Thesis

The utility of the ITS-1 region in assessing the genetic variability in the brine fly, *Ephydra gracilis*

Submitted by

Brian J. Oney

As a partial fulfillment of the requirements for
Thesis in Zoology
Weber State University
Ogden, Utah
May 2007

WEBER STATE UNIVERSITY

We hereby recommend that the thesis prepared under our supervision by Brian J. Oney entitled "The utility of the ITS-1 region in assessing the genetic variability in the brine fly, *Ephydra gracilis*" be accepted as fulfilling partial requirements for the Thesis in Zoology.

Thesis Committee

Advisor

Department Chair

Introduction

The socioeconomic value of the wetlands of the Great Salt Lake (GSL) ecosystem is priceless. The buffer zones created by the wetlands protect many areas used by humans for recreation, agriculture and housing. The wetlands are also an important nursery and staging location for many migratory bird populations, as well as a sink for pollutants. As the largest U.S. lake west of the Mississippi River, the GSL serves as one of the largest migratory bird refugia in North America.

Brine flies or shore flies (Diptera, family Ephydridae) are located in many regions throughout the United States. Blanketing the rocks and shore in black, the flies associated with the GSL can reach estimated densities of 220 million flies per kilometer of shoreline (UDNR, 2000). Brine flies perform a critical ecological function, processing an estimated 90 million kg of organic matter annually (UDNR, 2000). As algivorous primary consumers, they are a necessary link between the primary production and upper levels of the food web of the GSL. Most notably, they are an important food source for the millions of birds that depend on the GSL.

In 1959, the railroad rock causeway was completed, and thereby separated Gunnison Bay and Gilbert Bay with the exception of two culverts for water flow, (Fig. 1). Because all the major tributaries flow into the south arm, the hydrological properties of the two arms diverged, most importantly salinity. The amount of biota is much lower in the Gunnison Bay due to the higher brine concentrations. Brine fly larvae use algal bioherms and other benthic refugia to pupate and the diet of brine flies of the GSL consists chiefly of algae, bacteria and organic detritus (UDNR, 2000), of which all are much less abundant in Gunnison Bay.

There are two principle species of brine flies associated with the GSL, *Ephydra hians* and *E. gracilis* (Fig. 2). Little is known about the precise distribution of the species or the extent to which flies of one region can interbreed with members of the same species from different areas of the GSL. *Ephydra gracilis* far outnumbers the larger *E. hians* in both arms of the GSL, however both species occur in lower numbers in the north arm, likely due to the higher physiological stresses associated with the much higher salinity (~220 ppt) of Gunnison Bay (Post, 1980). The focus of this study was an analysis of the population genetic structure of *E. gracilis*, specifically to examine whether genetic differentiation has occurred between the brine flies of the north and south arms. Of critical importance to such an investigation is the specific region of DNA to be examined. Due to the intraspecific focus, it is important to concentrate on a segment of DNA that is changing relatively rapidly over time, because the rock causeway was finished in 1959. The DNA chosen for this study is the internal transcribed spacer region known as ITS-1.

This study examined sequence variation of the ITS-1 region within two groups of *E. gracilis*, one at Rozel Point in the north arm of the lake and one from Antelope Island in the south arm. Because of the potential for interbreeding, it is expected that flies from a given group would be more genetically similar to each other than they are to flies from different groups. The objective was to determine the feasibility of using ITS-1 for studying genetic variation within and between populations of *E. gracilis*. Specific objectives include: (1) to clone and sequence the ITS-1 region from individual flies; (2) to examine the extent of variation of ITS-1 within individuals; (3) to examine the extent of intra-population ITS-1 variation; and (4) to compare intra- and inter-population genetic variation.

Materials & Methods

Fly Collection and DNA amplification

Several thousand brine flies were collected from Rozel Point and Antelope Island (Fig. 1) during the summer of 2006. A simple swoop technique with a plastic bag was used to collect flies along the shoreline of the sample sites. The flies were stored frozen. An individual fly was crushed in 50 µl of homogenization buffer and then the mix was heated for 15-20 min at 95-100 °C in order to lyse the cells, inactivate any nucleases, and isolate DNA.

Internal Transcribed Spacer 1 (ITS-1)

ITS-1 lies between two regions that code for ribosomal RNA, the 18S and 5.8S genes (Fig. 3). The transcribed spacer region is eventually cleaved from the primary transcript and degraded. These rRNA genes themselves are highly conserved due to their importance in cellular function. However, the region between the genes, ITS-1, is noncoding and, therefore, changes relatively rapidly over time. Mutational changes include nucleotide insertions and deletions, which result in size variation of the ITS-1. For genetic studies, it is an additional, useful feature of the ITS-1 region that it is present in approximately 200 copies within the genome (Hillis and Dixon, 1998). Therefore, it is possible to assess the amount of genetic variation within an individual fly and compare it with the variation seen among different individuals. A final advantage to using this rRNA region is that it exists in nearly all eukaryotic taxa, and the ITS-1 region has been utilized in many studies of genetic variation in both vertebrates and invertebrates. Therefore, the results of this study can be compared with those of other ITS-1 studies, especially those pertaining to other insects.

Polymerase Chain Reaction

A 1 μL portion of the DNA isolate was used as a template in the polymerase chain reaction (PCR). One of the primers (7247) used was complementary to the region at the 3′ end of 18S rRNA gene and the second primer (7246) was complementary to the 5' end of the 5.8S gene (Fig. 3). The sequence of primer 7247 is 5′ - CCAAACAAGGTTTCCGTAGG – 3′ and of primer is 7246, CCTGCGTTCTTCATCGAC. Standard 25 μL reactions were performed in separate reactions with DNA from each individual fly.

Reaction volumes were as follows:

15.625 μL water
2.0 μL dNTPs (l0mM ea.)
1.25 μL 7247 primer (5 μM)
1.25 μL 7246 primer (5 μM)
2.5 μL10X Buffer
2.5 μL $MgCl_2$ [25 mM]
0.125 μL *Taq* polymerase [5U/μL]
1 μL DNA template
---
25 μL total reaction volume

Template DNA concentrations ranged from 5 ng/μL to 50 ng/μL. One drop of mineral oil was added to each tube and the reaction was placed into a pre-heated (94°C) thermocycler. A step-cycle file with 40 cycles was used. Each cycle consisted of: one minute at 94°; one minute at 50°; one minute at 72°. Following amplification, the products were analyzed by electrophoresis through a 1% agarose gel (Fig. 5).

Vector Ligation

The PCR products were added at concentrations of 10-20 ng/μL to the plasmid pCR®2.1-TOPO® (Fig. 6). The reaction was incubated at room temperature for five minutes according to the TOPO® cloning reaction protocol (Invitrogen).

Bacterial Transformation and Growth

One Shot® chemically competent *E. coli* cells were inoculated with 2 μL of TOPO® cloning reaction and were treated according to protocol from the Invitrogen TOPO® TA Cloning® instruction manual (version K2) to transform bacteria. The bacterial solution was spread onto selective growth media in amounts of 50 μL and 150 μL, in order to ensure properly spaced bacterial growth colonies. The growth medium was Luria-Bertani (LB) containing ampicillin in 1μL/1mL concentrations. X-gal was spread onto plates (40 μL) shortly before plating the transformation reaction. X-gal is a compound that will be converted into a blue product in the presence of the plasmid alone, but remains colorless if the plasmid was successfully ligated with the insert. The blue coloring is produced by reaction of the products from the expression of the LacZα region of pCR® 2.1-TOPO® and X-gal. The ligation of the PCR product into the vector disrupts the expression of this gene rendering the colony white. False positives occur due to the loss or removal of the 3′ deoxythymidine, also disrupting LacZα's expression. The agar plates of spread bacterial solution were grown 10-14 hours in an incubator at 37° C. A single colony stemming from a single transformed bacterium was inoculated into 2 mL of selective (ampicillin) liquid growth medium and cultured 16-20 hours in a G24 Environmental Incubator Shaker (New Brunswick Scientific Co., Inc) at 37°C.

Plasmid Isolation

Plasmids were isolated from the bacterial culture using a standard protocol and screened to see if the plasmid had indeed joined with the PCR product to form a recombinant plasmid. As an inexpensive screening for presence of the insert, plasmids were isolated and endonucleased with *Eco*R I as per protocol (Table 1).

After screening, a Wizard mini-prep kit (Promega Corp.) was used to isolate plasmids according to the manufacturer's protocol. The plasmids were purified, screened for the presence of adequate concentrations (Fig. 8) and subsequently sequenced with an ABI Prism automated system.

Sequence Analysis

Sequence analyses and manipulations were performed with the computer programs MacVector and PAUP 4.0 (Phylogenetic Analysis Using Parsimony). PCR®2.1-TOPO® has two promoter regions (T7 and SP6, Fig. 6) that allow the inserted DNA to be sequenced from either direction. Thus, a sequence obtained must be examined to determine its orientation. As sequences were received they were examined to determine which of the primers and accompanying DNA strand had been sequenced. If the sequenced DNA began with the 7247 primer, it and upstream DNA were deleted from the sequence. If the sequenced DNA began with the 7246, the reverse and complement of the sequence were acquired using the MacVector program. Sequence alignments were done using the online program, ClustalW. The resulting alignments were used to perform a quantitative sequence-difference analysis. This difference consisted collectively of insertions, deletions, substitutions, and haplotypic positions.

Haplotype Analysis

Through sequence analysis it is possible to identify different haplotypes. Comparing only nucleotides at haplotypic positions, sequence differences were tabulated (Table 2). Entire sequences were compared to each other using the PAUP 4.0 program. Gross sequence differences were changed to percent differences, averaged for each case comparison and plotted (Fig. 9).

Results

The ITS-1 region that was amplified and sequenced comprised an average of 666 nucleotide base pairs (bp) in all flies (Fig. 5). A total of 53 sequences were obtained, 31 from the Antelope Island population, and 22 from Rozel Point. When the sequences were aligned, occasional insertions and deletions were seen (see Appendix). These corresponded to mutational events that occurred in various copies of the ITS-1 region. Cumulatively among the ITS-1 sequences, there were six single nucleotide polymorphisms (SNPs), a double nucleotide polymorphism, and an adenosine microsatellite locus. There were numerous instances in the analyses, where two sequences were 0% different, that is 100% identical.

Through sequence analysis it is possible to identify different haplotypes. Each haplotype represents a unique pattern of nucleotides that are shared by individuals as a result of inheritance from a common ancestor. Comparing only nucleotides at haplotypic positions, sequence differences were tabulated (Table 2). These differences are not stochastic, but rather they represent a different haplotype. Occuring 32 times in the *E. gracilis* genome, the summarized haplotype sequence X was shared by 12 individuals

from Rozel Point and 20 individuals from Antelope Island (Table 2). As summarized in Fig. 9, there is very little discrimination within or among individuals from the two populations of *E. gracilis* sampled and the overall percentage of shared sequences is very high (77%) (Fig.10).

Discussion

This study reveals very little discrimination within or among individuals from the two populations of *E. gracilis* sampled. This lack of discrimination can be explained in three ways. (1) Along the eastern and northern shore of the Great Salt Lake, *E. gracilis* represents a single, continuous population that exhibits little genetic differentiation. Since there are few published data on brine fly biology, it remains to be determined if this hypothesis is consistent with brine fly distribution. (2) Due to a relatively small sample size, these preliminary results may not reflect the overall pattern of variation that exists in this population, which is quite large, consisting of billions of individuals. (3) The ITS-1 region is an inappropriate DNA sequence for trying to answer the questions that were the subject of this study, because it is too variable. Even though ITS-1 has proven useful in a number of other similar studies, in a variety of organisms, no DNA region is appropriate to assess genetic variation in all taxa.

Based on a number of previous studies, the ITS-1 region is useful when examining intraspecific genetic differentiation. Vogler and DeSalle (1994) also found a high degree of sequence variation within individuals, but using a similar sequence analysis program they were able to deduce phylogenetic relationships among populations of the tiger beetle, *Cicindela dorsalis*. Presa et al. (2002) examined the utility of the ITS-1 region at a very large geographic scale and found it to also be useful in describing

relationships between populations of the brown trout, *Salmo trutta*. Performing similar analyses on the ITS-1 region of grasshopper, *Chorthippus parallelus*, Parkin and Butlin (2004) suggested that the presence of variation of the ITS-1 region within individuals was mostly due to intrachromosomal gene conversion and they were able to postulate when two geographically distinct populations diverged. Parkin and Butlin (2004) sequenced a total of 30 ITS-1 clones from 4 populations of *Chorthippus parallelus* scattered throughout the Iberian Peninsula. Presa et al. (2002) sequenced 103 ITS-1 clones and found 37 haplotypes that is, an original sequence differing significantly not as a result of mutational changes but inherited differences in sequence composition. Each of these studies observed a relatively large amount of ITS-1 sequence variation.

All of the ITS-1 studies cited above sampled sites far removed from each other (>1000 km). In this large distance there were many formidable, insuperable barriers to movement of the study species. In contrast, the two populations of brine flies sampled for this study are approximately 80 km apart with no comparable barriers separating them. Based on what is known about their ecology and with an adult lifespan of 3-5 days, brine flies mate shortly after emerging from the puparium and it is likely that they would seek the nearest mate. It is highly unlikely that a fly would traverse the entire distance of the lake. However, it is possible that wind and water currents transport adults and larvae throughout the lake. The water flow between the two arms is relatively minimal (Gwynn, 2007) but could be sufficient to transport larvae to the Rozel Point site, although it is questionable whether the larvae would tolerate such a large salinity change. The most likely explanation for movement of individuals is wind currents transporting adult flies around the lake. Due to the density of the *Ephydra* along the shoreline, the flies most

likely constitute a continuous single population surrounding the GSL supporting hypothesis (1). Thus, any mutations or variation within the genome of one fly could be propagated throughout the entire population, thereby homogenizing DNA sequences over time.

When unique mutations were removed from the analysis, phylogenetically informative nucleotide positions are left, that is, positions where two or more flies share the same nucleotide, which is different from other sequences. Parkin and Butlin (2004) used the information given from these SNPs, calling them within individual polymorphisms (WIPs), to draw their conclusions on population structure. There were 6 of these positions observed in the ITS-1 region (Table 2 & Appendix). The sequence of these positions represent a haplotype, a strictly inherited sequence that excludes mutations. The large percentage of shared haplotypes (Fig.10) strongly suggests that the two sample groups have a common ancestor, again supporting hypothesis (1).

When compared to other studies, it can be said that the total number of ITS-1 regions sequenced in this study (53) should be sufficient to reveal any genetic differentiation between the two sample groups if indeed they are different populations. In contrast to this study other studies, which also found a high degree of intra-individual ITS-1 sequence variation, were able to uncover population structure. This is likely due to the extended geographic range over which samples were collected. Thus, hypothesis (2) above seems unlikely, although more extensive sampling encompassing other regions of the lake inhabited by *E. gracilis* may provide additional evidence for or against this idea.

Harris and Crandall (2000) found ITS-1 to be phylogenetically uninformative when compared previous data from mtDNA in different populations of the crayfishes

Orconectes spp. However, Presa et al. (2002) found ITS-1 phylogenies to be congruent with those constructed from mtDNA. To further examine hypothesis (3), the results of sequence comparisons and haplotype distribution from another gene could be compared to those obtained here with ITS-1. If both genes show the same pattern, i.e., no differentiation among groups, this would argue for hypothesis (1) and against hypothesis (3). Genes that could be used include mitochondrial genes, such as *cytochrome b* or *cytochrome oxidase* subunit I, which have been useful for examining within-species variation (Cognato and Sperling, 2000; Castresana, 2001).

When the overall results of this study are compared to those described above, the results are quite similar. There is considerable sequence variation among the different copies of the ITS-1 regions within the same individual. When these sequences are compared to those from other individuals, the extent of variation is of a similar magnitude. The large majority of shared haplotypes (77%) of the flies from both groups suggests that genetic exchange is taking place between the sample groups. Therefore, hypothesis (1) provides the most probable explanation of the data. If this is the case, the ITS-1 region appears to be a suitable DNA sequence for revealing genetic variation and any associated biological changes of the study organism. The two groups of brine flies that were sampled belong to the same population, which seems to be relatively unaffected by the human-induced changes to the GSL ecosystem. Obtaining sequences of the mtDNA genes *cytochrome b* or *cytochrome oxidase* subunit I would contribute more information about the extent of genetic differentiation among populations of *E. gracilis*.

Acknowledgments

I would like to thank Drs. Chris Hoagstrom and John Mull for their editorial assistance, advice, and service on my thesis committee. I thank Dr. Jonathan Clark for his guidance, technical expertise, and trust. Without him, this study would not have been possible. This study was funded by an undergraduate research fellowship from the Office of Undergraduate Research at Weber State University in Ogden, Utah.

Literature Cited

Castresana, J. 2001. Cytochrome b phylogeny and the taxonomy of great apes and mammals. Molecular Biology and Evolution **18**: 465-471

Cognato, A.I. and Sperling, F.A.H. 2000. Phylogeny of Ips DeGeer species (Coleoptera: Scolytidae) inferred from mitochondrial cytochrome oxidase I DNA sequence. Molecular Phylogenetics and Evolution **14**: 445-460

Felix, E.A. and Rushforth, S.R., 1980. Biology of the south arm of the Great Salt Lake, Utah. Utah Geological and Mineral Survey, Bulletin **116**

Great Salt Lake Planning Team, (UDNR) Utah Department of Natural Resources, 2001. Great Salt Lake Comprehensive Management Plan Resource Document.

Gwynn, J.W. 2007. A Lake Divided - A History of the Southern Pacific Railroad Causeway and Its Effect on Great Salt Lake, Utah. Utah Geological Survey. <http://geology.utah.gov/utahgeo/gsl/lakedivided.htm>

Hillis, D.M. and M.T. Dixon. 1991. Ribosomal DNA: Molecular evolution and phylogenetic inference. Quarterly Review of Biology **66**: 411-453

Harris, D.J. and Crandall, K.A. 2000. Intragenomic variation within ITS1 and ITS2 of freshwater crayfishes (Decapoda: Cambaridae): implications for phylogenetic and microsatellite studies. Molecular Biology and Evolution **17**: 284-291

Parkin, E.J. and Butlin, R.K. 2004. Within- and between-individual variation among ITS1 copies in the meadow grasshopper *Chorthippus parallelus* indicates frequent intrachromosomal gene conversion. Molecular Biology and Evolution **21**: 1595-1601

Post, F.J., 1980. Biology of the North Arm. Utah Geological and Mineral Survey, Bulletin **116**

Presa, P.; Pardo, B. G.; Martinez, P.; Bernatchez, L. 2002. Phylogeographic congruence between mtDNA and rDNA ITS markers in brown trout. Molecular Biology and Evolution **19**: 2161-2175

Simpson, K.W. 1976. Shore flies and brine flies (Diptera: Ephydridae). Marine Insects, Ed. Cheng, L. North-Holland Pub. Co. New York Pg: 465-495

Vogler, A.P. and DeSalle, R. 1994. Evolution and phylogenetic information content of the ITS-1 region in the tiger beetle *Cicindela dorsalis*. Molecular Biology and Evolution **11**: 393-405

Wirth, W.W., Mathis, W.N., and Vockeroth, J.R. 1987. Manual of nearctic Diptera: Ephydridae. Vol. 2: 1027-1047

Table 1

This table highlights the protocol used as an inexpensive plasmid DNA isolation.

1. Transfer 1.5 ml of the overnight culture to a microcentrifuge tube and pellet cells at 10,000 RPM, 2 min.
2. Remove the supernatant completely and resuspend the cells in 300 µL STET buffer. Drag the tube across a rack approximately ten times. It is important that the cells are completely resuspended.
3. Add 25 µL lysozyme (10 mg/ml) and mix by inversion.
4. Immediately place tube in a boiling water bath and incubate 45 sec.
5. Spin 10,000 RPM, 5 min. The pellet contains cellular debris and chromosomal DNA; the supernatant contains plasmid DNA and RNA. Remove debris with a toothpick.
6. Add 350 µL isopropanol to the supernatant, mix well, and spin 10,000 rpm for 5 min.
7. Wash pellet once with 70% ethanol. Pellet DNA again.
8. Remove ethanol completely and let pellet dry for 10 min.
9. Resuspend DNA in 20 µL TE buffer; store at 4°C (short term) or -20°C (long term).
To check for the presence of an insert in the plasmid, perform a restriction enzyme digest. Be sure to include a control plasmid with no insert for comparison.
10. Remove 2 µL of the each plasmid DNA and combine with 8 µL restriction mix (0.2 µL RE, 1 µL 10x buffer, 6.8 µL H_2O). Mix well and incubate 37°C for 1 hour.
11. Analyze the digested DNA on a 1% agarose mini-gel (**Figure 7**).

Table 2

The summarized nucleotides and their respective positions in the ITS-1 region represent different haplotypes. All nucleotide positions mark nucleotides that were inherited. Sequences sharing a letter to the right are identical.

Position in ITS1	78	102	121	122	184	259	428	Haplotype
AI1A	C	A	C	T	A	A	G	Y
AI1B	C	A	C	T	A	A	G	Y
AI1C	T	A		T	A	A	G	
AI2A	T	A		T	C	G	A	X
AI2B	T	A		T	C	G	A	X
AI3A	T	A		T	C	G	A	X
AI3B	C	A	C	T	C	G	A	
AI3C	T	A		T	C	G	G	Q
AI4A	T	A		T	C	G	A	X
AI4B	T	A		A	C	G	A	W
AI4C	T	A		T	C	G	A	X
AI5A	T	A		T	C	G	A	X
AI5B	T	A		T	C	G	A	X
AI5C	T	A		T	C	G	A	X
AI6A	T	A		T	C	G	G	Q
AI6B	T	A		T	C	G	A	X
AI7A	T	A		T	C	G	A	X
AI7B	T	A		T	C	G	G	Q
AI7C	T	A		T	C	G	A	X
AI8A	T	A		T	C	G	G	Q
AI8B	T	A		T	C	G	A	X
AI9A	T	A		T	C	G	A	X
AI10A	T	A		T	C	G	A	X
AI10B	T	A		A	C	G	A	W
AI10C	T	A		T	C	G	A	X
AI10D	T	A		T	C	G	A	X
AI11A	T	A		T	C	G	A	X
AI11B	T	A		T	C	G	A	X
AI11C	T	A		A	C	G	A	W
AI11D	T	A		T	C	G	A	X
AI11E	T	A		T	C	G	A	X
RP1A	T	A		T	C	G	G	Q
RP1B	T	A		T	C	G	A	X
RP2A	T	A		T	C	G	A	X
RP3A	T	A		T	C	G	A	X
RP3B	T	A		T	C	G	A	X
RP3C	C	T	C	T	A	A	G	
RP4A	T	A		T	C	G	A	X
RP4B	T	A		T	C	G	A	X
RP4C	T	A		T	C	G	A	X
RP5A	T	A		T	C	G	A	Z
RP5B	C	A	C	T	A	A	G	Y
RP5C	C	A	C	T	A	A	G	Y
RP5D	C	A	C	T	C	G	A	
RP6A	T	A		T	C	G	A	X
RP6B	T	A		T	C	G	A	Z
RP6C	T	A		T	C	G	A	X
RP7A	T	A		T	C	G	A	X
RP7B	C	T	C	T	A	A	A	U
RP7C	T	A		T	C	G	A	X
RP8A	C	T	C	T	A	A	A	U
RP8B	T	A		T	C	G	A	X
RP9A	T	A		T	C	A	A	

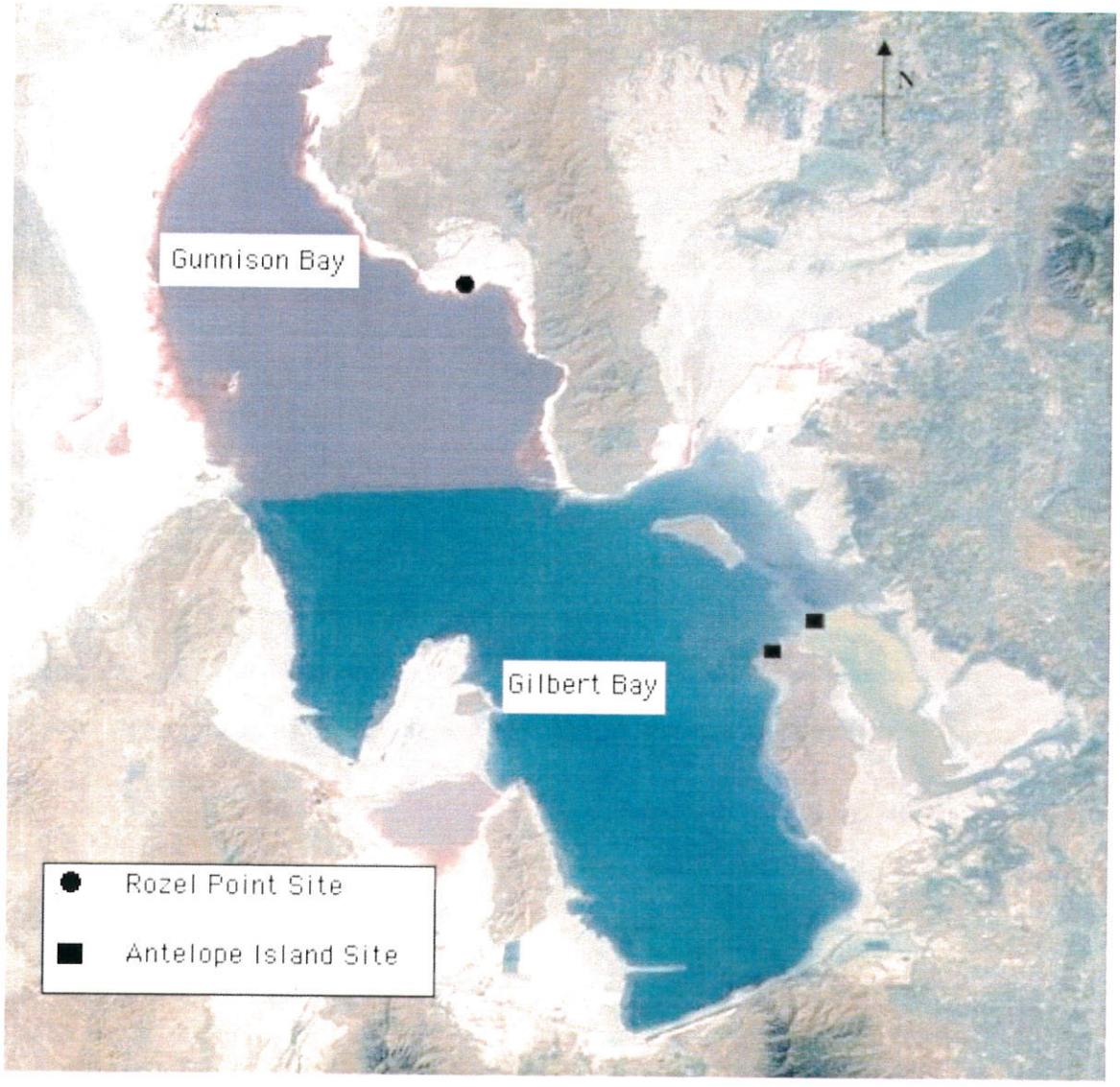


Fig. 1.—The sampling sites used in this study. Satellite picture taken from www.geology.utah.gov

Fig. 2.—The subject of this study is the gray brine fly, *Ephydra gracilis*.

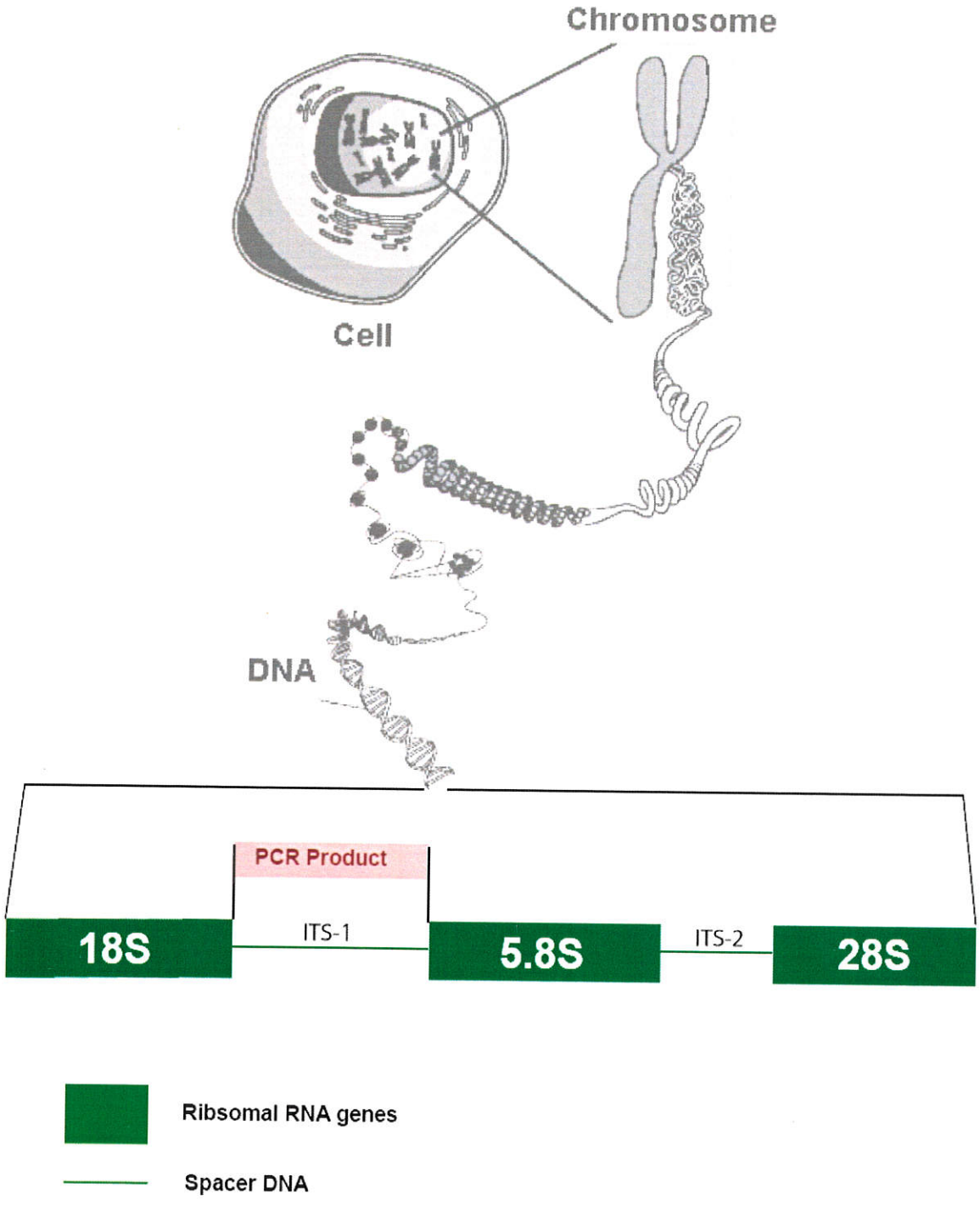


Fig. 3.—Schematic representation of the ITS-1 region.
The ITS-1 region lies between two genes that code for rRNAs. The ITS-1 region of the brine fly *E. cinerea* comprises, on average, 666 nucleotide base pairs, but varies greatly between species. The relative location of the PCR product is show in pink.

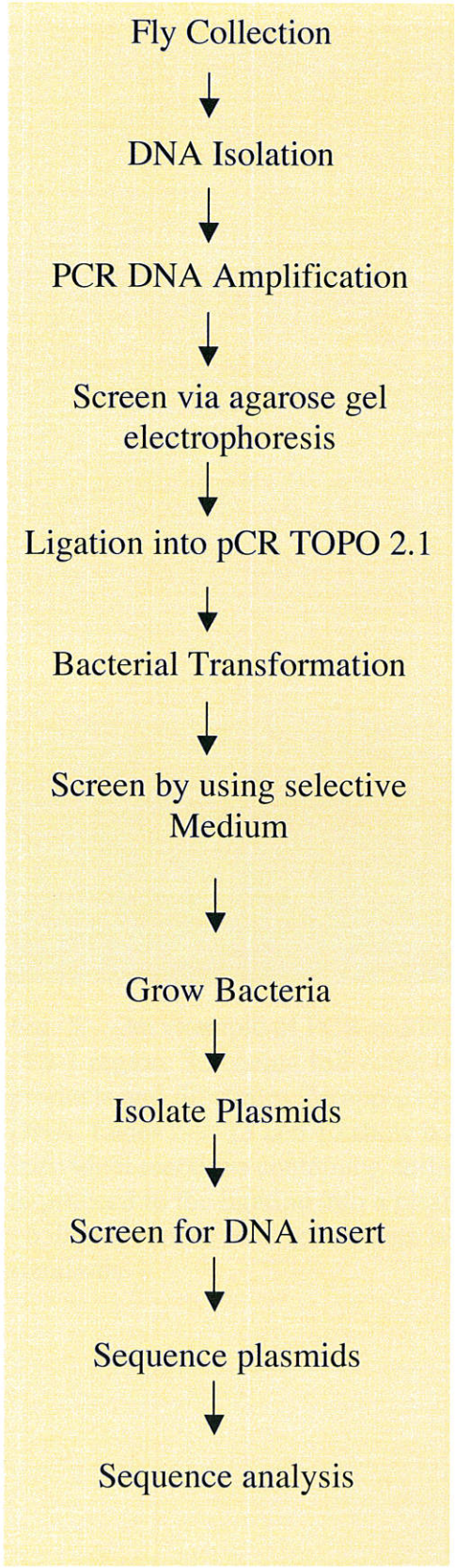


Fig. 4.—Schematic representation of the process undertaken to obtain the data in this study.

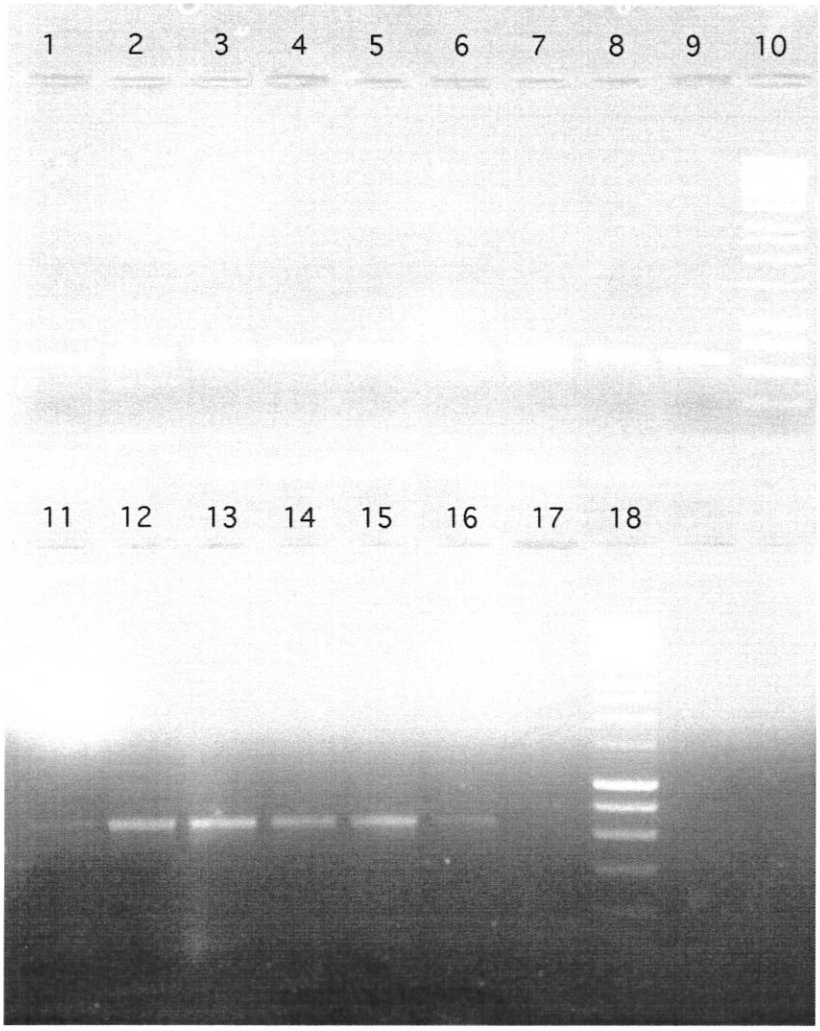
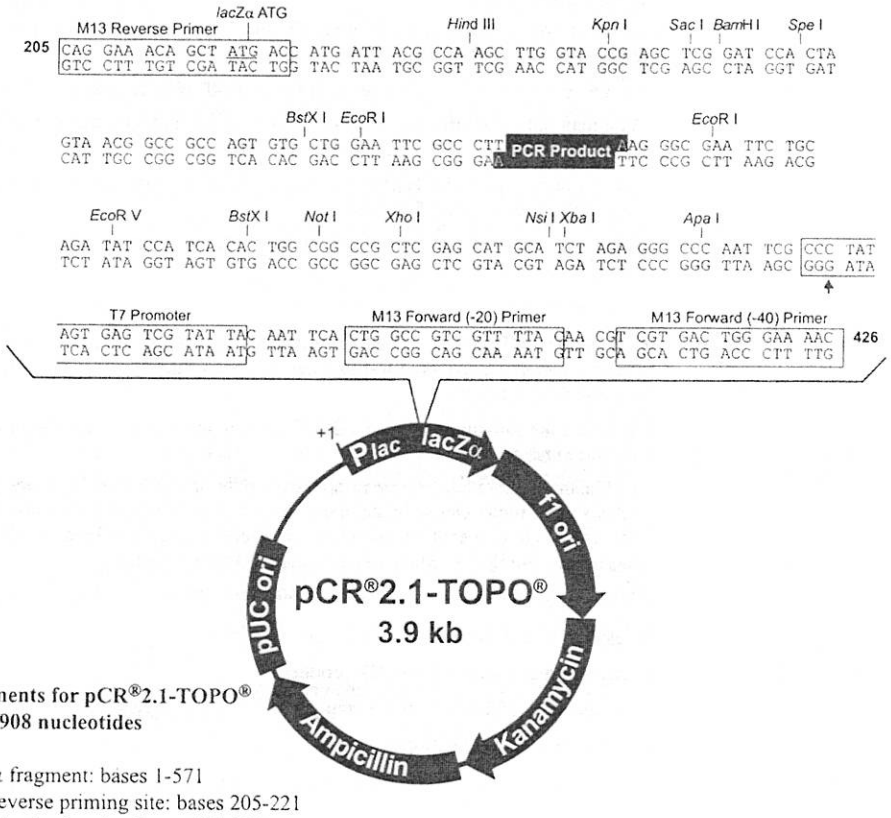


Fig. 5.—An example of PCR products associated with the ITS-1 region. Screening following PCR visualized the products and illustrates the approximate concentrations of DNA. Lanes 2-9, 13 and 15 show bands that came from individual reactions containing sufficient concentrations DNA to proceed to the ligation into a plasmid vector. The bands are ~650 nucleotides in length. Lanes 10 & 18 contain size standards.

Map of pCR®2.1-TOPO®

pCR®2.1-TOPO® Map

The map below shows the features of pCR®2.1-TOPO® and the sequence surrounding the TOPO® Cloning site. Restriction sites are labeled to indicate the actual cleavage site. The arrow indicates the start of transcription for T7 polymerase. For the full sequence of the vector, you may download it from our web site or call Technical Service (page 19).



Comments for pCR®2.1-TOPO®
3908 nucleotides

LacZα fragment: bases 1-571
M13 reverse priming site: bases 205-221
Multiple cloning site: bases 234-357
T7 promoter/priming site: bases 364-383
M13 Forward (-20) priming site: bases 391-406
M13 Forward (-40) priming site: bases 411-426
f1 origin: bases 548-962
Kanamycin resistance ORF: bases 1296-2090
Ampicillin resistance ORF: bases 2108-2968
pUC origin: bases 3113-3786

Fig. 6.—Map of the cloning vector, pCR®2.1-TOPO® vector plasmid (image taken from the Invitrogen K2 protocol) used for this study. The plasmids possesses an origin of replication, an ampicillin resistance region, an *Eco*R I endonuclease cleavage site, the LacZα operon, and the SP6 and T7 promoter regions.

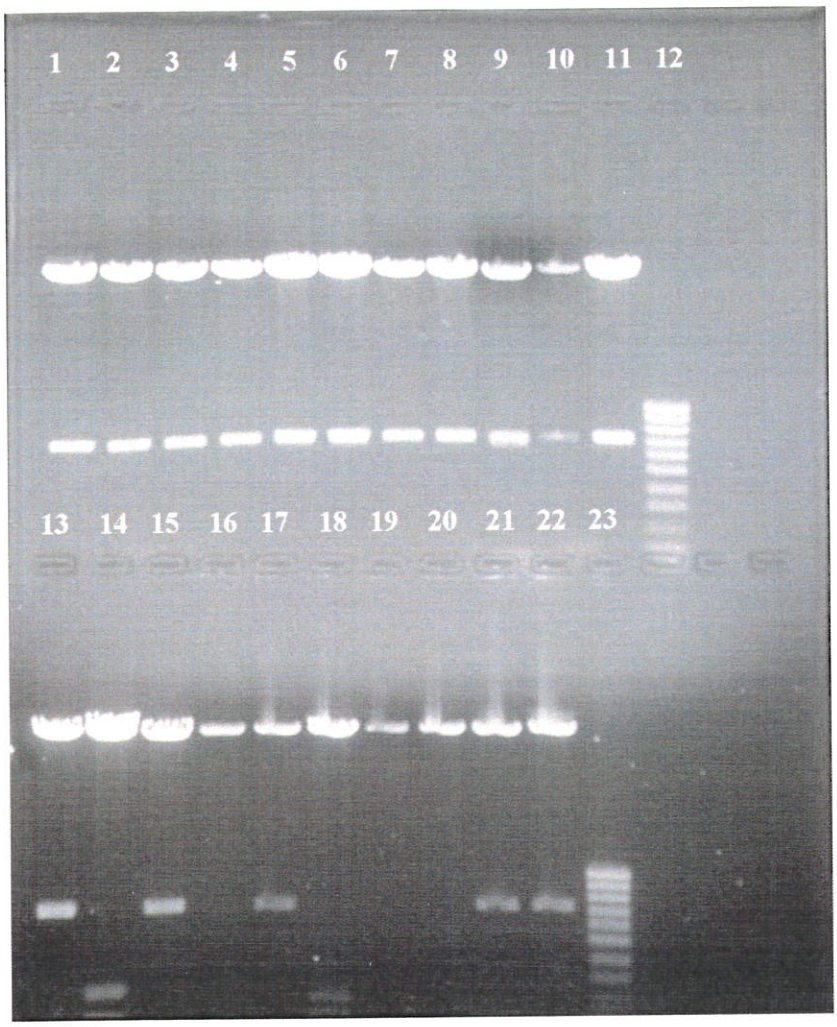


Fig. 7.—Example of recombinant plasmids containing ITS-1 inserts. Screening the plasmids (~3908 bp) for the inserted DNA (~650) revealed its presence (lanes 1-11, 13, 15, 17, 21, and 22) and its absence (lanes 16, 18, 19, and 20). The bands at ~200 bp in lanes 14 and 18 are artifacts. Lanes 12 and 23 contain size standards.

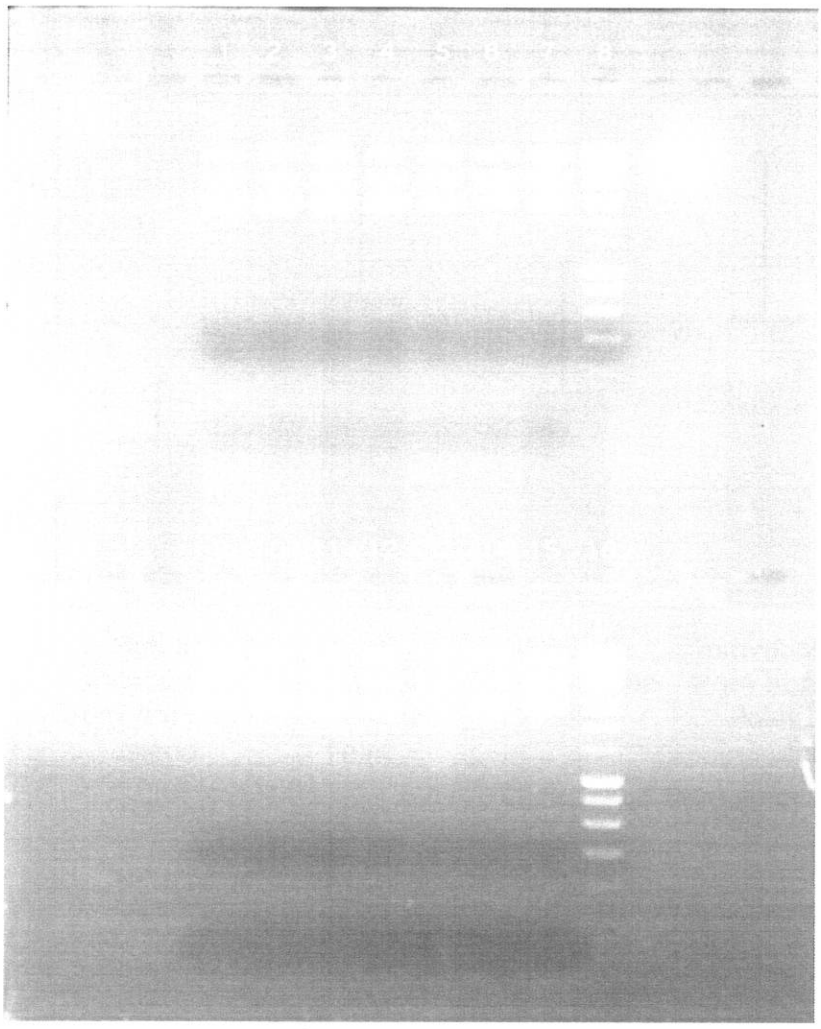


Fig. 8.—Determining the concentration of plasmids prior to sequencing. Lanes 1-7 and 9-15 show recombinant plasmids in concentrations adequate for sequencing. Lanes 8 and 16 contain a DNA size standard.

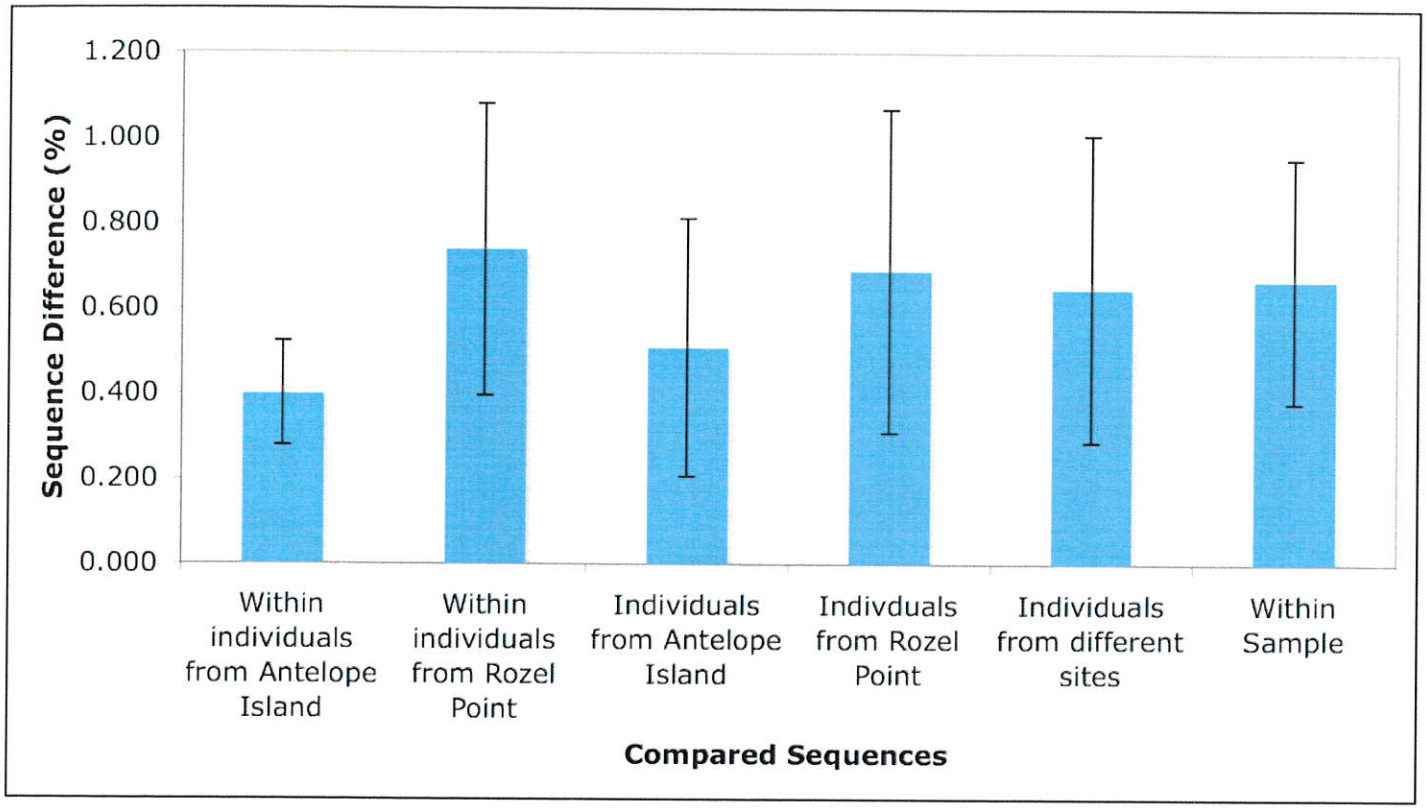


Fig. 9.—A summary of ITS-1 sequence differentiation with standard deviation bars. Complete sequences were included in each of the respective comparisons.

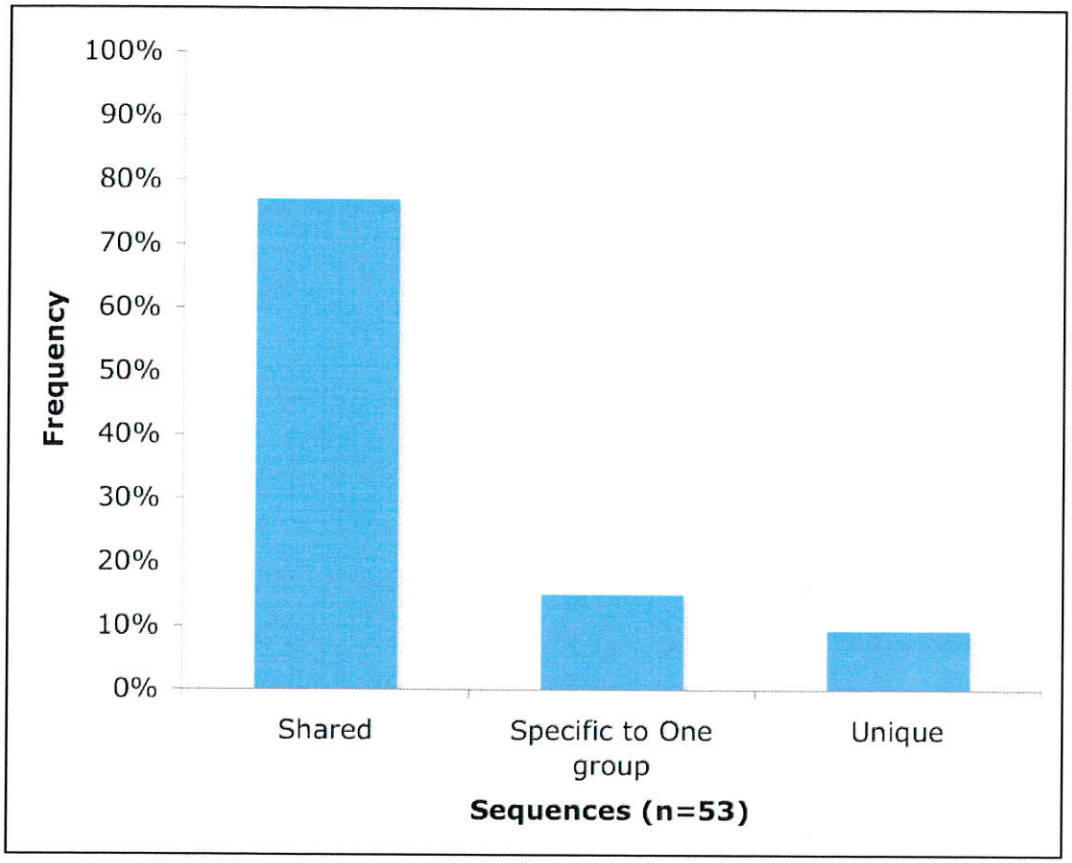


Fig. 10.— Distribution of ITS-1 haplotypes. The frequency of haplotypes common to both sampled groups is high (40 shared sequences). Some haplotypes were unique to a group (8 sequences total). Some haplotypes were unique only to an individual (5 sequences)

Appendix Alignment of ITS-1 sequences from *E. gracilis*. Sequence identities are marked by a *. Nucleotide positions that represent a different haplotype are marked by a #, whereas positions that are different and represent unique mutations lack a marking.

```
AI1A        TGAACCTGCGGAACGATCATTAATGTTTCTATCCGAAAATATTATTGAATATATTTTTTA
AI1B        TGAACCTGCGGAAGGATCATTAATGTTCCTATCCGAAAATATTATTGAATATATTTTTTA
AI1C        TGAACCTGCGGAAGGATCATTAATGTTCCTATCCGAAAATATTATTGAATATATTTTTTA
AI2A        TGAACCTGCGGAAGGATCATTAATGTTCCTATCCGAAAATATTATTGAATATATTTTTTA
AI2B        TGAACCTGCGGAAGGATCATTAATGTTCCTATCCGAAAATATTATTGAATATATTTTTTA
AI3A        TGAACCTGCGGAAGGATCATTATTGTTCCTATCCGAAAATATTATTGAATATATTTTTTA
AI3B        TGAACCTGCGGAAGGATCATTAATGTTCCTATCCGAAAATATTATTGAATATATTTTTTA
AI3C        TGAACCTGCGGAAGGATCATTAATGTTCCTATCCGAAAATATTATTGAATATATTTTTTA
AI4A        TGAACCTGCGGAAGGATCATTAATGTTCCTATCCGAAAATATTATTGAATATATTTTTTA
AI4B        TGAACCTGCGGAAGGATCATTAATGTTCCTATCCGAAAATATTATTGAATATATTTTTTA
AI4C        CGAACCTGCGGAAGGATCATTAATGTTCCTATCCGAAAATATTATTGAATATATTTTTTA
AI5A        TGAACCTGCGGAAGGATCATTAATGTTCCTATCCGAAAATATTATTGAATATATTTTTTA
AI5B        TGAACCTGCGGAAGGATCATTAATGTTCCTATCCGAAAATATTATTGAATATATTTTTTA
AI5C        TGAACCTGCGGAAGGATCATTAATGTTCCTATCCGAAAATATTATTGAATATATTTTTTA
AI6A        TGAACCTGCGGAAGGATCATTAATGTTCCTATCCGAAAATATTATTGAATATATTTTTTA
AI6B        TGAACCTGCGGAAGGATCATTAATGTTCCTATCCGAAAATATTATTGAATATATTTTTTA
AI7A        TGAACCTGCGGAAGGATCATTAATGTTCCTATCCGAAAATATTATTGAATATATTTTTTA
AI7B        TGAACCTGCGGAAGGATCATTAATGTTCCTATCCGAAAATATTATTGAATATATTTTTTA
AI7C        TGAACCTGCGGAAGGATCATTAATGTTCCTATCCGAAAATATTATTGAATATATTTTTTA
AI8A        TGAACCTGCGGAAGGATCATTAATGTTCCTATCCGAAAATATTATTGAATATATTTTTTA
AI8B        TGAACCTGCGGAAGGATCATTAATGTTCCTATCCGAAAATATTATTGAATATATTTTTTA
AI9A        TGAACCTGCGGAAGGATCATTAATNTTCCTATCCGAAAATATTATTGAATATATTTTTTA
AI10A       TGAACCTGCGGAAGGATCATTAATGTTCCTATCCGAAAATATTATTGAATATATTTTTTA
AI10B       TGAACCTGCGGAAGGATCATTAATGTTCCTATCCGAAAATATTATT-AATATATTTTTTA
AI10C       TGAACCTGCGGAAGGATCATTAATGTTCCTATCCGAAAATATTATTGAATATATTTTTTA
AI10D       TGAACCTGCGGAAGGATCATTAATGTTCCTATCCGAAAATATTATTGAATATATTTTTTA
AI11A       TGAACCTGCGGAAGGATCATTAATGTTCCTATCCGAAAATATTATTGAATATATTTTTTA
AI11B       TGAACCTGCGGAAGGATCATTAATGTTCCTATCCGAAAATATTATTGAATATATTTTTTA
AI11C       TGAACCTGCGGAAGGATCATTAATGTTCCTATCCGAAAATATTATTGAATATATTTTTTA
AI11D       TGAACCTGCGGTAGGATCATTAATGTTCCTATCCGAAAATATTATTGAATATATTTTTTA
AI11E       TGAACCTGCGGAAGGATCATTAATGTTCCTATCCGAAAATATTATTGAATATATTTTTTA
RP1A        TGAACCTGCGGAAGGATCATTAATGTTCCTATCCGAAAATATTATTGAATATATTTTTTA
RP1B        TGAACCTGCGGAAGGATCATTAATGTTCCTATCCGAAAATATTATTGAATATATTTTTTA
RP2A        TGAACCTGCGGAAGGATCATTAATGTTCCTATCCGAAAATATTATTGAATATATTTTTTA
RP3A        TGAACCTGCGGAAGGATCATTAATGTTCCTATCCGAAAATATTATTGAATATATTTTTTA
RP3B        TGAACCTGCGGAAGGATCATTAATGTTCCTATCCGAAAATATTATTGAATATATTTTTTA
RP3C        TGAACCTGCGGAAGGATCATTAATGTTCCTATCCGAAAATATTATTGAATATATTTTTTA
RP4A        TGAACCTGCGGAAGGATCATTAATGTTCCTATCCGAAAATATTATTGAATATATTTTTTA
RP4B        TGAACCTGCGGAAGGATCATTAATGTTCCTATCCGAAAATATTATTGAATATATTTTTTA
RP4C        TGAACCTGCGGAAGGATCATTAATGTTCCTATCCGAAAATATTATTGAATATATTTTTTA
RP5A        TGAACCTGCGGAAGGATCATTAATGTTCCTATCCGAAAATATTATTGAATATATTTTTTA
RP5B        TGAACCTGCGGAAGGATCATTAATGTTCCTATCCGAAAATATTATTGAATATATTTTTTA
RP5C        TGAACCTGCGGAAGGATCATTAATGTTCCTATCCGAAAATATTATTGAATATATTTTTTA
RP5D        TGAACCTGCGGAAGGATCATTAATGTTCCTATCCGAAAATATTATTGAATATATTTTTTA
RP6A        TGAACCTGCGGAAGGATCATTAATGTTCCTATCCGAAAATATTATTGAATATATTTTTTA
RP6B        TGAACCTGCGGAAGGATCATTAACGTTCCTATCCGAAAATATTATTGAATATATTTTTTA
RP6C        TGAACCTGCGGAAGGATCATTAATGTTCCTATCCGAAAATATTATTGAATATATTTTTTA
RP7A        TGAACCTGCGGAAGGATCATTAATGTTCCTATCCGAAAATATTATTGAATATATTTTTTA
RP7B        TGAACCTGCGGAAGGATCATTAATGTTCCTATCCGAAAATATTATTGAATATATTTTTTA
RP7C        TGAACCTGCGGAAGGATCATTAATGTTCCTATCCGAAAATATTATTGAATATATTTTTTA
RP8A        TGAACCTGCGGAAGGATCATTAATGTTCCTATCCGAAAATATTATTGAATATATTTTTTA
RP8B        TGAACCTGCGGAAGGATCATTAATGTTCCTATCCGAAAATATTATTGAATATATTTTTTA
RP9A        TGAACCTGCGGGAGGATCATTAATGTTCCTATCCGAAAATATTATTGAATATATTTTTTA
            ********** * ********   ** ****************** ************
```

```
AI1A            AAAAATATATTAATAATACAATTAAAAATGATGATTTAAGACAAGTTCTTAATCGTCATC
AI1B            AAAAATATATTAATAATACAATTAAAAATGATGATTTAAGACAAGTTCTTAATCGTCATC
AI1C            AAAAATATATTAATAATATAATTAAAAATGATGATTTAAGACAAGTTCTTAATCGTCATC
AI2A            AAAAATATATTAATAATATAATTAAAAATGATGATTTAAGACAAGTTCTTAATCGTCATC
AI2B            AAAAATATATTAATAATATAATTAAAAATGATGATTTAAGACAAGTTCTTAATCGTCATC
AI3A            AAAAATATATTAATAATATAATTAAAAATGATGATTTAAGACAAGTTCTTAATCGTCATC
AI3B            AAAAATATATTAATAATACAATTAAAAATGATGATTTAAGACAAGTTCTTAATCGTCATC
AI3C            AAAAATATATTAATAATATAATTAAAAATGATGATTTAAGACAAGTTCTTAATCGTCATC
AI4A            AAAAATATATTAATAATATAATTAAAAATGATGATTTAAGACAAGTTCTTAATCGTCATC
AI4B            AAAAATATATTAATAATATAATTAAAAATGATGATTTAAGACAAGTTCTTAATCGTCATC
AI4C            AAAAATATATTAATAATATAATTAAAAATGATGATTTAAGACAAGTTCTTAATCGTCATC
AI5A            AAAAATATATTAATAATATAATTAAAAATGATGATTTAAGACAAGTTCTTAATCGTCATC
AI5B            AAAAATATATTAATAATATAATTAAAAATGATGATTTAAGACAAGTTCTTAATCGTCATC
AI5C            AAAAATATATTAATAATATAATTAAAAATGATGATTTAAGACAAGTTCTTAATCGTCATC
AI6A            AAAAATATATTAATAATATAATTAAAAATGATGATTTAAGACAAGTTCTTAATCGTCATC
AI6B            AAAAATATATTAATAATATAATTAAAAATGATGATTTAAGACAAGTTCTTAATCGTCATC
AI7A            AAAAATATATTAATAATATGATTAAAAATGATGATTTAAGACAAGTTCTTAATCGTCATC
AI7B            AAAAATATATTAATAATATAATTAAAAATGATGATTTAAGACAAGTTCTTAATCGTCATC
AI7C            AAAAATATATTAATAATATAATTAAAAATGATGATTTAAGACAAGTTCTTAATCGTCATC
AI8A            AAAAATATATTAATAATATAATTAAAAATGATGATTTAAGACAAGTTCTTAATCGTCATC
AI8B            AAAAATATATTAATAATATAATTAAAAATGATGATTTAAGACAAGTTCTTAATCGTCATC
AI9A            AAAAATATATTAATAATATAATTAAAAATGACGATTTAAGACNAGTTCTTAATCGTCATC
AI10A           AAAAATATATTAATAATATAATTAAAAATGATGATTTAAGACAAGTTCTTAATCGTCATC
AI10B           AAAAATATATTAATAATATAATTAAAAATGATGATTTAAGACAAGTTCTTAATCGTCATC
AI10C           AAAAATATATTAATAATATAATTAAAAATGATGATTTAAGACAAGTTCTTAATCGTCATC
AI10D           AAAAATATATTAATAATATAATTAAAAATGATGATTTAAGACAAGTTCTTAATCGTCATC
AI11A           AAAAATATATTAATAATATAATTAAAAATGATGATTTAAGACAAGTTCTTAATCGTCATC
AI11B           AAAAATATATTAATAATATAATTAAAAATGATGATTTAAGACAAGTTCTTAATCGTCATC
AI11C           AAAAATATATTAATAATATAATTAAAAATGATGATTTAAGACAAGTTCTTAATCGTCATC
AI11D           AAAAATATATTAATAATATAATTAAAAATGATGATTTAAGACAAGTTCTTAATCGTCATC
AI11E           AAAAATATATTAATAATATAATTAAAAATGATGATTTAAGACAAGTTCTTAATCGTCATC
RP1A            AAAAATATATTAATAATATAATTAAAAATGATGATTTAAGACAAGTTCTTAATCGTCATC
RP1B            AAAAATATATTAATAATATAATTAAAAATGATGATTTAAGACAAGTTCTTAATCGTCATC
RP2A            AAAAATATATTAATAATATAATTAAAAATGATGATTTAAGACAAGTTCTTAATCGTCATC
RP3A            AAAAATATATTAATAATATAATTAAAAATGATGATTTAAGACAAGTTCTTAATCGTCATC
RP3B            AAAAATATATTAATAATATAATTAAAAATGATGATTTAAGACAAGTTCTTAATCGTCATC
RP3C            AAAAATATATTAATAATACAATTAAAAATGATGATTTAAGACTAGTTCTTAATCGTCATC
RP4A            GAAAATATATTAATAATATAATTAAAAATGATGATTTAAGACAAGTTCTTAATCGTCATC
RP4B            AAAAATATATTAATAATATAATTAAAAATGATGATTTAAGACAAGTTCTTAATCGTCATC
RP4C            AAAAATATATTAATAATATAATTAAAAATGATGATTTAAGACAAGTTCTTAATCGTCATC
RP5A            AAAAATATATTAATAATATAATTAAAAATGATGATTTAAGACAAGTTCTTAATCGTCATC
RP5B            AAAAATATATTAATAATACAATTAAAAATGATGATTTAAGACAAGTTCTTAATCGTCATC
RP5C            AAAAATATATTAATAATACAATTAAAAATGATGATTTAAGACAAGTTCTTAATCGTCATC
RP5D            AAAAATATATTAATAATACAATTAAAAATGATGATTTAAGACAAGTTCTTAATCGTCATC
RP6A            AAAAATATATTAATAATATAATTAAAAATGATGATTTAAGACAAGTTCTTAATCGTCATC
RP6B            AAAAATATATTAATAATATAATTAAAAATGATGATTTAAGACAAGTTCTTAATCGTCATC
RP6C            AAAAATATATTAATAATATAATTAAAAATGATGATTTAAGACAAGTTCTTAATCGTCATC
RP7A            AAAAATATATTAATAATATAATTAAAAATGATGATT-AAGACAAGTTCTTAATCGTCATC
RP7B            AAAAATATATTAATAATACAATTAAAAATGATGATTTAAGACTAGTTCTTAATCGTCATC
RP7C            AAAAATATATTAATAATATAATTAAAAATGATGATTTAAGACAAGTTCTTAATCGTCATC
RP8A            AAAAATATATTAATAATACAATTAAAAATGATGATTTAAGACTAGTTCTTAATCGTCATC
RP8B            AAAAATATATTAATAATATAATTAAAAATGATGATTTAAGACAAGTTCTTAATCGTCATC
RP9A            AAAAATATATTAATAATATAATTAAAAATGATGATTTAAGACAAGTTCTTAATCGTCATC
                 *****************# ********** **** *****#****************
```

```
AI1A        CTTAATAGTGTAACATTGTTACACTAAAAAGCTATATTATATATGGCAAAGAAAAATATC
AI1B        CTTAATAGTGTAACATTGTTACACTAAAAAGCTATATTATATATGGCAAAGAAAAATATC
AI1C        -TTAATAGTGTAACATTGTTACACTAAAAAGCTATATTATATATGGCAAAGAAAAATATC
AI2A        -TTAATAGTGTAACATTGTTACACTAAAAAGCTATATTATATATGGCAAAGAAAAATATC
AI2B        -TTAATAGTGTAACATTGTTACACTAAAAAGCTATATTATATATGGCAAAGAAAAATATC
AI3A        -TTAATAGTGTAACATTGTT-CACTAAAAAGCTATATTATATATGGCAAAGAAAAATATC
AI3B        CTTAATAGTGTAACATTGTTACACTAAAAAGCTATATTATATATGGCAAAGAAAAATATC
AI3C        -TTAATAGTGTAACATTGTTACACTAAAAAGCTATATTATATATGGCAAAGAAAAATATC
AI4A        -TTAATAGTGTAACATTGTTACACTAAAAAGCTATATTATATGTGGCAAAGAAAAATATC
AI4B        -ATAATAGTGTAACATTGTTACACTAAAAAGCTATATTATATATGGCAAAGAAAAATATC
AI4C        -TTAATAGTGTAACATTGTTACACTAAAAAGCTATATTATATATGGCAAAGAAAAATATC
AI5A        -TTAATAGTGTAACATTGTTACACTAAAAAGCTATATTATATATGGCAAAGAAAAATATC
AI5B        -TTAATAGTGTAACATTGTTACACTAAAAAGCTATATTATATATGGCAAGGAAAAATATC
AI5C        -TTAATAGTGTAACATTGTTACACTAAAAAGCTATATTATATATGGCAAAGAAAAATATC
AI6A        -TTAATAGTGTAACATTGTTACACTAAAAAGCTATATTATATATGGCAAAGAAAAATATC
AI6B        -TTAATAGTGTAACATTGTTACACTAAAAAGCTATATTATATATGGCAAAGAAAAATATC
AI7A        -TTAATAGTGTAACATTGTTACACTAAAAAGCTATATTATATATGGCAAAGAAAAATATC
AI7B        -TTAATAGTGTAACATTGTTACACTAAAAAGCTATATTATATATGGCAAAGAAAAATATC
AI7C        -TTAATAGTGTAACATTGTTACACTAAAAAGCTATATTATATATGGCAAAGAAAAATATC
AI8A        -TTAATAGTGTAACATCGTTACACTAAAAAGCTATATTATATATGGCAAAGAAAAATATC
AI8B        -TTANTAGTGTAACATTGTTACACTAAAAAGCTATATTATATATGGCAAAGAAAAATATC
AI9A        -TTAATAGTGTAACATTGTTACACTAAAAAGCTATATTATATATGGCAAAGAAAAATATC
AI10A       -TTAATAGTGTAACATTGTTACACTAAAAAGCTATATTATATATGGCAAAGAAAAATATC
AI10B       -ATAATAGTGTAACATTGTTACACTAAAAAGCTATATTATATATGGCAAAGAAAAATATC
AI10C       -TTAATAGTGTAACATTGTTACACTAAAAAGCTATATTATATATGGCAAAGAAAAATATC
AI10D       -TTAATAGTGTAACATTGTTACACTAAAAAGCTATATTATATATGGCAAAGAAAAATATC
AI11A       -TTAATAGTGTAACATTGTTACACTAAAAAGCTATATTATATATGGCAAAGAAAAATATC
AI11B       -TTAATAGTGTAACATTGTTACACTAAAAAGCTATATTATATATGGCAAAGAAAAATATC
AI11C       -ATAATAGTGTAACATTGTTACACTAAAAAGCTATATTATATATGGCAAAGAAAAATATC
AI11D       -TTAATAGTGTAACATTGTTACACTAAAAAGCTATATTATATATGGCAAAGAAAAATATC
AI11E       -TTAATAGTGTAACATTGTTACACTAAAAAGCTATATTATATATGGCAAAGAAAAATATC
RP1A        -TTAATAGTGTAACATTGTTACACTAAAAAGCTATATTATATATGGCAAAGAAAAATATC
RP1B        -TTAATAGTGTAACATTGTTACACTAAAAAGCTATATTATATATGGCAAAGAAAAATATC
RP2A        -TTAATAGTGTAGCATTGTTACACTAAAAAGCTATATTATATATGGCAAAGAAAAATATC
RP3A        -TTAATAGTGTAACATTGTTACACTAAAAAGCTATATTATATATGGCAAAGAAAAATATC
RP3B        -TTAATAGTGTAACATTGTTACACTAAAAAGCTATATTATATATGGCAAAGAAAAATATC
RP3C        CTTAATAGTGTAACATTGTTACACTAAAAAGCTATATTATATATGGCAAAGAAAAATATC
RP4A        -TTAATAGTGTAACATTGTTACACTAAAAAGCTATATTATATATGGCAAAGAAAAATATC
RP4B        -TTAATAGTGTAACATTGTTACACTAAAAAGCTATATTATATATGGCAAAGAAAAATATC
RP4C        -TTAATAGTGTAACATTGTTACACTAAAAAGCTATATTATATATGGCAAAGAAAAATATC
RP5A        -TTAATAGTGTAACATTGTTACACTAAAAAGCTATATTATATATGGCAAAGAAAAATATC
RP5B        CTTAATAGTGTAACATTGTTACACTAAAAAGCTATATTATATATGGCAAAGAAAAATATC
RP5C        CTTAATAGTGTAACATTGTTACACTAAAAAGCTATATTATATATGGCAAAGAAAAATATC
RP5D        CTTAATAGTGTAACATTGTTACACTAAAAAGCTATATTATATATGGCAAAGAAAAATATC
RP6A        -TTAATAGTGTAACATTGTTACACTAAAAAGCTATATTATATATGGCAAAGAAAAATATC
RP6B        -TTAATAGTGTAACATTGTTACACTAAAAAGCTATATTATATATGGCAAAGAAAAATATC
RP6C        -TTAATAGTGTAACATTGTTACACTAAAAAGCTATATTATATATGGCAAAGAAAAATATC
RP7A        -TTAATAGTGTAACATTGTTACACTAAAAAGCTATATTATATATGGCAAAGAAAAATATC
RP7B        CTTAATAGTGTAACATTGTTACACTAAAAAGCTATATTATATATGGCAAAGAAAAATATC
RP7C        -TTAATAGTGTAACATTGTTACACTAAAAAGCTATATTATATATGGCAAAGAAAAATATC
RP8A        CTTAATAGTGTAACATTGTTACACTAAAAAGCTATATTATATATGGCAAAGAAAAATATC
RP8B        -TTAATAGTGTAACATTGTTACACTAAAAAGCTATATTATATATGGCAAAGAAAAATATC
RP9A        -TTAATAGTGTAACATTGTTACACTAAAAAGCTATATTATATATGGCAAAGAAAAATATC
            ##** ******* *** *** ********************* ****** **********
```

```
AI1A        TATATTCATCTGGTTTAACTAATTCGCAATGCGCGAAGCAAATATAATTTGCAAACAATT
AI1B        TATATTCATCTGGTTTAACTAATTCGCAATGCGCGAAGCAAATATAATTTGCAAACAATT
AI1C        TATATTCATCTGGTTTAACTAATTCGCAATGCGCGAAGCAAATATAATTTGCAAACAATT
AI2A        TATCTTCATCTGGTTTAACTAATTCGCAATGCGCGAAGCAAATATAATTTGCAAACAATT
AI2B        TATCTTCATCTGGTTTAACTAATTCGCAATGCGCGAAGCAAATATAATTKGCAAACAATT
AI3A        TATCTTCATCTGGTTTAACTAATTCGCAATGCGCGAAGCAAATATAATTTGCAAACAATT
AI3B        TATCTTCATCTGGTTTAACTAATTCGCAATGCGCGAAGCAAATATAATTTGCAAACAATT
AI3C        TATCTTCATCTGGTTTAACTAATTCGCAATGCGCGAAGCAAATATAATTTGCAAACAATT
AI4A        TATCTTCATCTGGTTTAACTAATTCGCAATGCGCGAAGCAAATATAATTTGCAAACAATT
AI4B        TATCTTCATCTGGTTTAACTAATTCGCAATGCGCGAAGCAAATATAATTTGCAAACAATT
AI4C        TATCTTCATCTGGTTTAACTAATTCGCAATGCGCGAAGCAAATATAATTTGCAAACAATT
AI5A        TATCTTCATCTGGTTTAACTAATTCGCAATGCGCGAAGCAAATATAATTTGCAAACAATT
AI5B        TATCTTCATCTGGTTTAACTAATTCGCAATGCGCGAAGCAAATATAATTTGCAAACAATT
AI5C        TATCTTCATCTGGTTTAACTAATTCGCAATGCGCGAAGCAAATATAATTTGCAAACAATT
AI6A        TATCTTCATCTGGTTTAACTAATTCGCAATGCGCGAAGCAAATATAATTTGCAAACAATT
AI6B        TATCTTCATCTGGTTTAACTAATTCGCAATGCGCGAAGCAAATATAATTTGCAAACAATT
AI7A        TATCTTCATCTGGTTTAACTAATTCGCAATGCGCGAAGCAAATATAATTTGCAAACAATT
AI7B        TATCTTCATCTGGTTTAACTAATTCGCAATGCGCGAAGCAAATATAATTTGCAAACAATT
AI7C        TATCTTCATCTGGTTTAACTAATTCGCAATGCGCGAAGCAAATATAATTTGCAAACAATT
AI8A        TATCTTCATCTGGTTTAACTAATTCGCAATGCGCGAAGCAAATATAATTTGCAAACAATT
AI8B        TATCTTCATCTGGTTTAACTAATTCNCAATGCGCGAANCAAATATAATTTGCAAACAATT
AI9A        TATCTTCATCTGGNTTANCTAATTCGCNATGCGCGAAGCAAATATAATTTGCAAACNATT
AI10A       TATCTTCATCTGGTTTAACTAATTCGCAATGCGCGAAGCAAATATAATTTGCAAACAATT
AI10B       TATCTTCATCTGGTTTAACTAATTCGCAATGCGCGAAGCAAATATAATTTGCAAACAATT
AI10C       TATCTTCATCTGGTTTAACTAATTCGCAATGCGCGAAGCAAATATAATTTGCAAACAATT
AI10D       TATCTTCATCTGGTTTAACTAATTCGCAATGCGCGAAGCAAATATAATTTGCAAACAATT
AI11A       TATCTTCATCTGGTTTAACTAATTCGCAATGCGCGAAGCAAATATAATTTGCAAACAATT
AI11B       TATCTTCATCTGGTTTAACTAATTCGCAATGCGCGAAGCAAATATAATTTGCAAACAATT
AI11C       TATCTTCATCTGGTTTAACTAATTCGCAATGCGCGAAGCAAATATAATTTGCAAACAAYT
AI11D       TATCTTCATCTGGTTTAACTAATTCGCAATGCGCGAAGCAAATATAATTTGCAAACAATT
AI11E       TATCTTCATCTGGTTTAACTAATTCGCAATGCGCGAAGCAAATATAATTTGCAAACAATT
RP1A        TATCT--ATCTA-------TAATTCGCAATGCGCGAAGCAAATATAATTTGCAAACAATT
RP1B        TATCTTCATCTGGTTTAACTAATTCGCAATGCGCGAAGCAAATATAATTTGCAAACAATT
RP2A        TATCTTCATCTGGTTTAACTAATTCGCAATGCGCGAAGCAAATATAATTTGCAAACAATT
RP3A        TATCTTCATCTGGTTTAACTAATTCGCAATGCGCGAAGCAAATATAATTTGCAAACAATT
RP3B        TATCTTCATCTGGTTTAACTAATTCGCAATGCGCGAAGCAAATATAATTTGCAAACAATT
RP3C        TATATTCATCTGGTTTAACTAATTCGCAATGCGCGAAGCAAATATAATTTGCAAACAATT
RP4A        TATCTTCATCTGGTTTAACTAATTCGCAATGCGCGAAGCAAATATAATTTGCAAACAATT
RP4B        TATCTTCATCTGGTTTAACTAATTCGCAATGCGCGAAGCAAATATAATTTGCAAACAATT
RP4C        TATCTTCATCTGGTTTAACTAATTCGCAATGCGCGAAGCAAATATAATTTGCAAACAATT
RP5A        TATCTTCATCTGGTTTAACTAATTCGCAATGCGCGAAGCAAATATAATTTGCAAACAATT
RP5B        TATATTCATCTGGTTTAACTAATTCNCAATGCGCGAAGCAAATATAATTTGCAAACAATT
RP5C        TATATTCATCTGGTTTAACTAATTCGCAATGCGCGAAGCAAATATAATTTGCAAACAATT
RP5D        TATCTTCATCTGGTTTAACTAATTCGCAATGCGCGAAGCAAATATAATTTGCAAACAATT
RP6A        TATCTTCATCTGGTTTAACTAATTCGCAATGCGCGAAGCAAATATAATTTGCAAACAATT
RP6B        TATCTTCATCTGGTTTAACTAATTCGCAATGCGCGAAGCAAATATAATTTGCAAACAATT
RP6C        TATCTTCATCTGGTTTAACTAATTCGCAATGCGCGAAGCAAATATAATTTGCAAACAATT
RP7A        TATCTTCATCTGGTTTAACTAATTCGCAATGCGCGAAGCAAATATAATTTGCAAACAATT
RP7B        TATATTCATCTGGTTTAACTAATTCGCAATGCGCGAAGCAAATATAATTTGCAAACAATT
RP7C        TATCTTCATCTGGTTTAACTAATTCGCAATGCGCGAAGCAAATATAATTTGCAAACAATT
RP8A        TATATTCATCTGGTTTAACTAATTCGCAATGCGCGAAGCAAATATAATTTGCAAACAATT
RP8B        TATCTTCATCTGGTTTAACTAATTCGCAATGCGCGAAGCAAATATAATTTGCAAACAATT
RP9A        TATCTTCATCTGGTTTAACTAATTCGCAATGCGCGAAGCAAATATAATTTGCAAACAATT
            ***#*  ****        ****** * ********* ********** ****** * *
```

```
AI1A        TGATCCTTCAATATAAAAAAATTTAATATACAAATTCCTAAAATCTTAAATGGGTTTATT
AI1B        TGATCCTTCAATATAAAAAAATTTAATATACAAATTCCTAAAATCTTAAATGGGTTTATT
AI1C        TGATCCTTCAATATAAAAAAATTTAATATACAAATTCCTAAAATCTTAAATGGGTTTATT
AI2A        TGATCCTTCAATATAAAAAGATTTAATATACAAATTCCTAAAATCTTAAATGGGTTTATT
AI2B        TGATCCTTCAATATAAAAAGATTTAATATACAAATTCCTAAAATCTTAAATGGGTTTATT
AI3A        TGATCCTTCAATATAAAAAGATTTAATATACAAATTCCTAAAATCTTAAATGGGTTTATT
AI3B        TGATCCTTCAATATAAAAAGATTTAATATACAAATTCCTAAAATCTTAAATGGGTTTATT
AI3C        TGATCCTTCAATATAAAAAGATTTAATATACAAATTCCTAAAATCTTAAATGGGTTTATT
AI4A        TGATCCTTCAATATAAAAAGATTTAATATACAAATTCCTAAAATCTTAAATGGGTTTATT
AI4B        TGATCCTTCAATATAAAAAGATTTAATATACAAATTCCTAAAATCTTAAATTGGTTTATT
AI4C        TGATCCTTCAATATAAAAAGATTTAATATACAAATTCCTAAAATCTTAAATGGGTTTATT
AI5A        TGGTCCTTCAATATAAAAAGATTTAATATACAAATTCCTAAAATCTTAAATGGGTTTATT
AI5B        TGATCCTTCAATATAAAAAGATTTAATATACAAATTCCTAAAATCTTAAATGGGTTTATT
AI5C        TGATCCTTCAATATAAAAAGATTTAATATACAAATTCCTAAAATCTTAAATGGGTTTATT
AI6A        TGATCCTTCAATATAAAAAGATTTAATATACAAATTCCTAAAATCTTAAATGGGTTTATT
AI6B        TGATCCTTCAATATAAAAAGATTTAATATACAAATTCCTAAAATCTTAAATGGGTTTATT
AI7A        TGATCCTTCAATATAAAAAGATTTAATATACAAATTCCTAAAATCTTAAATGGGTTTATT
AI7B        TGATCCTTCAATATAAAAAGATTTAATATACAAATTCCTAAAATCTTAAATGGGTTTATT
AI7C        TGATCCTTCAATATAAAAAGATTTAATATACARATTCCTAAAATCTTAAATGGGTTTATT
AI8A        TGATCCTTCAATATAAAAAGATTTAATATACAAATTCCTAAAATCTTAAATGGGTTTATT
AI8B        TGATCCTTCAATATAAAAAGATTTAATATACAAATTCCTAAAATCTTAAATGGGTTTATT
AI9A        TGATCCTTCAATATAAAAAGATTTAATANACAAATTCCTAAAATCTTAAATGGGTTTATT
AI10A       TGATCCTTCAATATAAAAAGATTTAATATACAAATTCCTAAAATCTTAAATGGGTTTATT
AI10B       TGATCCTTCAATATAAAAAGATTTAATATACAAATTCCTAAAATCTTAAATGGGTTTATT
AI10C       TGATCCTTCAATATAAAAAGATTTAATATACAAATTCCTAAAATCTTAAATGGGTTTATT
AI10D       TGGTCCTTCAATATAAAAAGATTTAATATACAAATTCCTAAAATCTTAAATGGGTTTATT
AI11A       TGATCCTTCAATATAAAAAGATTTAATATACAAATTCCTAAAATCTTAAATGGGTTTATT
AI11B       TGATCCTTCAATATAAAAAGATTTAATATACAAATTCCTAAAATCTTAAATGGGTTTATT
AI11C       TGATCCTTCAATATAAAAAGATTTAATATACAAATTCCTAAAATCTTAAATGGGTTTATT
AI11D       TGATCCTTCAATATAAAAAGATTTAATATACAAATTCCTAAAATCTTAAATGGGTTTATT
AI11E       TGATCCTTCAATATAAAAAGATTTAATATACAAATTCCTAAAATCTTAAATGGGTTTATT
RP1A        TGATCCTTCAATATAAAAAGATTTAATATACAAATTCCTAAAATCTTAAATGGGTTTATT
RP1B        TGATCCTTCAATATAAAAAGATTTAATATACAAATTCCTAAAATCTTAAATGGGTTTATT
RP2A        TGATCCTTCAATATAAAAAGATTTAATATACAAATTCCTGAAATCTTAAATGGGTTTATT
RP3A        TGATCCTTCAATATAAAAAGATTTAATATACAAATTCCTAAAATCTTAAATGGGTTTATT
RP3B        TGATCCTTCAATATAAAAAGATTTAATATACAAATTCCTAAAATCTTAAATGGGTTTATT
RP3C        TGATCCTTCAATATAAAAAAATTTAATATACAAATTCCTAAAATCTTAAATGGGTTTATT
RP4A        TGATCCTTCAATATAAAAAGATTTAATATACAAATTCCTAAAATCTTAAATGGGTTTATT
RP4B        TGATCCTTCAATATAAAAAGATTTAATATACAAATTCCTAAAATCTTAAACGGGTTTATT
RP4C        TGATCCTTCAATATAAAAAGATTTAATATACAAATTCCTAAAATCTTAAATGGGTTTATT
RP5A        TGATCCTTCAATATAAAAAGATTTAATATACAAATTCCTAAAATCTTAAATGGGTTTATT
RP5B        TGATCCTTCAATATAAAAAAATTTAATATACAAATTCCTAAAATCTTAAATGGGTTTATT
RP5C        TGATCCTTCAATATAAAAAAATTTAATATACAAATTCCTAAAATCTTAAATGGGTTTATT
RP5D        TGATCCTTCAATATAAAAAGATTTAATATACAAATTCCTAAAATCTTAAATGGGTTTATT
RP6A        TGATCCTTCAATATAAAAAGATTTAATATACAAATTCCTAAAATCTTAAATGGGTTTATT
RP6B        TGATCCTTCAATATAAAAAGATTTAATATACAAATTCCTAAAATCTTAAATGGGTTTATT
RP6C        TGATCCTTCAATATAAAAAGATTTAATATACAAATTCCTAAAATCTTAAATGGGTTTATT
RP7A        TGATCCTTCAATATAAAAAGATTTAATATACAAATTCCTAAAATCTTAAATGGGTTTATT
RP7B        TGATCCTTCAATATAAAAAAATTTAATATACAAATTCCTAAAATCTTAAATGGGTTTATT
RP7C        TGATCCTTCAATATAAAAAGATTTAATATACAAATTCCTAAAATCTTAAATGGGTTTATT
RP8A        TGATCCTTCAATATAAAAAAATTTAATATACAAATTCCTAAAATCTTAAATGGGTTTATT
RP8B        TGATCCTTCAATATAAAAAGATTTAATATACAAATTCCTAAAATCTTAAATGGGTTTATT
RP9A        TGATCCTTCAATATAAAAAATTTAATATACAAATTCCTAAAATCTTAAATGGGTTTATTT
            ** ****************#******** *** ****** **********  ********
```

```
AI1A        TAAAACGGCAAAAATAAGGTAGCAAAATGCAGTCATATAAATAGTGACGTTCAAAATTGA
AI1B        TAAAACGGCAAAAATAAGGTAGCAAAATGCAGTCATATAAATAGTGACGTTCAAAATTGA
AI1C        TAAAACGGCAAAAATAAGGTAGCAAAATGCAGTCATATAAATAGTGACGTTCAAAATTGA
AI2A        TAAAACGGCAAAAATAAGGTAGCAAAATGCARTCATATAAATAGTGACGTTCAAAATTGA
AI2B        TAAAACGGCAAAAATAAGGTAGCAAAATGCAGTCATATAAATAGTGACGTTCAAAATTGA
AI3A        TAAAACGGCAAAAATAAGGTAGCAAAATGCAGTCATATAAATAGTGACGTTCAAAATTGA
AI3B        TAAAACGGCAAAAATAAGGTAGCAAAATGCAGTCATATAAATAGTGACGTTCAAAATTGA
AI3C        TAAAACGGCAAAAATAAGGTAGCAAAATGCAGTCATATAAATAGTGACGTTCAAAATTGA
AI4A        TAAAACGGCAAAAATAAGGTAGCAAAATGCAGTCATATAAATAGTGACGTTCAAAATTGA
AI4B        TAAAACGGCAAAAATAAGGTAGCAAAATGCAGTCATATAAATAGTGACGTTCAAAATTGA
AI4C        TAAAACGGCAAAAATAAGGTAGCAAAATGCAGTCATATAAATAGTGACGTTCAAAATTGA
AI5A        TAAAACGGCAAAAATAAGGTAGCAAAATGCAGTCATATAAATAGTGACGTTCAAAATTGA
AI5B        TAAAACGGCAAAAATAAGGTAGCAAAATGCAGTCATATAAATAGTGACGTTCAAAATTGA
AI5C        TAAAACGGCAAAAATAAGGTAGCAAAATGCAGTCATATAAATAGTGACGTTCAAAATTGA
AI6A        TAAAACGGCAAAAATAAGGTAGCAAAATGCAGTCATATAAATAGTGACGTTCAAAATTGA
AI6B        TAAAACGGCAAAAATAAGGTAGCAAAATGCAGTCATATAAATAGTGACGTTCAAAATTGA
AI7A        TAAAACGGCAAAAATAAGGTAGCAAAATGCAGTCATATAAATAGTGACGTTCAAAATTGA
AI7B        TAAAACGGCAAAAATAAGGTAGCAAAATGCAGTCATATAAATAGTGACGTTCAAAATTGA
AI7C        TAAAACGGCAAAAATAAGGTAGCAAAATGCAGTCATATAAATAGTGACGTTCAAAATTGA
AI8A        TAAAACGGCAAAAATAAGGTAGCAAAATGCAGTCATATAAATAGTGACGTTCAAAATTGA
AI8B        TANNACGGCAAAAATAAGGTAGCAAAATGCACTCATATAANTAGTGACGTTCAAAATTGA
AI9A        TAAAACGGCAAAAATAAGGTANCAAAATGCACTCATATAAATAGTGACGTTCAAAATTGA
AI10A       TAAAACGGCAAAAATAAGGTAGCAAAATGCAGTCATATAAATAGTGACGTTCAAAATTGA
AI10B       TAAAACGGCAAAAATAAGGTAGCAAAATGCAGTCATATAAATAGTGACGTTCAAAATTGA
AI10C       TAAAACGGCAAAAATAAGGTAGCAAAATGCAGTCATATAAATAGTGACGTTCAAAATTGA
AI10D       TAAAACGGCAAAAATAAGGTAGCAAAATGCAGTCATATAAATAGTGACGTTCAAAATTGA
AI11A       TAAAACGGCAAAAATAAGGTAGCAAAATGCAGTCATATAAATAGTGACGTTCAAAATTGA
AI11B       TAAAACGGCAAAAATAAGGTAGCAAAATGCAGTCATATAAATAGTGACGTTCAAAATTGA
AI11C       TAAAACGGCAAAAATAAGGTAGCAAAATGCAGTCATATAAATAGTGACGTTCAAAATTGA
AI11D       TAAAACGGCAAAAATAAGGTAGCAAAATGCAGTCATATAAATAGTGACGTTCAAAATTGA
AI11E       TAAAACGGCAAAAATAAGGTAGCAAAATGCAGTCATATAAATAGTGACGTTCAAAATTGA
RP1A        TAAAACGGCAAAAATAAGGTAGCAAAATGCAGTCATATAAATAGTGACGTTCAAAATTGA
RP1B        TAAAACGGCAAAAATAAGGTAGCAAAATGCAGTCATATAAATAGTGACGTTCAAAATTGA
RP2A        TAAAAAGGCAAAAATAAGGTAGCAAAATGCAGTCATATAAATAGTGACGTTCAAAATTGA
RP3A        TAAAACGGCAAAAATAAGGTAGCAAAATGCAGTCATATAAATAGTGACGTTCGAAATTGA
RP3B        TAAAACGGCAAAAATAAGGTAGCAAAATGCAGTCATATAAATAGTGACGTTCAAAATTGA
RP3C        TAAAACGGCAAAAATAAGGTAGCAAAATGCAGTCATATAAATAGTGATGTTCAAAATTGA
RP4A        TAAAACGGCAAAAATAAGGTAGCAAAATGCAGTCATATAAATAGTGACGTTCAAAATTGA
RP4B        TAAAACGGCAAAAATAAGGTAGCAAAATGCAGTCATATAAATAGTGACGTTCAAAATTGA
RP4C        TAAAACGGCAAAAATAAGTTAGCAAAATGCAGTCATATAAATAGTGACGTTCAAAATTGA
RP5A        TAAAACGGCAAAAATAAGGTAGCAAAATGCAGTCATATAAATAGTGACGTTCAAAATTGA
RP5B        TAAAACGGCAAAAATAAGGTANCAAAATGCAGTCATATAAATAGTGACGTTCAAAATTGA
RP5C        TAAAACGGCAAAAATAAGGTAGCAAAATGCAGTCATATAAATAGTGACGTTCAAAATTGA
RP5D        TAAAACGGCAAAAATAAGGTAGCAAAATGCAGTCATATAAATAGTGACGTTCAAAATTAA
RP6A        TAAAACGGCAAAAATAAGGTAGCAAAATGCAGTCATATAAATAGTGACGTTCAAAATTGA
RP6B        TAAAACGGCAAAAATAAGGTAGCAAAATGCAGTCATATAAATAGTGACGTTCAAAATTGA
RP6C        TAAAACGGCAAAAATAAGGTAGCAAAATGCAGTCATATAAATAGTGACGTTCAAAATTGA
RP7A        TAAAACGGCAAAAATAAGGTAGCAAAATGTAGTCATATAAATAGTGACGTTCAAAATTGA
RP7B        TAAAACGGCAAAAATAAGGTAGCAAAATGCAGTCATATAAATAGTGACGTTCAAAATTGA
RP7C        TAAAACGGCAAAAATAAGGTAGCAAAATGCAGTCATATAAATAGTGACGTTCAAAATTGA
RP8A        TAAAACGGCAAAAATAAGGTAGCAAAATGCAGTCATATAAATAGTGACGTTCAAAATTGA
RP8B        TAAAACGGCAAAAATAAGGTAGCAAAATGCAGTCATATAAATAGTGACGTTCAAAATTGA
RP9A        TAAAACGGCAAAAATAAGGTAGCAAAATGCAGTCATATAAATAGTGACGTTCAAAATTGA
            **  * ************ ** ******* * ******** ****** **** ***** *
```

```
AI1A        ATATTTGTGGAGTGTATTCAATGAAGAATCTATTGAGCTTGTTAAACATTTTCGTTGCAC
AI1B        ATATTTGTGGAGTGTATTCAATGAAGAATCTATTGAGCTTGTTAAACATTTTCGTTGCAC
AI1C        ATATTTGTGGAGTGTATTCAATGAAGAATCTATTGAGCTTGTTAAACATTTTCGTTGCAC
AI2A        ATATTTGTGGAGTGTATTCAATGAAGAATCTATTGAGCTTGTTAAACATTTTCGTTGCAC
AI2B        ATATTTGTGGAGTGTATTCAATGAAGAATCTATTGAGCTTGTTAAACATTTTCGTTGCAC
AI3A        ATATTTGTGGAGTGTATTCAATGAAGAATCTATTGAGCTTGTTAAACATTTTCGTTGCAC
AI3B        ATATTTGTGGAGTGTATTCAATGAAGAATCTATTGAGCTTGTTAAACATTTTCGTTGCAC
AI3C        ATATTTGTGGAGTGTATTCAATGAAGAATCTATTGAGCTTGTTAAACATTTTCGTTGCAC
AI4A        ATATTTGTGGAGTGTATTCAATGAAGAATCTATTGAGCTTGTTAAACATTTTCGTTGCAC
AI4B        ATATTTGTGGAGTGTATTCAATGAAGAATCTATTGAGCTTGTTAAACATTTTCGTTGCAC
AI4C        ATATTTGTGGAGTGTATTCAATGAAGAATCTATTGAGCTTGTTAAACATTTTCGTTGCAC
AI5A        ATATTTGTGGAGTGTATTCAATGAAGAATCTATTGAGCTTGTTAAACATTTTCGTTGCAC
AI5B        ATATTTGTGGAGTGTATTCAATGAAGAATCTATTGAGCTTGTTAAACATTTTCGTTGCAC
AI5C        ATATTTGTGGAGTGTATTCAATGAAGAATCTATTGAGCTTGTTAAACATTTTCGTTGCAC
AI6A        ATATTTGTGGAGTGTATTCAATGAAGAATCTATTGAGCTTGTTAAACATTTTCGTTGCAC
AI6B        ATATTTGTGGAGTGTATTCAATGAAGAATCTATTGAGCTTGTTAAACATTTTCGTTGCAC
AI7A        ATATTTGTGGAGTGTATTCAATGAAGAATCTATTGAGCTTGTTAAACATTTTCGTTGCAC
AI7B        ATATTTGTGGAGTGTATTCAATGAAGAATCTATTGAGCTTGTTAAACATTTTCGTTGCAC
AI7C        ATATTTGTGGAGCGTATTCAATGAAGAATCTATTGAGCTTGTTAAACATTTTCGTTGCAC
AI8A        ATATTTGTGGAGTGTATTCAATGAAGAATCTATTGAGCTTGTTAAACATTTTCGTTGCAC
AI8B        ATATTTGTGGAGTGNATTCANTGAANAATCTNTTGAGCTTGTTAAACATTTTCGTTGCAC
AI9A        ATATTTGTGGAGTGTATTCAATGAANAATCTATTGAGCTTGNTAAACATTTTCGTTGCNC
AI10A       ATATTTGTGGAGTGTATTCAATGAAGAATCTATTGAGCTTGTTAAACATTTTCGTTGCAC
AI10B       ATATTTGTGGAGTGTATTCAATGAAGAATCTATTGAGCTTGTTAAACATTTTCGTTGCAC
AI10C       ATATTTGTGGAGTGTATTCAATGAAGAATCTATTGAGCTTGTTAAACATTTTCGTTGCAC
AI10D       ATATTTGTGGAGTGTATTCAATGAAGAATCTATTGAGCTTGTTAAACATTTTCGTTGCAC
AI11A       ATATTTGTGGAGTGTATTCAATGAAGAATCTATTGAGCTTGTTAAACATTTTCGTTGCAC
AI11B       ATATTTGTGGAGTGTATTCAATGAAGAATCTATTGAGCTTGTCAAACATTTTCGTTGCAC
AI11C       ATATTTGTGGAGTGTATTCAATGAAGAATCTATTGAGCTTGTTAAACATTTTCGTTGCAC
AI11D       ATATTTGTGGAGTGTATTCAATGAAGAATCTATTGAGCTTGTTAAACATTTTCGTTGCAC
AI11E       ATATTTGTGGAGTGTATTCAATGAAGAATCTATTGAGCTTGTTAAACATTTTCGTTGCAC
RP1A        ATATTTGTGGAGTGTATTCAATGAAGAATCTATTGAGCTTGTTAAACATTTTCGTTGCAC
RP1B        ATATTTGTGGAGTGTATTCAATGAAGAATCTATTGAGCTTGTTAAACATTTTCGTTGCAC
RP2A        ATATTTGTGGAGTGTATTCAATGAAGAATCTATTGAGCTTGTTAAACATTTTCGTTGCAC
RP3A        ATATTTGTGGAGTGTATTCAATGAAGAATCTATTGAGCTTGTTAAACATTTTCGTTGCAC
RP3B        ATATTTGTGGAGTGTATTCAATGAAGAATCTATTGAGCTTGTTAAACATTTTCGTTGCAC
RP3C        ATATTTGTGGAGTGTATTCAATGAAGAATCTATTGAGCTTGTTAAACATTTTCGTTGCAC
RP4A        ATATTTGTGGAGTGTATGCAATGAAGAATCTATTGGGCTTGTTGAACATTTTCGTTGCAC
RP4B        ATATTTGTGGAGTGTATTCAATGAAGAATCTATTGAGCTTGTTAAACATTTTCGTTGCAC
RP4C        ATATTTGTGGAGTGTATTCAATGAAGAATCTATTGAGCTTGTTAAACATTTTCGTTGCAC
RP5A        ATATTTGTGGAGTGTATTCAATGGAGAATCTATTGAGCTTGTTAAACATTTTCGCTGCAC
RP5B        ATATTTGTGGAGTGTATTCAATGAAGAATCTATTGAGCTTGTTAAACATTTTCGTTGCAC
RP5C        ATATTTGTGGAGTGTATTCAATGAANAATCTATTGAGCTTGTTAAACATTTTCGTTGCAC
RP5D        ATATTTGTGGAGTGTATTCAATGAAGAATCTATTGAGCTTGTTAAACATTTTCGTTGCAC
RP6A        ATATTTGTGGAGTGTATTCAATGAAGAATCTATTGAGCTTGTTAAACATTTTCGTTGCAC
RP6B        ATATTTGTGGAGTGTATTCAATGAAGAATCTATTGAGCTTGTTAAACATTTTCGTTGCAC
RP6C        ATATTTGTGGAGTGTATTCAATGAAGAATCTATTGAGCTTGTTAAACATTTTCGTTGCAC
RP7A        ATATTTGTGGAGTGTATTCAATGAAGAATCTATTGAGCTTGTTAAACATTTTCGTTGCAC
RP7B        ATATTTGTGGAGTGTATTCGATGAAGAATCTATTGAGCTTGTTAAACATTTTCGTTGCAC
RP7C        ATATTTGTGGAGTGTATTCAATGAAGAATCTATTGAGCTTGTTAAACATTTTCGTTGCAC
RP8A        ATATTTGTGGAGTGTATTCAATGAAGAATCTATTGAGCTTGTTAAACATTTTCGTTGCAC
RP8B        ATATTTGTGGAGTGTATTCAATGAAGAATCTATTGAGCTTGTTAAACATTTTCGTTGCAC
RP9A        ATATTTGTGGAGTGTATTCAATGAAGAATCTATTGAGCTTGTTAAACATTTTCGTTGCAC
            ************ * ** *  ** * ***** *** *****   ********** *** *
```

```
AI1A            TTAGT-ATGCTTGGTGTATTCCAAGTATTAATTAAGTTAATGTTATTTTTAATAAAAAAA
AI1B            TTAGT-ATGCTTGGTGTATTCCAAGTATTAATTAAGTTAATGTTATTTTTAATAAAAAAA
AI1C            TTAGT-ATGCTTGGTGTATTCCAAGTATTAATTAAGTTAATGTTATTTTTAATAAAAAAA
AI2A            TTAKT-ATACTTGGTGTATTCCAAGTATTAATTAAGTTAATGTTATTTTTAATAAAAAAA
AI2B            TTAGT-ATACTTGGTGTATTCCAAGTATTAATTAAGTTAATGTTATTTTTAATAAAAAAA
AI3A            TTAGTTATACTTGGTGTATTCCAAGTATTAATTAAGTTAATGTTATTTTTAATAAAAAAA
AI3B            TTAGTTATACTTGGTGTATTCCAAGTATTAATTAAGTTAATGTTATTTTTAATAAAAAAA
AI3C            TTAGT-ATGCTTGGTGTATTCCAAGTATTAATTAAGTTAACGTTATTTTTAATAAAAAAA
AI4A            TTAGT-ATACTTGGTGTATTCCAAGTATTAATTAAGTTAATGTTATTTTTAATAAAAAAA
AI4B            TTAGT-ATACTTGGTGTATTCCAAGTATTAATTAAGTTAATGTTATTTTTAATAAAAAAA
AI4C            TTAGT-ATACTTGGTGTATTCCAAGTATTAATTAAGTTAATGTTATTTTTAATAAAAAAA
AI5A            TTAGT-ATACTTGGTGTATTCCAAGTATTAATTAAGTTAATGTTATTTTTAATAAAAAAA
AI5B            TTAGT-ATACTTGGTGTATTCCAAGTATTAA----GTTAATGTTATTTTTAATAAAAAAA
AI5C            TTAGT-ATACTTGGTGTATTCCAAGTATTAATTAAGTTAATGTTATTTTTAATAAAAAAA
AI6A            TTAGT-ATGCTTGGTGTATTCCAAGTATTAATTAAGTTAATGTTATTTTTAATAAAAAAA
AI6B            TTAGT-ATACTTGGTGTATTCCAAGTATTAATTAAGTTAATGTTATTTTTAATAAAAAAA
AI7A            TTAGT-ATACTTGGTGTATTCCAAGTATTAATTAAGTTAATGTTATTTTTAATAAAAAAA
AI7B            TTAGT-ATGCTTGGTGTATTCCAAGTATTAATTAAGTTAATGTTATTTTTAATAAAAAAA
AI7C            TTAGT-ATACTTGGTGTATTCCAAGTATTAATTAAGTTAATGTTATTTTTAATAAAAAAA
AI8A            TTAGT-ATGCTTGGTGTATTCCAAGTATTAATTAAGTTAATGTTATTTTTAATAAAAAAA
AI8B            TTAGT-ATACTTGGTGTATTCCAAGTATTAATTAANTTAATGTTATTTTNAATAAAAAAA
AI9A            TTACT-ATACTTGGTGTATTCCAAGTATTAATNANGTTAATGTTATTTTTAATAAAAAAA
AI10A           TTAGT-ATACTTGGTGTATTCCAAGTATTAATTAAGTTAATGTTATTTTTAATAAAAAAA
AI10B           TTAGT-ATACTTGGTGTATTCCAAGTATTAATTAAGTTAATGTTATTTTTAATAAAAAAA
AI10C           TTAGT-ATACTTGGTGTATTCCAAGTATTAATTAAGTTAATGTTATTTTTAATAAAAAAA
AI10D           TTAGT-ATACTTGGTGTATTCCAAGTATTAATTAAGTTAATGTTATTTTTAATAAAAAAA
AI11A           TTAGT-ATACTTGGTGTATTCCAAGTATTAATTAAGTTAATGTTATTTTTAATAAAAAAA
AI11B           TTAGT-ATACTTGGTGTATTCCAAGTATTAATTAAGTTAATGTTATTTTTAATAAAAAAA
AI11C           TTAGT-ATACTTGGTGTATTCCAAGTATTAATTAAGTTAATGTTATTTTTAATAAAAAAA
AI11D           TTAGT-ATACTTGGTGTATTCCAAGTATTAATTAAGTTAATGTTATTTTTAATAAAAAAA
AI11E           TTAGT-ATACTTGGTGTATTCCAAGTATTAATTAAGTTAATGTTATTTTTAATAAAAAAA
RP1A            TTAGT-ATGCTTGGTGTATTCCAAGTATTAATTAAGTTAATGTTATTTTTAATAAAGAAA
RP1B            TTAGT-ATACTTGGTGTATTCCAAGTATTAATTAAGTTAATGTTATTTTTAATAAAAAAA
RP2A            TTAGT-ATACTTGGTGTATTCCAAGTATTAATTAAGTTAATGTTATTTTTAATAAAAAAA
RP3A            TTAGT-ATACTTGGTGTATTCCAAGTATTAATTAAGTTAATGTTATTTCTAATAAAAAAA
RP3B            TTAGT-ATACTTGGTGTATTCCAAGTATTAATTAAGTTAATGTTATTTTTAATAAAAAAA
RP3C            TTAGT-ATGCTTGGTGTATTCCAAGTATTAATTAAGTTAATGTTATTTTTAATAAAAAAA
RP4A            TTAGT-ATACTTGGTGTATTCCAAGTATTAATTAAGTTAATGTTATTTTTAATAAAAAAA
RP4B            TTAGT-ATACTTGGTGTATTCCAAGTATTAATTAAGTTAATGTTATTTTTAATAAAAAAA
RP4C            TTAGT-ATACTTGGTGTATTCCAAGTATTAATTAAGTTAATGTTATTTTTAATAAAAAAA
RP5A            TTAGT-ATACTTGGTGTATTCCAAGTATTAATTAAGTTAATGTTATTTTTAATAAAAAAA
RP5B            TTAGT-ATGCTTGGTGTATTCCAAGTATTAATTAAGTTAATGTTATTTTTAATAAAAAAA
RP5C            CTAAT-ATGCTTGGTGTATTCCAAGTATTAATTAAGTTAATGTTATTTTTAATAAAAAAA
RP5D            TTAGT-ATACTTGGTGTATTCCAAGTATTAATTAAGTTAATGTTATTTTTAATAAAAAAA
RP6A            TTAGT-ATACTTGGTGTATTCCAAGTATTAATTAAGTTAATGTTATTTTTAATAAAAAAA
RP6B            TTAGT-ATACTTGGTGTATTCCAAGTATTAATTAAGTTAATGTTATTTTTAATAAAAAAA
RP6C            TTAGT-ATACTTGGTGTATTCCAAGTATTAATTAAGTTAATGTTATTTTTAATAAAAAAA
RP7A            TTAGT-ATACTTGGTGTATTCCAAGTATTAATTAAGTTAATGTTATTTTTAATAAAAAAA
RP7B            TTAGT-ATACTTGGTGTATTCCAAGTATTAATTAAGTTAATGTTATTTTTAATAAAAAAA
RP7C            TTAGT-ATACTTGGTGTATTCCAAGTATTAATTAAGTTAATGTTATTTTTAATAAAAAAA
RP8A            TTAGT-ATACTTGGTGTATTCCAAGTATTAATTAAGTTAATGTTATTTTTAATAAAAAAA
RP8B            TTAGT-ATACTTGGTGTATTCCAAGTATTAATTAAGTTAATGTTATTTTTAATAAAAAAA
RP9A            TTAGT-ATACTTGGTGTATTCCAAGTATTAATTAAGTTAATGTTATTTTTAATAAAAAAA
                 ** * **#**********************     **** *******  ****** ***
```

```
AI1A        TAACTAAGACATTCGCAACAAAAAAATGTTTAAAAAAAAAAA-TTTATTGAAGGAATTAA
AI1B        TAACTAAGACATTCGCAACAAAAAAATGTTTAAAAAAAAAAA-TTTATTGAAGGAATTAA
AI1C        TAACTAAGACATTCGCAACAAAAAAATGTTTAAAAAAAAAAA-TTTATTGAAGGAATTAA
AI2A        TAACTAAGACATTCGCAACAAAAAAATGTTTAAAAAAAAAA--TTTATTGAAGGAATTAA
AI2B        TAACTAAGACATTCGCAACAAAAAAATGTTTAAAAAAAAAA--TTTATTGAAGGAATTAA
AI3A        TAACTAAGACATTCGCAACAAAAAAATGTTTAAAAAAAAAA--TTTATTGAAGGAATTAA
AI3B        TAACTAAGACATTCGCAACAAAAAAATGTTTAAAAAAAAAA--TTTATTGAAGGAATTAA
AI3C        TAACTAAGACATTCGCAACAAAAAAATGTTTAAAAAAAAAAA-TTTATTGAAGGAATTAA
AI4A        TAACTAAGACATTCGCAACAAAAAAATGTTTAAAAAAAAAAA-ATTATTGAAGGAATTAA
AI4B        TAACTAAGACATTCGCAACAAAAAAATGTTTAAAAAAAAA---TTTATTGAAGGAATTGA
AI4C        TAACTAAGACATTCGCAACAAAAAAATGTTTAAAAAAAAAA--TTTATTGAAGGAATTAA
AI5A        TAACTAAGACATTCGCAACAAAAAAATGTTTAAAAAAAAAA--TTTATTGAAGGAATTAA
AI5B        TAACTAAGACACTCGCAACAAAAAAATGTTTAAAAAAAAAA--TTTATTGAAGGAATTAA
AI5C        TAACTAAGACATTCGCAACAAAAAAATGTTTAAAAAAAAAA--TTTATTGAAGGAATTAA
AI6A        TAACTAAGACATTCGCAACAAAAAAATGTTTAAAAAAAAAA--TTTATTGAAGGAATTAA
AI6B        TAACTAAGACATTCGCAACAAAAAAATGTTTAAAAAAAAAA--TTTATTGAAGGAATTAA
AI7A        TAACTAAGACATTCGCAACAAAAAAATGTTTAAAAAAAAAAA-TTTATTGAAGGAATTAA
AI7B        TAACTAAGACATTCGCAACAAAAAAATGTTTAAAAAAAAAAAATTTATTGAAGGAATTAA
AI7C        TAACTAAGACATTCGCAACAAAAAAATGTTTAAAAAAAAAAA-TTTATTGAAGGAATTAA
AI8A        TAACTAAGACATTCGCAACAAAAAAATGTTTAAAAAAAAAAA-TTTATTGAAGGAATTAA
AI8B        TAACTAAGACATTCNCAACAAAAAAATGTTTAAAAAAAAAA--TTTATTGAAGGAATTAN
AI9A        TAACTAAGACATTCGCAACAAAAAAATGTTTAAAAAAAAAA--TTTANTGAAGGAATTAA
AI10A       TAACTAAGACATTCGCAACAAAAAAATGTTTAAAAAAAAAA--TTTATTGAAGGAATTAA
AI10B       TAACTAAGACATTCGCAACAAAAAAATGTTTAAAAAAAAA---TTTATTGAAGGAATTAA
AI10C       TAACTAAGACATTCGCAACAAAAAAATGTTTAAAAAAAAAA--TTTATTGAAGGAATTAA
AI10D       TAACTAAGACATTCGCAACAAAAAAATGTT-AAAAAAAAAA--TTTATTGAAGGAATTAA
AI11A       TAACTAAGACATTCGCAACAAAAAAATGTTTAAAAAAAAAA--TTTATTGAAGGAATTAA
AI11B       TAACTAAGACATTCGCGACAAAAAAATGTTTAAAAAAAAAA--TTTATTGAAGGAATTAA
AI11C       TAACTAAGACATTCGCAACAAAAAAATGTTTAAAAAAAAAA--TTTATTGAAGGAATTAA
AI11D       TAACTAAGACATTCGCAACAAAAAAATGTTTAAAAAAAAAA--TTTATTGAAGGAATTAA
AI11E       TAACAAAGACATTCGCAACAAAAAAATGTTTAAAAAAAAAA--TTTATTGAAGGAATTAA
RP1A        TAACTAAGACATTCGCAACAAAAAAATGTTTAAAAAAAAAAA-TTTATTGAAGGAATTAA
RP1B        TAACTAAGACATTCGCAACAAAAAAATGTTTAAAAAAAAAAA-TTTATTGAAGGAATTAA
RP2A        TAACTAAGACATTCGCAACAAAAAAATGTTTAAAAAAAAAA--TTTATTGAAGGAATTAA
RP3A        TAACTAAGACATTCGCAACAAAAAAATGTTTAAAAAAAAAAA-TTTATTGAAGGAATTAA
RP3B        TAACTAAGACATTCGCAACAAAAAAATGTTTAAAAAAAAAA--TTTATTGAAGGAATTAA
RP3C        TAACTAAGACATTCGCAATAAAAAAATGTTTAAAAAAAAAAA-TTTATTGAAGGAATTAA
RP4A        TAACTAAGACATTCGCAACAAAAAAATGTTTAAAAAAAAAA--TTTATTGAAGGAATTAA
RP4B        TAACTAAGACATTCGCAACAAAAAAATGTTTAAAAAAAAAA--TTTATTGAAGGAATTAA
RP4C        TAACTAAGACATTCGCAACAAAAAAATGTTTAAAAAAAAAA--TTTATTGAAGGAATTAA
RP5A        TAACTAAGACATTCGCAACAAAAAAATGTTTAAAAAAAAA---TTTATTGAAGGAATTAA
RP5B        TAACTAAGACATTCGCAACAAAAAAATGTTTAAAAAAAAAAA-TTTATTGAAGGAATTAA
RP5C        TAACTAAGACATTCGCAACAAAAAAATGTTTAAAAAAAAAA--TTTATTGAAGGAATTAA
RP5D        TAACTAANACATTCGCAACAAAAAAATGTTTAAAAAAAAAAA-TTTATTGAAGGAATTAA
RP6A        TAACTAAGACATTCGCAACAAAAAAATGTTTAAAAAAAAAAA-TTTATTGAAGGAATTAA
RP6B        TAACTAAGACATTCGCAACAAAAAAATGTTTAAAAAAAAA---TTTATTGAAGGAATTAA
RP6C        TAACTAAGACATTCGCAACAAAAAAATGTTTAAAAAAAAAA--TTTATTGAAGGAATTAA
RP7A        TAACTAAGACATTCGCAACAAAAAAATGTTTAAAAAAAAAA--TTTATTGAAGGAATTAA
RP7B        TAACTAAGACATTCGCAACAAAAAAATGTTTAAAAAAAAA--A-TTTATTGAAGGAATTAA
RP7C        TAACTAAGACATTCGCAACAAAAAAATGTTTAAAAAAAAAA--TTTATTGAAGGAATTAA
RP8A        TAACTAAGACATTCGCAACAAAAAAATGTTTAAAAAAAAA-A-TTTATTGAAGGAATTAA
RP8B        TAACTAAGACATTCGCAACAAAAAAATGTTTAAAAAAAAAA--TTTATTGAAGGAATTAA
RP9A        TAACTAAGACATTCGCAACAAAAAAATGTTTAAAAAAAAAAAATTTATTGAAGGAATTAA
            **** ** *** ** * * ********** ********#### *** **********
```

```
AI1A        TTTATGCCAAATATTTATTAAAATTTTGGTGTATTTTTTAATTTCTTTCAATAATATAAA
AI1B        TTTATGCCAAATATTTATTAAAATTTTGGTGTATTTTTTAATTTCTTTCAATAATATAAA
AI1C        TTTATGCCAAATATTTATTAAAATTTTGGTGTATTTTTTAATTTCTTTCAATAATATAAA
AI2A        TTTATGCCAAATATTTATTAAAATTTTGGTGTATTTTTTAATTTCTTTCAATAATATAAA
AI2B        TTTATGCCAAATATTTATTAAAATTTTGGTGTATTTTTTAATTTCTTTCAATAATATAAA
AI3A        TTTATGCCAAATATTTATTAAAATTTTGGTGTATTTTTTAATTTCTTTCAATAATATAAA
AI3B        TTTATGCCAAATATTTATTAAAATTTTGGTGTATTTTTTAATTTCTTTCAATAATATAAA
AI3C        TTTATGCCAAATATTTATTAAAATTTTGGTGTATTTTTTAATTTCTTTCAATAATATAAA
AI4A        TTTATGCCAAATATTTATTAAAATTTTGGTGTATTTTTTAATTTCTTTCAATAATATAAA
AI4B        TTTATGCCAAATATTTATTAAAATTTTGGTGTATTTTTTAATTTCTTTCAATAATATAAA
AI4C        TTTATGCCAAATATTTATTAAAATTTTGGTGTATTTTTTAATTTCTTTCAATAATATAAA
AI5A        TTTATGCCAAATATTTATTAAAATTTTGGTGTATTTTTTAATTTCTTTCAATAATATAAA
AI5B        TTTATGCCAAATATTTATTAAAATTTTGGTGTATTTTTTAATTTCTTTCAATAATATAAA
AI5C        TTTATGCCAAATATTTATTAAAATTTTGGTGTATTTTTTAATTTCTTTCAATAATATAAA
AI6A        TTTATGCCAAATATTTATTAAAATTTTGGTGTATTTTTTAATTTCTTTCAATAATATAAA
AI6B        TTTATGCCAAATATTTATTAAAATTTTGGTGTATTTTTTAATTTCTTTCAATAATATAAA
AI7A        TTTATGCCAAATATTTATTAAAATTTTGGTGTATTTTTTAATTTCTTTCAATAATATAAA
AI7B        TTTATGCCAAATATTTATTAAAATTTTGGTGTATTTTTTAATTTCTTTCAATAATATAAA
AI7C        TTTATGCCAAATATTTATTAAAATTTTGGTGTATTTTTTAATTTCTTTCAATAATATAAA
AI8A        TTTATGCCAAATATTTATTAAAATTTTGGTGTATTTTTTAATTTCTTTCAATAATATAAA
AI8B        TTTATGCCNAATANTTATTAAAATTTTGGTGTATTTTTTA-TTTCTTTCNNTANTANAAN
AI9A        TTTATGCCNAATATTTATTAAAATTNTGGTGTANTTTTTAATTTCNTTCAN-AATANANA
AI10A       TTTATGCCAAATATTTATTAAAATTTTGGTGTATTTTTTAATTTCTTTCAANAATATAAA
AI10B       TTTATGCCAAATATTTATTAAAATTTTGGTGTATTTTTTAATTTCTTTCAATAATATAAA
AI10C       TTTATGCCAAATATTTATTAAAATTTTGGTGTATTTTTTAATTTCTTTCAATAATATAAA
AI10D       TTTATGCCAAATATTTATTAAAATTTTGGTGTATTTTTTAATTTCTTTCAATAATATAAA
AI11A       TTTATGCCAAATATTTATTAAAATTTTGGTGTATTTTTTAATTTCTTTCAATAATATAAA
AI11B       TTTATGCCAAATATTTATTAAAATTTTGGTGTATTTTTTAATTTCTTTCAATAATATAAA
AI11C       TTTATGCCAAATATTTATTAAAATTTTGGTGTATTTTTTAATTTCTTTCAATAATATAAA
AI11D       TTTATGCCAAATATTTATTAAAATTTTGGTGTATTTTTTAATTTCTTTCAATAATATAAA
AI11E       TTTATGCCAAATATTTATTAAAATTTTGGTGTATTTTTTAATTTCTTTCAATAATATAAA
RP1A        TTTATGCCAAATATTTATTAAAATTTTGGTGTATTTTTTAATTTCTTTCAATAATATAAA
RP1B        TTTATGCCAAATATTTATTAAAATTTTGGTGTATTTTTTAATTTCTTTCAATAATATAAA
RP2A        TTTATGCCAAATATTTATTAAAATTTTCGTGTATTTTTTAATTTCTTTCAATAATATAAA
RP3A        TTTATGCCAAATATTTATTAAAATTTTGGTGTATTTTTTAATTTCTTTCAATAATATAAA
RP3B        TTTATGCCAAATATTTATTAAAATTTTGGTGTATTTTTTAATTTCTTTCAATAATATAAA
RP3C        TTTATGCCAAATATTTATTAAAATTTTGGTGTATTTTTTAATTTCTTTCAATAATATAAA
RP4A        TTTATGCCAAATATTTATTAAAATTTTGGTGTATTTTTTAATTTCTTTCAATAATATAAA
RP4B        TTTATGCCAAATATTTATTAAAATTTTGGTGTATTTTTTAATTTCTTTCAATAATATAAA
RP4C        TTTATGCCAAATATTTATTAAAATTTTGGTGTATTTTTTAATTTCTTTCAATAATATAAA
RP5A        TTTATGCCAAATATTTATTAAAATTTTGGTGTGTTTTTTAATTTCTTTCAATAATATAAA
RP5B        TTTATGCCAAATATTTATTAAAATTTTGGTGTATTTTTTAATTTCTTTCAATAATATAAA
RP5C        TTTATGCCNNATATTTATTAAAATTTTGGTGTATTTTTTAATTTCTTTCAATAATATAAA
RP5D        TTTATGCCAAATATTTATTAAAATTTTGGTGTATTTTTTAATTTCTTTCAATAATATAAA
RP6A        TTTATGCCAAATATTTATTAAAATTTTGGTGTATTTTTTAATTTCTTTCAATAATATAAA
RP6B        TTTATGCCAAATATTTATTAAAATTTTGGTGTATTTTTTAATTTCTTTCAATAATATAAA
RP6C        TTTATGCCAAATATTTATTAAAATTTTGGTGTATTTTTTAATTTCTTTCAATAATATAAA
RP7A        TTTATGCCAAATATTTATTAAAATTTTGGCGTATTTTTTAATTTCTTTCAATAATATAAA
RP7B        TTTATGCCAAATATTTATTAAAATTTTGGTGTATTTTTTAATTTCTTTCAATAATATAAA
RP7C        TTTATGCCAAATATTTATTAAAATTTTAGTGTATTTTTTAATTTCTTTCAATAATATAAA
RP8A        TTTATGCCAAAT----ATTGAAATTTTGGTGTATTTTTTAATTTCTTTCAATAATATAAA
RP8B        TTTATGCCAAATATTTATTAAAATTTTGGTGTATTTTTTAATTTCTTTCAANAATATAAA
RP9A        TTTATGCCAAATATTTATTAAAATTTTGGTGTATTTTTTAATTTCTTTCAATAATATAAA
            ********  **    *** ***** * * **  ****** **** ***   * ** *
```

```
AI1A        AATTGACATTAATTAAAAACAAAAAACAAAAATATCACTCTAAGCGGTGGATCACTTGGC
AI1B        AATTGACATTAATTAAAAACAAAAAACAAAAATATCACTCTAAGCGGTGGATCACTTGGC
AI1C        AATTGACATTAATTAAAAACAAAAAACAAAAATATCACTCTAAGCGGTGGATCACTTGGC
AI2A        AATTGACATTAATTAAAAACAAAAAACAAAAATATCACTCTAAGCGGTGGATCACTTGGC
AI2B        AATTGACATTAATTAAAAACAAAAAACAAAAATATCACTMTAAGCGGTGGATCACTTGGC
AI3A        AATTGACATTAATTAAAAACAAAAAACAAAAATATCACTCTAAGCGGTGGATCACTTGGC
AI3B        AATTGACATTAATTAAAAACAAAAAACAAAAATATCACTCTAAGCGGTGGATCACTTGGC
AI3C        AATTGACATTAATTAAAAACAAAAAACAAAAATATCACTCTAAGCGGTGGATCACTTGGC
AI4A        AATTGACATTAATTAAAAACAAAAAACAAAAATATCACTCTAAGCGGTGGATCACTTGGC
AI4B        AATTGACATTAATTAAAAACAAAAAACAAAAATATCACTCTAAGCGGTGGATCACTTGGC
AI4C        AATTGACATTAATTAAAAACAAAAAACAAAAATATCACTCTAAGCGGTGGATCACTTGGC
AI5A        AATTGACATTAATTAAAAACAAAAAACAAAAATATCACTCTAAGCGGTGGATCACTTGGC
AI5B        AATTGACATTAATTAAAAACAAAAAACAAAAATATCACTCTAAGCGGTGGATCACTTGGC
AI5C        AATTGACATTAATTAAAAACAAAAAACAAAAATATCACTCTAAGCGGTGGATCACTTGGC
AI6A        AATTGACATTAATTAAAAACAAAAAACAAAAATATCACTWTAAGCGGTGGATCACTTGGC
AI6B        AATTGACATTAATTAAAAACAAAAAACAAAAATATCACTCTAAGCGGTGGATCACTTGGC
AI7A        AATTGACATTAATTAGAAACAAAAAACAAAAATATCACTCTAAGCGGTGGATCACTTGGC
AI7B        AATTGACATTAATTAAAAACAAAAAACAAAAATATCACTCTAAGCGGTGGATCACTTGGC
AI7C        AATTGACATTTATTAAAAACAAAAAACAAAAATATCACTCTAAGCGGTGGATCACTTGGC
AI8A        AATTGACATTAACTAAAAACAAAAAACAAAAATATCACTCTAAGCGGTGGATCACTTGGC
AI8B        AATTGACATTAATTCAGANCAAAAANCAAAAATATCNCNCTNANCGGNGGATCNCTTGNC
AI9A        AATTGACATTAATTACAN-CACAAAACANANATATCACTCNAAGCGGTGGATCACTTGGC
AI10A       AATTGACATAAATTAAAAACAAAAAACAAAAATATCACTCTAANCGGNGGATCACTTGGC
AI10B       AATTGACATTAATTAAAAACAAAAAACAAAAATATCACTCTAAGCGGTGGATCACTTGGC
AI10C       AATTGACATTAATTAAAAACAAAAAACAAAAATATCACTCTAAGCGGTGGATCACTTGGC
AI10D       AATTGACATTAATTAAAAACAAAAAACAAAAATATCACTCTAAGCGGTGGATCACTTGGC
AI11A       AATTGACATTAATTAAAAACAAAAAACAAAAATATCACTCTAAGCGGTGGATCACTTGGC
AI11B       AATTGACATTAATTAAAAACAAAAAACAAAAATATCACTCTAAGCGGTGGATCACTTGGC
AI11C       AATTGACATTAATTAAAAACAAAAAACAAAAATATCACTCTAAGCGGTGGATCACTTGGC
AI11D       AATTGACATTAATTAAAAACAAAAAACAAAAATATCACTCTAAGCGGTGGATCACTTGGC
AI11E       AATTGACATTAATTAAAAACAAAAAACAAAAATATCACTCTAAGCGGTGGATCACTTGGC
RP1A        AATTGACATTAATTAAAAACAAAAAACAAAAATATCACTCTAAGCGGTGGATCACTTGGC
RP1B        AATTGACATTAATTAAAAACAAAAAACAAAAATATCACTCTAAGCGGTGGATCACTTGGC
RP2A        AATCGACATTAATTAAAAACAAAAAACAAAAATATCACTYTAAGCGGTGGATCACTTGGC
RP3A        AATTGACATTAATTAAAAACAAAAAACAAAAATATCACTCTAAGCGGTGGATCACTTGGC
RP3B        AATTGACATTAATTAAAAACAAAAAACAAAAATATCACTCTAAGCGGTGGATCACTTGGC
RP3C        AATTGACATTAATTAAAAACAAAAAACAAAAATATCACTCTAAGCGGTGGATCACTTGGC
RP4A        AATTGACATTAATTAAAAACAAAAAACAAAAATATCACTCTAAGCGGTGGATCACTTGGC
RP4B        AATTGACATTAATTAAAAACAAAAAACAAAAATATCACTCTAAGCGGTGGATCACTTGGC
RP4C        AATTTACATTAATTAAAAACAAAAAACAAAAATATCACTCTAAGCGGTGGATCACTTGGC
RP5A        AATTGACATTAATTAAAAACAAAGAACAAAAATATCACTCTAAGCGGTGGATCACTTGGC
RP5B        AATTGACATTAATTAAAAACAAAAAACAAAAATATCACTCTAAGCGGTGGATCACTTGGC
RP5C        AATTGACATTAATTAAAAACAAAAAACAAAAATATCACTCTAAGCGGTGGATCACTTGGC
RP5D        AATTGACATTAATTAAAAACAAAAAACAAAAATATCACTCTAAGCGGTGGATCACTTGGC
RP6A        AATTGACATTAATTAAAAACAAAAAACAAAAATATCGCTCTAAGCGGTGGATCACTTGGC
RP6B        AATTGACATTAATTAAAAACAAAAAACAAAAATATCACTCTAAGCGGTGGATCACTTGGC
RP6C        AATTGACATTAATTAAAAACAAAAAACAAAAATATCACTCTAAGCGGTGGATCACTTGGC
RP7A        AATTGACATTAATTAAAAACAAAAAACAAAAATATCACTCTAAGCGGTGGATCACTTGGC
RP7B        AATTGACATTAATTAAAAACAAAAAACAAAAATATCACTCTAAGCGGTGGACCACTTGG-
RP7C        AATTGACATTAATTAAAAACAAAAAACAAAAATATCACTCTAAGCGGTGGATCACTTGGC
RP8A        AATTGACATTAATTAAAAACAAAAAACAAAAATATCACTCTAAGCGGTGGATCACTTGGC
RP8B        AATTGACATAAATTAAAAACAAAAAACAAAAATATCACTCTAAGCGGGGGATCACTTGGC
RP9A        AATTGACATTAATTAAAAACAAAAAACAAAAATATCACTCTAAGCGGTGGATCACTTGGC
            ***  ****  * *     ** * * ** * ***** *    * *** *** * ****
```

```
AI1A            TCATGG-------
AI1B            TCATGG-------
AI1C            TCATGG-------
AI2A            TCATGG-------
AI2B            TCATGG-------
AI3A            TCATGG-------
AI3B            TCATGGGTC----
AI3C            TCATGG-------
AI4A            TCATGG-------
AI4B            TCATGG-------
AI4C            TCATGG-------
AI5A            TCATGGGTCGAT-
AI5B            TCATGGGTCGATG
AI5C            TCATGG-------
AI6A            TCATGG-------
AI6B            TCATGG-------
AI7A            TCATGG-------
AI7B            TCATGG-------
AI7C            TCATGG-------
AI8A            TCATGG-------
AI8B            TCATGG-------
AI9A            TCATGG-------
AI10A           TCATGG-------
AI10B           TCATGGGTCGATG
AI10C           TCATGG-------
AI10D           TCATGG-------
AI11A           TCATGGGTCGATG
AI11B           TCATGG-------
AI11C           TCATGG-------
AI11D           TCATGGGTCGATG
AI11E           TCATGGGTCGATG
RP1A            TCATGG-------
RP1B            TCATGG-------
RP2A            TCATGG-------
RP3A            TCATGG-------
RP3B            TCATGG-------
RP3C            TCATGG-------
RP4A            TCATGG-------
RP4B            TCATGG-------
RP4C            TCATGG-------
RP5A            TCATGG-------
RP5B            TCATGG-------
RP5C            TCATGG-------
RP5D            TCATGG-------
RP6A            TCATGG-------
RP6B            TCATGG-------
RP6C            TCATGG-------
RP7A            TCATGG-------
RP7B            -------------
RP7C            TCATGG-------
RP8A            TCATGG-------
RP8B            TCATGG-------
RP9A            TCATGGGTCGATG
```