Proceedings of the Fourth International Crustacean Congress, 1998

STRUCTURE AND INTRASPECIFIC VARIABILITY OF THE CONTROL REGION mtDNA IN THE PINK SHRIMP, *FARFANTEPENAEUS DUORARUM* (DECAPODA, PENAEIDAE)

BY

MICHAŁ GRABOWSKI[1]) and KENNETH C. STUCK[2])

[1]) Laboratory of Polar Biology & Oceanobiology, University of Łódź, 12/16 Banacha str., PL-90-237 Łódź, Poland

[2]) Gulf Coast Research Laboratory, USM Institute of Marine Sciences, P.O. Box 7000, Ocean Springs, MS 39566-7000, U.S.A.

ABSTRACT

The entire control region (c. 989 nucleotides) of the mitochondrial genome of *Farfantepenaeus duorarum* with parts of its flanking genes (tRNAile and 12S-rRNA) has been amplified, sequenced, and compared to the published sequence of *F. notialis*. The sequence length and nucleotide composition appear well conserved. The estimated value of Kimura 2-parameter distance reaches 0.079 between both species of shrimp and 0.006-0.026 among conspecific individuals of *F. duorarum*. However, statistical evaluation of genetic distance data and parsimony analysis of sequences indicate significant differences between the two species. Parsimony analysis of sequences from six *F. duorarum* collected along the Atlantic coast and the coast of the Gulf of Mexico, does not reflect the geographic distribution of the specimens.

INTRODUCTION

Penaeid shrimps are among the most important crustaceans harvested and cultured throughout the world. However, despite their economic significance and heavy exploitation, they have been rarely studied at the population level (Benzie et al., 1993; Bouchon et al., 1994; Machado et al., 1992; Sunden & Davis, 1991). Recent implementation of the polymerase chain reaction (PCR) has enabled amplification and analysis of specific portions of the genome. In order to overcome the problem of insufficient sequence divergence within species, the AT-rich region, known also as a control region, has been employed in population studies. This major non-coding domain in the mtDNA molecule seems to be associated with high levels of variation in arthropods (Monnerot et al., 1990; Monforte et al., 1993) and in vertebrates (Meyer, 1993; Rosel et al., 1995).

Farfantepenaeus duorarum (Burkenroad, 1939) and *F. notialis* (Pérez-Farfante, 1967) have been considered to be closely related to each other: Burkenroad

(1939) described them as two morphological forms of *F. duorarum* and Pérez-Farfante (1967), based on biometric studies, considered each form to represent a geographical subspecies. *F. duorarum* ranges from lower Chesapeake Bay to southern Florida, through the Gulf of Mexico to Cape Catoche and the Isla Mujeres at the tip of the Yucatan Peninsula, while *F. notialis* ranges from Cuba to the Atlantic coast of Brazil (Pérez-Farfante, 1988). The studied species is commercially harvested along the coasts of the United States. The major pink shrimp fishery occurs in the Gulf of Mexico, where it comprises up to 10% of the total US shrimp landings (Klima et al., 1982; SEFSC, 1993).

In the current paper, a complete sequence of the pink shrimp *F. duorarum* AT-rich region, its structure, and intraspecific variability is presented. Also the questions concerning phylogenetic relationship between the species and *F. notialis*, as well as the rate of AT-rich region evolution are addressed.

MATERIALS AND METHODS

Pink shrimp specimens

Individuals of the species were collected along the Gulf of Mexico coasts: offshore Texas, 1 specimen; Davis Bayou (Mississippi), 1 spm.; Pascaguola River (Mississippi), 1 spm.; Fort Myers (southwestern Florida), 1 spm.; and on the Atlantic Coasts: Moorehead (North Carolina), 2 spms. Animals were instantly dissected and their 6th abdominal somites were preserved and kept in SED buffer (Amos & Hoelzel, 1991). A collection of samples is in the authors' possession.

DNA extraction

Samples of shrimp muscle tissue were digested overnight in a buffer containing 50 mM KCl, 10 mM Tris pH 8.5, 0.01% gelatin, 0.5% Nonidet P-40, 0.5% Tween 20, 80 μg/ml proteinase K, 1 M NaCl, and 0.6% CTAB (cetyltrimethylammonium bromide) followed by the standard phenol-chloroform extraction. After extraction the DNA was resuspended in 30-50 μl TE of pH = 8.0.

Amplification of mt AT-rich region

A primer set for the PCR reaction was designed based on mitochondrial genome partial sequence published (García-Machado et al., 1996) for *Farfantepenaeus notialis* (as *Penaeus notialis*). The primers were rooted in the tRNAile (forward) and 12S r-RNA (reverse) genes. The nucleotide sequences of the primers were, respectively: ctrgn1 — 5′TCAAGATAATCCTTTTTCAGGCAC3′ and ctrgn2 — 5′TGGCTCATTAAATTTTACATGTGC3′. The PCR was carried

out in a final volume of 25 μl in a reaction mixture containing all four dNTPs (each at 400 μM), 0.8 μM of primer concentration, 1.5 μM of $MgCl_2$, and 5 units of *Taq* polymerase (Amersham Life Science). The polymerase buffer already containing $MgCl_2$ was provided by the manufacturer. The thermal program was 94°C for 3 min. followed by 35 cycles of 30 sec. at 94°C, 30 sec. at 55°C, 60 sec. at 72°C, and 72°C for 5 min. as a final extension after the last cycle.

DNA cloning and sequencing

Electrophoresis of the amplified mixture along with size standard marker was performed in a 1.5% agarose gel in TBE buffer and the DNA was stained with ethidium bromide. A specific amplified fragment of $\sim$ 1000 bp was excised from the gel and purified using QIAquickTM Gel Extraction Kit (Qiagen®). The purified fragment was cloned using pGEM®-T Easy Vector System (Promega®). DNA from colonies containing inserts was used to inoculate 5 ml minipreps and the plasmid containing the PCR-generated insert was isolated with the WizardTM Plus Miniprep DNA Purification System (Promega®).

Purified plasmid DNA was sent to the University of Maine — DNA Sequencing Facility where it was sequenced with an ABI model 373A Stretch DNA automatic sequencer.

All obtained sequences were submitted to GenBank internet database (http://www.ncbi.nlm.nih.gov/). Accession codes for the sequences are: FD1 (AF100736), FD2 (AF104498), FD3 (AF104496), FD4 (AF104497), FD5 (AF104499), FD6 (AF104500).

Sequence analysis and estimation of genetic distance

Sequenced products were identified with BLASTN 2.0.3 searching utility on GenBank, EMBL, DDBJ, and PDB data bases (Altshul et al., 1997). All sequences were aligned with CLUSTAL V multiple alignment tool (Higgins et al., 1992).

Genetic distances between sequences were calculated according to Kimura's Two-Parameter Method (Kimura, 1980) using DNADIST program incorporated into PHYLIP package, version 3.5c (Felsenstein, 1993). A statistical support for the differences between the distances was provided by performing one-way ANOVA and post hoc Least Significance Difference test on the distance matrix.

The "Branch and Bound" parsimony method was used to construct the most parsimonious tree from the sequences. Obtained clades were statistically assessed by bootstrapping across 1000 replicates with the SEQBOOT program of the PHYLIP package.

RESULTS AND DISCUSSION

AT-rich region structure

Sequenced products amplified from six *Farfantepenaeus duorarum* individuals were identified to be homologous to published complete mitochondrial non-coding region and parts of enclosing tRNAile and 12S r-RNA genes of *F. notialis* (see García-Machado et al., 1996). Multiple alignment of all the sequences including *F. notialis* was done (App. 1). The length of the non-coding region of *F. duorarum* is 988-990 bp, which is very similar to the *F. notialis* (983 bp). Its AT content (c. 78.4%) is slightly lower than in *F. notialis* (79.3%) and intermediate between that of other crustaceans: 68% in *Artemia franciscana* Kellog, 1906 (see Valverde et al., 1994), *Daphnia pulex* Leydig, 1860 (see Van Raay & Crease, 1994) and insects: ~ 93% in *Drosophila yakuba* Burla, 1954 (see Clary & Wolstenholme, 1985) and *Anopheles gambiae* Giles, 1902 complex (see Caccone et al., 1996). All the sequences contain a high number of direct and inverted repeats of size 10-14 bp (from 7 to 12 per sequence), scattered evenly in the sequence. Among them there are only two conserved direct (dr1-5′TAAAAAATAGT3′ and dr2-5′TTATACAAAAA3′) and palindromic (pr1-3′TTTAATTAAA5′ and pr2-5′TAATTAATTA3′) repeats, occurring at the same positions in all analyzed sequences of both species. Two of them, dr1 and pr1, are placed between nucleotide 178 and nucleotide 298 (nucleotide numbers refer to the *F. notialis* sequence). This part of 120 bp appears to be the longest conserved domain within the non-coding region. It shows no intraspecific variablity and contains only 4 substituted positions (3 transitions and 1 transversion) differentiating the two *Farfantepenaeus* species. The other conserved domain is situated between nucleotide 685 and 739 showing neither intra- nor interspecific variability. In contrast, the most variable domain between nucleotide 440 and 595 contains 36 of 85 (including gaps) differences between *F. duorarum* and *F. notialis* (i.e., 42%).

Non-coding region variability

Data generated in the current study exhibit very little size variation and very high sequence identity among *F. duorarum* specimens (97.3-99.2%) and compared to *F. notialis* (~ 92%). The mean estimated Kimura 2-parameter distance between the species is approximately 0.079, while among conspecific *F. duorarum*, distances range from 0.006 to 0.026. One-way ANOVA and post hoc LSD test indicate that the distances between *F. notialis* and *F. duorarum* sequences are significantly higher than all the others in the matrix. Also, a parsimony analysis of non-coding region sequences of six specimens of *F. duorarum* and one of

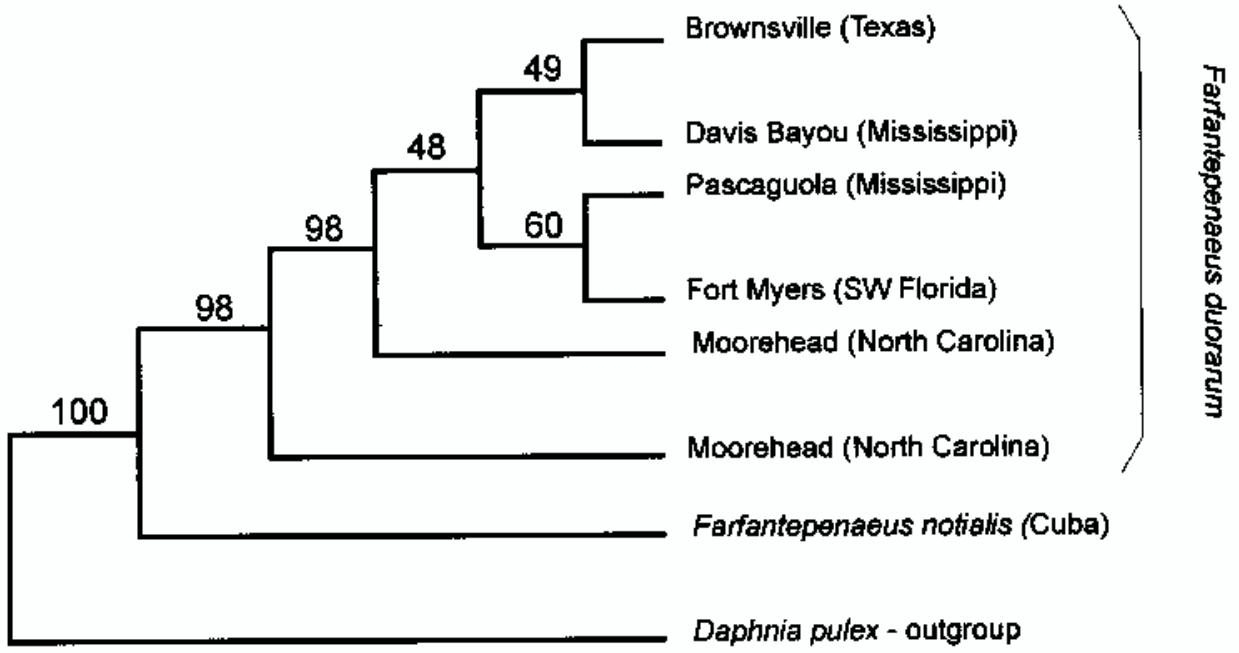

Fig. 1. Majority-rule consensus tree illustrating phylogenetic relationships among control region mtDNA sequences of *Farfantepenaeus duorarum* (Burkenroad, 1939) and *F. notialis* (Pérez-Farfante, 1969) estimated using the maximum parsimony method and the PHYLIP package. Numbers on branches represent bootstrap percentages (1000 replicates).

F. notialis (fig. 1) shows a branching order that corresponds with the biometry based conclusion by Pérez-Farfante (1967), that the two species form distinct phylogenetic units.

Rather low values of genetic divergence between the sequences contradict the observations of a high level of size and sequence variation of mtDNA AT-rich region in other organisms. For example, in *Drosophila* species, the region length varies from $\sim$ 1 kb in *D. virilis* Sturtevant, 1916 and *D. yakuba*, to 5.1 kb in *D. melanogaster* Meigen, 1830. Although of similar size, the regions of the first two species are very different, with only two small blocks showing sequence similarity (Clary & Wolstenholme, 1987). Also, an extensive intraspecific polymorphism for length of control regions within *D. melanogaster* was detected (Hale & Singh, 1986).

Large genetic distances were observed between other penaeid species (Palumbi & Benzie, 1991). However, calculation of genetic distance for the COI gene (sequences from Palumbi & Benzie, 1991) between *Litopenaeus stylirostris* (Stimpson, 1874) and *L. vannamei* (Boone, 1931) gives a value much higher (0.113) than the one estimated for *F. duorarum/F. notialis* for the faster evolving control region. That suggests the latter species separated more recently. The other possible explanation would be a slower than expected rate of control region evolution. Research upon the *Anopheles gambiae* complex (see Caccone et al., 1996), members of which are known to be very closely related, indicated several times lower

genetic distance based on AT-rich region compared to distance calculated using sequences of ND4 and ND5 genes. Similarly, in six species of the butterfly genus *Jalmenus* sp. the control region does almost not present any variation (Taylor et al., 1993). Also recent studies on amphibians (McKnight & Shaffer, 1997) and birds (Randi & Lucchini, 1998) show similar patterns. However, in the absence of sequences of other domains, it is difficult to explore the question of potentially slower rate of non-coding region evolution in the studied shrimp.

Branching order within the *F. duorarum* clade does not reflect well the geographic distribution of the specimens (fig. 1). A general phylogeographic pattern observed in a variety of marine taxa inhabiting the south-eastern region of the United States implies a deep phylogenetic division between Atlantic and Gulf of Mexico populations, in consequence of an almost complete isolation of the basins during the Pleistocene (Avise, 1992). Our data, supported by RFLP analysis of larger numbers of amplified sequences (unpublished), do not reflect well this division within the studied species. That could be attributed to gene flow between the populations or to a recent separation preventing the populations to reach a state of reciprocal monophyly with respect to the mtDNA gene tree (Avise, 1986). However, more extensive research is needed to elucidate the problem.

Concluding, PCR-generated sequence of mtDNA non-coding region is a useful marker for population analysis and speciation research in penaeid shrimp. It will be employed for further study in phylogeographic pattern of *Farfantepenaeus duorarum* and congeneric species.

REFERENCES

ALTSHUL, S. F., T. L. MADDEN, A. A. SCHÄFFER, J. ZHANG, Z. ZHANG, W. MILLER & D. J. LIPMAN, 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database programs. Nucleic Acids Res., **25**: 3389-3402.

AMOS, B. & A. R. HOELZEL, 1991. Long-term preservation of whale skin for DNA analysis. Rep. internatn. Whal. Comm., (Special Issue) **13**: 99-103.

AVISE, J. C., 1986. Mitochondrial DNA and the evolutionary genetics of higher animals. Phil. Trans. R. Soc. London, (B) **312**: 325-3424.

— —, 1992. Molecular population structure and the biogeographic history of a regional fauna: a case history with lessons for conservation biology. Oikos, **63**: 62-76.

BENZIE, J. A. H., E. BALLMENT & S. FRUSHER, 1993. Genetic structure of *Penaeus monodon* in Australia: concordant results from mtDNA and allozymes. Aquaculture, **111**: 89-93.

BOUCHON, D., C. SOUTY-GROSSET & R. RAIMOND, 1994. Mitochondrial DNA variation and markers of species identity in two penaeid shrimp species: *Penaeus monodon* Fabricius and *P. japonicus* Bate. Aquaculture, **127**: 131-144.

BURKENROAD, M. D., 1939. Further observations on Penaeidae of the northern Gulf of Mexico. Bull. Bingham oceanogr. Collect., Yale Univ., **6** (6): 1-62.

CACCONE, A., B. A. GARCIA & J. R. POWELL, 1996. Evolution of the mitochondrial DNA control region in the *Anopheles gambiae* complex. Insect mol. Biol., **5** (1): 51-59.

CLARY, D. O. & D. R. WOLSTENHOLME, 1985. The mitochondrial DNA molecule of *Drosophila yakuba*: nucleotide sequence, gene organisation, and genetic code. Journ. mol. Evol., **22**: 252-271.

FELSENSTEIN, J., 1993. PHYLIP (Phylogeny Inference Package) version 3.5c. (Distributed by the author. Department of Genetics, University of Washington, Seattle).

GARCÍA-MACHADO, E., N. DENNEBUOY, M. O. SUAREZ & J. C. MOUNOLOU, 1996. Partial sequence of the shrimp *Penaeus notialis* mitochondrial genome. Mol. Biol. Gen., **319**: 473-486.

HALE, L. & R. S. SINGH, 1986. Extensive variation and heteroplasmy in size of mitochondrial DNA among geographic populations of *D. melanogaster*. Proc. natn. Acad. Sci., U.S.A., **83**: 8813-8817.

HIGGINS, D. G., A. J. BLEASBY & R. FUCHS, 1992. Clustal V: improved software for multiple sequence alignment. Comput. Appl. Biosci., **8**: 189-191.

KIMURA, M., 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. Journ. mol. Evol., **16**: 111-120.

KLIMA, E. F., K. N. BAXTER & F. J. PATELLA, 1982. A review of the offshore shrimp fishery and the 1981 Texas closure. Mar. Fish. Rev., **44**: 16-30.

MACHADO, E. G., N. DENNEBUOY, M. O. SUAREZ, J. C. MOUNOLOU & M. MONNEROT, 1992. Mitochondrial 16S-rRNA gene of two species of shrimps: sequence variability and secondary structure. Crustaceana, **65** (3): 279-286.

MCKNIGHT, M. L. & H. B. SHAFFER, 1997. Large, rapidly evolving intergenic spacers in the mitochondrial DNA of the salamander family Ambystomatidae (Amphibia: Caudata). Mol. Biol. Evol., **14** (11): 1167-1176.

MEYER, A., 1993. Evolution of mitochondrial DNA in fishes. In: P. W. HOCHACHKA & T. P. MOMMSEN (eds.), Biochemistry and molecular biology of fishes, **2**: 1-38. (Molecular biology frontiers. Elsevier, Amsterdam).

MONFORTE, A., E. BARRIO & A. LATORRE, 1993. Characterization of the length polymorphism in the A + T-rich region of the *Drosophila obscura* group species. Journ. mol. Evol., **36**: 214-223.

MONNEROT, M., M. SOLIGNAC & D. R. WOLSTENHOLME, 1990. Discrepancy in divergence of the mitochondrial and nuclear genomes of *Drosophila teissieri* and *Drosophila yakuba*. Journ. mol. Evol., **30**: 500-508.

PALUMBI, S. R. & J. BENZIE, 1991. Large mitochondrial DNA differences between morphologically similar penaeid shrimp. Molec. mar. Biol. Biotech., **1** (1): 27-34.

PÉREZ-FARFANTE, I., 1967. A new species and two new subspecies of shrimp of the genus *Penaeus* from the western Atlantic. Proc. biol. Soc. Washington, **80**: 83-99.

— —, 1988. Illustrated key to penaeoid shrimps of commerce in the Americas. NOAA techn. Rep., NMFS-**64**: 1-32.

RANDI, E. & V. LUCCHINI, 1998. Organization and evolution of the mitochondrial DNA control region in the avian genus *Alectoris*. Journ. mol. Evol., **47** (4): 449-462.

ROSSEL, P. E., A. E. DIZON, M. G. HAYGOOD, 1995. Variability of the mitochondrial control region in populations of the harbour porpoise, *Phocoena phocoena*, on interoceanic and regional scales. Canadian Journ. Fish. aquat. Sci., **52**: 1210-1219.

SOUTHEAST FISHERIES SCIENCE CENTER (SEFSC), 1993. Status of fishery resources of the southeastern United States for 1992. NOAA tech. Memo., NMFS-SEFSC-**326**: 1-89.

SUNDEN, S. L. F. & S. K. DAVIS, 1991. Evaluation of genetic variation in a domestic population of *Penaeus vannamei* (Boone): a comparison with three natural populations. Aquaculture, **97**: 131-140.

TAYLOR, M. F. J., S. W. MCKECHNIE, N. PIERCE & M. KREITMAN, 1993. The lepidopteran mitochondrial control region: structure and evolution. Mol. Biol. Evol., **10**: 1259-1272.

UPHOLT, W. B., 1977. Estimation of DNA sequence divergence from comparison of restriction endonuclease digests. Nucleic Acids Res., **4**: 1257-1265.

VALVERDE, J. R., B. BATUECAS, C. MORATILLA, R. MARCO & R. GARESSE, 1994. The complete mitochondrial DNA sequence of the crustacean *Artemia franciscana*. Journ. mol. Evol., **4**: 400-408.

VAN RAAY, T. J. & T. J. CREASE, 1994. Partial mitochondrial DNA sequence of the crustacean *Daphnia pulex*. Curr. Genet., **25**: 66-72.

[For Appendix table 1, see pp. 341-344.]

First received 22 July 1998.

Final version accepted 31 December 1998.

APPENDIX TABLE 1

Comparison of six *Farfantepenaeus duorarum* (Burkenroad, 1939) and one *F. notialis* (Pérez-Farfante, 1969) mitochondrial control region sequences. Underlined characters indicate part of neighboring genes (tRNAile and 12S-rRNA). FNO = *F. notialis* from Cuba (sequence from García-Machado et al., 1996); FD1...6 = *F. duorarum*: 1, Moorehead (North Carolina); 2, Fort Myers (SW Florida); 3, Pascaguola (Mississippi); 4, Davis Bayou (Mississippi); 5, Brownsville (Texas); 6, Moorehead (North Carolina).

```
FN0 TTCATTTTAT ATAAAAAAgT AAAAAAAAgC AACATAACAA ACATTAAACA TACTACTTgC ATTAAATTgC gAgTTAgCCg gTTggTAAAg ATTCggggAg 100
FD1 .......... g......... .....g..A. .g.A...Ag. ........T. .gC...C... .........g. ...C...... A........A g......... 100
FD2 .......... g......... .....g..A. .g.A...Ag. ........T. .gC...C... .........g. ...C...... A........A g......... 100
FD3 .......... g......... .....g..A. .g.A...Ag. ........T. .gC...C... .........g. ...C...... A........A g......... 100
FD4 .......... g......... .....g..A. .g.g...Ag. ........T. .AC...C... .........C. ...C...... A........A g......... 100
FD5 .......... g......... .....g..A. .g.g...Ag. ........T. .AC...C... .........g. ...C...... A........A g......... 100
FD6 .......... g......... .....g..A. .g.g...Ag. ........T. .AT...C... .........g. ...C...... A........A g......... 100
FN0 TgAAAAAAAA ACTAATTTCA gTTATTAATA TATTAAATTA AATTTATATA AgAgAAAAAT AATTTTAAAT TTgTCTTTTT AATTAAATAg ATTgATTTTA 200
FD1 .......... .......... ....C..... .......... ...T.A.... .Agg...... A......... ..A...C... .......... .......... 200
FD2 .......... .......... ....C..... .......... ...T.g.... .AgA...... A......... ..A...C... .......... .......... 200
FD3 .......... .......... ....C..... .......... ...T.g.... .AgA...... A......... ..A...C... .......... .......... 200
FD4 .......... .......... ....C..... .......... ...T.g.... .AgA...... A......... ..A...C... .......... .......... 200
FD5 .......... .......... ....C..... .......... ...C.g.... .AgA...... A......... ..A...C... .......... .......... 200
FD6 .......... .......... ....C..... .......... ...T.A.... .AgA...... g......... ..A...T... .......... .......... 200
FN0 TTTTAAATgA TTggTATAAA AAATAgTgTC TTCTTTTTgT ATgAAACTCT AATCTTCTAg AATTTgAgTT TCAAATAAgA AgACgCTATT TTATTTAACA 300
FD1 .......... ..A....... .......... .........A ........T. .......... .....A.... .......... .......... .........CT 300
FD2 .......... ..A....... .......... .........A ........T. .......... .....A.... .......... .......... .........CT 300
FD3 .......... ..A....... .......... .........A ........T. .......... .....A.... .......... .......... .........CT 300
FD4 .......... ..A....... .......... .........A ........T. .......... .....A.... .......... .......... .........CT 300
FD5 .......... ..A....... .......... .........A ........T. .......... .....A.... .......... .......... .........CT 300
FD6 .......... ..A....... .......... .........A ........T. .......... .....A.... .......... .......... .........CT 300
```

APPENDIX TABLE 1
(Continued)

FN0	TTAATAATAg CAATATATAT CTACTATATA AgTCTTTTAT TTATATAATT AATTAAgTTA TATACTTCAA ATAgAATgCA TATCATATCA TgACATTCTA	695
FD1	....A..... ....g..... .T..... .....g..... ....C. ...-T...CTT	699
FD2	....A..... ....g..... .T..... .....g..... ....T. ...-C...CCT	699
FD3	....A..... ....C..... .T..... .....g..... ....C. ...-C...CCT	699
FD4	....A..... ....g..... .T..... .....A..... ....C. ...-T...CCT	699
FD5	....A..... ....A..... .C..... .....g..... ....C. ...-C...CCT	699
FD6	....A..... ....g..... .T..... .....g..... ....C. ...-C...CCT	699
FN0	AAAAATAgTC gAggggTCTT AAgAAATCCT ATAAATCAAg ATTTTTAATT TAAgAgggTT ACTTATAACC TATAATTAAA gTTCTAAAAC TAATAgTgTA	795
FD1	....T..... ..... C..... ....g..... ....g.....C.	799
FD2	....T..... ..... C..... ....g..... ....g.....C.	799
FD3	....T..... ..... C..... ....g..... ....g.....C.	799
FD4	....T..... ..... C..... ....g..... ....g.....C.	799
FD5	....T..... ..... C..... ....g..... ....g.....C.	799
FD6	....C..... ..... T..... ....g..... ....g.....C.	799
FN0	AggTTAgACC TTAgTAgTTA gTTTATTATg AgATACAAgT TTTATACAAA AAgggGAAAA AACCAAgggg g----TTAgA TTCTAACATC TTTATTTTAT	891
FD1	..g...A... ..... A..... .T.....g.. ....- -A...AA.gg --AgC.....	895
FD2	..g...A... ..... A..... .T.....g.. ....- -A...AA.gg ggAgT.....	897
FD3	..g...A... ..... A..... .T.....g.. ....- --...AA.gg g-AgT.....	895
FD4	..A...A... ..... A..... .T.....g.. ....- -A...AA.g- -gAgT.....	895
FD5	..g...A... ..... A..... .T.....g.. ....- -A...AA.g- -gAgT.....	895
FD6	..g...A... ..... A..... .T.....g.. ....- -A...--.Ag ggAgT.....	895

APPENDIX TABLE 1
(Continued)

```
FN0 TTAATAATAg CAATATATAT CTACTATATA AgTCTTTTAT TTATATAATT AATTAAgTTA TATACTTCAA ATAgAATgCA TATCATATCA TgACATTCTA 695
FD1 ....A..... ....g..... .T........ .......... ..g....... .......... .......... .........C. ...-T...CTT .......... 699
FD2 ....A..... ....g..... .T........ .......... ..g....... .......... .......... .........T. ...-C...CCT .......... 699
FD3 ....A..... ....C..... .T........ .......... ..g....... .......... .......... .........C. ...-C...CCT .......... 699
FD4 ....A..... ....g..... .T........ .......... ..A....... .......... .......... .........C. ...-T...CCT .......... 699
FD5 ....A..... ....A..... .C........ .......... ..g....... .......... .......... .........C. ...-C...CCT .......... 699
FD6 ....A..... ....g..... .T........ .......... ..g....... .......... .......... .........C. ...-C...CCT .......... 699
FN0 AAAAATAgTC gAggggTCTT AAgAAATCCT ATAAATCAAg ATTTTTAATT TAAgAgggTT ACTTATAACC TATAATTAAA gTTCTAAAAC TAATAgTgTA 795
FD1 .......... .......... .......... .......... ...T...... .......... C......... ....g..... .......... ..g.....C. 799
FD2 .......... .......... .......... .......... ...T...... .......... C......... ....g..... .......... ..g.....C. 799
FD3 .......... .......... .......... .......... ...T...... .......... C......... ....g..... .......... ..g.....C. 799
FD4 .......... .......... .......... .......... ...T...... .......... C......... ....g..... .......... ..g.....C. 799
FD5 .......... .......... .......... .......... ...T...... .......... C......... ....g..... .......... ..g.....C. 799
FD6 .......... .......... .......... .......... ...C...... .......... T......... ....g..... .......... ..g.....C. 799
FN0 AggTTAgACC TTAgTAgTTA gTTTATTATg AgATACAAgT TTTATACAAA AAggggAAAA AACCAAgggg g----TTAgA TTCTAACATC TTTATTTTAT 891
FD1 ..g...A... .......... A......... .T.....g.. .......... ..........- -A...AA.gg --AgC..... .......... .......... 895
FD2 ..g...A... .......... A......... .T.....g.. .......... ..........- -A...AA.gg ggAgT..... .......... .......... 897
FD3 ..g...A... .......... A......... .T.....g.. .......... ..........- --...AA.gg g-AgT..... .......... .......... 895
FD4 ..A...A... .......... A......... .T.....g.. .......... ..........- -A...AA.g- -gAgT..... .......... .......... 895
FD5 ..g...A... .......... A......... .T.....g.. .......... ..........- -A...AA.g- -gAgT..... .......... .......... 895
FD6 ..g...A... .......... A......... .T.....g.. .......... ..........- -A...--.Ag ggAgT..... .......... .......... 895
```

APPENDIX TABLE 1
(Continued)

```
FN0 gTTAATTAAT TAAATTAATT -TTTTAAgTA ATTATACAAA AACTACTACA TTgTTTTATg TgATTCAAAT TTTAgAAATA TATAAgAAAg TTACTTTAAT 990
FD1 ......T... .....T.... CT.....A.. .......... .....C.... .C........ .......... ....g....g ..g....... ...g...... 995
FD2 ......T... .....T.... CC.....A.. .......... .....C.... .C........ .......... ....g....g ..g....... ...g...... 997
FD3 ......T... .....T.... CC.....A.. .......... .....C.... .C........ .......... ....g....g ..g....... ...g...... 995
FD4 ......T... .....C.... CC.....A.. .......... .....C.... .C........ .......... ....g....g ..g....... ...g...... 995
FD5 ......T... .....T.... CC.....A.. .......... .....C.... .C........ .......... ....T....g ..g....... ...g...... 995
FD6 ......C... .....T.... CC.....A.. .......... .....T.... .C........ .......... ....g....g ..g....... ...g...... 995

FN0 AAAAgTTTTA TCCTAgCTgg TTCTTTCATT ATTAggTTAA Tg 1032
FD1 .......... .......... .......... A......... .. 1037
FD2 .......... .......... .......... A......... .. 1039
FD3 .......... .......... .......... g......... .. 1037
FD4 .......... .......... .......... A......... .. 1037
FD5 .......... .......... .......... A......... .. 1037
FD6 .......... .......... .......... A......... .. 1037
```