurait été possible durant plusieurs millions d'années pour les êtres vivants (Tiffney 2000). À partir de l'éocène, le climat s'est graduellement refroidi, ce qui aurait favorisé le développement de la végétation de mieux en mieux adaptée au froid (White *et al.* 1997). Aussi, durant le miocène, l'occupation d'une partie du corridor de Bering par des forêts semblables à celles de l'Asie de l'est en témoigne (Tiffney 2000). Fort possiblement, les *Picea* auraient également profité de la modification du climat pour s'établir sur d'autres continents. La direction de la colonisation des *Picea* dans cette région n'est cependant pas clairement établie: plusieurs auteurs tels que Wright (1955) ont opté pour la dispersion de l'Asie vers l'Amérique, tandis que deux études moléculaires (Sigurgeirsson et Szmidt 1993; Ran *et al.* 2006) basées sur l'ADN chloroplastique ont démontré une dispersion en sens inverse. Toutefois, tel que mentionné précédemment, les scénarios biogéographiques découlant des études qui utilisent l'ADN chloroplastique peuvent être biaisés (voir chapitre 3).

1.2.4.2.3 - Changements topographiques et climatiques globaux

A la limite de l'éocène et de l'oligocène, il y a environ 35 millions d'années, l'assèchement du détroit de Turgai en Asie et le refroidissement global ont provoqué des changements très importants dans la faune et la flore terrestres (LePage 2003a; LePage 2003b).

Entre l'Asie de l'ouest et l'Asie de l'est, s'étendant de l'océan Arctique jusqu'à la mer de Téthys, le détroit de Turgai aurait été un obstacle majeur aux échanges fauniques et floristiques entre les deux parties du continent depuis le jurassique, c'est-à-dire depuis au moins 145 millions d'années. L'assèchement de ce large détroit aurait permis l'établissement de la flore est-asiatique vers l'ouest de l'Asie et vers l'Europe. Cependant, pour les *Picea*, cette dispersion aurait été plutôt tardive, à cause du climat tropical ou sub-tropical qui régnait alors dans ces régions. Le climat aurait possiblement régulé la

progression de la végétation tempérée vers l'Europe jusqu'au miocène et au pliocène (LePage 2003a). Vraisemblablement pour ces raisons, les *Picea* sont toujours absents des données fossiles européennes jusqu'au miocène (LePage 2003a). Puisque cette période d'obstacle à la dispersion entre l'Asie et l'Europe comprend la période de divergence des genres *Picea* et *Pinus*, c'est-à-dire il y a 120 à 140 millions d'années (Florin 1963; Miller 1989), les *Picea* asiatiques n'auraient eu la chance de s'établir une première fois en Europe qu'après le retrait du détroit de Turgai, et seulement lorsque le climat a commencé à se refroidir.

Le refroidissement global aurait aussi influencé la présence, la persistance et la dispersion des conifères tels que les *Picea*. La planète était en général plus chaude et plus humide que maintenant avant le milieu de l'éocène, il y a environ 40 millions d'années. Par la suite, et spécialement durant les 15 derniers millions d'années, des climats plus froids sont apparus, avec des précipitations régionales plus marquées (Ruddiman et Kutzbach 1991). Ces changements étaient principalement attribuables à l'élévation spectaculaire des plateaux et des montagnes dans le sud de l'Asie et dans l'ouest de l'Amérique du Nord, et qui ont eu différents effets sur l'atmosphère terrestre (Ruddiman et Kutzbach 1991; Zhang *et al.* 2000; Herzschuh *et al.* 2005). En Amérique du Nord, la Cordillère américaine s'est élevée de plus de 3 kilomètres à partir du milieu du miocène (résumé des événements et références dans LePage 2003a); en Asie, surtout vers la fin du miocène, il y a 5 à 10 millions d'années, la chaîne des Himalayas et le Plateau du Tibet se sont élevés d'au moins 2 kilomètres. La formation de ces nouveaux habitats de montagne aurait aussi favorisé l'établissement des *Picea* dans ces régions plus froides présentant les stress environnementaux des hautes altitudes. Ces changements topographiques ont pu également constituer de nouveaux facteurs de vicariance favorisant l'isolement et la spéciation. En concordance avec ces modifications au relief terrestre et ces changements majeurs des biotypes, les données fossiles montrent que la diversification du genre *Picea* paraît notable partout dans l'hémisphère boréal pour la période du miocène. De plus, des études géologiques et biologiques ont révélé que les différences climatiques semblent plus marquées entre la Chine orientale et la Chine occidentale depuis le miocène (Zhang *et al.* 2000). La Chine hébergeant un bon nombre d'espèces de *Picea* contemporains, à tout le

moins une diversité certaine, l'apparition de différentes zones climatiques a très probablement provoqué une vague de spéciation pour les *Picea* distants géographiquement. Par conséquent, des distinctions plus claires devraient se retrouver entre les groupes d'espèces des deux régions de la Chine. Aussi, Wright (1955) a-t-il proposé que l'ensemble des espèces nord-eurasiennes avec l'ensemble des espèces du nord de la Chine et du Formosan constitueraient un groupe d'espèces distinct, à l'exception de *P. obovata*.

1.2.4.3 - La spéciation des *Picea*

1.2.4.3.1 - Facteurs et mécanismes de spéciation des *Picea*

De manière générale, le facteur principalement responsable de la spéciation des *Picea* serait l'isolement géographique (Wright 1955). Ceci impliquerait donc la spéciation allopatrique suite à la fragmentation des habitats (Mayr 1942), nommée « événement de vicariance », ou encore suite à la dispersion des individus, à l'expansion subséquente de l'aire de répartition globale de l'espèce, puis à la fragmentation de l'aire. Distinctement, les différences dans les moments de floraison, de même que d'autres facteurs génétiques internes présents chez les *Picea*, constitueraient plutôt des conséquences de l'isolement géographique (Wright 1955). Suivant l'isolement géographique sur de grandes distances, les différences de climat, répondant aux variations de latitude et aux changements topographiques, représenteraient un facteur d'isolement secondaire pour les *Picea* (Wright 1955), et pourraient, par exemple, modeler les moments de floraison.

Les espèces de *Picea* les plus nordiques possèdent les aires de distribution les plus vastes. On y retrouve les occurrences d'hybridation naturelle connues chez les *Picea*. Les *Picea* étant des conifères pionniers, leur colonisation des espaces libres après le retrait des glaciers aurait été rapide au nord de l'hémisphère boréal (Liu 1988; Shi *et al.* 1998 [*in* Liu *et al.* 2002]). Plus au sud, leurs besoins environnementaux spécifiques les ont possiblement confinés graduellement dans des habitats de montagne aujourd'hui disjoints (Wright 1955; e.g. Jaramillo-Correa *et al.* 2006). On pourrait donc s'attendre à ce que les espèces

méridionales démontrent une spéciation allopatrique plus avancée que les espèces nordiques car leur isolement géographique aurait été maintenu depuis plus longtemps.

En ce qui concerne les espèces sympatriques, le grand nombre d'aires de répartition qui se chevauchent complexifie l'établissement des liens de parenté entre les espèces, notamment chez les *Picea* asiatiques. La sympatrie des espèces de *Picea* morphologiquement distinctes et reproductivement isolées leur prête naturellement des origines différentes (Wright 1955). À titre d'exemple, ce serait le cas de plusieurs espèces japonaises aujourd'hui sympatriques ou parapatriques. Chacune de celles-ci sur les îles japonaises montre un lien étroit avec une espèce du reste du continent asiatique (Wright 1955). Les espèces qui sont aujourd'hui exclusivement japonaises auraient achevé leur spéciation dans leurs aires allopatriques respectives et vraisemblablement à l'intérieur du continent; elles auraient par la suite étendu leurs aires de répartition vers les îles du Japon. À cet endroit, elles auraient été disjointes assez longtemps de leur parente non insulaire pour que la spéciation crée une divergence entre ces espèces-soeurs. Par conséquent, chacune des six espèces japonaises impliquerait un mouvement de colonisation vers les régions des îles japonaises (Wright 1955).

1.2.4.3.2 - Liens entre les *Picea* eurasiatiques et hypothèses de spéciation géographique

En Europe, seulement deux espèces distinctes de *Picea* sont présentes: *P. abies* et *P. omorika*. Les quelques sites reliques hébergeant ce dernier près de la rivière Drina dans les Balkans sont entièrement inclus dans la vaste aire de répartition de *P. abies*. Ces deux espèces, aujourd'hui sympatriques et morphologiquement distinctes (Wright 1955), auraient donc deux origines biogéographiques distinctes. L'ancêtre commun aux lignées distinctes ayant mené à *P. abies* et à *P. omorika* aurait subi la spéciation allopatrique suite à l'isolement géographique, et les aires de répartition des lignées distinctes produites par cette spéciation se seraient éventuellement rejointes. Il est à noter que l'ancêtre commun aux deux lignées est possiblement très ancien, et qu'avant que les lignées « *abies* » et « *omorika* » ne produisent les espèces éponymes, plusieurs événements de divergence et de spéciation peuvent avoir eu lieu ailleurs sur divers

continents. Ainsi, la désignation d'« ancêtre lointain commun... » sera utilisée dorénavant pour identifier un ancêtre qui est commun aux sujets, mais qui semble plus lointain dans le temps.

Vers l'est, la distribution de *P. abies* se fond avec celle de *P. obovata* dans la région de l'Oural, sans qu'il n'y ait de frontière claire entre les deux espèces. Ces espèces couvrent la majorité des territoires nordiques de l'Eurasie et leur zone de parapatrie constitue une grande région d'hybridation naturelle et d'introgression (Vidacovic 1991). Certains auteurs considèrent *P. obovata* comme une sous-espèce de *P. abies* (Vidacovic 1991) parce que leur morphologie est semblable et qu'ils s'hybrident. Cependant, une étude récente indique que leurs hybrides possèdent des caractéristiques distinctes de celles des espèces parentales et qui sont particulières aux zones d'hybridation interspécifique (Krutovskii et Bergmann 1995). Ces deux espèces ne semblent pas proches parentes non plus sur la base de leur patron de distribution, ni sur la base des essais de croisements (Wright 1955). *P. abies* et *P. obovata* seraient donc des espèces allopatriques qui auraient achevé leur spéciation sur des territoires distincts, qui se seraient éventuellement rejointes, et qui possèdent aujourd'hui des populations sympatriques dans les Grandes Plaines de Russie (Pravdin 1975 [*in* Vidacovic 1991]). Les hybrides observés pourraient vraisemblablement être issus de leur contact secondaire.

Au sud-ouest de sa distribution, *P. obovata* ressemble morphologiquement à *P. orientalis* (Wright 1955). Ce dernier est endémique des montagnes du Caucase, dans le sud-ouest de la Russie, ainsi que des montagnes turques, et il est allopatrique avec tous les *Picea*. À partir de cette région vers l'est, les premières espèces de *Picea* asiatiques rencontrées sont allopatriques et distribuées en périphérie de la Chine, sur ses frontières nord, ouest et sud-ouest. Il s'agit de *P. schrenkiana*, *P. smithiana* et *P. spinulosa*, les deux derniers semblant très proches parents (Wright 1955). Pour sa part, *P. schrenkiana* semble plus près morphologiquement et géographiquement de *P. spinulosa*, puis de *P. obovata*.

Le reste des espèces asiatiques vivent dans l'est de l'Asie le long d'un transect nord-est, de l'Inde orientale jusqu'en Corée, et de là, vers le nord sur la côte asiatique ainsi que sur les îles japonaises, à part *P. morrisonicola* qui habite l'île de Taiwan plus au sud, dans la continuité des îles japonaises. Celui-ci est allopatrique et plus semblable morphologiquement à certaines espèces du sud-est de la Chine, telles que *P. brachytyla*, *P. purpurea* et *P. wilsonii* (Wright 1955). De plus, le fait que ces trois derniers soient parapatriques ou sympatriques les uns avec les autres et reproductivement isolés leur conférerait une origine relativement ancienne (Wright 1955). Les trois espèces paraissent peu ou pas affectées par l'introgression, mais elles formeraient un groupe d'espèces phylogénétiquement cohérent lié à *P. morrisonicola* (Wright 1955).

Les liens entre la douzaine d'espèces asiatiques restantes sont complexes parce que plusieurs sont sympatriques ou parapatriques avec plus d'une autre tout en demeurant reproductivement isolées, sauf aux extrêmes du transect continental décrit ci-haut. *P. farreri*, le plus méridional du transect continental, est une espèce allopatrique relique du sud de la Chine. Découvert seulement en 1980 en Birmanie, *P. farreri* n'apparaît donc pas dans le travail de Wright (1955), mais Farjon (1990) le rapproche de *P. likiangensis*. Ce dernier est sympatrique avec *P. montigena*, mais plusieurs auteurs donnent à *P. montigena* le statut de sous-espèce ou de variété de *P. likiangensis*. Ces deux taxons sont parapatriques avec *P. asperata*, sans croisement apparent, et ils auraient donc une origine différente. Parmi les espèces étroitement liées à ce dernier, on note *P. meyeri* et *P. retroflexa* (Wright 1955; Vidacovic 1991). Wright (1955) faisait observer que *P. asperata*, *P. montigena* et *P. retroflexa* sont des espèces morphologiquement très semblables, et que les deux dernières sont aussi morphologiquement semblables à *P. koyamai*, une espèce des îles japonaises. Parmi les taxons est-asiatiques avec lesquels *P. asperata* aurait un ancêtre commun plus lointain, il y a ses congénères sympatriques, *P. purpurea* et *P. wilsonii*. Plus au nord-ouest de ce transect, bouclant la boucle des espèces situées autour du plateau tibétain et des montagnes lui étant adjacentes du côté nord-est, *P. crassifolia* est considéré comme étant très près de *P. schrenkiana* et possiblement une variété de celui-ci (Farjon 1990). Enfin, à l'extrême nord-est du transect continental, l'espèce côtière *P. koraiensis* en Corée a été confondue par certains auteurs avec l'espèce nordique *P. obovata*, lequel étend

son aire vers le sud jusqu'à rejoindre celle de *P. koraiensis*. *P. koraiensis* a été confondu aussi avec *P. koyamai* qui croît de l'autre côté de la Mer du Japon sur l'île de Honshu (Wright 1955). Par ailleurs, *P. koraiensis* est sympatrique avec *P. jezoensis*, quoique possiblement séparé écologiquement de celui-ci (Fenzel 1933 [*in* Wright 1955]), et les espèces lui étant le plus près géographiquement sont *P. asperata* et *P. wilsonii*. Dans ces régions, les aires de *P. koraiensis*, *P. asperata* et *P. wilsonii* auraient été jointes durant le pléistocène, puisque les montagnes étaient plus basses d'environ un kilomètre et, par conséquent, la faible distance qui sépare ces espèces aujourd'hui aurait possiblement été inexistante (Ma 1947 [*in* Wright 1955]).

P. koyamai fait partie des espèces japonaises sympatriques sur l'île Honshu, dont *P. alcoquiana*, *P. jezoensis*, *P. maximowiczii* et *P. torano*. Wright (1955), qui incluait *P. koraiensis* dans *P. koyamai*, suggérait que ce taxon serait le plus primitif du genre *Picea*. Il mentionnait aussi que les espèces japonaises, incluant *P. glehnii* sur l'île japonaise Hokkaido, tireraient leur origine de l'intérieur du continent et qu'il y aurait eu six événements de migration distincts vers les îles. En effet, tel que mentionné précédemment, la sympatrie des six taxons de Honshu reproductivement isolés, de même que la présence de *P. glehnii* dans le paysage insulaire japonais en sympatrie avec *P. jezoensis*, amènent à penser que l'origine de ces espèces se situerait plutôt dans les régions non insulaires, avec de multiples occurences de spéciation géographique vers les îles (Wright 1955). Par ailleurs, cet auteur avançait que l'origine commune du groupe d'espèces *P. koyamai* (et *P. koraiensis*), *P. meyeri*, *P. retroflexa*, *P. asperata*, *P. likiangensis* et l'européen *P. abies* serait relativement récente, ce qui apparaît en quelque sorte paradoxal avec la situation de *P. koyamai* comme taxon le plus ancestral du genre.

L'espèce possédant la plus vaste distribution sur la côte et les îles est-asiatiques est *P. jezoensis*, possiblement proche parente de *P. brachytyla* (Wright 1955). Parapatrique avec *P. obovata* le long de la mer d'Okhotsk, sympatrique avec *P. koraiensis* (ou *P. koyamai*) en Corée, *P. glehnii* sur Hokkaido et avec toutes les autres espèces de *Picea* sur Honshu, *P. jezoensis* semble distinct de tous car aucun croisement interspécifique naturel ou artificiel

avec eux ne semble possible (Wright 1955). Il constituerait, selon Wright (1955), le point de jonction entre des espèces de l'est de l'Asie et du nord-ouest de l'Amérique.

Il y aurait donc eu différents mouvements des groupes d'espèces de *Picea* à l'intérieur même du continent asiatique. Les changements géomorphologiques, tels que l'élévation des plateaux asiatiques, ont apporté des remaniements des aires de répartition des *Picea* à travers l'Asie entière, divisant les aires continues et provoquant des occasions de spéciation allopatrique, puis superposant en sympatrie les aires des espèces s'étant différenciées préalablement sur leurs sites allopatriques. Ceci complique l'interprétation des liens entre les espèces car l'image actuelle présente une solution par la négative, en mettant en évidence les distinctions entre les espèces plutôt que les liens entre les espèces voisines. De là, par exemple, on ne peut instinctivement déduire quelle a été la première espèce de *Picea* ayant colonisé les îles japonaises ni son lien direct avec l'une des espèces de l'intérieur du continent.

1.2.4.3.3 - Liens entre les *Picea* nord-américains et hypothèses de spéciation géographique

De vastes régions de sympatrie entre des espèces de *Picea* sont également observables en Amérique du Nord. Dix taxons de *Picea* s'y partagent l'ensemble du territoire: *P. breweriana*, *P. chihuahuana*, *P. martinezii*, *P. mexicana*, *P. pungens*, *P. glauca*, *P. engelmannii*, *P. sitchensis*, *P. mariana* et *P. rubens*. Les cinq premiers taxons possèdent des aires allopatriques, tandis que les autres partagent leur aire de distribution en totalité ou en sections de tailles variées.

La sympatrie de *P. glauca* avec *P. mariana* sur toute leur aire de répartition transcontinentale, couplée au fait que ces deux taxons morphologiquement distincts demeurent reproductivement isolés, démontre que ces taxons constituent des espèces distinctes d'origines biogéographiques différentes. Il en est de même pour *P. glauca* et *P. rubens*, une espèce dérivée de *P. mariana* (Perron *et al.* 2000; Jaramillo-Correa *et al.* 2003). Le lien phylogénétique intime entre les espèces parapatriques morphologiquement semblables (Wright 1955) *P. mariana* et *P. rubens* n'est plus à discuter (Perron et Bousquet

1997; Jaramillo-Correa et Bousquet 2003). Quant au cas *glauca/mariana*, il indique qu'il y aurait au moins deux origines distinctes pour les espèces nord-américaines, et par conséquent, au moins deux groupes d'espèces en Amérique du Nord au niveau phylogénétique. L'ancêtre lointain commun aux grandes lignées distinctes ayant mené à *P. glauca* et *P. mariana* croissait-il sur ce même continent? Cet ancêtre lointain était-il plutôt sur un continent différent? Cette dernière hypothèse impliquerait au moins deux mouvements de colonisation différents vers l'Amérique si l'origine des *Picea* nord-américains contemporains se situe sur un autre continent que l'Amérique, et elle impliquerait des mouvements de sortie et de retour en Amérique si leur origine est américaine.

À l'ouest de l'Amérique du Nord, les espèces *P. glauca*, *P. engelmannii* et *P. sitchensis*, des espèces morphologiquement semblables, et possiblement *P. pungens* plus au sud, présentent des aires parapatriques ou sympatriques qui permettent un certain niveau d'introgression ou d'hybridation naturelle tel que mentionné plus haut. Selon Wright (1955), *P. glauca*, *P. engelmannii*, *P. sitchensis* et *P. pungens* auraient une origine commune assez récente liée à l'espèce est-asiatique *P. jezoensis*. Cet auteur avance aussi que les taxons du nord-est de l'Amérique du Nord, soit *P. mariana* et *P. rubens*, proviendraient d'une migration plus ancienne que les taxons du nord-ouest de l'Amérique du Nord du groupe *P. glauca*. Il mentionne également que plusieurs mouvements de colonisation auraient eu lieu à partir de l'Asie vers l'Amérique. Pour leur part, *P. breweriana* et *P. chihuahuana*, taxons reliques, seraient issus d'un autre mouvement de colonisation. Ainsi, toujours suivant Wright (1955), il y aurait eu trois dispersions distinctes de l'Asie et/ou l'Europe jusqu'en Amérique du Nord. Les résultats de l'étude de Sigurgeirsson et Szmidt (1993), basée sur le génome chloroplastique, vont dans le même sens que les constats de Wright (1955) pour ce qui est du nombre de mouvements de colonisation intercontinentaux. Ils montrent des phylogénies dans lesquelles les taxons américains sont entremêlés avec les taxons asiatiques, d'où on peut déduire un rapprochement au niveau génétique entre les *Picea* de différents continents. Cependant, contrairement à Wright (1955), cette étude infère une origine américaine pour toutes les espèces contemporaines du genre *Picea*, avec *P. breweriana* comme taxon ancestral. À

tout le moins en Amérique, *P. breweriana* semble être effectivement l'espèce la plus divergente sur la base de l'anatomie des aiguilles et d'autres études moléculaires (Ledig *et al.* 2004).

1.2.5 - Résumé de la problématique

Le problème de la classification et de la délimitation des espèces de *Picea* a été soulevé il y a plus de 60 ans par Bobrov (1944) (*in* Vidacovic 1991) suivant la complexité des liens morphologiques entre les *Picea*. En effet, tous les caractères morphologiques des *Picea*, tels que mentionnés par Wright (1955), sont partagés entre les continents. Plus tard, les hybrides artificiels facilement créés entre des espèces de *Picea* taxonomiquement éloignées ont ajouté à la confusion.

L'origine du genre *Picea* n'est toujours pas claire non plus: en Asie ou en Amérique? Pour chacune de ces hypothèses, certains ont déjà tenté d'expliquer la répartition actuelle des *Picea* à travers le monde, ainsi que leur histoire biogéographique et géomorphologique, à partir de l'âge minimal de la diversification du genre offert par les données fossiles. Aussi, comme le montre le Tableau 1.1, les grands groupes d'espèces de *Picea* établis jusqu'ici supposent tous au moins deux mouvements de colonisation d'un continent à l'autre.

Donc, avec ses caractéristiques morphologiques enchevêtrées, son manque de définition quant à l'isolement reproducteur des espèces et sa distribution sur plusieurs continents de l'hémisphère nord, cet ancien groupe de plantes possède une histoire évolutive qui gagnerait à être élucidée à l'aide d'outils plus discriminants pouvant fournir des caractères discrets et plus objectifs. C'est précisément l'orientation que nous avons choisi de poursuivre pour la présente thèse en voulant vérifier l'utilité des caractères moléculaires.

1.3 - La phylogénie moléculaire des *Picea* par séquençage de l'ADN

1.3.1 - L'avenue moléculaire: généralités

Les polymorphismes détectables au niveau de l'ADN se sont révélés grandement utiles pour les chercheurs désirant découvrir les liens phylogénétiques entre des espèces végétales à différents niveaux taxonomiques (Chase *et al.* 1993). À cet effet, le séquençage des nucléotides d'une région donnée de l'ADN s'avère être le moyen informatif le plus direct. En comparaison, les marqueurs RFLPs (*restriction fragment length polymorphisms*) et les marqueurs RAPDs (*randomly amplified polymorphic DNA*) impliquent une démarche expérimentale plus indirecte pour la détermination des polymorphismes d'ADN. De plus, l'homologie des fragments d'ADN est difficilement vérifiable avec ces méthodes, surtout dans le cadre d'une étude phylogénétique impliquant des espèces dont le niveau de divergence est peu documenté. Maintenant plus accessible, le séquençage d'ADN offre une plus grande précision d'information que ces méthodes puisque les renseignements recherchés sont acquis au niveau le plus fin de résolution, soit les nucléotides. Il s'agit là d'une qualité d'information et d'une précision qu'il est nécessaire d'obtenir dans le cadre de la présente étude.

Les séquences nucléotidiques homologues entre les organismes d'intérêt peuvent être déterminées directement après l'exécution des PCR (*polymerase chain reaction*, Mullis et Faloona [1987]), une technique de routine pour les études moléculaires d'aujourd'hui. Pour les études portant sur l'évolution ou la phylogénie des plantes, cette méthode est toujours très employée car elle est rapide et efficace pour amplifier différentes régions du génome choisies suivant le niveau taxonomique à traiter (*e.g.* Bousquet *et al.* 1990, White *et al.* 1990, Laroche *et al.* 1997; Wang *et al.* 1999, Wang *et al.* 2000).

Pour chaque étude, la région génomique pertinente est choisie suivant les objectifs spécifiques. Dans cette optique, il devient indispensable de connaître les particularités des génomes en termes de pertinence, de qualité et de quantité d'information qu'il est adéquat

d'obtenir. En effet, les trois génomes portés par chaque cellule végétale possèdent des taux évolutifs différents (Sytsma et Schaal 1985; Wolfe *et al.* 1987; Mitton 1994): en général, l'ADN mitochondrial évolue le plus lentement, suivi de l'ADN chloroplastique, puis de l'ADN nucléaire (Wolfe *et al.* 1987; Mitton 1994). Dans chaque cas, généralement, les gènes possèdent un mode d'évolution conservateur et évoluent relativement lentement, tandis que les régions non codantes ou intergéniques offrent davantage de sites informatifs par effort de séquençage, car elles évoluent plus rapidement. Aussi, la combinaison de diverses régions de l'ADN s'est déjà montrée valable, puisqu'elle améliore la résolution des phylogénies en apportant davantage de sites variables à comparer (*e.g.* Wang *et al.* 2003).

1.3.2 - Les études de phylogénie moléculaire chez les conifères

En ce qui a trait aux conifères, des études de phylogénie moléculaire par séquençage ont réussi à éclairer davantage les liens entre des familles, des genres et des espèces. À titre d'exemple au niveau interfamilial, Cheng *et al.* (2000) ont réalisé la phylogénie des genres de Taxaceae et Cephalotaxaceae en utilisant le gène *mat*K, une région de l'ADN chloroplastique, en combinaison avec la région intergénique transcrite de l'ADN ribosomique nucléaire *ITS*. Aussi, les régions chloroplastiques *rbc*L et *mat*K ont déjà été utilisées pour établir une phylogénie comprenant sept familles de Coniférales (Quinn *et al.* 2002), et une autre, les sept sous-familles des Cupressaceae *sensu lato* (Gadek *et al.* 2000). Au niveau intergénérique, la phylogénie des Podocarpaceae, la famille la plus diverse des conifères, a été réalisée à l'aide de la région relativement conservée *18S* de l'ADN ribosomique nucléaire (Kelch 1998). Par ailleurs, plusieurs études avaient pour objectif d'estimer la phylogénie de divers genres de la famille des Pinaceae.

La famille des Pinaceae comprend onze genres de conifères, dont le genre *Pinus* qui inclut le plus grand nombre d'espèces, soit un peu plus d'une centaine, et qui constitue ainsi près de la moitié de la famille. Entre autres à cause de son importance en nombre d'espèces, ce genre a davantage fait l'objet d'études moléculaires, telles certaines recourant au séquençage de la région nucléaire *ITS* pour près de la moitié des espèces de pins (Liston *et al.* 1999; Liston *et al.* 2003), une autre incluant cette région nucléaire en combinaison

avec quatre autres pour les sous-sections du genre *Pinus* (Syring *et al.* 2005), une troisième pour les pins européens avec quatre gènes du génome chloroplastique dont les gènes *mat*K et *rbc*L (Wang *et al.* 1999), et une autre incluant pratiquement tous les *Pinus* a été réalisée avec différentes régions du génome chloroplastique (Gernandt *et al.* 2005). Ces études ont utilisé des amorces spécifiques qui, dans plusieurs cas, avaient été développées pour d'autres groupes de plantes, et qui peuvent être appliquées également à d'autres genres de Pinaceae, vu la conservation relative des gènes impliqués.

Grâce à l'approche moléculaire utilisée pour les Pinaceae jusqu'ici, les phylogénies des genres *Pinus*, *Larix*, *Pseudotsuga* et *Tsuga* ont pu être estimées, et leur histoire biogéographique a pu être précisée (*e.g.* études citées dans Liston *et al.* 2003, et Gros-Louis *et al.* 2005). En ce qui concerne les *Picea*, seulement deux essais de phylogénie, l'un par Sigurgeirsson et Szmidt (1993) et l'autre par Ran *et al.* (2006), ont été réalisés, la première phylogénie ayant été faite à l'aide de RFLPs du génome chloroplastique et la seconde à l'aide de séquences nucléotidiques de deux gènes chloroplastiques. Cependant, nous verrons plus loin pourquoi ces études basées sur le génome chloroplastique ne répondent pas aux objectifs de la détermination de l'histoire biogéographique du genre *Picea*. Enfin, pour un autre essai de phylogénie pour sept espèces de *Picea* par Smith et Klein (1994), l'utilisation d'amorces universelles pour les *ITS* nucléaires s'est soldée en un échec, puisque les séquences d'*ITS* obtenues dans cette étude étaient plutôt d'origine fongique due à la contamination des échantillons (Camacho *et al.* 1997). Puisque la phylogénie moléculaire visant à reconstruire l'histoire biogéographique des espèces du genre *Picea* n'avait jamais été réalisée jusqu'à ce jour, la présente étude a visé cet objectif en commençant par un test de faisabilité avec le génome nucléaire, dont les gènes, rappelons-le, sont généralement plus variables que ceux des autres génomes.

1.3.3 - Le choix des régions génomiques

1.3.3.1 - L'ADN nucléaire

1.3.3.1.1 - Généralités

Pour estimer la phylogénie d'un groupe de plantes, le génome nucléaire est habituellement attrayant à tous les niveaux taxonomiques inférieurs à la famille (e.g. Jeandroz *et al.* 1997), puisqu'une variété de régions montrant des taux évolutifs différents (Schaal et Learn 1988; Baldwin 1993) le composent et sont assez variables pour permettre même certaines études de populations. De nombreuses études comparatives rapportent le succès de l'utilisation de ce génome en systématique des plantes à ces différents niveaux taxonomiques (voir Alvarez et Wendel 2003).

L'ADN ribosomique nucléaire correspond à la séquence génique nucléaire qui dirige la synthèse des sous-unités de l'ARN ribosomique (Sytsma et Schaal 1990). Cette région de l'ADN est moyennement à fortement répétitive et elle peut contenir de quelques centaines à plusieurs milliers de copies de séquences disposées en tandem chez la plupart des eucaryotes (Rodgers et Bendich 1987; Schaal *et al.* 1991; Won et Renner 2005). Ces répétitions sont composées d'une région codante, dont les trois sous-unités *18S* (ou *17S*), *5.8S* et *25S* (ou *26S*) sont séparées par des régions intergéniques transcrites, soit les *ITS-1* et les *ITS-2*. De chaque côté de la grande région *ITS* se situent d'autres régions intergéniques très longues, aussi appelées régions non transcrites *IGS* ou *NTS*, qui sont des régions très hétérogènes et très variables pouvant être longues de plusieurs milliers de kilobases (Sytsma et Schaal 1990; Kim et Mabry 1991). Les régions *ITS* sont souvent préférées aux *IGS* à cause leur longueur relativement courte, qui permet plus facilement l'ancrage des amorces universelles à l'intérieur des régions plus conservées et qui facilite l'amplification subséquente.

Les régions *ITS* de l'ADN ribosomique nucléaire se prêtent habituellement très bien aux phylogénies interspécifiques, une qualité qui en fait l'une des régions les plus

fréquemment séquencées (Won et Renner 2005). L'homogénéisation des multiples copies de ces régions est rapide dans le génome (Baldwin *et al.* 1995). Ce phénomène élimine ainsi les nouvelles mutations ponctuelles quand il y a maintien du flux génique entre les individus d'une même espèce. Il est donc extrêmement rare d'observer de la variation intraspécifique au sein de cette région (*e.g.* Jeandroz *et al.* 1997), d'où l'utilité de ces régions pour la phylogénie d'espèces. Chez les conifères, les régions *ITS* se sont révélées intéressantes au niveau intergénérique (*e.g.* Cheng *et al.* 2000) et interspécifique (*e.g.* Liston *et al.* 1999), de même que pour des études de populations (*e.g.* Quijada *et al.* 2004) et d'hybridation (*e.g.* Quijada *et al.* 1997).

1.3.3.1.2 - L'utilisation des régions nucléaires pour les *Picea*

De ce point de départ théorique, notre hypothèse de travail à l'origine des travaux de cette thèse était que les *ITS* devaient démontrer suffisamment de variation chez le genre *Picea* pour pouvoir en estimer la phylogénie et possiblement en retracer l'histoire biogéographique. De plus, ces régions sont assez courtes pour être facilement séquencées chez les angiospermes (e.g. Jeandroz *et al.* 1997). Les régions *ITS* sont aussi réputées être plus variables que les gènes *rbc*L (chloroplastique) ou *18S* (nucléaire) déjà disponibles pour certaines espèces de *Picea* (Savard *et al.* 1993). Des essais préliminaires d'amplification ont été menés pour les *Picea* avec des amorces universelles, avec succès pour la région *ITS-2*, mais sans succès pour la région *ITS-1* (M. Bouillé, résultats non publiés).

D'autres études qui avaient utilisé le clonage d'ADN ribosomique (*e.g.* Maggini *et al.* 1998) ou les fragments de restriction et l'hybridation Southern (*e.g.* Karvonen *et al.* 1994) ont démontré que certaines parties de la région *ITS-1* se répètent de façon contiguë avec des variations mineures chez les *Picea*, ce qui en augmente la longueur de façon significative et rend l'amplification et le séquençage subséquent difficiles. C'est ce qui expliquerait l'insuccès des essais préliminaires d'amplification de l'*ITS-1* que nous avons menés chez *Picea*. En effet, la recherche d'amorces internes pour séquencer cette région d'environ 3 000 paires de bases a été réalisée pour le présent travail, ainsi que des essais de

clonage de brins, mais sans succès, ce qui a confirmé l'impasse pour le séquençage rapide de la région *ITS-1* (M. Bouillé, résultats non publiés). Une étude fondamentale récente basée sur le clonage d'ADN ribosomique (Campbell *et al.* 2005) indique que cette région est d'une architecture très complexe chez les *Picea*. Aussi, elle suit un patron de duplication particulier qui complique singulièrement l'amplification et le séquençage pour l'estimation de la phylogénie du genre *Picea*.

Pour ce qui est de la région *ITS-2*, les amorces utilisées lors des tests préliminaires pour l'amplification et le séquençage étaient universelles pour les plantes et les champignons (White *et al.* 1990). Il a donc fallu valider que les séquences obtenues correspondaient bien à la région *ITS-2*. Ce test a été réalisé en comparant les séquences résultantes à celles déposées dans la banque mondiale de séquences GenBank: il a été démontré que ces séquences correspondaient bien à d'autres fragments *ITS-2* de la famille des Pinaceae. L'amplification des *ITS-2* a d'abord été effectuée pour plusieurs espèces de *Picea* divergentes pour vérifier l'importance de la variation interspécifique. Les résultats ont montré que cette courte région d'environ 300 paires de bases présentait malheureusement trop peu de variations, soit seulement quelques sites variables, pour réaliser une phylogénie détaillée du genre *Picea* en son entier (M. Bouillé, résultats non publiés).

Conséquemment, d'autres régions variables du génome nucléaire ont été considérées pour estimer la phylogénie du genre. Nous avons vérifié la variabilité de séquences pour plusieurs gènes nucléaires pour lesquels des amorces spécifiques étaient disponibles et avaient été élaborées à partir du séquençage d'ADNc de *Picea* (Perry et Bousquet 1998a, 1998b; Perry *et al.* 1999). Plus précisément, trois gènes nucléaires distincts se sont avérés suffisamment variables au niveau de leurs séquences. Cependant, les comparaisons des séquences obtenues pour trois taxons de *Picea* distincts taxonomiquement et génétiquement, ainsi que les analyses phylogénétiques subséquentes, ont démontré que les espèces n'étaient pas monophylétiques pour ces gènes, à cause de la persistance de polymorphismes trans-spécifiques ancestraux. En raison de l'originalité de cette

découverte, un manuscrit a été publié à ce sujet et fait partie intégrante du présent ouvrage (Bouillé et Bousquet 2005).

1.3.3.1.3 - La particularité des polymorphismes trans-spécifiques

En général, dans une population de plantes, la majorité des nouvelles mutations sont éliminées peu après leur émergence. Cependant, une petite minorité de mutations avantageuses ou neutres peut atteindre un certain équilibre et persister jusqu'à très longtemps dans la population (Klein *et al.* 1998). C'est cette persistance des mutations à grande échelle qui a été détectée dans le cadre de notre étude suite à l'analyse de trois loci nucléaires pour trois espèces divergentes du genre *Picea*, et qui ne pouvait être anticipée lors de l'amorce des travaux.

Les polymorphismes d'ADN sont dits « trans-spécifiques » lorsqu'ils sont partagés entre deux ou plusieurs espèces du même genre. En fait, outre qu'il puisse être produit par effet de convergence, une circonstance cependant plutôt rarissime, le polymorphisme trans-spécifique apparaît lors du passage d'un polymorphisme donné d'une espèce ancestrale à deux espèces dérivées. Le maintien de ce polymorphisme sur une longue période à l'intérieur des deux espèces dérivées constitue une séparation incomplète des lignées (voir Pamilo et Nei 1988, Rosenberg 2003, Maddison et Knowles 2006). Donc, un polymorphisme trans-spécifique date généralement d'avant la divergence des espèces. Le moyen de détection des polymorphismes trans-spécifiques est la coalescence, qui est le fait de retracer, dans une généalogie, les ancêtres communs aux lignées étudiées en remontant dans le passé. Dès lors, lorsque l'on compare les polymorphismes de différentes espèces, s'il existe des allèles différents à l'intérieur de l'une des espèces pour un locus donné, il est possible que la coalescence d'au moins l'un de ces allèles dépasse la divergence des espèces lorsque l'on remonte au-delà de leur origine. Ces polymorphismes ancestraux, partagés entre les espèces, seront donc au moins aussi âgés que la division des populations ayant mené aux espèces distinctes. De plus, vu leur âge encore plus avancé que la divergence des espèces, ils auront eu davantage le temps de recombiner avec d'autres polymorphismes à l'intérieur des espèces (Machado *et al.* 2002).

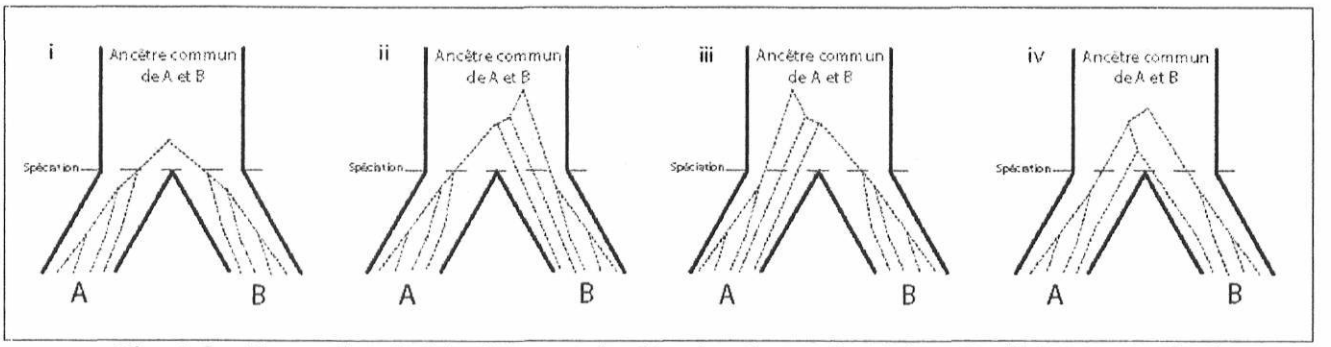

Fig. 1.2 - Types de généalogies pour les lignées des espèces A et B. (i) Monophylie de A et B; (ii) Paraphylie de B par rapport à A; (iii) Paraphylie de A par rapport à B; (iv) Polyphylie de A et B (adapté de Rosenberg [2003]).

À cause de l'existence de ce phénomène particulier, il faut distinguer les phylogénies de gènes (ou généalogies d'allèles) des phylogénies d'espèces. Pour les polymorphismes trans-spécifiques, les deux types de phylogénies seront la plupart du temps différentes, car les possibilités que les espèces n'apparaissent pas monophylétiques sont substantielles (Pamilo et Nei 1988). À l'opposé de la monophylie, qui est le regroupement de tous les allèles d'une espèce dans le même clade (Fig. 1.2 i), on distingue la paraphylie et la polyphylie. La paraphylie d'une espèce par rapport à une autre est l'absence de regroupement de tous ses allèles à l'intérieur d'un seul clade dans la phylogénie d'un gène donné, tandis que les allèles de l'autre espèce se retrouvent dans le même clade (Fig. 1.2 ii et iii). La polyphylie peut se définir comme étant la paraphylie réciproque des deux espèces, car les allèles des deux espèces se retrouvent dans les deux clades (Fig. 1.2 iv). Aussi, plus il y aura de polymorphismes trans-spécifiques entre plusieurs espèces dans une généalogie de gène, plus la polyphylie sera prononcée. Par conséquent, ni la période de divergence des espèces ni le plus récent ancêtre commun aux deux espèces ne pourraient être déterminés par l'étude de ces loci.

De là, on peut comprendre qu'une phylogénie d'espèces de *Picea* pourrait être erronée si des polymorphismes trans-spécifiques étaient présents pour les gènes nucléaires étudiés. La vérification de leur présence peut être réalisée en estimant l'importance et le type de variation intraspécifique pour le locus nucléaire concerné en comparant les séquences de plusieurs individus de chaque espèce. Les types de variations démontreront à

quel point les espèces apparaîtront de façon monophylétique, paraphylétique ou polyphylétique dans la phylogénie. Certains loci étudiés pourront en effet comporter uniquement des polymorphismes intraspécifiques fixés, ou, au mieux, aucun polymorphisme intraspécifique, et les regroupements des espèces apparaîtront réciproquement monophylétiques. Ce sont ces loci nucléaires neutres qu'il faut privilégier pour effectuer des phylogénies d'espèces à partir de peu de gènes nucléaires, mais leur recherche peut impliquer une démarche exhaustive, surtout dans le cas d'une phylogénie impliquant autant d'espèces que dans le genre *Picea*.

Le principe de la fixation des polymorphismes est au coeur même du mécanisme de l'évolution. Les changements qui s'accumulent et se fixent dans les génomes avec le temps participent à la divergence de ces derniers. Ultimement, pour les loci neutres ou presque neutres, en l'absence complet de flux génique interspécifique, comme dans le cas d'espèces distinctes reproductivement isolées, la dérive génétique à l'intérieur de chaque espèce provoque l'accumulation de polymorphismes fixés et la perte des polymorphismes partagés. Toutefois, ce phénomène étant graduel, un état transitionnel ou intermédiaire peut se traduire en différents niveaux de fixation/partage des polymorphismes. En théorie, au fil du temps, pour des loci nucléaires neutres ou presque neutres qui présentent des polymorphismes partagés, les populations en voie de spéciation paraîtront de moins en moins paraphylétiques dans une phylogénie. De fait, dès l'instant de divergence de deux espèces, les espèces montrent plus probablement de la polyphylie, puis, au fil du temps, de la paraphylie, et enfin, de la monophylie réciproque (Rosenberg 2003).

Néanmoins, certaines circonstances peuvent ralentir le processus de l'évolution en termes de spéciation. Outre certains cas provoqués par la sélection naturelle, l'une des causes de la conservation des polymorphismes à l'intérieur d'une espèce est la taille effective des populations. En effet, plus une population est grande, moins les effets de la dérive génétique se font sentir à l'intérieur de celle-ci, et les nouveaux polymorphismes sont plus facilement conservés dans toute la population. Dans le processus de spéciation, ce statisme relatif freine la divergence génique entre les populations d'une espèce qui

conservent un flux génique important, car la fixation des nouveaux polymorphismes, au détriment des polymorphismes plus anciens, est retardée. Cependant, le maintien de grands effectifs de population permet la conservation d'un plus grand nombre de polymorphismes dans l'espèce et, conséquemment, la diversité génétique est accrue à l'intérieur de l'espèce. D'ici, on rejoint donc l'idée des polymorphismes trans-spécifiques: pour deux espèces données, les polymorphismes ancestraux des gènes nucléaires neutres ou presque neutres seront conservés plus longtemps dans les populations de grande dimension comme celles des conifères. Les *Picea*, comme les autres Pinaceae, sont effectivement connus pour leurs grandes tailles de populations (Miksche 1985; Govindaraju 1988), et les faibles niveaux de différenciation de leurs populations ont déjà été évalués dans le passé (*e.g.* Boyle et Morgenstern 1987, Furnier *et al.* 1991, Isabel *et al.* 1995, Müller-Starck 1995, Jaramillo-Correa *et al.* 2001, Perry et Bousquet 2001, Collignon *et al.* 2002, Gamache *et al.* 2003). En ce qui concerne les estimations de tailles historiques de populations pour les *Picea*, le facteur relatif au temps moyen de génération a été établi à 50 ans dans cette thèse. Dans le cycle de vie du peuplement, de sa naissance jusqu'à sa mort, on considère qu'il y a un chevauchement des générations tout au moins partiel. Il s'agit donc d'un temps moyen de génération établi à l'intérieur de la durée de vie d'un peuplement. Cependant, ce temps de génération est plus élevé que l'atteinte de la maturité sexuelle des arbres. L'atteinte de la maturité sexuelle chez les conifères, telle que considérée à 25 ans par Brown *et al.* (2004) dans leurs calculs de tailles de populations de pins, correspondrait plutôt au début de la formation de la nouvelle population, et non au moment où la population est entièrement formée, avec son effectif maximal. Il apparaît donc nécessaire de considérer un délai additionnel, ce qui évite de surestimer la taille effective des populations.

La taille de population varie également selon le type de génome dont il est question. Tous paramètres neutres établis pour une population idéale, le caractère haploïde des organelles provoque la fixation d'une mutation beaucoup plus rapidement dans les organelles que dans le génome nucléaire (Birky *et al.* 1983; Chesser et Baker 1996), en empêchant toutefois la perpétuation de mutations délétères. Ce taux plus rapide est dû au fait que la taille effective de la population des gènes d'organelles est la moitié (N) de celle des gènes nucléaires (2N), chez les espèces diploïdes comme les *Picea*. En effet, à cause de

leurs plus petites tailles de populations, les gènes d'organelles sont plus sensibles à la dérive génétique que les gènes nucléaires (Gugerli *et al.* 2001). Aussi, les taux d'extinction des allèles sont plus élevés pour les gènes d'organelles que pour les gènes nucléaires, et les taux de séparation de lignées, soit la perte de polymorphismes trans-spécifiques pour une descendance, y sont beaucoup plus rapides (Zhang et Hewitt 2003). La présence de polymorphismes trans-spécifiques semble effectivement plus répandue pour les gènes nucléaires que pour les gènes d'organelles (Linder et Rieseberg 2004). C'est aussi ce qu'ont révélé les études de la présente thèse (chapitres 2 et 3).

L'utilisation des génomes d'organelles est donc toute désignée pour l'estimation d'une phylogénie d'espèces de *Picea*, car elle limite les problèmes de la coalescence des allèles précédant celle des espèces grâce à la fixation plus rapide des mutations dans ces génomes que dans les génomes nucléaires. Effectivement, comme il en sera question au troisième chapitre de cette thèse, aucun cas de polyphylétisme ou de paraphylétisme n'a été remarqué lors de l'analyse des séquences de gènes chloroplastiques ou mitochondriaux chez les espèces pour lesquelles plus d'un individu avait été séquencé.

1.3.3.2 - Les génomes d'organelles

1.3.3.2.1 - Transmission différentielle des génomes d'organelles: généralités

Lorsqu'il s'agit d'une étude biogéographique, le choix d'une région donnée de l'ADN ne doit pas être laissé au hasard. Les buts précis de l'étude justifieront leur choix parmi les génomes nucléaire, mitochondrial ou chloroplastique, et, naturellement, l'interprétation des résultats devra suivre en conséquence.

Il existe trois types de transmission des génomes végétaux: la transmission maternelle, paternelle ou biparentale. Selon qu'il s'agisse d'un genre de plante ou d'un autre, les deux organelles, soit les chloroplastes (plastides) et les mitochondries, ne sont pas toujours transmises par le même parent lors de la fécondation. De plus, chez les

angiospermes, plus du tiers des genres étudiés héritent parfois leurs plastides des deux parents (Mogensen 1996). En général, les plastides sont transmis maternellement, mais quelques plantes font exception à la règle, comme la majorité des conifères, qui affichent plutôt une transmission paternelle des chloroplastes (Neale et Sederoff 1989). Les études effectuées jusqu'à maintenant tendent à démontrer que les plantes à fleurs transmettent leurs mitochondries généralement de façon maternelle. Cependant, ce n'est pas le cas pour toutes les familles de conifères non plus (Mogensen 1996). Chez les Pinaceae, la plus grande famille de conifères qui inclut les *Picea*, la transmission est maternelle pour la mitochondrie (Sutton *et al.* 1991; Jaramillo-Correa *et al.* 2003) et paternelle pour le chloroplaste (Szmidt *et al.* 1988; Neale et Sederoff 1989). Cette différence de transmission s'explique par la mécanique de fécondation particulière pour ce genre de plantes. En effet, lors de la fécondation entre le pollen et la graine, les plastides maternels, c'est-à-dire ceux de la graine, se font modifier et demeurent dans le cytoplasme à l'écart du noyau zygotique. À ce moment, ce dernier est étroitement entouré de mitochondries maternelles qui seront incorporées dans le néocytoplasme avec les plastides paternels non modifiés (Mogensen 1996). De là, on peut comprendre que le pollen, gamète mâle, transmettra le génome chloroplastique paternel, et que la graine, partie reproductrice femelle des *Picea*, transmettra son génome mitochondrial à la descendance.

1.3.3.2.2 - Les génomes d'organelles pour la phylogénie des *Picea*

Cette situation particulière de la transmission distincte des organelles offre la possibilité d'étudier l'évolution des *Picea* sous différentes facettes, parce qu'elle permet de comparer les phylogénies obtenues de chaque génome d'organelle. Ainsi, les phylogénies tirées des informations chloroplastiques et mitochondriales pourront différer si le pollen et la graine ont des histoires géographiques divergentes, car le pollen peut être disséminé sur des distances beaucoup plus grandes que la graine chez des espèces telles les *Picea* (Comtois 1997). Pour les graines, les voies de spéciation sont plus faciles à suivre parce que les arbres répandent leurs graines sur des distances plus limitées, et peuvent former graduellement de nouvelles populations en périphérie de leur aire de répartition. Quoique les événements de dispersion des graines sur de longues distances

existent également, ils sont moins probables que les événements de dissémination à longue distance du pollen (Petit *et al.* 1997). Après l'installation de barrières géographiques, écologiques, climatiques ou autres qui les isoleront reproductivement, deux populations données viendront à se différencier car, ainsi isolées, elles auront accumulé des mutations qui leur seront propres. Par conséquent, dans l'optique d'une étude biogéographique du genre *Picea*, une phylogénie basée sur l'ADN mitochondrial établira entre les taxons des liens de parenté qui pourront mieux démontrer les voies de colonisation que les ancêtres des espèces contemporaines de *Picea* ont empruntées, découvrant du même coup les événements de spéciation liés à la dispersion géographique à l'intérieur du genre. En ce qui concerne le pollen des *Picea*, tout comme celui de la plupart des espèces anémophiles, il peut atterrir des centaines de kilomètres plus loin (Comtois 1997) et féconder des graines dans des populations qui sont géographiquement isolées et qui peuvent appartenir à un taxon distinct ou en voie de divergence au niveau du génome mitochondrial. Même si le flux génique mitochondrial a été interrompu depuis longtemps entre ces populations, le flux génique chloroplastique à longue distance peut donc s'effectuer pendant particulièrement longtemps suite à la séparation physique des populations, spécialement pour des espèces aux barrières reproductives lentes à s'établir comme chez les *Picea* (voir Wright 1955). Parallèlement, ce flux génique chloroplastique par pollen peut s'établir de nouveau suite à un contact secondaire entre les espèces qui ont précédemment divergé. Ce processus évolutif se nomme le transfert latéral du génome chloroplastique. Il implique qu'une espèce ait capturé le génome chloroplastique d'une espèce distincte, et ce n'est donc pas la transmission habituelle du génome d'un parent à la descendance au sein d'une même espèce. Le résultat anticipé sera tel que la phylogénie chloroplastique d'un genre de la famille des Pinaceae sera significativement moins structurée géographiquement que sa phylogénie mitochondriale, et elle apparaîtra plus réticulée.

Ainsi, sans prétendre que les phylogénies des *Picea* de Sigurgeirsson et Szmidt (1993) et de Ran *et al.* (2006) basées sur le génome chloroplastique soient erronées, on peut toutefois émettre l'hypothèse qu'elles peuvent représenter un portrait de l'évolution des *Picea* différent de celui de l'histoire biogéographique du genre, surtout dans le cas d'un tel genre dont l'hybridation ou l'introgression est possible entre de nombreuses espèces.

D'ailleurs, Sigurgeirsson et Szmidt (1993) préconisaient l'évaluation du transfert latéral du génome chloroplastique lors des études futures de phylogénie des *Picea*. En effet, parce que le transfert des gènes par le pollen peut brouiller ou homogénéiser les événements de différenciation (Ennos 1994), l'hybridation ou l'introgression, surtout ancienne, tend à donner l'image d'une évolution réticulée (Rieseberg et Soltis 1991). Cette particularité a été notée chez le genre conifèrien *Chamaecyparis* (Cupressaceae), qui a fait l'objet d'une étude biogéographique dans laquelle le transfert de l'ADN chloroplastique d'une espèce vers un hybride a été mis au jour (Wang *et al.* 2003). Parallèlement, des phylogénies interspécifiques différentes ont été obtenues pour les génomes nucléaire et chloroplastique (Semerikov *et al.* 2003) ou pour les génomes mitochondrial et chloroplastique (Gros-Louis *et al.* 2005) chez le genre *Larix*, de la famille des Pinaceae. L'interprétation la plus plausible de ces divergences était l'hybridation ancienne, suite à laquelle le génome d'une espèce aurait été transféré à une autre espèce proche parente.

Pour conclure sur ce thème, dans le cas des *Picea*, on peut résumer qu'une phylogénie mitochondriale pourrait mieux répondre aux besoins de la détermination de l'histoire biogéographique du genre et des phénomènes de la spéciation liée à la dispersion géographique; une phylogénie chloroplastique serait plutôt indicative du flux génique chloroplastique et des occurrences de transfert latéral du génome chloroplastique, dont les événements anciens d'hybridation ou d'introgression seraient responsables. Enfin, la comparaison des deux phylogénies d'organelles, mitochondriale et chloroplastique, s'impose chez les Pinaceae pour y révéler l'importance du transfert latéral et ainsi valider la pertinence de la phylogénie chloroplastique pour reconstituer l'histoire biogéographique dans cette famille. Il faut savoir qu'aucune phylogénie de Pinaceae publiée à ce jour n'a été estimée à partir du génome mitochondrial. Par conséquent, l'utilité des phylogénies chloroplastiques pour l'inférence biogéographique n'a jamais été validée jusqu'ici pour la famille des Pinaceae, pour lesquels le génome chloroplastique est transmis paternellement par le pollen.

1.3.3.2.3 - Le choix et la structure des régions de l'ADN chloroplastique

Les deux génomes d'organelles chez les plantes n'ont pas la même taille ni la même forme. Pour l'heure, le plus connu des deux est le génome chloroplastique, exclusif au monde végétal. Chez les conifères, le génome chloroplastique forme une molécule circulaire à double brin longue d'environ 120 kilobases, et il est par conséquent plus court que celui de la majorité des autres types de plantes (Ahuja 2001). Les gènes de ce génome sont conservés à l'intérieur des espèces, ce qui a permis l'estimation de phylogénies à partir de gènes chloroplastiques pour plusieurs genres et familles de conifères chez les Cupressaceae, Taxodiaceae et Pinaceae (études citées dans Ahuja 2001). Chez les *Picea*, Sigurgeirsson et Szmidt (1993) ont estimé une phylogénie en se basant sur les polymorphismes de longueurs de fragments de restriction du génome chloroplastique, et Ran *et al.* (2006) ont récemment utilisé les séquences des régions chloroplastiques *trn*C-*trn*D et *trn*T-*trn*F dans le même but.

Pour le présent travail, plusieurs régions de l'ADN chloroplastique ont été considérées, afin de déterminer celles qui affichaient les meilleurs taux de sites variables pour estimer la phylogénie du genre *Picea* à partir de séquences d'ADN chloroplastique. Malgré le fort niveau de conservation des gènes chloroplastiques, nous avons pu obtenir un niveau de polymorphisme jugé utile suite aux essais préliminaires de séquençage de trois grandes régions du génome chloroplastique: la région *rbc*L longue d'environ 1400 paires de bases, la région *trn*T-[*trn*L-intron]-*trn*F faisant environ la même longueur, et la région *trn*K d'une longueur totale d'environ 2 500 paires de bases incluant un intron et le gène *mat*K en entier. Pour la présente étude, quelques séquences partielles et quelques séquences complètes de ces régions déjà publiées dans GenBank pour les *Picea* ou d'autres Pinaceae (*e.g.* Wang *et al.* 1999, Wei et Wang 2003) ont servi de points de comparaison et de canevas pour l'élaboration des amorces d'amplification et de leurs sites d'ancrage.

Des études basées sur diverses régions de l'ADN chloroplastique ont montré que la région *rbc*L est particulièrement utile pour estimer les liens phylogénétiques entre les familles chez les gymnospermes (*e.g.* Quinn *et al.* 2002). Cependant, cette région est

beaucoup plus conservée à l'intérieur du genre *Picea* (voir chapitre 3). Elle fait partie d'un ensemble fonctionnel nommé RuBisCO, et constitue la plus grande de ses sous-unités. La région chloroplastique *rbc*L est de structure simple (Fig. 1.3 A) et de composition relativement conservée (Bousquet *et al.* 1992; Kellogg et Juliano 1997).

La région des gènes *trn*T-[*trn*L-intron]-*trn*F a été séquencée avec succès chez de nombreux groupes de plantes, car elle contient des sections relativement conservées qui permettent l'ancrage des amorces universelles (Borsch *et al.* 2003). Elle comprend deux gènes (*trn*T[UGU] et *trn*F[GAA]) très conservés et relativement courts, des espaceurs intergéniques plus variables, ainsi qu'un court exon conservé (*trn*L[UAA]), lui-même divisé en deux sections par un espaceur variable (Fig. 1.3 B).

Parmi les régions codantes connues dans le génome chloroplastique, *mat*K est l'une de celles qui évoluent le plus rapidement chez les plantes en général (Wolfe 1991 [*in* Johnson et Soltis 1995]), étant toutefois plus conservée que la plupart des gènes nucléaires. Ce gène est localisé à l'intérieur d'un intron, lui-même situé entre les deux exons (5' et 3') du gène *trn*K (Fig. 1.3 C) (Sugita *et al.* 1985; Hilu *et al.* 2003). Chez les gymnospermes, le gène *mat*K a un taux moyen de variabilité nucléotidique par site au moins trois fois plus élevé que le gène *rbc*L (Johnson et Soltis 1995; Hilu *et al.* 2003). Germano et Klein (1999) ont déjà fait état de l'importance de la variabilité du *trn*K (incluant le *mat*K) chez *Picea*. En se basant sur la comparaison de trois espèces américaines, soit *P. glauca*, *P. mariana* et *P. rubens*, cette étude mentionnait que la variabilité interspécifique pour cette région chloroplastique était trop faible en soi pour arriver à estimer une phylogénie bien résolue du genre *Picea* en utilisant seulement cette région. Par ailleurs, une phylogénie des genres de la famille des Pinaceae a déjà été estimée à l'aide du gène *mat*K (Wang *et al.* 2000), ce gène montrant alors une variabilité suffisante pour ce niveau de comparaison taxonomique.

Si chacune de ces régions chloroplastiques prise à elle seule ne peut garantir le pouvoir de résolution nécessaire pour une phylogénie claire des *Picea*, la prise en compte simultanée de plusieurs régions peut donner une résolution suffisante. Cependant, même

s'il a été recommandé par certains de recourir à la combinaison de différents ensembles de données (morphologiques, nucléaires, mitochondriales et chloroplastiques) dans le but d'acquérir une résolution plus fidèle au niveau de l'espèce (Doyle 1992), il faut s'assurer que les ensembles de données soient compatibles, et donc, qu'ils partagent la même histoire évolutive. Ceci s'avère nécessaire pour éviter de sérieuses erreurs d'inférence. Par conséquent, avant de considérer en bloc les différents ensembles de données, la compatibilité de ces derniers doit être mise à l'épreuve, que ce soit pour des ensembles de données de différents génomes ou de différentes régions à l'intérieur du même génome (*e.g.* Seelanan *et al.* 1997, Chat *et al.* 2004; de Queiroz et Gatesy 2007).

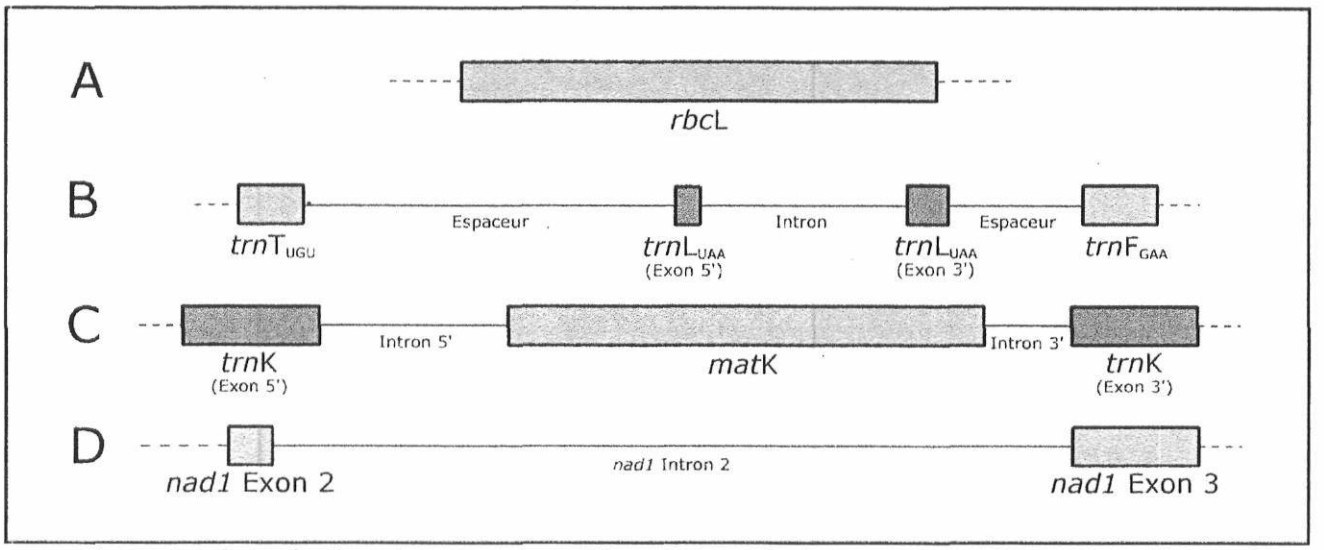

Fig. 1.3 - Organisation des gènes et régions intergéniques pour les loci étudiés (échelles variables). A) *rbc*L; B) *trn*TLF (adapté de Borsch *et al.* [2003]); C) *trn*K (adapté de Johnson et Soltis [1995]); D) Intron 2 de *nad1*.

1.3.3.2.4 - Le choix et la structure de la région de l'ADN mitochondrial

Le génome mitochondrial est relativement petit chez les animaux, mais de taille beaucoup plus grande chez les plantes supérieures (revue dans Ahuja 2001). La structure précise de ce génome est toujours inconnue chez les conifères, si ce n'est qu'il est constitué de diverses molécules linéaires et circulaires de tailles variables (Ahuja 2001).

L'ADN mitochondrial est considéré comme étant plus difficile à isoler et plus complexe que l'ADN chloroplastique (Sytsma et Schaal 1985). Ce génome aux gènes très conservés chez les plantes (Laroche *et al.* 1997) a longtemps été utile pour des études de niveaux taxonomiques égaux ou supérieurs à la famille (Hiesel *et al.* 1994), ou encore dans le cas de sujets très spécifiques autres que la taxonomie ou la phylogénétique. Auparavant, l'ADN mitochondrial était moins valorisé pour les études phylogénétiques, parce que le fort degré de réarrangement du génome rend difficile le transfert des amorces d'amplification d'une espèce à l'autre, et aussi à cause du faible taux de substitution observé chez la majorité des plantes étudiées (Wolfe *et al.* 1987; Laroche *et al.* 1997; Chat *et al.* 2004). Cependant, les récents développements de détection de régions plus variables du génome mitochondrial ont rendu ce génome de plus en plus utilisé pour les études phylogénétiques interspécifiques (*e.g.* Watano *et al.* 1996, Soranzo *et al.* 1999, Mitton *et al.* 2000). Les introns et les espaceurs intergéniques mitochondriaux, comme ceux des deux autres génomes, offrent davantage de variations que les régions codantes, et certaines études phylogénétiques les ont utilisés avec succès (*e.g.* Demesure *et al.* 1995).

Chez les gymnospermes, déjà quelques études génétiques ont abordé le génome mitochondrial, plus particulièrement la région *nad1*, soit la première sous-unité du gène pour la NADH déshydrogénase, sans pour autant viser l'estimation de la phylogénie des espèces. Chez les plantes, cette région compte souvent cinq exons, deux introns simples et deux doubles introns. Représentant les introns simples, l'intron 2 est bordé par l'exon 2 et l'exon 3, et l'intron 4 est bordé par l'exon 4 et l'exon 5 (Walheithner *et al.* 1990). L'intron 2 (Fig. 1.3 D) est relativement petit, ce qui facilite l'ancrage des amorces d'amplification à l'intérieur des exons, lesquels sont généralement plus conservés. En plus des substitutions dans cette région, de nombreux indels peuvent être inclus dans un contexte phylogénétique. Ces derniers semblent former une forte proportion de la variation pour certains groupes de plantes (Freudenstein et Chase 2001). Certains auteurs ont avancé que les indels offrent d'ailleurs une meilleure estimation des liens phylogénétiques, car ils sont plus rares (Sültmann *et al.* 1995) et moins réversibles (Meyer *et al.* 1986) que les substitutions. Les analyses préliminaires de la présente étude ont d'ailleurs démontré que l'insertion des

indels codés dans les matrices de données augmentait effectivement de façon marquée la résolution de la phylogénie des *Picea*.

À la lumière des arguments concernant la présence possible de polymorphismes intraspécifiques, puisque divers conifères montrent de tels polymorphismes pour l'intron 2 de *nad1* (*e.g.* Sinclair *et al.* 1999, Mitton *et al.* 2000, Sperisen *et al.* 2001), il est nécessaire de vérifier si les polymorphismes détectés à l'intérieur des espèces forment une structure monophylétique dans la phylogénie. Comme le séquençage de la région *nad1* représente, dans cette thèse, une première mondiale pour estimer la phylogénie d'un genre conifèrien, la présence de polymorphismes intraspécifiques a été vérifiée en analyse préliminaire pour un certain nombre d'espèces à large distribution, à l'instar de l'examen de la variation nucléotidique au niveau des loci chloroplastiques. Comme pour ces derniers, il est apparu que les polymorphismes intraspécifiques étaient plutôt rares, et que, lorsque présents, ils n'influençaient pas le caractère monophylétique de l'espèce dans la phylogénie. La plupart des polymorphismes observés pour les organelles seraient donc apparus suite à la divergence des espèces de *Picea*.

Jusqu'ici, les séquences mitochondriales n'ont jamais été utilisées pour estimer la phylogénie d'un genre de gymnospermes. Cependant, la transmission maternelle du gène *nad1* avait déjà été démontrée pour les *Picea* par Jaramillo-Correa *et al.* (2003), et les amorces spécifiques ont été développées lors de travaux préparatifs dans le cadre de la présente thèse. La comparaison de plus de 3 200 paires de bases et de nombreux indels de diverses longueurs ont pu servir à estimer la phylogénie mitochondriale d'un genre conifèrien pour la première fois et de retracer l'histoire biogéographique des *Picea*. Cette phylogénie mitochondriale a pu être comparée aux phylogénies chloroplastiques dans le but d'estimer l'importance des transferts latéraux du génome chloroplastique.

1.4 - Objectifs de la thèse

La présente thèse vise à approfondir les connaissances fondamentales sur les facteurs influençant l'évolution des *Picea* et les liens phylogénétiques qui unissent les espèces de ce genre. Pour y parvenir, les paramètres induisant la variabilité de différentes régions des génomes ont dû tout d'abord être identifiés pour déterminer les patrons d'évolution chez *Picea*. Le génome nucléaire, toujours très utilisé pour les études phylogénétiques, comportait un obstacle majeur au séquençage direct de la région *ITS-1* des espèces de *Picea* (voir Campbell *et al.* 2005). Le séquençage d'une seconde région nucléaire fort étudiée chez les angiospermes, l'*ITS-2*, n'a révélé que très peu de polymorphismes chez le genre *Picea*. Par conséquent, d'autres régions nucléaires ont été considérées, en ayant à l'esprit que certains polymorphismes intraspécifiques pouvaient s'avérer « trans-spécifiques ». L'hypothèse initiale voulait que certains loci pouvaient effectivement présenter des polymorphismes précédant la divergence des espèces, et, par conséquent, n'être d'aucune utilité pour estimer la phylogénie d'espèces chez *Picea*. Cependant, si ces polymorphismes étaient rares, ces régions du génome nucléaire pouvaient vraisemblablement être utiles pour l'estimation de la phylogénie d'espèces chez le genre *Picea*.

Ainsi, la première étude, qui fait l'objet du second chapitre, a consisté en une généalogie d'allèles pour différents loci nucléaires. Des marqueurs moléculaires avaient été développés précédemment pour des régions principalement non codantes de différents gènes du génome nucléaire (Perry et Bousquet 1998a, 1998b; Perry *et al.* 1999). Trois de ces loci avaient été retenus pour trois espèces distinctes et divergentes de *Picea* (*P. abies*, *P. glauca*, et *P. mariana*) parce qu'ils avaient montré un niveau de polymorphisme appréciable lors des tests préliminaires. Après le séquençage des allèles, les phylogénies d'allèles ont été estimées pour vérifier, d'une part, si ces marqueurs nucléaires neutres ou presque neutres pouvaient être utilisés pour estimer la phylogénie des espèces chez *Picea*. D'autre part, l'âge de ces polymorphismes ancestraux a été déterminé et les tailles de populations ancestrales ont été inférées. Cette partie de la thèse a fait l'objet d'un article publié dans la revue *American Journal of Botany* (Bouillé et Bousquet 2005).

La seconde étude, correspondant ici au troisième chapitre, visait à reconstituer la phylogénie et l'histoire biogéographique du genre *Picea* en son entier, à identifier les principaux modes de spéciation, et à tester l'hypothèse de la présence de l'évolution réticulée chez les *Picea*. L'importance de ce type particulier d'évolution moléculaire devait être évaluée par la comparaison des phylogénies chloroplastique et mitochondriale. Quant à l'origine des espèces contemporaines de *Picea*, deux régions hypothétiques avaient déjà été considérées: l'Asie ou l'Amérique. À partir de l'origine du genre contemporain, la dispersion des espèces et les événements de spéciation qui y sont associés pouvaient être inférés. Les patrons observés pouvaient possiblement corroborer certaines des classifications morphologiques actuelles, soit en confirmant le regroupement des espèces formant des hybrides naturels, ou encore par la cohérence géographique relevée à travers la phylogénie. Cette cohérence géographique pouvait vraisemblablement être le reflet du phénomène de spéciation géographique. La méthodologie impliquait de prime abord que les séquences de différents marqueurs du génome chloroplastique et du génome mitochondrial susmentionnées soient déterminées pour les 35 espèces de *Picea* ainsi que quelques taxons de référence à l'extérieur du genre *Picea*, et que l'importance de la divergence intraspécifique et interspécifique soit évaluée, puis que les phylogénies estimées soient comparées. Ces résultats multiples ont été consolidés en un troisième chapitre, qui a été soumis récemment pour publication dans la revue *Molecular Phylogenetics and Evolution.*

Le présent ouvrage rassemble donc deux grandes études distinctes unies par le point commun de l'évolution des gènes et des espèces: l'une qui ouvre une fenêtre sur la coalescence des allèles précédant la divergence des lignées menant aux espèces contemporaines, et l'autre qui se penche sur l'évolution des lignées de *Picea* suite à la divergence induite par la spéciation.

1.5 - Références

Ahuja MR. 2001. Recent advances in molecular genetics of forest trees. Euphytica 121: 173-195.

Alvarez I et Wendel JF. 2003. Ribosomal *ITS* sequences and plant phylogenetic inference. Mol. Phylogenet. Evol. 29: 417-434.

Baldwin BG. 1993. Molecular phylogenetics of *Calycadenia* (Compositae) based on *ITS* sequences of nuclear ribosomal DNA: chromosomal and morphological evolution reexamined. Amer. J. Bot. 80: 222-238.

Baldwin B, Sanderson MJ, Porter JM, Wojciechowski MF, Campbell CS et Donoghue MJ. 1995. The *ITS* region of nuclear ribosomal DNA: a valuable source of evidence on Angiosperm phylogeny. Ann. MO. Bot. Gard. 82: 247-277.

Bennuah SY, Wang T et Aiken SN. 2004. Genetic analysis of the *Picea sitchensis* x *glauca* introgression zone in British Columbia. For. Ecol. Manage. 197: 65-77.

Birky CW, Maruyama T et Fuerst P. 1983. An approach to population and evolutionary genetic theory for genes in mitochondria and chloroplasts, and some results. Genetics 103: 513-527.

Borsch T, Hilu KW, Quandt D, Wilde V, Neinhuis C et Barthlott W. 2003. Noncoding plastid *trnT-trnF* sequences reveal a well resolved phylogeny of basal angiosperms. J. Evol. Biol. 16: 558-576.

Bouillé M et Bousquet J. 2005. Trans-species shared polymorphisms at orthologous nuclear gene loci among distant species in the conifer *Picea* (Pinaceae): implications for the long-term maintenance of genetic diversity in trees. Amer. J. Bot. 92: 63-73.

Boulter MC et Manum SB. 1996. Oligocene and Miocene vegetation in high latitudes of the North Atlantic: palynological evidence from the Hovgård Ridge in the Greenland sea (site 908). Proceedings of the Ocean Drilling Program, Scientific Results 151: 289-296.

Bousquet J, Simon L, Lalonde M. 1990. DNA amplification from vegetative and sexual tissues of trees using polymerase chain reaction. Can. J. For. Res. 20: 254-257.

Bousquet J, Strauss SH, Doerksen AH et Price RA. 1992. Extensive variation in evolutionary rate of *rbcL* gene sequences among seed plants. Proc. Natl. Acad. Sci. USA 89: 7844-7848.

Boyle TJB et Morgenstern EK. 1987. Some aspects of the population structure of black spruce in central New Brunswick. Silv. Genet. 36: 53-60.

Brown GR, Gill GP, Kuntz RJ, Langley CH et Neale DB. 2004. Nucleotide diversity and linkage disequilibrium in loblolly pine. Proc. Natl. Acad. Sci. USA 101: 15255-15260.

Brown MD, Hosseini SH, Torroni A, Bandelt H-J, Allen JC, Schurr TG, Scozzari R, Cruciani F et Wallace DC. 1998. mtDNA haplogroup X: an ancient link between Europe/Western Asia and North America? Am. J. Hum. Genet. 63: 1852-1861.

Camacho FJ, Gernandt DS, Liston A, Stone JK et Klein A. 1997. Endophytic fungal DNA, the source of contamination in spruce needle DNA. Mol. Ecol. 6: 983-987.

Campbell CS, Wright WA, Cox M, Vining TF, Major CS et Arsenault MP. 2005. Nuclear ribosomal DNA internal transcribed spacer 1 (*ITS1*) in *Picea* (Pinaceae): sequence divergence and structure. Mol. Phylogenet. Evol. 35: 165-185.

CCMF. Conseil canadien des ministres des forêts. 2005.

http://nfdp.ccfm.org/compendium/silviculture/tables_index_f.php/.

Chase M, Soltis DE, Olmstead RG, Morgan D, Les DH, Mishler BD, Duvall MR, Price RA, Hills HG, Qiu Y-L, Kron KA, Rettig JH, Conti E, Palmer JD, Manhart JR, Sytsma KJ, Michaels HJ, Kress WJ, Karol KG, Clark WD, Hedrén M, Gaut BS, Jansen RK, Kim K-J, Wimpee CF, Smith JF, Furnier GR, Strauss SH, Xiang Q-Y, Plunkett GM, Soltis PS, Swensen SM, Williams SE, Gadek PA, Quinn CJ, Eguiarte LE, Golenberg E, Learn Jr GH, Graham SW, Barrett SCH, Dayanandan S et Albert VA. 1993. Phylogenetics of seed plants: an analysis of nucleotide sequences from the plastid gene *rbc*L. Ann. MO. Bot. Gard. 80: 528-580.

Chat J, Jauregui B, Petit RJ et Nadot S. 2004. Reticulate evolution in kiwifruit (*Actinidia*, Actinidiaceae) identified by comparing their maternal and paternal phylogenies. Amer. J. Bot. 91: 736-747.

Cheng Y, Nicolson RG, Tripp K et Chaw S-M. 2000. Phylogeny of Taxaceae and Cephalotaxaceae genera inferred from chloroplast *matK* gene and nuclear rDNA *ITS* region. Mol. Phylogenet. Evol. 14: 353-365.

Chesser RK et Baker RJ. 1996. Effective sizes and dynamics of uniparentally and diparentally inherited genes. Genetics 144: 1225-1235.

Collignon A-M, Van de Sype H et Favre J-M. 2002. Geographical variation in random amplified DNA and quantitative traits in Norway spruce. Can. J. For. Res. 32: 266-282.

Comtois P. 1997. Pollen dispersal and long distance transport: the case of thermophilic pollen in subarctic Canada. Aerobiologia 13: 37-42.

Davidson JF. 1952. The use of taxonomy in ecology. Ecology 33: 297-299.

Demesure B, Sodzi N et Petit RJ. 1995. A set of universal primers for amplification of polymorphic non-coding regions of mitochondrial and chloroplast DNA in plants. Mol. Ecol. 4: 129-131.

de Queiroz A et Gatesy J. 2007. The supermatrix approach to systematics. Trends Ecol. Evol. 22: 34-41.

Dobzhansky T. 1951. Genetics and the origin of species. 3^{e} édition, révisée. Columbia University Press, New York. pp. 257, 264.

Doyle JJ. 1992. Gene trees and species trees: molecular systematics as one-character taxonomy. Syst. Bot. 17: 144-163.

Ennos R. 1994. Estimating the relative rates of pollen and seed migration among plant populations. Heredity 72: 250-259.

EPICA community members. 2006. One-to-one coupling of glacial climate variability in Greenland and Antarctica. Nature 444: 195-198.

Farjon A. 1990. Pinaceae: drawings and descriptions of the genera *Abies*, *Cedrus*, *Pseudolarix*, *Keteleeria*, *Nothotsuga*, *Tsuga*, *Cathaya*, *Pseudotsuga*, *Larix* and *Picea*. Regnum Veg. 121: 1-330.

Florin R. 1963. The distribution of conifer and taxad genera in time and space. Acta Horti. Bergiani 20: 121-312.

Fowler DP. 1966. A new spruce hybrid - *Picea schrenkiana* x *P. glauca*. Proc. Second Genetics Workshop, Soc. Amer. For. 7th Lake states forest tree improvement Conf., Hickory Corners. Michigan, 1965: 44-47.

Freudenstein JV et Chase MW. 2001. Analysis of mitochondrial *nad*1b-c intron sequences in Orchidaceae: utility and coding of length-change characters. Syst. Bot. 26: 643-657.

Furnier GR, Stine M, Mohn CA et Clyde MA. 1991. Geographic patterns of variation in allozymes and height growth in white spruce. Can. J. For. Res. 21: 707-712.

Gadek PA, Alpers DL, Heslewood MM et Quinn CJ. 2000. Relationships within Cupressaceae *sensu lato*: a combined morphological and molecular approach. Amer. J. Bot. 87: 1044-1057.

Gamache I, Jaramillo-Correa JP, Payette S et Bousquet J. 2003. Diverging patterns of mitochondrial and nuclear DNA diversity in subarctic black spruce: imprint of a founder effect associated with postglacial colonization. Mol. Ecol. 12: 891-901.

Génome Canada. 2006. http://www.genomecanada.ca/.

Germano J et Klein AS. 1999. Species-specific nuclear and chloroplast single nucleotide polymorphisms to distinguish *Picea glauca*, *P. mariana* and *P. rubens*. Theor. Appl. Genet. 99: 37-49.

Gernandt DS, Geada Lopez G, Garcia SO et Liston A. 2005. Phylogeny and classification of *Pinus*. Taxon 54: 29-42.

Gordon AG. 1976. The taxonomy and genetics of *Picea rubens* and its relationship to *Picea mariana*. Can. J. Bot. 54: 781-813.

Gordon AG. 1986. Breeding, genetics and genecological studies in spruce for tree improvement in 1983 and 1984, Sault Ste.Marie, Ontario. *In*: Proc. 20th Meet. Can. Tree Improve. Assoc., Quebec City, Quebec, 1985, Part I. pp. 112-116.

Govindaraju DR. 1988. Life histories, neighbourhood sizes, and variance structure in some North American conifers. Biol. J. Linn. Soc. 35: 69-78.

Gros-Louis M-C, Bousquet J, Pâques LE et Isabel N. 2005. Species-specific diagnostic markers in *Larix* spp. based on RAPDs and nuclear, cpDNA and mtDNA gene sequences, and their phylogenetic implications. Tree Genet. Genomes 1: 50-63.

Gugerli F, Senn J, Anzidei M, Madaghiele A, Büchler U, Sperisen C et Vendramin GG. 2001. Chloroplast microsatellites and mitochondrial *nad1* intron 2 sequences indicate congruent phylogenetic relationships among Swiss stone pine (*Pinus cembra*), Siberian stone pine (*Pinus sibirica*), and Siberian dwarf pine (*Pinus pumila*). Mol. Ecol. 10: 1489-1497.

Hennig W. 1966. Phylogenetic systematics. 3^{e} édition. Trad. par DD Davis et R Zanderl. Urbana: Univ. of Illinois Press. pp. 84-90.

Herzschuh U, Zhang C, Mischke S, Herzschuh R, Mohammadi F, Mingram B, Kürschner H et Riedel F. 2005. A late Quaternary lake record from the Qilian Mountains (NW China): evolution of the primary production and the water depth reconstructed from macrofossil, pollen, biomarker, and isotope data. Global Planet. Change 46: 361-379.

Hiesel R, von Haesler A et Brennicke A. 1994. Plant mitochondrial nucleic acid sequences as a tool for phylogenetic analysis. Proc. Natl. Acad. Sci. USA 91: 634-638.

Hills LV et Ogilvie RT. 1970. *Picea banksii n. sp.* Beaufort formation (Tertiary), northwestern Banks Island, Arctic Canada. Can. J. Bot. 48:457-464.

Hilu KW, Borsch T, Müller K, Soltis DE, Soltis PS, Savolainen V, Chase MW, Powell MP, Alice LA, Evans R, Sauquet H, Neinhuis C, Slotta TAB, Rohwer JG, Campbell CS et Chatrou LW. 2003. Angiosperm phylogeny based on *matK* sequence information. Amer. J. Bot. 90: 1758-1776.

Holyoak DT. 1983. The identity and origins of *Picea abies* (L.) Karsten from the Chelford interstadial (Late Pleistocene) of England. New Phytol. 95: 153-157.

Isabel N, Beaulieu J et Bousquet J. 1995. Complete congruence between gene diversity estimates derived from genotypic data at enzyme and random amplified polymorphic loci in black spruce. Proc. Natl. Acad. Sci. USA 92: 6369-6373.

Jaramillo-Correa JP, Beaulieu J et Bousquet J. 2001. Contrasting evolutionary forces driving population structure at expressed sequence tag polymorphisms, allozymes and quantitative traits in white spruce. Mol. Ecol. 10: 2729-2740.

Jaramillo-Correa JP, Beaulieu J, Ledig FT et Bousquet J. 2006. Decoupled mitochondrial and chloroplast DNA population structure reveals Holocene collapse and population isolation in a threatened Mexican-endemic conifer. Mol. Ecol. 15: 2787-2800.

Jaramillo-Correa JP et Bousquet J. 2003. New evidence from mitochondrial DNA of a progenitor-derivative species relationship between black spruce and red spruce (Pinaceae). Amer. J. Bot. 90: 1801-1806.

Jaramillo-Correa JP, Bousquet J, Beaulieu J, Isabel N, Perron M et Bouillé M. 2003. Cross-species amplification of mitochondrial DNA sequence-tagged-site markers in conifers: the nature of polymorphism and variation within and among species in *Picea*. Theor. Appl. Genet. 106: 1353-1367.

Jeandroz S, Roy A et Bousquet J. 1997. Phylogeny and phylogeography of the circumpolar genus *Fraxinus* (Oleaceae) based on internal transcribed spacer sequences of nuclear ribosomal DNA. Mol. Phylogenet. Evol. 7: 241-251.

Johnson LA et Soltis DE. 1995. Phylogenetic inference in Saxifragaceae *sensu stricto* and *Gilia* (Polemoniaceae) using *matK* sequences. Ann. MO. Bot. Gard. 82: 149-175.

Karvonen P, Szmidt AE et Savolainen O. 1994. Length variation in the internal transcribed spacers of ribosomal DNA in *Picea abies* and related species. Theor. Appl. Genet. 89: 969-974.

Kelch DG. 1998. Phylogeny of Podocarpaceae: comparison of evidence from morphology and *18S* rDNA. Amer. J. Bot. 85: 986-996.

Kellogg EA et Juliano ND. 1997. The structure and function of RuBisCO and their implications for systematic studies. Amer. J. Bot. 84: 413-428.

Kim K-J et Mabry TJ. 1991. Phylogenetic and evolutionary implications of nuclear ribosomal DNA variation in dwarf dandelions (*Krigia*, Lactuceae, Asteraceae). Plant Syst. Evol. 177: 53-69.

Klein J, Sato A, Nagl S et O'hUigin C. 1998. Molecular trans-species polymorphisms. Ann. Rev. Ecol. Syst. 29: 1-21, C1.

Kobayashi K, Yoshikawa J et Suzuki M. 2000. DNA identification of *Picea* species of the last glacial age in northern Japan. Jpn. J. Histor. Bot. 8: 67-80.

Krutovskii KV et Bergmann F. 1995. Introgressive hybridization and phylogenetic relationships between Norway, *Picea abies* (L.) Karst., and Siberian, *P. obovata* Ledeb., spruce species studied by isozyme loci. Heredity 74: 464-480.

Laroche J, Li P, Maggia L et Bousquet J. 1997. Molecular evolution of angiosperm mitochondrial introns and exons. Proc. Natl. Acad. Sci. USA 94: 5722-5727.

Ledig FT, Hodgskiss PD, Krutovskii KV, Neale DB et Eguiluz-Piedra T. 2004. Relationships among the spruces (*Picea*, Pinaceae) of Southwestern North America. Syst. Bot. 29: 275-295.

LePage BA. 2001. New species of *Picea* A. Dietrich (Pinaceae) from the middle Eocene of Axel Heiberg Island, Arctic Canada. Bot. J. Linn. Soc. 135: 137-167.

LePage BA. 2003a. The evolution, biogeography and palaeoecology of the Pinaceae based on fossil and extant representatives. Acta Hort. 615: 29-52.

LePage BA. 2003b. A new species of *Tsuga* (Pinaceae) from the middle Eocene of Axel Heiberg Island, Canada, and an assessment of the evolution and biogeographical history of the genus. Bot. J. Linn. Soc. 141: 257-296.

Li HL. 1953. Present distribution and habitats of the conifers and taxads. Evolution 7: 245-261.

Linder CR et Rieseberg LH. 2004. Reconstructing patterns of reticulate evolution in plants. Amer. J. Bot. 91: 1700-1708.

Liston A, Robinson A, Piñero D et Alvarez-Buylla ER. 1999. Phylogenetics of *Pinus* (Pinaceae) based on nuclear ribosomal DNA internal transcribed spacer region sequences. Mol. Phylogenet. Evol. 11: 95-109.

Liston A, Gernandt DS, Vining TF, Campbell CS et Piñero D. 2003. Molecular phylogeny of Pinaceae and *Pinus*. Acta Hortic. 615: 107-114.

Liu J-Q, Gao T-G, Chen Z-D et Lu A-M. 2002. Molecular phylogeny and biogeography of the Qinghai-Tibet Plateau endemic *Nannoglottis* (Asteraceae). Mol. Phylogenet. Evol. 23: 307-325.

Liu K-B. 1988. Quaternary history of the temperate forests of China. Quaternary Science Reviews 7: 1-20.

Liu T-S. 1982. A new proposal for the classification of the genus *Picea*. Acta Phytotax. Geobot. 33: 227-244.

Machado CA, Kliman RM, Markert JA et Hey J. 2002. Inferring the history of speciation from multilocus DNA sequence data: the case of *Drosophila pseudoobscura* and close relatives. Mol. Biol. Evol. 19: 472-488.

Maddison WP et Knowles LL. 2006. Inferring phylogeny despite incomplete lineage sorting. Syst. Biol. 55: 21-30.

Maggini F, Marrocco R, Gelati MT et Ilio De Dominicis R. 1998. Lengths and nucleotide sequences of the internal spacers of nuclear ribosomal DNA in gymnosperms and pteridophytes. Plant Syst. Evol. 213: 199-205.

Mayr E. 1942. Systematics and the origin of species. Columbia University Press, New York. p. 21.

Mayr E. 1963. Animal species and evolution. Belknap Press. Cambridge, MA. 797 pp.

Meyer TE, Cusanovich MA et Kamen MD. 1986. Evidence against the use of bacterial amino acid sequence data for construction of all-inclusive phylogenetic trees. Proc. Natl. Acad. Sci. USA 83: 217-220.

Mikkola L. 1969. Observations on interspecific sterility in *Picea*. Ann. Bot. Fenn. 6: 285-339.

Miksche JP. 1985. Recent advances of biotechnology and forest trees. Forest Chron. 61: 449-453.

Miller Jr CN. 1988. The origin of modern conifer families. *In*: Beck, C.B., ed. Origin and evolution of gymnosperms. Columbia University Press, New York. pp. 448-486.

Miller Jr CN. 1989. A new species of *Picea* based on silicified seed cones from the Oligocene of Washington. Amer. J. Bot. 76: 747-754.

Mitton JB. 1994. Molecular approaches to population biology. Ann. Rev. Ecol. Sys. 25:

45-69.

Mitton JB, Kreiser BR et Rehfeldt GE. 2000. Primers designed to amplify a mitochondrial *nad1* intron in ponderosa pine, *Pinus ponderosa*, limber pine, *P. flexilis*, and Scots pine, *P. sylvestris*. Theor. Appl. Genet. 101: 1269-1272.

Mogensen HL. 1996. The hows and whys of cytoplasmic inheritance in seed plants. Amer. J. Bot. 83: 383-404.

Müller-Starck G. 1995. Genetic variation in high elevated populations of Norway spruce (*Picea abies* [L.] Karst.) in Switzerland. Silv. Genet. 44: 356-362.

Mullis KB et Faloona FA. 1987. Specific synthesis of DNA *in vitro* via a polymerase catalyzed chain reaction. Methods Enzymol. 155: 333-350.

Neale DB et Sederoff RR. 1989. Paternal inheritance of chloroplast DNA and maternal inheritance of mitochondrial DNA in loblolly pine. Theor. Appl. Genet. 77: 212-216.

Ohsawa T. 1994. Anatomy and relationships of petrified seed cones of the Cupressaceae, Taxodiaceae, and Sciadopityaceae. J. Plant Res. 107: 503-512.

Page CN et Hollands RC. 1987. The taxonomic and biogeographic position of Sitka spruce. Proc. R. Soc. Edinburgh Sect. B Biol. Sci. 93B: 13-24.

Pamilo P et Nei M. 1988. Relationships between gene trees and species trees. Mol. Biol. Evol. 5: 568-583.

Perron M et Bousquet J. 1997. Natural hybridization between black spruce and red spruce. Mol. Ecol. 6: 725-734.

Perron M, Perry DJ, Andalo C et Bousquet J. 2000. Evidence from sequence-tagged-site markers of a recent progenitor-derivative species pair in conifers. Proc. Natl. Acad. Sci. USA 97: 11331-11336.

Perry DJ et Bousquet J. 1998a. Sequence-tagged-site (STS) markers of arbitrary genes: development, characterization and analysis of linkage in black spruce. Genetics 149: 1089-1098.

Perry DJ et Bousquet J. 1998b. Sequence-tagged-site (STS) markers of arbitrary genes: the utility of black spruce-derived STS primers in other conifers. Theor. Appl. Genet. 97: 735-743.

Perry DJ, Isabel N et Bousquet J. 1999. Sequence-tagged-site (STS) markers of arbitrary genes: the amount and nature of variation revealed in Norway spruce. Heredity 83: 239-248.

Perry DJ et Bousquet J. 2001. Genetic diversity and mating system of post-fire and post-harvest black spruce: an investigation using codominant sequence-tagged-sites (STS) markers. Can. J. For. Res. 31: 32-40.

Petit RJ, Pineau E, Demesure D, Bacilieri R, Ducousso A et Kremer A. 1997. Chloroplast DNA footprints of postglacial recolonization by oaks. Proc. Natl. Acad. Sci. USA 94: 9996-10001.

Prager EM, Fowler DP et Wilson AC. 1976. Rates of evolution in conifers (Pinaceae). Evolution 30: 637-649.

Quijada A, Liston A, Delgado P, Vazquez-Lobo A et Alvarez-Buylla ER. 2004. Variation in the nuclear ribosomal DNA internal transcribed spacer (*ITS*) region of *Pinus rzedowskii* revealed by PCR-RFLP. Theor. Appl. Genet. 96: 539-544.

Quijada A, Liston A, Robinson W et Alvarez-Buylla E. 1997. The ribosomal *ITS* region as a marker to detect hybridization in pines. Mol. Ecol. 6: 995-996.

Quinn CJ, Price RA et Gadek PA. 2002. Familial concepts and relationships in the conifers based on *rbcL* and *matK* comparisons. Kew Bull. 57: 513-531.

Rabineau M, Berné S, Olivet J-L, Aslanian D, Guillocheau F et Joseph P. 2006. Paleo sea levels reconsidered from direct observation of paleoshoreline position during Glacial Maxima (for the last 500,000 years). Earth Planet. Sci. Lett. 252: 119-137.

Rajora OP et Dancik BP. 2000. Population genetic variation, structure, and evolution in Engelmann spruce, white spruce, and their natural hybrid complex in Alberta. Can. J. Bot. 78: 768-780.

Ran J-H, Wei X-X et Wang X-Q. 2006. Molecular phylogeny and biogeography of *Picea* (Pinaceae): implications for phylogeographical studies using cytoplasmic haplotypes. Mol. Phylogenet. Evol. 41: 405-419.

Reidla M, Kivisild T, Metspalu E, Kaldma K, Tambets K, Tolk H-V, Parik J, Loogväli E-L, Derenko M, Malyarchuk B, Bermisheva M, Zhadanov S, Pennarun E, Gubina M, Golubenko M, Damba L, Fedorova S, Gusar V, Grechanina E, Mikerezi I, Moisan J-P, Chaventré A, Khusnutdinova E, Osipova L, Stepanov V, Voevoda M, Achilli A, Rengo C, Rickards O, De Stefano GF, Papiha S, Beckman L, Janicijevic B, Rudan P, Anagnou N, Michalodimitrakis E, Koziel S, Usanga E, Geberhiwot T, Herrnstadt C, Howell N, Torroni A et Villems R. 2003. Origin and diffusion of mtDNA haplogroup X. Am. J. Hum. Genet. 73: 1178-1190.

Rieseberg LH et Soltis DE. 1991. Phylogenetic consequences of cytoplasmic gene flow in plants. Evol. Trends Plants 5: 65-84.

Rieseberg LH et Wendel JF. 1993. Introgression and its consequences in plants. *In* RG Harrison (ed.), Hybrid zones and the evolutionary process. Oxford University Press, Oxford, UK. p. 98.

Rodgers SO et Bendich AJ. 1987. Ribosomal RNA genes in plants: variability in copy number and in the intergenic spacer. Plant Mol. Biol. 9: 509-520.

Rosenberg NA. 2003. The shapes of neutral gene genealogies in two species: probabilities of monophyly, paraphyly, and polyphyly in a coalescent model. Evolution 57: 1465-1477.

Roucoux KH, Tzedakis PC, de Abreu L et Shackleton NJ. 2006. Climate and vegetation changes 180,000 to 345,000 years ago recorded in a deep-sea core off Portugal. Earth Planet. Sci. Lett. 249: 307-325.

Ruddiman WF et Kutzbach JE. 1991. Plateau uplift and climatic change. Sci. Am. March: 66-75.

Savard L, Michaud M et Bousquet J. 1993. Genetic diversity and phylogenetic relationships between birches and alders using *ITS*, *18S* rRNA and *rbcL* gene sequences. Mol. Phylogenet. Evol. 2: 112-118.

Schaal BA et Learn Jr GH. 1988. Ribosomal DNA variation within and among plant populations. Ann. MO. Bot. Gard. 75: 1207-1216.

Schaal BA, O'Kane SL et Rogstad SH. 1991. DNA variation in plant populations. Trends Ecol. Evol. 6: 323-333.

Schmidt PA. 1989. Contributions to the taxonomy and evolution of the genus *Picea* A. Dietr. Flora 182: 435-461.

Seelanan T, Schnabel A et Wendel JF. 1997. Congruence and consensus in the Cotton tribe (Malvaceae). Syst. Bot. 22: 259-290.

Semerikov VL, Zhang H, Sun M et Lascoux M. 2003. Conflicting phylogenies of *Larix*

(Pinaceae) based on cytoplasmic and nuclear DNA. Mol. Phylogenet. Evol. 27: 173-184.

Sigurgeirsson A et Szmidt AE. 1993. Phylogenetic and biogeographic implications of chloroplast DNA variation in *Picea*. Nord. J. Bot. 13: 233-246.

Sinclair WT, Morman JD et Ennos RA. 1999. The postglacial history of Scots pine (*Pinus sylvestris* L.) in western Europe: evidence from mitochondrial DNA variation. Mol. Ecol. 8: 83-88.

Smith DE et Klein AS. 1994. Phylogenetic inferences on the relationships of North American and European *Picea* species based on nuclear ribosomal *18S* sequences and the internal transcribed spacer 1 region. Mol. Phylogenet. Evol. 3: 17-26.

Soranzo N, Provan J et Powell W. 1999. An example of microsatellite length variation in the mitochondrial genome of conifers. Genome 42: 158-161.

Sperisen C, Büchler U, Gugerli F, Mátyás G, Geburek T et Vendramin GG. 2001. Tandem repeats in plant mitochondrial genomes: application to the analysis of population differentiation in the conifer Norway spruce. Mol. Ecol. 10: 257-263.

Steig EJ. 2006. The south-north connection. Nature 444: 152-153.

Sugita M, Shinozaki K et Sugiura M. 1985. Tobacco chloroplast $tRNA^{Lys}$ (UUU) gene contains a 2.5-kilobase-pair intron: an open reading frame and a conserved boundary sequence in the intron. Proc. Natl. Acad. Sci. USA 82: 3557-3561.

Sültmann H, Mayer WE, Figueroa F, Tichy H et Klein J. 1995. Phylogenetic analysis of cichlid fishes using nuclear DNA markers. Mol. Biol. Evol. 12: 1033-1047.

Sutton BCS, Flanagan DJ, Gawley JR, Newton CH, Lester DT et El-Kassaby YA. 1991. Inheritance of chloroplast and mitochondrial DNA in *Picea* and composition of hybrids from introgression zones. Theor. Appl. Genet. 82: 242-248.

Syring J, Willyard A, Cronn R et Liston A. 2005. Evolutionary relationships among *Pinus* (Pinaceae) subsections inferred from multiple low-copy nuclear loci. Amer. J. Bot. 92: 2086-2100.

Sytsma KJ et Schaal BA. 1985. Phylogenetics of the *Lisianthius skinneri* (Gentianaceae) species complex in Panama utilizing DNA restriction fragment analysis. Evolution 39: 594-608.

Sytsma KJ et Schaal BA. 1990. Ribosomal DNA variation within and among individuals of *Lisianthius* (Gentianaceae) populations. Plant Syst. Evol. 170: 97-106.

Szmidt AE, El-Kassaby YA, Sigurgeirsson A et Alden T. 1988. Classifying seedlots of *Picea sitchensis* and *P. glauca* in zones of introgression, using restriction analysis of chloroplast DNA. Theor. Appl. Genet. 76: 841-845.

Tiffney BH. 2000. Geographic and climatic influences on the Cretaceous and Tertiary history of Euramerican floristic similarity. Acta Univ. Carol. Geol. 44: 5-16.

Turna I. 2004. Variation of morphological characters of Oriental spruce (*Picea orientalis*) in Turkey. Biologia, Bratislava 59: 519-526.

Vidacovic M. 1991. Conifers - morphology and variation. USDA, Forest Service. US-Yugoslav Joint Board on Scientific and Technological Cooperation. Ed. Boro Brerkalo. Republic of Croatia. Graficki zavod Hrvatske. pp. 286-372.

Walheithner JA, Macfarlane JL et Wolstenholme DR. 1990. A sequence encoding a maturase-related protein in a group II intron of a plant mitochcondrial *nad*1 gene. Proc. Natl. Acad. Sci. USA 87: 548-552.

Wang WP, Hwang CY, Lin TP et Hwang SY. 2003. Historical biogeography and phylogenetic relationships of the genus *Chamaecyparis* (Cupressaceae) inferred from chloroplast DNA polymorphism. Plant Syst. Evol. 241: 13-28.

Wang X-Q, Tank DC et Sang T. 2000. Phylogeny and divergence times in Pinaceae: evidence from three genomes. Mol. Biol. Evol. 17: 773-781.

Wang X-R, Tsumura Y, Yoshimaru H, Nagasaka K et Szmidt AE. 1999. Phylogenetic relationships of Eurasian pines (*Pinus*, Pinaceae) based on chloroplast *rbcL*, *matK*, *rpl20-rps18* spacer, and *trnV* intron sequences. Amer. J. Bot. 86: 1742-1753.

Watano Y, Imazu M et Shimizu T. 1996. Spatial distribution of cpDNA and mtDNA haplotypes in a hybrid zone between *Pinus pumila* and *P. parviflora* var. *pentaphylla* (Pinaceae). J. Plant Res. 109: 403-408.

Wei X-X et Wang X-Q. 2003. Phylogenetic split of *Larix*: evidence from paternally inherited cpDNA *trnT-trnF* region. Plant Syst. Evol. 239: 67-77.

Weng C et Jackson ST. 2000. Species differentiation of North American spruce (*Picea*) based on morphological and anatomical characteristics of needles. Can. J. Bot. 78: 1367-1383.

White JM, Ager TA, Adam DP, Leopold EB, Liu G, Jette H et Schweger CE. 1997. An 18 million year record of vegetation and climate change in northwestern Canada and Alaska: tectonic and global correlates. Palaeogeo. P. 130: 293-306.

White T-J, Bruns T, Lee S et Taylor J. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. *In* "PCR Protocols, a guide to methods and applications" (Innis MA, Gelfang DH, Sninsky JJ et White TJ, édit.), Academic Press, San Diego, California. pp. 315-322.

Wolfe JA. 1964. Miocene floras from Fingerrock Wash, southwestern Nevada. US Geol. Survey Prof. Paper 454N: 1-36.

Wolfe KH, Li WH et Sharp PM. 1987. Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. Proc. Natl. Acad. Sci. USA 84: 9054-9058.

Won H et Renner SS. 2005. The internal transcribed spacer of nuclear ribosomal DNA in the gymnosperm *Gnetum*. Mol. Phylogenet. Evol. 36: 581-597.

Wright JW. 1955. Species crossability in spruce in relation to distribution and taxonomy. For. Sci. 1: 319-349.

Zhang D, Fengquan L et Jianmin B. 2000. Eco-environmental effects of the Qinghai-Tibet Plateau uplift during the Quaternary in China. Envir. Geol. 39: 1352-1358.

Zhang DX et Hewitt GM. 2003. Nuclear DNA analyses in genetic studies of populations: practice, problems and prospects. Mol. Ecol. 12: 563-584.

Chapitre 2

Les polymorphismes trans-spécifiques partagés entre des espèces coniférienes distinctes chez le genre *Picea* (Pinaceae): implications pour le maintien à long terme de la diversité génétique des arbres

Ce chapitre a été publié sous forme d'article scientifique dans la revue *American journal of Botany*, et peut être trouvé sous la référence « Bouillé M. et Bousquet J. 2005. Trans-species shared polymorphisms at orthologous nuclear gene loci among distant species in the conifer *Picea* (Pinaceae): implications for the long-term maintenance of genetic diversity in trees. Amer. J. Bot. 92: 63-73 ».

2.1 - Résumé

Des polymorphismes intraspécifiques et trans-spécifiques partagés ont été détectés pour trois loci nucléaires pour trois espèces distinctes du genre *Picea*. La période de la divergence entre les espèces a été estimé à 13 à 20 millions d'années (mA) et les temps de coalescence des allèles étaient de l'ordre de 10 à 18 mA, ce qui implique que les polymorphismes puissent être partagés entre les espèces depuis leur ancêtre commun. La fixation des polymorphismes aurait été retardée par de grandes tailles historiques de populations, lesquelles ont été inférées à plus de 100 000 individus. Ces résultats suggèrent que les polymorphismes trans-spécifiques partagés peuvent être fréquents pour les loci nucléaires de plantes et qu'ils peuvent contribuer significativement à la diversité allélique, surtout chez les plantes anémophiles et allogames possédant de grands effectifs de populations. Ces polymorphismes appellent à la prudence quant à l'utilisation de gènes nucléaires pour estimer les phylogénies de tels groupes de plantes.

Trans-species shared polymorphisms at orthologous nuclear gene loci among distant species in the conifer *Picea* (Pinaceae): implications for the long-term maintenance of genetic diversity in trees

2.2 - Abstract

For each of three nuclear gene loci, intraspecific as well as trans-specific shared polymorphisms were detected in DNA among three distantly related species in the genus *Picea*. Few fixed interspecific polymorphisms were observed. Allele genealogies did not match species phylogenies, and species lineages were not reciprocally monophyletic. Based on molecular clocks and morphological evidence from the fossil record, the divergence time between species was estimated at 13–20 million years (my), and a mutation rate of 2.23×10^{-10} to 3.42×10^{-10} per site per year was estimated. Large historical population sizes in excess of 100 000 were inferred, which would have delayed the fixation of polymorphisms. These numbers translated into allele coalescence times in the order of 10 to 18 my, which implies the sharing of polymorphisms since common ancestry. These results suggest that trans-species shared polymorphisms might be frequent at plant nuclear gene loci, leading to high allelic diversity. Such a trend is more likely in trees and plants characterized by ecological and life-history determinants favoring large population sizes such as an outcrossing mating system, wind pollination, and a dominant position in ecosystem. These polymorphisms also call for caution in estimating congeneric species phylogenies from nuclear gene sequences in such plant groups.

2.3 - Introduction

In plants, evidence of allele coalescence preceding ancient lineage divergence has been reported among widely divergent taxa in the Solanaceae for the self-incompatibility locus (Ioerger et al., 1990). The maintenance of trans-species shared polymorphisms among taxa such as *Nicotiana*, *Petunia*, and *Solanum* is thought to be locus-specific, involving frequency-dependent selection. Trans-species shared polymorphisms have also been observed between closely related taxa with incomplete reproductive isolation such as in *Drosophila* (Wang et al., 1997; Machado et al., 2002) and in *Abies* (Isoda and Shiraishi, 2001), which suggests the involvement of recent lateral gene flow. The maintenance of trans-species shared polymorphisms at orthologous nuclear gene loci could also depend on the fixation time of polymorphisms and thus, factors related to population size (Tajima, 1983; Rosenberg, 2003). For neutral or nearly neutral polymorphisms, large population sizes would retard their fixation and favor their persistence in populations (Kimura and Ohta, 1969; Hudson, 1990).

Most conifers including *Picea* sp. are characterized by an open-pollinated mating system and wind-dispersed pollen, which appear to promote extensive nuclear gene flow and large population sizes (Hamrick et al., 1992). This trend is best illustrated by the little differentiation in allozymes or nuclear DNA markers usually observed among spruce populations (Boyle and Morgenstern, 1987; Furnier et al., 1991; Isabel et al., 1995; Müller-Starck, 1995; Jaramillo-Correa et al., 2001; Perry and Bousquet, 2001; Collignon et al., 2002; Gamache et al., 2003). In the long-term, these factors should retard the fixation of neutral or nearly neutral polymorphisms, resulting in increased genetic diversity. Such factors have been suggested to account, at least partly, for the maintenance of high levels of allozyme diversity observed in conifers and trees in general, as compared to most other plants and organisms (e.g., Hamrick et al., 1992; Hamrick and Godt, 1996; Ledig, 1998).

Thus, an investigation of the genealogy of orthologous alleles at nuclear gene loci among distant conifer species might reveal deep coalescence, perhaps preceding species and lineage divergence. If so, shared ancestral polymorphisms would lead to allele

genealogies that conflict with species phylogenies. While fixed polymorphisms accumulated because lineage divergence would favor the monophyly of species, ancient shared polymorphisms would tend to break up such monophyletic assemblages.

Codominant markers of expressed nuclear genes have been developed for genetic mapping and population genetics studies for a number of spruce species (Perry and Bousquet, 1998a, b; Perry et al., 1999). These markers correspond to segregating alleles at arbitrarily chosen orthologous gene loci. Sampling allele diversity at the DNA sequence level for these gene loci offers the possibility to investigate allele coalescence in a set of reproductively isolated and divergent species characterized by high genetic diversity and reputed large effective population sizes. In this study, alleles for each of three nuclear gene loci were sampled, sequenced, and compared among three distantly related biological species in the genus *Picea*.

2.4 - Material and methods

2.4.1 - Locus and taxa sampling

The core of this study is based on the sampling of three transcribed regions of the nuclear genome previously described: *Sb16*, part of a putative gene for the 60S ribosomal protein L13a spanning DNA sequences from two introns and three exons; *Sb29*, part of a putative gene for the "No Apical Meristem" (NAM) protein fragment comprising DNA sequences from one exon and one intron; and *Sb62*, part of a putative gene for the ribosomal protein L15 spanning DNA sequences from two introns and two exons (Perry and Bousquet, 1998a). For each of these loci, the presence of length polymorphisms caused by insertions-deletions (indels) was previously observed among segregating alleles in at least one *Picea* species, and some of the length variant alleles were shared between species (Perry and Bousquet, 1998a, b; Perry et al., 1999).

The three spruce species sampled, the North American, sympatric *Picea glauca* (Moench) Voss and *P. mariana* (Mill.) B.S.P., and the North Eurasian *P. abies* (L.) Karst.

are dominant in their ecosystems with large natural ranges (Wright, 1955). They harbor an open-pollinated mating system and extensive intraspecific gene flow for nuclear genes (e.g., Perry and Bousquet, 2001; Gamache et al., 2003). These various parameters should concur in large effective population sizes. On the other hand, they are not closely related: contrary to closely related taxa in the genus, they do not cross naturally and show little compatibility with each other (Wright, 1955; Mikkola, 1969), so they warrant recognition as distinct biological species. They also represent divergent taxa in terms of morphology (Wright, 1955; Weng and Jackson, 2000), with deep phylogenetic branching in the genus, as shown by phylogenies based on paternally inherited cpDNA (Sigurgeirsson and Szmidt, 1993; Bouillé and Bousquet, unpublished data). Hence, it is highly unlikely that if trans-species shared polymorphisms were observed, they could be attributed to recent lateral gene flow between species.

For each species and each nuclear locus, around 20 alleles (haplotypes) were sampled from a panel of individuals representative of multiple populations from diverse areas of the species ranges so that most common alleles would be sampled for each locus in each species. To ensure that only one allele was sequenced at a time, haploid megagametophytes were used. Sequences from outgroup taxa belonging to the Pinaceae family were also determined for rooting the three gene trees: *Abies lasiocarpa*, *Pinus cembra*, *Pseudotsuga mensiezii*, and *Tsuga canadensis*. In order to estimate the divergence time between the three spruce species (see next), the chloroplast gene *rbcL* (large subunit of ribulose-1,5-bisphosphate carboxylase) was also sequenced for each of *Picea abies*, *P. glauca*, and *P. mariana*. Other conifer *rbcL* sequences were retrieved from Genbank (www.ncbi.nlm.nih.gov/Genbank).

2.4.2 - Biochemical methods

Genomic DNA was extracted with the DNeasy Plant mini kit (Qiagen, Mississauga, Ontario). PCR and sequencing were carried out using the following primers: for *Sb16*, forward primer 5'-GTTCCGCCACCATATGAC-3' and reverse 5'-GCTCATTCAGCTACAAAAGC-3'; for *Sb29* forward 5'-AGCGGCATTGAACAGAGTAAC-3' and reverse 5'-AATGGAAATGAAGGCAGACTC-3';

and for *Sb62*, forward 5'-TGAGATCCGTGGCTGAAGAG-3' and reverse 5'-GATAACGCCGGAGAGATAGAG-3'. PCR was performed with Platinum *Taq* DNA polymerase (Gibco BRL, Carlsbad, California, USA) according to the manufacturer's recommendations: 4 min at 95°C for denaturation; 40 cycles consisting of 30 s at 95°C, 30 s at annealing temperatures adjusted to 60°C for *Sb16*, and 55°C for *Sb29* and *Sb62*, and 1 min at 72°C for extension; and final elongation for 10 min at 72°C. For *rbcL*, the primers and PCR protocol followed previously published procedures (Wang et al., 1999). Amplified fragments were purified and concentrated with Amicon Microcon-PCR filter units (Millipore, Bedford, Massachusetts, USA). Both DNA strands were directly sequenced with Perkin-Elmer ABI DNA sequencers (3100 and 3700, Applied Biosystems, Foster City, California, USA), using BigDye Terminator cycle sequencing kits. Genbank accession numbers are as follows: for *Sb16*, AY606806 to AY606815; for *Sb29*, AY6117037 to AY611050; for *Sb62*, AY611051 to AY611064; for *rbcL*, AY611034 to AY611036.

2.4.3 - Data analysis

For each nuclear locus, indels as well as substitutions were identified among alleles within and across species. Because of the small number of substitutions observed between alleles within and among species, correcting for multiple substitutions had no effect on the observed numbers of substitutions. In addition, all indels were treated as single events because there was no indication that the observed indels were associated with simple sequence repeats or other potential source of homoplasious variation. Nucleotide diversity (π) was estimated for each species using the program DnaSP (Rozas and Rozas, 1999). The average numbers of pairwise differences per site, including substitutions and indels, were estimated between alleles within species (d_w) and between species (d_b). Net between-species divergence values per site were also obtained by correcting for within-species diversity according to Nei (1987, equation 10.21). For each nuclear locus, a site-by-site analysis was also conducted, and sites were classified as those carrying (1) shared polymorphisms between *Picea* species (trans-species shared polymorphisms, those derived polymorphisms shared by at least two species which do not support species monophyly),

(2) reciprocally fixed differences between *Picea* species (interspecific fixed polymorphisms, which support species monophyly), and (3) apomorphies (intraspecifically derived polymorphisms limited to one species). The derived or ancestral nature of each state at a polymorphic site was determined by comparison with outgroup sequences from other genera of the Pinaceae. A state was assumed as derived when it was different than that observed in the outgroup sequences.

Gene/allele genealogies were estimated for each locus using parsimony analysis and the neighbor-joining method as implemented in PAUP* 4.0b10 (Swofford, 2002). For both methods, pairwise deletion of indel sites was enforced so the number of characters corresponded to the length of the consensus sequence between any two taxa, plus a number of additional binary characters for scoring the presence or absence of each indel. If a substitution and an indel were superimposed on a same site, they were considered independently, and a missing value was assigned to the substitution site(s) for those sequences with the deletion. Parsimony analyses consisted of heuristic searches with 100 replicates of random additions of sequences, Tree-bisection-reconnection branch swapping and saving multiple trees. Scenarios constraining alleles of the same species to monophyly (reciprocal monophyly) were tested: differences in tree length between constrained trees and minimum trees were tested for statistical significance using the T-PTP test of Faith (1991) as implemented in PAUP* with 5000 permutations for each gene and by considering both substitutions and indels. For neighbor-joining, pairwise numbers of differences per site were used, including substitutions and indels. The robustness of internal nodes was assessed by means of 500 bootstrap replicates for each method. Nodes supported by a frequency equal or larger than 50% were retained.

For each locus, intragenic recombination was tested among *Picea* sequences following the "four-gamete test" (Hudson and Kaplan, 1985) as implemented in DnaSP (Rozas and Rozas, 1999). The homoplasy test (Maynard Smith and Smith, 1998) was also conducted with the START 1.0.6 software (Jolley et al., 2001) and corroborated the results obtained earlier in all cases. If recombinant sites were identified for a given gene, the

putative recombinant types were identified as follows. The ancestral state of character at each recombining position was determined from the outgroup sequences, thus allowing the identification of the allele of the ancestral type. Then, both allele types, each different from the ancestral allele at only one of the two recombinant sites, were assumed derived, thus giving rise to two derived types. The recombinant type was identified as the one resulting from recombination between these two derived types. Recombination between the ancestral type and a derived type at the two recombinant sites is also possible but less likely, especially if the double-derived type is rare or unique, such as an apomorphy. Most likely, this double-derived type is rather the recombinant type. Once a putative recombinant type allele was identified for a given gene, it was substracted from the data set, and the gene tree was estimated again to verify the robustness of tree topology.

A mutation rate (per site per year) was estimated for the three nuclear gene loci studied by averaging the divergence values per site (substitutions only) estimated between the three spruce species and then dividing by $2T$ where T is the divergence time between the spruce species under study. The first approach to estimate T relied on an analysis of the fossil record, where the earliest fossils morphologically representative of extant taxa could be interpreted as evidence for the diversification of the genus *Picea* in its major extant lineages. The second approach relied on estimating a molecular clock from rates of protein divergence, as estimated by antigenic distances between the genus *Picea* and both subgenera of *Pinus* in the Pinaceae (Prager et al., 1976) (Fig. 2.1). The lineage leading to *Picea* split early from the lineage leading to *Pinus* in the history of the Pinaceae (e.g., Magallón and Sanderson, 2002). Thus, the average distance between *Picea* and *Pinus* was calibrated by using as landmark the date of early diversification of the Pinaceae in the Early Cretaceous, 120–140 my ago (mya) (Florin, 1963; Miller, 1988). No test of rate constancy could be performed and the relationship between time and antigenic distance was assumed to be linear (Prager et al., 1976). The third approach was based on a molecular clock constructed using 32 *rbcL* sequences from both subgenera of *Pinus* (Wang et al., 1999) and *rbcL* sequences determined herein for *Picea abies*, *P. glauca*, and *P. mariana*. To construct the molecular clock, the average pairwise number of substitutions was estimated from all pairwise sequence comparisons between *Pinus* and *Picea* using both synonymous and

nonsynonymous substitutions. The same landmark as before (120–140 mya) was used to calibrate the clock (Fig. 2.1). Before estimating the *rbcL* clock, lineage relative rate tests (Li and Bousquet, 1992) were used to assess rate homogeneity between sequences of *Pinus* and *Picea*, using as outgroups *rbcL* sequences from three sister taxa of the Pinaceae: *Cupressus corneyana* (Cupressaceae), *Podocarpus gracilior* (Podocarpaceae), and *Taxus baccata* (Taxaceae) (Wang et al., 1999). The lineage relative rate test follows approximately the standardized normal distribution. The tests were conducted with an application developed by J. Laroche (Centre for Bioinformatics, Univ. Laval). The *rbcL* rates were corrected for multiple substitutions following the two-parameter method of Kimura (1980) because divergence rates ranged from 5 to 10% between conifer families.

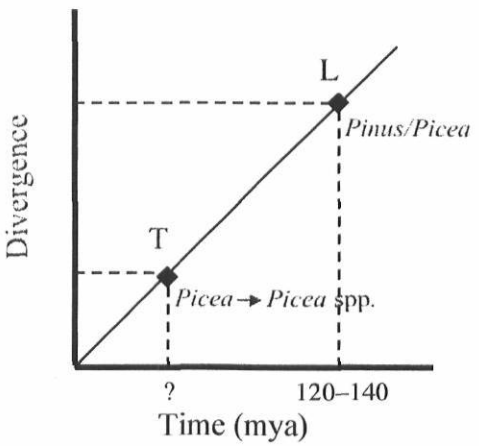

Fig. 2.1. Estimating the divergence time between *Picea abies*, *P. glauca*, and *P. mariana* (node T). A first estimate was directly obtained from analysis of the fossil record, and two other estimates were obtained from molecular clocks (antigenic distances and *rbcL* sequences; see Materials and Methods), which were both calibrated using the landmark L, the divergence time between the genera *Pinus* and *Picea*, as derived from analysis of the fossil record. Date estimates at node T were obtained by linear intrapolation for each of the antigenic distances and *rbcL* clock. For *rbcL*, relative rate tests were conducted, and no significant rate heterogeneity was detected (see Results). Figure not drawn to scale.

2.5 - Results

2.5.1 - Locus *Sb16*

This locus harbored the lowest haplotype/allelic diversity with eight distinct haplotypes/alleles observed among 23 haploid megagametophytes sampled (Table 2.1). Of 845 positions sampled, 28 sites were variable (Table 2.1). Neither substitutions nor indel polymorphisms were observed in *P. abies* in spite of repetitive sampling (Table 2.2, Table 2.3). *Picea glauca* was the most diverse species for this gene, in terms of the number of alleles detected and number of polymorphic sites observed along the gene (Table 2.2). The largest number of pairwise differences per site (0.0198) was noted between alleles 847-*P. glauca* and 802-*P. mariana*2, and the smallest number (0.0025) was observed between the same size alleles 804-*P. abies*, 804-*P. glauca*1, and 804-*P. glauca*2 from different species (see Table 2.3). Except for *P. abies*, the numbers of pairwise differences per site within species (values d_w, Table 2.2) were in the same range as those observed between species (values d_b, Table 2.4). The net divergence values between species (Table 2.4) were smaller, on average, than within-species divergence values (Table 2.2), indicating that allele divergence would precede species divergence, on average. Trans-species shared polymorphisms were observed at several sites (sites S, Table 2.3), which does not support the hypothesis of monophyly of *P. glauca* alleles.

Table 2.1. Parameters derived from the sequence alignment of alleles from three distantly related *Picea* species at three nuclear gene loci (excluding outgroup sequences from other Pinaceae)

Locus	GC (%)	Length of consensus sequence (bp)	Haploid sampling size	Number of haplotypes/ alleles detected	Number of indels detected	Indel length variation (bp)	Number of substitution sites	Number of variable sites[a]	Number of parsimony informative sites
Sb16	50	845	23	8	5	2 to 43	23	28	15
Sb29	43	518	19	11	6	1 to 21	22	28	14
Sb62	47	650	20	12	8	1 to 33	21	29	10
Total	-	-	62	31	19	-	66	85	39

[a] Including indels and substitution sites.

A total of two equally parsimonious gene topologies of 30 steps were found, with alleles from *P. glauca* not regrouped in a monophyletic assemblage (Fig. 2.2A). The general pattern was one of *P. glauca* alleles coalescing before the divergence of lineages leading to extant species. The neighbor-joining tree was similar (Fig. 2.2B). Constraining the tree topology to render *P. glauca* monophyletic resulted in a statistically significant increase ($P < 0.01$) of nine steps in tree length, as compared to the minimum trees. Tests of intragenic recombination among *Picea* alleles were positive for two pairs of sites (Table 2.3). Both cases involved a trans-species shared polymorphism at site 734. Thus, intragenic recombination might have contributed to trans-species shared polymorphism in at least one case. As in the Methods, allele 802-*P. mariana*1 was identified as a potential recombinant, likely from ancient recombination preceding species divergence. However, excluding this sequence from the phylogenetic analyses did not change the basic pattern of *P. glauca* alleles coalescing before the divergence of lineages leading to extant species.

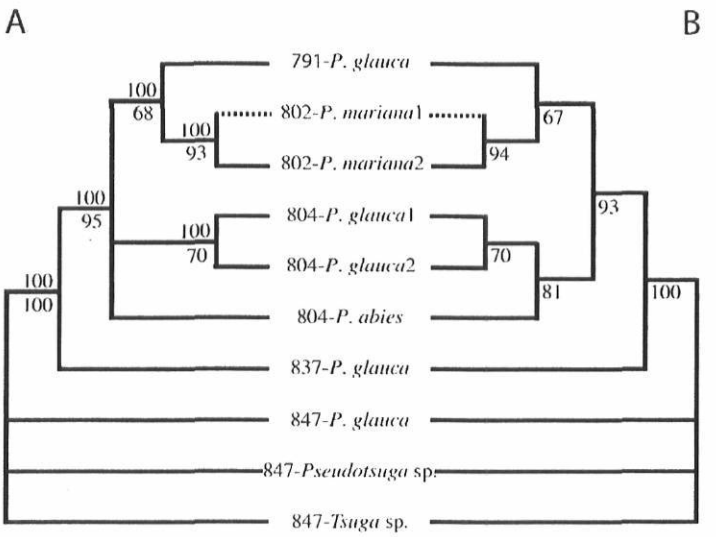

Fig. 2.2. Phylogeny of spruce alleles at locus *Sb16* (60S ribosomal protein L13A). (A) majority-rule consensus tree obtained by parsimony. (B) Tree obtained by the neighbor-joining method. Only nodes minimally supported by 50% bootstrap estimates are shown. Numbers above the branches indicate majority-rule consensus, and numbers below the branches, bootstrap estimates from 500 replicates. Dashed lines indicate putative recombinant alleles.

Table 2.2. Intraspecific diversity parameters for three nuclear gene loci in three distantly related *Picea* species.

Species	*Sb16*						*Sb29*						*Sb62*					
	n	A	S_s	S_{is}	π	d_w	n	A	S_s	S_{is}	π	d_w	n	A	S_s	S_{is}	π	d_w
P. abies	7	1	0	0	0	0	5	3	10	13	0.0134	0.0173	5	3	5	9	0.0064	0.0054
P. glauca	11	5	16	20	0.0092	0.0109	8	4	10	12	0.0103	0.0126	7	3	2	5	0.0023	0.0058
P. mariana	5	2	6	6	0.0075	0.0075	6	4	7	9	0.0075	0.0093	8	6	15	18	0.0094	0.0118

Note: n, haploid sampling size; A, number of alleles detected; S_s, number of substitution sites; S_{is}, number of polymorphic sites (indels + substitutions); π, nucleotide diversity per site within species; d_w, pairwise divergence per site within species (indels + substitutions).

Table 2.3. Variable sites for the locus *Sb16* (60S ribosomal protein L13A) in *Picea* spp.

		Variable positions[a]		
		Intron-1	Intron-2	Exon-3
Alleles detected (bp length – species)	Number of sequences sampled	01112222333333334 31692679023444690 27540386192015399	555666 345003 959236	77777 03478 14065
Picea				
791-*P. glauca*	4	C1CA1CTGAG0GGTAT1	T0GCAT	ACGAG
802-*P. mariana*1	3	T1CA0CTGAG1GGTAT0	A0GCAA	ATGAA
802-*P. mariana*2	2	T1CT0CTGTG1GTTAT0	A0GCAA	GCGAG
804-*P. glauca*1	1	C1CA0CCGAG0AGTAC0	T0GCAA	ACGGG
804-*P. glauca*2	2	C1CA0CTGAG0AGTAC0	T0GCGA	ACGGG
804-*P. abies*	7	C1CA0CTGAG0GGTAC0	T0GCGA	ACGAG
837-*P. glauca*	1	C0CA0CTAAG0GGTAC0	T1AAAA	ATAAG
847-*P. glauca*	3	C0TA0TTAAA0GGGGC0	T0GAAA	ACGAG
Putative recombination[b]		-R----------R--- ----------R-----R---		
Outgroups				
847-*Pseudotsuga* sp.	2	C0TA0TTAAA0GGGGC0	T0GAAA	ACGAG
847-*Tsuga* sp.	1	C0TA0TTAAA0GGGGC0	T0GAAA	ACGAG
Types of polymorphism[c]		FSSIISISISFIISSSI	FIISSI	ISIII

[a]Indels are coded as binary characters (0, 1).

[b]Putative recombinant sites are identified in pairs (R--R).

[c]Types of polymorphism: S, trans-species shared polymorphisms between *Picea* spp., including derived polymorphisms shared by more than one species; F, reciprocally fixed interspecific polymorphisms between *Picea* spp.; I, intraspecific derived polymorphisms (apomorphies). The derived or ancestral nature of each state at a polymorphic site was determined by comparison with outgroup sequences from other genera of the Pinaceae (see Methods for more details).

Table 2.4. Pairwise divergence per site and net pairwise divergence per site between species for each of three nuclear gene loci.

Locus	Pairwise divergence per site (indels and substitutions) between species (d_b)			Pairwise divergence per site (substitutions only) between species		
	abies-glauca	*abies-mariana*	*glauca-mariana*	*abies-glauca*	*abies-mariana*	*glauca-mariana*
Sb16	0.0068	0.0100	0.0140	0.0055	0.0087	0.0115
	(0.0013)	(0.0062)	(0.0048)	(0.0009)	(0.0049)	(0.0031)
Sb29	0.0179	0.0163	0.0148	0.0135	0.0121	0.0114
	(0.0029)	(0.0030)	(0.0039)	(0.0017)	(0.0016)	(0.0025)
Sb62	0.0078	0.0124	0.0106	0.0042	0.0074	0.0059
	(0.0022)	(0.0038)	(0.0018)	(0.0001)	(0.0005)	(0.0001)

Note: In parentheses, net pairwise divergence values corrected for within-species diversity (Nei, 1987).

2.5.2 - Locus *Sb29*

The haplotype/allelic diversity observed for *Sb29* was higher than that detected for *Sb16* (Table 2.1). The gene was also more variable, with a total of 28 variables sites of the 518 positions sampled (Table 2.1). Contrary to the trend observed for *Sb16*, *P. abies* appeared to be the most diverse species in terms of haplotype/allelic diversity, together with *P. glauca* (Table 2.2). The largest number of pairwise differences per site (0.0238) was between same-species alleles 498-*P. abies* and 511-*P. abies*, and the smallest number (0.0059) between alleles 503-*P. glauca* and 512-*P. glauca*1. The numbers of pairwise differences per site within species (values d_w, Table 2.2) were in the same range as those observed between species (values d_b, Table 2.4). As for *Sb16*, the net divergence values between species (Table 2.4) were smaller than within-species divergence values (Table 2.2), indicating that allele divergence would precede species divergence, on average. Trans-species shared polymorphisms were observed at several sites between haplotypes from different *Picea* taxa (sites S, Table 2.5). No sites with fixed interspecific differences favoring reciprocal species monophyly (sites F) were noted (Table 2.5).

A total of 20 most parsimonious trees of 38 steps was obtained. Alleles from *P. abies* and *P. glauca* did not form reciprocally monophyletic groups (Fig. 2.3A). The pattern was again one of allele coalescence preceding the divergence of lineages leading to extant species. The neighbor-joining tree was similar (Fig. 2.3B). Gene genealogies constraining alleles from the same species to monophyly (reciprocal monophyly) were four steps longer than the minimum trees, which was a statistically significant increase ($P < 0.01$). Tests of intragenic recombination within the genus *Picea* recovered five putative recombinant sites, of which three involved trans-species shared polymorphisms. Based on the Methods, two recombinant alleles were tentatively identified: 503-*P. glauca* and 511-*P. abies*. Both were likely the result of ancient recombination at or before species divergence. Removing these putative recombinants from the phylogenetic analyses did not alter the basic pattern of allele coalescence preceding the divergence of lineages leading to extant species.

Table 2.5. Variable sites for the locus *Sb29* (NAM protein fragment) in *Picea* spp.

Alleles detected (bp length – species)	Number of sequences sampled	Variable positions[a] Exon	Intron
		000000011111112233344444	4555
		013478904567784979900034	9111
		891033447172701881906888	3345
Picea			
490-*P. mariana*	1	GAG0TT1-TGA0CT0CTTAATATT	C00A
498-*P. abies*	1	GAG1TT0ATGC1CC0CTTAAAAGT	C10A
501-*P. abies*	2	AAG1TT0ATGC0CC0CTTAATAGT	C10A
502-*P. glauca*	1	GAG1TT0ATGC0CC0GCTAATTTC	C00A
503-*P. glauca*	2	GAG1TT0CTGC0CC0CTTGATTTT	T01C
511-*P. mariana*1	2	GAT0TT0CTGC0GT0CTTAATATT	C00A
511-*P. mariana*2	2	GAG0TT0CTAC0CC0CTGAATTTT	C00A
511-*P. abies*	2	GGG0GT0CCGC0CC0CCTAATTTC	C00A
512-*P. glauca*1	2	GAG0GT0CTGC0CC0CTTGATTTT	C01C
512-*P. glauca*2	3	GAG0TA0CTGC0CC0CTTGGTTTT	C01A
517-*P. mariana*	1	GAG0TT0CTGC0CT1CTTAATTTT	C00A
Putative recombination[b]		----R----------R----------	
		---R-------------------------R	
		---R-------------------------R-	
Outgroups			
512-*Pseudotsuga* sp.	2	GAG0TA0CTGC0CC0CTTGGTTTT	C00A
512-*Tsuga* sp.	1	GAG0TA0CTGC0CC0CTTGGTTTT	C00A
512-*Pinus* sp.	1	GAG0TA0CTGC0CC0CTTGGTTTT	C00A
Types of polymorphism[c]		IIISSSISIIIIIIIISISSISIS	IIII

[a]Indels are coded as binary characters (0, 1).

[b]Putative recombinant sites are identified in pairs (R--R).

[c]Types of polymorphism: S, trans-species shared polymorphisms between *Picea* spp., including derived polymorphisms shared by more than one species; F, reciprocally fixed interspecific polymorphisms between *Picea* spp.; I, intraspecific derived polymorphisms (apomorphies). The derived or ancestral nature of each state at a polymorphic site was determined by comparison with outgroup sequences from other genera of the Pinaceae (see Methods for more details).

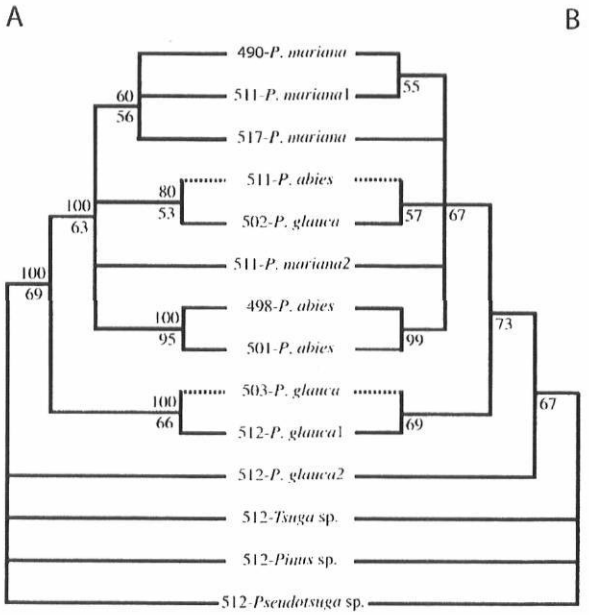

Fig. 2.3. Phylogeny of spruce alleles at locus *Sb29* (NAM protein fragment) in *Picea* spp. (A) Majority-rule consensus tree obtained by parsimony (B) Tree obtained by the neighbor-joining method. Only nodes minimally supported by 50% bootstrap estimates are shown. Numbers above the branches indicate majority-rule consensus, and numbers below the branches, bootstrap estimates from 500 replicates. Dashed lines indicate putative recombinant alleles.

2.5.3 - Locus *Sb62*

This locus harbored the highest haplotype/allelic diversity (Table 2.1). As for variation along the gene, of 650 positions sampled, a total of 29 variables sites were detected. Only 10 of these were informative, those with polymorphisms shared by at least two alleles (Table 2.1). *Picea mariana* was by far the most diverse species for this gene, in terms of number of alleles detected and number of polymorphic sites along the gene (Table 2.2). The largest number of pairwise differences per site (0.0231) was between alleles 556-*P. abies* and 578-*P. mariana*1, and the smallest number (0.0017) was between alleles 577-*P. glauca* and 578-*P. abies*, therefore from two different species. The numbers of pairwise differences per site within species (values d_w, Table 2.2) were in the same range as those observed between species (values d_b, Table 2.4). As for the two previous loci, the net divergence values between species (Table 2.4) were smaller than within-species divergence values (Table 2.2), again indicating that allele divergence would precede species

divergence, on average. Trans-species shared polymorphisms involving alleles from different *Picea* species were at several sites (sites S, Table 2.6). As for *Sb29*, no sites with fixed interspecific polymorphism favoring reciprocal species monophyly (sites F) were observed (Table 2.6).

Table 2.6. Variable sites for the locus *Sb62* (60S ribosomal protein L15) in *Picea* spp.

		Variable positions[a]		
		Exon-1	Intron	Exon-2
	Number	0	00011111111222222222333	445
Alleles detected	of sequences	2	78802356789144556788990 12	145
(bp length – species)	sampled	4	78977187234789136745345 78	777
Picea				
544-*P. abies*	1	G	A1--CA1ACA00GT0CATTGC0G00	CTA
553-*P. mariana*1	1	G	A0CTCA0ACA01AT0CATTGC1G00	CTT
553-*P. mariana*2	1	G	A0CTCA0ACA01GT0TATTGC1G00	CTA
556-*P. abies*	1	G	A0CTCC0GCA00GT1--TTCT0T01	CTA
561-*P. mariana*	1	G	A0CTCA0ACA11GT0CATTGC1G00	CTA
563-*P. mariana*	3	G	A0CTCA0ACA01GT0CATGGC0G00	CTA
575-*P. glauca*	1	G	A0CTCA0ACA00GT0CATTGC0G01	TTA
577-*P. glauca*	5	G	A0CTCA1ACA00GT0CATTGC0G00	CTA
578-*P. abies*	3	G	A0CTCA0ACA00GT0CATTGC0G00	CTA
578-*P. mariana*1	1	A	T0CTTA0ACT00GG0CGTTGC0G00	CCA
578-*P. mariana*2	1	G	A0TCCA0ACA00GG0CACTGC0A00	CTA
579-*P. glauca*	1	G	A0CTCA1ATA00GT0CATTGC0G10	CTA
Outgroups				
586-*Abies* sp.	1	G	A0TTCA0ACT10GG0CGTTGC0-00	CCA
584-*Tsuga* sp.	1	G	A0TTCA0ACT11GT0CGTTGC0C00	CCA
Types of polymorphism[b]		I	IISIIISIISS?I?IISIIIIIIIS	ISI

Note: No intragenic recombination was detected for this locus.
[a]Indels are coded as binary characters (0, 1). Dashes indicate missing sites due to a deletion. [b]Types of polymorphism: S, trans-species shared polymorphisms between *Picea* spp., including derived polymorphisms shared by more than one species; F, reciprocally fixed interspecific polymorphisms between *Picea* spp.; I, intraspecifically derived polymorphisms (apomorphies). The derived or ancestral nature of each state at a polymorphic site was determined by comparison with outgroup sequences from other genera of the Pinaceae (see Methods for more details). "?" indicates an ambiguous site due to polymorphism in the outgroup.

A single most parsimonious gene tree of 35 steps was obtained, with alleles from the three species coalescing before the divergence of lineages leading to extant species (Fig. 2.4A). The neighbor-joining tree was essentially the same (Fig. 2.4B). Gene genealogies constraining species to monophyly (reciprocal monophyly) were five steps longer than the minimum trees, which was a statistically significant increase ($P < 0.01$). No intragenic

recombination was detected for this locus among *Picea* sequences.

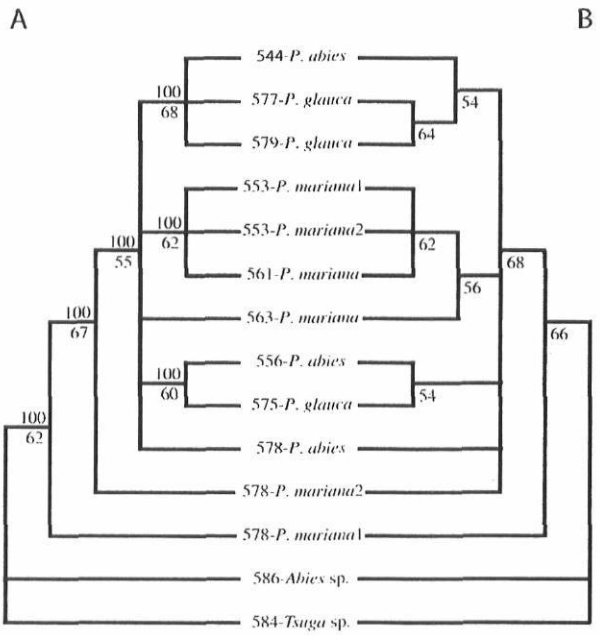

Fig. 2.4. Phylogeny of spruce alleles at locus *Sb62* (60S ribosomal protein L15). (A) Majority-rule consensus tree obtained by parsimony. (B) Tree obtained by the neighbor-joining method. Only nodes minimally supported by 50% bootstrap estimates are shown. Numbers above the branches indicate majority-rule consensus, and numbers below the branches are bootstrap estimates from 500 replicates.

2.5.4 - Trans-species shared polymorphisms

The number of polymorphisms per type (S, F, I) may vary between Tables 2.3, 2.5, and 2.6 when all taxa were considered simultaneously, and Table 2.7 where mutually exclusive pairwise taxa comparisons are considered. Trans-species shared polymorphisms (sites S in Tables 2.3, 2.5, 2.6) did not show any distribution pattern suggestive of maintenance by selection. These polymorphisms were mostly found in introns and involved substitutions as well as indels without a clear clustering pattern, and they were found in exons mainly for *Sb29*, at synonymous, nonsynonymous, and indel sites. Tajima's *D* test values (Tajima, 1989) as implemented in DnaSP (Rozas and Rozas, 1999) did not deviate significantly from neutral expectations for any of the three gene regions surveyed (D = −0.90 for *Sb16*, −1.01 for *Sb29* and −1.67 for *Sb62*, all $P > 0.05$).

Table 2.7. Types of polymorphisms detected between *Picea* species at each of three nuclear gene loci.

	Types of polymorphisms											
	abies-glauca				*abies-mariana*				*glauca-mariana*			
Locus	S	F	I_1	I_2	S	F	I_1	I_2	S	F	I_1	I_2
Sb16	9	0	0	11	0	5	0	6	10	3	10	5
Sb29	8	0	8	4	1	0	12	8	3	0	9	9
Sb62	2	0	7	3	5	0	9	13	5	0	5	13
Total	19	0	15	18	6	5	21	27	18	3	24	27

Note: S, number of sites with trans-species shared polymorphisms between *Picea* spp. (sites not supporting species monophyly), including derived polymorphisms shared by more than one species; F, number of sites with reciprocally fixed interspecific polymorphisms between *Picea* spp. (sites supporting species monophyly); I_1 and I_2, number of sites with intraspecific-derived polymorphisms (apomorphies) in first and second species, respectively. The derived or ancestral nature of each state at a polymorphic site was determined by comparison with outgroup sequences from other genera of the Pinaceae (see Methods for more details).

In pairwise comparisons of species (Table 2.7), trans-species shared polymorphisms were the most prevalent between *P. abies* and *P. glauca* with 19 instances. Also, these taxa had the fewest apomorphies and no fixed interspecific polymorphisms, suggesting a higher phylogenetic proximity between them. Locus-to-locus variation in the diverse types of polymorphism was high, reinforcing the need to span a larger number of nuclear loci or scanning cpDNA and mtDNA regions before reaching any definitive conclusion about their higher phylogenetic proximity.

Overall, *Sb16* contributed 11 sites to trans-species shared polymorphisms (Table 2.3), while *Sb29* contributed nine such sites (Table 2.5) and *Sb62* seven such sites (Table 2.6). If we were to remove putative recombinant sites involving trans-species shared polymorphisms for loci *Sb16* and *Sb29*, there would remain eight, six, and seven sites involving trans-species shared polymorphisms, respectively, for *Sb16*, *Sb29*, and *Sb62*.

Fixed interspecific polymorphisms are indicative of a separate lineage evolution. There were no fixed interspecific polymorphisms favoring species monophyly for loci *Sb29* and *Sb62*, while there were only three such sites for *Sb16*, all favoring the monophyly of *P. mariana* alleles.

2.5.5 - Divergence time between the three *Picea* species and mutation rate at nuclear gene loci

Because of uncertainty regarding the exact divergence history among the lineages leading to *Picea abies*, *P. glauca*, *and P. mariana*, a trichotomy was assumed for estimating their divergence time *T*. A first approximation of *T* is given by the fossil record for which the earliest fossils morphologically representative of extant taxa could be interpreted as evidence for the diversification of the genus in its major extant lineages. These fossils date back to the middle Miocene, around 15 mya, and more questionable fossils were reported from the early Miocene, back to 23 mya (Wolfe, 1964). While *P. abies*, *P. glauca*, and *P. mariana* represent morphologically divergent taxa (Wright, 1955; Weng and Jackson, 2000) with deep phylogenetic branching (Sigurgeirsson and Szmidt, 1993; Bouillé and Bousquet, unpublished data), other extant taxa might have branched out earlier and thus, this period should be considered as a lower bound estimate of the divergence between the lineages leading to extant taxa. This period is much latter than the first reliable occurrence of *Picea* fossils, that is, from the middle Eocene, around 45 mya (Axelrod, 1998; LePage, 2001).

A second estimate of divergence time *T* was obtained from antigenic distance data (Prager et al., 1976). With an average intergeneric distance of 3.5 units for the divergence between *Picea* and *Pinus* and a landmark divergence time of 120–140 mya between these two genera, the maximum distance of 0.5 unit reported between *Picea mariana*, *P. glauca*, and *P. abies* (Prager et al., 1976) converts to a divergence time *T* of 17–20 mya, assuming a linear relationship (Fig. 2.1).

A third estimate of divergence time *T* could be obtained from *rbcL* sequences, which behaved essentially as a molecular clock for the set of sequences sampled: lineage relative rate tests indicated that there was no rate heterogeneity between the *Picea* and *Pinus* lineages using three distinct sister groups to the Pinaceae (test value = 1.16 with outgroup *Cupressus*, test value = 1.92 with outgroup *Podocarpus*, test value = 0.97 with outgroup *Taxus*, all not significant at $P = 0.05$). With an average rate (K) of 0.0276 substitution per

site (synonymous and nonsynonymous) for the divergence between *Picea* and *Pinus* and a landmark divergence time of 120–140 mya (L), and by dividing K by $2L$, an overall rate of substitution per site per year of 0.99×10^{-10} to 1.15×10^{-10} was obtained. Applying these rates to half of the average pairwise rate of 0.0031 substitution per site estimated between *Picea abies*, *P. glauca*, and *P. mariana* translates to a divergence time T of 13 to 16 mya (Fig. 2.1).

An estimate of mutation rate μ for the three nuclear gene loci was obtained by calculating the average divergence per site between species (substitutions only) (Table 2.4) and then dividing by $2T$ where T was estimated between 13 to 20 mya, the range of possible values determined earlier for the split between *Picea abies*, *P. glauca*, and *P. mariana*. With an average substitution rate per site of 0.0089 between spruce species, μ was estimated as 2.23×10^{-10} to 3.42×10^{-10} per site per year. This mutation rate is used later for estimating parameters of historical demography.

2.6 - Discussion

2.6.1 - Diversity at nuclear gene loci within spruce taxa

Low levels of nucleotide diversity (Table 2.2, π values) were observed within spruce species. However, the values of nucleotide diversity were in the same range as those reported for synonymous sites of protein coding sequences in the conifer *Pinus sylvestris* L. and in *Arabidopsis* (Dvornyk et al., 2002). In contrast, haplotype/allelic diversity was quite high for the three spruce species studied at each of the three gene loci investigated, except for the locus *Sb16* in *Picea abies*, which was monomorphic. A similar trend of high haplotype diversity was observed for the gene *pal1* in the conifer *Pinus sylvestris*, which was mostly caused by synonymous substitutions (Dvornyk et al., 2002). Altogether, the evidence suggests that large effective population sizes are necessary to maintain such important cohorts of neutral or nearly neutral alleles.

There was no clear trend in the current study as to which of the three *Picea* species analyzed was the most diverse in terms of haplotype/allelic diversity: *P. glauca* was the most diverse for *Sb16* while *P. abies* was fixed, but *P. abies* and *P. glauca* were the most diverse for *Sb29*, and *P. mariana* was the most diverse for *Sb62*. When analyzing diploid genotypes for a dozen of expressed nuclear gene loci, average estimates of observed heterozygosity were similar among the three spruce species analyzed herein, but locus-to-locus variance was high (Perry and Bousquet, 1998a, b; Perry et al., 1999). Overall, our results suggest that the gene-to-gene variance is high and that more loci will need to be investigated in order to compare adequately the diversity residing in the various species. Such gene-to-gene variance is expected because of stochastic factors (Tajima, 1983; Arbogast et al., 2002; Rosenberg and Nordborg, 2002).

2.6.2 - Allele coalescence and historical population size

For each of the three nuclear gene loci analyzed, the average numbers of pairwise differences between alleles from different species (d_b values) were small and in the same range as those observed between alleles within species (d_w values). When correcting for within-species diversity, the resulting net divergence values between species were smaller than those observed within-species, indicating that, on average, allele divergence would precede species divergence. This trend is surprising, given a divergence time of at least 13 mya between these biological species. This unexpected trend is also supported by the analysis of gene sequences site by site in which no fixed interspecific polymorphisms were found for two of the three nuclear gene loci analyzed, *Sb29* and *Sb62*. These polymorphisms would support reciprocal species monophyly. On the other hand, several trans-species shared polymorphisms were observed for each of the three nuclear gene loci. The trend towards trans-species shared polymorphisms was confirmed by a phylogenetic analysis in which both distance and character-state approaches indicated that there were no instances where alleles from the various species were reciprocally monophyletic. Rather, alleles from the various species were intermingled for *Sb29* or *Sb62*. For *Sb16*, for which no allelic variation was observed for *P. abies* and the two alleles detected in *P. mariana*

were monophyletic, alleles of *P. glauca* were not monophyletic and many instances of trans-species shared polymorphisms were identified.

While introgressive hybridization represents a possible cause for the presence of shared polymorphisms among closely related taxa (e.g., Wang et al., 1997; Isoda and Shiraishi, 2001; Machado et al., 2002), and while natural hybridization has been reported between closely related spruce taxa (e.g., between *P. mariana* and *P. rubens* Sarg., Perron and Bousquet, 1997), more or less recent interspecific gene exchanges appear highly unlikely between the three spruce species analyzed herein. Contrary to closely related taxa in the genus, these species do not cross naturally, and they are hardly compatible with each other (Wright, 1955; Mikkola, 1969), which warrant their actual recognition as distinct biological species. In such a slowly evolving genus as *Picea*, achieving reproductive isolation is generally indicative of large divergence (Wright, 1955). These species are also divergent with respect to several morphological characters, which is notable for this rather morphologically uniform genus (Wright, 1955; Weng and Jackson, 2000). In phylogenies based on cpDNA, these species are also present in distinct lineages with deep phylogenetic branching in the genus (Sigurgeirsson and Szmidt, 1993; Bousquet and Bouillé unpublished data). Thus, it is highly unlikely that these biological species have exchanged genes in the recent past. If they had done so, cpDNA phylogenies would reveal high phylogenetic affinity between these taxa, because cpDNA is paternally transmitted by pollen in conifers (e.g., Stine et al., 1989).

Hence, under such conditions, trans-species shared polymorphisms are likely to be of shared ancestry. For all three loci, Tajima's test for selection was inconclusive and trans-specific shared polymorphisms did not show any distribution pattern suggestive of maintenance by selection. Many shared polymorphisms were observed in introns, and there was an excess of shared polymorphisms over fixed polymorphisms for the three loci investigated (Table 2.7), which is not suggestive of locus-specific selective effects (Wang et al., 1997). Even if neutral tests are known for their lack of power, it seems reasonable to assume, given the overall evidence at hand, that the observed polymorphisms are neutral or

nearly neutral and that allele coalescence time would be governed largely by demographic factors.

Historical population size (N_e) was estimated by solving the equation $\pi = 4N_e\mu$ (Tajima, 1983), where μ is the mutation rate per generation for the nuclear gene loci examined. The mutation rate μ was estimated earlier as 2.23×10^{-10} to 3.42×10^{-10} per site per year. This estimate is in the same range as the estimated rate of synonymous substitutions per site per year obtained from comparing spruce and pine nuclear phytochrome *PHYO* gene sequences ($\mu = 4.8 \times 10^{-10}$, Garcia-Gil et al., 2003). By assuming 50 years as the average generation time in these species, which is about three to five times the age at first reproduction but which is less than the maximum life expectancy in these species (Burns and Honkala, 1990), μ becomes 1.11 to 1.71×10^{-8} per site per generation. Then, by using average π values deduced from Table 2.2 for each species, N_e estimates from 96 000 to 182 000 were obtained by applying the formula above, depending of the species (Table 2.8). The N_e estimates are large and well above estimates for animal species for which numbers in the range of 10 000 to 50 000 have been estimated (e.g., Sherry et al., 1997; Hare et al., 2002). Even if N_e estimates were underestimated or overestimated due to a number of factors, including the value of generation time used and the fact that trees have overlapping generations (Caballero, 1994), the order of magnitude of the estimates obtained appears robust. More precise estimates for each species would require better estimates of generation time that would take into account temporal and spatial heterogeneity as well as generation overlap. As well, they would require allele frequencies estimated from several more loci in order to stabilize the gene-to-gene variance in nucleotide diversity (Tajima, 1983; Arbogast et al., 2002).

The average coalescence time of two randomly picked alleles at the same gene locus is estimated by $2N_e$ generations (Tajima, 1983). Hence, the previously estimated values of N_e apply to a $2N_e$ period of around 200 000 to 360 000 generations or 10 to 18 mya, assuming a generation time of 50 years (Table 2.8). These estimates are in the same range as those of species divergence time (13 to 20 mya). They indicate the ancient nature of the

allelic polymorphisms observed in these species and the sharing of neutral or nearly neutral variation since common ancestry. These estimates are in agreement with the coalescence of many alleles before species divergence on the gene trees and the lack of reciprocal monophyly between alleles from different species.

Table 2.8. Estimates of historical population size and allele coalescence time from data at three nuclear gene loci in three distantly related *Picea* species.

Species	$\pi_{AVERAGE}$	N_e [a]	$2N_e$ (generations)	$2N_e$ (mya)
P. abies	0.0066	96 491–148 649	192 982–297 298	9.6–14.9
P. glauca	0.0073	106 725–164 414	213 450–328 828	10.7–16.4
P. mariana	0.0081	118 421–182 432	236 842–364 864	11.8–18.2

Note: $\pi_{AVERAGE}$, the average nucleotide diversity per site from three nuclear gene loci for each species (estimated from Table 2.2); N_e, the estimated historical population size; $2N_e$, the estimated coalescence time of two randomly picked alleles, in number of generations and in millions of years ago (mya), assuming a generation time of 50 years.
[a] N_e was derived from the formula $\pi = 4N_e\mu$, where μ, the mutation rate per site per generation (substitutions only) was obtained as $\mu = K/2T$, where K is the average of pairwise divergence values per site between species (substitutions only, Table 2.3) and T is the estimated divergence time between species (13 to 20 mya, see text): $\mu = 2.23 \times 10^{-10}$ to 3.42×10^{-10} substitution/site/year and $\mu = 1.11 \times 10^{-8}$ to 1.71×10^{-8} substitution/site/generation assuming a generation time of 50 years.

2.6.3 - Implications for the long-term maintenance of genetic diversity in trees

The order of magnitude of these values of historical population size (N_e) reconciles well with much lower estimates obtained at the time of speciation for *Picea rubens*, a recent derivative of *P. mariana* during the Pleistocene, which harbors reduced gene diversity at expressed gene loci (in terms of observed heterozygosity, proportion of polymorphic loci and number of alleles, Perron et al., 2000) compared to that of *P. abies* (Perry et al., 1999), *P. glauca* (Perry and Bousquet, 1998b), or *P. mariana* (Perry and Bousquet, 1998a; Perron et al., 2000). To account for the lower genetic diversity observed in *P. rubens*, a mild bottleneck of N_e in the range of 10 000 was estimated from genetic drift simulations (Perron et al., 2000). Hence, the estimates of historical population size in this study, which are more than an order of magnitude higher than those inferred for the genetically depauperate *P. rubens*, support the idea that large historical population sizes in

P. abies, *P. glauca*, and *P. mariana* led at least partly to the long-term maintenance of higher levels of gene diversity. Indeed, these three species do not seem to have suffered much from the Pleistocene glaciations, even if they were displaced repeatedly (e.g., Davis, 1983).

Such large historical population sizes also reconcile well with our knowledge of the ecological determinants and reproductive biology of these species. Spruces harbor archaic wind-pollinated mating systems with high outcrossing rates (e.g., Perry and Bousquet, 2001). It is unlikely that these traits have changed during their history. These taxa are also abundant and occupy a dominant position in their ecosystems (Wright, 1955). All these factors contribute to maintaining large effective population sizes. Such population sizes and extensive gene flow are also suggested by the low levels of population differentiation in spruces, as measured by nuclear markers of a molecular or biochemical nature (Boyle and Morgenstern, 1987; Furnier and Stine, 1991; Isabel et al., 1995; Müller-Starck, 1995; Jaramillo-Correa et al., 2001; Perry and Bousquet, 2001; Collignon et al., 2002; Gamache et al., 2003). Exceptions to this rule exist in which spruce taxa with more scattered populations and restricted ranges show more population differentiation and overall smaller genetic diversity as a result of more or less recent bottlenecks, such as for *P. rubens* (Hawley and DeHayes, 1994; Perron et al., 2000; Rajora et al., 2000) or for *P. martinezii* T. F. Patterson (Ledig et al., 2000).

These results bear serious implications for our understanding of the maintenance of neutral or nearly neutral genetic diversity at nuclear genes in spruces and, more generally, in conifers and other tree species harboring ecological and reproductive strategies promoting large population sizes. On average, one would expect large population sizes to favor the maintenance of high levels of neutral or nearly neutral genetic diversity. Under such conditions, novel alleles would be less likely to be lost through drift and their time to fixation would be longer (Kimura and Ohta, 1969).

If effective, this process would lead to the disproportionate coexistence of neutral or nearly neutral polymorphisms at nuclear gene loci and high levels of haplotype/allelic diversity, such as observed in this study and in previous studies (Perry and Bousquet, 1998a, b; Perry et al., 1999). The comparative analysis of allozyme diversity among plant taxa with contrasting life-history and ecological determinants appears to support such a neutral/demographic process (Hamrick et al., 1992; Hamrick and Godt, 1996). On average, taxa with higher levels of heterozygosity and higher number of alleles per locus are characterized by ecological, population, and demographic features conferring large population sizes, such as outcrossing, wind pollination or dominant position in the ecosystem. The three spruce species analyzed in this study are good examples, with high levels of diversity at the population level (H_e typically in the range of 0.20–0.30 and above) for allozymes (Boyle and Morgenstern, 1987; Bergmann and Ruetz, 1991; Furnier et al., 1991; Hamrick et al., 1992; Isabel et al., 1995; Müller-Starck 1995; Jaramillo-Correa et al., 2001). Such a link between population determinants and diversity at allozyme loci is not unexpected, because of the essentially neutral or nearly neutral nature of allozyme polymorphisms (e.g., in black spruce, Isabel et al., 1995; in white spruce, Jaramillo-Correa et al., 2001). Recently, this diversity trend has further been confirmed at the DNA level, for which the highest levels of within-population gene diversity for RAPD and microsatellite markers were detected in outcrossing plant taxa and long-lived perennials, which include tree species (Nybom 2004).

2.6.4 - Implications for estimating species phylogenies from nuclear genes

The detection of trans-species shared polymorphisms at orthologous gene loci calls for caution in estimating congeneric species phylogenies from nuclear genes in plants with life-history and reproductive determinants favoring large effective population sizes, even when gene orthologues can be unambiguously distinguished from putative paralogs. In this study, well-characterized regions of the nuclear genome were used and paralogous sequences were avoided: primers were designed to be specific to a single gene region (Perry and Bousquet, 1998a), the mendelian segregation of allelic variants had been previously documented (Perry and Bousquet, 1998a, b; Perry et al., 1999), and orthology

was further validated by sequencing alleles from haploid tissues. Even so, allele orthology did not appear to be a sufficient criterion to guarantee a gene tree matching a species tree. The detection of several trans-species shared polymorphisms at the three gene loci investigated in this study warns about potential pitfalls in estimating species phylogeny from nuclear genes in such taxa. It emphasizes the need for data from many nuclear loci (Arbogast et al., 2002; Rosenberg and Nordborg, 2002). Ideally, validation should be sought with phylogenies derived from chloroplast and mitochondrial genes for which allele coalescence time and the frequency of trans-species shared polymorphisms would be reduced as compared to nuclear genes. This is because of the uniparental transmission of organellar genomes for which the average coalescence time of two randomly picked cpDNA or mtDNA alleles is not $2N_e$ but N_e generations. Such recommendations of caution appear even more appropriate when little is known about the demography and the history of a species or group of species.

Trans-species shared polymorphisms of two different types were detected. The ones shared by ancestry were the most abundant. But for two of the three nuclear genes analyzed, a few trans-species shared polymorphisms were detected, which were likely the result of ancient intragenic recombination. The genome-wide frequency of intragenic recombination is unknown for plant nuclear genes. But intragenic recombination in itself constitutes another factor that can lead to biased phylogenies (Shierup and Hein, 2000; Posada and Crandall, 2002; Rosenberg and Nordborg, 2002). In our study, this factor was taken into account by estimating gene trees with and without the recombinant alleles.

The presence of trans-species shared polymorphisms also calls for caution in interpreting gene genealogies. In plant and tree species suspected of large historical population sizes, nonmonophyletic patterns of allelic variation should persist for tens or hundreds of thousands generations. In such cases, it has been pointed out that extreme caution should be exercised when inferring reticulate evolution from polyphyletic patterns at nuclear loci (Hare et al., 2002). Such patterns could also be caused by the long-term maintenance of trans-species polymorphisms of shared ancestry. More generally, our study

emphasizes the limitations of single or oligo-gene genealogies at the nuclear level to circumscribe useful taxonomical or ecological units when large historical population sizes are suspected.

2.7 - Acknowledgments

We thank S. Senneville for support in the laboratory, J. Laroche (Bioinformatics Centre, Univ. Laval) for relative rate tests, B. Lepage (Univ. Pennsylvania) for discussions about the spruce fossil record, D. J. Perry (Canadian Grain Commission) for fruitful discussions at the onset of the study, and O. Savolainen (Univ. of Oulu), J. Hey (Rutgers Univ.), and two anonymous reviewers for constructive comments on previous drafts of this manuscript. This study was supported by a discovery grant from the National Science and Engineering Research Council of Canada to J.B.

2.8 - References

Arbogast, B. S., S. V. Edwards, J. Wakeley, P. Beerli, and J. B. Slowinski. 2002. Estimating divergence times from molecular data on phylogenetic and population genetic timescales. *Annual Review of Ecology and Systematics* 33: 707–740.

Axelrod, D. I. 1998. The Eocene Thunder Mountain flora of central Idaho. *University of California Publications in Geological Sciences* 142: 1–61.

Bergmann, F., and W. Ruetz. 1991. Isozyme variation and heterozygosity in random tree samples and selected orchard clones from the same Norway spruce populations. *Forest Ecology and Management* 46: 39–47.

Boyle, T. J. B., and E. K. Morgenstern. 1987. Some aspects of the population structure of black spruce in central New Brunswick. *Silvae Genetica* 36: 53–60.

Burns, R. M., and B. H. Honkala. 1990. Silvics of North America, vol. 1, Conifers. Agriculture Handbook 654. US Department of Agriculture, Forest Service, Washington, D.C., USA

Caballero, A. 1994. Developments in the prediction of effective population size. *Heredity* 73: 657–679.

Collignon, A.-M., H. Van de Sype, and J.-M. Favre. 2002. Geographical variation in random amplified DNA and quantitative traits in Norway spruce. *Canadian Journal of Forest Research* 32: 266–282.

Davis, M. B. 1983. Quaternary history of deciduous forests of eastern North America and Europe. *Annals of the Missouri Botanical Garden* 70: 550–563.

Dvornyk, V., A. Sirviö, M. Mikkonen, and O. Savolainen. 2002. Low nucleotide diversity at the *pal1* locus in the widely distributed *Pinus sylvestris*. *Molecular Biology and Evolution* 19: 179–188.

Faith, D.P. 1991. Cladistic permutation tests for monophyly and nonmonophyly. *Systematic Zoology* 40: 366–375.

Florin, R. 1963. The distribution of conifer and taxad genera in time and space. *Acta Horti Bergiani* 20: 121–312.

Furnier, G. R., M. Stine, C. A. Mohn, and M. A. Clyde. 1991. Geographic patterns of variation in allozymes and height growth in white spruce. *Canadian Journal of Forest Research* 21: 707–712.

Gamache, I., J. P. Jaramillo-Corea, S. Payette, and J. Bousquet. 2003. Diverging patterns of mitochondrial and nuclear DNA diversity in subarctic black spruce: imprint of a founder effect associated with postglacial colonization. *Molecular Ecology* 12: 891–901.

Garcia-Gil, M. R., M. Mikkonen, and O. Savolainen. 2003. Nucleotide diversity at two phytochrome loci along a latitudinal cline in *Pinus sylvestris*. *Molecular Ecology* 12: 1195–1206.

Hamrick, J. L., M. J. W. Godt, and S. L. Sherman-Broyles. 1992. Factors influencing levels of genetic diversity in woody plant species. *New Forests* 6: 95–124.

Hamrick, J. L, and M. J. W. Godt. 1996. Effects of life history traits on genetic diversity in plant species. *Philosophical Transactions of the Royal Society of London, B, Biological Sciences* 351: 1291–1298.

Hare, M. P., F. Cipriano, and S. R. Palumbi. 2002. Genetic evidence on the demography of speciation in allopatric dolphin species. *Evolution* 56: 804–816.

Hawley, G. J., and D. H. DeHayes. 1994. Genetic diversity and population structure of red spruce (*Picea rubens*). *Canadian Journal of Botany* 72: 1778–1786.

Hudson, R. R. 1990. Gene genealogies and the coalescent process. *In* D. Gutuyama and J. Antonivics [eds.], *Oxford Surveys in Evolutionary Biology*, vol. 7, 1–44. Oxford University Press, Oxford, UK.

Hudson, R. R., and N. L. Kaplan. 1985. Statistical properties of the number of recombination events in the history of a sample of DNA samples. *Genetics* 111: 147–164.

Ioerger, T. R., A. G. Clark, and T.-H. Kao. 1990. Polymorphism at the self-incompatibility locus in Solanaceae predates speciation. *Proceedings of the National Academy of Sciences, USA* 87: 9732–9735.

Isabel, N., J. Beaulieu, and J. Bousquet. 1995. Complete congruence between gene diversity estimates derived from genotypic data at enzyme and random amplified polymorphic DNA loci in black spruce. *Proceedings of the National Academy of Sciences, USA* 92: 6369–6373.

Isoda, K., and S. Shiraishi. 2001. Allelic sequence polymorphisms in the intron region of the nuclear-encoded *Gap*C gene preceded the speciation of three closely related *Abies* species (Pinaceae). *Theoretical and Applied Genetics* 102: 244–250.

Jaramillo-Correa, J. P., J. Beaulieu, and J. Bousquet. 2001. Contrasting evolutionary forces driving population structure at expressed sequence tag polymorphisms, allozymes and quantitative traits in white spruce. *Molecular Ecology* 10: 2729–2740.

Jolley, K. A., E. J. Feil, M.-S. Chan, and M. C. J. Maiden. 2001. Sequence type analysis and recombinational tests (START). *Bioinformatics* 17: 1230–1231.

Kimura, M., and T. Ohta. 1969. The average number of generations until fixation of a mutant gene in a finite population. *Genetics* 61: 763–771.

Kimura, M. 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution* 16: 111–120.

Ledig, F. T. 1998. Genetic variation in *Pinus*. *In* D.M. Richardson [ed.], Ecology and biogeography of *Pinus*, 251–280. Cambridge University Press, Cambridge, UK.

Ledig, F. T., B. Bermejo-Velazquez, P. D. Hodgskiss, D. R. Johnson, C. Flores-Lopez, and V. Jacob-Cervantes. 2000. The mating system and genus diversity in Martinez spruce, an extremely rare endemic of Mexico's Sierra Madre Oriental: an example of facultative selfing and survival in interglacial refugia. *Canadian Journal of Forest Research* 30: 1156–1164.

LePage, B. A. 2001. New species of *Picea* A. Dietrich (Pinaceae) from the middle Eocene of Axel Heiberg Island, Arctic Canada. *Botanical Journal of the Linnean Society* 135: 137–167.

Li, P., and J. Bousquet. 1992. Relative-rate test for nucleotide substitutions between two lineages. *Molecular Biology and Evolution* 9: 1185–1189.

Machado, C. A., R. M. Kliman, J. A. Markert, and J. Hey. 2002. Inferring the history of speciation from multilocus DNA sequence data: the case of *Drosophila pseudoobscura* and close relatives. *Molecular Biology and Evolution* 19: 472–488.

Magallón, S., and M. J. Sanderson. 2002. Relationships among seed plants inferred from highly conserved genes: sorting conflicting phylogenetic signals among ancient lineages. *American Journal of Botany* 89: 1991–2006.

Maynard Smith, J., and N. H. Smith. 1998. Detecting recombination from gene trees. *Molecular Biology and Evolution* 15: 590–599.

Mikkola, L. 1969. Observations on interspecific sterility in *Picea*. *Annales Botanici Fennici* 6: 285–339.

Miller, C. N. 1988. The origin of modern conifer families. *In* C. B. Beck [ed.]. Origin and evolution of the gymnosperms, 448–486. Columbia University Press, New York, USA.

Müller-Starck, G. 1995. Genetic variation in high elevated populations of Norway spruce (*Picea abies* (L.) Karst.) in Switzerland. *Silvae Genetica* 44: 356–362.

Nei, M. 1987. Molecular evolutionary genetics. Columbia University Press, New York, USA.

Nybom, H. 2004. Comparison of different nuclear DNA markers for estimating intraspecific genetic diversity in plants. *Molecular Ecology* 13: 1143–1155.

Perron, M., and J. Bousquet. 1997. Natural hybridization between black spruce and red spruce. *Molecular Ecology* 6: 725–734.

Perron, M., D. J. Perry, C. Andalo, and J. Bousquet. 2000. Evidence from sequence-tagged-site markers of a recent progenitor-derivative species pair in conifers. *Proceedings of the National Academy of Sciences, USA* 97: 11331–11336.

Perry, D. J., and J. Bousquet. 1998a. Sequence-tagged-site (STS) markers of arbitrary genes: development, characterization and analysis of linkage in black spruce. *Genetics* 149: 1089–1098.

Perry, D. J., and J. Bousquet. 1998b. Sequence-tagged-site (STS) markers of arbitrary genes: the utility of black spruce-derived STS primers in other conifers. *Theoretical and Applied Genetics* 97: 735–743.

Perry, D. J., and J. Bousquet. 2001. Genetic diversity and mating system of post-fire and post-harvest black spruce: An investigation using codominant sequence-tagged-site (STS) markers. *Canadian Journal of Forest Research* 31: 32–40.

Perry, D. J., N. Isabel, and J. Bousquet. 1999. Sequence-tagged-site (STS) markers of arbitrary genes: the amount and nature of variation revealed in Norway spruce. *Heredity* 83: 239–248.

Posada, D., and K. A. Crandall. 2002. The effect of recombination on the accuracy of phylogeny estimation. *Journal of Molecular Evolution* 54: 396–402.

Prager, E. M., D. P. Fowler, and A. C. Wilson. 1976. Rates of evolution in conifers (Pinaceae). *Evolution* 30: 637–649.

Rajora, O. P., A. Mosseler, and J. E. Major. 2000. Indicators of population viability in red spruce, *Picea rubens*. II. Genetic diversity, population structure, and mating behavior. *Canadian Journal of Botany* 78: 941–956.

Rosenberg, N. A. 2003. The shapes of neutral gene genealogies in two species: probabilities of monophyly, paraphyly, and polyphyly in a coalescent model. *Evolution* 57: 1465–1477.

Rosenberg, N. A., and M. Nordborg. 2002. Genealogical trees, coalescent theory and the analysis of genetic polymorphisms. *Nature Review Genetics* 3: 380–390.

Rozas, J., and R. Rozas. 1999. DnaSP version 3: an integrated program for molecular population genetics and molecular evolution analysis. *Bioinformatics* 15: 174–175.

Schierup, M. H., and J. Hein. 2000. Consequences of recombination on traditional phylogenetic analysis. *Genetics* 156: 879–891.

Sherry, S. T., H. C. Harpending, M. A. Batzer, and M. Stoneking. 1997. *Alu* evolution in human populations: using the coalescent to estimate effective population size. *Genetics* 147: 1977–1982.

Sigurgeirsson, A., and A. E. Szmidt. 1993. Phylogenetic and biogeographic implications of chloroplast DNA variation in *Picea*. *Nordic Journal of Botany* 13: 233–246.

Stine, M., B. B. Sears, and D. E. Keathley. 1989. Inheritance of plastids in interspecific hybrids of blue spruce and white spruce. *Theoretical and Applied Genetics* 78: 768–774.

Swofford, D. L. 2002. PAUP*. Phylogenetic analysis using parsimony (*and other methods), version 4, Sinauer Associates, Sunderland, Massachusetts, USA.

Tajima, F. 1983. Evolutionary relationship of DNA sequences in finite populations. *Genetics* 105: 437–460.

Tajima, F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123: 585–595.

Wang, R. L., J. Wakeley, and J. Hey. 1997. Gene flow and natural selection in the origin of *Drosophila pseudoobscura* and close relatives. *Genetics* 147: 1091–1106.

Wang, X.-R., Y. Tsumura, H. Yoshimaru, K. Nagasaka, and A. E. Szmidt. 1999. Phylogenetic relationships of Eurasian pines (*Pinus*, Pinaceae) based on chloroplast *rbcL*, *matK*, *rpl20-rps18* spacer, and *trnV* intron sequences. *American Journal of Botany* 86: 1742–1753.

Weng, C., and S. T. Jackson. 2000. Species differentiation of North American spruce (*Picea*) based on morphological and anatomical characteristics of needles. *Canadian Journal of Botany* 78: 1367–1383.

Wolfe, J. A. 1964. Miocene floras from Fingerrrock Wash, southwestern Nevada. US Geological Survey Professional Paper 454-N: 1–36.

Wright, J. W. 1955. Species crossability in spruce in relation to distribution and taxonomy. *Forest Science* 1: 319–349.

Chapitre 3

Incompatibilité des phylogénies chloroplastiques et mitochondriale chez *Picea* : la spéciation géographique et la réticulation comme mécanismes de l'évolution à long terme chez les genres de Pinaceae

Ce chapitre a été soumis sous forme d'article scientifique et, si accepté, pourra être retrouvé sous la référence « Bouillé M., S. Senneville, and J. Bousquet. 2007. Discordant mtDNA and cpDNA spruce phylogenies indicate geographic speciation and reticulation as driving factors for the long-term evolution in the Pinaceae. » Les alignements des séquences de ce chapitre sont présentés à l'appendice 2.

3.1 - Résumé

La phylogénie des *Picea* a été estimée suite au séquençage de trois régions du génome chloroplastique, transmis paternellement, ainsi qu'une région du génome mitochondrial, transmis maternellement. Des différences significatives ont été observées entre les phylogénies chloroplastiques, et entre les phylogénies chloroplastiques et la phylogénie mitochondriale. Aucune des phylogénies ne correspondait aux classifications actuelles. La phylogénie mitochondriale était géographiquement plus structurée que les phylogénies chloroplastiques. Sur l'arbre mitochondrial, la plupart des taxons nord-américains formaient un groupe monophylétique. La cohérence géographique des divers regroupements sur cet arbre implique que des phénomènes de spéciation de nature géographique se soient produits. Les incompatibilités entre les phylogénies d'organelles suggéraient l'occurrence de transferts latéraux anciens du génome chloroplastique entre les lignées mitochondriales. Les incompatibilités entre les partitions chloroplastiques suggéraient en outre la présence de recombinaison présumément liée à l'évolution réticulée ancienne. Les marqueurs chloroplastiques seraient donc de moindre valeur phylogénétique et phylogéographique que les marqueurs mitochondriaux pour les genres ayant une transmission paternelle du génome chloroplastique, tel que chez les Pinaceae.

Discordant mtDNA and cpDNA spruce phylogenies indicate geographic speciation and reticulation as driving factors for the long-term evolution in the Pinaceae

3.2 - Abstract

The phylogeny of *Picea* was determined by sequencing three regions from the paternally inherited chloroplast genome (*trnK*, *rbcL*, and *trnTLF*), and the intron 2 of the maternally transmitted mitochondrial gene *nad1*. Significant differences were found between the *trnK* tree and the *rbcL* and *trnTLF* phylogenetic trees, and between cpDNA and mtDNA phylogenies. None of the phylogenies matched morphological classifications. The mtDNA phylogeny was geographically more structured than cpDNA phylogenies. Most North American taxa formed a monophyletic group on the mtDNA tree, indicating geographic speciation by range fragmentation or by dispersal and isolation. Similar patterns, also found among Asian taxa, likely represent a major trend in the Pinaceae because of similar life-history, autecology, and reproduction system. Incongruences between organelle phylogenies also indicated ancient lateral transfers of the chloroplast genome between mtDNA lineages. Incongruences between cpDNA partitions further suggested heterologous recombination presumably linked to ancient reticulate evolution. These results indicate the reduced value of cpDNA as a phylogenetic and phylogeographical marker in genera with paternal transmission of the chloroplast genome such as in the Pinaceae.

3.3 - Introduction

The lateral transfer of genetic material and subsequent reticulate evolution resulting from interspecific hybridization is thought to have played a major role in the evolution of many plant groups (Okuyama et al., 2004). Evidence for ancient lateral transfer has been reported between plant genera (e.g. Soltis et al., 1991), and it has been identified between Angiosperm species harboring different modes of organelle inheritance (e.g. Chat et al., 2004). Such a process complicates the definition of species entities from a genomic perspective because of its blending effect. Organelle genomes are usually uniparentally inherited in plants, and for most Angiosperms, the chloroplast and the mitochondrial genome are maternally transmitted (Mogensen, 1996). However, in spruce and other Pinaceae, the cpDNA lineage is paternally transmitted through pollen while the mtDNA lineage is maternally inherited through ovules (Mogensen, 1996; Jaramillo-Correa et al., 2003), which provides a unique opportunity to evaluate unambiguously instances of reticulate evolution.

Spruces and other Pinaceae are mostly outcrossing species with wind-dispersed pollen, which results in typically high amounts of gene flow among populations (e.g. Hamrick et al., 1992; Jaramillo-Correa et al., 2001). Spruces appear to represent an old lineage with fossils dating back to the Eocene (LePage, 2001). However, they bear little morphological novelties (Frankis, 1989), are morphologically weakly differentiated (Wright, 1955; Weng and Jackson, 2000), and are not well isolated reproductively (e.g. Wright, 1955; Mikkola, 1969; Gordon, 1976). Frequent introgressive hybridization between parapatric taxa (e.g. Krutovskii and Bergmann, 1995; Perron and Bousquet, 1997 ; Rajora and Dancik, 2000) and successful crosses between geographically distant species have been noted (e.g. Wright, 1955; Mikkola, 1969; Gordon, 1976). Given such promiscuity and the ability of wind-pollinated trees to disperse their pollen over long distances (Comtois, 1997), it is hypothesized that reticulate evolution incurred by lateral gene flow and cpDNA transfer across mtDNA lineages may have played a significant role in the evolutionary history of the genus.

Because mtDNA-carrying seeds are usually dispersed over smaller distances than cpDNA-carrying pollen, one would expect the phylogeny derived from mtDNA to be geographically more cohesive than that derived from cpDNA in the Pinaceae, especially in presence of reticulate evolution. By analysing patterns of distribution, Wright (1955) hypothetized that allopatric and parapatric speciation, together termed geographic speciation, might represent a main process of species birth in spruces. In general, the central mechanism to geographic speciation is dispersal followed by isolation and/or range fragmentation (see Brown and Lomolino, 1998). In spruce, it would have been promoted by colonization of new lands and by displacements of natural ranges induced by cycles of climate change, such as those incurred by the Pleistocene glaciations in the Northern hemisphere (Wright, 1955; Perron et al., 2000). A number of closely related spruce species have parapatric ranges with zones of contact characterized by introgressive hybridization (e.g. *P. mariana* and *P. rubens* in eastern North America, *P. glauca* and *P. englemannii* in western North America), as also observed in other Pinaceae genera such as *Larix* and *Pinus*, which indicates that these species have not completed their speciation process and are likely to represent more or less recent products of geographic speciation (Perron et al., 2000).

The objectives of this study were twofold: 1) to estimate the mtDNA phylogeny of an old Pinaceae genus with large geographic coverage and infer the driving speciation processes; 2) to evaluate the concordance of the corresponding cpDNA phylogeny and assess the occurrence of more or less ancient reticulate evolution. To achieve these goals, complete sequences for the mitochondrial *nad1* intron 2 and for the chloroplast genes *rbcL*, *trnK*, and the *trnTLF* regions were determined for the 35 taxa representative of the whole circumpolar genus *Picea* and for appropriate outgroups.

3.4 - Material and methods

3.4.1 - Taxa sampling

Material from 35 *Picea* representative of the whole genus and two *Pinus* outgroups were sampled from trees of documented natural origin in various arboreta, botanical gardens and seed banks (described in Appendix 1).

3.4.2 - DNA regions sampling

The *rbcL*, *trnK*, and *trnTLF* regions of the paternally-inherited chloroplast genome were used herein and in previous conifer phylogenetic studies (e.g. Wang et al. [1999] and Quinn et al. [2002] for the *rbcL* and *trnK* regions; Wei and Wang [2003] and Qiao et al. [2007] for *trnTLF*). The second intron of the *nad1* region of the maternally-inherited mitochondrial genome has been used previously for phylogenetic inference in the Angiosperms (e.g. Bakker et al., 2000; Freudenstein and Chase, 2001) and in the Gymnosperms (e.g. Qiao et al., 2007).

3.4.3 - Biochemical methods

DNA was extracted from needles or megagametophytes with the DNeasy Plant Mini Kit (Qiagen, Mississauga, Ontario, Canada). Information about PCR primers is reported in Table 3.1. The position of newly designed primers for the *nad1* region and conditions for PCR can be found respectively in Fig. 3.1 and Table 3.2. Both DNA strands were sequenced with Perkin-Elmer ABI DNA Sequencers (3100 and 3700, Applied Biosystems, Foster City, California, USA), using BigDye Terminator cycle sequencing kits. DNA sequences have been deposited in Genbank (see Appendix 1).

Table 3.1 - PCR primers used in this study.

DNA region	Primer name	Sequence (forward)	Primer name	Sequence (reverse)
rbcL	1F, 2F, 3F	(in Wang et al., 1999)	1R, 2R, 3R	(in Wang et al., 1999)
trnK	1F, 2F	(in Wang et al., 1999)	1R, 2R	(in Wang et al., 1999)
	3914F	(in Johnson and Soltis, 1995)	R3, R4, R6	(in Germano and Klein, 1999)
	F2, F4	(in Germano and Klein, 1999)		
trnT-[L-intron]-F	b, d, f	(in Taberlet et al., 1991)	a, c, e	(in Taberlet et al., 1991)
nad1 intron 1[1]	*nad1*.1-F	5'-GGTCGGATGTTACCCTAACC-3'	*nad1*.1-R	5'-ACCTACCCCTCGCTACTATCTC-3'
	nad1.2-F	5'-GGGTGCCGCAAGGTTATAC-3'	*nad1*.2-R	5'-GGATAAGGATAAGGCCAAGG-3'
	nad1.2b-F[2]	5'-GATGATGATGCCCCTATTGAT-3'	*nad1*.2b-R	5'-GATCTCCTCTCAAGGCCAAC-3'
	nad1.3-F	5'-ACTTACCAGTAAAGGCCCGA-3'	*nad1*.3-R	5'-GCCATGAAAAGACTTCCCTG-3'
	nad1.4-F	5'-CCCTACCCCTTCGGATACT-3'	*nad1*.4-R	5'-GTAGCGAGGTTACCGCTTAGT-3'
	nad1.5-F	5'-GCCATGAAAAGCCCTGTA-3'	*nad1*.5-R	5'-GTGTTGCTCTATCTCGGCAT-3'
	nad1.6-F	5'-CCACATGAGTAGTTCGCTCC-3'	*nad1*.6-R	5'-TACGGGGAACAAGGGAATAC-3'
	nad1.e-F[3]	5'-AAGGATAGGATAATAAAAAGACCAAA-3'	*nad*1.4b-R[4]	5'-TCTAAGGTGTTGTGTCCTCTCC-3'
	nad1.i-F[3]	5'-AAAGTATGGGCCGCCATTAA-3'		

Notes: [1]Newly designed in this study. [2]Added to obtain the sequence of *P. orientalis*. [3]Added to obtain the sequence of *P. omorika*. [4]Added to obtain the sequences of *P. orientalis* and *P. morrisonicola*.

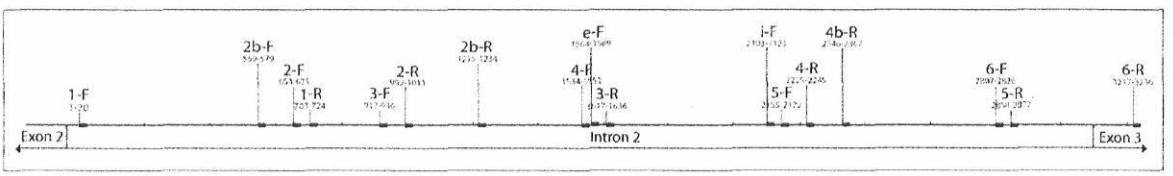

Fig. 3.1 - Position of PCR primers designed in this study to amplify the mitochondrial *nad1* intron 2.

Table 3.2 - PCR cycling parameters for gene amplification.

Locus or primer pairs	Cycles parameters
rbcL; *trnK* 1F-1R; *trnK* 2F-2R; *nad1*.1 (F-R); *nad1*.6 (F-R)	30 cycles of [94° 45 sec, 58° 50 sec, 72° 1 min 20 sec][1]
nad1.2 (F-R); *nad1*.2b (F-R)	35 cycles of [94° 45 sec, 57° 50 sec, 72° 1 min 20 sec][1]
nad1.3 (F-R); *nad1*.4 (F-R); *nad1*.4b (R); *nad1*.5 (F-R); *nad1*.e (F); *nad1*.i (F)	40 cycles of [94° 45 sec, 54° 50 sec, 72° 1 min 20 sec][1,4]
3914F-R6; F2-R4; F4-R3	30 cycles of [94° 30 sec, 55° 1 min 30 sec, 72° 2 min 30 sec][2,5]
trnT-[L-intron]-F	40 cycles of [94° 1 min, 56° 1 min, 72° 1 min 20 sec][3,5]

Notes: [1]Modified from Wang et al. (1999). [2]Modified from Germano and Klein (1999). [3]Modified from Taberlet et al. (1991). [4]With 40% more TAQ and 30% more primers. [5]With final extension of 10 min instead of 5 min.

3.4.4 - Phylogenetic analyses

Intraspecific variation was checked for each of seven species by sequencing from 2 to 5 samples : *P. abies, P. engelmannii, P. glauca, P. jezoensis, P. mariana, P. pungens*, and *P. sitchensis*. Intraspecific polymorphisms were detected for chloroplast genes in *P. abies* (1 nucleotidic site in *trnTLF*), *P. glauca* (1 nucleotidic site in *trnTLF* - an apomorphy), and in *P. sitchensis* (1 nucleotidic site in *trnK*). They were present in three taxa for the mitochondrial locus *nad1*: *P. abies* (36 nucleotide sites including 13 apomorphies), *P. jezoensis* (9 nucleotide sites including 3 apomorphies) and *P. sitchensis* (1 indel shared with *P. engelmannii*). For each locus, a preliminary analysis based on maximum parsimony (MP) showed monophyly for each of these taxa at each locus. To minimize the number of sequences in phylogenetic analyses, multiple sequences for each of these species were merged into a single consensus sequence where intraspecific variations were noted as multistate characters following the standard ambiguity code.

The possibility of concatenating the three chloroplast loci was assessed because all regions of the presumably nonrecombining chloroplastic genome should share a common evolutionary history. Topologies obtained from MP were compared with the «incongruence length difference» (ILD, Farris et al., 1995) test estimated by the «partition homogeneity test» in PAUP 4.0b10 (Swofford, 2001). One thousand replicates were performed, with multistate characters defined as polymorphisms, uninformative sites excluded from analysis, and the maximum number of trees set to 10. Overall, the three chloroplastic datasets could not be combined ($P = 0.017$). Results precluded the concatenation of the *trnK* and *trnTLF* chloroplast datasets ($P = 0.007$), but the *rbcL* and the *trnTLF* partitions could be combined ($P = 0.611$). Therefore, two separate partitions representing the chloroplast loci were delineated, *rbcL-trnTLF* and *trnK*.

Phylogenetic trees were estimated with MP as implemented in PAUP 4.0b10. Phylogenies were obtained after heuristic searches were conducted with multistate characters considered as polymorphisms, equal character weights, tree-bisection-reconnection (TBR) branch-swapping, 100 random sequence additions, ACCTRAN branch

length optimization, MULTREES option in effect, swapping on best trees only, and gaps treated as missing characters. However, since indels accounted for 14% of the informative sites, they were also considered and coded as additional binary characters. Branch support was computed from 500 bootstrap resamplings, using the same parameters as described above. Branch lengths on resulting strict consensus trees were computed with the «Branch-length table » option in PAUP 4.0b10. Using PAUP 4.0b10, minimum length topologies were also obtained with the neighbor-joining (NJ) method (Saitou and Nei, 1987) based on a matrix of pairwise divergence values taking into account both substitutions and indel characters (all equally weighted) as in MP, with 500 bootstrap resamplings.

Topological differences between gene partitions were evaluated using the nonparametric «Wilcoxon signed-ranks» test (Templeton, 1983) encoded in PAUP, following the basic analysis principles described in Seelanan et al. (1997). Parsimony strict consensus trees that arose from the different data partitions were compared as follows to identify the conflicting nodes. Each node of interest in the tree derived from the chloroplast locus *trnK* was edited as a constraint in a most parsimonious tree search (see above for search parameters) using the mitochondrial dataset (*nad1*). The resulting trees and the trees obtained in a search without constraint were compared (pairwise tree-to-tree distances) to find the pair of closest trees to be used in the Wilcoxon signed-ranks test. This test identified the characters in the mitochondrial dataset accounting for the length difference between the constrained and unconstrained trees for each node of interest, and calculated the significance ($P < 0.05$) of the topological difference accounted for by that node. Then, each node of interest in the strict consensus topology derived for the mitochondrial locus *nad1* was used as a constraint in analyses with the chloroplast *trnK* dataset in a similar manner. Differences in topologies were also evaluated as described above between the chloroplast *rbcL-trnTLF* and the mitochondrial *nad1* strict consensus trees, and between both chloroplast *trnK* and *rbcL-trnTLF* strict consensus trees.

3.4.5 - Mapping of morphological classifications onto phylogenetic trees

For each estimated phylogeny, morphological classifications of the genus were mapped onto the phylogenetic trees as a qualitative measure of congruence. The most recent morphological classifications of Liu (1982) and Farjon (1990), based primarily on vegetative and reproductive characters, respectively, were used and are depicted herein by their first level of grouping (Table 3.3 and Fig. 3.2). Earlier classifications (for a review, see Schmidt-Vogt [1977]) were also mapped with little congruence obtained with the diverse phylogenies estimated herein (results not shown).

Table 3.3. Names of taxa analyzed and their subspecific classification according to Liu (1982) and Farjon (1990).

Taxon	Liu (1982)		Farjon (1990)	
	Subspecific classification	Symbol in Fig. 3.2	Subspecific classification	Symbol in Fig. 3.2
P. abies	*Picea*	▽	*Picea*	○
P. alcoquiana	*Picea*	▽	*Picea*	○
P. asperata	*Picea*	▽	*Picea*	○
P. brachytyla	*Omorika*	■	*Picea*	○
P. breweriana	*Omorika*	■	*Picea*	○
P. chihuahuana	*Omorika*	■	*Picea*	○
P. crassifolia	*Picea*	▽	*Picea*	○
P. engelmannii	*Picea*	▽	*Casicta*	✱
P. farreri	-	n. m.[3]	*Picea*	○
P. glauca	*Picea*	▽	*Picea*	○
P. glehnii	*Omorika*	■	*Picea*	○
P. jezoensis	*Picea*	▽	*Casicta*	✱
P. koraiensis	*Omorika*	■	*Picea*	○
P. koyamai	*Omorika*	■	*Picea*	○
P. likiangensis	*Picea*	▽	*Casicta*	✱
P. mariana	*Picea*	▽	*Picea*	○
P. martinezii	-	n. m.[3]	-	n. m.[3]
P. maximowiczii	*Omorika*	■	*Picea*	○
P. mexicana	*Picea*	▽	*Casicta*	✱[1]
P. meyeri	*Picea*	▽	*Picea*	○
P. montigena	*Picea*	▽	*Casicta*	✱[2]
P. morrisonicola	*Omorika*	■	*Picea*	○
P. obovata	*Picea*	▽	*Picea*	○
P. omorika	*Omorika*	■	*Picea*	○
P. orientalis	*Omorika*	■	*Picea*	○
P. pungens	*Picea*	▽	*Casicta*	✱

(continued)

Table 3.3 - (continued)

P. purpurea	*Picea*	▽	*Casicta*	*
P. retroflexa	*Picea*	▽	*Picea*	○
P. rubens	*Picea*	▽	*Picea*	○
P. schrenkiana	*Omorika*	■	*Picea*	○
P. sitchensis	*Picea*	▽	*Casicta*	*
P. smithiana	*Omorika*	■	*Picea*	○
P. spinulosa	*Omorika*	■	*Picea*	○
P. torano	*Picea*	▽	*Picea*	○
P. wilsonnii	*Omorika*	■	*Picea*	○

Notes: [1]*P. engelmannii* ssp. *mexicana*. [2]*P. likiangensis* ssp. *montigena*. [3]n. m.: not mentioned in the classification.

3.5 - Results

3.5.1 - Sequence variation

Types and amounts variations found for each region and average pairwise sequence divergence values are given in Table 3.4. To our knowledge, the present study reports the first instance of a genus phylogeny based exclusively on maternally inherited mtDNA gene sequences in the Gymnosperms. With an average length of 2 898 bp (2 862 to 3 078 bp), the mtDNA *nad1* intron 2 showed six times more informative sites than all three cpDNA regions combined (*rbcL*, *trnK*, and *trnTLF*), and about twice more informative indels relative to the number of informative substitution sites.

Table 3.4 - Summary of sequence variation and DNA divergence found in the genus *Picea*.

Locus	Consensus sequence length (bp)	Number of substitutions/ coded indels [length in bp]	Number of parsimony informative sites		Average pairwise divergence (%)		
			Subst.	Indels	Intraspecific	Interspecific within *Picea*	Interspecific between *Picea* and *Pinus*[1]
cpDNA							
rbcL	1351	19 / 0	11	0	nil	0.25	2.64
trnTLF	1456	44 / 7 [1 to 20]	20	2	0.020	0.44	7.46
rbcL/trnTLF[2]	-	-	-	-	0.010	0.35	5.11
trnK	2527	50 / 5 [5 to 18]	25	3	0.006	0.33	4.38
All cp genes	-	-	-	-	0.008	0.34	4.77
mtDNA							
nad1 intron 2	3184	221 / 44 [1 to 114]	151	25	0.206	1.23	1.93

Notes: [1] using outgroups *Pinus banksiana* and *Pinus koraiensis*. [2] group of two genes based on tests of topological congruence between data partitions (see Results).

3.5.2 - Phylogenies based on cpDNA

Intraspecific polymorphism was checked for seven species. While some polymorphisms were observed, they did not affect the topologies obtained (see Materials & Methods). Parsimony analyses were first conducted for each cpDNA locus separately. Tests of topological incongruence («Incongruence Length Difference» [ILD, Farris et al., 1995]) revealed two significantly different cpDNA partitions ($P < 0.05$) : *rbcL/trnTLF* and *trnK*. Hence, for the remaining of the phylogenetic analyses, these two datasets were considered separately. The *rbcL/trnTLF* parsimony analysis resulted in 192 minimum trees of 87 steps when considering only the ingroup *Picea*. The *rbcL/trnTLF* strict consensus tree (Fig. 3.2A) showed *P. breweriana* branching out near the root of the tree, followed by *P. sitchensis*, then a well supported group of three other North American taxa including *P. glauca* and *P. engelmannii*. The next group (well supported node III) included three clades in which the rest of the North American taxa were spread: *P. pungens*, *P. mariana* and *P. rubens* nested within the *P. abies* clade, and *P. chihuahuana* and *P. martinezii* nested within the *P. alcoquiana* clade. Taxa groups exhibiting no sequence difference for the *rbcL/trnTLF* partition were the following: *P. engelmannii* and *P. glauca*; *P. abies* and *P. koyamai*; *P. montigena*, *P. retroflexa* and *P. torano*; *P. farreri* and *P. likiangensis*; *P. purpurea* and *P. wilsonii*. The *trnK* parsimony analysis yielded 64 minimum trees of 67 steps when considering only the ingroup *Picea*. The *trnK* strict consensus tree shows an important topological difference with the *rbcL/trnTLF* strict consensus: the well supported clade (node XI) of the North American taxa *P. engelmannii*, *P. glauca* and *P. mexicana* was located next to the *P. abies* clade instead of being at the base of the phylogeny (Fig. 3.2B). In both cpDNA phylogenies, *P. engelmannii*, *P. glauca* and *P. sitchensis*, a group of North American taxa morphologically weakly differentiated and forming natural hybrids («H3» in Fig. 3.2A-B) were not regrouped together. For the *trnK* partition, no divergence was found between *P. abies* and *P. koyamai*; *P. asperata*, *P. crassifolia*, *P. montigena*, *P. retroflexa* and *P. torano*; *P. engelmannii*, *P. glauca* and *P. mexicana*; *P. farreri* and *P. likiangensis*; *P. mariana* and *P. omorika*; *P. purpurea* and *P. wilsonii*. Also, neither of the different morphological classifications matched the major clades in either cpDNA phylogenies (see Fig. 3.2A-B for those of Liu [1982] and Farjon [1990]). Although

significant topological heterogeneity was observed between cpDNA data partitions, sequences from the three chloroplast loci were concatenated for the sake of considering total evidence, but no inferences were drawn from that tree (Fig. 3.2C). Overall, for all cpDNA loci, no sequence differences were observed between *P. abies* and *P. koyamai*; nor between *P. engelmannii* and *P. glauca*; *P. farreri* and *P. likiangensis*; *P. montigena*, *P. retroflexa* and *P. torano*; *P. purpurea* and *P. wilsonii*.

3.5.3 - Phylogeny based on mtDNA

The parsimony analysis of *nad1* intron 2 yielded 372 minimum trees of 403 steps when considering only the ingroup *Picea*. In the mtDNA strict consensus tree (Fig. 3.2D), the base of the phylogeny (below node XII) consisted of clades of Asian taxa. The largest resolved clade of the tree (node XII, with moderate bootstrap support) included all the North American taxa. The North American *P. breweriana* was regrouped with the Asian *P. obovata* with strong bootstrap support, and had less sequence divergence with it (0.24 %) than with the other North American taxa (from 0.55 to 0.62 %). The other taxa of clade XII were regrouped into two major subclades: an Asian subclade of seven taxa including *P. schrenkiana*, and a subclade including nine of the 10 North American taxa. Bootstrap support for these clades was moderate. This subclade was further divided in four groups with moderate bootstrap support, however differently than for the cpDNA trees: 1) *P. mariana*, *P. rubens* and the European *P. omorika*; 2) *P. chihuahuana*, *P. mexicana* and *P. pungens*; 3) *P. glauca*, *P. engelmannii* and *P. sitchensis*; and 4) *P. martinezii*. *P. martinezii* was closer to the *P. chihuahuana* group than to the other North American taxa, but with sizeable sequence divergence : while the average sequence divergence within the *P. chihuahuana* group was 0.07%, the average divergence between *P. martinezii* and taxa from the *P. chihuahuana* group was 0.27%. The mtDNA phylogeny regrouped closely all taxa forming natural hybrids (H groups) except *P. abies* and *P. obovata* (Fig. 3.2D). For this locus, no sequence differences were observed between *P. asperata* and *P. torano*, nor between *P. crassifolia* and *P. glehnii*. As for cpDNA, the morphological classifications did not match with the major mtDNA clades (Fig. 3.2D).

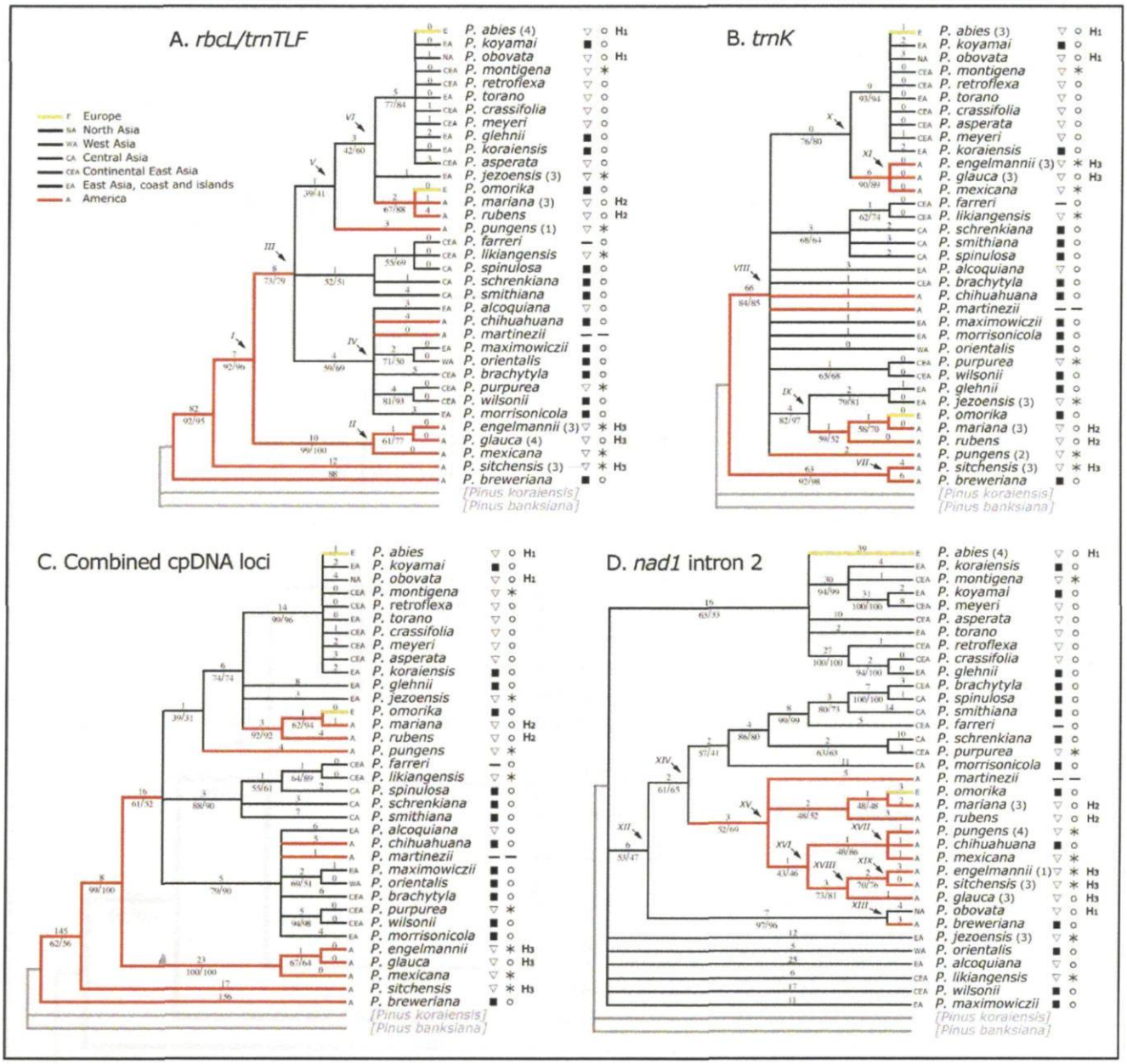

Fig. 3.2 Strict consensus trees. (A) *trnK* partition (CI=0.881, RI=0.935, RC=0.824, HI=0.179). (B) *rbcL/trnTLF* partition (CI=0.816, RI=0.894, RC=0.730, HI=0.195). (C) combined cpDNA data sets (CI=0.802, RI=0.884, RC=0.709, HI=0.228). (D) mtDNA *nad1* intron 2 (CI=0.753, RI=0.762, RC=0.574, HI=0.309). Parentheses: number of samples per species for assessment of intraspecific variation (see Methods). Numbers above branches: branch length; below branches: parsimony bootstrap values (left) and NJ bootstrap values (right). Arrows indicate significantly different nodes between topologies (Wilcoxon signed-ranks test) details in Table 5). H. species pairs with introgressive hybridization. Morphological classifications according to Liu 1982): subgenus *Picea* = ▽, subgenus *Omorika* = ■ , and according to Farjon 1990): section *Picea* = ❍ , section *Casicta* = ✱

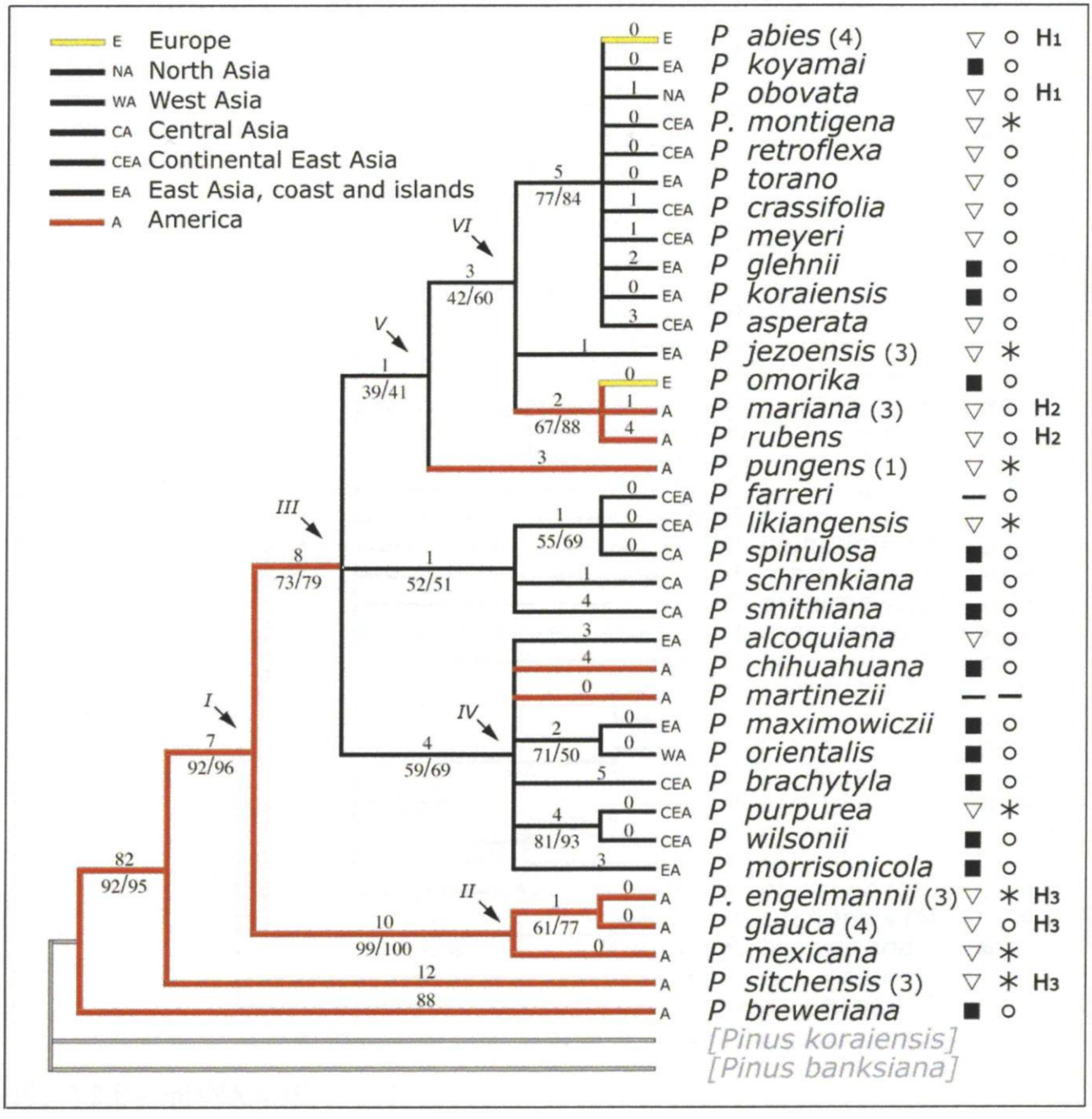

Fig. 3.2 A cpDNA *rbcL* and *trnT-L-F*

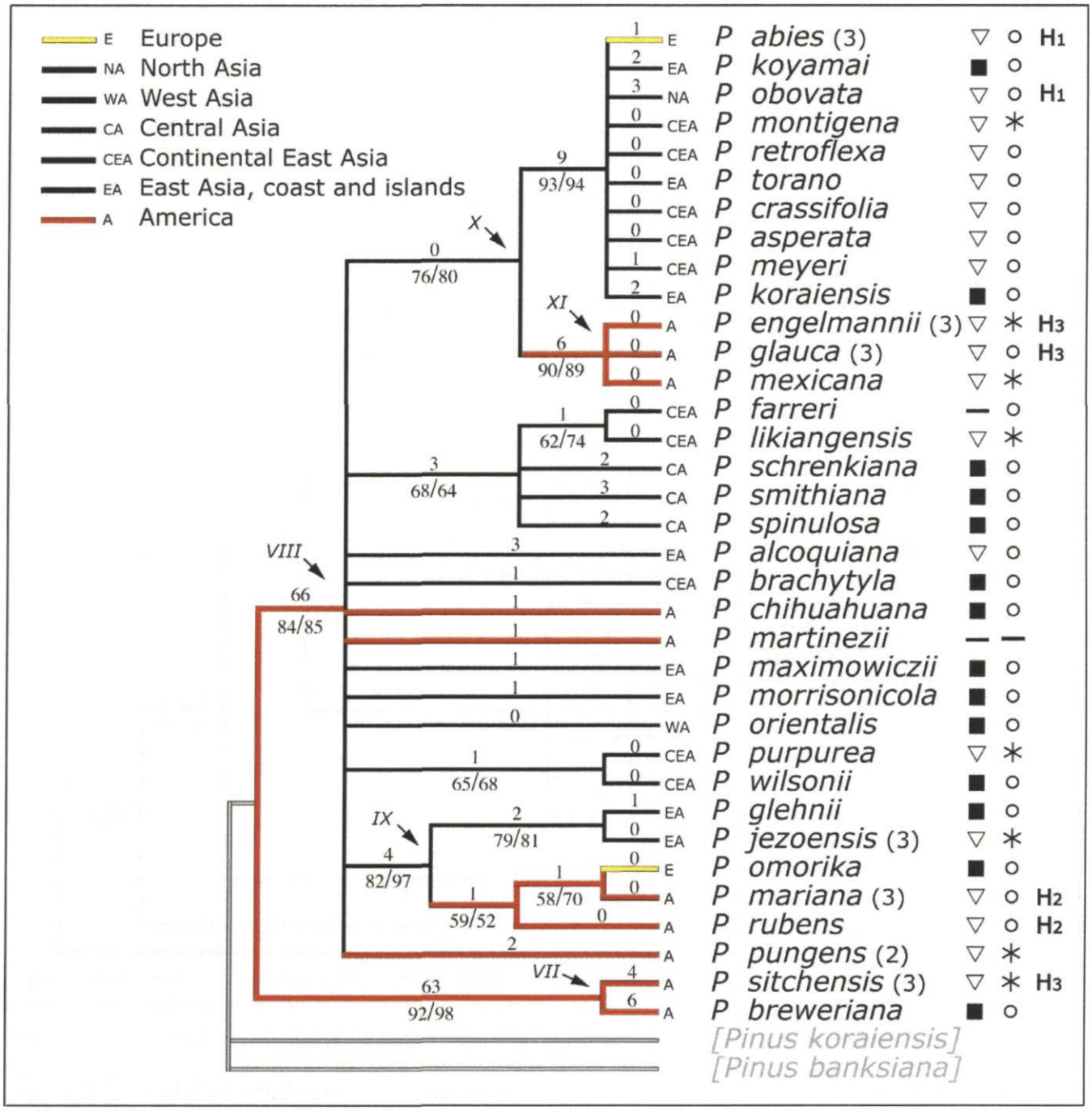

Fig. 3.2 B cpDNA *trnK*

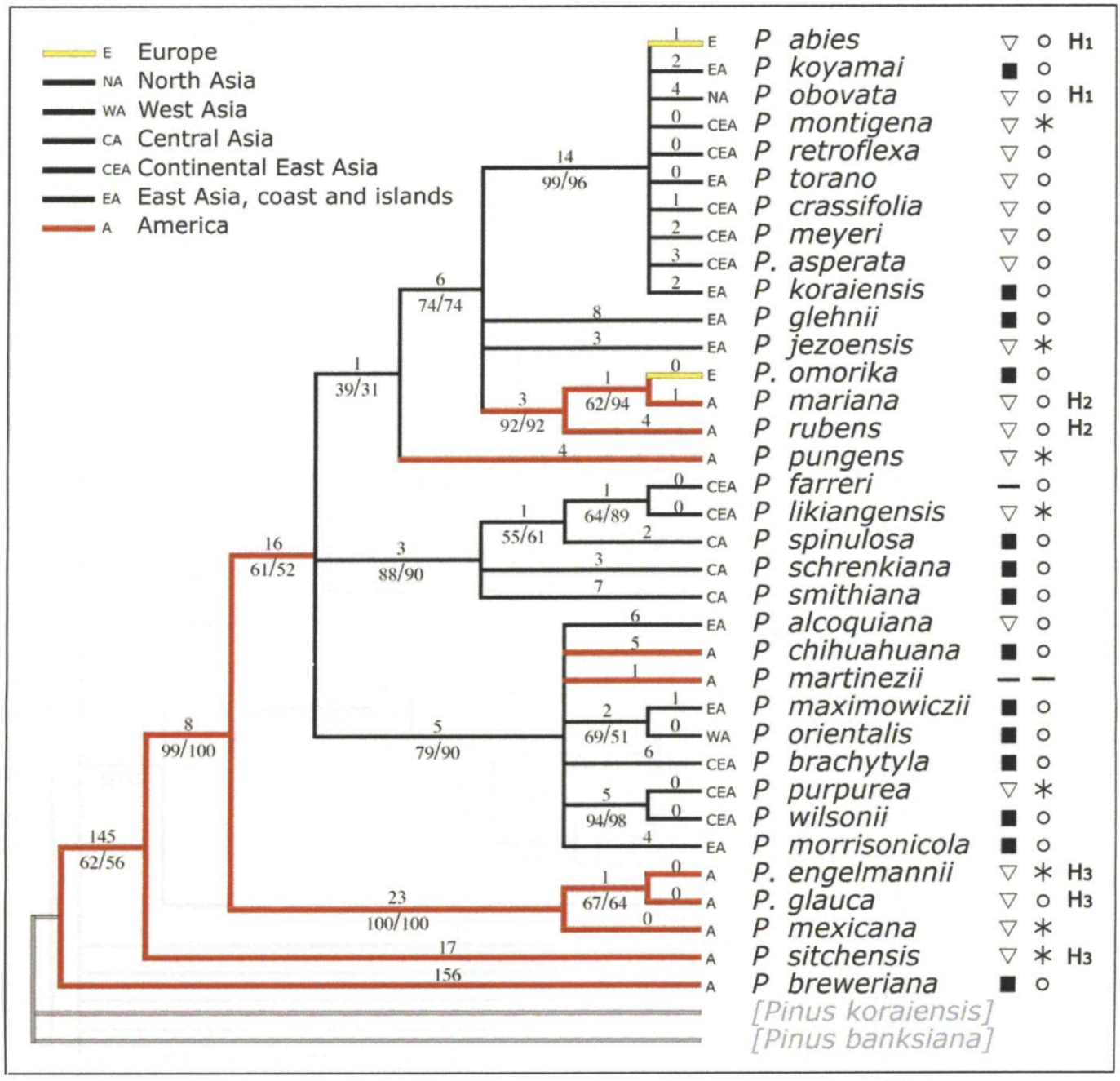

Fig. 3.2 C cpDNA *rbcL*, *trnT-L-F* and *trnK*

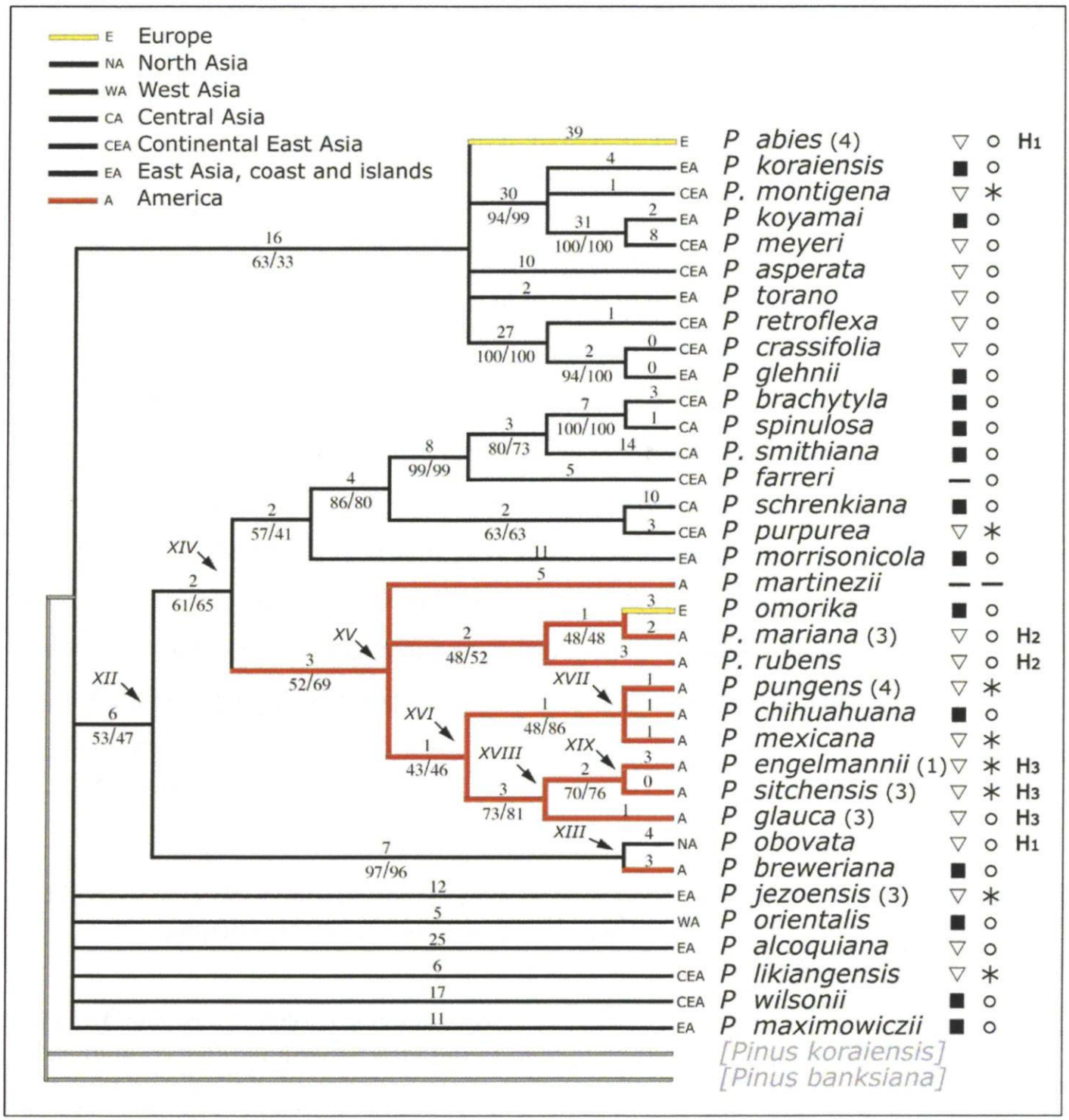

Fig. 3.2 D mtDNA *nad1* intron2

3.5.4 - Topological comparisons

Incongruence length difference between cpDNA topologies arising from various partitions as well as between cpDNA and mtDNA topologies was detected (see exploratory analyses in Material and Methods). To identify the sources of topological incongruence between data partitions, each conflicting phylogenetic node was tested using the Wilcoxon signed-ranks test (Templeton, 1983). Several significant differences ($P < 0.05$) involving North American taxa were observed (Table 3.5). In comparing topologies from the two cpDNA partitions delineated (*rbcL/trnTLF* versus *trnK*), two nodes with North American taxa accounted for the differences between cpDNA trees: the North American clades of nodes IX and XI (Fig. 3.2B) were transposed between the two cpDNA phylogenies. Comparisons between the mtDNA phylogeny and each of the two cpDNA topologies involved many significant topological differences (Table 3.5). About the same clades involving North American taxa appeared inconsistent between the mtDNA topology and each of the two cpDNA partitions tested. For North American taxa, only the grouping of *P. mariana*, *P. rubens*, and the European *P. omorika* was congruent between the mtDNA and the two cpDNA topologies.

Table 3.5 - Significant differences ($P < 0.05$) involving North American taxa in topological comparisons (Wilcoxon signed-ranks test).

	Length	Nb. of trees	Gain	Loss	Net	$P < 0.05$
rbcL/trnTLF and *trnK* topological comparison						
-Constraint on the *rbcL/trnTLF* data set						
trnK strict consensus tree (all clades)	95	168	36	0	36	0.0047
node *IX* (clade *P. glehnii* to *P. rubens*)	91	240	10	0	10	0.0455
node *X* (clade *P. abies* to *P. mexicana*)	94	60	40	5	35	0.0196
-Constraint on the *trnK* data set						
rbcL/trnTLF strict consensus tree (all clades)	73	9	31.5	4.5	27	0.0339
trnK and *nad1* topological comparison						
-Constraint on the *trnK* data set						
nad1 strict consensus tree (all clades)	124	2	378	0	378	< 0.0001
node *XII* (clade *P. schrenkiana* to *P. breweriana*)	81	285	105	0	105	0.0002
node *XIII* (clade *P. obovata* and *P. breweriana*)	82	492	128	8	120	0.0006
node *XIV* (clade *P. schrenkiana* to *P. glauca*)	84	2 664	136	0	136	0.0001
node *XV* (clade *P. martinezii* to *P. glauca*)	80	192	91	0	91	0.0003
node *XVI* (clade *P. pungens* to *P. glauca*)	78	126	84	7	77	0.0023
node *XVII* (clade *P. pungens* to *P. mexicana*)	73	483	21	0	21	0.0143
node *XVIII* (clade *P. engelmannii* to *P. glauca*)	80	588	91	0	91	0.0003
node *XIX* (clade *P. engelmannii* and *P. sitchensis*)	79	336	78	0	78	0.0005

Table 3.5 - (continued)

-Constraint on the *nad1* data set							
trnK strict consensus tree (all clades)		529	3 561	4 119	0	3 741	< 0.0001
node *VII*	(clade *P. breweriana* and *P. sitchensis*)	416	372	91	0	91	0.0003
node *VIII*	(clade *P. abies* to *P. pungens*)	416	372	91	0	91	0.0003
node *IX*	(clade *P. glehnii* to *P. rubens*)	456	2 532	1 378	0	1 431	< 0.0001
node *X*	(clade *P. abies* to *P. mexicana*)	468	18 108	1 924	0	1 891	< 0.0001
node *XI*	(clade *P. engelmannii* to *P. mexicana*)	408	1 488	15	0	15	0.0253
rbcL/trnTLF and *nad1* topological comparison							
-Constraint on the *rbcL/trnTLF* data set							
nad1 strict consensus tree (all clades)		141	6	378	0	378	< 0.0001
node *XII*	(clade *P. schrenkiana* to *P. breweriana*)	103	342	120	0	120	0.0002
node *XIII*	(clade *P. obovata* and *P. breweriana*)	99	1 152	97.5	7.5	90	0.0013
node *XIV*	(clade *P. schrenkiana* to *P. glauca*)	102	1 560	128	8	120	0.0006
node *XVI*	(clade *P. pungens* to *P. glauca*)	95	1 887	65	13	52	0.0209
node *XVII*	(clade *P. pungens* to *P. mexicana*)	97	911	71.5	6.5	65	0.0039
node *XVIII*	(clade *P. engelmannii* to *P. glauca*)	95	96	49.5	5.5	44	0.0114
node *XIX*	(clade *P. engelmannii* and *P. sitchensis*)	96	278	60	6	54	0.0067
-Constraint on the *nad1* data set							
rbcL/trnTLF strict consensus tree (all clades)		492	340	1 454	31	1 423	< 0.0001
node *I*	(clade *P. abies* to *P. mexicana*)	416	372	91	0	91	0.0003
node *II*	(clade *P. engelmannii* to *P. mexicana*)	408	1 488	15	0	15	0.0253
node *III*	(clade *P. abies* to *P. morrisonicola*)	415	744	78	0	78	0.0005
node *IV*	(clade *P. alcoquiana* to *P. morrisonicola*)	442	4 199	840	21	819	< 0.0001
node *V*	(clade *P. abies* to *P. pungens*)	420	45	308	70	238	0.0011
node *VI*	(clade *P. abies* to *P. rubens*)	419	1 124	161.5	9.5	152	0.0002

3.6 - Discussion

3.6.1 - cpDNA phylogenies

Three major outcomes regarding the cpDNA phylogeny of a large Pinaceae genus arose from this study: statistically different data partitions among the cpDNA regions analyzed, incongruence between cpDNA with mtDNA phylogenies, and lack of congruence with taxonomical treatments, whether from vegetative (e.g. Liu, 1982) or reproductive structures (e.g. Farjon, 1990). CpDNA RFLPs were used early on to infer the phylogenetic relationships among *Picea* taxa (Sigurgeirsson and Szmidt, 1993). Potential problems of DNA fragment homoplasy in using RFLPs were then pointed out, as well as other pitfalls in establishing a reliable phylogeny, including introgressive hybridization and the paternal inheritance of cpDNA in spruce (Sigurgeirsson and Szmidt, 1993). Recently, cpDNA sequences were used to study the phylogeography of *Picea* but without considering possible ancient reticulate evolution from paternal inheritance of cpDNA (Ran et al., 2006),

which may lead to erroneous biogeographical interpretations. The implications will be discussed below. However, the cpDNA clades observed in our study agree to a large extent with the major cpDNA clades previously delineated (Sigurgeirsson and Szmidt, 1993; Ran et al., 2006), in concurrence with Wright (1955): the «*glauca* alliance» (nodes II and XI), the «*brachytyla* alliance» (node IV) except *P. torano* [syn. *P. polita*], and the «*abies* alliance» except *P. glehnii* (the large polytomy including *P. abies*).

A major and statistically significant topological incongruence was noted between the two cpDNA partitions in the present study: the displacement of the «*glauca* alliance». The basal position of this group in the *rbcL/trnTLF* phylogeny is congruent with other cpDNA results from RFLPs (Sigurgeirsson and Szmidt, 1993) or from combined analysis of *trnTLF* and *trnC-trnD* sequence datasets (Ran et al., 2006). The differential positioning of this group on the *trnK* tree likely indicates cpDNA heterologous recombination from more or less ancient hybridization, which has also been suggested in some pine cpDNA (e.g. Marshall et al., 2001). This is paralleled nowadays by the promiscuity of *P. glauca* and its ability to produce viable artificial hybrids with many species throughout the genus (e.g. *P. engelmannii*, *P. sitchensis*, *P. jezoensis*, *P. koyamai*, *P. pungens*, *P. schrenkiana*) (reviewed in OECD 1999). Further specific investigations are needed to fully understand the nature and extent of this likely rearrangement.

3.6.2 - mtDNA phylogeny

Although bootstrap support was moderate for several clades, underlining the need to sequence more mtDNA loci, there was generally good congruence between the MP and NJ trees, and numerous major incongruences were noted between cpDNA and mtDNA phylogenies. In the Pinaceae, cpDNA is paternally-inherited through pollen, which can travel on much longer distances than seeds, as shown by long-range dispersal following wind updraft (Comtois, 1997). Given the observed promiscuity within several groups of extant taxa, long-range lateral transfers of cpDNA and reticulate evolution are likely on the geological time scale, which would obliterate ancient relationships by descent. Such a trend would render spruce cpDNA phylogenies, and, more generally, Pinaceae cpDNA

phylogenies less reliable than their mtDNA couterparts for phylogeographic inferences at the genus level. Interspecific lateral transfer of the chloroplast genome has been inferred in Angiosperms (Soltis et al., 1991; Rieseberg and Soltis, 1991), including in plants presenting different modes of organelle inheritance (e.g. Chat et al., 2004), and especially in genera with a history of extensive natural hybridization still prevailing today (e.g. Hamzeh and Dayanandan, 2004). In *Picea*, artificial crossing experiments (e.g. Wright, 1955; Mikkola, 1969; Gordon, 1976; Gordon, 1986) have resulted in high interfertility between most closely related taxa and, to a lesser extent, between some taxonomically distant taxa (e.g. Wright, 1955). Such weak reproductive isolation seen in various groups throughout the genus parallels the weak morphological differentiation observed between these taxa. Most likely, it indicates a slow rate of speciation, given that lineages leading to extant members of divergent clades would have split many million years ago (Bouillé and Bousquet, 2005). Such promiscuity between many spruces and apparently slow speciation rates may be related to large historical population sizes, as suggested for many lineages of perennial plant species (Bousquet et al., 1992). Large historical populations sizes in excess of 10^5 have recently been estimated in spruces (Bouillé and Bousquet, 2005) and in pines (Brown et al., 2004; Ma et al., 2006). The maintenance of large hybrid zones between morphologically distinct taxa (Wright, 1955; Perron et al., 2000; Sutton et al., 1991) and the various locations of these zones of contact following repetitive range displacements and glacial cycles (Wright, 1955) are a distinct signature of weak reproductive isolation, which would have promoted opportunities for interspecific gene exchange. On the geological time scale, long-range lateral gene exchange between allopatric taxa, even widely distant, should also be considered as probable. Consequently, phylogeographical hypotheses regarding the centre of origin of the genus or subsequent taxa dispersions could not be tested reliably using cpDNA in *Picea*.

Because of the similarity of life history and reproductive characters between *Picea* and other members of the Pinaceae, it is likely that most cpDNA phylogenies of the genera of the Pinaceae are affected by reticulate evolution and do not fully reflect evolution by descent. A recent although limited comparison of the cpDNA and mtDNA phylogenies of the smaller Pinaceae genus *Larix* suggested a case of chloroplast capture involving *L.*

sibirica (Gros-Louis et al., 2005). As for the genus *Picea*, introgressive hybridization has also been reported in *Larix* and crossability between taxa is generally high (e.g. Semerikov et al., 2003).

3.6.3 - Incongruence with morphological classifications

cpDNA and mtDNA clades do not agree with classifications primarily based on vegetative characters (e.g. Liu, 1982), nor with the commonly accepted classifications based on reproductive characters (e.g. Farjon, 1990). However, in Pinaceae, morphological similarity, be it vegetative or reproductive, is not necessarily the result of close ancestry (e.g. Mitton and Andalora, 1981). Along some tree ranges, especially those covering relatively large areas, phenotypic plasticity can lead to significant morphological variation (e.g. Gordon, 1976; Rajora and Dancik, 2000; Liu, 1982; Konishi and Suzuki (1997) *in* Kobayashi et al., 2000). Sympatric non-hybridizing species can also exhibit character convergence in specific conditions such as harsh environments (e.g. Gordon, 1976). If lateral transfer of cpDNA has occurred repeatedly in the history of the genus, as indicated by the multiple incongruences between the phylogenies derived from the differentially inherited mtDNA and cpDNA, the nuclear genome would have been affected too. In such cases, it is likely to be mixed at various degrees, depending of the amount and direction of introgression, leading to altered phenotypes. Taxonomic considerations at the subspecific or varietal level may serve to illustrate the pitfalls in tracing relationships based on morphological affinities. For instance, *P. mexicana* is currently considered as a subspecies of *P. engelmannii* (Taylor and Patterson, 1980; Farjon, 2001). The present mtDNA phylogeny is in conflict with this classification scheme, but there is a closer relationship between these taxa according to the cpDNA phylogenies where reticulate evolution is rampant. Also, *P. montigena* is currently treated as a subspecies of *P. likiangensis*, and *P. purpurea* is thought as closely linked to *P. likiangensis* (Farjon, 2001), but little mtDNA or cpDNA affinity was found between *P. likiangensis* and *P. montigena*, or between *P. likiangensis* and *P. purpurea*. Thus, the taxonomical treatment of spruces still falls onto difficult grounds because morphologically or phenotypically similar entities may not reflect genetic nor genealogical relationships.

3.6.4 - Evidence for geographic speciation and reticulate evolution

Unlike cpDNA phylogenies, the mtDNA phylogeny estimated herein clearly shows more geographical coherence, which is indicative of geographic speciation by range fragmentation or dispersal and isolation. Such a spatially more structured pattern is expected for a genome that is maternally transmitted through ovules and dispersed by sizeable seeds instead of pollen. This trend is clearly illustrated with respect to the North American taxa: all of these taxa except *P. breweriana* are found into a single subclade supported by both MP and NJ analyses.

In disagreement with cpDNA-based views that the origin of the genus would turn around the North American ancestor of the relictual *P. breweriana* (Sigurgeirsson and Szmidt, 1993; Ran et al., 2006 ; and as suggested by the cpDNA phylogenies herein), the mtDNA phylogeny supports Wright's (1955) biogeographical contentions placing the origin of the genus in Asia, because the base of the mtDNA phylogeny is mostly represented by Asian taxa. Wright also hypothetized that *P. jezoensis* could be the link between the Asian and West American taxa because it is morphologically similar to *P. sitchensis* and *P. engelmannii*, and crosses reasonably well with *P. glauca* (Wright, 1955). However, the mtDNA data placed *P. jezoensis* rather distantly from this group. More mtDNA data, for instance from other regions of the *nad1* gene and additional loci, will be necessary to resolve the phylogenetic relationships among the basal Asian taxa.

When tracing back the past dissemination of lineages leading to extant *Picea* taxa, the geographically heterogeneous and well supported clade formed by the West American *P. breweriana* and the North Asian *P. obovata* suggests a migration of the genus from East Asia towards Western North America, likely via the Beringian land bridge. Such a dispersal pattern has been inferred for *Tsuga* (LePage, 2003). This dispersal scenario, distinct and presumably earlier than that involving the rest of the North American taxa, is in agreement with *P. breweriana* considered as the most ancestral and most divergent taxon in North America, based on crossing experiments (Gordon, 1986) and molecular evidence (Ledig et al., 2004), and because of some distinct morphological, anatomical and genetic features

(Weng and Jackson, 2000). Furthermore, in the present mtDNA data, the highest pairwise sequence divergence values found between North American taxa were those involving *P. breweriana*: from 0.55 % with *P. chihuahuana* to 0.62 % with *P. rubens*/*P. mariana*, compared to 0.25% average pairwise divergence within the large North American clade. The closest relative to *P. breweriana*, *P. obovata*, is cosmopolitan in the northern part of the Asian continent. It hybridizes naturally with the North European *P. abies* in their broad zone of contact (Krutovskii and Bergmann, 1995), but such a trend was not reflected in the mtDNA phylogeny (taxa noted «H1» in Fig. 3.2). When considering mtDNA sequence divergence values, *P. obovata* and *P. abies* were both more closely related to *P. breweriana*, with respective divergence values of 0.24 % and 1.48 %, than to each other (1.52 %). In addition, *P. obovata* had its second lowest mtDNA sequence divergence with all other North American taxa and the European *P. omorika* (from 0.59 to 0.66 %) and thirdly, with all Asian taxa (from 0.69 to 2.22 %). This trend suggests a common ancestral origin for *P. obovata* and *P. breweriana,* presumably in northeastern Asia and in western North America, followed by geographic speciation during a period with cooler climate. Later stepwise speciation in Eurasia could involve the westward continental spread of this ancestor, perhaps leading to *P. abies*, which naturally hybridizes with *P. obovata*, and other Asian taxa located more south.

All other North American taxa formed a monophyletic group highly indicative of geographical speciation, and involving a second intercontinental migration most presumably from eastern Asia via the Beringian land bridge. This group includes both the largely distributed *P. mariana* and *P. glauca*. They have sympatric transcontinental geographical ranges and are considered as distinct biological entities as interspecific crosses have only been met with very limited success (Wright, 1955; Gordon, 1986). In northeastern North America, the geographically proximal *P. mariana* and *P. rubens* were closely related. They share an extensive zone of contact in the Appalachians where introgressive hybridization has been reported repeatedly (e.g. Perron and Bousquet, 1997). The present data supports the hypothesis of *P. rubens* being a recent derivative of *P. mariana*, presumably arising from allopatric speciation during the Pleistocene (Perron et al., 2000; Jaramillo-Correa and Bousquet, 2003). The phylogenetic evidence also shows

these taxa to be closely related to the central European *P. omorika*, a trend confirmed by cpDNA as well as morphological and cross data (Wright, 1955; Gordon, 1976). *P. omorika* is a relict species sympatric to *P. abies* in the Tara Mountains in the Balkans, but *P. abies* clearly has a different and more ancient origin. Whether *P. omorika* is an allopatric derivative of *P. mariana* or vice versa will need comparing genetic diversity at many nuclear loci, with the hypothesis that the progenitor harbors more genetic diversity (Perron et al., 2000). However, the high genetic identity between the European *P. omorika* and the North American *P. mariana* and *P. rubens* suggests a relatively recent dispersal of the lineage across a land bridge in the North Atlantic region, a pathway evidenced by the presence of *P. mariana* fossils in Greenland from the Pliocene/Pleistocene era (Bennike, 1990). As for the high cpDNA affinity between the *P. mariana*/*P. rubens*/*P. omorika* group and the East Asian *P. jezoensis*, it likely represents a case of lateral transfer of the chloroplast genome, since these two allopatric entities are dispersed on different continents and are distantly related on the mtDNA phylogeny.

In western North America, the geographically proximal *P. glauca*, *P. engelmannii* and *P. sitchensis* were also tightly knitted together on the mtDNA tree, which supports the idea of a common recent origin (Wright, 1955) and geographic speciation involving range fragmentation or dispersion and isolation prior to secondary contacts and demonstrated introgressive hybridization (Rajora and Dancik, 2000; Page and Hollands, 1987; Bennuah et al., 2004). As for the odd proximity of *P. sitchensis* and *P. breweriana* on the *trnK* tree, it would suggest past lateral transfer of cpDNA between these lineages.

Geographical speciation is also suggested on the mtDNA tree by the closely related *P. mexicana*, *P. chihuahuana* and *P. pungens*, which have disjunct ranges all located in the Middle-West and South of North America (described in Ledig et al., 2004). Although these three taxa exhibited with one another the lowest pairwise sequence divergence values (0.07 %) among North American taxa for mtDNA, their cpDNA sequence divergence was relatively higher (in the range of the average interspecific divergence for North American taxa), which resulted in their positioning in three distant clades on the *rbcL*/*trnTLF*

phylogeny. This pattern suggests ancient cpDNA lateral transfer that would imply the lineage leading to *P. engelmannii*, as evidenced by close cpDNA affinity between *P. mexicana* and *P. engelmannii* (Fig. 3.2A-B) as well as by morphological similarity (Taylor and Patterson, 1980). As for *P. martinezii*, mtDNA data set it apart from the other Mexican taxa. Its average sequence divergence with the *P. chihuahuana* clade was 0.27 %, which was in the range of the average divergence observed among the other taxa of the large North American clade (0.25 %). Thus, mtDNA evidence supports the controversial notion of *P. martinezii* being distinct from other Mexican species, as recently suggested from the study of another mtDNA locus in Mexican spruces (Jaramillo-Correa et al., 2006).

Processes of range fragmentation or dispersion and isolation to achieve geographic speciation could be inferred in several instances for pairs of Asian taxa each involving one taxon confined to the coast or islands and the other located inland: *P. koraiensis* and *P. montigena*, *P. koyamai* and *P. meyeri*, *P. maximowiczii* and *P. wilsonii*, and *P. crassifolia* and *P. glehnii* (Fig. 3.2). These pairs would represent four events of geographical speciation at different times because of their current superimposed distributions. As for ancient lateral transfer of cpDNA, at least four Eurasian cases seem likely: for the parapatric hybridizing *P. obovata* and *P. abies* in northern Eurasia; for the allopatric *P. brachytyla* and *P. wilsonii* in China; for the allopatric *P. likiangensis* with *P. farreri* in Asia (no differences in cpDNA sequences); and for the parapatric *P. purpurea* with *P. wilsonii* (no differences in cpDNA sequences). In a pairwise fashion, these taxa were distantly related on the mtDNA tree, but were closely knitted on cpDNA trees. The most striking case is that of *P. likiangensis* and *P. farreri*, along with taxa from the large *P. schrenkiana* mtDNA clade sister to the North American clade, which are spread around the mountainous regions surrounding the Taklimakan Desert and Qinghai-Tibet Plateau in central Asia. While two long-recognized areas of plant endemism are recognized in Asia, 1- that of southern China and west to the Himalayas and 2- that of Japan, Korea, and northeastern China (Donoghue et al., 2001), the *schrenkiana* mtDNA clade corresponds to the first area. The lateral transfer of cpDNA among taxa of this clade would be relatively recent, after the raising of the Qinghai-Tibet plateau region and the establishment of deserts in central Asia during the Quaternary period, less than 1.6 million years ago (Zhang et al., 2000). These

topographical changes significantly influenced the climate (Ruddiman and Kutzbach, 1991) and direction of winds in central Asia. Consequently, pollen-mediated introgressive hybridization and resulting reticulate evolution of cpDNA would have been favoured by the relatively new circulation of summer air caused by the south-west monsoon (see Chang, 1983).

3.6.5 - Implications for other Pinaceae and conifers

Phylogenetic results indicate that repetitive reticulate evolution of the paternally inherited chloroplast genome of spruce is a major impediment for phylogenetical inference. Thus, for conifers possessing similar differential inheritance of their organelles such as in the Pinaceae, phylogenies inferred from the mitochondrial genome are likely to best approximate the true biogeographical history of the genus, beyond the intermingling history of characters derived from the nuclear or the chloroplast genome. Potent cases of geographical speciation at the intercontinental, continental, and regional scales could also be delineated from the spruce mtDNA phylogeny. Geographic speciation has also been suggested for some closely-related pines (e.g. Wang and Szmidt, 1994; Gernandt et al., 2005; Syring et al., 2007) and the trend appears to hold true for the smaller genus *Larix* when re-interpreting the phylogenies derived from cpDNA and mtDNA sequences (Gros-Louis et al., 2005) in the light of the present findings on spruce. Because of similar life-history and reproductive biology among conifers, it is anticipated that geographical speciation and reticulate evolution, as observed in the present study, should represent major driving forces in the evolution of most conifer genera.

3.7 - Acknowledgments

We thank T. Ward and J. Li (Arnold Arboretum, Boston, Massachusetts), E. Johnson (Holden Arboretum, Kirtland, Ohio), M. Chase and L. Csiba (Royal Botanic Garden, Kew, UK), F. T. Ledig (Univ. of California at Davis, Institute of Forest Genetics, USDA), J. Beaulieu and M. Deslauriers (Canadian Forest Service, Québec City), E. Morin (Botanical Garden of Montréal, Québec), D. Simpson and B. Daigle (National Tree Seed Centre,

Fredericton, New Brunswick), P. Brownless (Royal Botanic Gardens, Edinburgh, UK), and S. Toomer (Westonbirt National Arboretum, Tetbury, UK) for providing plant materials, M. Lamothe, and N. Tremblay (Canada Research Chair in Forest Genomics, Univ. Laval) for support in the laboratory, F. Larochelle (Center for Bioinformatics and Computational Biology, Univ. Laval) for support with softwares, and the Natural Sciences and Engineering Research Council of Canada for a grant to J. B.

3.8 - Litterature cited

BAKKER, F. T., A. CULHAM, C. E. PANKHURST, AND M. GIBBY. 2000. Mitochondrial and chloroplast DNA-based phylogeny of *Pelargonium* (Geraniaceae). Amer. J. Bot. 87, 727-734.

BENNIKE, O. 1990. The Kap København formation: stratigraphy and paleobotany of a Plio-Pleistocene sequence in Peary Land, North Greenland. Medd. Groenl. Geosci. 23, 1-85.

BENNUAH, S. Y., T. WANG, AND S. N. AITKEN. 2004. Genetic analysis of the *Picea sitchensis* x *glauca* introgression zone in British Columbia. For. Ecol. Manage. 197, 65-77.

BOUILLE, M., AND J. BOUSQUET. 2005. Trans-species shared polymorphisms at orthologous nuclear gene loci among distant species in the conifer *Picea* (Pinaceae): implications for the long-term maintenance of genetic diversity in trees. Amer. J. Bot. 92, 63-73.

BOUSQUET, J., S. H. STRAUSS, A. H. DOERKSEN, AND R. A. PRICE. 1992. Extensive variation in evolutionary rate of *rbcL* gene sequences among seed plants. Proc. Natl. Acad. Sci. USA. 89, 7844-7848.

BROWN, G. R., G. P. GILL, R. J. KUNTZ, C. H. LANGLEY, AND D. B. NEALE. 2004. Nucleotide diversity and linkage disequilibrium in loblolly pine. Proc. Natl. Acad. Sci. USA. 101, 15255-15260.

BROWN, J. H., AND M. V. LOMOLINO. 1998. Biogeography: second edition. Sinauer Associates, Inc. (Eds.). Sunderland, Massachusetts, pp. 232-243.

CHANG, D. H. S. 1983. The Tibetan Plateau in relation to the vegetation of China. Ann. MO. Bot. Gard. 70, 564-570.

CHAT, J., B. JAUREGUI, R .J. PETIT, AND S. NADOT. 2004. Reticulate evolution in kiwifruit (*Actinidia*, Actinidiaceae) identified by comparing their maternal and paternal phylogenies. Amer. J. Bot. 91, 736-747.

COMTOIS, P. 1997. Pollen dispersal and long distance transport: the case of thermophilic pollen in subarctic Canada. Aerobiologia 13, 37-42.

DONOGHUE, M. J., C. D. BELL, AND J. LI. 2001. Phylogenetic patterns in northern hemisphere plant geography. Int. J. Plant Sci. 162, S41–S52.

FARJON, A. 1990. Pinaceae: drawings and descriptions of the genera *Abies, Cedrus, Pseudolarix, Keteleeria, Nothotsuga, Tsuga, Cathaya, Pseudotsuga, Larix* and *Picea*. Regnum Veg. 121, 1-330.

FARJON, A. 2001. World checklist and bibliography of conifers, 2nd ed. Kew: Royal Botanical Gardens (Eds.), United Kingdom, pp. 150-167.

FARRIS, J. S., M. KÄLLERSJÖ, A. G. KLUGE, AND C. BULT. 1995. Testing significance of incongruence. Cladistics 10, 315-319.

FRANKIS, M. P. 1989. Genetic inter-relationships in Pinaceae. Notes R. Bot. Gard. (Edinb.) 45, 527-548.

FREUDENSTEIN, J. V., AND M. W. CHASE. 2001. Analysis of mitochondrial *nad1*b-c intron sequences in Orchidaceae: utility and coding of length-change characters. Syst. Bot. 26, 643-657.

GERMANO, J., AND A. S. KLEIN. 1999. Species-specific nuclear and chloroplast single nucleotide polymorphisms to distinguish *Picea glauca*, *P. mariana* and *P. rubens*. Theor. Appl. Genet. 99, 37-49.

GERNANDT, D. S., G. G. LOPEZ, S. O. GARCIA, AND A. LISTON. 2005. Phylogeny and classification of *Pinus*. Taxon 54, 29-42.

GORDON, A. G. 1976. The taxonomy and genetics of *Picea rubens* and its relationship to *Picea mariana*. Can. J. Bot. 54, 781–813.

GORDON, A. G. 1986. Breeding, genetics and genecological studies in spruce for tree improvement in 1983 and 1984, Sault Ste.Marie, Ontario. In: Proc. 20th Meet. Can. Tree Improve. Assoc., Quebec City, Quebec, 1985, Part I.

GROS-LOUIS, M.-C., J. BOUSQUET, L. E. PAQUES, AND N. ISABEL. 2005. Species-specific diagnostic markers in *Larix* spp. based on RAPDs and nuclear, cpDNA and mtDNA gene sequences, and their phylogenetic implications. Tree Genet. Genomes 1, 50-63.

HAMRICK, J. L., M. J. W. GODT, AND S. L. SHERMAN-BROYLES. 1992. Factors influencing levels of genetic diversity in woody plant species. New Forests 6, 95-124.

HAMZEH, M., AND S. DAYANANDAN. 2004. Phylogeny of *Populus* (Salicaceae) based on nucleotide sequences of chloroplast *tnrT-trnF* region and nuclear rDNA. Amer. J. Bot. 91, 1398-1408.

HUELSENBECK, J. P., AND F. RONQUIST. 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17, 754-755.

JARAMILLO-CORREA, J. P., J. BEAULIEU, AND J. BOUSQUET. 2001. Contrasting evolutionary forces driving population structure at expressed sequence tag polymorphisms, allozymes and quantitative traits in white spruce. Mol. Ecol. 10, 2729-2740.

JARAMILLO-CORREA, J. P., AND J. BOUSQUET. 2003. New evidence from mitochondrial DNA of a progenitor-derivative species relationship between black spruce and red spruce (Pinaceae). Amer. J. Bot. 90, 1801-1806.

JARAMILLO-CORREA, J. P., J. BOUSQUET, J. BEAULIEU, N. ISABEL, M. PERRON, AND M. BOUILLE. 2003. Cross-species amplification of mitochondrial DNA sequence-tagged-site markers in conifers: the nature of polymorphism and variation within and among species in *Picea*. Theor. Appl. Genet. 106, 1353-1367.

JARAMILLO-CORREA, J. P., J. BEAULIEU, F. T. LEDIG, AND J. BOUSQUET. 2006. Decoupled mitochondrial and chloroplast DNA population structure reveals Holocene collapse and population isolation in a threatened Mexican-endemic conifer. Mol. Ecol. 15, 2787-2800.

JOHNSON, L. A., AND D. E. SOLTIS. 1995. Phylogenetic inference in Saxifragaceae *sensu stricto* and *Gilia* (Polemoniaceae) using *matK* sequences. Ann. MO. Bot. Gard. 82, 149-175.

KONISHI, S., AND M. SUZUKI. 1997. Analysis of cone morphology of *Picea glehnii* (Pinaceae). Jpn. J. of Histor. Bot. 5, 67-76 (in Japanese), *in* Kobayashi, K., Yoshikawa, J., Suzuki, M., 2000. DNA identification of *Picea* species of the Last Glacial Age in northern Japan. Jpn. J. Hist. Bot. 8, 67-80.

KRUTOVSKII, K. V., AND F. BERGMANN. 1995. Introgressive hybridization and phylogenetic relationships between Norway, *Picea abies* (L.) Karst., and Siberian, *P. obovata* Ledeb., spruce species studied by isozyme loci. Hered. 74, 464-480.

LEDIG, F. T., P. D. HODGSKISS, K. V. KRUTOVSKII, D. B. NEALE, AND T. EGUILUZ-PIEDRA. 2004. Relationships among the spruces (*Picea*, Pinaceae) of Southwestern North America. Syst. Bot. 29, 275-295.

LEPAGE, B. A. 2001. New species of *Picea* A. Dietrich (Pinaceae) from the middle Eocene of Axel Heiberg Island, Arctic Canada. Bot. J. Linn. Soc. 135, 137-167.

LEPAGE, B. A. 2003. A new species of *Tsuga* (Pinaceae) from the middle Eocene of Axel Heiberg Island, Canada, and an assessment of the evolution and biogeographical history of the genus. Bot. J. Linn. Soc. 141, 257-296.

LIU, T. S. 1982. A new proposal for the classification of the genus *Picea*. Acta Phytotax. Geobot. 33, 227-244.

MA, X.-F., A. E. SZMIDT, AND X.-R. WANG. 2006. Genetic structure and evolutionary history of a diploid hybrid pine *Pinus densata* inferred from the nucleotide vaiation at seven gene loci. Mol. Biol. Evol. 23, 807-816.

MARSHALL, H.D., C. NEWTON, AND K. RITLAND. 2001. Sequence-repeat polymorphisms exhibit the signature of recombination in lodgepole pine chloroplast DNA. Mol. Biol. Evol. 18, 2136-2138.

MIKKOLA, L. 1969. Observations on interspecific sterility in *Picea*. Ann. Bot. Fenn. 6, 285-339.

MITTON, J. B., AND R. ANDALORA. 1981. Genetic and morphological relationships between blue spruce, *Picea pungens*, and Engelmann spruce, *Picea engelmannii*, in the Colorado Front Range. Can. J. Bot. 59, 2088-2094.

MOGENSEN, H. L. 1996. The hows and whys of cytoplasmic inheritance in seed plants. Amer. J. Bot. 83, 383-404.

NYLANDER, J. A. A. 2004. MrModeltest v2. Program distributed by the author. Evolutionary Biology Centre, Uppsala Univ.

O. E. C. D. 1999. Consensus document on the biology of *Picea glauca* (Moench) Voss (white spruce). Series on Harmonization of Regulatory Oversight in Biotechnology No. 13, eds. Joint Meet. Chemicals Committee and Working Party on Chemicals (Environ. Direct., OECD Environ. Health and Safety Publ.), Paris, France, 45 pp.

OKUYAMA, Y., N. FUJII, M. WAKABAYASHI, A. KAWAKITA, M. ITO, M. WATANABE, N. MURAKAMI, AND M. KATO. 2004. Nonuniform concerted evolution and chloroplast capture: heterogeneity of observed introgression patterns in three molecular data partition phylogenies of Asian *Mitella* (Saxifragaceae). Mol. Biol. Evol. 22, 285-296.

PAGE, C. N., AND R. C. HOLLANDS. 1987. The taxonomic and biogeographic position of Sitka spruce. Proc. Roy. Soc. Edinburgh 93B, 13-24.

PERRON, M., AND J. BOUSQUET. 1997. Natural hybridization between black spruce and red spruce. Mol. Ecol. 6, 725-734.

PERRON, M., D. J. PERRY, C. ANDALO, AND J. BOUSQUET. 2000. Evidence from sequence-tagged-site markers of a recent progenitor-derivative species pair in conifers. Proc. Natl. Acad. Sci. USA. 97, 11331- 11336.

POSADA, D., AND K. A CRANDALL. 1998. Modeltest: testing the model of DNA substitution. Bioinformatics 14, 817-818.

QIAO, C.-Y., J.-H. RAN, Y. LI, AND X.-Q. WANG. 2007. Phylogeny and biogeography of *Cedrus* (Pinaceae) inferred from sequences of seven paternal chloroplast and maternal mitochondrial DNA regions. Ann. Bot. 100, 573-580.

QUINN, C. J., R. A. PRICE, AND P. A. GADEK. 2002. Familial concepts and relationships in the conifers based on *rbcL* and *matK* sequence comparisons. Kew Bull. 57, 513-531.

RAJORA, O. P., AND B. P. DANCIK. 2000. Population genetic variation, structure, and evolution in Engelmann spruce, white spruce, and their natural hybrid complex in Alberta. Can. J. Bot. 78, 768-780.

RAN, J.-H., X.-X. WEI, AND X.-Q. WANG. 2006. Molecular phylogeny and biogeography of *Picea* (Pinaceae): Implications for phylogeographical studies using cytoplasmic haplotypes. Mol. Phylogenet. Evol. 41, 405-419.

RIESEBERG, L. H., AND D. E. SOLTIS. 1991. Phylogenetic consequences of cytoplasmic gene flow in plants. Evol. Trends Plants 5, 65-84.

RUDDIMAN, W. F., AND J. E. KUTZBACH. 1991. Plateau uplift and climatic change. Sci. Amer. March, 66-75.

SAITOU, N., AND M. NEI. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4, 406-425.

SCHMIDT-VOGT, H. 1977. Die Fichte. Book 1. Taxonomie, Verbreitung, Morphologie, Ökologie, Waldgesellschaften. Paul Parey Ed. Hamburg and Berlin, Germany. 647 pp.

SEELANAN, T., A. SCHNABEL, AND J. F. WENDEL. 1997. Congruence and consensus in the Cotton tribe (Malvaceae). Syst. Bot. 22, 259-290.

SEMERIKOV, V. L., H. ZHANG, M. SUN, AND M. LASCOUX. 2003. Conflicting phylogenies of *Larix* (Pinaceae) based on cytoplasmic and nuclear DNA. Mol. Phylogenet. Evol. 27, 173-184.

SIGURGEIRSSON, A., AND A. E. SZMIDT. 1993. Phylogenetic and biogeographic implications of chloroplast DNA variation in *Picea*. Nord. J. Bot. 13, 233-246.

SOLTIS, D. E., P. S. SOLTIS, T. G. COLLIER, AND M. L. EDGERTON. 1991. Chloroplast DNA variation within and among genera of the *Heuchera* group (Saxifragaceae): evidence for chloroplast transfer and paraphyly. Amer. J. Bot. 78, 1091-1112.

SORENSON, M. D. 1999. TreeRot, version 2. Boston Univ., Boston, Massachusetts.

SUTTON, B. C. S., D. J. FLANAGAN, J. R. GAWLEY, C. H. NEWTON, D. T. LESTER, AND Y. A. EL-KASSABY. 1991. Inheritance of chloroplast and mitochondrial DNA in *Picea* and composition of hybrids from introgression zones. Theor. Appl. Genet. 82, 242-248.

SWOFFORD, D. L. 2001. PAUP*. Phylogenetic Analysis Using Parsimony (* and Other Methods). Version 4.0b8. Sinauer Associates, Inc., Sunderland, Massachusetts.

SYRING, J., R. F. DEL CASTILLO, R. CRONN, AND A. LISTON. 2007. Multiple nuclear loci reveal the distinctiveness of the threatened, neotropical *Pinus chiapensis*. Syst. Bot. 32, 703-717.

TABERLET, P., L. GIELLY, G. PAUTOU, AND J. BOUVET. 1991. Universal primers for amplification of three non-coding regions of chloroplast DNA. Plant Mol. Biol. 17, 1105-1109.

TAYLOR, R. J., AND T. F. PATTERSON. 1980. Biosystematics of Mexican spruce species and populations. Taxon 29, 421-469.

TEMPLETON, A. R. 1983. Phylogenetic inference from restriction endonuclease cleavage site maps with particular reference to the evolution of humans and apes. Evol. 37, 221-244.

WANG, X.-R., AND A. E. SZMIDT. 1994. Hybridization in chloroplast DNA variation in a *Pinus* species complex from Asia. Evolution 48, 1020-1031.

WANG, X.-R., Y. TSUMURA, H. YOSHIMARU, K. NAGASAKA, AND A. E. SZMIDT. 1999. Phylogenetic relationships of Eurasian pines (*Pinus*, Pinaceae) based on chloroplast *rbcL, matK, rpl20-rps18* spacer, and *trnV* intron sequences. Amer. J. Bot. 86, 1742-1753.

WENG, C., AND S. T. JACKSON. 2000. Species differentiation of North American spruce (*Picea*) based on morphological and anatomical characteristics of needles. Can. J. Bot. 78, 1367-1383.

WEI, X.-X., AND X.-Q WANG. 2003. Phylogenetic split of *Larix*: evidence from paternally inherited cpDNA *trn*T-*trn*F region. Plant Syst. Evol. 239, 67-77.

WRIGHT, J. W. 1955. Species crossability in spruce in relation to distribution and taxonomy. For. Sci. 1, 319-349.

ZHANG, D., L. FENGQUAN, AND B. JIANMIN. 2000. Eco-environmental effects of the Qinghai-Tibet Plateau uplift during the Quaternary in China. Envir. Geol. 39, 1352-1358.

3.9 - Appendix 1

Voucher information and GenBank accession numbers for taxa used in this study. Sample origin: AA: Arnold Arboretum (Massachusetts, U.S.A.); CFS: Canadian Forest Service (Québec [CFL], Canada); CEF: Centre d'étude de la forêt (Québec, Canada); HA: Holden Arboretum (Ohio, U.S.A.); IFG: Institute of Forest Genetics, U.S.D.A. Forest Service (U.S.A.); JBM: Jardin Botanique de Montréal (Montréal, Canada); KEW: Royal Botanic Gardens, Kew (U.K.); LUS: Lushan Mountain (China); NTSC: National Tree Seed Centre (New Brunswick, Canada); RBG: Royal Botanic Gardens, Edinburgh (U.K.); WNA: The Westonbirt National Arboretum (Tetbury, U.K.); WNS: Western Native Seeds (U.S.A.).

Taxon—GenBank accessions: *rbcL*, *trnK*, *trnTLF*, *nad1* intron 2 **(in process)**; natural origin of specimen or natural distribution of the species for unprecise locality; source of sample; *voucher number*. Dash in voucher number indicate collected in the wild.

Picea abies (L.) Karsten—rbcL, trnK, trnTLF, EF440449; Poland; CFS; *5433*.

P. abies— EF440570, EF440492, EF440530, EF440450; Romania; CFS; *5390*.

P. abies— - , - , trnTLF, EF440451; distrib.: Europe; CEF; *F8789*.

P. alcoquiana (Veitch *ex* Lindley) Carrière— EF440571, EF440493, EF440532, EF440452; distrib.: Japan; AA; *5317*A*.

P. asperata Masters— EF440572, EF440494, EF440533, EF440453; China: Szechwan; LUS; —.

P. brachytyla (Franchet) Pritzel— EF440573, EF440495, EF440534, EF440454; China; KEW; *1980-365*.

P. breweriana S. Watson— EF440574, EF440496, EF440535, EF440455; U.S.A.: Oregon; RBG; *19651207*.

P. chihuahuana Martínez— EF440575, EF440497, EF440536, EF440456; Mexico; AA; *315-88*G*.

P. crassifolia Komarov— EF440576, EF440498, EF440537, EF440457; China: Qinghai; RBG; *19802500*.

P. engelmannii Parry ex Engelmann— EF440577, EF440499, trnTLF, EF440458; U.S.A.: Montana; NSS; *7186160.0*.

P. engelmannii—rbcL, trnK, EF440538, -; Canada: Alberta; NSS; *9160010.0*.

P. farreri Page & Rushforth— EF440578, EF440500, EF440539, EF440459; Burma; WNA; *1990/367*.

P. glauca (Moench) Voss— EF440579, EF440501, trnTLF, EF440460; Canada: Ontario; CFS; *2856-100*.

P. glauca—rbcL, trnK, trnTLF, EF440461; Canada: Manitoba; CFS; *P445-B3-1*.

P. glauca—rbcL, trnK, trnTLF, EF440462; Canada: Newfoundland; CFS; *P8-B3-1*.

P. glauca— - , - , EF440542, - ; Canada: British-Columbia; CFS.

P. glehnii (Fr. Schmidt) Masters— EF440580, EF440502, EF440543, EF440463; Japan: Hokkaido; AA; *392-86*A*.

P. jezoensis (Siebold & Zuccarini) Carrière— EF440581, EF44053, EF440544, EF440464; Japan: Hokkaido; AA; *925-86*.

P. jezoensis—*rbcL*, *trnK*, *trnTLF*, -; Central Korea; AA; *502-77*A*.

P. jezoensis—*rbcL*, *trnK*, *trnTLF*, EF440465; China: Liaoning; AA; *1275-79*A*.

P. koraiensis Nakai— EF440582, EF440504, EF440545, EF440466; China: Liaoning; AA; *1276-79*A*.

P. koyamai Shirasawa— EF440583, EF440505, EF440546, EF440467; distrib.: Japan, Honshu; HA; *77-272*.

P. likiangensis (Franchet) Pritzel— EF440584, EF440506, EF440547, EF440468; China; KEW; *1995-3850*.

P. mariana (Miller) Britton, Sterns, & Poggenberg— EF440585, EF440507, EF440548, EF440469; U.S.A.: Wisconsin; CFS; *3293-5*.

P. mariana—*rbcL*, *trnK*, *trnTLF*, *nad1* intron 2; Canada: Québec; CFS; *340-6*.

P. mariana—*rbcL*, *trnK*, *trnTLF*, *nad1* intron 2; Canada: Western Ontario; CFS; *6927-1*.

P. martinezii T. F. Patterson— EF440586, EF440508, EF440549, EF440470; Mexico: Nuevo Leon; IFG; *Lot C 1994*.

P. maximowiczii Regel ex Masters— EF440587, EF440509, EF440550, EF440471 [a]; distrib.: Japan; RBG; *19795115*.

P. mexicana Martínez— EF440588, EF440510, EF440551, EF440472; Mexico: Kohahuila; IFG; *Lot DW 1997*.

P. meyeri Rehder & Wilson— EF440589, EF440511, EF440552, EF440473; China: Szechwan; LUS; —.

P. montigena Masters— EF440590, EF440512, EF440553, EF440474 [a]; China; KEW; *1957-32503.*

P. morrisonicola Hayata— EF440591, EF440513, EF440554, EF440475; Taiwan; RBG; *19924278.*

P. obovata Ledebour— EF440592, EF440514, EF440555, EF440476; distrib.: Northern Eurasia; HA; *60-235.*

P. omorika (Pancic) Purkyne— EF440593, EF440515, EF440556, EF440477; distrib.: Serbia/Bosnia; AA; *1496-1*A.*

P. orientalis (L.) Link— EF440594, EF440516, EF440557, EF440478 [a]; Georgia: Caucasus Mnts; AA; *17639*C.*

P. pungens Engelmann—*rbcL, trnK*, - , EF440481; U.S.A.: New Mexico; NSS; *8285450.*

P. pungens— EF440597, EF440519, EF440560, *nad1* intron 2; U.S.A.: Colorado; WNS; —.

P. pungens— - , - , - , *nad1* intron 2; U.S.A.: Arizona; NSS; *6080054.*

P. purpurea Masters— EF440598, EF440520, EF440561, EF440482; China: W. Szechwan; AA; *6995*H.*

P. retroflexa Masters— EF440599, EF440521, EF440562, EF440483; distrib.: China; RBG; *19698147.*

P. rubens Sargent— EF440600, EF440522, EF440563, EF440484; Appalachian Mnts; CEF; *691-8.*

P. schrenkiana Fisher & Meyer— EF440601, EF440523, EF440564, EF440485; China: Xinjiang; AA; *1278-79*D.*

P. sitchensis (Bongard) Carrière— EF440602, EF440524, EF440565, EF440486; U.S.A.: Alaska; NSS; *7083840.*

P. sitchensis—*rbcL, trnK, trnTLF, nad1* intron 2; Canada: British-Columbia; NSS; *9070011.*

P. sitchensis— - , EF440525, *trnTLF*, EF440487; U.S.A.: Oregon; NSS; *6881140.*

P. smithiana (Wallich) Boissier— EF440603, EF440526, EF440566, EF440488; Afghanistan; RBG; *19711566.*

P. spinulosa (Griffith) Henry— EF440604, EF440527, EF440567, EF440489; Bhutan; RBG; *19792551.*

P. torano (Siebold ex Koch) Koehne— EF440605, EF440528, EF440568, EF440490; Japan; HA; *77-270.*

P. wilsonnii Masters— EF440606, EF440529, EF440569, EF440491; China: Hubei; LUS; —.

Pinus banksiana Lamb. (outgroup)— EF440595, EF440517, EF440558, EF440479; distrib.: Canada & U.S.A.; CEF; —.

Pinus koraiensis Siebold & Zuccarini (outgroup)— EF440596, EF440518, EF440559, EF440480 2[a]; distrib.: Eastern Asia; CFS; —.

---

[a] Partial sequence

Chapitre 4

Conclusions et perspectives

Il n'existe pas de consensus absolu entre les divers spécialistes quant à l'histoire évolutive du genre *Picea*. Les études basées sur la morphologie, l'anatomie et la reproduction de ces espèces ont déjà été décortiquées, comparées, discutées et critiquées par ceux qui ont bâti les nombreuses classifications du genre *Picea*. D'ailleurs, il est toujours ardu d'effectuer des regroupements taxonomiques à l'intérieur de ce genre, notamment à cause de la méconnaissance des variations morphologiques intraspécifiques et de l'évolution réticulée.

Étant donné la confusion, de plus récentes études moléculaires ont voulu participer à éclairer l'histoire évolutive des *Picea* (*e.g.* Sigurgeirsson et Szmidt 1993, Ran *et al.* 2006). En théorie, les caractères moléculaires peuvent apporter une vision objective des liens de parenté entre les *Picea*, de la hiérarchie évolutive qui sous-tend ces liens, et de l'histoire biogéographique de ce groupe de conifères. En effet, les phylogénies moléculaires corroborent habituellement les essais de classification des plantes et facilitent les études comparatives sur le développement, le métabolisme, la reproduction, la pathologie et la génomique des plantes (Kuzoff et Gasser 2000). C'est dans autant de domaines que l'histoire évolutive particulière des *Picea* peut donner des assises supplémentaires pour mieux conserver les espèces et, ultimement, permettre d'améliorer la quantité et la qualité des produits forestiers découlant des *Picea*.

4.1 - Phylogénie des *Picea* basée sur les gènes nucléaires

Pour tracer l'histoire évolutive des *Picea* dans le cadre de notre étude, le choix des marqueurs phylogénétiques s'est d'abord arrêté sur le génome nucléaire. Les régions les plus utilisées pour estimer les phylogénies de genres d'angiospermes, soit les régions nucléaires *ITS-1* et *ITS-2*, ont présenté des obstacles à l'estimation de la

phylogénie du genre *Picea* en entier, à cause de la complexité des patrons d'évolution des *Picea* pour la région *ITS-1* (Campbell *et al.* 2005), et du peu de variations interspécifiques pour la région *ITS-2* (M. Bouillé, résultats non publiés).

Devant l'impasse avec ces deux régions, d'autres loci nucléaires ont été considérés en priorité, car les régions connues des génomes d'organelles ne présentaient en général que peu de variations chez *Picea* (*e.g.* Germano et Klein 1999). Les résultats relatifs à cette considération ont été présentés au deuxième chapitre de la présente thèse. Les trois loci choisis pour leur variabilité interspécifique adéquate présentaient aussi des polymorphismes intraspécifiques qui se sont avérés trans-spécifiques. Deux types de polymorphismes trans-spécifiques ont été remarqués, dont ceux qui proviendraient de l'ancêtre commun aux trois espèces de *Picea* étudiées et ceux, plus rares, qui tireraient leur origine de la recombinaison intragénique ancienne.

Les diverses méthodes d'estimation de la divergence, soit des extrapolations à partir des données fossiles et de certains taux de divergence, ont permis de déterminer l'âge minimal de coalescence des allèles nucléaires séquencés. Cet âge s'est avéré correspondre à un temps de coalescence qui précédait la divergence des trois espèces *P. abies*, *P. glauca* et *P. mariana*. En effet, les données fossiles ramènent à au moins 15 à 23 mA la période de divergence des lignées ancestrales menant aux espèces contemporaines, tandis que les calculs basés sur des données moléculaires et biochimiques ont établi ici cette divergence à 17 à 20 mA d'une part, et 13 à 16 mA d'autre part. Donc, c'était un point de comparaison englobant une période de 13 à 20 mA qui avait été établi, et la frontière inférieure du résultat de l'âge minimal de coalescence, estimé à 10 à 18 mA, se situait même au-delà de la médiane de cette période. Donc, la coalescence des allèles pouvait précéder celle des trois espèces.

Aussi, l'hypothèse selon laquelle les polymorphismes ancestraux des gènes neutres ou presque neutres seraient plus facilement conservés dans les grandes tailles de population a été corroborée, car la taille effective historique des populations ancestrales pour ces loci et

ces espèces de *Picea* a été évaluée minimalement à plus de 100 000 individus. Cette mise en évidence de très longues périodes de coalescence pour les allèles de gènes nucléaires chez *Picea* constitue l'un des premiers cas les mieux documentés chez les plantes. De manière générale, ces résultats seraient prévisibles chez d'autres groupes de plantes de même type et qui affichent de grandes tailles de populations, tels que d'autres genres à l'intérieur de la famille des Pinaceae et certains groupes de feuillus parmi les angiospermes anémophiles et allogames.

Enfin, à cause de la présence de ces polymorphismes trans-spécifiques qui ont été maintenus dans de grandes populations ancestrales, ces gènes nucléaires se sont révélés inadéquats pour estimer une phylogénie des espèces de *Picea* qui représenterait l'histoire biogéographique du genre. Par conséquent, les génomes d'organelles, malgré leur fort degré de conservation au niveau de la séquence des gènes, ont dû être envisagés pour atteindre cet objectif. Les résultats ont été rapportés au troisième chapitre de la thèse.

4.2 - L'histoire biogéographique du genre *Picea*

Le génome chloroplastique avait déjà été utilisé lors de deux études phylogénétiques du genre *Picea*: celle de Sigurgeirsson et Szmidt (1993) et la toute récente étude de Ran *et al.* (2006). Cette dernière n'était pas encore publiée lors de l'amorce des travaux de cette partie de la thèse. Sans le démontrer hors de tout doute, Sigurgeirsson et Szmidt (1993) mentionnaient déjà certaines inquiétudes associées aux études phylogéographiques du genre *Picea* à partir du génome chloroplastique, telles que la possibilité de recombinaison durant l'évolution des ADN chloroplastiques des *Picea*, l'homologie difficilement démontrée pour les fragments RFLP de même longueur, ainsi que les événements d'introgression qui peuvent provoquer la réticulation en raison de la transmission paternelle du génome chloroplastique chez la majorité des Pinaceae. La méthodologie employée par Ran *et al.* (2006) a corrigé le problème de l'homologie des fragments RFLP, puisque des séquences d'ADN de quelques régions du génome chloroplastique ont été utilisées. Cependant, comme l'indiquaient Sigurgeirsson et Szmidt (1993), la validité des phylogénies chloroplastiques des *Picea* devrait être vérifiée en

comparant celles-ci avec des phylogénies découlant de régions du génome mitochondrial, lequel est moins sujet à l'évolution réticulée.

Les résultats du présent travail sur le génome chloroplastique s'apparentent aux résultats des études de Sigurgeirsson et Szmidt (1993) et de Ran *et al.* (2006) en ce que les phylogénies résultantes montrent un enchevêtrement complexe entre taxons américains et asiatiques. De là, on peut déduire un rapprochement au niveau chloroplastique entre les *Picea* de différents continents, et on pourrait en inférer de multiples migrations intercontinentales. Aussi, la comparaison des différentes partitions chloroplastiques a nouvellement démontré ici la possibilité de recombinaison intragénomique ancienne pour les *Picea*, laquelle avait déjà été observée chez les pins par Marshall *et al.* (2001).

Puisque l'hybridation et l'introgression anciennes auraient pu jouer un rôle important dans l'évolution réticulée du genre *Picea*, des discordances étaient attendues entre les phylogénies chloroplastiques et la phylogénie mitochondriale. Aussi, contrairement aux phylogénies chloroplastiques, la phylogénie mitochondriale a montré une très forte cohérence géographique. Par conséquent, les histoires biogéographiques se sont avérées discordantes entre les deux génomes, tel qu'anticipé. Les résultats de la comparaison des deux types de phylogénies suggèrent que le transfert latéral du génome chloroplastique serait survenu entre plusieurs taxons tout au long de l'histoire du genre *Picea*, créant ainsi un patron clair d'évolution réticulée au niveau des phylogénies chloroplastiques.

Outre les patrons d'évolution réticulée, certaines conclusions ont pu être tirées quant à l'histoire biogéographique des *Picea* en tant que telle. Tout d'abord, la spéciation de type géographique des *Picea*, ainsi qu'elle avait été suggérée par Wright (1955), a été clairement démontrée, puisqu'elle se reflétait par la cohérence géographique à l'intérieur de la phylogénie mitochondriale. De plus, à partir d'une origine asiatique du genre, appuyée par les présents résultats et les hypothèses de Wright (1955), il y aurait possiblement eu deux événements distincts de dispersion des *Picea* contemporains de l'Asie vers l'Amérique. Ensuite, un autre événement de colonisation beaucoup plus récent serait survenu entre

l'Amérique et l'Europe, et qui pourrait impliquer une dispersion de l'ouest vers l'est. En ce qui a trait au coeur du continent asiatique, la proximité phylogénétique de certains taxons insulaires avec des taxons continentaux tend à démontrer que les événements de migration et de spéciation géographique auraient été nombreux, ce qui rejoint également les hypothèses de Wright (1955).

En ce qui concerne les regroupements phylogénétiques, certains résultats offrent un regard nouveau sur les liens entre les espèces de *Picea*. Par exemple, la situation phylogénétique relative de *Picea pungens* ne semble correspondre à aucune hypothèse établie jusqu'ici, et son rapprochement phylogénétique avec certains taxons du sud de l'Amérique du Nord représente une nouvelle inférence. Par ailleurs, le cas particulier du contact secondaire de *P. abies* et *P. obovata*, lesquels présentent une vaste zone d'hybridation, a pu être vérifié grâce à la comparaison des phylogénies d'organelles. Sauf en ce qui a trait à ce couple d'espèces qui forment des hybrides naturels, les autres paires de taxons de *Picea* qui s'hybrident ou qui présentent de l'introgression se retrouvaient dans les mêmes clades dans la phylogénie mitochondriale, tel qu'attendu.

Enfin, aucune correspondance n'a été relevée entre les phylogénies résultantes et les classifications consultées. À ce sujet, les taxons de *Picea* qui présentent de l'hybridation naturelle ou de l'introgression ne se retrouvent pas dans les mêmes subdivisions des classifications consultées. En fait, aucune des classifications traditionnelles consultées ne semblait montrer des regroupements géographiques naturels qu'on pourrait s'attendre de voir pour un groupe d'arbres de si vaste répartition géographique que les *Picea*. Cette constatation porte à croire que les caractères morphologiques utilisés ne reflètent possiblement pas les lignées ancestrales et les phénomènes de spéciation du genre, mais qu'ils reflètent plutôt l'évolution réticulée. Celle-ci caractériserait non seulement le génome chloroplastique, mais aussi le génome nucléaire. Ceci signifie que plusieurs caractères morphologiques pourraient être convergents. En conséquence, ceux-ci seraient peu indicatifs de la divergence ancestrale telle qu'imaginée dans un contexte d'identification des lignées évolutives. De là, on peut identifier ce phénomène comme une source potentielle du

manque de consensus entre les spécialistes de la taxonomie.

Les constatations générales faites ici pour le genre *Picea* sont probablement applicables aux autres genres de Pinaceae, puisque tous les genres de cette famille présentent une transmission maternelle du génome mitochondrial, alors que le génome chloroplastique est transmis paternellement via le pollen (Mogensen 1996).

4.3 - Les perspectives

4.3.1 - Les études phylogénétiques

L'importance de la variabilité morphologique entre les espèces de *Picea* et à l'intérieur de celles-ci demeure inconnue. Aussi, comme pour les autres genres de plantes, parmi les causes de variabilité morphologique figurent la convergence de certains caractères ainsi que l'influence de l'environnement. En général, on pourrait s'attendre à ce qu'une proximité génétique plus importante sous-tende une moins grande divergence de phénotype en milieu contrôlé (Rieseberg et Wendel 1993). Quoiqu'il serait envisageable de tenter de déterminer l'influence des facteurs abiotiques sur le phénotype des *Picea* en milieu contrôlé, il est peu probable que cette méthodologie offre une solution rapide à court terme. Il serait donc plus réaliste d'accumuler les informations supplémentaires directement sur le terrain (*e.g.* Turna 2004) pour évaluer l'importance de la variation morphologique ainsi que les variations morphologiques reliées aux facteurs abiotiques. Il serait ainsi possible de déterminer, entre autres, si l'importance de la plasticité phénotypique varie proportionnellement à la taille de l'aire de répartition des espèces de *Picea* (voir Gordon 1976, Liu 1982, Konishi et Suzuki 1997 [*in* Kobayashi *et al.* 2000], Rajora et Dancik 2000) et de mieux préciser la valeur taxonomique des caractères morphologiques distincts pour des taxons sympatriques. Parallèlement, la recherche des marqueurs moléculaires indicatifs des variations morphologiques peu influencées par l'environnement pourrait aussi être utile au niveau phylogénétique.

Également, l'ajout de marqueurs phylogénétiques mitochondriaux additionnels, par exemple d'autres introns de la région *nad1*, permettrait d'améliorer la résolution de la phylogénie mitochondriale décrite ci-haut, surtout en ce qui concerne le groupe des lignées asiatiques à la base de l'arbre phylogénétique. De cette phylogénie mieux résolue, la pertinence des caractères morphologiques traditionnels pourrait être établie quant à leur valeur indicative de la phylogéographie (*e.g.* Jeandroz *et al.* 1997).

La divergence génétique relative des taxons de *Picea* pourrait également être évaluée plus précisément avec des marqueurs phylogénétiques mitochondriaux supplémentaires. Les degrés de divergence relatifs permettraient de préciser la taxonomie des *Picea* en ce qui concerne les liens phylogénétiques plus fins entre les espèces proches parentes en Asie, à savoir la révision de leur statut d'espèce, de sous-espèce ou de variété, tel que suggéré par Vogellehner (1977 [*in* Vidacovic 1991]). À partir des lignées ancestrales en Asie déterminées avec plus de précision, les patrons de spéciation entre les taxons asiatiques insulaires pourraient être mieux inférés.

Avec une phylogénie plus précise, basée sur davantage de caractères moléculaires du génome mitochondrial, il serait aussi possible de dater les périodes de divergence des principales lignées ancestrales ainsi que l'époque des migrations intercontinentales. En effet, les estimations d'horloges moléculaires peuvent aider à corroborer les estimations paléontologiques et mieux déterminer l'époque de chacune de ces dispersions (*e.g.* Xiang *et al.* 2000, Xiang *et al.* 2005). À ce titre, les approches bayesiennes apparaissent des plus appropriées, car elles permettent d'analyser simultanément la topologie, la longueur relative des branches d'une phylogénie ainsi que l'information temporelle, afin d'estimer de manière plus juste les périodes de divergence (Renner 2005). Les données moléculaires et fossiles doivent être simultanément prises en compte pour qu'un tel exercice soit fructueux, et elles sont maintenant plus facilement accessibles.

Pour ce qui entoure la grande majorité des taxons de *Picea*, l'assignation d'un individu à une espèce est maintenant rendue possible par l'utilisation des marqueurs

développés ici. C'est donc dire que les individus situés *ex situ*, comme dans les jardins botaniques ou les arboretums, peuvent maintenant être identifiés de manière plus sûre à l'aide des marqueurs mitochondriaux de la région *nad1*. Possiblement qu'avec des marqueurs mitochondriaux supplémentaires, l'assignation des individus à leur sous-espèce ou variété pourra même être effectuée.

De nouvelles informations moléculaires pourraient être découvertes en ciblant le génome nucléaire. Une phylogénie bien résolue d'origine nucléaire pourrait permettre de mieux évaluer le degré de divergence ancestrale relative entre les grands groupes de *Picea*, par exemple entre les groupes nord-américains d'origine distincte. Globalement, de nouveaux marqueurs pourraient être identifiés parmi les gènes en faible nombre de copies (*e.g.* Syring *et al.* 2005), ou en consultant les bases de données de séquençage de masse (*e.g.* projet ForestTree DB du groupe Arborea 2007, Pavy *et al.* 2007), ou encore en développant une approche basée sur des marqueurs AFLPs nucléaires desquels seraient identifiées les variations interspécifiques. Il s'agit d'ailleurs d'une méthode qui a déjà été employée avec succès pour des espèces de proche parenté ou encore au niveau intraspécifique (*e.g.* Després *et al.* 2003).

En ce qui a trait à la systématique, c'est l'ajout de nouvelles connaissances sur le genre *Picea* qui permet généralement de confirmer les balises de délimitation des espèces avec plus de certitude, et qui permet de découvrir pourquoi et en quoi les liens sont complexes entre ces espèces. Miroir de l'évolution naturelle des espèces, la classification des *Picea*, comme celle de tout autre groupe de plante, devrait être basée sur les deux dimensions rattachées à la systématique évolutive: les liens généalogiques (la cladistique), et les liens établis entre les espèces contemporaines en termes de degrés de similitudes et de différences (la taxonomie classique) (Mayr 1995, 1997 [*in* Grant 1998]). Les liens généalogiques ont été traités ici, quoique partiellement puisque le génome nucléaire n'a pas encore apporté sa contribution cladistique. Certainement, de futurs traitements taxonomiques suivront pour établir les liens entre les espèces contemporaines, grâce à de nouvelles études comparatives plus élaborées qui fourniront davantage d'indices de

classification. L'évolution réticulée à l'intérieur du genre *Picea* étant une réalité, tel qu'observé lors du présent travail, il faudra savoir composer avec ses effets globaux sur les génotypes et les phénotypes observables aujourd'hui. C'est donc dire qu'il faut aussi considérer la taxonomie des *Picea* dans un contexte global, qui inclut les indices pertinents accumulés jusqu'ici, en assouplissant la rigidité méthodologique de la caractérisation morphologique seule, ou de la caractérisation cladistique seule. L'une des conciliations devant être faite relève des phénomènes de l'hybridation et de l'introgression. En effet, l'hybridation et l'introgression, qui ont vraisemblablement beaucoup participé à l'évolution du genre, contribuent à augmenter la diversité génétique et taxonomique d'un groupe de plantes (Rieseberg et Wendel 1993; Perron et Bousquet 1997). Ces phénomènes évolutifs façonnent de nouveaux génotypes et phénotypes, car les hybrides ou les introgressants portent un mélange de caractères parentaux, intermédiaires et extrêmes (Rieseberg et Wendel 1993). Ces types particuliers gagneront à être davantage étudiés, afin que les caractères phénotypiques qu'ils portent puissent obtenir consensus quant à leur pertinence dans la classification morphologique. Des études élaborées basées sur des gènes nucléaires pourront toutefois offrir un portrait global des liens entre les *Picea*, puisque les effets de l'hybridation et de l'introgression, observables lors de l'étude des organelles, paraîtront plus dilués dans le génome nucléaire. En effet, les gènes nucléaires sembleraient généralement moins vulnérables aux effets de l'hybridation et de l'introgression que les gènes chloroplastiques, principalement à cause de leur taille effective de population qui est deux fois plus importante (Rieseberg et Soltis 1991). Les événements d'introgression, de même que leurs effets, pourront être plus facilement détectés dans le génome nucléaire, parce que les deux types de polymorphismes dus à l'introgression, le type d'origine et le type introgressant, persistent le plus souvent ensemble dans ce génome (Rieseberg et Soltis 1991). Enfin, dans une perspective globale, l'effet des polymorphismes trans-spécifiques observés précédemment (Bouillé et Bousquet 2005) pourrait également être dilué par l'utilisation d'un grand nombre de gènes nucléaires (Rosenberg et Nordborg 2002; Rokas *et al.* 2003). Ainsi, le génome nucléaire apporterait une contribution certaine à la systématique du genre *Picea*.

4.3.2 - Les études de génétique des populations

En ce qui concerne les taxons qui s'hybrident ou qui sont de proche parenté, deux complexes distincts pourraient être davantage étudiés en Amérique du Nord, étant donné l'origine commune et la proximité génétique de leurs composantes. Les variations mitochondriales régionales des zones hybrides à l'intérieur du complexe *P. engelmannii/glauca/sitchensis* pourraient être comparées à celles, vraisemblablement un peu plus anciennes, à l'intérieur du complexe *P. pungens/mexicana/chihuahuana*. Il serait également pertinent de comparer les flux géniques ancestraux propagés distinctement par le pollen et la graine, afin d'identifier la part de variabilité intraspécifique et la part interspécifique dans les aires d'hybridation naturelle, par exemple en complément d'études telles que celle de Rajora et Dancik (2000). Les études de flux génique entre les populations à l'intérieur de ces complexes et entre ces complexes dans un contexte comparatif pourraient montrer des patrons de dispersion et une histoire biogéographique mieux résolus. Enfin, une étude spécifique impliquant *P. mariana*, *P. rubens* et *P. omorika* aiderait à déterminer si ce dernier est dérivé de *P. mariana* ou si c'est l'inverse. L'hypothèse de départ prônerait que l'espèce dérivée possède un sous-ensemble de la diversité génétique observée dans l'espèce ancestrale. À ce sujet, des marqueurs génétiques au niveau de différents loci nucléaires sont déjà disponibles (voir Perron *et al.* 2000).

En Asie, les variations intraspécifiques détectées lors du séquençage de quelques individus de *P. jezoensis* pourraient être étudiées à travers toute l'aire de répartition de cette espèce, dans le but de déterminer le patron de distribution régional des mutations mitochondriales, ainsi que la présence de sous-espèces ou variétés d'origine génétique.

Par ailleurs, de plus amples études seraient nécessaires afin de comprendre la nature et l'importance de la recombinaison intragénomique détectée à l'intérieur du génome chloroplastique. Pour ce faire, les niveaux d'homoplasie causés par la recombinaison pourraient être estimés par différentes méthodes statistiques, par exemple à partir de la comparaison d'haplotypes tirés de marqueurs hypervariables (voir Marshall *et al.* 2001).

4.3.3 - La cartographie des génomes de *Picea*

Les récents développements dans le domaine de la cartographie génétique des *Picea* aideront à identifier éventuellement les gènes liés aux caractères d'adaptation ou aux caractères à valeur économique. Les nouvelles cartes génétiques ciblant ces gènes candidats suite au criblage de masse d'innombrables marqueurs, tel que dans les projets Arborea (Arborea 2007), Dendrome (Dendrome 2007) et Treenomix (Treenomix 2007), permettront dans un avenir rapproché d'accélérer la détection des régions génomiques les plus divergentes ou fonctionnellement les plus pertinentes (Bousquet *et al.* 2007). Par exemple, le séquençage de régions ainsi identifiées pour l'incompatibilité reproductive pourrait permettre d'identifier le degré de faisabilité des croisements entre les nombreuses espèces de *Picea*. De là, également, les avenues d'amélioration génétique entre les espèces de *Picea* seraient pavées plus rapidement en ce qui a trait à la sélection par croisement de caractères désirables en foresterie.

Actuellement, grâce à certaines études comparant le génome de différentes espèces du genre *Picea*, il est anticipé que le nombre de gènes placés sur les cartes génétiques augmentera rapidement (Pelgas *et al.* 2006). Ces études démontrent notamment un fort degré de conservation de la macrostructure du génome entre des espèces de *Picea* phylogénétiquement distantes, ce qui n'est pas sans rappeler la relative homogénéité morphologique et génétique du genre.

4.4 - Conclusion

Quoique la présente étude repose sur une recherche quasi fondamentale, elle constitue en quelque sorte un nouveau point de départ. Elle visait, d'une part, à jeter quelques nouveaux photons sur la confusion qui a régné jusqu'ici quant aux liens interspécifiques entre les *Picea*, en vue, entre autres, d'éclairer les futurs travaux de systématique sur ce groupe de conifères. D'autre part, elle a permis d'exposer les processus principaux qui sous-tendent l'évolution des *Picea*, tout en fournissant des informations sur les marqueurs moléculaires à préconiser et sur les pièges à éviter lors des études

phylogénétiques et phylogéographiques sur les *Picea*. Nous sommes confiante que les informations qui en sont tirées pourront s'appliquer à d'autres conifères de caractéristiques semblables, tels que ceux à l'intérieur de la famille des Pinaceae.

4.5 - Références

Arborea. 2007. http://www.arborea.ulaval.ca/.

Bouillé M et Bousquet J. 2005. Trans-species shared polymorphisms at orthologous nuclear gene loci among distant species in the conifer *Picea* (Pinaceae): implications for the long-term maintenance of genetic diversity in trees. Amer. J. Bot. 92: 63-73.

Bousquet J, Isabel N, Pelgas B, Cottrell J, Rungis D et Ritland K. 2007. Chapter 3, Spruce. *In* Genome mapping and molecular breeding in plants, Volume 7. Éd. C. Kole. Springer-Verlag, Berlin. pp. 93-114.

Campbell CS, Wright WA, Cox M, Vining TF, Major CS et Arsenault MP. 2005. Nuclear ribosomal DNA internal transcribed spacer 1 (*ITS1*) in *Picea* (Pinaceae): sequence divergence and structure. Mol. Phylogenet. Evol. 35: 165-185.

Dendrome. 2006. http://dendrome.ucdavis.edu/.

Després L, Gielly L, Redoutet B et Taberlet P. 2003. Using AFLP to resolve phylogenetic relationships in a morphologically diversified plant species complex when nuclear and chloroplast sequences fail to reveal variability. Mol. Phylogenet. Evol. 27: 185-196.

ForestTreeDB. 2006. http://ccgb.umn.edu:8088/ForestTreeDB/.

Germano J et Klein AS. 1999. Species-specific nuclear and chloroplast single nucleotide polymorphisms to distinguish *Picea glauca*, *P. mariana* and *P. rubens*. Theor. Appl. Genet. 99: 37-49.

Gordon AG. 1976. The taxonomy and genetics of *Picea rubens* and its relationship to *Picea mariana*. Can. J. Bot. 54: 781-813.

Grant V. 1998. Primary classification and phylogeny of the Polemoniaceae, with comments on molecular cladistics. Amer. J. Bot. 85: 741-752.

Jeandroz S, Roy A et Bousquet J. 1997. Phylogeny and phylogeography of the circumpolar genus *Fraxinus* (Oleaceae) based on internal transcribed spacer sequences of nuclear ribosomal DNA. Mol. Phylogenet. Evol. 7: 241-251.

Kobayashi K, Yoshikawa J et Suzuki M. 2000. DNA identification of *Picea* species of the last glacial age in northern Japan. Jpn. J. Histor. Bot. 8: 67-80.

Kuzoff RK et Gasser CS. 2000. Recent progress in reconstructing angiosperm phylogeny. Trends Plant Sci. 5: 330-336.

Liu T-S. 1982. A new proposal for the classification of the genus *Picea*. Acta Phytotax. Geobot. 33: 227-244.

Marshall HD, Newton C et Ritland K. 2001. Sequence-repeat polymorphisms exhibit the signature of recombination in lodgepole pine chloroplast DNA. Mol. Biol. Evol. 18: 2136-2138.

Mogensen HL. 1996. The hows and whys of cytoplasmic inheritance in seed plants. Amer. J. Bot. 83: 383-404.

Pavy N, Johnson JJ, Crow JA, Paule C, Kunau T, MacKay J, Retzel E. 2007. ForestTreeDB: a database dedicated to the mining of tree transcriptomes, Nucleic Acids Res. NAR Database Issue, Jan 2007.

Pelgas B, Beauseigle S, Acheré V, Jeandroz S, Bousquet J et Isabel N. 2006. Comparative genome mapping among *Picea glauca*, *P. abies* and *P. mariana* x *rubens*, and correspondance with other Pinaceae. Theor. Appl. Genet. 113: 1371-1393.

Perron M et Bousquet J. 1997. Natural hybridization between black spruce and red spruce. Mol. Ecol. 6: 725-734.

Perron M, Perry DJ, Andalo C et Bousquet J. 2000. Evidence from sequence-tagged-site markers of a recent progenitor-derivative species pair in conifers. Proc. Natl. Acad. Sci. USA 97: 11331-11336.

Rajora OP et Dancik BP. 2000. Population genetic variation, structure, and evolution in Engelmann spruce, white spruce, and their natural hybrid complex in Alberta. Can. J. Bot. 78: 768-780.

Ran J-H, Wei X-X et Wang X-Q. 2006. Molecular phylogeny and biogeography of *Picea* (Pinaceae): implications for phylogeographical studies using cytoplasmic haplotypes. Mol. Phylogenet. Evol. 41: 405-419.

Renner SS. 2005. Relaxed molecular clocks for dating historical plant dispersal events. Trends Plant Sci. 10: 550-558.

Rieseberg LH et Soltis DE. 1991. Phylogenetic consequences of cytoplasmic gene flow in plants. Evol. Trends Plants 5: 65-84.

Rieseberg LH et Wendel JF. 1993. Introgression and its consequences in plants. *In* RG Harrison (ed.), Hybrid zones and the evolutionary process. Oxford University Press, Oxford, UK. p. 99.

Rokas A, Williams BL, King N et Carroll SB. 2003. Genome-scale approaches to resolving incongruence in molecular phylogenies. Nature 425: 798-804.

Rosenberg NA et Nordborg M. 2002. Genealogical trees, coalescent theory and the analysis of genetic polymorphisms. Nature Rev. Genet. 3: 380–390.

Sigurgeirsson A et Szmidt AE. 1993. Phylogenetic and biogeographic implications of chloroplast DNA variation in *Picea*. Nord. J. Bot. 13: 233-246.

Syring J, Willyard A, Cronn R et Liston A. 2005. Evolutionary relationships among *Pinus* (Pinaceae) subsections inferred from multiple low-copy nuclear loci. Amer. J. Bot. 92: 2086-2100.

Treenomix. 2006. http://www.treenomix.ca/.

Turna I. 2004. Variation of morphological characters of Oriental spruce (*Picea orientalis*) in Turkey. Biologia, Bratislava 59: 519-526.

Vidacovic M. 1991. Conifers - morphology and variation. USDA, Forest Service. US-Yugoslav Joint Board on Scientific and Technological Cooperation. Ed. Boro Brerkalo. Republic of Croatia. Graficki zavod Hrvatske. pp. 286-372.

Wright JW. 1955. Species crossability in spruce in relation to distribution and taxonomy. For. Sci. 1: 319-349.

Xiang Q-Y, Manchester SR, Thomas DT, Zhang W et Fan C. 2005. Phylogeny, biogeography, and molecular dating of cornelian cherries (*Cornus*, Cornaceae): tracking tertiary plant migration. Evolution 59: 1685-1700.

Xiang Q-Y, Soltis DE, Soltis PS, Manchester SR et Crawford DJ. 2000. Timing the Eastern Asian-Eastern North American floristic disjunction: molecular clock corroborates paleontological estimates. Mol. Phylogenet. Evol. 15: 462-472.

Appendice 2

Alignements de séquences (programme informatique)

```
#NEXUS

Begin data;
        Dimensions ntax=37 nchar=5433;
        Format datatype=dna gap=. missing=? interleave symbols="01";
        Options gapmode=missing;
        Matrix

[Ci-dessous: Matrice rbcL avec indels (0, 1)  en plus. Aucun indel. Donc, ntax=50 et nchar=1351]

          [1                                                  50]
 ABIno1   TCAGGACTCC ATTTACTAGC TTCACGGATC ACTTCATTAC CTTCACGAGC
    ALC   TCAGGACTCC ATTTACTAGC TTCACGGATC ACTTCATTAC CTTCACGAGC
    ASP   TCAGGACTCC ATTTACTAGC TTCACGGATC ACTTCATTAC CTTCACGAGC
    BRA   TCAGGACTCC ATTTACTAGC TTCACGGATC ACTTCATTAC CTTCACGAGC
    BRE   TCAGGACTCC ATTTAGAAGC TTCACGGATC ACTTCATTAC CTTCACGAGC
    CHI   TCAGGACTCC ATTTACTAGC TTCACGGATC ACTTCATTAC CTTCACGAGC
    CRA   TCAGGACTCC ATTTACTAGC TTCACGGATC ACTTCATTAC CTTCACGAGC
 ENGno3   TCAGGACTCC ATTTACTAGC TTCACGGATC ACTTCATTAC CTTCACGAGC
    FAR   TCAGGACTCC ATTTACTAGC TTCACGGATT ACTTCATTAC CTTCACGAGC
 GLAno2   TCAGGACTCC ATTTACTAGC TTCACGGATC ACTTCATTAC CTTCACGAGC
    GLE   TCAGGACTCC ATTTACTAGC TTCACGGATC ACTTCATTAC CTTCACGAGC
 JEZno1   TCAGGACTCC ATTTACTAGC TTCACGGATC ACTTCATTAC CTTCACGAGC
    KOR   TCAGGACTCC ATTTACTAGC TTCACGGATC ACTTCATTAC CTTCACGAGC
    KOY   TCAGGACTCC ATTTACTAGC TTCACGGATC ACTTCATTAC CTTCACGAGC
    LIK   TCAGGACTCC ATTTACTAGC TTCACGGATT ACTTCATTAC CTTCACGAGC
 MARno2   TCAGGACTCC ATTTACTAGC TTCACGGATC ACTTCATTAC CTTCACGAGC
    MAT   TCAGGACTCC ATTTACTAGC TTCACGGATC ACTTCATTAC CTTCACGAGC
    MAX   TCAGGACTCC ATTTACTAGC TTCACGGATC ACTTCATTAC CTTCACGAGC
    MEX   TCAGGACTCC ATTTACTAGC TTCACGGATC ACTTCATTAC CTTCACGAGC
    MEY   TCAGGACTCC ATTTACTAGC TTCACGGATC ACTTCATTAC CTTCACGAGC
    MON   TCAGGACTCC ATTTACTAGC TTCACGGATC ACTTCATTAC CTTCACGAGC
    MOR   TCAGGACTCC ATTTACTAGC TTCACGGATC ACTTCATTAC CTTCACGAGC
    OBO   TCAGGACTCC ATTTACTAGC TTCACGGATC ACTTCATTAC CTTCACGAGC
    OMO   TCAGGACTCC ATTTACTAGC TTCACGGATC ACTTCATTAC CTTCACGAGC
    ORI   TCAGGACTCC ATTTACTAGC TTCACGGATC ACTTCATTAC CTTCACGAGC
PINBANK   TCAGGACTCC ATTTACAAGC TTCACGGATC ACTTCATTAC CTTCACGAGC
 PINKOR   TCAGGACTCC ATTTACAAGC TTCACGGATC ACTTCATTAC CTTCACGAGC
    PUN   TCAGGACTCC ATTTACTAGC TTCACGGATC ACTTCATTAC CTTCACGAGC
    PUR   TCAGGACTCC ATTTACTAGC TTCACGGATC ACTTCATTAC CTTCACGAGC
    RET   TCAGGACTCC ATTTACTAGC TTCACGGATC ACTTCATTAC CTTCACGAGC
    RUB   TCAGGACTCC ATTTACTAGC TTCACGGATC ACTTCATTAC CTTCACGAGC
    SCH   TCAGGACTCC ATTTACTAGC TTCACGGATC ACTTCATTAC CTTCACGAGC
 SITno1   TCAGGACTCC ATTTACTAGC TTCACGGATC ACTTCATTAC CTTCACGAGC
    SMI   TCAGGACTCC ATTTACTAGC TTCACGGATC ACTTCATTAC CTTCACGAGC
    SPI   TCAGGACTCC ATTTACTAGC TTCACGGATT ACTTCATTAC CTTCACGAGC
    TOR   TCAGGACTCC ATTTACTAGC TTCACGGATC ACTTCATTAC CTTCACGAGC
    WIL   TCAGGACTCC ATTTACTAGC TTCACGGATC ACTTCATTAC CTTCACGAGC

          [51                                                 100]
 ABIno1   AAGATCACGT CCTTCATTAC GAGCTTGTAC ACAAGCTTCT AGAGCAACTC
    ALC   AAGATCACGT CCTTCATTAC GAGCTTGTAC ACAAGCTTCT AGAGCAACTC
    ASP   AAGATCACGT CCTTCATTAC GAGCTTGTAC ACAAGCTTCT AGAGCAACTC
    BRA   AAGATCACGT CCTTCATTAC GAGCTTGTAC ACAAGCTTCT AGAGCAACTC
    BRE   AAGATCACGT CCTTCATTAC GAGCTTGTAC ACAAGCTTCT AGAGCAACTC
    CHI   AAGATCACGT CCTTCATTAC GAGCTTGTAC ACAAGCTTCT AGAGCAACTC
    CRA   AAGATCACGT CCTTCATTAC GAGCTTGTAC ACAAGCTTCT AGAGCAACTC
 ENGno3   AAGATCACGT CCTTCATTAC GAGCTTGTAC ACAAGCTTCT AGAGCAACTC
    FAR   AAGATCACGT CCTTCATTAC GAGCTTGTAC ACAAGCTTCT AGAGCAACTC
 GLAno2   AAGATCACGT CCTTCATTAC GAGCTTGTAC ACAAGCTTCT AGAGCAACTC
    GLE   AAGATCACGT CCTTCATTAC GAGCTTGTAC ACAAGCTTCT AGAGCAACTC
 JEZno1   AAGATCACGT CCTTCATTAC GAGCTTGTAC ACAAGCTTCT AGAGCAACTC
    KOR   AAGATCACGT CCTTCATTAC GAGCTTGTAC ACAAGCTTCT AGAGCAACTC
    KOY   AAGATCACGT CCTTCATTAC GAGCTTGTAC ACAAGCTTCT AGAGCAACTC
    LIK   AAGATCACGT CCTTCATTAC GAGCTTGTAC ACAAGCTTCT AGAGCAACTC
 MARno2   AAGATCACGT CCTTCATTAC GAGCTTGTAC ACAAGCTTCT AGAGCAACTC
    MAT   AAGATCACGT CCTTCATTAC GAGCTTGTAC ACAAGCTTCT AGAGCAACTC
```

```
    MAX  AAGATCACGT CCTTCATTAC GAGCTTGTAC ACAAGCTTCT AGAGCAACTC
    MEX  AAGATCACGT CCTTCATTAC GAGCTTGTAC ACAAGCTTCT AGAGCAACTC
    MEY  AAGATCACGT CCTTCATTAC GAGCTTGTAC ACAAGCTTCT AGAGCAACTC
    MON  AAGATCACGT CCTTCATTAC GAGCTTGTAC ACAAGCTTCT AGAGCAACTC
    MOR  AAGATCACGT CCTTCATTAC GAGCTTGTAC ACAAGCTTCT AGAGCAACTC
    OBO  AAGATCACGT CCTTCATTAC GAGCTTGTAC ACAAGCTTCT AGAGCAACTC
    OMO  AAGATCACGT CCTTCATTAC GAGCTTGTAC ACAAGCTTCT AGAGCAACTC
    ORI  AAGATCACGT CCTTCATTAC GAGCTTGTAC ACAAGCTTCT AGAGCAACTC
PINBANK  AAGATCACGT CCTTCATTAC GAGCTTGTAC ACAAGCTTCT AGAGCAACCC
 PINKOR  AAGATCACGT CCTTCATTAC GAGCTTGTAC ACAAGCTTCT AGAGCAACCC
    PUN  AAGATCACGT CCTTCATTAC GAGCTTGTAC ACAAGCTTCT AGAGCAACTC
    PUR  AAGATCACGT CCTTCATTAC GAGCTTGTAC ACAAGCTTCT AGAGCAACTC
    RET  AAGATCACGT CCTTCATTAC GAGCTTGTAC ACAAGCTTCT AGAGCAACTC
    RUB  AAGATCACGT CCTTCATTAC GAGCTTGTAC ACAAGCTTCT AGAGCAACTC
    SCH  AAGATCACGT CCTTCATTAC GAGCTTGTAC ACAAGCTTCT AGAGCAACTC
 SITno1  AAGATCACGT CCTTCATTAC GAGCTTGTAC ACAAGCTTCT AGAGCAACTC
    SMI  AAGATCACGT CCTTCATTAC GAGCTTGTAC ACAAGCTTCT AGAGCAACTC
    SPI  AAGATCACGT CCTTCATTAC GAGCTTGTAC ACAAGCTTCT AGAGCAACTC
    TOR  AAGATCACGT CCTTCATTAC GAGCTTGTAC ACAAGCTTCT AGAGCAACTC
    WIL  AAGATCACGT CCTTCATTAC GAGCTTGTAC ACAAGCTTCT AGAGCAACTC

         [101                                                150]
ABIno1  GATTAGCTAC TGCACCAGGT GCATTTCCCC AAGGGTGTCC CAAAGTTCCC
    ALC  GATTAGCTAC TGCACCAGGT GCATTTCCCC AAGGGTGTCC CAAAGTTCCC
    ASP  GATTAGCTAC TGCACCAGGT GCATTTCCCC AAGGGTGTCC CAAAGTTCCC
    BRA  GATTAGCTAC TGCACCAGGT GCATTTCCCC AAGGGTGTCC CAAAGTTCCC
    BRE  GATTAGCTAC TGCACCAGGT GCATTTCCCC AAGGGTGTCC CAAAGTTCCC
    CHI  GATTAGCTAC TGCACCAGGT GCATTTCCCC AAGGGTGTCC CAAAGTTCCC
    CRA  GATTAGCTAC TGCACCAGGT GCATTTCCCC AAGGGTGTCC CAAAGTTCCC
ENGno3  GATTAGCTAC TGCACCAGGT GCATTTCCCC AAGGGTGTCC CAAAGTTCCC
    FAR  GATTAGCTAC TGCACCAGGT GCATTTCCCC AAGGGTGTCC CAAAGTTCCC
GLAno2  GATTAGCTAC TGCACCAGGT GCATTTCCCC AAGGGTGTCC CAAAGTTCCC
    GLE  GATTAGCTAC TGCACCAGGT GCATTTCCCC AAGGGTGTCC CAAAGTTCCC
JEZno1  GATTAGCTAC TGCACCAGGT GCATTTCCCC AAGGGTGTCC CAAAGTTCCC
    KOR  GATTAGCTAC TGCACCAGGT GCATTTCCCC AAGGGTGTCC CAAAGTTCCC
    KOY  GATTAGCTAC TGCACCAGGT GCATTTCCCC AAGGGTGTCC CAAAGTTCCC
    LIK  GATTAGCTAC TGCACCAGGT GCATTTCCCC AAGGGTGTCC CAAAGTTCCC
MARno2  GATTAGCTAC TGCACCAGGT GCATTTCCCC AAGGGTGTCC CAAAGTTCCC
    MAT  GATTAGCTAC TGCACCAGGT GCATTTCCCC AAGGGTGTCC CAAAGTTCCC
    MAX  GATTAGCTAC TGCACCAGGT GCATTTCCCC AAGGGTGTCC CAAAGTTCCC
    MEX  GATTAGCTAC TGCACCAGGT GCATTTCCCC AAGGGTGTCC CAAAGTTCCC
    MEY  GATTAGCTAC TGCACCAGGT GCATTTCCCC AAGGGTGTCC CAAAGTTCCC
    MON  GATTAGCTAC TGCACCAGGT GCATTTCCCC AAGGGTGTCC CAAAGTTCCC
    MOR  GATTAGCTAC TGCACCAGGT GCATTTCCCC AAGGGTGTCC CAAAGTTCCC
    OBO  GATTAGCTAC TGCACCAGGT GCATTTCCCC AAGGGTGTCC CAAAGTTCCC
    OMO  GATTAGCTAC TGCACCAGGT GCATTTCCCC AAGGGTGTCC CAAAGTTCCC
    ORI  GATTAGCTAC TGCACCAGGT GCATTTCCCC AAGGGTGTCC CAAAGTTCCC
PINBANK  GATTAGCTAC TGCACCAGGT GCATTTCCCC AAGGATGTCC CAAAGTTCCC
 PINKOR  GATTAGCTGC TGCACCAGGT GCATTTCCCC AAGGATGTCC CAAAGTTCCC
    PUN  GATTAGCTAC TGCACCAGGT GCATTTCCCC AAGGGTGTCC CAAAGTTCCC
    PUR  GATTAGCTAC TGCACCAGGT GCATTTCCCC AAGGGTGTCC CAAAGTTCCC
    RET  GATTAGCTAC TGCACCAGGT GCATTTCCCC AAGGGTGTCC CAAAGTTCCC
    RUB  GATTAGCTAC TGCACCAGGT GCATTTCCCC AAGGGTGTCC CAAAGTTCCC
    SCH  GATTAGCTAC TGCACCAGGT GCATTTCCCC AAGGGTGTCC CAAAGTTCCC
 SITno1  GATTAGCTAC TGCACCAGGT GCATTTCCCC AAGGGTGTCC CAAAGTTCCC
    SMI  GATTAGCTAC TGCACCAGGT GCATTTCCCC AAGGGTGTCC CAAAGTTCCC
    SPI  GATTAGCTAC TGCACCAGGT GCATTTCCCC AAGGGTGTCC CAAAGTTCCC
    TOR  GATTAGCTAC TGCACCAGGT GCATTTCCCC AAGGGTGTCC CAAAGTTCCC
    WIL  GATTAGCTAC TGCACCAGGT GCATTTCCCC AAGGGTGTCC CAAAGTTCCC

         [151                                                200]
ABIno1  CCACCAAACT GTAGTACGGA ATCATCCCCA AAGATCTCGG TCAGAGCAGG
    ALC  CCACCAAACT GTAGTACGGA ATCATCCCCA AAGATCTCGG TCAGAGCAGG
    ASP  CCACCAAACT GTAGTACGGA ATCATCCCCA AAGATCTCGG TCAGAGCAGG
    BRA  CCACCAAACT GTAGTACGGA ATCATCCCCA AAGATCTCGG TCAGAGCAGG
    BRE  CCACCAAACT GTAGTACGGA ATCATCCCCA AAGATCTCGG TCAGAGCAGG
    CHI  CCACCAAACT GTAGTACGGA ATCATCCCCA AAGATCTCGG TCAGAGCAGG
    CRA  CCACCAAACT GTAGTACGGA ATCATCCCCA AAGATCTCGG TCAGAGCAGG
ENGno3  CCACCAAACT GTAGTACGGA ATCATCCCCA AAGATCTCGG TCAGAGCAGG
    FAR  CCACCAAACT GTAGTACGGA ATCATCCCCA AAGATCTCGG TCAGAGCAGG
GLAno2  CCACCAAACT GTAGTACGGA ATCATCCCCA AAGATCTCGG TCAGAGCAGG
    GLE  CCACCAAACT GTAGTACGGA ATCATCCCCA AAGATCTCGG TCAGAGCAGG
JEZno1  CCACCAAACT GTAGTACGGA ATCATCCCCA AAGATCTCGG TCAGAGCAGG
    KOR  CCACCAAACT GTAGTACGGA ATCATCCCCA AAGATCTCGG TCAGAGCAGG
    KOY  CCACCAAACT GTAGTACGGA ATCATCCCCA AAGATCTCGG TCAGAGCAGG
    LIK  CCACCAAACT GTAGTACGGA ATCATCCCCA AAGATCTCGG TCAGAGCAGG
MARno2  CCACCAAACT GTAGTACGGA ATCATCCCCA AAGATCTCGG TCAGAGCAGG
    MAT  CCACCAAACT GTAGTACGGA ATCATCCCCA AAGATCTCGG TCAGAGCAGG
    MAX  CCACCAAACT GTAGTACGGA ATCATCCCCA AAGATCTCGG TCAGAGCAGG
    MEX  CCACCAAACT GTAGTACGGA ATCATCCCCA AAGATCTCGG TCAGAGCAGG
    MEY  CCACCAAACT GTAGTACGGA ATCATCCCCA AAGATCTCGG TCAGAGCAGG
    MON  CCACCAAACT GTAGTACGGA ATCATCCCCA AAGATCTCGG TCAGAGCAGG
    MOR  CCACCAAACT GTAGTACGGA ATCATCCCCA AAGATCTCGG TCAGAGCAGG
    OBO  CCACCAAACT GTAGTACGGA ATCATCCCCA AAGATCTCGG TCAGAGCAGG
```

```
   OMO  CCACCAAACT GTAGTACGGA ATCATCCCCA AAGATCTCGG TCAGAGCAGG
   ORI  CCACCAAACT GTAGTACGGA ATCATCCCCA AAGATCTCGG TCAGAGCAGG
PINBANK CCACCAAACT GTAGTACGGA ATCATCCCCA AAGATCTCGG TCAGAGCAGG
PINKOR  CCACCAAACT GTAGTACGGA ATCATCCCCA AAGATCTCGG TCAGAGCAGG
   PUN  CCACCAAACT GTAGTACGGA ATCATCCCCA AAGATCTCGG TCAGAGCAGG
   PUR  CCACCAAACT GTAGTACGGA ATCATCCCCA AAGATCTCGG TCAGAGCAGG
   RET  CCACCAAACT GTAGTACGGA ATCATCCCCA AAGATCTCGG TCAGAGCAGG
   RUB  CCACCAAACT GTAGTACGGA ATCATCTCCA AAGATCTCGG TCAGAGCAGG
   SCH  CCACCAAACT GTAGTACGGA ATCATCCCCA AAGATCTCGG TCAGAGCAGG
SITno1  CCACCAAACT GTAGTACGGA ATCATCCCCA AAGATCTCGG TCAGAGCAGG
   SMI  CCACCAAACT GTAGTACGGA ATCATCCCCA AAGATCTCGG TCAGAGCAGG
   SPI  CCACCAAACT GTAGTACGGA ATCATCCCCA AAGATCTCGG TCAGAGCAGG
   TOR  CCACCAAACT GTAGTACGGA ATCATCCCCA AAGATCTCGG TCAGAGCAGG
   WIL  CCACCAAACT GTAGTACGGA ATCATCCCCA AAGATCTCGG TCAGAGCAGG

        [201                                                250]
ABIno1  CATATGCCAA ACGTGAATAC CTCCTGAAGC TACGGGCAGG ACACCTGGCA
   ALC  CATATGCCAA ACGTGAATAC CTCCTGAAGC TACGGGCAGG ACACCTGGCA
   ASP  CATATGCCAA ACGTGAATAC CTCCTGAAGC TACGGGCAGG ACACCTGGCA
   BRA  CATATGCCAA ACGTGAATAC CTCCTGAAGC TACGGGCAGG ACACCTGGCA
   BRE  CATATGCCAA ACGTGAATAC CTCCTGAAGC TACGGGCAGG ACACCTGGCA
   CHI  CATATGCCAA ACGTGAATAC CTCCTGAAGC TACGGGCAGG ACACCTGGCA
   CRA  CATATGCCAA ACGTGAATAC CTCCTGAAGC TACGGGCAGG ACACCTGGCA
ENGno3  CATATGCCAA ACGTGAATAC CTCCTGAAGC TACGGGCAGG ACACCTGGCA
   FAR  CATATGCCAA ACGTGAATAC CTCCTGAAGC TACGGGCAGG ACACCTGGCA
GLAno2  CATATGCCAA ACGTGAATAC CTCCTGAAGC TACGGGCAGG ACACCTGGCA
   GLE  CATATGCCAA ACGTGAATAC CTCCTGAAGC TACGGGCAGG ACACCTGGCA
JEZno1  CATATGCCAA ACGTGAATAC CTCCTGAAGC TACGGGCAGG ACACCTGGCA
   KOR  CATATGCCAA ACGTGAATAC CTCCTGAAGC TACGGGCAGG ACACCTGGCA
   KOY  CATATGCCAA ACGTGAATAC CTCCTGAAGC TACGGGCAGG ACACCTGGCA
   LIK  CATATGCCAA ACGTGAATAC CTCCTGAAGC TACGGGCAGG ACACCTGGCA
MARno2  CATATGCCAA ACGTGAATAC CTCCTGAAGC TACGGGCAGG ACACCTGGCA
   MAT  CATATGCCAA ACGTGAATAC CTCCTGAAGC TACGGGCAGG ACACCTGGCA
   MAX  CATATGCCAA ACGTGAATAC CTCCTGAAGC TACGGGCAGG ACACCTGGCA
   MEX  CATATGCCAA ACGTGAATAC CTCCTGAAGC TACGGGCAGG ACACCTGGCA
   MEY  CATATGCCAA ACGTGAATAC CTCCTGAAGC TACGGGCAGG ACACCTGGCA
   MON  CATATGCCAA ACGTGAATAC CTCCTGAAGC TACGGGCAGG ACACCTGGCA
   MOR  CATATGCCAA ACGTGAATAC CTCCTGAAGC TACGGGCAGG ACACCTGGCA
   OBO  CATATGCCAA ACGTGAATAC CTCCTGAAGC TACGGGCAGG ACACCTGGCA
   OMO  CATATGCCAA ACGTGAATAC CTCCTGAAGC TACGGGCAGG ACACCTGGCA
   ORI  CATATGCCAA ACGTGAATAC CTCCTGAAGC TACGGGCAGG ACACCTGGCA
PINBANK CATATGCCAA ACGTGAATAC CTCCTGAAGC TACGGGCAGA ACACCTGGCA
PINKOR  CATATGCCAA ACGTGAATAC CTCCTGAAGC TACGGGCAGG ACGCCTGGCA
   PUN  CATATGCCAA ACGTGAATAC CTCCTGAAGC TACGGGCAGG ACACCTGGCA
   PUR  CATATGCCAA ACGTGAATAC CTCCTGAAGC TACGGGCAGG ACACCTGGCA
   RET  CATATGCCAA ACGTGAATAC CTCCTGAAGC TACGGGCAGG ACACCTGGCA
   RUB  CATATGCCAA ACGTGAATAC CTCCTGAAGC TACGGGCAGG ACACCTGGCA
   SCH  CATATGCCAA ACGTGAATAC CTCCTGAAGC TACGGGCAGG ACACCTGGCA
SITno1  CATATGCCAA ACGTGAATAC CTCCTGAAGC TACGGGCAGG ACACCTGGCA
   SMI  CATATGCCAA ACGTGAATAC CTCCTGAAGC TACGGGCAGG ACACCTGGCA
   SPI  CATATGCCAA ACGTGAATAC CTCCTGAAGC TACGGGCAGG ACACCTGGCA
   TOR  CATATGCCAA ACGTGAATAC CTCCTGAAGC TACGGGCAGG ACACCTGGCA
   WIL  CATATGCCAA ACGTGAATAC CTCCTGAAGC TACGGGCAGG ACACCTGGCA

        [251                                                300]
ABIno1  TAGATACCCA ATCTTGAGTG AAGTAAATAC CACGACTTCG ATCTTTTTCG
   ALC  TAGATACCCA ATCTTGAGTG AAGTAAATAC CACGACTTCG ATCTTTTTCG
   ASP  TAGATACCCA ATCTTGAGTG AAGTAAATAC CACGACTTCG ATCTTTTTCG
   BRA  TAGATACCCA ATCTTGAGTG AAGTAAATAC CACGACTTCG ATCTTTTTCG
   BRE  TAGATACCCA ATCTTGAGTG AAGTAAATAC CACGACTTCG ATCTTTTTCG
   CHI  TAGATACCCA ATCTTGAGTG AAGTAAATAC CACGACTTCG ATCTTTTTCG
   CRA  TAGATACCCA ATCTTGAGTG AAGTAAATAC CACGACTTCG ATCTTTTTCG
ENGno3  TAGATACCCA ATCTTGAGTG AAGTAAATAC CACGACTTCG ATCTTTTTCG
   FAR  TAGATACCCA ATCTTGAGTG AAGTAAATAC CACGACTTCG ATCTTTTTCG
GLAno2  TAGATACCCA ATCTTGAGTG AAGTAAATAC CACGACTTCG ATCTTTTTCG
   GLE  TAGATACCCA ATCTTGAGTG AAGTAAATAC CACGACTTCG ATCTTTTTCG
JEZno1  TAGATACCCA ATCTTGAGTG AAGTAAATAC CACGACTTCG ATCTTTTTCG
   KOR  TAGATACCCA ATCTTGAGTG AAGTAAATAC CACGACTTCG ATCTTTTTCG
   KOY  TAGATACCCA ATCTTGAGTG AAGTAAATAC CACGACTTCG ATCTTTTTCG
   LIK  TAGATACCCA ATCTTGAGTG AAGTAAATAC CACGACTTCG ATCTTTTTCG
MARno2  TAGATACCCA ATCTTGAGTG AAGTAAATAC CACGACTTCG ATCTTTTTCG
   MAT  TAGATACCCA ATCTTGAGTG AAGTAAATAC CACGACTTCG ATCTTTTTCG
   MAX  TAGATACCCA ATCTTGAGTG AAGTAAATAC CACGACTTCG ATCTTTTTCG
   MEX  TAGATACCCA ATCTTGAGTG AAGTAAATAC CACGACTTCG ATCTTTTTCG
   MEY  TAGATACCCA ATCTTGAGTG AAGTAAATAC CACGACTTCG ATCTTTTTCG
   MON  TAGATACCCA ATCTTGAGTG AAGTAAATAC CACGACTTCG ATCTTTTTCG
   MOR  TAGATACCCA ATCTTGAGTG AAGTAAATAC CACGACTTCG ATCTTTTTCG
   OBO  TAGATACCCA ATCTTGAGTG AAGTAAATAC CACGACTTCG ATCTTTTTCG
   OMO  TAGATACCCA ATCTTGAGTG AAGTAAATAC CACGACTTCG ATCTTTTTCG
   ORI  TAGATACCCA ATCTTGAGTG AAGTAAATAC CACGACTTCG ATCTTTTTCG
PINBANK TAGATACCCA ATCTTGAGTG AAGTAAATAC CACGACTTCG ATCTTTTTCG
PINKOR  TAGATACCCA ATCTTGAGTG AAGTAAACAC CACGACTTCG ATCTTTTTCG
   PUN  TAGATACCCA ATCTTGAGTG AAGTAAATAC CACGACTTCG ATCTTTTTCG
   PUR  TAGATACCCA ATCTTGAGTG AAGTAAATAC CACGACTTCG ATCTTTTTCG
```

```
   RET  TAGATACCCA ATCTTGAGTG AAGTAAATAC CACGACTTCG ATCTTTTTCG
   RUB  TAGATACCCA ATCTTGAGTG AAGTAAATAC CACGACTTCG ATCTTTTTCG
   SCH  TAGATACCCA ATCTTGAGTG AAGTAAATAC CACGACTTCG ATCTTTTTCG
SITno1  TAGATACCCA ATCTTGAGTG AAGTAAATAC CACGACTTCG ATCTTTTTCG
   SMI  TAGATACCCA ATCTTGAGTG AAGTAAATAC CACGACTTCG ATCTTTTTCG
   SPI  TAGATACCCA ATCTTGAGTG AAGTAAATAC CACGACTTCG ATCTTTTTCG
   TOR  TAGATACCCA ATCTTGAGTG AAGTAAATAC CACGACTTCG ATCTTTTTCG
   WIL  TAGATACCCA ATCTTGAGTG AAGTAAATAC CACGACTTCG ATCTTTTTCG

        [301                                                350]
ABIno1  ATAAAATCAT CACGCAGTAG ATCAACAAAC CCTAAAGTGA TTTCTCGTTC
   ALC  ATAAAATCAT CACGCAGTAG ATCAACAAAC CCTAAAGTGA TGTCTCGTTC
   ASP  ATAAAATCAT CACGCAGTAG ATCAACAAAC CCTAAAGTGA TTTCTCGTTC
   BRA  ATAAAATCAT CACGCAGTAG ATCAACAAAC CCTAAAGTGA TGTCTCGTTC
   BRE  ATAAAATCAT CACGCAGTAG ATCAACAAAC CCTAAAGTGA CGTCTCGTTC
   CHI  ATAAAATCAT CACGCAGTAG ATCAACAAAC CCTAAAGTGA TGTCTCGTTC
   CRA  ATAAAATCAT CACGCAGTAG ATCAACAAAC CCTAAAGTGA TTTCTCGTTC
ENGno3  ATAAAATCAT CACGCAGTAG ATCAACAAAC CCTAAAGTGA CGTCTCGTTC
   FAR  ATAAAATCAT CACGCAGTAG ATCAACAAAC CCTAAAGTGA TGTCTCGTTC
GLAno2  ATAAAATCAT CACGCAGTAG ATCAACAAAC CCTAAAGTGA CGTCTCGTTC
   GLE  ATAAAATCAT CACGCAGTAG ATCAACAAAC CCTAAAGTGA TTTCTCGTTC
JEZno1  ATAAAATCAT CACGCAGTAG ATCAACAAAC CCTAAAGTGA TTTCTCGTTC
   KOR  ATAAAATCAT CACGCAGTAG ATCAACAAAC CCTAAAGTGA TTTCTCGTTC
   KOY  ATAAAATCAT CACGCAGTAG ATCAACAAAC CCTAAAGTGA TTTCTCGTTC
   LIK  ATAAAATCAT CACGCAGTAG ATCAACAAAC CCTAAAGTGA TGTCTCGTTC
MARno2  ATAAAATCAT CACGCAGTAG ATCAACAAAC CCTAAAGTGA TTTCTCGTTC
   MAT  ATAAAATCAT CACGCAGTAG ATCAACAAAC CCTAAAGTGA TGTCTCGTTC
   MAX  ATAAAATCAT CACGCAGTAG ATCAACAAAC CCTAAAGTGA TGTCTCGTTC
   MEX  ATAAAATCAT CACGCAGTAG ATCAACAAAC CCTAAAGTGA CGTCTCGTTC
   MEY  ATAAAATCAT CACGCAGTAG ATCAACAAAC CCTAAAGTGA TTTCTCGTTC
   MON  ATAAAATCAT CACGCAGTAG ATCAACAAAC CCTAAAGTGA TTTCTCGTTC
   MOR  ATAAAATCAT CACGCAGTAG ATCAACAAAC CCTAAAGTGA TGTCTCGTTC
   OBO  ATAAAATCAT CACGCAGTAG ATCAACAAAC CCTAAAGTGA TTTCTCGTTC
   OMO  ATAAAATCAT CACGCAGTAG ATCAACAAAC CCTAAAGTGA TTTCTCGTTC
   ORI  ATAAAATCAT CACGCAGTAG ATCAACAAAC CCTAAAGTGA TGTCTCGTTC
PINBANK  ATAAAATCAT CACGCAGTAG ATCAACAAAC CCTAAAGTTA TGTCTCGTTC
PINKOR  ATAAAATCAT CACGCAGTAG ATCAACAAAC CCTAAAGTGA CGTCTCGTTC
   PUN  ATAAAATCAT CACGCAGTAG ATCAACAAAC CCTAAAGTGA TGTCTCGTTC
   PUR  ATAAAATCAT CACGCAGTAG ATCAACAAAC CCTAAAGTGA TGTCTCGTTC
   RET  ATAAAATCAT CACGCAGTAG ATCAACAAAC CCTAAAGTGA TTTCTCGTTC
   RUB  ATAAAATCAT CACGCAGTAG ATCAACAAAC CCTAAAGTGA TTTCTCGTTC
   SCH  ATAAAATCAT CACGCAGTAG ATCAACAAAC CCTAAAGTGA TGTCTCGTTC
SITno1  ATAAAATCAT CACGCAGTAG ATCAACAAAC CCTAAAGTGA CGTCTCGTTC
   SMI  ATAAAATCAT CACGCAGTAG ATCAACAAAC CCTAAAGTGA TGTCTCGTTC
   SPI  ATAAAATCAT CACGCAGTAG ATCAACAAAC CCTAAAGTGA TGTCTCGTTC
   TOR  ATAAAATCAT CACGCAGTAG ATCAACAAAC CCTAAAGTGA TTTCTCGTTC
   WIL  ATAAAATCAT CACGCAGTAG ATCAACAAAC CCTAAAGTGA TGTCTCGTTC

        [351                                                400]
ABIno1  CCCTTCAAGT TTACCTACTA CAGTACCGCC GTGAATATGA TCTCCACCGG
   ALC  CCCTTCAAGT TTACCTACTA CAGTACCGCC GTGAATATGA TCTCCACCGG
   ASP  CCCTTCAAGT TTACCTACTA CAGTACCGCC GTGAATATGA TCTCCACCGG
   BRA  CCCTTCAAGT TTACCTACTA CAGTACCGCC GTGAATATGA TCTCCACCGG
   BRE  CCCTTCAAGT TTACCTACTA CAGTACCGGA GTGAATATGA TCTCCACCGG
   CHI  CCCTTCAAGT TTACCTACTA CAGTACCGCC GTGAATATGA TCTCCACCGG
   CRA  CCCTTCAAGT TTACCTACTA CAGTACCGCC GTGAATATGA TCTCCACCGG
ENGno3  CCCTTCAAGT TTACCTACTA CAGTACCGCC GTGAATATGA TCTCCACCGG
   FAR  CCCTTCAAGT TTACCTACTA CAGTACCGCC GTGAATATGA TCTCCACCGG
GLAno2  CCCTTCAAGT TTACCTACTA CAGTACCGCC GTGAATATGA TCTCCACCGG
   GLE  CCCTTCAAGT TTACCTACTA CAGTACCGCC GTGAATATGA TCTCCACCGG
JEZno1  CCCTTCAAGT TTACCTACTA CAGTACCGCC GTGAATATGA TCTCCACCGG
   KOR  CCCTTCAAGT TTACCTACTA CAGTACCGCC GTGAATATGA TCTCCACCGG
   KOY  CCCTTCAAGT TTACCTACTA CAGTACCGCC GTGAATATGA TCTCCACCGG
   LIK  CCCTTCAAGT TTACCTACTA CAGTACCGCC GTGAATATGA TCTCCACCGG
MARno2  CCCTTCAAGT TTACCTACTA CAGTACCGCC GTGAATATGA TCTCCACCGG
   MAT  CCCTTCAAGT TTACCTACTA CAGTACCGCC GTGAATATGA TCTCCACCGG
   MAX  CCCTTCAAGT TTACCTACTA CAGTACCGCC GTGAATATGA TCTCCACCGG
   MEX  CCCTTCAAGT TTACCTACTA CAGTACCGCC GTGAATATGA TCTCCACCGG
   MEY  CCCTTCAAGT TTACCTACTA CAGTACCGCC GTGAATATGA TCTCCACCGG
   MON  CCCTTCAAGT TTACCTACTA CAGTACCGCC GTGAATATGA TCTCCACCGG
   MOR  CCCTTCAAGT TTACCTACTA CAGTACCGCC GTGAATATGA TCTCCACCGG
   OBO  CCCTTCAAGT TTACCTACTA CAGTACCGCC GTGAATATGA TCTCCACCGG
   OMO  CCCTTCAAGT TTACCTACTA CAGTACCGCC GTGAATATGA TCTCCACCGG
   ORI  CCCTTCAAGT TTACCTACTA CAGTACCGCC GTGAATATGA TCTCCACCGG
PINBANK  CCCTTCAAGT TTACCTACTA CAGTACCGGC GTGAATATGA TCTCCACCGG
PINKOR  CCCTTCAAGT TTACCTACTA CAGTACCGGC GTGAACATGA TCTCCACCGG
   PUN  CCCTTCAAGT TTACCTACTA CAGTACCGCC GTGAATATGA TCTCCACCGG
   PUR  CCCTTCAAGT TTACCTACTA CAGTACCGCC GTGAATATGA TCTCCACCGG
   RET  CCCTTCAAGT TTACCTACTA CAGTACCGCC GTGAATATGA TCTCCACCGG
   RUB  CCCTTCAAGT TTACCTACTA CAGTACCGCC GTGAATATGA TCTCCACCGG
   SCH  CCCTTCAAGT TTACCTACTA CAGTACCGCC GTGAATATGA TCTCCACCGG
SITno1  CCCTTCAAGT TTACCTACTA CAGTACCGGC GTGAATATGA TCTCCACCGG
   SMI  CCCTTCAAGT TTACCTACTA CAGTACCGCC GTGAATATGA TCTCCACCGG
   SPI  CCCTTCAAGT TTACCTACTA CAGTACCGCC GTGAATATGA TCTCCACCGG
```

```
   TOR  CCCTTCAAGT TTACCTACTA CAGTACCGCC GTGAATATGA TCTCCACCGG
   WIL  CCCTTCAAGT TTACCTACTA CAGTACCGCC GTGAATATGA TCTCCACCGG

        [401                                                450]
ABIno1  ACATACGCAA TGCTTTAGCC AGTACACGGA AGTGCATACC ATGATTTTTT
   ALC  ACATACGCAA TGCTTTAGCC AGTACACGGA AGTGCATACC ATGATTTTTT
   ASP  ACATACGCAA TGCTTTAGCC AGTACACGGA AGTGCATACC ATGATTTTTT
   BRA  ACATACGCAA TGCTTTAGCC AGTACACGGA AGTGCATACC ATGATTTTTT
   BRE  ACATACGCAA TGCTTTAGCC AGTACACGGA AGTGCATACC ATGATTTTTT
   CHI  ACATACGCAA TGCTTTAGCC AGTACACGGA AGTGCATACC ATGATTTTTT
   CRA  ACATACGCAA TGCTTTAGCC AGTACACGGA AGTGCATACC ATGATTTTTT
ENGno3  ACATACGCAA TGCTTTAGCC AGTACACGGA AGTGCATACC ATGATTTTTT
   FAR  ACATACGCAA TGCTTTAGCC AGTACACGGA AGTGCATACC ATGATTTTTT
GLAno2  ACATACGCAA TGCTTTAGCC AGTACACGGA AGTGCATACC ATGATTTTTT
   GLE  ACATACGCAA TGCTTTAGCC AGTACACGGA AGTGCATACC ATGATTTTTT
JEZno1  ACATACGCAA TGCTTTAGCC AGTACACGGA AGTGCATACC ATGATTTTTT
   KOR  ACATACGCAA TGCTTTAGCC AGTACACGGA AGTGCATACC ATGATTTTTT
   KOY  ACATACGCAA TGCTTTAGCC AGTACACGGA AGTGCATACC ATGATTTTTT
   LIK  ACATACGCAA TGCTTTAGCC AGTACACGGA AGTGCATACC ATGATTTTTT
MARno2  ACATACGCAA TGCTTTAGCC AGTACACGGA AGTGCATACC ATGATTTTTT
   MAT  ACATACGCAA TGCTTTAGCC AGTACACGGA AGTGCATACC ATGATTTTTT
   MAX  ACATACGCAA TGCTTTAGCC AGTACACGGA AGTGCATACC ATGATTTTTT
   MEX  ACATACGCAA TGCTTTAGCC AGTACACGGA AGTGCATACC ATGATTTTTT
   MEY  ACATACGCAA TGCTTTAGCC AGTACACGGA AGTGCATACC ATGATTTTTT
   MON  ACATACGCAA TGCTTTAGCC AGTACACGGA AGTGCATACC ATGATTTTTT
   MOR  ACATACGCAA TGCTTTAGCC AGTACACGGA AGTGCATACC ATGATTTTTT
   OBO  ACATACGCAA TGCTTTAGCC AGTACACGGA AGTGCATACC ATGATTTTTT
   OMO  ACATACGCAA TGCTTTAGCC AGTACACGGA AGTGCATACC ATGATTTTTT
   ORI  ACATACGCAA TGCTTTAGCC AGTACACGGA AGTGCATACC ATGATTTTTT
PINBANK  ACATACGCAA TGCTTTAGCC AGTACCCGGA AATGCATACC ATGAATTTTT
PINKOR  ACATACGCAA TGCTTTAGCC AGTACACGGA AATGCATACC ATGATTTCTT
   PUN  ACATACGCAA TGCTTTAGCC AGTACACGGA AATGCATACC ATGATTTTTT
   PUR  ACATACGCAA TGCTTTAGCC AGTACACGGA AGTGCATACC ATGATTTTTT
   RET  ACATACGCAA TGCTTTAGCC AGTACACGGA AGTGCATACC ATGATTTTTT
   RUB  ACATACGCAA TGCTTTAGCC AGTACACGGA AGTGCATACC ATGATTTTTT
   SCH  ACATACGCAA TGCTTTAGCC AGTACACGGA AGTGCATACC ATGATTTTTT
SITno1  ACATACGCAA TGCTTTAGCC AGTACACGGA AGTGCATACC ATGATTTTTT
   SMI  ACATACGCAA TGCTTTAGCC AGTACACGGA AGTGCATACC ATGATTTTTT
   SPI  ACATACGCAA TGCTTTAGCC AGTACACGGA AGTGCATACC ATGATTTTTT
   TOR  ACATACGCAA TGCTTTAGCC AGTACACGGA AGTGCATACC ATGATTTTTT
   WIL  ACATACGCAA TGCTTTAGCC AGTACACGGA AGTGCATACC ATGATTTTTT

        [451                                                500]
ABIno1  TGTCTGTCAA TCACTGCATG CATCGCGCGG TGAATGTGAA GAAGTAGGCC
   ALC  TGTCTGTCAA TCACTGCATG CATCGCGCGG TGAATGTGAA GAAGTAGGCC
   ASP  TGTCTGTCAA TCACTGCATG CATCGCGCGG TGAATGTGAA GAAGTAGGCC
   BRA  TGTCTGTCAA TCACTGCATG CATCGCGCGG TGAATGTGAA GAAGTAGGCC
   BRE  TGTCTGTCAA TCACTGCATG CATCGCGCGG TGAATGTGAA GAAGTAGGCC
   CHI  TGTCTGTCAA TCACTGCATG CATCGCGCGG TGAATGTGAA GAAGTAGGCC
   CRA  TGTCTGTCAA TCACTGCATG CATCGCGCGG TGAATGTGAA GAAGTAGGCC
ENGno3  TGTCTGTCAA TCACTGCATG CATCGCGCGG TGAATGTGAA GAAGTAGGCC
   FAR  TGTCTGTCAA TCACTGCATG CATCGCGCGG TGAATGTGAA GAAGTAGGCC
GLAno2  TGTCTGTCAA TCACTGCATG CATCGCGCGG TGAATGTGAA GAAGTAGGCC
   GLE  TGTCTGTCAA TCACTGCATG CATCGCGCGG TGAATGTGAA GAAGTAGGCC
JEZno1  TGTCTGTCAA TCACTGCATG CATCGCGCGG TGAATGTGAA GAAGTAGGCC
   KOR  TGTCTGTCAA TCACTGCATG CATCGCGCGG TGAATGTGAA GAAGTAGGCC
   KOY  TGTCTGTCAA TCACTGCATG CATCGCGCGG TGAATGTGAA GAAGTAGGCC
   LIK  TGTCTGTCAA TCACTGCATG CATCGCGCGG TGAATGTGAA GAAGTAGGCC
MARno2  TGTCTGTCAA TCACTGCATG CATCGCGCGG TGAATGTGAA GAAGTAGGCC
   MAT  TGTCTGTCAA TCACTGCATG CATCGCGCGG TGAATGTGAA GAAGTAGGCC
   MAX  TGTCTGTCAA TCACTGCATG CATCGCGCGG TGAATGTGAA GAAGTAGGCC
   MEX  TGTCTGTCAA TCACTGCATG CATCGCGCGG TGAATGTGAA GAAGTAGGCC
   MEY  TGTCTGTCAA TCACTGCATG CATCGCGCGG TGAATGTGAA GAAGTAGGCC
   MON  TGTCTGTCAA TCACTGCATG CATCGCGCGG TGAATGTGAA GAAGTAGGCC
   MOR  TGTCTGTCAA TCACTGCATG CATCGCGCGG TGAATGTGAA GAAGTAGGCC
   OBO  TGTCTGTCAA TCACTGCATG CATCGCGCGG TGAATGTGAA GAAGTAGGCC
   OMO  TGTCTGTCAA TCACTGCATG CATCGCGCGG TGAATGTGAA GAAGTAGGCC
   ORI  TGTCTGTCAA TCACTGCATG CATCGCGCGG TGAATGTGAA GAAGTAGGCC
PINBANK  TGTCTGTCAA TAACTGCATG CATCGCGCGG TGAATGTGAA GAAGTAGGCC
PINKOR  TGTCTGTCAA TAACTGCATG CATCGCGCGG TGAATGTGAA GAAGTAGGCC
   PUN  TGTCTGTCAA TCACTGCATG CATCGCGCGG TGAATGTGAA GAAGTAGGCC
   PUR  TGTCTGTCAA TCACTGCATG CATCGCGCGG TGAATGTGAA GAAGTAGGCC
   RET  TGTCTGTCAA TCACTGCATG CATCGCGCGG TGAATGTGAA GAAGTAGGCC
   RUB  TGTCTGTCAA TCACTGCATG CATCGCGCGG TGAATGTGAA GAAGTAGGCC
   SCH  TGTCTGTCAA TCACTGCATG CATCGCGCGG TGAATGTGAA GAAGTAGGCC
SITno1  TGTCTGTCAA TCACTGCATG CATCGCGCGG TGAATGTGAA GAAGTAGGCC
   SMI  TGTCTGTCAA TCACTGCATG CATCGCGCGG TGAATGTGAA GAAGTAGGCC
   SPI  TGTCTGTCAA TCACTGCATG CATCGCGCGG TGAATGTGAA GAAGTAGGCC
   TOR  TGTCTGTCAA TCACTGCATG CATCGCGCGG TGAATGTGAA GAAGTAGGCC
   WIL  TGTCTGTCAA TCACTGCATG CATCGCGCGG TGAATGTGAA GAAGTAGGCC

        [501                                                550]
ABIno1  GTTGTCTCGG CAATAATGAG CCAAAGAAGT ATTTGCGGTA AAACCTCCCG
   ALC  GTTGTCTCGG CAATAATGAG CCAAAGAAGT ATTTGCGGTA AAACCTCCCG
```

```
   ASP  GTTGTCTCGG CAATAATGAG CCAAAGAAGT ATTTGCGGTA AAACCTCCCG
   BRA  GTTGTCTCGG CAATAATGAG CCAAAGAAGT ATTTGCGGTA AAACCTCCCG
   BRE  GTTGTCTCGG CAATAATGAG CCAAAGAAGT ATTTGCGGTA AAACCTCCCG
   CHI  GTTGTCTCGG CAATAATGAG CCAAAGAAGT ATTTGCGGTA AAACCTCCCG
   CRA  GTTGTCTCGG CAATAATGAG CCAAAGAAGT ATTTGCGGTA AAACCTCCCG
ENGno3  GTTGTCTCGG CAATAATGAG CCAAAGAAGT ATTTGCGGTA AAACCTCCCG
   FAR  GTTGTCTCGG CAATAATGAG CCAAAGAAGT ATTTGCGGTA AAACCTCCCG
GLAno2  GTTGTCTCGG CAATAATGAG CCAAAGAAGT ATTTGCGGTA AAACCTCCCG
   GLE  GTTGTCTCGG CAATAATGAG CCAAAGAAGT ATTTGCGGTA AAACCTCCCG
JEZno1  GTTGTCTCGG CAATAATGAG CCAAAGAAGT ATTTGCGGTA AAACCTCCCG
   KOR  GTTGTCTCGG CAATAATGAG CCAAAGAAGT ATTTGCGGTA AAACCTCCCG
   KOY  GTTGTCTCGG CAATAATGAG CCAAAGAAGT ATTTGCGGTA AAACCTCCCG
   LIK  GTTGTCTCGG CAATAATGAG CCAAAGAAGT ATTTGCGGTA AAACCTCCCG
MARno2  GTTGTCTCGG CAATAATGAG CCAAAGAAGT ATTTGCGGTA AAACCTCCCG
   MAT  GTTGTCTCGG CAATAATGAG CCAAAGAAGT ATTTGCGGTA AAACCTCCCG
   MAX  GTTGTCTCGG CAATAATGAG CCAAAGAAGT ATTTGCGGTA AAACCTCCCG
   MEX  GTTGTCTCGG CAATAATGAG CCAAAGAAGT ATTTGCGGTA AAACCTCCCG
   MEY  GTTGTCTCGG CAATAATGAG CCAAAGAAGT ATTTGCGGTA AAACCTCCCG
   MON  GTTGTCTCGG CAATAATGAG CCAAAGAAGT ATTTGCGGTA AAACCTCCCG
   MOR  GTTGTCTCGG CAATAATGAG CCAAAGAAGT ATTTGCGGTA AAACCTCCCG
   OBO  GTTGTCTCGG CAATAATGAG CCAAAGAAGT ATTTGCGGTA AAACCTCCCG
   OMO  GTTGTCTCGG CAATAATGAG CCAAAGAAGT ATTTGCGGTA AAACCTCCCG
   ORI  GTTGTCTCGG CAATAATGAG CCAAAGAAGT ATTTGCGGTA AAACCTCCCG
PINBANK  GTTGTCTCGG CAATAATGAG CCAAAGAAGT ATTTGCGGTA AAACCTCCCG
PINKOR  GTTGTCTCGG CAATAATGAG CCAAAGAAGT ATTTGCGGTA AAACCTCCCG
   PUN  GTTGTCTCGG CAATAATGAG CCAAAGAAGT ATTTGCGGTA AAACCTCCCG
   PUR  GTTGTCTCGG CAATAATGAG CCAAAGAAGT ATTTGCGGTA AAACCTCCCG
   RET  GTTGTCTCGG CAATAATGAG CCAAAGAAGT ATTTGCGGTA AAACCTCCCG
   RUB  GTTGTCTCGG CAATAATGAG CCAAAGAAGT ATTTGCGGTA AAACCTCCCG
   SCH  GTTGTCTCGG CAATAATGAG CCAAAGAAGT ATTTGCGGTA AAACCTCCCG
SITno1  GTTGTCTCGG CAATAATGAG CCAAAGAAGT ATTTGCGGTA AAACCTCCCG
   SMI  GTTGTCTCGG CAATAATGAG CCAAAGAAGT ATTTGCGGTA AAACCTCCCG
   SPI  GTTGTCTCGG CAATAATGAG CCAAAGAAGT ATTTGCGGTA AAACCTCCCG
   TOR  GTTGTCTCGG CAATAATGAG CCAAAGAAGT ATTTGCGGTA AAACCTCCCG
   WIL  GTTGTCTCGG CAATAATGAG CCAAAGAAGT ATTTGCGGTA AAACCTCCCG

        [551                                                600]
ABIno1  TCAGATAGTC ATGCATAACG ATAGGAACTC CCAATTCTCT TGCAAATACT
   ALC  TCAGATAGTC ATGCATAACG ATAGGAACTC CCAATTCTCT TGCAAATACT
   ASP  TCAG?TAGTC ATGCATAACG ATAGGAACTC CCAATTCTCT TGCAAATATT
   BRA  TCAGATAGTC ATGCATAACG ATAGGAACTC CCAATTCTCT TGCAAATACT
   BRE  TCAGATAGTC ATGCATAACG ATAGGAACTC CCAATTCTCT TGCAAATACT
   CHI  TCAGATAGTC ATGCATAACG ATAGGAACTC CCAATTCTCT TGCAAATACT
   CRA  TCAGATAGTC ATGCATAACG ATAGGAACTC CCAATTCTCT TGCAAATACT
ENGno3  TCAGATAGTC ATGCATAACG ATAGGAACTC CCAATTCTCT TGCAAATATT
   FAR  TCAGATAGTC ATGCATAACG ATAGGAACTC CCAATTCTCT TGCAAATACT
GLAno2  TCAGATAGTC ATGCATAACG ATAGGAACTC CCAATTCTCT TGCAAATATT
   GLE  TCAGATAGTC ATGCATAACG ATAGGAACTC CCAATTCTCT TGCAAATACT
JEZno1  TCAGATAGTC ATGCATAACG ATAGGAACTC CCAATTCTCT TGCAAATACT
   KOR  TCAGATAGTC ATGCATAACG ATAGGAACTC CCAATTCTCT TGCAAATACT
   KOY  TCAGATAGTC ATGCATAACG ATAGGAACTC CCAATTCTCT TGCAAATACT
   LIK  TCAGATAGTC ATGCATAACG ATAGGAACTC CCAATTCTCT TGCAAATACT
MARno2  TCAGATAGTC ATGCATAACG ATAGGAACTC CCAATTCTCT TGCAAATATT
   MAT  TCAGATAGTC ATGCATAACG ATAGGAACTC CCAATTCTCT TGCAAATACT
   MAX  TCAGATAGTC ATGCATAACG ATAGGAACTC CCAATTCTCT TGCAAATATT
   MEX  TCAGATAGTC ATGCATAACG ATAGGAACTC CCAATTCTCT TGCAAATATT
   MEY  TCAGATAGTC ATGCATAACG ATAGGAACTC CCAATTCTCT TGCAAATACT
   MON  TCAGATAGTC ATGCATAACG ATAGGAACTC CCAATTCTCT TGCAAATACT
   MOR  TCAGATAGTC ATGCATAACG ATAGGAACTC CCAATTCTCT TGCAAATA?T
   OBO  TCAGATAGTC ATGCATAACG ATAGGAACTC CCAATTCTCT TGCAAATACT
   OMO  TCAGATAGTC ATGCATAACG ATAGGAACTC CCAATTCTCT TGCAAATATT
   ORI  TCAGATAGTC ATGCATAACG ATAGGAACTC CCAATTCTCT TGCAAATATT
PINBANK  TCAGATAGTC ATGCATAACG ATAGGAACTC CCAATTCTCT TGCAAATATT
PINKOR  TCAGGTAGTC ATGCATAACA ATAGGAACTC CCAATTCTCT TGCAAATACT
   PUN  TCAGATAGTC ATGCATAACG ATAGGAACTC CCAATTCTCT TGCAAATACT
   PUR  TCAGATAGTC ATGCATAACG ATAGGAACTC CCAATTCTCT TGCAAATATT
   RET  TCAGATAGTC ATGCATAACG ATAGGAACTC CCAATTCTCT TGCAAATACT
   RUB  TCAGATAGTC ATGCATAACG ATAGGAACTC CCAATTCTCT TGCAAATATT
   SCH  TCAGATAGTC ATGCATAACG ATAGGAACTC CCAATTCTCT TGCAAATACT
SITno1  TCAGATAGTC ATGCATAACG ATAGGAACTC CCAATTCTCT TGCAAATATT
   SMI  TCAGATAGTC ATGCATAACG ATAGGAACTC CCAATTCTCT TGCAAATACT
   SPI  TCAGATAGTC ATGCATAACG ATAGGAACTC CCAATTCTCT TGCAAATACT
   TOR  TCAGATAGTC ATGCATAACG ATAGGAACTC CCAATTCTCT TGCAAATACT
   WIL  TCAGATAGTC ATGCATAACG ATAGGAACTC CCAATTCTCT TGCAAATATT

        [601                                                650]
ABIno1  GCCCTTTTCA TCATTTCTTC ACATGTACCT GCAGTAGCAT TCAAGTAATG
   ALC  GCCCTTTTCA TCATTTCTTC ACATGTACCT GCAGTAGCAT TCAAGTAATG
   ASP  GCCCTTTTCA TCATTTCTTC ACATGTACCT GCAGTAGCAT TCAAGTAATG
   BRA  GCCCTTTTCA TCATTTCTTC ACATGTACCT GCAGTAGCAT TCAAGTAATG
   BRE  GCCCTTTTCA TCATTTCTTC ACATGTACCT GCAGTAGCAT TCAAGTAATG
   CHI  GCCCTTTTCA TCATTTCTTC ACATGTACCT GCAGTAGCAT TCAAGTAATG
   CRA  GCCCTTTTCA TCATTTCTTC ACATGTACCT GCAGTAGCAT TCAAGTAATG
ENGno3  GCCCTTTTCA TCATTTCTTC ACATGTACCT GCAGTAGCAT TCAAGTAATG
```

```
   FAR  GCCCTTTTCA TCATTTCTTC ACATGTACCT GCAGTAGCAT TCAAGTAATG
GLAno2  GCCCTTTTCA TCATTTCTTC ACATGTACCT GCAGTAGCAT TCAAGTAATG
   GLE  GCCCTTTTCA TCATTTCTTC ACATGTACCT GCAGTAGCAT TCAAGTAATG
JEZno1  GCCCTTTTCA TCATTTCTTC ACATGTACCT GCAGTAGCAT TCAAGTAATG
   KOR  GCCCTTTTCA TCATTTCTTC ACATGTACCT GCAGTAGCAT TCAAGTAATG
   KOY  GCCCTTTTCA TCATTTCTTC ACATGTACCT GCAGTAGCAT TCAAGTAATG
   LIK  GCCCTTTTCA TCATTTCTTC ACATGTACCT GCAGTAGCAT TCAAGTAATG
MARno2  GCCCTTTTCA TCATTTCTTC ACATGTACCT GCAGTAGCAT TCAAGTAATG
   MAT  GCCCTTTTCA TCATTTCTTC ACATGTACCT GCAGTAGCAT TCAAGTAATG
   MAX  GCCCTTTTCA TCATTTCTTC ACATGTACCT GCAGTAGCAT TCAAGTAATG
   MEX  GCCCTTTTCA TCATTTCTTC ACATGTACCT GCAGTAGCAT TCAAGTAATG
   MEY  GCCCTTTTCA TCATTTCTTC ACATGTACCT GCAGTAGCAT TCAAGTAATG
   MON  GCCCTTTTCA TCATTTCTTC ACATGTACCT GCAGTAGCAT TCAAGTAATG
   MOR  GCCCTTTTCA TCATTTCTTC ACATGTACCT GCAGTAGCAT TCAAGTAATG
   OBO  GCCCTTTTCA TCATTTCTTC ACATGTACCT GCAGTAGCAT TCAAGTAATG
   OMO  GCCCTTTTCA TCATTTCTTC ACATGTACCT GCAGTAGCAT TCAAGTAATG
   ORI  GCCCTTTTCA TCATTTCTTC ACATGTACCT GCAGTAGCAT TCAAGTAATG
PINBANK  GCCCTTTTCA TCATTTCTTC ACATGTACCT GCAGTAGCAT TCAAATAATG
PINKOR  GCCCTTTTCA TCATTTCTTC ACATGTACCT GCAGTAGCAT TCAAATAATG
   PUN  GCCCTTTTCA TCATTTCTTC ACATGTACCT GCAGTAGCAT TCAAGTAATG
   PUR  GCCCTTTTCA TCATTTCTTC ACATGTACCT GCAGTAGCAT TCAAGTAATG
   RET  GCCCTTTTCA TCATTTCTTC ACATGTACCT GCAGTAGCAT TCAAGTAATG
   RUB  GCCCTTTTCA TCATTTCTTC ACATGTACCT GCAGTAGCAT TCAAGTAATG
   SCH  GCCCTTTTCA TCATTTCTTC ACATGTACCT GCAGTAGCAT TCAAGTAATG
SITno1  GCCCTTTTCA TCATTTCTTC ACATGTACCT GCAGTAGCAT TCAAGTAATG
   SMI  GCCCTTTTCA TCATTTCTTC ACATGTACCT GCAGTAGCAT TCAAGTAATG
   SPI  GCCCTTTTCA TCATTTCTTC ACATGTACCT GCAGTAGCAT TCAAGTAATG
   TOR  GCCCTTTTCA TCATTTCTTC ACATGTACCT GCAGTAGCAT TCAAGTAATG
   WIL  GCCCTTTTCA TCATTTCTTC ACATGTACCT GCAGTAGCAT TCAAGTAATG

        [651                                                700]
ABIno1  TCCCTTAATT TCACCCGTCT CAGCCTGAGC TTTATAAATT GCTTCCGCAC
   ALC  TCCCTTAATT TCACCCGTCT CAGCCTGAGC TTTATAAATT GCTTCCGCAC
   ASP  TCCCTTAATT TCACCCGTCT CAGCCTGAGC TTTATAAATT GCTTCCGCAC
   BRA  TCCCTTAATT TCACCCGTCT CAGCCTGAGC TTTATAAATT GCTTCCGCAC
   BRE  TCCCTTAATT TCACCCGTCT CAGCCTGAGC TTTATTAATT GCTTCCGCAC
   CHI  TCCCTTAATT TCACCCGTCT CAGCCTGAGC TTTATAAATT GCTTCCGCAC
   CRA  TCCCTTAATT TCACCCGTCT CAGCCTGAGC TTTATAAATT GCTTCCGCAC
ENGno3  TCCCTTAATT TCACCCGTCT CAGCCTGAGC TTTATAAATT GCTTCCGCAC
   FAR  TCCCTTAATT TCACCCGTCT CAGCCTGAGC TTTATAAATT GCTTCCGCAC
GLAno2  TCCCTTAATT TCACCCGTCT CAGCCTGAGC TTTATAAATT GCTTCCGCAC
   GLE  TCCCTTAATT TCACCCGTCT CAGCCTGAGC TTTATAAATT GCTTCCGCAC
JEZno1  TCCCTTAATT TCACCCGTCT CAGCCTGAGC TTTATAAATT GCTTCCGCAC
   KOR  TCCCTTAATT TCACCCGTCT CAGCCTGAGC TTTATAAATT GCTTCCGCAC
   KOY  TCCCTTAATT TCACCCGTCT CAGCCTGAGC TTTATAAATT GCTTCCGCAC
   LIK  TCCCTTAATT TCACCCGTCT CAGCCTGAGC TTTATAAATT GCTTCCGCAC
MARno2  TCCCTTAATT TCACCCGTCT CAGCCTGAGC TTTATAAATT GCTTCCGCAC
   MAT  TCCCTTAATT TCACCCGTCT CAGCCTGAGC TTTATAAATT GCTTCCGCAC
   MAX  TCCCTTAATT TCACCCGTCT CAGCCTGAGC TTTATAAATT GCTTCCGCAC
   MEX  TCCCTTAATT TCACCCGTCT CAGCCTGAGC TTTATAAATT GCTTCCGCAC
   MEY  TCCCTTAATT TCACCCGTCT CAGCCTGAGC TTTATAAATT GCTTCCGCAC
   MON  TCCCTTAATT TCACCCGTCT CAGCCTGAGC TTTATAAATT GCTTCCGCAC
   MOR  TCCCTTAATT TCACCCGTCT CAGCCTGAGC TTTATAAATT GCTTCCGCAC
   OBO  TCCCTTAATT TCACCCGTCT CAGCCTGAGC TTTATAAATT GCTTCCGCAC
   OMO  TCCCTTAATT TCACCCGTCT CAGCCTGAGC TTTATAAATT GCTTCCGCAC
   ORI  TCCCTTAATT TCACCCGTCT CAGCCTGAGC TTTATAAATT GCTTCCGCAC
PINBANK  TCCCTTAATT TCACCCGTCT CAGCCTGAGC CTTATAAATT GCTTCTGCAC
PINKOR  TCCCTTAATT TCACCCGTCT CAGCCTGAGC CTTATTAATT GCTTCCGCAC
   PUN  TCCCTTAATT TCACCCGTCT CAGCCTGAGC TTTATAAATT GCTTCCGCAC
   PUR  TCCCTTAATT TCACCCGTCT CAGCCTGAGC TTTATAAATT GCTTCCGCAC
   RET  TCCCTTAATT TCACCCGTCT CAGCCTGAGC TTTATAAATT GCTTCCGCAC
   RUB  TCCCTTAATT TCACCCGTCT CAGCCTGAGC TTTATTAATG GCTTCCGCAC
   SCH  TCCCTTAATT TCACCCGTCT CAGCCTGAGC TTTATAAATT GCTTCCGCAC
SITno1  TCCCTTAATT TCACCCGTCT CAGCCTGAGC TTTATAAATT GCTTCCGCAC
   SMI  TCCCTTAATT TCACCCGTCT CAGCCTGAGC TTTATAAATT GCTTCCGCAC
   SPI  TCCCTTAATT TCACCCGTCT CAGCCTGAGC TTTATAAATT GCTTCCGCAC
   TOR  TCCCTTAATT TCACCCGTCT CAGCCTGAGC TTTATAAATT GCTTCCGCAC
   WIL  TCCCTTAATT TCACCCGTCT CAGCCTGAGC TTTATAAATT GCTTCCGCAC

        [701                                                750]
ABIno1  AAAAGACAAA ACGATCTCTC CAGCGCATGA ATGGTTGGGA ATTTACGTTC
   ALC  AAAAGACAAA ACGATCTCTC CAGCGCATGA ATGGTTGGGA ATTTACGTTC
   ASP  AAAAGACAAA ACGATCTCTC CAGCGCATGA ATGGTTGGGA ATTTACGTTC
   BRA  AAAAGACAAA ACGATCTCTC CAGCGCATGA ATGGTTGGGA ATTTACGTTC
   BRE  AAAAGACAAA ACGATCTCTC CAGCGCATGA ATGGTTGGGA ATTTACGTTC
   CHI  AAAAGACAAA ACGATCTCTC CAGCGCATGA ATGGTTGGGA ATTTACGTTC
   CRA  AAAAGACAAA ACGATCTCTC CAGCGCATGA ATGGTTGGGA ATTTACGTTC
ENGno3  AAAAGACAAA ACGATCTCTC CAGCGCATGA ATGGTTGGGA ATTTACGTTC
   FAR  AAAAGACAAA ACGATCTCTC CAGCGCATGA ATGGTTGGGA ATTTACGTTC
GLAno2  AAAAGACAAA ACGATCTCTC CAGCGCATGA ATGGTTGGGA ATTTACGTTC
   GLE  AAAAGACAAA ACGATCTCTC CAGCGCATGA ATGGTTGGGA ATTTACGTTC
JEZno1  AAAAGACAAA ACGATCTCTC CAGCGCATGA ATGGTTGGGA ATTTACGTTC
   KOR  AAAAGACAAA ACGATCTCTC CAGCGCATGA ATGGTTGGGA ATTTACGTTC
   KOY  AAAAGACAAA ACGATCTCTC CAGCGCATGA ATGGTTGGGA ATTTACGTTC
```

```
   LIK  AAAAGACAAA ACGATCTCTC CAGCGCATGA ATGGTTGGGA ATTTACGTTC
MARno2  AAAAGACAAA ACGATCTCTC CAGCGCATGA ATGGTTGGGA ATTTACGTTC
   MAT  AAAAGACAAA ACGATCTCTC CAGCGCATGA ATGGTTGGGA ATTTACGTTC
   MAX  AAAAGACAAA ACGATCTCTC CAGCGCATGA ATGGTTGGGA ATTTACGTTC
   MEX  AAAAGACAAA ACGATCTCTC CAGCGCATGA ATGGTTGGGA ATTTACGTTC
   MEY  AAAAGACAAA ACGATCTCTC CAGCGCATGA ATGGTTGGGA ATTTACGTTC
   MON  AAAAGACAAA ACGATCTCTC CAGCGCATGA ATGGTTGGGA ATTTACGTTC
   MOR  AAAAGACAAA ACGATCTCTC CAGCGCATGA ATGGTTGGGA ATTTACGTTC
   OBO  AAAAGACAAA ACGATCTCTC CAGCGCATGA ATGGTTGGGA ATTTACGTTC
   OMO  AAAAGACAAA ACGATCTCTC CAGCGCATGA ATGGTTGGGA ATTTACGTTC
   ORI  AAAAGACAAA ACGATCTCTC CAGCGCATGA ATGGTTGGGA ATTTACGTTC
PINBANK AAAAGACAAA ACGATCTCTC CAGCGCATGA ATGGTTGGGA ATTTACGTTC
PINKOR  AAAAGACAAA ACGGTCTCTC CAGCGCATGA ATGGTTGGGA ATTTACGTTC
   PUN  AAAAGACAAA ACGATCTCTC CAGCGCATGA ATGGTTGGGA ATTTACGTTC
   PUR  AAAAGACAAA ACGATCTCTC CAGCGCATGA ATGGTTGGGA ATTTACGTTC
   RET  AAAAGACAAA ACGATCTCTC CAGCGCATGA ATGGTTGGGA ATTTACGTTC
   RUB  AAAAGACAAA ACGATCTCTC CAGCGCATGA ATGGTTGGGA ATTTACGTTC
   SCH  AAAAGACAAA ACGATCTCTC CAGCGCATGA ATGGTTGGGA ATTTACGTTC
SITno1  AAAAGACAAA ACGATCTCTC CAGCGCATGA ATGGTTGGGA ATTTACGTTC
   SMI  AAAAGACAAA ACGATCTCTC CAGCGCATGA ATGGTTGGGA ATTTACGTTC
   SPI  AAAAGACAAA ACGATCTCTC CAGCGCATGA ATGGTTGGGA ATTTACGTTC
   TOR  AAAAGACAAA ACGATCTCTC CAGCGCATGA ATGGTTGGGA ATTTACGTTC
   WIL  AAAAGACAAA ACGATCTCTC CAGCGCATGA ATGGTTGGGA ATTTACGTTC

       [751                                                800]
ABIno1  TCATCATCCT TGGTAAAATC GAGTCCACCA CGGAGACATT CGTAAACTGC
   ALC  TCATCATCCT TGGTAAAATC GAGTCCACCA CGGAGACATT CGTAAACTGC
   ASP  TCATCATCCT TGGTAAAATC GAGTCCACCA CGGAGACATT CGTAAACTGC
   BRA  TCATCATCCT TGGTAAAATC GAGTCCACCA CGGAGACATT CGTAAACTGC
   BRE  TCATCATCCT TGGTAAAATC AAGTCCACCA CGGAGACATT CGTAAACTGC
   CHI  TCATCATCCT TGGTAAAATC GAGTCCACCA CGGAGACATT CGTAAACTGC
   CRA  TCATCATCCT TGGTAAAATC GAGTCCACCA CGGAGACATT CGTAAACTGC
ENGno3  TCATCATCCT TGGTAAAATC GAGTCCACCA CGGAGACATT CGTAAACTGC
   FAR  TCATCATCCT TGGTAAAATC GAGTCCACCA CGGAGACATT CGTAAACTGC
GLAno2  TCATCATCCT TGGTAAAATC GAGTCCACCA CGGAGACATT CGTAAACTGC
   GLE  TCATCATCCT TGGTAAAATC GAGTCCACCA CGGAGACATT CGTAAACTGC
JEZno1  TCATCATCCT TGGTAAAATC GAGTCCACCA CGGAGACATT CGTAAACTGC
   KOR  TCATCATCCT TGGTAAAATC GAGTCCACCA CGGAGACATT CGTAAACTGC
   KOY  TCATCATCCT TGGTAAAATC GAGTCCACCA CGGAGACATT CGTAAACTGC
   LIK  TCATCATCCT TGGTAAAATC GAGTCCACCA CGGAGACATT CGTAAACTGC
MARno2  TCATCATCCT TGGTAAAATC GAGTCCACCA CGGAGACATT CGTAAACTGC
   MAT  TCATCATCCT TGGTAAAATC GAGTCCACCA CGGAGACATT CGTAAACTGC
   MAX  TCATCATCCT TGGTAAAATC GAGTCCACCA CGGAGACATT CGTAAACTGC
   MEX  TCATCATCCT TGGTAAAATC GAGTCCACCA CGGAGACATT CGTAAACTGC
   MEY  TCATCATCCT TGGTAAAATC GAGTCCACCA CGGAGACATT CGTAAACTGC
   MON  TCATCATCCT TGGTAAAATC GAGTCCACCA CGGAGACATT CGTAAACTGC
   MOR  TCATCATCCT TGGTAAAATC GAGTCCACCA CGGAGACATT CGTAAACTGC
   OBO  TCATCATCCT TGGTAAAATC GAGTCCACCA CGGAGACATT CGTAAACTGC
   OMO  TCATCATCCT TGGTAAAATC GAGTCCACCA CGGAGACATT CGTAAACTGC
   ORI  TCATCATCCT TGGTAAAATC GAGTCCACCA CGGAGACATT CGTAAACTGC
PINBANK TCATCATCCT TAGTAAAATC GAGTCCACCA CGGAGACATT CGTAAACTGC
PINKOR  TCATCATCCT TGGTAAAATC GAGTCCACCA CGGAGACATT CGTAAACTGC
   PUN  TCATCATCCT TGGTAAAATC GAGTCCACCA CGGAGACATT CGTAAACTGC
   PUR  TCATCATCCT TGGTAAAATC GAGTCCACCA CGGAGACATT CGTAAACTGC
   RET  TCATCATCCT TGGTAAAATC GAGTCCACCA CGGAGACATT CGTAAACTGC
   RUB  TCATCATCCT TGGTAAAATC GAGTCCACCA CGGAGACATT CGTAAACTGC
   SCH  TCATCATCCT TGGTAAAATC GAGTCCACCA CGGAGACATT CGTAAACTGC
SITno1  TCATCATCCT TGGTAAAATC GAGTCCACCA CGGAGACATT CGTAAACTGC
   SMI  TCATCATCCT TGGTAAAATC AAGTCCACCA CGGAGACATT CGTAAACTGC
   SPI  TCATCATCCT TGGTAAAATC GAGTCCACCA CGGAGACATT CGTAAACTGC
   TOR  TCATCATCCT TGGTAAAATC GAGTCCACCA CGGAGACATT CGTAAACTGC
   WIL  TCATCATCCT TGGTAAAATC GAGTCCACCA CGGAGACATT CGTAAACTGC

       [801                                                850]
ABIno1  TCTACCATAG TTCTTGGCCG ATAGACCCAA TTTTGGTTTG ATAGTACATC
   ALC  TCTACCATAG TTCTTGGCCG ATAGACCCAA TTTTGGTTTG ATAGTACATC
   ASP  TCTACCATAG TTCTTGGCCG ATAGACCCAA TTTTGGTTTG ATAGTACATC
   BRA  TCTACCATAG TTCTTGGCCG ATAGACCCAA TTTTGGTTTG ATAGTACATC
   BRE  TCTACCATAG TTCTTGGCCG ATAGACCCAA TTTTGGTTTG ATAGTACATC
   CHI  TCTACCATAG TTCTTGGCCG ATAGACCCAA TTTTGGTTTG ATAGTACATC
   CRA  TCTACCATAG TTCTTGGCCG ATAGACCCAA TTTTGGTTTG ATAGTACATC
ENGno3  TCTACCATAG TTCTTGGCCG ATAGACCCAA TTTTGGTTTG ATAGTACATC
   FAR  TCTACCATAG TTCTTGGCCG ATAGACCCAA TTTTGGTTTG ATAGTACATC
GLAno2  TCTACCATAG TTCTTGGCCG ATAGACCCAA TTTTGGTTTG ATAGTACATC
   GLE  TCTACCATAG TTCTTGGCCG ATAGACCCAA TTTTGGTTTG ATAGTACATC
JEZno1  TCTACCATAG TTCTTGGCCG ATAGACCCAA TTTTGGTTTG ATAGTACATC
   KOR  TCTACCATAG TTCTTGGCCG ATAGACCCAA TTTTGGTTTG ATAGTACATC
   KOY  TCTACCATAG TTCTTGGCCG ATAGACCCAA TTTTGGTTTG ATAGTACATC
   LIK  TCTACCATAG TTCTTGGCCG ATAGACCCAA TTTTGGTTTG ATAGTACATC
MARno2  TCTACCATAG TTCTTGGCCG ATAGACCCAA TTTTGGTTTG ATAGTACATC
   MAT  TCTACCATAG TTCTTGGCCG ATAGACCCAA TTTTGGTTTG ATAGTACATC
   MAX  TCTACCATAG TTCTTGGCCG ATAGACCCAA TTTTGGTTTG ATAGTACATC
   MEX  TCTACCATAG TTCTTGGCCG ATAGACCCAA TTTTGGTTTG ATAGTACATC
   MEY  TCTACCATAG TTCTTGGCCG ATAGACCCAA TTTTGGTTTG ATAGTACATC
```

```
    MON  TCTACCATAG TTCTTGGCCG ATAGACCCAA TTTTGGTTTG ATAGTACATC
    MOR  TCTACCATAG TTCTTGGCCG ATAGACCCAA TTTTGGTTTG ATAGTACATC
    OBO  TCTACCATAG TTCTTGGCCG ATAGACCCAA TTTTGGTTTG ATAGTACATC
    OMO  TCTACCATAG TTCTTGGCCG ATAGACCCAA TTTTGGTTTG ATAGTACATC
    ORI  TCTACCATAG TTCTTGGCCG ATAGACCCAA TTTTGGTTTG ATAGTACATC
PINBANK  TCTACCATAG TTTTTGGCTG ATAGACCCAA TTTTGGTTTG ATAGTACATC
 PINKOR  TCTACCATAG TTTTTGGCTG ATAGACCCAA TTTTGGTTTG ATAGTACATC
    PUN  TCTACCATAG TTCTTGGCCG ATAGACCCAA TTTTGGTTTG ATAGTACATC
    PUR  TCTACCATAG TTCTTGGCCG ATAGACCCAA TTTTGGTTTG ATAGTACATC
    RET  TCTACCATAG TTCTTGGCCG ATAGACCCAA TTTTGGTTTG ATAGTACATC
    RUB  TCTACCATAG TTCTTGGCCG ATAGACCCAA TTTTGGTTTG ATAGTACATC
    SCH  TCTACCATAG TTCTTGGCCG ATAGACCCAA TTTTGGTTTG ATAGTACATC
 SITno1  TCTACCATAG TTCTTGGCCG ATAGACCCAA TTTTGGTTTG ATAGTACATC
    SMI  TCTACCATAG TTCTTGGCCG ATAGACCCAA TTTTGGTTTT ATAGTACATC
    SPI  TCTACCATAG TTCTTGGCCG ATAGACCCAA TTTTGGTTTG ATAGTACATC
    TOR  TCTACCATAG TTCTTGGCCG ATAGACCCAA TTTTGGTTTG ATAGTACATC
    WIL  TCTACCATAG TTCTTGGCCG ATAGACCCAA TTTTGGTTTG ATAGTACATC

         [851                                                900]
 ABIno1  CCAATAAAGG ACGGCCATAT TTGTTCAATT TATCTCTTTC AACTTGGATA
    ALC  CCAATAAAGG ACGGCCATAT TTGTTCAATT TATCTCTTTC AACTTGGATA
    ASP  CCAATAAAGG ACGGCCATAT TTGTTCAATT TATCTCTTTC AACTTGGATA
    BRA  CCAATAAAGG ACGGCCATAT TTGTTCAATT TATCTCTTTC AACTTGGATA
    BRE  CCAATAAAGG ACGGCCATAT TTGTTCAATT TATCTCTTTC AACTTGGATA
    CHI  CCAATAAAGG ACGGCCATAT TTGTTCAATT TATCTCTTTC AACTTGGATA
    CRA  CCAATAAAGG ACGGCCATAT TTGTTCAATT TATCTCTTTC AACTTGGATA
 ENGno3  CCAATAAAGG ACGGCCATAT TTGTTCAATT TATCTCTTTC AACTTGGATA
    FAR  CCAATAAAGG ACGGCCATAT TTGTTCAATT TATCTCTTTC AACTTGGATA
 GLAno2  CCAATAAAGG ACGGCCATAT TTGTTCAATT TATCTCTTTC AACTTGGATA
    GLE  CCAATAAAGG ACGGCCATAT TTGTTCAATT TATCTCTTTC AACTTGGATA
 JEZno1  CCAATAAAGG ACGGCCATAT TTGTTCAATT TATCTCTTTC AACTTGGATA
    KOR  CCAATAAAGG ACGGCCATAT TTGTTCAATT TATCTCTTTC AACTTGGATA
    KOY  CCAATAAAGG ACGGCCATAT TTGTTCAATT TATCTCTTTC AACTTGGATA
    LIK  CCAATAAAGG ACGGCCATAT TTGTTCAATT TATCTCTTTC AACTTGGATA
 MARno2  CCAATAAAGG ACGGCCATAT TTGTTCAATT TATCTCTTTC AACTTGGATA
    MAT  CCAATAAAGG ACGGCCATAT TTGTTCAATT TATCTCTTTC AACTTGGATA
    MAX  CCAATAAAGG ACGGCCATAT TTGTTCAATT TATCTCTTTC AACTTGGATA
    MEX  CCAATAAAGG ACGGCCATAT TTGTTCAATT TATCTCTTTC AACTTGGATA
    MEY  CCAATAAAGG ACGGCCATAT TTGTTCAATT TATCTCTTTC AACTTGGATA
    MON  CCAATAAAGG ACGGCCATAT TTGTTCAATT TATCTCTTTC AACTTGGATA
    MOR  CCAATAAAGG ACGGCCATAT TTGTTCAATT TATCTCTTTC AACTTGGATA
    OBO  CCAATAAAGG ACGGCCATAT TTGTTCAATT TATCTCTTTC AACTTGGATA
    OMO  CCAATAAAGG ACGGCCATAT TTGTTCAATT TATCTCTTTC AACTTGGATA
    ORI  CCAATAAAGG ACGGCCATAT TTGTTCAATT TATCTCTTTC AACTTGGATA
PINBANK  CCAATAATGG ACGACCATAT TTGTTCAATT TATCTCTTTC AACTTGGATA
 PINKOR  CCAATAAAGG ACGACCGTAT TTGTTCAATT TATCTCTTTC AACTTGGATA
    PUN  CCAATAAAGG ACGGCCATAT TTGTTCAATT TATCTCTTTC AACTTGGATA
    PUR  CCAATAAAGG ACGGCCATAT TTGTTCAATT TATCTCTTTC AACTTGGATA
    RET  CCAATAAAGG ACGGCCATAT TTGTTCAATT TATCTCTTTC AACTTGGATA
    RUB  CCAATAAAGG ACGGCCATAT TTGTTCAATT TATCTCTTTC AACTTGGATA
    SCH  CCAATAAAGG ACGGCCATAT TTGTTCAATT TATCTCTTTC AACTTGGATA
 SITno1  CCAATAAAGG ACGGCCATAT TTGTTCAATT TATCTCTTTC AACTTGGATA
    SMI  CCAATAAAGG ACGGCCATAT TTGTTCAATT TATCTCTTTC AACTTGGATA
    SPI  CCAATAAAGG ACGGCCATAT TTGTTCAATT TATCTCTTTC AACTTGGATA
    TOR  CCAATAAAGG ACGGCCATAT TTGTTCAATT TATCTCTTTC AACTTGGATA
    WIL  CCAATAAAGG ACGGCCATAT TTGTTCAATT TATCTCTTTC AACTTGGATA

         [901                                                950]
 ABIno1  CCATGAGGTG GACCTTGAAA TGTTTTGGAA TAAGCAGGGG GAATCCGCAA
    ALC  CCATGAGGTG GACCTTGAAA TGTTTTGGAA TAAGCAGGGG GAATCCGCAA
    ASP  CCATGAGGTG GACCTTGAAA TGTTTTGGAA TAAGCAGGGG GAATCCGCAA
    BRA  CCATGAGGTG GACCTTGAAA TGTTTTGGAA TAAGCAGGGG GAATCCGCAA
    BRE  CCATGAGGTG GACCTTGAAA TGTTTTGGAA TAAGCAGGGG GAATCCGCAA
    CHI  CCATGAGGTG GACCTTGAAA TGTTTTGGAA TAAGCAGGGG GAATCCGCAA
    CRA  CCATGAGGTG GACCTTGAAA TGTTTTGGAA TAAGCAGGGG GAATCCGCAA
 ENGno3  CCATGAGGTG GACCTTGAAA TGTTTTGGAA TAAGCAGGGG GAATCCGCAA
    FAR  CCATGAGGTG GACCTTGAAA TGTTTTGGAA TAAGCAGGGG GAATCCGCAA
 GLAno2  CCATGAGGTG GACCTTGAAA TGTTTTGGAA TAAGCAGGGG GAATCCGCAA
    GLE  CCATGAGGTG GACCTTGAAA TGTTTTGGAA TAAGCAGGGG GAATCCGCAA
 JEZno1  CCATGAGGTG GACCTTGAAA TGTTTTGGAA TAAGCAGGGG GAATCCGCAA
    KOR  CCATGAGGTG GACCTTGAAA TGTTTTGGAA TAAGCAGGGG GAATCCGCAA
    KOY  CCATGAGGTG GACCTTGAAA TGTTTTGGAA TAAGCAGGGG GAATCCGCAA
    LIK  CCATGAGGTG GACCTTGAAA TGTTTTGGAA TAAGCAGGGG GAATCCGCAA
 MARno2  CCATGAGGTG GACCTTGAAA TGTTTTGGAA TAAGCAGGGG GAATCCGCAA
    MAT  CCATGAGGTG GACCTTGAAA TGTTTTGGAA TAAGCAGGGG GAATCCGCAA
    MAX  CCATGAGGTG GACCTTGAAA TGTTTTGGAA TAAGCAGGGG GAATCCGCAA
    MEX  CCATGAGGTG GACCTTGAAA TGTTTTGGAA TAAGCAGGGG GAATCCGCAA
    MEY  CCATGAGGTG GACCTTGAAA TGTTTTGGAA TAAGCAGGGG GAATCCGCAA
    MON  CCATGAGGTG GACCTTGAAA TGTTTTGGAA TAAGCAGGGG GAATCCGCAA
    MOR  CCATGAGGTG GACCTTGAAA TGTTTTGGAA TAAGCAGGGG GAATCCGCAA
    OBO  CCATGAGGTG GACCTTGAAA TGTTTTGGAA TAAGCAGGGG GAATCCGCAA
    OMO  CCATGAGGTG GACCTTGAAA TGTTTTGGAA TAAGCAGGGG GAATCCGCAA
    ORI  CCATGAGGTG GACCTTGAAA TGTTTTGGAA TAAGCAGGGG GAATCCGCAA
PINBANK  CCATGAGGTG GACCCTGAAA TGTTTTGGAA TAAGCAGGGG GAATCCGCAA
```

```
PINKOR  CCATGAGGTG GACCCTGAAA TGTTTTGGAA TAAGCAGGGG GAATCCGCAA
   PUN  CCATGAGGTG GACCTTGAAA TGTTTTGGAA TAAGCAGGGG GAATCCGCAA
   PUR  CCATGAGGTG GACCTTGAAA TGTTTTGGAA TAAGCAGGGG GAATCCGCAA
   RET  CCATGAGGTG GACCTTGAAA TGTTTTGGAA TAAGCAGGGG GAATCCGCAA
   RUB  CCATGAGGTG GACCTTGAAA TGTTTTGGAA TAAGCAGGGG GAATCCGCAA
   SCH  CCATGAGGTG GACCTTGAAA TGTTTTGGAA TAAGCAGGGG GAATCCGCAA
SITno1  CCATGAGGTG GACCTTGAAA TGTTTTGGAA TAAGCAGGGG GAATCCGCAA
   SMI  CCATGAGGTG GACCTTGAAA TGTTTTGGAA TAAGCAGGGG GAATCCGCAA
   SPI  CCATGAGGTG GACCTTGAAA TGTTTTGGAA TAAGCAGGGG GAATCCGCAA
   TOR  CCATGAGGTG GACCTTGAAA TGTTTTGGAA TAAGCAGGGG GAATCCGCAA
   WIL  CCATGAGGTG GACCTTGAAA TGTTTTGGAA TAAGCAGGGG GAATCCGCAA

        [951                                                1000]
ABIno1  ATCTTCCAAA CGTAGAGCCC GTAGGGCCTT GAATCCAAAT ACATTACCTA
   ALC  ATCTTCCAAA CGTAGAGCCC GTAGGGCCTT GAATCCAAAT ACATTACCTA
   ASP  ATCTTCCAAA CGTAGAGCCC GTAGGGCCTT GAATCCAAAT ACATTACCTA
   BRA  ATCTTCCAAA CGTAGAGCCC GTAGGGCCTT GAATCCAAAT ACATTACCTA
   BRE  ATCTTCCAAA CGTAGAGCCC GTAGGGCCTT GAATCCAAAT ACATTACCTA
   CHI  ATCTTCCAAA CGTAGAGCCC GTAGGGCCTT GAATCCAAAT ACATTACCTA
   CRA  ATCTTCCAAA CGTAGAGCCC GTAGGGCCTT GAATCCAAAT ACATTACCTA
ENGno3  ATCTTCCAAA CGTAGAGCCC GTAGGGCCTT GAATCCAAAT ACATTACCTA
   FAR  ATCTTCCAAA CGTAGAGCCC GTAGGGCCTT GAATCCAAAT ACATTACCTA
GLAno2  ATCTTCCAAA CGTAGAGCCC GTAGGGCCTT GAATCCAAAT ACATTACCTA
   GLE  ATCTTCCAAA CGTAGAGCCC GTAGGGCCTT GAATCCAAAT ACATTACCTA
JEZno1  ATCTTCCAAA CGTAGAGCCC GTAGGGCCTT GAATCCAAAT ACATTACCTA
   KOR  ATCTTCCAAA CGTAGAGCCC GTAGGGCCTT GAATCCAAAT ACATTACCTA
   KOY  ATCTTCCAAA CGTAGAGCCC GTAGGGCCTT GAATCCAAAT ACATTACCTA
   LIK  ATCTTCCAAA CGTAGAGCCC GTAGGGCCTT GAATCCAAAT ACATTACCTA
MARno2  ATCTTCCAAA CGTAGAGCCC GTAGGGCCTT GAATCCAAAT ACATTACCTA
   MAT  ATCTTCCAAA CGTAGAGCCC GTAGGGCCTT GAATCCAAAT ACATTACCTA
   MAX  ATCTTCCAAA CGTAGAGCCC GTAGGGCCTT GAATCCAAAT ACATTACCTA
   MEX  ATCTTCCAAA CGTAGAGCCC GTAGGGCCTT GAATCCAAAT ACATTACCTA
   MEY  ATCTTCCAAA CGTAGAGCCC GTAGGGCCTT GAATCCAAAT ACATTACCTA
   MON  ATCTTCCAAA CGTAGAGCCC GTAGGGCCTT GAATCCAAAT ACATTACCTA
   MOR  ATCTTCCAAA CGTAGAGCCC GTAGGGCCTT GAATCCAAAT ACATTACCTA
   OBO  ATCTTCCAAA CGTAGAGCCC GTAGGGCCTT GAATCCAAAT ACATTACCTA
   OMO  ATCTTCCAAA CGTAGAGCCC GTAGGGCCTT GAATCCAAAT ACATTACCTA
   ORI  ATCTTCCAAA CGTAGAGCCC GTAGGGCCTT GAATCCAAAT ACATTACCTA
PINBANK  ATCTTCCAAA CGTAGAGCCC GTAGGGCCTT GAATCCAAAT ACATTACCTA
PINKOR  ATCTTCCAAA CGTAGAGCCC GTAGGGCCTT GAATCCAAAT ACATTACCTA
   PUN  ATCTTCCAAA CGTAGAGCCC GTAGGGCCTT GAATCCAAAT ACATTACCTA
   PUR  ATCTTCCAAA CGTAGAGCCC GTAGGGCCTT GAATCCAAAT ACATTACCTA
   RET  ATCTTCCAAA CGTAGAGCCC GTAGGGCCTT GAATCCAAAT ACATTACCTA
   RUB  ATCTTCCAAA CGTAGAGCCC GTAGGGCCTT GAATCCAAAT ACATTACCTA
   SCH  ATCTTCCAAA CGTAGAGCCC GTAGGGCCTT GAATCCAAAT ACATTACCTA
SITno1  ATCTTCCAAA CGTAGAGCCC GTAGGGCCTT GAATCCAAAT ACATTACCTA
   SMI  ATCTTCCAAA CGTAGAGCCC GTAGGGCCTT GAATCCAAAT ACATTACCTA
   SPI  ATCTTCCAAA CGTAGAGCCC GTAGGGCCTT GAATCCAAAT ACATTACCTA
   TOR  ATCTTCCAAA CGTAGAGCCC GTAGGGCCTT GAATCCAAAT ACATTACCTA
   WIL  ATCTTCCAAA CGTAGAGCCC GTAGGGCCTT GAATCCAAAT ACATTACCTA

        [1001                                               1050]
ABIno1  CAATGGAAGT GAACAAGTTA GTAACAGAAC CTTCTTCGAA AAGGTCTAAG
   ALC  CAATGGAAGT GAACAAGTTA GTAACAGAAC CTTCTTCGAA AAGGTCTAAG
   ASP  CAATGGAAGT GAACAAGTTA GTAACAGAAC CTTCTTCGAA AAGGTCTAAG
   BRA  CAATGGAAGT GAACAAGTTA GTAACAGAAC CTTCTTCGAA AAGGTCTAAG
   BRE  CAATGGAAGT GAACAAGTTA GTAACAGAAC CTTCTTCGAA AAGGTCTAAG
   CHI  CAATGGAAGT GAACAAGTTA GTAACAGAAC CTTCTTCGAA AAGGTCTAAG
   CRA  CAATGGAAGT GAACAAGTTA GTAACAGAAC CTTCTTCGAA AAGGTCTAAG
ENGno3  CAATGGAAGT GAACAAGTTA GTAACAGAAC CTTCTTCGAA AAGGTCTAAG
   FAR  CAATGGAAGT GAACAAGTTA GTAACAGAAC CTTCTTCGAA AAGGTCTAAG
GLAno2  CAATGGAAGT GAACAAGTTA GTAACAGAAC CTTCTTCGAA AAGGTCTAAG
   GLE  CAATGGAAGT GAACAAGTTA GTAACAGAAC CTTCTTCGAA AAGGTCTAAG
JEZno1  CAATGGAAGT GAACAAGTTA GTAACAGAAC CTTCTTCGAA AAGGTCTAAG
   KOR  CAATGGAAGT GAACAAGTTA GTAACAGAAC CTTCTTCGAA AAGGTCTAAG
   KOY  CAATGGAAGT GAACAAGTTA GTAACAGAAC CTTCTTCGAA AAGGTCTAAG
   LIK  CAATGGAAGT GAACAAGTTA GTAACAGAAC CTTCTTCGAA AAGGTCTAAG
MARno2  CAATGGAAGT GAACAAGTTA GTAACAGAAC CTTCTTCGAA AAGGTCTAAG
   MAT  CAATGGAAGT GAACAAGTTA GTAACAGAAC CTTCTTCGAA AAGGTCTAAG
   MAX  CAATGGAAGT GAACAAGTTA GTAACAGAAC CTTCTTCGAA AAGGTCTAAG
   MEX  CAATGGAAGT GAACAAGTTA GTAACAGAAC CTTCTTCGAA AAGGTCTAAG
   MEY  CAATGGAAGT GAACAAGTTA GTAACAGAAC CTTCTTCGAA AAGGTCTAAG
   MON  CAATGGAAGT GAACAAGTTA GTAACAGAAC CTTCTTCGAA AAGGTCTAAG
   MOR  CAATGGAAGT GAACAAGTTA GTAACAGAAC CTTCTTCGAA AAGGTCTAAG
   OBO  CAATGGAAGT GAACAAGTTA GTAACAGAAC CTTCTTCGAA AAGGTCTAAG
   OMO  CAATGGAAGT GAACAAGTTA GTAACAGAAC CTTCTTCGAA AAGGTCTAAG
   ORI  CAATGGAAGT GAACAAGTTA GTAACAGAAC CTTCTTCGAA AAGGTCTAAG
PINBANK  CAATGGAAGT GAACAAGTTA GTAACAGAAC CTTCTTCGAA AAGGTCTAAG
PINKOR  CAATGGAAGT GAACAGGTTA GTAACAGAAC CTTCTTCGAA AAGGTCTAAG
   PUN  CAATGGAAGT GAACAAGTTA GTAACAGAAC CTTCTTCGAA AAGGTCTAAG
   PUR  CAATGGAAGT GAACAAGTTA GTAACAGAAC CTTCTTCGAA AAGGTCTAAG
   RET  CAATGGAAGT GAACAAGTTA GTAACAGAAC CTTCTTCGAA AAGGTCTAAG
   RUB  CAATGGAAGT GAACAAGTTA GTAACAGAAC CTTCTTCGAA AAGGTCTAAG
   SCH  CAATGGAAGT GAACAAGTTA GTAACAGAAC CTTCTTCGAA AAGGTCTAAG
```

```
SITno1  CAATGGAAGT GAACAAGTTA GTAACAGAAC CTTCTTCGAA AAGGTCTAAG
   SMI  CAATGGAAGT GAACAAGTTA GTAACAGAAC CTTCTTCGAA AAGGTCTAAG
   SPI  CAATGGAAGT GAACAAGTTA GTAACAGAAC CTTCTTCGAA AAGGTCTAAG
   TOR  CAATGGAAGT GAACAAGTTA GTAACAGAAC CTTCTTCGAA AAGGTCTAAG
   WIL  CAATGGAAGT GAACAAGTTA GTAACAGAAC CTTCTTCGAA AAGGTCTAAG

        [1051                                              1100]
ABIno1  GGGTAAGCTA CAAAGGCAAT AAATTGACTT TCCTCTCCAG CAACGGGCTC
   ALC  GGGTAAGCTA CAAAGGCAAT AAATTGAGTT TCCTCTCCAG GAACGGGCTC
   ASP  GGGTAAGCTA CAAAGGCAAT AAATTGACTT TCCTCTCCAG ?AACGGGCTC
   BRA  GGGTAAGCTA CAAAGGCAAT AAATTGACTT TCCTCTCCAG GAACGGGCTC
   BRE  GGGTAAGCTA CATAGGCAAT AAATTGACTT TCCTCTCCAG GAACGGGCTC
   CHI  GGGTAAGCTA CAAAGGCAAT AAATTGAGTT TCCTCTCCAG GAACGGGCTC
   CRA  GGGTAAGCTA CAAAGGCAAT AAATTGACTT TCCTCTCCAG CAACGGGCTC
ENGno3  GGGTAAGCTA CATAGGCAAT AAATTGAGTT TCCTCTCCAG GAACGGGCTC
   FAR  GGGTAAGCTA CAAAGGCAAT AAATTGACTT TCCTCTCCAG GAACGGGCTC
GLAno2  GGGTAAGCTA CATAGGCAAT AAATTGAGTT TCCTCTCCAG GAACGGGCTC
   GLE  GGGTAAGCTA CAAAGGCAAT AAATTGACTT TCCTCTCCAG CAACGGGCTC
JEZno1  GGGTAAGCTA CAAAGGCAAT AAATTGACTT TCCTCTCCAG CAACGGGCTC
   KOR  GGGTAAGCTA CA?AGGCAAT AAATTGACTT TCCTCTCCAG CAACGGGCTC
   KOY  GGGTAAGCTA CAAAGGCAAT AAATTGACTT TCCTCTCCAG CAACGGGCTC
   LIK  GGGTAAGCTA CAAAGGCAAT AAATTGACTT TCCTCTCCAG GAACGGGCTC
MARno2  GGGTAAGCTA CAAAGGCAAT AAATTGACTT TCCTCTCCAG CAACGGGCTC
   MAT  GGGTAAGCTA CAAAGGCAAT AAATTGAGTT TCCTCTCCAG GAACGGGCTC
   MAX  GGGTAAGCTA CAAAGGCAAT AAATTGAGCT TCCTCTCCAG GAACGGGCTC
   MEX  GGGTAAGCTA CATAGGCAAT AAATTGAGTT TCCTCTCCAG GAACGGGCTC
   MEY  GGGTAAGCTA CAAAGGCAAT AAATTGACTT TCCTCTCCAG CAACGGGCTC
   MON  GGGTAAGCTA CAAAGGCAAT AAATTGACTT TCCTCTCCAG CAACGGGCTC
   MOR  GGGTAAGCTA CAAAGGCAAT AAATTGACTT TCCTCTCCAG GAACGGGCTC
   OBO  GGGTAAGCTA CAAAGGCAAT AAATTGACTT TCCTCTCCAG CAACGGGCTC
   OMO  GGGTAAGCTA CAAAGGCAAT AAATTGACTT TCCTCTCCAG CAACGGGCTC
   ORI  GGGTAAGCTA CAAAGGCAAT AAATTGAGCT TCCTCTCCAG GAACGGGCTC
PINBANK  GGGTAAGCTA CATAGGCAAT AAATTGAGCC TCCTCTCCAG GAATGGGTTC
PINKOR  GGGTAAGCTA CATAGGCAAT AAATTGATTC TCCTCTCCAG GAACGGGCTC
   PUN  GGGTAAGCTA CAAAGGCAAT AAATTGAGTT TCCTCTCCAG CAACGGGCTC
   PUR  GGGTAAGCTA CAAAGGCAAT AAATTGACTT TCCTCTCCAG GAACGGGCTC
   RET  GGGTAAGCTA CAAAGGCAAT AAATTGACTT TCCTCTCCAG CAACGGGCTC
   RUB  GGGTAAGCTA CAAAGGCAAT AAATTGACTT TCCTCTCCAG CAACGGGCTC
   SCH  GGGTAAGCTA CAAAGGCAAT AAATTGACTT TCCTCTCCAG GAACGGGCTC
SITno1  GGGTAAGCTA CATAGGCAAT AAATTGACTT TCCTCTCCAG GAACGGGCTC
   SMI  GGGTAAGCTA CA?AGGCAAT AAATTGA?TT TCCTCTCCAG GAACGGGCTC
   SPI  GGGTAAGCTA CAAAGGCAAT AAATTGACTT TCCTCTCCAG GAACGGGCTC
   TOR  GGGTAAGCTA CAAAGGCAAT AAATTGACTT TCCTCTCCAG CAACGGGCTC
   WIL  GGGTAAGCTA CAAAGGCAAT AAATTGACTT TCCTCTCCAG GAACGGGCTC

        [1101                                              1150]
ABIno1  GATGTCATAG CATCGCCCCT TGTAACGATC AAGACTAGTA AGTCCATCGG
   ALC  GATGTCATAG CATCGCCCCT TGTAACGATC AAGACTAGTA AGTCCATCGG
   ASP  GATGTCATAG CATCGCCCCT TGTAACGATC AAGACTAGTA AGTCCATCGG
   BRA  GATGTCATAG CATCGCCCCT TGTAACGATC AAGACTAGTA AGTCCATCGG
   BRE  GATGTCATAG CATCGCCCCT TGTAACGATC AAGACTAGTA AGTCCATCGG
   CHI  GATGTCATAG CATCGCCCCT TGTAACGATC AAGACTAGTA AGTCCATCGG
   CRA  GATGTCATAG CATCGCCCCT TGTAACGATC AAGACTAGTA AGTCCATCGG
ENGno3  GATGTCATAG CATCGCCCCT TGTAACGATC AAGACTAGTA AGTCCATCGG
   FAR  GATGTCATAG CATCGCCCCT TGTAACGATC AAGACTAGTA AGTCCATCGG
GLAno2  GATGTCATAG CATCGCCCCT TGTAACGATC AAGACTAGTA AGTCCATCGG
   GLE  GATGTCATAG CATCGCCCCT TGTAACGATC AAGACTAGTA AGTCCATCGG
JEZno1  GATGTCATAG CATCGCCCCT TGTAACGATC AAGACTAGTA AGTCCATCGG
   KOR  GATGTCATAG CATCGCCCCT TGTAACGATC AAGACTAGTA AGTCCATCGG
   KOY  GATGTCATAG CATCGCCCCT TGTAACGATC AAGACTAGTA AGTCCATCGG
   LIK  GATGTCATAG CATCGCCCCT TGTAACGATC AAGACTAGTA AGTCCATCGG
MARno2  GATGTCATAG CATCGCCCCT TGTAACGATC AAGACTAGTA AGTCCATCGG
   MAT  GATGTCATAG CATCGCCCCT TGTAACGATC AAGACTAGTA AGTCCATCGG
   MAX  GATGTCATAG CATCGCCCCT TGTAACGATC AAGACTAGTA AGTCCATCGG
   MEX  GATGTCATAG CATCGCCCCT TGTAACGATC AAGACTAGTA AGTCCATCGG
   MEY  GATGTCATAG CATCGCCCCT TGTAACGATC AAGACTAGTA AGTCCATCGG
   MON  GATGTCATAG CATCGCCCCT TGTAACGATC AAGACTAGTA AGTCCATCGG
   MOR  GATGTCATAG CATCGCCCCT TGTAACGATC AAGACTAGTA AGTCCATCGG
   OBO  GATGTCATAG CATCGCCCCT TGTAACGATC AAGACTAGTA AGTCCATCGG
   OMO  GATGTCATAG CATCGCCCCT TGTAACGATC AAGACTAGTA AGTCCATCGG
   ORI  GATGTCATAG CATCGCCCCT TGTAACGATC AAGACTAGTA AGTCCATCGG
PINBANK  GATGTCATAG CATCGCCCCT TGTAACGATC GAGACTAGTA AGTCCATCGG
PINKOR  GATGTCATAG CATCGACCCT TGTAACGATC GAGACTAGTA AGTCCATCGG
   PUN  GATGTCATAG CATCGCCCCT TGTAACGATC AAGACTAGTA AGTCCATCGG
   PUR  GATGTCATAG CATCGCCCCT TGTAACGATC AAGACTAGTA AGTCCATCGG
   RET  GATGTCATAG CATCGCCCCT TGTAACGATC AAGACTAGTA AGTCCATCGG
   RUB  GATGTCATAG CATCGCCCCT TGTAACGATC AAGACTAGTA AGTCCATCGG
   SCH  GATGTCATAG CATCGCCCCT TGTAACGATC AAGACTAGTA AGTCCATCGG
SITno1  GATGTCATAG CATCGCCCCT TGTAACGATC AAGACTAGTA AGTCCATCGG
   SMI  GATGTCATAG CATCGCCCCT TGTAACGATC AAGACTAGTA AGTCCATCGG
   SPI  GATGTCATAG CATCGCCCCT TGTAACGATC AAGACTAGTA AGTCCATCGG
   TOR  GATGTCATAG CATCGCCCCT TGTAACGATC AAGACTAGTA AGTCCATCGG
   WIL  GATGTCATAG CATCGCCCCT TGTAACGATC AAGACTAGTA AGTCCATCGG
```

```
          [1151                                                1200]
ABIno1    TCCAAACAGT GGTCCATGTA CCGGTGGAAG ATTCAGCAGC TACTGCTGCT
   ALC    TCCAAACAGT GGTCCATGTA CCGGTGGAAG ATTCAGCAGC TACTGCTGCT
   ASP    TCCAAACAGT GGTCCATGTA CCGGTGGAAG ATTCAGCAGC TACTGCTGCT
   BRA    TCCAAACAGT GGTCCATGTA CCGGTGGAAG ATTCAGCAGC TACTGCTGCT
   BRE    TCCAAACAGT GGTCCATGTA CCGGTGGAAG ATTCAGCAGC TACTGCTGCT
   CHI    TCCAAACAGT GGTCCATGTA CCGGTGGAAG ATTCAGCAGC TACTGCTGCT
   CRA    TCCAAACAGT GGTCCATGTA CCGGTGGAAG ATTCAGCAGC TACTGCTGCT
ENGno3    TCCAAACAGT GGTCCATGTA CCGGTGGAAG ATTCAGCAGC TACTGCTGCT
   FAR    TCCAAACAGT GGTCCATGTA CCGGTGGAAG ATTCAGCAGC TACTGCTGCT
GLAno2    TCCAAACAGT GGTCCATGTA CCGGTGGAAG ATTCAGCAGC TACTGCTGCT
   GLE    TCCAAACAGT GGTCCATGTA CCGGTGGAAG ATTCAGCAGC TACTGCTGCT
JEZno1    TCCAAACAGT GGTCCATGTA CCGGTGGAAG ATTCAGCAGC TACTGCTGCT
   KOR    TCCAAACAGT GGTCCATGTA CCGGTGGAAG ATTCAGCAGC TACTGCTGCT
   KOY    TCCAAACAGT GGTCCATGTA CCGGTGGAAG ATTCAGCAGC TACTGCTGCT
   LIK    TCCAAACAGT GGTCCATGTA CCGGTGGAAG ATTCAGCAGC TACTGCTGCT
MARno2    TCCAAACAGT GGTCCATGTA CCGGTGGAAG ATTCAGCAGC TACTGCTGCT
   MAT    TCCAAACAGT GGTCCATGTA CCGGTGGAAG ATTCAGCAGC TACTGCTGCT
   MAX    TCCAAACAGT GGTCCATGTA CCGGTGGAAG ATTCAGCAGC TACTGCTGCT
   MEX    TCCAAACAGT GGTCCATGTA CCGGTGGAAG ATTCAGCAGC TACTGCTGCT
   MEY    TCCAAACAGT GGTCCATGTA CCGGTGGAAG ATTCAGCAGC TACTGCTGCT
   MON    TCCAAACAGT GGTCCATGTA CCGGTGGAAG ATTCAGCAGC TACTGCTGCT
   MOR    TCCAAACAGT GGTCCATGTA CCGGTGGAAG ATTCAGCAGC TACTGCTGCT
   OBO    TCCAAACAGT GGTCCATGTA CCGGTGGAAG ATTCAGCAGC TACTGCTGCT
   OMO    TCCAAACAGT GGTCCATGTA CCGGTGGAAG ATTCAGCAGC TACTGCTGCT
   ORI    TCCAAACAGT GGTCCATGTA CCGGTGGAAG ATTCAGCAGC TACTGCTGCT
PINBANK   TCCAAACAGT GGTCCATGTA CCGGTGGAAG ATTCAGCAGC TACTGCTGCT
PINKOR    TCCAAACAGT GGTCCATGTA CCGGTGGAAG ATTCAGCAGC TACTGCTGCA
   PUN    TCCAAACAGT GGTCCATGTA CCGGTGGAAG ATTCAGCAGC TACTGCTGCT
   PUR    TCCAAACAGT GGTCCATGTA CCGGTGGAAG ATTCAGCAGC TACTGCTGCT
   RET    TCCAAACAGT GGTCCATGTA CCGGTGGAAG ATTCAGCAGC TACTGCTGCT
   RUB    TCCAAACAGT GGTCCATGTA CCGGTGGAAG ATTCAGCAGC TACTGCTGCT
   SCH    TCCAAACAGT GGTCCATGTA CCGGTGGAAG ATTCAGCAGC TACTGCTGCT
SITno1    TCCAAACAGT GGTCCATGTA CCGGTGGAAG ATTCAGCAGC TACTGCTGCT
   SMI    TCCAAACAGT GGTCCATGTA CCGGTGGAAG ATTCAGCAGC TACTGCTGCT
   SPI    TCCAAACAGT GGTCCATGTA CCGGTGGAAG ATTCAGCAGC TACTGCTGCT
   TOR    TCCAAACAGT GGTCCATGTA CCGGTGGAAG ATTCAGCAGC TACTGCTGCT
   WIL    TCCAAACAGT GGTCCATGTA CCGGTGGAAG ATTCAGCAGC TACTGCTGCT

          [1201                                                1250]
ABIno1    CCCGCTTCCT CGGGCGGCAC CCCTGGTTGA GGAGTTACTC GGAATGCTGC
   ALC    CCCGCTTCCT CGGGCGGCAC CCCTGGTTGA GGAGTTACTC GGAATGCTGC
   ASP    CCCGCTTCCT CGGGCGGCAC CCCTGGTTGA GGAGTTACTC GGAATGCTGC
   BRA    CCCGCTTCCT CGGGCGGCAC CCCTGGTTGA GGAGTTACTC GGAATGCTGC
   BRE    CCCGCTTCCT CGGGCGGCAC CCCTGGTTGA GGAGTTACTC GGAATGCTGC
   CHI    CCCGCTTCCT CGGGCGGCAC CCCTGGTTGA GGAGTTACTC GGAATGCTGC
   CRA    CCCGCTTCCT CGGGCGGCAC CCCTGGTTGA GGAGTTACTC GGAATGCTGC
ENGno3    CCCGCTTCCT CGGGCGGCAC CCCTGGTTGA GGAGTTACTC GGAATGCTGC
   FAR    CCCGCTTCCT CGGGCGGCAC CCCTGGTTGA GGAGTTACTC GGAATGCTGC
GLAno2    CCCGCTTCCT CGGGCGGCAC CCCTGGTTGA GGAGTTACTC GGAATGCTGC
   GLE    CCCGCTTCCT CGGGCGGCAC CCCTGGTTGA GGAGTTACTC GGAATGCTGC
JEZno1    CCCGCTTCCT CGGGCGGCAC CCCAGGTTGA GGAGTTACTC GGAATGCTGC
   KOR    CCCGCTTCCT CGGGCGGCAC CCCTGGTTGA GGAGTTACTC GGAATGCTGC
   KOY    CCCGCTTCCT CGGGCGGCAC CCCTGGTTGA GGAGTTACTC GGAATGCTGC
   LIK    CCCGCTTCCT CGGGCGGCAC CCCTGGTTGA GGAGTTACTC GGAATGCTGC
MARno2    CCCGCTTCCT CGGGCGGCAC CCCTGGTTGA GGAGTTACTC GGAATGCTGC
   MAT    CCCGCTTCCT CGGGCGGCAC CCCTGGTTGA GGAGTTACTC GGAATGCTGC
   MAX    CCCGCTTCCT CGGGCGGCAC CCCTGGTTGA GGAGTTACTC GGAATGCTGC
   MEX    CCCGCTTCCT CGGGCGGCAC CCCTGGTTGA GGAGTTACTC GGAATGCTGC
   MEY    CCCGCTTCCT CGGGCGGCAC CCCTGGTTGA GGAGTTACTC GGAATGCTGC
   MON    CCCGCTTCCT CGGGCGGCAC CCCTGGTTGA GGAGTTACTC GGAATGCTGC
   MOR    CCCGCTTCCT CGGGCGGCAC CCCTGGTTGA GGAGTTACTC GGAATGCTGC
   OBO    CCCGCTTCCT CGGGCGGCAC CCCTGGTTGA GGAGTTACTC GGAATGCTGC
   OMO    CCCGCTTCCT CGGGCGGCAC CCCTGGTTGA GGAGTTACTC GGAATGCTGC
   ORI    CCCGCTTCCT CGGGCGGCAC CCCTGGTTGA GGAGTTACTC GGAATGCTGC
PINBANK   CCCGCTTCTT CAGGTGGCAC CCCGGGTTGA GGAGTTACTC GGAATGCTGC
PINKOR    CCCGCTTCCT CGGCTGGCAC CCCAGGTTGA GGAGTTACTC GGAATGCTGC
   PUN    CCCGCTTCCT CGGGCGGCAC CCCTGGTTGA GGAGTTACTC GGAATGCTGC
   PUR    CCCGCTTCCT CGGGCGGCAC CCCTGGTTGA GGAGTTACTC GGAATGCTGC
   RET    CCCGCTTCCT CGGGCGGCAC CCCTGGTTGA GGAGTTACTC GGAATGCTGC
   RUB    CCCGCTTCCT CGGGCGGCAC CCCTGGTTGA GGAGTTACTC GGAATGCTGC
   SCH    CCCGCTTCCT CGGGCGGCAC CCCTGGTTGA GGAGTTACTC GGAATGCTGC
SITno1    CCCGCTTCCT CGGGCGGCAC CCCTGGTTGA GGAGTTACTC GGAATGCTGC
   SMI    CCCGCTTCCT CGGGCGGCAC CCCTGGTTGA GGAGTTACTC GGAATGCTGC
   SPI    CCCGCTTCCT CGGGCGGCAC CCCTGGTTGA GGAGTTACTC GGAATGCTGC
   TOR    CCCGCTTCCT CGGGCGGCAC CCCTGGTTGA GGAGTTACTC GGAATGCTGC
   WIL    CCCGCTTCCT CGGGCGGCAC CCCTGGTTGA GGAGTTACTC GGAATGCTGC

          [1251                                                1300]
ABIno1    CAAAATATCC GTATCTTTGG TCTGATATTC AGGAGTATAA TAAGTTAATC
   ALC    CAAAATATCC GTATCTTTGG TCTGATATTC AGGAGTATAA TAAGTTAATC
   ASP    CAAAATATCC GTATCTTTGG TCTGATATTC AGGAGTATAA TAAGTTAATC
   BRA    CAAAATATCC GTATCTTTGG TCTGATATTC AGGAGTATAA TAAGTTAATC
   BRE    CAAAATATCC GTATCTTTGG TCTGATATTC AGGAGTATAA TAAGTTAATC
```

```
   CHI  CAAAATATCC GTATCTTTGG TCTGATATTC AGGAGTATAA TAAGTTAATC
   CRA  CAAAATATCC GTATCTTTGG TCTGATATTC AGGAGTATAA TAAGTTAATC
ENGno3  CAAAATATCC GTATCTTTGG TCTGATATTC AGGAGTATAA TAAGTTAATC
   FAR  CAAAATATCC GTATCTTTGG TCTGATATTC AGGAGTATAA TAAGTTAATC
GLAno2  CAAAATATCC GTATCTTTGG TCTGATATTC AGGAGTATAA TAAGTTAATC
   GLE  CAAAATATCC GTATCTTTGG TCTGATATTC AGGAGTATAA TAAGTTAATC
JEZno1  CAAAATATCC GTATCTTTGG TCTGATATTC AGGAGTATAA TAAGTTAATC
   KOR  CAAAATATCC GTATCTTTGG TCTGATATTC AGGAGTATAA TAAGTTAATC
   KOY  CAAAATATCC GTATCTTTGG TCTGATATTC AGGAGTATAA TAAGTTAATC
   LIK  CAAAATATCC GTATCTTTGG TCTGATATTC AGGAGTATAA TAAGTTAATC
MARno2  CAAAATATCC GTATCTTTGG TCTGATATTC AGGAGTATAA TAAGTTAATC
   MAT  CAAAATATCC GTATCTTTGG TCTGATATTC AGGAGTATAA TAAGTTAATC
   MAX  CAAAATATCC GTATCTTTGG TCTGATATTC AGGAGTATAA TAAGTTAATC
   MEX  CAAAATATCC GTATCTTTGG TCTGATATTC AGGAGTATAA TAAGTTAATC
   MEY  CAAAATATCC GTATCTTTGG TCTGATATTC AGGAGTATAA TAAGTTAATC
   MON  CAAAATATCC GTATCTTTGG TCTGATATTC AGGAGTATAA TAAGTTAATC
   MOR  CAAAATATCC GTATCTTTGG TCTGATATTC AGGAGTATAA TAAGTTAATC
   OBO  CAAAATATCC GTATCTTTGG TCTGATATTC AGGAGTATAA TAAGTTAATC
   OMO  CAAAATATCC GTATCTTTGG TCTGATATTC AGGAGTATAA TAAGTTAATC
   ORI  CAAAATATCC GTATCTTTGG TCTGATATTC AGGAGTATAA TAAGTTAATC
PINBANK  CAAGATATCT GTATCTTTGG TCTGATATTC AGGAGTATAA TAAGTTAATC
PINKOR  CAAGATATCC GTATCTTTGG TCTGATATTC AGGAGTATAA TAAGTTAATC
   PUN  CAAAATATCC GTATCTTTGG TCTGATATTC AGGAGTATAA TAAGTTAATC
   PUR  CAAAATATCC GTATCTTTGG TCTGATATTC AGGAGTATAA TAAGTTAATC
   RET  CAAAATATCC GTATCTTTGG TCTGATATTC AGGAGTATAA TAAGTTAATC
   RUB  CAAAATATCC GTATCTTTGG TCTGATATTC AGGAGTATAA TAAGTTAATC
   SCH  CAAAATATCC GTATCTTTGG TCTGATATTC AGGAGTATAA TAAGTTAATC
SITno1  CAAAATATCC GTATCTTTGG TCTGATATTC AGGAGTATAA TAAGTTAATC
   SMI  CAAAATATCC GTATCTTTGG TCTGATATTC AGGAGTATAA TAAGTTAATC
   SPI  CAAAATATCC GTATCTTTGG TCTGATATTC AGGAGTATAA TAAGTTAATC
   TOR  CAAAATATCC GTATCTTTGG TCTGATATTC AGGAGTATAA TAAGTTAATC
   WIL  CAAAATATCC GTATCTTTGG TCTGATATTC AGGAGTATAA TAAGTTAATC

        [1301                                              1350]
ABIno1  TGTAATCTTT AACACCAGCT TTAAATCCGA CACTAGCTTT AGTCTCTGTT
   ALC  TGTAATCTTT AACACCAGCT TTAAATCCGA CACTAGCTTT AGTCTCTGTT
   ASP  TGTAATCTTT AACACCAGCT TTAAATCCGA CACTAGCTTT AGTCTCTGTT
   BRA  TGTAATCTTT AACACCAGCT TTAAATCCGA CACTAGCTTT AGTCTCTGTT
   BRE  TGTAATCTTT AACACCAGCT TTAAATCCGA CACTAGCTTT AGTCTCTGTT
   CHI  TGTAATCTTT AACACCAGCT TTAAATCCGA CACTAGCTTT AGTCTCTGTT
   CRA  TGTAATCTTT AACACCAGCT TTAAATCCGA CACTAGCTTT AGTCTCTGTT
ENGno3  TGTAATCTTT AACACCAGCT TTAAATCCGA CACTAGCTTT AGTCTCTGTT
   FAR  TGTAATCTTT AACACCAGCT TTAAATCCGA CACTAGCTTT AGTCTCTGTT
GLAno2  TGTAATCTTT AACACCAGCT TTAAATCCGA CACTAGCTTT AGTCTCTGTT
   GLE  TGTAATCTTT AACACCAGCT TTAAATCCGA CACTAGCTTT AGTCTCTGTT
JEZno1  TGTAATCTTT AACACCAGCT TTAAATCCGA CACTAGCTTT AGTCTCTGTT
   KOR  TGTAATCTTT AACACCAGCT TTAAATCCGA CACTAGCTTT AGTCTCTGTT
   KOY  TGTAATCTTT AACACCAGCT TTAAATCCGA CACTAGCTTT AGTCTCTGTT
   LIK  TGTAATCTTT AACACCAGCT TTAAATCCGA CACTAGCTTT AGTCTCTGTT
MARno2  TGTAATCTTT AACACCAGCT TTAAATCCGA CACTAGCTTT AGTCTCTGTT
   MAT  TGTAATCTTT AACACCAGCT TTAAATCCGA CACTAGCTTT AGTCTCTGTT
   MAX  TGTAATCTTT AACACCAGCT TTAAATCCGA CACTAGCTTT AGTCTCTGTT
   MEX  TGTAATCTTT AACACCAGCT TTAAATCCGA CACTAGCTTT AGTCTCTGTT
   MEY  TGTAATCTTT AACACCAGCT TTAAATCCGA CACTAGCTTT AGTCTCTGTT
   MON  TGTAATCTTT AACACCAGCT TTAAATCCGA CACTAGCTTT AGTCTCTGTT
   MOR  TGTAATCTTT AACACCAGCT TTAAATCCGA CACTAGCTTT AGTCTCTGTT
   OBO  TGTAATCTTT AACACCAGCT TTAAATCCGA CACTAGCTTT AGTCTCTGTT
   OMO  TGTAATCTTT AACACCAGCT TTAAATCCGA CACTAGCTTT AGTCTCTGTT
   ORI  TGTAATCTTT AACACCAGCT TTAAATCCGA CACTAGCTTT AGTCTCTGTT
PINBANK  TGTAATCTTT AACACCAGCT TTGAATCCGA CACTAGCTTT AGTCTCTGTT
PINKOR  TGTAATCTTT AACACCAGCT TTGAACCCGA CACTAGCTTT AGTCTCTGTT
   PUN  TGTAATCTTT AACACCAGCT TTAAATCCGA CACTAGCTTT AGTCTCTGTT
   PUR  TGTAATCTTT AACACCAGCT TTAAATCCGA CACTAGCTTT AGTCTCTGTT
   RET  TGTAATCTTT AACACCAGCT TTAAATCCGA CACTAGCTTT AGTCTCTGTT
   RUB  TGTAATCTTT AACACCAGCT TTAAATCCGA CACTAGCTTT AGTCTCTGTT
   SCH  TGTAATCTTT AACACCAGCT TTAAATCCGA CACTAGCTTT AGTCTCTGTT
SITno1  TGTAATCTTT AACACCAGCT TTAAATCCGA CACTAGCTTT AGTCTCTGTT
   SMI  TGTAATCTTT AACACCAGCT TTAAATCCGA CACTAGCTTT AGTCTCTGTT
   SPI  TGTAATCTTT AACACCAGCT TTAAATCCGA CACTAGCTTT AGTCTCTGTT
   TOR  TGTAATCTTT AACACCAGCT TTAAATCCGA CACTAGCTTT AGTCTCTGTT
   WIL  TGTAATCTTT AACACCAGCT TTAAATCCGA CACTAGCTTT AGTCTCTGTT

        [1351 ]
ABIno1  T
   ALC  T
   ASP  T
   BRA  T
   BRE  T
   CHI  T
   CRA  T
ENGno3  T
   FAR  T
GLAno2  T
   GLE  T
```

```
 JEZno1  T
    KOR  T
    KOY  T
    LIK  T
 MARno2  T
    MAT  T
    MAX  T
    MEX  T
    MEY  T
    MON  T
    MOR  T
    OBO  T
    OMO  T
    ORI  T
PINBANK  T
 PINKOR  T
    PUN  T
    PUR  T
    RET  T
    RUB  T
    SCH  T
 SITno1  T
    SMI  T
    SPI  T
    TOR  T
    WIL  T
```

[Ci-dessous: Matrice trnK avec indels (0, 1) en plus. 8 indels. Donc, ntax=50 et nchar=2529+8-1=2536]
[Début à 1352 et fin à 3887]

```
         [1           +1                                          50]
 ABIno1  CTGCTCCTTC AGTA0TCAGTA AGAGCGAGAA ATCTCTCCCC AAACCGTGCT
    ALC  CTGCTCCTTC AGTA0TCAGTA AGAGCGAGAA ATCTCTCCCC AAACCGTGCT
    ASP  CTGCTCCTTC AGTA0TCAGTA AGAGCGAGAA ATCTCTCCCC AAACCGTGCT
    BRA  CTGCTCCTTC AGTA0TCAGTA AGAGCGAGAA ATCTCTCCCC AAACCGTGCT
    BRE  CTGCTCCT.. ....1TCAGTA AGAGCGAGAA ATCTCTCCCC AAACCGTGCT
    CHI  CTGCTCCTTC AGTA0TCAGTA AGAGCGAGAA ATCTCTCCCC AAACCGTGCT
    CRA  CTGCTCCTTC AGTA0TCAGTA AGAGCGAGAA ATCTCTCCCC AAACCGTGCT
 ENGno3  CTGCTCCTTC AGTA0TCAGTA AGAGCGAGAA ATCTCTCCCC AAACCGTGCT
    FAR  CTGCTCCTTC AGTA0TCAGTA AGAGCGAGAA ATCTCTCCCC AAACCGTGCT
 GLAno2  CTGCTCCTTC AGTA0TCAGTA AGAGCGAGAA ATCTCTCCCC AAACCGTGCT
    GLE  CTGCTCCTTC AGTA0TCAGTA AGAGCGAGAA ATCTCTCCCC AAACCGTGCT
 JEZno1  CTGCTCCTTC AGTA0TCAGTA AGAGCGAGAA ATCTCTCCCC AAACCGTGCT
    KOR  CTGCTCCTTC AGTA0TCAGTA AGAGCGAGAA ATCTCTCCCC AAACCGTGCT
    KOY  CTGCTCCTTC AGTA0TCAGTA AGAGCGAGAA ATCTCTCCCC AAACCGTGCT
    LIK  CTGCTCCTTC AGTA0TCAGTA AGAGCGAGAA ATCTCTCCCC AAACCGTGCT
 MARno2  CTGCTCCTTC AGTA0TCAGTA AGAGCGAGAA ATCTCTCCCC AAACCGTGCT
    MAT  CTGCTCCTTC AGTA0TCAGTA AGAGCGAGAA ATCTCTCCCC AAACCGTGCT
    MAX  CTGCTCCTTC AGTA0TCAGTA AGAGCGAGAA ATCTCTCCCC AAACCGTGCT
    MEX  CTGCTCCTTC AGTA0TCAGTA AGAGCGAGAA ATCTCTCCCC AAACCGTGCT
    MEY  CTGCTCCTTC AGTA0TCAGTA AGAGCGAGAA ATCTCTCCCC AAACCGTGCT
    MON  CTGCTCCTTC AGTA0TCAGTA AGAGCGAGAA ATCTCTCCCC AAACCGTGCT
    MOR  CTGCTCCTTC AGTA0TCAGTA AGAGCGAGAA ATCTCTCCCC AAACCGTGCT
    OBO  CTGCTCCTTC AGTA0TCAGTA AGAGCGAGAA ATCTCTCCCC AAACCGTGCT
    OMO  CTGCTCCTTC AGTA0TCAGTA AGAGCGAGAA ATCTCTCCCC AAACCGTGCT
    ORI  CTGCTCCTTC AGTA0TCAGTA AGAGCGAGAA ATCTCTCCCC AAACCGTGCT
PINBANK  CTGCTCCTTC AGTA0TCAGTA AGAGCGAAAA ATCTCTCCCC AAACCGTGCT
 PINKOR  CTGCTCCTTC AGTA0TCAGTA AGAGCGAGAA ATCTCTCCCC AAACCGTGCT
    PUN  CTGCTCCTTC AGTA0TCAGTA AGAGCGAGAA ATCTCTCCCC AAACCGTGCT
    PUR  CTGCTCCTTC AGTA0TCAGTA AGAGCGAGAA ATCTCTCCCC AAACCGTGCT
    RET  CTGCTCCTTC AGTA0TCAGTA AGAGCGAGAA ATCTCTCCCC AAACCGTGCT
    RUB  CTGCTCCTTC AGTA0TCAGTA AGAGCGAGAA ATCTCTCCCC AAACCGTGCT
    SCH  CTGCTCCTTC AGTA0TCAGTA AGAGCGAGAA ATCTCTCCCC AAACCGTGCT
 SITno1  CTGCTCCTTC AGTA0TCAGTA AGAGCGAGAA ATCTCTCCCC AAACCGTGCT
    SMI  CTGCTCCTTC AGTA0TCAGTA AGAGCGAGAA ATCTCTCCCC AAACCGTGCT
    SPI  CTGCTCCTTC AGTA0TCAGTA AGAGCGAGAA ATCTCTCCCC AAACCGTGCT
    TOR  CTGCTCCTTC AGTA0TCAGTA AGAGCGAGAA ATCTCTCCCC AAACCGTGCT
    WIL  CTGCTCCTTC AGTA0TCAGTA AGAGCGAGAA ATCTCTCCCC AAACCGTGCT

         [51                                        +1          100]
 ABIno1  TGCAATTCTC ATCGCACACG GCTTTCCCCA TGTAATGTAT CTATTTCCT.0
    ALC  TGCAATTCTC ATCGCACACG GCTTTCCCCA TGTAATGTAT CTATTTCCT.0
    ASP  TGCAATTCTC ATCGCACACG GCTTTCCCCA TGTAATGTAT CTATTTCCT.0
    BRA  TGCAATTCTC ATCGCACACG GCTTTCCCCA TGTAATGTAT CTATTTCCT.0
    BRE  TGCAATTCTC ATCGCACACG GCTTTCCCCA TGTAATGTAT CTATTTCCT.0
    CHI  TGCAATTCTC ATCGCACACG GCTTTCCCCA TGTAATGTAT CTATTTCCT.0
    CRA  TGCAATTCTC ATCGCACACG GCTTTCCCCA TGTAATGTAT CTATTTCCT.0
 ENGno3  TGCAATTCTC ATCGCACACG GCTTTCCCCA TGTAATGTAT CTATTTCCT.0
    FAR  TGCAATTCTC ATCGCACACG GCTTTCCCCA TGTAATGTAT CTATTTCCT.0
 GLAno2  TGCAATTCTC ATCGCACACG GCTTTCCCCA TGTAATGTAT CTATTTCCT.0
    GLE  TGCAATTCTC ATCGCACACG GCTTTCCCCA TGTAATGTAT CTATTTCCT.0
 JEZno1  TGCAATTCTC ATCGCACACG GCTTTCCCCA TGTAATGTAT CTATTTCCT.0
```

```
   KOR  TGCAATTCTC ATCGCACACG GCTTTCCCCA TGTAATGTAT CTATTTCCT.0
   KOY  TGCAATTCTC ATCGCACACG GCTTTCCCCA TGTAATGTAT CTATTTCCT.0
   LIK  TGCAATTCTC ATCGCACACG GCTTTCCCCA TGTAATGTAT CTATTTCCT.0
MARno2  TGCAATTCTC ATCGCACACG GCTTTCCCCA TGTAATGTAT CTATTTCCT.0
   MAT  TGCAATTCTC ATCGCACACG GCTTTCCCCA TGTAATGTAT CTATTTCCT.0
   MAX  TGCAATTCTC ATCGCACACG GCTTTCCCCA TGTAATGTAT CTATTTCCT.0
   MEX  TGCAATTCTC ATCGCACACG GCTTTCCCCA TGTAATGTAT CTATTTCCT.0
   MEY  TGCAATTCTC ATCGCACACG GCTTTCCCCA TGTAATGTAT CTATTTCCT.0
   MON  TGCAATTCTC ATCGCACACG GCTTTCCCCA TGTAATGTAT CTATTTCCT.0
   MOR  TGCAATTCTC ATCGCACACG GCTTTCCCCA TGTAATGTAT CTATTTCCT.0
   OBO  TGCAATTCTC ATCGCACACG GCTTTCCCCA TGTAATGTAT CTATTTCCT.0
   OMO  TGCAATTCTC ATCGCACACG GCTTTCCCCA TGTAATGTAT CTATTTCCT.0
   ORI  TGCAATTCTC ATCGCACACG GCTTTCCCCA TGTAATGTAT CTATTTCCT.0
PINBANK  TGCAATTCTC ATTGCACACG GCTTTCCCCA TGTAATGTAT CTATTTCCT.0
PINKOR  TGCAATTCTC ATTGCACACG GCTTTCC... .......TAT ...TTTCCTT1
   PUN  TGCAATTCTC ATCGCACACG GCTTTCCCCA TGTAATGTAT CTATTTCCT.0
   PUR  TGCAATTCTC ATCGCACACG GCTTTCCCCA TGTAATGTAT CTATTTCCT.0
   RET  TGCAATTCTC ATCGCACACG GCTTTCCCCA TGTAATGTAT CTATTTCCT.0
   RUB  TGCAATTCTC ATCGCACACG GCTTTCCCCA TGTAATGTAT CTATTTCCT.0
   SCH  TGCAATTCTC ATCGCACACG GCTTTCCCCA TGTAATGTAT CTATTTCCT.0
SITno1  TGCAATTCTC ATCGCACACG GCTTTCCCCA TGTAATGTAT CTATTTCCT.0
   SMI  TGCAATTCTC ATCGCACACG GCTTTCCCCA TGTAATGTAT CTATTTCCT.0
   SPI  TGCAATTCTC ATCGCACACG GCTTTCCCCA TGTAATGTAT CTATTTCCT.0
   TOR  TGCAATTCTC ATCGCACACG GCTTTCCCCA TGTAATGTAT CTATTTCCT.0
   WIL  TGCAATTCTC ATCGCACACG GCTTTCCCCA TGTAATGTAT CTATTTCCT.0

        [101                                                150]
ABIno1  AATCATTTTA GTTCTCGGAT GGGAAAATAA AAATGATTCT GAATCGAGGC
   ALC  AATCATTTTA GTTCTCGGAT GGCAAAATAA AAATGATTCT GAATCGAGGC
   ASP  AATCATTTTA GTTCTCGGAT GGGAAAATAA AAATGATTCT GAATCGAGGC
   BRA  AATCATTTTA GTTCTCGGAT GGCAAAATAA AAATGATTCT GAATCGAGGC
   BRE  AATCATTTTA GTTCTCGGAT GGCAAAATAA AAATGATTCT GAATCGAGGC
   CHI  AATCATTTTA GTTCTCGGAT GGCAAAATAA AAATGATTCT GAATCGAGGC
   CRA  AATCATTTTA GTTCTCGGAT GGGAAAATAA AAATGATTCT GAATCGAGGC
ENGno3  AATCATTTTA GTTCTCGGAT GGCAAAATAA AAATGATTCT GAATCGAGGC
   FAR  AATCATTTTA GTTCTCGGAT GGCAAAATAA AAATGATTCT GAATCGAGGC
GLAno2  AATCATTTTA GTTCTCGGAT GGCAAAATAA AAATGATTCT GAATCGAGGC
   GLE  AATCATTTTA GTTCTCGGAT GGCAAAATAA AAATGATTCT GAATCGAGGC
JEZno1  AATCATTTTA GTTCTCGGAT GGCAAAATAA AAATGATTCT GAATCGAGGC
   KOR  AATCATTTTA GTTCTCGGAT GGCAAAATAA AAATGATTCT GAATCGAGGC
   KOY  AATCATTTTA GTTCTCGGAT GGGAAAATAA AAATGATTCT GAATCGAGGC
   LIK  AATCATTTTA GTTCTCGGAT GGCAAAATAA AAATGATTCT GAATCGAGGC
MARno2  AATCATTTTA GTTCTCGGAT GGCAAAATAA AAATGATTCT GAATCGAGGC
   MAT  AATCATTTTA GTTCTCGGAT GGCAAAATAA AAATGATTCT GAATCGAGGC
   MAX  AATCATTTTA GTTCTCGGAT GGGAAAATAA AAATGATTCT GAATCGAGGC
   MEX  AATCATTTTA GTTCTCGGAT GGCAAAATAA AAATGATTCT GAATCGAGGC
   MEY  AATCATTTTA GTTCTCGGAT GGCAAAATAA AAATGATTCT GAATCGAGGC
   MON  AATCATTTTA GTTCTCGGAT GGGAAAATAA AAATGATTCT GAATCGAGGC
   MOR  AATCATTTTA GTTCTCGGAT GGCAAAATAA AAATGATTCT GAATCGAGGC
   OBO  AATCATTTTA GTTCTCGGAT GGGAAAATAA AAATGATTCT GAATCGAGGC
   OMO  AATCATTTTA GTTCTCGGAT GGCAAAATAA AAATGATTCT GAATCGAGGC
   ORI  AATCATTTTA GTTCTCGGAT GGCAAAATAA AAATGATTCT GAATCGAGGC
PINBANK  AATCATTTTA GTTCTCGGAT AGAAAAATAA AAATGATTCT GAATCGAGGC
PINKOR  AATACTTTTA GTTCTCGGAT AGAAAAAGAA AAATGATTCT GAATCGAGGC
   PUN  AATCATTTTA GTTCTCGGAT GGCAAAATAA AAATGATTCT GAATCGAGGC
   PUR  AATCATTTTA GTTCTCGGAT GGCAAAATAA AAATGATTCT GAATCGAGGC
   RET  AATCATTTTA GTTCTCGGAT GGGAAAATAA AAATGATTCT GAATCGAGGC
   RUB  AATCATTTTA GTTCTCGGAT GGCAAAATAA AAATGATTCT GAATCGAGGC
   SCH  AATCATTTTA GTTCTCGGAT GGCAAAATAA AAATGATTCT GAATCGAGGC
SITno1  AATCATTTTA GTTCTCGGAT GGCAAAATAA AAATGATTCT GAATCGAGGC
   SMI  AATCATTTTA GTTCTCGGAT GGCAAAATAA AAATGATTCT GAATCGAGGC
   SPI  AATCATTTTA GTTCTCGGAT GGCAAAATAA AAATGATTCT GAATCGAGGC
   TOR  AATCATTTTA GTTCTCGGAT GGGAAAATAA AAATGATTCT GAATCGAGGC
   WIL  AATCATTTTA GTTCTCGGAT GGCAAAATAA AAATGATTCT GAATCGAGGC

        [151                                                200]
ABIno1  AATGACTCAA AGAGCATTTC ATTGGTCAAA TAAGTTTTCT ATTTTCAGAT
   ALC  AATGACTCAA AGAGCATTTC ATTGGTCAAA TAAGTTTTCT ATTTTCAGAT
   ASP  AATGACTCAA AGAGCATTTC ATTGGTCAAA TAAGTTTTCT ATTTTCAGAT
   BRA  AATGACTCAA AGAGCATTTC ATTGGTCAAA TAAGTTTTCT ATTTTCAGAT
   BRE  AATGACTCAA AGAGCATTTC ATTGGTCAAA TAAGTTTTCT ATTTTCAGAT
   CHI  AATGACTCAA AGAGCATTTC ATTGGTCAAA TAAGTTTTCT ATTTTCAGAT
   CRA  AATGACTCAA AGAGCATTTC ATTGGTCAAA TAAGTTTTCT ATTTTCAGAT
ENGno3  AATGACTCAA AGAGCATTTC ATTGGTCAAA TAAGTTTTCT ATTTTCAGAT
   FAR  AATGACTCAA AGAGCATTTC ATTGGTCAAA TAAGTTTTCT ATTTTCAGAT
GLAno2  AATGACTCAA AGAGCATTTC ATTGGTCAAA TAAGTTTTCT ATTTTCAGAT
   GLE  AATGACTCAA AGAGCATTTC ATTGGTCAAA TAAGTTTTCT ATTTTCAGAT
JEZno1  AATGACTCAA AGAGCATTTC ATTGGTCAAA TAAGTTTTCT ATTTTCAGAT
   KOR  AATGACTCAA AGAGCATTTC ATTGGTCAAA TAAGTTTTCT ATTTTCAGAT
   KOY  AATGACTCAA AGAGCATTTC ATTGGTCAAA TAAGTTTTCT ATTTTCAGAT
   LIK  AATGACTCAA AGAGCATTTC ATTGGTCAAA TAAGTTTTCT ATTTTCAGAT
MARno2  AATGACTCAA AGAGCATTTC ATTGGTCAAA TAAGTTTTCT ATTTTCAGAT
   MAT  AATGACTCAA AGAGCATTTC ATTGGTCAAA TAAGTTTTCT ATTTTCAGAT
   MAX  AATGACTCAA AGAGCATTTC ATTGGTCAAA TAAGTTTTCT ATTTTCAGAT
```

```
   MEX  AATGACTCAA AGAGCATTTC ATTGGTCAAA TAAGTTTTCT ATTTTCAGAT
   MEY  AATGACTCAA AGAGCATTTC ATTGGTCAAA TAAGTTTTCT ATTTTCAGAT
   MON  AATGACTCAA AGAGCATTTC ATTGGTCAAA TAAGTTTTCT ATTTTCAGAT
   MOR  AATGACTCAA AGAGCATTTC ATTGGTCAAA TAAGTTTTCT ATTTTCAGAT
   OBO  AATGACTCAA AGAGCATTTC ATTGGTCAAA TAAGTTTTCT ATTTTCAGAT
   OMO  AATGACTCAA AGAGCATTTC ATTGGTCAAA TAAGTTTTCT ATTTTCAGAT
   ORI  AATGACTCAA AGAGCATTTC ATTGGTCAAA TAAGTTTTCT ATTTTCAGAT
PINBANK AATTACTCAA AGAGCATTTC ATTGGTCAAA TAAGTTTTCT ATTTTAAGAT
PINKOR AATTACTCAA AGAGCATTTC ATTGATCAAA TAAGTTTTCT ATTTTAAGAT
   PUN  AATGACTCAA AGAGCATTTC ATTGGTCAAA TAAGTTTTCT ATTTTCAGAT
   PUR  AATGACTCAA AGAGCATTTC ATTGGTCAAA TAAGTTTTCT ATTTTCAGAT
   RET  AATGACTCAA AGAGCATTTC ATTGGTCAAA TAAGTTTTCT ATTTTCAGAT
   RUB  AATGACTCAA AGAGCATTTC ATTGGTCAAA TAAGTTTTCT ATTTTCAGAT
   SCH  AATGACTCAA AGAGCATTTC ATTGGTCAAA TAAGTTTTCT ATTTTAAGAT
SITno1 AATGACTCAA AGAGCATTTC ATTGGTCAAA TAAGTTTTCT ATTTTCAGAT
   SMI  AATGACTCAA AGAGCATTTC ATTGGTCAAA TAAGTTTTCT ATTTTCAGAT
   SPI  AATGACTCAA AGAGCATTTC ATTGGTCAAA TAAGTTTTCT ATTTTCAGAT
   TOR  AATGACTCAA AGAGCATTTC ATTGGTCAAA TAAGTTTTCT ATTTTCAGAT
   WIL  AATGACTCAA AGAGCATTTC ATTGGTCAAA TAAGTTTTCT ATTTTCAGAT

        [201                                                250]
ABIno1 CCTGTATCTT TTGCCAGGAA TGGGCTAGGG GATTTATCTG GGGAATATCT
   ALC  CCTGTATCTT TTGCCAGGAA TGGGCTAGGG GATTTATCTG GGGAATATCT
   ASP  CCTGTATCTT TTGCCAGGAA TGGGCTAGGG GATTTATCTG GGGAATATCT
   BRA  CCTGTATCTT TTGCCAGGAA TGGGCTAGGG GATTTATCTG GGGAATATCT
   BRE  CCTGTATCTT TTGCCAGGAA TGGGCTAGGG GATTTATCTG GGGAATATCT
   CHI  CCTGTATCTT TTGCCAGGAA TGGGCTAGGG GATTTATCTG GGGAATATCT
   CRA  CCTGTATCTT TTGCCAGGAA TGGGCTAGGG GATTTATCTG GGGAATATCT
ENGno3 CCTGTATCTT TTGCCAGGAA TGGGCTAGGG GATTTATCTG GGGAATATCT
   FAR  CCTGTATCTT TTGCCAGGAA TGGGCTAGGG GATTTATCTG GGGAATATCT
GLAno2 CCTGTATCTT TTGCCAGGAA TGGGCTAGGG GATTTATCTG GGGAATATCT
   GLE  CCTGTATCTT TTGCCAGGAA TGGGCTAGGG GATTTATCTG GGGAATATCT
JEZno1 CCTGTATCTT TTGCCAGGAA TGGGCTAGGG GATTTATCTG GGGAATATCT
   KOR  CCTGTATCTT TTGCCAGGAA TGGGCTAGGG GATTTATCTG GGGAATATCT
   KOY  CCTGTATCTT TTGCCAGGAA TGGGCTAGGG GATTTATCTG GGGAATATCT
   LIK  CCTGTATCTT TTGCCAGGAA TGGGCTAGGG GATTTATCTG GGGAATATCT
MARno2 CCTGTATCTT TTGCCAGGAA TGGGCTAGGG GATTTATCTG GGGAATATCT
   MAT  CCTGTATCTT TTGCCAGGAA TGGGCTAGGG GATTTATCTG GGGAATATCT
   MAX  CCTGTATCTT TTGCCAGGAA TGGGCTAGGG GATTTATCTG GGGAATATCT
   MEX  CCTGTATCTT TTGCCAGGAA TGGGCTAGGG GATTTATCTG GGGAATATCT
   MEY  CCTGTATCTT TTGCCAGGAA TGGGCTAGGG GATTTATCTG GGGAATATCT
   MON  CCTGTATCTT TTGCCAGGAA TGGGCTAGGG GATTTATCTG GGGAATATCT
   MOR  CCTGTATCTT TTGCCAGGAA TGGGCTAGGG GATTTATCTG GGGAATATCT
   OBO  CCTGTATCTT TTGCCAGGAA TGGGCTAGGG GATTTATCTG GGGAATATCT
   OMO  CCTGTATCTT TTGCCAGGAA TGGGCTAGGG GATTTATCTG GGGAATATCT
   ORI  CCTGTATCTT TTGCCAGGAA TGGGCTAGGG GATTTATCTG GGGAATATCT
PINBANK CCTGTATCTT TTGCCAGGAA TTAGCTAGGG GATTTATCTG GGGAATATCT
PINKOR CCTGTATCTT TTGCCAGGAA TTAACTAGGG GATTTATCTG GGGAATATCT
   PUN  CCTGTATCTT TTGCCAGGAA TGGGCTAGGG GATTTATCTG GGGAATATCT
   PUR  CCTGTATCTT TTGCCAGGAA TGGGCTAGGG GATTTATCTG GGGAATATCT
   RET  CCTGTATCTT TTGCCAGGAA TGGGCTAGGG GATTTATCTG GGGAATATCT
   RUB  CCTGTATCTT TTGCCAGGAA TGGGCTAGGG GATTTATCTG GGGAATATCT
   SCH  CCTGTATCTT TTGCCAGGAA TGGGCTAGGG GATTTATCTG GGGAATATCT
SITno1 CCTGTATCTT TTGCCAGGAA TGGGCTAGGG GATTTATCTG GGGAATATCT
   SMI  CCTGTATCTT TTGCCAGGAA TGGGCTAGGG GATTTATCTG GGGAATATCT
   SPI  CCTGTATCTT TTGCCAGGAA TGGGCTAGGG GATTTATCTG GGGAATATCT
   TOR  CCTGTATCTT TTGCCAGGAA TGGGCTAGGG GATTTATCTG GGGAATATCT
   WIL  CCTGTATCTT TTGCCAGGAA TGGGCTAGGG GATTTATCTG GGGAATATCT

        [251                                                300]
ABIno1 GAATGCCAAA TTCGTTCTCT CTGTGAACGG GCCGCCGCTT TTGATGAAAA
   ALC  GAATGCCAAA TTCGTTCTCT CTGTGAACGG GCCGCCGCTT TTGATGAAAA
   ASP  GAATGCCAAA TTCGTTCTCT CTGTGAACGG GCCGCCGCTT TTGATGAAAA
   BRA  GAATGCCAAA TTCGTTCTCT CTGTGAACGG GCCGCCGCTT TTGATGAAAA
   BRE  GAATGCCAAA TTCGTTCTCT CTGTGAACGG GCCGCCGCTT TTGATGAAAA
   CHI  GAATGCCAAA TTCGTTCTCT CTGTGAACGG GCCGCCGCTT TTGATGAAAA
   CRA  GAATGCCAAA TTCGTTCTCT CTGTGAACGG GCCGCCGCTT TTGATGAAAA
ENGno3 GAATGCCAAA TTCGTTCTCT CTGTGAACGG GCCGCCGCTT TTGATGAAAA
   FAR  GAATGCCAAA TTCGTTCTCT CTGTGAACGG GCCGCCGCTT TTGATGAAAA
GLAno2 GAATGCCAAA TTCGTTCTCT CTGTGAACGG GCCGCCGCTT TTGATGAAAA
   GLE  GAATGCCAAA TTCGTTCTCT CTGTGAACGG GCCGCCGCTT TTGATGAAAA
JEZno1 GAATGCCAAA TTCGTTCTCT CTGTGAACGG GCCGCCGCTT TTGATGAAAA
   KOR  GAATGCCAAA TTCGTTCTCT CTGTGAACGG GCCGCCGCTT TTGATGAAAA
   KOY  GAATGCCAAA TTCGTTCTCT CTGTGAACGG GCCGCCGCTT TTGATGAAAA
   LIK  GAATGCCAAA TTCGTTCTCT CTGTGAACGG GCCGCCGCTT TTGATGAAAA
MARno2 GAATGCCAAA TTCGTTCTCT CTGTGAACGG GCCGCCGCTT TTGATGAAAA
   MAT  GAATGCCAAA TTCGTTCTCT CTGTGAACGG GCCGCCGCTT TTGATGAAAA
   MAX  GAATGCCAAA TTCGTTCTCT CTGTGAACGG GCCGCCGCTT TTGATGAAAA
   MEX  GAATGCCAAA TTCGTTCTCT CTGTGAACGG GCCGCCGCTT TTGATGAAAA
   MEY  GAATGCCAAA TTCGTTCTCT CTGTGAACGG GCCGCCGCTT TTGATGAAAA
   MON  GAATGCCAAA TTCGTTCTCT CTGTGAACGG GCCGCCGCTT TTGATGAAAA
   MOR  GAATGCCAAA TTCGTTCTCT CTGTGAACGG GCCGCCGCTT TTGATGAAAA
   OBO  GAATGCCAAA TTCGTTCTCT CTGTGAACGG GCCGCCGCTT TTGATGAAAA
   OMO  GAATGCCAAA TTCGTTCTCT CTGTGAACGG GCCGCCGCTT TTGATGAAAA
```

```
   ORI  GAATGCCAAA TTCGTTCTCT CTGTGAACGG GCCGCCGCTT TTGATGAAAA
PINBANK GAATGCCAAA TTCGTTCTCT CTGCGAACGG GCCGCCGCTT TTGATGAAAA
 PINKOR GAATGCCAAA TTCGTTCTCT CTGCGAACGG GCCGCCGCTT TTGACGAAAA
   PUN  GAATGCCAAA TTCGTTCTCT CTGTGAACGG GCCGCCGCTT TTGATGAAAA
   PUR  GAATGCCAAA TTCGTTCTCT CTGTGAACGG GCCGCCGCTT TTGATGAAAA
   RET  GAATGCCAAA TTCGTTCTCT CTGTGAACGG GCCGCCGCTT TTGATGAAAA
   RUB  GAATGCCAAA TTCGTTCTCT CTGTGAACGG GCCGCCGCTT TTGATGAAAA
   SCH  GAATGCCAAA TTCGTTCTCT CTGTGAACGG GCCGCCGCTT TTGATGAAAA
 SITno1 GAATGCCAAA TTCGTTCTCT CTGTGAACGG GCCGCCGCTT TTGATGAAAA
   SMI  GAATGCCAAA TTCGTTCTCT CTGTGAACGG GCCGCCGCTT TTGATGAAAA
   SPI  GAATGCCAAA TTCGTTCTCT CTGTGAACGG GCCGCCGCTT TTGATGAAAA
   TOR  GAATGCCAAA TTCGTTCTCT CTGTGAACGG GCCGCCGCTT TTGATGAAAA
   WIL  GAATGCCAAA TTCGTTCTCT CTGTGAACGG GCCGCCGCTT TTGATGAAAA

        [301                                                350]
 ABIno1 AGGCGGAGAA TCCAATTCTC GTTCTTTTGA AAACGATTTT TTAAAGAGTT
   ALC  AGGCGGAGAA TCCAATTCTC GTTCTTTTGA AAACGATTTT TTAAAGAGTT
   ASP  AGGCGGAGAA TCCAATTCTC GTTCTTTTGA AAACGATTTT TTAAAGAGTT
   BRA  AGGCGGAGAA TCCAATTCTC GTTCTTTTGA AAACGATTTT TTAAAGAGTT
   BRE  AGGCGGAGAA TCCAATTCTC GTTCTTTTGA AAACGATTTT TTAAAGAGTT
   CHI  AGGCGGAGAA TCCAATTCTC GTTCTTTTGA AAACGATTTT TTAAAGAGTT
   CRA  AGGCGGAGAA TCCAATTCTC GTTCTTTTGA AAACGATTTT TTAAAGAGTT
 ENGno3 AGGCGGAGAA TCCAATTCTC GTTCTTTTGA AAACGATTTT TTAAAGAGTT
   FAR  AGGCGGAGAA TCCAATTCTC GTTCTTTTGA AAACGATTTT TTAAAGAGTT
 GLAno2 AGGCGGAGAA TCCAATTCTC GTTCTTTTGA AAACGATTTT TTAAAGAGTT
   GLE  AGGCGGAGAA TCCAATTCTC GTTCTTTTGA AAACGATTTT TTAAAGAGTT
 JEZno1 AGGCGGAGAA TCCAATTCTC GTTCTTTTGA AAACGATTTT TTAAAGAGTT
   KOR  AGGCGGAGAA TCCAATTCTC GTTCTTTTGA AAACGATTTT TTAAAGAGTT
   KOY  AGGCGGAGAA TCCAATTCTC GTTCTTTTGA AAACGATTTT TTAAAGAGTT
   LIK  AGGCGGAGAA TCCAATTCTC GTTCTTTTGA AAACGATTTT TTAAAGAGTT
 MARno2 AGGCGGAGAA TCCAATTCTC GTTCTTTTGA AAACGATTTT TTAAAGAGTT
   MAT  AGGCGGAGAA TCCAATTCTC GTTCTTTTGA AAACGATTTT TTAAAGAGTT
   MAX  AGGCGGAGAA TCCAATTCTC GTTCTTTTGA AAACGATTTT TTAAAGAGTT
   MEX  AGGCGGAGAA TCCAATTCTC GTTCTTTTGA AAACGATTTT TTAAAGAGTT
   MEY  AGGCGGAGAA TCCAATTCTC GTTCTTTTGA AAACGATTTT TTAAAGAGTT
   MON  AGGCGGAGAA TCCAATTCTC GTTCTTTTGA AAACGATTTT TTAAAGAGTT
   MOR  AGGCGGAGAA TCCAATTCTC GTTCTTTTGA AAACGATTTT TTAAAGAGTT
   OBO  AGGCGGAGAA TCCAATTCTC GTTCTTTTGA AAACGATTTT TTAAAGAGTT
   OMO  AGGCGGAGAA TCCAATTCTC GTTCTTTTGA AAACGATTTT TTAAAGAGTT
   ORI  AGGCGGAGAA TCCAATTCTC GTTCTTTTGA AAACGATTTT TTAAAGAGTT
PINBANK GCGCAGAGAA TCAAATTCTC GTTCTTTTGA AAACGATTTT TTAAAGAGTT
 PINKOR AGGCAGAGAA TCAAATTCTC GTTCTTTTGA AAACGATTTT TGAAAGAGTT
   PUN  AGGCGGAGAA TCCAATTCTC GTTCTTTTGA AAACGATTTT TTAAAGAGTT
   PUR  AGGCGGAGAA TCCAATTCTC GTTCTTTTGA AAACGATTTT TTAAAGAGTT
   RET  AGGCGGAGAA TCCAATTCTC GTTCTTTTGA AAACGATTTT TTAAAGAGTT
   RUB  AGGCGGAGAA TCCAATTCTC GTTCTTTTGA AAACGATTTT TTAAAGAGTT
   SCH  AGGCGGAGAA TCCAATTCTC GTTCTTTTGA AAACGATTTT TTCAAGAGTT
 SITno1 AGGCGGAGAA TCCAATTCTC GTTCTTTTGA AAACGATTTT TTAAAGAGTT
   SMI  AGGCGGAGAA TCCAATTCTC GTTCTTTTGA AAACGATTTT TTAAAGAGTT
   SPI  AGGCGGAGAA TCCAATTCTC GTTCTTTTGA AAACGATTTT TTAAAGAGTT
   TOR  AGGCGGAGAA TCCAATTCTC GTTCTTTTGA AAACGATTTT TTAAAGAGTT
   WIL  AGGCGGAGAA TCCAATTCTC GTTCTTTTGA AAACGATTTT TTAAAGAGTT

        [351                                                400]
 ABIno1 CTGGACCTAA TTCCTTCCGA ACTACACGTA TCGTACTTTT ATGTTTACAG
   ALC  CTGGACCTAA TTCCTTCCGA ACTACACGTA TCGTACTTTT ATGTTTACAG
   ASP  CTGGACCTAA TTCCTTCCGA ACTACACGTA TCGTACTTTT ATGTTTACAG
   BRA  CTGGACCTAA TTCCTTCCGA ACTACACGTA TCGTACTTTT ATGTTTACAG
   BRE  CTGGACCTAA TTCCTTCCGA ACTACACGTA TCGTACTTTT ATGTTTACAG
   CHI  CTGGACCTAA TTCCTTCCGA ACTACACGTA TCGTACTTTT ATGTTTACAG
   CRA  CTGGACCTAA TTCCTTCCGA ACTACACGTA TCGTACTTTT ATGTTTACAG
 ENGno3 CTGGACCTAA TTCCTTCCGA ACTACACGTA TCGTACTTTT ATGTTTACAG
   FAR  CTGGACCTAA TTCCTTCCGA ACTACACGTA TCGTACTTTT ATGTTTACAG
 GLAno2 CTGGACCTAA TTCCTTCCGA ACTACACGTA TCGTACTTTT ATGTTTACAG
   GLE  CTGGACCTAA TTCCTTCCGA ACTACACGTA TCGTACTTTT ATGTTTACAG
 JEZno1 CTGGACCTAA TTCCTTCCGA ACTACACGTA TCGTACTTTT ATGTTTACAG
   KOR  CTGGACCTAA TTCCTTCCGA ACTACACGTA TCGTACTTTT ATGTTTACAG
   KOY  CTGGACCTAA TTCCTTCCGA ACTACACGTA TCGTACTTTT ATGTTTACAG
   LIK  CTGGACCTAA TTCCTTCCGA ACTACACGTA TCGTACTTTT ATGTTTACAG
 MARno2 CTGGACCTAA TTCCTTCCGA ACTACACGTA TCGTACTTTT ATGTTTACAG
   MAT  CTGGACCTAA TTCCTTCCGA ACTACACGTA TCGTACTTTT ATGTTTACAG
   MAX  CTGGACCTAA TTCCTTCCGA ACTACACGTA TCGTACTTTT ATGTTTACAG
   MEX  CTGGACCTAA TTCCTTCCGA ACTACACGTA TCGTACTTTT ATGTTTACAG
   MEY  CTGGACCTAA TTCCTTCCGA ACTACACGTA TCGTACTTTT ATGTTTACAG
   MON  CTGGACCTAA TTCCTTCCGA ACTACACGTA TCGTACTTTT ATGTTTACAG
   MOR  CTGGACCTAA TTCCTTCCGA ACTACACGTA TCGTACTTTT ATGTTTACAG
   OBO  CTGGACCTAA TTCCTTCCGA ACTACACGTA TCGTACTTTT ATGTTTACAG
   OMO  CTGGACCTAA TTCCTTCCGA ACTACACGTA TCGTACTTTT ATGTTTACAG
   ORI  CTGGACCTAA TTCCTTCCGA ACTACACGTA TCGTACTTTT ATGTTTACAG
PINBANK CTGGACCTAA TTCCTTCCGA ACTACACGTA TCGTACTTTT ATGTTTACAG
 PINKOR CTGGACCTAA TTCCTTCCGA ACTACACGTA TCGTACTTTT ATGTTTACAG
   PUN  CTGGACCTAA TTCCTTCCGA ACTACACGTA TCGTACTTTT ATGTTTACAG
   PUR  CTGGACCTAA TTCCTTCCGA ACTACACGTA TCGTACTTTT ATGTTTACAG
   RET  CTGGACCTAA TTCCTTCCGA ACTACACGTA TCGTACTTTT ATGTTTACAG
```

```
   RUB  CTGGACCTAA TTCCTTCCGA ACTACACGTA TCGTACTTTT ATGTTTACAG
   SCH  CTGGACCTAA TTCCTTCCGA ACTACACGTA TCGTACTTTT ATGTTTACAG
SITno1  CTGGACCTAA TTCCTTCCGA ACTACACGTA TCGTACTTTT ATGTTTACAG
   SMI  CTGGACCTAA TTCCTTCCGA ACTACACGTA TCGTACTTTT ATGTTTACAG
   SPI  CTGGACCTAA TTCCTTCCGA ACTACACGTA TCGTACTTTT ATGTTTACAG
   TOR  CTGGACCTAA TTCCTTCCGA ACTACACGTA TCGTACTTTT ATGTTTACAG
   WIL  CTGGACCTAA TTCCTTCCGA ACTACACGTA TCGTACTTTT ATGTTTACAG

       [401                                                450]
ABIno1  GCTAAAGTTT TAGCACATGA TAATGAAAGT ATATACTTTA TACGATATAA
   ALC  GCTAAAGTTT TAGCACATGA TAATGAAAGT ATATACTTTA TACGATATAA
   ASP  GCTAAAGTTT TAGCACATGA TAATGAAAGT ATATACTTTA TACGATATAA
   BRA  GCTAAAGTTT TAGCACATGA TAATGAAAGT ATATACTTTA TACGATATAA
   BRE  GCTAAAGTTT TAGCACATGA TAATGAAAGT ATATACTTTA TACGATATAA
   CHI  GCTAAAGTTT TAGCACATGA TAATGAAAGT ATATACTTTA TACGATATAA
   CRA  GCTAAAGTTT TAGCACATGA TAATGAAAGT ATATACTTTA TACGATATAA
ENGno3  GCTAAAGTTT TAGCACATGA TAATGAAAGT ATATACTTTA TACGATATAA
   FAR  GCTAAAGTTT TAGCACATGA TAATGAAAGT ATATACTTTA TACGATATAA
GLAno2  GCTAAAGTTT TAGCACATGA TAATGAAAGT ATATACTTTA TACGATATAA
   GLE  GCTAAAGTTT TAGCACATGA TAATGAAAGT ATATACTTTA TACGATATAA
JEZno1  GCTAAAGTTT TAGCACATGA TAATGAAAGT ATATACTTTA TACGATATAA
   KOR  GCTAAAGTTT TAGCACATGA TAATGAAAGT ATATACTTTA TACGATATAA
   KOY  GCTAAAGTTT TAGCACATGA TAATGAAAGT ATATACTTTA TACGATATAA
   LIK  GCTAAAGTTT TAGCACATGA TAATGAAAGT ATATACTTTA TACGATATAA
MARno2  GCTAAAGTTT TAGCACATGA TAATGAAAGT ATATACTTTA TACGATATAA
   MAT  GCTAAAGTTT TAGCACATGA TAATGAAAGT ATATACTTTA TACGATATAA
   MAX  GCTAAAGTTT TAGCACATGA TAATGAAAGT ATATACTTTA TACGATATAA
   MEX  GCTAAAGTTT TAGCACATGA TAATGAAAGT ATATACTTTA TACGATATAA
   MEY  GCTAAAGTTT TAGCACATGA TAATGAAAGT ATATACTTTA TACGATATAA
   MON  GCTAAAGTTT TAGCACATGA TAATGAAAGT ATATACTTTA TACGATATAA
   MOR  GCTAAAGTTT TAGCACATGA TAATGAAAGT ATATACTTTA TACGATATAA
   OBO  GCTAAAGTTT TAGCACATGA TAATGAAAGT ATATACTTTA TACGATATAA
   OMO  GCTAAAGTTT TAGCACATGA TAATGAAAGT ATATACTTTA TACGATATAA
   ORI  GCTAAAGTTT TAGCACATGA TAATGAAAGT ATATACTTTA TACGATATAA
PINBANK  GCTAAAGTTT TAGCACATGA TAATGAAAGT ATATACTTTA TACGATATAA
PINKOR  GCTAAAGTTT TAGCACATGA TAATGAAAGT ATATACTTTA TACGATATAA
   PUN  GCTAAAGTTT TAGCACATGA TAATGAAAGT ATATACTTTA TACGATATAA
   PUR  GCTAAAGTTT TAGCACATGA TAATGAAAGT ATATACTTTA TACGATATAA
   RET  GCTAAAGTTT TAGCACATGA TAATGAAAGT ATATACTTTA TACGATATAA
   RUB  GCTAAAGTTT TAGCACATGA TAATGAAAGT ATATACTTTA TACGATATAA
   SCH  GCTAAAGTTT TAGCACATGA TAATGAAAGT ATATACTTTA TACGATATAA
SITno1  GCTAAAGTTT TAGCACATGA TAATGAAAGT ATATACTTTA TACGATATAA
   SMI  GCTAAAGTTT TAGCACATGA TAATGAAAGT ATATACTTTA TACGATATAA
   SPI  GCTAAAGTTT TAGCACATGA TAATGAAAGT ATATACTTTA TACGATATAA
   TOR  GCTAAAGTTT TAGCACATGA TAATGAAAGT ATATACTTTA TACGATATAA
   WIL  GCTAAAGTTT TAGCACATGA TAATGAAAGT ATATACTTTA TACGATATAA

       [451                                                500]
ABIno1  ACCATCTCGA CCAAAGGATC CACTGTAGTA ATGAAAAAGA TTTCTCCAAA
   ALC  ACCATCTCGA CCAAAGGATC CACTGTAGTA ATGAAAAAGA TTTCTCCAAA
   ASP  ACCATCTCGA CCAAAGGATC CACTGTAGTA ATGAAAAAGA TTTCTCCAAA
   BRA  ACCATCTCGA CCAAAGGATC CACTGTAGTA ATGAAAAAGA TTTCTCCAAA
   BRE  ACCATCTCGA CCAAAGGATC CACTGTAGTA ATGAAAAAGA TTTCTCCAAA
   CHI  ACCATCTCGA CCAAAGGATC CACTGTAGTA ATGAAAAAGA TTTCTCCAAA
   CRA  ACCATCTCGA CCAAAGGATC CACTGTAGTA ATGAAAAAGA TTTCTCCAAA
ENGno3  ACCATCTCGA CCAAAGGATC CACTGTAGTA ATGAAAAAGA TTTCTCCAAA
   FAR  ACCATCTCGA CCAAAGGATC CACTGTAGTA ATGAAAAAGA TTTCTCCAAA
GLAno2  ACCATCTCGA CCAAAGGATC CACTGTAGTA ATGAAAAAGA TTTCTCCAAA
   GLE  ACCATCTCGA CCAAAGGATC CACTGTAGTA ATGAAAAATA TTTCTCCAAA
JEZno1  ACCATCTCGA CCAAAGGATC CACTGTAGTA ATGAAAAAGA TTTCTCCAAA
   KOR  ACCATCTCGA CCAAAGGATC CACTGTAGTA ATGAAAAAGA TTTCTCCAAA
   KOY  ACCATCTCGA CCAAAGGATC CACTGTAGTA ATGAAAAAGA TTTCTCCAAA
   LIK  ACCATCTCGA CCAAAGGATC CACTGTAGTA ATGAAAAAGA TTTCTCCAAA
MARno2  ACCATCTCGA CCAAAGGATC CACTGTAGTA ATGAAAAATA TTTCTCCAAA
   MAT  ACCATCTCGA CCAAAGGATC CACTGTAGTA ATGAAAAAGA TTTCTCCAAA
   MAX  ACCATCTCGA CCAAAGGATC CACTGTAGTA ATGAAAAAGA TTTCTCCAAA
   MEX  ACCATCTCGA CCAAAGGATC CACTGTAGTA ATGAAAAAGA TTTCTCCAAA
   MEY  ACCATCTCGA CCAAAGGATC CACTGTAGTA ATGAAAAAGA TTTCTCCAAA
   MON  ACCATCTCGA CCAAAGGATC CACTGTAGTA ATGAAAAAGA TTTCTCCAAA
   MOR  ACCATCTCGA CCAAAGGATC CACTGTAGTA ATGAAAAAGA TTTCTCCAAA
   OBO  ACCATCTCGA CCAAAGGATC CACTGTAGTA ATGAAAAAGA TTTCTCCAAA
   OMO  ACCATCTCGA CCAAAGGATC CACTGTAGTA ATGAAAAATA TTTCTCCAAA
   ORI  ACCATCTCGA CCAAAGGATC CACTGTAGTA ATGAAAAAGA TTTCTCCAAA
PINBANK  ACCATCTCGA TCAAAGGATC CACTGTAGTA ATGAAAAAGA TTTCTCCAAA
PINKOR  ACCATCTCGA CCAAAGGATC CACTGTAGTA ATGAAAAAGA TTTCTCCAAA
   PUN  ACCATCTCGA CCAAAGGATC CACTGTAGTA ATGAAAAAGA TTTCTCCAAA
   PUR  ACCATCTCGA CCAAAGGATC CACTGTAGTA ATGAAAAAGA TTTCTCCAAA
   RET  ACCATCTCGA CCAAAGGATC CACTGTAGTA ATGAAAAAGA TTTCTCCAAA
   RUB  ACCATCTCGA CCAAAGGATC CACTGTAGTA ATGAAAAAGA TTTCTCCAAA
   SCH  ACCATCTCGA CCAAAGGATC CACTGTAGTA ATGAAAAAGA TTTCTCCAAA
SITno1  ACCATCTCGA CCAAAGGATC CACTGTAGTA ATGAAAAAGA TTTCTCCAAA
   SMI  ACCATCTCGA CCAAAGGATC CACTGTAGTA ATGAAAAAGA TTTCTCCAAA
   SPI  ACCATCTCGA CCAAAGGATC CACTGTAGTA ATGAAAAAGA TTTCTCCAAA
   TOR  ACCATCTCGA CCAAAGGATC CACTGTAGTA ATGAAAAAGA TTTCTCCAAA
```

```
   WIL  ACCATCTCGA CCAAAGGATC CACTGTAGTA ATGAAAAAGA TTTCTCCAAA

        [501                                                 550]
ABIno1  TTCGATCGAA TCGATCCAGG ATATCATCAT CTGTTAGACT GGTCCAAGAC
   ALC  TTCGATCGAA TCGATCCAGG ATATCATCAT CTGTTAGACT GGTCCAAGAC
   ASP  TTCGATCGAA TCGATCCAGG ATATCATCAT CTGTTAGACT GGTCCAAGAC
   BRA  TTCGATCGAA TCGATCCAGG ATATCATCAT CTGTTAGACT GGTCCAAGAC
   BRE  TTCGATCGAA TCGATCCAGG ATATCATCAT CTGTTAGACT GGTCCAAGAC
   CHI  TTCGATCGAA TCGATCCAGG ATATCATCAT CTGTTAGACT GGTCCAAGAC
   CRA  TTCGATCGAA TCGATCCAGG ATATCATCAT CTGTTAGACT GGTCCAAGAC
ENGno3  TTCGATCGAA TCGATCCAGG ATATCATCAT CTGTTAGACT GGTCCAAGAC
   FAR  TTCGATCGAA TCGATCCAGG ATATCATCAT CTGTTAGACT GGTCCAAGAC
GLAno2  TTCGATCGAA TCGATCCAGG ATATCATCAT CTGTTAGACT GGTCCAAGAC
   GLE  TTCGATCGAA TCGATCCAGG ATATCATCAT CTGTTAGACT GGTCCAAGAC
JEZno1  TTCGATCGAA TCGATCCAGG ATATCATCAT CTGTTAGACT GGTCCAAGAC
   KOR  TTCGATCGAA TCGATCCAGG ATATCATCAT CTGTTAGACT GGTCCAAGAC
   KOY  TTCGATCGAA TCGATCCAGG ATATCATCAT CTGTTAGACT GGTCCAAGAC
   LIK  TTCGATCGAA TCGATCCAGG ATATCATCAT CTGTTAGACT GGTCCAAGAC
MARno2  TTCGATCGAA TCGATCCAGG ATATCATCAT CTGTTAGACT GGTCCAAGAC
   MAT  TTCGATCGAA TCGATCCAGG ATATCATCAT CTGTTAGACT GGTCCAAGAC
   MAX  TTCGATCGAA TCGATCCAGG ATATCATCAT CTGTTAGACT GGTCCAAGAC
   MEX  TTCGATCGAA TCGATCCAGG ATATCATCAT CTGTTAGACT GGTCCAAGAC
   MEY  TTCGATCGAA TCGATCCAGG ATATCATCAT CTGTTAGACT GGTCCAAGAC
   MON  TTCGATCGAA TCGATCCAGG ATATCATCAT CTGTTAGACT GGTCCAAGAC
   MOR  TTCGATCGAA TCGATCCAGG ATATCATCAT CTGTTAGACT GGTCCAAGAC
   OBO  TTCGATCGAA TCGATCCAGG ATATCATCAT CTGTTAGACT GGTCCAAGAC
   OMO  TTCGATCGAA TCGATCCAGG ATATCATCAT CTGTTAGACT GGTCCAAGAC
   ORI  TTCGATCGAA TCGATCCAGG ATATCATCAT CTGTTAGACT GGTCCAAGAC
PINBANK  TTTGATCGAA TCGATCGAGG ATATCATCAT CTGTTAGACT GGTCCAAGAC
PINKOR  TTTGATCAAA TCGATTGAGG ATATCATCAT CTGTTAGATT GGTCCAAGAC
   PUN  TTCGATCGAA TCGATCCAGG ATATCATCAT CTGTTAGACT GGTCCAAGAC
   PUR  TTCGATCGAA TCGATCCAGG ATATCATCAT CTGTTAGACT GGTCCAAGAC
   RET  TTCGATCGAA TCGATCCAGG ATATCATCAT CTGTTAGACT GGTCCAAGAC
   RUB  TTCGATCGAA TCGATCCAGG ATATCATCAT CTGTTAGACT GGTCCAAGAC
   SCH  TTCGATCGAA TCGATCCAGG ATATCATCAT CTGTTAGACT GGTCCAAGAC
SITno1  TTCGATCGAA TCGATCCAGG ATATCATCAT CTGTTAGACT GGTCCAAGAC
   SMI  TTCGATCGAA TCGATCCAGG ATATCATCAT CTGTTAGACT GGTCCAAGAC
   SPI  TTCGATCGAA TCGATCCAGG ATATCATCAT CTGTTAGACT GGTCCAAGAC
   TOR  TTCGATCGAA TCGATCCAGG ATATCATCAT CTGTTAGACT GGTCCAAGAC
   WIL  TTCGATCGAA TCGATCCAGG ATATCATCAT CTGTTAGACT GGTCCAAGAC

        [551                                                 600]
ABIno1  AATTTACTAA TTGGCCGCCC TGATATATCA CAGAATTTTT CTCTAGACAA
   ALC  AATTTACTAA TTGGCCGCCC TGATATATCA CAGAATTTTT CTCTAGACAA
   ASP  AATTTACTAA TTGGCCGCCC TGATATATCA CAGAATTTTT CTCTAGACAA
   BRA  AATTTACTAA TTGGCCGCCC TGATATATCA CAGAATTTTT CTCTAGACAA
   BRE  AATTTACTAA TTGGCCGCCC TGATATATCA CAGAATTTTT CTCTAGACAA
   CHI  AATTTACTAA TTGGCCGCCC TGATATATCA CAGAATTTTT CTCTAGACAA
   CRA  AATTTACTAA TTGGCCGCCC TGATATATCA CAGAATTTTT CTCTAGACAA
ENGno3  AATTTACTAA TTGGCCGCCC TGATATATCA CAGAATTTTT CTCTAGACAA
   FAR  AATTTACTAA TTGGCCGCCC TGATATATCA CAGAATTTTT CTCTAGACAA
GLAno2  AATTTACTAA TTGGCCGCCC TGATATATCA CAGAATTTTT CTCTAGACAA
   GLE  AATTTACTAA TTGGCCGCCC TGATATATCA CAGAATTTTT CTCTAGACAA
JEZno1  AATTTACTAA TTGGCCGCCC TGATATATCA CAGAATTTTT CTCTAGACAA
   KOR  AATTTACTAA TTGGCCGCCC TGATATATCA CAGAATTTTT CTCTAGACAA
   KOY  AATTTACTAA TTGGCCGCCC TGATATATCA CAGAATTTTT CTCTAGACAA
   LIK  AATTTACTAA TTGGCCGCCC TGATATATCA CAGAATTTTT CTCTAGACAA
MARno2  AATTTACTAA TTGGCCGCCC TGATATATCA CAGAATTTTT CTCTAGACAA
   MAT  AATTTACTAA TTGGCCGCCC TGATATATCA CAGAATTTTT CTCTAGACAA
   MAX  AATTTACTAA TTGGCCGCCC TGATATATCA CAGAATTTTT CTCTAGACAA
   MEX  AATTTACTAA TTGGCCGCCC TGATATATCA CAGAATTTTT CTCTAGACAA
   MEY  AATTTACTAA TTGGCCGCCC TGATATATCA CAGAATTTTT CTCTAGACAA
   MON  AATTTACTAA TTGGCCGCCC TGATATATCA CAGAATTTTT CTCTAGACAA
   MOR  AATTTACTAA TTGGCCGCCC TGATATATCA CAGAATTTTT CTCTAGACAA
   OBO  AATTTACTAA TTGGCCGCCC TGATATATCA CAGAATTTTT CTCTAGACAA
   OMO  AATTTACTAA TTGGCCGCCC TGATATATCA CAGAATTTTT CTCTAGACAA
   ORI  AATTTACTAA TTGGCCGCCC TGATATATCA CAGAATTTTT CTCTAGACAA
PINBANK  AATTTACTAA TTGGCCGCCC TGATATGTCA CAGAATTTTT CTGTAGCCAA
PINKOR  AATTTACTAA TTGGCCGTCC TGATACGTCA CAGAATTTTT CTCTAGCCAA
   PUN  AATTTACTAA TTGGCCGCCC TGATATATCA CAGAATTTTT CTCTAGACAA
   PUR  AATTTACTAA TTGGCCGCCC TGATATATCA CAGAATTTTT CTCTAGACAA
   RET  AATTTACTAA TTGGCCGCCC TGATATATCA CAGAATTTTT CTCTAGACAA
   RUB  AATTTACTAA TTGGCCGCCC TGATATATCA CAGAATTTTT CTCTAGACAA
   SCH  AATTTACTAA TTGGCCGCCC TGATATATCA CAGAATTTTT CTCTAGACAA
SITno1  AATTTACTAA TTGGCCGCCC TGATATATCA CAGAATTTTT CTCTAGACAA
   SMI  AATTTACTAA TTGGCCGCCC TGATATATCA CAGAATTTTT CTCTAGACAA
   SPI  AATTTACTAA TTGGCCGCCC TGATATATCA CAGAATTTTT CTCTAGACAA
   TOR  AATTTACTAA TTGGCCGCCC TGATATATCA CAGAATTTTT CTCTAGACAA
   WIL  AATTTACTAA TTGGCCGCCC TGATATATCA CAGAATTTTT CTCTAGACAA

        [601                                                 650]
ABIno1  TAATCCAATT ATTGGTACAA TCGGAACTAT TGGATCAAAT TCATCGGTAA
   ALC  TAATCCAATT ATTGGTACAA TCGGAACTAT TGGATCAAAT TCATCAGTAA
   ASP  TAATCCAATT ATTGGTACAA TCGGAACTAT TGGATCAAAT TCATCGGTAA
```

```
   BRA  TAATCCAATT ATTGGTACAA TCGGAACTAT TGGATCAAAT TCATCAGTAA
   BRE  TAATCCAATT ATTGGTACAA TCGGAACTAT TGGATCAATT TCATCGGTAA
   CHI  TAATCCAATT ATTGGTACAA TCGGAACTAT TGGATCAAAT TCATCAGTAA
   CRA  TAATCCAATT ATTGGTACAA TCGGAACTAT TGGATCAAAT TCATCGGTAA
ENGno3  TAATCCAATT ATTGGTACAA TCGGAACTAT TGGATCAAAT TCATCGGTAA
   FAR  TAATCCAATT ATTGGTACAA TCGGAACTAT TGGATCAAAT TCATCGGTAA
GLAno2  TAATCCAATT ATTGGTACAA TCGGAACTAT TGGATCAAAT TCATCGGTAA
   GLE  TAATCCAATT ATTGGTACAA TCGGAACTAT TGGATCAAAT TCATCGGTAA
JEZno1  TAATCCAATT ATTGGTACAA TCGGAACTAT TGGATCAAAT TCATCGGTAA
   KOR  TAATCCAATT ATTGGTACAA TCGGAACTAT TGGATCAAAT TCATCGGTAA
   KOY  TAATCCAATT ATTGGTACAA TCGGAACTAT TGGATCAAAT TCATCGGTAA
   LIK  TAATCCAATT ATTGGTACAA TCGGAACTAT TGGATCAAAT TCATCGGTAA
MARno2  TAATCCAATT ATTGGTACAA TCGGAACTAT TGGATCAAAT TCATCGGTAA
   MAT  TAATCCAATT ATTGGTACAA TCGGAACTAT TGGATCAAAT TCATCAGTAA
   MAX  TAATCCAATT ATTGGTACAA TCGGAACTAT TGGATCAAAT TCATCAGTAA
   MEX  TAATCCAATT ATTGGTACAA TCGGAACTAT TGGATCAAAT TCATCGGTAA
   MEY  TAATCCAATT ATTGGTACAA TCGGAACTAT TGGATCAAAT TCATCGGTAA
   MON  TAATCCAATT ATTGGTACAA TCGGAACTAT TGGATCAAAT TCATCGGTAA
   MOR  TAATCCAATT ATTGGCACAA TCGGAACTAT TGGATCAAAT TCATCAGTAA
   OBO  TAATCCAATT ATTGGTACAA TCGGAACTAT TGGATCAAAT TCATCGGTAA
   OMO  TAATCCAATT ATTGGTACAA TCGGAACTAT TGGATCAAAT TCATCGGTAA
   ORI  TAATCCAATT ATTGGTACAA TCGGAACTAT TGGATCAAAT TCATCAGTAA
PINBANK  TAATCCAATT ATTGGTACAA TCGGAACTAT TGGATCAATT TCATCGGTAA
PINKOR  TAATCCAAGT ATTGGTACAA TCGGAACTAT TGGATCAAAT TCATTGGTAA
   PUN  TAATCCAATT ATTGGTACAA TCGGAACTAT TGGATCAAAT TCATCGGTAA
   PUR  TAATCCAATT ATTGGTACAA TCGGAACTAT TGGATCAAAT TCATCAGTAA
   RET  TAATCCAATT ATTGGTACAA TCGGAACTAT TGGATCAAAT TCATCGGTAA
   RUB  TAATCCAATT ATTGGTACAA TCGGAACTAT TGGATCAAAT TCATCGGTAA
   SCH  TAATCCAATT ATTGGTACAA TCGGAACTAT TGGATCAAAT TCATCGGTAA
SITno1  TAATCCAATT ATTGGTACAA TCGGAACTAT TGGATCAAAT TCATCGGTAA
   SMI  TAATCCAATT ATTGGTACAA TCGGAACTAT TGGATCAAAT TCATCGGTAA
   SPI  TAATCCAATT ATTGGTACAA TCGGAACTAT TGGATCAAAT TCATCAGTAA
   TOR  TAATCCAATT ATTGGTACAA TCGGAACTAT TGGATCAAAT TCATCGGTAA
   WIL  TAATCCAATT ATTGGTACAA TCGGAACTAT TGGATCAAAT TCATCAGTAA

        [651                                              700]
ABIno1  TAAGATCGGC AATGAATAAC TCATCTAGCA TTTTAGTCTT GACCAAAAGA
   ALC  TAAGATCGGC AATGAATAAC TCATCTAGCA TTTTAGTCTT GACCAAAAGA
   ASP  TAAGATCGGC AATGAATAAC TCATCTAGCA TTTTAGTCTT GACCAAAAGA
   BRA  TAAGATCGGC AATGAATAAC TCATCTAGCA TTTTAGTCTT GACCAAAAGA
   BRE  TAAGATCGGC AATGAATAAC TCATCTAGCA TTTTAGTCTT GACCAAAAGA
   CHI  TAAGATCGGC AATGAATAAC TCATCTAGCA TTTTAGTCTT GACCAAAAGA
   CRA  TAAGATCGGC AATGAATAAC TCATCTAGCA TTTTAGTCTT GACCAAAAGA
ENGno3  TAAGATCGGC AATGAATAAC TCATCTAGCA TTTTAGTCTT GACCAAAAGA
   FAR  TAAGATCGGC AATGAATAAC TCATCTAGCA TTTTAGTCTT GACCAAAAGA
GLAno2  TAAGATCGGC AATGAATAAC TCATCTAGCA TTTTAGTCTT GACCAAAAGA
   GLE  TAAGATCGGC AATGAATAAC TCATCTAGCA TTTTAGTCTT GACCAAAAGA
JEZno1  TAAGATCGGC AATGAATAAC TCATCTAGCA TTTTAGTCTT GACCAAAAGA
   KOR  TAAGATCGGC AATGAATAAC TCATCTAGCA TTTTAGTCTT GACCAAAAGA
   KOY  TAAGATCGGC AATGAATAAC TCATCTAGCA TTTTAGTCTT GACCAAAAGA
   LIK  TAAGATCGGC AATGAATAAC TCATCTAGCA TTTTAGTCTT GACCAAAAGA
MARno2  TAAGATCGGC AATGAATAAC TCATCTAGCA TTTTAGTCTT GACCAAAAGA
   MAT  TAAGATCGGC AATGAATAAC TCATCTAGCA TTTTAGTCTT GACCAAAAGA
   MAX  TAAGATCGGC AATGAATAAC TCATCTAGCA TTTTAGTCTT GACCAAAAGA
   MEX  TAAGATCGGC AATGAATAAC TCATCTAGCA TTTTAGTCTT GACCAAAAGA
   MEY  TAAGATCGGC AATGAATAAC TCATCTAGCA TTTTAGTCTT GACCAAAAGA
   MON  TAAGATCGGC AATGAATAAC TCATCTAGCA TTTTAGTCTT GACCAAAAGA
   MOR  TAAGATCGGC AATGAATAAC TCATCTAGCA TTTTAGTCTT GACCAAAAGA
   OBO  TAAGATCGGC AATGAATAAC TCATCTAGCA TTTTAGTCTT GACCAAAAGA
   OMO  TAAGATCGGC AATGAATAAC TCATCTAGCA TTTTAGTCTT GACCAAAAGA
   ORI  TAAGATCGGC AATGAATAAC TCATCTAGCA TTTTAGTCTT GACCAAAAGA
PINBANK  TAAGATCGGC GATGAATAAC TCATCTAGCA TTTTGGTTCT GACCAAAATA
PINKOR  TAAGATCGGC GATGAATAAC TCATCTAGCA TTTTTGTTCT GACAAAAAGA
   PUN  TAAGATCGGC AATGAATAAC TCATCTAGCA TTTTAGTCTT GACCAAAAGA
   PUR  TAAGATCGGC AATGAATAAC TCATCTAGCA TTTTAGTCTT GACCAAAAGA
   RET  TAAGATCGGC AATGAATAAC TCATCTAGCA TTTTAGTCTT GACCAAAAGA
   RUB  TAAGATCGGC AATGAATAAC TCATCTAGCA TTTTAGTCTT GACCAAAAGA
   SCH  TAAGATCGGC AATGAATAAC TCATCTAGCA TTTTAGTCTT GACCAAAAGA
SITno1  TAAGATCGGC AATGAATAAC TCATCTAGCA TTTTAGTCTT GACCAAAAGA
   SMI  TAAGATCGGC AATTAATAAC TCATCTAGCA TTTTAGTCTT GACCAAAAGA
   SPI  TAAGATCGGC AATGAATAAC TCATCTAGCA TTTTAGTCTT GACCAAAAGA
   TOR  TAAGATCGGC AATGAATAAC TCATCTAGCA TTTTAGTCTT GACCAAAAGA
   WIL  TAAGATCGGC AATGAATAAC TCATCTAGCA TTTTAGTCTT GACCAAAAGA

        [701                                              750]
ABIno1  GGTTTCATCC GAACCCTCAG AAAATAACCT AAAGAAGAAG AACAATTCTT
   ALC  GGTTTCATCC GAACCCTCAG AAAATAACCT AAAGAAGAAG AACAATTCTT
   ASP  GGTTTCATCC GAACCCTCAG AAAATAACCT AAAGAAGAAG AACAATTCTT
   BRA  GGTTTCATCC GAAACCTCAG AAAATAACCT AAAGAAGAAG AACAATTCTT
   BRE  GGTTTCATCC GAACCCTCAG AAAATAACCT AAAGAAGAAG AACAATTCTT
   CHI  GGTTTCATCC GAACCCTCAG AAAATAACCT AAAGAAGAAG AACAATTCTT
   CRA  GGTTTCATCC GAACCCTCAG AAAATAACCT AAAGAAGAAG AACAATTCTT
ENGno3  GGTTTCATCC GAAACCTCAG AAAATAACCT AAAGAAGAAG AACAATTCTT
   FAR  GGTTTCATCC GAACCCTCAG AAAATAACCT AAAGAAGAAG AACAATTCTT
```

```
GLAno2  GGTTTCATCC GAAACCTCAG AAAATAACCT AAAGAAGAAG AACAATTCTT
   GLE  GGTTTCATCC GAACCCTCAG AAAATAACCT AAAGAAGAAG AACAATTCTT
JEZno1  GGTTTCATCC GAACCCTCAG AAAATAACCT AAAGAAGAAG AACAATTCTT
   KOR  GGTTTCATCC GAACCCTCAG AAAATAACCT AAAGAAGAAG AACAATTCTT
   KOY  GGTTTCATCC GAACCCTCAG AAAATAACCT AAAGAAGAAG AACAATTCTT
   LIK  GGTTTCATCC GAACCCTCAG AAAATAACCT AAAGAAGAAG AACAATTCTT
MARno2  GGTTTCATCC GAACCCTCAG AAAATAACCT AAAGAAGAAG AACAATTCTT
   MAT  GGTTTCATCC GAACCCTCAG AAAATAACCT AAAGAAGAAG AAAAATTCTT
   MAX  GGTTTCATCC GAACCCTCAG AAAATAACCT AAAGAAGAAG AACAATTCTT
   MEX  GGTTTCATCC GAAACCTCAG AAAATAACCT AAAGAAGAAG AACAATTCTT
   MEY  GGTTTCATCC GAACCCTCAG AAAATAACCT AAAGAAGAAG AACAATTCTT
   MON  GGTTTCATCC GAACCCTCAG AAAATAACCT AAAGAAGAAG AACAATTCTT
   MOR  GGTTTCATCC GAACCCTCAG AAAATAACCT AAAGAAGAAG AACAATTCTT
   OBO  GGTTTCATCC GAACCCTCAG AAAATAACCT AAAGAAGAAG AACAATTCTT
   OMO  GGTTTCATCC GAACCCTCAG AAAATAACCT AAAGAAGAAG AACAATTCTT
   ORI  GGTTTCATCC GAACCCTCAG AAAATAACCT AAAGAAGAAG AACAATTCTT
PINBANK  GGGTTCATCC GAACCCTCAA AAAATAACCT GGAGAAGAAG AACAATTCTT
PINKOR  GGGTTCATCC GAACCCTCAG AAAATAACCT GGAGAAGAAG AACAATTCTT
   PUN  GGTTTCATCC GAACCCTCAG AAAATAACCT AAAGAAGAAG AACAATTCTT
   PUR  GGTTTCATCC GAACCCTCAG AAAATAACCT AAAGAAGAAG AACAATTCTT
   RET  GGTTTCATCC GAACCCTCAG AAAATAACCT AAAGAAGAAG AACAATTCTT
   RUB  GGTTTCATCC GAACCCTCAG AAAATAACCT AAAGAAGAAG AACAATTCTT
   SCH  GGTTTCATCC GAACCCTCAG AAAATAACCT AAAGAAGAAG AACAATTCTT
SITno1  GGTTTCATCC GAACCCTCAT AAAATAACCT AAAGAAGAAG AACAATTCTT
   SMI  GGTTTCATCC GAACCCTCAG AAAATAACCT AAAGAAGAAG AACAATTCTT
   SPI  GGTTTCATCC GAACCCTCAG AAAATAACCT AAAGAAGAAG AACAATTCTT
   TOR  GGTTTCATCC GAACCCTCAG AAAATAACCT AAAGAAGAAG AACAATTCTT
   WIL  GGTTTCATCC GAACCCTCAG AAAATAACCT AAAGAAGAAG AACAATTCTT

        [751                                                800]
ABIno1  GGATAATTGA TGAGAACATA YCCTATATGG TTCGTTCCAA AGATGGAAAT
   ALC  GGATAATTGA TGAGAACATA CCCTATATGG TTCGTTCCAA AGATGGAAAT
   ASP  GGATAATTGA TGAGAACATA CCCTATATGG TTCGTTCCAA AGATGGAAAT
   BRA  GGATAATTGA TGAGAACATA CCCTATATGG TTCGTTCCAA AGATGGAAAT
   BRE  GGATAATTGA TGAGAACATA CCCTATATGG TTCGTTCCAA AGATGGAAAT
   CHI  GGATAATTGA TGAGAACATA CCCTATATGG TTCGTTCCAA AGATGGAAAT
   CRA  GGATAATTGA TGAGAACATA CCCTATATGG TTCGTTCCAA AGATGGAAAT
ENGno3  GGATAATTGA TGAGAACATA CCCTATATGG TTCGTTCCAA AGATGGAAAT
   FAR  GGATAATTGA TGAGAACATA CCCTATATGG TTCGTTCCAA AGATGGAAAT
GLAno2  GGATAATTGA TGAGAACATA CCCTATATGG TTCGTTCCAA AGATGGAAAT
   GLE  GGATAATTGA TGAGAACATA CCCTATAAGG TTCGTTCCAA AGATGGAAAT
JEZno1  GGATAATTGA TGAGAACATA CCCTATAAGG TTCGTTCCAA AGATGGAAAT
   KOR  GGATAATTGA TGAGAACATA CCCTATATGG TTCGTTCCAA AGATGGAAAT
   KOY  GGATAATTGA TGAGAACATA CCCTATATGG TTCGTTCCAA AGATGGAAAT
   LIK  GGATAATTGA TGAGAACATA CCCTATATGG TTCGTTCCAA AGATGGAAAT
MARno2  GGATAATTGA TGAAAACATA CCCTATAAGG TTCGTTCCAA AGATGGAAAT
   MAT  GGATAATTGA TGAGAACATA CCCTATATGG TTCGTTCCAA AGATGGAAAT
   MAX  GGATAATTGA TGAGAACATA CCCTATATGG TTCGTTCCAA AGATGGAAAT
   MEX  GGATAATTGA TGAGAACATA CCCTATATGG TTCGTTCCAA AGATGGAAAT
   MEY  GGATAATTGA TGAGAACATA CCCTATATGG TTCGTTCCAA AGATGGAAAT
   MON  GGATAATTGA TGAGAACATA CCCTATATGG TTCGTTCCAA AGATGGAAAT
   MOR  GGATAATTGA TGAGAACATA CCCTATATGG TTCGTTCCAA AGATGGAAAT
   OBO  GGATAATTGA TGAGAACATA CCCTATATGG TTCGTTCCAA AGATGGAAAT
   OMO  GGATAATTGA TGAAAACATA CCCTATAAGG TTCGTTCCAA AGATGGAAAT
   ORI  GGATAATTGA TGAGAACATA CCCTATATGG TTCGTTCCAA AGATGGAAAT
PINBANK  GGATAATTGA TGAGAACAGA CCCTATACGG TTCGGACCAA AGATGGAAAT
PINKOR  GGATAATTGA TGAGAACAGA CCCTATACGG TTCGGACCAA AGATGGAAAT
   PUN  GGATAATTGA TGAGAACATA CCCTATATGG TTCGTTCCAA AGATGGAAAT
   PUR  GGATAATTGA TGAGAACATA CCCTATATGG TTCGTTCCAA AGATGGAAAT
   RET  GGATAATTGA TGAGAACATA CCCTATATGG TTCGTTCCAA AGATGGAAAT
   RUB  GGATAATTGA TGAAAACATA CCCTATAAGG TTCGTTCCAA AGATGGAAAT
   SCH  GGATAATTGA TGAGAACATA CCCTATATGG TTCGTTCCAA AGATGGAAAT
SITno1  GGATAATTGA TGARAACATA CCCTATATGG TTCGTTCCAA AGATGGAAAT
   SMI  GGATAATTGA TGAGAACATA CCCTATATGG TTCGTTCCAA AGATGGAAAT
   SPI  GGATAATTGA TGAGAACATA CCCTATATGG TTCGTTCCAA AGATGGAAAT
   TOR  GGATAATTGA TGAGAACATA CCCTATATGG TTCGTTCCAA AGATGGAAAT
   WIL  GGATAATTGA TGAGAACATA CCCTATATGG TTCGTTCCAA AGATGGAAAT

        [801                                                850]
ABIno1  AACATTGCCG AAAAATTGGA AGATAATATC TATATTTTTT CACTAGGAGA
   ALC  AACATTGCCG AAAAATTGGA AGATAATATC TATATTTTTT CACTAGGAGA
   ASP  AACATTGCCG AAAAATTGGA AGATAATATC TATATTTTTT CACTAGGAGA
   BRA  AACATTGCCG AAAAATTGGA AGATAATATC TATATTTTTT CACTAGGAGA
   BRE  AACATTGCCG AAAAATTGGA AGATAATATC TATATTTTTT CACTAGGAGA
   CHI  AACATTGCCG AAAAATTGGA AGATAATATC TATATTTTTT CACTAGGAGA
   CRA  AACATTGCCG AAAAATTGGA AGATAATATC TATATTTTTT CACTAGGAGA
ENGno3  AACATTGCCG AAAAATTGGA AGATAATATC TATATTTTTT CACTAGGAGA
   FAR  AACATTGCCG AAAAATTGGA AGATAATATC TATATTTTTT CACTAGGAGA
GLAno2  AACATTGCCG AAAAATTGGA AGATAATATC TATATTTTTT CACTAGGAGA
   GLE  AACATTGCCG AAAAATTGGA AGATAATATC TATATTTTTT CACTAGGAGA
JEZno1  AACATTGCCG AAAAATTGGA AGATAATATC TATATTTTTT CACTAGGAGA
   KOR  AACATTGCCG AAAAATTGGA AGATAATATC TATATTTTTT CACTAGGAGA
   KOY  AACATTGCCG AAAAATTGGA AGATAATATC TATATTTTTT CACTAGGAGA
   LIK  AACATTGCCG AAAAATTGGA AGATAATATC TATATTTTTT CACTAGGAGA
```

```
MARno2  AACATTGCCG AAAAATTGGA AGATAATATC TATATTTTTT CACTAGGAGA
   MAT  AACATTGCCG AAAAATTGGA AGATAATATC TATATTTTTT CACTAGGAGA
   MAX  AACATTGCCG AAAAATTGGA AGATAATATC TATATTTTTT CACTAGGAGA
   MEX  AACATTGCCG AAAAATTGGA AGATAATATC TATATTTTTT CACTAGGAGA
   MEY  AACATTGCCG AAAAATTGGA AGATAATATC TATATTTTTT CACTAGGAGA
   MON  AACATTGCCG AAAAATTGGA AGATAATATC TATATTTTTT CACTAGGAGA
   MOR  AACATTGCCG AAAAATTGGA AGATAATATC TATATTTTTT CACTAGGAGA
   OBO  AACATTGCCG AAAATTTGGA AGATAATATC TATATTTTTT CACTAGGAGA
   OMO  AACATTGCCG AAAAATTGGA AGATAATATC TATATTTTTT CACTAGGAGA
   ORI  AACATTGCCG AAAAATTGGA AGATAATATC TATATTTTTT CACTAGGAGA
PINBANK AAAATTGCCG AAAAATTAGA AGATAATATC TACATTTTTT AACTAGGAGA
PINKOR  AACATTGCCG AAAAAGTAGA AGATAATATC TACATTTTTT CACTAGGAGA
   PUN  AACATTGCCG AAAAATTGGA AGATAATATC TATATTTTTT CACTAGGAGA
   PUR  AACATTGCCG AAAAATTGGA AGATAATATC TATATTTTTT CACTAGGAGA
   RET  AACATTGCCG AAAAATTGGA AGATAATATC TATATTTTTT CACTAGGAGA
   RUB  AACATTGCCG AAAAATTGGA AGATAATATC TATATTTTTT CACTAGGAGA
   SCH  AACATTGCCG AAAAATTGGA AGATAATATC TATATTTTTT CACTAGGAGA
SITno1  AACATTGCCG AAAAATTGGA AGATAATATC TATATTTTTT CACTAGGAGA
   SMI  AACATTGCCG AAAAATTGGA AGATAATATC TATATTTTTT CACTAGGAGA
   SPI  AACATTGCCG AAAAATTGGA AGATAATATC TATATTTTTT CACTAGGAGA
   TOR  AACATTGCCG AAAAATTGGA AGATAATATC TATATTTTTT CACTAGGAGA
   WIL  AACATTGCCG AAAAATTGGA AGATAATATC TATATTTTTT CACTAGGAGA

        [851                                              900]
ABIno1  TGAGTACCCT TTATAGCTAT AATAGATCTT TCTCCATATC TAACATAGTG
   ALC  TGAGTACCCT TTATAGCTAT AATAGATCTT TCTCCATATC TAACATAGTG
   ASP  TGAGTACCCT TTATAGCTAT AATAGATCTT TCTCCATATC TAACATAGTG
   BRA  TGAGTACCCT TTATAGCTAT AATAGATCTT TCTCCATATC TAACATAGTG
   BRE  TGAGTACCCT TTATAGCTAA AATAGATCTT TCTCCATATC TAACATAGTG
   CHI  TGAGTACCCT TTATAGCTAT AATAGATCTT TCTCCATATC TAACATAGTG
   CRA  TGAGTACCCT TTATAGCTAT AATAGATCTT TCTCCATATC TAACATAGTG
ENGno3  TTAGTACCCT TTATAGCTAT AATAGATCTT TCTCCATATC TAACATAGTG
   FAR  TGAGTACCCT TTATAGCTAT AATAGATCTT TCTCCATATC TAACATAGTG
GLAno2  TTAGTACCCT TTATAGCTAT AATAGATCTT TCTCCATATC TAACATAGTG
   GLE  TGAGTACCCT TTATAGCTAT AATAGATCTT TCTCCATATC TAACATAGTG
JEZno1  TGAGTACCCT TTATAGCTAT AATAGATCTT TCTCCATATC TAACATAGTG
   KOR  TGAGTACCCT TTATAGCTAT AATAGATCTT TCTCCATATC TAACATAGTG
   KOY  TGAGTACCCT TTATAGCTAT AATAGATCTT TCTCCATATC TAACATAGTG
   LIK  TGAGTACCCT TTATAGCTAT AATAGATCTT TCTCCATATC TAACATAGTG
MARno2  TGAGTACCCT TTATAGCTAT AATAGATCTT TCTCCATATC TAACATAGTG
   MAT  TGAGTACCCT TTATAGCTAT AATAGATCTT TCTCCATATC TAACATAGTG
   MAX  TGAGTACCCT TTATAGCTAT AATAGATCTT TCTCCATATC TAACATAGTG
   MEX  TTAGTACCCT TTATAGCTAT AATAGATCTT TCTCCATATC TAACATAGTG
   MEY  TGAGTACCCT TTATAGCTAT AATAGATCTT TCTCCATATC TAACATAGTG
   MON  TGAGTACCCT TTATAGCTAT AATAGATCTT TCTCCATATC TAACATAGTG
   MOR  TGAGTACCCT TTATAGCTAT AATAGATCTT TCTCCATATC TAACATAGTG
   OBO  TGAGTACCCT TTATAGCTAT AATAGATCTT TCTCCATATC TAACATAGTG
   OMO  TGAGTACCCT TTATAGCTAT AATAGATCTT TCTCCATATC TAACATAGTG
   ORI  TGAGTACCCT TTATAGCTAT AATAGATCTT TCTCCATATC TAACATAGTG
PINBANK TGAGCACCCT TTATAGCTAT AATAGGTCTT TCGCCATATC TAACATAGTG
PINKOR  TGAGTACCTT TTATAGCTAT AATAGATCTT TCACCATATC TAACATAGTT
   PUN  TGAGTACCCT TTATAGCTAT AATAGATCTT TCTCCATATC TAACATAGTG
   PUR  TGAGTACCCT TTATAGCTAT AATAGATCTT TCTCCATATC TAACATAGTG
   RET  TGAGTACCCT TTATAGCTAT AATAGATCTT TCTCCATATC TAACATAGTG
   RUB  TGAGTACCCT TTATAGCTAT AATAGATCTT TCTCCATATC TAACATAGTG
   SCH  TGAGTACCCT TTATAGCTAT AATAGATCTT TCTCCATATC TAACATAGTG
SITno1  TGAGTACCCT TTATAGCTAT AATAGATCTT TCTCCATATC TAACATAGTG
   SMI  TGAGTACCCT TTATAGCTAT AATAGATCTT TCTCCATATC TAACATAGTG
   SPI  TGAGTACCCT TTATAGCTAT AATAGATCTT TCTCCATATC TAACATAGTG
   TOR  TGAGTACCCT TTATAGCTAT AATAGATCTT TCTCCATATC TAACATAGTG
   WIL  TGAGTACCCT TTATAGCTAT AATAGATCTT TCTCCATATC TAACATAGTG

        [901                                              950]
ABIno1  AATGTTAGGA TCCTTCAACG ACCAGATACT TTGAAATGAA TTTCGACGAG
   ALC  AATGTTAGGA TCCTTCAACG ACCAGATACT TTGAAATGAA TTTCGACGAG
   ASP  AATGTTAGGA TCCTTCAACG ACCAGATACT TTGAAATGAA TTTCGACGAG
   BRA  AATGTTAGGA TCCTTCAACG ACCAGATACT TTGAAATGAA TTTCGACGAG
   BRE  AATGTTAGGA TCCTTCAACG ACCAGATACT TTGAAGTGAA TTTCGACGAG
   CHI  AATGTTAGGA TCCTTCAACG ACCAGATACT TTGAAATGAA TTTCGACGAG
   CRA  AATGTTAGGA TCCTTCAACG ACCAGATACT TTGAAATGAA TTTCGACGAG
ENGno3  AATGTTAGGA TCCTTCAACG ACCAGATACT TTGAAATGAA TTTCGACGAG
   FAR  AATGTTAGGA TCCTTCAACG ACCAGATACT TTGAAATGAA TTTCGACGAG
GLAno2  AATGTTAGGA TCCTTCAACG ACCAGATACT TTGAAATGAA TTTCGACGAG
   GLE  AATGTTAGGA TCCTTCAACG ACCAGATACT TTGAAATGAA TTTCGACGAG
JEZno1  AATGTTAGGA TCCTTCAACG ACCAGATACT TTGAAATGAA TTTCGACGAG
   KOR  AATGTTAGGA TCCTTCAACG ACCAGATACT TTGAAATGAA TTTCGACGAG
   KOY  AATGTTAGGA TCCTTCAACG ACCAGATACT TTGAAATGAA TTTCGACGAG
   LIK  AATGTTAGGA TCCTTCAACG ACCAGATACT TTGAAATGAA TTTCGACGAG
MARno2  AATGTTAGGA TCCTTCAACG ACCAGATACT TTGAAATGAA TTTCGACGAG
   MAT  AATGTTAGGA TCCTTCAACG ACCAGATACT TTGAAATGAA TTTCGACGAG
   MAX  AATGTTAGGA TCCTTCAACG ACCAGATACT TTGAAATGAA TTTCGACGAG
   MEX  AATGTTAGGA TCCTTCAACG ACCAGATACT TTGAAATGAA TTTCGACGAG
   MEY  AATGTTAGGA TCCTTCAACG ACCAGATACT TTGAAATGAA TTTCGACGAG
   MON  AATGTTAGGA TCCTTCAACG ACCAGATACT TTGAAATGAA TTTCGACGAG
```

```
   MOR  AATGTTAGGA TCCTTCAACG ACCAGATACT TTGAAATGAA TTTCGACGAG
   OBO  AATGTTAGGA TCCTTCAACG ACCAGATACT TTGAAATGAA TTTCGACGAG
   OMO  AATGTTAGGA TCCTTCAACG ACCAGATACT TTGAAATGAA TTTCGACGAG
   ORI  AATGTTAGGA TCCTTCAACG ACCAGATACT TTGAAATGAA TTTCGACGAG
PINBANK AATTTTAGGA TCCTTCAACG ACCAGATACT TTTCAGTGAA TTTCGACGAG
PINKOR AATTTTAGGA TCCTTCAACA ACCAGATACT TTTCAGTGAA TTTCGACGAG
   PUN  AATGTTAGGA TCCTTCAACG ACCAGATACT TTGAAATGAA TTTCGACGAG
   PUR  AATGTTAGGA TCCTTCAACG ACCAGATACT TTGAAATGAA TTTCGACGAG
   RET  AATGTTAGGA TCCTTCAACG ACCAGATACT TTGAAATGAA TTTCGACGAG
   RUB  AATGTTAGGA TCCTTCAACG ACCAGATACT TTGAAATGAA TTTCGACGAG
   SCH  AATGTTAGGA TCCTTCAACG ACCAGATACT TTGAAATGAA TTTCGACGAG
SITno1 AATGTTAGGA TCCTTCAACG ACCAGATACT TTGAAGTGAA TTTCGACGAG
   SMI  AATGTTAGGA TCCTTCAACG ACCAGATACT TTGAAATGAA TTTCGACGAG
   SPI  AATGTTAGGA TCCTTCAACG ACCAGATACT TTGAAATGAA TTTCGACGAG
   TOR  AATGTTAGGA TCCTTCAACG ACCAGATACT TTGAAATGAA TTTCGACGAG
   WIL  AATGTTAGGA TCCTTCAACG ACCAGATACT TTGAAATGAA TTTCGACGAG

        [951                                                1000]
ABIno1 AAAATAGAAA AATATTTTTT ATTTTTCGAT GAAAATGAGT TCGCTGAGGG
   ALC  AAAATAGAAA AATATTTTTT ATTTTTCGAT GAAAATGAGT TCGCTGAGGG
   ASP  AAAATAGAAA AATATTTTTT ATTTTTCGAT GAAAATGAGT TCGCTGAGGG
   BRA  AAAATAGAAA AATATTTTTT ATTTTTCGAT GAAAATGAGT TCGCTGAGGG
   BRE  AAAATAGAAA AATATTTTTT ATTTTTCGAT GAAAATGAGT TCGCTGAGGG
   CHI  AAAATAGAAA AATATTTTTT ATTTTTCGAT GAAAATGAGT TCGCTGAGGG
   CRA  AAAATAGAAA AATATTTTTT ATTTTTCGAT GAAAATGAGT TCGCTGAGGG
ENGno3 AAAATAGAAA AATATTTTTT ATTTTTCGAT GAAAATGAGT TCGCTGAGGG
   FAR  AAAATAGAAA AATATTTTTT ATTTTTCGAT GAAAATGAGT TCGCTGAGGG
GLAno2 AAAATAGAAA AATATTTTTT ATTTTTCGAT GAAAATGAGT TCGCTGAGGG
   GLE  AAAATAGAAA AATATTTTTT ATTTTTCGAT GAAAATGAGT TCGCTGAGGG
JEZno1 AAAATAGAAA AATATTTTTT ATTTTTCGAT GAAAATGAGT TCGCTGAGGG
   KOR  AAAATAGAAA AATATTTTTT ATTTTTCGAT GAAAATGAGT TCGCTGAGGG
   KOY  AAAATAGAAA AATATTTTTT ATTTTTCGAT GAAAATGAGT TCGCTGAGGG
   LIK  AAAATAGAAA AATATTTTTT ATTTTTCGAT GAAAATGAGT TCGCTGAGGG
MARno2 AAAATAGAAA AATATTTTTT ATTTTTCGAT GAAAATGAGT TCGCTGAGGG
   MAT  AAAATAGAAA AATATTTTTT ATTTTTCGAT GAAAATGAGT TCGCTGAGGG
   MAX  AAAATAGAAA AATATTTTTT ATTTTTCGAT GAAAATGAGT TCGCTGAGGG
   MEX  AAAATAGAAA AATATTTTTT ATTTTTCGAT GAAAATGAGT TCGCTGAGGG
   MEY  AAAATAGAAA AATATTTTTT ATTTTTCGAT GAAAATGAGT TCGCTGAGGG
   MON  AAAATAGAAA AATATTTTTT ATTTTTCGAT GAAAATGAGT TCGCTGAGGG
   MOR  AAAATAGAAA AATATTTTTT ATTTTTCGAT GAAAATGAGT TCGCTGAGGG
   OBO  AAAATAGAAA AATATTTTTT ATTTTTCGAT GAAAATGAGT TCGCTGAGGG
   OMO  AAAATAGAAA AATATTTTTT ATTTTTCGAT GAAAATGAGT TCGCTGAGGG
   ORI  AAAATAGAAA AATATTTTTT ATTTTTCGAT GAAAATGAGT TCGCTGAGGG
PINBANK AAAATATGAT AATATGTTTT ATCTTTCGAT GAAAATGAGT TCGCTGAGGG
PINKOR AAAATATGAT AATATGTTTT ATCTTTCGAT GAAAATGAGT TCGCTGAGGG
   PUN  AAAATAGAAA AATATTTTTT ATTTTTCGAT GAAAATGAGT TCGCTGAGGG
   PUR  AAAATAGAAA AATATTTTTT ATTTTTCGAT GAAAATGAGT TCGCTGAGGG
   RET  AAAATAGAAA AATATTTTTT ATTTTTCGAT GAAAATGAGT TCGCTGAGGG
   RUB  AAAATAGAAA AATATTTTTT ATTTTTCGAT GAAAATGAGT TCGCTGAGGG
   SCH  AAAATAGAAA AATATTTTTT ATTTTTCGAT GAAAATGAGT TCGCTGAGGG
SITno1 AAAATAGAAA AATATTTTTT ATTTTTCGAT GAAAATGAGT TCGCTGAGGG
   SMI  AAAATAGAAA AATATTTTTC ATTTTTCGAT GAAAATGAGT TCGCTGAGGG
   SPI  AAAATAGAAA AATATTTTTT ATTTTTCGAT GAAAATGAGT TCGCTGAGGG
   TOR  AAAATAGAAA AATATTTTTT ATTTTTCGAT GAAAATGAGT TCGCTGAGGG
   WIL  AAAATAGAAA AATATTTTTT ATTTTTCGAT GAAAATGAGT TCGCTGAGGG

        [1001                                               1050]
ABIno1 AAAGACCCAT GAGATAACGA TCGTGAATGA GAGGATCGTT TAAGAAGGGA
   ALC  AAAGACCCAT GAGATAACGA TCGTGAATGA GAGGATCGTT TAAGAAGGGA
   ASP  AAAGAGCCAT GAGATAACGA TCGTGAATGA GAGGATCGTT TAAGAAGGGA
   BRA  AAAGACCCAT GAGATAACGA TCGTGAATGA GAGGATCGTT TAAGAAGGGA
   BRE  AAAGACCCAT GAGATAACGA TCGTGAATGA GAGGATCGTT TAAGAAGGGA
   CHI  AAAGACCCAT GAGATAACGA TCGTGAATGA GAGGATCGTT TAAGAAGGGA
   CRA  AAAGAGCCAT GAGATAACGA TCGTGAATGA GAGGATCGTT TAAGAAGGGA
ENGno3 AAAGACCCAT GAGATAACGA TCGTGAATGA GAGGATCGTT TAAGAAGGGA
   FAR  AAAGACCCAT GAGATAACGA TCGTGAATGA GAGGATCGTT TAAGAAGGGA
GLAno2 AAAGACCCAT GAGATAACGA TCGTGAATGA GAGGATCGTT TAAGAAGGGA
   GLE  AAAGACCCAT GAGATAACGA TCGTGAATGA GAGGATCGTT TAAGAAGGGA
JEZno1 AAAGACCCAT GAGATAACGA TCGTGAATGA GAGGATCGTT TAAGAAGGGA
   KOR  AAAGAGCCAT GAGATAACGA TCGTGAATGA GAGGATCGTT TAAGAAGGGA
   KOY  AAAGACCCAT GAGATAACGA TCGTGAATGA GAGGATCGTT TAAGAAGGGA
   LIK  AAAGACCCAT GAGATAACGA TCGTGAATGA GAGGATCGTT TAAGAAGGGA
MARno2 AAAGACCCAT GAGATAACGA TCGTGAATGA GAGGATCGTT TAAGAAGGGA
   MAT  AAAGACCCAT GAGATAACGA TCGTGAATGA GAGGATCGTT TAAGAAGGGA
   MAX  AAAGACCCAT GAGATAACGA TCGTGAATGA GAGGATCGTT TAAGAAGGGA
   MEX  AAAGACCCAT GAGATAACGA TCGTGAATGA GAGGATCGTT TAAGAAGGGA
   MEY  AAAGAGCCAT GAGATAACGA TCGTGAATGA GAGGATCGTT TAAGAAGGGA
   MON  AAAGAGCCAT GAGATAACGA TCGTGAATGA GAGGATCGTT TAAGAAGGGA
   MOR  AAAGACCCAT GAGATAACGA TCGTGAATGA GAGGATCGTT TAAGAAGGGA
   OBO  AAAGACCCAT GAGATAACGA TCGTGAATGA GAGGATCGTT TAAGAAGGGA
   OMO  AAAGACCCAT GAGATAACGA TCGTGAATGA GAGGATCGTT TAAGAAGGGA
   ORI  AAAGACCCAT GAGATAACGA TCGTGAATGA GAGGATCGTT TAAGAAGGGA
PINBANK AAAGATCCAT GAGACAACGA TCGTGAATGA GAGGATCGTT TAAGACGGGA
PINKOR AAAGGTCTAT GAGACAACGA TCGTGAATGA GAGGATCGTT TAAGAAGGGA
```

```
   PUN  AAAGACCCAT GAGATAACGA TCGTGAATGA GAGGATCGTT TAAGAAGGGA
   PUR  AAAGACCCAT GAGATAACGA TCGTGAATGA GAGGATCGTT TAAGAAGGGA
   RET  AAAGAGCCAT GAGATAACGA TCGTGAATGA GAGGATCGTT TAAGAAGGGA
   RUB  AAAGACCCAT GAGATAACGA TCGTGAATGA GAGGATCGTT TAAGAAGGGA
   SCH  AAAGACCCAT GAGATAACGA TCGTGAATGA GAGGATCGTT TAAGAAGGGA
SITno1  AAAGACCCAT GAGATAACGA TCGTGAATGA GAGGATCGTT TAAGAAGGGA
   SMI  AAAGACCCAT GAGATAACGA TCGTGAATGA GAGGATCGTT TAAGAAGGGA
   SPI  AAAGACCCAT GAGATAACGA TCGTGAATGA GAGGATCGTT TAAGAAGGGA
   TOR  AAAGAGCCAT GAGATAACGA TCGTGAATGA GAGGATCGTT TAAGAAGGGA
   WIL  AAAGACCCAT GAGATAACGA TCGTGAATGA GAGGATCGTT TAAGAAGGGA

        [1051                                              1100]
ABIno1  AACTAAAATG GATTCGCATT CATAGACATA ATTATTCCAC AGGAACAGGA
   ALC  AACTAAAATG GATTCGCATT CATAGACATA ATTATTCCAC AGGAACAGGA
   ASP  AACTAAAATG GATTCGCATT CATAGACATA ATTATTCCAC AGGAACAGGA
   BRA  AACTAAAATG GATTCGCATT CATAGACATA ATTATTCCAC AGGAACAGGA
   BRE  AACTAAAATG GATTCGCATT CATAGACATA ATTATTCCAC AGGAACAGGA
   CHI  AACTAAAATG GATTCGCATT CATAGACATA ATTATTCCAC AGGAACAGGA
   CRA  AACTAAAATG GATTCGCATT CATAGACATA ATTATTCCAC AGGAACAGGA
ENGno3  AACTAAAATG GATTCGCATT CATAGACATA ATTATTCCAC AGGAACAGGA
   FAR  AACTAAAATG GATTCGCATT CATAGACATA ATTATTCCAC AGGAACAGGA
GLAno2  AACTAAAATG GATTCGCATT CATAGACATA ATTATTCCAC AGGAACAGGA
   GLE  AACTAAAAAT GATTCGCATT CATAGACATA ATTATTCCAC AGGAACAGGA
JEZno1  AACTAAAAAT GATTCGCATT CATAGACATA ATTATTCCAC AGGAACAGGA
   KOR  AACTAAAATG GATTCGCATT CATAGACATA ATTATTCCAC AGGAACAGGA
   KOY  AACTAAAATG GATTCGCATT CATAGACATA ATTATTCCAC AGGAACAGGA
   LIK  AACTAAAATG GATTCGCATT CATAGACATA ATTATTCCAC AGGAACAGGA
MARno2  AACTAAAATG GATTCGCATT CATAGACATA ATTATTCCAC AGGAACAGGA
   MAT  AACTAAAATG GATTCGCATT CATAGACATA ATTATTCCAC AGGAACAGGA
   MAX  AACTAAAATG GATTCGCATT CATAGACATA ATTATTCCAC AGGAACAGGA
   MEX  AACTAAAATG GATTCGCATT CATAGACATA ATTATTCCAC AGGAACAGGA
   MEY  AACTAAAATG GATTCGCATT CATAGACATA ATTATTCCAC AGGAACAGGA
   MON  AACTAAAATG GATTCGCATT CATAGACATA ATTATTCCAC AGGAACAGGA
   MOR  AACTAAAATG GATTCGCATT CATAGACATA ATTATTCCAC AGGAACAGGA
   OBO  AACTAAAATG GATTCGCATT CATAGACATA ATTATTCCAC AGGAACAGGA
   OMO  AACTAAAATG GATTCGCATT CATAGACATA ATTATTCCAC AGGAACAGGA
   ORI  AACTAAAATG GATTCGCATT CATAGACATA ATTATTCCAC AGGAACAGGA
PINBANK  AAATAAAATG GATTCGCATT CACAGACATA ATAATTCCAC AGGAACAGGA
PINKOR  AAATAAAATG GATTCGCATT CATAGACATA ATAATTCCAC AGGAACAGGA
   PUN  AACTAAAATG GATTCGCATT CATAGACATA ATTATTCCAC AGGAACAGGA
   PUR  AACTAAAATG GATTCGCATT CATAGACATA ATTATTCCAC AGGAACAGGA
   RET  AACTAAAATG GATTCGCATT CATAGACATA ATTATTCCAC AGGAACAGGA
   RUB  AACTAAAATG GATTCGCATT CATAGACATA ATTATTCCAC AGGAACAGGA
   SCH  AACTAAAATG GATTCGCATT CATAGACATA ATTATTCCAC AGGAACAGGA
SITno1  AACTAAAATG GATTCGCATT CATAGACATA ATTATTCCAC AGGAACAGGA
   SMI  AACTAAAATG GATTCGCATT CATAGACATA ATTATTCCAC AGGAACAGGA
   SPI  AACTAAAATG GATTCGCATT CATAGACATA ATTATTCCAC AGGAACAGGA
   TOR  AACTAAAATG GATTCGCATT CATAGACATA ATTATTCCAC AGGAACAGGA
   WIL  AACTAAAATG GATTCGCATT CATAGACATA ATTATTCCAC AGGAACAGGA

        [1101                                              1150]
ABIno1  AGAATCTCGT ATTTACTTTC GGGACGACAA TAATCGATCT TTGTAAACTC
   ALC  AGAATCTCGT ATTTACTTTC GGGACGACAA TAATCGATCT TTGTAAACTC
   ASP  AGAATCTCGT ATTTACTTTC GGGACGACAA TAATCGATCT TTGTAAACTC
   BRA  AGAATCTCGT ATTTACTTTC GGGACGACAA TAATCGATCT TTGTAAACTC
   BRE  AGAATCTCGT ATTTACTTTC GGGACAACAA TAATCGATCT TTGTAAACTC
   CHI  AGAATCTCGT ATTTACTTTC GGGACGACAA TAATCGATCT TTGTAAACTC
   CRA  AGAATCTCGT ATTTACTTTC GGGACGACAA TAATCGATCT TTGTAAACTC
ENGno3  AGAATCTCGT ATTTACTTTC GGGACGACAA TAATCGATCT TTGTAAACTC
   FAR  AGAATCTCGT ATTTACTTTC GGGACGACAA TAATCGATCT TTGTAAACTC
GLAno2  AGAATCTCGT ATTTACTTTC GGGACGACAA TAATCGATCT TTGTAAACTC
   GLE  AGAATCTCGT ATTTACTTTC GGGACGACAA TAATCGATCT TTGTAAACTC
JEZno1  AGAATCTCGT ATTTACTTTC GGGACGACAA TAATCGATCT TTGTAAACTC
   KOR  AGAATCTCGT ATTTACTTTC GGGACGACAA TAATCGATCT TTGTAAACTC
   KOY  AGAATCTCGT ATTTACTTTC GGGACGACAA TAATCGATCT TTGTAAACTC
   LIK  AGAATCTCGT ATTTACTTTC GGGACGACAA TAATCGATCT TTGTAAACTC
MARno2  AGAATCTCGT ATTTACTTTC GGGACGACAA TAATCGATCT TTGTAAACTC
   MAT  AGAATCTCGT ATTTACTTTC GGGACGACAA TAATCGATCT TTGTAAACTC
   MAX  AGAATCTCGT ATTTACTTTC GGGACGACAA TAATCGATCT TTGTAAACTC
   MEX  AGAATCTCGT ATTTACTTTC GGGACGACAA TAATCGATCT TTGTAAACTC
   MEY  AGAATCTCGT ATTTACTTTC GGGACGACAA TAATCGATCT TTGTAAACTC
   MON  AGAATCTCGT ATTTACTTTC GGGACGACAA TAATCGATCT TTGTAAACTC
   MOR  AGAATCTCGT ATTTACTTTC GGGACGACAA TAATCGATCT TTGTAAACTC
   OBO  AGAATCTCGT ATTTACTTTC GGGACGACAA TAATCGATCT TTGTAAACTC
   OMO  AGAATCTCGT ATTTACTTTC GGGACGACAA TAATCGATCT TTGTAAACTC
   ORI  AGAATCTCGT ATTTACTTTC GGGACGACAA TAATCGATCT TTGTAAACTC
PINBANK  AGAATCTCGT ATTTACTTTC GGGACGACAA TAATCGATCT TTGTAAACTC
PINKOR  AGAATCTCGT ATTTACTCTT GGGACGACAA TAATTGATCT TTGTAAATTC
   PUN  AGAATCTCGT ATTTACTTTC GGGACGACAA TAATCGATCT TTGTAAACTC
   PUR  AGAATCTCGT ATTTACTTTC GGGACGACAA TAATCGATCT TTGTAAACTC
   RET  AGAATCTCGT ATTTACTTTC GGGACGACAA TAATCGATCT TTGTAAACTC
   RUB  AGAATCTCGT ATTTACTTTC GGGACGACAA TAATCGATCT TTGTAAACTC
   SCH  AGAATCTCGT ATTTACTTTC GGGACGACAA TAATCGATCT TTGTAAACTC
SITno1  AGAATCTCGT ATTTACTTTC GGGACGACAA TAATCGATCT TTGTAAACTC
```

```
   SMI  AGAATCTCGT ATTTACTTTC GGGACGACAA TAATCGATCT TTGTAAACTC
   SPI  AGAATCTCGT ATTTACTTTC GGGACGACAA TAATCGATCT TTGTAAACTC
   TOR  AGAATCTCGT ATTTACTTTC GGGACGACAA TAATCGATCT TTGTAAACTC
   WIL  AGAATCTCGT ATTTACTTTC GGGACGACAA TAATCGATCT TTGTAAACTC

        [1151      1 site enlevé              +1        1200]
ABIno1  TCTGAACTAT TTCGATATTC ..........  ........0AT AGAGAATAG
   ALC  TCTGAACTAT TTCGATTCTC TGAACTATTT CGATATTC1AT AGAGAATAG
   ASP  TCTGAACTAT TTCGATATTC ..........  ........0AT AGAGAATAG
   BRA  TCTGAACTAT TTCGATATTC ..........  ........0AT AGAGAATAG
   BRE  TCTGAACTAT TTCGATATTC ..........  ........0AT AGAGAATAG
   CHI  TCTGAACTAT TTCGATATTC ..........  ........0AT AGAGAATAG
   CRA  TCTGAACTAT TTCGATATTC ..........  ........0AT AGAGAATAG
ENGno3  TCTGAACTAT TTCGATATTC ..........  ........0AT AGAGAATAG
   FAR  TCTGAACTAT TTCGATATTC ..........  ........0AT AGAGAATAG
GLAno2  TCTGAACTAT TTCGATATTC ..........  ........0AT AGAGAATAG
   GLE  TCTGAACTAT TTCGATATTC ..........  ........0AT AGAGAATAG
JEZno1  TCTGAACTAT TTCGATATTC ..........  ........0AT AGAGAATAG
   KOR  TCTGAACTAT TTCGATATTC ..........  ........0AT AGAGAATAG
   KOY  TCTGAACTAT TTCGATATTC ..........  ........0AT AGAGAATAG
   LIK  TCTGAACTAT TTCGATATTC ..........  ........0AT AGAGAATAG
MARno2  TCTGAACTAT TTCGATATTC ..........  ........0AT AGAGAATAG
   MAT  TCTGAACTAT TTCGATATTC ..........  ........0AT AGAGAATAG
   MAX  TCTGAACTAT TTCGATATTC ..........  ........0AT AGAGAATAG
   MEX  TCTGAACTAT TTCGATATTC ..........  ........0AT AGAGAATAG
   MEY  TCTGAACTAT TTCGATATTC ..........  ........0AT AGAGAATAG
   MON  TCTGAACTAT TTCGATATTC ..........  ........0AT AGAGAATAG
   MOR  TCTGAACTAT TTCGATATTC ..........  ........0AT AGAGAATAG
   OBO  TCTGAACTAT TTCGATATTC ..........  ........0AT AGAGAATAG
   OMO  TCTGAACTAT TTCGATATTC ..........  ........0AT AGAGAATAG
   ORI  TCTGAACTAT TTCGATATTC ..........  ........0AT AGAGAATAG
PINBANK  TCTGAACTAT TTCGATATTC ..........  ........0AT AGAGAATAG
PINKOR  TCTGGACTAT TTCGATATTC ..........  ........0AT AGAGAACAG
   PUN  TCTGAACTAT TTCGATATTC ..........  ........0AT AGAGAATAG
   PUR  TCTGAACTAT TTCGATATTC ..........  ........0AT AGAGAATAG
   RET  TCTGAACTAT TTCGATATTC ..........  ........0AT AGAGAATAG
   RUB  TCTGAACTAT TTCGATATTC ..........  ........0AT AGAGAATAG
   SCH  TCTGAACTAT TTCGATATTC ..........  ........0AT AGAGAATAG
SITno1  TCTGAACTAT TTCGATATTC ..........  ........0AT AGAGAATAG
   SMI  TCTGAACTAT TTCGATATTC ..........  ........0AT AGAGAATAG
   SPI  TCTGAACTAT TTCGATATTC ..........  ........0AT AGAGAATAG
   TOR  TCTGAACTAT TTCGATATTC ..........  ........0AT AGAGAATAG
   WIL  TCTGAACTAT TTCGATATTC ..........  ........0AT AGAGAATAG

        [1201                                              1250]
ABIno1  ATCGTAATGG GTGCAAGGAG GGAGCATCTC CGATCCAGCG ACGAAAGGTT
   ALC  ATCGTAATGG GTGCAAGGAG GGAGCATCTC CGATCCAGCG ACGAAAGGTT
   ASP  ATCGTAATGG GTGCAAGGAG GGAGCATCTC CGATCCAGCG ACGAAAGGTT
   BRA  ATCGTAATGG GTGCAAGGAG GGAGCATCTC CGATCCAGCG ACGAAAGGTT
   BRE  ATCGTAATGG GTGCAAGGAG GGAGCATCTC CGATCCAGCG ACGAAAGGTT
   CHI  ATCGTAATGG GTGCAAGGAG GGAGCATCTC CGATCCAGCG ACGAAAGGTT
   CRA  ATCGTAATGG GTGCAAGGAG GGAGCATCTC CGATCCAGCG ACGAAAGGTT
ENGno3  ATCGTAATGG GTGCAAGGAG GGAGCATCTC CGATCCAGCG ACGAAAGGTT
   FAR  ATCGTAATGG GTGCAAGGAG GGAGCATCTC CGATCCAGCG ACGAAAGGTT
GLAno2  ATCGTAATGG GTGCAAGGAG GGAGCATCTC CGATCCAGCG ACGAAAGGTT
   GLE  ATCGTAATGG GTGCAAGGAG GGAGCATCTC CGATCCAGCG ACGAAAGGTT
JEZno1  ATCGTAATGG GTGCAAGGAG GGAGCATCTC CGATCCAGCG ACGAAAGGTT
   KOR  ATCGTAATGG GTGCAAGGAG GGAGCATCTC CGATCCAGCG ACGAAAGGTT
   KOY  ATCGTAATGG GTGCAAGGAG GGAGCATCTC CGATCCAGCG ACGAAAGGTT
   LIK  ATCGTAATGG GTGCAAGGAG GGAGCATCTC CGATCCAGCG ACGAAAGGTT
MARno2  ATCGTAATGG GTGCAAGGAG GGAGCATCTC CGATCCAGCG ACGAAAGGTT
   MAT  ATCGTAATGG GTGCAAGGAG GGAGCATCTC CGATCCAGCG ACGAAAGGTT
   MAX  ATCGTAATGG GTGCAAGGAG GGAGCATCTC CGATCCAGCG ACGAAAGGTT
   MEX  ATCGTAATGG GTGCAAGGAG GGAGCATCTC CGATCCAGCG ACGAAAGGTT
   MEY  ATCGTAATGG GTGCAAGGAG GGAGCATCTC CGATCCAGCG ACGAAAGGTT
   MON  ATCGTAATGG GTGCAAGGAG GGAGCATCTC CGATCCAGCG ACGAAAGGTT
   MOR  ATCGTAATGG GTGCAAGGAG GGAGCATCTC CGATCCAGCG ACGAAAGGTT
   OBO  ATCGTAATGG GTGCAAGGAG GGAGCATCTC CGATCCAGCG ACGAAAGGTT
   OMO  ATCGTAATGG GTGCAAGGAG GGAGCATCTC CGATCCAGCG ACGAAAGGTT
   ORI  ATCGTAATGG GTGCAAGGAG GGAGCATCTC CGATCCAGCG ACGAAAGGTT
PINBANK  ATCGTAATGG GTGCAAGGAG GGAGCATCTC CGATCCAGCG ACGAAAGGTT
PINKOR  AGCGTAATGG GTGCAAGGAG GGAGCATCTC GGATCAAGCG ACGAAAGGTT
   PUN  ATCGTAATGG GTGCAAGGAG GGAGCATCTC CGATCCAGCG ACGAAAGGTT
   PUR  ATCGTAATGG GTGCAAGGAG GGAGCATCTC CGATCCAGCG ACGAAAGGTT
   RET  ATCGTAATGG GTGCAAGGAG GGAGCATCTC CGATCCAGCG ACGAAAGGTT
   RUB  ATCGTAATGG GTGCAAGGAG GGAGCATCTC CGATCCAGCG ACGAAAGGTT
   SCH  ATCGTAATGG GTGCAAGGAG GGAGCATCTC CGATCCAGCG ACGAAAGGTT
SITno1  ATCGTAATGG GTGCAAGGAG GGAGCATCTC CGATCCAGCG ACGAAAGGTT
   SMI  ATCGTAATGG GTGCAAGGAG GGAGCATCTC CGATCCAGCG ACGAAAGGTT
   SPI  ATCGTAATGG GTGCAAGGAG GGAGCATCTC CGATCCAGCG ACGAAAGGTT
   TOR  ATCGTAATGG GTGCAAGGAG GGAGCATCTC CGATCCAGCG ACGAAAGGTT
   WIL  ATCGTAATGG GTGCAAGGAG GGAGCATCTC CGATCCAGCG ACGAAAGGTT

        [1251                                              1300]
```

```
ABIno1  CGAACCAAAA TTTCTGGATG AATAGAATAG GGTATTCGTG TATCTGATAC
   ALC  CGAACCAAAA TTTCTGGATG AATAGAATAG GGTATTCGTG TATCTGATAC
   ASP  CGAACCAAAA TTTCTGGATG AATAGAATAG GGTATTCGTG TATCTGATAC
   BRA  CGAACCAAAA TTTCTGGATG AATAGAATAG GGTATTCGTG TATCTGATAC
   BRE  CGAACCAAAA TTTCCGGATG AATAGAATAG GGTATTCGTG TATCTGATAC
   CHI  CGAACCAAAA TTTCTGGATG AATAGAATAG GGTATTCGTG TATCTGATAC
   CRA  CGAACCAAAA TTTCTGGATG AATAGAATAG GGTATTCGTG TATCTGATAC
ENGno3  CGAACCAAAA TTTCTGGATG AATAGAATAG GGTATTCGTG TATCTGATAC
   FAR  CGAACCAAAA TTTCTGGATG AATAGAATAG GGTATTCGTG TATCTGATAC
GLAno2  CGAACCAAAA TTTCTGGATG AATAGAATAG GGTATTCGTG TATCTGATAC
   GLE  CGAACCAAAA TTTCTGGATG AATAGAATAG GGTATTCGTG TATCTGATAC
JEZno1  CGAACCAAAA TTTCTGGATG AATAGAATAG GGTATTCGTG TATCTGATAC
   KOR  CGAACCAAAA TTTCTGGATG AATAGAATAG GGTATTCGTG TATCTGATAC
   KOY  CGAACCAAAA TTTCTGGATG AATAGAATAG GGTATTCGTG TATCTGATAC
   LIK  CGAACCAAAA TTTCTGGATG AATAGAATAG GGTATTCGTG TATCTGATAC
MARno2  CGAACCAAAA TTTCTGGATG AATAGAATAG GGTATTCGTG TATCTGATAC
   MAT  CGAACCAAAA TTTCTGGATG AATAGAATAG GGTATTCGTG TATCTGATAC
   MAX  CGAACCAAAA TTTCTGGATG AATAGAATAG GGTATTCGTG TATCTGATAC
   MEX  CGAACCAAAA TTTCTGGATG AATAGAATAG GGTATTCGTG TATCTGATAC
   MEY  CGAACCAAAA TTTCTGGATG AATAGAATAG GGTATTCGTG TATCTGATAC
   MON  CGAACCAAAA TTTCTGGATG AATAGAATAG GGTATTCGTG TATCTGATAC
   MOR  CGAACCAAAA TTTCTGGATG AATAGAATAG GGTATTCGTG TATCTGATAC
   OBO  CGAACCAAAA TTTCTGGATG AATAGAATAG GGTATTCGTG TATCTGATAC
   OMO  CGAACCAAAA TTTCTGGATG AATAGAATAG GGTATTCGTG TATCTGATAC
   ORI  CGAACCAAAA TTTCTGGATG AATAGAATAG GGTATTCGTG TATCTGATAC
PINBANK  CGAACCAAAA TTTCTGGATG AATAGAATAG GGTATTCGTG TATCTGATAC
PINKOR  CGAACCAAAA TTTCCGGATG AATAGAATAG GGTATTCGTG CATCTGATAT
   PUN  CGAACCAAAA TTTCTGGATG AATAGAATAG GGTATTCGTG TATCTGATAC
   PUR  CGAACCAAAA TTTCTGGATG AATAGAATAG GGTATTCGTG TATCTGATAC
   RET  CGAACCAAAA TTTCTGGATG AATAGAATAG GGTATTCGTG TATCTGATAC
   RUB  CGAACCAAAA TTTCTGGATG AATAGAATAG GGTATTCGTG TATCTGATAC
   SCH  CGAACCAAAA TTTCTGGATG AATAGAATAG GGTATTCGTG TATCTGATAC
SITno1  CGAACCAAAA TTTCCGGATG AATAGAATAG GGTATTCGTG TATCTGATAC
   SMI  CGAACCAAAA TTTCTGGATG AATAGAATAG GGTATTCGTG TATCTGATAC
   SPI  CGAACCAAAA TTTCTGGATG AATAGAATAG GGTATTCGTG TATCTGATAC
   TOR  CGAACCAAAA TTTCTGGATG AATAGAATAG GGTATTCGTG TATCTGATAC
   WIL  CGAACCAAAA TTTCTGGATG AATAGAATAG GGTATTCGTG TATCTGATAC

        [1301                                              1350]
ABIno1  ATAATTTGAA TGCGGGAATT TATCCTCCAA GAAGGGAAAT ATTGAATGGA
   ALC  ATAATTTGAA TGCGGGAATT TATCCTCCAA GAAGGGAAAT ATTGAATGGA
   ASP  ATAATTTGAA TGCGGGAATT TATCCTCCAA GAAGGGAAAT ATTGAATGGA
   BRA  ATAATTTGAA TGCGGGAATT TATCCTCCAA GAAGGGAAAT ATTGAATGGA
   BRE  ATAATTTGAA TGCGGGAATT TATCCTCCAA GAAGGGAAAT ATTGAATGGA
   CHI  ATAATTTGAA TGCGGGAATT TATCCTCCAA GAAGGGAAAT ATTGAATGGA
   CRA  ATAATTTGAA TGCGGGAATT TATCCTCCAA GAAGGGAAAT ATTGAATGGA
ENGno3  ATAATTTGAA TGCGGGAATT TATCCTCCAA GAAGGGAAAT ATTGAATGGA
   FAR  ATAATTTGAA TGCGGGAATT TATCCTCCAA GAAGGGAAAT ATTGAATGGA
GLAno2  ATAATTTGAA TGCGGGAATT TATCCTCCAA GAAGGGAAAT ATTGAATGGA
   GLE  ATAATTTGAA TGCGGGAATT TATCCTCCAA GAAGGGAAAT ATTGAATGGA
JEZno1  ATAATTTGAA TGCGGGAATT TATCCTCCAA GAAGGGAAAT ATTGAATGGA
   KOR  ATAATTTGAA TGCGGGAATT TATCCTCCAA GAAGGGAAAT ATTGAATGGA
   KOY  ATAATTTGAA TGCGGGAATT TATCCTCCAA GAAGGGAAAT ATTGAATGGA
   LIK  ATAATTTGAA TGCGGGAATT TATCCTCCAA GAAGGGAAAT ATTGAATGGA
MARno2  ATAATTTGAA TGCGGGAATT TATCCTCCAA GAAGGGAAAT ATTGAATGGA
   MAT  ATAATTTGAA TGCGGGAATT TATCCTCCAA GAAGGGAAAT ATTGAATGGA
   MAX  ATAATTTGAA TGCGGGAATT TATCCTCCAA GAAGGGAAAT ATTGAATGGA
   MEX  ATAATTTGAA TGCGGGAATT TATCCTCCAA GAAGGGAAAT ATTGAATGGA
   MEY  ATAATTTGAA TGCGGGAATT TATCCTCCAA GAAGGGAAAT ATTGAATGGA
   MON  ATAATTTGAA TGCGGGAATT TATCCTCCAA GAAGGGAAAT ATTGAATGGA
   MOR  ATAATTTGAA TGCGGGAATT TATCCTCCAA GAAGGGAAAT ATTGAATGGA
   OBO  ATAATTTGAA TGCGGGAATT TATCCTCCAA GAAGGGAAAT ATTGAATGGA
   OMO  ATAATTTGAA TGCGGGAATT TATCCTCCAA GAAGGGAAAT ATTGAATGGA
   ORI  ATAATTTGAA TGCGGGAATT TATCCTCCAA GAAGGGAAAT ATTGAATGGA
PINBANK  ATAATTTGAA TGCGGGAATT TATCCTCCAA GAAGGGAAAT ATTGAATGGA
PINKOR  ATAATTTGAA TGCGGGAATT TATCCTCCAA GAAGGGAAAT ATTGAATGGA
   PUN  ATAATTTGAA TGCGGGAATT TATCCTCCAA GAAGGGAAAT ATTGAATGGA
   PUR  ATAATTTGAA TGCGGGAATT TATCCTCCAA GAAGGGAAAT ATTGAATGGA
   RET  ATAATTTGAA TGCGGGAATT TATCCTCCAA GAAGGGAAAT ATTGAATGGA
   RUB  ATAATTTGAA TGCGGGAATT TATCCTCCAA GAAGGGAAAT ATTGAATGGA
   SCH  ATAATTTGAA TGCGGGAATT TATCCTCCAA GAAGGGAAAT ATTGAATGGA
SITno1  ATAATTTGAA TGCGGGAATT TATCCTCCAA GAAGGGAAAT ATTGAATGGA
   SMI  ATAATTTGAA TGCGGGAATT TATCCTCCAA GAAGGGAAAT ATTGAATGGA
   SPI  ATAATTTGAA TGCGGGAATT TATCCTCCAA GAAGGGAAAT ATTGAATGGA
   TOR  ATAATTTGAA TGCGGGAATT TATCCTCCAA GAAGGGAAAT ATTGAATGGA
   WIL  ATAATTTGAA TGCGGGAATT TATCCTCCAA GAAGGGAAAT ATTGAATGGA

        [1351                                              1400]
ABIno1  TCGACCGGAA ACTCTTCCAT TCATTCATCC CTTCCACAGA ATATTTTGAC
   ALC  TCGACCGGAA ACTCTTCCAT TCATTCATCC CTTCCACAGA ATATTTTGAC
   ASP  TCGACCGGAA ACTCTTCCAT TCATTCATCC CTTCCACAGA ATATTTTGAC
   BRA  TCGACCGGAA ACTCTTCCAT TCATTCATCC CTTCCACAGA ATATTTTGAC
   BRE  TCGACCGGAA ACTCTTCCAT TCATTCATCC CTTCCACAGA ATATTTTGAC
   CHI  TCGACCGGAA ACTCTTCCAT TCATTCATCC CTTCCACAGA ATATTTTGAC
```

```
   CRA  TCGACCGGAA ACTCTTCCAT TCATTCATCC CTTCCACAGA ATATTTTGAC
ENGno3  TCGACCGGAA ACTCTTCCAT TCATTCATCC CTTCCACAGA ATATTTTGAC
   FAR  TCGACCGGAA ACTCTTCCAT TCATTCATCC CTTCCACAGA ATATTTTGAC
GLAno2  TCGACCGGAA ACTCTTCCAT TCATTCATCC CTTCCACAGA ATATTTTGAC
   GLE  TCGACCGGAA ACTCTTCCAT TCATTCATCC CTTCCACAGA ATATTTTGAC
JEZno1  TCGACCGGAA ACTCTTCCAT TCATTCATCC CTTCCACAGA ATATTTTGAC
   KOR  TCGACCGGAA ACTCTTCCAT TCATTCATCC CTTCCACAGA ATATTTTGAC
   KOY  TCGACCGGAA ACTCTTCCAT TCATTCATCC CTTCCACAGA ATATTTTGAC
   LIK  TCGACCGGAA ACTCTTCCAT TCATTCATCC CTTCCACAGA ATATTTTGAC
MARno2  TCGACCGGAA ACTCTTCCAT TCATTCATCC CTTCCACAGA ATATTTTGAC
   MAT  TCGACCGGAA ACTCTTCCAT TCATTCATCC CTTCCACAGA ATATTTTGAC
   MAX  TCGACCGGAA ACTCTTCCAT TCATTCATCC CTTCCACAGA ATATTTTGAC
   MEX  TCGACCGGAA ACTCTTCCAT TCATTCATCC CTTCCACAGA ATATTTTGAC
   MEY  TCGACCGGAA ACTCTTCCAT TCATTCATCC CTTCCACAGA ATATTTTGAC
   MON  TCGACCGGAA ACTCTTCCAT TCATTCATCC CTTCCACAGA ATATTTTGAC
   MOR  TCGACCGGAA ACTCTTCCAT TCATTCATCC CTTCCACAGA ATATTTTGAC
   OBO  TCGACCGGAA ACTCTTCCAT TCATTCATCC CTTCCACAGA ATATTTTGAC
   OMO  TCGACCGGAA ACTCTTCCAT TCATTCATCC CTTCCACAGA ATATTTTGAC
   ORI  TCGACCGGAA ACTCTTCCAT TCATTCATCC CTTCCACAGA ATATTTTGAC
PINBANK  TCGACCGGAA ACTCTTCCAT TCATTCATCC CTTCCACAGA ATATTTTGAC
PINKOR  TCGACCGGAA ACTCTTCCAT TCATTCATCC CTTCCACAGA ATATTTTGAC
   PUN  TCGACCGGAA ACTCTTCCAT TCATTCATCC CTTCCACAGA ATATTTTGAC
   PUR  TCGACCGGAA ACTCTTCCAT TCATTCATCC CTTCCACAGA ATATTTTGAC
   RET  TCGACCGGAA ACTCTTCCAT TCATTCATCC CTTCCACAGA ATATTTTGAC
   RUB  TCGACCGGAA ACTCTTCCAT TCATTCATCC CTTCCACAGA ATATTTTGAC
   SCH  TCGACCGGAA ACTCTTCCAT TCATTCATCC CTTCCACAGA ATATTTTGAC
SITno1  TCGACCGGAA ACTCTTCCAT TCATTCATCC CTTCCACAGA ATATTTTGAC
   SMI  TCGACCGGAA ACTCTTCCAT TCATTCATCC CTTCCACAGA ATATTTTGAC
   SPI  TCGACCGGAA ACTCTTCCAT TCATTCATCC CTTCCACAGA ATATTTTGAC
   TOR  TCGACCGGAA ACTCTTCCAT TCATTCATCC CTTCCACAGA ATATTTTGAC
   WIL  TCGACCGGAA ACTCTTCCAT TCATTCATCC CTTCCACAGA ATATTTTGAC

       [1401                                              1450]
ABIno1  CGTATAGAGA ACGGAACTTC CAGGACCAAT GTCAGTCCTT CTAGTACCGA
   ALC  CGTATAGAGA ACGGAACTTC CAGGACCAAT GTCAGTCCTT CTAGTACCGA
   ASP  CGTATAGAGA ACGGAACTTC CAGGACCAAT GTCAGTCCTT CTAGTACCGA
   BRA  CGTATAGAGA ACGGAACTTC CAGGACCAAT GTCAGTCCTT CTAGTACCGA
   BRE  CGTATAGAGA ACGGAACTTC CAGGACCAAT GTCAGTCCTT CTAGTACCGA
   CHI  CGTATAGAGA ACGGAACTTC CAGGACCAAT GTCAGTCCTT CTAGTACCGA
   CRA  CGTATAGAGA ACGGAACTTC CAGGACCAAT GTCAGTCCTT CTAGTACCGA
ENGno3  CGTATAGAGA ACGGAACTTC CAGGACCAAT GTCAGTCCTT CTAGTACCGA
   FAR  CGTATAGAGA ACGGAACTTC CAGGACCAAT GTCAGTCCTT CTAGTACCGA
GLAno2  CGTATAGAGA ACGGAACTTC CAGGACCAAT GTCAGTCCTT CTAGTACCGA
   GLE  CGTATAGAGA ACGGAACTTC CAGGACCAAT GTCAGTCCTT CTAGTACCGA
JEZno1  CGTATAGAGA ACGGAACTTC CAGGACCAAT GTCAGTCCTT CTAGTACCGA
   KOR  CGTATAGAGA ACGGAACTTC CAGGACCAAT GTCAGTCCTT CTAGTACCGA
   KOY  CGTATAGAGA ACGGAACTTC CAGGACCAAT GTCAGTCCTT CTAGTACCGA
   LIK  CGTATAGAGA ACGGAACTTC CAGGACCAAT GTCAGTCCTT CTAGTACCGA
MARno2  CGTATAGAGA ACGGAACTTC CAGGACCAAT GTCAGTCCTT CTAGTACCGA
   MAT  CGTATAGAGA ACGGAACTTC CAGGACCAAT GTCAGTCCTT CTAGTACCGA
   MAX  CGTATAGAGA ACGGAACTTC CAGGACCAAT GTCAGTCCTT CTAGTACCGA
   MEX  CGTATAGAGA ACGGAACTTC CAGGACCAAT GTCAGTCCTT CTAGTACCGA
   MEY  CGTATAGAGA ACGGAACTTC CAGGACCAAT GTCAGTCCTT CTAGTACCGA
   MON  CGTATAGAGA ACGGAACTTC CAGGACCAAT GTCAGTCCTT CTAGTACCGA
   MOR  CGTATAGAGA ACGGAACTTC CAGGACCAAT GTCAGTCCTT CTAGTACCGA
   OBO  CGTATAGAGA ACGGAACTTC CAGGACCAAT GTCAGTCCTT CTAGTACCGA
   OMO  CGTATAGAGA ACGGAACTTC CAGGACCAAT GTCAGTCCTT CTAGTACCGA
   ORI  CGTATAGAGA ACGGAACTTC CAGGACCAAT GTCAGTCCTT CTAGTACCGA
PINBANK  CGTATAGAGA ACGGAACTTC CAGGACCAAT GTCAGTCCTT CTAGTACCGA
PINKOR  CGTATAGAGA ACGGAACTTC CAGGACCAAT GTAAGTCCTT CTAGTACCGA
   PUN  CGTATAGAGA ACGGAACTTC CAGGACCAAT GTCAGTCCTT CTAGTACCGA
   PUR  CGTATAGAGA ACGGAACTTC CAGGACCAAT GTCAGTCCTT CTAGTACCGA
   RET  CGTATAGAGA ACGGAACTTC CAGGACCAAT GTCAGTCCTT CTAGTACCGA
   RUB  CGTATAGAGA ACGGAACTTC CAGGACCAAT GTCAGTCCTT CTAGTACCGA
   SCH  CGTATAGAGA ACGGAACTTC CAGGACCAAT GTCAGTCCTT CTAGTACCGA
SITno1  CGTATAGAGA ACGGAACTTC CAGGACCAAT GTCAGTCCTT CTAATACCGA
   SMI  CGTATAGAGA ACGGAACTTC CAGGACCAAT GTCAGTCCTT CTAGTACCGA
   SPI  CGTATAGAGA ACGGAACTTC CAGGACCAAT GTCAGTCCTT CTAGTACCGA
   TOR  CGTATAGAGA ACGGAACTTC CAGGACCAAT GTCAGTCCTT CTAGTACCGA
   WIL  CGTATAGAGA ACGGAACTTC CAGGACCAAT GTCAGTCCTT CTAGTACCGA

       [1451                                              1500]
ABIno1  TTCAGAATAG GAACTCTTCT TGCGATCAAC TAATGGATTT GGATCACAAT
   ALC  TTCAGAATAG GAACTCTTCT TGCGATCAAC TAATGGATTT GGATCGCAAT
   ASP  TTCAGAATAG GAACTCTTCT TGCGATCAAC TAATGGATTT GGATCACAAT
   BRA  TTCAGAATAG GAACTCTTCT TGCGATCAAC TAATGGATTT GGATCGCAAT
   BRE  TTCAGAATAG GAACTCTTCT TGCGATCAAC TAATGGATTT GGATCGCAAT
   CHI  TTCAGAATAG GAACTCTTCT TGCGATCAAC TAATGGATTT GGATCGCAAT
   CRA  TTCAGAATAG GAACTCTTCT TGCGATCAAC TAATGGATTT GGATCACAAT
ENGno3  TTCAGAATAG GAACTCTTCT TGCGATCAAC TAATGGATTT GGATCACAAT
   FAR  TTCAGAATAG GAACTCTTCT TGCGATCAAC TAATGGATTT GGATCGCAAT
GLAno2  TTCAGAATAG GAACTCTTCT TGCGATCAAC TAATGGATTT GGATCACAAT
   GLE  TTCAGAATAG GAACTCTTCT TGCGATCAAC TAATGGATTT GGATCGCAAT
JEZno1  TTCAGAATAG GAACTCTTCT TGCGATCAAC TAATGGATTT GGATCGCAAT
```

```
   KOR  TTCAGAATAG GAACTCTTCT TGCGATCAAC TAATGGATTT GGATCACAAT
   KOY  TTCAGAATAG GAACTCTTCT TGCGATCAAC TAATGGATTT GGATCACAAT
   LIK  TTCAGAATAG GAACTCTTCT TGCGATCAAC TAATGGATTT GGATCGCAAT
MARno2  TTCAGAATAG GAACTCTTCT TGCGATCAAC TAATGGATTT GGATCGCAAT
   MAT  TTCAGAATAG GAACTCTTCT TGCGATCAAC TAATGGATTT GGATCGCAAT
   MAX  TTCAGAATAG GAACTCTTCT TGCGATCAAC TAATGGATTT GGATCGCAAT
   MEX  TTCAGAATAG GAACTCTTCT TGCGATCAAC TAATGGATTT GGATCACAAT
   MEY  TTCAGAATAG GAACTCTTCT TGCGATCAAC TAATGGATTT GGATCACAAT
   MON  TTCAGAATAG GAACTCTTCT TGCGATCAAC TAATGGATTT GGATCACAAT
   MOR  TTCAGAATAG GAACTCTTCT TGCGATCAAC TAATGGATTT GGATCGCAAT
   OBO  TTCAGAATAG GAACTCTTCT TGCGATCAAC TAATGGATTT GGATCACAAT
   OMO  TTCAGAATAG GAACTCTTCT TGCGATCAAC TAATGGATTT GGATCGCAAT
   ORI  TTCAGAATAG GAACTCTTCT TGCGATCAAC TAATGGATTT GGATCGCAAT
PINBANK  TTCAGAATAG GAACTCTTCT TGCGATCAAC TAATGGATTT GGATCACAAT
PINKOR  TTCAGAATAG GAACTCTTCT TGCAATCAGC TAATGGATTT GGAGCGCAAT
   PUN  TTCAGAATAG GAACTCTTCT TGCGATCAAC TAATGGATTT GGATCGCAAT
   PUR  TTCAGAATAG GAACTCTTCT TGCGATCAAC TAATGGATTT GGATCGCAAT
   RET  TTCAGAATAG GAACTCTTCT TGCGATCAAC TAATGGATTT GGATCACAAT
   RUB  TTCAGAATAG GAACTCTTCT TGCGATCAAC TAATGGATTT GGATCGCAAT
   SCH  TTCAGAATAG GAACTCTTCT TGCGATCAAC TAATGGATTT GGATCGCAAT
SITno1  TTCAGAATAG GAACTCTTCT TGCGATCAAC TAATGGATTT GGATCGCAAT
   SMI  TTCAGAATAG GAACTCTTCT TGCGATCAAC TAATGGATTT GGATCGCAAT
   SPI  TTCAGAATAG GAACTCTTCT TGCGATCAAC TAATGGATTT GGATCGCAAT
   TOR  TTCAGAATAG GAACTCTTCT TGCGATCAAC TAATGGATTT GGATCACAAT
   WIL  TTCAGAATAG GAACTCTTCT TGCGATCAAC TAATGGATTT GGATCGCAAT

        [1501                                              1550]
ABIno1  TCACGAATAA AACAATTGAA TGATTCTGTT GACGTATTTG ACCAATCAAA
   ALC  TCACGAATAA AACAATTGAA TGATTCTGTT GACGTATTTG ACCAATCAAA
   ASP  TCACGAATAA AACAATTGAA TGATTCTGTT GACGTATTTG ACCAATCAAA
   BRA  TCACGAATAA AACAATTGAA TGATTCTGTT GACGTATTTG ACCAATCAAA
   BRE  TCACGAATAA AAAAATTGAA TGATTCTGTT GACGTATTTG ACCAATCAAA
   CHI  TCACGAATAA AACAATTGAA TGATTCTGTT GACGTATTTG ACCAATCAAA
   CRA  TCACGAATAA AACAATTGAA TGATTCTGTT GACGTATTTG ACCAATCAAA
ENGno3  TCACGAATAA AACAATTGAA TGATTCTGTT GACGTATTTG ACCAATCAAA
   FAR  TCACGAATAA AACAATTGAA TGATTCTGTT GACGTATTTG ACCAATCAAA
GLAno2  TCACGAATAA AACAATTGAA TGATTCTGTT GACGTATTTG ACCAATCAAA
   GLE  TCACGAATAA AACAATTGAA TGATTCTGTT GACGTATTTG ACCAATCAAA
JEZno1  TCACGAATAA AACAATTGAA TGATTCTGTT GACGTATTTG ACCAATCAAA
   KOR  TCACGAATAA AACAATTGAA TGATTCTGTT GACGTATTTG ACCAATCAAA
   KOY  TCACGAATAA AACAATTGAA TGATTCTGTT GACGTATTTG ACCAATCAAA
   LIK  TCACGAATAA AACAATTGAA TGATTCTGTT GACGTATTTG ACCAATCAAA
MARno2  TCACGAATAA AACAATTGAA TGATTCTGTT GACGTATTTG ACCAATCAAA
   MAT  TCACGAATAA AACAATTGAA TGATTCTGTT GACGTATTTG ACCAATCAAA
   MAX  TCACGAATAA AACAATTGAA TGATTCTGTT GACGTATTTG ACCAATCAAA
   MEX  TCACGAATAA AACAATTGAA TGATTCTGTT GACGTATTTG ACCAATCAAA
   MEY  TCACGAATAA AACAATTGAA TGATTCTGTT GACGTATTTG ACCAATCAAA
   MON  TCACGAATAA AACAATTGAA TGATTCTGTT GACGTATTTG ACCAATCAAA
   MOR  TCACGAATAA AACAATTGAA TGATTCTGTT GACGTATTTG ACCAATCAAA
   OBO  TCACGAATAA AACAATTGAA TGATTCTGTT GACGTATTTG ACCAATCAAA
   OMO  TCACGAATAA AACAATTGAA TGATTCTGTT GACGTATTTG ACCAATCAAA
   ORI  TCACGAATAA AACAATTGAA TGATTCTGTT GACGTATTTG ACCAATCAAA
PINBANK  TCACGAATAA AACAATTGAA TGATTCTGTT GACGTATTTG ACCAATCAAA
PINKOR  TCACGAATAA AACAATTGAA TGATTTTGTT GACGTATTTG ACCAATCAAA
   PUN  TCACGAATAA AACAATTGAA TGATTCTGTT GACGTATTTG ACCAATCAAA
   PUR  TCACGAATAA AACAATTGAA TGATTCTGTT GACGTATTTG ACCAATCAAA
   RET  TCACGAATAA AACAATTGAA TGATTCTGTT GACGTATTTG ACCAATCAAA
   RUB  TCACGAATAA AACAATTGAA TGATTCTGTT GACGTATTTG ACCAATCAAA
   SCH  TCACGAATAA AACAATTGAA TGATTCTGTT GACGTATTTG ACCAATCAAA
SITno1  TCACGAATAA AACAATTGAA TGATTCTGTT GACGTATTTG ACCAATCAAA
   SMI  TCACGAATAA AACAATTGAA TGATTCTGTT GACGTATTTG ACCAATCAAA
   SPI  TCACGAATAA AACAATTGAA TGATTCTGTT GACGTATTTG ACCAATCAAA
   TOR  TCACGAATAA AACAATTGAA TGATTCTGTT GACGTATTTG ACCAATCAAA
   WIL  TCACGAATAA AACAATTGAA TGATTCTGTT GACGTATTTG ACCAATCAAA

        [1551                                              1600]
ABIno1  CGTTTTACAG TTAGGAAACT GAACTGATCA TTGGAACTTA AATGTTCCAT
   ALC  CGTTTTACAG TTAGGAAACT GAACTGATCA TTGGAACTTA AATGTTCCAT
   ASP  CGTTTTACAG TTAGGAAACT GAACTGATCA TTGGAACTTA AATGTTCCAT
   BRA  CGTTTTACAG TTAGGAAACT GAACTGATCA TTGGAACTTA AATGTTCCAT
   BRE  CGTTTTACAG TTAGGAAACT GAACTGATCA TTGGAACTTA AATGTTCCAT
   CHI  CGTTTTACAG TTAGGAAACT GAACTGATCA TTGGAACTTA AATGTTCCAT
   CRA  CGTTTTACAG TTAGGAAACT GAACTGATCA TTGGAACTTA AATGTTCCAT
ENGno3  CGTTTTACAG TTAGGAAACT GAACTGATCA TTGGAACTTA AATGTTCCAT
   FAR  CGTTTTACAG TTAGGAAACT GAACTGATCA TTGGAACTTA AATGTTCCAT
GLAno2  CGTTTTACAG TTAGGAAACT GAACTGATCA TTGGAACTTA AATGTTCCAT
   GLE  CGTTTTACAG TTAGGAAACT GAACTGATCA TTGGAACTTA AATGTTCCAT
JEZno1  CGTTTTACAG TTAGGAAACT GAACTGATCA TTGGAACTTA AATGTTCCAT
   KOR  CGTTTTACAG TTAGGAAACT GAACTGATCA TTGGAACTTA AATGTTCCAT
   KOY  CGTTTTACAG TTAGGAAACT GAACTGATCA TTGGAACTTA AATGTTCCAT
   LIK  CGTTTTACAG TTAGGAAACT GAACTGATCA TTGGAACTTA AATGTTCCAT
MARno2  CGTTTTACAG TTAGGAAACT GAACTGATCA TTGGAACTTA AATGTTCCAT
   MAT  CGTTTTACAG TTAGGAAACT GAACTGATCA TTGGAACTTA AATGTTCCAT
   MAX  CGTTTTACAG TTAGGAAACT GAACTGATCA TTGGAACTTA AATGTTCCAT
```

```
   MEX  CGTTTTACAG TTAGGAAACT GAACTGATCA TTGGAACTTA AATGTTCCAT
   MEY  CGTTTTACAG TTAGGAAACT GAACTGATCA TTGGAACTTA AATGTTCCAT
   MON  CGTTTTACAG TTAGGAAACT GAACTGATCA TTGGAACTTA AATGTTCCAT
   MOR  CGTTTTACAG TTAGGAAACT GAACTGATCA TTGGAACTTA AATGTTCCAT
   OBO  CGTTTTACAG TTAGGAAACT GAACTGATCA TTGGAACTTA AATGTTCCAT
   OMO  CGTTTTACAG TTAGGAAACT GAACTGATCA TTGGAACTTA AATGTTCCAT
   ORI  CGTTTTACAG TTAGGAAACT GAACTGATCA TTGGAACTTA AATGTTCCAT
PINBANK CGTTTTACAG TTAGGAAACT GAACTGATCA TTGGAACTTA AATGTTCCAT
 PINKOR CGTTTTACAG TTAGGAAACT GAATTGATCA TTGGAACTTA AATGTTCCAT
   PUN  CGTTTTACAG TTAGGAAACT GAACTGATCA TTGGAACTTA AATGTTCCAT
   PUR  CGTTTTACAG TTAGGAAACT GAACTGATCA TTGGAACTTA AATGTTCCAT
   RET  CGTTTTACAG TTAGGAAACT GAACTGATCA TTGGAACTTA AATGTTCCAT
   RUB  CGTTTTACAG TTAGGAAACT GAACTGATCA TTGGAACTTA AATGTTCCAT
   SCH  CGTTTTACAG TTAGGAAACT GAACTGATCA TTGGAACTTA AATGTTCCAT
 SITno1 CGTTTTACAG TTAGGAAACT GAACTGATCA TTGGAACTTA AATGTTCCAT
   SMI  CGTTTTACAG TTAGGAAACT GAACTGATCA TTGGAACTGA AATGTTCCAT
   SPI  CGTTTTACAG TTAGGAAACT GAACTGATCA TTGGAACTTA AATGTTCCAT
   TOR  CGTTTTACAG TTAGGAAACT GAACTGATCA TTGGAACTTA AATGTTCCAT
   WIL  CGTTTTACAG TTAGGAAACT GAACTGATCA TTGGAACTTA AATGTTCCAT

        [1601                                              1650]
ABIno1 CGGTTCGGAG GAACTCGATC CATCCAAATA ATGATCATGA GAAATTGCGT
   ALC  CGGTTCGGAG GAACTCGATC CATCCAAATA ATGATCATGA GAAATTGCGT
   ASP  CGGTTCGGAG GAACTCGATC CATCCAAATA ATGATCATGA GAAATTGCGT
   BRA  CGGTTCGGAG GAACTCGATC CATCCAAATA ATGATCATGA GAAATTGCGT
   BRE  CGGTTCGGAG GAACTCGATC CATCCAAATA ATGATCATGA GAAATTGCGT
   CHI  CGGTTCGGAG GAACTCGATC CATCCAAATA ATGATCATGA GAAATTGCGT
   CRA  CGGTTCGGAG GAACTCGATC CATCCAAATA ATGATCATGA GAAATTGCGT
ENGno3 CGGTTCGGAG GAACTCGATC CATCCAAATA ATGATCATGA GAAATTGCGT
   FAR  CGGTTCGGAG GAACTCGATC CATCCAAATA ATGATCATGA GAAATTGCGT
GLAno2 CGGTTCGGAG GAACTCGATC CATCCAAATA ATGATCATGA GAAATTGCGT
   GLE  CGGTTCGGAG GAACTCGATC CATCCAAATA ATGATCATGA GAAATTGCGT
JEZno1 CGGTTCGGAG GAACTCGATC CATCCAAATA ATGATCATGA GAAATTGCGT
   KOR  CGGTTCGGAG GAACTCGATC CATCCAAATA ATGATCATGA GAAATTGCGT
   KOY  CGGTTCGGAG GAACTCGATC CATCCAAATA ATGATCATGA GAAATTGCGT
   LIK  CGGTTCGGAG GAACTCGATC CATCCAAATA ATGATCATGA GAAATTGCGT
MARno2 CGGTTCGGAG GAACTCGATC CATCCAAATA ATGATCATGA GAAATTGCGT
   MAT  CGGTTCGGAG GAACTCGATC CATCCAAATA ATGATCATGA GAAATTGCGT
   MAX  CGGTTCGGAG GAACTCGATC CATCCAAATA ATGATCATGA GAAATTGCGT
   MEX  CGGTTCGGAG GAACTCGATC CATCCAAATA ATGATCATGA GAAATTGCGT
   MEY  CGGTTCGGAG GAACTCGATC CATCCAAATA ATGATCATGA GAAATTGCGT
   MON  CGGTTCGGAG GAACTCGATC CATCCAAATA ATGATCATGA GAAATTGCGT
   MOR  CGGTTCGGAG GAACTCGATC CATCCAAATA ATGATCATGA GAAATTGCGT
   OBO  CGGTTCGGAG GAACTCGATC CATCCAAATA ATGATCATGA GAAATTGCGT
   OMO  CGGTTCGGAG GAACTCGATC CATCCAAATA ATGATCATGA GAAATTGCGT
   ORI  CGGTTCGGAG GAACTCGATC CATCCAAATA ATGATCATGA GAAATTGCGT
PINBANK CGGTTCGGAG GAACTCGATC CATCCAAATA ATGATCATGA GAAATTGCGT
 PINKOR CGGTTCGGAA GAACTTGATC CATCCAAATA ATGATCATGA GAAATTGCGT
   PUN  CGGTTCGGAG GAACTCGATC CATCCAAATA ATGATCATGA GAAATTGCGT
   PUR  CGGTTCGGAG GAACTCGATC CATCCAAATA ATGATCATGA GAAATTGCGT
   RET  CGGTTCGGAG GAACTCGATC CATCCAAATA ATGATCATGA GAAATTGCGT
   RUB  CGGTTCGGAG GAACTCGATC CATCCAAATA ATGATCATGA GAAATTGCGT
   SCH  CGGTTCGGAG GAACTCGATC CATCCAAATA ATGATCATGA GAAATTGCGT
 SITno1 CGGTTCGGAG GAACTCGATC CATCCAAATA ATGATCATGA GAAATTGCGT
   SMI  CGGTTCGGAG GAACTCGATC CATCCAAATA ATGATCATGA GAAATTGCGT
   SPI  CGGTTCGGAG GAACTCGATC CATCCAAATA ATGATCATGA GAAATTGCGT
   TOR  CGGTTCGGAG GAACTCGATC CATCCAAATA ATGATCATGA GAAATTGCGT
   WIL  CGGTTCGGAG GAACTCGATC CATCCAAATA ATGATCATGA GAAATTGCGT

        [1651                                              1700]
ABIno1 AAAGATCTTC CTGAAAAAAG AGTGGATATA AAAAGCATTG TTGCCAAGAG
   ALC  AAAGATCTTC CTGAAAAAAG AGTGGATATA AAAAGCATTG TTGCCAAGAG
   ASP  AAAGATCTTC CTGAAAAAAG AGTGGATATA AAAAGCATTG TTGCCAAGAG
   BRA  AAAGATCTTC CTGAAAAAAG AGTGGATATA AAAAGCATTG TTGCCAAGAG
   BRE  AAAGATCTTC CTGAAAAAAG AGTGGATATA AAAAGCATTG TTGCCAAGAG
   CHI  AAAGATCTTC CTGAAAAAAG AGTGGATATA AAAAGCATTG TTGCCAAGAG
   CRA  AAAGATCTTC CTGAAAAAAG AGTGGATATA AAAAGCATTG TTGCCAAGAG
ENGno3 AAAGATCTTC CTGAAAAAAG AGTGGATATA AAAAGCATTG TTGCCAAGAG
   FAR  AAAGATCTTC CTGAAAAAAG AGTGGATATA AAAAGCATTG TTGCCAAGAG
GLAno2 AAAGATCTTC CTGAAAAAAG AGTGGATATA AAAAGCATTG TTGCCAAGAG
   GLE  AAAGATCTTC CTGAAAAAAG AGTGGATATA AAAAGCATTG TTGCCAAGAG
JEZno1 AAAGATCTTC CTGAAAAAAG AGTGGATATA AAAAGCATTG TTGCCAAGAG
   KOR  AAAGATCTTC CTGAAAAAAG AGTGGATATA AAAAGCATTG TTGCCAAGAG
   KOY  AAAGATCTTC CTGAAAAAAG AGTGGATATA AAAAGCATTG TTGCCAAGAG
   LIK  AAAGATCTTC CTGAAAAAAG AGTGGATATA AAAAGCATTG TTGCCAAGAG
MARno2 AAAGATCTTC CTGAAAAAAG AGTGGATATA AAAAGCATTG TTGCCAAGAG
   MAT  AAAGATCTTC CTGAAAAAAG AGTGGATATA AAAAGCATTG TTGCCAAGAG
   MAX  AAAGATCTTC CTGAAAAAAG AGTGGATATA AAAAGCATTG TTGCCAAGAG
   MEX  AAAGATCTTC CTGAAAAAAG AGTGGATATA AAAAGCATTG TTGCCAAGAG
   MEY  AAAGATCTTC CTGAAAAAAG AGTGGATATA AAAAGCATTG TTGCCAAGAG
   MON  AAAGATCTTC CTGAAAAAAG AGTGGATATA AAAAGCATTG TTGCCAAGAG
   MOR  AAAGATCTTC CTGAAAAAAG AGTGGATATA AAAAGCATTG TTGCCAAGAG
   OBO  AAAGATCTTC CTGAAAAAAG AGTGGATATA AAAAGCATTG TTGCCAAGAG
   OMO  AAAGATCTTC CTGAAAAAAG AGTGGATATA AAAAGCATTG TTGCCAAGAG
```

```
   ORI  AAAGATCTTC CTGAAAAAAG AGTGGATATA AAAAGCATTG TTGCCAAGAG
PINBANK AAAGATCTTC CTGAAAAAAG AGTGGATATA AAAAGCATTG TTGCCAAGAG
 PINKOR AAAGATCTTC CTGAAAAAAG AGTGGATATA AAAAGCATTG TTGCCGAGAG
   PUN  AAAGATCTTC CTGAAAAAAG AGTGGATATA AAAAGCATTG TTGCCAAGAG
   PUR  AAAGATCTTC CTGAAAAAAG AGTGGATATA AAAAGCATTG TTGCCAAGAG
   RET  AAAGATCTTC CTGAAAAAAG AGTGGATATA AAAAGCATTG TTGCCAAGAG
   RUB  AAAGATCTTC CTGAAAAAAG AGTGGATATA AAAAGCATTG TTGCCAAGAG
   SCH  AAAGATCTTC CTGAAAAAAG AGTGGATATA AAAAGCATTG TTGCCAAGAG
 SITno1 AAAGATCTTC CTGAAAAAAG AGTGGATATA AAAAGCATTG TTGCCAAGAG
   SMI  AAAGATCTTC CTGAAAAAAG AGTGGATATA AAAAGCATTG TTGCCAAGAG
   SPI  AAAGATCTTC CTGAAAAAAG AGTGGATATA AAAAGCATTG TTGCCAAGAG
   TOR  AAAGATCTTC CTGAAAAAAG AGTGGATATA AAAAGCATTG TTGCCAAGAG
   WIL  AAAGATCTTC CTGAAAAAAG AGTGGATATA AAAAGCATTG TTGCCAAGAG

        [1701                                              1750]
 ABIno1 CTATCTTCCT TTCCGTATCK ATGGAACTCA TCCATTTTCA AACCT.....
   ALC  CTATCTTCCT TTCCGTATCT ATGGAACTCA TCCATTTTCA AACCTCAGTT
   ASP  CTATCTTCCT TTCCGTATCT ATGGAACTCA TCCATTTTCA AACCT.....
   BRA  CTATCTTCCT TTCCGTATCT ATGGAACTCA TCCATTTTCA AACCTCAGTT
   BRE  CTATCTTCCT TTCCGTATCT ATGGAACTCA TCCATTTGAA AACCTCAGTT
   CHI  CTATCTTCCT TTCCGTATCT ATGGAACTCA TCCATTTTCA AACCTCAGTT
   CRA  CTATCTTCCT TTCCGTATCT ATGGAACTCA TCCATTTTCA AACCT.....
 ENGno3 CTATCTTCCT TTCCGTATCT ATGGAACTCA TCCATTTTCA AACCTCAGTT
   FAR  CTATCTTCCT TTCCGTATCT ATGGAACTCA TCCATTTTCA AACCTCAGTT
 GLAno2 CTATCTTCCT TTCCGTATCT ATGGAACTCA TCCATTTTCA AACCTCAGTT
   GLE  CTATCTTCCT TTCCGTATCT ATGGAACTCA TCCATTTTAA AACCTCAGTT
 JEZno1 CTATCTTCCT TTCCGTATCT ATGGAACTCA TCCATTTTAA AACCTCAGTT
   KOR  CTATCTTCCT TTCCGTATCT ATGGAACTCA TCCATTTTAA AACCT.....
   KOY  CTATCTTCCT TTCCGTATCG ATGGAACTCA TCCATTTTCA AACCT.....
   LIK  CTATCTTCCT TTCCGTATCT ATGGAACTCA TCCATTTTCA AACCTCAGTT
 MARno2 CTATCTTCCT TTCCGTATCT ATRGAACTCA TCCATTTKAA AACCTCAGTT
   MAT  CTATCTTCCT TTCCGTATCT ATGGAACTCA TCCATTTTCA AACCTCAGTT
   MAX  CTATCTTCCT TTCCGTATCT ATGGAACTCA TCCATTTTCA AACCTCAGTT
   MEX  CTATCTTCCT TTCCGTATCT ATGGAACTCA TCCATTTTCA AACCTCAGTT
   MEY  CTATCTTCCT TTCCGTATCT ATGGAACTCA TCCATTTTCA AACCT.....
   MON  CTATCTTCCT TTCCGTATCT ATGGAACTCA TCCATTTTCA AACCT.....
   MOR  CTATCTTCCT TTCCGTATCT ATGGAACTCA TCCATTTTCA AACCTCAGTT
   OBO  CTATCTTCCT TTCCGTATCT ATGGAACTCA TCCATTTTCA AACCT.....
   OMO  CTATCTTCCT TTCCGTATCT ATGGAACTCA TCCATTTTAA AACCTCAGTT
   ORI  CTATCTTCCT TTCCGTATCT ATGGAACTCA TCCATTTTCA AACCTCAGTT
PINBANK CTATCTTCCT TTCCGTATCT ATGGAACTCA TCCATTTTCA AACCTCAGTT
 PINKOR TTATCTTCCT TTCCGTATCT ATGGAATTCA TCCATTTCCA AACCTCAGCT
   PUN  CTATCTTCCT TTCCGTATCT ATGGAACTCA TCCATTTTCA AACCTCAGTT
   PUR  CTATCTTCCT TTCCGTATCT ATGGAACTCA TCCATTTTCA AACCTCAGTT
   RET  CTATCTTCCT TTCCGTATCT ATGGAACTCA TCCATTTTCA AACCT.....
   RUB  CTATCTTCCT TTCCGTATCT ATGGAACTCA TCCATTTTAA AACCTCAGTT
   SCH  CTATCTTCCT TTCCGTATCT ATGGAACTCA TCCATTTTCA AACCTCAGTT
 SITno1 CTATCTTCCT TTCCGTATCT ATGGAACTCA TCCATTTTAA AACCTCAGTT
   SMI  CTATCTTCCT TTCCGTATCT ATGGAACTCA TCCATTTTCA AACCTCAGTT
   SPI  CTATCTTCCT TTCCGTATCT ATGGAACTCA TCCATTTTCA AACCTCAGTT
   TOR  CTATCTTCCT TTCCGTATCT ATGGAACTCA TCCATTTTCA AACCT.....
   WIL  CTATCTTCCT TTCCGTATCT ATGGAACTCA TCCATTTTCA AACCTCAGTT

        [1751    +1                                    +1    1800]
 ABIno1 ..1GGCCCTCG TTCATGAGAA TAACCTCTTA AATTCTGAGA TAATACAAT0A
   ALC  AT0GGCCCTCG TTCATGAGAA TAACCTCTTA AATTCTGAGA TAATACAAT0A
   ASP  ..1GGCCCTCG TTCATGAGAA TAACCTCTTA AATTCTGAGA TAATACAAT0A
   BRA  AT0GGCCCTCG TTCATGAGAA TAACCTCTTA AATTCTGAGA TAATACAAT0A
   BRE  AT0GGCCCTCG TTCATGAGAA TAACCTCTTA AATTCTGAGA TAATACAAT0A
   CHI  AT0GGCCCTCG TTCATGAGAA TAACCTCTTA AATTCTGAGA TAATACAAT0A
   CRA  ..1GGCCCTCG TTCATGAGAA TAACCTCTTA AATTCTGAGA TAATACAAT0A
 ENGno3 AT0GGCCCTCG TTCATGAGAA TAACCTCTTA AATTCTGAGA TAATACAAT0A
   FAR  AT0GGCCCTCG TTCATGAGAA TAACCTCTTA AATTCTGAGA TAAT.....1A
 GLAno2 AT0GGCCCTCG TTCATGAGAA TAACCTCTTA AATTCTGAGA TAATACAAT0A
   GLE  AT0GGCCCTCG TTCATGAGAA TAACCTCTTA AATTCTGAGA TAATACAAT0A
 JEZno1 AT0GGCCCTCG TTCATGAGAA TAACCTCTTA AATTCTGAGA TAATACAAT0A
   KOR  ..1GGCCCTCG TTCATGAGAA TAACCTCTTA AATTCTGAGA TAATACAAT0A
   KOY  ..1GGCCCTCG TTCATGAGAA TAACCTCTTA AATTCTGAGA TAATACAAT0A
   LIK  AT0GGCCCTCG TTCATGAGAA TAACCTCTTA AATTCTGAGA TAAT.....1A
 MARno2 AT0GGCCCTCG TTCATGAGAA TAACCTCTTA AATTCTGAGA TAATACAAT0A
   MAT  AT0GGCCCTCG TTCATGAGAA TAACCTCTTA AATTCTGAGA TAATACAAT0A
   MAX  AT0GGCCCTCG TTCATGAGAA TAACCTCTTA AATTCTGAGA TAATACAAT0A
   MEX  AT0GGCCCTCG TTCATGAGAA TAACCTCTTA AATTCTGAGA TAATACAAT0A
   MEY  ..1GGCCCTCG TTCATGAGAA TAACCTCTTA AATTCTGAGA TAATACAAT0A
   MON  ..1GGCCCTCG TTCATGAGAA TAACCTCTTA AATTCTGAGA TAATACAAT0A
   MOR  AT0GGCCCTCG TTCATGAGAA TAACCTCTTA AATTCTGAGA TAATACAAT0A
   OBO  ..1GGCCCTCG TTCATGAGAA TAACCTCTTA AATTCTGAGA TAATACAAT0A
   OMO  AT0GGCCCTCG TTCATGAGAA TAACCTCTTA AATTCTGAGA TAATACAAT0A
   ORI  AT0GGCCCTCG TTCATGAGAA TAACCTCTTA AATTCTGAGA TAATACAAT0A
PINBANK AT0GGCCCTCG TTCATGAGAA TAACCTCTTC ATTTCTGATA AAAT.....1A
 PINKOR AT0GGCCCTCG TTCATGAGAA TAACCTCTTC ATTTCTGAGA AAAT.....1A
   PUN  AT0GGCCCTCG TTCATGAGAA TAACCTCTTA AATTCTGAGA TAATACAAT0A
   PUR  AT0GGCCCTCG TTCATGAGAA TAACCTCTTA AATTCTGAGA TAATACAAT0A
   RET  ..1GGCCCTCG TTCATGAGAA TAACCTCTTA AATTCTGAGA TAATACAAT0A
```

```
   RUB  AT0GGCCCTCG TTCATGAGAA TAACCTCTTA AATTCTGAGA TAATACAAT0A
   SCH  AT0GGCCCTCG TTCATGAGAA TAACCTCTTA AATTCTGAGA TAATACAAT0A
SITno1  AT0GGCCCTCG TTCATGAGAA TAACCTCTTA AATTCTGAGA TAATACAAT0A
   SMI  AT0GGCCCTCG TTCATGAGAA TAACCTCTTA AATTCTGAGA TAATACAAT0A
   SPI  AT0GGCCCTCG TTCATGAGAA TAACCTCTTA AATTCTGAGA TAATACAAT0A
   TOR  ..1GGCCCTCG TTCATGAGAA TAACCTCTTA AATTCTGAGA TAATACAAT0A
   WIL  AT0GGCCCTCG TTCATGAGAA TAACCTCTTA AATTCTGAGA TAATACAAT0A

        [1801                                          +1     1850]
ABIno1  CATGGTGCGA TCTAGTCGGG ACAAGATAGG AAAAA.GATA GAT.....0GA
   ALC  CATGGTGCGA TCTAGTCGGG ACAAGATAGG AAAAA.GATA GAT.....0GA
   ASP  CATGGTGCGA TCTAGTCGGG ACAAGATAGG AAAAA.GATA GAT.....0GA
   BRA  CATGGTGCGA TCTAGTCGGG ACAAGATAGG AAAAA.GATA GAT.....0GA
   BRE  CATGGTGCGA TCTAGTCGGG ACAAGATAGG AAAAA.GATA GATATCCG1GA
   CHI  CATGGTGCGA TCTAGTCGGG ACAAGATAGG AAAAA.GATA GAT.....0GA
   CRA  CATGGTGCGA TCTAGTCGGG ACAAGATAGG AAAAA.GATA GAT.....0GA
ENGno3  CATGGTGCGA TCTAGTCGGG ACAAGATAGG AAAAA.GATA GAT.....0GA
   FAR  CATGGTGCGA TCTAGTCGGG ACAAGATAGG AAAAA.GATA GAT.....0GA
GLAno2  CATGGTGCGA TCTAGTCGGG ACAAGATAGG AAAAA.GATA GAT.....0GA
   GLE  CATGGTGCGA TCTAGTCGGG ACAAGATAGG AAAAA.GATA GAT.....0GA
JEZno1  CATGGTGCGA TCTAGTCGGG ACAAGATAGG AAAAA.GATA GAT.....0GA
   KOR  CATGGTGCGA TCTAGTCGGG ACAAGATAGG AAAAA.GATA GAT.....0GA
   KOY  CATGGTGCGA TCTAGTCGGG ACAAGATAGG AAAAA.GATA GAT.....0GA
   LIK  CATGGTGCGA TCTAGTCGGG ACAAGATAGG AAAAA.GATA GAT.....0GA
MARno2  CATGGTGCGA TCTAGTCGGG ACAAGATAGG AAAAA.GATA GAT.....0GA
   MAT  CATGGTGCGA TCTAGTCGGG ACAAGATAGG AAAAA.GATA GAT.....0GA
   MAX  CATGGTGCGA TCTAGTCGGG ACAAGATAGG AAAAA.GATA GAT.....0GA
   MEX  CATGGTGCGA TCTAGTCGGG ACAAGATAGG AAAAA.GATA GAT.....0GA
   MEY  CATGGTGCGA TCTAGTCGGG ACAAGATAGG AAAAA.GATA GAT.....0GA
   MON  CATGGTGCGA TCTAGTCGGG ACAAGATAGG AAAAA.GATA GAT.....0GA
   MOR  CATGGTGCGA TCTAGTCGGG ACAAGATAGG AAAAA.GATA GAT.....0GA
   OBO  CATGGTGCGA TCTAGTCGGG ACAAGATAGG AAAAA.GATA GAT.....0GA
   OMO  CATGGTGCGA TCTAGTCGGG ACAAGATAGG AAAAA.GATA GAT.....0GA
   ORI  CATGGTGCGA TCTAGTCGGG ACAAGATAGG AAAAA.GATA GAT.....0GA
PINBANK  CATGGTGCGA TCTAGTCGGA ACAAGATAGG AAAAAAGAGA TATATCCG1GA
PINKOR  CATGGTGCGA TCTAGTCGGA ACAAGATAGG AAAAAAGAGA TATATCCA1AA
   PUN  CATGGTGCGA TCTAGTCGGG ACAAGATAGG AAAAA.GATA GAT.....0GA
   PUR  CATGGTGCGA TCTAGTCGGG ACAAGATAGG AAAAA.GATA GAT.....0GA
   RET  CATGGTGCGA TCTAGTCGGG ACAAGATAGG AAAAA.GATA GAT.....0GA
   RUB  CATGGTGCGA TCTAGTCGGG ACAAGATAGG AAAAA.GATA GAT.....0GA
   SCH  CATGGTGCGA TCTAGTCGGG ACAAGATAGG AAAAA.GATA GAT.....0GA
SITno1  CATGGTGCGA TCTAGTCGGG ACAAGATAGG AAAAA.GATA GATATCCG1GA
   SMI  CATGGTGCGA TCTAGTCGGG ACAAGATAGG AAAAA.GATA GAT.....0GA
   SPI  CATGGTGCGA TCTAGTCGGG ACAAGATAGG AAAAA.GATA GAT.....0GA
   TOR  CATGGTGCGA TCTAGTCGGG ACAAGATAGG AAAAA.GATA GAT.....0GA
   WIL  CATGGTGCGA TCTAGTCGGG ACAAGATAGG AAAAA.GATA GAT.....0GA

        [1851                          +1                1900]
ABIno1  TATAAAGATT CTATCTTCTA TTCAGC0AGAT TTACTACCCA AGGATCTCGT
   ALC  TATAAAGATT CTATCTTCTA TTCAGC0AGAT TTACTACCCA AGGATCTCGT
   ASP  TATAAAGATT CTATCTTCTA TTCAGC0AGAT TTACTACCCA AGGATCTCGT
   BRA  TATAAAGATT CTATCTTCTA TTCAGC0AGAT TTACTACCCA AGGATCTCGT
   BRE  TATAAAGATT CTATCTTCTA TTCAGC0AGAT TTACTACCCA AGGATCTCGT
   CHI  TATAAAGATT CTATCTTCTA TTCAGC0AGAT TTACTACCCA AGGATCTCGT
   CRA  TATAAAGATT CTATCTTCTA TTCAGC0AGAT TTACTACCCA AGGATCTCGT
ENGno3  TATAAAGATT CTATCTTCTA TTCAGC0AGAT TTACTACCCA AGGATCTCGT
   FAR  TATAAAGATT CTATCTTCTA TTCAGC0AGAT TTACTACCCA AGGATCTCGT
GLAno2  TATAAAGATT CTATCTTCTA TTCAGC0AGAT TTACTACCCA AGGATCTCGT
   GLE  TATAAAGATT CTATCTTCTA TTCAGC0AGAT TTACTACCCA AGGATCTCGT
JEZno1  TATAAAGATT CTATCTTCTA TTCAGC0AGAT TTACTACCCA AGGATCTCGT
   KOR  TATAAAGATT CTATCTTCTA TTCAGC0AGAT TTACTACCCA AGGATCTCGT
   KOY  TATAAAGATT CTATCTTCTA TTCAGC0AGAT TTACTACCCA AGGATCTCGT
   LIK  TATAAAGATT CTATCTTCTA TTCAGC0AGAT TTACTACCCA AGGATCTCGT
MARno2  TATAAAGATT CTATCTTCTA TTCAGC0AGAT TTACTACCCA AGGATCTCGT
   MAT  TATAAAGATT CTATCTTCTA TTCAGC0AGAT TTACTACCCA AGGATCTCGT
   MAX  TATAAAGATT CTATCTTCTA TTCAGC0AGAT TTACTACCCA AGGATCTCGT
   MEX  TATAAAGATT CTATCTTCTA TTCAGC0AGAT TTACTACCCA AGGATCTCGT
   MEY  TATAAAGATT CTATCTTCTA TTCAGC0AGAT TTACTACCCA AGGATCTCGT
   MON  TATAAAGATT CTATCTTCTA TTCAGC0AGAT TTACTACCCA AGGATCTCGT
   MOR  TATAAAGATT CTATCTTCTA TTCAGC0AGAT TTACTACCCA AGGATCTCGT
   OBO  TATAAAGATT CTATCTTCTA TTCAGC0AGAT TTACTACCCA AGGATCTCGT
   OMO  TATAAAGATT CTATCTTCTA TTCAGC0AGAT TTACTACCCA AGGATCTCGT
   ORI  TATAAAGATT CTATCTTCTA TTCAGC0AGAT TTACTACCCA AGGATCTCGT
PINBANK  TATCAATATT CTATCTTCTA T.....1AGAT TTACTACCCA AGGATCTCAT
PINKOR  TATAAAGATT CTATCTTCTA T.....1AGAT TTACTACCCA AGGATCTCAT
   PUN  TATAAAGATT CTATCTTCTA TTCAGC0AGAT TTACTACCCA AGGATCTCGT
   PUR  TATAAAGATT CTATCTTCTA TTCAGC0AGAT TTACTACCCA AGGATCTCGT
   RET  TATAAAGATT CTATCTTCTA TTCAGC0AGAT TTACTACCCA AGGATCTCGT
   RUB  TATAAAGATT CTATCTTCTA TTCAGC0AGAT TTACTACCCA AGGATCTCGT
   SCH  TATAAAGATT CTATCTTCTA TTCAGC0AGAT TTACTACCCA AGGATCTCGT
SITno1  TATAAAGATT CTATCTTCTA TTCAGC0AGAT TTACTACCCA AGGATCTCGT
   SMI  TATAAAGATT CTATCTTCTA TTCAGC0AGAT TTACTACCCA AGGATCTCGT
   SPI  TATAAAAATT CTATCTTCTA TTCAGC0AGAT TTACTACCCA AGGATCTCGT
   TOR  TATAAAGATT CTATCTTCTA TTCAGC0AGAT TTACTACCCA AGGATCTCGT
```

```
   WIL  TATAAAGATT CTATCTTCTA TTCAGC0AGAT TTACTACCCA AGGATCTCGT

        [1901                                                1950]
ABIno1  TCGTCATGAG GAAGAACGAA AATCTTTTAT CCTGGCAACC GATCGCTCTC
   ALC  TCGTCATGAG GAAGAACGAA AATCTTTTAT CCTGGCAACC GATCGCTCTC
   ASP  TCGTCATGAG GAAGAACGAA AATCTTTTAT CCTGGCAACC GATCGCTCTC
   BRA  TCGTCATGAG GAAGAACGAA AATCTTTTAT CCTGGCAACC GATCGCTCTC
   BRE  TCGTCATGAG GAAGAACGAA AATCTTTTAT CCTGGCAACC GATCGCTCTC
   CHI  TCGTCATGAG GAAGAACGAA AATCTTTTAT CCTGGCAACC GATCGCTCTC
   CRA  TCGTCATGAG GAAGAACGAA AATCTTTTAT CCTGGCAACC GATCGCTCTC
ENGno3  TCGTCATGAG GAAGAACGAA AATCTTTTAT CCTGGCAACC GATCGCTCTC
   FAR  TCGTCATGAG GAAGAACGAA AATATTTTAT CCTGGCAACC GATCGCTCTC
GLAno2  TCGTCATGAG GAAGAACGAA AATCTTTTAT CCTGGCAACC GATCGCTCTC
   GLE  TCGTCATGAG GAAGAACGAA AATCTTTTAT CCTGGCAACC GATCGCTCTC
JEZno1  TCGTCATGAG GAAGAACGAA AATCTTTTAT CCTGGCAACC GATCGCTCTC
   KOR  TCGTCATGAG GAAGAACGAA AATCTTTTAT CCTGGCAACC GATCGCTCTC
   KOY  TCGTCATGAG GAAGAACGAA AATCTTTTAT CCTGGCAACC GATCGCTCTC
   LIK  TCGTCATGAG GAAGAACGAA AATATTTTAT CCTGGCAACC GATCGCTCTC
MARno2  TCGTCATGAG GAAGAACGAA AATCTTTTAT CCTGGCAACC GATCGCTCTC
   MAT  TCGTCATGAG GAAGAACGAA AATCTTTTAT CCTGGCAACC GATCGCTCTC
   MAX  TCGTCATGAG GAAGAACGAA AATCTTTTAT CCTGGCAACC GATCGCTCTC
   MEX  TCGTCATGAG GAAGAACGAA AATCTTTTAT CCTGGCAACC GATCGCTCTC
   MEY  TCGTCATGAG GAAGAACGAA AATCTTTTAT CCTGGCAACC GATCGCTCTC
   MON  TCGTCATGAG GAAGAACGAA AATCTTTTAT CCTGGCAACC GATCGCTCTC
   MOR  TCGTCATGAG GAAGAACGAA AATCTTTTAT CCTGGCAACC GATCGCTCTC
   OBO  TCGTCATGAG GAAGAACGAA AATCTTTTAT CCTGGCAACC GATCGCTCTC
   OMO  TCGTCATGAG GAAGAACGAA AATCTTTTAT CCTGGCAACC GATCGCTCTC
   ORI  TCGTCATGAG GAAGAACGAA AATCTTTTAT CCTGGCAACC GATCGCTCTC
PINBANK  TCGTCATGAG GAAGAACGAA AATCTTTTAT CCTGGCAACC AATCGCTCTC
PINKOR  TCGTCATGAG GAAGAACAAA ATTTTTTTAT CCTGGCAACC AATCGCTCTC
   PUN  TCGTCATGAG GAAGAACGAA AATCTTTTAT CCTGGCAACC GATCGCTCTC
   PUR  TCGTCATGAG GAAGAACGAA AATCTTTTAT CCTCGCAACC GATCGCTCTC
   RET  TCGTCATGAG GAAGAACGAA AATCTTTTAT CCTGGCAACC GATCGCTCTC
   RUB  TCGTCATGAG GAAGAACGAA AATCTTTTAT CCTGGCAACC GATCGCTCTC
   SCH  TCGTCATGAG GAAGAACGAA AATATTTTAT CCTGGCAACC GATCGCTCTC
SITno1  TCGTCATGAG GAAGAACGAA AATATTTTAT CCTGGCAACC GATCGCTCTC
   SMI  TCGTCATGAG GAAGAACGAA AATATTTTAT CCTGGCAACC GATCGCTCTC
   SPI  TCGTCATGAG GAAGAACGAA AATATTTTAT CCTGGCAACC GATCGCTCTC
   TOR  TCGTCATGAG GAAGAACGAA AATCTTTTAT CCTGGCAACC GATCGCTCTC
   WIL  TCGTCATGAG GAAGAACGAA AATCTTTTAT CCTCGCAACC GATCGCTCTC

        [1951                                                2000]
ABIno1  CTGACTCATG GAATAAATAA ACCTGGTCAG TACACTATTT ATCTTACACA
   ALC  CTGACTCATG GAATAAATAA ACCTGGTCAG TACACTATTT ATCTTACACA
   ASP  CTGACTCATG GAATAAATAA ACCTGGTCAG TACACTATTT ATCTTACACA
   BRA  CTGACTCATG GAATAAATAA ACCTGGTCAG TACACTATTT ATCTTACACA
   BRE  CTGACTCATG GAATAAATCA ACCTGGTCAG TACACTATTT ATCTTACACA
   CHI  CTGACTCATG GAATAAATAA ACCTGGTCAG TACACTATTT ATCTTACACA
   CRA  CTGACTCATG GAATAAATAA ACCTGGTCAG TACACTATTT ATCTTACACA
ENGno3  CTGACTCATG GAATAAATAA ACCTGGTCAG TACACTATTT ATCTTACACA
   FAR  CTGACTCATG GAATAAATAA ACCTGGTCAG TACACTATTT ATCTTACACA
GLAno2  CTGACTCATG GAATAAATAA ACCTGGTCAG TACACTATTT ATCTTACACA
   GLE  CTGACTCATG GAATAAATAA ACCTGGTCAG TACACTATTT ATCTTACACA
JEZno1  CTGACTCATG GAATAAATAA ACCTGGTCAG TACACTATTT ATCTTACACA
   KOR  CTGACTCATG GAATAAATAA ACCTGGTCAG TACACTATTT ATCTTACACA
   KOY  CTGACTCATG GAATAAATAA ACCTGGTCAG TACACTATTT ATCTTACACA
   LIK  CTGACTCATG GAATAAATAA ACCTGGTCAG TACACTATTT ATCTTACACA
MARno2  CTGACTCATG GAATAAATAA ACCTGGTCAG TACACTATTT ATCTTACACA
   MAT  CTGACTCATG GAATAAATAA ACCTGGTCAG TACACTATTT ATCTTACACA
   MAX  CTGACTCATG GAATAAATAA ACCTGGTCAG TACACTATTT ATCTTACACA
   MEX  CTGACTCATG GAATAAATAA ACCTGGTCAG TACACTATTT ATCTTACACA
   MEY  CTGACTCATG GAATAAATAA ACCTGGTCAG TACACTATTT ATCTTACACA
   MON  CTGACTCATG GAATAAATAA ACCTGGTCAG TACACTATTT ATCTTACACA
   MOR  CTGACTCATG GAATAAATAA ACCTGGTCAG TACACTATTT ATCTTACACA
   OBO  CTGACTCATG GAATAAATAA ACCTGGTCAG TACACTATTT ATCTTACACA
   OMO  CTGACTCATG GAATAAATAA ACCTGGTCAG TACACTATTT ATCTTACACA
   ORI  CTGACTCATG GAATAAATAA ACCTGGTCAG TACACTATTT ATCTTACACA
PINBANK  CTGACTCATG GAATTAATAA ACCTGGTCAG TACACTATTT ATCTTACACA
PINKOR  CTGACTCATG GAATAAATAA ACCTGGTCAG TACACTATTT ATCTTACACA
   PUN  CTGACTCATG GAATAAATAA ACCTGGTCAG TACACTATTT ATCTTACACA
   PUR  CTGACTCATG GAATAAATAA ACCTGGTCAG TACACTATTT ATCTTACACA
   RET  CTGACTCATG GAATAAATAA ACCTGGTCAG TACACTATTT ATCTTACACA
   RUB  CTGACTCATG GAATAAATAA ACCTGGTCAG TACACTATTT ATCTTACACA
   SCH  CTGACTCATG GAATAAATAA ACCTGGTCAG TACACTATTT ATCTTACACA
SITno1  CTGACTCATG GAATAAATCA ACCTGGTCAG TACACTATTT ATCTTACACA
   SMI  CTGACTCATG GAATAAATAA ACCTGGTCAG TACACTATTT ATCTTACACA
   SPI  CTGACTCATG GAATAAATAA ACCTGGTCAG TACACTATTT ATCTTACACA
   TOR  CTGACTCATG GAATAAATAA ACCTGGTCAG TACACTATTT ATCTTACACA
   WIL  CTGACTCATG GAATAAATAA ACCTGGTCAG TACACTATTT ATCTTACACA

        [2001                                                2050]
ABIno1  TCTGTACTCG AGTACTTAGG ACTCATTGTG AGTGTATAAA TCCCTGATCT
   ALC  TCTGTACTCG AGTACTTAGG ACTCATTGTG AGTGTATAAA TCCCTGATCT
   ASP  TCTGTACTCG AGTACTTAGG ACTCATTGTG AGTGTATAAA TCCCTGATCT
```

```
   BRA  TCTGTACTCG AGTACTTAGG ACTCATTGTG AGTGTATAAA TCCCTGATCT
   BRE  TCTGTACTCG AGTACTTAGG ACTCATTGTG AGTGTATAAA TCCCTGATCT
   CHI  TCTGTACTCG AGTACTTAGG ACTCATTGTG AGTGTATAAA TCCCTGATCT
   CRA  TCTGTACTCG AGTACTTAGG ACTCATTGTG AGTGTATAAA TCCCTGATCT
ENGno3  TCTGTACTCG AGTACTTAGG ACTCATTGTG AGTGTATAAA TCCCTGATCT
   FAR  TCTGTACTCG AGTACTTAGG ACTCATTGTG AGTGTATAAA TCCCTGATCT
GLAno2  TCTGTACTCG AGTACTTAGG ACTCATTGTG AGTGTATAAA TCCCTGATCT
   GLE  TCTGTACTCG AGTACTTAGG ACTCATTGTG AGTGTATAAA TCCCTGATCT
JEZno1  TCTGTACTCG AGTACTTAGG ACTCATTGTG AGTGTATAAA TCCCTGATCT
   KOR  TCTGTACTCG AGTACTTAGG ACTCATTGTG AGTGTATAAA TCCCTGATCT
   KOY  TCTGTACTCG AGTACTTAGG ACTCATTGTG AGTGTATAAA TCCCTGATCT
   LIK  TCTGTACTCG AGTACTTAGG ACTCATTGTG AGTGTATAAA TCCCTGATCT
MARno2  TCTGTACTCG AGTACTTAGG ACTCATTGTG AGTGTATAAA TCCCTGATCT
   MAT  TCTGTACTCG AGTACTTAGG ACTCATTGTG AGTGTATAAA TCCCTGATCT
   MAX  TCTGTACTCG AGTACTTAGG ACTCATTGTG AGTGTATAAA TCCCTGATCT
   MEX  TCTGTACTCG AGTACTTAGG ACTCATTGTG AGTGTATAAA TCCCTGATCT
   MEY  TCTGTACTCG AGTACTTAGG ACTCATTGTG AGTGTATAAA TCCCTGATCT
   MON  TCTGTACTCG AGTACTTAGG ACTCATTGTG AGTGTATAAA TCCCTGATCT
   MOR  TCTGTACTCG AGTACTTAGG ACTCATTGTG AGTGTATAAA TCCCTGATCT
   OBO  TCTGTACTCG AGTACTTAGG ACTCATTGTG AGTGTATAAA TCCCTGATCT
   OMO  TCTGTACTCG AGTACTTAGG ACTCATTGTG AGTGTATAAA TCCCTGATCT
   ORI  TCTGTACTCG AGTACTTAGG ACTCATTGTG AGTGTATAAA TCCCTGATCT
PINBANK  TCTGTACTCG AGTATTTAGG ACTCATTGTG AGTGTATAAA TCCCTGATCT
PINKOR  TCTGTACTCG AGTACTTAGG ACTCATTGTG AGTGTATAAA TCCCTGATCT
   PUN  TCTGTACTCG AGTACTTAGG ACTCATTGTT AGTGTATAAA TCCCTGATCT
   PUR  TCTGTACTCG AGTACTTAGG ACTCATTGTG AGTGTATAAA TCCCTGATCT
   RET  TCTGTACTCG AGTACTTAGG ACTCATTGTG AGTGTATAAA TCCCTGATCT
   RUB  TCTGTACTCG AGTACTTAGG ACTCATTGTG AGTGTATAAA TCCCTGATCT
   SCH  TCTGTACTCG AGTACTTAGG ACTCATTGTG AGTGTATAAA TCCCTGATCT
SITno1  TCTGTACTCG AGTACTTAGG ACTCATTGTG AGTGTATAAA TCCCTGATCT
   SMI  TCTGTACTCG AGTACTTAGG ACTCATTGTG AGTGTATAAA TCCCTGATCT
   SPI  TCTGTACTCG AGTACTTAGG ACTCATTGTG AGTGTATAAA TCCCTGATCT
   TOR  TCTGTACTCG AGTACTTAGG ACTCATTGTG AGTGTATAAA TCCCTGATCT
   WIL  TCTGTACTCG AGTACTTAGG ACTCATTGTG AGTGTATAAA TCCCTGATCT

        [2051                                              2100]
ABIno1  ACTCAGGGAA AACCTCCCGA TATCGGGTAC TAAAATGCTC TTAACTTCTG
   ALC  ACTCAGGGAA AACCTCCCGA TATCGGGTAC TAAAATGCTC TTAACTTCTG
   ASP  ACTCAGGGAA AACCTCCCGA TATCGGGTAC TAAAATGCTC TTAACTTCTG
   BRA  ACTCAGGGAA AACCTCCCGA TATCGGGTAC TAAAATGCTC TTAACTTCTG
   BRE  ACTCAGGAAA AACCTCCCGA TATCGGGTAC TAAAATGCTC TTAACTTCTG
   CHI  ACTCAGGGAA AACCTCCCGA TATCGGGTAC TAAAATGCTC TTAACTTCTG
   CRA  ACTCAGGGAA AACCTCCCGA TATCGGGTAC TAAAATGCTC TTAACTTCTG
ENGno3  ACTCAGGGAA AACCTCCCGA TATCGGGTAC TAAAATGCTC TTAACTTCTG
   FAR  ACTCAGGGAA AACCTCCCGA TATCGGGTAC TAAAATGCTC TTAACTTCTG
GLAno2  ACTCAGGGAA AACCTCCCGA TATCGGGTAC TAAAATGCTC TTAACTTCTG
   GLE  ACTCAGGGAA AACCTCCCGA TATCGGGTAC TAAAATGCTC TTAACTTCTG
JEZno1  ACTCAGGGAA AACCTCCCGA TATCGGGTAC TAAAATGCTC TTAACTTCTG
   KOR  ACTCAGGGAA AACCTCCCGA TATCGGGTAC TAAAATGCTC TTAACTTCTG
   KOY  ACTCAGGGAA AACCTCCCGA TATCGGGTAC TAAAATGCTC TTAACTTCTG
   LIK  ACTCAGGGAA AACCTCCCGA TATCGGGTAC TAAAATGCTC TTAACTTCTG
MARno2  ACTCAGGGAA AACCTCCCGA TATCGGGTAC TAAAATGCTC TTAACTTCTG
   MAT  ACTCAGGGAA AACCTCCCGA TATCGGGTAC TAAAATGCTC TTAACTTCTG
   MAX  ACTCAGGGAA AACCTCCCGA TATCGGGTAC TAAAATGCTC TTAACTTCTG
   MEX  ACTCAGGGAA AACCTCCCGA TATCGGGTAC TAAAATGCTC TTAACTTCTG
   MEY  ACTCAGGGAA AACCTCCCGA TATCGGGTAC TAAAATGCTC TTAACTTCTG
   MON  ACTCAGGGAA AACCTCCCGA TATCGGGTAC TAAAATGCTC TTAACTTCTG
   MOR  ACTCAGGGAA AACCTCCCGA TATCGGGTAC TAAAATGCTC TTAACTTCTG
   OBO  ACTCAGGGAA AACCTCCCGA TATCGGGTAC TAAAATGCTC TTAACTTCTG
   OMO  ACTCAGGGAA AACCTCCCGA TATCGGGTAC TAAAATGCTC TTAACTTCTG
   ORI  ACTCAGGGAA AACCTCCCGA TATCGGGTAC TAAAATGCTC TTAACTTCTG
PINBANK  ACTCAGGGAA AACCTCCCGA TATCGGGTAC TAAACTGCTC TTAACTTCTG
PINKOR  ACTCAGGGAA AACCTCCCGA TATCAGGTAC TAAACTGCTC TTAACTTCTG
   PUN  ACTCAGGGAA AACCTCCCGA TATCGGGTAC TAAAATGCTC TTAACTTCTG
   PUR  ACTCAGGGAA AACCTCCCGA TATCGGGTAC TAAAATGCTC TTAACTTCTG
   RET  ACTCAGGGAA AACCTCCCGA TATCGGGTAC TAAAATGCTC TTAACTTCTG
   RUB  ACTCAGGGAA AACCTCCCGA TATCGGGTAC TAAAATGCTC TTAACTTCTG
   SCH  ACTCAGGGAA AACCTCCCGA TATCGGGTAC TAAAATGCTC TTAACTTCTG
SITno1  ACTCAGGAAA AACCTCCCGA TATCGGGTAC TAAAATGCTC TTAACTTCTG
   SMI  ACTCAGGGAA AACCTCCCGA TATCGGGTAC TAAAATGCTC TTAACTTCTG
   SPI  ACTCAGGGAA AACCTCCCGA TATCGGGTAC TAAAATGCTC TTAACTTCTG
   TOR  ACTCAGGGAA AACCTCCCGA TATCGGGTAC TAAAATGCTC TTAACTTCTG
   WIL  ACTCAGGGAA AACCTCCCGA TATCGGGTAC TAAAATGCTC TTAACTTCTG

        [2101                                              2150]
ABIno1  ATCAGTAARG GGAATTCCTC GCTCGTTTAA TTGGAACGAG GAACAGAAGC
   ALC  ATCAGTAAAG GGAATTCCTC GCTCGTTTAA TTGGAACGAG GAACAGAAGC
   ASP  ATCAGTAAAG GGAATTCCTC GCTCGTTTAA TTGGAACGAG GAACAGAAGC
   BRA  ATCAGTAAAG GGAATTCCTC GCTCGTTTAA TTGGAACGAG GAACAGAAGC
   BRE  ATCAGTAAAG GGAATTCCTC GCTCGTTTAA TTGGAACGAG GAACAGAAGC
   CHI  ATCAGTAAAG GGAATTCCTC GCTCGTTTAA TTGGAACGAG GAACAGAAGC
   CRA  ATCAGTAAAG GGAATTCCTC GCTCGTTTAA TTGGAACGAG GAACAGAAGC
ENGno3  ATCAGTAAAG GGAATTCCTC GCTCGTTTAA TTGGAACGAG GAACAGAAGC
   FAR  ATCAGTAAAG GGAATTCCTC GCTCGTTTAA TTGGAACGAG GAACAGAAGC
```

```
GLAno2  ATCAGTAAAG GGAATTCCTC GCTCGTTTAA TTGGAACGAG GAACAGAAGC
   GLE  ATCAGTAAAG GGAATTCCTC GCTCGTTTAA TTGGAACGAG GAACAGAAGC
JEZno1  ATCAGTAAAG GGAATTCCTC GCTCGTTTAA TTGGAACGAG GAACAGAAGC
   KOR  ATCAGTAAAG GGAATTCCTC GCTCGTTTAA TTGGAACGAG GAACAGAAGC
   KOY  ATCAGTAAAG GGAATTCCTC GCTCGTTTAA TTGGAACGAG GAACAGAAGC
   LIK  ATCAGTAAAG GGAATTCCTC GCTCGTTTAA TTGGAACGAG GAACAGAAGC
MARno2  ATCAGTAAAG GGAATTCCTC GCTCGTTTAA TTGGAACGAG GAACAGAAGC
   MAT  ATCAGTAAAG GGAATTCCTC GCTCGTTTAA TTGGAACGAG GAACAGAAGC
   MAX  ATCAGTAAAG GGAATTCCTC GCTCGTTTAA TTGGAACGAG GAACAGAAGC
   MEX  ATCAGTAAAG GGAATTCCTC GCTCGTTTAA TTGGAACGAG GAACAGAAGC
   MEY  ATCAGTAAAG GGAATTCCTC GCTCGTTTAA TTGGAACGAG GAACAGAAGC
   MON  ATCAGTAAAG GGAATTCCTC GCTCGTTTAA TTGGAACGAG GAACAGAAGC
   MOR  ATCAGTAAAG GGAATTCCTC GCTCGTTTAA TTGGAACGAG GAACAGAAGC
   OBO  ATCAGTAAGG GGAATTCCTC GCTCGTTTAA TTGGAACGAG GAACAGAAGC
   OMO  ATCAGTAAAG GGAATTCCTC GCTCGTTTAA TTGGAACGAG GAACAGAAGC
   ORI  ATCAGTAAAG GGAATTCCTC GCTCGTTTAA TTGGAACGAG GAACAGAAGC
PINBANK  ATCAGTAAAG GGAATTCCTC GCTCGCTTAA TTGGAACGAG GAACAGAAGC
PINKOR  ATCAGTAAAG GGAATTCCTC GCTCGTTAAA TTGGAACGAG GAACAGAAGC
   PUN  ATCAGTAAAG GGAATTCCTC GCTCGTTTAA TTGGAACGAG GAACAGAAGC
   PUR  ATCAGTAAAG GGAATTCCTC GCTCGTTTAA TTGGAACGAG GAACAGAAGC
   RET  ATCAGTAAAG GGAATTCCTC GCTCGTTTAA TTGGAACGAG GAACAGAAGC
   RUB  ATCAGTAAAG GGAATTCCTC GCTCGTTTAA TTGGAACGAG GAACAGAAGC
   SCH  ATCAGTAAAG GGAATTCCTC GCTCGTTTAA TTGGAACGAG GAACAGAAGC
SITno1  ATCAGTAAAG GGAATTCCTC GCTCGTTTAA TTGGAACGAG GAACAGAAGC
   SMI  ATCAGTAAAG GGAATTCCTC GCTCGTTTAA TTGGAACGAG GAACAGAAGC
   SPI  ATCAGTAAAG GGAATTCCTC GCTCGTTTAA TTGGAACGAG GAACAGAAGC
   TOR  ATCAGTAAAG GGAATTCCTC GCTCGTTTAA TTGGAACGAG GAACAGAAGC
   WIL  ATCAGTAAAG GGAATTCCTC GCTCGTTTAA TTGGAACGAG GAACAGAAGC

        [2151                                              2200]
ABIno1  TCGTTTCTCT GGGTCTACTT ATGATTCAAG TCATATAGCT TTCTCCATTG
   ALC  TCGTTTCTCT GGGTCTACTT ATGATTCAAG TCATATAGCT TTCTCCATTG
   ASP  TCGTTTCTCT GGGTCTACTT ATGATTCAAG TCATATAGCT TTCTCCATTG
   BRA  TCGTTTCTCT GGGTCTACTT ATGATTCAAG TCATATAGCT TTCTCCATTG
   BRE  TCGTTTCTCT GGGTCTACTT ATGATTCAAG TCATATAGCT TTCTCCATTG
   CHI  TCGTTTCTCT GGGTCTACTT ATGATTCAAG TCATATAGCT TTCTCCATTT
   CRA  TCGTTTCTCT GGGTCTACTT ATGATTCAAG TCATATAGCT TTCTCCATTG
ENGno3  TCGTTTCTCT GGGTCTACTT ATGATTCAAG TCATATAGCT TTCTCCATTG
   FAR  TCGTTTCTCT GGGTCTACTT ATGATTCAAG TCATATAGCT TTATCCATTG
GLAno2  TCGTTTCTCT GGGTCTACTT ATGATTCAAG TCATATAGCT TTCTCCATTG
   GLE  TCGTTTCTCT GGGTCTACTT ATGATTCAAG TCATATAGCT TTCTCCATTG
JEZno1  TCGTTTCTCT GGGTCTACTT ATGATTCAAG TCATATAGCT TTCTCCATTG
   KOR  TCGTTTCTCT GGGTCTACTT ATGATTCAAG TCATATAGCT TTCTCCATTG
   KOY  TCGTTTCTCT GGGTCTACTT ATGATTCAAG TCATATAGCT TTCTCCATTG
   LIK  TCGTTTCTCT GGGTCTACTT ATGATTCAAG TCATATAGCT TTATCCATTG
MARno2  TCGTTTCTCT GGGTCTACTT ATGATTCAAG TCATATAGCT TTCTCCATTG
   MAT  TCGTTTCTCT GGGTCTACTT ATGATTCAAG TCATATAGCT TTCTCCATTG
   MAX  TCGTTTCTCT GGGTCTACTT ATGATTCAAG TCATATAGCT TTCTCCATTG
   MEX  TCGTTTCTCT GGGTCTACTT ATGATTCAAG TCATATAGCT TTCTCCATTG
   MEY  TCGTTTCTCT GGGTCTACTT ATGATTCAAG TCATATAGCT TTCTCCATTG
   MON  TCGTTTCTCT GGGTCTACTT ATGATTCAAG TCATATAGCT TTCTCCATTG
   MOR  TCGTTTCTCT GGGTCTACTT ATGATTCAAG TCATATAGCT TTCTCCATTG
   OBO  TCGTTTCTCT GGGTCTACTT ATGATTCAAG TCATATAGCT TTCTCCATTG
   OMO  TCGTTTCTCT GGGTCTACTT ATGATTCAAG TCATATAGCT TTCTCCATTG
   ORI  TCGTTTCTCT GGGTCTACTT ATGATTCAAG TCATATAGCT TTCTCCATTG
PINBANK  TCGTTTCTCT GGGTCTACTT ATGATTCAAG TCATATAGCT TTCTCCATTG
PINKOR  TCATTTCTCT GGGTCTACTT ATGATTCAAG TCATATAGCT TTCTCCATTG
   PUN  TCGTTTCTCT GGGTCTACTT ATGATTCAAG TCATATAGCT TTCTCCATTG
   PUR  TCGTTTCTCT GGGTCTACTT ATGATTCAAG TCATATAGCT TTCTCCATTG
   RET  TCGTTTCTCT GGGTCTACTT ATGATTCAAG TCATATAGCT TTCTCCATTG
   RUB  TCGTTTCTCT GGGTCTACTT ATGATTCAAG TCATATAGCT TTCTCCATTG
   SCH  TCGTTTCTCT GGGTCTACTT ATGATTCAAG TCATATAGCT TTATCCATTG
SITno1  TCGTTTCTCT GGGTCTACTT ATGATTCAAG TCATATAGCT TTCTCCATTG
   SMI  TCGTTTCTCT GGGTCTACTT ATGATTCAAG TCATATAGCT TTATCCATTG
   SPI  TCGTTTCTCT GGGTCTACTT ATGATTCAAG TCATATAGCT TTATCCATTG
   TOR  TCGTTTCTCT GGGTCTACTT ATGATTCAAG TCATATAGCT TTCTCCATTG
   WIL  TCGTTTCTCT GGGTCTACTT ATGATTCAAG TCATATAGCT TTCTCCATTG

        [2201                                              2250]
ABIno1  ACTCATTCGA CTTAACCGCG AATTGATTCG ATCCTATTCA CAATTATCAC
   ALC  ACTCATTCGA CTTAACCGCG AATTGATTCG ATCCTATTCA CAATTATCAC
   ASP  ACTCATTCGA CTTAACCGCG AATTGATTCG ATCCTATTCA CAATTATCAC
   BRA  ACTCATTCGA CTTAACCGCG AATTGATTCG ATCCTATTCA CAATTATCAC
   BRE  ACTCATTCGA CTTAACCGCG AATTGATTCG ATCCTATTCA CAATTATCAC
   CHI  ACTCATTCGA CTTAACCGCG AATTGATTCG ATCCTATTCA CAATTATCAC
   CRA  ACTCATTCGA CTTAACCGCG AATTGATTCG ATCCTATTCA CAATTATCAC
ENGno3  ACTCATTCGA CTTAACCGCG AATTGATTCG ATCCTATTCA CAATTATCAC
   FAR  ACTCATTCGA CTTAACCGCG AATTGATTCG ATCCTATTCA CAATTATCAC
GLAno2  ACTCATTCGA CTTAACCGCG AATTGATTCG ATCCTATTCA CAATTATCAC
   GLE  ACTCATTCGA CTTAACCGCG AATTGATTCG ATCCTATTCA CAATTATCAC
JEZno1  ACTCATTCGA CTTAACCGCG AATTGATTCG ATCCTATTCA CAATTATCAC
   KOR  ACTCATTCGA CTTAACCGCG AATTGATTCG ATCCTATTCA CAATTATCAC
   KOY  ACTCATTCGA CTTAACCGCG AATTGATTCG ATCCTATTCA CAATTATCAC
   LIK  ACTCATTCGA CTTAACCGCG AATTGATTCG ATCCTATTCA CAATTATCAC
```

```
MARno2  ACTCATTCGA CTTAACCGCG AATTGATTCG ATCCTATTCA CAATTATCAC
   MAT  ACTCATTCGA CTTAACCGCG AATTGATTCG ATCCTATTCA CAATTATCAC
   MAX  ACTCATTCGA CTTAACCGCG AATTGATTCG ATCCTATTCA CAATTATCAC
   MEX  ACTCATTCGA CTTAACCGCG AATTGATTCG ATCCTATTCA CAATTATCAC
   MEY  ACTCATTCGA CTTAACCGCG AATTGATTCG ATCCTATTCA CAATTATCAC
   MON  ACTCATTCGA CTTAACCGCG AATTGATTCG ATCCTATTCA CAATTATCAC
   MOR  ACTCATTCGA CTTAACCGCG AATTGATTCG ATCCTATTCA CAATTATCAC
   OBO  ACTCATTCGA CTTAACCGCG AATTGATTCG ATCCTATTCA CAATTATCAC
   OMO  ACTCATTCGA CTTAACCGCG AATTGATTCG ATCCTATTCA CAATTATCAC
   ORI  ACTCATTCGA CTTAACCGCG AATTGATTCG ATCCTATTCA CAATTATCAC
PINBANK ACTCATTCGA CTTAACCGCG AATTGATTCA ATCCTATTCA CAATTATCAC
 PINKOR ACTCATTCGA CTTAACCGCG AATTGATGAA ATCCTATTCA CAATTATCAC
   PUN  ACTCATTCGA CTTAACCGCG AATTGATTCG ATCCTATTCA CAATTATCAC
   PUR  ACTCATTCGA CTTAACCGCG AATTGATTCG ATCCTATTCA CAATTATCAC
   RET  ACTCATTCGA CTTAACCGCG AATTGATTCG ATCCTATTCA CAATTATCAC
   RUB  ACTCATTCGA CTTAACCGCG AATTGATTCG ATCCTATTCA CAATTATCAC
   SCH  ACTCATTCGA CTTAACCGCG AATTGATTCG ATCCTATTCA CAATTATCAC
SITno1  ACTCATTCGA CTTAACCGCG AATTGATTCG ATCCTATTCA CAATTATCAC
   SMI  ACTCATTCGA CTTAACCGCG AATTGATTCG ATCCTATTCA CAATTATCAC
   SPI  ACTCATTCGA CTTAACCGCG AATTGATTCG ATCCTATTCA CAATTATCAC
   TOR  ACTCATTCGA CTTAACCGCG AATTGATTCG ATCCTATTCA CAATTATCAC
   WIL  ACTCATTCGA CTTAACCGCG AATTGATTCG ATCCTATTCA CAATTATCAC

        [2251                                              2300]
ABIno1  ATTTGATCCG AAAGGATGAT CATATTATAC ATCCACCTAG GTATATAATA
   ALC  ATTTGATCCG AAAGGATGAT CATATTATAC ATCCACCTAG GTATATAATA
   ASP  ATTTGATCCG AAAGGATGAT CATATTATAC ATCCACCTAG GTATATAATA
   BRA  ATTTGATCCG AAAGGATGAT CATATTATAC ATCCACCTAG GTATATAATA
   BRE  ATTTGATCCG AAAGGATGAT CATATTATAC ATCCACCTAG GTATATAATA
   CHI  ATTTGATCCG AAAGGATGAT CATATTATAC ATCCACCTAG GTATATAATA
   CRA  ATTTGATCCG AAAGGATGAT CATATTATAC ATCCACCTAG GTATATAATA
ENGno3  ATTTGATCCG AAAGGATGAT CATATTATAC ATCCACCTAG GTATATAATA
   FAR  ATTTGATCCG AAAGGATGAT CATATTATAC ATCCACCTAG GTATATAATA
GLAno2  ATTTGATCCG AAAGGATGAT CATATTATAC ATCCACCTAG GTATATAATA
   GLE  ATTTGATCCG AAAGGATGAT CATATTATAC ATCCACCTAG GTATATAATA
JEZno1  ATTTGATCCG AAAGGATGAT CATATTATAC ATCCACCTAG GTATATAATA
   KOR  ATTTGATCCG AAAGGATGAT CATATTATAC ATCCACCTAG GTATATAATA
   KOY  ATTTGATCCG AAAGGATGAT CATATTATAC ATCCACCTAG GTATATAATA
   LIK  ATTTGATCCG AAAGGATGAT CATATTATAC ATCCACCTAG GTATATAATA
MARno2  ATTTGATCCG AAAGGATGAT CATATTATAC ATCCACCTAG GTATATAATA
   MAT  ATTTGATCCG AAAGGATGAT CATATTATAC ATCCACCTAG GTATATAATA
   MAX  ATTTGATCCG AAAGGATGAT CATATTATAC ATCCACCTAG GTATATAATA
   MEX  ATTTGATCCG AAAGGATGAT CATATTATAC ATCCACCTAG GTATATAATA
   MEY  ATTTGATCCG AAAGGATGAT CATATTATAC ATCCACCTAG GTATATAATA
   MON  ATTTGATCCG AAAGGATGAT CATATTATAC ATCCACCTAG GTATATAATA
   MOR  ATTTGATCCG AAAGGATGAT CATATTATAC ATCCACCTAG GTATATAATA
   OBO  ATTTGATCCG AAAGGATGAT CATATTATAC ATCCACCTAG GTATATAATA
   OMO  ATTTGATCCG AAAGGATGAT CATATTATAC ATCCACCTAG GTATATAATA
   ORI  ATTTGATCCG AAAGGATGAT CATATTATAC ATCCACCTAG GTATATAATA
PINBANK ATTTGATCCG AAAGGATGAG CATATTATAC ATCCATCTAG GTATATAATA
 PINKOR ATTTGATCCG AAAGGATGAG CATATTATAC ATCCATCCAG GTATATAATA
   PUN  ATTTGATCCG AAAGGATGAT CATATTATAC ATCCACCTAG GTATATAATA
   PUR  ATTTGATCCG AAAGGATGAT CATATTATAC ATCCACCTAG GTATATAATA
   RET  ATTTGATCCG AAAGGATGAT CATATTATAC ATCCACCTAG GTATATAATA
   RUB  ATTTGATCCG AAAGGATGAT CATATTATAC ATCCACCTAG GTATATAATA
   SCH  ATTTGATCCG AAAGGATGAT CATATTATAC ATCCACCTAG GTATATAATA
SITno1  ATTTGATCCG AAAGGATGAT CATATTATAC ATCCACCTAG GTATATAATA
   SMI  ATTTGATCCG AAAGGATGAT CATATTATAC ATCCACCTAG GTATATAATA
   SPI  ATTTGATCCG AAAGGATGAT CATATTATAC ATCCACCTAG GTATATAATA
   TOR  ATTTGATCCG AAAGGATGAT CATATTATAC ATCCACCTAG GTATATAATA
   WIL  ATTTGATCCG AAAGGATGAT CATATTATAC ATCCACCTAG GTATATAATA

        [2301                                              2350]
ABIno1  GACATTTCCC ACCCATTTGA TTCCTTGAAT TAGATCCGGT CCTGATCGAA
   ALC  GACATTTCCC ACCCATTTGA TTCCTTGAAT TAGATCCGGT CCTGATCGAA
   ASP  GACATTTCCC ACCCATTTGA TTCCTTGAAT TAGATCCGGT CCTGATCGAA
   BRA  GACATTTCCC ACCCATTTGA TTCCTTGAAT TAGATCCGGT CCTGATCGAA
   BRE  GACATTTCCC ACCCATTTGA TTCCTTGAAT CAGATCCGGT CCTGATCGAA
   CHI  GACATTTCCC ACCCATTTGA TTCCTTGAAT TAGATCCGGT CCTGATCGAA
   CRA  GACATTTCCC ACCCATTTGA TTCCTTGAAT TAGATCCGGT CCTGATCGAA
ENGno3  GACATTTCCC ACCCATTTGA TTCCTTGAAT TAGATCCGGT CCTGATCGAA
   FAR  GACATTTCCC ACCCATTTGA TTCCTTGAAT TAGATCCGGT CCTGATCGAA
GLAno2  GACATTTCCC ACCCATTTGA TTCCTTGAAT TAGATCCGGT CCTGATCGAA
   GLE  GACATTTCCC ACCCATTTGA TTCCTTGAAT TAGATCCGGT CCTGATCGAA
JEZno1  GACATTTCCC ACCCATTTGA TTCCTTGAAT TAGATCCGGT CCTGATCGAA
   KOR  GACATTTCCC ACCCATTTGA TTCCTTGAAT TAGATCCGGT CCTGATCGAA
   KOY  GACATTTCCC ACCCATTTGA TTCCTTGAAT TAGATCCGGT CCTGATCGAA
   LIK  GACATTTCCC ACCCATTTGA TTCCTTGAAT TAGATCCGGT CCTGATCGAA
MARno2  GACATTTCCC ACCCATTTGA TTCCTTGAAT TAGATCCGGT CCTGATCGAA
   MAT  GACATTTCCC ACCCATTTGA TTCCTTGAAT TAGATCCGGT CCTGATCGAA
   MAX  GACATTTCCC ACCCATTTGA TTCCTTGAAT TAGATCCGGT CCTGATCGAA
   MEX  GACATTTCCC ACCCATTTGA TTCCTTGAAT TAGATCCGGT CCTGATCGAA
   MEY  GACATTTCCC ACCCATTTGA TTCCTTGAAT TAGATCCGGT CCTGATCGAA
   MON  GACATTTCCC ACCCATTTGA TTCCTTGAAT TAGATCCGGT CCTGATCGAA
```

```
   MOR  GACATTTCCC ACCCATTTGA TTCCTTGAAT TAGATCCGGT CCTGATCGAA
   OBO  GACATTTCCC ACCCATTTGA TTCCTTGAAT TAGATCCGGT CCTGATCGAA
   OMO  GACATTTCCC ACCCATTTGA TTCCTTGAAT TAGATCCGGT CCTGATCGAA
   ORI  GACATTTCCC ACCCATTTGA TTCCTTGAAT TAGATCCGGT CCTGATCGAA
PINBANK GACATTTCCC ACCCATTTTA TTCCTTAAAT AAGATCCGGT CCTGATCGAA
 PINKOR GACATTTTCC ACCCATTTGA TTCCTTGAAT CAGATCCGGT CCTGATCGAA
   PUN  GACATTTCCC ACCCATTTGA TTCCTTGAAT TAGATCCGGT CCTGATCGAA
   PUR  GACATTTCCC ACCCATTTGA TTCCTTGAAT TAGATCCGGT CCTGATCGAA
   RET  GACATTTCCC ACCCATTTGA TTCCTTGAAT TAGATCCGGT CCTGATCGAA
   RUB  GACATTTCCC ACCCATTTGA TTCCTTGAAT TAGATCCGGT CCTGATCGAA
   SCH  GACATTTCCC ACCCATTTGA TTCCTTGAAT TAGATCCGGT CCTGATCGAA
 SITno1 GACATTTCCC ACCCATTTGA TTCCTTGAAT CAGATCCGGT CCTGATCGAA
   SMI  GACATTTCCC ACCCATTTGA TTCCTTGAAT TAGATCCGGT CCTGATCGAA
   SPI  GACATTTCCC ACCCATTTGA TTCCTTGAAT TAGATCCGGT CCTGATCGAA
   TOR  GACATTTCCC ACCCATTTGA TTCCTTGAAT TAGATCCGGT CCTGATCGAA
   WIL  GACATTTCCC ACCCATTTGA TTCCTTGAAT TAGATCCGGT CCTGATCGAA

        [2351                                              2400]
 ABIno1 TCATATCAGG TATCCTAAAA ATAATATCAG GTATCATAAA GATCATATGT
   ALC  TACAATCAGG TATCCTAAAA ATCATATCAG GTATCATAAA GATCATATGT
   ASP  TCATATCAGG TATCCTAAAA AT?ATATCAG GTATCATAAA GATCATATGT
   BRA  TACAATCAGG TATCCTAAAA ATCATATCAG GTATCATAAA GATCATATGT
   BRE  TACAATCAGG TATCCTAAAA ATCATATCAG GTATCATAAA GATCATATGT
   CHI  TACAATCAGG TATCCTAAAA ATCATATCAG GTATCATAAA GATCATATGT
   CRA  TCATATCAGG TATCCTAAAA AT?ATATCAG GTATCATAAA GATCATATGT
 ENGno3 TA?TATCAGG TATCCTAAAA ATCATATCAG GTATCATAAA GATCATATGT
   FAR  TACAATCAGG TATCCTAAAA ATCATATCAG GTATCATAAA GATCATATGT
 GLAno2 TAATATCAGG TATCCTAAAA ATCATATCAG GTATCATAAA GATCATATGT
   GLE  TACAATCAGG TATCCTAAAA ATAATATCAG GTATCATAAA GATCATATGT
 JEZno1 TACAATCAGG TATCCTAAAA ATAATATCAG GTATCATAAA GATCATATGT
   KOR  TCATATCAGG TATCCTAAAA AT?ATATCAG GTATCATAAA GATCATATGT
   KOY  TCATATCAGG TATCCTAAAA ATAATATCAG GTATCATAAA GATCATATGT
   LIK  TACAATCAGG TATCCTAAAA ATCATATCAG GTATCATAAA GATCATATGT
 MARno2 TACAATCAGG TATCCTAAAA ATAATATCAG GTATCATAAA GATCATATGT
   MAT  TACAATCAGG TATCCTAAAA ATCATATCAG GTATCATAAA GATCATATGT
   MAX  TACAATCAGG TATCCTAAAA ATCATATCAG GTATCATAAA GATCATATGT
   MEX  TAATATCAGG TATCCTAAAA ATCATATCAG GTATCATAAA GATCATATGT
   MEY  TCATATCAGG TATCCTAAAA ATAATATCAG GTATCATAAA GATCATATGT
   MON  TCATATCAGG TATCCTAAAA AT?ATATCAG GTATCATAAA GATCATATGT
   MOR  TACAATCAGG TATCCTAAAA ATCATATCAG GTATCATAAA GATCATATGT
   OBO  TCATATCAGG TATCCTAAAA ATAATATCAG GTATCATAAA GATCATATGT
   OMO  TACAATCAGG TATCCTAAAA ATAATATCAG GTATCATAAA GATCATATGT
   ORI  TACAATCAGG TATCCTAAAA ATCATATCAG GTATCATAAA GATCATATGT
PINBANK TACAATCAGG TATCCGAGAA ATAATATCAG GTATAATAAA GATCATATGT
 PINKOR TACAATCAGG TATCCGAGAA ATAATATCAG GTATCATAAA GATCATATGT
   PUN  TACAATCAGG TATCCTAAAA ATCATATCAG GTATCATAAA GATCATATGT
   PUR  TACAATCAGG TATCCTAAAA ATCATATCAG GTATCATAAA GATCATATGT
   RET  TCATATCAGG TATCCTAAAA AT?ATATCAG GTATCATAAA GATCATATGT
   RUB  TACAATCAGG TATCCTAAAA ATAATATCAG GTATCATAAA GATCATATGT
   SCH  TACAATCAGG TATCCTAAAA ATCATATCAG GTATCATAAA GATCATATGT
 SITno1 TACAATCAGG TATCCTAAAA ATCATATCAG GTATCATAAA GATCATATGT
   SMI  TACAATCAGG TATCCTAAAA ATCATATCAG GTATCATAAA GATCATATGT
   SPI  TACAATCAGG TATCCTAAAA ATCATATCAG GTATCATAAA GATCATATGT
   TOR  TCATATCAGG TATCCTAAAA AT?ATATCAG GTATCATAAA GATCATATGT
   WIL  TACAATCAGG TATCCTAAAA ATCATATCAG GTATCATAAA GATCATATGT

        [2401                                    +1        2450]
 ABIno1 TGGAACGACA TGCTGTTTTT TCCGTTAATT CTACTTCGA0G GATCAGTCGT
   ALC  TGGAACGACA TGCTGTTTTT TCCGTTAATT CTACTTCGA0G GATCAGTCGT
   ASP  TGGAACGACA TGCTGTTTTT TCCGTTAATT CTACTTCGA0G GATCAGTCGT
   BRA  TGGAACGACA TGCTGTTTTT TCCGTTAATT CTACTTCGA0G GATCAGTCGT
   BRE  TGGAACGACA TGCTGTTTTT TCCGTTAATT CTACTTCGA0G GATCAGTCGT
   CHI  TGGAACGACA TGCTGTTTTT TCCGTTAATT CTACTTCGA0G GATCAGTCGT
   CRA  TGGAACGACA TGCTGTTTTT TCCGTTAATT CTACTTCGA0G GATCAGTCGT
 ENGno3 TGGAACGACA TGCTGTTTTT TCCGTTAATT CTACTTCGA0G GATCAGTCGT
   FAR  TGGAACGACA TGCTGTTTTT TCCGTTAATT CTACTTCGA0G GATCAGTCGT
 GLAno2 TGGAACGACA TGCTGTTTTT TCCGTTAATT CTACTTCGA0G GATCAGTCGT
   GLE  TGGAACGACA TGCTGTTTTT TCCGTTAATT CTACTTCGA0G GATCAGTCGT
 JEZno1 TGGAACGACA TGCTGTTTTT TCCGTTAATT CTACTTCGA0G GATCAGTCGT
   KOR  TGGAACGACA TGCTGTTTTT TCCGTTAATT CTACTTCGA0G GATCAGTCGT
   KOY  TGGAACGACA TGCTGTTTTT TCCGTTAATT CTACTTCGA0G GATCAGTCGT
   LIK  TGGAACGACA TGCTGTTTTT TCCGTTAATT CTACTTCGA0G GATCAGTCGT
 MARno2 TGGAACGACA TGCTGTTTTT TCCGTTAATT CTACTTCGA0G GATCAGTCGT
   MAT  TGGAACGACA TGCTGTTTTT TCCGTTAATT CTACTTCGA0G GATCAGTCGT
   MAX  TGGAACGACA TGCTGTTTTT TCCGTTAATT CTACTTCGA0G GATCAGTCGT
   MEX  TGGAACGACA TGCTGTTTTT TCCGTTAATT CTACTTCGA0G GATCAGTCGT
   MEY  TGGAACGACA TGCTGTTTTT TCCGTTAATT CTACTTCGA0G GATCAGTCGT
   MON  TGGAACGACA TGCTGTTTTT TCCGTTAATT CTACTTCGA0G GATCAGTCGT
   MOR  TGGAACGACA TGCTGTTTTT TCCGTTAATT CTACTTCGA0G GATCAGTCGT
   OBO  TGGAACGACA TGCTGTTTTT TCCGTTAATT CTACTTCGA0G GATCAGTCGT
   OMO  TGGAACGACA TGCTGTTTTT TCCGTTAATT CTACTTCGA0G GATCAGTCGT
   ORI  TGGAACGACA TGCTGTTTTT TCCGTTAATT CTACTTCGA0G GATCAGTCGT
PINBANK TGGAACGACA TGCTTTTTTT TCCGTTCATT ATACTTCGA0G GATCAGTCGT
 PINKOR TGGAACGACA TGCTTTTTTT TCCGTTCATT A........1G GATCAGTCGT
```

```
   PUN  TGGAACGACA TGCTGTTTTT TCCGTTAATT CTACTTCGA0G GATCAGTCGT
   PUR  TGGAACGACA TGCTGTTTTT TCCGTTAATT CTACTTCGA0G GATCAGTCGT
   RET  TGGAACGACA TGCTGTTTTT TCCGTTAATT CTACTTCGA0G GATCAGTCGT
   RUB  TGGAACGACA TGCTGTTTTT TCCGTTAATT CTACTTCGA0G GATCAGTCGT
   SCH  TGGAACGACA TGCTGTTTTT TCCGTTAATT CTACTTCGA0G GATCAGTCGT
SITno1  TGGAACGACA TGCTGTTTTT TCCGTTAATT CTACTTCGA0G GATCAGTCGT
   SMI  TGGAACGACA TGCTGTTTTT TCCGTTAATT CTACTTCGA0G GATCAGTCGT
   SPI  TGGAACGACA TGCTGTTTTT TCCGTTAATT CTACTTCGA0G GATCAGTCGT
   TOR  TGGAACGACA TGCTGTTTTT TCCGTTAATT CTACTTCGA0G GATCAGTCGT
   WIL  TGGAACGACA TGCTGTTTTT TCCGTTAATT CTACTTCGA0G GATCAGTCGT

        [2451                                              2500]
ABIno1  AGTCTTCCAA ACTTTACCGA TGGTATCGAC GAGTTTTCTA CTTCATTCAA
   ALC  AGTCTTCCAA ACTTTACCGA TGGTATCGAC GAGTTTTCTA CTTCATTCAA
   ASP  AGTCTTCCAA ACTTTACCGA TGGTATCGAC GAGTTTTCTA CTTCATTCAA
   BRA  AGTCTTCCAA ACTTTACCGA TGGTATCGAC GAGTTTTCTA CTTCATTCAA
   BRE  AGTCTTCCGA ACTTTACCGA TGGTATCGAC GAGTTTTCTA CTTCATTCAA
   CHI  AGTCTTCCAA ACTTTACCGA TGGTATCGAC GAGTTTTCTA CTTCATTCAA
   CRA  AGTCTTCCAA ACTTTACCGA TGGTATCGAC GAGTTTTCTA CTTCATTCAA
ENGno3  AGTCTTCCAA ACTTTACCGA TGGTATCGAC GAGTTTTCTA CTTCATTCAA
   FAR  AGTCTTCCAA ACTTTACCGA TGGTATCGAC GAGTTTTCTA CTTCATTCAA
GLAno2  AGTCTTCCAA ACTTTACCGA TGGTATCGAC GAGTTTTCTA CTTCATTCAA
   GLE  AGTCTTCCAA ACTTTACCGA TGGTATCGAC GAGTTTTCTA CTTCATTCAA
JEZno1  AGTCTTCCAA ACTTTACCGA TGGTATCGAC GAGTTTTCTA CTTCATTCAA
   KOR  AGTCTTCCAA ACTTTACCGA TGGTATCGAC GAGTTTTCTA CTTCATTCAA
   KOY  AGTCTTCCAA ACTTTACCGA TGGTATCGAC GAGTTTTCTA CTTCATTCAA
   LIK  AGTCTTCCAA ACTTTACCGA TGGTATCGAC GAGTTTTCTA CTTCATTCAA
MARno2  AGTCTTCCAA ACTTTACCGA TGGTATCGAC GAGTTTTCTA CTTCATTCAA
   MAT  AGTCTTCCAA ACTTTACCGA TGGTATCGAC GAGTTTTCTA CTTCATTCAA
   MAX  AGTCTTCCAA ACTTTACCGA TGGTATCGAC GAGTTTTCTA CTTCATTCAA
   MEX  AGTCTTCCAA ACTTTACCGA TGGTATCGAC GAGTTTTCTA CTTCATTCAA
   MEY  AGTCTTCCAA ACTTTACCGA TGGTATCGAC GAGTTTTCTA CTTCATTCAA
   MON  AGTCTTCCAA ACTTTACCGA TGGTATCGAC GAGTTTTCTA CTTCATTCAA
   MOR  AGTCTTCCAA ACTTTACCGA TGGTATCGAC GAGTTTTCTA CTTCATTCAA
   OBO  AGTCTTCCAA ACTTTACCGA TGGTATCGAC GAGTTTTCTA CTTCATTCAA
   OMO  AGTCTTCCAA ACTTTACCGA TGGTATCGAC GAGTTTTCTA CTTCATTCAA
   ORI  AGTCTTCCAA ACTTTACCGA TGGTATCGAC GAGTTTTCTA CTTCATTCAA
PINBANK  AGTCTTCCGA ACTTTACCGA TGGTATCGAC GAGTTTTCTA CTTCATTCAA
PINKOR  AGTCTTCCGA ACTTTACCGA TGGTATCGAC GAGTTTTCTA CTTCATTCAA
   PUN  AGTCTTCCAA ACTTTACCGA TGGTATCGAC GAGTTTTCTA CTTCATTCAA
   PUR  AGTCTTCCAA ACTTTACCGA TGGTATCGAC GAGTTTTCTA CTTCATTCAA
   RET  AGTCTTCCAA ACTTTACCGA TGGTATCGAC GAGTTTTCTA CTTCATTCAA
   RUB  AGTCTTCCAA ACTTTACCGA TGGTATCGAC GAGTTTTCTA CTTCATTCAA
   SCH  AGTCTTCCAA ACTTTACCGA TGGTATCGAC GAGTTTTCTA CTTCATTCAA
SITno1  AGTCTTCCGA ACTTTACCGA TGGTATCGAC GAGTTTTCTA CTTCATTCAA
   SMI  AGTCTTCCAA ACTTTACCGA TGGTATCGAC GAGTTTTCTA CTTCATTCAA
   SPI  AGTCTTCCAA ACTTTACCGA TGGTATCGAC GAGTTTTCTA CTTCATTCAA
   TOR  AGTCTTCCAA ACTTTACCGA TGGTATCGAC GAGTTTTCTA CTTCATTCAA
   WIL  AGTCTTCCAA ACTTTACCGA TGGTATCGAC GAGTTTTCTA CTTCATTCAA

        [2501                2528 car 1 site enlevé; fin à 3887 dans la matrice]
ABIno1  ATGTGTAAAA GACCTTAGTC GCACTTAAA
   ALC  ATGTGTAAAA GACCTTAGTC GCACTTAAA
   ASP  ATGTGTAAAA GACCTTAGTC GCACTTAAA
   BRA  ATGTGTAAAA GACCTTAGTC GCACTTAAA
   BRE  ATGTGTAAAA GACCTTAGTC GCACTTAAA
   CHI  ATGTGTAAAA GACCTTAGTC GCACTTAAA
   CRA  ATGTGTAAAA GACCTTAGTC GCACTTAAA
ENGno3  ATGTGTAAAA GACCTTAGTC GCACTTAAA
   FAR  ATGTGTAAAA GACCTTAGTC GCACTTAAA
GLAno2  ATGTGTAAAA GACCTTAGTC GCACTTAAA
   GLE  ATGTGTAAAA GACCTTAGTC GCACTTAAA
JEZno1  ATGTGTAAAA GACCTTAGTC GCACTTAAA
   KOR  ATGTGTAAAA GACCTTAGTC GCACTTAAA
   KOY  ATGTGTAAAA GACCTTAGTC GCACTTAAA
   LIK  ATGTGTAAAA GACCTTAGTC GCACTTAAA
MARno2  ATGTGTAAAA GACCTTAGTC GCACTTAAA
   MAT  ATGTGTAAAA GACCTTAGTC GCACTTAAA
   MAX  ATGTGTAAAA GACCTTAGTC GCACTTAAA
   MEX  ATGTGTAAAA GACCTTAGTC GCACTTAAA
   MEY  ATGTGTAAAA GACCTTAGTC GCACTTAAA
   MON  ATGTGTAAAA GACCTTAGTC GCACTTAAA
   MOR  ATGTGTAAAA GACCTTAGTC GCACTTAAA
   OBO  ATGTGTAAAA GACCTTAGTC GCACTTAAA
   OMO  ATGTGTAAAA GACCTTAGTC GCACTTAAA
   ORI  ATGTGTAAAA GACCTTAGTC GCACTTAAA
PINBANK  ATGTGTAAAA GATCTTAGTC GCACTTAAA
PINKOR  ATGTGTAAAA GATCTTAGTC GCACTTAAA
   PUN  ATGTGTAAAA GACCTTAGTC GCACTTAAA
   PUR  ATGTGTAAAA GACCTTAGTC GCACTTAAA
   RET  ATGTGTAAAA GACCTTAGTC GCACTTAAA
   RUB  ATGTGTAAAA GACCTTAGTC GCACTTAAA
   SCH  ATGTGTAAAA GACCTTAGTC GCACTTAAA
SITno1  ATGTGTAAAA GACTTTAGTC GCACTTAAA
```

```
   SMI  ATGTGTAAAA GACCTTAGTC GCACTTAA?
   SPI  ATGTGTAAAA GACCTTAGTC GCACTTAAA
   TOR  ATGTGTAAAA GACCTTAGTC GCACTTAAA
   WIL  ATGTGTAAAA GACCTTAGTC GCACTTAAA
```

[Ci-dessous: Matrice trnTLF avec indels (0, 1) en plus. 26 indels. Donc, ntax=50 et nchar=1519+27= 1546]
[Début à 3888 et fin à 5433]

```
          [1                                                  50]
 ABIno1   CACGAGGATT TTCAGTCCTC TGCTCTACCA ACTGAGCTAT CCCGGCTCTT
    ALC   CACGAGGATT TTCAGTCCTC TGCTCTACCA ACTGAGCTAT CCCGGCTCTT
    ASP   CACGAGGATT TTCAGTCCTC TGCTCTACCA ACTGAGCTAT CCCGGCTCTT
    BRA   CACGAGGATT TTCAGTCCTC TGCTCTACCA ACTGAGCTAT CCCGGCTCTT
    BRE   CACGAGGATT TTCAGTCCTC TGCTCTACCA ACTGAGCTAT CCCGGCTCTT
    CHI   CACGAGGATT TTCAGTCCTC TGCTCTACCA ACTGAGCTAT CCCGGCTCTT
    CRA   CACGAGGATT TTCAGTCCTC TGCTCTACCA ACTGAGCTAT CCCGGCTCTT
 ENGno3   CACGAGGATT TTCAGTCCTC TGCTCTACCA ACTGAGCTAT CCCGGCTCTT
    FAR   CACGAGGATT TTCAGTCCTC TGCTCTACCA ACTGAGCTAT CCCGGCTCTT
 GLAno2   CACGAGGATT TTCAGTCCTC TGCTCTACCA ACTGAGCTAT CCCGGCTCTT
    GLE   CACGAGGATT TTCAGTCCTC TGCTCTACCA ACTGAGCTAT CCCGGCTCTT
 JEZno1   CACGAGGATT TTCAGTCCTC TGCTCTACCA ACTGAGCTAT CCCGGCTCTT
    KOR   CACGAGGATT TTCAGTCCTC TGCTCTACCA ACTGAGCTAT CCCGGCTCTT
    KOY   CACGAGGATT TTCAGTCCTC TGCTCTACCA ACTGAGCTAT CCCGGCTCTT
    LIK   CACGAGGATT TTCAGTCCTC TGCTCTACCA ACTGAGCTAT CCCGGCTCTT
 MARno2   CACGAGGATT TTCAGTCCTC TGCTCTACCA ACTGAGCTAT CCCGGCTCTT
    MAT   CACGAGGATT TTCAGTCCTC TGCTCTACCA ACTGAGCTAT CCCGGCTCTT
    MAX   CACGAGGATT TTCAGTCCTC TGCTCTACCA ACTGAGCTAT CCCGGCTCTT
    MEX   CACGAGGATT TTCAGTCCTC TGCTCTACCA ACTGAGCTAT CCCGGCTCTT
    MEY   CACGAGGATT TTCAGTCCTC TGCTCTACCA ACTGAGCTAT CCCGGCTCTT
    MON   CACGAGGATT TTCAGTCCTC TGCTCTACCA ACTGAGCTAT CCCGGCTCTT
    MOR   CACGAGGATT TTCAGTCCTC TGCTCTACCA ACTGAGCTAT CCCGGCTCTT
    OBO   CACGAGGATT TTCAGTCCTC TGCTCTACCA ACTGAGCTAT CCCGGCTCTT
    OMO   CACGAGGATT TTCAGTCCTC TGCTCTACCA ACTGAGCTAT CCCGGCTCTT
    ORI   CACGAGGATT TTCAGTCCTC TGCTCTACCA ACTGAGCTAT CCCGGCTCTT
PINBANK   CACGAGGATT TTCAGTCCTC TGCTCTACCA ACTGAGCTAT CCCGGCTCTT
 PINKOR   CACGAGGATT TTCAGTCCTC TGCTCTACCA ACTGAGCTAT CCCGGCTCTT
    PUN   CACGAGGATT TTCAGTCCTC TGCTCTACCA ACTGAGCTAT CCCGGCTCTT
    PUR   CACGAGGATT TTCAGTCCTC TGCTCTACCA ACTGAGCTAT CCCGGCTCTT
    RET   CACGAGGATT TTCAGTCCTC TGCTCTACCA ACTGAGCTAT CCCGGCTCTT
    RUB   CACGAGGATT TTCAGTCCTC TGCTCTACCA ACTGAGCTAT CCCGGCTCTT
    SCH   CACGAGGATT TTCAGTCCTC TGCTCTACCA ACTGAGCTAT CCCGGCTCTT
 SITno1   CACGAGGATT TTCAGTCCTC TGCTCTACCA ACTGAGCTAT CCCGGCTCTT
    SMI   CACGAGGATT TTCAGTCCTC TGCTCTACCA ACTGAGCTAT CCCGGCTCTT
    SPI   CACGAGGATT TTCAGTCCTC TGCTCTACCA ACTGAGCTAT CCCGGCTCTT
    TOR   CACGAGGATT TTCAGTCCTC TGCTCTACCA ACTGAGCTAT CCCGGCTCTT
    WIL   CACGAGGATT TTCAGTCCTC TGCTCTACCA ACTGAGCTAT CCCGGCTCTT

          [51                                                 100]
 ABIno1   CCCTGTGGAT CATCCAGGTA CAGGGTTTTC ACTTGTGTCA ACTAAAAATA
    ALC   CCCTGTGGAT CATCCAGGTA CAGGGTTTTC ACTTGTGTCA ACTAAAAATA
    ASP   CCCTGTGGAT CATCCAGGTA CAGGGTTTTC ACTTGTGTCA ACTAAAAATA
    BRA   CCCTGTGGAT CATCCAGGTA CAGGGTTTTC ACTTGTATCA ACTAAAAATA
    BRE   CCCTGTGGAT CATCCAGGTA CAGGGTTTTC ACTTGTGTCA ACTAAAAATA
    CHI   CCCTGTGGAT CATCCAGGTA CAGGGTTTTC ACTTGTGTCA ACTAAAAATA
    CRA   CCCTGTGGAT CATCCAGGTA CAGGGTTTTC ACTTGTGTCA ACTAAAAATA
 ENGno3   CCCTGTGGAT CATCCAGGTA CAGGGTTTTC ACTTGTGTCA ACTAAAAATA
    FAR   CCCTGTGGAT CATCCAGGTA CAGGGTTTTC ACTTGTGTCA ACTAAAAATA
 GLAno2   CCCTGTGGAT CATCCAGGTA CAGGGTTTTC ACTTGTGTCA ACTAAAAATA
    GLE   CCCTGTGGAT CATCCAGGTA CAGGGTTTTC ACTTGTGTCA ACTAAAAATA
 JEZno1   CCCTGTGGAT CATCCAGGTA CAGGGTTTTC ACTTGTGTCA ACTAAAAATA
    KOR   CCCTGTGGAT CATCCAGGTA CAGGGTTTTC ACTTGTGTCA ACTAAAAATA
    KOY   CCCTGTGGAT CATCCAGGTA CAGGGTTTTC ACTTGTGTCA ACTAAAAATA
    LIK   CCCTGTGGAT CATCCAGGTA CAGGGTTTTC ACTTGTGTCA ACTAAAAATA
 MARno2   CCCTGTGGAT CATCCAGGTA CAGGGTTTTC ACTTGTGTCA ACTAAAAATA
    MAT   CCCTGTGGAT CATCCAGGTA CAGGGTTTTC ACTTGTGTCA ACTAAAAATA
    MAX   CCCTGTGGAT CATCCAGGTA CAGGGTTTTC ACTTGTGTCA ACTAAAAATA
    MEX   CCCTGTGGAT CATCCAGGTA CAGGGTTTTC ACTTGTGTCA ACTAAAAATA
    MEY   CCCTGTGGAT CATCCAGGTA CAGGGTTTTC ACTTGTGTCA ACTAAAAATA
    MON   CCCTGTGGAT CATCCAGGTA CAGGGTTTTC ACTTGTGTCA ACTAAAAATA
    MOR   CCCTGTGGAT CATCCAGGTA CAGGGTTTTC ACTTGTGTCA ACTAAAAATA
    OBO   CCCTGTGGAT CATCCAGGTA CAGGGTTTTC ACTTGTGTCA ACTAAAAATA
    OMO   CCCTGTGGAT CATCCAGGTA CAGGGTTTTC ACTTGTGTCA ACTAAAAATA
    ORI   CCCTGTGGAT CATCCAGGTA CAGGGTTTTC ACTTGTGTCA ACTAAAAATA
PINBANK   CCCTGTGGAT CATCCTGGTA CAGGGTTTTA ACTTGTGTCA ACTAAAAAGA
 PINKOR   CCCTGTGGAT CATCCTGGTA CAGGTTCTTA ACTTATGTCA ACTAAAAATA
    PUN   CCCTGTGGAT CATCCAGGTA CAGGGTTTTC ACTTGTGTCA ACTAAAAATA
    PUR   CCCTGTGGAT CATCCAGGTA CAGGGTTTTC ACTTGTATCA ACTAAAAATA
    RET   CCCTGTGGAT CATCCAGGTA CAGGGTTTTC ACTTGTGTCA ACTAAAAATA
    RUB   CCCTGTGGAT CATCCAGGTA CAGGGTTTTC ACTTGTGTCA ACTAAAAATA
    SCH   CCCTGTGGAT CATCCAGGTA CAGGGTTTTC ACTTGTGTCA ACTAAAAATA
 SITno1   CCCTGTGGAT CATCCAGGTA CAGGGTTTTC ACTTGTGTCA ACTAAAAATA
    SMI   CCCTGTGGAT CATCCAGGTA CAGGGTTTTC ACTTGTGTCA ACTAAAAATA
```

```
   SPI  CCCTGTGGAT CATCCAGGTA CAGGGTTTTC ACTTGTGTCA ACTAAAAATA
   TOR  CCCTGTGGAT CATCCAGGTA CAGGGTTTTC ACTTGTGTCA ACTAAAAATA
   WIL  CCCTGTGGAT CATCCAGGTA CAGGGTTTTC ACTTGTATCA ACTAAAAATA

        [101      +2          +1                                150]
ABInol  AGGAAAAAA.0 C0TATTTTT.0C CAGTTTTGAG AATGATCTTT CTTATTTTCA
   ALC  AGGAAAAAA.0 C0TATTTTT.0C CAGTTTTGAG AATGATCTTT CTTATTTTCA
   ASP  AGGAAAAAA.0 C0TATTTTT.0C CAGTTTTGAG AATGATCTTT CTTATTTTCA
   BRA  AGGAAAAAA.0 C0TATTTTT.0C CAGTTTTGAG AATGATCTTT CTTATTTTCA
   BRE  AGGAAAAAA.0 C0TATTTTT.0C CAGTTTTTAG AATGATCTTT CTTATTTTCA
   CHI  AGGAAAAAC.0 C0TATTTTT.0C CAGTTTTGAG AATGATCTTT CTTATTTTCA
   CRA  AGGAAAAAA.0 C0TATTTTT.0C CAGTTTTGAG AATGATCTTT CTTATTTTCA
ENGno3  AGGAAAAAA.0 C0TATTTTT.0C CAGTTTTGAG AATGATCTTT CTTATTTTCA
   FAR  AGGAAAAAA.0 C0TATTTTT.0C CAGTTTTGAG AATGATCTTT CTTATTTTCA
GLAno2  AGGAAAAAA.0 C0TATTTTT.0C CAGTTTTGAG AATGATCTTT CTTATTTTCA
   GLE  AGGAAAAAA.0 C0TATTTTT.0C CAGTTTTGAG AATGATCTTT CTTATTTTCA
JEZno1  AGGAAAAAA.0 C0TATTTTT.0C CAGTTTTGAG AATGATCTTT CTTATTTTCA
   KOR  AGGAAAAAA.0 C0TATTTTT.0C CAGTTTTGAG AATGATCTTT CTTATTTTCA
   KOY  AGGAAAAAA.0 C0TATTTTT.0C CAGTTTTGAG AATGATCTTT CTTATTTTCA
   LIK  AGGAAAAAA.0 C0TATTTTT.0C CAGTTTTGAG AATGATCTTT CTTATTTTCA
MARno2  AGGAAAAAA.0 C0TATTTTT.0C CAGTTTTGAG AATGATCTTT CTTATTTTCA
   MAT  AGGAAAAAA.0 C0TATTTTT.0C CAGTTTTGAG AATGATCTTT CTTATTTTCA
   MAX  AGGAAAAAA.0 C0TATTTTT.0C CAGTTTTGAG AATGATCTTT CTTATTTTCA
   MEX  AGGAAAAAA.0 C0TATTTTT.0C CAGTTTTGAG AATGATCTTT CTTATTTTCA
   MEY  AGGAAAAAA.0 C0TATTTTT.0C CAGTTTTGAG AATGATCTTT CTTATTTTCA
   MON  AGGAAAAAA.0 C0TATTTTT.0C CAGTTTTGAG AATGATCTTT CTTATTTTCA
   MOR  AGGAAAAAA.0 C0TATTTTT.0C CAGTTTTGAG AATGATCTTT CTTATTTTCA
   OBO  AGGAAAAAA.0 C0TATTTTT.0C CAGTTTTGAG AATGATCTTT CTTATTTTCA
   OMO  AGGAAAAAA.0 C0TATTTTT.0C CAGTTTTGAG AATGATCTTT CTTATTTTCA
   ORI  AGGAAAAAA.0 C0TATTTTT.0C CAGTTTTGAG AATGATCTTT CTTATTTTCA
PINBANK  AGTAAAAAA.0 G0TATTTTTA1C TAGTTTTTAG AATGATCTTT ATTATTTTCG
PINKOR  AGGAAAAAG.0 .1TATTTTTC1C CAGTTTTTAC AATGATCTTT ATTATTTTCG
   PUN  AGGAAAAAA.0 C0TATTTTT.0C CAGTTTTGAG AATGATCTTT CTTATTTTCA
   PUR  AGGAAAAAA.0 C0TATTTTT.0C CAGTTTTGAG AATGATCTTT CTTATTTTCA
   RET  AGGAAAAAA.0 C0TATTTTT.0C CAGTTTTGAG AATGATCTTT CTTATTTTCA
   RUB  AGGAAAAAA.0 C0TATTTTT.0C CAGTTTTGAG AATGATCTTT CTTATTTTCA
   SCH  AGGAAAAAA.0 C0TATTTTT.0C CAGTTTTGAG AATGATCTTT CTTATTTTCA
SITno1  AGGAAAAAAA1 C0TATTTTT.0C CAGTTTTGAG AATGATCTTT CTTATTTTCA
   SMI  AGGAAAAAA.0 C0GATTTTT.0C CAGTTTTGAG AATGATCTTT CTTATTTTCA
   SPI  AGGAAAAAA.0 C0TATTTTT.0C CAGTTTTGAG AATGATCTTT CTTATTTTCA
   TOR  AGGAAAAAA.0 C0TATTTTT.0C CAGTTTTGAG AATGATCTTT CTTATTTTCA
   WIL  AGGAAAAAA.0 C0TATTTTT.0C CAGTTTTGAG AATGATCTTT CTTATTTTCA

        [151                      +1                      +1  200]
ABInol  ATCTG..... .......... .....0GAAGT CACTAATATG ATAAAAAA0TG
   ALC  ATCTG..... .......... .....0GAAGT CACTAATATG AGAAAAAA0TG
   ASP  ATCTG..... .......... .....0GAAGT CACTAATATG ATAAAAAA0TG
   BRA  ATCTGTCTTT CTTATTTTCA ATCTG1GAAGT CACTAATATG AGAAAAAA0TG
   BRE  ATCTG..... .......... .....0GAAGT CACTAATATG AGAAAAAA0TG
   CHI  ATCTG..... .......... .....0GAAGT CACTAATATG AGAAAAAA0TG
   CRA  ATCTG..... .......... .....0GAAGT CACTAATATG ATAAAAAA0TG
ENGno3  ATCTG..... .......... .....0GAAGT CACTAATATG AGAAAAAA0TG
   FAR  ATCTG..... .......... .....0GAAGT CACTAATATG AGAAAAAA0TG
GLAno2  ATCTG..... .......... .....0GAAGT CACTAATATG AGAAAAAA0TG
   GLE  ATCTG..... .......... .....0GAAGT CACTAATATG ATAAAAAA0TG
JEZno1  ATCTG..... .......... .....0GAAGT CACTAATATG AGAAAAAA0TG
   KOR  ATCTG..... .......... .....0GAAGT CACTAATATG A?AAAAAA0TG
   KOY  ATCTG..... .......... .....0GAAGT CACTAATATG ATAAAAAA0TG
   LIK  ATCTG..... .......... .....0GAAGT CACTAATATG AGAAAAAA0TG
MARno2  ATCTG..... .......... .....0GAAGT CACTAATATG AGAAAAAA0TG
   MAT  ATCTG..... .......... .....0GAAGT CACTAATATG AGAAAAAA0TG
   MAX  ATCTG..... .......... .....0GAAGT CACTAATATG AGAAAAAA0TG
   MEX  ATCTG..... .......... .....0GAAGT CACTAATATG AGAAAAAA0TG
   MEY  ATCTG..... .......... .....0GAAGT CACTAATATG ATAAAAAA0TG
   MON  ATCTG..... .......... .....0GAAGT CACTAATATG ATAAAAAA0TG
   MOR  ATCTG..... .......... .....0GAAGT CACTAATATG AGAAAAAA0TG
   OBO  ATCTG..... .......... .....0GAAGT CACTAATATG ATAAAAAA0TG
   OMO  ATCTG..... .......... .....0GAAGT CACTAATATG AGAAAAAA0TG
   ORI  ATCTG..... .......... .....0GAAGT CACTAATATG A?AAAAAA0TG
PINBANK  ATCTG..... .......... .....0GAAGT CACTAATATG ATAAAAA.1TG
PINKOR  ATCTG..... .......... .....0GAAGT CACTAATATG AGAAAAAA0TG
   PUN  ATCTG..... .......... .....0GAAGT AACTAATATG AGAAAAAA0TG
   PUR  ATCTG..... .......... .....0GAAGT CACTAATATG AGAAAAAA0TG
   RET  ATCTG..... .......... .....0GAAGT CACTAATATG ATAAAAAA0TG
   RUB  ATCTG..... .......... .....0GAAGT CACTAATATG AGAAAAAA0TG
   SCH  ATCTG..... .......... .....0GAAGT CACTAATATG AGAAAAAA0TG
SITno1  ATCTG..... .......... .....0GAAGT CACTAATATG AGAAAAAA0TG
   SMI  ATCTG..... .......... .....0GAAGT CACTAATATG AGAAAAAA0TG
   SPI  ATCTG..... .......... .....0GAAGT CACTAATATG AGAAAAAA0TG
   TOR  ATCTG..... .......... .....0GAAGT CACTAATATG ATAAAAAA0TG
   WIL  ATCTG..... .......... .....0GAAGT CACTAATATG AGAAAAAA0TG

        [201      +1          +1          +1   +1          250]
ABInol  GACTGCGACT G0AATAATTTC C0AAATGAT.. ...0T0TCATCT ATGTGGATCA
```

```
   ALC  GACTGCGACT G0AATAATTTC C0AAATGAT.. ...0T0TCATCT ATGTGGATCA
   ASP  GACTGCGACT G0AATAATTTC C0AAATGAT.. ...0T0TCATCT ATGTGGATCA
   BRA  GACTGCGACT G0AATAATTTC C0AAATGAT.. ...0.1TCATCT ATGTGGATCA
   BRE  GACTGCGACT G0AATAATTTC C0AAATGAT.. ...0T0TCATCT ATGTGGATCA
   CHI  GACTGCGACT G0AATAATTTC C0AAATGAT.. ...0T0TCATCT ATGTGGATCA
   CRA  GACTGCGACT G0AATAATTTC C0AAATGAT.. ...0T0TCATCT ATGTGGATCA
ENGno3  GACTGCGATT G0AATAATTTC G0AAATGAT.. ...0T0TCATCT ATGTGGATCA
   FAR  GACTGCGACT G0AATAATTTC C0AAATGAT.. ...0T0TCATCT ATGTGGATCA
GLAno2  GACTGCGATT G0AATAATTTC G0AAATGAT.. ...0T0TCATCT ATGTGGATCA
   GLE  GACTG..... .1AATAATTTC C0AAATGAT.. ...0T0TCATCT ATGTGGATCA
JEZno1  GACTGCGACT G0AATAATTTC C0AAATGAT.. ...0T0TCATCT ATGTGGATCA
   KOR  GACTGCGACT G0AATAATTTC C0AAATGAT.. ...0T0TCATCT ATGTGGATCA
   KOY  GACTGCGACT G0AATAATTTC C0AAATGAT.. ...0T0TCATCT ATGTGGATCA
   LIK  GACTGCGACT G0AATAATTTC C0AAATGAT.. ...0T0TCATCT ATGTGGATCA
MARno2  GACTGCGACT G0AATAATTTC C0AAATGAT.. ...0T0TCATCT ATGTGGATCA
   MAT  GACTGCGACT G0AATAATTTC C0AAATGAT.. ...0T0TCATCT ATGTGGATCA
   MAX  GACTGCGACT G0AATAATTTC C0AAATGAT.. ...0T0TCATCT ATGTGGATCA
   MEX  GACTGCGACT G0AATAATTTC G0AAATGAT.. ...0T0TCATCT ATGTGGATCA
   MEY  GACTGCGACT G0AATAATTTC C0AAATGAT.. ...0T0TCATCT ATGTGGATCA
   MON  GACTGCGACT G0AATAATTTC C0AAATGAT.. ...0T0TCATCT ATGTGGATCA
   MOR  GACTGCGACT G0AATAATTTC C0AAATGAT.. ...0T0TCATCT ATGTGGATCA
   OBO  GACTGCGACT G0AATAATTTC C0AAATGAT.. ...0T0TCATCT ATGTGGATCA
   OMO  GACTGCGACT G0AATAATTTC C0AAATGAT.. ...0T0TCATCT ATGTGGATCA
   ORI  GACTGCGACT G0AATAATTTC C0AAATGAT.. ...0T0TCATCT ATGTGGATCA
PINBANK  GACTGCAATT T0AATAATTTA A0AAATGATCT ACT1C0TATTCT ATGTAGATCA
PINKOR  GACTGCAATT G0AATAATTTC .1AAACGAT.. ...0C0TAATAT ATGTAGATCA
   PUN  GACTGCGACT G0AATAATTTC C0AAATGAT.. ...0T0TCATCT ATGTGGATCA
   PUR  GACTGCGACT G0AATAATTTC C0AAATGAT.. ...0T0TCATCT ATGTGGATCA
   RET  GACTGCGACT G0AATAATTTC C0AAATGAT.. ...0T0TCATCT ATGTGGATCA
   RUB  GACTGCGACT G0AATAATTTC C0AAATGAT.. ...0T0TAATCT ATGTGGATCA
   SCH  GACTGCGACT G0AATAATTTC C0AAATGAT.. ...0T0TCATCT ATGTGGATCA
SITno1  GACTGCGACT G0AATAATTTC C0AAATGAT.. ...0T0TCATCT ATGTGGATCA
   SMI  GACTGCGACT G0AATAATTTC C0AAATGAT.. ...0T0TCATCT ATGTGGATCA
   SPI  GACTGCGACT G0AATAATTTC C0AAATGAT.. ...0T0TCATCT ATGTGGATCA
   TOR  GACTGCGACT G0AATAATTTC C0AAATGAT.. ...0T0TCATCT ATGTGGATCA
   WIL  GACTGCGACT G0AATAATTTC C0AAATGAT.. ...0T0TCATCT ATGTGGATCA

        [251                                                  300]
ABIno1  TCTATCACAT AATGAATTGA TAATATGATC AACTTATTAT CTAATCAACT
   ALC  TCTATCACAT AATGAATTGA TAATATGATC AACTTATTAT CTAATCAACT
   ASP  TCTATCACAT AATGAATTGA TAATATGATC AACTTATTAT CTAATCAACT
   BRA  TCTATCACAT AATGAATTGA TAATATGATC AACTTATTAT CTAATCAACT
   BRE  TCTATCACAT AATGAATTGA TAATATGATC AACTTATTAT CTAATCAACT
   CHI  TCTATCACAT AATGAATTGA TAATATGATC AACTTATTAT CTAATCAACT
   CRA  TCTATCACAT AATGAATTGA TAATATGATC AACTTATTAT CTAATCAACT
ENGno3  TCTATCACAT AATGAATTGA TAATATGATC AACTTATTAT CTAATCAACT
   FAR  TCTATCACAT AATGAATTGA TAATATGATC AACTTATTAT CTAATCAACT
GLAno2  TCTATCACAT AATGAATTGA TAATATGATC AACTTATTAT CTAATMAACT
   GLE  TCTATCACAT AATGAATTGA TAATATGATC AACTTATTAT CTAATCAACT
JEZno1  TCTATCACAT AATGAATTGA TAATATGATC AACTTATTAT CTAATCAACT
   KOR  TCTATCACAT AATGAATTGA TAATATGATC AACTTATTAT CTAATCAACT
   KOY  TCTATCACAT AATGAATTGA TAATATGATC AACTTATTAT CTAATCAACT
   LIK  TCTATCACAT AATGAATTGA TAATATGATC AACTTATTAT CTAATCAACT
MARno2  TCTATCACAT AATGAATTGA TAATATGATC AACTTATTAT CTAATCAACT
   MAT  TCTATCACAT AATGAATTGA TAATATGATC AACTTATTAT CTAATCAACT
   MAX  TCTATCACAT AATGAATTGA TAATATGATC AACTTATTAT CTAATCAACT
   MEX  TCTATCACAT AATGAATTGA TAATATGATC AACTTATTAT CTAATCAACT
   MEY  TCTATCACAT AATGAATTGA TAATATGATC AACTTATTAT CTAATCAACT
   MON  TCTATCACAT AATGAATTGA TAATATGATC AACTTATTAT CTAATCAACT
   MOR  TCTATCACAT AATGAATTGA TAATATGATC AACTTATTAT CTAATCAACT
   OBO  TCTATCACAT AATGAATTGA TAATATGATC AACTTATTAT CTAATCAACT
   OMO  TCTATCACAT AATGAATTGA TAATATGATC AACTTATTAT CTAATCAACT
   ORI  TCTATCACAT AATGAATTGA TAATATGATC AACTTATTAT CTAATCAACT
PINBANK  TATATCACAA AATAAATTGA TCATATGATC AACTTATTAA CAGATCAACT
PINKOR  TATATCACAA AATGAATTGA TCATATGATC AACTCATTAA CGGATCAACT
   PUN  TCTATCACAT AATGAATTGA TAATATGATC AACTTATTAT CTAATCAACT
   PUR  TCTATCACAT AATGAATTGA TAATATGATC AACTTATTAT CTAATCAACT
   RET  TCTATCACAT AATGAATTGA TAATATGATC AACTTATTAT CTAATCAACT
   RUB  TCTATCACAT AATGAATTGA TAATATGATC AACTTATTAT CTAATCAACT
   SCH  TCTATCACAT AATGAATTGA TAATATGATC AACTTATTAT CTAATCAACT
SITno1  TCTATCACAT AATGAATTGA TAATATGATC AACTTATTAT CTAATCAACT
   SMI  TCTATCACAT AATGAATTGA TAATATGATC AACTTATTAT CTAATCAACT
   SPI  TCTATCACAT AATGAATTGA TAATATGATC AACTTATTAT CTAATCAACT
   TOR  TCTATCACAT AATGAATTGA TAATATGATC AACTTATTAT CTAATCAACT
   WIL  TCTATCACAT AATGAATTGA TAATATGATC AACTTATTAT CTAATCAACT

        [301                                                  350]
ABIno1  CAACTTGTCA TAAAATGGAT GGACAGATTC ATGTTCTAAT TTATTCTGTT
   ALC  CAACTTGTCA TAAAATGGAT GGACAGATTC ATGTTCTAAT TTCTTCTGTT
   ASP  CAACTTGTCA TAAAATGGAT GGACAGATTC ATGTTCTAAT TTATTCTGTT
   BRA  CAACTTGTCA TAAAATGGAT GGACAGATTC ATGTTCTAAT TTCTTCTGTT
   BRE  CAACTTGTCA TAAAATGGAT GGACAGATTC ATGTTCTAAT TTCTTCTGTT
   CHI  CAACTTGTCA TAAAATGGAT GGACAGATTC ATGTTCTAAT TTCTTCTGTT
   CRA  CAACTTGTCA TAAAATGGAT GGACAGATTC ATGTTCTAAT TTATTCTGTT
```

```
ENGno3  CAACTTGTCA TAAAATGGAT GGACAGATTC ATGTTCTAAT TTCTTCTGTT
   FAR  CAACTTGTCA TAAAATGGAT GGACAGATTC ATGTTCTAAT TTCTTCTGTT
GLAno2  CAACTTGTCA TAAAATGGAT GGACAGATTC ATGTTCTAAT TTCTTCTGTT
   GLE  CAACTTGTCA TAAAATGGAT GGACAGATTC ATGTTCTAAT TTATTCTGTT
JEZno1  CAACTTGTCA TAAAATGGAT GGACAGATTC ATGTTCTAAT TTCTTCTGTT
   KOR  CAACTTGTCA TAAAATGGAT GGACAGATTC ATGTTCTAAT TT?TTCTGTT
   KOY  CAACTTGTCA TAAAATGGAT GGACAGATTC ATGTTCTAAT TTATTCTGTT
   LIK  CAACTTGTCA TAAAATGGAT GGACAGATTC ATGTTCTAAT TTCTTCTGTT
MARno2  CAACTTGTCA TAAAATGGAT GGACAGATTC ATGTTCTAAT TTCTTCTGTT
   MAT  CAACTTGTCA TAAAATGGAT GGACAGATTC ATGTTCTAAT TTCTTCTGTT
   MAX  CAACTTGTCA TAAAATGGAT GGACAGATTC ATGTTCTAAT TTCTTCTGTT
   MEX  CAACTTGTCA TAAAATGGAT GGACAGATTC ATGTTCTAAT TTCTTCTGTT
   MEY  CAACTTGTCA TAAAATGGAT GGACAGATTC ATGTTCTAAT TTATTCTGTT
   MON  CAACTTGTCA TAAAATGGAT GGACAGATTC ATGTTCTAAT TTATTCTGTT
   MOR  CAACTTGTCA TAAAATGGAT GGACAGATTC ATGTTCTAAT TTCTTCTGTT
   OBO  CAACTTGTCA TAAAATGGAT GGACAGATTC ATGTTCTAAT TTATTCTGTT
   OMO  CAACTTGTCA TAAAATGGAT GGACAGATTC ATGTTCTAAT TTCTTCTGTT
   ORI  CAACTTGTCA TAAAATGGAT GGACAGATTC ATGTTCTAAT TT?TTCTGTT
PINBANK CAACTTGTTA TAAGATGGAT GAAAAGATTC ATGTTCTAAT TTCTTCTCTT
PINKOR  CAACTTGTCA TAAGATGGAT GAAAAGATTC ATGTTCTAAT TTCTTCTCTT
   PUN  CAACTTGTCA TAAAATGGAT GGACAGATTC ATGTTCTAAT TTCTTCTGTT
   PUR  CAACTTGTCA TAAAATGGAT GGACAGATTC ATGTTCTAAT TTCTTCTGTT
   RET  CAACTTGTCA TAAAATGGAT GGACAGATTC ATGTTCTAAT TTATTCTGTT
   RUB  CAACTTGTCA TAAAATGGAT GGACAGATTC ATGTTCTAAT TTCTTCTGTT
   SCH  CAACTTGTCA TAAAATGGAT GGACAGATTC ATGTTCTAAT TTCTTCTGTT
SITno1  CAACTTGTCA TAAAATGGAT GGACAGATTC ATGTTCTAAT TTCTTCTGTT
   SMI  CAACTTGTCA TAAAATGGAT GGACAGATTC ATGTTCTAAT TTATTCTGTT
   SPI  CAACTTGTCA TAAAATGGAT GGACAGATTC ATGTTCTAAT TTCTTCTGTT
   TOR  CAACTTGTCA TAAAATGGAT GGACAGATTC ATGTTCTAAT TTATTCTGTT
   WIL  CAACTTGTCA TAAAATGGAT GGACAGATTC ATGTTCTAAT TTCTTCTGTT

        [351          +1    +1                          +1         400]
ABIno1  CATGAAGAAA TA0AAATA0GTG AGGTAAAATA ATCGAGAAG0T GAGAATGGAT
   ALC  CATGAAGAAA TC0AAATA0GTG AGGTAAAATA ATCGAGAAG0T GAGAATGGAT
   ASP  CATGAAGAAA TA0AAATA0GTG AGGTAAAATA ATCGAGAAG0T GAGAATGGAT
   BRA  CATGAAGAAA TC0AAATA0GTG AGGTAAAATA ATCGAGAAG0T GAGAATGGAT
   BRE  CATGAAGAAA TC0AAATA0GTG AGGTAAAATA ATCGAGAAG0T GAGAATGGAT
   CHI  CATGAAGAAA TC0AAATA0GTG AGGTAAAATA ATCAAGAAG0T GAGAATGGAT
   CRA  CATGAAGAAA TA0AAATA0GTG AGGTAAAATA ATCGAGAAG0T GAGAATGGAT
ENGno3  CATGAAGAAA KA0AAATA0GTG AGGTAAAATA ATCGAGAAG0T GAGAATGGAT
   FAR  CATGAAGAAA TA0.....1GTG AGGTAAAATA ATCGAGAAG0T GAGAATGGAT
GLAno2  CATGAAGAAA TA0AAATA0GTG AGGTAAAATA ATCGAGAAG0T GAGAATGGAT
   GLE  CATGAAGAAA TA0AAATA0GTG AGGTAAAATA ATCGAGAAG0T GAGAATGGAT
JEZno1  CATGAAGAAA TA0AAATA0GTG AGGTAAAATA ATCGAGAAG0T GAGAATGGAT
   KOR  CATGAAGAAA TA0AAATA0GTG AGGTAAAATA ATCGAGAAG0T GAGAATGGAT
   KOY  CATGAAGAAA TA0AAATA0GTG AGGTAAAATA ATCGAGAAG0T GAGAATGGAT
   LIK  CATGAAGAAA TA0.....1GTG AGGTAAAATA ATCGAGAAG0T GAGAATGGAT
MARno2  CATGAAGAAA TA0AAATA0GTG AGGTAAAATA ATCGAGAAG0T GAGAATGGAT
   MAT  CATGAAGAAA TC0AAATA0GTG AGGTAAAATA ATCGAGAAG0T GAGAATGGAT
   MAX  CATGAAGAAA TC0AAATA0GTG AGGTAAAATA ATCGAGAAG0T GAGAATGGAT
   MEX  CATGAAGAAA TA0AAATA0GTG AGGTAAAATA ATCGAGAAG0T GAGAATGGAT
   MEY  CATGAAGAAA TA0AAATA0GTG AGGTAAAATA ATCGAGAAG0T GAGAATGGAT
   MON  CATGAAGAAA TA0AAATA0GTG AGGTAAAATA ATCGAGAAG0T GAGAATGGAT
   MOR  CATGAAGAAA TA0AAATA0GTG AGGTAAAATA ATCGAGAAG0T GAGAATGGAT
   OBO  CATGAAGAAA TA0AAATA0GTG AGGTAAAATA ATCGAGAAG0T GAGAATGGAT
   OMO  CATGAAGAAA TA0AAATA0GTG AGGTAAAATA ATCGAGAAG0T GAGAATGGAT
   ORI  CATGAAGAAA T?0AAATA0GTG AGGTAAAATA ATCGAGAAG0T GAGAATGGAT
PINBANK CATGAAAAAA ..1AAATA0GCG AGGTAAAAGA ATCGATAAA0T TAGAATGGAT
PINKOR  AGAGAAGAAA ..?.....?... .......... .........1T TAGAATGGAT
   PUN  CATGAAGAAA TA0AAATA0GTG AGGTAAAATA ATCGAGAAG0T GAGAATGGAT
   PUR  CATGAAGAAA TC0AAATA0GTG AGGTAAAATA ATCGAGAAG0T GAGAATGGAT
   RET  CATGAAGAAA TA0AAATA0GTG AGGTAAAATA ATCGAGAAG0T GAGAATGGAT
   RUB  CATGAAGAAA TA0AAATA0GTG AGGTAAAATA ATCGAGAAG0T GAGAATGGAT
   SCH  CATGAAGAAA TA0.....1GTG AGGTAAAATA ATCGAGAAG0T GAGAATGGAT
SITno1  CATGAAGAAA TA0AAATA0GTG AGGTAAAAGA ATCGAGAAG0T GAGAATGGAT
   SMI  CATGAAGAAA TA0.....1GTG AGGTAAAATA ATCGAGAAG0T GAGAATGGAT
   SPI  CATGAAGAAA TA0.....1GTG AGGTAAAATA ATCGAGAAG0T GAGAATGGAT
   TOR  CATGAAGAAA TA0AAATA0GTG AGGTAAAATA ATCGAGAAG0T GAGAATGGAT
   WIL  CATGAAGAAA TC0AAATA0GTG AGGTAAAATA ATCGAGAAG0T GAGAATGGAT

        [401                                  +1                450]
ABIno1  TCGAACTAAT GGAATTGGAG AAAATAGATC AGT.0CGTTCG GGAACGAACC
   ALC  TCAAACTAAT GGAATTGGAG AAAATAGATC AGT.0CGTTCG GGAACGAACC
   ASP  TCGAACTAAT GGAATTGGAG AAAATAGATC AGT.0CGTTCG GGAACGAACC
   BRA  TCAAACTAAT GGAATTGGAG AAAATAGATC AGT.0CGTTCG GGAACGAACC
   BRE  TCGAACTAAT GGAATTGGAG AAAATAGATC AGT.0CGTTCG GGAACGAACC
   CHI  TCGAACTAAT GGAATTGGAG AAAATAGATC AGT.0CGTTCG GGAACGAACC
   CRA  TCGAACTAAT GGAATTGGAG AAAATAGATC AGT.0CGTTCG GGAACGAACC
ENGno3  TCGAACTAAT GGAATTGGAG AAAATAGATC AGT.0CGTTCG GGAACGAACC
   FAR  TCGAACTAAT GGAATTGGAG AAAATAGATC AGT.0CGTTCG GGAACGAACC
GLAno2  TCGAACTAAT GGAATTGGAG AAAATAGATC AGT.0CGTTCG GGAACGAACC
   GLE  TCGAACTAAT GGAATTGGAG AAAATAGATC AGT.0CGTTCG GGAACGAACC
JEZno1  TCGAACTAAT GGAATTGGAG AAAATAGATC AGT.0CGTTCG GGAACGAACC
   KOR  TCGAACTAAT GGAATTGGAG AAAATAGATC AGT.0CGTTCG GGAACGAACC
```

```
   KOY  TCGAACTAAT GGAATTGGAG AAAATAGATC AGT.0CGTTCG GGAACGAACC
   LIK  TCGAACTAAT GGAATTGGAG AAAATAGATC AGT.0CGTTCG GGAACGAACC
MARno2  TCGAACTAAT GGAATTGGAG AAAATAGATC AGT.0CGTTCG GGAACGAACC
   MAT  TCGAACTAAT GGAATTGGAG AAAATAGATC AGT.0CGTTCG GGAACGAACC
   MAX  TCGAACTAAT GGAATTGGAG AAAATAGATC AGT.0CGTTCG GGAACGAACC
   MEX  TCGAACTAAT GGAATTGGAG AAAATAGATC AGT.0CGTTCG GGAACGAACC
   MEY  TCAAACTAAT GGAATTGGAG AAAATAGATC AGT.0CGTTCG GGAACGAACC
   MON  TCGAACTAAT GGAATTGGAG AAAATAGATC AGT.0CGTTCG GGAACGAACC
   MOR  TCAAACTAAT GGAATTGGAG AAAATAGATC AGT.0CGTTCG GGAACGAACC
   OBO  TCGAACTAAT GGAATTGGAG AAAATAGATC AGT.0CGTTCG GGAACGAACC
   OMO  TCGAACTAAT GGAATTGGAG AAAATAGATC AGT.0CGTTCG GGAACGAACC
   ORI  TC?AACTAAT GGAATTGGAG AAAATAGATC AGT.0CGTTCG GGAACGAACC
PINBANK  TCGAACCAAT GGAATTGGAG AAGATAGATC AATC1CGTTCG GGAACGAACC
PINKOR  TTGAACCAAT GGAATTGGAG AAAATAGATC AGTC1CGTTCG GGAACGAACT
   PUN  TCGAACTAAT GGAATTGGAG AAAATAGATC AGT.0CGTTCG GGAACGAACC
   PUR  TCGAACTAAT GGAATTGGAG AAAATAGATC AGT.0CGGTCG GGAACGAACC
   RET  TCGAACTAAT GGAATTGGAG AAAATAGATC AGT.0CGTTCG GGAACGAACC
   RUB  TCGAACTAAT GGAATTGGAG AAAATAGATC AGT.0CGTTCG GGAACGAACC
   SCH  TCGAACTAAT GGAATTGGAG AAAATAGATC AGT.0CGTTCG GGAACGAACC
SITno1  TCGAACTAAT GGAATTGGAG AAAATAGATC AGT.0CGTTCG GGAACGAACC
   SMI  TCGAACTAAT GGAATTGGAG AAAATAGATC AGT.0CGTTCG GGAACGAACC
   SPI  TCGAACTAAT GGAATTGGAG AAAATAGATC AGT.0CGTTCG GGAACGAACC
   TOR  TCGAACTAAT GGAATTGGAG AAAATAGATC AGT.0CGTTCG GGAACGAACC
   WIL  TCGAACTAAT GGAATTGGAG AAAATAGATC AGT.0CGGTCG GGAACGAACC

        [451                                                500]
ABIno1  TAGGTGGGGA TAGAGGGACT TGAACCCTCA CGGTCTATAA AGCCAACGGA
   ALC  TAGGTGGGGA TAGAGGGACT TGAACCCTCA CGGTCTATAA AGCCAACGGA
   ASP  TAGGTGGGGA TAGAGGGACT TGAACCCTCA CGGTCTATAA AGCCAACGGA
   BRA  TAGGTGGGGA TAGAGGGACT TGAACCCTCA CGGTCTATAA AGCCAACGGA
   BRE  TAGGTGGGGA TAGAGGGACT TGAACCCTCA CGGTCTATAA AGCCAACGGA
   CHI  TAGGTGGGGA TAGAGGGACT TGAACCCTCA CGGTCTATAA AGCCAACGGA
   CRA  TAGGTGGGGA TAGAGGGACT TGAACCCTCA CGGTCTATAA AGCCAACGGA
ENGno3  TAGGTGGGGA TAGAGGGACT TGAACCCTCA CGGTCTATAA AGCCAACGGA
   FAR  TAGGTGGGGA TAGAGGGACT TGAACCCTCA CGGTCTATAA AGCCAACGGA
GLAno2  TAGGTGGGGA TAGAGGGACT TGAACCCTCA CGGTCTATAA AGCCAACGGA
   GLE  TAGGTGGGGA TAGAGGGACT TGAACCCTCA CGGTCTATAA AGCCAACGGA
JEZno1  TAGGTGGGGA TAGAGGGACT TGAACCCTCA CGGTCTATAA AGCCAACGGA
   KOR  TAGGTGGGGA TAGAGGGACT TGAACCCTCA CGGTCTATAA AGCCAACGGA
   KOY  TAGGTGGGGA TAGAGGGACT TGAACCCTCA CGGTCTATAA AGCCAACGGA
   LIK  TAGGTGGGGA TAGAGGGACT TGAACCCTCA CGGTCTATAA AGCCAACGGA
MARno2  TAGGTGGGGA TAGAGGGACT TGAACCCTCA CGGTCTATAA AGCCAACGGA
   MAT  TAGGTGGGGA TAGAGGGACT TGAACCCTCA CGGTCTATAA AGCCAACGGA
   MAX  TAGGTGGGGA TAGAGGGACT TGAACCCTCA CGGTCTATAA AGCCAACGGA
   MEX  TAGGTGGGGA TAGAGGGACT TGAACCCTCA CGGTCTATAA AGCCAACGGA
   MEY  TAGGTGGGGA TAGAGGGACT TGAACCCTCA CGGTCTATAA AGCCAACGGA
   MON  TAGGTGGGGA TAGAGGGACT TGAACCCTCA CGGTCTATAA AGCCAACGGA
   MOR  TAGGTGGGGA TAGAGGGACT TGAACCCTCA CGGTCTATAA AGCCAACGGA
   OBO  TAGGTGGGGA TAGAGGGACT TGAACCCTCA CGGTCTATAA AGCCAACGGA
   OMO  TAGGTGGGGA TAGAGGGACT TGAACCCTCA CGGTCTATAA AGCCAACGGA
   ORI  TAGGTGGGGA TAGAGGGACT TGAACCCTCA CGGTCTATAA AGCCAACGGA
PINBANK  TGGGTGGGGA TAGAGGGACT TGAACCCTCA CGGTCTATAA AGCCAACGGA
PINKOR  TGGGTGGGGA TAGAGGGACT TGAACCCTCA CGGTCTATAA AGCCAACGGA
   PUN  TAGGTGGGGA TAGAGGGACT TGAACCCTCA CGGTCTATAA AGCCAACGGA
   PUR  TAGGTGGGGA TAGAGGGACT TGAACCCTCA CGGTCTATAA AGCCAACGGA
   RET  TAGGTGGGGA TAGAGGGACT TGAACCCTCA CGGTCTATAA AGCCAACGGA
   RUB  TAGGTGGGGA TAGAGGGACT TGAACCCTCA CGGTCTATAA AGCCAACGGA
   SCH  TAGGTGGGGA TAGAGGGACT TGAACCCTCA CGGTCTATAA AGCCAACGGA
SITno1  TAGGTGGGGA TAGAGGGACT TGAACCCTCA CGGTCTATAA AGCCAACGGA
   SMI  TAGGTGGGGA TAGAGGGACT TGAACCCTCA CGGTCTATAA AGCCAACGGA
   SPI  TAGGTGGGGA TAGAGGGACT TGAACCCTCA CGGTCTATAA AGCCAACGGA
   TOR  TAGGTGGGGA TAGAGGGACT TGAACCCTCA CGGTCTATAA AGCCAACGGA
   WIL  TAGGTGGGGA TAGAGGGACT TGAACCCTCA CGGTCTATAA AGCCAACGGA

        [501                                                550]
ABIno1  TTTTCCTCCT ACTGCAATTT GCATTGTTGT TGACATTGAC ACGTAGAATT
   ALC  TTTTCCTCCT ACTGCAATTT GCATTGTTGT TGACATTGAC ACGTAGAATT
   ASP  TTTTCCTCCT ACTGCAATTT GCATTGTTGT TGACATTGAC ACGTAGAATT
   BRA  TTTTCCTCCT ACTGCAATTT GCATTGTTGT TGACATTGAC ACGTAGAATT
   BRE  TTTTCCTCCT ACTGCAATTT GCATTGTTGT TGACATTGAC ACGTAGAATT
   CHI  TTTTCCTCCT ACTGCAATTT GCATTGTTGT TGACATTGAC ACGTAGAATT
   CRA  TTTTCCTCCT ACTGCAATTT GCATTGTTGT TGACATTGAC ACGTAGAATT
ENGno3  TTTTCCTCCT ACTGCAATTT GCATTGTAGT TGACATTGAC ACGTAGAATT
   FAR  TTTTCCTCCT ACTGCAATTT GCATTGTTGT TGACATTGAC ACGTAGAATT
GLAno2  TTTTCCTCCT ACTGCAATTT GCATTGTAGT TGACATTGAC ACGTAGAATT
   GLE  TTTTCCTCCT ACTGCAATTT GCATTGTTGT TGACATTGAC ACGTAGAATT
JEZno1  TTTTCCTCCT ACTGCAATTT GCATTGTTGT TGACATTGAC ACGTAGAATT
   KOR  TTTTCCTCCT ACTGCAATTT GCATTGTTGT TGACATTGAC ACGTAGAATT
   KOY  TTTTCCTCCT ACTGCAATTT GCATTGTTGT TGACATTGAC ACGTAGAATT
   LIK  TTTTCCTCCT ACTGCAATTT GCATTGTTGT TGACATTGAC ACGTAGAATT
MARno2  TTTTCCTCCT ACTGCAATTT GCATTGTTGT TGACATTGAC ACGTAGAATT
   MAT  TTTTCCTCCT ACTGCAATTT GCATTGTTGT TGACATTGAC ACGTAGAATT
   MAX  TTTTCCTCCT ACTGCAATTT GCATTGTTGT TGACATTGAC ACGTAGAATT
   MEX  TTTTCCTCCT ACTGCAATTT GCATTGTAGT TGACATTGAC ACGTAGAATT
```

```
   MEY  TTTTCCTCCT ACTGCAATTT GCATTGTTGT TGACATTGAC ACGTAGAATT
   MON  TTTTCCTCCT ACTGCAATTT GCATTGTTGT TGACATTGAC ACGTAGAATT
   MOR  TTTTCCTCCT ACTGCAATTT GCATTGTTGT TGACATTGAC ACGTAGAATT
   OBO  TTTTCCTCCT ACTGCAATTT GCATTGTTGT TGACATTGAC ACGTAGAATT
   OMO  TTTTCCTCCT ACTGCAATTT GCATTGTTGT TGACATTGAC ACGTAGAATT
   ORI  TTTTCCTCCT ACTGCAATTT GCATTGTTGT TGACATTGAC ACGTAGAATT
PINBANK TTTTCCTCCT ACTGCAATTT GCATTGTTGT TGACATTGAC ATGTAGAATT
 PINKOR TTTTCCTCCT ACTGCAATTT GCATTGTTGT TGACATTGAC ATGTAGAATT
   PUN  TTTTCCTCCT ACTGCAATTT GCATTGTTGT TGACATTGAC ACGTAGAATT
   PUR  TTTTCCTCCT ACTGCAATTT GCATTGTTGT TGACATTGAC ACGTAGAATT
   RET  TTTTCCTCCT ACTGCAATTT GCATTGTTGT TGACATTGAC ACGTAGAATT
   RUB  TTTTCCTCCT ACTGCAATTT GCATTGTTGT TGACATTGAC ACGTAGAATT
   SCH  TTTTCCTCCT ACTGCAATTT GCATTGTTGT TGACATTGAC ACGTAGAATT
 SITno1 TTTTCCTCCT ACTGCAATTT GCATTGTTGT TGACATTGAC ACGTAGAATT
   SMI  TTTTCCTCCT ACTGCAATTT GCATTGTTGT TGACATTGAC ACGTAGAATT
   SPI  TTTTCCTCCT ACTGCAATTT GCATTGTTGT TGACATTGAC ACGTAGAATT
   TOR  TTTTCCTCCT ACTGCAATTT GCATTGTTGT TGACATTGAC ACGTAGAATT
   WIL  TTTTCCTCCT ACTGCAATTT GCATTGTTGT TGACATTGAC ACGTAGAATT

        [551                          +1                      600]
 ABIno1 GGACTCTATC TTTATTCTCG TCCAA0CCATT TATTCCAAAA ACTGATTCAA
   ALC  GGACTCTATC TTTATTCTCG TCCAA0CCATT TATTCCAAAA ACTGATTCAA
   ASP  GGACTCTATC TTTATTCTCG TCCAA0CCATT TATTCCAAAA ACTGATTCAA
   BRA  GGACTCTATC TTTATTCTCG TCCAA0CCATT TATTCCAAAA ACTGATTCAA
   BRE  GGACTCTATC TTTATTCTCG TCCAA0CCATT TATTCCAAAA ACTGATTCAA
   CHI  GGACTCTATC TTTATTCTCG TCCAA0CCATT TATTCCAAAA ACTGATTCAA
   CRA  GGACTCTATC TTTATTCTCG TCCAA0CCATT TATTCCAAAA ACTGATTCAA
 ENGno3 GGACTCTATC TTTATTCTCG TCCAA0CCATT TATTCCAAAA ACTGATTCAA
   FAR  GGACTCTATC TTTATTCTCG TCCAA0CCATT TATTCCAAAA ACTGATTCAA
 GLAno2 GGACTCTATC TTTATTCTCG TCCAA0CCATT TATTCCAAAA ACTGATTCAA
   GLE  GGACTCTATC TTTATTCTCG TCCAA0CCATT TATTCCAAAA ACTGATTCAA
 JEZno1 GGACTCTATC TTTATTCTCG TCCAA0CCATT TATTCCAAAA ACTGATTCAA
   KOR  GGACTCTATC TTTATTCTCG TCCAA0CCATT TATTCCAAAA ACTGATTCAA
   KOY  GGACTCTATC TTTATTCTCG TCCAA0CCATT TATTCCAAAA ACTGATTCAA
   LIK  GGACTCTATC TTTATTCTCG TCCAA0CCATT TATTCCAAAA ACTGATTCAA
 MARno2 GGACTCTATC TTTATTCTCG TCCAA0CCATT TATTCCAAAA ACTGATTCAA
   MAT  GGACTCTATC TTTATTCTCG TCCAA0CCATT TATTCCAAAA ACTGATTCAA
   MAX  GGACTCTATC TTTATTCTCG TCCAA0CCATT TATTCCAAAA ACTGATTCAA
   MEX  GGACTCTATC TTTATTCTCG TCCAA0CCATT TATTCCAAAA ACTGATTCAA
   MEY  GGACTCTATC TTTATTCTCG TCCAA0CCATT TATTCCAAAA ACTGATTCAA
   MON  GGACTCTATC TTTATTCTCG TCCAA0CCATT TATTCCAAAA ACTGATTCAA
   MOR  GGACTCTATC TTTATTCTCG TCCAA0CCATT TATTCCAAAA ACTGATTCAA
   OBO  GGACTCTATC TTTATTCTCG TCCAA0CCATT TATTCCAAAA ACTGATTCAA
   OMO  GGACTCTATC TTTATTCTCG TCCAA0CCATT TATTCCAAAA ACTGATTCAA
   ORI  GGACTCTATC TTTATTCTCG TCCAA0CCATT TATTCCAAAA ACTGATTCAA
 PINBANK GGACTCTATC TTTATCCTCG TCCAA0CCATT TATTCCAAAA AAGAATTAAA
 PINKOR GGACTCTATC TTTATCCTCG TCCA.1CCATT TATTCCAAAA ACTGATTCAA
   PUN  GGACTCTATC TTTATTCTCG TCCAA0CCATT TATTCCAAAA ACTGATTCAA
   PUR  GGACTCTATC TTTATTCTCG TCCAA0CCATT TATTCCAAAA ACTGATTCAA
   RET  GGACTCTATC TTTATTCTCG TCCAA0CCATT TATTCCAAAA ACTGATTCAA
   RUB  GGACTCTATC TTTATTCTCG TCCAA0CCATT TATTCCAAAA ACTGATTCAA
   SCH  GGACTCTATC TTTATTCTCG TCCAA0CCATT TATTCCAAAA ACTGATTCAA
 SITno1 GGACTCTATC TTTATTCTCG TCCAA0CCATT TATTCCAAAA ACTGATTCAA
   SMI  GGACTCTATC TTTATTCTCG TCCAA0CCATT TATTCCAAAA ACTGATTCAA
   SPI  GGACTCTATC TTTATTCTCG TCCAA0CCATT TATTCCAAAA ACTGATTCAA
   TOR  GGACTCTATC TTTATTCTCG TCCAA0CCATT TATTCCAAAA ACTGATTCAA
   WIL  GGACTCTATC TTTATTCTCG TCCAA0CCATT TATTCCAAAA ACTGATTCAA

        [601                                                650]
 ABIno1 TTCTCTATCT AAAGTAGAGA AATTCATAAT TGGATTACTT AATGTAAAAT
   ALC  TTCTCTATCT AAAGTAGAGA AGTTCATAAT TGGATTACTT AATGTAAAAT
   ASP  TTCTCTATCT AAAGTAGAGA AATTCATAAT TGGATTACTT AATGTAAAAT
   BRA  TTCTCTATCT AAAGTAGAGA AGTTCATAAT TGGATTACTT AATGTAAAAT
   BRE  TTCTCTATCT AGAGTAGAGA AGTTCATAAT TGGATTGCTT AATGTCAAAT
   CHI  TTCTCTATCT AAAGTAGAGA AGTTCATAAT TGGATTACTT AATGTAAAAT
   CRA  TTCTCTATCT AAAGTAGAGA AATTCATAAT TGGATTACTT AATGTAAAAT
 ENGno3 TTCTCTATCT AAAGTAGAGA AGTTCATAAT TGGATTACTT AATGTAAAAT
   FAR  TTCTCTATCT AAAGTAGAGA AGTTCATAAT TGGATTACTT AATGTAAAAT
 GLAno2 TTCTCTATCT AAAGTAGAGA AGTTCATAAT TGGATTACTT AATGTAAAAT
   GLE  TTCTCTATCT AAAGTAGAGA AATTCATAAT TGGATTACTT AATGTAAAAT
 JEZno1 TTCTCTATCT AAAGTAGAGA AGTTCATAAT TGGATTACTT AATGTAAAAT
   KOR  TTCTCTATCT AAAGTAGAGA AATTCATAAT TGGATTACTT AATGTAAAAT
   KOY  TTCTCTATCT AAAGTAGAGA AATTCATAAT TGGATTACTT AATGTAAAAT
   LIK  TTCTCTATCT AAAGTAGAGA AGTTCATAAT TGGATTACTT AATGTAAAAT
 MARno2 TTCTCTATCT AAAGTAGAGA AGTTCATAAT TGGATTACTT AATGTAAAAT
   MAT  TTCTCTATCT AAAGTAGAGA AGTTCATAAT TGGATTACTT AATGTAAAAT
   MAX  TTCTCTATCT AAAGTAGAGA AGTTCATAAT TGGATTACTT AATGTAAAAT
   MEX  TTCTCTATCT AAAGTAGAGA AGTTCATAAT TGGATTACTT AATGTAAAAT
   MEY  TTCTCTATCT AAAGTAGAGA AATTCATAAT TGGATTACTT AATGTAAAAT
   MON  TTCTCTATCT AAAGTAGAGA AATTCATAAT TGGATTACTT AATGTAAAAT
   MOR  TTCTCTATCT AAAGTAGAGA AGTTCATAAT TGGATTACTT AATGTAAAAT
   OBO  TTCTCTATCT AAAGTAGAGA AATTCATAAT TGGATTACTT AATGTAAAAT
   OMO  TTCTCTATCT AAAGTAGAGA AGTTCATAAT TGGATTACTT AATGTAAAAT
   ORI  TTCTCTATCT AAAGTAGAGA AGTTCATAAT TGGATTACTT AATGTAAAAT
```

```
PINBANK  TTATCCATCT AGAGTAGATA AGTTCATAAT TGGATTACTT AATGTCAAAT
 PINKOR  TTTTCCATCT AGAGTAGATA AGTTCATAAT TGGATTACTT AATGTCAAAT
    PUN  TTCTCTATCT AAAGTAGAGA AGTTCATAAT TGGATTACTT AATGTAAAAT
    PUR  TTCTCTATCT AAAGTAGAGA AGTTCATAAT TGGATTACTT AATGTAAAAT
    RET  TTCTCTATCT AAAGTAGAGA AATTCATAAT TGGATTACTT AATGTAAAAT
    RUB  TTCTCTATCT AAAGTAGAGA AGTTCATAAT TGGATTACTT AATGTAAAAT
    SCH  TTCTCTATCT AAAGTAGAGA AGTTCATAAT TGGATTACTT AATGTAAAAT
 SITno1  TTCTCTATCT AAAGTAKAGA AGTTCATAAT TGGATTACTT AATGTCAAAT
    SMI  TTCTCTATCT AAAGTAGAGA AGTTCATAAT TGGATTACTT AATGTAAAAT
    SPI  TTCTCTATCT AAAGTAGAGA AGTTCATAAT TGGATTACTT AATGTAAAAT
    TOR  TTCTCTATCT AAAGTAGAGA AATTCATAAT TGGATTACTT AATGTAAAAT
    WIL  TTCTCTATCT AAAGTAGAGA AGTTCATAAT TGGATTACTT AATGTAAAAT

         [651                                +1     +1        700]
 ABIno1  CAGTACTTCA ACTCGAATCT GGCATCTATC TTA0CGAAT.0A GATGCTTGGA
    ALC  CAGTACTTCA ACTCGAATCT GGCATCTATC TTA0CGAAT.0A GATGCTTGGA
    ASP  CAGTACTTCA ACTCGAATCT GGCATCTATC TTA0CGAAT.0A GATGCTTGGA
    BRA  CAGTACTTCA ACTCGAATCT GGCATCTATC TTA0CGAAT.0A GATGCTTGGA
    BRE  CAGTACTTCA ACTCGAATCT GGCATCTATC TTA0CGAAT.0A GCTGCTTGGA
    CHI  CAGTACTTCA ACTCGAATCT GGCATCTATC TTA0CGAAT.0A GATGCTTGGA
    CRA  CAGTACTTCA ACTCGAATCT GGCATCTATC TTA0CGAAT.0A GATGCTTGGA
 ENGno3  CAGTACTTCA ACTCGAATCT GGCATCTATC TTA0CGAAT.0A GATGCTTGGA
    FAR  CAGTACTTCA ACTCGAATCT GGCATCTATC TTA0CGAAT.0A GATGCTTGGA
 GLAno2  CAGTACTTCA ACTCGAATCT GGCATCTATC TTA0CGAAT.0A GATGCTTGGA
    GLE  CAGTACTTCA ACTCGAATCT GGCATCTATC TTA0CGAAT.0A GATGCTTGGA
 JEZno1  CAGTACTTCA ACTCGAATCT GGCATCTATC TTA0CGAAT.0A GATGCTTGGA
    KOR  CAGTACTTCA ACTCGAATCT GGCATCTATC TTA0CGAAT.0A GATGCTTGGA
    KOY  CAGTACTTCA ACTCGAATCT GGCATCTATC TTA0CGAAT.0A GATGCTTGGA
    LIK  CAGTACTTCA ACTCGAATCT GGCATCTATC TTA0CGAAT.0A GATGCTTGGA
 MARno2  CAGTACTTCA ACTCGAATCT GGCATCTATC TTA0CGAAT.0A GATGCTTGGA
    MAT  CAGTACTTCA ACTCGAATCT GGCATCTATC TTA0CGAAT.0A GATGCTTGGA
    MAX  CAGTACTTCA ACTCGAATCT GGCATCTATC TTA0CGAAT.0A GATGCTTGGA
    MEX  CAGTACTTCA ACTCGAATCT GGCATCTATC TTA0CGAAT.0A GATGCTTGGA
    MEY  CAGTACTTCA ACTCGAATCT GGCATCTATC TTA0CGAAT.0A GATGCTTGGA
    MON  CAGTACTTCA ACTCGAATCT GGCATCTATC TTA0CGAAT.0A GATGCTTGGA
    MOR  CAGTACTTCA ACTCGAATCT GGCATCTATC TTA0CGAAT.0A GATGCTTGGA
    OBO  CAGTACTTCA ACTCGAATCT GGCATCTATC TTA0CGAAT.0A GATGCTTGGA
    OMO  CAGTACTTCA ACTCGAATCT GGCATCTATC TTA0CGAAT.0A GATGCTTGGA
    ORI  CAGTACTTCA ACTCGAATCT GGCATCTATC TTA0CGAAT.0A GATGCTTGGA
PINBANK  CAGTACTTCA ACTTGAATCT GGCATCTATC TTA0TGAATA1A AATGCTTGGA
 PINKOR  CAGTACTTCA ACTCGAATCT GGCATCTATC TTA0TGAATA1A AATGCTTGGA
    PUN  CAGTACTTCA ACTCGAATCT GGCATCTATC TTA0CGAAT.0A GATGCTTGGA
    PUR  CAGTACTTCA ACTCGAATCT GGCATCTATC TTA0CGAAT.0A GATGCTTGGA
    RET  CAGTACTTCA ACTCGAATCT GGCATCTATC TTA0CGAAT.0A GATGCTTGGA
    RUB  CAGTACTTCA ACTCGAATCT GGCATCTATC TTA0CGAAT.0A GATGCTTGGA
    SCH  CAGTACTTCA ACTCGAATCT GGCATCT...  ...1CGAAT.0A GATGCTTGGA
 SITno1  CAGTACTTCA ACTCGAATCT GGCATCTATC TTA0CGAAT.0A GATGCTTGGA
    SMI  CAGTACTTCA ACTCGAATCT GGCATCTATC TTA0CGAAT.0A GATGCTTGGA
    SPI  CAGTACTTCA ACTCGAATCT GGCATCTATC TTA0CGAAT.0A GATGCTTGGA
    TOR  CAGTACTTCA ACTCGAATCT GGCATCTATC TTA0CGAAT.0A GATGCTTGGA
    WIL  CAGTACTTCA ACTCGAATCT GGCATCTATC TTA0CGAAT.0A GATGCTTGGA

         [701                                 +1                  750]
 ABIno1  ACGATTCAAG TTCTGATCGC GAGT.....0T TTGTCTGATG TTATATCACA
    ALC  ACGATTCAAG TTCTGATCGC GAGT.....0T TTGTCTGATG TTATATCACA
    ASP  ACGATTCAAG TTCTGATCGC GAGT.....0T TTGTCTGATG TTATATCACA
    BRA  ACGATTCAAG TTCTGATCGC GAGT.....0T TTGTCTGATG TTATATCACA
    BRE  ACGATTCAAG TTCTGATCGC GAGT.....0T TTGTCTGATG TTATATCACA
    CHI  ACGATTCAAG TTCTGATCGC GAGT.....0T TTGTCTGATG TTATATCACA
    CRA  ACGATTCAAG TTCTGATCGC GAGT.....0T TTGTCTGATG TTATATCACA
 ENGno3  ACGATTCAAG TTCTGATCGC GAGT.....0T TTGTCTGATG TTATATCACA
    FAR  ACGATTCAAG TTCTGATCGC GAGT.....0T TTGTCTGATG TTATATCACA
 GLAno2  ACGATTCAAG TTCTGATCGC GAGT.....0T TTGTCTGATG TTATATCACA
    GLE  ACGATTCAAG TTCTGATCGC GAGT.....0T TTGTCTGATG TTATATCACA
 JEZno1  ACGATTCAAG TTCTGATCGC GAGT.....0T TTGTCTGATG TTATATCACA
    KOR  ACGATTCAAG TTCTGATCGC GAGT.....0T TTGTCTGATG TTATATCACA
    KOY  ACGATTCAAG TTCTGATCGC GAGT.....0T TTGTCTGATG TTATATCACA
    LIK  ACGATTCAAG TTCTGATCGC GAGT.....0T TTGTCTGATG TTATATCACA
 MARno2  ACGATTCAAG TTCTGATCGC GAGT.....0T TTGTCTGATG TTATATCACA
    MAT  ACGATTCAAG TTCTGATCGC GAGT.....0T TTGTCTGATG TTATATCACA
    MAX  ACGATTCAAG TTCTGATCGC GAGT.....0T TTGTCTGATG TTATATCACA
    MEX  ACGATTCAAG TTCTGATCGC GAGT.....0T TTGTCTGATG TTATATCACA
    MEY  ACGATTCAAG TTCTGATCGC GAGT.....0T TTGTCTGATG TTATATCACA
    MON  ACGATTCAAG TTCTGATCGC GAGT.....0T TTGTCTGATG TTATATCACA
    MOR  ACGATTCAAG TTCTGATCGC GAGT.....0T TTGTCTGATG TTATATCACA
    OBO  ACGATTCAAG TTCTGATCGC GAGT.....0T TTGTCTGATG TTATATCACA
    OMO  ACGATTCAAG TTCTGATCGC GAGT.....0T TTGTCTGATG TTATATCACA
    ORI  ACGATTCAAG TTCTGATCGC GAGT.....0T TTGTCTGATG TTATATCACA
PINBANK  ACGATTCAAG TTCTGATCGC CAGT.....0T TTGTCTGGTG TTATATAACA
 PINKOR  ACGATTCAAG TTCTGATCGC CAGTCCAGT1T TTGTCTGATG TTATATAAAA
    PUN  ACGATTCAAG TTCTGATCGC GAGT.....0T TTGTCTGATG TTATATCACA
    PUR  ACGATTCAAG TTCTGATCGC GAGT.....0T TTGTCTGATG TTATATCACA
    RET  ACGATTCAAG TTCTGATCGC GAGT.....0T TTGTCTGATG TTATATCACA
    RUB  ACGATTCAAG TTCTGATCGC GAGT.....0T TTGTCTGATG TTATATCACA
```

```
   SCH  ACGATTCAAG TTCTGATCGC GAGT.....0T TTGTCTGATG TTATATCACA
SITno1  ACGATTCAAG TTCTGATCGC GAGT.....0T TTGTCTGATG TTATATCACA
   SMI  ACGATTCAAG TTCTGATCGC GAGT.....0T TTGTCTGATG TTATATCACA
   SPI  ACGATTCAAG TTCTGATCGC GAGT.....0T TTGTCTGATG TTATATCACA
   TOR  ACGATTCAAG TTCTGATCGC GAGT.....0T TTGTCTGATG TTATATCACA
   WIL  ACGATTCAAG TTCTGATCGC GAGT.....0T TTGTCTGATG TTATATCACA

        [751  +1             +1                              800]
ABIno1  T.....0TCCC ACTTTTT0AGG .TGTAAATAG AGCGTTCTAT AAATACAGTA
   ALC  T.....0TCCC ACTTTTT0AGG .TGTAAATAG AGCGTTCTAT AAATACAGTA
   ASP  T.....0TCCC ACTTTTT0AGG .TGTAAATAG AGCGTTCTAT AAATACAGTA
   BRA  T.....0TCCC ACTTTTT0AGG .TGTAAATAG AGCGTTCTAT AAATACAGTA
   BRE  T.....0TCCC ACTTTTT0AGG .TGTAAATAG AGCGTTCTAT AAATACAGTA
   CHI  T.....0TCCC ACTTTTT0AGG .TGTAAATAG AGCGTTCTAT AAATACAGTA
   CRA  T.....0TCCC ACTTTTT0AGG .TGTAAATAG AGCGTTCTAT AAATACAGTA
ENGno3  T.....0TCCC ACTTTTT0AGG .TGTAAATAG AGCGTTCTAT AAATACAGTA
   FAR  T.....0TCCC ACTTTTT0AGG .TGTAAATAG AGCGTTCTAT AAATACAGTA
GLAno2  T.....0TCCC ACTTTTT0AGG .TGTAAATAG AGCGTTCTAT AAATACAGTA
   GLE  T.....0TCCC ACTTTTT0AGG .TGTAAATAG AGCGTTCTAT AAATACAGTA
JEZno1  T.....0TCCC ACTTTTT0AGG .TGTAAATAG AGCGTTCTAT AAATACAGTA
   KOR  T.....0TCCC ACTTTTT0AGG .TGTAAATAG AGCGTTCTAT AAATACAGTA
   KOY  T.....0TCCC ACTTTTT0AGG .TGTAAATAG AGCGTTCTAT AAATACAGTA
   LIK  T.....0TCCC ACTTTTT0AGG .TGTAAATAG AGCGTTCTAT AAATACAGTA
MARno2  T.....0TCCC ACTTTTT0AGG .TGTAAATAG AGCGTTCTAT AAATACAGTA
   MAT  T.....0TCCC ACTTTTT0AGG .TGTAAATAG AGCGTTCTAT AAATACAGTA
   MAX  T.....0TCCC ACTTTTT0AGG .TGTAAATAG AGCGTTCTAT AAATACAGTA
   MEX  T.....0TCCC ACTTTTT0AGG .TGTAAATAG AGCGTTCTAT AAATACAGTA
   MEY  T.....0TCCC ACTTTTT0AGG .TGTAAATAG AGCGTTCTAT AAATACAGTA
   MON  T.....0TCCC ACTTTTT0AGG .TGTAAATAG AGCGTTCTAT AAATACAGTA
   MOR  T.....0TCCC ACTTTTT0AGG .TGTAAATAG AGCGTTCTAT AAATACAGTA
   OBO  T.....0TCCC ACTTTTT0AGG .TGTAAATAG AGCGTTCTAT AAATACAGTA
   OMO  T.....0TCCC ACTTTTT0AGG .TGTAAATAG AGCGTTCTAT AAATACAGTA
   ORI  T.....0TCCC ACTTTTT0AGG .TGTAAATAG AGCGTTCTAT AAATACAGTA
PINBANK  TCTCTC1TCTC CATTTTT0GAG GTGTAAATAG ATCGTTCTAT AACTACAGTA
 PINKOR  TATCTC1TC.C ACTTTT.1GAG GTGTAAATAG AGCGTTCTAT AACTACAGTA
   PUN  T.....0TCCC ACTTTTT0AGG .TGTAAATAG AGCGTTCTAT AAATACAGTA
   PUR  T.....0TCCC ACTTTTT0AGG .TGTAAATAG AGCGTTCTAT AAATACAGTA
   RET  T.....0TCCC ACTTTTT0AGG .TGTAAATAG AGCGTTCTAT AAATACAGTA
   RUB  T.....0TCCC ACTTTTT0AGG .TGTAAATAG AGCGTTCTAT AAATACAGTA
   SCH  T.....0TCCC ACTTTTT0AGG .TGTAAATAG AGCGTTCTAT AAATACAGTA
SITno1  T.....0TCCC ACTTTTT0AGG .TGTAAATAG AGCGTTCTAT AAATACAGTA
   SMI  T.....0TCCC ACTTTTT0AGG .TGTAAATAG AGCGTTCTAT AAATACAGTA
   SPI  T.....0TCCC ACTTTTT0AGG .TGTAAATAG AGCGTTCTAT AAATACAGTA
   TOR  T.....0TCCC ACTTTTT0AGG .TGTAAATAG AGCGTTCTAT AAATACAGTA
   WIL  T.....0TCCC ACTTTTT0AGG .TGTAAATAG AGCGTTCTAT AAATACAGTA

        [801                                                 850]
ABIno1  TTGGACCAAA TGAGATTCAT TCGTTAGAAT AGCTTCCATT GAGTCTCTGC
   ALC  TTGGACCAAA TGAGATTCAT TCGTTAGAAT AGCTTCCATT GAGTCTCTGC
   ASP  TTGGACCAAA TGAGATTCAT TCGTTAGAAT AGCTTCCATT GAGTCTCTGC
   BRA  TTGGACCAAA TGAGATTCAT TCGTTAGAAT AGCTTCCATT GAGTCTCTGC
   BRE  TTGGACCAAA TGAGATTCAT TCGTTAGAAT AGCTTCCATT GAGTCTCTGC
   CHI  TTGGACCAAA TTAGATTCAT TCGTTAGAAT AGCTTCCATT GAGTCTCTGC
   CRA  TTGGACCAAA TTAGATTCAT TCGTTAGAAT AGCTTCCATT GAGTCTCTGC
ENGno3  TTGGACCAAA TGAKATTCAT TCGTTAGAAT AGCTTCCATT GAGTCTCTGC
   FAR  TTGGACCAAA TGAGATTCAT TCGTTAGAAT AGCTTCCATT GAGTCTCTGC
GLAno2  TTGGACCAAA TGAGATTCAT TCGTTAGAAT AGCTTCCATT GAGTCTCTGC
   GLE  TTGGACCAAA TGAGATTCAT TCGTTAGAAT AGCTTCCATT GAGTCTCTGC
JEZno1  TTGGACCAAA TGAGATTCAT TCGTTAGAAT AGCTTCCATT GAGTCTCTGC
   KOR  TTGGACCAAA TGAGATTCAT TCGTTAGAAT AGCTTCCATT GAGTCTCTGC
   KOY  TTGGACCAAA TGAGATTCAT TCGTTAGAAT AGCTTCCATT GAGTCTCTGC
   LIK  TTGGACCAAA TGAGATTCAT TCGTTAGAAT AGCTTCCATT GAGTCTCTGC
MARno2  TTGGACCAAA TGAGATTCAT TCGTTAGAAT AGCTTCCATT GAGTCTCTGC
   MAT  TTGGACCAAA TGAGATTCAT TCGTTAGAAT AGCTTCCATT GAGTCTCTGC
   MAX  TTGGACCAAA TGAGATTCAT TCGTTAGAAT AGCTTCCATT GAGTCTCTGC
   MEX  TTGGACCAAA TGAGATTCAT TCGTTAGAAT AGCTTCCATT GAGTCTCTGC
   MEY  TTGGACCAAA TGAGATTCAT TCGTTAGAAT AGCTTCCATT GAGTCTCTGC
   MON  TTGGACCAAA TGAGATTCAT TCGTTAGAAT AGCTTCCATT GAGTCTCTGC
   MOR  TTGGACCAAA TGAGATTCAT TCGTTAGAAT AGCTTCCATT GAGTCTCTGC
   OBO  TTGGACCAAA TGAGATTCAT TCGTTAGAAT AGCTTCCATT GAGTCTCTGC
   OMO  TTGGACCAAA TGAGATTCAT TCGTTAGAAT AGCTTCCATT GAGTCTCTGC
   ORI  TTGGACCAAA TGAGATTCAT TCGTTAGAAT AGCTTCCATT GAGTCTCTGC
PINBANK  TTGGACCAAA TGAGATTCAT TCGTTAGAAT AGCTTCCATT GAGTCTCTGC
 PINKOR  TTGGACCAAA TGAGATTCAT TCGTTAGAAT AGCTTCCATT GAGTCTCTGC
   PUN  TTGGACCAAA TGAGATTCAT TCGTTAGAAT AGCTTCCATT GAGTCTCTGC
   PUR  TTGGACCAAA TGAGATTCAT TCGTTAGAAT AGCTTCCATT GAGTCTCTGC
   RET  TTGGACCAAA TGAGATTCAT TCGTTAGAAT AGCTTCCATT GAGTCTCTGC
   RUB  TTGGACCAAA TGAGATTCAT TCGTTAGAAT AGCTTCCATT GAGTCTCTGC
   SCH  TTGGACCAAA TGAGATTCAT TCGTTAGAAT AGCTTCCATT GAGTCTCTGC
SITno1  TTGGACCAAA TGAGATTCAT TCGTTAGAAT AGCTTCCATT GAGTCTCTGC
   SMI  TTGGACCAAA TGAGATTCAT TCGTTAGAAT AGCTTCCATT GAGTCTCTGC
   SPI  TTGGACCAAA TGAGATTCAT TCGTTAGAAT AGCTTCCATT GAGTCTCTGC
   TOR  TTGGACCAAA TGAGATTCAT TCGTTAGAAT AGCTTCCATT GAGTCTCTGC
   WIL  TTGGACCAAA TGAGATTCAT TCGTTAGAAT AGCTTCCATT GAGTCTCTGC
```

```
          [851                                                +1    900]
ABIno1    ACCTATCCCC TTCCTATCTT AGGAGAATAA AAGAAGAAAC T0ATTGTCTCC
   ALC    ACCTATCCCC TTCCTATCTT AGGAGAATAA AAGAAGAAAC T0ATTGTCTCC
   ASP    ACCTATCCCC TTCCTATCTT AGGAGAATAA AAGAAGAAAC T0ATTGTCTCC
   BRA    ACCTATCCCC TTCCTATCTT AGGAGAATAA AAGAAGAAAC T0ATTGTCTCC
   BRE    ACCTATCCCC TTCCTATCTT AGGAGAAGAA AAGAAGAAAC T0ATTGTCTCC
   CHI    ACCTATCCCC TTCCTATCTT AGGAGAATAA AAGAAGAAAC T0ATTGTCTCC
   CRA    ACCTATCCCC TTCCTATCTT AGGAGAATAA AAGAAGAAAC T0ATTGTCTCC
ENGno3    ACCTATCCCC TTCCTATCTT AGGAGAATAA AAGAAGAAAC T0ATTGTCTCC
   FAR    ACCTATCCCC TTCCTATCTT AGGAGAATAA AAGAAGAAAC T0ATTGTCTCC
GLAno2    ACCTATCCCC TTCCTATCTT AGGAGAATAA AAGAAGAAAC T0ATTGTCTCC
   GLE    ACCTATCCCC TTCCTATCTT AGGAGAATAA AAGAAGAAAC T0ATTGTCTCC
JEZno1    ACCTATCCCC TTCCTATCTT AGGAGAATAA AAGAAGAAAC T0ATTGTCTCC
   KOR    ACCTATCCCC TTCCTATCTT AGGAGAATAA AAGAAGAAAC T0ATTGTCTCC
   KOY    ACCTATCCCC TTCCTATCTT AGGAGAATAA AAGAAGAAAC T0ATTGTCTCC
   LIK    ACCTATCCCC TTCCTATCTT AGGAGAATAA AAGAAGAAAC T0ATTGTCTCC
MARno2    ACCTATCCCC TTCCTATCTT AGGAGAATAA AAGAAGAAAC T0ATTGTCTCC
   MAT    ACCTATCCCC TTCCTATCTT AGGAGAATAA AAGAAGAAAC T0ATTGTCTCC
   MAX    ACCTATCCCC TTCCTATCTT AGGAGAATAA AAGAAGAAAC T0ATTGTCTCC
   MEX    ACCTATCCCC TTCCTATCTT AGGAGAATAA AAGAAGAAAC T0ATTGTCTCC
   MEY    ACCTATCCCC TTCCTATCTT AGGAGAATAA AAGAAGAAAC T0ATTGTCTCC
   MON    ACCTATCCCC TTCCTATCTT AGGAGAATAA AAGAAGAAAC T0ATTGTCTCC
   MOR    ACCTATCCCC TTCCTATCTT AGGAGAATAA AAGAAGAAAC T0ATTGTCTCC
   OBO    ACCTATCCCC TTCCTATCTT AGGAGAATAA AAGAAGAAAC T0ATTGTCTCC
   OMO    ACCTATCCCC TTCCTATCTT AGGAGAATAA AAGAAGAAAC T0ATTGTCTCC
   ORI    ACCTATCCCC TTCCTATCTT AGGAGAATAA AAGAAGAAAC T0ATTGTCTCC
PINBANK   ACCTATCCCC TTCCTATATT AGGAGAAGAA A........C .1ATTGTCTTC
PINKOR    ACCTATCCCC TTCCTATCTT AGGAGAAGAA A........C .1ATTGTCTTC
   PUN    ACCTATCCCC TTCCTATCTT AGGAGAATAA AAGAAGAAAC T0ATTGTCTCC
   PUR    ACCTATCCCC TTCCTATCTT AGGAGAATAA AAGAAGAAAC T0ATTGTCTCC
   RET    ACCTATCCCC TTCCTATCTT AGGAGAATAA AAGAAGAAAC T0ATTGTCTCC
   RUB    ACCTATCCCC TTCCTATCTT AGGAGAATAA AAGAAGAAAC T0ATTGTCTCC
   SCH    ACCTATCCCC TTCCTATCTT AGGAGAATAA AAGAAGAAAC T0ATTGTCTCC
SITno1    ACCTATCCCC TTCCTATCTT AGGAGAAGAA AAGAAGAAAC T0ATTGTCTCA
   SMI    ACCTATCCCC TTCCTATCTT AGGAGAATAA AAGAAGAAAC T0ATTGTCTCC
   SPI    ACCTATCCCC TTCCTATCTT AGGAGAATAA AAGAAGAAAC T0ATTGTCTCC
   TOR    ACCTATCCCC TTCCTATCTT AGGAGAATAA AAGAAGAAAC T0ATTGTCTCC
   WIL    ACCTATCCCC TTCCTATCTT AGGAGAATAA AAGAAGAAAC T0ATTGTCTCC

          [901                                                950]
ABIno1    ATGAACCGGA TTTGGCTCAG GATTACCCAT TCAAAATATC CCAGGGTTCC
   ALC    ATGAACCGGA TTTGGCTCAG GATTACCCAT TCAAAATATC CCAGGGTTCC
   ASP    ATGAACCGGA TTTGGCTCAG GATTACCCAT TCAAAATATC CCAGGGTTCC
   BRA    ATGAACCGGA TTTGGCTCAG GATTACCCAT TCAAAATATC CCAGGGTTCC
   BRE    ATGAACCGGA TTTGGCTCAG GATTACCCAT TCAAAATATC CCAGGGTTCC
   CHI    ATGAACCGGA TTTGGCTCAG GATTACCCAT TCAAAATATC CCAGGGTTCC
   CRA    ATGAACCGGA TTTGGCTCAG GATTACCCAT TCAAAATATC CCAGGGTTCC
ENGno3    ATGAACCGGA TTTGGCTCAG GATTACCCAT TCAAAATATC CCAGGGTTCC
   FAR    ATGAACCGGA TTTGGCTCAG GATTACCCAT TCAAAATATC CCAGGGTTCC
GLAno2    ATGAACCGGA TTTGGCTCAG GATTACCCAT TCAAAATATC CCAGGGTTCC
   GLE    ATGAACCGGA TTTGGCTCAG GATTACCCAT TCAAAATATC CCAGGGTTCC
JEZno1    ATGAACCGGA TTTGGCTCAG GATTACCCAT TCAAAATATC CCAGGGTTCC
   KOR    ATGAACCGGA TTTGGCTCAG GATTACCCAT TCAAAATATC CCAGGGTTCC
   KOY    ATGAACCGGA TTTGGCTCAG GATTACCCAT TCAAAATATC CCAGGGTTCC
   LIK    ATGAACCGGA TTTGGCTCAG GATTACCCAT TCAAAATATC CCAGGGTTCC
MARno2    ATGAACCGGA TTTGGCTCAG GATTACCCAT TCAAAATATC CCAGGGTTCC
   MAT    ATGAACCGGA TTTGGCTCAG GATTACCCAT TCAAAATATC CCAGGGTTCC
   MAX    ATGAACCGGA TTTGGCTCAG GATTACCCAT TCAAAATATC CCAGGGTTCC
   MEX    ATGAACCGGA TTTGGCTCAG GATTACCCAT TCAAAATATC CCAGGGTTCC
   MEY    ATGAACCGGA TTTGGCTCAG GATTACCCAT TCAAAATATC CCAGGGTTCC
   MON    ATGAACCGGA TTTGGCTCAG GATTACCCAT TCAAAATATC CCAGGGTTCC
   MOR    ATGAACCGGA TTTGGCTCAG GATTACCCAT TCAAAATATC CCAGGGTTCC
   OBO    ATGAACCGGA TTTGGCTCAG GATTACCCAT TCAAAATATC CCAGGGTTCC
   OMO    ATGAACCGGA TTTGGCTCAG GATTACCCAT TCAAAATATC CCAGGGTTCC
   ORI    ATGAACCGGA TTTGGCTCAG GATTACCCAT TCAAAATATC CCAGGGTTCC
PINBANK   ATGAACCGGA TTTGGCTCAG GATTACCCAT TCAAAATATC CCAGGGTTCC
PINKOR    ATGAACCGGA TTTGGCTCAG GATTACCCAT TCAAAATATC CCAGGGTTCC
   PUN    ATGAACCGGA TTTGGCTCAG GATTACCCAT TCAAAATATC CCAGGGTTCC
   PUR    ATGAACCGGA TTTGGCTCAG GATTACCCAT TCAAAATATC CCAGGGTTCC
   RET    ATGAACCGGA TTTGGCTCAG GATTACCCAT TCAAAATATC CCAGGGTTCC
   RUB    ATGAACCGGA TTTGGCTCAG GATTACCCAT TCAAAATATC CCAGGGTTCC
   SCH    ATGAACCGGA TTTGGCTCAG GATTACCCAT TCAAAATATC CCAGGGTTCC
SITno1    ATGAACCGGA TTTGGCTCAG GATTACCCAT TCAAAATATC CCAGGGTTCC
   SMI    ATGAACCGGA TTTGGCTCAG GATTACCCAT TCAAAATATC CCAGGGTTCC
   SPI    ATGAACCGGA TTTGGCTCAG GATTACCCAT TCAAAATATC CCAGGGTTCC
   TOR    ATGAACCGGA TTTGGCTCAG GATTACCCAT TCAAAATATC CCAGGGTTCC
   WIL    ATGAACCGGA TTTGGCTCAG GATTACCCAT TCAAAATATC CCAGGGTTCC

          [951                                               1000]
ABIno1    CTGGATTTGG AAGCTATCAC TTGGTAGGTT TCCATACCAA GGCTCAATTC
   ALC    CTGGATTTGG AAGCTATCAC TTGGTAGGTT TCCATACCAA GGCTCAATTC
   ASP    CTGGATTTGG AAGCTATCAC TTGGTAGGTT TCCATACCAA GGCTCAATTC
   BRA    CTGGATTTGG AAGCTATCAC TTGGTAGGTT TCCATACCAA GGCTCAATTC
```

```
   BRE  CTGGATTTGG AAGCTATCAC TTGGTAGGTT TCCATACCAA GGCTCAATTC
   CHI  CTGGATTTGG AAGCTATCAC TTGGTAGGTT TCCATACCAA GGCTCAATTC
   CRA  CTGGATTTGG AAGCTATCAC TTGGTAGGTT TCCATACCAA GGCTCAATTC
ENGno3  CTGGATTTGG AAGCTATCAC TTGGTAGGTT TCCATACCAA GGCTCAATTC
   FAR  CTGGATTTGG AAGCTATCAC TTGGTAGGTT TCCATACCAA GGCTCAATTC
GLAno2  CTGGATTTGG AAGCTATCAC TTGGTAGGTT TCCATACCAA GGCTCAATTC
   GLE  CTGGATTTGG AAGCTATCAC TTGGTAGGTT TCCATACCAA GGCTCAATTC
JEZno1  CTGGATTTGG AAGCTATCAC TTGGTAGGTT TCCATACCAA GGCTCAATTC
   KOR  CTGGATTTGG AAGCTATCAC TTGGTAGGTT TCCATACCAA GGCTCAATTC
   KOY  CTGGATTTGG AAGCTATCAC TTGGTAGGTT TCCATACCAA GGCTCAATTC
   LIK  CTGGATTTGG AAGCTATCAC TTGGTAGGTT TCCATACCAA GGCTCAATTC
MARno2  CTGGATTTGG AAGCTATCAC TTGGTAGGTT TCCATACCAA GGCTCAATTC
   MAT  CTGGATTTGG AAGCTATCAC TTGGTAGGTT TCCATACCAA GGCTCAATTC
   MAX  CTGGATTTGG AAGCTATCAC TTGGTAGGTT TCCATACCAA GGCTCAATTC
   MEX  CTGGATTTGG AAGCTATCAC TTGGTAGGTT TCCATACCAA GGCTCAATTC
   MEY  CTGGATTTGG AAGCTATCAC TTGGTAGGTT TCCATACCAA GGCTCAATTC
   MON  CTGGATTTGG AAGCTATCAC TTGGTAGGTT TCCATACCAA GGCTCAATTC
   MOR  CTGGATTTGG AAGCTATCAC TTGGTAGGTT TCCATACCAA GGCTCAATTC
   OBO  CTGGATTTGG AAGCTATCAC TTGGTAGGTT TCCATACCAA GGCTCAATTC
   OMO  CTGGATTTGG AAGCTATCAC TTGGTAGGTT TCCATACCAA GGCTCAATTC
   ORI  CTGGATTTGG AAGCTATCAC TTGGTAGGTT TCCATACCAA GGCTCAATTC
PINBANK  CTGGATTTGG AAGCTATCAC TTGGTAGGTT TCCATACCAA GGCTCAATTC
PINKOR  CTGGATTTGG AAGCTATCAC TTGGTAGGTT TCCATACCAA GGCTCAATTC
   PUN  CTGGATTTGG AAGCTATCAC TTGGTAGGTT TCCATACCAA GGCTCAATTC
   PUR  CTGGATTTGG AAGCTATCAC TTGGTAGGTT TCCATACCAA GGCTCAATTC
   RET  CTGGATTTGG AAGCTATCAC TTGGTAGGTT TCCATACCAA GGCTCAATTC
   RUB  CTGGATTTGG AAGCTATCAC TTGGTAGGTT TCCATACCAA GGCTCAATTC
   SCH  CTGGATTTGG AAGCTATCAC TTGGTAGGTT TCCATACCAA GGCTCAATTC
SITno1  CTGGATTTGG AAGCTATCAC TTGGTAGGTT TCCATACCAA GGCTCAATTC
   SMI  CTGGATTTGG AAGCTATCAC TTGGTAGGTT TCCATACCAA GGCTCAATTC
   SPI  CTGGATTTGG AAGCTATCAC TTGGTAGGTT TCCATACCAA GGCTCAATTC
   TOR  CTGGATTTGG AAGCTATCAC TTGGTAGGTT TCCATACCAA GGCTCAATTC
   WIL  CTGGATTTGG AAGCTATCAC TTGGTAGGTT TCCATACCAA GGCTCAATTC

        [1001                                              1050]
ABIno1  GATCAAGTCC GTAGCGTCTA CCGATTTCGC CATATCCCCT TCTCGKAGAA
   ALC  GATCAAGTCC GTAGCGTCTA CCGATTTCGC CATATCCCCT TCTCGGAGAA
   ASP  GATCAAGTCC GTAGCGTCTA CCGATTTCGC CATATCCCCT TCTCGGAGAA
   BRA  GATCAAGTCC GTAGCGTCTA CCGATTTCGC CATATCCCCT TCTCGGAGAA
   BRE  GATCAAGTCC GTAGCGTCTA CCGATTTCGC CATATCCCCT TCTCGGAGAA
   CHI  GATCAAGTCC GTAGCGTCTA CCGATTTCGC CATATCCCCT TCTCGGAGAA
   CRA  GATCAAGTCC GTAGCGTCTA CCGATTTCGC CATATCCCCT TCTCGGAGAA
ENGno3  GATCAAGTCC GTAGCGTCTA CCGATTTCGC CATATCCCCT TCTCGGAGAA
   FAR  GATCAAGTCC GTAGCGTCTA CCGATTTCGC CATATCCCCT TCTCGGAGAA
GLAno2  GATCAAGTCC GTAGCGTCTA CCGATTTCGC CATATCCCCT TCTCGGAGAA
   GLE  GATCAAGTCC GTAGCGTCTA CCGATTTCGC CATATCCCCT TCTCGGAGAA
JEZno1  GATCAAGTCC GTAGCGTCTA CCGATTTCGC CATATCCCCT TCTCGGAGAA
   KOR  GATCAAGTCC GTAGCGTCTA CCGATTTCGC CATATCCCCT TCTCGGAGAA
   KOY  GATCAAGTCC GTAGCGTCTA CCGATTTCGC CATATCCCCT TCTCGGAGAA
   LIK  GATCAAGTCC GTAGCGTCTA CCGATTTCGC CATATCCCCT TCTCGGAGAA
MARno2  GATCAAGTCC GTAGCGTCTA CCGATTTCGC CATATCCCCT TCTCGGAGAA
   MAT  GATCAAGTCC GTAGCGTCTA CCGATTTCGC CATATCCCCT TCTCGGAGAA
   MAX  GATCAAGTCC GTAGCGTCTA CCGATTTCGC CATATCCCCT TCTCGGAGAA
   MEX  GATCAAGTCC GTAGCGTCTA CCGATTTCGC CATATCCCCT TCTCGGAGAA
   MEY  GATCAAGTCC GTAGCGTCTA CCGATTTCGC CATATCCCCT TCTCGGAGAA
   MON  GATCAAGTCC GTAGCGTCTA CCGATTTCGC CATATCCCCT TCTCGGAGAA
   MOR  GATCAAGTCC GTAGCGTCTA CCGATTTCGC CATATCCCCT TCTCGGAGAA
   OBO  GATCAAGTCC GTAGCGTCTA CCGATTTCGC CATATCCCCT TCTCGTAGAA
   OMO  GATCAAGTCC GTAGCGTCTA CCGATTTCGC CATATCCCCT TCTCGGAGAA
   ORI  GATCAAGTCC GTAGCGTCTA CCGATTTCGC CATATCCCCT TCTCGGAGAA
PINBANK  GATCAAGTCC GTAGCGTCTA CCGATTTCGC CATATCCCCT TCTCGAAGAG
PINKOR  GATCAAGTCC GTAGCGTCTA CCAATTCCGC CATATCCCCT TCTCGAAGAA
   PUN  GATCAAGTCC GTAGCGTCTA CCGATTTCGC CATATCCCCT TCTCGGAGAA
   PUR  GATCAAGTCC GTAGCGTCTA CCGATTTCGC CATATCCCCT TCTCGGAGAA
   RET  GATCAAGTCC GTAGCGTCTA CCGATTTCGC CATATCCCCT TCTCGGAGAA
   RUB  GATCAAGTCC GTAGCGTCTA CCGATTTCGC CATATCCCCT TCTCGGAGAA
   SCH  GATCAAGTCC GTAGCGTCTA CCGATTTCGC CATATCCCCT TCTCGGAGAA
SITno1  GATCAAGTCC GTAGCGTCTA CCGATTTCGC CATATCCCCT TCTCGGAGAA
   SMI  GATCAAGTCC GTAGCGTCTA CCGATTTCGC CATATCCCCT TCTCGGAGAA
   SPI  GATCAAGTCC GTAGCGTCTA CCGATTTCGC CATATCCCCT TCTCGGAGAA
   TOR  GATCAAGTCC GTAGCGTCTA CCGATTTCGC CATATCCCCT TCTCGGAGAA
   WIL  GATCAAGTCC GTAGCGTCTA CCGATTTCGC CATATCCCCT TCTCGGAGAA

        [1051              +1                              1100]
ABIno1  CGAGATCATA ..........0 ATCATAGGAT CTAATATCCA TCAGTGTTAT
   ALC  CGAGATCATA ..........0 ATCATAGGAT CTAATATCCA TCAGTGTTAT
   ASP  CGAGATCATA ..........0 ATCATAGGAT CTAATATCCA TCAGTGTTAT
   BRA  CGAGATCATA ..........0 ATCATAGGAT CTAATATCCA TCAGTGTTAT
   BRE  CGAGATCATA ..........0 ATCATAGGAT CTAATATCCA TCAGTGTTAT
   CHI  CGAGATCATA ..........0 ATCATAGGAT CTAATATCCA TCAGTGTTAT
   CRA  CGAGATCATA ..........0 ATCATAGGAT CTAATATCCA TCAGTGTTAT
ENGno3  CGAGATCATA ..........0 ATCATAGGAT CTAATATCCA TCAGTGTTAT
   FAR  CGAGATCATA ..........0 ATCATAGGAT CTAATATCCA TCAGTGTTAT
GLAno2  CGAGATCATA ..........0 ATCATAGGAT CTAATATCCA TCAGTGTTAT
```

```
   GLE  CGAGATCATA ..........0 ATCATAGGAT CTAATATCCA TCAGTGTTAT
 JEZno1 CGAGATCATA ..........0 ATCATAGGAT CTAATATCCA TCAGTGTTAT
   KOR  CGAGATCATA ..........0 ATCATAGGAT CTAATATCCA TCAGTGTTAT
   KOY  CGAGATCATA ..........0 ATCATAGGAT CTAATATCCA TCAGTGTTAT
   LIK  CGAGATCATA ..........0 ATCATAGGAT CTAATATCCA TCAGTGTTAT
 MARno2 CGAGATCATA ..........0 ATAATAGGAT CTAATATCCA TCAGTGTTAT
   MAT  CGAGATCATA ..........0 ATCATAGGAT CTAATATCCA TCAGTGTTAT
   MAX  CGAGATCATA ..........0 ATCATAGGAT CTAATATCCA TCAGTGTTAT
   MEX  CGAGATCATA ..........0 ATCATAGGAT CTAATATCCA TCAGTGTTAT
   MEY  CGAGATCATA ..........0 ATCATAGGAT CTAATATCCA TCAGTGTTAT
   MON  CGAGATCATA ..........0 ATCATAGGAT CTAATATCCA TCAGTGTTAT
   MOR  CGAGATCATA ..........0 ATCATAGGAT CTAATATCCA TCAGTGTTAT
   OBO  CGAGATCATA ..........0 ATCATAGGAT CTAATATCCA TCAGTGTTAT
   OMO  CGAGATCATA ..........0 ATCATAGGAT CTAATATCCA TCAGTGTTAT
   ORI  CGAGATCATA ..........0 ATCATAGGAT CTAATATCCA TCAGTGTTAT
PINBANK ACAGATCATA CCTTGATCTA1 ATCCTATTAT GTAATATCCA TCAGCATTAT
 PINKOR CAAGATCATA CCTTGTTCTA1 ATCCTATTAT GTAATATCCA TCAGCATTAT
   PUN  CGAGATCATA ..........0 ATCATAGGAT CTAATATCCA TCAGTGTTAT
   PUR  CGAGATCATA ..........0 ATCATAGGAT CTAATATCCA TCAGTGTTAT
   RET  CGAGATCATA ..........0 ATCATAGGAT CTAATATCCA TCAGTGTTAT
   RUB  CGAGATCATA ..........0 ATCATAGGAT CTAATATCCA TCAGTGTTAT
   SCH  CGAGATCATA ..........0 ATCATAGGAT CTAATATCCA TCAGTGTTAT
 SITno1 CGAGATCATA ..........0 ATCATAGGAT CTAATATCCA TCAGTGTTAT
   SMI  CGAGATCATA ..........0 ATCATAGGAT CTAATATCCA TCAGTGTTAT
   SPI  CGAGATCATA ..........0 ATCATAGGAT CTAATATCCA TCAGTGTTAT
   TOR  CGAGATCATA ..........0 ATCATAGGAT CTAATATCCA TCAGTGTTAT
   WIL  CGAGATCATA ..........0 ATCATAGGAT CTAATATCCA TCAGTGTTAT

        [1101                                    +1     +1    1150]
ABIno1 TCATTGGATA TTTGCTCAAT ATTGGGTGGG AGA....0TTC TGCCC.0TACT
   ALC  TAATTGGATA TTAGCTCAAT ATTGGGTGGG AGA....0TTC TGCCC.0TACT
   ASP  TCATTGGATA TTTGCTCAAT ATTGGGTGGG AGA....0TTC TGCCC.0TACT
   BRA  TCATTGGATA TTAGCTCAAT ATTGGGTGGG AGA....0TTC TGCCC.0TACT
   BRE  TCATTGGATA TTTGCTCAAT ATTGGGTGGG AGA....0TTC TGCCC.0TACT
   CHI  TCATTGGATA TTAGCTCAAA ATTGGGTGGG AGA....0TTC TGCCC.0TACT
   CRA  TCATTGGATA TTTGCTCAAT ATTGGGTGGG AGA....0TTC TGCCC.0TACT
ENGno3 TCATTGGATA TTTGCTCAAT ATTGGGTGGG AGA....0TTC TGCCCC1TACT
   FAR  TCATTGGATA TTTGCTCAAT ATTGGGTGGG AGA....0TTC TGCCC.0TACT
GLAno2 TCATTGGATA TTTGCTCAAT ATTGGGTGGG AGA....0TTC TGCCCC1TACT
   GLE  TCATTGGATA TTTGCTCAAT ATTGGGTGGG AGA....0TTC TGCCC.0TACT
 JEZno1 TCATTGGATA TTTGCTCAAT ATTGGGTGGG AGA....0TTC TGCCC.0TACT
   KOR  TCATTGGATA TTTGCTCAAT ATTGGGTGGG AGA....0TTC TGCCC.0TACT
   KOY  TCATTGGATA TTTGCTCAAT ATTGGGTGGG AGA....0TTC TGCCC.0TACT
   LIK  TCATTGGATA TTTGCTCAAT ATTGGGTGGG AGA....0TTC TGCCC.0TACT
 MARno2 TCATTGGATA TTTGCTCAAT ATTGGGTGGG AGA....0TTC TGCCC.0TACT
   MAT  TCATTGGATA TTAGCTCAAT ATTGGGTGGG AGA....0TTC TGCCC.0TACT
   MAX  TCATTGGATA TTAGCTCAAT ATTGGGTGGG AGA....0TTC TGCCC.0TACT
   MEX  TCATTGGATA TTTGCTCAAT ATTGGGTGGG AGA....0TTC TGCCCC1TACT
   MEY  TCATTGGATA TTTGCTCAAT ATTGGGTGGG AGA....0TTC TGCCC.0TACT
   MON  TCATTGGATA TTTGCTCAAT ATTGGGTGGG AGA....0TTC TGCCC.0TACT
   MOR  TCATTGGATA TTAGCTCAAT ATTGGGTGGG AGA....0TTC TGCCC.0TACT
   OBO  TCATTGGATA TTTGCTCAAT ATTGGGTGGG AGA....0TTC TGCCC.0TACT
   OMO  TCATTGGATA TTTGCTCAAT ATTGGGTGGG AGA....0TTC TGCCC.0TACT
   ORI  TCATTGGATA TT?GCTCAAT ATTGGGTGGG AGA....0TTC TGCCC.0TACT
PINBANK TCATTGGATA TTTGCTCAAT ATTGGGTGGG AGAAATA1TCC TCCCC.0TACT
 PINKOR TCATTGGATA TTTGCTCAAT ATTGGGTGGT AGAAATA1TCC TCCCC.0TACT
   PUN  TCATTGGATA TTTGCTCAAT ATTGGGTGGG AGA....0TTC TGCCC.0TACT
   PUR  TCATTGGATA TTAGCTCAAT ATTGGGTGGG AGA....0TTC TGCCC.0TACT
   RET  TCATTGGATA TTTGCTCAAT ATTGGGTGGG AGA....0TTC TGCCC.0TACT
   RUB  TCATTGGATA TTTGCTCAAT ATTGGGTGGG AGA....0TTC TGCCC.0TACT
   SCH  TCATTGGATA TTTGCTCAAT ATTGGGTGGG AGA....0TTC TGCCC.0TACT
 SITno1 TCATTGGATA TTTGCTCAAT ATTGGGTGGG AGA....0TTC TGCCC.0TACT
   SMI  TCATTGGATA TTTGCTCAAT ATTGGGTGGG AGA....0TTC TGCCC.0TACT
   SPI  TCATTGGATA TTTGCTCAAT ATTGGGTGGG AGA....0TTC TGCCC.0TACT
   TOR  TCATTGGATA TTTGCTCAAT ATTGGGTGGG AGA....0TTC TGCCC.0TACT
   WIL  TCATTGGATA TTAGCTCAAT ATTGGGTGGG AGA....0TTC TGCCC.0TACT

        [1151                                        +1       1200]
ABIno1 CCCCTTCTCT CGCCATCTCT .......... .......... ..........0
   ALC  CCCCTTCTCT CGCCATCTCT .......... .......... ..........0
   ASP  CCCCTTCTCT CGCCATCTCT .......... .......... ..........0
   BRA  CCCCTTCTCT CGCCATCTCT .......... .......... ..........0
   BRE  CCCCTTCTCT CGCCATCTCT .......... .......... ..........0
   CHI  CCCCTTCTCT CGCCATCTCT .......... .......... ..........0
   CRA  CCCCTTCTCT CGCCATCTCT .......... .......... ..........0
ENGno3 CCCCTTCTCT CGCCATCTCT .......... .......... ..........0
   FAR  CCCCTTCTCT CGCCATCTCT .......... .......... ..........0
GLAno2 CCCCTTCTCT CGCCATCTCT .......... .......... ..........0
   GLE  CCCCTTCTCT CGCCATCTCT .......... .......... ..........0
 JEZno1 CCCCTTCTCT CGCCATCTCT .......... .......... ..........0
   KOR  CCCCTTCTCT CGCCATCTCT .......... .......... ..........0
   KOY  CCCCTTCTCT CGCCATCTCT .......... .......... ..........0
   LIK  CCCCTTCTCT CGCCATCTCT .......... .......... ..........0
 MARno2 CCCCTTCTCT CGCCATCTCT .......... .......... ..........0
```

```
   MAT  CCCCTTCTCT CGCCATCTCT .......... .......... ..........0
   MAX  CCCCTTCTCT CGCCATCTCT .......... .......... ..........0
   MEX  CCCCTTCTCT CGCCATCTCT .......... .......... ..........0
   MEY  CCCCTTCTCT CGCCATCTCT .......... .......... ..........0
   MON  CCCCTTCTCT CGCCATCTCT .......... .......... ..........0
   MOR  CCCCTTCTCT CGCCATCTCT .......... .......... ..........0
   OBO  CCCCTTCTCT CGCCATCTCT .......... .......... ..........0
   OMO  CCCCTTCTCT CGCCATCTCT .......... .......... ..........0
   ORI  CCCCTTCTCT CGCCATCTCT .......... .......... ..........0
PINBANK CCCCTTCTCT CCCCTTCTCT CCCCTTCTCT CCCCTTCTCT CCCCTTCTCT1
PINKOR  CCCCTTCTCT CGCCATCTCT .......... .......... ..........0
   PUN  CCCCTTCTCT CGCCATCTCT .......... .......... ..........0
   PUR  CCCCTTCTCT CGCCATCTCT .......... .......... ..........0
   RET  CCCCTTCTCT CGCCATCTCT .......... .......... ..........0
   RUB  CCCCTTCTCT CGCCATCTCT .......... .......... ..........0
   SCH  CCCCTTCTCT CGCCATCTCT .......... .......... ..........0
SITno1  CCCCTTCTCT CGCCATCTCT .......... .......... ..........0
   SMI  CCCCTTCTCT CGCCATCTCT .......... .......... ..........0
   SPI  CCCCTTCTCT CGCCATCTCT .......... .......... ..........0
   TOR  CCCCTTCTCT CGCCATCTCT .......... .......... ..........0
   WIL  CCCCTTCTCT CGCCATCTCT .......... .......... ..........0

        [1201                                  +1            1250]
ABIno1  ATCTCTACCC CAATCGAATA TTACTGGCTG GT0AATCATTA TATTATTTCT
   ALC  ATCTCTACCC CAATCGAATA TTACTGGCTG GT0AATCATTA TATTATTTCT
   ASP  ATCTCTACCC CAATCGAATA TTACTGGCTG GT0AATCATTA TATTATTTCT
   BRA  ATCTCTACCC CAATCGAATA TTACTGGCTG GT0AATCATTA TATTATTTCT
   BRE  ATCTCTACCC CAATCGAATA TGACTGGCTG GT0AATCATTA TATTATTTCT
   CHI  ATCTCTACCC CAATCGAATA TTACTGGCTG GT0AATCATTA TATTATTTCT
   CRA  ATCTCTACCC CAATCGAATA TTACTGGCTG GT0AATCATTA TATTATTTCT
ENGno3  ATCTCTACCC CAATCGAATA TTACTGGCTG GT0AATCATTA TATTATTTCT
   FAR  ATCTCTACCC CAATCGAATA TTACTGGCTG GT0AATCATTA TATTATTTCT
GLAno2  ATCTCTACCC CAATCGAATA TTACTGGCTG GT0AATCATTA TATTATTTCT
   GLE  ATCTCTACCC CAATCGAATA TTACTGGCTG GT0AATCATTA TATTATTTCT
JEZno1  ATCTCTACCC CAATCGAATA TTACTGGCTG GT0AATCATTA TATTATTTCT
   KOR  ATCTCTACCC CAATCGAATA TTACTGGCTG GT0AATCATTA TATTATTTCT
   KOY  ATCTCTACCC CAATCGAATA TTACTGGCTG GT0AATCATTA TATTATTTCT
   LIK  ATCTCTACCC CAATCGAATA TTACTGGCTG GT0AATCATTA TATTATTTCT
MARno2  ATCTCTACCC CAATCGAATA TTACTGGCTG GT0AATCATTA TATTATTTCT
   MAT  ATCTCTACCC CAATCGAATA TTACTGGCTG GT0AATCATTA TATTATTTCT
   MAX  ATCTCTACCC CAATCGAATA TTACTGGCTG GT0AATCATTA TATTATTTCT
   MEX  ATCTCTACCC CAATCGAATA TTACTGGCTG GT0AATCATTA TATTATTTCT
   MEY  ATCTCTACCC CAATCGAATA TTACTGGCTG GT0AATCATTA TATTATTTCT
   MON  ATCTCTACCC CAATCGAATA TTACTGGCTG GT0AATCATTA TATTATTTCT
   MOR  ATCTCTACCC CAATCGAATA TTACTGGCTG GT0AATCATTA TATTATTTCT
   OBO  ATCTCTACCC CAATCGAATA TTACTGGCTG GT0AATCATTA TATTATTTCT
   OMO  ATCTCTACCC CAATCGAATA TTACTGGCTG GT0AATCATTA TATTATTTCT
   ORI  ATCTCTACCC CAATCGAATA TTACTGGCTG GT0AATCATTA TATTATTTCT
PINBANK ATCTCTACCC CAATCGAATA TGACTGG... ..1AATCATTA TATTATTTAT
PINKOR  ATCTCTATCC CAATCGAATA TGACTGG... .G1AATCATTA TATTATTTCT
   PUN  ATCTCTACCC CAATCGAATA TTACTGGCTG GT0AATCATTA TATTATTTCT
   PUR  ATCTCTACCC CAATCGAATA TTACTGGCTG GT0AATCATTA TATTATTTCT
   RET  ATCTCTACCC CAATCGAATA TTACTGGCTG GT0AATCATTA TATTATTTCT
   RUB  ATCTCTACCC CAATCGAATA TTACTGGCTG GT0AATCATTA TATTATTTCT
   SCH  ATCTCTACCC CAATCGAATA TTACTGGCTG GT0AATCATTA TATTATTTCT
SITno1  ATCTCTACCC CAATCGAATA TTACTGACTG GT0AATCATTA TATAATTTCT
   SMI  ATCTCTACCC CAATCGAATA TTACTGGCTG GT0AATCATTA TATTATTTCT
   SPI  ATCTCTACCC CAATCGAATA TTACTGGCTG GT0AATCATTA TATTATTTCT
   TOR  ATCTCTACCC CAATCGAATA TTACTGGCTG GT0AATCATTA TATTATTTCT
   WIL  ATCTCTACCC CAATCGAATA TTACTGGCTG GT0AATCATTA TATTATTTCT

        [1251                                          1300]
ABIno1  GCATTTCAAA TGCAATGTTA TTATCCTCCC CTAGTCGATT TGGAAGAGTG
   ALC  G?ATTTGAAA TGCAATGTTA TTATCCTCCC CTAGTCGATT TGGAAGAGTG
   ASP  GCATTTGAAA TGCAATGTTA TTATCCTCCC CTAGTCGATT TGGAAGAGTG
   BRA  GTATTTGAAA TGCAATGTTA TTATCCTCCC CTAGTCGATT TGGAAGAGTG
   BRE  GCATTTGAAA TGCAATGTTA TTATCCTCCC CTAGTCGATT TGGAAGAGTG
   CHI  GTATTTGAAA TGCAATGTTA TTATCCTCCC CTAGTCGATT TGGAAGAGTG
   CRA  GCATTTCAAA TGCAATGTTA TTATCCTCCC CTAGTCGATT TGGAAGAGTG
ENGno3  GCATTTGAAA TGCAATGTTA TTATCCTCCC CTAGTCGATT TGGAAGAGTG
   FAR  GCATTTGAAA TGCAATGTTA TTATCCTCCC CTAGTCGATT TGGAAGAGTG
GLAno2  GCATTTGAAA TGCAATGTTA TTATCCTCCC CTAGTCGATT TGGAAGAGTG
   GLE  GCATTTGAAA TGCAATGTTA TTATCCTCCC CTAGTCGATT TGGAAGAGTG
JEZno1  GCATTTGAAA TGCAATGTTA TTATCCTCCC CTAGTCGATT TGGAAGAGTG
   KOR  GCATTT?AAA TGCAATGTTA TTATCCTCCC CTAGTCGATT TGGAAGAGTG
   KOY  GCATTTCAAA TGCAATGTTA TTATCCTCCC CTAGTCGATT TGGAAGAGTG
   LIK  GCATTTGAAA TGCAATGTTA TTATCCTCCC CTAGTCGATT TGGAAGAGTG
MARno2  GCATTTGAAA TGCAATGTTA TTATCCTCCC CTAGTCGATT TGGAAGAGTG
   MAT  GTATTTGAAA TGCAATGTTA TTATCCTCCC CTAGTCGATT TGGAAGAGTG
   MAX  GTATTTGAAA TGCAATGTTA TTATCCTCCC CTAGTCGATT TGGAAGAGTG
   MEX  GCATTTGAAA TGCAATGTTA TTATCCTCCC CTAGTCGATT TGGAAGAGTG
   MEY  GCATTTCAAA TGCAATGTTA TTATCCTCCC CTAGTCGATT TGGAAGAGTG
   MON  GCATTTCAAA TGCAATGTTA TTATCCTCCC CTAGTCGATT TGGAAGAGTG
   MOR  GTATTTGAAA TGCAATGTTA TTATCCTCCC CTAGTCGATT TGGAAGAGTG
```

```
   OBO  GCATTTCAAA TGCAATGTTA TTATCCTCCC CTAGTCGATT TGGAAGAGTG
   OMO  GCATTTGAAA TGCAATGTTA TTATCCTCCC CTAGTCGATT TGGAAGAGTG
   ORI  G?ATTTGAAA TGCAATGTTA TTATCCTCCC CTAGTCGATT TGGAAGAGTG
PINBANK CCATTTGAAA TGCAATGTTA TTATCCTCCC CTAGTCGATT TGGAAGAGTG
 PINKOR CCATTTGAAA TGCAATGTTA TTATCCTCCC CTAGTCGATT TGGAAGAGTG
   PUN  GCATTTGAAA TGCAATGTTA TTATCCTCCC CTAGTCGATT TGGAAGAGTG
   PUR  GTATTTGAAA TGCAATGTTA TTATCCTCCC CTAGTCGATT TGGAAGAGTG
   RET  GCATTTCAAA TGCAATGTTA TTATCCTCCC CTAGTCGATT TGGAAGAGTG
   RUB  GCATTTGAAA TGCAATGTTA TTATCCTCCC CTAGTCGATT TGGAAGAGTG
   SCH  GCATTTGAAA TGCAATGTTA TTATCCTCCC CTAGTCGATT TGGAAGAGTG
 SITno1 GCATTTGAAA TGCAATGTTA TTATCCTCCC CTAGTCGATT TGGAAGAGTG
   SMI  GCATTTGAAA TGCAATGTTA TTATCCTCCC CTAGTCGATT TGGAAGAGTG
   SPI  GCATTTGAAA TGCAATGTTA TTATCCTCCC CTAGTCGATT TGGAAGAGTG
   TOR  GCATTTCAAA TGCAATGTTA TTATCCTCCC CTAGTCGATT TGGAAGAGTG
   WIL  GTATTTGAAA TGCAATGTTA TTATCCTCCC CTAGTCGATT TGGAAGAGTG

        [1301                                          +1       1350]
 ABIno1 AGTAAAACGA TCGATAGAAA TTCATCCTTA CTTCCCTTT0T CTTTCCCTTG
   ALC  AGTAAAACGA TCGATATAAA TTAATCCTTA CTTCCCTTT0T CTTTCCCTTG
   ASP  AGTAAAACGA TCGATATAAA TTCATCCTTA CTTCCCTTT0T CTTTCCCTTG
   BRA  AGTAAAACGA TCGATATAAA TTCATCCTTA CTTCCCTTT0T CTTTCCCTTG
   BRE  AGTAAAACGA TCGATATCAA TTCATCCTTA CTTCCCTTT0T CTTTCCCTTG
   CHI  AGTAAAACGA TCGATATAAA TTCATCCTTA CTTCCCTTT0T CTTTCCCTTG
   CRA  AGTAAAACGA TCGATAGAAA TTCATCCTTA CTTCCCTTT0T CTTTCCCTTG
 ENGno3 AGTAAAACGA TCGATATAAA TTAATCCTTA CTTCCCTTT0T CTTTCCCTTG
   FAR  AGTAAAACGA TCGATATAAA TTCATCCTTA CTTCCCTTT0T CTTTCCCTTG
 GLAno2 AGTAAAACGA TCGATATAAA TTAATCCTTA CTTCCCTTT0T CTTTCCCTTG
   GLE  AGTAAAACGA TCGATAGAAA TTCATCCTTA CTTCCCTTT0T CTTTCCCTTG
 JEZno1 AGTAAAACGA TCGATATAAA TTCATCCTTA CTTCCCTTT0T CTTTCCCTTG
   KOR  AGTAAAACGA TCGATA?AAA TTCATCCTTA CTTCCCTTT0T CTTTCCCTTG
   KOY  AGTAAAACGA TCGATAGAAA TTCATCCTTA CTTCCCTTT0T CTTTCCCTTG
   LIK  AGTAAAACGA TCGATATAAA TTCATCCTTA CTTCCCTTT0T CTTTCCCTTG
 MARno2 AGTAAAACGA TCGATATAAA TTCATCCTTA CTTCCCTTT0T CTTTCCCTTG
   MAT  AGTAAAACGA TCGATATAAA TTCATCCTTA CTTCCCTTT0T CTTTCCCTTG
   MAX  AGTAAAACGA TCGATATAAA TTCATCCTTA CTTCCCTTT0T CTTTCCCTTG
   MEX  AGTAAAACGA TCGATATAAA TTAATCCTTA CTTCCCTTT0T CTTTCCCTTG
   MEY  AGTAAAACGA TCGATAGAAA TTCATCCTTA CTTCCCTTT0T CTTTCCCTTG
   MON  AGTAAAACGA TCGATAGAAA TTCATCCTTA CTTCCCTTT0T CTTTCCCTTG
   MOR  AGTAAAACGA TCGATATAAA TTCATCCTTA CTTCCCTTT0T CTTTCCCTTG
   OBO  AGTAAAACGA TCGATAGAAA TTCATCCTTA CTTCCCTTT0T CTTTCCCTTG
   OMO  AGTAAAACGA TCGATATAAA TTCATCCTTA CTTCCCTTT0T CTTTCCCTTG
   ORI  AGTAAAACGA TCGATATAAA TTCATCCTTA CTTCCCTTT0T CTTTCCCTTG
PINBANK AGTAAAACGA TCGATATCAA TTGACCCTTA CTTCTCCTT0T CTTTCCCTTG
 PINKOR AGTAAAACGA TCGATATCAA TTGACCCTT. .........1T CTTTCCCTTG
   PUN  AGTAAAACGA TCGATATAAA TTCATCCTTA CTTCCCTTT0T CTTTCCCTTG
   PUR  AGTAAAACGA TCGATATAAA TTCATCCTTA CTTCCCTTT0T CTTTCCCTTG
   RET  AGTAAAACGA TCGATAGAAA TTCATCCTTA CTTCCCTTT0T CTTTCCCTTG
   RUB  AGTAAAACGA TCGATATAAA TTCATCCTTA CTTCCCTTT0T CTTTCCCTTG
   SCH  AGTAAAACGA TCGATATAAA TTCATCCTTA CTTCCCTTT0T CTTTCCCTTG
 SITno1 AGTAAAACGA TCGATATCAA TTCATCCTTA CTTCCCTTT0T CTTTCCCTTG
   SMI  AGTAAAACGA TCGATATAAA TTCATCCTTA CTTCCCTTT0T CTTTCCCTTG
   SPI  AGTAAAACGA TCGATA?AAA TTCATCCTTA CTTCCCTTT0T CTTTCCCTTG
   TOR  AGTAAAACGA TCGATAGAAA TTCATCCTTA CTTCCCTTT0T CTTTCCCTTG
   WIL  AGTAAAACGA TCGATATAAA TTCATCCTTA CTTCCCTTT0T CTTTCCCTTG

        [1351              +1                                  1400]
 ABIno1 AAGGAATCTA CAATCTACA0T TCAAACAATG TTCCGTTGTA ATTGTAATCT
   ALC  AAGGAATCTA CAATCTACA0T TCAAACAATG TTCCGTTGTA ATTGTAATCT
   ASP  AAGGAATCTA CAATCTACA0T TCAAACAATG TTCCGTTGTA ATTGTAATCT
   BRA  AAGGAATCTA CAATCTACA0T TCAAACAATG TTCCGTTGTA ATTGTAATCT
   BRE  AAGGAATCTA CAATCTACA0T TCAAACAATG TTCCGTTGTA ATTGTAATCT
   CHI  AAGGAATCTA CAATCTACA0T TCAAACAATG TTCCGTTGTA ATTGTAATCT
   CRA  AAGGAATCTA CAATCTACA0T TCAAACAATG TTCCGTTGTA ATTGTAATCT
 ENGno3 AAGGAATCTA CAATCTACA0T TCAAACAATG TTCCGTTGTA ATTGTAATCT
   FAR  AAGGAATCTA CAATCTACA0T TCAAACAATG TTCCGTTGTA ATTGTAATCT
 GLAno2 AAGGAATCTA CAATCTACA0T TCAAACAATG TTCCGTTGTA ATWGTAATCT
   GLE  AAGGAATCTA CAATCTACA0T TCAAACAATG TTCCGTTGTA ATTGTAATCT
 JEZno1 AAGGAATCTA CAATCTACA0T TCAAACAATG TTCCGTTGTA ATTGTAATCT
   KOR  AAGGAATCTA CAATCTACA0T TCAAACAATG TTCCGTTGTA ATTGTAATCT
   KOY  AAGGAATCTA CAATCTACA0T TCAAACAATG TTCCGTTGTA ATTGTAATCT
   LIK  AAGGAATCTA CAATCTACA0T TCAAACAATG TTCCGTTGTA ATTGTAATCT
 MARno2 AAGGAATCTA CAATCTACA0T TCAAACAATG TTCCGTTGTA ATTGTAATCT
   MAT  AAGGAATCTA CAATCTACA0T TCAAACAATG TTCCGTTGTA ATTGTAATCT
   MAX  AAGGAATCTA CAATCTACA0T TCAAACAATG TTCCGTTGTA ATTGTAATCT
   MEX  AAGGAATCTA CAATCTACA0T TCAAACAATG TTCCGTTGTA ATTGTAATCT
   MEY  AAGGAATCTA CAATCTACA0T TCAAACAATG TTCCGTTGTA ATTGTAATCT
   MON  AAGGAATCTA CAATCTACA0T TCAAACAATG TTCCGTTGTA ATTGTAATCT
   MOR  AAGGAATCTA CAATCTACA0T TCAAACAATG TTCCGTTGTA ATTGTAATCT
   OBO  AAGGAATCTA CAATCTACA0T TCAAACAATG TTCCGTTGTA ATTGTAATCT
   OMO  AAGGAATCTA CAATCTACA0T TCAAACAATG TTCCGTTGTA ATTGTAATCT
   ORI  AAGGAATCTA CAATCTACA0T TCAAACAATG TTCCGTTGTA ATTGTAATCT
PINBANK AAAGAATCTA CA.......1T TCAGACAATG TTCCGTTGTA ATCGTAATCT
 PINKOR AAAGAATCTA CA.......1T TCAAACAATG TTCCTTTGTA ATCGTAATCT
   PUN  AAGGAATCTA CAATCTACA0T TCAAACAATG TTCCGTTGTA ATTGTAATCT
```

```
   PUR  AAGGAATCTA CAATCTACA0T TCAAACAATG TTCCGTTGTA ATTGTAATCT
   RET  AAGGAATCTA CAATCTACA0T TCAAACAATG TTCCGTTGTA ATTGTAATCT
   RUB  AAGGAATCTA CAATCTACA0T TCAAACAATG TTCCGTTGTA ATTGTAATCT
   SCH  AAGGAATCTA CAATCTACA0T TCAAACAATG TTCCGTTGTA ATTGTAATCT
SITno1  AAGGAATCTA CAATCTACA0T TCAAACAATG TTCCGTTGTA ATTGTAATCT
   SMI  AAGGAATCTA CAATCTACA0T TCAAACAATG TTCCGTTGTA ATTGTAATCT
   SPI  AAGGAATCTA CAATCTACA0T TCAAACAATG TTCCGTTGTA ATTGTAATCT
   TOR  AAGGAATCTA CAATCTACA0T TCAAACAATG TTCCGTTGTA ATTGTAATCT
   WIL  AAGGAATCTA CAATCTACA0T TCAAACAATG TTCCGTTGTA ATTGTAATCT

        [1401                                                1450]
ABIno1  GGATTGCGTC ATTGAATCTG ATTCGCGCTA TAATCGTTAG GATTCCAAAA
   ALC  GGATTGCGTC ATTGAATCTG ATTCGCGCTA TAATCGTTAG GATTCCAAAA
   ASP  GGATTGCGTC ATTGAATCTG ATTCGCGCTA TAATCGTTAG GATTCCAAAA
   BRA  GGATTGCGTC ATTGAATCTG ATTCGCGCTA TAATCGTTAG GATTCCAAAA
   BRE  GGATTGCGTC ATTGAATCTG ATTCGCGCTA TAATCGTTAG GATTCCAAAA
   CHI  GGATTGCGTC ATTGAATCTG ATTCGCGCTA TAATCGTTAG GATTCCAAAA
   CRA  GGATTGCGTC ATTGAATCTG ATTCGCGCTA TAATCGTTAG GATTCCAAAA
ENGno3  GGATTGCGTC ATTGAATCTG ATTCGCGCTA TAATCGTTAG GATTCCAAAA
   FAR  GGATTGCGTC ATTGAATCTG ATTCGCGCTA TAATCGTTAG GATTCCAAAA
GLAno2  GGATTGCGTC ATTGAATCTG ATTCGCGCTA TAATCGTTAG GATTCCAAAA
   GLE  GGATTGCGTC ATTGAATCTG ATTCGCGCTA TAATCGTTAG GATTCCAAAA
JEZno1  GGATTGCGTC ATTGAATCTG ATTCGCGCTA TAATCGTTAG GATTCCAAAA
   KOR  GGATTGCGTC ATTGAATCTG ATTCGCGCTA TAATCGTTAG GATTCCAAAA
   KOY  GGATTGCGTC ATTGAATCTG ATTCGCGCTA TAATCGTTAG GATTCCAAAA
   LIK  GGATTGCGTC ATTGAATCTG ATTCGCGCTA TAATCGTTAG GATTCCAAAA
MARno2  GGATTGCGTC ATTGAATCTG ATTCGCGCTA TAATCGTTAG GATTCCAAAA
   MAT  GGATTGCGTC ATTGAATCTG ATTCGCGCTA TAATCGTTAG GATTCCAAAA
   MAX  GGATTGCGTC ATTGAATCTG ATTCGCGCTA TAATCGTTAG GATTCCAAAA
   MEX  GGATTGCGTC ATTGAATCTG ATTCGCGCTA TAATCGTTAG GATTCCAAAA
   MEY  GGATTGCGTC ATTGAATCTG ATTCGCGCTA TAATCGTTAG GATTCCAAAA
   MON  GGATTGCGTC ATTGAATCTG ATTCGCGCTA TAATCGTTAG GATTCCAAAA
   MOR  GGATTGCGTC ATTGAATCTG ATTCGCGCTA TAATCGTTAG GATTCCAAAA
   OBO  GGATTGCGTC ATTGAATCTG ATTCGCGCTA TAATCGTTAG GATTCCAAAA
   OMO  GGATTGCGTC ATTGAATCTG ATTCGCGCTA TAATCGTTAG GATTCCAAAA
   ORI  GGATTGCGTC ATTGAATCTG ATTCGCGCTA TAATCGTTAG GATTCCAAAA
PINBANK  GGATTGCGTC ATTGAATCTG ATTCGCGCTA TAATCGTTAG GATTCCAAAG
PINKOR  GGATTGCGTC ATTGAATCCG ATTCGCGCTA TAATCGTTAG GATTCCAAAG
   PUN  GGATTGCGTC ATTGAATCTG ATTCGCGCTA TAATCGTTAG GATTCCAAAA
   PUR  GGATTGCGTC ATTGAATCTG ATTCGCGCTA TAATCGTTAG GATTCCAAAA
   RET  GGATTGCGTC ATTGAATCTG ATTCGCGCTA TAATCGTTAG GATTCCAAAA
   RUB  GGATTGCGTC ATTGAATCTG ATTCGCGCTA TAATCGTTAG GATTCCAAAA
   SCH  GGATTGCGTC ATTGAATCTG ATTCGCGCTA TAATCGTTAG GATTCCAAAA
SITno1  GGATTGCGTC ATTGAATCTG ATTCACGCTA TAATCGTTAG GATTCCAAAA
   SMI  GGATTGCGTC ATTGAATCTG ATTCGCGCTA TAATCGTTAG GATTCCAAAA
   SPI  GGATTGCGTC ATTGAATCTG ATTCGCGCTA TAATCGTTAG GATTCCAAAA
   TOR  GGATTGCGTC ATTGAATCTG ATTCGCGCTA TAATCGTTAG GATTCCAAAA
   WIL  GGATTGCGTC ATTGAATCTG ATTCGCGCTA TAATCGTTAG GATTCCAAAA

        [1451                                                1500]
ABIno1  GAATCTATGG ATCTCACACC AATGAAATGA GGAGTTATAT TCCATTGAGC
   ALC  GAATCTATGG ATCTCACACC AATGAAATGA GGAGTTATAT TCCATTGAGC
   ASP  GAATCTATGG ATCTCACACC AATGAAATGA GGAGTTATAT TCCATTGAGC
   BRA  GAATCTATGG ATCTCACACC AATGAAATGA GGAGTTATAT TCCATTGAGC
   BRE  GAATCTATGG ATCTCACACC AATGAAATGA GGAGTTATAT TCCATTGAGC
   CHI  GAATCTATGG ATCTCACACC AATGAAATGA GGAGTTATAT TCCATTGAGC
   CRA  GAATCTATGG ATCTCACACC AATGAAATGA GGAGTTATAT TCCATTGAGC
ENGno3  GAATCTATGG ATCTCACACC AATGAAATGA GGAGTTATAT TCCATTGAGC
   FAR  GAATCTATGG ATCTCACACC AATGAAATGA GGAGTTATAT TCCATTGAGC
GLAno2  GAATCTATGG ATCTCACACC AATGAAATGA GGAGTTATAT TCCATTGAGC
   GLE  GAATCTATGG ATCTCACACC AATGAAATGA GGAGTTATAT TCCATTGAGC
JEZno1  GAATCTATGG ATCTCACACC AATGAAATGA GGAGTTATA? TCCATTGAGC
   KOR  GAATCTATGG ATCTCACACC AATGAAATGA GGAGTTATAT TCCATTGAGC
   KOY  GAATCTATGG ATCTCACACC AATGAAATGA GGAGTTATAT TCCATTGAGC
   LIK  GAATCTATGG ATCTCACACC AATGAAATGA GGAGTTATAT TCCATTGAGC
MARno2  GAATCTATGG ATCTCACACC AATGAAATGA GGAGTTATAG TCCATTGAGC
   MAT  GAATCTATGG ATCTCACACC AATGAAATGA GGAGTTATAT TCCATTGAGC
   MAX  GAATCTATGG ATCTCACACC AATGAAATGA GGAGTTATAT TCCATTGAGC
   MEX  GAATCTATGG ATCTCACACC AATGAAATGA GGAGTTATAT TCCATTGAGC
   MEY  GAATCTATGG ATCTCACACC AATGAAATGA GGAGTTATAT TCCATTGAGC
   MON  GAATCTATGG ATCTCACACC AATGAAATGA GGAGTTATAT TCCATTGAGC
   MOR  GAATCTATGG ATCTCACACC AATGAAATGA GGAGTTATAT TCCATTGAGC
   OBO  GAATCTATGG ATCTCACACC AATGAAATGA GGAGTTATAT TCCATTGAGC
   OMO  GAATCTATGG ATCTCACACC AATGAAATGA GGAGTTATAG TCCATTGAGC
   ORI  GAATCTATGG ATCTCACACC AATGAAATGA GGAGTTATAT TCCATTGAGC
PINBANK  GAATCTATGG ATCTCACCCC AATGAAATGG GGAATGATAT TCCATCGAGC
PINKOR  GAATCTATGG ATCTCACACC AATGAAATGA GGAGTTATAT TCCATCGAGC
   PUN  GAATCTATGG ATCTCACACC AATGAAATGA GGAGTTATAT TCCATTGAGC
   PUR  GAATCTATGG ATCTCACACC AATGAAATGA GGAGTTATAT TCCATTGAGC
   RET  GAATCTATGG ATCTCACACC AATGAAATGA GGAGTTATAT TCCATTGAGC
   RUB  GAATCTATGG ATCTCACACC AATGAAATGA GGAGTTATAG TCCATTGAGC
   SCH  GAATCTATGG ATCTCACACC AATGAAATGA GGAGTTATAT TCCATTGAGC
SITno1  GAATCTATGG ATCTCACACC AATGAAATGA GGAGTTATAT TCCATTGAGC
   SMI  GAATCTATGG ATCTCACACC AATGAAATGA GGAGTTATAT TCCATTGAGC
```

```
    SPI  GAATCTATGG ATCTCACACC AATGAAATGA GGAGTTATAT TCCATTGAGC
    TOR  GAATCTATGG ATCTCACACC AATGAAATGA GGAGTTATAT TCCATTGAGC
    WIL  GAATCTATGG ATCTCACACC AATGAAATGA GGAGTTATAT TCCATTGAGC

         [1501           1519]
 ABIno1  CTGCTTAGCT CAGAGGTTA
    ALC  CTGCTTAGCT CAGAGGTTA
    ASP  CTGCTTAGCT CAGAGGTTA
    BRA  CTGCTTAGCT CAGAGGTTA
    BRE  CTGCTTAGCT CAGAGGTTA
    CHI  CTGCTTAGCT CAGAGGTTA
    CRA  CTGCTTAGCT CAGAGGTTA
 ENGno3  CTGCTTAGCT CAGAGGTTA
    FAR  CTGCTTAGCT CAGAGGTTA
 GLAno2  CTGCTTAGCT CAGAGGTTA
    GLE  CTGCTTAGCT CAGAGGTTA
 JEZno1  CTGCTTAGCT CAGAGGTTA
    KOR  CTGCTTAGCT CAGAGGTTA
    KOY  CTGCTTAGCT CAGAGGTTA
    LIK  CTGCTTAGCT CAGAGGTTA
 MARno2  CTGCTTAGCT CAGAGGTTA
    MAT  CTGCTTAGCT CAGAGGTTA
    MAX  CTGCTTAGCT CAGAGGTTA
    MEX  CTGCTTAGCT CAGAGGTTA
    MEY  CTGCTTAGCT CAGAGGTTA
    MON  CTGCTTAGCT CAGAGGTTA
    MOR  CTGCTTAGCT CAGAGGTTA
    OBO  CTGCTTAGCT CAGAGGTTA
    OMO  CTGCTTAGCT CAGAGGTTA
    ORI  CTGCTTAGCT CAGAGGTTA
PINBANK  CTGCTTAGCT CAGAGGTTA
 PINKOR  CTGCTTAGCT CAGAGGTTA
    PUN  CTGCTTAGCT CAGAGGTTA
    PUR  CTGCTTAGCT CAGAGGTTA
    RET  CTGCTTAGCT CAGAGGTTA
    RUB  CTGCTTAGCT CAGAGGTTA
    SCH  CTGCTTAGCT CAGAGGTTA
 SITno1  CTGCTTAGCT CAGAGGTTA
    SMI  CTGCTTAGCT CAGAGGTTA
    SPI  CTGCTTAGCT CAGAGGTTA
    TOR  CTGCTTAGCT CAGAGGTTA
    WIL  CTGCTTAGCT CAGAGGTTA
;
endblock;

begin sets;
        charset rbcL = 1-1351;
        charset trnK = 1352-3887;
        charset trnTLF = 3888-.;
        charpartition rbcLtrnK = rbcL: 1-1351, trnK: 1352-3887;
        charpartition rbcLtrnTLF = rbcL: 1-1351, trnTLF: 3888-.;
        charpartition trnKtrnTLF = trnK: 1352-3887, trnTLF: 3888-.;
        charpartition rbcLtrnKtrnTLF = rbcL: 1-1351, trnK: 1352-3887, trnTLF: 3888-.;
end;


-------------------------------------------------


#NEXUS

Begin data;
        Dimensions ntax=37 nchar=3224;
        Format datatype=dna gap=. missing=? interleave symbols="01";
        Options gapmode=missing;
        Matrix

[nad1 nombre de sites d'indels en plus: 44; termine à 3222]
        [1                                                  50]
 ABIno1  CGCAGCATGG GTAATGTCCG CGTCTCGCCG CAAGGCTAAT TTTGAGTTGC
    ALC  CGCAGCATGG GTAATGTCCG CGTCTCGCCG CAAGGCTAAT TTTGAGTTGC
    ASP  CGCAGCATGG GTAATGTCCG CGTCTCGCCG CAAGGCTAAT TTTGAGTTGC
    BRA  CGCAGCATGG GTAATGTCCG CGTCTCGCCG CAAGGCTAAT TTTGAGTTGC
    BRE  CGCAGCATGG GTAATGTCCG CGTCTCGCCG CAAGGCTAAT TTTGAGTTGC
    CHI  CGCAGCATGG GTAATGTCCG CGTCTCGCCG CAAGGCTAAT TTTGAGTTGC
    CRA  CGCAGCATGG GTAATGTCCG CGTCTCGCCG CAAGGCTAAT TTTGAGTTGC
 ENGno1  CGCAGCATGG GTAATGTCCG CGTCTCGCCG CAAGGCTAAT TTTGAGTTGC
    FAR  CGCAGCATGG GTAATGTCCG CGTCTCGCCG CAAGGCTAAT TTTGAGTTGC
 GLAno1  CGCAGCATGG GTAATGTCCG CGTCTCGCCG CAAGGCTAAT TTTGAGTTGC
    GLE  CGCAGCATGG GTAATGTCCG CGTCTCGCCG CAAGGCTAAT TTTGAGTTGC
 JEZno3  CGCAGCATGG GTAATGTCCG CGTCTCGCCG CAAGGCTAAT TTTGAGTTGC
    KOR  CGCAGCATGG GTAATGTCCG CGTCTCGCCG CAAGGCTAAT TTTGAGTTGC
```

```
   KOY  CGCAGCATGG GTAATGTCCG CGTCTCGCCG CAAGGCTAAT TTTGAGTTGC
   LIK  CGCAGCATGG GTAATGTCCG CGTCTCGCCG CAAGGCTAAT TTTGAGTTGC
MARno4  CGCAGCATGG GTAATGTCCG CGTCTCGCCG CAAGGCTAAT TTTGAGTTGC
   MAT  CGCAGCATGG GTAATGTCCG CGTCTCGCCG CAAGGCTAAT TTTGAGTTGC
   MAX  CGCAGCATGG GTAATGTCCG CGTCTCGCCG CAAGGCTAAT TTTGAGTTGC
   MEX  CGCAGCATGG GTAATGTCCG CGTCTCGCCG CAAGGCTAAT TTTGAGTTGC
   MEY  CGCAGCATGG GTAATGTCCG CGTCTCGCCG CAAGGCTAAT TTTGAGTTGC
   MON  CGCAGCATGG GTAATGTCCG CGTCTCGCCG CAAGGCTAAT TTTGAGTTGC
   MOR  CGCAGCATGG GTAATGTCCG CGTCTCGCCG CAAGGTTAAT TTTGAGTTGC
   OBO  CGCAGCATGG GTAATGTCCG CGTCTCGCCG CAAGGCTAAT TTTGAGTTGC
   OMO  CGCAGCATGG GTAATGTCCG CGTCTCGCCG CAAGGCTAAT TTTGAGTTGC
   ORI  CGCAGCATGG GTAATGTCCG CGTCTCGCCG CAAGGCTAAT TTTGAGTTGC
PINBANK  CGCAGCATGG GTAATGTCCG CGTCTCGCCG CAAGGCTAAT TTTGAGTTGC
PINKOR  CGCAGCATGG GTAATGTCCG CGTCTCGCCG CAAGGCTAAT TTTGAGTTGC
PUNno1  CGCAGCATGG GTAATGTCCG CGTCTCGCCG CAAGGCTAAT TTTGAGTTGC
   PUR  CGCAGCATGG GTAATGTCCG CGTCTCGCCG CAAGGCTAAT TTTGAGTTGC
   RET  CGCAGCATGG GTAATGTCCG CGTCTCGCCG CAAGGCTAAT TTTGAGTTGC
   RUB  CGCAGCATGG GTAATGTCCG CGTCTCGCCG CAAGGCTAAT TTTGAGTTGC
   SCH  CGCAGCATGG GTAATGTCCG CGTCTCGCCG CAAGGCTAAT TTTGAGTTGC
SITno1  CGCAGCATGG GTAATGTCCG CGTCTCGCCG CAAGGCTAAT TTTGAGTTGC
   SMI  CGCAGCATGG GTAATGTCCG CGTCTCGCCG CAAGGCTAAT TTTGAGTTGC
   SPI  CGCAGCATGG GTAATGTCCG CGTCTCGCCG CAAGGCTAAT TTTGAGTTGC
   TOR  CGCAGCATGG GTAATGTCCG CGTCTCGCCG CAAGGCTAAT TTTGAGTTGC
   WIL  CGCAGCATGG GTAATGTCCG CGTCTCGCCG CAAGGCTAAT TTTGAGTTGC

       [51                                                 100]
ABInol  AACAACAGGG TTGAGGCCCC TTCCGGAAAA GGCAGTCCGG CGGCACCAAG
   ALC  AACAACAGGG TTGAGGCCCC TTCCGGAAAA GGCAGTCCGG CGGCACCAAG
   ASP  AACAACAGGG TTGAGGCCCC TTCCGGAAAA GGCAGTCCGG CGGCACCAAG
   BRA  AACAACAGGG TTGAGGCCCC TTCCGGAAAA GGCAGTCCGG CGGCACCAAG
   BRE  AACAACAGGG TTGAGGCCCC TTCCGGAAAA GGCAGTCCGG CGGCACCAAG
   CHI  AACAACAGGG TTGAGGCCCC TTCCGGAAAA GGCAGTCCGG CGGCACCAAG
   CRA  AACAACAGGG TTGAGGCCCC TTCCGGAAAA GGCAGTCCGG CGGCACCAAG
ENGno1  AACAACAGGG TTGAGGCCCC TTCCGGAAAA GGCAGTCCGG CGGCACCAAG
   FAR  AACAACAGGG TTGAGGCCCC TTCCGGAAAA GGCAGTCCGG CGGCACCAAG
GLAno1  AACAACAGGG TTGAGGCCCC TTCCGGAAAA GGCAGTCCGG CGGCACCAAG
   GLE  AACAACAGGG TTGAGGCCCC TTCCGGAAAA GGCAGTCCGG CGGCACCAAG
JEZno3  AACAACAGGG TTGAGGCCCC TTCCGGAAAA GGCAGTCCGG CGGCACCAAG
   KOR  AACAACAGGG TTGAGGCCCC TTCCGGAAAA GGCAGTCCGr CGGCACCAAG
   KOY  AACAACAGGG TTGAGGCCCC TTCCGGAAAA GGCAGTCCGG CGGCACCAAG
   LIK  AACAACAGGG TTGAGGCCCC TTCCGGAAAA GGCAGTCCGG CGGCACCAAG
MARno4  AACAACAGGG TTGAGGCCCC TTCCGGAAAA GGCAGTCCGG CGGCACCAAG
   MAT  AACAACAGGG TTGAGGCCCC TTCCGGAAAA GGCAGTCCGG CGGCACCAAG
   MAX  AACAACAGGG TTGAGGCCCC TTCCGGAAAA GGCAGTCCGG CGGCACCAAG
   MEX  AACAACAGGG TTGAGGCCCC TTCCGGAAAA GGCAGTCCGG CGGCACCAAG
   MEY  AACAACAGGG TTGAGGCCCC TTCCGGAAAA GGCAGTCCGG CGGCACCAAG
   MON  AACAACAGGG TTGAGGCCCC TTCCGGAAAA GGCAGTCCGG CGGCACCAAG
   MOR  AACAACAGGG TTGAGGCCCC TTCCGGAAAA GGCAGTCCGG CGGCACCAAG
   OBO  AACAACAGGG TTGAGGCCCC TTCCGGAAAA GGCAGTCCGG CGGCACCAAG
   OMO  AACAACAGGG TTGAGGCCCC TTCCGGAAAA GGCAGTCCGG CGGCACCAAG
   ORI  AACAACAGGG TTGAGGCCCC TTCCGGAAAA GGCAGTCCGG CGGCACCAAG
PINBANK  AACAACAGGG TTGAGGCCCC TTCCGGAAAA GGCAGTCCGG CGGCACCAAG
PINKOR  AACAACAGGG TTGAGGCCCC TTCCGGAAAA GGCAGTCCGG CGGCACCAAG
PUNno1  AACAACAGGG TTGAGGCCCC TTCCGGAAAA GGCAGTCCGG CGGCACCAAG
   PUR  AACAACAGGG TTGAGGCCCC TTCCGGAAAA GGCAGTCCGG CGGCACCAAG
   RET  AACAACAGGG TTGAGGCCCC TTCCGGAAAA GGCAGTCCGG CGGCACCAAG
   RUB  AACAACAGGG TTGAGGCCCC TTCCGGAAAA GGCAGTCCGG CGGCACCAAG
   SCH  AACAACAGGG TTGAGGCCCC TTCCGGAAAA GGCAGTCCGG CGGCACCAAG
SITno1  AACAACAGGG TTGAGGCCCC TTCCGGAAAA GGCAGTCCGG CGGCACCAAG
   SMI  AACAACAGGG TTGAGGCCCC TTCCGGAAAA GGCAGTCCGG CGGCACCAAG
   SPI  AACAACAGGG TTGAGGCCCC TTCCGGAAAA GGCAGTCCGG CGGCACCAAG
   TOR  AACAACAGGG TTGAGGCCCC TTCCGGAAAA GGCAGTCCGG CGGCACCAAG
   WIL  AACAACAGGG TTGAGGCCCC TTCCGGAAAA GGCAGTCCGG CGGCACCAAG

       [101                                                150]
ABInol  GGTCCTGGTA CTATCCAGGT GCGAATAACC CCCGGGGGTG GCTGCAATGG
   ALC  GGTCCTGGTA CTATCCAGGT GCGAATAACC CCCGGGGGTG GCTGCAATGG
   ASP  GGTCCTGGTA CTATCCAGGT GCGAATAACC CCCGGGGGTG GCTGCAATGG
   BRA  GGTCCTGGTA CTATCCAGGT GCGAATAACC CCCGGGGGTG GCTGCAATGG
   BRE  GGTCCTGGTA CTATCCAGGT GCGAATAACC CCCGGGGGTG GCTGCAATGG
   CHI  GGTCCTGGTA CTATCCAGGT GCGAATAACC CCCGGGGGTG GCTGCAATGG
   CRA  GGTCCTGGTA CTATCCAGGT GCGAATAACC CCCGGGGGTG GCTGCAATGG
ENGno1  GGTCCTGGTA CTATCCAGGT GCGAATAACC CCCGGGGGTG GCTGCAATGG
   FAR  GGTCCTGGTA CTATCCAGGT GCGAATAACC CCCGGGGGTG GCTGCAATGG
GLAno1  GGTCCTGGTA CTATCCAGGT GCGAATAACC CCCGGGGGTG GCTGCAATGG
   GLE  GGTCCTGGTA CTATCCAGGT GCGAATAACC CCCGGGGGTG GCTGCAATGG
JEZno3  GGTCCTGGTA CTATCCAGGT GCGAATAACC CCCGGGGGTG GCTGCAATGG
   KOR  GGTCCTGGTA CTATCCAGGT GCGAATAACC CCCGGGGGTs GCTGCAATGG
   KOY  GGTCCTGGTA CTATCCAGGT GCGAATAACC CCCGGGGGTG GCTGCAATGG
   LIK  GGTCCTGGTA CTATCCAGGT GCGAATAACC CCCGGGGGTG GCTGCAATGG
MARno4  GGTCCTGGTA CTATCCAGGT GCGAATAACC CCCGGGGGTG GCTGCAATGG
   MAT  GGTCCTGGTA CTATCCAGGT GCGAATAACC CCCGGGGGTG GCTGCAATGG
   MAX  GGTCCTGGTA CTATCCAGGT GCGAATAACC CCCGGGGGTG GCTGCAATGG
   MEX  GGTCCTGGTA CTATCCAGGT GCGAATAACC CCCGGGGGTG GCTGCAATGG
```

```
   MEY  GGTCCTGGTA CTATCCAGGT GCGAATAACC CCCGGGGGTG GCTGCAATGG
   MON  GGTCCTGGTA CTATCCAGGT GCGAATAACC CCCGGGGGTG GCTGCAATGG
   MOR  GGTCCTGGTA CTATCCAGGT GCGAATAACC CCCGGGGGTG GCTGCAATGG
   OBO  GGTCCTGGTA CTATCCAGGT GCGAATAACC CCCGGGGGTG GCTGCAATGG
   OMO  GGTCCTGGTA CTATCCAGGT GCGAATAACC CCCGGGGGTG GCTGCAATGG
   ORI  GGTCCTGGTA CTATCCAGGT GCGAATAACC CCCGGGGGTG GCTGCAATGG
PINBANK GGTCCTGGTA CTATCCAGGT GCGAATAACC CCCGGGGGTG GCTGCAATGG
 PINKOR GGTCCTGGTA CTATCCAGGT GCGAATAACC CCCGGGGGTG GCTGCAATGG
 PUNno1 GGTCCTGGTA CTATCCAGGT GCGAATAACC CCCGGGGGTG GCTGCAATGG
   PUR  GGTCCTGGTA CTATCCAGGT GCGAATAACC CCCGGGGGTG GCTGCAATGG
   RET  GGTCCTGGTA CTATCCAGGT GCGAATAACC CCCGGGGGTG GCTGCAATGG
   RUB  GGTCCTGGTA CTATCCAGGT GCGAATAACC CCCGGGGGTG GCTGCAATGG
   SCH  GGTCCTGGTA CTATCCAGGT GCGAATAACC CCCGGGGGTG GCTGCAATGG
 SITno1 GGTCCTGGTA CTATCCAGGT GCGAATAACC CCCGGGGGTG GCTGCAATGG
   SMI  GGTCCTGGTA CTATCCAGGT GCGAATAACC CCCGGGGGTG GCTGCAATGG
   SPI  GGTCCTGGTA CTATCCAGGT GCGAATAACC CCCGGGGGTG GCTGCAATGG
   TOR  GGTCCTGGTA CTATCCAGGT GCGAATAACC CCCGGGGGTG GCTGCAATGG
   WIL  GGTCCTGGTA CTATCCAGGT GCGAATAACC CCCGGGGGTG GCTGCAATGG

        [151                                                200]
 ABIno1 GCAGAGATGC CACTCACCGG CCTAAACGAC GAGCAAACAC TCGAACGTGA
   ALC  GCAGAGATGC CACTCACCGG CCTAAACGAC GAGCAAACAC TCGAACGTGA
   ASP  GCAGAGATGC CACTCACCGG CCTAAACGAC GAGCAAACAC TCGAACGTGA
   BRA  GCAGAGATGC CACTCACCGG CCTAAACGAC GAGCAAACAC TCGAACGTGA
   BRE  GCAGAGATGC CACTCACCGG CCTAAACGAC GAGCAAACAC TCGAACGTGA
   CHI  GCAGAGATGC CACTCACCGG CCTAAACGAC GAGCAAACAC TCGAACGTGA
   CRA  GCAGAGATGC CACTCACCGG CCTAAACGAC GAGCAAACAC TCGAACGTGA
 ENGno1 GCAGAGATGC CACTCACCGG CCTAAACGAC GAGCAAACAC TCGAACGTGA
   FAR  GCAGAGATGC CACTCACCGG CCTAAACGAC GAGCAAACAC TCGAACGTGA
 GLAno1 GCAGAGATGC CACTCACCGG CCTAAACGAC GAGCAAACAC TCGAACGTGA
   GLE  GCAGAGATGC CACTCACCGG CCTAAACGAC GAGCAAACAC TCGAACGTGA
 JEZno3 GCAGAGATGC CACTCACCGG CCTAAACGAC GAGCAAACAC TCGAACGTGA
   KOR  GCAGAGATGC CACTCACCGG CCTAAACGAC GAGCAAACAC TCGAACGTGA
   KOY  GCAGAGATGC CACTCACCGG CCTAAACGAC GAGCAAACAC TCGAACGTGA
   LIK  GCAGAGATGC CACTCACCGG CCTAAACGAC GAGCAAACAC TCGAACGTGA
 MARno4 GCAGAGATGC CACTCACCGG CCTAAACGAC GAGCAAACAC TCGAACGTGA
   MAT  GCAGAGATGC CACTCACCGG CCTAAACGAC GAGCAAACAC TCGAACGTGA
   MAX  GCAGAGATGC CACTCACCGG CCTAAACGAC GAGCAAACAC TCGAACGTGA
   MEX  GCAGAGATGC CACTCACCGG CCTAAACGAC GAGCAAACAC TCGAACGTGA
   MEY  GCAGAGATGC CACTCACCGG CCTAAACGAC GAGCAAACAC TCGAACGTGA
   MON  GCAGAGATGC CACTCACCGG CCTAAACGAC GAGCAAACAC TCGAACGTGA
   MOR  GCAGAGATGC CACTCACCGG CCTAAACGAC GAGCAAACAC TCGAACGTGA
   OBO  GCAGAGATGC CACTCACCGG CCTAAACGAC GAGCAAACAC TCGAACGTGA
   OMO  GCAGAGATGC CACTCACCGG CCTAAACGAC GAGCAAACAC TCGAACGTGA
   ORI  GCAGAGATGC CACTCACCGG CCTAAACGAC GAGCAAACAC TCGAACGTGA
PINBANK GCAGAGATGC CACTCATCGG CCTAAACGAC GAGCAAACAC TCGAACGTGA
 PINKOR GCAGAGATGC CACTCACCGG CCTAAACGAC GAGCAAACAC TCGAACGTGA
 PUNno1 GCAGAGATGC CACTCACCGG CCTAAACGAC GAGCAAACAC TCGAACGTGA
   PUR  GCAGAGATGC CACTCACCGG CCTAAACGAC GAGCAAACAC TCGAACGTGA
   RET  GCAGAGATGC CACTCACCGG CCTAAACGAC GAGCAAACAC TCGAACGTGA
   RUB  GCAGAGATGC CACTCACCGG CCTAAACGAC GAGCAAACAC TCGAACGTGA
   SCH  GCAGAGATGC CACTCACCGG CCTAAACGAC GAGCAAACAC TCGAACGTGA
 SITno1 GCAGAGATGC CACTCACCGG CCTAAACGAC GAGCAAACAC TCGAACGTGA
   SMI  GCAGAGATGC CACTCACCGG CCTAAACGAC GAGCAAACAC TCGAACGTGA
   SPI  GCAGAGATGC CACTCACCGG CCTAAACGAC GAGCAAACAC TCGAACGTGA
   TOR  GCAGAGATGC CACTCACCGG CCTAAACGAC GAGCAAACAC TCGAACGTGA
   WIL  GCAGAGATGC CACTCACCGG CCTAAACGAC GAGCAAACAC TCGAACGTGA

        [201                                                250]
 ABIno1 GAGCGAGGGA TCACCCAACG AATGGACTCT ACAGCGACAG GAACCATG..
   ALC  GAGCGAGGGA TCACCCAACG AATGGACTCT ACAGCGACAG GAACCATG..
   ASP  GAGCGAGGGA TCACCCAACG AATGGACTCT ACAGCGACAG GAACCATG..
   BRA  GAGCGAGGGA TCACCCAACG AATGGACTCT ACAGCGACAG GAACCATGTG
   BRE  GAGCGAGGGA TCACCCAACG AATGGACTCT ACAGCGACAG GAACCATG..
   CHI  GAGCGAGGGA TCACCCAACG AATGGACTCT ACAGCGACAG GAACCATG..
   CRA  GAGCGAGGGA TCACCCAACG AATGGACTCT ACAGCGACAG GAACCATG..
 ENGno1 GAGCGAGGGA TCACCCAACG AATGGACTCT ACAGCGACAG GAACCATG..
   FAR  GAGCGAGGGA TCACCCAACG AATGGACTCT ACAGCGACAG GAACCATG..
 GLAno1 GAGCGAGGGA TCACCCAACG AATGGACTCT ACAGCGACAG GAACCATG..
   GLE  GAGCGAGGGA TCACCCAACG AATGGACTCT ACAGCGACAG GAACCATG..
 JEZno3 GAGCGAGGGA TCACCCAACG AATGGACTCT ACAGCGACAG GAACCATG..
   KOR  GAGCGAGGGA TCACCCAACG AATGGACTCT ACAGCGACAG GAACCATG..
   KOY  GAGCGAGGGA TCACCCAACG AATGGACTCT ACAGCGACAG GAACCATG..
   LIK  GAGCGAGGGA TCACCCAACG AATGGACTCT ACAGCGACAG GAACCATG..
 MARno4 GAGCGAGGGA TCACCCAACG AATGGACTCT ACAGCGACAG GAACCATG..
   MAT  GAGCGAGGGA TCACCCAACG AATGGACTCT ACAGCGACAG GAACCATG..
   MAX  GAGCGAGGGA TCACCCAACG AATGGACTCT ACAGCGACAG GAACCATG..
   MEX  GAGCGAGGGA TCACCCAACG AATGGACTCT ACAGCGACAG GAACCATG..
   MEY  GAGCGAGGGA TCACCCAACG AATGGACTCT ACAGCGACAG GAACCATG..
   MON  GAGCGAGGGA TCACCCAACG AATGGACTCT ACAGCGACAG GAACCATG..
   MOR  GAGCGAGGGA TCACCCAACG AATGGACTCT ACAGCGACAG GAACCATG..
   OBO  GAGCGAGGGA TCACCCAACG AATGGACTCT ACAGCGACAG GAACCATG..
   OMO  GAGCGAGGGA TCACCCAACG AATGGACTCT ACAGCGACAG GAACCATG..
   ORI  GAGCGAGGGA TCACCCAACG AATGGACTCT ACAGCGACAG GAACCATG..
```

```
PINBANK  GAGCGAGGGA TCACCCAACG AATGGACTCT ACAGCGACAG GAACCATG..
 PINKOR  GAGCGAGGGA TCACCCAACG AATGGACTCT ACAGCGACAG GAACCATG..
 PUNnol  GAGCGAGGGA TCACCCAACG AATGGACTCT ACAGCGACAG GAACCATG..
    PUR  GAGCGAGGGA TCACCCAACG AATGGACTCT ACAGCGACAG GAACCATG..
    RET  GAGCGAGGGA TCACCCAACG AATGGACTCT ACAGCGACAG GAACCATG..
    RUB  GAGCGAGGGA TCACCCAACG AATGGACTCT ACAGCGACAG GAACCATG..
    SCH  GAGCGAGGGA TCACCCAACG AATGGACTCT ACAGCGACAG GAACCATG..
 SITnol  GAGCGAGGGA TCACCCAACG AATGGACTCT ACAGCGACAG GAACCATG..
    SMI  GAGCGAGGGA TCACCCAACG AATGGACTCT ACAGCGACAG GAACCATG..
    SPI  GAGCGAGGGA TCACCCAACG AATGGACTCT ACAGCGACAG GAACCATGTG
    TOR  GAGCGAGGGA TCACCCAACG AATGGACTCT ACAGCGACAG GAACCATG..
    WIL  GAGCGAGGGA TCACCCAACG AATGGACTCT ACAGCGACAG GAACCATG..

         [251                      +1                        300]
 ABInol  .......... .......... ....0CTTTTG GATGGAGAAG TGGGGACCTT
    ALC  .......... .......... ....0CTTTTG GATGGAGAAG TGGGGACCTT
    ASP  .......... .......... ....0CTTTTG GATGGAGAAG TGGGGACCTT
    BRA  GACTCTACAG CGACAGGAAC CATG1CTTTTG GATGGAGAAG TGGGGACCTT
    BRE  .......... .......... ....0CTTTTG GATGGAGAAG TGGGGACCTT
    CHI  .......... .......... ....0CTTTTG GATGGAGAAG TGGGGACCTT
    CRA  .......... .......... ....0CTTTTG GATGGAGAAG TGGGGACCTT
 ENGnol  .......... .......... ....0CTTTTG GATGGAGAAG TGGGGACCTT
    FAR  .......... .......... ....0CTTTTG GATGGAGAAG TGGGGACCTT
 GLAnol  .......... .......... ....0CTTTTG GATGGAGAAG TGGGGACCTT
    GLE  .......... .......... ....0CTTTTG GATGGAGAAG TGGGGACCTT
 JEZno3  .......... .......... ....0CTTTTG GATGGAGAAG TGGGGACCTT
    KOR  .......... .......... ....0CTTTTG GATGGAGAAG TGGGGACCTT
    KOY  .......... .......... ....0CTTTTG GATGGAGAAG TGGGGACCTT
    LIK  .......... .......... ....0CTTTTG GATGGAGAAG TGGGGACCTT
 MARno4  .......... .......... ....0CTTTTG GATGGAGAAG TGGGGACCTT
    MAT  .......... .......... ....0CTTTTG GATGGAGAGG TGGGGACCTT
    MAX  .......... .......... ....0CTTTTG GATGGAGAAG TGGGGACCTT
    MEX  .......... .......... ....0CTTTTG GATGGAGAAG TGGGGACCTT
    MEY  .......... .......... ....0CTTTTG GATGGAGAAG TGGGGACCTT
    MON  .......... .......... ....0CTTTTG GATGGAGAAG TGGGGACCTT
    MOR  .......... .......... ....0CTTTTG GATGGAGAAG TGGGGACCTT
    OBO  .......... .......... ....0CTTTTG GATGGAGAAG TGGGGACCTT
    OMO  .......... .......... ....0CTTTTG GATGGAGAAG TGGGGACCTT
    ORI  .......... .......... ....0CTTTTG GATGGAGAAG TGGGGACCTT
PINBANK  .......... .......... ....0CTTTTG GATGGAGAAG TGGGAACCTT
 PINKOR  .......... .......... ....0CTTTTG GATGGAGAAG TGGGGACCTT
 PUNnol  .......... .......... ....0CTTTTG GATGGAGAAG TGGGGACCTT
    PUR  .......... .......... ....0CTTTTG GATGGAGAAG TGGGGACCTT
    RET  .......... .......... ....0CTTTTG GATGGAGAAG TGGGGACCTT
    RUB  .......... .......... ....0CTTTTG GATGGAGAAG TGGGGACCTT
    SCH  .......... .......... ....0CTTTTG GATGGAGAAG TGGGGACCTT
 SITnol  .......... .......... ....0CTTTTG GATGGAGAAG TGGGGACCTT
    SMI  .......... .......... ....0CTTTTG GATGGAGAAG TGGGGACCTT
    SPI  GACTCTACAG CGACAGGAAC CATG1CTTTTG GATGGAGAAG TGGGGACCTT
    TOR  .......... .......... ....0CTTTTG GATGGAGAAG TGGGGACCTT
    WIL  .......... .......... ....0CTTTTG GATGGAGAAG TGGGGACCTT

         [301                                                  350]
 ABInol  AATGATGAAT AGTAATAGGC CGAATCTGAA CTGCGGGAAT ATCTATTCGG
    ALC  AATGAT?AAT AGTAATAGGC CGAATCTGAA CTGCGGGAAT ATCTATTCGG
    ASP  AATGATGAAT AGTAATAGGC CGAATCTGAA CTGCGGGAAT ATCTATTCGG
    BRA  AATGATGAAT AGTAATAGGC CGAATCTGAA CTGCGGGAAT ATCTATTCGG
    BRE  AATGATGAAT AGTAATAGGC CGAATCTGAA CTGCGGGAAT ATCTATTCGG
    CHI  AATGATGAAT AGTAATAGGC CGAATCTGAA CTGCGGGAAT ATCTATTCGG
    CRA  AATGATGAAT AGTAATAGGC CGAATCTGAA CTGCGGGAAT ATCTATTCGG
 ENGnol  AATGATGAAT AGTAATAGGC CGAATCTGAA CTGCGGGAAT ATCTATTCGG
    FAR  AATGATGAAT AGTAATAGGC CGAATCTGAA CTGCGGGAAT ATCTATTCGG
 GLAnol  AATGATGAAT AGTAATAGGC CGAATCTGAA CTGCGGGAAT ATCTATTCGG
    GLE  AATGATGAAT AGTAATAGGC CGAATCTGAA CTGCGGGAAT ATCTATTCGG
 JEZno3  AATGATGAAT AGTAATAGGC CGAATCTGAA CTGCGGGAAT ATCTATTCGG
    KOR  AATGATGAAT AGTAATAGGC CGAATCTGAA CTGCGGGAAT ATCTATTCGG
    KOY  AATGATGAAT AGTAATAGGC CGAATCTGAA CTGCGGGAAT ATCTATTCGG
    LIK  AATGATGAAT AGTAATAGGC CGAATCTGAA CTGCGGGAAT ATCTATTCGG
 MARno4  AATGATGAAT AGTAATAGGC CGAATCTGAA CTGCGGGAAT ATCTATTCGG
    MAT  AATGATGAAT AGTAATAGGC CGAATCTGAA CTGCGGGAAT ATCTATTCGG
    MAX  AATGATGAAT AGTAATAGGC CGAATCTGAA CTGCGGGAAT ATCTATTCGG
    MEX  AATGATGAAT AGTAATAGGC CGAATCTGAA CTGCGGGAAT ATCTATTCGG
    MEY  AATGATGAAT AGTAATAGGC CGAATCTGAA CTGCGGGAAT ATCTATTCGG
    MON  AATGATGAAT AGTAATAGGC CGAATCTGAA CTGCGGGAAT ATCTATTCGG
    MOR  AATGATGAAT AGTAATAGGC CGAATCTGAA CTGCGGGAAT ATCTATTCGG
    OBO  AATGATGAAT AGTAATAGGC CGAATCTGAA CTGCGGGAAT ATCTATTCGG
    OMO  AATGATGAAT AGTAATAGGC CGAATCTGAA CTGCGGGAAT ATCTATTCGG
    ORI  AATGATGAAT AGTAATAGGC CGAATCTGAA CTGCGGGAAT ATCTATTCGG
PINBANK  AATGATGAAT AGTAATAGGC CGAATCTGAA CTGCGGGAAT ATATATTCGG
 PINKOR  AATGATGAAT AGTAATAGGC CGAATCTGAA CTGCGGGAAT ATCTATTCGG
 PUNnol  AATGATGAAT AGTAATAGGC CGAATCTGAA CTGCGGGAAT ATCTATTCGG
    PUR  AATGATGAAT AGTAATAGGC CGAATCTGAA CTGCGGGAAT ATCTATTCGG
    RET  AATGATGAAT AGTAATAGGC CGAATCTGAA CTGCGGGAAT ATCTATTCGG
    RUB  AATGATGAAT AGTAATAGGC CGAATCTGAA CTGCGGGAAT ATCTATTCGG
```

```
   SCH  AATGATGAAT AGTAATAGGC CGAATCTGAA CTGCGGGAAT ATCTATTCGG
SITno1  AATGATGAAT AGTAATAGGC CGAATCTGAA CTGCGGGAAT ATCTATTCGG
   SMI  AATGATGAAT AGTAATAGGC CGAATCTGAA CTGCGGGAAT ATCTATTCGG
   SPI  AATGATGAAT AGTAATAGGC CGAATCTGAA CTGCGGGAAT ATCTATTCGG
   TOR  AATGATGAAT AGTAATAGGC CGAATCTGAA CTGCGGGAAT ATCTATTCGG
   WIL  AATGATGAAT AGTAATAGGC CGAATCTGAA CTGCGGGAAT ATCTATTCGG

       [351            +1                                  400]
ABIno1 AACCTAGGGG ACTAATATT0A ATAGGGGCCA GGAATGTATG GGTGACAGAT
   ALC  AACCTAGGGG ACTAATATT0A ATAGGGGCCA GGAATGTATG GGTGACAGAT
   ASP  AACCTAGGGG ACTAATATT0A ATAGGGGCCA GGAATGTATG GGTGACAGAT
   BRA  AACCTAGGGG ACT......1A ATAGGGGCCA GGAATGTATG GGTGACAGAT
   BRE  AACCTAGGGG ACTAATATT0A ATAGGGGCCA GGAATGTATG GGTGACAGAT
   CHI  AACCTAGGGG ACTAATATT0A ATAGGGGCCA GGAATGTATG GGTGACAGAT
   CRA  AACCTAGGGG ACTAATATT0A ATAGGGGCCA GGAATGTATG GGTGACAGAT
ENGno1 AACCTAGGGG ACTAATATT0A ATAGGGGCCA GGAATGTATG GGTGACAGAT
   FAR  AACCTAGGGG ACT......1A ATAGGGGCCA GGAATGTATG GGTGACAGAT
GLAno1 AACCTAGGGG ACTAATATT0A ATAGGGGCCA GGAATGTATG GGTGACAGAT
   GLE  AACCTAGGGG ACTAATATT0A ATAGGGGCCA GGAATGTATG GGTGACAGAT
JEZno3 AACCTAGGGG ACTAATATT0A ATAGGGGCCA GGAATGTATG GGTGACAGAT
   KOR  AACCTAGGGG ACTAATATT0A ATAGGGGCCA GGAATGTATG GGTGACAGAT
   KOY  AACCTAGGGG ACTAATATT0A ATAGGGGCCA GGAATGTATG GGTGACAGAT
   LIK  AACCTAGGGG ACT......1A ATAGGGGCCA GGAATGTATG GGTGACAGAT
MARno4 AACCTAGGGG ACTAATATT0A ATAGGGGCCA GGAATGTATG GGTGACAGAT
   MAT  AACCTAGGGG ACT.....T1A ATAGGGGCCA GGAATGTATG GGTGACAGAT
   MAX  AACCTAGGGG ACTAATATT0A ATAGGGGCCA GGAATGTATG GGTGACAGAT
   MEX  AACCTAGGGG ACTAATATT0A ATAGGGGCCA GGAATGTATG GGTGACAGAT
   MEY  AACCTAGGGG ACTAATATT0A ATAGGGGCCA GGAATGTATG GGTGACAGAT
   MON  AACCTAGGGG ACTAATATT0A ATAGGGGCCA GGAATGTATG GGTGACAGAT
   MOR  AACCTAGGGG ACTAATATT0A ATAGGGGCCA GGAATGTATG GGTGACAGAT
   OBO  AACCTAGGGG ACTAATATT0A ATAGGGGCCA GGAATGTATG GGTGACAGAT
   OMO  AACCTAGGGG ACTAATATT0A ATAGGGGCCA GGAATGTATG GGTGACAGAT
   ORI  AACCTAGGGG ACTAATATT0A ATAGGGGCCA GGAATGTATG GGTGACAGAT
PINBANK AACCTAAGGG ACTAATATG0A ATAGGGGCCA GGAATGTATG GGTGACAGAT
PINKOR AACCTAGGGG ACTAATATT0A ATAGGGGCCA GGAATGTATG GGTGACAGAT
PUNno1 AACCTAGGGG ACTAATATT0A ATAGGGGCCA GGAATGTATG GGTGACAGAT
   PUR  AACCTAGGGG ACT......1A ATAGGGGCCA GGAATGTATG GGTGACAGAT
   RET  AACCTAGGGG ACTAATATT0A ATAGGGGCCA GGAATGTATG GGTGACAGAT
   RUB  AACCTAGGGG ACTAATATT0A ATAGGGGCCA GGAATGTATG GGTGACAGAT
   SCH  AACCTAGGGG ACT......1A ATAGGGGCCA GGAATGTATG GGTGACAGAT
SITno1 AACCTAGGGG ACTAATATT0A ATAGGGGCCA GGAATGTATG GGTGACAGAT
   SMI  AACCTAGGGG ACT......1A ATAGGGGCCA GGAATGTATG GGTGACAGAT
   SPI  AACCTAGGGG ACT......1A ATAGGGGCCA GGAATGTATG GGTGACAGAT
   TOR  AACCTAGGGG ACTAATATT0A ATAGGGGCCA GGAATGTATG GGTGACAGAT
   WIL  AACCTAGGGG ACTAATATT0A ATAGGGGCCA GGAATGTATG GGTGACAGAT

         [401                                              450]
ABIno1 CGGCCATAAA TGTACTCCGG AATATACACC AGGGCAACAA ATGTCGGTC.
   ALC  CGGCCATAAA TGTACTCCGG AATATACACC AGGGCAACAA ATGTCGGTC.
   ASP  CGGCCATAAA TGTACTCCGG AATATACACC AGGGCAACAA ATGTCGGTC.
   BRA  CGGCCATAAA TGTACTCCGG AATATACACC AGGGCAACAA ATGTCGGTC.
   BRE  CGGCCATAAA TGTACTCCGG AATATACACC AGGGCAACAA ATGTCGGTCC
   CHI  CGGCCATAAA TGTACTCCGG AATATACACC AGGGCAACAA ATGTCGGTC.
   CRA  CGGCCATAAA TGTACTCCGG AATATACACC AGGGCAACAA ATGTCGGTC.
ENGno1 CGGCCATAAA TGTACTCCGG AATATACACC AGGGCAACAA ATGTCGGTC.
   FAR  CGGCCATAAA TGTACTCCGG AATATACACC AGGGCAACAA ATGTCGGTC.
GLAno1 CGGCCATAAA TGTACTCCGG AATATACACC AGGGCAACAA ATGTCGGTC.
   GLE  CGGCCATAAA TGTACTCCGG AATATACACC AGGGCAACAA ATGTCGGTC.
JEZno3 CGGCCATAAA TGTACTCCGG AATATACACC AGGGCAACAA ATGTCGGTC.
   KOR  CGGCCATAAA TGTACTCCGG AATATACACC AGGGCAACAA ATGTCGGTC.
   KOY  CGGCCATAAA TGTACTCCGG AATATACACC AGGGCAACAA ATGTCGGTC.
   LIK  CGGCCATAAA TGTACTCCGG AATATACACC AGGGCAACAA ATGTCGGTC.
MARno4 CGGCCATAAA TGTACTCCGG AATATACACC AGGGCAACAA ATGTCGGTC.
   MAT  CGGCCATAAA TGTACTCCGG AATATACACC AGGGCAACAA ATGTCGGTC.
   MAX  CGGCCATAAA TGTACTCCGG AATATACACC AGGGCAACAA ATGTCGGTC.
   MEX  CGGCCATAAA TGTACTCCGG AATATACACC AGGGCAACAA ATGTCGGTC.
   MEY  CGGCCATAAA TGTACTCCGG AATATACACC AGGGCAACAA ATGTCGGTC.
   MON  CGGCCATAAA TGTACTCCGG AATATACACC AGGGCAACAA ATGTCGGTC.
   MOR  CGGCCATAAA TGTACTCCGG AATATACACC AGGGCAACAA ATGTCGGTC.
   OBO  CGGCCATAAA TGTACTCCGG AATATACACC AGGGCAACAA ATGTCGGTC.
   OMO  CGGCCATAAA TGTACTCCGG AATATACACC AGGGCAACAA ATGTCGGTC.
   ORI  CGGCCATAAA TGTACTCCGG AATATACACC AGGGCAACAA ATGTCGGTC.
PINBANK CGGCCATAAA TGTACTCCGG AATATACACC AGGGCAACAA ATGTCGGTC.
PINKOR CGGCCATAAA TGTACTCCGG AATATACACC AGGGCAACAA ATGTCGGTC.
PUNno1 CGGCCATAAA TGTACTCCGG AAGATACACC AGGGCAACAA ATGTCGGTC.
   PUR  CG?CCATAAA TGTACTCCGG AATATACACC AGGGCAACAA ATGTCGGTC.
   RET  CGGCCATAAA TGTACTCCGG AATATACACC AGGGCAACAA ATGTCGGTC.
   RUB  CGGCCATAAA TGTACTCCGG AATATACACC AGGGCAACAA ATGTCGGTC.
   SCH  CGGCCATAAA TGTACTCCGG AATATACACC AGGGCAACAA ATGTCGGTC.
SITno1 CGGCCATAAA TGTACTCCGG AATATACACC AGGGCAACAA ATGTCGGTC.
   SMI  CGGCCATAAA TGTACTCCGG AATATACACC AGGGCAACAA ATGTCGGTC.
   SPI  CGGCCATAAA TGTACTCCGG AATATACACC AGGGCAACAA ATGTCGGTC.
   TOR  CGGCCATAAA TGTACTCCGG AATATACACC AGGGCAACAA ATGTCGGTC.
   WIL  CGGCCATAAA TGTACTCCGG AATATACACC AGGGCAACAA ATGTCGGTC.
```

```
        [451  +1                                                    500]
 ABInol  ....0GACGAT GCCGCCCGTT TCCATTTCGT GGAAGTGCTC GGCAGAGGAA
    ALC  ....0GACGAT GCCGCCCGTT TCCATTTCGT GGAAGTGCTC GGCAGAGGAA
    ASP  ....0GACGAT GCCGCCCGTT TCCATTTCGT GGAAGTGCTC GGCAGAGGAA
    BRA  ....0GACGAT GCCGCCCGTT TCCATTTCGT GGAAGTGCTC GGCAGAGGAA
    BRE  GGTC1GACGAT GCCGCCCGTT TCCATTTCGT GGAAGTGCTC GGCAGAGGAA
    CHI  ....0GACGAT GCCGCCCGTT TCCATTTCGT GGAAGTGCTC GGCAGAGGAA
    CRA  ....0GACGAT GCCGCCCGTT TCCATTTCGT GGAAGTGCTC GGCAGAGGAA
 ENGnol  ....0GACGAT GCCGCCCGTT TCCATTTCGT GGAAGTGCTC GGCAGAGGAA
    FAR  ....0GACGAT GCCGCCCGTT TCCATTTCGT GGAAGTGCTC GGCAGAGGAA
 GLAnol  ....0GACGAT GCCGCCCGTT TCCATTTCGT GGAAGTGCTC GGCAGAGGAA
    GLE  ....0GACGAT GCCGCCCGTT TCCATTTCGT GGAAGTGCTC GGCAGAGGAA
 JEZno3  ....0GACGAT GCCGCCCGTT TCCATTTCGT GGAAGTGCTC GGCAGAGGAA
    KOR  ....0GACGAT GCCGCCCGTT TCCATTTCGT GGAAGTGCTC GGCAGAGGAA
    KOY  ....0GACGAT GCCGCCCGTT TCCATTTCGT GGAAGTGCTC GGCAGAGGAA
    LIK  ....0GACGAT GCCGCCCGTT TCCATTTCGT GGAAGTGCTC GGCAGAGGAA
 MARno4  ....0GACGAT GCCGCCCGTT TCCATTTCGT GGAAGTGCTC GGCAGAGGAA
    MAT  ....0GACGAT GCCGCCCGTT TCCATTTCGT GGAAGTGCTC GGCAGAGGAA
    MAX  ....0GACGAT GCCGCCCGTT TCCATTTCGT GGAAGTGCTC GGCAGAGGAA
    MEX  ....0GACGAT GCCGCCCGTT TCCATTTCGT GGAAGTGCTC GGCAGAGGAA
    MEY  ....0GACGAT GCCGCCCGTT TCCATTTCGT GGAAGTGCTC GGCAGAGGAA
    MON  ....0GACGAT GCCGCCCGTT TCCATTTCGT GGAAGTGCTC GGCAGAGGAA
    MOR  ....0GACGAT GCCGCCCGTT TCCATTTCGT GGAAGTGCTC GGCAGAGGAA
    OBO  ....0GACGAT GCCGCCCGTT TCCATTTCGT GGAAGTGCTC GGCAGAGGAA
    OMO  ....0GACGAT GCCGCCCGTT TCCATTTCGT GGAAGTGCTC GGCAGAGGAA
    ORI  ....0GACGAT GCCGCCCGTT TCCATTTCGT GGAAGTGCTC GGCAGAGGAA
PINBANK  ....0GACGAT GCCGCCCGTT TCCATTTCGT GGAAGTGCTC GGCAGAGGAA
 PINKOR  ....0GACGAT GCCGCCCGTT TCCATTTCGT GGAAGTGCTC GGCAGAGGAA
 PUNnol  ....0GACGAT GCCGCCCGTT TCCATTTCGT GGAAGTGCTC GGCAGAGGAA
    PUR  ....0GACGAT GCCGCCCGTT TCCATTTCGT GGAAGTGCTC GGCAGAGGAA
    RET  ....0GACGAT GCCGCCCGTT TCCATTTCGT GGAAGTGCTC GGCAGAGGAA
    RUB  ....0GACGAT GCCGCCCGTT TCCATTTCGT GGAAGTGCTC GGCAGAGGAA
    SCH  ....0GACGAT GCCGCCCGTT TCCATTTCGT GGAAGTGCTC GGCAGAGGAA
 SITnol  ....0GACGAT GCCGCCCGTT TCCATTTCGT GGAAGTGCTC GGCAGAGGAA
    SMI  ....0GACGAT GCCGCCCGTT TCCATTTCGT GGAAGTGCTC GGCAGAGGAA
    SPI  ....0GACGAT GCCGCCCGTT TCCATTTCGT GGAAGTGCTC GGCAGAGGAA
    TOR  ....0GACGAT GCCGCCCGTT TCCATTTCGT GGAAGTGCTC GGCAGAGGAA
    WIL  ....0GACGAT GCCGCCCGTT TCCATTTCGT GGAAGTGCTC GGCAGAGGAA

        [501                                                    550]
 ABInol  GGGGCTGTAG GTGATGATGA TGCCCCTATT GATGAAGGAA TCGCTGGGCA
    ALC  GGGGCTGTAG GTGATGATGA TGCCCCTATT GATGAAGGAA TCGCTGGGCA
    ASP  GGGGCTGTAG GTGATGATGA TGCCCCTATT GATGAAGGAA TCGCTGGGCA
    BRA  GGGGCTGTAG GTGATGATGA TGCCCCTATT GATGAAGGAA TCGCTGGGCA
    BRE  GGGGCTGTAG GTGATGATGA TGCCCCTATT GATGAAGGAA TCGCTGGGCA
    CHI  GGGGCTGTAG GTGATGATGA TGCCCCTATT GATGAAGGAA TCGCTGGGCA
    CRA  GGGGCTGTAG GTGATGATGA TGCCCCTATT GATGAAGGAA TCGCTGGGCA
 ENGnol  GGGGCTGTAG GTGATGATGA TGCCCCTATT GATGAAGGAA TCGCTGGGCA
    FAR  GGGGCTGTAG GTGATGATGA TGCCCCTATT GATGAAGGAA TCGCTGGGCA
 GLAnol  GGGGCTGTAG GTGATGATGA TGCCCCTATT GATGAAGGAA TCGCTGGGCA
    GLE  GGGGCTGTAG GTGATGATGA TGCCCCTATT GATGAAGGAA TCGCTGGGCA
 JEZno3  GGGGCTGTAG GTGATGATGA TGCCCCTATT GATGAAGGAA TCGCTGGGCA
    KOR  GGGGCTGTAG GTGATGATGA TGCCCCTATT GATGAAGGAA TCGCTGGGCA
    KOY  GGGGCTGTAG GTGATGATGA TGCCCCTATT GATGAAGGAA TCGCTGGGCA
    LIK  GGGGCTGTAG GTGATGATGA TGCCCCTATT GATGAAGGAA TCGCTGGGCA
 MARno4  GGGGCTGTAG GTGATGATGA TGCCCCTATT GATGAAGGAA TCGCTGGGCA
    MAT  GGGGCTGTAG GTGATGATGA TGCCCCTATT GATGAAGGAA TCGCTGGGCA
    MAX  GGGGCTGTAG GTGATGATGA TGCCCCTATT GATGAAGGAA TCGCTGGGCA
    MEX  GGGGCTGTAG GTGATGATGA TGCCCCTATT GATGAAGGAA TCGCTGGGCA
    MEY  GGGGCTGTAG GTGATGATGA TGCCCCTATT GATGAAGGAA TCGCTGGGCA
    MON  GGGGCTGTAG GTGATGATGA TGCCCCTATT GATGAAGGAA TCGCTGGGCA
    MOR  GGGGCTGTAG GTGATGATGA TGCCCCTATT GATGAAGGAA TCGCTGGGCA
    OBO  GGGGCTGTAG GTGATGATGA TGCCCCTATT GATGAAGGAA TCGCTGGGCA
    OMO  GGGGCTGTAG GTGATGATGA TGCCCCTATT GATGAAGGAA TCGCTGGGCA
    ORI  GGGGCTGTAG GTGATGATGA TGCCCCTATT GATGAAGGAA TCGCTGGGCA
PINBANK  GGGGCTGTAG GTGATGATGA TGCCCCTATT GATGAAGGAA TCGCTGGGCA
 PINKOR  GGGGCTGTAG GTGATGATGA TGCCCCTATT GATGAAGGAA TCGCTGGGCA
 PUNnol  GGGGCTGTAG GTGATGATGA TGCCCCTATT GATGAAGGAA TCGCTGGGCA
    PUR  GGGGCTGTAG GTGATGATGA TGCCCCTATT GATGAAGGAA TCGCTGGGCA
    RET  GGGGCTGTAG GTGATGATGA TGCCCCTATT GATGAAGGAA TCGCTGGGCA
    RUB  GGGGCTGTAG GTGATGATGA TGCCCCTATT GATGAAGGAA TCGCTGGGCA
    SCH  GGGGCTGTAG GTGATGATGA TGCCCCTATT GATGAAGGAA TCGCTGGGCA
 SITnol  GGGGCTGTAG GTGATGATGA TGCCCCTATT GATGAAGGAA TCGCTGGGCA
    SMI  GGGGCTGTAG GTGATGATGA TGCCCCTATT GATGAAGGAA TCGCTGGGCA
    SPI  GGGGCTGTAG GTGATGATGA TGCCCCTATT GATGAAGGAA TCGCTGGGCA
    TOR  GGGGCTGTAG GTGATGATGA TGCCCCTATT GATGAAGGAA TCGCTGGGCA
    WIL  GGGGCTGTAG GTGATGATGA TGCCCCTATT GATGAAGGAA TCGCTGGGCA

        [551                                              +1      600]
 ABInol  CCTATTTCTT TCTCTCTTAT TTCTTTCATA GGCAGCGGGT CC0TGGTCAAA
    ALC  CCTATTTCTT TCTCTCTTAT TTCTTTCATA GGCAGCGGGT CC0TGGTCAAA
    ASP  CCTATTTCTT TCTCTCTTAT TTCTTTCATA GGCAGCGGGT CC0TGGTCAAA
    BRA  CCTATTTCTT TATCTCTTAT TTCTTTCATA GGCAGCGGGT CC0TGGTAAAA
```

```
   BRE  CCTATTTCTT TCTCTCTTAT TTCTTTCATA GGCAGCGGGT CC0TGGTAAAA
   CHI  CCTATTTCTT TCTCTCTTAT TTCTTTCATA GGCAGCGGGT CC0TGGTCAAA
   CRA  CCTATTTCTT TCTCTCTTAT TTCTTTCATA GGCAGCGGGT CC0TGGTCAAA
ENGno1  CCTATTTCTT TCTCTCTTAT TTCTTTCATA GGCAGCGGGT CC0TGGTCAAA
   FAR  CCTATTTCTT TCTCTCTTAT TTCTTTCATA GGCAGCGGGT CC0TGGTCAAA
GLAno1  CCTATTTCTT TCTCTCTTAT TTCTTTCATA GGCAGCGGGT CC0TGGTCAAA
   GLE  CCTATTTCTT TCTCTCTTAT TTCTTTCATA GGCAGCGGGT CC0TGGTCAAA
JEZno3  CCTATTTCTT TCTCTCTTAT TTCTTTCATA GGCAGCGGGT CC0TGGTCAAA
   KOR  CCTATTTCTT TCTCTCTTAT TTCTTTCATA GGCAGCGGGT CC0TGGTCAAA
   KOY  CCTATTTCTT TCTCTCTTAT TTCTTTCATA GGCAGCGGGT CC0TGGTCAAA
   LIK  CCTATTTCTT TATCTCTTAT TTCTTTCATA GGCAGCGGGT CC0TGGTAAAA
MARno4  CCTATTTCTT TCTCTCTTAT TTCTTTCATA GGCAGCGGGT CC0TGGTCAAA
   MAT  CCTATTTCTT TCTCTCTTAT TTCTTTCATA GGCAGCGGGT CC0TGGTCAAA
   MAX  CCTATTTCTT TCTCTCTTAT TTCTTTCATA GGCAGCGGGT CC0TGGTCAAA
   MEX  CCTATTTCTT TCTCTCTTAT TTCTTTCATA GGCAGCGGGT CC0TGGTCAAA
   MEY  CCTATTTCTT TCTCTCTTAT TTCTTTCATA GGCAGCGGGT CC0TGGTCAAA
   MON  CCTATTTCTT TCTCTCTTAT TTCTTTCATA GGCAGCGGGT CC0TGGTCAAA
   MOR  CCTATTTCTT TCTCTCTTAT TTCTTTCATA GGCAGCG... ..1TGGTCAAA
   OBO  CCTATTTCTT TCTCTCTTAT TTCTTTCATA GGCAGCGGGT CC0TGGTCAAA
   OMO  CCTATTTCTT TCTCTCTTAT TTCTTTCATA GGCAGCGGGT CC0TGGTCAAA
   ORI  CCTATTTCTT TCTCTCTTAT TTCTTTCATA GGCAGCGGGT CC0TGGTCAAA
PINBANK  CCTATTTCTT TCTCTCTTAT TTATTTCATA GGCAGCAGGT CC0TGGTCAAA
 PINKOR  CCTATTTCTT TCTCTCTTAT TTCTTTCATA GGCAGCGGGT CC0TGGTCAAA
 PUNno1  CCTATTTCTT TCTCTCTTAT TTCTTTCATA GGCAGCGGGT CC0TGGTCAAA
   PUR  CCTATTTCTT TCTCTCTTAT TTCTTTCATA GGCAGC?GGT CC0TGGTCAAA
   RET  CCTATTTCTT TCTCTCTTAT TTCTTTCATA GGCAGCGGGT CC0TGGTCAAA
   RUB  CCTATTTCTT TCTCTCTTAT TTCTTTCATA GGCAGCGGGT CC0TGGTCAAA
   SCH  CCTATTGATT TCTCTCTTAT TTCTTTCATA GGCAGCGGGT CC0TGGTCAAA
SITno1  CCTATTTCTT TCTCTCTTAT TTCTTTCATA GGCAGCGGGT CC0TGGTCAAA
   SMI  CCTATTTATT TCTCTCTTAT TTCTTTCATA GGCAGCGGGT CC0TGGTCAAA
   SPI  CCTATTTCTT TATCTCTTAT TTCTTTCATA GGCAGCGGGT CC0TGGTAAAA
   TOR  CCTATTTCTT TCTCTCTTAT TTCTTTCATA GGCAGCGGGT CC0TGGTCAAA
   WIL  CCTATTTCTT TCTCTCTTAT TTCTTTCATA GGCAGCGGGT CC0TGGTCAAA

        [601                    +1       indels ignorés    650]
ABIno1  TCCGGTGGGT GCCGCAAGGT TATA0CCCCCC [CC....]TATC TATCTATAGA
   ALC  TCCGGTGGGT GCCGCAAGGT TATA0CCCCCC [CCCC..]TATC TATCTATAGA
   ASP  TCCGGTGGGT GCCGCAAGGT TATA0CCCCCC [CCC...]TATC TATCTATAGA
   BRA  TCCGGTGGGT GCCGCAAGGT TATA0CCCCCC [C.....]TATC TATCTATAGA
   BRE  TCCGGTGGGT GCCGCAAGGT TATA0CCCCCC [CCCC..]TATC TATCTATAGA
   CHI  TCCGGTGGGT GCCGCAAGGT TATA0CCCCCC [CCC...]TATC TATCTATAGA
   CRA  TCCGGTGGGT GCCGCAAGGT TATA0CCCCCC [CCC...]TATC TATCTATAGA
ENGno1  TCCGGTGGGT GCCGCAAGGT TATA0CCCCCC [CCCC..]TATC TATCTATAGA
   FAR  TCCGGTGGGT GCCGCAAGGT TATA0CCCCCC [CC....]TCTC TATCTATAGA
GLAno1  TCCGGTGGGT GCCGCAAGGT TATA0CCCCCC [CCCCC.]TATC TATCTATAGA
   GLE  TCCGGTGGGT GCCGCAAGGT TATA0CCCCCC [CCC...]TATC TATCTATAGA
JEZno3  TCCGGTGGGT GCCGCAAGGT TATA0CCCCCC [CCC...]TATC TATCTATAGA
   KOR  TCCGGTGGGT GCCGCAAGGT TATA0CCCCCC [CCC...]TATC TATCTATAGA
   KOY  TCCGGTGGGT GCCGCAAGGT TATA0CCCCCA [CCC...]TATC TATCTATAGA
   LIK  TCCGGTGGGT GCCGCAAGGT TATA0CCCCCC [C.....]TATC TATCTATAGA
MARno4  TCCGGTGGGT GCCGCAAGGT TATA0CCCCCC [CCC...]TATC TATCTATAGA
   MAT  TCCGGTGGGT GCCGCAAGGT TATA0CCCCCC [CCCC..]TATC TATCTATAGA
   MAX  TCCGGTGGGT GCCGCAAGGT TATA0CCCCCC [CCC...]TATC TATCTATAGA
   MEX  TCCGGTGGGT GCCGCAAGGT TATA0CCCCCC [CCC...]TATC TATCTATAGA
   MEY  TCCGGTGGGT GCCGCAAGGT TATA0CCCCCA [CCC...]TATC TATCTATAGA
   MON  TCCGGTGGGT GCCGCAAGGT TATA0CCCCCA [CCC...]TATC TATCTATAGA
   MOR  TCCGGTGGGT GCCGCAAGGT ....1CCCCCC [CC....]TATC TATCTATAGA
   OBO  TCCGGTGGGT GCCGCAAGGT TATA0CCCCCC [CCC...]TATC TATCTATAGA
   OMO  TCCGGTGGGT GCCGCAAGGT TATA0CCCCCC [CCC...]TATC TATCTATAGA
   ORI  TCCGGTGGGT GCCGCAAGGT TATA0CCCCCC [CCC...]TATC TATCTATAGA
PINBANK  TCCGGTGGGT GCCGCAAGGT TATA0CCCCCC [CC....]TATC TATCTATAGA
 PINKOR  TCCGGTGGGT GCCGCAAGGT TATA0CCCCCC [CC....]TATC TATCTATAGA
 PUNno1  TCCGGTGGGT GCCGCAAGGT TATA0CCCCCC [CCC...]TATC TATCTATAGA
   PUR  ?CCGGTGGGT GCCGCAAGGT TATA0CCCCCC [CCCC..]TATC TAT?TATAGA
   RET  TCCGGTGGGT GCCGCAAGGT TATA0CCCCCC [CCC...]TATC TATCTATAGA
   RUB  TCCGGTGGGT GCCGCAAGGT TATA0CCCCCC [CCC...]TATC TATCTATAGA
   SCH  TCCGGTGGGT GCCGCAAGGT TATA0CCCCCC [CCC...]TATC TATCTATAGA
SITno1  TCCGGTGGGT GCCGCAAGGT TATA0CCCCCC [CCCC..]TATC TATCTATAGA
   SMI  TCCGGTGGGT GCCGCAAGGT TATA0CCCCCC [CC....]TATC TATCTATAGA
   SPI  TCCGGTGGGT GCCGCAAGGT TATA0CCCCCC [C.....]TATC TATCTATAGA
   TOR  TCCGGTGGGT GCCGCAAGGT TATA0CCCCCC [CCC...]TATC TATCTATAGA
   WIL  TCCGGTGGGT GCCGCAAGGT TATA0CCCCCC [CCC...]TATC TATCTATAGA

        [651                                  +1          700]
ABIno1  TAGAGAGATA GTAGCGAGGG GTAGGTATCG ATCCGATA0GG GATAACCTTA
   ALC  TAGAGAGATA GTAGCGAGGG GTAGGTATCG GTCCGATA0GG GATAACCTTA
   ASP  TAGAGAGATA GTAGCGAGGG GTAGGTATC? GTCCGATA0?G GATAACCTTA
   BRA  TAGAGAGATA GTAGCGAGGG GTAGGTATCG GTCCGATA0GG GATAACCTTA
   BRE  TAGAGAGATA GTAGCGAGGG GTAGGTATCG GTCCGATA0GG GATAACCTTA
   CHI  TAGAGAGATA GTAGCGAGGG GTAGGTATCG GTCCGATA0GG GATAACCTTA
   CRA  TAGAGAGATA GTAGCGAGGG GTAGGTATCT ATCCGATA0AG GATAACCTTA
ENGno1  TAGAGAGATA GTAGCGAGGG GTAGGTATCG GTCCGATA0GG GATAACCTTA
   FAR  TAGAGAGATA GTAGCGAGGG GTAGGTATCG GTCCGATA0GG GATAACCTTA
GLAno1  TAGAGAGATA GTAGCGAGGG GTAGGTATCG GTCCGATA0GG GATAACCTTA
```

```
   GLE  TAGAGAGATA GTAGCGAGGG GTAGGTATCT ATCCGATA0AG GATAACCTTA
JEZno3  TAGAGAGATA GTAGCGAGGG GTAGGTATCG GTCCGATA0GG GATAACCTTA
   KOR  TAGAGAGATA GTAGCGAGGG GTA?GTATCG GTCCGATA0GG GATAACCTTA
   KOY  TAGAGAGATA GTAGCGAGGG GTAGGTATCG GTCCGATA0GG GATAACCTTA
   LIK  TAGAGAGATA GTAGCGAGGG GTAGGTATCG GTCCGATA0GG GATAACCTTA
MARno4  TAGAGAGATA GTAGCGAGGG GTAGGTATCG GTCCGATA0GG GATAACCTTA
   MAT  TAGAGAGATA GTAGCGAGGG GTAGGTATCG GTCCGATA0GG GATAACCTTA
   MAX  TAGAGAGATA GTAGCGAGGG GTAGGTATCG GTCCGATA0GG GATAACCTTA
   MEX  TAGAGAGATA GTAGCGAGGG GTAGGTATCG GTCCGATA0GG GATAACCTTA
   MEY  TAGAGAGATA GTAGCGAGGG GTAGGTATCG GTCTGATA0GG GATAACCTTG
   MON  TAGAGAGATA GTAGCGAGGG GTAGGTATC? GTC?GATA0?G GATAACCTT?
   MOR  TAGAGAGATA GTAGCGAGGG GTAGGTAT?? ?T????????G GATAAC?TT?
   OBO  TAGAGAGATA GTAGCGAGGG GTAGGTATCG GTCCGATA0GG GATAACCTTA
   OMO  TAGAGAGATA GTAGCGAGGG GTAGGTATCG GTCCGATA0GG GATAACCTTA
   ORI  TAGAGAGATA GTAGCGAGGG GTAGGTATcg gTccgata0gG GATAACgTTa
PINBANK  TAGAGAGATA GTAGCGAGGG GTAGGTAT?G GT?CGATA0GG GATAACCTTA
PINKOR  TAGAGAGATA GTAGCGAGGG GTAGGTATCG GTCCGATA0GG GATAACCTTA
PUNno1  TAGAGAGATA GTAGCGAGGG GTAGGTATCG GTCCGATA0GG GATAACCTTA
   PUR  TAGAGAGATA GTAGCGAGGG GTAGGTATCG GTCCGATA0GG GATAACCTTA
   RET  TAGAGAGATA GTAGCGAGGG GTAGGTATCT GTCCGATA0AG GATAACCTTA
   RUB  TAGAGAGATA GTAGCGAGGG GTAGGTATCG GTC.....1GG GATAACCTTA
   SCH  TAGAGAGATA GTAGCGAGGG GTAGGTATTG GTCCGATA0GG GATAACCTTA
SITno1  TAGAGAGATA GTAGCGAGGG GTAGGTATCG GTCCGATA0GG GATAACCTTA
   SMI  TAGAGAGATA GTAGCGAGGG GTAGGTATCG GTCCGATA0GG GATAACCTTA
   SPI  TAGAGAGATA GTAGCGAGGG GTAGGTATCG GTCCGATA0GG GATAACCTTA
   TOR  TAGAGAGATA GTAGCGAGGG GTAGGTATCG GTCCGATA0GG GATAACCTTA
   WIL  TAGAGAGATA GTAGCGAGGG GTAGGTAT?G GTCCGATA0GG GATAAC?TTA

        [701                                                750]
ABInol  ATTCTTAATA AAAAGGTCCT TCAAGGCCCT CACCCTCGAA GGGCCAGCGT
   ALC  ATTCTTAATA AAAAGGTCCT TCAAGGCCCT CGCCCTCGGA GGGCGAGCGT
   ASP  ATTCTTAATA AAAAGGTCCT TCAAGGCCCT C?CCCTCGGA GGG?GAGC?T
   BRA  ATTCTTAATA AAAAGGTCCT TCAAGGCCCT CGCCCTCGGA GGGCGAGCGT
   BRE  ATTCTTAATA AAAAGGTCCT TCAAGGCCCT CGCCCTCGGA GGGCGAGCGT
   CHI  ATTCTTAATA AAAAGGTCCT TCAAGGCCCT CGCCCTCGGA GGGCGAGCGT
   CRA  ATTCTTAATA AAAAGGTCCT TCAAGGCCCT CGCCCTCGGA GGGAGAGCGT
ENGno1  ATTCTTAATA AAAAGGTCCT TCAAGGCCCT CGCCCTCGGA GGGCGAGCGT
   FAR  ATTCTTAATA AAAAGGTCCT TCAAGGCCCT CGCCCTCGGA GGGCGAGCGT
GLAno1  ATTCTTAATA AAAAGGTCCT TCAAGGCCCT CGCCCTCGGA GGGCGAGCGT
   GLE  ATTCTTAATA AAAAGGTCCT TCAAGGCCCT CGCCCTCGGA GGGAGAGCGT
JEZno3  ATTCTTAATA AAAAGGTCCT TCAAGGCCCT CGCCCTCGGA GGGCGAGCGT
   KOR  ATTCTTAATA AAAAGGTCCT TCAAGGCCCT CGCCCTCgGA GGGCGAGCgT
   KOY  ATTCTTAATA AAAAGGTCCT TCAAGGCCCT CGCCCTCGGA GGGCGAGCGT
   LIK  ATTCTTAATA AAAAGGTCCT TCAAGGCCCT CGCCCTCGGA GGGCGAGCGT
MARno4  ATTCTTAATA AAAAGGTCCT TCAAGGCCCT CGCCCTCGGA GGGCGAGCGT
   MAT  ATTCTTAATA AAAAGGTCCT TCAAGGCCCT CGCCCTCGGA GGGCGAGCGT
   MAX  ATTCTTAATA AAAAGGTCCT TCAAGGCCCT CGCCCTCGGA GGGCGAGCGT
   MEX  ATTCTTAATA AAAAGGTCCT TCAAGGCCCT CGCCCTCGGA GGGCGAGCGT
   MEY  ATTCTTAATA AAAAGTTCCT TCAAGGCCCT CGCCCTCGGA GGGCGAGCGT
   MON  ATTCTTAATA AAAAG?TCCT TCAAGGCCCT CGCCCTCGGA GGG?GAGCGT
   MOR  ATTCTTAATA AAAAG?TCCT TCAAGG?CCT C?CCCTC??A GG???AGC?T
   OBO  ATTCTTAATA AAAAGGTCCT TCAAGGCCCT CGCCCTCGGA GGGCGAGCAT
   OMO  ATTCTTAATA AAAAGGTCCT TCAAGGCCCT CGCCCTCGGA GGGCGAGCGT
   ORI  ATTCTTAATA AAAAGgTCCT TCAAGGcCCT CgCCCTCggA GGgcgAGCgT
PINBANK  ATTCTTAATA AAAAGGTCCT TCAAGG?CCT CGCCCTCGGA GGGCGAGCGT
PINKOR  ATTCTTAATA AAAAGGTCCT TCAAGGCCCT CGCCCTCGGA GGGCGAGCGT
PUNno1  ATTCTTAATA AAAAGGTCCT TCAAGGCCCT CGCCCTCGGA GGGCGAGCGT
   PUR  ATTCTTAATA AAAAGGTCCT TCAAGGCCCT CGCCCTCGGA GGGCGAGC?T
   RET  ATTCTTAATA AAAAGGTCCT TCAAGGCCCT CGCCCTCGGA GGGAGAGCGT
   RUB  ATTCTTAATA AAAAGGTCCT TCAAGGCCCT CGCCCTCGGA GGGCGAGCGT
   SCH  ATTCTTAATA AAAAGGTCCT TCAAGGCCCT CGCCCTCGGA GGGCGAGCGT
SITno1  ATTCTTAATA AAAAGGTCCT TCAAGGCCCT CGCCCTCGGA GGGCGAGCGT
   SMI  ATTCTTAATA AAAAGGTCCT TCAAGGCCCT CGCCCTCGGA GGGCGAGCGT
   SPI  ATTCTTAATA AAAAGGTCCT TCAAGGCCCT CGCCCTCGGA GGGCGAGCGT
   TOR  ATTCTTAATA AAAAGGTCCT TCAAGGCCCT CACCCTCGGA GGGCGAGCAT
   WIL  ATTCTTAATA AAAAGGTCCT TCAAGGCCCT CGCCCTCGGA GGGCGAGCGT

        [751                                          +1    800]
ABInol  TCGTTTATTA CCCTCTCCCT TTTGGTC... .......... ....0GAGCAT
   ALC  TCGTTTATTA CCCTCTCCCT TTTGGTC... .......... ....0GAGCGT
   ASP  TCGTTTATTA CCCTCTCCCT TTTGGT?... .......... ....0GAGC?T
   BRA  TCGTTTATTA CCCTCTCCCT TTTGGTC... .......... ....0GAGCGT
   BRE  TCGTTTATTA CCCTCTCCCT TTTGGTC... .......... ....0GAGCGT
   CHI  TCGTTTATTA CCCTCTCCCT TTTGGTC... .......... ....0GAGCGT
   CRA  TCGTTTATTA CCCTCTCCCT TTTGGTT... .......... ....0GAGCAT
ENGno1  TCGTTTATTA CCCTCTCCCT TTTGGTC... .......... ....0GAGCGT
   FAR  TCGTTTATTA CCCTCTCCCT TTTGGTC... .......... ....0GAGCGT
GLAno1  TCGTTTATTA CCCTCTCCCT TTTGGTC... .......... ....0GAGCGT
   GLE  TCGTTTATTA CCCTCTCCCT TTTGGTT... .......... ....0GAGCAT
JEZno3  TC?TTTATTA CCCTCTCCCT TTTGGTC... .......... ....0GAGCGT
   KOR  TCGTTTATTm CCCTCTCCCT TTTGGTC... .......... ....0GAGCAT
   KOY  TCGTTTATTA CCCTCTCCCT TTTGGTC... .......... ....0GAGCGT
   LIK  TCGTTTATTA CCCTCTCCCT TTTGGTC... .......... ....0GAGCGT
MARno4  TCGTTTATTA CCCTCTCCCT TTTGGTC... .......... ....0GAGCGT
```

```
   MAT  TCGTTTATTA CCCTCTCCCT TTTGGTC... .......... ....0GAGCGT
   MAX  TCGTTTATTA CCCTCTCCCT TTTGGTc... .......... ....0gAGC?T
   MEX  TCGTTTATTA CCCTCTCCCT TTTGGTC... .......... ....0GAGCGT
   MEY  TTGTTTATTA CCCTCTCCCT TTTGGTT... .......... ....0GAGCGT
   MON  T?GTTTATTA CCCTCTCCCT TTTGGTT... .......... ....0GAGC?T
   MOR  T??TTTATTA CCCTCTCCCT TTTGGT???? ?????????? ??????AGC?T
   OBO  TCGTTTATTA CCCTCTCCCT TTTGGTC... .......... ....0AAGCGT
   OMO  TCGTTTATTA CCCTCTCCCT TTTGGTC... .......... ....0GAGCGT
   ORI  TcgTTTATTA CCCTCTCCCT TTTGGTc... .......... ....0gAGCgT
PINBANK TCGTTTATTA CCCTCTCCCT TTTGGTC... .......... ....0GAGCGT
 PINKOR TCGTTTATTA CCCTCTCCCT TTTGGTC... .......... ....0GAGCGT
 PUNnol TCGTTTATTA CCCTCTCCCT TTTGGTC... .......... ....0GAGCGT
   PUR  TCGTTTATTA CCCTCTCCCT TTTGGTC... .......... ....0GAGCGT
   RET  TCGTTTATTA CCCTCTCCCT TTTGGTT... .......... ....0GAGCAT
   RUB  TCGTTTATTA CCCTCTCCCT TTTGGTC... .......... ....0GAGCGT
   SCH  TCGTTTATTA CCCTCTCCCT TTTGGTCTCC CTTTTGGTTT GGTClGAGCGT
 SITnol TCGTTTATTA CCCTCTCCCT TTTGGTC... .......... ....0GAGCGT
   SMI  TCGTTTATTA CCCTCTCCCT TTTGGTC... .......... ....0GAGCGT
   SPI  TCGTTTATTA CCCTCTCCCT TTTGGTC... .......... ....0GAGCGT
   TOR  TCGTTTATTA CCCTCTCCCT TTTGGTC... .......... ....0GAGCGT
   WIL  TCGTTTATTA CCCTCTCCCT TTTGGTC... .......... ....0GAGCGT

        [801                      +1                  +1 +1    850]
 ABInol TCATTTATCA CCCTTTATTT .....0CTTCC CCTCAAAGGG ...00CTAGCTA
   ALC  TCATTTATCA CCCTTTATTT .....0?TTCC CCTCAAAGGG ...00CTAGCTA
   ASP  TCATTTATCA CCCTTTATTT .....0CTTCC CCTCAAAGGG ...00?TAGCTA
   BRA  TCATTTATCA CCCTTTATTT .....0CTTCC CCTCAAAGGG GGG11CTAGCTA
   BRE  TCATTTATCA CCCTTTATTT .....0CTTCC CCTCAAAGGG ...00CTAGCTA
   CHI  TCATTTATCA CCCTTTATTT .....0CTTCC CCTCAAAGGG ...00CTAGCTA
   CRA  TCATTTATCA CCCTTTATTT .....0CTTCC CCTCAAAGGG ...00ATAGCTA
 ENGnol TCATTTATCA CCCTTTATTT .....0CTTCC CCTCAAAGGG ...00CTAGCTA
   FAR  TCATTTATCA CCCTTTATTT CATTT1CTTCC CCTCAAAGGG ...00CTAGCTA
 GLAnol TCATTTATCA CCCTTTATTT .....0CTTCC CCTCAAAGGG ...00CTAGCTA
   GLE  TCATTTATCA CCCTTTATTT .....0CTTCC CCTCAAAGGG ...00ATAGCTA
 JEZno3 TCATTTATCA CCCTTTATTT .....0CTTCC CCTCAAAGGG ...00CTAGCTA
   KOR  TCATTTATCA CCCTTTATTT .....0CTTCC CCTCAAAGGG ...00CTAGCTA
   KOY  TCATTTATCA CCCTTTATTT .....0CTTCC CCTCAAAGGG ...00CTAGCTA
   LIK  TCATTTATCA CCCTTTATTT .....0CTTCC CCTCAAAGGG ...00CTAGCTA
 MARno4 TCATTTATCA CCCTTTATTT .....0CTTCC CCTCAAAGGG ...00CTAGCTA
   MAT  TCATTTATCA CCCTTTATTT .....0CTTCC CCTCAAAGGG ...00CTAGCTA
   MAX  TCATTTATCA CCCTTTATTT ???????TTCC CCTCAAA?GG ????0?TAGCTA
   MEX  TCATTTATCA CCCTTTATTT .....0CTTCC CCTCAAAGGG ...00CTAGCTA
   MEY  TCATTTATCA CCCTTTATTT .....0CTTCC CCTCAAAGGG ...00CTAGCTA
   MON  TCATTTATCA CCCTTTATTT .....0CTTCC CCTCAAAGGG ...00?TAGCTA
   MOR  TCATTTATCA CCCTTTATTT ???????TTCC CCTCAAA?GG ????0?TAGCTA
   OBO  TCATTTATCA CCCTTTATTT .....0ATTCC CCTCAAAGGG ...00CTAGCTA
   OMO  TCATTTATCA CCCTTTATTT .....0CTTCC CCTCAAAGGG ...00CTAGCTA
   ORI  TCATTTATCA CCCTTTATTT .attt1cTTCC CCTCAAAgGG ..g01?TAGCTA
PINBANK TCATTTATCA CCCTTTATTT .....0CTTCC CCTCAAAGGG ...00CTAGCTA
 PINKOR TCATTTATCA CCCTTTATTT .....0CTTCC CCTCAAAGGG ...00CTAGCTA
 PUNnol TCATTTATCA CCCTTTATTT .....0CTTCC CCTCAAAGGG ...00CTAGCTA
   PUR  TCATTTATCA CCCTTTATTT .....0CTTCC CCTCAAAGGG ...00CTAGCTA
   RET  TCATTTATCA CCCTTTATTT .....0CTTCC CCTCAAAGGG ...00ATAGCTA
   RUB  TCATTTATCA CCCTTTATTT .....0CTTCC CCTCAAAGGG ...00CTAGCTA
   SCH  TCATTTATCA CCCTTTATTT .....0ATTCC CCTCAAAGGG ...00CTAGCTA
 SITnol TCATTTATCA CCCTTTATTT .....0CTTCC CCTCAAAGGG ...00CTAGCTA
   SMI  TCATTTATCA CCCTTTATTT .....0CTTCC CCTCAAAGGG ...00CTAGCTA
   SPI  TCATTTATCA CCCTTTATTT .....0CTTCC CCTCAAAGGG GGG11CTAGCTA
   TOR  TCATTTATCA CCCTTTATTT .....0CTTCC CCTCAAAGGG ...00CTAGCTA
   WIL  TCATTTATCA CCCTTTATTT .....0CTTCC CCTCAAAGGG ...00CTAGCTA

        [851                                                 900]
 ABInol GAAGGTAAGG TTATCACTTA CCAGTAAAG? CCCGAKAAAC CCTTCTTCAC
   ALC  GAAGGTAAGG TTATCACTTA CCAGTAAAGG CCCGAGAAAC CCTTCTTCAC
   ASP  GAAGGTAAGG TTATCACTTA CCAGTAAAGG CCCGAGAAAC CCTTCTTCAC
   BRA  GAAGGTAAGG TTATCACTTA CCAGTAAAGG CCCGAGAAAC CCTTCTTCAC
   BRE  GAAGGTAAGG TTATCACTTA CCAGTAAAGG CCCGAGAAAC CCTTCTTCAC
   CHI  GAAGGTAAGG TTATCACTTA CCAGTAAAGG CCCGAGAAAC CCTTCTTCAC
   CRA  GAAGGTAAGG TTATCACTTA CCAGTAAAGG CCCGAGAAAC CCTTCTTCAC
 ENGnol GAAGGTAAGG TTATCACTTA CCAGTAAAGG CCCGAGAAAC CCTTCTTCAC
   FAR  GAAGGTAAGG TTATCACTTA CCAGTAAAGG CCCGAGAAAC CCTTCTTCAC
 GLAnol GAAGGTAAGG TTATCACTTA CCAGTAAAGG CCCGAGAAAC CCTTCTTCAC
   GLE  GAAGGTAAGG TTATCACTTA CCAGTAAAGG CCCGAGAAAC CCTTCTTCAC
 JEZno3 GAAGGTAAGG TTATCACTTA CCAGTAAAGG CCCGAGAAAC CCTTCTTCAC
   KOR  GAAGGTAAGG TTATCACTTA CCAGTAAAGG CC?GAGAAAC CCTTCTTCAC
   KOY  GAAGGTAAGG TTATCACTTA CCAGTAAAGG CCCGAGAAAC CCTTCTTCAC
   LIK  GAAGGTAAGG TTATCACTTA CCAGTAAAGG CCCGAGAAAC CCTTCTTCAC
 MARno4 GAAGGTAAGG TTATCACTTA CCAGTAAAGG CCCGAGAAAC CCTTCTTCAC
   MAT  GAAGGTAAGG TTATCACTTA CCAGTAAAGG CCCGAGAAAC CCTTCTTCAC
   MAX  GAAGGTAAGG TTATCACT?A ?CAGTAAAG? C??GA?AAAC CCTTCTTCAC
   MEX  GAAGGTAAGG TTATCACTTA CCAGTAAAGG CCCGAGAAAC CCTTCTTCAC
   MEY  GAAGGTAAGG TTATCACTTA CCAGTAAAGG CC?GAGAAAC CCTTCTTCAC
   MON  GAAGGTAAGG TTATCACTTA CCAGTAAAGG CC?GAGAAAC CCTTCTTCAC
   MOR  GAAGGTAAGG TTATCACT?A ?CAGTAAAGG CCCGA?AAAC CCTTCTTCAC
```

```
   OBO  GAAGGTAAGG TTATCACTTA ACAGTAAAGG CCCGAGAAAC CCTTCTTCAC
   OMO  GAAGGTAAGG TTATCACTTA CCAGTAAAGG CCCGAGAAAC CCTTCTTCAC
   ORI  GAAGGTAAGG TTATCACT?A ?CAGTAAAG? C??GA?AAAC CCTTCTTCAC
PINBANK GAAGGTAAGG TTATCACTGA CCAGTAAAGG CCCGAGAAAC CCTTCTTCAC
 PINKOR GAAGGTAAGG TTATCACTTA CCAGTAAAGG CCCGAGAAAC CCTTCTTCAC
 PUNno1 GAAGGTAAGG TTATCACTTA CCAGTAAAGG CCCGAGAAAC CCTTCTTCAC
   PUR  GAAGGTAAGG TTATCACTTA CCAGTAAAGG CCCGAGAAAC CCTTCTTCAC
   RET  GAAGGTAAGG TTATCACTTA CCAGTAAAGG CCCGAGAAAC CCTTCTTCAC
   RUB  GAAGGTAAGG TTATCACTTA CCAGTAAAGG CCCGAGAAAC CCTTCTTCAC
   SCH  GAAGGTAAGG TTATCACTTA CCAGTAAAGG CCCGAGAAAC CCTTCTTCAC
 SITno1 GAAGGTAAGG TTATCACTTA CCAGTAAAGG CCCGAGAAAC CCTTCTTCAC
   SMI  GAAGGTAAGG TTATCACTTA CCAGTAAAGG CCCGAGAAAC CCTTCTTCAC
   SPI  GAAGGTAAGG TTATCACTTA CCAGTAAAGG CCCGAGAAAC CCTTCTTCAC
   TOR  GAAGGTAAGG TTATCACTTA CCAGTAAAGG CCCGAGAAAC CCTTCTTCAC
   WIL  GAAGGTAAGG TTATCACTGA CCAGTAAAGG CCCGAGAAAC CCTTCTTCAC

        [901                                                950]
ABIno1 CTGGGTTGTT CCTTGTTCTA TGGTGTTGTG CTCTTATTAT CCTTGGCCTT
   ALC  CCGGGTTGTT CCTTGTTCTA TGGTGTTGTG CTCTTATTAT CCTTGGCCTT
   ASP  CCGGGTTGTT CCTTGTTCTA TGGTGTTGTG CTCTTATTAT CCTTGGCCTT
   BRA  CCGGGTTGTT CCTTGTTCTA TGGTGTTGTG CTCTTATTAT CCTTGGCCTT
   BRE  CCGGGTTGTT CCTTGTTCTA TGGTGTTGTG CTCTTATTAT CCTTGGCCTT
   CHI  CCGGGTTGTT CCTTGTTCTA TGGTGTTGTG CTCTTATTAT CCTTGGCCTT
   CRA  CCGGGTTGTT CCTTGTTCTA TGGTGTTGTG CTCTTATTAT CCTTGGCCTT
ENGno1 CCGGGTTGTT CCTTGTTCTA TGGTGTTGTG CTCTTATTAT CCTTGGCCTT
   FAR  CCGGGTTGTT CCTTGTTCTA TGGTGTTGTG CTCTTATTAT CCTTGGCCTT
GLAno1 CCGGGTTGTT CCTTGTTCTA TGGTGTTGTG CTCTTATTAT CCTTGGCCTT
   GLE  CCGGGTTGTT CCTTGTTCTA TGGTGTTGTG CTCTTATTAT CCTTGGCCTT
JEZno3 CCAGGTTGTT CCTTGTTCTA TGGTGTTGTG CTCTTATTAT CCTTGGCCTT
   KOR  CC?GGTTGTT CCTTGTTCTA TGGTGTTGTG CTCTTATTAT CCTTGGCCTT
   KOY  CCGGGTTTTT CCTTGTTCTA TGGTGTTGTG CTCTTATTAT CCTTGGCCTT
   LIK  CCGGGTTGTT CCTTGTTCTA TGGTGTTGTG CTCTTATTAT CCTTGGCCTT
MARno4 CCGGGTTGTT CCTTGTTCTA TGGTGTTGTG CTCTTATTAT CCTTGGCCTT
   MAT  CCGGGTTGTT CCTTATTCTA TGGTGTTGTG CTCTTATTAT CCTTGGCCTT
   MAX  C???GTT?TT ?CTT?TTCTA TGGTGTTGTG CTCTTATTAT CCTTGGCCTT
   MEX  CCGGGTTGTT CCTTGTTCTA TGGTGTTGTG CTCTTATTAT CCTTGGCCTT
   MEY  CCGGGTTGTT CCTTGTTCTA TGGTGTTGTG CTCTTATTAT CCTTGGCCTT
   MON  CCGGGTTGTT CCTTGTTCTA TGGTGTTGTG CTCTTATTAT CCTTGGCCTT
   MOR  C???GTT?TT CCTTGTTCTA TGGTGTTGTG CTCTTATTAT CCTTGGCCTT
   OBO  CCGGGTTGTT CCTTGTTCTA TGGTGTTGTG CTCTTATTAT CCTTGGCCTT
   OMO  CCGGGTTGTT CCTTGTTCTA TGGTGTTGTG CTCTTATTAT CCTTGGCCTT
   ORI  C???GTT?TT ?CTT?TTCTA TGGTGTTGTG CTCTTATTAT CCTTGGCCTT
PINBANK CTGGGTTGTT CCTTGTTCTA TGGTGTTGTG CTCTTATTAT CCTTGGCCTT
 PINKOR CCGGGTTGTT CCTTGTTCTA TGGTGTTGTG CTCTTATTAT CCTTGGCCTT
 PUNno1 CCGGGTTGTT CCTTGTTCTA TGGTGTTGTG CTCTTATTAT CCTTGGCCTT
   PUR  CCGGGTTGTT CCTTGTTCTA TGGTGTTGTG CTCTTATTAT CCTTGGCCTT
   RET  CCGGGTTGTT CCTTGTTCTA TGGTGTTGTG CTCTTATTAT CCTTGGCCTT
   RUB  CCGGGTTGTT CCTTGTTCTA TGGTGTTGTG CTCTTATTAT CCTTGGCCTT
   SCH  CCGGGTTGTT CCTTGTTCTA TGGTGTTGTG CTCTTATTAT CCTTGGCCTT
 SITno1 CCGGGTTGTT CCTTGTTCTA TGGTGTTGTG CTCTTATTAT CCTTGGCCTT
   SMI  C??GGTTGTT CCTTGTTCTA TGGTGTTGTG CTCTTATTAT CCTTGGCCTT
   SPI  CCGGGTTGTT CCTTGTTCTA TGGTGTTGTG CTCTTATTAT CCTTGGCCTT
   TOR  CCGGGTTGTT CCTTGTTCTA TGGTGTTGTG CTCTTATTAT CCTTGGCCTT
   WIL  CTGGGTTGTT CCTTGTTCTA TGGTGTTGTG CTCTTATTAT CCTTGGCCTT

        [951                                               1000]
ABIno1 ATCCTTATCC TTATTATCGG CCCTGGCTTG CTATATTTAG TCTCAAGGAA
   ALC  ATCCTTATCC TTATTATCGG CCCTGGCCTG CTATCTTTAG TCTCAAGGAA
   ASP  ATCCTTATCC TTATTATCGG CCC?GGCCTG CTATCTTTAG TCTCAAGGAA
   BRA  ATCCTTATCC TTATTATCGG CCCTGGCCTG CTATCTTTAG TCTCAAGGAA
   BRE  ATCCTTATCC TTATTATCGG CCCTGGCCTG CTATCTTTAG TCTCAAGGAA
   CHI  ATCCTTATCC TTATTATCGG CCCTGGCCTG CTATCTTTAG TCTCAAGGAA
   CRA  ATCCTTATCC TTATTATCGG CCCCGGCCT? CTATCTTTAG TCCCAAGGAA
ENGno1 ATCCTTATCC TTATTATCGG CCCTGGCCTG CTATCTTTAG TCTCAAGGAA
   FAR  ATCCTTATCC TTATTATCGG CCCTGGCCTG CTATCTTTAG TCTCAAGGAA
GLAno1 ATCCTTATCC TTATTATCGG CCCTGGCCTG CTATCTTTAG TCTCAAGGAA
   GLE  ATCCTTATCC TTATTATCGG CCCCGGCCTG CTATCTTTAG TCCCAAGGAA
JEZno3 ATCCTTATCC TTATTATCGG CCCTGGCCTG CTATCTTTAG TCTCAAGGAA
   KOR  ATCCTTATCC TTATTAT?GG CCC?GGCCTG CTATCTTTAG TCTCAAGGAA
   KOY  ATCCTTATAC TTATTATAGG CCCTGGCCTG CTATCTTTAG TCTCAAGGAA
   LIK  ATCCTTATCC TTATTATCGG CCCTGGCCTG CTATCTTTAG TCTCAAGGAA
MARno4 ATCCTTATCC TTATTATCGG CCCTGGCCTG CTATCTTTAG TCTCAAGGAA
   MAT  ATCCTTATCC TTATTATCGG CCCTGGCCTG CTATCTTTAG TCTCAAGGAA
   MAX  ATCCTTAT?C TTATTAT?GG CC??GGC?T? CTAT?TTTAG T??CAAGG?A
   MEX  ATCCTTATCC TTATTATCGG CCCTGGCCTG CTATCTTTAG TCTCAAGGAA
   MEY  ATCCTTAT?C TTATTATAGG CCCTGGCCTG CTATCTTTAG TCTCAAGGAA
   MON  ATCCTTATCC TTATTAT?GG CCC?GGC?TG CTATCTTTAG TC?CAAGGAA
   MOR  ATCCTTATCC TTATTATCGG CCCTGGCCTG CTATCTTTAG TATCAAGGAA
   OBO  ATCCTTATCC TTATTATCGG CCCTGGCCTG CTATCTTTAG TCTCAAGGAA
   OMO  ATCCTTATCC TTATTATCGG CCCTGGCCTG CTATCTTTAG TcTCAAGGAA
   ORI  ATCCTTAT?C TTATTAT?GG CC??GGC?T? CTAT?TTTAG T??CAAGG?A
PINBANK ATCCTTATCC TTATTATCGG CCCTGGCCTG CTATCTTTAG TCTCAAGGAA
 PINKOR ATCCTTAT?C TTATTAT?GG CC??GGC?T? CTAT?TTTAG T??CAAGG?A
 PUNno1 ATCCTTATCC TTATTATCGG CCCTGGCCTG CTATCTTTAG TCTCAAGGAA
```

```
   PUR  ATCCTTATCC TTATTATCGG CCCTGGCCTG CTATCTTTAG TCTCAAGGAA
   RET  ATCCTTATCC TTATTATCGG CCC?GGCCTG CTATCTTTAG TCTCAAGGAA
   RUB  ATCCTTATCC TTATTATCGG CCCTGGCCTG CTATCTTTAG TCTCAAGGAA
   SCH  ATCCTTATCC TTATTATCGG CCCTGGCCTG CTATCTTTAG TCTCAAGGAA
SITno1  ATCCTTATCC TTATTATCGG CCCTGGCCTG CTATCTTTAG TCTCAAGGAA
   SMI  ATCCTTATCC TTATTAT?GG CCCTGGCCTG CTATCTTTAG TCTCAAGGAA
   SPI  ATCCTTATCC TTATTATCGG CCCTGGCCTG CTATCTTTAG TCTCAAGGAA
   TOR  ATCCTTATCC TTATTATCGG CCC?GGCCTG CTATCTTTAG TCTCAAGGAA
   WIL  ATCCTTATCC TTATTATCGG CCCTGGCCTG CTATCTTTAG TCTCAAGGAA

        [1001                                  +2          1050]
ABIno1  TTAAG..... .......... .......... ......00TCGT TAATAAGAGA
   ALC  TTAGG..... .......... .......... ......00TCGT TAATAAGAGA
   ASP  TTAAG..... .......... .......... ......00TCGT TAATAAGAGA
   BRA  TTAAG..... .......... .......... ......00TCGT TAATAAGAGA
   BRE  TTAAG..... .......... .......... ......00TCGT TAATAAGAGA
   CHI  TTAAG..... .......... .......... ......00TCGT TAATAAGAGA
   CRA  TTAAG..... .......... .......... ......00TCGT TAATAAGAGA
ENGno1  TTAAGGGCCT GCTATCTTTA GTCTCAAGGA ATTAAG11TCGT TAATAAGAGA
   FAR  TTAAG..... .......... .......... ......00TAGT TAATAAGAGA
GLAno1  TTAAG..... .......... .......... ......00TCGT TAATAAGAGA
   GLE  TTAAG..... .......... .......... ......00TCGT TAATAAGAGA
JEZno3  TTAAG..... .......... .......... ......00TCGT TAATAAGAGA
   KOR  TTAAG..... .......... .......... ......00TCGT TAATAAGAGA
   KOY  TTAAG..... .......... .......... ......00TCGT TAATAAGAGA
   LIK  TTAAG..... .......... .......... ......00TCGT TAATAAGAGA
MARno4  TTAAG..... .......... .......... ......00TCGT TAATAAGAGA
   MAT  TTAAG..... .......... .......... ......00TCGT TAATAAGAGA
   MAX  TTA?G????? ?????????? ?????????? ????????T??T TAATAAGA?A
   MEX  TTAAG..... .......... .......... ......00TCGT TAATAAGATA
   MEY  TTAAG..... .......... .......... ......00TCGT TAATAAGAGA
   MON  TTAAG..... .......... .......... ......00TC?T TAATAAGAGA
   MOR  TTAAG..... .......... .......... ......00TCGT TAATAAGAGA
   OBO  TTAAG..... .......... .......... ......00TCGT TAATAAGAGA
   OMO  TTAAG..... .......... .......... ......00TCGT TAATAAGAGA
   ORI  TTA?G????? ?????????? ?????????? ????????T??T TAATAAGA?A
PINBANK TTAAG..... .......... .......... ......00TCGT TAATAAGAGA
PINKOR  TTA?G????? ?????????? ?????????? ????????T??T TAATAAGA?A
PUNno1  TTAAG..... .......... .......... ......00TCGT TAATAAGAGA
   PUR  TTAAG..... .......... .......... ......00TCGT TAATAAGAGA
   RET  TTAAG..... .......... .......... ......00TC?T TAATAAGAGA
   RUB  TTAAG..... .......... .......... ......00TCGT TAATAAGAGA
   SCH  TTAAG..... .......... .......... ......00TCGT TAATAAGAGA
SITno1  TTAAG..... .......... .......... ......10TCGT TAATAAGAGA
   SMI  TTAAG..... .......... .......... ......00TCGT TAATAAGAGA
   SPI  TTAAG..... .......... .......... ......00TCGT TAATAAGAGA
   TOR  TTAAG..... .......... .......... ......00TC?T TAATAAGAGA
   WIL  TTAAG..... .......... .......... ......00TC?T TAATAAGAGA

        [1051                                              1100]
ABIno1  CTTYRTTACA GTAGGGAAAC AACCATAAGT CAAATCTGGT CAATACTTCG
   ALC  CTTCGTTACA GTAGGGAAAC AACCATAAGT CAAATCTGGT CAATACTTCG
   ASP  CTTCATTACA GTAGGGAAAC AACCATAAGT CAAATCTGGT CAATACTTC?
   BRA  CTTCGTTACA GTAGGGCAAC AACCATAAGT CAAATCTGGT CAATACTTCG
   BRE  CTTCGTTACA GTAGGGCAAC AACCATAAGT CAAATCTGGT CAATACTTCG
   CHI  CTTCGTTACA GTAGGGCAAC AACCATAAGT AAAATCTGGT CAATACTTCG
   CRA  CTTCATTACA GTAGGGAAAC AACCATAAGT CAAATCTGGT CAATACTTCG
ENGno1  CTTCGTTACA GTAGGGCAAC AACCATAAGT AAAATCTGGT CAATACTTCG
   FAR  CTTCGTTACA GTAGGGCAAC AACCATAAGT CAAATCTGGT CAATACTTCG
GLAno1  CTTCGTTACA GTAGGGCAAC AACCATAAGT AAAATCTGGT CAATACTTCG
   GLE  CTTCATTACA GTAGGGAAAC AACCATAAGT CAAATCTGGT CAATACTTCG
JEZno3  CTTTGTTACA GTAGGGCAAC AACCATAAGT CAAATCTAGT CAATACTTCG
   KOR  CTTCGTTACA GTAGGG?AAC AACCATAAGT CAAAT?TGGT CAATACTTCG
   KOY  CTTCGTTACA GTAGGGCAAC AACCATAAGT CAAATTTGGT CAATACTTCG
   LIK  CTTCGTTACA GTAGGG?AAC AACCATAAGT CAAAT?TGGT CAATACTTCG
MARno4  CTTCGTTACA GTAGGGCAAC AACCATAAGT AAAATCTGGT CAATACTTCG
   MAT  CTTCGTTACA GTAGGGCAAC AACCATAAGT AAAATCTGGT CAATACTTCG
   MAX  CTT??TTA?A GTAGGG?AAC AACCATAAGT ?AAAT?T?GT CAATACTTC?
   MEX  CTTCGTTACA GTAGGGCAAC AACCATAAGT AAAATCTGGT CAATACTTCG
   MEY  CTTCGTTACA GTAGGGCAAC AACCATAAGT CAAATTTGGT CAATACTTCG
   MON  CTTC?TTACA GTAGGG?AAC AACCATAAGT CAAAT?TGGT CAATACTTCG
   MOR  CTTCGTTACA GTAGGGCAAC AACCATAAGT CAAATCTGGT CAATACTTCG
   OBO  CTTCGTTACA GTAGGGCAAC AACCATAAGT CAAATCTGGT CAATACTTCG
   OMO  CTTCGTTACA GTAGGGCAAC AACCATAAGT aAAATCTGGT CAATACTTCG
   ORI  CTT??TTA?A GTAGGG?AAC AACCATAAGT ?AAAT?T?GT CAATACTTC?
PINBANK CTTCGTTACA GTAGGGCAAC AACCATAAGT CAAATCTGGT CAATACTTCG
PINKOR  CTT??TTA?A GTAGGG?AAC AACCATAAGT ?AAAT?T?GT CAATACTTC?
PUNno1  CTTCGTTACA GTAGGGCAAC AACCATAAGT AAAATCTGGT CAATACTTCG
   PUR  CTTCGTTACA GTAGGGCAAC AACCATAAGT CAAATCTGGT CAATACTTCG
   RET  CTTC?TTACA GTAGGG?AAC AACCATAAGT CAAATCTGGT CAATACTTCG
   RUB  CTTCGTTACA GTAGGGCAAC AACCATAAGT AAAATCTGGT CAATACTTCG
   SCH  CTTCGTTACA GTAGGGCAAC AACCATAAGT AAAATCTGGT CAATACTTCG
SITno1  CTTCGTTACA GTAGGGCAAC AACCATAAGT AAAATCTGGT CAATACTTCG
   SMI  CTTCGTTACA GTAGGGCAAC AACCATAAGT CAAATCTGGT CAATACTTCG
```

```
   SPI  CTTCGTTACA GTAGGGCAAC AACCATAAGT CAAATCTGGT CAATACTTCG
   TOR  CTTCATTACA GTAGGGAAAC AACCATAAGT CAAATCTGGT CAATACTTC?
   WIL  CTT?GTTACA GTAGGGCAAC AACCATAAGT CAAAT?TGGT CAATACTTCG

        [1101  +1                                          +1  1150]
ABInol  AAGAGCC0CTT CCTGATTAAC TCTAGCGAAG AAGGTTTCTT CTAAA0GATCA
   ALC  AAGAGCC0CTT CCTGATTGAC TCTAGCGAAG AAGGTTTCTT CTAAA0GATCA
   ASP  AAGAGCC0CTT CCTGATTGAC TCTAGCGAAG AAGGTTTCTT ?TAAA0GATCA
   BRA  .......lCTT CCTGATTGAC TCTTTAGAAG AAGGTTTCTT CTAAA0GATCA
   BRE  AAGAGCC0CTT CCTGATTGAC TCTAGCGAAG AAGGTTTCTT CTAAA0GAGAA
   CHI  AAGAGCC0CTT CCTGATTGAC TCTAGCGAAG AAGGTTTCTT CTAA.1GATCA
   CRA  AAGAGCC0CTT CCTGATTGAC TCTAGCAAAG AAGGTTTCTT TTAAA0GATCA
ENGnol  AAGAGCC0CTT CCTGATTGAC TCTAGCGAAG AAGGTTTCTT CTAA.1GATCA
   FAR  AAGAGCC0CTT CCTGATTGAC TCTTTAGAAG AAGGTTTCTT CTAAA0GATCA
GLAnol  AAGAGCC0CTT CCTGATTGAC TCTAGCGAAG AAGGTTTCTT CTAA.1GATCA
   GLE  AAGAGCC0CTT CCTGATTGAC TCTAGCAAAG AAGGTTTCTT TTAAA0GATCA
JEZno3  AAGAGCC0CTT CCTGATTGAC TCTAGCGAAG AAGGTTTCTT CTAAA0GATCA
   KOR  AAGAGCC0CTT CCT?ATTGAC TCTAGC?AAG AAGGTTT?TT CTAAA0GATCA
   KOY  AAGAGCC0CTT CCTAATTGAC TCTAGCGAAG AAGGTTTTTT CTAAA0GATCA
   LIK  AAGAGCC0CTT CCTGATTGAC TCTAGCGAAG AAGGTTTCTT CTAAA0GATCA
MARno4  AAGAGCC0CTT CCTGATTGAC TCTAGCGAAG AAGGTTTCTT CTAAA0GATCA
   MAT  AAGAGCC0CTT CCTGATTGAC TCTAGCGAAG AAGGTTTCTT CTAAA0GATCA
   MAX  ????????CTT CC??ATT?AC TCT????AAG AAGGTTT?TT ?T?A??GA??A
   MEX  AAGAGCC0CTT CCTGATTGAC TCTAGCGAAG AAGGTTTCTT CTAA.1GATCA
   MEY  AAGAGCC0CTT CCTAATTGAC TCTAGCGAAG AAGGTTTTTT CTAAA0GATCA
   MON  AAGA?CC0CTT CCTGATTGAC TCTAGC?AAG AAGGTTTCTT CTAAA0GATCA
   MOR  AAGAGCC0CTT CCTGATTGAC TCTAGCGAAG AAGGTTTCTT CTAAA0GATCA
   OBO  AAGAGCC0CTT CCTGATTGAC TCTAGCGAAG AAGGTTTCTT CTAAA0GAGAA
   OMO  AAGAGCC0CTT CCTGATTGAC TCTAGCGAAG AAGGTTTCTT CTAAA0GATCA
   ORI  ????????CTT CC??ATT?AC TCT????AAG AAGGTTT?TT ?T?A??GA??A
PINBANK  AAGAGCC0CTT CCTGATTGAC TCTAGTGAAG AAGGTTTCTT CTAAA0GATCA
PINKOR  AAGA?CC?CTT CC??ATT?AC TCT????AAG AAGGTTT?TT ?T?AA?GA??A
PUNnol  AAGAGCC0CTT CCTGATTGAC TCTAGCGAAG AAGGTTTCTT CTAA.1GATCA
   PUR  AAGAGCC0CTT CCTGATTGAC TCTAGCGAAG AAGGTTTCTT CTAAA0GATCA
   RET  AAGAGCC0CTT CCTGATTGAC TCTAGC?AAG AAGGTTTCTT ?TAAA0GATCA
   RUB  AAGAGCC0CTT CCTGATTGAC TCTAGCGAAG AAGGTTTCTT CTAAA0GATCA
   SCH  AAGAGCC0CTT CCTGATTGAC TCTAGCGAAG AAGGTTTCTT CTAAA0GATCA
SITnol  AAGAGCC0CTT CCTGATTGAC TCTAGCGAAG AAGGTTTCTT CTAA.1GATCA
   SMI  AAGAGCC0CTT CCTGATTGAC TCTTTAGAAG AAGGTTTCTT CTAAA0GATCA
   SPI  .......lCTT CCTGATTGAC TCTTTAGAAG AAGGTTTCTT CTAAA0GATCA
   TOR  AAGAGCC0CTT CC?GATTGAC TCTAGC?AAG AAGGTTTCTT ?TAAA0GATCA
   WIL  AAGAGCC0CTT CCTGATTGAC TCTAGCGAAG AAGGTTTCTT CTAAA0GATCA

        [1151                                                1200]
ABInol  ATACTTCTAG AAGGTTGGCC TTGAGAGGAG ATCTTATGGA GGGTTTACTC
   ALC  ATACTTCTAG AAGGTTGGCC TTGAGAGGAG ATCTTATGGA GGGTTGACTC
   ASP  ATACTTCTAG AAGGTTGGCC TTGAGAGGAG ATCTTATGGA GGGTTGACTC
   BRA  ATACTTCTAG AAGGTTGGCC TTGAGAGGAG ATCTTATGGA GGGTTGACTC
   BRE  ATACTTCTAG AAGGTTGGCC TTGAGAGGAG ATCTTATGGA GGGTTGACTC
   CHI  ATACTTCTAG AAGGTTGGCC TTGAGAGGAG ATCTTATGGA GGGTTGACTC
   CRA  ATACTTCTAG AAGGTTGGCC TTGAGAGGAG ATCTTATGGA GGGTTGACTC
ENGnol  ATACTTCTAG AAGGTTGGCC TTGAGAGGAG ATCTTATGGA GGGTTGACTC
   FAR  ATACTTCTAG AAGGTTGGCC TTGAGAGGAG ATCTTATGGA GGGTTGACTC
GLAnol  ATACTTCTAG AAGGTTGGCC TTGAGAGGAG ATCTTATGGA GGGTTGACTC
   GLE  ATACTTCTAG AAGGTTGGCC TTGAGAGGAG ATCTTATGGA GGGTTGACTC
JEZno3  ATACTTCTAG AAGGTTGGCC TTGAGAGGAG ATCTTATGGA GGGTTGACTC
   KOR  AT?CTTCTAG AAGGTTGGCC TTGAGAGGAG ATCTTATGGA GGGTTGACTC
   KOY  ATGCTTCTAG AAGGTTGGCC TTGAGAGGAG ATCTTATGGA GGGTTGACTC
   LIK  ATACTTCTAG AAGGTTGGCC TTGAGAGGAG ATCTTATGGA GGGTTGACTC
MARno4  ATACTTCTAG AAGGTTGGCC TTGAGAGGAG ATCTTATGGA GGGTTGACTC
   MAT  ATACTTCTAG AAGGTTGGCC TTGAGAGGAG ATCTTATGGA GGGTTGACTC
   MAX  AT?CTTCTA? AAGGTTGGCC TTGAGAGGAG AT?TT?TGGA GGGTT?ACTC
   MEX  ATACTTCTAG AAGGTTGGCC TTGAGAGGAG ATCTTATGGA GGGTTGACTC
   MEY  ATGCTTCTAG AAGGTTGGCC TTGAGAGGAG ATCTTATGGA GGGTTGACTC
   MON  AT?CTTCTAG AAGGTTGGCC TTGAGAGGAG ATCTTATGGA GGGTTGACTC
   MOR  ATACTTCTAG AAGGTTGGCC TTGAGAGGAG ATCTTATGGA GGGTTGACTC
   OBO  ATACTTCTAG AAGGTTGGCC TTGAGAGGAG ATCTTATGGA GGGTTGACTC
   OMO  ATACTTCTAG AAGGTTGGCC TTGAGAGGAG ATCTTATGGA GGGTTGACTC
   ORI  AT?CTTCTA? AAGGTTGGCC TTGAGAGGAG AT?TT?TGGA GGGTT?ACTC
PINBANK  ATACTTCTAT AAGGTTGGCC TTGAGAGGAG ATCTTATGGA GGGTTGACTC
PINKOR  AT?CTTCTA? AAGGTTGGCC TTGAGAGGAG AT?TT?TGGA GGGTT?ACTC
PUNnol  ATACTTCTAG AAGGTTGGCC TTGAGAGGAG ATCTTATGGA GGGTTGACTC
   PUR  ATACTTCTAG AAGGTTGGCC TTGAGAGGAG ATCTTATGGA GGGTTGACTC
   RET  ATACTTCTAG AAGGTTGGCC TTGAGAGGAG ATCTTATGGA GGGTTGACTC
   RUB  ATACTTCTAG AAGGTTGGCC TTGAGAGGAG ATCTTATGGA GGGTTGACTC
   SCH  ATACTTCTAG AAGGTTGGCC TTGAGAGGAG ATCTTATGGA GGGTTGACTC
SITnol  ATACTTCTAG AAGGTTGGCC TTGAGAGGAG ATCTTATGGA GGGTTGACTC
   SMI  ATACTTCTAG AAGGTTGGCC TTGAGAGGAG ATCTTATGGA GGGTTGACTC
   SPI  ATACTTCTAG AAGGTTGGCC TTGAGAGGAG ATCTTATGGA GGGTTGACTC
   TOR  ATACTTCTAG AAGGTTGGCC TTGAGAGGAG ATCTTATGGA GGGTTGACTC
   WIL  ATACTTCTAG AAGGTTGGCC TTGAGAGGAG AT?TT?TGGA GGGTTGACTC

        [1201           +1                                   1250]
ABInol  TGACCTGATT AA.....0CCC TTCTCATTTT CACCCGAAGT AATATTTTAC
```

```
   ALC  TGACCTGATT AA.....0CCC TTCTCATTTT CACCCGAAGT AATATTTTAC
   ASP  TGACCTGATT AA.....0CCC TTCTCATTTT CACCCG?AG? AATATTTTAC
   BRA  TGACCTGATT AAATTAA1CCC TTCTCATTTT CACCCGAAGT AATATTTTAC
   BRE  TGACCTGATT AA.....0CCC TTCTCATTTT CACCCGAAGT AATATTTTAC
   CHI  TGACCTGATT AA.....0CCC TTCTCATTTT CACCCGAAGT AATATTTTAC
   CRA  TGACCTGATT AA.....0CCC TTCTCATTTT CACCCACA?C AATATTTTAC
ENGno1  TGACCTGATT AA.....0CCC TTCTCATTTT CACCCGAAGT AATATTTTAC
   FAR  TGACCTGATT AA.....0CCC TTCTCATTTT CACCCGAAGT AATATTTTAC
GLAno1  TGACCTGATT AA.....0CCC TTCTCATTTT CACCCGAAGT AATATTTTAC
   GLE  TGACCTGATT AA.....0CCC TTCTCATTTT CACCCACAGC AATATTTTAC
JEZno3  TGACCTGATT AA.....0CCC TTCTCATTTT CACCCGAAGT AATATTTTAC
   KOR  TGAC?TGATT AA.....0CCC TTCTCATTTT CACCC??AG? AATATTTTAC
   KOY  TGACATGATT AA.....0CCC TTCTCATTTT CACCCGAAGT AATATTTTAC
   LIK  TGACCTGATT AA.....0CCC TTCTCATTTT CACCCGAAGT AATATTTTAC
MARno4  TGACCTGATT AA.....0CCC TTCTCATTTT CACCCGAAGT AATATTTTAC
   MAT  TGACCTGATT AA.....0CCC TTCTCATTTT CACCCGAAGT AATATTTTAC
   MAX  TGAC?TGATT AA??????CCC TTCTCATTTT ?ACCC??A?? AATATTTTAC
   MEX  TGACCTGATT AA.....0CCC TTCTCATTTT CACCCGAAGT AATATTTTAC
   MEY  TGACATGATT AA.....0CCC TTCTCATTTT CACCCGAAGT AATATTTTAC
   MON  TGAC?TGATT AA.....0CCC TTCTCATTTT CACCC??AG? AATATTTTAC
   MOR  TGACCTGATT AA.....0CCC TTCTCATTTT AACCCGAAGT AATATTTTAC
   OBO  TGACCTGATT AA.....0CCC TTCTCATTTT CACCCGAAGT AATATTTTAC
   OMO  TGACCTGATT AA.....0CCC TTCTCATTTT cACCCGAAGT AATATTTTAC
   ORI  TGAC?TGATT AA??????CCC TTCTCATTTT ?ACCC??A?? AATATTTTAC
PINBANK  TGACCTGATT AA.....0CCC TTCTCATTTT CACCCGAAGT AATATTTTAC
 PINKOR  TGAC?TGATT AA??????CCC TTCTCATTTT ?ACCC??A?? AATATTTTAC
 PUNno1  TGACCTGATT AA.....0CCC TTCTCATTTT CACCCGAAGT AATATTTTAC
   PUR  TGACCTGATT AA.....0CCC TTCTCATTTT CACCCGAAGT AATATTTTAC
   RET  TGACCTGATT AA.....0CCC TTCTCATTTT CACCCA?AG? AATATTTTAC
   RUB  TGACCTGATT AA.....0CCC TTCTCATTTT CACCCGAAGT AATATTTTAC
   SCH  TGACCTGATT AA.....0CCC TTCTCATTTT CACCCGAAGT AATATTTTAC
SITno1  TGACCTGATT AA.....0CCC TTCTCATTTT CACCCGAAGT AATATTTTAC
   SMI  TGACCTGATT AA.....0CCC TTCTCATTTT CACCCAAAGT AATATTTTAC
   SPI  TGACCTGATT AAATTAA1CCC TTCTCATTTT CACCCGAAGT AATATTTTAC
   TOR  TGACCTGATT AA.....0CCC TTCTCATTTT CACCC??AG? AATATTTTAC
   WIL  TGACCTGATT AA.....0CCC TTCTCATTTT CACCC?AAGT AATATTTTAC

        [1251                                    +3 +1       1300]
ABIno1  GGTCAATTCT TCGAAAGACC CCATCCTTAT TAA.....000C0C CTTATCTGAC
   ALC  GGTCAATTCT TCGAAAGAGC CCATCCTTAT TAA.....000C0C CTTATCTGAC
   ASP  GGTCAATTCT TCGAAAGAG? CCATCCTTAT TAA.....000C0C CTTATCTGAC
   BRA  GGTAAATTCT TCGAAAGAGC CCATCCTTAT TAA.....000C0C CTTATCTGAC
   BRE  GGTCAATTCT TCGAAAGAGC CCATCCTTAT TAA.....000C0C CTTATCTGAC
   CHI  GGTAAATTCT TCGAAAGAGC CCATCCTTAT TAA.....000C0C CTTATCTGAC
   CRA  GGTCAATTCT TCGAAAGAGC CCATCCTTAT TAA.....000C0C CTTATCTGAC
ENGno1  GGTAAATTCT TCGAAAGAGC CCATCCTTAT TAA.....000C0C CTTATCTGAC
   FAR  GGTAAATTCT TCGAAAGAGC CCATCCTTAT TAA.....000C0C CTTATCTGAC
GLAno1  GGTAAATTCT TCGAAAGAGC CCATCCTTAT TAA.....000C0C CTTATCTGAC
   GLE  GGTCAATTCT TCGAAAGAGC CCATCCTTAT TAA.....000C0C CTTATCTGAC
JEZno3  GGTCAATTCT TCGAAAGAGC CCATCCTTAT TAA.....000C0C CTTATCTGAC
   KOR  GGTCAATTCT TCGAAAGAG? CCATCCTTAT TAA.....000C0C CTTATCTGAC
   KOY  GGTCAATTCT TCAAAAGAGA CCATCCTTAT TAA.....000C0C CTTATCTGAC
   LIK  GGTCAATTCT TCGAAAGAGC CCATCCTTAT TAA.....000C0C CTT?TCTGAC
MARno4  GGTAAATTCT TCGAAAGAGC CCATCCTTAT TAA.....000C0C CTTATCTGAC
   MAT  GGTCAATTCT TCGAAAGAGC CCATCCTTAT TAA.....000C0C CTTATCTGAC
   MAX  ??T?AATTCT TC?AAAGA?? C?ATCCTTAT TA???????????C CTT?TCT?AC
   MEX  GGTAAATTCT TCGAAAGAGC CCATCCTTAT TAA.....000C0C CTTATCTGAC
   MEY  GGTCAATTCT TCGAAAGAGA CCATCCTTAT TAA.....000C0C CTTATCTGAC
   MON  G?T?AATTCT TC?AAAGAG? CCATCCTTAT TAA.....000C0C CTTATCT?AC
   MOR  GGTAAATTCT TCGAAAGAGC CCATCCTTAT TAAATTAA110C0C CTTATCTGAC
   OBO  GGTCAATTCT TCGAAAGAGC CCATCCTTAT TAA.....000C0C CTTATCTGAC
   OMO  GGTAAATTCT TCGAAAGAGC CCATCCTTAT TAAattaa100C0C CTTATCTGAC
   ORI  ??T?AATTCT TC?AAAGA?? C?ATCCTTAT TA???????????C CTT?TCT?AC
PINBANK  GGTCAATTCT TCGAAAGAGC CTATCCTTAT TAAA....000.1C CTTATCTGAC
 PINKOR  ??T?AATTCT TC?AAAGA?? C?ATCCTTAT TAA??????????C CTT?TCT?AC
 PUNno1  GGTAAATTCT TCGAAAGAGC CCATCCTTAT TAA.....000C0C CTTATCTGAC
   PUR  GGTAAATTCT TCGAAAGAGC CCATCCTTAT TAA.....000C0C CTTATCTGAC
   RET  G?T?AATTCT TC?AAAGAGC CCATCCTTAT TAA.....000C0C CTTATCT?AC
   RUB  GGTAAATTCT TCGAAAGAGC CCATCCTTAT TAA.....000C0C CTTATCTGAC
   SCH  GGTAAATTCT TCGAAAGAGC CCATCCTTAT TA......001C0C CTTATCTGAC
SITno1  GGTAAATTCT TCGAAAGAGC CCATCCTTAT TAA.....000C0C CTTATCTGAC
   SMI  GGTAAATTCT TCGAAAGAGC CCATCCTTAT TAA.....000C0C CTTATCTGAC
   SPI  GGTAAATTCT TCGAAAGAGC CCATCCTTAT TAA.....000C0C CTTATCTGAC
   TOR  GGTCAATTCT TCGAAAGAG? CCATCCTTAT TAA.....000C0C CTTATCT?AC
   WIL  AGTCAATTCT TCGAAAGAGC CCATCCTTAT TAA.....000C0C CTTATCTGAC

        [1301                                                1350]
ABIno1  ARATTTTATC TTGATCAATC TTTCCTTATC TAATGAAGAG TTGTCCTTTT
   ALC  AGATTTTATC TTGATCAATC CTTCCTTATC TAATGAAGAG TTGGCCTTTT
   ASP  AGATTTTATC TTGATCAATC CTTCCTTATC TAATGAAGAG TTGGCCTTTT
   BRA  AGATTTGATC TTGATCAATC CTTCCTTATC TAATGAAGAG TTGGCCTTTT
   BRE  AGATTTGATC TTGATCAATC CTTCCTTATC TAATGAAGAG TTGGCCTTTT
   CHI  AGATTTTATC TTGATCAATC CTTCCTTATC TAATGAAGAG TTGGCCTTTT
   CRA  AGATTTTATC TTGATCAATC CTTCCTTATC TAATGAAGAG TTGGCCTTTT
```

```
ENGno1  AGATTTGATC TTGATCAATC CTTCCTTATC TAATGAAGAG TTGGCCTTTT
   FAR  AGATTTGATC TTGATCAATC CTTCCTTATC TAATGAAGAG TTGGCCTTTT
GLAno1  AGATTTGATC TTGATCAATC CTTCCTTATC TAATGAAGAG TTGGCCTTTT
   GLE  AGATTTTATC TTGATCAATC CTTCCTTATC TAATGAAGAG TTGGCCTTTT
JEZno3  AGATTTTATC TTGATCAATC CTTCCTTATC TAATGAAGAG TTGGCCTTTT
   KOR  A?ATTTTATC TTGATCAATC CTTCCTTATC TAAT?AAGAG TTGGCCTTTT
   KOY  AAATTTTATC TTGATCAATC CTTCCTTATC TAATTAAGAG TTGGCCTTTT
   LIK  AGATTTTATC TTGATCAATC CTTCCTTATC TAATGAAGAG TTGGCCTTTT
MARno4  AGATTTTATC TTGATCAATC CTTCCTTATC TAATGAAGAG TTGGCCTTTT
   MAT  AGATTTTATC TTGATCAATC CTTCCTTATC TAATGAAGAG TTGGCCTTTT
   MAX  A?ATTT?ATC TTGA??AAT? ?TTC?TTATC TAAT?AAGAG TTG?CCTTTT
   MEX  AGATTTTATC TTGATCAATC CTTCCTTATC TAATGAAGAG TTGGCCTTTT
   MEY  AAATTTTATC TTGATCAATC CTTCCTTATC TAATTAAGAG TTGTCCTTTT
   MON  AGATTTTATC TTGATCAATC CTTCCTTATC TAATGAAGAG TTGGCCTTTT
   MOR  AGATTTTATC TTGATCAATC CTTCCTTATC TAATGAAGAG TTGGCCTTTT
   OBO  AGATTTGATC TTGATCAATC CTTCCTTATC TAATGAAGAG TTGGCCTTTT
   OMO  AGATTTTATC TTGATCAATC CTTCCTTATC TAATGAAGAG TTGGCCTTTT
   ORI  A?ATTT?ATC TTGA??AAT? ?TTC?TTATC TAAT?AAGAG TTG?CCTTTT
PINBANK  AGATTTTATC TTGATCAATC CTTCCTTATC TAATGAAGAG TTGGCCTTTT
 PINKOR  A?ATTT?ATC TTGA??AAT? ?TTC?TTATC TAAT?AAGAG TTG?CCTTTT
 PUNno1  AGATTTTATC TTGATCAATC CTTCCTTATC TAATGAAGAG TTGGCCTTTT
   PUR  AGATTTTATC TTGA??AATC CTTCCTTATC TAATGAAGAG TTGGCCTTTT
   RET  AGATTTTATC TTGATCAATC CTTCCTTATC TAATGAAGAG TTGGCCTTTT
   RUB  AGATTTTATC TTGATCAATC CTTCCTTATC TAATGAAGAG TTGGCCTTTT
   SCH  AGATTTTATC TTGATCAATC CTTCCTTATC TAATGAAGAG TTGGCCTTTT
 SITno1  AGATTTGATC TTGATCAATC CTTCCTTATC TAATGAAGAG TTGGCCTTTT
   SMI  AGATTTGATC TTGATCAATC CTTC?TTATC TAATGAAGAG TTGGCCTTTT
   SPI  AGATTTGATC TTGATCAATC CTTCCTTATC TAATGAAGAG TTGGCCTTTT
   TOR  AGATTTTATC TTGATCAATC CTTCCTTATC TAATGAAGAG TTGGCCTTTT
   WIL  AGATTTTATC TTGATCAAT? CTTCCTTATC TAATGAAGAG TTGGCCTTTT

         [1351                                          +2      1400]
ABIno1  CAGGTTAATT CTTAAAAAAA AYTGTACTAA TTAAT..... .000AACGAACGA
   ALC  CAGGTGAATT CTTAAAAAAA ACTGTACTAA TTAAT..... .000ATCGAACGA
   ASP  CAGGTGAATT CTTAAAAAAA A?TGTACTAA TTAAT..... .000ATCGAACGA
   BRA  CAGGTGAATT CTTAAAAAAA ACTGTACTAA TTAATATTAA T111ATCGAACGA
   BRE  CAGGTGAATT CTTAAAAAAA ACTGTACTAA TTAAT..... .000ATCGAACGA
   CHI  CAGGTGAATT CTTAAAAAAA ACTGTACTAA TTAAT..... .000ATCGAACGA
   CRA  CAGGTGAATT CTTAAAAAAA ATTGTACTAA TTAAT..... .000ATCGAACGA
ENGno1  CAGGTGAATT CTTAAAAAAA ACTGTACTAA TTAAT..... .000ATCGAACGA
   FAR  CAGGTGAATT CTTAAAAAAA ACTGTACTAA TTAAT..... .000ATCGAACGA
GLAno1  CAGGTGAATT CTTAAAAAAA ACTGTACTAA TTAAT..... .000ATCGAACGA
   GLE  CAGGTGAATT CTTAAAAAAA ATTGTACTAA TTAAT..... .000ATCGAACGA
JEZno3  CAGGTGAATT CTTAAAAAAA AATGTACTAA TTAAT..... .000ATCGAACGA
   KOR  CAGGTGAATT CTTAAAAAAA A?TGTACTAA TTAAT..... .000ATCGAACGA
   KOY  CAGGTGAATT CTTAAAAAAA ACTGTACTAA TTAAT..... .000ATCGAACGA
   LIK  CAGGTGAATT CTTAAAAAAA A?TGTACTAA TTAAT..... .000ATC?AACGA
MARno4  CAGGTGAATT CTTAAAAAAA ACTGTACTAA TTAATATTAA T111ATCGAACGA
   MAT  CAGGTGAATT CTTAAAAAAA ACTGTACTAA TTAAT..... .000ATCGAACGA
   MAX  C?G?T?AATT CTTAAAAAAA A?TGTACTAA TT?AT????? ??????C?AAC?A
   MEX  CAGGTGAATT CTTAAAAAAA ACTGTACTAA TTAAT..... .000ATCGAACGA
   MEY  CAGGTGAATT CTTAAAAAAA ACTGTACTAA TTAAT..... .000ATCGAACGA
   MON  CAGGTGAATT CTTAAAAAAA A?TGTACTAA TTAAT..... .000ATCGAACGA
   MOR  CAGGTGAATT CTTAAAAAAA ACTGTACTAA TTAAT..... .000ATCGAACGA
   OBO  CAGGTGAATT CTTAAAAAAA ACTGTACTAA TTAAT..... .000ATCGAACGA
   OMO  CAGGTGAATT CTTAAAAAAA ACTGTACTAA TTAAT..... .001ATCGAACGA
   ORI  C?G?T?AATT CTTAAAAAAA A?TGTACTAA TT?AT????? ??????CGAAC?A
PINBANK  CAGGTGAATT CTTAAAAAAA ACTGTACTAA TTAAT..... .000ATCGAACGA
 PINKOR  C?G?T?AATT CTTAAAAAAA A?TGTACTAA TT?AT????? ??????C?AAC?A
 PUNno1  CAGGTGAATT CTTAAAAAAA ACTGTACTAA TTAAT..... .000ATCGAACGA
   PUR  CAGGTGAATT CTTAAAAAAA ACTGTACTAA TTAAT..... .000ATCGAACGA
   RET  CAGGTGAATT CTTAAAAAAA A?TGTACTAA TTAAT..... .000ATCGAACGA
   RUB  CAGGTGAATT CTTAAAAAAA ACTGTACTAA TTAAT..... .000ATCGAACGA
   SCH  CAGGTGAATT CTTAAAAAAA ACTGTACTAA TTAAT.TTAA T100ATCGAACGA
 SITno1  CAGGTGAATT CTTAAAAAAA ACTGTACTAA TTAAT..... .000ATCGAACGA
   SMI  C?GGTGAATT CTTAAAAAAA ACTGTACTAA TTCATATTAA T111ATCGAACGA
   SPI  CAGGTGAATT CTTAAAAAAA ACTGTACTAA TTAATATTAA T111ATCGAACGA
   TOR  CAGGTGAATT CTTAAAAAAA A?TGTACTAA TTAAT..... .000ATCGAACGA
   WIL  CAGGTGAATT CTTAAAAAAA A?TGTACTAA TTAAT..... .000ATCGAACGA

         [1401                      +1                      1450]
ABIno1  AGAAAATCGA CTCAAATTGT ......0CAGA ATGAAACCTA AAAATCCTTA
   ALC  AGAAAATCGA CTCAAATTGT ......0CAGA ATGAAACCTA AAAATCCTTA
   ASP  AGAAAATCGA CTCAAATTGT ......0CAGA ATGAAACCTA AAAATCCTTA
   BRA  AGAAAATCGA CTCAAATTGT ......0CAGA ATGAAACCTA AAAATCCTTA
   BRE  AGAAAATCGA CTCAAATTGT ......0CAGA ATGAAACCTA AAAATCCTTA
   CHI  AGAAAATCGA CTCAAATTGT ......0CAGA ATGAAACCTA AAAATCCTTA
   CRA  AGAAAATCGA CTCAAATTGT ......0CAGA ATGAAACCTA AAAATCCTTA
ENGno1  AGAAAATCGA CTCAAATTGT ......0CAGA ATGAAACCTA AAAATCCTTA
   FAR  AGAAAATCGA CTCAAATTGT ......0CAGA ATGAAACCTA AAAATCCTTA
GLAno1  AGAAAATCGA CTCAAATTGT ......0CAGA ATGAAACCTA AAAATCCTTA
   GLE  AGAAAATCGA CTCAAATTGT ......0CAGA ATGAAACCTA AAAATCCTTA
JEZno3  AGAAAATCGA CTCAAATTGT ......0CAGA ATGAAACCTA AAAATCCTTA
   KOR  AGAAAAT?GA CTCAAATTGT ......0CAGA ATGAAACCTA AAAATCCTTA
```

```
   KOY  AGAAAATTGA CTCAAATTGT ......0CAGA ATGAAACCTA AAAATCCTTA
   LIK  AGAAAATCGA CTCAAATTGT ......0CAGA ATGAAACCTA ?AAATCCTTA
MARno4  AGAAAATCGA CTCAAATTGT ......0CAGA ATGAAACCTA AAAATCCTTA
   MAT  AGAAAATCGA CTCAAATTGT ......0CAGA ATGAAACCTA AAAATCCTTA
   MAX  AGAAAAT??A CTCAAATTGT ???????CA?A ATGAAACCTA ?AAATCCTTA
   MEX  AGAAAATCGA CTCAAATTGT ......0CAGA ATGAAACCTA AAAATCCTTA
   MEY  AGAAAATTGA CTCAAATTGT ......0CAGA ATGAAACCTA AAAATCCTTA
   MON  AGAAAATC?A CTCAAATTGT ......0CAGA ATGAAACCTA AAAATCCTTA
   MOR  AGAAAATCGA CTCAAATTGT ......0CAGA ATGAAACCTA AAAATCCTTA
   OBO  AGAAAATCGA CTCAAATTGT ......0CAGA ATGAAACCTA AAAATCCTTA
   OMO  AGAAAATCGA CTCAAATTGT ......0CAGA ATGAAACCTA AAAATCCTTA
   ORI  AGAAAAT??A CTCAAATTGT ???????CA?A ATGAAACCTA ?AAATCCTTA
PINBANK AGAAAATCGA CTCAAATTGT ......0CAAA ATGAAACCTA AAAATCCTTA
PINKOR  AGAAAAT??A CTCAAATTGT ???????CA?A ATGAAACCTA ?AAATCCTTA
PUNno1  AGAAAATCGA CTCAAATTGT ......0CAGA ATGAAACCTA AAAATCCTTA
   PUR  AGAAAATCGA CTCAAATTGT ......0CAGA ATGAAACCTA AAAATCCTTA
   RET  AGAAAATC?A CTCAAATTGT ......0CAGA ATGAAACCTA AAAATCCTTA
   RUB  AGAAAATCGA CTCAAATTGT AATTGT1CAGA ATGAAACCTA AAAATCCTTA
   SCH  AGAAAATCGA CTCAAATTGT ......0CAGA ATGAAACCTA AAAATCCTTA
SITno1  AGAAAATCGA CTCAAATTGT ......0CAGA ATGAAACCTA AAAATCCTTA
   SMI  AGAAAATCGA CTCAAATTGT ......0CAGA ATGAAACCTA AAAATCCTTA
   SPI  AGAAAATCGA CTCAAATTGT ......0CAGA ATGAAACCTA AAAATCCTTA
   TOR  AGAAAATCGA CTCAAATTGT ......0CAGA ATGAAACCTA AAAATCCTTA
   WIL  AGAAAATCGA CTCAAATTGT ......0CAGA ATGAAACCTA AAAATCCTTA

        [1451                                                1500]
ABInol  AATATCCCTT GYTGTAACGA ACTTGTGAAT CTCCCTACCC CTTCGGATAC
   ALC  AATATCGCTT GTTGTAACGA ACTTGTGAAT CTCCCTACCC CTTCGGATAC
   ASP  AATATCCCTT G?TGTAA?GA ACTTGTGAAT CTCCCTACCC CTTCGGATAC
   BRA  AATATCCCTT GCTGTAACGA ACTTGTGAAT CTCCCTACCC CTTCGGATAC
   BRE  AATATCCCTT GCTGTAACGA ACTTGTGAAT CTCCCTACCC CTTCGGATAC
   CHI  AATATCCCTT GCTGTAACGA ACTTGTGAAT CTCCCTACCC CTTCGGATAC
   CRA  AATATCCCTT GTTGTAATGA ACTTGTGAAT CTCCCTACCC CTTCGGATAC
ENGno1  AATATCCCTT GCTGTAACGA ACTTGTGAAT CTCCCTACCC CTTCGGATAC
   FAR  AATATCCCTT GCTGTAACGA ACTTGTGAAT CTCCCTACCC CTTCGGATAC
GLAno1  AATATCCCTT GCTGTAACGA ACTTGTGAAT CTCCCTACCC CTTCGGATAC
   GLE  AATATCCCTT GTTGTAATGA ACTTGTGAAT CTCCCTACCC CTTCGGATAC
JEZno3  AATATCCCTT GTTGTAACGA ACTTGTGAAT CTCCCTACCC CTTCGGATAC
   KOR  AATATCCCTT G?T?TAA?GA ACTTGTGAAT CTCCCTACCC CTTCGGATAC
   KOY  AATATCCCTT GCTATAACGA ACTTGTGAAT CTCCCTACCC CTTCGGATAC
   LIK  AATATCCCTT G?TGTAACGA ACTTGTGAAT CTCCCTACCC CTTCGGATAC
MARno4  AATATCCCTT GCTGTAACGA ACTTGTGAAT CTCCCTACCC CTTCGGATAC
   MAT  AATATCCCTT GCTGTAACGA ACTTGTGAAT CTCCCTACCC CTTCGGATAC
   MAX  AATA?C?CTT G?T?TAA?GA ACTT?TGAAT ?TCCCTACCC CTTCGGATAC
   MEX  AATATCCCTT GCTGTAACGA ACTTGTGAAT CTCCCTACCC CTTCGGATAC
   MEY  AATATCCCTT GCTATAACGA ACTTGTGAAT CTCCCTACCC CTTCGGATAC
   MON  AATATCCCTT G?TGTAA?GA ACTTGTGAAT CTCCCTACCC CTTCGGATAC
   MOR  AATATCCCTT GCTGTAACGA ACTTGTGAAT CTCCCTACCC CTTCGGATAC
   OBO  AATATCCCTT GCTGTAACGA ACTTGTGAAT CTCCCTACCC CTTCGGATAC
   OMO  AATATCCCTT GCTGTAACGA ACTTGTGAAT CTCCCTACCC CTTCGGATAC
   ORI  AATA?C?CTT G?T?TAA?GA ACTT?TGAAT ?TCCCTACCC CTTCGGATAC
PINBANK AATACCCCTT GCTGTAACGA ACTTGTGAAT CTCCCTACCC CTTCGGATAC
PINKOR  AATA?C?CTT G?T?TAA?GA ACTT?TGAAT ?TCCCTACCC CTTCGGATAC
PUNno1  AATATCCCTT GCTGTAACGA ACTTGTGAAT CTCCCTACCC CTTCGGATAC
   PUR  AATATCCCTT GCTGTAACGA ACTTGTGAAT CTCCCTACCC CTTCGGATAC
   RET  AATATCCCTT G?TGTAA?GA ACTTGTGAAT CTCCCTACCC CTTCGGATAC
   RUB  AATATCCCTT GCTGTAACGA ACTTGTGAAT CTCCCTACCC CTTCGGATAC
   SCH  AATATCCCTT GCTGTAACGA ACTTGTGAAT CTCCCTACCC CTTCGGATAC
SITno1  AATATCCCTT GCTGTAACGA ACTTGTGAAT CTCCCTACCC CTTCGGATAC
   SMI  AATATCCCTT GCTGTAACGA ACTTGTGAAT CTCCCTACCC CTTCGGATAC
   SPI  AATATCCCTT GCTGTAACGA ACTTGTGAAT CTCCCTACCC CTTCGGATAC
   TOR  AATATCCCTT GTTGTAA?GA ACTTGTGAAT CTCCCTACCC CTTCGGATAC
   WIL  AATATCCCTT G?TGTAACGA ACTTGTGAAT CTCCCTACCC CTTCGGATAC

        [1501                                                1550]
ABInol  TAAGGATATA ATAAGGATAG GATAATAAAA AGACCAAATA GACCAGGAAA
   ALC  TAAGGATATA ATAAGGATAG GATAATAAAA AGACCAAATA GACCAAGAAA
   ASP  TAAGGATATA ATAAGGATAG GATAATAAAA AGACCAAATA GACCAAGAAA
   BRA  TAAGGATATA ATAAGGATAG GATAATAAAA AGACCAAATA GACCAGGAAA
   BRE  TAAGGATATA ATAAGGATAG GATAATAAAA AGACCAAATA GACCAGGAAA
   CHI  TAAGGATATA ATAAGGATAG GATAATAAAA AGACCAAATA GACCAGGAAA
   CRA  TAAGGATATA ATAAGGATAG GATAATAAAA AGACCAAATA GACCAGGAAG
ENGno1  TAAGGATATA ATAAGGATAG GATAATAAAA AGACCAAATA GACCAGGAAA
   FAR  TAAGGATATA ATAAGGATAG GATAATAAAA AGACCAAATA GACCAGGAAA
GLAno1  TAAGGATATA ATAAGGATAG GATAATAAAA AGACCAAATA GACCAGGAAA
   GLE  TAAGGATATA ATAAGGATAG GATAATAAAA AGACCAAATA GACCAGGAAG
JEZno3  TAAGGATATA ATAAGGATAG GATAATAAAA AGACCAAATA GACCAGGAAA
   KOR  TAAGGATATA ATAAGGATAG GATAATAAAA AGACCAAATA GACCAGGAAA
   KOY  TAAGGATATA ATAAGGATAG GATAATAAAA AGACCAAATA GACCAGGAAA
   LIK  TAAGGATATA ATAAGGATAG GATAATAAAA AGACCATATA GACCAAGAAA
MARno4  TAAGGATATA ATAAGGATAG GATAATAAAA AGACCAAATA GACCAGGAAA
   MAT  TAAGGATATA ATAAGGATAG GATAATAAAA AGACCAAATA GACCAGGAAA
   MAX  TAAGGATATA A?AAGGATAG GATA?TAAAA AGACCA?ATA GACCA?GAA?
   MEX  TAAGGATATA ATAAGGATAG GATAATAAAA AGACCAAATA GACCAGGAAA
```

```
   MEY  TAAGGATATA ATAAGGATAG GATAATAAAA AGACCAAATA GACCAGGAAA
   MON  TAAGGATATA ATAAGGATAG GATAATAAAA AGACCAAATA GACCAGGAAA
   MOR  TAAGGATATA ATAAGGATAG GATAATAAAA AGACCAAATA GACCAGGAAA
   OBO  TAAGGATATA ATAAGGATAG GATAATAAAA AGACCAAATA GACCAGGAAA
   OMO  TAAGGATATA ATAAGGATAG GATAATAAAA AGACCAAATA GACCAGGAAA
   ORI  TAAGGATATA ATAAGGATAG GATAATAAAA AGACCAAATA GACCAGGAAA
PINBANK TAAGGATATA AAAAGGATAG GATAATAAAA AGACCAAATA GACCAAGAAA
 PINKOR TAAGGATATA A?AAGGATAG GATA?TAAAA AGACCA?ATA GACCA?GAA?
 PUNnol TAAGGATATA ATAAGGATAG GATAATAAAA AGACCAAATA GACCAGGAAA
   PUR  TAAGGATATA ATAAGGATAG GATAATAAAA AGACCAAATA GACCAGGAAA
   RET  TAAGGATATA ATAAGGATAG GATAATAAAA AGACCAAATA GACCAGGAAG
   RUB  TAAGGATATA ATAAGGATAG GATAATAAAA AGACCAAATA GACCAGGAAA
   SCH  TAAGGATATA ATAAGGATAG GATAATAAAA AGACCAAATA GACCAGGAAA
 SITnol TAAGGATATA ATAAGGATAG GATAATAAAA AGACCAAATA GACCAGGAAA
   SMI  TAAGGATATA ATAAGGATAG GATAGTAAAA AGACCAAATA GACCAGGAAA
   SPI  TAAGGATATA ATAAGGATAG GATAATAAAA AGACCAAATA GACCAGGAAA
   TOR  TAAGGATATA ATAAGGATAG GATAATAAAA AGACCAAATA GACCA?GAAA
   WIL  TAAGGATATA AAAAGGATAG GATAATAAAA AGACCAAATA GACCAAGAAA

        [1551                                                1600]
 ABInol GTACTTACAA AGCCCCAGGG AAGTATTTTC ATGGCCCATA CTTTATTTAT
   ALC  GTACTTAAAA AGCCCCAGGG AAGTCTTTTC ATGGCCCATA CTTTATTTAC
   ASP  GTACTTACAA AGCCCCAGGG AAGTCTTTTC ATGGCCCATA CTTTATTTAC
   BRA  GTACTTACAA AGCCCCAGGG AAGTATTTTC ATGGCCCATA CTTTATTTAC
   BRE  GTACTTACAA AGCCCCAGGG AAGTCTTTTC ATGGCCCATA CTTTATTTAC
   CHI  GTACTTACAA AGCCCCAGGG AAGTCTTTTC ATGGCCCATA CTTTATTTAC
   CRA  GTACTTACAA AGCCCCAGGG AAGTCTTTTC ATGGCCCATA CTTTATTTAC
 ENGnol GTACTTACAA AGCCCCAGGG AAGTCTTTTC ATGGCCCATA CTTTATTTAC
   FAR  GTACTTACAA AGCCCCAGGG AAGTCTTTTC ATGGCCCATA CTTTATTTAC
 GLAnol GTACTTACAA AGCCCCAGGG AAGTCTTTTC ATGGCCCATA CTTTATTTAC
   GLE  GTACTTACAA AGCCCCAGGG AAGTCTTTTC ATGGCCCATA CTTTATTTAC
 JEZno3 GTACTTACAA AKCCCCAGGG AAGTCTTTTC ATGGCCCATA MTTTATTTAC
   KOR  GTACTTACAA AGCCCCAGGG AAGTCTTTTC ATGGCCCATA CTTTATTTAT
   KOY  GTACTTACAA AGCCCCAGGG AAGTCTTTTC ATGGCCCATA CTTTATTTAT
   LIK  GTACTT?CAA AGCCCCAGGG AAGTCTTTTC ATGGCCCATA CTTTATTTAC
 MARno4 GTACTTACAA AGCCCCAGGG AAGTCTTTTC ATGGCCCATA CTTTATTTAC
   MAT  GTACTTACAA AGCCCCAGGG AAGTCTTTTC ATGGCCCATA CTTTATTTAC
   MAX  GTACTT?CAA AGCCCCAGGG AAGTCTTTTC ATGGCCCATA CTTTATTTAC
   MEX  GTACTTACAA AGCCCCAGGG AAGTCTTTTC ATGGCCCATA CTTTATTTAC
   MEY  GTACTTACAA AGCCCCAGGG AAGTCTTTTC ATGGCCCATA CTTTATTTAT
   MON  GTACTTACAA AGCCCCAGGG AAGTCTTTTC ATGGCCCATA CTTTATTTAT
   MOR  GTACTTACAA AGCCCCAGGG AAGTCTTTTC ATGGCCCATA CTTTATTTAC
   OBO  GTACTTACAA AGCCCCAGGG AAGTCTTTTC ATGGCCCATA CTTTATTTAC
   OMO  GTACTTACAA AGCCCCAGGG AAGTcTTTTC ATGGCCCATA cTTTATTTAc
   ORI  GTACTTACAA AGCCCCAGGG AAGTCTTTTC ATGGCCCATA CTTTATTTAC
PINBANK GTACTTACAA AGCCCCAGGG AAGTCTTTTC ATGGCCCATA ATTTATTTAC
 PINKOR GTACTTACAA AGCCCCAGGG AAGTCTTTTC ATGGCCCATA CTTTATTTAC
 PUNnol GTACTTACAA AGCCCCAGGG AAGTCTTTTC ATGGCCCATA CTTTATTTAC
   PUR  GTACTTACAA AGCCCCAGGG AAGTCTTTTC ATGGCCCATA CTTTATTTAC
   RET  GTACTTACAA AGCCCCAGGG AAGTCTTTTC ATGGCCCATA CTTTATTTAC
   RUB  GTACTTACAA AGCCCCAGGG AAGTCTTTTC ATGGCCCATA CTTTATTTAC
   SCH  GTACTTACAA AGCCCCAGGG AAGTCTTTTC ATGGCCCATA CTTTATTTAC
 SITnol GTACTTACAA AGCCCCAGGG AAGTCTTTTC ATGGCCCATA CTTTATTTAC
   SMI  GTACTTACAA AGCCCCAGGG AAGTCTTTTC ATGGCCCATA CTTTATTTAC
   SPI  GTACTTACAA AGCCCCAGGG AAGTCTTTTC ATGGCCCATA CTTTATTTAC
   TOR  GTACTTACAA AGCCCCAGGG AAGTCTTTTC ATGGCCCATA CTTTATTTAC
   WIL  GTACTTGCAA AG??CCAGGG AAGTCTTTTC ATGGCCCATA CTTTATTTA?

        [1601                                                1650]
 ABInol GAATGAAKAC CTTCATTGTC TCAAATTTGA GGATCTTTAT ACGCTTGACC
   ALC  GAATGAAGAC CTTCACTGTC TCAAATTTGA GGATCTTTAT ACGCTTGACC
   ASP  GAATGAAGAC CTTCACTGTC TCAAATTTGA GGATCCTTAT ACGCTTGACC
   BRA  GAATGAAGAC CTTCACTGTC TCAAATTTGA GGATCTTGAT ACGCTTGACC
   BRE  GAATGAAGAC CTTCACTGTC TCAAATTTGA GGATCTTGAT ACGCTTGACC
   CHI  GAATGAAGAC CTTCACTGTC TCAAATTTGA GGATCTTGAT ACGCTTGACC
   CRA  GAATGAAGAC CTTCACTGTC TCAAATTTGA GGATCTTTAT ACGCTTGACC
 ENGnol GAATGAAGAC CTTCACTGTC TCAAATTTGA GGATCTTGAT ACGCTT....
   FAR  GAATGAAGAC CTTCACTGTC TCAAATTTGA GGATCTTGAT ACGCTTGACC
 GLAnol GAATGAAGAC CTTCACTGTC TCAAATTTGA GGATCTTGAT ACGCTT....
   GLE  GAATGAAGAC CTTCACTGTC TCAAATTTGA GGATCTTTAT ACGCTTGACC
 JEZno3 GAATGAAGAC CTTCAYTGTC TCAAATTTGA GGATCTTTAT ACGCTTGACC
   KOR  GAATGAAGAC CTTCACTGTC TCAAATTTGA GGATCTTTAT ATGCTTGACC
   KOY  GAATGAAGAC CTTCACTGTC TCAAATTTGA GGATCTTTAT ATGCTTGACC
   LIK  GAATGAAGAC CTTCACTGTC TCAAATTTGA GGATCTTTAT ACGCTTGACC
 MARno4 GAATGAAGAC CTTCACTGTC TCAAATTTGA GGATCTTGAT ACGCTTGACC
   MAT  GAATGAAGAC CTCCACTGTC TCAAATTTGA GGATCTTGAT ACGCTTGACC
   MAX  GAATGAAGAC CTTCACTGTC TCAAATTTGA GGATCTTTAT ACGCTTGACC
   MEX  GAATGAAGAC CTTCACTGTC TCAAATTTGA GGATCTTGAT ACGCTTGACC
   MEY  GAATGAAGAC CTTCACTGTC TCAAATTTGA GGATCTTTAT ATGCTTGACC
   MON  GAATGAAGAC CTTCACTGTC TCAAATTTGA GGATCTTTAT ATGCTTGACC
   MOR  GAATGAAGAC CTTCACTGTC TCAAATTTGA GGATCTTGAT ACGCTTGACC
   OBO  GAATGAAGAC CTTCACTGTC TCAAATTTGA GGATCTTGAT ACGCTTGACC
   OMO  GAATGAAgAC CTtCAcTGTC TCAAATTTGA GGATCtTgAT AcGCTTgacc
   ORI  GAATGAAGAC CTTCACTGTC TCAAATTTGA GGATCTTTAT ACGCTTGACC
```

```
PINBANK  GAATGAAGAC CTTCACTGTC TCAAATTTGA GGATCTTTAT ACGCTTGACC
 PINKOR  GAATGAAGAC CTTCACTGTC TCAAATTTGA GGATCTTGAT ACGCTT....
 PUNno1  GAATGAAGAC CTTCACTGTC TCAAATTTGA GGATCTTGAT ACGCTTGACC
    PUR  GAATGAAGAC CTTCACTGTC TCAAATTTGA GGATCTTGAT ACGCTTGACC
    RET  GAATGAAGAC CTTCACTGTC TCAAATTTGA GGATCTTTAT ACGCTTGACC
    RUB  GAATGAAGAC CTTCACTGTC TCAAATTTGA GGATCTTGAT ACGCTTGACC
    SCH  GAATGAAGAC CTTCACTGTC TCAAATTTGA GGATCTTGAT ACGCTTGACC
 SITno1  GAATGAAGAC CTTCACTGTC TCAAATTTGA GGATCTTGAT ACGCTT....
    SMI  GAATGAAGAC CTTCACTGTC TCAAATTTGA GGATCTTGAT ACGCTTGACC
    SPI  GAATGAAGAC CTTCACTGTC TCAAATTTGA GGATCTTGAT ACGCTTGACC
    TOR  GAATGAAGAC CTTCACTGTC TCAAATTTGA GGATC?TTAT ACGCTTGACC
    WIL  GAATGAAGAC CTTCACTGTC TCAAATTTGA GGATCTTTAT ACGCTTGACC

         [1651   +1                                    +1        1700]
 ABIno1  A0ACGGGCCRC CATCGAAGAA AATAGGCTAT AACGAAAGTA0 GTCACCTTTA
    ALC  A0ACGGGCCAC CATCGAAGAA AATAGGCTAT AACGAAAGTT0 ATCACCTTTA
    ASP  A0ACGGGCCGC CATCGAAGAA AATAGGCTAT AACGAAAGTA0 GTCACCTTTA
    BRA  A0ACGGGCCGC CATCGAAGAA AATAGGCTAT AACGAAAGTA0 GTCGCCTTTA
    BRE  A0ACGGGCCGC CATCGAAGAA AATAGGCTAT AACGAAAGTA0 GTCGCCTTTA
    CHI  A0ACGGGCCGC CATCGAAGAA AATAGGCTAT AACGAAAGTA0 GTCGCCTTTA
    CRA  A0ATGGGCCTC CATCAAAGAA AATAGGCTAT AACGAAAGTA0 GTCGCCTTTA
 ENGno1  .1ACGGGCCGC CATCGAAGAA AATAGGCTAT AACGAAAGTA0 GTCGCCTTTA
    FAR  A0ACGGGCCGC CATCGAAGAA AATAGGCTAT AACGA.....1 GTCGCCTTTA
 GLAno1  .1ACGGGCCGC CATCGAAGAA AATAGGCTAT AACGAAAGTA0 GTCGCCTTTA
    GLE  A0ATGGGCCTC CATCAAAGAA AATAGGCTAT AACGAAAGTA0 GTCGCCTTTA
 JEZno3  A0ACGGGCCGC CATCGAAGAA AATAGGCTAT AACGAAAGTA0 GTCGCCTTTA
    KOR  A0ACAGGCCGA CATCGAAGAA AATAGGCTAT AACGAAAGTA0 GTCGCCTTTA
    KOY  A0ACAGGCCGA CATCGAAGAA AATAGGCTAT AACGAAAGTA0 GTCGCCTTTA
    LIK  A0ACGGGCCGC CATCGAAGAA AATA?GCTAT AACGAAAGTA0 GTCGCCTTTA
 MARno4  A0ACGGGCCGC CATCGAAGAA AATAGGCTAT AACGAAAGTA0 GTCGCCTTTA
    MAT  A0ACGGGCCGC CATCGAAGAA AATAGGCTAT AACGAAAGTA0 GTCGCCTTTA
    MAX  A0ACGGGCCGC CATCGAA?AA AATAGGCTAT AACGAAAGTA0 GTCGCCTTTA
    MEX  A0ACGGGCCGC CATCGAAGAA AATAGGCTAT AACGAAAGTA0 GTCGCCTTTA
    MEY  A0ACAGGCCGA CATCGAAGAA AATAGGCTAT AACGAAAGTA0 GTCGCCTTTA
    MON  A0ACAGGCCGA CATCGAAGAA AATAGGCTAT AACGAAAGTA0 GTCGCCTTTA
    MOR  A0ACGGGCCGC CATCGAAGAA AATAGGCTAT AACGAAAGTA0 GTCGCCTTTA
    OBO  A0ACGGGCCGC CATCGAAGAA AATAGGCTAT AACGAAAGTA0 GTCGCCTTTA
    OMO  a0AcggGCCgc CATCgAAGAA AATAGGCTAT AACGAAAGTA0 GTCGCCTTTA
    ORI  A0ACG?GCCGC CATCGAAGAA AATAGGC?AT AA?GAAAGTA0 GT?GCCTTTA
PINBANK  A0ACGTGCCGC CATCGAAGAA AATAGGCCAT AATGAAAGTA0 GTTGCCTTTA
 PINKOR  .1ACGGGCCGC CATCGAAGAA AATAGGCTAT AACGAAAGTA0 GTCGCCTTTA
 PUNno1  A0ACGGGCCGC CATCGAAGAA AATAGGCTAT AACGAAAGTA0 GTCGCCTTTA
    PUR  A0ACGGGCCGC CATCGAAGAA AATAGGCTAT AACGAAAGTA0 GTCGCCTTTA
    RET  A0ATGGGCCTC CATCAAAGAA AATAGGCTAT AACGAAAGTA0 GTCGCCTTTA
    RUB  A0ACGGGCCGC CATCGAAGAA AATAGGCTAT AACGAAAGTA0 GTCGCCTTTA
    SCH  A0ACGGGCCGC CATCGAAGAA AATAGGCTAT AACGAAAGTA0 GTCGCCTTTA
 SITno1  .1ACGGGCCGC CATCGAAGAA AATAGGCTAT AACGAAAGTA0 GTCGCCTTTA
    SMI  A0ACGGGCCGC CATCGAAGAA AATAGGCTAT AACGAAAGTA0 GTCACCTTTA
    SPI  A0ACGGGCCGC CATCGAAGAA AATAGGCTAT AACGAAAGTA0 GTCGCCTTTA
    TOR  A0A?GGGCC?C CATC?AAGAA AATAGGCTAT AACGAAAGTA0 GTC?CCTTTA
    WIL  A0A?GGGCCGC CATCGAAGAA AATAGGCTAT AACGAAAGTA0 GTCGCCTTTA

         [1701                                        +1      1750]
 ABIno1  CCCTTCAATG AGTGGACTAA A..... ..... .0GTAGTGGTG
    ALC  GCCTTCAATG AGTGGACTAA A..... ..... .0GTAGTGGTG
    ASP  GCCTTCAATG AGTGGACTAA A..... ..... .0GTAGTGGTG
    BRA  GCCTTCAATG AGTGGACTAA A..... ..... .0GTAGTGGTG
    BRE  GCCTTCAATG AGTGGACTAA A..... ..... .0GTAGTGGTG
    CHI  GCCTTCAATG AGTGGACTAA A..... ..... .0GTAGTGGTG
    CRA  GCCTTCAATG AGTGGACTAA A..... ..... .0GTAGTGGTG
 ENGno1  GCCTTCAATG AGTGGACTAA A..... ..... .0GTAGTGGTG
    FAR  GCCTTCAATG AGTGGACTAA A..... ..... .0GTAGTGGTG
 GLAno1  GCCTTCAATG AGTGGACTAA ACCTTCAATG AGTGGACTAA A1GTAGTGGTG
    GLE  GCCTTCAATG AGTGGACTAA A..... ..... .0GTAGTGGTG
 JEZno3  GCCTTCAATG AGTGGACTAA A..... ..... .0GTAGTGGTG
    KOR  GCCTTCAATG AGTGGACTAA A..... ..... .0GTAGTGGTG
    KOY  GCCTTCAATG AGTGGACTAA A..... ..... .0GTAGTGGTG
    LIK  GCCTTCAATG AGTGGACTAA A..... ..... .0GTAGTGGTG
 MARno4  GCCTTCAATG AGTGGACTAA A..... ..... .0GTAGTGGTG
    MAT  GCCTTCAATG AGTGGACTAA A..... ..... .0GTAGTGGTG
    MAX  ?CCTTCAATG AGTGGACTAA A..... ..... .0GTAGTGGTG
    MEX  GCCTTCAATG AGTGGACTAA A..... ..... .0GTAGTGGTG
    MEY  GCCTTCAATG AGTGGACTAA A..... ..... .0GTAGTGGTG
    MON  GCCTTCAATG AGTGGACTAA A..... ..... .0GTAGTGGTG
    MOR  GCCTTCAATG AGTGGACTAA A..... ..... .0GTAGTGGTG
    OBO  GCCTTCAATG AGTGGACTAA A..... ..... .0GTAGTGGTG
    OMO  GCCTTCAATG AGTGGACTAA A..... ..... .0GTAGTGGTG
    ORI  GCCTTCAATG AGTGGACTAA A..... ..... .0GTAGTGGTG
PINBANK  GCCTTCAATG AGTGGACTAA A..... ..... .0GTAGTGTTG
 PINKOR  ?CCTTCAATG AGTGGACTAA A????????? ?????????? ??GTAGTG?TG
 PUNno1  GCCTTCAATG AGTGGACTAA A..... ..... .0GTAGTGGTG
    PUR  GCCTTCAATG AGTGGACTAA A..... ..... .0GTAGTGGTG
    RET  GCCTTCAATG AGTGGACTAA A..... ..... .0GTAGTGGTG
    RUB  GCCTTCAATG AGTGGACTAA A..... ..... .0GTAGTGGTG
```

```
   SCH  GCCTTCAATG AGTGGACTAA A......... .......... .0GTAGTGGTG
SITnol  GCCTTCAATG AGTGGACTAA A......... .......... .0GTAGTGGTG
   SMI  GCCTTCAATG AGTGGACTAA A......... .......... .0GTAGTGGTG
   SPI  GCCTTCAATG AGTGGACTAA A......... .......... .0GTAGTGGTG
   TOR  GCCTTCAATG AGTGGACTAA A......... .......... .0GTAGTGGTG
   WIL  ?CCTTCAATG AGTGGACTAA A......... .......... .0GTAGTGGTG

        [1751                indel important    +1         1800]
ABInol  GAACATTGTT ATGGAGTGGG CGTCTTTTGG GG.....0GGT TATGKTCCCA
   ALC  GAACATTGTT ATGGAGTGGG CGTCTTTTGG GG.....0GGT TATGGTCCCG
   ASP  GAACATTGTT ATGGAGTGGG AGTCTTTTGG GG.....0GGT TATGGTCCCG
   BRA  GAACATTGTT ATGGAGTGGG CGTCTTTTGG GGTATCG1GGT TATGGTCCCG
   BRE  GAACATTGTT ATGGAGTGGG CGTCTTTTGG GGTATCG1GGT TATGGTCCCG
   CHI  GAACATTGTT ATGGAGTGGG CGTCTTTTGG GGTATCG1GGT TATGGTCCCG
   CRA  GAACATTGTT ATGGAGTGGG TGTCTTTTGG GG.....0GGT TATGGTCCCG
ENGnol  GAACATTGTT ATGGAGTGGG CGTCTTTTGG GGTATCG1GGT TATGGTCCCG
   FAR  GAACATTGTT ATGGAGTGGG CGTCTTTTGG GGTATCG1GGT TATGGTCCCG
GLAnol  GAACATTGTT ATGGAGTGGG CGTCTTTTGG GGTATCG1GGT TATGGTCCCG
   GLE  GAACATTGTT ATGGAGTGGG TGTCTTTTGG GG.....0GGT TATGGTCCCG
JEZno3  GAACATTGTT ATGGAGTGGG CRTCTTTTGG GG.....0GGT TATGGTCCCG
   KOR  GAACATTGTT ATGGAGTGGG CATCTTTTGG GG.....0GGT TATGGTCCCA
   KOY  GAACATTGTT ATGGAGTGGG CATCTTTTGG GG.....0GGT TATGGTCCCA
   LIK  GAACATTGTT ATGGAGTGGG CGTCTTTTGG GG.....0GGT TATGGTCCC?
MARno4  GAACATTGTT ATGGAGTGGG CGTCTTTTGG GGTATCG1GGT TATGGTCCCG
   MAT  GAACATTGTT ATGGAGTGGG CGTCTTTTGG GGTATCG1GGT TATGGTCCCG
   MAX  GAACATTGTT ATGGAGTGGG C?TCTTTTGG GGTATCA1GGT TATGGTCCC?
   MEX  GAACATTGTT ATGGAGTGGG CGTCTTTTGG GGTATCG1GGT TATGGTCCCG
   MEY  GAACATTGTT ATGGAGTGGG CATCTTTTGG GG.....0GGT TATGGTCCCA
   MON  GAACATTGTT ATGGAGTGGG CATCTTTTGG GG.....0GGT TATGGTCCCA
   MOR  GAACATTGTT ATGGAGTGGG CGTCTTTTGG GGTATCG1GGT TATGGTCCCG
   OBO  GAACATTGTT ATGGAGTGGG CGTCTTTTGG GGTATCG1GGT TATGGTCCCG
   OMO  GAACATTGTT ATGGAGTGGG CGTCTTTTGG GGTATCG1GGT TATGGTCCCG
   ORI  GAACATTGTT ATGGAGTGGG CGTCTTTTGG GG.....0GGT TATGGTCCCG
PINBANK  GAACATTGTT ATGGAGTGGG CGTCTTTTGG GG.....0GGT TATGGTCCCG
PINKOR  GAACATTGTT AT?GAGTGGG ??T?TTTTGG GG???????GT TAT???????
PUNnol  GAACATTGTT ATGGAGTGGG CGTCTTTTGG GGTATCG1GGT TATGGTCCCG
   PUR  GAACATTGTT ATGGAGTGGG CGTCTTTTGG GGTATCG1?GT TATGGTCCCG
   RET  GAACATTGTT ATGGAGTGGG TGTCTTTTGG GG.....0GGT TATGGTCCCG
   RUB  GAACATTGTT ATGGAGTGGG CGTCTTTTGG GGTATCG1GGT TATGGTCCCG
   SCH  GAACATTGTT ATGGAGTGGG CGTCTTTTGG GGTATCG1GGT TATGGTCCCG
SITnol  GAACATTGTT ATGGAGTGGG CGTCTTTTGG GGTATCG1GGT TATGGTCCCG
   SMI  GAACATTGTT ATGGAGTGGG C?T?TTTTGG GGTATTG1GGT TATGGTCCC?
   SPI  GAACATTGTT ATGGAGTGGG CGTCTTTTGG GGTATCG1GGT TATG.T....
   TOR  GAACATTGTT ATGGAGTGGG ?GTCTTTTGG GG.....0GGT TATGGTCCCG
   WIL  GAACATTGTT ATGGAGTGGG CGTCTTTTGG GG.....0GGT TATGGTCCCG

        [1801 +1                        +1   +1               1850]
ABInol  T0TGGTCTTCG TGTCTAAA.. .......0GAA0 CCTTTCCCAA TGCGGACGAG
   ALC  T0TGGTCTTCG TGTCTAAA.. .......0GAA0 CCTCGCCCGA TGTGGACGAG
   ASP  T0TGGTCTTCG TGTCTAAA.. .......0GAA0 CCTCGCCTGA TGCGGACGAA
   BRA  T0TGGTCTTCG TGTCTAAAGT GTCTAAA1GAA0 CCTCGCCCGA TGCGGACGAG
   BRE  T0TGGTCTTCG TGTCTAAA.. .......0GAA0 CCTCGCCCGA TGCGGACGAG
   CHI  T0TGGTCTTCG TGTCTAAA.. .......0...1 CCTCGCCCGA TGCGGACGAG
   CRA  T0TGGTCTTCA TGTCTAAA.. .......0TAA0 CCTCGCCCGA TGCGGACGAG
ENGnol  T0TGGTCTTCG TGTCTAAA.. .......0...1 CCTCGCCCGA TGCGGACGAG
   FAR  T0TGGTCTTCG TGTCTAAAGT GTCTAAA1GAA0 CCTCGCCCGA TGCGGACGAG
GLAnol  T0TGGTCTTCG TGTCTAAA.. .......0GAA0 CCTCGCCCGA TGCGGACGAG
   GLE  T0TGGTCTTCA TGTCTAAA.. .......0TAA0 CCTCGCCCGA TGCGGACGAG
JEZno3  T0TGGTCTTCR TGTCTAAA.. .......0...1 CCTCGCCCGA TGCGGACGAG
   KOR  T0TGGTCTTCG TGTCTAAA.. .......0...1 CCTCGTCCGA TGCGGATGAG
   KOY  T0TGGTCTTCG TGTCTAAA.. .......0...1 CCTCGTCCGA TGCGGATGAG
   LIK  T0TGGTCTTCG TGTCTAAA.. .......0GAA0 CCT?GCCCGA TG??GACGAG
MARno4  T0TGGTCTTCG TGTCTAAA.. .......0GAA0 CCTCGCCCGA TGCGGACGAG
   MAT  T0TGGTCTTCG TGTCTAAA.. .......0GAA0 CCTCGCCCGA TGCGGACGAG
   MAX  T0TGGTCTTC? TGTCTAAA?? ?????????AA? CCT???C??A TG??GA?GA?
   MEX  T0TGGTCTTCG TGTCTAAA.. .......0...1 CCTCGCCCGA TGCGGACGAG
   MEY  T0TGGTCTTCG TGTCTAAA.. .......0...1 CCTCGTCCGA TGCGGATGAG
   MON  T0TGGTCTTCG TGTCTAAA.. .......0...1 CCTCGTCCGA TGCGGATGAG
   MOR  T0TGGTCTTCG TGTCTAAA.. .......0GAA0 CCTCGCCCGA TGCGGACGAG
   OBO  T0TGGTCTTCG TGTCTAAA.. .......0GAA0 CCTCGCCCGA TGCGGACGAG
   OMO  T0TGGTCTTCG TGTCTAAA.. .......0gaa0 CCTCGCCCGA TGCGGACGAG
   ORI  T0TGGTCTTCG TGTCTAAA.. .......0GAA0 CCTCGCCCGA TGCGGACGAG
PINBANK  T0TGGTCTTCA TGTCTAAA.. .......0GAA0 CCTCGCCCGA TGCGGACGAG
PINKOR  T0TGGTCTTC? TGTCTAAA?? ?????????AA? CCT???C??A TG??GA?GA?
PUNnol  T0TGGTCTTCG TGTCTAAA.. .......0...1 CCTCGCCCGA TGCGGACGAG
   PUR  T0TGGTCTTCG TGTCTAAA.. .......0...1 CCTCGCCCGA TGCGGACGAG
   RET  T0TGGTCTTCA TGTCTAAA.. .......0...1 CCTCGCCCGA TGCGGACGAG
   RUB  T0TGGTCTTCG TGTCTAAA.. .......0...1 CCTCGCCCGA TGCGGACGAG
   SCH  T0TGGTCTTCG TGTCTAAA.. .......0GAA0 CCTCGCCCGA TGCGGACGAG
SITnol  T0TGGTCTTCG TGTCTAAA.. .......0GAA0 CCTCGCCCGA TGCGGACGAG
   SMI  T0TGGTCTTCG TGTCTAAAGT GTCTAAA1GAA0 CCTCGCCCGA TGCGGACGAG
   SPI  .1TGGTCTTCG TGTCTAAAGT GTCTAAA1GAA0 CCTCGCCCGA TGCGGACGAG
   TOR  T0TGGTCTTC? TGTCTAAA.. .......0?AA0 CCTCGCC?GA TGCGGACGA?
   WIL  T0TGGTCTTCG TGTCTAAA.. .......0GAA0 CCTCGCCCGA TGCTGACGAG
```

```
          [1851                                                    1900]
ABInol    CCATCCTTCC TCGCCCTCCT CTATCAGATG AGAAATTCTT CCTCTGGTCT
   ALC    CCATCCTTCC TCGCCCTCCT CTGTCGGATG AGCAATTCTT CCTCTGGTCT
   ASP    CCATCCTTCC TCGCCCTCCT CTGTCGGACA AGAAATTCTT CCTCTGGTCT
   BRA    CCATCCTTCC TCGCCCTCCT CTGTCGGACG AGCAATTATT CCTCTGGTCT
   BRE    CCATCCTTCC TCGCCCG... .......... .......... ..........
   CHI    CCATCCTTCC TCGCCCTCCT CTGTCGGACG AACAATTATT CCTCTGGTCT
   CRA    CCATCCTTCA TCG?CCTCCT CTGTCGGACG AGCAATTCTT CCTCTGGTCT
ENGnol    CCATCCTTCC TCGCCCTCCT CTGTCGGACG AACAATTCTT CCTCTGGTCT
   FAR    CCATCCTTCC TCGCCCTCCT CTGTCGGACG AGCAATTATT CCTCTGGTCT
GLAnol    CCATCCTTCC TCGCCCTCCT CTGTCGGACG AACAATTCTT CCTCTGGTCT
   GLE    CCATCCTTCA TCGCCCTCCT CTGTCGGACG AGCAATTCTT CCTCTGGTCT
JEZno3    CTATACTTCC TCGCCATCCT CTGTCGGACG AGCAATTCTT CCTCTAGTCT
   KOR    CCATCCTTCC TCGCCCTCCT TTGTCGAACG ATCAATTCTT CCTCTGGTCT
   KOY    CCATCCTTCC TCGCCCTCCT TTGTCGAACG ATCAATTCTT CCTCTGGTCT
   LIK    CCATCCTTCC TCGCCCTCCT CTGTCGGA?G AGCAATTCTT CCTCTGGTCT
MARno4    CCATCCTTCC TCGCCCTCCT CTGTCGGACG AACAATTCTT CCTCTGGTCT
   MAT    CCATCCTTCC TCGCCCTCCT CTGTCGGACG AACAATTCTT CCTCTGGTCT
   MAX    C?AT?CTTC? ????C???CT ?T?TC??A?? A??AATT?TT CCT?T?GT?T
   MEX    CCATCCTTCC TCGCCCTCCT CTGTCGGACG AACAATTCTT CCTCTGGTCT
   MEY    CCATCCTTCC TCGCCCTCCT TTGTCGAACG ATCAATTCTT CCTCTGGTCT
   MON    CCATCCTTCC TCGCCCTCCT TTGTCGAACG ATCAATTCTT CCTCTGGTCT
   MOR    CCATCCTTCC TCGCCCTCCT CTGTCGGACG AGCAATTATT CCTCTGGTCT
   OBO    CCATCCTTCC TCGCCCG... .......... .......... ..........
   OMO    CCATCCTTCC TCGCCCTCCT CTGTCGGACG AACAATTCTT CCTCTGGTCT
   ORI    CCATCCTTCC TCGCCCTCCT CTGTCGGACG AGCAATTCTT CCTCTGGT?T
PINBANK   CCATCCTTCC CCCCCCTCCT CTGTCGGACG AGCAATTCTT CCTTTGGTAT
PINKOR    C?AT?CTTC? ????C???CT ?T?TC??A?? A??AATT?TT CCT?T?GT?T
PUNnol    CCATCCTTCC TCGCCCTCCT CTGTCGGACG AACAATTCTT CCTCTGGTCT
   PUR    CCATCCTTCC TCGCCCTCCT CTGTCGGACG AGCAATTCTT CCTCTGGTCT
   RET    CCATCCTTCA TCGCCCTCCT CTGTCGGACG AGCAATTCTT CCTCTGGTCT
   RUB    CCATCCTTCC TCGCCCTCCT CTGTCGGACG AACAATTCTT CCTCTGGTCT
   SCH    CCATCCTTCC TCGCCCTCCT CTGTCGGACG AGCAATTCTT CCTCTGGTCT
SITnol    CCATCCTTCC TCGCCCTCCT CTGTCGGACG AACAATTCTT CCTCTGGTCT
   SMI    CCATCCTTCC TT?CCCTCCT CTGTCGGACG AGCAATTCTT CCTCTGGTCT
   SPI    CCATCCTTCC TCGCCCTCCT CTGTCGGACG AGCAATTATT CCTCTGGTCT
   TOR    CCATCCTTCC TCGCCCTCCT CTGTCGGAC? AG?AATTCTT CCTCTGGTCT
   WIL    CCATCCTTCC TCGCCCTCCT CTGTCGGATG AG?AATTCTT CCTCTGGTCT

          [1901                                +1          1950]
ABInol    TTAGAGCGGT TCGGTTCCCT CACCCTCACC CTC0CKGCGTC GGGCGAGTCA
   ALC    TTAGAGCGGT TCGGTTCCCT CACCCTCACC CTC0CTGCATC GGGCGATTCA
   ASP    TTAGAGCGGT TCGGTTCCCT CACCCTCACC CTC0CTGCATC GGGCGATTCA
   BRA    TTAGAGCGGT TCGGTTCCCT CACCCTC... ...1CTGCGTC GGGCGAGTCA
   BRE    .......... .......... .......... ...?....... ..........
   CHI    TTAGAGCGGT TCGGTTCCCT CACCCTCGCC CTC0CTGCGTC GGGCGAGTCA
   CRA    TTAGAGCGGT TCGGTTCCCT CACCCTCGCC CTC0CTGCATC GGGAAATTCA
ENGnol    TTAGAGCGGT TCGGTTCCCT CACCCTCGCC CTC0CTGCGTC GGGCGAGTCA
   FAR    TTAGAGCGGT TCGGTTCCCT CACCCTC... ...1CTGCGTC GGGCGAGTCA
GLAnol    TTAGAGCGGT TCGGTTCCCT CACCCTCGCC CTC0CTGCGTC GGGCGAGTCA
   GLE    TTAGAGCGGT TCGGTTCCCT CACCCTCGCC CTC0CTGCATC GGGAAATTCA
JEZno3    TTAGAGCGGT TCGGTTCCCT CACCCTCGCC CTC0CTGCRTC GGGCGAGTCA
   KOR    TTAGAGTGGT TCGGTTCCCT CACCCTCGCC CTC0CTGCGTT GGGCGAGTCA
   KOY    TTAGAGTGGT TCGGTTCCCT CACCCTCGCC CTC0CTGCGTT GGGCGAGTCA
   LIK    TTAGAGCGGT TC?GTTCCCT CA?CCTCACC CTC0CTGC?TC GGGCGATTCA
MARno4    TTAGAGCGGT TCGGTTCCCT CACCCTCGCC CTC0CTGCGTC GGGCGAGTCA
   MAT    TTAGAGCGGT TCGGTTCCCT CACCCTCGCC CTC0CTGCGTC GGGCGAGTCA
   MAX    TTAGAG?GGT TC??TTCCCT CACCCTC?CC CTC?C?GC?T? ?GG??A?TCA
   MEX    TTAGAGCGGT TCGGTTCCCT CACCCTCGCC CTC0CTGCGTC GGGCGAGTCA
   MEY    TTAGAGTGGT TCGGTTCCCT CACCCTCGCC CTC0CTGCGTT GGGCGAGTCA
   MON    TTAGAGTGGT TCGGTTCCCT CACCCTCGCC CTC0CTGCGTT GGGCGAGTCA
   MOR    TTAGAGCGGT TCGGTTCCCT CACCCTCGCC CTC0CTGCGTC GGGCGAGTCA
   OBO    .......... .......... .......... ...?....... ..........
   OMO    TTAGAGCGGT TCGGTTCCCT CACCCTCGCC CTC0CTGCGTC GGGCGAGTCA
   ORI    TTAGAGCGGT TCGGTTCCCT CACCCTCGCC CTC0CTGCGTC GGGCGAGTCA
PINBANK   TTAGAGCGGT TCGGTTCCCT CACCCTCGCC CTC0CTGCGTC GGGCGAGTCA
PINKOR    TTAGAG?GGT TC??TTCCCT CA?????CC CTC?C?GC?T? ?GG??A?TCA
PUNnol    TTAGAGCGGT TCGGTTCCCT CACCCTCGCC CTC0CTGCGTC GGGCGAGTCA
   PUR    TTAGAGCGGT TCGGTTCCCT CACCCTC... ...1CTGCGTC GGGCGAGTCA
   RET    TTAGAGCGGT TCGGTTCCCT CACCCTCGCC CTC0CTGCATC GGGAAATTCA
   RUB    TTAGAGCGGT TCGGTTCCCT CACCCTCGCC CTC0CTGCGTC GGGCGAGTCA
   SCH    TTAGAGCGGT TCGGTTCCCT CACCCTC... ...1CTGCGTC GGGCGAGTCA
SITnol    TTAGAGCGGT TCGGTTCCCT CACCCTCGCC CTC0CTGCGTC GGGCGAGTCA
   SMI    TTAGAGCGGT TCGGTTCCCT CACCCTC... ...1CTGCGTC GGGCGAGTCA
   SPI    TTAGAGCGGT TCGGTTCCCT CACCCTC... ...1CTGCGTC GGGCGAGTCA
   TOR    TTAGAGCGGT TCGGTTCCCT CACCCTC?CC CTC0CTGCATC GGG??ATTCA
   WIL    TTAGAGCGGT TCGGTTCCCT CACCCTCACC CTC0CTGCGTC GGGCAATTCA

          [1951                      +1    +1              2000]
ABInol    CTTAAAAGTA ACTTATCTGG TGGGGGG.1CC 0TTGGCTAGAA TTAGTCAATT
   ALC    CTTAAAAGTA ACTTATCTGG TGGGGGGG0CC 0TTGGCCGGAA TTCGTCAATT
   ASP    CTTAAAAGTA ACTTATCTGG TGGGGGGG0CC 0TTGGCCAGAA TTCGTCAATT
   BRA    CTTAAAAGTA ACTTATCTGG TGGGGGGG0CC 0TTGGCCGGAA TTCGTCAATT
```

```
   BRE  .......... .......... ........?.. 1ATGGCCGGAA TTCGTCAATT
   CHI  CTTAAAAGTA ACTTATCTGG TGGGGGGG0CC 0TTGGCCGGAA TTCGTCAATT
   CRA  CTTAAAAGTA ACTTATCTGG TGGGGGGG0CC 0TTGGCCGGAA TTCATCAATT
ENGno1  CTTAAAAGTA ACTTATCTGG TGGGGGGG0CC 0TTGGCCGGAA TTCGTCAATT
   FAR  CTTAAAAGTA ACTTATCTGG TGGGGGGG0CC 0TTGGCCGGAA TTCGTCAATT
GLAno1  CTTAAAAGTA ACTTATCTGG TGGGGGGG0CC 0TTGGCCGGAA TTCGTCAATT
   GLE  CTTAAAAGTA ACTTATCTGG TGGGGGGG0CC 0TTGGCCGGAA TTCATCAATT
JEZno3  CTTAAAAGTA ACTTATCTGG TGGGGGGG0CC 0TTGGCCRGAA TTCGTCAATT
   KOR  CTTAAAAGTA ACTTATCTGG TGGGGGGG0CC 0TTGGCTAGAA TTCATCAATT
   KOY  CTTAAAAGTA ACTTATCTGG TGGGGGGG0CC 0TTGGCTAGAA TTCATCAATT
   LIK  CTTAAAAGTA ACTTATCTGG TGGGGGGG0CC 0TTGGCCGGAA TTCGTCAATT
MARno4  CTTAAAAGTA ACTTATCCGG TGGGGGGG0CC 0TTGGCCGGAA TTCGTCAATT
   MAT  CTTAAAAGTA ACTTATCTGG TGGGGGGG0CC 0TTGGCCGGAA TTCGTCAATT
   MAX  CTTAAAA?TA ACTTATC??G TGGGGGGG?CC ??TGGC???AA TT??TCAATT
   MEX  CTTAAAAGTA ACTTATCTGG TGGGGGGG0CC 0TTGGCCGGAA TTCGTCAATT
   MEY  CTTAAAAGTA ACTTATCTGG TGGGGGGG0CC 0TTGGCTAGAA TTCATCAATT
   MON  CTTAAAAGTA ACTTATCTGG TGGGGGGG0CC 0TTGGCTAGAA TTCATCAATT
   MOR  CTTAAAAGTA ACTTATCTGG TGGGGGGG0CC 0TTGGCCGGAA TTCGTCAATT
   OBO  .......... .......... ........?.. 1ATGGCCGGAA TTCGTCAATT
   OMO  CTTAAAAGTA ACTTATC?GG TGGGGGGG0CC 0TTGGCCGGAA TTCGTCAATT
   ORI  CTTAAAAGTA ACTTATCTGG TGGGGGGG0CC 0TTGGCCGGAA TTCGTCAATT
PINBANK  CTTAAAAGTA ACTTATCTGG TGGGGGGG0CC 0TTGGCCGAAA TTCGTCAATT
PINKOR  CTTAAAA?TA ACTTATC??G TGGGGGG??CC ??TGGC???AA TT??TCAATT
PUNno1  CTTAAAAGTA ACTTATCTGG TGGGGGGG0CC 0TTGGCCGGAA TTCGTCAATT
   PUR  CTTAAAAGTA ACTTATCTGG TGGGGGGG0CC 0TTGGCCGGAA TTCGTCAATT
   RET  CTTAAAAGTA ACTTATCTGG TGGGGGGG0CC 0TTGGCCGGAA TTCATCAATT
   RUB  CTTAAAAGTA ACTTATCCGG TGGGGGGG0CC 0TTGGCCGGAA TTCGTCAATT
   SCH  CTTAAAAGTA ACTTATCTGG TGGGGGGG0CC 0TTGGCCGGAA TTCGTCAATT
SITno1  CTTAAAAGTA ACTTATCTGG TGGGGGGG0CC 0TTGGCCGGAA TTCGTCAATT
   SMI  CTTAAAAGTA ACTTATCTGG TGGGGGGG0CC 0TTGGCTGGAA TTCGTCAATT
   SPI  CTTAAAAGTA ACTTATCTGG TGGGGGGG0CC 0TTGGCCGGAA TTCGTCAATT
   TOR  CTTAAAAGTA ACTTATCTGG TGGGGGGG0CC 0TTGGCC?GAA TTC?TCAATT
   WIL  CTTAAAAGTA ACTTATCTGG TGGGGGGG0CC 0TTGGC?GGAA TTCGTCAATT

        [2001                                    +2        2050]
ABIno1  CAGAGACTAT TTCAAGCCTT GTATTTAT.. .........00T CCCCTAAAAA
   ALC  AAAAGACTAT TTCAAGCCCT ATATTTAT.. .........00T CCCCTAAAAA
   ASP  CAGAGACTAT TTCAAGCCCT GTATTTAT.. .........00T CCCCTAAAAA
   BRA  CAGAGACTAT TTCAAGCCCT GTATTTAT.. .........00T CCCCTAAAAA
   BRE  CAGAGACTAT TTCAAGCCCT GTATTTAT.. .........00T CCCCTAAAAA
   CHI  CAGAGACTAT TTCAAGCCCT GTATTTAT.. .........00T CCCCTAAAAA
   CRA  CAGAGACTAT TTCAAGCCCT GTATTTAT.. .........00T CCCCTAAAAA
ENGno1  CAGAGACTAT TTAAAGCCCT GTATTTAT.. .........00T CCCCTAAAAA
   FAR  CAGAGACTAT TTCAAGCCCT GTATTTAT.. .........00T CCCCTAAAAA
GLAno1  CAGAGACTAT TTAAAGCCCT GTATTTAT.. .........00T CCCCTAAAAA
   GLE  CAGAGACTAT TTCAAGCCCT GTATTTAT.. .........00T CCCCTAAAAA
JEZno3  CAGAGACTAT TTCAAGCCCT GTATTTAT.. .........00T CCCCTAAAAA
   KOR  CAAAGACTAT TTCAAGCCCT GTATTTAT.. .........00T CCCCTAAAAA
   KOY  CAAAGACTAT TTCAAGCCCT GTATTTAT.. .........00T CCCCTAAAAA
   LIK  CAGAGACTAT TTCAAGCCCT GTATTTAT.. .........00T CCCCTAAAAA
MARno4  CAGAGACTAT TTCAAGCCCT GTATTTAT.. ...ATTTAT01T CCCCTAAAAA
   MAT  CAGAGACTAT TTCAAGCCCT GTATTTAT.. .........00T CCCCTAAAAA
   MAX  ?A?A?ACTAT TT?AAGCC?T ?TATTTAT?? ??????????T ?CCCTAAAAA
   MEX  CAGAGACTAT TTCAAGCCCT GTATTTAT.. .........00T CCCCTAAAAA
   MEY  CAAAGACTAT TTCAAGCCCT ATATTTAT.. .........00T CCCCTAAAAA
   MON  CAAAGACTAT TTCAAGCCCT GTATTTAT.. .........00T CCCCTAAAAA
   MOR  CAGAGACTAT TTAAAGCCCT GTATTTAT.. .........00T CCCCTAAAAA
   OBO  CAGAGACTAT TTCAAGCCCT GTATTTAT.. .........00T CCCCTAAAAA
   OMO  CAGAGACTAT TT?AAGCCCT GTATTTAT.. ...ATTTAT01T CCCCTAAAAA
   ORI  CAGAGACTAT TTCAAGCCCT GTATTTAT.. .........00T CCCCTAAAAA
PINBANK  CAGAGACTAT TTCAAGCCCT GTATTTAT.. .........00T CCCCTAAAAA
PINKOR  ?A?A?ACTAT TT?AAGCC?T ?TATTTAT?? ??????????T ?CCCTAAAAA
PUNno1  CAGAGACTAT TTCAAGCCCT GTATTTAT.. .........00T CCCCTAAAAA
   PUR  CAGAGACTAT TTCAAGCCCT GTATTTAT.. .........00T CCCCTAAAAA
   RET  CAGAGACTAT TTCAAGCCCT GTATTTAT.. .........00T CCCCTAAAAA
   RUB  CAGAGACTAT TTCAAGCCCT GTATTTAT.. ...ATTTAT01T CCCCTAAAAA
   SCH  CAGAGACTAT TTCAAGCCCT GTATTTAT.. .........00T CCCCTAAAAA
SITno1  CAGAGACTAT TTAAAGCCCT GTATTTAT.. .........00T CCCCTAAAAA
   SMI  CAGAGACTAT TTCAAGCCCT GTATTTATTA ATAATTTAT11T ACCCTAAAAA
   SPI  CAGAGACTAT TTCAAGCCCT GTATTTAT.. .........00T CCCCTAAAAA
   TOR  CAGAGACTAT TTCAAGCCCT GTATTTAT.. .........00T CCCCTAAAAA
   WIL  CAGAGACTAT TTCAAGCCCT GTATTTAT.. .........00T CCCCTAAAAA

        [2051  +1                    +2                        2100]
ABIno1  AAAAA0GTCTG GGCYGCCATT AAATT0AAAA0G GGGCCATGAA AATACTTTGC
   ALC  AAAAA0GTCTG GGCTGCCATT AAATT0AAAA0G GGGCCATGAA AATACTTTCC
   ASP  AAAAA0GTCTG GGCCGCCATT AAATT0AAAA0G GGGCCATGAA AATACTTTCC
   BRA  AAAAA0GTCTG GGCCGCCAT. .....1AAAA0G GGGCCATGAA AATACTTTCC
   BRE  AAAAA0GTCTG GGCCGCCATT AAATT0AAAA0G GGGCCATGAA AATACTTTCC
   CHI  AAAAA0GTCTG GGCCGCCATT AAATT0AAAA0G GGGCCATGAA AAGACTTTCC
   CRA  AAAAA0GTTTG GGCCGCCATT AAATT0AAAA0G GGGCCATGAA AATACTTTCC
ENGno1  AAAAA0GTCTG GGCCGCCATT AAATT0AAAA0G GGGCCATGAA AATACTTTCC
   FAR  AAAAA0GTATG GGCCGCCAT. .....1AAAA0G GGGCCATGAA AAGACTTTCC
GLAno1  AAAAA0GTCTG GGCCGCCATT AAATT0AAAA0G GGGCCATGAA AAGACTTTCC
```

```
   GLE  AAAAA0GTTTG GGCCGCCATT AAATT0AAAA0G GGGCCATGAA AATACTTTCC
JEZno3  AAAAA0GTYTG GGCCGCCATT AAATT0AAAA0G GGGCCATGAA AATACTTTCC
   KOR  AAAAA0GTCTG GGCCGCCATT AAATT0AAAA0G GGGCCATGAA AATACTTTCC
   KOY  AAAAA0GTCTG GGCCGCCATT AAATT0AAAA0G GGGCCATGAA AATACTTTCC
   LIK  AAAAA0GTC?G GGCCGCCATT AAATT0AAAA0G GGGCCATGAA AATACTTTCC
MARno4  AAAAA0GTCTG GGCCGCCATT AAATT0AAAA0G GGGCCATGAA AAGACTTTCC
   MAT  AAAAA0GTCTG GGCCGCCATT AAATT0AAAA0G GGGCCATGAA AAGACTTTCC
   MAX  AAAA??GT??G G?C?GCCATT AAATT?AAAA?G GGGCCAT?AA AA?A?TTT?C
   MEX  AAAAA0GTCTG GGCCGCCATT AAATT0AAAA0G GGGCCATGAA AAGACTTTCC
   MEY  AAAAA0GTCTG GGCCGCCATT AAATT0AAAA0G GGGCCATGAA AATACTTTCC
   MON  AAAAA0GTCTG GGCCGCCATT AAATT0AAAA0G GGGCCATGAA AATACTTTCC
   MOR  AAAA.1GTATG GGCCGCCATT AAATT0AAAA0G GGGCCATGAA AAGACTTTCC
   OBO  AAAAA0GTCTG GGCCGCCATT AAATT0AAAA0G GGGCCATGAA AATACTTTCC
   OMO  AAAAA0GTCTG GGCCGCCATT AAATT0AAAA0G GGGCCATGAA AAgACTTTCC
   ORI  AAAAA0GTCTG GGCCGCCATT AAATT0AAAA0G GGGCCATGAA AATA?TTTCC
PINBANK  AAAAA0GTTTG GGCCGCCATT AAATT0AAAA0G GGGCCATGAA AATAGTTTCC
PINKOR  AAAA??GT??G G?C?GCCAT? ??????AAAA?G GGGCCAT?AA AA?A?TTT?C
PUNno1  AAAAA0GTCTG GGCCGCCATT AAATT0AAAA0G GGGCCATGAA AAGACTTTCC
   PUR  AAAAA0GTCTG GGCCGCCATT AAATT0AAAA0G GGGCCATTAA AAGACTTTCC
   RET  AAAAA0GTTTG GGCCGCCATT AAATT0AAAA0G GGGCCATGAA AATACTTTCC
   RUB  AAAAA0GTCTG GGCCGCCATT AAATT0AAAA0G GGGCCATGAA AAGACTTTCC
   SCH  AAAAA0GTCTG GGCCGCCATT AAATT0AAAA0G GGGCCATGAA AAGACTTTCC
SITno1  AAAAA0GTCTG GGCCGCCATT AAATT0AAAA0G GGGCCATGAA AAGACTTTCC
   SMI  AAAAA0GTCTG GGCCG..... .....1....1G GGGCCATGAA AAGACTTTCC
   SPI  AAAAA0GTCTG GGCCGCCAT. .....1AAAA0G GGGCCATGAA AAGACTTTCC
   TOR  AAAAA0GTCTG GGCCGCCATT AAATT0AAAA0G GGGCCATGAA AATACTTTCC
   WIL  AAAAA0GTC?G GGCCGCCATT AAATT0AAAA0G GGGCCATGAA AATACTTTCC

      [2101                                                  2150]
ABInol  AGGGCCATGA AAAGCCTTGT ATTTATTCCC CTCAAAGGGG CTTTGTAAGG
   ALC  AGGGCCATGA AAAGCCCTGT ATTTATTCCC CTCAAAGGGG CTTTGTAAGG
   ASP  AGGGCCATGA AAAGCCCTGT ATTTATTCCC CTCAAAGGGG CTTTGTAAGG
   BRA  AGGGCCATGA AAAGCCCTGT ATTTATTCCC CTCAAAGGGG CTTTGTAAGG
   BRE  AGGGCCATGA AAAGCCCTGT ATTTATTCCC CTCAAAGGGG CTTTGTAAGG
   CHI  AGGGCCATGA AAAGCCCTGT ATTTATTCCC CTCAAAGGGG CTTTGTAAGG
   CRA  AGGGCCATGA AAAGCCCTGT ATTTATTCCC CTCAAAGGGG CTTTGTAAGG
ENGno1  AGGGCCATGA AAAGCCCTGT ATTTATTCCC CTCAAAGGGG CTTTGTAAGG
   FAR  AGGGCCATGA AAAGCCCTGT ATTTATTCCC CTCAAAGGGG CTTTGTAAGG
GLAno1  AGGGCCATGA AAAGCCCTGT ATTTATTCCC CTCAAAGGGG CTTTGTAAGG
   GLE  AGGGCCATGA AAAGCCCTGT ATTTATTCCC CTCAAAGGGG CTTTGTAAGG
JEZno3  AGGGCCATGA AAAGCCCTGT ATTTATTCCC CTCAAAGGGG CTTTGTAAGG
   KOR  AGGGCCATGA AAAGCCCTgT ATTTATTCCC CTCAAAGGGG CTTTGTAAGG
   KOY  AGGGCCATGA AAAGCCCTAT ATTTATTCCC CTCAAAGGGG CTTTGTAAGG
   LIK  AGGGCCATGA AAA?CCCTGT ATTTATTCCC CTCAAAGGGG CTTTGTAAGG
MARno4  AGGGCCATGA AAAGCCCTGT ATTTATTCCC CTCAAAGGGG CTTTGTAAGG
   MAT  AGGGCCATGA AAAGCCCTGT ATTTATTCCC CTCAAAGGGG CTTTGTAAGG
   MAX  AGGGCCATGA AAA?CCCTGT ATTTATTCCC CTCAAAGGGG CTTTGTAAGG
   MEX  AGGGCCATGA AAAGCCCTGT ATTTATTCCC CTCAAAGGGG CTTTGTAAGG
   MEY  AGGGCCATGA AAAGCCCTAT ATTTATTCCC CTCAAAGGGG CTTTGTAAGG
   MON  AGGGCCATGA AAAGCCCT?T ATTTATTCCC CTCAAAGGGG CTTTGTAAGG
   MOR  AGGGCCATGA AAAGCCCTGT ATTTATTCCC CTCAAAGGGG CTTTGTAAGG
   OBO  AGGGCCATGA AAAGCCCTGT ATTTATTCCC CTCAAAGGGG CTTTGTAAGG
   OMO  AGGGCCATGA AAAgCCCTGT ATTTATTCCC CTCAAAGGGG CTTTGTAAGG
   ORI  AGGGCCATGA AAAGCCCTGT ATTTATTCCC CTCAAAGGGG CTTTGTAAGG
PINBANK  AGGGCCATGA AAATCCCTGT ATTTATTCCC CTCAAAGGGG CTTTGTAAGG
PINKOR  AGGGCCATGA AAA?CC?T?T ATTTATTCCC CT?AAAGGGG ?TTTGTAAGG
PUNno1  AGGGCCATGA AAAGCCCTGT ATTTATTCCC CTCAAAGGGG CTTTGTAAGG
   PUR  AGGGCCATGA AAAGCCCTGT ATTTATTCCC CTAAAAGGGG CTTTGTAAGG
   RET  AGGGCCATGA AAAGCCCTGT ATTTATTCCC CTCAAAGGGG CTTTGTAAGG
   RUB  AGGGCCATGA AAAGCCCTGT ATTTATTCCC CTCAAAGGGG CTTTGTAAGG
   SCH  AGGGCCATGA AAAGCCCTGT ATTTATTCCC CTCAAAGGGG CTTTGTAAGG
SITno1  AGGGCCATGA AAAGCCCTGT ATTTATTCCC CTCAAAGGGG CTTTGTAAGG
   SMI  AGGGCCATGA AAAGCCCTGT ATTTATTCCC CTCAAAGGGG ?TTTGTAAGG
   SPI  AGGGCCATGA AAAGCCCTGT ATTTATTCCC CTCAAAGGGG CTTTGTAAGG
   TOR  AGGGCCATGA AAAGCCCTGT ATTTATTCCC CTCAAAGGGG ?TTTGTAAGG
   WIL  AGGGCCATGA AAA?CCCTGT ATTTATTCCC CTCAAAGGGG CTTTGTAAGG

      [2151   +1                  +1                      2200]
ABInol  ACTTGTCT0GG AGATATCTGT CTCACTAA0GM GGTAAYCTCG CTACCTCACC
   ALC  ACTTGTCT0GG AGATATCTGT CTCACTAA0GC GGTAACCTCG CTACCTCACC
   ASP  ACTTGTCT0GG AGATATCTGT CTCACTAA0G? GGTAACCTCG CTACCTCACC
   BRA  ACT.....1GG AGATATCTGT CTCACTAA0GC GGTAACCTCG CTACCTCACC
   BRE  ACTTGTCT0GG AGATATCTGT CTCACTAA0GC GGTAACCTCG CTACCTCACC
   CHI  ACTTGTCT0GG AGATATCTGT CTCACTAA0GC GGTAACCTCG CTACCTCACC
   CRA  ACTTGTCT0GG AGATATCTGT CTCACTAA0G? GGTAACCTCG CTACCTCACC
ENGno1  ACTTGTCT0GG AGATATCTGT CTCACTAA0GC GGTAACCTCG CTACCTCACC
   FAR  ACT.....1GG AGATATCTGT CTCACTAA0GC GGTAACCTCG CTACCTCACC
GLAno1  ACTTGTCT0GG AGATATCTGT CTCACTAA0GC GGTAACCTCG CTACCTCACC
   GLE  ACTTGTCT0GG AGATATCTGT CTCACTAA0G? GGTAACCTCG CTACCTCACC
JEZno3  ACTTGTCT0GG AGATATCTGT CTCACTAA0GT GGTAACCTCG CTACCTCACC
   KOR  ACTTGTCT0GG AGATATCTGT CTCACTAA0Ga gGTAACCTCG CTACCTCACC
   KOY  ACTTGTCT0GG AGATATCTGT CTCACTAA0G? ?GTAACCTCG CTACCTCACC
   LIK  ACTTGTCT0GG AGATATCTGT CTCACTAA0GC GGTAACCTCG CTACCTCACC
MARno4  ACTTGTCT0GG AGATATCTGT CTCACTAA0GC GGTAACCTCG CTACCTCACC
```

```
   MAT  ACTTGTCT0GG AGATATCTGT CTCACTAA0GC GGTAACCTCG CTACCTCACC
   MAX  AcTTGTCT0GG AGATATCTGT CTCACTAA0GC GGTAACCTCg CTACCTCACC
   MEX  ACTTGTCT0GG AGATATCTGT CTCACTAA0GC GGTAACCTCG CTACCTCACC
   MEY  ACTTGTCT0GG AGATATCTGT CTCACTAA0G? ?GTAACCTCG CTACCTCACC
   MON  ACTTGTCT0GG AGATATCTGT CTCACTAA0G? ?GTAACCTCG CTACCTCACC
   MOR  ACTTGTCT0GG ATATATCTGT CTCACTAA0GC GGTAACCTCG CTACCTCACC
   OBO  ACTTGTCT0GG AGATATCTGT CTCACTAA0GC GGTAACCTCG CTACCTCACC
   OMO  ACTTGTCT0GG AGATATCTGT CTCACTAA0GC GGTAACCT?G CTACCTCaCC
   ORI  ACTTGTCT0GG AGATATCTGT CTCACTAA0GC GGTAACCTCG CTACCTCaCC
PINBANK ACTTGTCT0GG AGATATCTGT CTCACTAA0GC GGTAACCTTT CTACCTCGCC
 PINKOR A?????????G A?ATAT?TGT CTCACTAA0G? ?GTAA?CT?? CTACCTC?CC
 PUNno1 ACTTGTCT0GG AGATATCTGT CTCACTAA0GC GGTAACCTCG CTACCTCACC
   PUR  ACTTGTCT0GG AGATATCTGT CTCACTAA0GC GGTAACCTCG CTACCTCACC
   RET  ACTTGTCT0GG AGATATCTGT CTCACTAA0G? GGTAACCTCG CTACCTCACC
   RUB  ACTTGTCT0GG AGATATCTGT CTCACTAA0GC GGTAACCTCG CTACCTCACC
   SCH  ACTTGTCT0GG AGATATATGT CTCACTAA0GC GGTAACCTCG CTACCTCACC
 SITno1 ACTTGTCT0GG AGATATCTGT CTCACTAA0GC GGTAACCTCG CTACCTCACC
   SMI  ACT.....1GG AGATATCTGT CTCACTAA0GC GGTAACCTCG CTACCTCACC
   SPI  ACT.....1GG AGATATCTGT CTCACTAA0GC GGTAACCTCG CTACCTCACC
   TOR  ACTTGTCT0GG AGATATCTGT CTCACTAA0G? GGTAACCTCG CTACCTCACC
   WIL  AATTGTCT0GG AGATATCTGT CTCACTAA0GC GGTAACCTCA CTACCTCACC

        [2201              +1                                2250]
 ABIno1 CCCTAAYGAC TRCTCTATCT0 TAGTGTTYRT TCGCRYGAYA AGGATACAAT
   ALC  CCCTAACGAC TGCTCTATCT0 TAGTGTTCAT TCGCGTGACA AGGATACAAT
   ASP  CCCTAACGAC TGCTCTATCT0 TAGTGTT??T TCG??TGACA AGGATACAAT
   BRA  CCCTAACGAC TGCTCT....1 TAGTGTTCGT TCGCGCGACA AGGATACAAT
   BRE  CCCTAACGAC TGCTCTATCT0 TAGTGTTCGT TCGCGCGACA AGGATACAAT
   CHI  CCCTAACGAC TGCTCTATCT0 TAGTGTTCGT TCGCGCGACA AGGATACAAT
   CRA  CCCTAACGAC TGCTCTATCT0 TAGTGTTTAT TCGCATGACA AGGATACAAT
 ENGno1 CCCTAACGAC TGCTCTATCT0 TAGTGTTCGT TCGCGCGACA AGGATACAAT
   FAR  CCCTAACGAC TGCTCT....1 TAGTGTTCGT TCGCGCGACA AGGATACAAT
 GLAno1 CCCTAACGAC TGCTCTATCT0 TAGTGTTCGT TCGCGCGACA AGGATACAAT
   GLE  CCCTAACGAC TGCTCTATCT0 TAGTGTTTAT TCGCATGACA AGGATACAAT
 JEZno3 CCCTAACGAC TGCTCTATCT0 TAGTGTTCGT TCGCGCGACA AGGATACAAT
   KOR  CCCTAAcGAc TGCTCTATCT0 TAGTGTTtaT TCgCaTGACA AGGATACAAT
   KOY  CCCTAATGAA TGCTCTATCT0 TAGTGTTCGT TCACGTGACA AGGATACAAT
   LIK  CCCTAACGAC TGCTCTATCT0 TAGTGTTC?T TCGCGCGACA AGGATACAAT
 MARno4 CCCTAACGAC TGCTCTATCT0 TAGTGTTCGT TCGCGCGACA AGGATACAAT
   MAT  CCCTAACGAC TGCTCTATCT0 TAGTGTTCGT TCGCGCGACA AGGATACAAT
   MAX  CCCTAACGAC TGCTCTATCT0 TAGTGTTCgT TCGCGCgACA AGGATACAAT
   MEX  CCCTAACGAC TGCTCTATCT0 TAGTGTTCGT TCGCGCGACA AGGATACAAT
   MEY  CCCTAATGAA TGCTCTATCT0 TAGTGTTCGT TCACGTGACA AGGATACAAT
   MON  CCCTAA?GA? TGCTCTATCT0 TAGTGTT??T TC?C?TGACA AGGATACAAT
   MOR  CCCTAACGAC TGCTCT....1 TAGTGTTCGT TCGCGCGACA AGGATACAAT
   OBO  CCCTAACGAC TGCTCTATCT0 TAGTGTTCGT TCGCGCGACA AGGATACAAT
   OMO  CCCTAACGAC TGCTCTCTCT0 TAGTGTTCgT TCGCGCGACA AGGATACAAT
   ORI  CCCTAACGAC TGCTCTATCT0 TAGTGTTCgT TCGCGCGACA AGGATACAAT
PINBANK CCCTAACGAC TGCTCTATCT0 TAGTGTTCAT TCGCGCGACA AGGATACA?T
 PINKOR CCCTAA?GA? T?CTC?????? TAGTGTT??T TC?????A?A AGGATACA?T
 PUNno1 CCCTAACGAC TGCTCTATCT0 TAGTGTTCGT TCGCGCGACA AGGATACAAT
   PUR  CCCTAACGAC TGCTCT....1 TAGTGTTCGT TCGCGCGACA AGGATACAAT
   RET  CCCTAACGAC TGCTCTATCT0 TAGTGTTTAT TCGCATGACA AGGATACAAT
   RUB  CCCTAACGAC TGCTCTATCT0 TAGTGTTCGT TCGCGCGACA AGGATACAAT
   SCH  CCCTAACGAC TGCTCT....1 TAGTGTTCGT TCGCGCGACA AGGATACAAT
 SITno1 CCCTAACGAC TGCTCTATCT0 TAGTGTTCGT TCGCGCGACA AGGATACAAT
   SMI  CCCTAACGAC TGCTCT....1 TAGTGTTCGT TCGCGCGACA AGGATACAAT
   SPI  CCCTAACGAC TGCTCT....1 TAGTGTTCGT TCGCGCGACA AGGATACAAT
   TOR  CCCTAACGAC TGCTCTATCT0 TAGTGTTTAT TCGC?TGACA AGGATACAAT
   WIL  CCCTAACGAC TGCTCTATCT0 TAGTGTTCAT TC?C?TGACA AGGATACAAT

        [2251                                                2300]
 ABIno1 AAAGAGMATT TATGYGCTCG ACCGACGTGA AATTCATAGT CTTAGGAGAG
   ALC  AAAGAGCATT TATGCGCTTG ACCGACGTGA AATTCATAGT CTTAGGAGAG
   ASP  AAAGAGCATT TATGTGCTCG ACCGACGTGA AATTCATAGT CTTAGGAGAG
   BRA  AAAGAGCATT TATGCGCTCG ACCGACGTGA AATTCATAGT CTTAGGAGAG
   BRE  AAAGAGCATT TATGCGCTCG ACCGACGTGA AATTCATAGT CTTAGGAGAG
   CHI  AAAGAGCATT TATGCGCTCG ACCGACGTGA AATTCATAGT CTTAGGAGAG
   CRA  AAAGAGCATT TATGTGCTCG ACCGACGTGA AATTCATAGT CTTAGGAGAG
 ENGno1 AAAGAGCATT TATGCGCTCG ACCGACGTGA AATTCATAGT CTTAGGAGAG
   FAR  AAAGAGCATT TATGCGCTCG ACCGACGTGA AATTCATAGT CTTAGGAGAG
 GLAno1 AAAGAGCATT TATGCGCTCG ACCGACGTGA AATTCATAGT CTTAGGAGAG
   GLE  AAAGAGCATT TATGTGCTCG ACCGACGTGA AATTCATAGT CTTAGGAGAG
 JEZno3 AAAGAGCATT TATGCGCTCG ACCGACGTGA AATTCATAGT CTTAGGAGAG
   KOR  AAAGAGCATT TATGtGCTCG ACCGACGTGA AATTCATAGT CTTAGGAGAG
   KOY  AAAGAGCATT TATGTGCTCG ACCGACGTGA AATTCATAGT CTTAGGAGAG
   LIK  AAAGAGCATT TATGCGCTCG ACCGACGTGA AATTCATAGT CTTAGGAGAG
 MARno4 AAAGAGCATT TATGCGCTCG ACCGACGTGA AATTCATAGT CTTAGGAGAG
   MAT  AAAGAGCATT TATGCGCTCG ACCGACGTGA AATTCATAGT CTTAGGAGAG
   MAX  AAAGAGCATT TATGCGCTCG ACcGACGTGA AATTCATAGT CTTAGGAGAG
   MEX  AAAGAGCATT TATGCGCTCG ACCGACGTGA AATTCATAGT CTTAGGAGAG
   MEY  AAAGAGCATT TATGCGCTCG ACCGACGTGA AATTCATAGT CTTAGGAGAG
   MON  AAAGAGCATT TATG?GCTCG ACCGACGTGA AATTCATAGT CTTAGGAGAG
   MOR  AAAGAGCATT TATGCGCTCG ACCGACGTGA AATTCATAGT CTTAGGAGAG
```

```
    OBO  AAAGAGCATT TATGCGCTCG ACCGACGTGA AATTCATAGT CTTAGGAGAG
    OMO  AAAGAGcATT TATGcgCTcG ACcGACGTGA AATTCATAGT CTTAGGAGAG
    ORI  AAAGAGCATT TATGCGCTCG ACCGACGTGA AATTCATAGT CTTAGGAGAG
PINBANK  AAAGAGCATT TATGCGCTCG ACCGACGTGA AATTCATAGT CTTAGGAGAG
 PINKOR  AAAGAG?ATT TATG??CT?G AC?GACGTGA AATTCATAGT CTTAGGAGAG
 PUNno1  AAAGAGCATT TATGCGCTCG ACCGACGTGA AATTCATAGT CTTAGGAGAG
    PUR  AAAGAGCATT TATGCGCTCG ACCGACGTGA AATTCATAGT CTTAGGAGAG
    RET  AAAGAGCATT TATGTGCTCG ACCGACGTGA AATTCATAGT CTTAGGAGAG
    RUB  AAAGAGCATT TATGCGCTCG ACCGACGTGA AATTCATAGT CTTAGGAGAG
    SCH  AAAGAGCATT TATGCGCTCG ACCGACGTGA AATTCATAGT CTTAGGAGAG
 SITno1  AAAGAGCATT TATGCGCTCG ACCGACGTGA AATTCATAGT CTTAGGAGAG
    SMI  AAAGAGCATT TATGCGCTCG ACCGACGTGA AATTCATAGT CTTAGGAGAG
    SPI  AAAGAGCATT TATGCGCTCG ACCGACGTGA AATTCATAGT CTTAGGAGAG
    TOR  AAAGAGCATT TATGTGCTCG ACCGACGTGA AATTCATAGT CTTAGGAGAG
    WIL  AAAGAGCATT TATGC?CTCG ACCGACGTGA AATTCATAGT CTTAGGAGAG

         [2301                                                2350]
 ABIno1  GACACAACAC CTTAGAACTT GAACATTTTC ACATYCTTTC ARGACCTTTT
    ALC  GACACAACAC CTTAGAACTT GAACATTTTC ACATCCTTTC AGGACCTTTT
    ASP  GACACAACAC CTTAGAACTT GAACATTTTC ACATTCTTTC AAGACCTTTT
    BRA  GACACAACAC CTTAGAACTT GAACATTTTC ACATCCTTTC AGGACCTTTT
    BRE  GACACAACAC CTTAGAACTT GAACATTTTC ACATCCTTTC AGGACCTTTT
    CHI  GACACAACAC CTTAGAACTT GAACATTTTC ACATCCTTTC AGGACCTTTT
    CRA  GACACAACAC CTTAGAACTT GAACATTTTC ACATTCTTTC AAGACCTTTT
 ENGno1  GACACAACAC CTTAGAACTT GAACATTTTC ACATCCTTTC AGGACCTTTT
    FAR  GACACAACAC CTTAGAACTT GAACATTTTC ACATCCTTTC AGGACCTTTT
 GLAno1  GACACAACAC CTTAGAACTT GAACATTTTC ACATCCTTTC AGGACCTTTT
    GLE  GACACAACAC CTTAGAACTT GAACATTTTC ACATTCTTTC AAGACCTTTT
 JEZno3  GACACAACAC CTTAGAACTT GAACATTTTC ACATCCTTTC AGGACCTTTT
    KOR  GACACAACAC CTTAGAACTT GAACATTTTC ACATTCTTTC AAGACCTTTT
    KOY  GACACAACAC CTTAGAACTT GAAAATTTTC ACATCCTTTC AGGACCTTTT
    LIK  GACACAACAC CTTAGAACTT GAACATTTTC ACATCCTTTC AGGACCTTTT
 MARno4  GACACAACAC CTTAGAACTT GAACATTTTC ACATCCTTTC AGGACCTTTT
    MAT  GACACAACAC CTTAGAACTT GAACATTTTC ACATCCTTTC AGGACCTTTT
    MAX  GACACAACAC CTTAGAACTT GAACATTTTC ACATCCTTTC AGGACCTTTT
    MEX  GACACAACAC CTTAGAACTT GAACATTTTC ACATCCTTTC AGGACCTTTT
    MEY  GACACAACAC CTTAGAACTT GAAAATTTTC ACATCCTTTC AGGACCTTTT
    MON  GACACAACAC CTTAGAACTT GAACATTTTC ACATTCTTTC AAGACCTTTT
    MOR  GACACAACAC CTTAGAACTT GAACATTTTC ACATCCTTTC AGGACCTTTT
    OBO  GACACAACAC CTTAGAACTT GAACATTTTC ACATCCTTTC AGGACCTTTT
    OMO  GACACAACAC CTTAGAACTT gAAcATTTTC ACATcCTTTC AgGACCTTTT
    ORI  GACACAACAC CTTAGAACTT GAACATTTTC ACATCCTTTC AGGACCTTTT
PINBANK  GACACAACAC CTTAGAACTT GAACATTTTC ACATCCTTTC AGGACCTTTT
 PINKOR  GACACAACAC CTTAGAACTT ?AA?ATTTTC ACAT?CTTTC A?GACCTTTT
 PUNno1  GACACAACAC CTTAGAACTT GAACATTTTC ACATCCTTTC AGGACCTTTT
    PUR  GACACAACAC CTTAGAACTT GAACATTTTC ACATCCTTTC AGGACCTTTT
    RET  GACACAACAC CTTAGAACTT GAACATTTTC ACATTCTTTC AAGACCTTTT
    RUB  GACACAACAC CTTAGAACTT GAACATTTTC ACATCCTTTC AGGACCTTTT
    SCH  GACACAACAC CTTAGAACTT GAACATTTTC ACATCCTTTC AGGACCTTTT
 SITno1  GACACAACAC CTTAGAACTT GAACATTTTC ACATCCTTTC AGGACCTTTT
    SMI  GACACAACAC CTTAGAACTT TAACATTTTC ACATCCTTTC AGGACCTTTT
    SPI  GACACAACAC CTTAGAACTT GAACATTTTC ACATCCTTTC AGGACCTTTT
    TOR  GACACAACAC CTTAGAACTT GAACATTTTC ACATTCTTTC AAGACCTTTT
    WIL  GACACAACAC CTTAGAACTT GAACATTTTC ACATCCTTTC AGGACCTTTT

         [2351                +1                          2400]
 ABIno1  AATGAAGGGA AGCCCYGTAG C0GAAATCCCT TTCCTTTGCG TAGCTCAGGT
    ALC  AATGAAGGGA AGCCCCGTAG C0GAAATCCAT TTCCTTTGCG TAGCTCAGGT
    ASP  AATGAAGGGA AGCCCCGTAG C0GAAATCCCT TTCCTTTGCG TAGCTCAGGT
    BRA  AATGAAGGGA AGCCCCGTAG C0GAAATCCCT TTCCTTTGCG TAGCTCAGGT
    BRE  AATGAAGGGA AGCCCC.... .1GAAATCCCT TTCCTTTGCG TAGCTCAGGT
    CHI  AATGAAGGGA AGCCCCGTAG C0GAAATCCCT TTCCTTTGCG TAGCTCAGGT
    CRA  AATGAAGGGA AGCCCCGTAG C0GAAATCCCT TTCCTTTGCG TAGCTCAGGT
 ENGno1  AATGAAGGGA AGCCCCGTAG C0GAAATCCCT TTCCTTTGCG TAGCTCAGGT
    FAR  AATGAAGGGA AGCCCCGTAG C0GAAATCCCT TTCCTTTGCG TAGCTCAGGT
 GLAno1  AATGAAGGGA AGCCCCGTAG C0GAAATCCCT TTCCTTTGCG TAGCTCAGGT
    GLE  AATGAAGGGA AGCCCCGTAG C0GAAATCCCT TTCCTTTGCG TAGCTCAGGT
 JEZno3  AATGAAGGGA AGC?CCGTAG ?0GAAATCCCT TTCCTTTGC? TAGCTCAGGT
    KOR  AATGAAGGGA AGCCCCGTAG C0GAAATCCCT TTCCTTTGCG TAGCTCAGGT
    KOY  AATGAAGGGA AGCCCCGTAG C0GAAATCCCT TTCCTTTGTG TAGCTTAGGT
    LIK  AATGAAGGGA AGCCCCGTAG C0GAAATCCCT TTCCTTTGC? TAGCTCAGGT
 MARno4  AATGAAGGGA AGCCCCGTAG C0GAAATCCCT TTCCTTTGCG TAGCTCAGGT
    MAT  AATGAAGGGA AGCCCCGTAG C0GAAATCCCT TTCCTTTGCG TAGCTCAGGT
    MAX  AATGAAGGGA AGCCTCGTAG A0GAAATCCCT TTCCTTTGCG TAGCTCAAGT
    MEX  AATGAAGGGA AGCCCCGTAG C0GAAATCCCT TTCCTTTGCG TAGCTCAGGT
    MEY  AATGAAGGGA AGCCCCGTAG C0GAAATCCCT TTCCTTTGTG TAGCTTAGGT
    MON  AATGAAGGGA AGCCCCGTAG C0GAAATCCCT TTCCTTTGCG TAGCTCAGGT
    MOR  AATGAAGGGA AGCCCCGTAG C0GAAATCCCT TTCCTTTGCG TAGCTCAGGT
    OBO  AATGAAGGGA AGCCCC.... .1GAAATCCCT TTCCTTTGCG TAGCTCAGGT
    OMO  AATGAAGGGA AGCcccgtag c0GAAATCCcT TTCCTTTGcg TAGCTcAgGT
    ORI  AATGAAGGGA AGCCCC.... .1GAAATCCCT TTCCTTTGCG TAGCTCAGGT
PINBANK  AATGAAGGGA AGCCCCGTAG C0GAAATCCCT TTCCTTTGCG TAGCTCAGGT
 PINKOR  AATGAAGGGA AGC??????? ??GAAATCC?T TTCCTTTG?? TAGCT?A?GT
 PUNno1  AATGAAGGGA AGCCCCGTAG C0GAAATCCCT TTCCTTTGCG TAGCTCAGGT
```

```
   PUR  AATGAAGGGA AGCCCCGTAG C0GAAATCCCT TTCCTTTGCG TAGCTCAGGT
   RET  AATGAAGGGA AGCCCCGTAG C0GAAATCCCT TTCCTTTGCG TAGCTCAGGT
   RUB  AATGAAGGGA AGCCCCGTAG C0GAAATCCCT TTCCTTTGCG TAGCTCAGGT
   SCH  AATGAAGGGA AGCCCCGTAG C0GAAATCCCT TTCCTTTGCG TAGCTCAGGT
SITno1  AATGAAGGGA AGCCCCGTAG C0GAAATCCCT TTCCTTTGCG TAGCTCAGGT
   SMI  AATGAAGGGA AGCCCCGTAG C0GAAATCCCT TTCCTTTGCG TAGCTCAGGT
   SPI  AATGAAGGGA AGCCCCGTAG C0GAAATCCCT TTCCTTTGCG TAGCTCAGGT
   TOR  AATGAAGGGA AGCCCCGTAG C0GAAATCCCT TTCCTTTGCG TAGCTCAGGT
   WIL  AATGAAGGGA AGCCCCGTAG C0GAAATCCCT TTCCTTTGC? TAGCTCAGGT

        [2401                                              2450]
ABIno1  CAGGTTATCC TTACCTTCGC RTTCAGGCTT TGAGG..... ..........
   ALC  CAGGTTATCC TTACCTTCGT GTTCAGGCTT TGAGG..... ..........
   ASP  CAGGTTATCC TTACCTTTGC GTTCAGGCTT TGAGG..... ..........
   BRA  CAGGTTATCC TTACCTTCGC GTTCAGGCTT TGAGGAACGT AGTCGCTCGA
   BRE  CAGGTTATCC TTACCTTCGC GTTCAGGCTT TGAGG..... ..........
   CHI  CAGGTTATCC TTACCTTCGC GTTCAGGCTT TGAGG..... ..........
   CRA  CAGGTTATCC TTACCTTTGC GTTCAGGCTT TGAGG..... ..........
ENGno1  CAGGTTATCC TTACCTTCGC GTTCAGGCTT TGAGG..... ..........
   FAR  CAGGTTATCC TTACCTTCGC GTTCAGGCTT TGAGGAACGT AGTCGCTCGA
GLAno1  CAGGTTATCC TTACCTTCGC GTTCAGGCTT TGAGG..... ..........
   GLE  CAGGTTATCC TTACCTTTGC GTTCAGGCTT TGAGG..... ..........
JEZno3  CAGGTTATCC TTACCTTCGC GTT?A?GCTT TGAGG..... ..........
   KOR  CAGGTTATCC TTACCTTTGC GTTCAGGCTT TGAGG..... ..........
   KOY  CAGATTATCC TTACCTTCGT GTTCAGGCTT TGAGG..... ..........
   LIK  CAGGTTATCC TTACCTTCGC GTTCAGGCTT TGAGG..... ..........
MARno4  CAGGTTATCC TTACCTTCGC GTTCAGGCTT TGAGG..... ..........
   MAT  CAGGTTATCC TTACCTTCGC GTTCAGGCTT TGAGG..... ..........
   MAX  CAGGTTATCC TTACCTTCGC GTTCAGGCTT TGAGG..... ..........
   MEX  CAGGTTATCC TTACCTTCGC GTTCAGGCTT TGAGG..... ..........
   MEY  CAGATTATCC TTACCTTCGT GTTCAGGCTT TGAGG..... ..........
   MON  CAGGTTATCC TTACCTTTGC GTTCAGGCTT TGAGG..... ..........
   MOR  CAGGTTATCC TTACCTTCGC GTTCAGGCTT TGAGG..... ..........
   OBO  CAGGTTATCC TTACCTTCGC GTTCAGGCTT TGAGG..... ..........
   OMO  CAGgTTATCC TTACCTTcGc gTTcAgGCTT TGAGG..... ..........
   ORI  CAGGTTATCC TTACCTTCGC GTTCAGGCTT TGAGG..... ..........
PINBANK  CAGGTTATCC TTACCTTCGC GTTCAAGCTT TGAGG..... ..........
PINKOR  CAG?TTATCC TTACCTT?G? ?TT?A?GCTT TGAGG????? ??????????
PUNno1  CAGGTTATCC TTACCTTCGC GTTCAGGCTT TGAGG..... ..........
   PUR  CAGGTTATCC TTACCTTCGC GTTCAGGCTT TGAGG..... ..........
   RET  CAGGTTATCC TTACCTTTGC GTTCAGGCTT TGAGG..... ..........
   RUB  CAGGTTATCC TTACCTTCGC GTTCAGGCTT TGAGG..... ..........
   SCH  CAGGTTATCC TTACCTTCGC GTTCAGGCTT TGAGG..... ..........
SITno1  CAGGTTATCC TTACCTTCGC GTTCAGGCTT TGAGG..... ..........
   SMI  CAGGTTATCC TTACCTTCGC GTTCAGGCTT TGAGGAACGT AGTCGCTCGA
   SPI  CAGGTTATCC TTACCTTCGC GTTCAGGCTT TGAGGAACGT AGTCGCTCGA
   TOR  CAGGTTATCC TTACCTTTGC GTTCAGGCTT TGAGG..... ..........
   WIL  CAG?TTATCC TTACCTTCGC GTTCAGGCTT TGAGG..... ..........

        [2451         +1                    +1             2500]
ABIno1  .......... ....0GGAAKA AATATA.... ..0GGGACCTT CCCTTCGCTA
   ALC  .......... ....0GGAATA AATATA.... ..0GGGACCTT CCCTTCGCTA
   ASP  .......... ....0GGAATA AATATA.... ..0GGGACCTT CCCTTCGCTA
   BRA  CCATAAGGGA GAGG1GGAATA AATATAAATA TA1GGGACCTT CCCTTCGCTA
   BRE  .......... ....0GGAATA AATATA.... ..0GGGACCTT CCCTTCGCTA
   CHI  .......... ....0GGAAGA AATATA.... ..0GGGACCTT CCCTTCGCTA
   CRA  .......... ....0GGAATA AATATA.... ..0GGGACCTT CCCTTCGCTA
ENGno1  .......... ....0GGAATA AATATA.... ..0GGGACCTT CCCTTCGCTA
   FAR  CCATAAGGGA GAGG1GGAATA AATATAAATA TA1GGGACCTT CCCTTCGCTA
GLAno1  .......... ....0GGAAGA AATATA.... ..0GGGACCTT CCCTTCGCTA
   GLE  .......... ....0GGAATA AATATA.... ..0GGGACCTT CCCTTCGCTA
JEZno3  .......... ....0GGAATA AATATA.... ..0GGGACCTT CCCTTCG?TA
   KOR  .......... ....0GGAATA AATATA.... ..0GGGACCTT CCCTTCGCTA
   KOY  .......... ....0GGAATA AATATA.... ..0GGGACCTT CCATTCGCTA
   LIK  .......... ....0GGAATA AATATA.... ..0GGGACCTT CCCTTCG?TA
MARno4  .......... ....0GGAAGA AATATA.... ..0GGGACCTT CCCTTCGCTA
   MAT  .......... ....0GGAAGA AATATA.... ..0GGGACCTT CCCTTCGCTA
   MAX  .......... ....0GGAATA AATATA.... ..0GGGACCTT CCCGTCGCTA
   MEX  .......... ....0GGAAGA AATATA.... ..0GGGACCTT CCCTTCGCTA
   MEY  .......... ....0GGAATA AATATA.... ..0GGGACCTT CCATTCGCTA
   MON  .......... ....0GGAATA AATATA.... ..0GGGACCTT CCCTTCGCTA
   MOR  .......... ....0GGAATA AATATA.... ..0GGGACCTT CCCTTCGCTA
   OBO  .......... ....0GGAATA AATATA.... ..0GGGACCTT CCCTTCGCTA
   OMO  .......... ....0GGAAtA AATATA.... ..0GGGACCTT CCctTcGcTA
   ORI  .......... ....0GGAATA AATATA.... ..0GGGACCTT CCCTTCGCTA
PINBANK  .......... ....0GGAATA AATATA.... ..0GGGACCTT CCCTT?G?TA
PINKOR  ?????????? ?????GGAA?A AATATA???? ???GGGACCTT CC??T?G?TA
PUNno1  .......... ....0GGAAGA AATATA.... ..0GGGACCTT CCCTTCGCTA
   PUR  .......... ....0GGAATA AATATAAATA TA1GGGACCTT CCCTTCGCTA
   RET  .......... ....0GGAATA AATATA.... ..0GGGACCTT CCCTTCGCTA
   RUB  .......... ....0GGAAGA AATATA.... ..0GGGACCTT CCCTTCGCTA
   SCH  .......... ....0GGAATA AATATAAATA TA1GGGACCTT CCCTTCGCTA
SITno1  .......... ....0GGAATA AATATA.... ..0GGGACCTT CCCTTCGCTA
   SMI  CCATAAGGGA GAGG1GGAATA AATATAAATA TA1GGGACCTT CCCTTCGCTA
```

```
   SPI  CCATAAGGGA GAGG1GGAATA AATATAAATA TA1GGGACCTT CCCTTCGCTA
   TOR  .......... ....0GGAATA AATATA.... ..0GGGACCTT CCCTTCGCTA
   WIL  .......... ....0GGAATA AATATA.... ..0GGGACCTT CCCTTCGCTA

        [2501                                                2550]
ABInol  TGCGTAAGAC CTAGACCACC GTAGCGAAGG GAAGGCAAGA TMTATAGTTA
   ALC  TGCATAATAC CTAGACCACC GTAGCGAAGG GAAGGCAAGA TCTATAGTTA
   ASP  TGCGTAAGAC CTAGACCACC GTAGCGAAGG GAAGGCAAGA TCTATAGTTA
   BRA  TGCGTAAGAC CTAGACCACC GTAGCGAAGG GAAGGCAAGA TCTATAGTTA
   BRE  TGCGTAAGAC CTAGACCACC GTAGCGAAGG GAAGGCAAGA TCTATAGTTA
   CHI  TGCGTAAGAC CTAGACCACC GTAGCGAAGG GAAGGCAAGA TCTATAGTTA
   CRA  TGCGTAAGAC CTAGACCACC GTAGCGAAGG GAAGGCAAGA TCTATAGTTA
ENGnol  TGCGTAAGAC CTAGACCACC GTAGCGAAGG GAAGGCAAGA TCTATAGTTA
   FAR  TGCGTAAGAC CTAGACCACC GTAGCGAAGG GAAGGCAAGA TCTATAGTTA
GLAnol  TGCGTAAGAC CTAGACCACC GTAGCGAAGG GAAGGCAAGA TCTATAGTTA
   GLE  TGCGTAAGAC CTAGACCACC GTAGCGAAGG GAAGGCAAGA TCTATAGTTA
JEZno3  TGCGTAAGAC CTAGACCAC? GTAGCGAAGG GAAGGCAAGA TCTATAGTTA
   KOR  TGCGTAAGAC CTAGACCACT GTAGCGAAGG GAAGGCAAGA TCTATAGTTA
   KOY  TGCGTAAGAC CTAGACCACC GTAGCGAAGG GAAGGCAAGA TCTATAGTTA
   LIK  TGCGTAAGAC CTAGACCACC GTAGCGAAGG GAAGGCAAGA TCTATAGTTA
MARno4  TGCGTAAGAC CTAGACCACC GTAGCGAAGG GAAGGCAAGA TCTATAGTTA
   MAT  TGCGTAAGAC CTAGACCACC GTAGCGAAGG GAAGGCAAGA TCTATAGTTA
   MAX  TGCGTAAGAC CTAGACCACC GTAGCAAAGG GAAGGCAAGA TCTATAGTTA
   MEX  TGCGTAAGAC CTAGACCACC GTAGCGAAGG GAAGGCAAGA TCTATAGTTA
   MEY  TGCGTAAGAC CTAGACCACC GTAGCGAAGG GAAGGCAAGA TCTATAGTTA
   MON  TGCGTAAGAC CTAGACCACC GTAGCGAAGG GAAGGCAAGA TCTATAGTTA
   MOR  TGCGTAAGAC CTAGACCACC GTAGCGAAGG GAAGGCAAGA TCTATAGTTA
   OBO  TGCGTAAGAC CTAGACCACC GTAGCGAAGG GAAGGCAAGA TCTATAGTTA
   OMO  TGCgTAAgAC CTAGAcCACc GTAGCgAAGG GAAGGCAAgA TcTATAGTTA
   ORI  TGCGTAAGAC CTAGACCACC GTAGCGAAGG GAAGGCAAGA TCTATAGTTA
PINBANK  TGCGTAAGAC CTAGACCACT GTAGCGAAGG GAAGGCAAGA TCTATAGTTA
PINKOR  TGC?TAA?AC CTAGA?CAC? GTAGC?AAGG GAAGGCAA?A T?TATAGTTA
PUNnol  TGCGTAAGAC CTAGACCACC GTAGCGAAGG GAAGGCAAGA TCTATAGTTA
   PUR  TGCGTAAGAC CTAGACCACC GTAGCGAAGG GAAGGCAATA TCTATAGTTA
   RET  TGCGTAAGAC CTAGACCACC GTAGCGAAGG GAAGGCAAGA TCTATAGTTA
   RUB  TGCGTAAGAC CTAGACCACC GTAGCGAAGG GAAGGCAAGA TCTATAGTTA
   SCH  TGCGTAAGAC CTAGACCACC GTAGCGAAGG GAAGGCAATA TCTATAGTTA
SITnol  TGCGTAAGAC CTAGACCACC GTAGCGAAGG GAAGGCAAGA TCTATAGTTA
   SMI  TGCGTAAGAC CTAGACCACC GTAGCGAAGG GAAGGCAAGA TCTATAGTTA
   SPI  TGCGTAAGAC CTAGACCACC GTAGCGAAGG GAAGGCAAGA TCTATAGTTA
   TOR  TGCGTAAGAC CTAGACCACC GTAGCGAAGG GAAGGCAAGA TCTATAGTTA
   WIL  TGCGTAAGAC CTAGA?CACC GTAGCAAAGG GAAGGCAAGA TCTATAGTTA

        [2551                                                2600]
ABInol  CAKAATTACR RCACTATGGA TCRATCTAYG TTGCCGTAAC AAATTCACAA
   ALC  CAGAATTACG ?CACTATGGA TCGATCTACG TTGTCGTAAC AAATTCACAA
   ASP  CATAATTACA GCACTATGGA TCGATCTATG TTGCCGTAAC AAATTCACAA
   BRA  CAGAATTACG GCACTATGGA TCGATCTACG TTGCCGTAAC AAATTCACAA
   BRE  CAGAATTACG GCACTATGGA TCGATCTACG TTGCCGTAAC AAATTCACAA
   CHI  CAGAATTACG GCACTATGGA TCGATCTACG TTGCCGTAAC AAATTCACAA
   CRA  CATAATTACA ?CACTATGGA TCGATCTATG TTGCCGTAAC AAATTCACAA
ENGnol  CAGAATTACG GCACTATGGA TCGATCTACG TTGCCGTAAC AAATTCACAA
   FAR  CAGAATTACG GCACTATGGA TCGATCTACG TTGCCGTAAC AAATTCACAA
GLAnol  CAGAATTACG GCACTATGGA TCGATCTACG TTGCCGTAAC AAATTCACAA
   GLE  CATAATTACA GCACTATGGA TCGATCTATG TTGCCGTAAC AAATTCACAA
JEZno3  CAGAATTACG ?CACTATGGA TC?ATCTAC? TTGCCGTAAC AAATTCACAA
   KOR  CATAATTACA GCACTATGGA TCGATCTATG TTGCCGTAAC AAATTCACAA
   KOY  CAGAATTACA GCACTATGGA TCGATCTACG TTGTCATAAA AAATTCACAA
   LIK  CAGAATTACG GCACTATGGA TCGATCTACG TTGCCGTAAC AAATTCACAA
MARno4  CAGAATTACG GCACTATGGA TCGATCTACG TTGCCGTAAC AAATTCACAA
   MAT  CAGAATTACG GCACTATGGA TCGATCTACG TTGCCGTAAC AAATTCACAA
   MAX  CAGAATTACG ACACTATGGA TCGATCTACG TTGTCGTAAC AAATTCACAA
   MEX  CAGAATTACG GCACTATGGA TCGATCTACG TTGCCGTAAC AAATTCACAA
   MEY  CAGAATTACA GCACTATGGA TCGATCTACG TTGTCATAAA AAATTCACAA
   MON  CATAATTACA GCACTATGGA TCGATCTATG TTGCCGTAAC AAATTCACAA
   MOR  CAGAATTACG GCACTATGGA TCGATCTACG TTGCCGTAAC AAATTCACAA
   OBO  CAGAATTACG GCACTATGGA TCGATCTACG TTGCCGTAAC AAATTCACAA
   OMO  CAgAATTACg gCACTATGGA TCgATcTAcg TTGcCgTAAc AAATTCACAA
   ORI  CAGAATTACG GCACTATGGA TCGATCTACG TTGCCGTAAC AAATTCACAA
PINBANK  CAGAATTACG ACACTATGGA TC?ATCTAC? TTGCCGTAAC AAATTCACAA
PINKOR  CA?AATTAC? ?CACTATGGA TC?AT?TA?? TTG?C?TAA? AAATTCACAA
PUNnol  CAGAATTACG GCACTATGGA TCGATCTACG TTGCCGTAAC AAATTCACAA
   PUR  CAGAATTACG GCACTATGGA TCGATCTACG TTGCCGTAAC AAATTCACAA
   RET  CATAATTACA GCACTATGGA TCGAT?TATG TTGCCGTAAC AAATTCACAA
   RUB  CAGAATTACG GCACTATGGA TCGATCTACG TTGCCGTAAC AAATTCACAA
   SCH  CAGAATTACG GCACTATGGA TCGATCTACG TTGCCGTAAC AAATTCACAA
SITnol  CAGAATTACG GCACTATGGA TCGATCTACG TTGCCGTAAC AAATTCACAA
   SMI  CAGAATTACG GCACTATGGA TCGATCTACG TTGCCGTAAC AAATTCACAA
   SPI  CAGAATTACG GCACTATGGA TCGATCTACG TTGCCGTAAC AAATTCACAA
   TOR  CATAATTACA GCACTATGGA TCGATCTATG TTGCCGTAAC AAATTCACAA
   WIL  CAGAATTACG ?CACTATGGA TCGATCTACG TTGTCGTAAC AAATTCACAA

        [2601           +1                                   2650]
ABInol  ATGAATTCTA G......0ATA GTGCCCAATT TAGAACCTTG CGGAGGGACC
```

```
   ALC  ATGAATTCTA G......0ATA GTGTCCAATT TAGAACCTTG CGGAGGGACC
   ASP  ATGAATTCTA G......0ATA GTGCCCAATT TAGAACCTTG CGGAGGGACC
   BRA  ATGAATTCTA GTTCTAG1ATA GTGCCCAATT TAGAACCTTG CGGAGGGACC
   BRE  ATGAATTCTA G......0ATA GTGCCCAATT TAGAACCTTG CGGAGGGACC
   CHI  ATGAATTCTA G......0ATA GTGCCCAATT TAGAACCTTG CGGAGGGACC
   CRA  ATGAATTCTA G......0ATA GTGCCCAATT TAGAACCTTG CGGAGGGACC
ENGno1  ATGAATTCTA G......0ATA GTGCCCAATT TAGAACCTTG CGGAGGGACC
   FAR  ATGAATTCTA G......0ATA GTGCCCAATT TAGAACCTTG CGGAGGGACC
GLAno1  ATGAATTCTA G......0ATA GTGCCCAATT TAGAACCTTG CGGAGGGACC
   GLE  ATGAATTCTA G......0ATA GTGCCCAATT TAGAACCTTG CGGAGGGACC
JEZno3  ATGAATT?TA G......0ATA GTGCCCAATT TAGAACCTTG ?G?AGGGACC
   KOR  ATGAATTCTA G......0ATA GTGCCCAATT TAGAACCTTG CGGAGGGACC
   KOY  ATGAATTCTA G......0ATA GTGCCCAATT TAGAACCTTG CGGAGGGACC
   LIK  ATGAATTCTA G......0ATA GTGCCCAATT TAGAACCTTG CGGAGGGACC
MARno4  ATGAATTCTA G......0ATA GTGCCCAATT TAGAACCTTG CGGAGGGACC
   MAT  ATGAATTCTA G......0ATA GTGCCCAATT TAGAACCTTG CGGAGGGACC
   MAX  ATAAATTCTA G......0ATA GTGCCCAATT TAGAACCTTG CGGAGGGACC
   MEX  ATGAATTCTA G......0ATA GTGCCCAATT TAGAACCTTG CGGAGGGACC
   MEY  ATGAATTCTA G......0ATA GTGCCCAATT TAGAACCTTG CGGAGGGACC
   MON  ATGAATTCTA G......0ATA GTGCCCAATT TAGAACCTTG CGGAGGGACC
   MOR  ATGAATTCTA G......0ATA GTGCCCAATT TAGAACCTTG CGGAGGGACC
   OBO  ATGAATTCTA G......0ATA GTGCCCAATT TAGAACCTTG CGGAGGGACC
   OMO  ATgAATTcTA G......0ATA GTGcCCAATT TAGAACCTTG cGgAGGGACC
   ORI  ATGAATTCTA G......0ATA GTGCCCAATT TAGAACCTTG CGGAGGGACC
PINBANK  ATGAATTCTA G......0ATA GTGCCCAATT TAGAACCTTG CGTAGGGACC
PINKOR  AT?AATT?TA G???????ATA GTG?CCAATT TAGAACCTTG ?G?AGGGACC
PUNno1  ATGAATTCTA G......0ATA GTGCCCAATT TAGAACCTTG CGGAGGGACC
   PUR  ATGAATTCTA G......0ATA GTGCCCAATT TAGAACCTTG CGGAGGGACC
   RET  ATGAATTCTA G......0ATA GTGCCCAATT TAGAACCTTG CGGAGGGACC
   RUB  ATGAATTCTA G......0ATA GTGCCCAATT TAGAACCTTG CGGAGGGACC
   SCH  ATGAATTCTA G......0ATA GTGCCCAATT TAGAACCTTG CGGAGGGACC
SITno1  ATGAATTCTA G......0ATA GTGCCCAATT TAGAACCTTG CGGAGGGACC
   SMI  ATGAATTCTA G......0ATA GTGCCCAATT TAGAACCTTG CGGAGGGACC
   SPI  ATGAATTCTA G......0ATA GTGCCCAATT TAGAACCTTG CGGAGGGACC
   TOR  ATGAATTCTA G......0ATA GTGCCCAATT TAGAACCTTG CGGAGGGACC
   WIL  ATGAATTCTA G......0ATA GTGCCCAATT TAGAACCTTG TGGAGGGACC

        [2651                                                2700]
ABIno1  GTAGCTCCRG GAATATATGG GGATCTGMCG AACAACCAAC CCCCTAACCC
   ALC  GTAGCTCCGG GAATATATGG GGATCTGCCG AACAACCAAC CCCCTAACCC
   ASP  GTAGCTCCGG GAATATATGG GGATCTGACG AACAACCAAC CCCCTAACCC
   BRA  GTAGCTCCGG GAATATATGG GGATCTGCCG AACAACCAAC CCCCTAACCC
   BRE  GTAGCTCCGG GAATATATGG GGATCTGCCG AACAACCAAC CCCCTAACCC
   CHI  GTAGCTCCGG GAATATATGG GGATCTGCCG AACAACCAAC CCCCTAACCC
   CRA  GTAGCTCCGG GAATATATGG GGATCTGACG AACAACCAAC CCCCTAACCC
ENGno1  GTAGCTCCGG GAATATATGG GGATCTGCCG AACAACCAAC CCCCTAACCC
   FAR  GTAGCTCCGG GAATATATGG GGATCTGCCG AACAACCAAC CCCCTAACCC
GLAno1  GTAGCTCCGG GAATATATGG GGATCTGCCG AACAACCAAC CCCCTAACCC
   GLE  GTAGCTCCGG GAATATATGG GGATCTGACG AACAACCAAC CCCCTAACCC
JEZno3  GTAG?TCCGG GAATATATGG GGATCTGCCG AACAACCAAC CCCCTAACCC
   KOR  GTAGCTCCGG GAATATATGG GGATCTGACG AACAACCAAC CCCCTAACCC
   KOY  GTAGCTCCGG GAATATATGG GGATTTTCCA AACAACCAAC CCCCTAACCC
   LIK  GTAGCTCCGG GAATATATGG GGATCTGCCG AACAACCAAC CCCCTAACCC
MARno4  GTAGCTCCGG GAATATATGG GGATCTGCCG AACAACCAAC CCCCTAACCC
   MAT  GTAGCTCCGG GAATATATGG GGATCTGCCG AACAACCAAC CCCCTAACCC
   MAX  GTAGCTCCGG AAATATATGG GGATCTGCCG AACAACCAAC CCCCTAACCC
   MEX  GTAGCTCCGG GAATATATGG GGATCTGCCG AACAACCAAC CCCCTAACCC
   MEY  GTAGCTCCGG GAATATATGG GGATTTTCCA AACAACCAAC CCCCTAACCC
   MON  GTAGCTCCGG GAATATATGG GGATCTGACG AACAACCAAC CCCCTAACCC
   MOR  GTAGCTCCGG GAATATATGG GGATCTGCCG AACAACCAAC CCCCTAACCC
   OBO  GTAGCTCCGG GAATATATGG GGATCTGCCG AACAACCAAC CCCCTAACCC
   OMO  GTAGcTCCgG gAATATATGG GGATcTgcCg AACAACCAAC CCCCTAACCC
   ORI  GTAGCTCCGG GAATATATGG GGATCTGCCG AACAACCAAC CCCCTAACCC
PINBANK  GTAGGTCCGG GAATATATGG GGATCTGCCG AACAACCAAC CCCCTAACCC
PINKOR  GTAG?TCC?G ?AATATATGG GGAT?T??C? AACAACCAAC CCCCTAACCC
PUNno1  GTAGCTCCGG GAATATATGG GGATCTGCCG AACAACCAAC CCCCTAACCC
   PUR  GTAGCTCCGG GAATATATGG GGATCTGCCG AACAACCAAC CCCCTAACCC
   RET  GTAGCTCCGG GAATATATGG GGATCTGACG AACAACCAAC CCCCTAACCC
   RUB  GTAGCTCCGG GAATATATGG GGATCTGCCG AACAACCAAC CCCCTAACCC
   SCH  GTAGCTCCGG GAATATATGG GGATCTGCCG AACAACCAAC CCCCTAACCC
SITno1  GTAGCTCCGG GAATATATGG GGATCTGCCG AACAACCAAC CCCCTAACCC
   SMI  GTAGCTCCGG GAATATATGG GGATCTGCCG AACAACCAAC CCCCTAACCC
   SPI  GTAGCTCCGG GAATATATGG GGATCTGCCG AACAACCAAC CCCCTAACCC
   TOR  GTAGCTCCGG GAATATATGG GGATCTGACG AACAACCAAC CCCCTAACCC
   WIL  GTAGCTCCGG GAATATATGG GGATCTGCCG AACAACCAAC CCCCTAACCC

        [2701        +1                                      2750]
ABIno1  T......... ..0TTTGCAGC AAATGCTTTT ATTCACAGTG AAAGGGGGMA
   ALC  T......... ..0TTTGCAGC AAATGCTTTT ATTCACAGTG AAAGGGGGCA
   ASP  T......... ..0TTTGCAGC AAATGCTTTT ATTCACAGTG AAAGGGGGAA
   BRA  T......... ..0TTTGCAGC AAATGCTTTT ATTCACAGTG AAAGGGGGCA
   BRE  T......... ..0TTTGCAGC AAATGCTTTT ATTCACAGTG AAAGGGGGCA
   CHI  T......... ..0TTTGCAGC AAATGCTTTT ATTCACAGTG AAAGGGGGCA
   CRA  T......... ..0TTTGCAGC AAATGCTTTT ATTCACAGTG AAAGGGGGAA
```

```
ENGno1  T......... ..0TTTGCAGC AAATGCTTTT ATTCACAGTG AAAGGGGGCA
   FAR  T......... ..0TTTGCAGC AAATGCTTTT ATTCACAGTG AAAGGGGGCA
GLAno1  T......... ..0TTTGCAGC AAATGCTTTT ATTCACAGTG AAAGGGGGCA
   GLE  T......... ..0TTTGCAGC AAATGCTTTT ATTCACAGTG AAAGGGGGAA
JEZno3  T......... ..0TTTGCAGC AAATGCTTTT ATTCACAGTG AAAGGGGG?A
   KOR  T......... ..0TTTGCAGC AAATGCTTTT ATTCACAGTG AAAGGGGGAA
   KOY  T......... ..0TTTGCAGC AAATGCTTTT ATTCACAGTG AAAGGGGGAA
   LIK  T......... ..0TTTGCAGC AAATGCTTTT ATTCACAGTG AAAGGGGGCA
MARno4  T......... ..0TTTGCAGC AAATGCTTTT ATTCACAGTG AAAGGGGGCA
   MAT  T......... ..0TTTGCAGC AAATGCTTTT ATTCACAGTG AAAGGGGGCA
   MAX  T......... ..0TTTGCAGC AAATGCTTGT ATTCACAGTG AAAGGGGGCA
   MEX  T......... ..0TTTGCAGC AAATGCTTTT ATTCACAGTG AAAGGGGGCA
   MEY  T......... ..0TTTGCAGC AAATGCTTTT ATTCACAGTG AAAGGGGGAA
   MON  T......... ..0TTTGCAGC AAATGCTTTT ATTCACAGTG AAAGGGGGAA
   MOR  T......... ..0TTTGCAGC AAATGCTTTT ATTCACAGTG AAAGGGGGCA
   OBO  T......... ..0TTTGCAGC AAATGCTTTT ATTCACAGTG AAAGGGGGCA
   OMO  T......... ..0TTTGCAGc AAATGcTTtT ATTCACAGTG AAAGGGGGcA
   ORI  T......... ..0TTTGCAGC AAATGCTTTT ATTCACAGTG AAAGGGGGCA
PINBANK T......... ..0TTTGCAGA AAATGCTTTT ATTCACAGTG AAAGGGGGAA
PINKOR  T......... ..?TTTGCAG? AAATG?TT?T ATTCACAGTG AAAGGGGG?A
PUNno1  T......... ..0TTTGCAGC AAATGCTTTT ATTCACAGTG AAAGGGGGCA
   PUR  T......... ..0TTTGCAGC AAATGCTTTT ATTCACAGTG AAAGGGGGCA
   RET  T......... ..0TTTGCAGC AAATGCTTTT ATTCACAGTG AAAGGGGGAA
   RUB  T......... ..0TTTGCAGC AAATGCTTTT ATTCACAGTG AAAGGGGGCA
   SCH  TCCCCTAACC CT1TTTGCAGC AAATGCTTTT ATTCACAGTG AAAGGGGGCA
SITno1  T......... ..0TTTGCAGC AAATGCTTTT ATTCACAGTG AAAGGGGGCA
   SMI  T......... ..0TTTGCAGC AAATGCTTTT ATTCACAGTG AAAGGGGGCA
   SPI  T......... ..0TTTGCAGC AAATGCTTTT ATTCACAGTG AAAGGGGGCA
   TOR  T......... ..0TTTGCAGC AAATGCTTTT ATTCACAGTG AAAGGGGGAA
   WIL  T......... ..0TTTGCAGC AAATGCTTGT ATTCACAGTG AAAGGGGGCA

          [2751                                                2800]
ABIno1  ACCAACCACA TGAGTAGTTC GCTCCAAGAC CCAACCAATT GCTTGCCTCC
   ALC  ACCAACCACA TGAGTAGTTC GCTCCAAGAC CCAACCAATT GCTTGCCTCC
   ASP  ACCAACCACA TGAGTAGTTC GCTCCAAGAC CCAACCAATT GCTTGCCTCC
   BRA  ACCAACCACA TGAGTAGTTC GCTCCAAGAC CCAACCAATT GCTTGCCTCC
   BRE  ACCAACCACA TGAGTAGTTC GCTCCAAGAC CCAACCAATT GCTTGCCTCC
   CHI  ACCAACCACA TGAGTAGTTC GCTCCAAGAC CCAACCAATT GCTTGCCTCC
   CRA  ACCAACCACA TGAGTAGTTC GCTCCAAGAC CCAACCAATT ?CTTGCCTCC
ENGno1  ACCAACCACA TGAGTAGTTC GCTCCAAGAC CCAACCAATT GCTTGCCTCC
   FAR  ACCAACCACA TGAGTAGTTC GCTCCAAGAC CCAACCAATT GCTTGCCTCC
GLAno1  ACCAACCACA TGAGTAGTTC GCTCCAAGAC CCAACCAATT GCTTGCCTCC
   GLE  ACCAACCACA TGAGTAGTTC GCTCCAAGAC CCAACCAATT ?CTTGCCTCC
JEZno3  ACCAACCACA TGAGTAGTTC GCTCCAAGAC CCAACCAATT GCTTGCCTCC
   KOR  ACCAACCACA TGAGTAGTTC GCTCCAAGAC CCAACCAATT ?CTTGCCTCC
   KOY  ACCAACCACA TGAGTAGTTC GCTCCAAGAC CCAACCAATT GCTTGCCTCC
   LIK  ACCAACCACA TGAGTAGTTC GCTCCAAGAC CCAACCAATT GCTTGCCTCC
MARno4  ACCAACCACA TGAGTAGTTC GCTCCAAGAC CCAACCAATT GCTTGCCTCC
   MAT  ACCAACCACA TGAGTAGTTC GCTCCAAGAC CCAACCAATT GCTTGCCTCC
   MAX  ACCAACCACA TGAGTAGTTC GCTCCAAGAC CCAACCAATT GCTTGCCTCC
   MEX  ACCAACCACA TGAGTAGTTC GCTCCAAGAC CCAACCAATT GCTTGCCTCC
   MEY  ACCAACCACA TGAGTAGTTC GCTCCAAGAC CCAACCAATT GCTTGCCTCC
   MON  ACCAACCACA TGAGTAGTTC GCTCCAAGAC CCAACCAATT ?CTTGCCTCC
   MOR  ACCAACCACA TGAGTAGTTC GCTCCAAGAC CCAACCAATT GCTTGCCTCC
   OBO  ACCAACCACA TGAGTAGTTC GCTCCAAGAC CCAACCAATT GCTTGCCTCC
   OMO  ACCAACCACA TGAGTAGTTC gCTCCAAGAc CCAACCAATT gCTTGCCTCC
   ORI  ACCAACCACA TGAGTAGTTC GCTCCAAGAC CCAACCAATT GCTTGCCTCC
PINBANK ACCAACCACA TGAGTAGTTC GCTCCAAGAC CCAACCAATT ?CTTGCCTCC
PINKOR  ACCAACCACA TGAGTAGTTC ?CTCCAAGA? CCAACCAATT ?CTTGCCTCC
PUNno1  ACCAACCACA TGAGTAGTTC GCTCCAAGAC CCAACCAATT GCTTGCCTCC
   PUR  ACCAACCACA TGAGTAGTTC GCTCCAAGAC CCAACCAATT GCTTGCCTCC
   RET  ACCAACCACA TGAGTAGTTC GCTCCAAGAC CCAACCAATT ?CTTGCCTCC
   RUB  ACCAACCACA TGAGTAGTTC GCTCCAAGAC CCAACCAATT GCTTGCCTCC
   SCH  ACCAACCACA TGAGTAGTTC GCTCCAAGAC CCAACCAATT GCTTGCCTCC
SITno1  ACCAACCACA TGAGTAGTTC GCTCCAAGAC CCAACCAATT GCTTGCCTCC
   SMI  ACCAACCACA TGAGTAGTTC GCTCCAAGAC CCAACCAATT GCTTGCCTCC
   SPI  ACCAACCACA TGAGTAGTTC GCTCCAAGAC CCAACCAATT GCTTGCCTCC
   TOR  ACCAACCACA TGAGTAGTTC GCTCCAAGAC CCAACCAATT ?CTTGCCTCC
   WIL  ACCAACCACA TGAGTAGTTC GCTCCAAGAC CCAACCAATT GCTTGCCTCC

          [2801                                                2850]
ABIno1  CTCCGGATGC CGAGATAGAG CAACACAACC TATACGCCCC AACACCGATA
   ALC  CTCCGGATGC CGAGATAGAG CAACACAACC TATACGCCCC AACACCGATA
   ASP  CTCCGGATGC CGAGATAGAG CAACACAACC TATACGCCCC AACACCGATA
   BRA  CTCCGGATGC CGAGATAGAG CAACACAACC TATACGCCCC AACACCGATA
   BRE  CTCCGGATGC CGAGATAGAG CAACACAACC TATACGCCCC AACACCGATA
   CHI  CTCCGGATGC CGAGATAGAG CAACACAACC TATACGCCCC AACACCGATA
   CRA  CTCCGGATGC CGAGATAGAG CAACACAACC TATACGCCCC AACACCGATA
ENGno1  CTCCGGATGC CGAGATAGAG CAACACAACC TATACGCCCC AACACCGATA
   FAR  CTCCGGATGC CGAGATAGAG CAACACAACC TATACGCCCC AACACCGATA
GLAno1  CTCCGGATGC CGAGATAGAG CAACACAACC TATACGCCCC AACACCGATA
   GLE  CTCCGGATGC CGAGATAGAG CAACACAACC TATACGCCCC AACACCGATA
JEZno3  CTCCGGATGC CGAGATAGAG CAACACAACC TATACGCCCC AACACCGATA
   KOR  CTCCGGATGC CGAGATAGAG CAACACAACC TATACGCCCC AACACCGATA
```

```
   KOY  CTCCGGATGC CGAGATAGAG CAACACAACC TATACGCCCC AACACCGATA
   LIK  CTCCGGATGC CGAGATAGAG CAACACAACC TATACGCCCC AACACCGATA
MARno4  CTCCGGATGC CGAGATAGAG CAACACAACC TATACGCCCC AACACCGATA
   MAT  CTCCGGATGC CGAGATAGAG CAACACAACC TATACGCCCC AACACCGATA
   MAX  CTYCGGATGC CGAGATAGAG CAACACAACC TATACGCCCC AACACCGATA
   MEX  CTCCGGATGC CGAGATAGAG CAACACAACC TATACGCCCC AACACCGATA
   MEY  CTCCGGATGC CGAGATAGAG CAACACAACC TATACGCCCC AACACCGATA
   MON  CTCCGGATGC CGAGATAGAG CAACACAACC TATACGCCCC AACACCGATA
   MOR  CTCCGGATGC CGAGATAGAG CAACACAACC TATACGCCCC AACACCGATA
   OBO  CTCCGGATGC CGAGATAGAG CAACACAACC TATACGCCCC AACACCGATA
   OMO  CTCCGGATGC CGAGATAGAG CAACACAACC TATACGCCCC AACACCGATA
   ORI  CTCCGGATGC CGAGATAGAG CAACACAACC TATACGCCCC AACACCGATA
PINBANK  CTCCG?ATGC CGAGATAGAG ?AACACAACC TATAT?CCCC AACA?CGATA
PINKOR  CTCCGGATGC CGAGATAGAG CAACACAACC TATATGCCCC AACACCGATA
PUNno1  CTCCGGATGC CGAGATAGAG CAACACAACC TATACGCCCC AACACCGATA
   PUR  CTCCGGATGC CGAGATAGAG CAACACAACC TATACGCCCC AACACCGATA
   RET  CTCCGGATGC CGAGATAGAG CAACACAACC TATACGCCCC AACACCGATA
   RUB  CTCCGGATGC CGAGATAGAG CAACACAACC TATACGCCCC AACACCGATA
   SCH  CTCCGGATGC CGAGATAGAG CAACACAACC TATACGCCCC AACACCGATA
SITno1  CTCCGGATGC CGAGATAGAG CAACACAACC TATACGCCCC AACACCGATA
   SMI  CTCCGGATGC CGAGATAGAG CAACACAACC TATACGCCCC AACACCGATA
   SPI  CTCCGGATGC CGAGATAGAG CAACACAACC TATACGCCCC AACACCGATA
   TOR  CTCCGGATGC CGAGATAGAG CAACACAACC TATACGCCCC AACACCGATA
   WIL  CT?CGGATGC CGAGATAGAG CAACACAACC TATACGCCCC AACACCGATA

        [2851                                                2900]
ABInol  CTTCCAACAA CCTGCTTAAG GGCAAAAAGC TGAAGCGAGC GCTTGGTTGG
   ALC  CTTCCAACAA CCTGCTTAAG GGCAAAAAGC TGAAGCGAGC GCTTGGTTGG
   ASP  CTTCCAACAA CCTGCTTAAG GGCAAAAAGC TGAAGCGAGC GCTTGGTTGG
   BRA  CTTCCAACAA CCTGCTTAAG GGCAAAAAGC TGAAGCGAGC GCTTGGTTGG
   BRE  CTTCCAACAA CCTGCTTAAG GGCAAAAAGC TGAAGCGAGC GCTTGGTTGG
   CHI  CTTCCAACAA CCTGCTTAAG GGCAAAAAGC TGAAGCGAGC GCTTGGTTGG
   CRA  CTTCCAACAA CCTGCTTAAG GGCAAAAAGC TGAAGCGAGC GCTTGGTTGG
ENGno1  CTTCCAACAA CCTGCTTAAG GGCAAAAAGC TGAAGCGAGC GCTTGGTTGG
   FAR  CTTCCAACAA CCTGCTTAAG GGCAAAAAGC TGAAGCGAGC GCTTGGTTGG
GLAno1  CTTCCAACAA CCTGCTTAAG GGCAAAAAGC TGAAGCGAGC GCTTGGTTGG
   GLE  CTTCCAACAA CCTGCTTAAG GGCAAAAAGC TGAAGCGAGC GCTTGGTTGG
JEZno3  CTTCCAACAA CCTGCTTAAG GGCAAAAAGC TGAAGCGAGC GCTTGGTTGG
   KOR  CTTCCAACAA CCTGCTTAAG GGCAAAAAGC TGAAGCGAGC GCTTGGTTGG
   KOY  CTTCCAACAA CCTGCTTAAG GGCAAAAAGC TGAAGCGAGC GCTTGGTTGG
   LIK  CTTCCAACAA CCTGCTTAAG GGCAAAAAGC TGAAGCGAGC GCTTGGTTGG
MARno4  CTTCCAACAA CCTGCTTAAG GGCAAAAAGC TGAAGCGAGC GCTTGGTTGG
   MAT  CTTCCAACAA CCTGCTTAAG GGCAAAAAGC TGAAGCGAGC GCTTGGTTGG
   MAX  CTTCCAACAA CCTGCTTAAG GGCAAAAAGC TGAAGCGAGC GCTTGGTTGG
   MEX  CTTCCAACAA CCTGCTTAAG GGCAAAAAGC TGAAGCGAGC GCTTGGTTGG
   MEY  CTTCCAACAA CCTGCTTAAG GGCAAAAAGC TGAAGCGAGC GCTTGGTTGG
   MON  CTTCCAACAA CCTGCTTAAG GGCAAAAAGC TGAAGCGAGC GCTTGGTTGG
   MOR  CTTCCAACAA CCTGCTTAAG GGCAAAAAGC TGAAGCGAGC GCTTGGTTGG
   OBO  CTTCCAACAA CCTGCTTAAG GGCAAAAAGC TGAAGCGAGC GCTTGGTTGG
   OMO  CTTCCAACAA CCTGCTTAAG GGCAAAAAGC TGAAGCGAGC GCTTGGTTGG
   ORI  CTTCCAACAA CCTGCTTAAG GGCAAAAAGC TGAAGCGAGC GCTTGGTTGG
PINBANK  CTTCCAACAA CCTGCTTAAG GG?AAAAAGC TGAAGCGAGC GCTTGGTCGG
PINKOR  CTTCCAACAA CCTGCTTAAG GGCAAAAAGC TGAAGCGAGC GCTTGGTCGG
PUNno1  CTTCCAACAA CCTGCTTAAG GGCAAAAAGC TGAAGCGAGC GCTTGGTTGG
   PUR  CTTCCAACAA CCTGCTTAAG GGCAAAAAGC TGAAGCGAGC GCTTGGTTGG
   RET  CTTCCAACAA CCTGCTTAAG GGCAAAAAGC TGAAGCGAGC GCTTGGTTGG
   RUB  CTTCCAACAA CCTGCTTAAG GGCAAAAAGC TGAAGCGAGC GCTTGGTTGG
   SCH  CTTCCAACAA CCTGCTTAAG GGCAAAAAGC TGAAGCGAGC GCTTGGTTGG
SITno1  CTTCCAACAA CCTGCTTAAG GGCAAAAAGC TGAAGCGAGC GCTTGGTTGG
   SMI  CTTCCAACAA CCTGCTTAAG GGCAAAAAGC TGAAGCGAGC GCTTGGTTGG
   SPI  CTTCCAACAA CCTGCTTAAG GGCAAAAAGC TGAAGCGAGC GCTTGGTTGG
   TOR  CTTCCAACAA CCTGCTTAAG GGCAAAAAGC TGAAGCGAGC GCTTGGTTGG
   WIL  CTTCCAACAA CCTGCTTAAG GGCAAAAAGC TGAAGCGAGC GCTTGGTTGG

        [2901                                                2950]
ABInol  TAGGCCTATG CACCCCGATC AGCTAGCTAG GGATGGATAT GGCGAATGCC
   ALC  TAGGCCTATG CACCCCGATC AGCTAGCTAG GGATGGATAT GGCGAATGCC
   ASP  TAGGCCTATG CACCCCGATC AGCTAGCTAG GGATGGATAT GGCGAATGCC
   BRA  TAGGCCTATG CACCCCGATC AGCTAGCTAG GGATGGATAT GGCGAATGCC
   BRE  TAGGCCTATG CACCCCGATC AGCTAGCTAG GGATGGATAT GGCGAATGCC
   CHI  TAGGCCTATG CACCCCGATC AGCTAGCTAG GGATGGATAT GGCGAATGCC
   CRA  TAGGCCTATG CACCCCGATC AGCTAGCTAG GGATGGATAT GGCGAATGCC
ENGno1  TAGGCCTATG CACCCCGATC AGCTAGCTAG GGATGGATAT GGCGAATGCC
   FAR  TAGGCCTATG CACCCCGATC AGCTAGCTAG GGATGGATAT GGCGAATGCC
GLAno1  TAGGCCTATG CACCCCGATC AGCTAGCTAG GGATGGATAT GGCGAATGCC
   GLE  TAGGCCTATG CACCCCGATC AGCTAGCTAG GGATGGATAT GGCGAATGCC
JEZno3  TAGGCCTATG CACCCCGATC AGCTAGCTAG GGATGGATAT GGCGAATGCC
   KOR  TAGGCCTATG CACCCCGATC AGCTAGCTAG GGATGGATAT GGCGAATGCC
   KOY  TAGGCCTATG CACCCCGATC AGCTAGCTAG GGATGGATAT GGCGAATGCC
   LIK  TAGGCCTATG CACCCCGATC AGCTAGCTAG GGATGGATAT GGCGAATGCC
MARno4  TAGGCCTATG CACCCCGATC AGCTAGCTAG GGATGGATAT GGCGAATGCC
   MAT  TAGGCCTATG CACCCCGATC AGCTAGCTAG GGATGGATAT GGCGAATGCC
   MAX  TAGGCCTATG CACCCCGATC AGCTAGCTAG GGATGGATAT GGCGAATGCC
   MEX  TAGGCCTATG CACCCCGATC AGCTAGCTAG GGATGGATAT GGCGAATGCC
```

```
   MEY  TAGGCCTATG CACCCCGATC AGCTAGCTAG GGATGGATAT GGCGAATGCC
   MON  TAGGCCTATG CACCCCGATC AGCTAGCTAG GGATGGATAT GGCGAATGCC
   MOR  TAGGCCTATG CACCCCGATC AGCTAGCTAG GGATGGATAT GGCGAATGCC
   OBO  TAGGCCTATG CACCCCGATC AGCTAGCTAG GGATGGATAT GGCGAATGCC
   OMO  TAGGCCTATG CACCCCGATC AGCTAGCTAG GGATGGATAT GGCGAATGCC
   ORI  TAGGCCTATG CACCCCGATC AGCTAGCTAG GGATGGATAT GGCGAATGCC
PINBANK TAGGCCTATG CACCCCGATC AGCTAGCTAG GGATGGATAT GGCGAATGCC
PINKOR  TAGGCCTATG CACCCCGATA AGCTAGCTAG GGATGGATAG GGCGAATGCC
PUNno1  TAGGCCTATG CACCCCGATC AGCTAGCTAG GGATGGATAT GGCGAATGCC
   PUR  TAGGCCTATG CACCCCGATC AGCTAGCTAG GGATGGATAT GGCGAATGCC
   RET  TAGGCCTATG CACCCCGATC AGCTAGCTAG GGATGGATAT GGCGAATGCC
   RUB  TAGGCCTATG CACCCCGATC AGCTAGCTAG GGATGGATAT GGCGAATGCC
   SCH  TAGGCCTATG CACCCCGATC AGCTAGCTAG GGATGGATAT GGCGAATGCC
SITno1  TAGGCCTATG CACCCCGATC AGCTAGCTAG GGATGGATAT GGCGAATGCC
   SMI  TAGGCCTATG CACCCCGATC AGCTAGCTAG GGATGGATAT GGCGAATGCC
   SPI  TAGGCCTATG CACCCCGATC AGCTAGCTAG GGATGGATAT GGCGAATGCC
   TOR  TAGGCCTATG CACCCCGATC AGCTAGCTAG GGATGGATAT GGCGAATGCC
   WIL  TAGGCCTATG CACCCCGATC AGCTAGCTAG GGATGGATAT GGCGAATGCC

        [2951                                                3000]
ABInol  TGCCATAGTC TCTGAACACG GTGAAGTTCC CCTCCATCTA TTTATTTAAT
   ALC  TGCCATAGTC TCTGAACACG GTGAAGTTCC CCTCCATCTA TTTATTTAAT
   ASP  TGCCATAGTC TCTGAACACG GTGAAGTTCC CCTCCATCTA TTTATTTAAT
   BRA  TGCCATAGTC TCTGAACACG GTGAAGTTCC CCTCCATCTA TTTATTTCAT
   BRE  TGCCATAGTC TCTGAACACG GTGAAGTTCC CCTCCATCTA TTTATTTCAT
   CHI  TGCCATAGTC TCTGAACACG GTGAAGTTCC CCTCCATCTA TTTATTTAAT
   CRA  TGCCATAGTC TCTGAACACG GTGAAGTTCC CCTCCATCTA TTTATTTAAT
ENGno1  TGCCATAGTC TCTGAACACG GTGAAGTTCC CCTCCATCTA TTTATTTAAT
   FAR  TGCCATAGTC TCTGAACACG GTGAAGTTCC CCTCCATCTA TTTATTTCAT
GLAno1  TGCCATAGTC TCTGAACACG GTGAAGTTCC CCTCCATCTA TTTATTTAAT
   GLE  TGCCATAGTC TCTGAACACG GTGAAGTTCC CCTCCATCTA TTTATTTAAT
JEZno3  TGCCATAGTC TCTGAACACG GTGAAGTTCC CCTCCATCTA TTTATTTAAT
   KOR  TGCCATAGTC TCTGAACACG GTGAAGTTCC CCTCCATCTA TTTATTTAAT
   KOY  TGCCATAGTC TCTGAACACG GTGAAGTTCC CCTCCATCTA TTTATTTAAT
   LIK  TGCCATAGTC TCTGAACACG GTGAAGTTCC CCTCCATCTA TTTATTTCAT
MARno4  TGCCATAGTC TCTGAACACG GTGAAGTTCC CCTCCATCTA TTTATTTAAT
   MAT  TGCCATAGTC TCTGAACACG GTGAAGTTCC CCTCCATCTA TTTATTTAAT
   MAX  TGCCATAGTC TCTGAACACG GTGAAGTTCC CCTCCATCTA TTTATTTAAT
   MEX  TGCCATAGTC TCTGAACACG GTGAAGTTCC CCTCCATCTA TTTATTTAAT
   MEY  TGCCATAGTC TCTGAACACG GTGAAGTTCC CCTCCATCTA TTTATTTAAT
   MON  TGCCATAGTC TCTGAACACG GTGAAGTTCC CCTCCATCTA TTTATTTAAT
   MOR  TGCCATAGTC TCTGAACACG GTGAAGTTCC CCTCCATCTA TTTATTTAAT
   OBO  TGCCATAGTC TCTGAACACG GTGAAGTTCC CCTCCATCTA TTTATTTAAT
   OMO  TGCCATAGTC TCTGAACACG GTGAAGTTCC CCTCCATCTA TTTATTTAAT
   ORI  TGCCATAGTC TCTGAACACG GTGAAGTTCC CCTCCATCTA TTTATTTAAT
PINBANK TGCCATAGTC TCTTAACACG GTGAAGTTCC CCTCCATCTA TTCATTTCAT
PINKOR  TGCCATAGTA TCTGAACACG GTGAAGTTCC CCTCCATCTA TTCATTTCAT
PUNno1  TGCCATAGTC TCTGAACACG GTGAAGTTCC CCTCCATCTA TTTATTTAAT
   PUR  TGCCATAGTC TCTGAACACG GTGAAGTTCC CCTCCATCTA TTTATTTCAT
   RET  TGCCATAGTC TCTGAACACG GTGAAGTTCC CCTCCATCTA TTTATTTAAT
   RUB  TGCCATAGTC TCTGAACACG GTGAAGTTCC CCTCCATCTA TTTATTTAAT
   SCH  TGCCATAGTC TCTGAACACG GTGAAGTTCC CCTCCATCTA TTTATTTCAT
SITno1  TGCCATAGTC TCTGAACACG GTGAAGTTCC CCTCCATCTA TTTATTTAAT
   SMI  TGCCATAGTC TCTGAACACG GTGAAGTTCC CCTCCATCTA TTTATTTCAT
   SPI  TGCCATAGTC TCTGAACACG GTGAAGTTCC CCTCCATCTA TTTATTTCAT
   TOR  TGCCATAGTC TCTGAACACG GTGAAGTTCC CCTCCATCTA TTTATTTAAT
   WIL  TGCCATAGTC TCTGAACACG GTGAAGTTCC CCTCCATCTA TTTATTTAAT

        [3001                                                3050]
ABInol  TCAAGGGGCC GGGGGCGCGC CCTTGGAAAG CAAAGCGCGG ACGAGCCACA
   ALC  TCAAGGGGCC GGGGGCGCGC CCTTGGAAAG CAAAGCGCGG ACGAGCCACA
   ASP  TCAAGGGGCC GGGGGCGCGC CCTTGGAAAG CAAAGCGCGG ACGAGCCACA
   BRA  TCAAGGGGCC GGGGGCGCGC CCTTGGAAAG CAAAGCGCGG ACGAGCCACA
   BRE  TCAAGGGGCC GGGGGCGCGC CCTTGGAAAG CAAAGCGCGG ACGAGCCACA
   CHI  TCAAGGGGCC GGGGGCGCGC CCTTGGAAAG CAAAGCGCGG ACGAGCCACA
   CRA  TCAAGGGGCC GGGGGCGCGC CCTTGGAAAG CAAAGCGCGG ACGAGCCACA
ENGno1  TCAAGGGGCC GGGGGCGCGC CCTTGGAAAG CAAAGCGCGG ACGAGCCACA
   FAR  TCAAGGGGCC GGGGGCGCGC CCTTGGAAAG CAAAGCGCGG ACGAGCCACA
GLAno1  TCAAGGGGCC GGGGGCGCGC CCTTGGAAAG CAAAGCGCGG ACGAGCCACA
   GLE  TCAAGGGGCC GGGGGCGCGC CCTTGGAAAG CAAAGCGCGG ACGAGCCACA
JEZno3  TCAAGGGGCC GGGGGCGCGC CCTTGGAAAG CAAAGCGCGG ACGAGCCACA
   KOR  TCAAGGGGCC GGGGGCGCGC CCTTGGAAAG CAAAGCGCGG ACGAGCCACA
   KOY  TCAAGGGGCC GGGGGCGCGC CCTTGGAAAG CAAAGCGCGG ACGAGCCACA
   LIK  TCAAGGGGCC GGGGGCGCGC CCTTGGAAAG CAAAGCGCGG ACGAGCCACA
MARno4  TCAAGGGGCC GGGGGCGCGC CCTTGGAAAG CAAAGCGCGG ACGAGCCACA
   MAT  TCAAGGGGCC GGGGGCGCGC CCTTGGAAAG CAAAGCGCGG ACGAGCCACA
   MAX  TCAAGGGGCC GGGGGCGCGC CCTTGGAAAG CAAAGCGCGG ACGAGCCACA
   MEX  TCAAGGGGCC GGGGGCGCGC CCTTGGAAAG CAAAGCGCGG ACGAGCCACA
   MEY  TCAAGGGGCC GGGGGCGCGC CCTTGGAAAG CAAAGCGCGG ACGAGCCACA
   MON  TCAAGGGGCC GGGGGCGCGC CCTTGGAAAG CAAAGCGCGG ACGAGCCACA
   MOR  TCAAGGGGCC GGGGGCGCGC CCTTGGAAAG CAAAGCGCGG ACGAGCCACA
   OBO  TCAAGGGGCC GGGGGCGCGC CCTTGGAAAG CAAAGCGCGG ACGAGCCACA
   OMO  TCAAGGGGCC GGGGGCGCGC CCTTGGAAAG CAAAGCGCGG ACGAGCCACA
   ORI  TCAAGGGGCC GGGGGCGCGC CCTTGGAAAG CAAAGCGCGG ACGAGCCACA
```

```
PINBANK  TCAAGGGGCC GGGGGCGCGC CCTTGGAAAG CAAAGCACGG ACGAGCCACA
 PINKOR  TCAAGGGGCC GGGGGCGCGC CCTTGGAAAG CAAAGCACGG ACGAGCCACA
 PUNno1  TCAAGGGGCC GGGGGCGCGC CCTTGGAAAG CAAAGCGCGG ACGAGCCACA
    PUR  TCAAGGGGCC GGGGGCGCGC CCTTGGAAAG CAAAGCGCGG ACGAGCCACA
    RET  TCAAGGGGCC GGGGGCGCGC CCTTGGAAAG CAAAGCGCGG ACGAGCCACA
    RUB  TCAAGGGGCC GGGGGCGCGC CCTTGGAAAG CAAAGCGCGG ACGAGCCACA
    SCH  TCAAGGGGCC GGGGGCGCGC CCTTGGAAAG CAAAGCGCGG ACGAGCCACA
 SITno1  TCAAGGGGCC GGGGGCGCGC CCTTGGAAAG CAAAGCGCGG ACGAGCCACA
    SMI  TCAAGGGGCC GGGGGCGCGC CCTTGGAAAG CAAAGCGCGG ACGAGCCACA
    SPI  TCAAGGGGCC GGGGGCGCGC CCTTGGAAAG CAAAGCGCGG ACGAGCCACA
    TOR  TCAAGGGGCC GGGGGCGCGC CCTTGGAAAG CAAAGCGCGG ACGAGCCACA
    WIL  TCAAGGGGCC GGGGGCGCGC CCTTGGAAAG CAAAGCGCGG ACGAGCCACA

         [3051                                              3100]
 ABIno1  TGCAGGGAAA CTTGCACGTG TGGTCCTGAC CGGGGACCCC GGTATACTGT
    ALC  TGCAGGGAAA CTTGCACGTG TGGTCCTGAC CGGGGACCCC GGTATACTGT
    ASP  TGCAGGGAAA CTTGCACGTG TGGTCCTGAC CGGGGACCCC GGTATACTGT
    BRA  TGCAGGGAAA CTTGCACGTG TGGTCCTGAC CGGGGACCCC GGTATACTGT
    BRE  TGCAGGGAAA CTTGCACGTG TGGTCCTGAC CGGGGACCCC GGTATACTGT
    CHI  TGCAGGGAAA CTTGCACGTG TGGTCCTGAC CGGGGACCCC GGTATACTGT
    CRA  TGCAGGGAAA CTTGCACGTG TGGTCCTGAC CGGGGACCCC GGTATACTGT
 ENGno1  TGCAGGGAAA CTTGCACGTG TGGTCCTGAC CGGGGACCCC GGTATACTGT
    FAR  TGCAGGGAAA CTTGCACGTG TGGTCCTGAC CGGGGACCCC GGTATACTGT
 GLAno1  TGCAGGGAAA CTTGCACGTG TGGTCCTGAC CGGGGACCCC GGTATACTGT
    GLE  TGCAGGGAAA CTTGCACGTG TGGTCCTGAC CGGGGACCCC GGTATACTGT
 JEZno3  TGCAGGGAAA CTTGCACGTG TGGTCCTGAC CGGGGACCCC GGTATACTGT
    KOR  TGCAGGGAAA CTTGCACGTG TGGTCCTGAC CGGGGACCCC GGTATACTGT
    KOY  TGCAGGGAAA CTTGCACGTG TGGTCCTGAC CGGGGACCCC GGTATACTGT
    LIK  TGCAGGGAAA CTTGCACGTG TGGTCCTGAC CGGGGACCCC GGTATACTGT
 MARno4  TGCAGGGAAA CTTGCACGTG TGGTCCTGAC CGGGGACCCC GGTATACTGT
    MAT  TGCAGGGAAA CTTGCACGTG TGGTCCTGAC CGGGGACCCC GGTATACTGT
    MAX  TGCAGGGAAA CTTGCACGTG TGGTCCTGAC CGGGGACCCC GGTATACTGT
    MEX  TGCAGGGAAA CTTGCACGTG TGGTCCTGAC CGGGGACCCC GGTATACTGT
    MEY  TGCAGGGAAA CTTGCACGTG TGGTCCTGAC CGGGGACCCC GGTATACTGT
    MON  TGCAGGGAAA CTTGCACGTG TGGTCCTGAC CGGGGACCCC GGTATACTGT
    MOR  TGCAGGGAAA CTTGCACGTG TGGTCCTGAC CGGGGACCCC GGTATACTGT
    OBO  TGCAGGGAAA CTTGCACGTG TGGTCCTGAC CGGGGACCCC GGTATACTGT
    OMO  TGCAGGGAAA CTTGCACGTG TGGTCCTGAC CGGGGACCCC GGTATACTGT
    ORI  TGCAGGGAAA CTTGCACGTG TGGTCCTGAC CGGGGACCCC GGTATACTGT
PINBANK  TGCAGGGAAA CTTGCACGTG TGGCCCTGAC CGGGGACCCC GGTATACTGT
 PINKOR  TGCAGGGAAA CTTGCACGTG TGGCCCTGAC CGGGGACCCC GGTATACTGT
 PUNno1  TGCAGGGAAA CTTGCACGTG TGGTCCTGAC CGGGGACCCC GGTATACTGT
    PUR  TGCAGGGAAA CTTGCACGTG TGGTCCTGAC CGGGGACCCC GGTATACTGT
    RET  TGCAGGGAAA CTTGCACGTG TGGTCCTGAC CGGGGACCCC GGTATACTGT
    RUB  TGCAGGGAAA CTTGCACGTG TGGTCCTGAC CGGGGACCCC GGTATACTGT
    SCH  TGCAGGGAAA CTTGCACGTG TGGTCCTGAC CGGGGACCCC GGTATACTGT
 SITno1  TGCAGGGAAA CTTGCACGTG TGGTCCTGAC CGGGGACCCC GGTATACTGT
    SMI  TGCAGGGAAA CTTGCACGTG TGGTCCTGAC CGGGGACCCC GGTATACTGT
    SPI  TGCAGGGAAA CTTGCACGTG TGGTCCTGAC CGGGGACCCC GGTATACTGT
    TOR  TGCAGGGAAA CTTGCACGTG TGGTCCTGAC CGGGGACCCC GGTATACTGT
    WIL  TGCAGGGAAA CTTGCACGTG TGGTCCTGAC CGGGGACCCC GGTATACTGT

         [3101                                              3150]
 ABIno1  ACCAATATGT GTAGGTCCCC GTAATTCGAG TGAGATTGTC ACGGCGCAAA
    ALC  ACCAATATGT GTAGGTCCCC GTAATTCGAG TGAGATTGTC ACGGCGCAAA
    ASP  ACCAATATGT GTAGGTCCCC GTAATTCGAG TGAGATTGTC ACGGCGCAAA
    BRA  ACCAATATGT GTAGGTCCCC GTAATTCGAG TGAGATTGTC ACGGCGCAAA
    BRE  ACCAATATGT GTAGGTCCCC GTAATTCGAG TGAGATTGTC ACGGCGCAAA
    CHI  ACCAATATGT GTAGGTCCCC GTAATTCGAG TGAGATTGTC ACGGCGCAAA
    CRA  ACCAATATGT GTAGGTCCCC GTAATTCGAG TGAGATTGTC ACGGCGCAAA
 ENGno1  ACCAATATGT GTAGGTCCCC GTAATTCGAG TGAGATTGTC ACGGCGCAAA
    FAR  ACCAATATGT GTAGGTCCCC GTAATTCGAG TGAGATTGTC ACGGCGCAAA
 GLAno1  ACCAATATGT GTAGGTCCCC GTAATTCGAG TGAGATTGTC ACGGCGCAAA
    GLE  ACCAATATGT GTAGGTCCCC GTAATTCGAG TGAGATTGTC ACGGCGCAAA
 JEZno3  ACCAATATGT GTAGGTCCCC GTAATTCGAG TGAGATTGTC ACGGCGCAAA
    KOR  ACCAATATGT GTAGGTCCCC GTAATTCGAG TGAGATTGTC ACGGCGCAAA
    KOY  ACCAATATGT GTAGGTCCCC GTAATTCGAG TGAGATTGTC ACGGCGCAAA
    LIK  ACCAATATGT GTAGGTCCCC GTAATTCGAG TGAGATTGTC ACGGCGCAAA
 MARno4  ACCAATATGT GTAGGTCCCC GTAATTCGAG TGAGATTGTC ACGGCGCAAA
    MAT  ACCAATATGT GTAGGTCCCC GTAATTCGAG TGAGATTGTC ACGGCGCAAA
    MAX  ACCAATATGT GTAGGTCCCC GTAATTCGAG TGAGATTGTC ACGGCGCAAA
    MEX  ACCAATATGT GTAGGTCCCC GTAATTCGAG TGAGATTGTC ACGGCGCAAA
    MEY  ACCAATATGT GTAGGTCCCC GTAATTCGAG TGAGATTGTC ACGGCGCAAA
    MON  ACCAATATGT GTAGGTCCCC GTAATTCGAG TGAGATTGTC ACGGCGCAAA
    MOR  ACCAATATGT GTAGGTCCCC GTAATTCGAG TGAGATTGTC ACGGCGCAAA
    OBO  ACCAATATGT GTAGGTCCCC GTAATTCGAG TGAGATTGTC ACGGCGCAAA
    OMO  ACCAATATGT GTAGGTCCCC GTAATTCGAG TGAGATTGTC ACGGCGCAAA
    ORI  ACCAATATGT GTAGGTCCCC GTAATTCGAG TGAGATTGTC ACGGCGCAAA
PINBANK  ACCAATATGT GTAGGTCCCC GTAATTTGAG TGAGATTGTC ATGGCGCAAA
 PINKOR  ACCAATATGT GTAGGTCCCC GTAATTTGAG TGAGATTGTC ATGGCGCAAA
 PUNno1  ACCAATATGT GTAGGTCCCC GTAATTCGAG TGAGATTGTC ACGGCGCAAA
    PUR  ACCAATATGT GTAGGTCCCC GTAATTCGAG TGAGATTGTC ACGGCGCAAA
    RET  ACCAATATGT GTAGGTCCCC GTAATTCGAG TGAGATTGTC ACGGCGCAAA
    RUB  ACCAATATGT GTAGGTCCCC GTAATTCGAG TGAGATTGTC ACGGCGCAAA
```

```
   SCH  ACCAATATGT GTAGGTCCCC GTAATTCGAG TGAGATTGTC ACGGCGCAAA
SITno1  ACCAATATGT GTAGGTCCCC GTAATTCGAG TGAGATTGTC ACGGCGCAAA
   SMI  ACCAATATGT GTAGGTCCCC GTAATTCGAG TGAGATTGTC ACGGCGCAAA
   SPI  ACCAATATGT GTAGGTCCCC GTAATTCGAG TGAGATTGTC ACGGCGCAAA
   TOR  ACCAATATGT GTAGGTCCCC GTAATTCGAG TGAGATTGTC ACGGCGCAAA
   WIL  ACCAATATGT GTAGGTCCCC GTAATTCGAG TGAGATTGTC ACGGCGCAAA

         [3151                              3184]
ABIno1  GGCAGATATG GTTTGGTATT CCCTTGTTCC CCGT
   ALC  GGCAGATATG GTTTGGTATT CCCTTGTTCC CCGT
   ASP  GGCAGATATG GTTTGGTATT CCCTTGTTCC CCGT
   BRA  GGCAGATATG GTTTGGTATT CCCTTGTTCC CCGT
   BRE  GGCAGATATG GTTTGGTATT CCCTTGTTCC CCGT
   CHI  GGCAGATATG GTTTGGTATT CCCTTGTTCC CCGT
   CRA  GGCAGATATG GTTTGGTATT CCCTTGTTCC CCGT
ENGno1  GGCAGATATG GTTTGGTATT CCCTTGTTCC CCGT
   FAR  GGCAGATATG GTTTGGTATT CCCTTGTTCC CCGT
GLAno1  GGCAGATATG GTTTGGTATT CCCTTGTTCC CCGT
   GLE  GGCAGATATG GTTTGGTATT CCCTTGTTCC CCGT
JEZno3  GGCAGATATG GTTTGGTATT CCCTTGTTCC CCGT
   KOR  GGCAGATATG GTTTGGTATT CCCTTGTTCC CCGT
   KOY  GGCAGATATG GTTTGGTATT CCCTTGTTCC CCGT
   LIK  GGCAGATATG GTTTGGTATT CCCTTGTTCC CCGT
MARno4  GGCAGATATG GTTTGGTATT CCCTTGTTCC CCGT
   MAT  GGCAGATATG GTTTGGTATT CCCTTGTTCC CCGT
   MAX  GGCAGATATG GTTTGGTATT CCCTTGTTCC CCGT
   MEX  GGCAGATATG GTTTGGTATT CCCTTGTTCC CCGT
   MEY  GGCAGATATG GTTTGGTATT C@CTTGTTCC CCGT
   MON  GGCAGATATG GTTTGGTATT CCCTTGTTCC CCGT
   MOR  GGCAGATATG GTTTGGTATT CCCTTGTTCC CCGT
   OBO  GGCAGATATG GTTTGGTATT CCCTTGTTCC CCGT
   OMO  GGCAGATATG GTTTGGTATT CCCTTGTTCC CCGT
   ORI  GGCAGATATG GTTTGGTATT CCCTTGTTCC CCGT
PINBANK  GGCAGATATG GTTTGGTATT CCCTTGTTCC CCGT
 PINKOR  GGCAGATATG GTTTGGTATT CCCTTGTTCC CCGT
 PUNno1  GGCAGATATG GTTTGGTATT CCCTTGTTCC CCGT
   PUR  GGCAGATATG GTTTGGTATT CCCTTGTTCC CCGT
   RET  GGCAGATATG GTTTGGTATT CCCTTGTTCC CCGT
   RUB  GGCAGATATG GTTTGGTATT CCCTTGTTCC CCGT
   SCH  GGCAGATATG GTTTGGTATT CCCTTGTTCC CCGT
SITno1  GGCAGATATG GTTTGGTATT CCCTTGTTCC CCGT
   SMI  GGCAGATATG GTTTGGTATT CCCTTGTTCC CCGT
   SPI  GGCAGATATG GTTTGGTATT CCCTTGTTCC CCGT
   TOR  GGCAGATATG GTTTGGTATT CCCTTGTTCC CCGT
   WIL  GGCAGATATG GTTTGGTATT CCCTTGTTCC CCGT
;
endblock;
```

L'évolution des espèces :
L'être humain a la fâcheuse manie d'altérer la mémoire de la nature.

L'être humain domine, extermine,
La nature décline,
L'évolution chagrine...

L'être humain se raffine,
La nature imagine,
L'évolution fascine...

Marie Bouillé
19 décembre 2007