GENETIC VARIATION IN THE *GILA ROBUSTA* COMPLEX (TELEOSTEI: CYPRINIDAE) IN THE LOWER COLORADO RIVER

by

Michael R. Schwemm

A Thesis Presented in Partial Fulfillment
of the Requirements for the Degree
Master of Science

ARIZONA STATE UNIVERSITY

December 2006

ABSTRACT

Mitochondrial (ND2) and nuclear (Tpi-B, S7) DNA sequences were used to investigate the relative influences of species morphology, geographic isolation and recent hybridization on patterns of genetic variation in imperiled species of cyprinid fishes of the taxonomically complex genus *Gila* in the lower Colorado River basin. Despite consistent morphological differences among the three species (*G. robusta, G. nigra* and *G. intermedia*), low levels of genetic divergence exist among alleles at all loci examined. Genetic variation was not partitioned according to morphology or by hydrography; however, there were significant differences among subpopulations within drainages or species for each marker, evidenced by significant frequency differences (Tpi-B) or fixation for private alleles (ND2, S7). The results of nested clade analysis and analysis of molecular variance consistently indicate that fragmentation of individual subpopulations has occurred throughout the Gila and Bill Williams rivers, leading to individuality of many subpopulations. Tests for association of nuclear and mtDNA provided no evidence of recent hybridization; however, ancient admixture of these populations and species is not inconsistent with observed patterns of genetic variation. Patterns of morphological and genetic variation identify a complex history for these vulnerable species, making conservation of these forms difficult. Management objectives and implementation must strive to maintain individuality of distinct subpopulations under a single management plan that considers the entire species complex.

ACKNOWLEDGEMENTS

I would like to thank my major advisor, T. E. Dowling, for persistent encouragement throughout the progress and completion of this thesis. His patience and efforts were greatly appreciated. I also my thank committee members P. C. Marsh and S. Kumar for their time and critical review of this work. I am particularly grateful to P. C. Marsh for useful suggestions regarding planning and implementation of this project. Additionally, I thank A. T. Kelsen for technical advice in the laboratory. Funding for this project was made possible through grants from U. S. Bureau of Reclamation, U. S. Forest Service, and the New Mexico Game and Fish Department.

TABLE OF CONTENTS

Page

LIST OF TABLES……...……...v

LIST OF FIGURES……..……..vi

INTRODUCTION…….....1

MATERIALS AND METHODS……5

RESULTS AND DISCUSSION……11

LITERATURE CITED……..37

APPENDIX 1. Localities, Locality Abbreviations, Taxonomic Identity and Number of Individuals Sampled per Locality…….……...…94

APPENDIX 2. List of Haplotype and Allele Sequences for ND2, Tpi-B, and S7…….....97

LIST OF TABLES

Table Page

1. Primers and amplification parameters used for SSCP and sequencing reactions…..55
2. Geographic distribution of 39 haplotypes from the mitochondrial gene ND2……..56
3. Geographic distribution of 10 genotypes (a) and alleles (b) from the triosephosphate isomerase-B gene……..60
4. Geographic distribution of 25 genotypes (a) and alleles (b) from the S7 ribosomal gene……....63
5. MtDNA genetic diversity within 32 populations of *Gila* …….……..……..….69
6. Tpi-B genetic diversity within 32 populations of *Gila*……..70
7. S7 genetic diversity within 32 populations of *Gila*…….…..71
8. Significant associations of allelic and genotypic cytonuclear tests for mitochondrial ND2 with two nuclear loci Tpi-B and the S7 ribosomal protein and deviations from Hardy-Weinberg (H-W) expectations.....…….72
9. Analysis of molecular variance (AMOVA) for ND2, Tpi-B and S7 loci……..…73
10. Results and description of nested clade analysis (NCA) of ND2 haplotypes……...74
11. Results and description of nested clade analysis (NCA) of Tpi-B alleles……...….76
12. Results and description of nested clade analysis (NCA) of S7 alleles…….77
13. Pairwise F_{ST} values for 32 populations of *Gila* from ND2 sequences…….…….78
14. Pairwise F_{ST} values for 32 populations of *Gila* from Tpi-B sequences...…….83
15. Pairwise F_{ST} values for 32 populations of *Gila* from S7 sequences…….88
16. Management units of *Gila* populations in the lower Colorado River basin…….93

LIST OF FIGURES

Figure Page

1. Historic and current records of the *G. robusta* complex in the lower Colorado River, including populations from the Bill Williams and Gila drainages ….……....44
2. Map of 32 sampled localities of the *Gila robusta* complex of the lower Colorado River basin, AZ and NM……45
3. Maximum likelihood tree generated from mtDNA sequences. ……....46
4. Bootstrap consensus tree (1000 replicates) of 14,688 MP trees generated from mtDNA sequences ……....47
5. Maximum likelihood (a) and bootstrap consensus of maximum parsimony trees (b) generated from Tpi-B sequences …..…...…….48
6. Maximum likelihood (a) and bootstrap consensus of maximum parsimony (b) trees for S7 sequences ……...49
7. Statistical parsimony network and nesting design of ND2 haplotypes.....…...50
8. Statistical parsimony networks and nesting design of Tpi-B (a) and S7 (b) alleles..51
9. Population tree of lower Colorado River *Gila* constructed by F_{ST} estimates of sequence divergence at ND2 using the neighbor-joining method…….52
10. Population tree of lower Colorado River *Gila* constructed by F_{ST} estimates of sequence divergence at Tpi-B using the neighbor-joining method …….....53
11. Population tree of lower Colorado River *Gila* constructed by F_{ST} estimates of sequence divergence at S7 using the neighbor-joining method.....…...54

INTRODUCTION

Barriers to gene flow subdivide populations, providing opportunity for divergence and the evolution of reproductive isolation. In aquatic systems, previously well-connected populations often become isolated at both historical and recent time scales (Avise 2000). Fluctuations in amounts of surface water and resulting impacts on connectivity make arid environments unique opportunities to examine the impacts of fragmentation in aquatic systems (e.g., Dowling et al. 1996; Tibbets and Dowling 1996; Hughes et al. 2004; Hurt and Hedrick 2004). Additionally, understanding the evolutionary history of a species is required for informed management of declining taxa (reviewed in Soulé and Wilcox 1980; Vrijenhoek 1996; Avise 2004) as describing patterns of genetic variation and evolutionary mechanisms shaping them is necessary to maintain existing diversity and evolutionary potential (Avise 2004).

The lower Colorado River drainage of the southwestern United States provides an excellent opportunity to evaluate the relative importance of multiple historical influences in governing the distribution of genetic variation and phylogeographic relationships of its many endemic fishes. This drainage is typified by diverse aquatic habitats, ranging from high gradient mountain streams to slow pooling marshes, which eventually feed the same large rivers. In addition to a diversity of aquatic habitats, this region has also experienced fluctuations in connectivity among habitat patches, and as a whole, has become increasingly discontinuous since the Tertiary (Axelrod 1979). Obligate aquatic fauna are necessarily impacted through vicariant hydrogeographic events (Minckley et al. 1986), and many fishes from this region are endemic and exhibit high levels of phenotypic

variation (e.g., DeMarais 1986, 1992; Minckley et al. 1986; Tibbets and Dowling 1996; Douglas et al. 1999; Gerber et al. 2001). The relatively well understood geological history of the region, coupled with increased isolation due to aridification, has made these fishes ideal natural systems to explore mechanisms of evolution.

Seven species of chub inhabit the Colorado River basin, including *G. cypha, G. elegans, G. intermedia, G. jordani, G. nigra, G. robusta and G. seminuda* (Minckley and DeMarais 2000). All species are threatened by the introduction of predatory non-native fishes and loss of habitat due to re-routing of waterways (Minckley and Deacon 1991). To date, *G. intermedia, G. seminuda, G. jordani, G. cypha, and G. elegans* are listed as endangered species by the US Fish and Wildlife Service and *G. nigra* is a candidate species (Fed. Reg. 71(89), 26007-260017). *Gila robusta* remains the only species not listed or proposed for listing as endangered.

Three of these species, *G. intermedia*, *G. nigra*, and *G. robusta* (hereafter the *robusta* complex), form a phylogenetically unresolved complex. Here I focus on the *robusta* complex occurring within the Gila and Bill Williams rivers of the lower Colorado River basin. *Gila robusta* was once widespread and found throughout the entire Colorado River basin. In the lower Colorado River, this species was historically found in small to large streams of the Gila and Bill Williams drainages (Rinne 1969), but at present is restricted to several disjunct populations of small order and mainstream reaches of these rivers (Minckley and DeMarais 2000) (Figure 1). *Gila intermedia* and *G. nigra* only occur in the Gila and Bill Williams rivers systems, with *G. intermedia* more often found in cienega-like habitats while *G. nigra* occurs in headwater streams (Minckley and

DeMarais 2000; Voeltz 2002). Like most other fishes of the lower Colorado River basin, these chubs have been negatively impacted by human activities and have been the target of conservation efforts. Knowledge of patterns of genetic variation and evolutionary history are critical for informed management of these imperiled forms.

The *robusta* complex represents a spectrum of morphological diversity from slender and elongate (*G. robusta*) to deep-bodied and chunky in aspect (*G. intermedia)* (Rinne 1969, 1976; DeMarais 1986). *Gila robusta* and *G. intermedia* represent relative morphological extremes, while *G. nigra* (formerly *G. robusta grahami* prior to Minckley and DeMarais 2000) is intermediate in morphometric and meristic traits (Rinne 1969, 1976; DeMarais 1986; Minckley and DeMarais 2000). Accounts of hybridization among externally fertilized North American fishes (Hubbs 1955; Smith 1966; Minckley et al. 1986) and apparent morphological intermediacy of *G. nigra* led DeMarais (1986) and Minckley and DeMarais (2000) to suggest hybridization as a possible mechanism contributing to patterns of variation in *Gila*. These three taxa rarely co-occur but are distributed in a mosaic fashion throughout most tributaries of the Gila River basin (Figure 1). The apparent mosaic nature of this distribution may be partly due to difficulties discriminating among members of the *robusta* complex, with current taxonomic identification based largely on morphometric population means and geographic location. Difficulties diagnosing individuals complicates management as conservation authorities often rely on measures of population distinctiveness to justify and guide their actions.

These taxa have been studied extensively, yielding considerable preliminary data necessary for investigation of evolutionary processes (Rinne 1969, 1976; DeMarais 1986,

1992; Douglas et al. 1999; Minckley and DeMarais 2000; Gerber et al. 2001). Previous studies of these taxa indicated that specimens from the lower Colorado River are genetically distinct from *Gila* in other regions (Gerber et al. 2001); however, further subdivision of the *robusta* complex has not been supported by molecular data. In a study of allozymic variation, DeMarais (1992) failed to find diagnostic alleles that allowed discrimination among species, but some populations possessed unique alleles at high frequencies. Moreover, populations classified as *G. intermedia* showed much higher levels of variation than either *G. robusta* or *G. nigra*. However, his study did not reveal patterns consistent with morphology or concordant with geography, but instead identified populations of dissimilar morphology with similar allozymes. Consequently, the evolutionary history of lower Colorado River *Gila* remains unresolved and affords excellent opportunity for evaluating evolutionary processes of a closely related species group.

This study further investigates the evolutionary history of the *robusta* complex using mtDNA and nuclear sequence data, focusing on the distinct group from the lower Colorado River basin. The relatively rapid rate of evolution, lack of recombination, and four-fold reduction in effective population size makes mtDNA an ideal marker for the study of phylogeographic and population structure. Nuclear markers were used to provide additional markers of population structure, plus evaluate the potential for admixture among taxa (reviewed in Avise 2000). Here I employ phylogenetic and population genetic methods to characterize patterns of variation within and among populations, allowing for assessment of the role of historical and evolutionary factors in shaping

observed patterns of variation within the *G. robusta* complex. Not only is this information valuable for understanding the process of evolution, but it is critical for conservation and management of these imperiled forms.

MATERIALS and METHODS

Sampling and DNA extraction

Sampling of the Gila and Bill Williams river drainages encompassed 32 sites in seven subbasins in Arizona and New Mexico (i.e., Verde River, Salt River, San Pedro River, Santa Cruz River, Agua Fria River, Gila River mainstem and Bill Williams River - Figure. 2). Morphological classification and species status follow designations presented in Minckley and DeMarais (2000). Eleven to thirty individuals were collected from each locality (Appendix 1). Fish were captured using a combination of seining, hoop nets and electrofishing. Muscle or fin clips were removed and immediately preserved in 95% ethanol, or frozen in the field with dry ice. Frozen samples were either retained frozen (-20° C or -80° C) or transferred to 95% ethanol in the laboratory.

Characterization of molecular markers

Populations were screened for sequence variation at one mitochondrial DNA (mtDNA) and two nuclear (nDNA) gene loci. The mtDNA gene was represented by two pieces of NADH dehydrogenase subunit 2 (ND2, Gerber et al. 2001). Genomic DNA was extracted from muscle or fin clips by standard proteinase K/phenol/chloroform protocol as modified by Tibbets and Dowling (1996) and resuspended in 20-100 μL double-

distilled water. The first fragment (400 bp) was located in the middle of the gene and amplified using primers AA_{GILA} and C_{GILA} (Gerber et al. 2001) while the second fragment (388 bp) was located at the 3' end and amplified using primers F_{ROB} (Table 1) and E_{GILA} (Gerber et al. 2001). Each fragment was scored separately and concatenated later for statistical analysis. Nuclear genes were represented by exon 3 and introns 2 and 3 of triosephosphate isomerase-B (Tpi-B, 328 bp) and intron 1 of the S7 ribosomal protein (358 bp). Tpi-B was amplified using primers TPIB4-F2 (Merritt and Quattro 2001) and TPIB-R3 (Johnson et al. 2004). Characterization of the S7 fragment required nested reactions. The initial product was generated with primers S7RPEX1F and S7RPEX2R (Chow and Hazama 1998), diluted 1:100, and used as template for the SSCP reaction using S7RPEX1F and internal $S72R_{GILA}$ (designed for this study). Variation in molecular markers was characterized through analysis of single-stranded conformational polymorphisms (SSCPs) as described previously (Gerber et al. 2001; Dowling et al. 2005). PCR temperatures and cycle parameters for each target fragment are listed in Table 1. Final concentrations for PCR reagents were identical for both mitochondrial and nuclear fragments in 5 µL reaction volumes containing: 1X Taq buffer, 0.2 µM dNTPs, 0.05 µM α^{32}P-labeled dATPs, 2.5 mM MgCl, 0.5µM per primer and 1.25 units Taq polymerase (purified from a clone and used following conditions described in Pluthero (1993). Following PCR, 4 µL of stop solution (95% formamide, 0.05% bromophenol blue, 0.05% xylene cyanol, and 0.5M EDTA) were added to PCR products. These were denatured at 94°C for 5 min and held on ice briefly until use. Each product (3 µL) was loaded on a 6% acrylamide gel (37.5:1 acrylamide:bis-acrylamide, 1X TBE) and the

remainder stored frozen (-20°C) for subsequent analysis. Single-stranded PCR products were subjected to electrophoresis at room temperature at 8 W for 18 h. After electrophoresis, gels were dried, exposed to X-ray film for 12-24 h and developed with an automatic film developer. SSCP variants on each gel were scored, rearranged by mobility, and stored samples electrophoresed again to confirm mobility patterns both within and among gels. Alleles and genotypes were designated by letter in order of discovery. Presumptive nuclear and mitochondrial mobility groups were provided a multi-letter designation corresponding to genotype or composite haplotype.

For Tpi-B, SSCP failed to discriminate among two alleles that differed at one nucleotide position. Examination of these sequences indicated that these alleles could be differentiated using the restriction enzyme *Tsp45I*, resulting in one full length fragment (366 bp) or two pieces (307 bp and 59 bp). Restriction digestion was performed using 5µL of PCR product and 0.5 units of enzyme following manufacturer's recommendations. Digests were incubated at 65° for 3.5 h and fragments were separated by electrophoresis through 1.5% agarose gels prestained with ethidium bromide.

Sequencing

At least one representative of each presumptive SSCP variant was amplified for sequencing using the same PCR parameters as SSCP analysis but scaled to 50 µL reaction volumes. Mitochondrial fragments were amplified for sequencing using additional upstream primers (ND2-AA_{robseq} or ND2-F_{robseq}, Table 1) designed to generate readable sequence for the entire SSCP fragment. Nuclear loci were sequenced in both

directions to obtain sequence for the entire fragment and verify heterozygotes. PCR products were cleaned for sequencing in either of two ways, by centrifugation using Ultrafree-MC® filters (according to manufacturer's specifications), or by using Exonuclease I, shrimp alkaline phosphatase, and dilution buffer (EXOSAP, USB Corporation) using the following temperature profile (modified from the manufacturer): 37° C for 30 min, 80° C for 15 min and 12° C for 5 min. DNA sequences were obtained with ABI models 377 or 3730 sequencers with BigDye® kit version 3.1 (ABI Perkin-Elmer). Sequences were aligned manually using MEGA3.1 (Kumar et al. 2004) and assigned to allelic class by parsimony analysis with PAUP* 4.0 ver. 10b (Swofford 1998) as described below. When a representative sequence matched a known haplotype, all members of that SSCP group were scored to reflect this designation. New haplotypes were designated alphabetically in order of discovery, not mobility, and added to the template file.

Phylogenetic and Population Genetic Methods

Phylogenetic analysis.--Genetic diversity was initially examined using a phylogenetic approach before more fine-scale analyses were performed. Relatedness among mtDNA haplotypes was characterized using maximum-parsimony (MP) and maximum-likelihood (ML) methods as implemented in PAUP* ver. 4.10b (Swofford 1998). Parsimony analyses were completed by heuristic search using random stepwise-addition (50 replicates) and TBR branch swapping. Maximum likelihood analysis was performed using the HKY+G (Hasegawa et. al 1985) model of nucleotide substitution with gamma distribution ($\alpha = 0.434$) selected by Modeltest 3.7 (Posada and Crandall

1998). Topologies from both likelihood and parsimony methods were evaluated by bootstrapping (1000 replicates).

Population genetic analyses.--For each locality, standard population genetic parameters were estimated using Arlequin 2.0 (Schneider et al. 2000). The number of haplotypes (N_h) plus haplotype (*h*, Nei 1987) and nucleotide (π, Tajima 1983) diversities and their variances were used to compare populations for all loci. Nuclear markers (Tpi-B, S7) were additionally tested for departures from Hardy-Weinberg proportions, evaluated by Markov-chain exact tests (Guo and Thompson 1992) with statistical significance corrected for multiple loci using the standard Bonferroni technique (described in Weir 1996).

Analysis of molecular variance (AMOVA, Excoffier et al. 1992) was used to partition genetic variance within and among populations to test two alternative hypotheses that could explain patterns of genetic variation in the *robusta* complex. In the first approach, genetic variation was partitioned by morphological (species) designation (Minckley and DeMarais 2000). Alternatively, I tested to see if geography was a better predictor of the distribution of genetic variation. F-statistics and their significance were estimated by a non-parametric permutation approach (Excoffier et al. 1992) implemented in the software package Arlequin 2.0 (Schneider et al. 2000).

Pairwise F_{ST} values were generated to provide estimates of short term genetic divergence among populations for both mtDNA and nuclear DNA. Matrices of pairwise F_{ST} values were compared across loci by Mantel tests (Mantel 1967, Smouse et al. 1986). Tests of partial correlations among matrices employ a nonparametric permutation

procedure (1000 replicates) with significance levels corrected using the Bonferroni technique for the number of matrices compared. Additionally, matrices of pairwise F_{ST} were generated from sequence data and used to visualize genetic similarity among populations using the neighbor-joining method in MEGA 3.1 (Kumar et al. 2004).

Nested Clade Analysis.--Evolutionary history of haplotypes was inferred using the statistical parsimony (SP) approach of Templeton et al. (1992) as implemented in the software package TCS 1.21 (Clement et al. 2000). Ambiguous evolutionary relationships ($P<0.95$) were resolved using the coalescent-based frequency criterion, preferring connections through the most common haplotype (empirically evaluated by Crandall and Templeton 1993). More frequent haplotypes are expected to be older, more likely to give rise to other haplotypes, and have a greater number of connections.

Nested Clade Analysis (NCA) was used to test for association of genetic divergence and geography of each marker, using the SP network to identify nested haplotype clades as described by Templeton et al. (1992). Clades, sample sizes and river distances between populations were then entered and analyzed in the software package GeoDis 2.4 (Posada et al. 2000). For each nested clade GeoDis performs a χ^2 test based on haplotype frequencies and localities and then geographic distances between localities are used to calculate average distances between clade members. The distributions of clade distances are compared to that of a random distribution (10,000 permutations) and each clade tested for significance of the association of genetic divergence and geographic distance. Evolutionary mechanisms responsible for significant geographic patterns were inferred using the updated (2005) inference key of Templeton (1992).

Cytonuclear tests.--Tests for the association of mtDNA and nuclear DNA genes were performed to assess the potential of recent hybridization among members of the complex. The program CytoNuclear Disequilibria (Asmussen and Basten 1996; Basten and Asmussen 1997) was used to calculate the overall probability that each subpopulation is in equilibrium using a Markov-chain simulation (10,000 replicates) for both allele-haplotype and genotype-haplotype associations. Because multiple nuclear loci were tested for the same subpopulation, levels of significance were corrected using the standard Bonferroni technique (described in Weir 1996).

RESULTS AND DISCUSSION

Characterization of mtDNA variation within populations

Nucleotide sequences from ND2 (a total of 788 bp from two fragments) were examined in 678 specimens from 32 populations, identifying 42 polymorphic sites and 39 haplotypes. Most substitutions occurred at the 3rd codon position (27, 62%), with only (10, 24%) and (5, 12%) occurring at the 2nd and 1st positions, respectively. Visual inspection of haplotypes (Table 2) showed that two haplotypes (AA and AB) are widespread and relatively common throughout the basin (21 of 32 populations, 23% and 14% of all individuals, respectively). Most other haplotypes were relatively rare and often found within single river drainages or locations. Seven populations (Boulder Creek-Bill Williams, Blue River-San Carlos, Cienega Creek-Santa Cruz, East Verde River-Verde, Fossil River-Verde, Sheehy Spring-Santa Cruz and Trout Creek-Bill Williams) were fixed for unique haplotypes. Salt, Verde, San Pedro, Middle Gila, and Bill Williams

River subbasins exhibited haplotypes restricted to each subbasin. Generally, more haplotypes occurred in the Salt and Verde river tributaries compared to other Gila River tributaries. The Bill Williams subbasin, comprised of two *robusta* populations, were distinct from each other and shared no haplotypes with any Gila River drainage samples.

Relatedness among haplotypes was inferred using ML and MP methods. Both likelihood and parsimony analyses yielded near identical starburst patterns with centrally located haplotypes (AA and AB) surrounded by many terminal nodes with short branch lengths, typically single nucleotide substitutions (Figure 3; 4). Parsimony analysis recovered 14,688 equally parsimonious trees with mean tree length of 53 steps (CI=0.79, RI= 0.76). Divergence between the 39 haplotypes was low, and no haplotype differed by more than five nucleotide substitutions from the most common AA haplotype. Bootstrap analysis using parsimony resolved only 10 subclades with more than 50% bootstrap support while more inclusive grouping of most haplotypes was unsupported (Figure 4). The ML tree identified similar clustering, with additional clustering of an HB-HO group, EB-MB group, and a CA-CB-CK group (Figure 3). This result was not unexpected given the low level of divergence among haplotypes.

Nucleotide sequences identified varying but consistent patterns within subpopulations at both haplotype and nucleotide levels (Table 5). Gene diversity ranged from 0.00 - 0.91 (average population diversity = 0.38). Samples from nine localities (including samples of each of the three species) were monomorphic (Cienega Creek, East Verde River, Fossil Creek, Harden-Cienega Creek, Sabino Canyon, Sheehy Spring, Silver Creek, Turkey Creek-AZ and Walker Creek). Cherry Creek, Spring Creek-Verde

River and Aravaipa Creek exhibited the highest number of haplotypes (nine, seven and five, respectively).

Nucleotide diversity was generally low (average π for all populations = 0.001170) and indicated shallow divergence within samples. This pattern results from the repeated subpopulation structure composed of a single common haplotype with additional low frequency haplotypes differing by only a single nucleotide substitution. However, some populations, specifically those with more allelic diversity (e.g., Cherry Creek-Salt, Black River-Salt, Spring Creek-Verde, Williamson Valley Wash-Verde), exhibited much higher levels of π.

Characterization of Tpi-B variation within populations.--There was less variation at Tpi-B for the same individuals than evidenced in ND2 or S7 (see below). This survey identified only five alleles, resulting from six segregating sites and a two bp insertion in intron 2 (for allele E) in the 328 bp fragment. All alleles were closely related and differed by less than seven base-pairs from any other allele. Three alleles (A, B, and C) were widely distributed across all drainages, with four populations (Blue River, East Verde River, Turkey Creek-NM and Cienega Creek) monomorphic for one of these common alleles (Table 3). Two additional alleles, D and E, occurred infrequently, with D found only once from a single individual collected at Cherry Creek (Salt River) and E present in multiple individuals from the mainstem Verde River at Packard Ranch.

Phylogenetic analysis of the five Tpi-B alleles by ML and MP resulted in similar topologies with low levels of divergence (Figure 5). MP identified two equally parsimonious trees (TL= 7 – not including the insertion, CI=0.86, RI=0.67), one of which

was identical to the ML tree. Differences in topology were due to the location of allele C, which was either aligned more closely with A or the B-D group. Bootstrap analysis of the MP tree (unavailable for ML model GTR+SS) identified support for a relationship between alleles B and D group in 71% of 1000 replicates. Phylogenetic divergence and association of private allele D (found only in Cherry Creek) with the common B suggests this allele likely arose in that region. Similarly, private allele E (Verde River at Packard Ranch) differs by three bp (one substitution and a two bp insertion) removed from the most common allele (A). This result is somewhat surprising as both of the above localities with private alleles were historically connected to larger mainstem habitats and subsequently should be less affected by genetic drift. Additional sampling will likely identify these alleles in smaller tributaries of these systems.

Gene diversity within populations at Tpi-B was variable and ranged from 0.00 - 0.75 (Table 6). Most populations exhibited two or three alleles (more in Cherry Creek-Salt River and Verde River at Packard Ranch). Only four of the 32 populations (Blue River, Cienega Creek, East Verde River and Turkey Creek-NM) were monomorphic. Nucleotide diversity ranged from 0.0000 - 0.0064 (overall = 0.0024), with the highest levels also occurring in the Salt and Verde River subbasins.

Characterization of S7 variation within populations.--SSCP variation at the S7 locus revealed 13 alleles (A-M). The 358 bp fragment contained 10 polymorphic sites, with two alleles (E and M) differing from the most common allele only by non-overlapping two and eight bp deletions, respectively. Unlike Tpi-B, where nearly all populations were characterized by combinations of the three common alleles, multiple

localities at S7 exhibited private alleles that ranged in occurrence from rare to fixed (e.g., Cienega Creek and Sabino Canyon-San Pedro River; Boulder and Trout Creeks-Bill Williams River; and East Verde and West Clear Creeks-Verde River). These localities are typical of small and isolated headwater streams.

Relationships among S7 alleles were evaluated by MP and ML methods. Parsimony analysis identified four equally parsimonious trees (TL = 14, CI=0.85, RI=0.60) with two clades (D-K and F-G) receiving moderate bootstrap support (61% and 62%, respectively)(Figure 6). Maximum likelihood analysis (Figure 6) using the HKY+I (Hasegawa et al. 1985) model of evolution resulted in two trees similar to those estimated by MP. The first tree differed from the MP bootstrap consensus tree by uniting alleles C and J. The second ML tree differed from the former ML tree with the clustering of allele B and the D-K group.

Gene diversity at the S7 locus ranged from 0.00 to 0.75 (Table 7). Eight of 32 subpopulations were monomorphic (Blue River-middle Gila River, Boulder Creek-Bill Williams, East Verde River and Fossil Creek-Verde River, Marsh, Tonto and Spring creeks-Salt River, and Sheehy Spring-Santa Cruz), including populations of all three species. Most populations exhibited intermediate levels of *h* (0.26-0.58), with the exception being high levels found in East and upper Eagle creeks of the middle Gila River (0.73 and 0.75, respectively). In contrast, nucleotide diversity (overall π = 0.001425) varied across the entire region, including spectrum of populations ranging from none or low (π = 0.000096) to relatively higher (π = 0.004311) levels.

Patterns of sequence variation at one mtDNA and two nDNA loci indicate populations of the *robusta* complex are genetically similar, reflecting a shallow pattern of divergence for these markers. Alleles from nuclear loci and mtDNA differed by fewer than five base substitutions from the central haplotype, suggesting a relatively recent ancestry among *robusta*, *intermedia* and *nigra*. No fixed or allele frequency differences allowed discrimination among species for all populations. These results are consistent with previous allozymic studies (DeMarais 1992), where 16 polymorphic loci similarly failed to identify species specific markers, with some localities possessing diagnostic alleles while others had alleles found in several populations.

The greater resolution of sequence data revealed a more pronounced pattern than allozymes, identifying unique and private alleles at both ND2 and S7 in many populations. Nine populations (28%) were fixed for unique alleles at ND2 (Cienega Creek, East Verde River, Fossil Creek, Harden-Cienega Creek, Sabino Canyon, Sheehy Spring, Silver Creek, Turkey Creek-AZ, and Walker Creek) and eight populations (25%) fixed for S7 (Blue River, Boulder Creek, Fossil Creek, Marsh Creek, Sheehy Spring, Spring Creek-Salt River, Tonto Creek). Variation at Tpi-B was less pronounced with only two subpopulations exhibiting private alleles. Few exceptions occur where mainstream populations (with presumed increases in population size and gene flow) slow the effects of drift and retain high genotypic diversity.

While limited divergence among alleles and populations is indicative recent ancestry among species, the finding that many populations are either fixed for unique

variants or exhibit private variants of common alleles is indicative of significant past fragmentation and continued evolutionary independence throughout this complex.

Disequilibrium and Hybridization

Hybridization between *G. robusta* and *G. intermedia* has been proposed as a possible explanation for morphological intermediacy and the origin of *G. nigra* (DeMarais 1986, 1992). Covariation of nuclear, mtDNA, and/or morphological traits would be expected if recent contact and hybridization is responsible for observed patterns of variation. At Tpi-B, two of 32 populations (Williamson Valley Wash-Verde River and Bonita Creek-Middle Gila River) differed significantly from Hardy-Weinberg (HW) expectations, exhibiting heterozygote deficiencies, while no deviations from HW expectations were observed for S7. Tests for association of mtDNA with both Tpi-B and S7 alleles found only one population at each nuclear locus that deviates significantly from random. Given the small number of deviations scattered among samples, these results are likely due to chance. In addition, failure to detect association between any combinations of molecular markers is inconsistent with recent hybridization among distinct and formerly isolated forms.

Despite this result, the possibility remains that ancient hybridization may be involved. Using allozymic variation in samples from many of the same geographic populations, DeMarais (1992) also reported that recent hybridization was not supported. However, he proposed that patterns of genetic variation from allozymes (and similarly reported here with sequences) may be generated by ancient hybridization and admixture

leading to homogenization of genetic variation across the species complex. Past hybridization events with subsequent fragmentation is consistent with genetic patterns evidenced in both nuclear and mtDNA and has been observed in other members of the genus *Gila* in the Colorado River. Unlike the parental forms of *G. seminuda* and *G. jordani* of the upper Colorado River (e.g., *G. cypha*, *G. elegans*, and *G. robusta*), putative parental species from the lower basin lack diagnostic nuclear and mtDNA alleles making it very difficult to assess the incongruence of nuclear and mtDNA markers and the potential impact of hybridization.

The effects of local selection also likely play an important role in the evolution of this complex. Support for local selection requires that morphological variation is genetically determined and that independent populations with similar phenotypes (i.e., species) occur in similar environments. Experimental matings involving *G. robusta* other members of the genus (*G. cypha* and *G. elegans*) indicate that similar polymorphisms in closely related species are genetically based (Hamman 1981). Likewise, previous morphological examination (Rinne 1969, 1976; DeMarais 1986) identified strong correlation among morphological traits relative to the three species. This finding, coupled with a likely recent divergence among these species, suggests that relatively few loci of major effect may be responsible for morphological differences. The loci examined here are not likely linked to these morphological traits allowing them to move freely among species when they come into contact (Gerber et al. 2001).

Habitat preferences have been reported by Minckley (1973); however, Rinne (1976) noted that stream size and temperature correlate poorly with morphology and

rejected local selection due temporal instability of desert watercourses. Determining the relative significance of potential selective forces associated with habitat may not be simple or straightforward and will necessitate specific experiments designed to assess the association between morphology and environmental conditions. Note, however, the effects of local selection do not exclude other factors from affecting variation in *Gila* and combinations of forces must be considered in managing the complex.

Distribution of variation within and among populations

Population structure was characterized by partitioning nucleotide variation within and among various hierarchical units for each of the markers characterized (Table 9). To test for correspondence between mitochondrial and morphological variation, samples were partitioned by species following identifications provided in Minckley and DeMarais (2000). Variation among-populations was highly significant for mtDNA variation (F_{ST} = 0.672, $P < 0.001$). Further partitioning of among-population variance into within and among species components yielded no appreciable relationship between patterns of mtDNA variation and morphology (F_{CT} = .0199, P = 0.210), with nearly all among sample variance attributable to differences among populations within species (F_{SC} = 0.665, $P < 0.001$). The significance of geographic structure on mtDNA was examined by partitioning sequence variation by river basin. This approach indicated that variance among populations within subbasins was highly significant (F_{SC} = 0.647, $P<0.001$) while that among subbasins was not (F_{CT} = 0.067, P = 0.109).

Overall, genetic structuring at Tpi-B was consistent, although less pronounced than that provided by mtDNA. Results of AMOVA indicated significant variation in allele frequencies among samples (F_{ST} = 0.325, P<0.001). When total variation among samples at this locus was partitioned according to species, differences among species did not explain a significant portion of variation (F_{CT} = 0.040, P = 0.063), but high and significant variation existed across samples within species (F_{SC} = 0.300, P < 0.001). The role of geographic structuring between drainage subbasins was also tested for Tpi-B. Partitioning F_{ST} across subbasins and samples within subbasins identified relatively low but significant association with subbasins (F_{CT} = 0.071, P = 0.043), while most variation was attributable to differences among samples within subbasins (F_{SC} = 0.273, P<0.001).

Examination of sequence variation at the S7 locus revealed a similar pattern to that found at Tpi-B (Table 9), as a significant portion of variation was distributed among populations (F_{ST} = 0.441, P<0.001). Partitioning F_{ST} by species indicated that allele frequency differences among species accounted for a small but significant fraction of the total variance (F_{CT} = 0.102, P=0.010) with the majority attributable to differences among samples within species (F_{SC} = 0.377, P<0.001). Grouping according to drainage subbasin was also relatively small but significant (F_{CT} = 0.130, P=0.021) while a substantial proportion of genetic variation was attributable to differences among populations within subbasins (F_{SC} = 0.345, P<0.001).

Overall, partitioning of sequence variation among species by AMOVA explained only 5.6% (averaged across loci) of the total variation, with nearly all of the among species variation attributable to variation among populations within species (43.5%)

(Table 9). The same trend was consistent across all loci; however, geography (subbasin structure) consistently explained more variation than species designation. This pattern was most apparent with fine-scaled genetic markers, as visual inspection of mtDNA haplotypes within samples (Table 2) identified unique haplotypes in many populations, often diagnostic for those samples (e.g., Boulder Creek and Trout Creek from the Bill Williams, East Verde River and Fossil Creek from the Verde, Blue River from the San Carlos, and Cienega Creek and Sheehy Spring from the Santa Cruz). This indicates that many local populations have been sufficiently isolated to accumulate unique variants (and accordingly contribute to high F_{SC} values). At Tpi-B, the least variable marker, common alleles were widely dispersed with some frequency differences across the entire basin, and AMOVA results based on current hydrogeographic patterns only marginally better explained population substructure than did species (7% and 4.5% respectively).

Mantel tests.--Mantel tests conducted to examine the association of F_{ST} values between loci among all subpopulations produced varied results. Results from the S7/ND2 and S7/Tpi-B tests significantly reject the hypothesis of independent evolution of genes ($P = 0.005$ and $P = 0.022$, respectively). Alternatively, tests for association of Tpi-B/ND2 failed to reject the null hypothesis of no correlation between markers, however, only marginally ($P = 0.066$). Therefore, there was a consistent pattern of correlation of F_{ST} values among loci. This pattern results from the independent evolution of most populations, with many fixed for unique alleles or exhibiting strong frequency differences, driving high F_{ST} values for all loci, especially the more variable ND2 and S7.

Nested Clade Analysis

Geographic structure was also examined by Nested Clade Analysis (NCA). Nesting structure of the ND2 statistical parsimony network (Figure 7) divided the entire cladogram into four two-step clades composed of 19 single-step clades. Despite shallow divergence of this network, nested contingency tests revealed that all but two clades tested (from all levels) showed significant associations for haplotype frequency and geographic location (Table 10). The major clades 2-1 and 2-4 were each dominated by one of two common haplotypes (AA or AB) that were widely dispersed among 18 and 24 sample localities, respectively. Major clade 2-1, composed of five subclades, and both minor two-step clades (2-2 and 2-3), revealed significant clade distances, implicating a significant impact for allopatric fragmentation. Clade 2-4 showed a similar pattern, but lack of significant clade distances indicated that there was insufficient genetic variation to discriminate between range expansion/colonization events and restricted dispersal/gene flow.

Of the 19 single step clades, many were composed of haplotypes unique to specific localities and did not provide information relative to geographic dispersion. The remaining seven (1-1, 1-3, 1-10, 1-11, 1-13, 1-15 and 1-17) exhibited both genetic and geographic variation. Subclade 1-1 exhibited both high levels of diversity (four haplotypes) and widespread geographic distribution (16 localities). This clade included the most common haplotype AA and is found in populations of all three species. Evolutionary inferences for this clade indicated insufficient genetic resolution to

discriminate between range expansion/colonization and restricted dispersal/gene flow. Clade 1-13 shows a similar pattern to clade 1-1, with a central widespread interior haplotype as well as six low frequency tip haplotypes. Here, the inference key identified the processes responsible for this distribution as restricted gene flow/dispersal but with some long distance dispersal. The small clade 1-3 contained two haplotypes, one of which occurred in two distant locations (D_n= 340 km, Eagle Creek and lower Salt River) whereas the other closely related tip haplotype occurred only the upper Salt River drainage. The inference key identified long distance colonization or past fragmentation followed by range expansion. Because the only other haplotype in this clade occurs solely in the Salt River subbasin, such extremely long distance dispersal seems unlikely for a tip haplotype. This occurrence may be the result of an anthropogenic connection between the upper Salt River drainage and Eagle Creek, possibility due to a permanent connection of the Black River (Salt River subbasin) to Eagle Creek created for mining in 1945 (Phelps Dodge Corporation, Morenci, AZ).

Clade 1-10 again displays a similar pattern of haplotype distribution. Interior haplotype AG, occurring in all three species basin-wide, contains two other related variants (BG and QG) that were unique to the Verde River. The evolutionary processes responsible for shaping this clade were identified as restricted gene flow with isolation by distance. Clade 1-11 is composed of three haplotypes unique to the Salt River drainage that occur in geographically proximate populations (Cherry Creek and Black River). Although this nesting clade is geographically restricted, the inference key indicated that the sampling design was inadequate to discriminate among contiguous range expansion,

long distance dispersal and past fragmentation. Clades 1-15 and 1-17 failed to reject the permutational χ^2 tests thus failing to identify association between genetic variation and geographic distribution.

Nested clade analysis of mtDNA was generally concordant with the results from AMOVA as nearly all clades of the parsimony network are significantly associated with current fine-scaled hydrogeographic structure and indicate a pattern of restricted gene flow with isolation by distance overall (Table 10). Several high level (more inclusive) clades are consistent with complete allopatric fragmentation. However, several individual single-step and tip clades indicate two outcomes not involving fragmentation. In the first, genetic resolution was insufficient to infer process (due to one randomly dispersed widespread haplotype in Clade 1-1 and 2-1) or second, the lack of significant D_c, D_n or I-T values precluded inference of evolutionary process (complete coupling of haplotype phylogeny and geographically proximate localities).

Both clade inferences at high and low levels are also consistent with AMOVA results and patterns of genetic diversity. Few exceptions occur where mainstream populations with presumed increases in population size and increased potential for gene flow would slow the effects of drift and retain high genotypic diversity. Most populations are either fixed for unique variants or exhibit private variants of common alleles. Therefore, although divergence is relatively recent on a phylogenetic scale, fragmentation and evolutionary independence is the rule for populations across the *robusta* complex.

Following the same nesting rules used above I identified clades from the statistical parsimony (SP) network of both Tpi-B and S7 and tested for geographic

association of nuclear alleles. The five alleles at the Tpi-B locus were nested into two hierarchical levels including two second-level clades (2-1 and 2-2) and three single-step clades (1-1, 1-2, and 1-3)(Figure 8a). Despite little divergence and small number of alleles at this locus, clades at all levels with genetic and geographic variation showed significant associations for the χ^2 test ($P < 0.001$). The only exception was clade 2-1, which was nearly significant ($P = 0.07$), exhibiting a pattern likely resulting from restricted gene flow with isolation by distance. This clade contains all three common alleles at Tpi-B (A, B, and C) which are dispersed widely across the basin with significant frequency differences among several subpopulations. Clade 2-2 contains only the private allele E and did not provide any information on geographic structure. Clade 1-1 contains two alleles (A and C) whose distribution implicates restricted gene flow with isolation by distance, consistent with that of the higher level clade (2-1) in which they occur. Single-step clade 1-2 is also highly significant for geographic association of alleles, however, lack of significant clade distance values within the clade indicate that this variation was not informative. The final single-step for Tpi-B, clade 1-3, contains only the singleton allele E and was not evaluated.

S7 exhibited more allelic variation than Tpi-B. The SP network (Figure 8b) was nested similarly to both Tpi-B and mtDNA, showing a shallow divergence comprised of a series of mainly single-step clades (1-1, 1-2, 1-3, 1-4, and 1-5). Every clade at this locus demonstrated highly significant geographic structure ($P < 0.001$). The largest clade, 1-1, contains six alleles (the most common allele A and variants unique to several populations) with patterns of variation consistent with contiguous range expansion. Since

allele A is widespread and centrally located and variant alleles differ by single substitutions, the ancestral A allele and unique variants coexist, accounting for range expansion and geographic structure. Clade 1-2 contains two alleles (D and K), each unique to Cienega and the Eagle Creeks, respectively. Here, the process inferred is allopatric fragmentation. Next, clade 1-3 contains allele C unique to Trout Creek and the widespread allele J. One explanation that could be inferred from this pattern is that long distance colonization of Trout Creek with subsequent fragmentation may be involved. Alternatively, this pattern may reflect fragmentation and subsequent range expansion. Further evidence is required to differentiate between these hypotheses. Last, clade 1-4 (G and H) and clade 1-5 (I) contained no geographic or no genetic variation respectively and no evolutionary processes were inferred.

Across all three loci, NCA identified clusters of populations with genetic patterns likely caused by restricted gene flow. As similarly indicated by AMOVA, many alleles/haplotype at ND2 and S7 were private and precluded analysis based on geography as single-step clades. These alleles revealed complete fragmentation in higher level clades composed of populations fixed for unique alleles. In other cases at ND2 and S7, restricted gene flow was still detected but distinguishing among past fragmentation, range expansion, colonization or isolation by distance was not possible. Tpi-B, where typically only frequency differences exist among populations, still identified restricted gene flow with isolation by distance across all populations for two clades.

Phylogenetic Relationships Among Populations

Pairwise estimates of F_{ST} based on frequencies and sequence data were used as short-term genetic distances to generate phylogenetic trees of lower Colorado River subpopulations (Figures 9; 10; 11). For all loci, individual subpopulations were scattered across the trees, regardless of species or drainage relationships, consistent with the importance of local processes in driving the evolution of this complex. As noted above, results from the Mantel tests of F_{ST} values showed similarly that relationships between pairs of subpopulations exist across loci and suggests differences between subpopulations are meaningful. Of the three loci, the phylogram for the highly variable ND2 provides the most resolution for identifying populations with unique sequences. Both ND2 and S7 trees additionally show that many *G. robusta* subpopulations occupy positions central to the topology. This is due in part to the many shared alleles which is likely the result of more recent connectivity associated with their habitat.

Evolutionary implications and conservation of the Gila robusta complex

Members of the *robusta* complex exhibit complicated patterns of morphological and genetic variation, leading to confusion over assignment of individual populations to species and effective management of this endangered complex of fishes. Levels of divergence among alleles are low; however, population subdivision is strong. Despite significant differentiation among populations, there is little association between morphology, geography, and genetic structure, indicating that local processes are driving

evolution within this complex. Consistent with this result is the presence of many populations with unique alleles at high frequency.

Cryptic variation among closely related taxa presents an excellent application of molecular markers to aid conservation efforts, particularly where intermediate morphological variation is involved (Avise 2000). Because levels of genetic variation are typically associated with effective population size, neutral molecular markers allow for assessment of relative population sizes. Small population sizes can lead to inbreeding depression and loss of evolutionary potential, and increase the importance of stochastic events.

Care must also be taken to avoid human-induced changes in genetic composition of distinctive groups and allow for continuance of evolutionary processes. These markers allow for examination of levels and distribution of genetic variation within and among populations. Assessment of genetic distinctiveness in this manner allows for the identification of evolutionary significant units (ESUs) and management units (MUs)(Moritz 1994). ESUs are groups of populations that are reciprocally monophyletic for mtDNA alleles and show significant allele frequency differences at nuclear loci. MUs are less strongly defined, and are identified through frequency differences at either mtDNA or nuclear loci. The more inclusive ESU was designed to capture groups with deeper divergence but reserve the potential for gene flow despite contemporary fragmentation.

Long-standing difficulties in morphological discrimination and taxonomic distinction among members from the lower Colorado *robusta* complex have plagued

conservation efforts. The three species, *G. intermedia*, *G. nigra*, and *G. robusta*, are morphologically distinct and therefore each merits protection in its own right. Here molecular markers were used to characterize levels and patterns of genetic variation and provide valuable information for informed management of these three dwindling species. While many populations were found to be distinct, levels of divergence are indicative of relatively recent (but pre-human influence) shared common ancestry throughout the complex; genetic structure is not associated with species designation or even hydrography. For example, examination of mtDNA variation of *G. nigra* from the "Forks" region (East, West and Middle forks) of the Gila River in New Mexico and *G. intermedia* from nearby Turkey Creek shows the two are indistinguishable using these markers; however, *G. intermedia* from nearby Harden-Cienaga Creek is quite distinct (Figures 9; 11). Therefore, few populations or groups of populations satisfy reciprocal monophyly, although significant frequency differences and/or fixation at mtDNA and S7 nuclear loci are not uncommon.

The rigidity of political interpretations of species concepts and enforcement of the Endangered Species Act has made it difficult to manage groups like the *G. robusta* complex, with their recent evolutionary history and probable exchange of genetic material during pluvial periods when these organisms were most assuredly more abundant and widespread. Unfortunately, this limited perspective has forced managers into the difficult (if not impossible) task of assigning each population to its appropriate species. Because speciation is a continuous, complex process, with each species being shaped by a multitude of evolutionary forces (e.g., selection, drift, and gene flow),

assignment of populations to their appropriate species may not always be possible. This difficulty is especially compounded in situations where divergence is recent and exchange of genetic variation among units is intermittent as appears to be the case here. In such situations, it may not be prudent to force populations into discrete categories as this may generate unnatural isolation of species within the complex, reducing evolutionary potential and eliminating natural processes involved in evolution of the complex. Instead, the *complex* should be considered as a single evolutionary unit, and the management strategy designed to protect *all* the diversity represented throughout.

It is important to consider variation throughout the entire Colorado River basin to place variation and divergence in the lower basin *G. robusta* complex in appropriate context. Two isolated species of hybrid origin (involving *G. robusta*, *G. cypha*, and *G. elegans*) can be found in the Virgin and White river drainages (*G. seminuda* – DeMarais et al. 1992; *G. jordani* – Gerber et al. 2001, respectively). *Gila robusta* is relatively abundant in the mainstem Colorado River and tributaries above the Grand Canyon; however, all individuals from the headwaters of the Little Colorado River and above the Grand Canyon possess *G. cypha* or *G. elegans* mtDNA (Dowling and DeMarais 1993, Gerber et al. 2001). Populations of the *G. robusta* complex of the lower basin (including *G. robusta*, *G. intermedia*, and *G. nigra*) are disappearing rapidly and these species are on the verge of extinction. This is especially problematic as all individuals from the lower basin possess a unique divergent mtDNA lineage that has never been found outside the lower basin (Dowling and DeMarais 1993, Gerber et al. 2001, this study). Therefore,

populations of *Gila robusta* from below the Grand Canyon are clearly divergent and should be managed as a distinct unit from those above.

Within the lower basin, there are three distinct species described on the basis of morphological traits; however, molecular characters fail to reliably discriminate among them. For management purposes, it would be prudent to take a multi-tiered approach, using the morphological variation as the first level and then identifying appropriate units for management within each of the three forms (Table 16); therefore, each species will be discussed in turn below. Because of the significance of local processes on patterns of variation within these species, it is difficult to predict where unsampled populations might fall out, and these need examination before they become the targets of specific management actions. For existing populations examined here, focus was placed upon the existence of unique alleles, especially those at high frequency. Such alleles likely reflect past isolation and a unique evolutionary trajectory for those populations, indicative of their significance as management units.

Gila robusta was once distributed throughout much of the mainstem Gila River basin and the Bill Williams River drainage. In this study, eight samples of this species were examined, six from the Gila River drainage (representing four subbasins) and two from the Bill Williams. The latter two samples, Boulder Creek and Trout Creek were distinct and reciprocally monophyletic from each other and all other samples of *G. robusta*. Each of these samples possesses two unique, common mtDNA haplotypes that differ from other haplotypes by three substitutions, and unique S7 alleles at high frequency (fixed in Boulder Creek). These results indicate that populations from the Bill

Williams River have been isolated for a sufficient amount of time to accumulate distinct alleles and therefore should be managed separately from those of the Gila River. Given relatively high levels of difference among the two samples, it is likely that each subdrainage of the Bill Williams should also be managed separately; however, additional sampling is necessary to more fully address this issue. Unfortunately for many populations, especially for those in mainstream habitats of the Bill Williams basin, *Gila* is rare and in some cases extirpated.

The remaining Gila River drainage populations of *robusta* cluster around the center of the population topology (Figure 9), indicating that they share similar genetic features. The three most common mtDNA haplotypes (AA, AB and AG) and two most common S7 alleles (A, B) are common across all of these populations. In addition, some rare variants (e.g., AF, AH, AK for mtDNA, J for S7) are also shared across populations, consistent with relatively recent gene exchange. This is consistent with the preference of *G. robusta* for mainstem reaches, resulting in greater potential for gene exchange due to increased connectivity. Note, however, there is some evidence for independent evolution as well and these samples represent two management units. Cherry Creek is particularly noteworthy as it is somewhat distinct from the other Gila River populations of *G. robusta*, with 14 of 21 individuals sampled possessing six mtDNA haplotypes unique to this location. The upper Salt River sample of *G. robusta* from Black River also shows a high level of distinctiveness (13 of 24 individuals possessing three unique mtDNA haplotypes). This result, coupled with high levels of diversity, indicate that additional

sampling from remaining populations in tributaries of the upper Salt River is necessary for understanding evolution of this species and managing the complex.

Gila nigra is found in a mosaic distribution of headwater and moderate-sized reaches of the Gila River drainage. As for *G. robusta*, eight samples were characterized, representing three subbasins. All of these populations are divergent, scattered throughout the population topology (Figure 9) due to the existence of unique alleles at high frequency. Variation among the three samples from the Tonto drainage of the Salt River (Marsh, Tonto, and Spring creeks) and Cherry Creek *G. robusta* (also from the Salt River) are phenetically similar due to the prevalence of AB and related haplotypes. Despite this similarity, Marsh and Tonto creeks share a unique mtDNA haplotype (AK) at high frequency, only found elsewhere in a single individual from Cherry Creek, while Spring Creek has a unique haplotype found at high frequency (CB). This same pattern was repeated for Tpi-B, where Marsh and Tonto creeks had allele B at high frequency while Spring Creek was predominantly A. Despite some shared mtDNA alleles, Cherry Creek is distinctive at both Tpi-B (allele C predominant) and S7 (includes B that is not found in Salt River subbasin samples of *G. nigra*). Given this pattern, two units of management concern are represented in the Salt River subbasin, Tonto-Marsh and Spring Creek. This division within *G. nigra* is geographically concordant and likely reflects the geographical proximity of Tonto and Marsh creek collections relative to Spring creek.

Three samples of *G. nigra* from the Verde River are especially divergent for mtDNA. Samples from East Verde River and Fossil Creek are fixed for unique haplotypes (DB and RA, respectively) and West Clear has widespread but relatively rare

haplotype AG at high frequency. There is also some variation in allele frequencies at nuclear genes as well, with a unique S7 allele at high frequency in the East Verde River and in a single individual in West Clear Creek, indicating some recent connection between these geographically proximate locations. Based on fixation of unique haplotypes, East Verde River and Fossil Creek require special consideration as unique entities and thus managed separately from nearby West Clear Creek. The last population of *G. nigra* from the "Forks" region of the upper Gila River in New Mexico is similar to the latter sample and could be grouped with West Clear Creek; however, because of its great distance from this region it is most appropriate to maintain them separate entities.

Gila intermedia was the best represented of the three species, with 17 samples collected from seven subbasins across the Gila River basin. This species epitomizes the difficulties in developing management strategies for evolutionary complexes. There are several populations fixed (or nearly so) for unique alleles at one or more loci: Blue River (mtDNA), Cienega Creek (mtDNA and S7), Sabino Canyon (S7), and Sheehy Spring (mtDNA). Each of these should be managed independently due to their unique genetic composition. Turkey Creek, AZ was fixed for a distinct mtDNA haplotype that was also found in geographically proximal samples from Bass, Redfield and O'Donnell canyons, indicating that these could also be considered a management unit. The three intermedia samples from the Verde River subbasin, Spring Creek, Walker Creek and Williamson Valley Wash, each have distinct sets of mtDNA haplotypes and should be considered separately in management actions. Note, however, that the unique mtDNA haplotypes at Williamson Valley Wash are also found in *G. robusta* from the Verde River at

Perkinsville sample, and additionally Spring and Walker creeks have similar allele profiles with *G. nigra* from nearby West Clear Creek, consistent with recent gene exchange and/or divergence.

All remaining samples (with the exception of Silver Creek from the Agua Fria) are from the middle and upper Gila River tributaries. Samples from these sections of the Gila basin are generally central in the population topology due to the presence of mostly common, shared alleles. An interesting exception is provided by the Eagle Creek samples (East Fork, upper and lower) that are especially problematic and illustrate the pitfalls of using a small number of markers. For S7, samples from East Fork Eagle Creek, upper Eagle Creek (both *G. intermedia*), and lower Eagle Creek (*G. robusta*) share very similar allele profiles (including two unique S7 alleles, F and G), suggestive of recent contact and gene exchange; however, the distinctive morphologies and different mtDNA and Tpi-B frequencies between lower Eagle Creek and the two upper Eagle samples indicate restriction to gene exchange. In addition, these unique S7 alleles are found in other *G. intermedia* samples from the upper Gila River ("Forks" region of New Mexico) and occur in low frequency in Aravaipa Creek of the San Pedro subbasin near the Gila River confluence, indicating recent contact. Therefore, all middle and upper Gila River *G. intermedia* (Bonita, East and Upper Eagle, Harden-Cienega and Turkey creeks) may be considered a single management unit. Finally, the sample from Silver Creek is geographically distant from all other samples, yet it exhibits common alleles at all loci except for a single individual with a unique variant at S7. Because of its isolation and

lack of additional samples from that region, however, it would be prudent at this point to consider it separately until the region can be studied further.

In conclusion, the *Gila robusta* complex from the lower Colorado River basin exhibits a complicated mosaic of morphological and molecular characters that challenge our ability to conserve the tremendous diversity represented by the group. Because of its involved evolutionary history and the importance of local processes in its evolution, conservation and management of this group will require creative solutions and careful consideration of diversity represented throughout its range. Of special importance is development of a single, comprehensive plan to coordinate management actions. Results reported here identify key factors to consider; however, before specific management actions are undertaken it will be necessary to carefully consider all demographic, ecological, and genetic factors to insure protection of this imperiled group. Until such careful consideration has been afforded and specific management actions are identified, it is recommended that all populations or management units identified here (or newly collected) are protected and conserved to maintain the evolutionary potential and enhance survivorship of all members of the complex.

LITERATURE CITED

Asmussen, M. A. and C. J. Basten. 1996. Constraints and normalized measures for cytonuclear disequilibria. Heredity 76:207-214.

Avise, J. C. 2004. Molecular markers, natural history, and evolution. Sinauer Associates, Sunderland, Mass.

Avise, J. C. 2000. Phylogeography: the history and formation of species. Sinauer Associates, Sunderland, Mass.

Axelrod, D. I. 1979. Age and origin of the Sonoran desert vegetation. Occ. Cal. Acad. Sci. 132:1-74.

Basten, C. J. and M. A. Asmussen. 1997. The exact test for cytonuclear disequilibria. Genetics 146:1165-1171.

Chow, S. and K. Hazama. 1998. Universal PCR primers for S7 ribosomal protein gene introns in fish. Mol. Ecol. 7:1255-1256.

Clement, M., D. Posada and K. A. Crandall. 2000. TCS: a computer program to estimate gene genealogies. Mol. Ecol. 9:1657-1659.

Crandall, K. A., and A. R. Templeton. 1993. Empirical tests of some predictions from coalescent theory with applications to intraspecific phylogeny reconstruction. Genetics 134:959-969.

DeMarais, B. D. 1986. Morphological variation in *Gila* (Pisces: Cyprinidae) and geologic history: lower Colorado river basin. Unpub. M.S. Thesis, Arizona State University, Tempe.

DeMarais, B. D. 1992. Genetic relationships among fishes allied to the genus *Gila* (Teleostei: Cyprinidae) from the American Southwest. Unpub. Ph.D. Diss., Arizona State University, Tempe.

DeMarais B. D., T. E. Dowling, M. E. Douglas, W. L. Minckley, and P. C. Marsh. 1992. Origin of Gila-Seminuda (Teleostei, Cyprinidae) through introgressive hybridization - Implications for evolution and conservation. Proc. Natl. Acad. Sci. 89:2747-2751.

Douglas, M. E., W. L. Minckley, and B. D. DeMarais. 1999. Did vicariance mold phenotypes of western North American fishes? Evidence from Gila River cyprinids. Evolution 53:238-246.

Dowling, T. E., and B. D. DeMarais. 1993. Evolutionary significance of introgressive hybridization in cyprinid fishes. Nature 362:444-446.

Dowling, T. E., P. C. Marsh, A. T. Kelsen, and C. A. Tibbets. 2005. Genetic monitoring of wild and repatriated populations of endangered razorback sucker (*Xyrauchen texanus*, Catostomidae, Teleostei) in Lake Mohave, Arizona-Nevada. Mol. Ecol. 14:123-135.

Dowling, T. E., W. L. Minckley, P. C. Marsh, and E. S. Goldstein. 1996. Mitochondrial DNA variability in the endangered razorback sucker (*Xyrauchen texanus*): Analysis of hatchery stocks and implications for captive propagation. Conserv. Biol. 10:120-127.

Excoffier, L., P. Smouse, J. M. Quattro. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131:470–491.

Gerber, A. S., C. A. Tibbets, and T. E. Dowling. 2001. The role of introgressive hybridization in the evolution of the *Gila robusta* complex (Teleostei: Cyprinidae). Evolution 55:2028-2039.

Guo, S. and E. Thompson. 1992. Performing the exact test of Hardy-Weinberg proportion for multiple alleles. Biometrics 48:361-372.

Hamman, R. L. 1981. Hybridization of the three species of chub in a hatchery. Prog. Fish Cult. 43:140-141.

Hasegawa, M., H. Kishino, and T. Yano. 1985. Dating the human-ape spliting by a molecular clock of mitochondrial DNA. J. Mol. Evol. 22:160-174.

Hubbs, C. L. 1955. Hybridization between fish species in nature. Syst. Zool. 4:1-21.

Hughes, J., A. M. Baker, C. Bartlett, S. Bunn, K. Goudkamp, and J. Somerville. 2004. Past and present patterns of connectivity among populations of four cryptic species of freshwater mussels *Velesunio* spp. (Hyriidae) in central Australia. Mol. Ecol. 13:3197-3212.

Hurt, C., and P. Hedrick. 2004. Conservation genetics in aquatic species: General approaches and case studies in fishes and springsnails of arid lands. Aquat. Sci. 66:402-413.

Johnson, J. B., T. E. Dowling, and M. C. Belk. 2004. Neglected taxonomy of rare desert fishes: Congruent evidence for two species of leatherside chub. Syst. Biol. 53:841–855.

Kumar, S., K Tamura, and M. Nei. 2004. MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. Brief. in Bioinform. 5:150-163.

Mantel, N. A. 1967. The detection of disease clustering and a generalized regression approach. Cancer Res. 27:209-220.

Merritt, T. J. S. and J. M. Quattro. 2001. Evidence for a period of directional selection following gene duplication in a neutrally expressed locus of triosephosphate isomerase. Genetics 159:689-697.

Minckley, W. L. 1973. Fishes of Arizona. Arizona Game and Fish Dept., Phoenix.

Minckley, W. L., and J. E. Deacon (eds). 1991. Battle against extinction: Native fish management in the American West. University of Arizona Press, Tucson.

Minckley, W. L. and B. D. DeMarais. 2000. Taxonomy of chubs (Teleostei, Cyprinidae), genus *Gila* in the American southwest with comments on conservation. Copeia 1:251-256.

Minckley, W. L, D. A. Hendrickson, and C. E. Bond. 1986. Geography of western North American freshwater fishes: Description and relationships to intercontinental tectonism. Pp. 519-613. C. H. Hocutt and E. O. Wiley, eds. *In* The zoogeography of North American freshwater fishes. John Wiley, NY.

Moritz, C. 1994. Defining 'evolutionary signifcant units' for conservation. Trends Ecol. Evol. 9:373-375.

Nei, M. 1997. Molecular evolutionary genetics. Columbia University Press, NewYork, NY.

Pluthero, F. G. 1993. Rapid purification of high-activity Taq DNA polymerase. Nuc. Acid Res. 21:4850-4851.

Posada, D. and K. A. Crandall. 1998. Modeltest: testing the model of DNA substitution. Bioinformatics 14:817-818.

Rinne, J. N. 1969. Cyprinid fishes of the genus *Gila* from the lower Colorado River basin. Unpub. M.S. Thesis, Arizona State University, Tempe.

Rinne, J. N. 1976. Cyprinid fishes of the genus *Gila* from the lower Colorado River basin. Wasman J. of Biol. 34:65-107.

Schneider, S., D. Roessli, and L. Excoffier. 2000. Arlequin: A software for population genetics data analysis. Ver 2.00. Genetics and Biometry Lab, Dept. of Anthropology, University of Geneva.

Smith, G. R. 1966. Distribution and evolution of the North American catostomid fishes of the subgenus, *Pantosteus*, genus *Catostomus*. Misc. Publ. Mus. Zool., University of Michigan, 29:1-32.

Smouse, P. E., J. C. Long, and R. R. Sokal. 1986. Multiple regression and correlation extensions of the mantel test of matrix correspondence. Syst. Zool. 35:627.

Soulé, M. E., and B. A. Wilcox (eds). 1980. Conservation biology: An evolutionary-ecological perspective. Sinauer Associates, Sunderland, Mass.

Swofford, D. L. 1998. PAUP*:Phylogenetic Analysis Using Parsimony (*and Other Methods), Version 4. Sinauer, Sunderland, Mass.

Tajima, F. 1983. Measurement of DNA polymorphism. Pp. 37-59. N. Takahata and A. G. Clark, eds. *In*: Mechanisms of Molecular Evolution. Introduction to Molecular Paleopopulation Biology. Tokyo, Sunderland, MA: Japan Scientific Societies Press, Sinauer Associates.

Templeton, A. R., K. A. Crandall, and C. F. Sing. 1992. A cladistic-analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA-sequence data. III. Cladogram estimation. Genetics 132: 619-633.

Tibbets, C. A., and T. E. Dowling. 1996. Effects of intrinsic and extrinsic factors on population fragmentation in three species of North American minnows (Teleostei: Cyprinidae). Evolution 50:1280-1292.

U.S. Fish and Wildlife Service. 2006. Endangered and threatened species. Review of plant and animal taxa: proposed rule. Fed. Reg. 71(89), 26007-260017.

Voeltz, J. B. 2002. Roundtail Chub *(Gila robusta)* status survey of the Lower Colorado River Basin. AZGFD Non-Game Technical Report #186. pp. 221.

Vrijenhoek, R. C. 1996. Conservation genetics of North American desert fishes. Pp. 367-397. J. C. Avise and J. L. Hamrick, eds. *In* Conservation genetics: Case histories from nature. Chapman and Hall, New York.

Weir, B. S. 1996. Genetic data analysis II: Methods for discrete population genetic data. Sinauer Associates, Sunderland, Mass.

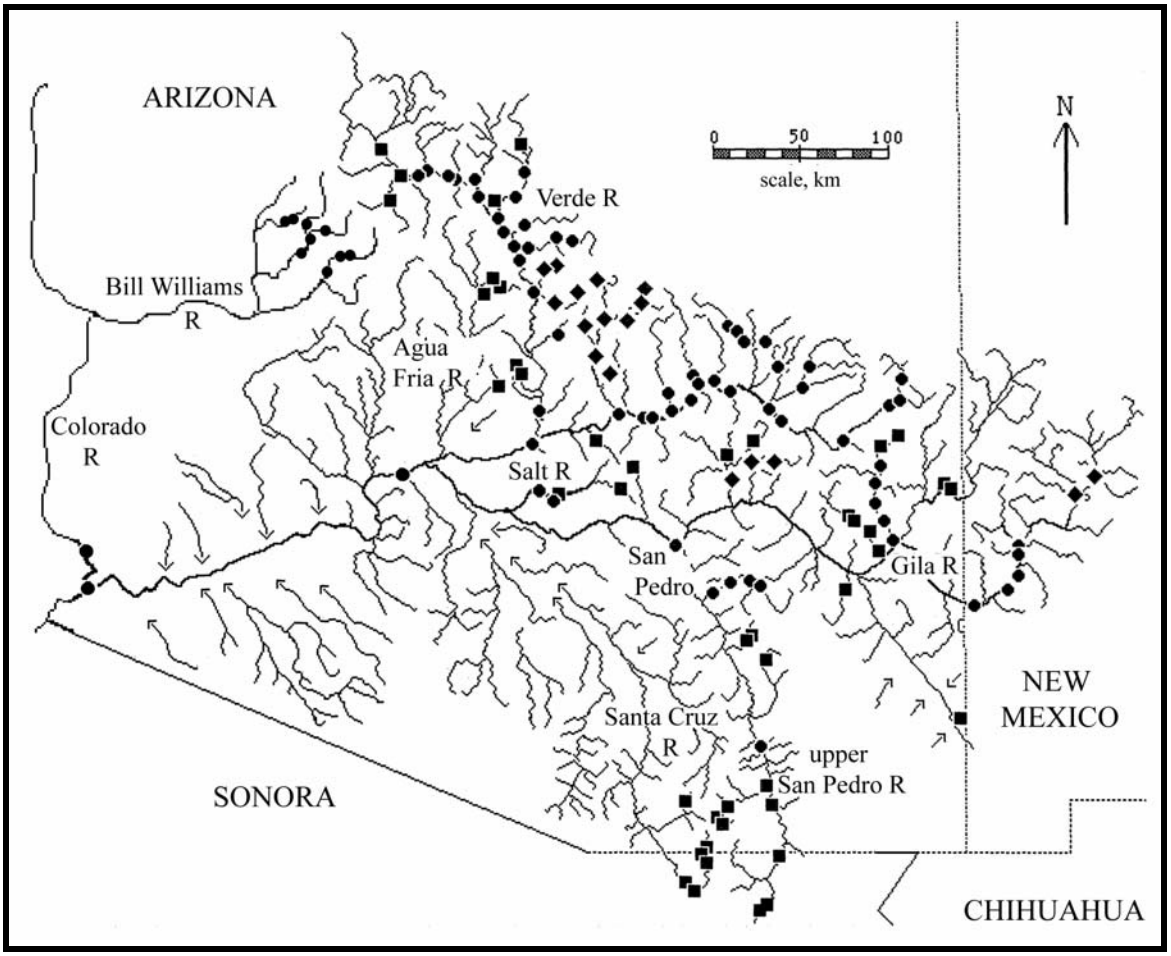

FIG. 1. Historic and current records of the *G. robusta* complex in the lower Colorado River basin, including populations from the Bill Williams and Gila drainages. Circles, diamonds and squares represent populations identified as *G. robusta*, *G. nigra* and *G. intermedia* respectively. (Figure modified from Douglas et al. 1999 and Minckley and DeMarais 2000).

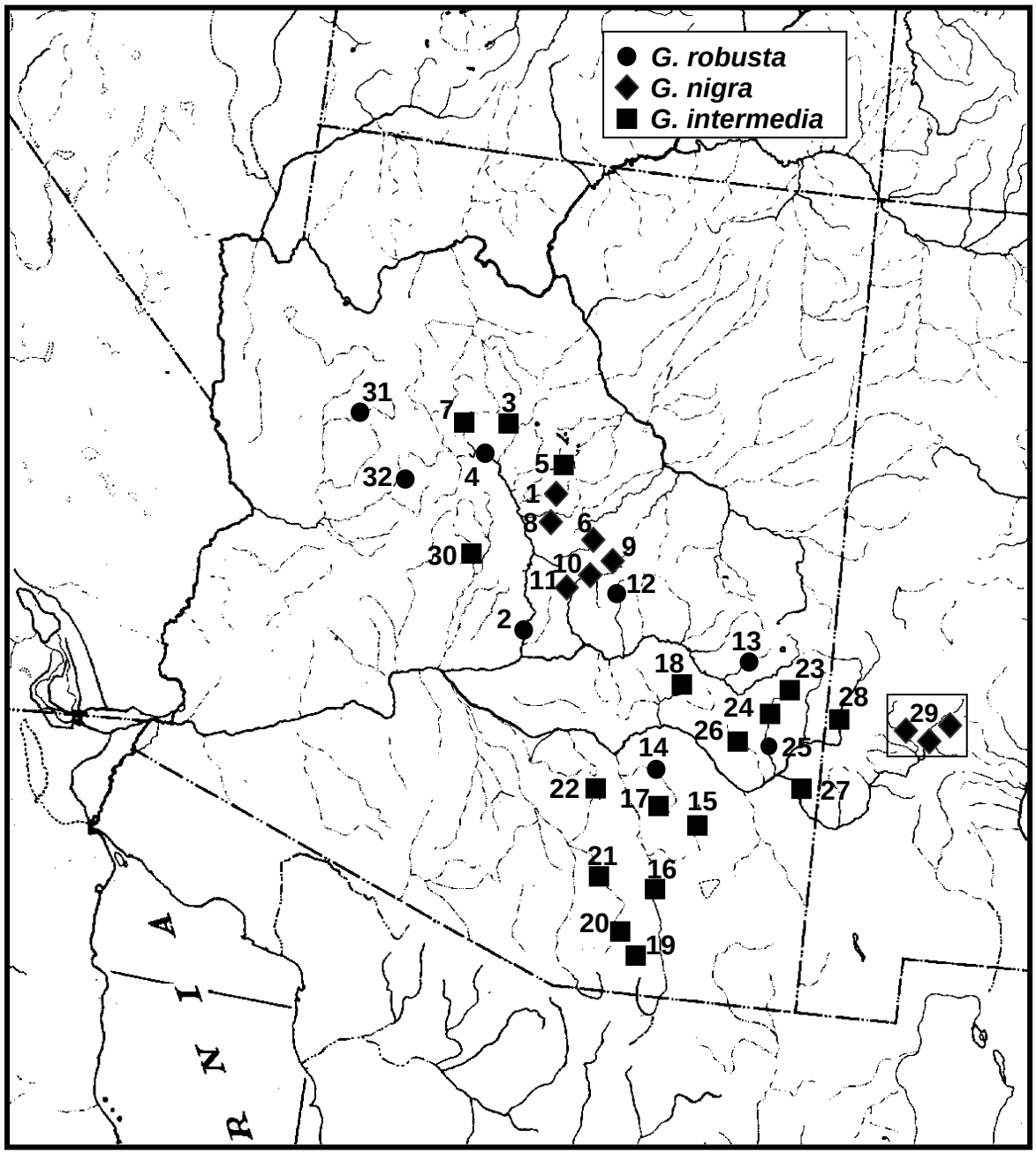

FIG. 2. Map of 32 sampled localities of the *Gila robusta* species of the lower Colorado River basin, AZ and NM. *Gila robusta, G. nigra and G. intermedia* are identified by circles, diamonds and squares, respectively. Site names and sample sizes for numbered localities shown below are listed in Appendix 1. The enclosed box of locality 29 (NMGILA) represents a pooled collection from the Gila River headwaters due to geographic proximity and small sample size.

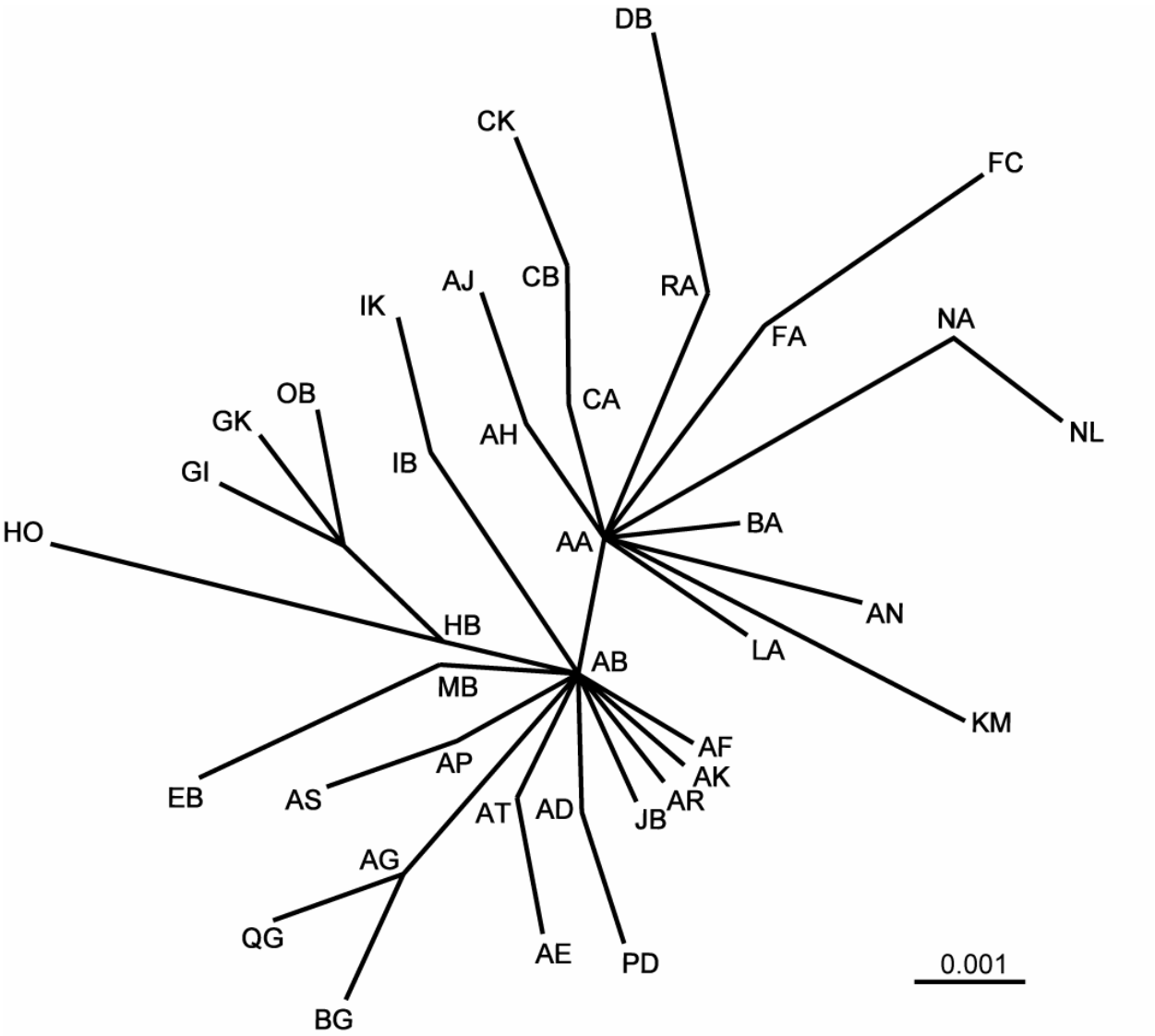

FIG. 3. Maximum likelihood tree generated from mtDNA sequences. Haplotypes identified in Table 2.

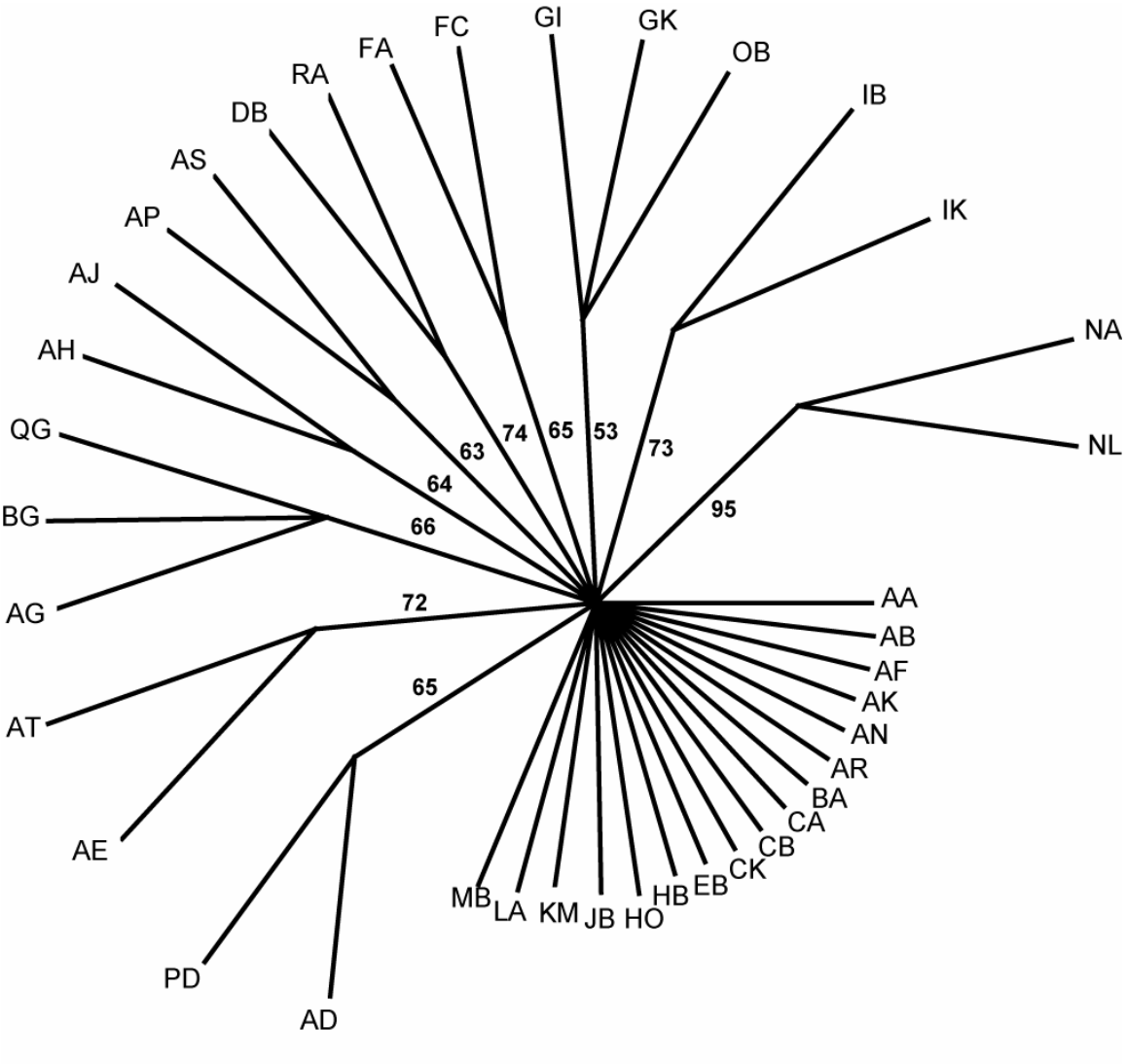

FIG. 4. Bootstrap consensus tree (1000 replicates) of 14,688 MP trees generated from mtDNA sequences. Values on branches are the % of 1000 bootstrap replicates recovered in the bootstrap consensus tree (only provided for values >50%).

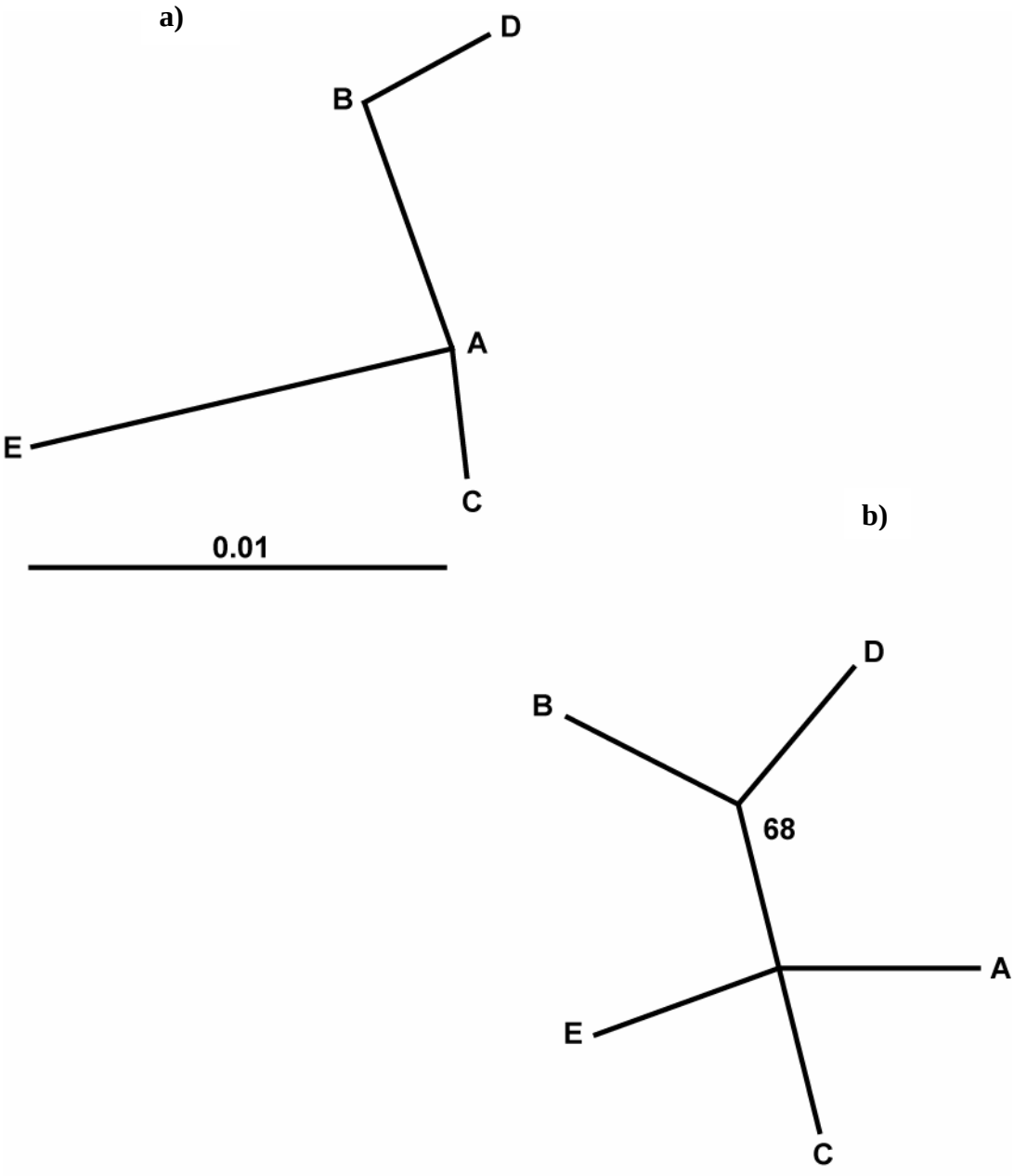

FIG. 5. Maximum likelihood (a) and bootstrap consensus of maximum parsimony trees (b) generated from Tpi-B sequences. Alleles identified in Table 3. Values on branches are the % of bootstrap replicates (1000) recovered in the bootstrap consensus tree (only provided for values >50%).

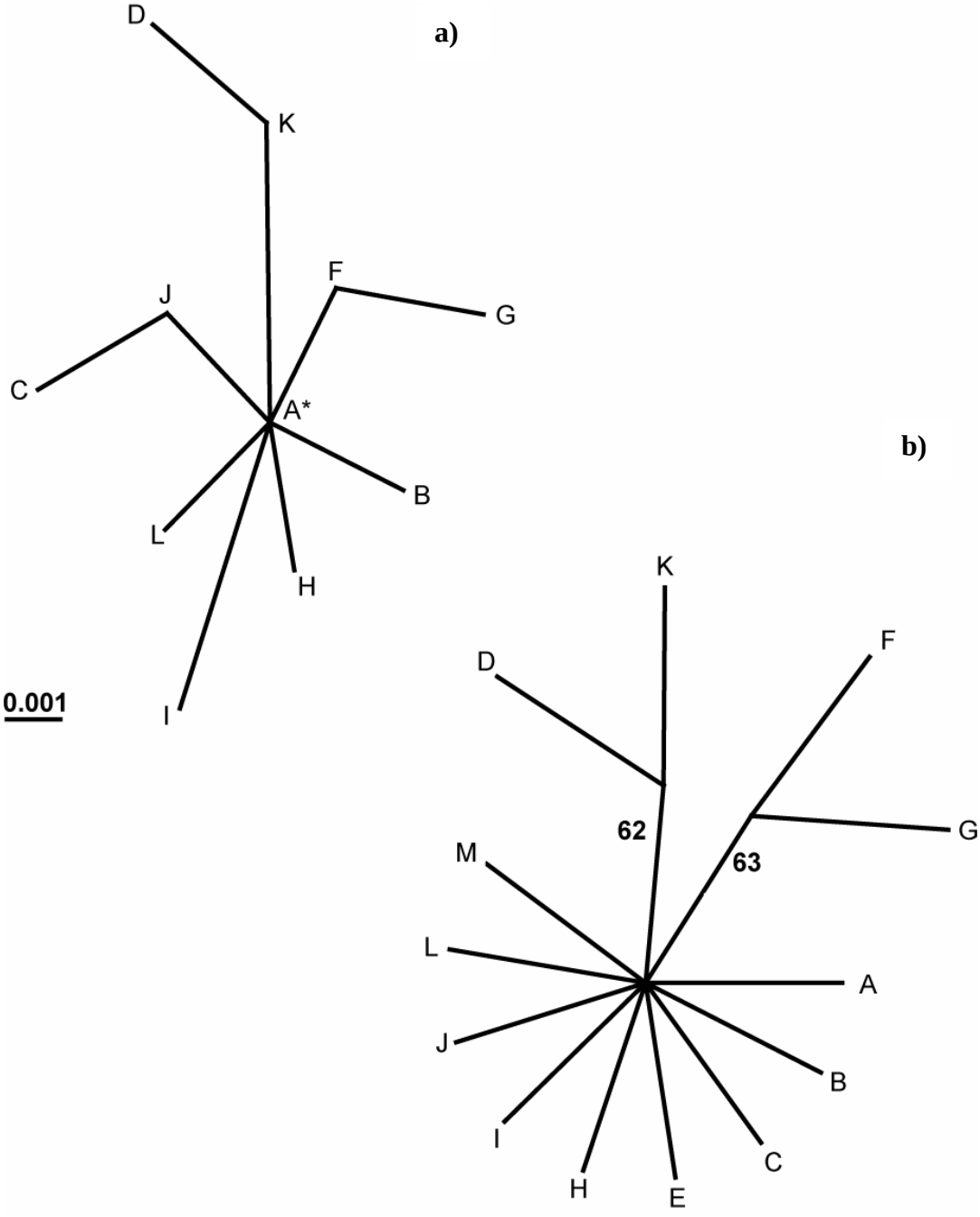

FIG. 6. Maximum likelihood (a) and bootstrap consensus of maximum parsimony (b) trees for S7 sequences. Alleles identified in Table 4. Allele A* in ML tree (a) additionally denotes alleles with deletions (E and M) from A. For MP, values on branches are the % of bootstrap replicates (1000) recovered in the bootstrap consensus tree (only provided for values >50%).

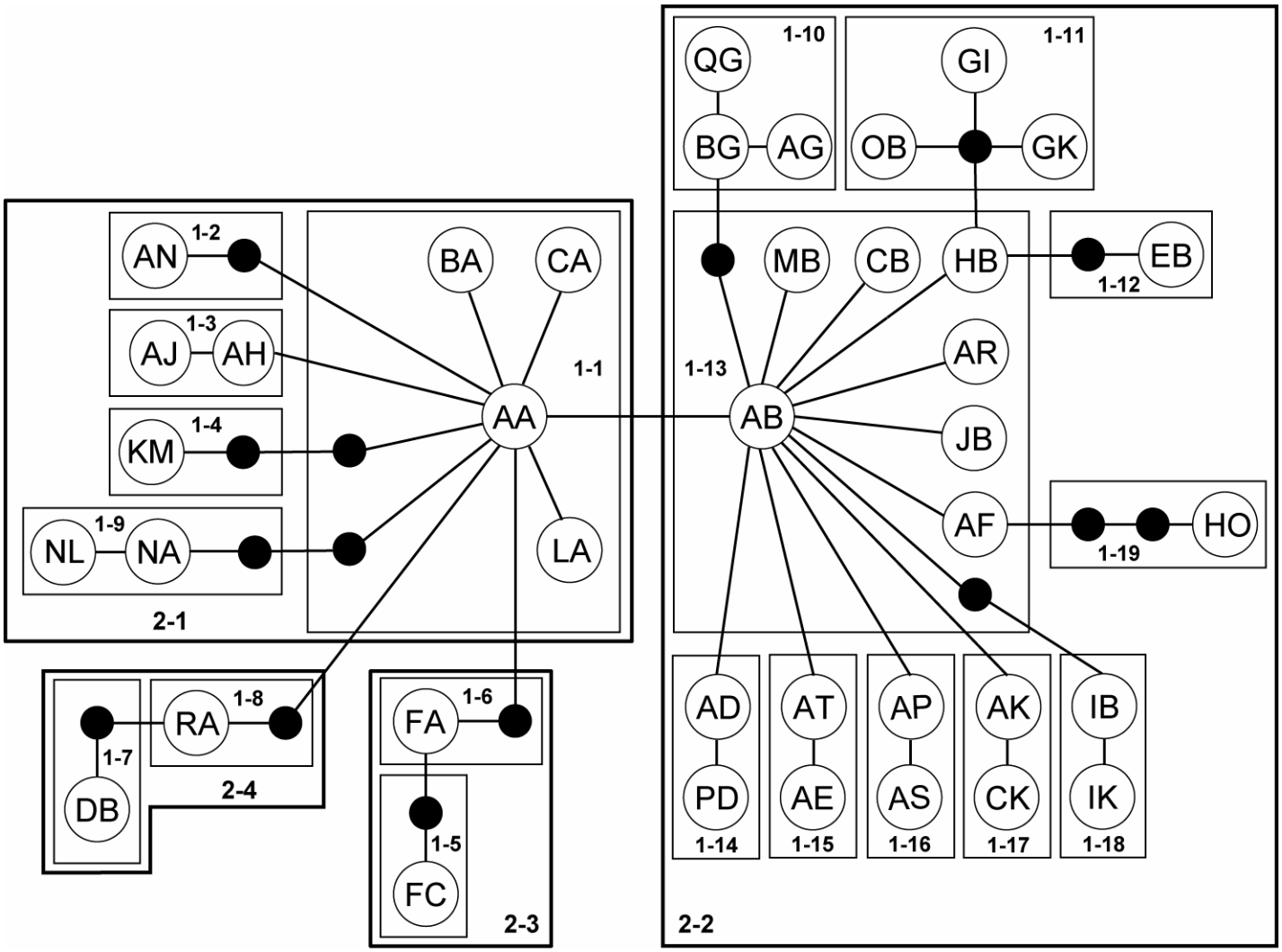

FIG. 7. Statistical parsimony network and nesting design of ND2 haplotypes used in nested clade analysis. Sample sizes and population are identified in Table 2. Black solid circles indicate inferred haplotypes.

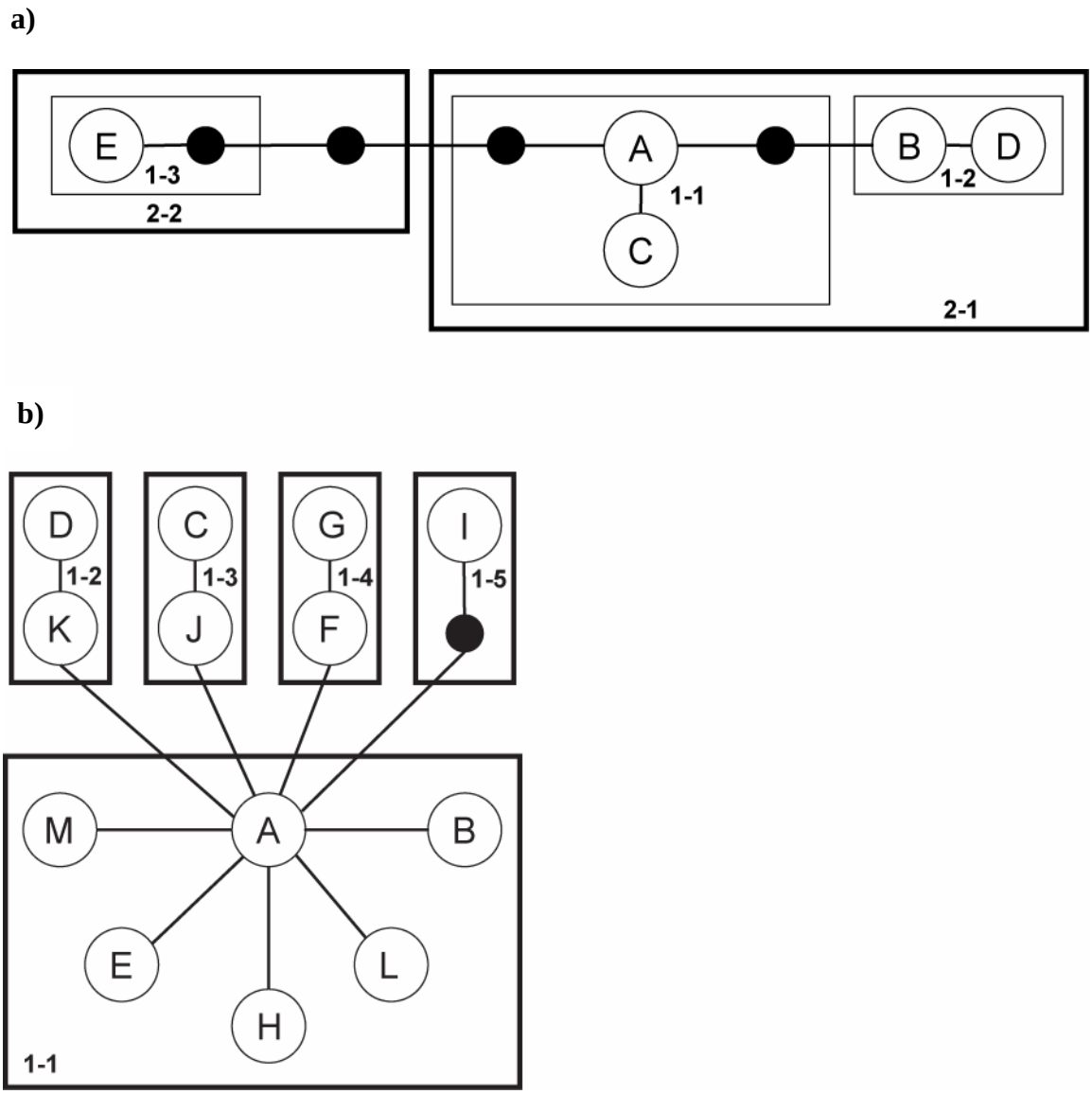

FIG. 8. Statistical parsimony networks and nesting design of Tpi-B (a) and S7 (b) alleles used in nested clade analysis. Sample sizes and populations composed of each allele are listed in Tables 3 and 4. Black solid circles indicate inferred alleles.

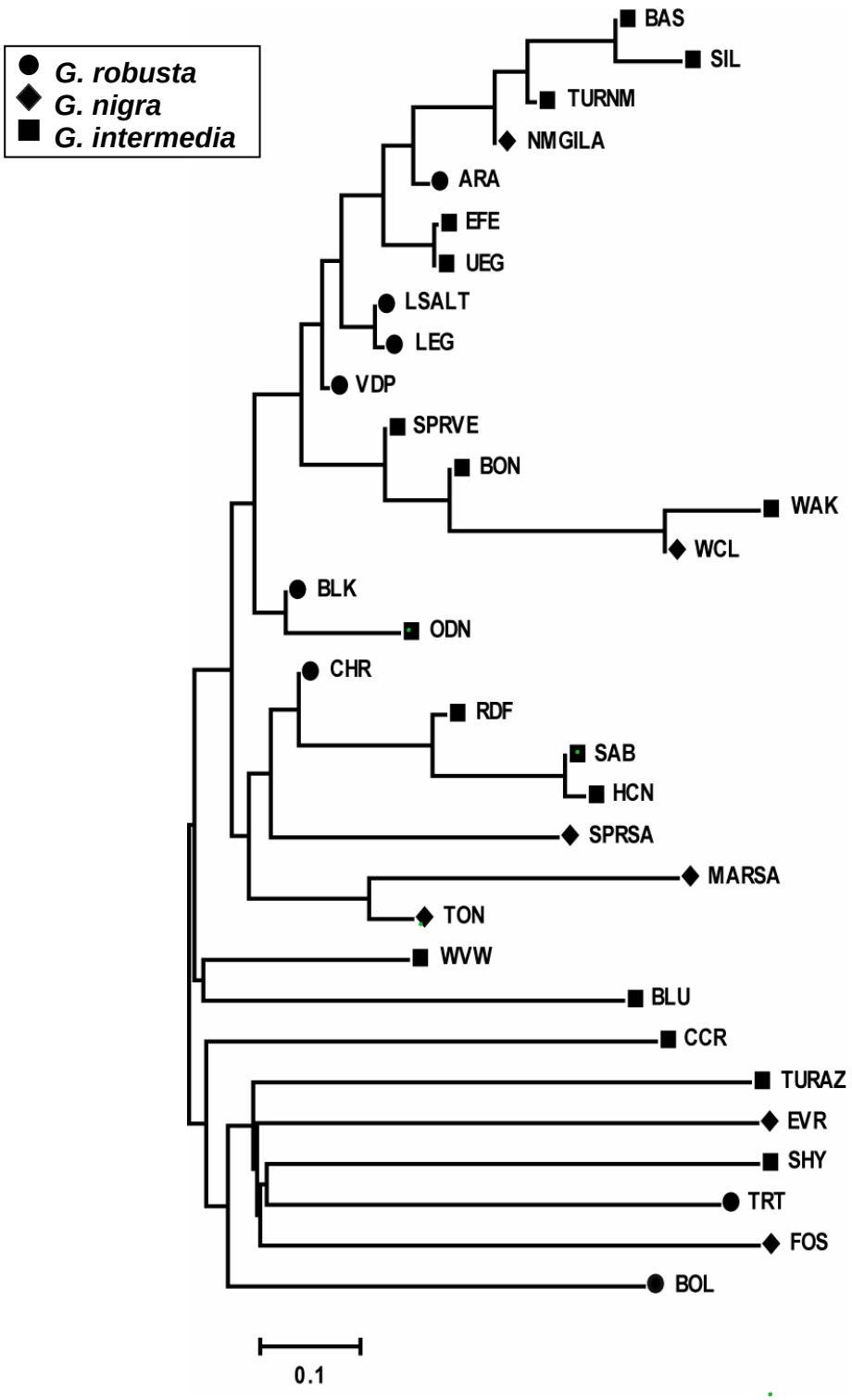

FIG. 9. Population tree of lower Colorado River basin *Gila* constructed by F_{ST} estimates of sequence divergence at ND2 using the neighbor-joining method. Circles, diamonds and squares represent taxonomic classification as *Gila robusta, G. nigra and G. intermedia*, respectively. (see Appendix 1.)

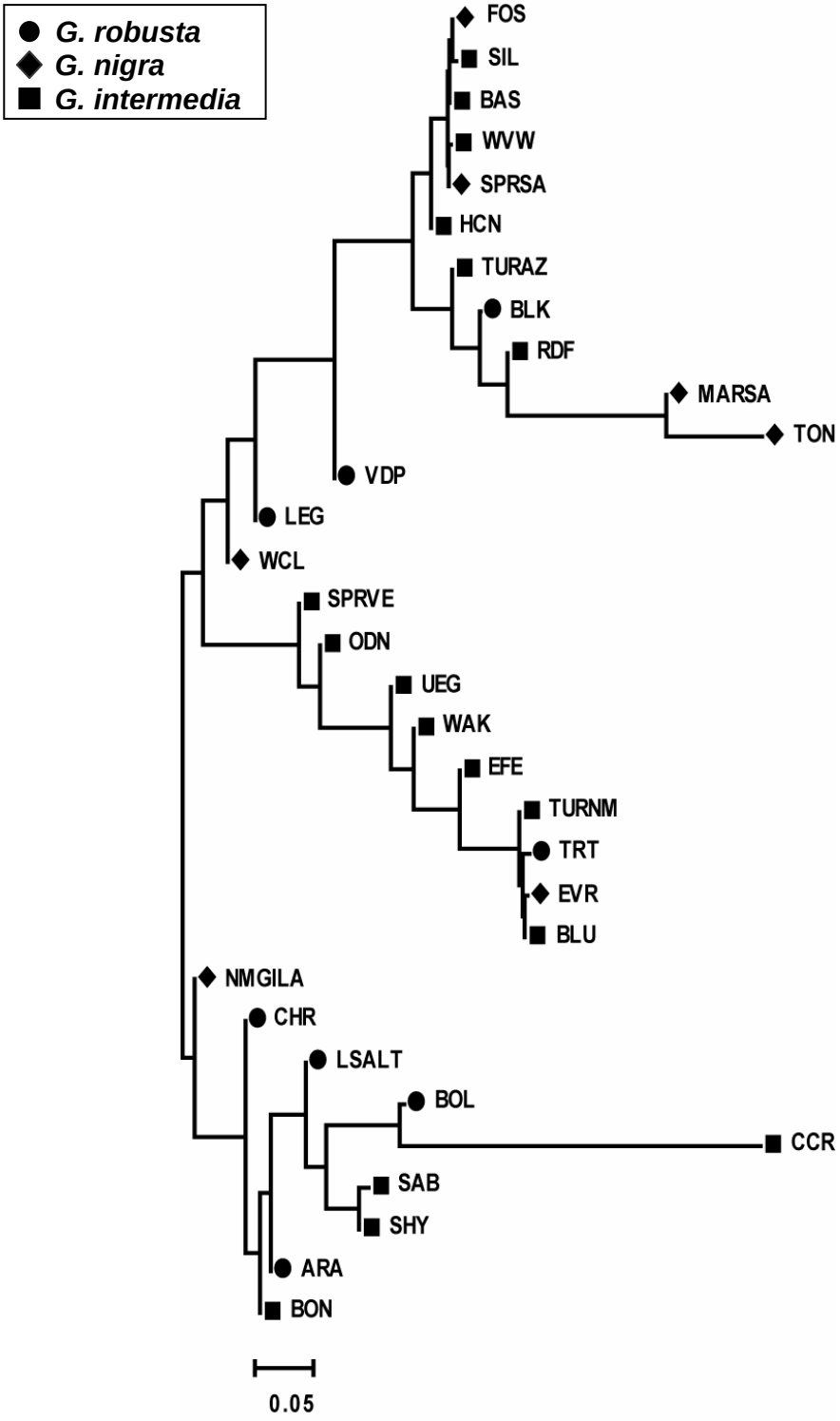

FIG. 10. Population tree of lower Colorado River basin *Gila* constructed by F_{ST} estimates of sequence divergence at Tpi-B using the neighbor-joining method. Circles, diamonds and squares represent taxonomic classification as *Gila robusta, G. nigra and G. intermedia*, respectively. (see Appendix 1.

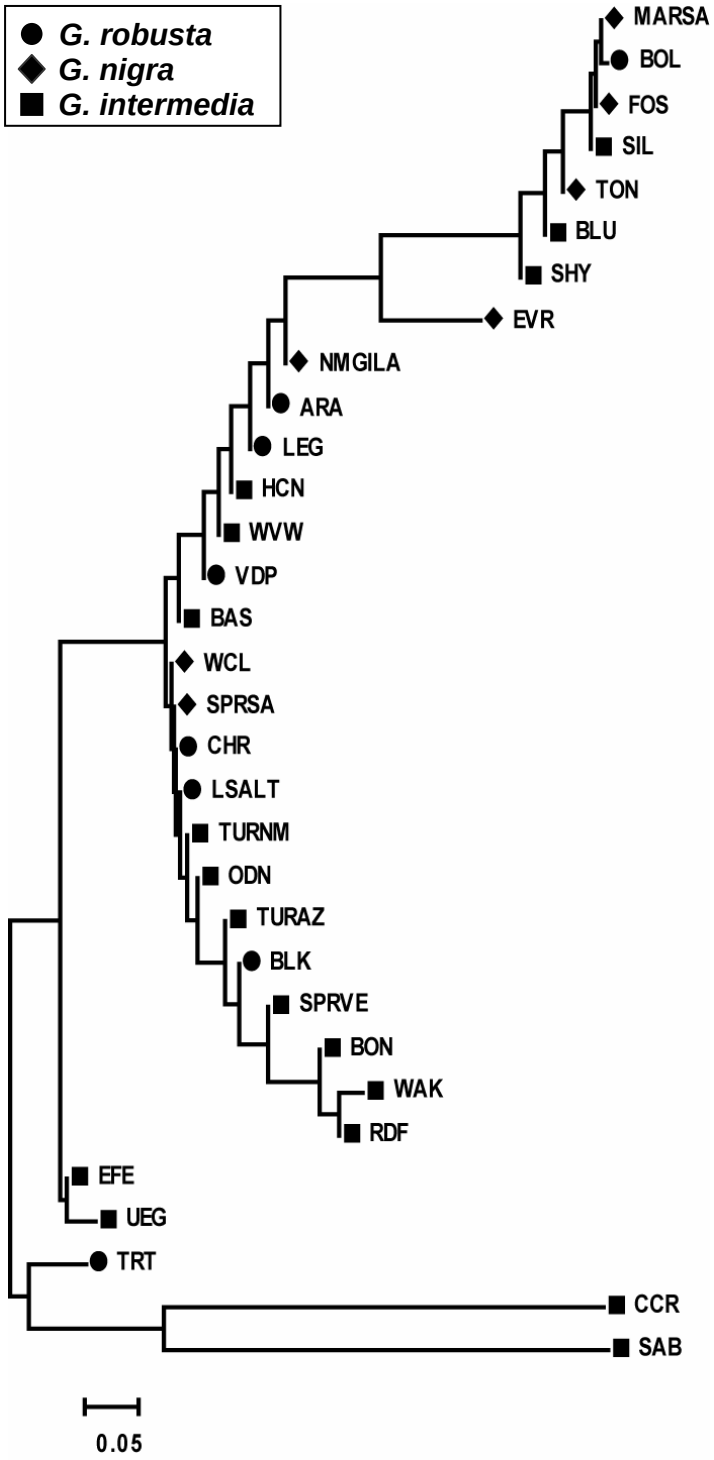

FIG. 11. Population tree of lower Colorado River basin *Gila* constructed by F_{ST} estimates of sequence divergence at S7 using the neighbor-joining method. Circles, diamonds and squares represent which taxonomic classification as *Gila robusta, G. nigra and G. intermedia*, respectively. (see Appendix 1.)

TABLE 1. Primers and amplification parameters used for SSCP and sequencing reactions.

gene	*primer name*	*primer sequence*	*PCR: denaturing, annealing, and extension (°C, time)*	*primer references*
ND2	ND2-F_{ROBSEQ}	5' GAGGAGGGCTAAACCAGACC 3' (external primer for sequencing)	94, 1 m; 48, 1 m; 72, 2 m	this study
	ND2-F_{ROB}	5' CTCCGCCACAAAGGTTAGCA 3'		this study
	ND2-E_{GILA}	5' TTCTGCTTAGAGCTTTGAAGGC 3'		Gerber et al. 2001
ND2	ND2-AA_{ROBSEQ}	5' CACTGGCTACTGGCTTGAAT 3' (external primer for sequencing)	94, 1 m; 48, 1 m; 72, 2 m	this study
	ND2-AA_{GILA}	5' TATCCCCCTAATGGCACAG 3'		Gerber et al. 2001
	ND2-C_{GILA}	5' ATGTGCAATTGAGGAGTAGG 3'		Gerber et al. 2001
S7	S7-$2R_{GILA}$	5' AACTCCAAGGGTGTTCTTAGCA 3'	94, 30 s; 53, 30 s; 72, 1 m, 30 s	this study
	S7RPEX1F	5' TGGCCTCTTCCTTGGCCGTC 3'		Chow and Hazama 1998
	S7RPEX2R	5' AACTCGTCTGGCTTTTCGCC 3' (external primer for nesting)		Chow and Hazama 1998
TPI-B	TPIB-F2	5' GCATYGGGGAGAAGCTRGAT 3'	94, 30 s; 57, 30 s; 72, 1 m, 30 s	Merritt and Quattro 2001
	TPIB-R3	5' TCCCGGAGCTTGTCATGCAC 3'		Johnson et al. 2004

TABLE 2. Geographic distribution of 39 haplotypes from the mitochondrial gene ND2 (788 base pairs). Species are identified as *Gila robusta* (R), *G. nigra* (N) and *G. intermedia* (I). (see Appendix 1. for population abbreviations)

	population																
	ara	bas	blk	blu	bol	bon	ccr	chr	efe	evr	fos	hcn	leg	lsal	mar	nmgila	odn
Haplotype	R	I	R	I	R	I	I	R	I	N	N	I	R	R	N	N	I
AA	11	19	8	.	.	6	.	2	5	.	.	.	9	13	.	14	14
AB	8	.	.	.	.	.	.	4	10	.	.	22	1	7	.	3	.
AD	.	.	.	.	11	.	.	.	.	.	.	.	.	.	.	.	.
AE	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
AF	1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
AG	.	.	3	.	.	12	.	.	.	.	.	.	7	8	.	1	.
AH	.	.	.	.	.	.	.	.	5	.	.	.	.	1	.	.	.
AJ	.	.	5	.	.	.	.	.	.	.	.	.	.	.	.	.	.
AK	.	.	.	.	.	.	.	1	.	.	.	.	.	.	24	.	.
AN	.	.	.	.	.	.	.	3	.	.	.	.	.	.	.	.	.
AP	.	.	.	18	.	.	.	.	.	.	.	.	.	.	.	.	.
AR	2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
AS	.	.	.	1	.	.	.	.	.	.	.	.	.	.	.	.	.
AT	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
BA	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
BG	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
CA	2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
CB	.	.	.	.	.	.	.	.	.	.	.		.	.	.	.	
CK	.	.	.	.	.	.	.	.	.	.	.	.	.	.	3	.	.
DB	.	.	.	.	.	.	.	.	.	20	.	.	.	.	.	.	.

TABLE 2. (continued), *mtDNA Haplotypes*

	population																
	ara	bas	blk	blu	bol	bon	ccr	chr	efe	evr	fos	hcn	leg	lsal	mar	nmgila	odn
Haplotype	R	I	R	I	R	I	I	R	I	N	N	I	R	R	N	N	I
EB	.	.	.	.	.	.	.	.	.	.		.	.	.	.	.	.
FA	.	.	3	.	.	.	.	.	.	.	.	.	.	.	.	.	.
FC	.	1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	6
GI	.	.	.		.	.	.	3	.	.	.	.	.	.	.	.	.
GK	.	.	.	.	.	.	.	3	.	.	.	.	.	.	.	.	.
HB	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HO	.	.	.	.	.	.	.	3	.	.	.	.	.	.	.	.	.
IB	.	.	.	.	.	.	.	1	.	.	.	.	.	.	.	.	.
IK	.	.	.	.	.	.	.	1	.	.	.	.	.	.	.	.	.
JB	.	.	.	.	.	.	20	.	.	.	.	.	.	.	.	.	.
KM	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
LA	.	.	.	.	.	2	.	.	.	.	.	.	3	.	.	.	.
MB	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
NA	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
NL	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
OB	.	.	5	.	.	.	.	.	.		.	.	.	.	.	.	.
PD	.	.	.	.	19	.	.	.	.	.	.	.	.	.	.	.	.
QG	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
RA	.	.	.	.	.	.	.	.	.	.	26	.	.	.	.	.	.
total	24	20	24	19	30	20	20	21	20	20	26	22	20	29	27	18	20

TABLE 2. (continued), *mtDNA Haplotypes*

	population															
	rdf	sab	shy	sil	sprsa	sprve	ton	trt	turaz	turnm	ueg	vdp	wak	wcl	wvw	
Haplotype	I	I	I	I	N	I	N	R	I	I	I	R	I	N	I	*total*
AA	.	.	.	29	.	5	.	.	.	13	2	7	.	4	.	161
AB	16	14	.	.	7	1	5	.	.	.	9	.	.	.	.	107
AD	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	11
AE	.	.	.	.	.	.	.	.	.	.	.	2	.	.	12	14
AF	.	.	.	.	.	.	.	.	.	.	.	5	.	.	1	7
AG	.	.	.	.	.	3	3	.	.	1	.	6	15	25	.	84
AH	.	.	.	.	.	.	.	.	.	.	6	.	.	.	.	12
AJ	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	5
AK	.	.	.	.	.	.	8	.	.	.	.	.	.	.	.	33
AN	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	3
AP	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	18
AR	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	2
AS	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1
AT	.	.	.	.	.	.	.	.	.	.	.	.	.	.	2	2
BA	.	.	.	.	.	3	.	.	.	.	.	.	.	.	.	3
BG	.	.	.	.	.	2	.	.	.	.	.	.	.	.	.	2
CA	.	.	.	.	.	.	.	.	.	.	.		.	.	.	2
CB	.		.	.	13	.	.		.	.	.	.	.	.	.	13
CK	.	.	.	.	.	.	.	.	.	.	.		.	.	.	3
DB	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	20

TABLE 2. (continued), *mtDNA Haplotypes*

	population															
	rdf	sab	shy	sil	sprsa	sprve	ton	trt	turaz	turnm	ueg	vdp	wak	wcl	wvw	
Haplotype	I	I	I	I	N	I	N	R	I	I	I	R	I	N	I	*total*
EB	.	.	.	.	.	.	.	.	.	.	.	.	.	.	5	5
FA	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	3
FC	4	.	.	.	.	.	.	.	18	.	.	.	.	.	.	29
GI	.	.	.	.	.	.	.	.	.	.	.		.	.	.	3
GK	.	.		.	.	.	.	.	.	.	.	.	.	.	.	3
HB	.	.	.	.	.	1	.	.	.	.	.	.	.	.	.	1
HO	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	3
IB	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1
IK	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1
JB	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	20
KM	.	.	11	.	.	.	.	.	.	.	.	.	.	.	.	11
LA	.	.	.	.	.	.	.	.	.	4	.	.	.	.	.	9
MB	.	.	.	.	.	.	.	.	.	.	1	.	.	.	.	1
NA	.	.	.	.	.	.	.	16	.	.	.	.	.	.	.	16
NL	.	.	.	.	.	.	.	14	.	.	.	.	.	.	.	14
OB		.	.	.	.	.	.	.	.	.	.	.	.		.	5
PD	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	19
QG	.	.	.	.	.	5	.	.	.	.	.	.	.	.	.	5
RA	.	.	.	.	.	.	.	.	.	.		.	.	.	.	26
total	20	14	11	29	20	20	16	30	18	18	18	20	15	29	20	678

TABLE 3. Geographic distribution of 10 genotypes (a) and alleles (b) from the triosephosphate isomerase-B gene (Tpi-B, 326 bp). Species are identified as *Gila robusta* (R), *G. nigra* (N) and *G. intermedia* (I). (see Appendix 1. for population abbreviations)

a)

	population															
	ara	bas	blk	blu	bol	bon	ccr	chr	efe	evr	fos	hcn	leg	lsal	mar	nmgila
Genotype	R	I	R	I	R	I	I	R	I	N	N	I	R	R	N	*I*
AA	7	5	.	19	4	9	.	2	17	20	9	5	2	2	2	3
AB	1	13	2	.	.	1	.	2	2	.	10	10	2	1	7	2
AC	7	.	2	.	14	1	.	6	1	.	.	3	6	11	.	3
AE	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
BB	1	2	7	.	.	.	.	1	.	.	7	3	2	.	15	1
BC	2	.	14	.	.	2	.	2	.	.	.	1	4	8	3	4
BE	.	.		.	.	.	.	.	.	.	.	.	.	.	.	.
CC	7	.	1	.	12	7	20	7	.	.	.	.	3	7	.	5
CD	.	.	.	.	.	.	.	1	.	.	.	.	.	.	.	.
CE	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
total	25	20	26	19	30	20	20	21	20	20	26	22	19	29	27	18

TABLE 3. (continued), *Tpi-B Genotypes*

a)

	population																
	odn	rdf	sab	shy	sil	sprsa	sprve	ton	trt	tur-az	tur-nm	ueg	vdp	wak	wcl	wvw	
Genotype	R	I	R	I	R	I	I	R	I	N	N	I	R	R	N	*I*	*total*
AA	10	1	1	1	10	7	10	.	28	.	18	12	1	12	5	1	223
AB	2	4	.	.	16	9	2	1	.	7	.	1	4	.	5	15	119
AC	6	2	10	8	.	.	6	.	.	2	.	4	1	3	7	1	104
AE	.	.	.	.	.	.	.	.	.	.	.	.	1	.	.	.	1
BB	.	7	.	.	3	4	.	13	.	5	.	.	4	.	.	3	78
BC	1	6	.	.	.	.	1	3	.	3	.	.	2	.	9	.	65
BE	.	.	.	.	.	.	.	.	.	.	.	.	1	.	.	.	1
CC	1	.	3	2	.	.	1	.	1	1	.	.	3	.	3	.	84
CD	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1
CE	.	.	.	.	.	.	.	.	.	.	.	.	3	.	.	.	3
total	20	20	14	11	29	20	20	17	29	18	18	17	20	15	29	20	679

TABLE 3. (continued), *Tpi-B Alleles*

b)

	population															
	ara	bas	blk	blu	bol	bon	ccr	chr	efe	evr	fos	hcn	leg	lsal	mar	nmgila
Allele	R	I	R	I	R	I	I	R	I	N	N	I	R	R	N	I
A	22	23	4	38	22	20	.	12	37	40	28	23	12	16	11	11
B	5	17	30	.	.	3	.	6	2	.	24	17	10	9	40	8
C	23	.	18	.	38	17	40	23	1	.	.	4	16	33	3	17
D	.	.	.	.	.	.	.	1	.	.	.	.	.	.	.	.
E	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
total	50	40	52	38	60	40	40	42	40	40	52	44	38	58	54	36

	population																
	odn	rdf	sab	shy	sil	sprsa	sprve	ton	trt	tur-az	tur-nm	ueg	vdp	wak	wcl	wvw	
Allele	I	I	I	I	I	N	I	N	R	I	I	I	R	I	N	I	*total*
A	28	8	12	10	36	23	28	1	56	9	36	29	8	27	22	18	670
B	3	24	.	.	22	17	3	30	.	20	.	1	15	.	14	21	341
C	9	8	16	12	.	.	9	3	2	7	.	4	12	3	22	1	341
D	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1
E	.	.	.	.	.	.	.	.	.	.	.	.	5	.	.	.	5
total	40	40	28	22	58	40	40	34	58	36	36	34	40	30	58	40	1358

TABLE 4. Geographic distribution of 25 genotypes (a) and alleles (b) from the S7 ribosomal gene (358 bp). Species are identified as *Gila robusta* (R), *G. nigra* (N) and *G. intermedia* (I). (see Appendix 1. for population abbreviations)

a)

	population															
	ara	bas	blk	blu	bol	bon	ccr	chr	efe	evr	fos	hcn	leg	lsalt	mar	nmgila
Genotype	I	I	I	I	I	N	I	N	R	I	I	I	R	I	N	I
AA	12	7	4	15	.	.	.	8	2	10	26	12	5	8	27	11
AB	8	11	12	.	.	9	.	8	5	.	.	9	7	12	.	5
AD	.	.	.	.	.	.	6	.	.	.	.	.	.	.	.	.
AF	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
AG	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
AI	.	.	.	.	.	.	.	.	.	9	.	.	.	.	.	.
AJ	3	.	.	.	.	.	.	.	1	.	.	.	4	1	.	1
AK	.	.	.	.	.	.	.	.	2	.	.	.	2	.	.	.
AL	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
AM	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
BB	1	2	8	.	.	6	.	5	2	.	.	1	1	7	.	1
BC	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

TABLE 4. (continued), S7 *Genotypes*

a)

	population															
	ara	bas	blk	blu	bol	bon	ccr	chr	efe	evr	fos	hcn	leg	lsalt	mar	nmgila
Genotype	I	I	I	I	I	N	I	N	R	I	I	I	R	I	N	I
BH	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
BJ	.	.	.	.	.	.	.	.	5	.	.	.	.	.	.	.
BK	.	.	.	.	.	.	.	.	1	.	.	.	.	.	.	.
BM	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
CC	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
DD	.	.	.	.	.	.	14	.	.	.	.	.	.	.	.	.
EE	.	.	.	.	30	.	.	.	.	.	.	.	.	.	.	.
FG	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
GG	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
II	.	.	.	.	.	.	.	.	.	1	.	.	.	.	.	.
JJ	.	.	.	.	.	.	.	.	1	.	.	.	.	.	.	.
JK	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
KK	.	.	.	.	.	.	.	.	1	.	.	.	.	.	.	.
total	24	20	24	15	30	15	20	21	20	20	26	22	19	28	27	18

TABLE 4. (continued), S7 *Genotypes*

a)

								population									
	odn	rdf	sab	shy	sil	sprsa	sprve	ton	trt	tur-az	tur-nm	ueg	vdp	wak	wcl	wvw	
Genotype	I	I	I	I	I	N	I	N	R	I	I	I	R	I	N	I	*total*
AA	6	1	.	11	28	20	3	16	.	4	5	1	10	1	8	12	273
AB	7	9	.	.	.	.	8	.	.	7	8	4	8	5	15	5	162
AD	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	6
AF	.	.	1	.	.	.	.	.	.	.	.	.	.	.	.	.	1
AG	.	.	4	.	.	.	.	.	.	.	.	.	.	.	.	.	4
AI	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	.	10
AJ	.	.	.	.	.	.	.	.	.	.	.	4	.	.	.	.	14
AK	.	.	.		.	.	.	.	.	.	.	3	.	.	.	.	7
AL	.	.	.	.	1	.	.	.	.	.	.	.	.	.	.	.	1
AM	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	.	1
BB	7	9	.	.	.	.	7	.	12	7	5	1	2	9	2	3	98
BC	.	.	.	.	.	.	.	.	15	.	.	.	.	.	.	.	15

TABLE 4. (continued), *S7 Genotype*

a)

	population																
	odn	rdf	sab	shy	sil	sprsa	sprve	ton	trt	tur-az	tur-nm	ueg	vdp	wak	wcl	wvw	
Genotype	I	I	I	I	I	N	I	N	R	I	I	I	R	I	N	I	*total*
BH	.		.	.	.	.	1	.	.	.	.	.	.	.	.	.	1
BJ	.	.	.	.	.	.	.	.	.	.	.	1	.	.	.	.	6
BK	.	.	.	.	.	.	.	.	.	.	.	1	.	.		.	2
BM	.	.	.	.	.	.	.	.	.	.	.	.	.	.	2	.	2
CC	.	.	.	.	.	.	.	.	2	.	.	.	.	.	.	.	2
DD	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	14
EE	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	30
FG	.	.	1	.	.	.	.	.	.	.	.	.	.	.	.	.	1
GG	.	.	8	.	.	.	.	.	.	.	.	.	.	.	.	.	8
II	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1
JJ	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1
JK	.	.	.	.	.	.	.	.	.	.	.	1	.	.	.	.	1
KK	.	.	.	.	.	.	.	.	.	.	.	1	.	.	.	.	2
total	20	19	14	11	29	20	19	16	29	18	18	17	20	15	29	20	663

TABLE 4. (continued), *S7 Alleles*

b)

	population															
	ara	bas	blk	blu	bol	bon	ccr	chr	efe	evr	fos	hcn	leg	lsalt	mar	nmgila
Allele	R	I	R	I	R	I	I	R	I	N	N	I	R	R	N	I
A	35	25	20	30	.	9	6	24	12	29	52	33	23	29	54	28
B	10	15	28	.	.	21	.	18	15	.	.	11	9	26	.	7
C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
D	.	.	.	.	.	.	34	.	.	.	.	.	.	.	.	.
E	.	.	.	.	60	.	.		.	.	.	.	.	.	.	.
F	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
H	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
I	.	.	.	.	.	.	.	.	.	11	.	.	.	.	.	.
J	3	.	.	.	.	.	.	.	8	.	.	.	4	1	.	1
K	.	.	.	.	.	.	.	.	5	.	.	.	2	.	.	.
L	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
M	.	.	.	.		.	.	.	.	.	.	.	.	.	.	.
total	48	40	48	30	60	30	40	42	40	40	52	44	38	56	54	36

TABLE 4. (continued), *S7 Alleles*

b)

									population								
	odn	rdf	sab	shy	sil	sprsa	sprve	ton	trt	tur-az	tur-nm	ueg	vdp	wak	wcl	wvw	
Allele	I	I	I	I	I	N	I	N	R	I	I	I	R	I	N	I	*total*
A	19	11	5	22	57	40	14	32	.	15	18	13	28	7	33	29	343
B	21	27	.	.	.	.	23	.	39	21	18	8	12	23	21	11	224
C		.	.	.	.	.	.	.	19	.	.	.	.	.	.	.	19
D	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0
E	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0
F	.	.	2	.	.	.	.	.	.	.	.	.	.	.	.	.	2
G	.	.	21	.	.	.	.	.	.	.	.	.	.	.	.	.	21
H	.	.	.	.	.	.	1	.	.	.	.	.	.		.	.	1
I	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	.	1
J	.	.	.	.	.	.	.	.	.	.	.	6	.	.	.	.	6
K		.	.	.	.	.	.	.	.	.	.	7	.	.	.	.	7
L	.	.	.	.	1	.	.	.	.		.		.	.	.	.	1
M		.	.	.	.		.	.	.	.	.	.	.	.	3	.	3
total	40	38	28	22	58	40	38	32	58	36	36	34	40	30	58	40	618

TABLE 5. MtDNA genetic diversity within 32 populations of *Gila* in the lower Colorado River basin. Standard errors are provided in parentheses. (see Appendix 1. for population abbreviations)

site code	*sample size*	N_h	*gene diversity (h)*	*nucleotide diversity (π)*
ARA	24	5	0.6920 (0.0646)	0.001168 (0.000932)
BAS	20	2	0.1000 (0.0880)	0.000508 (0.000549)
BLK	24	5	0.8043 (0.0410)	0.003931 (0.002358)
BLU	19	2	0.1053 (0.0920)	0.000134 (0.000259)
BON	20	3	0.5684 (0.0863)	0.002164 (0.001473)
BOL	30	2	0.4805 (0.0519)	0.000610 (0.000604)
CHR	21	9	0.9095 (0.0298)	0.004581 (0.002701)
CCR	20	1	0.0 (0.0)	0.0 (0.0)
LEG	20	4	0.6842 (0.0636)	0.002197 (0.001491)
UEG	18	4	0.6601 (0.0781)	0.001402 (0.001075)
EFE	20	3	0.6579 (0.0648)	0.001169 (0.000940)
EVR	20	1	0.0 (0.0)	0.0 (0.0)
FOS	26	1	0.0 (0.0)	0.0 (0.0)
NMGILA	18	3	0.3856 (0.1280)	0.000746 (0.00070)
HCN	22	1	0.0 (0.0)	0.0 (0.0)
MAR	27	2	0.2051 (0.0947)	0.000260 (0.000367)
ODN	20	2	0.4421 (0.0875)	0.002244 (0.001515)
RED	20	2	0.3368 (0.1098)	0.001282 (0.001003)
SAB	14	1	0.0 (0.0)	0.0 (0.0)
SHY	11	1	0.0 (0.0)	0.0 (0.0)
SIL	29	1	0.0 (0.0)	0.0 (0.0)
SPRSA	20	2	0.4789 (0.0720)	0.000608 (0.000613)
SPRVE	20	7	0.8579 (0.0418)	0.003106 (0.001959)
TON	16	3	0.6583 (0.0748)	0.001502 (0.001138)
TRT	30	2	0.5149 (0.0268)	0.000653 (0.000630)
TURAZ	18	1	0.0 (0.0)	0.0 (0.0)
TURNM	18	3	0.4510 (0.1174)	0.000887 (0.000785)
VDP	20	4	0.7526 (0.0455)	0.002712 (0.001757)
LSALT	29	4	0.6872 (0.0462)	0.001794 (0.001258)
WAK	15	1	0.0 (0.0)	0.0 (0.0)
WCL	29	2	0.2463 (0.0935)	0.000938 (0.000797)
WVW	20	4	0.5947 (0.0977)	0.002832 (0.001819)

TABLE 6. Tpi-B genetic diversity within 32 populations of *Gila* in the lower Colorado River basin. Standard errors are provided in parentheses. (see Appendix 1. for population abbreviations)

site code	*No. of gene copies*	N_a	*gene diversity (h)*	*nucleotide diversity (π)*
ARA	50	3	0.5967 (0.0327)	0.002694 (0.002132)
BAS	40	2	0.5013 (0.0293)	0.003090 (0.002357)
BLK	52	3	0.5520 (0.0444)	0.004496 (0.003074)
BLU	38	1	0.0 (0.0)	0.0 (0.0)
BON	40	3	0.5782 (0.0384)	0.002421 (0.001993)
BOL	60	2	0.4723 (0.0350)	0.001453 (0.001422)
CHR	42	4	0.6121 (0.0541)	0.003301 (0.002466)
CCR	40	1	0.0 (0.0)	0.0 (0.0)
LEG	38	3	0.6714 (0.0279)	0.004000 (0.002840)
UEG	24	3	0.2656 (0.0921)	0.001021 (0.001163)
EFE	40	3	0.1449 (0.0737)	0.000754 (0.000972)
EVR	40	1	0.0 (0.0)	0.0 (0.0)
FOS	52	2	0.5068 (0.0173)	0.003124 (0.002361
NMGILA	36	3	0.6524 (0.0395)	0.003773 (0.002726)
HCN	44	3	0.5825 (1.0406)	0.003512 (0.002575)
MAR	54	3	0.4144 (0.0689)	0.002743 (0.002156)
ODN	40	3	0.4654 (0.0763)	0.001978 (0.001744)
RED	40	3	0.5744 (0.0637)	0.004049 (0.002862)
SAB	28	2	0.5079 (0.0367)	0.001563 (0.001517)
SHY	22	2	0.5195 (0.0379)	0.001598 (0.001554)
SIL	58	2	0.4791 (0.0329)	0.002953 (0.002266)
SPRSA	40	2	0.5013 (0.0293)	0.003090 (0.002357)
SPRVE	40	3	0.4654 (0.0763)	0.001978 (.001744)
TON	34	3	0.2193 (0.0897)	0.001832 (0.001667)
TRT	58	2	0.0678 (0.0446)	0.000208 (0.000480)
TURAZ	36	3	0.6079 (0.0576)	0.004126 (0.002910)
TURNM	36	1	0.0 (0.0)	0.0 (0.0)
VDP	40	4	0.7321 (0.0319)	0.006381 (0.004048)
LSALT	58	3	0.5862 (0.0461)	0.003183 (0.002388)
WAK	30	2	0.1862 (0.0881)	0.000573 (0.000839)
WCL	58	3	0.6655 (0.0193)	0.003775 (0.002697)
WVW	40	3	0.5346 (0.0311)	0.003307 (0.002473)

TABLE 7. S7 genetic diversity within 32 populations of *Gila* in the lower Colorado River basin. Standard errors are provided in parentheses. (see Appendix 1. for population abbreviations)

site code	*No. of gene copies*	N_a	*gene diversity (h)*	*nucleotide diversity (π)*
ARA	48	3	0.4300 (0.0725)	0.001275 (0.001271)
BAS	40	2	0.4808 (0.0418)	0.001343 (0.001318)
BLK	48	2	0.4965 (0.0278)	0.001387 (0.001339)
BLU	30	1	0.0 (0.0)	0.0 (0.0)
BON	30	2	0.4598 (0.0612)	0.001214 (0.001249)
BOL	60	1	0.0 (0.0)	0.0 (0.0)
CHR	42	2	0.5017 (0.0275)	0.001402 (0.001352)
CCR	40	2	0.2615 (0.0796)	0.002192 (0.001808)
LEG	38	4	0.5789 (0.0711)	0.002118 (0.001770)
UEG	34	4	0.7469 (0.0348)	0.003545 (0.002544)
EFE	40	4	0.7321 (0.0319)	0.003370 (0.002441)
EVR	40	1	0.0 (0.0)	0.002285 (0.001860)
FOS	52	1	0.0 (0.0)	0.0 (0.0)
NMGILA	36	3	0.3667 (0.0851)	0.001055 (0.001141)
HCN	44	2	0.3837 (0.0664)	0.001072 (0.001145)
MAR	54	1	0.0 (0.0)	0.0 (0.0)
ODN	40	2	0.5115 (0.0195)	0.001429 (0.001370)
RED	38	2	0.4225 (0.0639)	0.001180 (0.001219)
SAB	28	3	0.4153 (0.1003)	0.001936 (0.001684)
SHY	22	1	0.0 (0.0)	0.0 (0.0)
SIL	58	2	0.0345 (0.0330)	0.000096 (0.000308)
SPRSA	40	1	0.0 (0.0)	0.001400 (0.001353)
SPRVE	38	3	0.5107 (0.0480)	0.001518 (0.001425)
TON	32	1	0.0 (0.0)	0.0 (0.0)
TRT	58	2	0.4483 (0.0439)	0.003757 (0.002621)
TURAZ	36	2	0.5000 (0.0335)	0.001397 (0.001354)
TURNM	36	2	0.5143 (0.0199)	0.001437 (0.001378)
VDP	40	2	0.4308 (0.0599)	0.001203 (0.001232)
LSALT	56	3	0.5253 (0.0221)	0.001515 (0.001244)
WAK	30	2	0.3701 (0.0841)	0.001034 (0.001133)
WCL	58	4	0.5517 (0.0407)	0.004311 (0.002915)
WVW	40	2	0.4090 (0.0651)	0.001142 (0.001193)

TABLE 8. Significant associations of allelic and genotypic cytonuclear tests for mitochondrial ND2 with two nuclear loci triosephosphate isomerase-B (Tpi-B) and the S7 ribosomal protein and deviations from Hardy-Weinberg (H-W) expectations for 32 localities of *Gila* (site code provided in Appendix 1.). Significant associations (Bonferroni corrected for two loci) are marked with asterisks (P<0.025).

	Tpi-B			*S7*		
	Cytonuclear Disequil.		*H-W*	*Cytonuclear Disequil.*		*H-W*
site code	*P -genotypic*	*P -allelic*	*P*	*P -genotypic*	*P -allelic*	*P*
ARA	0.640	0.772	0.124	0.9158	0.658	1.000
BAS	0.857	1.000	0.358	0.4516	0.132	0.647
BLK	0.203	0.133	0.290	0.6839	0.234	1.000
BLU	n/a	n/a	n/a	n/a	n/a	n/a
BON	0.603	0.193	0.000*	0.7353	0.774	1.000
BOL	0.144	0.144	1.000	n/a	n/a	n/a
CHR	1.000	0.493	0.730	0.9017	0.577	0.376
CCR	n/a	n/a	n/a	n/a	n/a	1.000
LEG	0.831	0.997	0.825	0.1711	0.946	0.719
UEG	0.834	0.818	1.000	0.1217	0.494	0.880
EFE	0.111	0.325	1.000	0.8584	0.962	0.394
EVR	n/a	n/a	n/a	n/a	n/a	n/a
FOS	n/a	n/a	0.259	n/a	n/a	n/a
NMGILA	0.984	1.000	0.502	0.008*	0.003*	0.632
HCN	n/a	n/a	0.879	n/a	n/a	1.000
MAR	0.715	0.445	0.621	n/a	n/a	n/a
ODN	0.087	0.135	1.000	0.4105	0.485	0.199
RED	0.326	0.145	0.916	1.000	0.669	1.000
SAB	n/a	n/a	0.268	n/a	n/a	0.620
SHY	n/a	n/a	0.260	n/a	n/a	n/a
SIL	n/a	n/a	0.457	n/a	n/a	1.000
SPRSA	0.428	1.000	0.676	n/a	n/a	n/a
SPRVE	0.611	0.543	1.000	0.1595	0.242	1.000
TON	0.211	0.102	1.000	n/a	n/a	n/a
TRT	0.031	0.001*	0.017*	0.0589	0.095	0.671
TURAZ	n/a	n/a	0.542	n/a	n/a	0.380
TURNM	n/a	n/a	n/a	0.4975	0.067	0.656
VDP	0.730	0.343	0.322	0.8773	0.769	1.000
LSALT	0.469	0.801	0.341	0.5422	0.189	0.703
WAK	n/a	n/a	1.000	n/a	n/a	1.000
WCL	0.025	0.182	0.253	0.3447	0.190	0.493
WVW	0.000*	0.032	0.022*	0.2711	0.042	0.107

TABLE 9. Analysis of molecular variance (AMOVA) for ND2, Tpi-B and S7 loci. Groups for species AMOVA (*Gila robusta, G. nigra* and *G. intermedia*) are defined in Minckley and DeMarais 2000. Subbasin groups for geographical AMOVA (Verde River, Salt River, San Pedro River, Santa Cruz River, Agua Fria River, Gila River mainstem and Bill Williams River) are defined in Appendix 1. Terms in parentheses provides % of variance accounted for by this measure and superscripted note indicates significance level.

AMOVA by species (morphology)

	between species	*among pop. within species*	*total among populations*
gene	F_{CT}	F_{SC}	F_{ST}
ND2	0.020 ^{NS} (2.0%)	0.666* (65.2%)	0.672* (67.2%)
Tpi-B	0.045 ^{NS} (4.5%)	0.328* (31.3%)	0.358* (35.8%)
S7	0.102 ^ (10.2%)	0.378* (33.9%)	0.441* (44.8%)

* = $P < 0.001$; ^ = $P < 0.050$; ^{NS} = not significant

AMOVA by geography (subbasin)

	between subbasins	*among pop. within subbasin*	*among all populations*
gene	F_{CT}	F_{SC}	F_{ST}
ND2	0.071^{NS} (7.1%)	0.648* (60.2%)	0.672* (67.2%)
Tpi-B	0.071^ (7.1.0%)	0.273* (25.3%)	0.324* (32.4.%)
S7	0.128^ (12.8%)	0.344* (30.1%)	0.429* (42.9%)

* = $P < 0.001$; ^ = $P < 0.050$; ^{NS} = not significant

TABLE 10. Results and description of nested clade analysis (NCA) of ND2 haplotypes showing clade structure, number of haplotypes (N_a), and evolutionary processes for each clade (Fig. 7). Locality abbreviations provided in Appendix 1. Clades without genetic or geographic variation are omitted.

Nesting Clade	N_a	*localities (n=)*	*NCA inference (Templeton 2005)*
one-step			
1-1	1	16: ARA, BAS, BLK, BON, CHR, EFE, LEG, LSAT, NMGILA, ODN, SIL, SPRVE, TURNM, UEG, VDP, WCL	Insufficient genetic resolution to discriminate between Range Expansion/Colonization and Restricted Dispersal/Gene Flow
1-3	2	4: BLK, EFE, LSALT, UEG	Long Distance Colonization or Past Fragmentation Followed by Range Expansion
1-10	3	11: BLK, BON, LEG, LSALT, NMGILA, SPRVE, TON, TURNM, VDP, WAK, WCL	Restricted Gene Flow with Isolation by Distance
1-11	3	2: CHR, BLK	Sampling design inadequate to discriminate between Contiguous Range Expansion, Long Distance Dispersal and Past Fragmentation
1-13	7	16: ARA, CCR, CHR, EFE, HCN, LEG, LSALT, NMGILA, RDF, SAB, SPRSA, SPRVE, TON, UEG, VDP, WVW	Restricted Gene Flow/Dispersal but with some Long Distance Dispersal
1-15	2	2: WVW,VDP	No significant values for Dc, Dn, or I-T
1-17	2	3: CHR, MAR, TON	No significant values for Dc, Dn, or I-T
two-step			
2-1	10	20: ARA, BAS, BLK, BON, CHR, EFE, LEG, LSAT, NMGILA, ODN, SHY, SIL, SPRVE, TON, TRT, TURNM, UEG, VDP, WCL, WVW	Allopatric Fragmentation

TABLE 10. (continued), *ND2 NCA results*

2-2	2	5: BAS, BLK, ODN, RDF, TURAZ	Allopatric Fragmentation
2-3	2	2: FOS, EVR	Allopatric Fragmentation
2-4	25	25: ARA, BLK, BOL, BON, BLU, CCR, CHR, EFE, HCN, LEG, LSALT, MAR, NMGILA, RDF, SAB, SPRSA, SPRVE, TON, TURNM, UEG, VDP, WCL, WVW	Insufficient genetic resolution to discriminate between Range Expansion/Colonization and Restricted Dispersal/Gene Flow
total cladogram	39	32: all populations	Restricted Gene Flow with Isolation by Distance

Clades without genetic or geographic variation: 1-2, 1-4, 1-5, 1-6, 1-7, 1-8, 1-9, 1-12, 1-14, 1-16, 1-18, 1-19

TABLE 11. Results and description of nested clade analysis (NCA) of Tpi-B haplotypes showing clade structure, number of haplotypes (N_a), and evolutionary processes for each clade (Fig. 8). Locality abbreviations provided in Appendix 1. Clades without genetic or geographic variation are omitted.

Nesting Clade	N_a	*localities (n=)*	*NCA inference (Templeton 2005)*
one-step			
1-1	2	32: all populations	No significant values for Dc, Dn, or I-T
1-2	2	23: ARA, BAS, BLK, BON, CHR, EFE, FOS, HCN, LEG, LSALT, MAR, NMGILA, ODN, RDF, SIL, SPRSA, SPRVE, TON, TURAZ, UEG, VDP, WCL, WVW	Restricted Gene Flow with Isolation by Distance
2-1	4	32: all populations	Restricted Gene Flow with Isolation by Distance
total cladogram	5	32: all populations	Inconclusive outcome

Clades without genetic or geographic variation: 1-3, 2-1

TABLE 12. Results and description of nested clade analysis (NCA) of S7 haplotypes showing clade structure, number of haplotypes (N_a), and evolutionary processes for each clade (Fig. 8). Locality abbreviations provided in Appendix 1. Clades without genetic or geographic variation are omitted.

Nesting Clade	N_a	*localities (n=)*	*NCA inference (Templeton 2005)*
one-step			
1-1	6	32: all populations	Range Expansion
1-2	2	4: CCR, EFE, LEG, UEG	Allopatric Fragmentation
1-3	2	7: ARA, EFE, LEG, LSALT, NMGILA, TRT, UEG	Past Fragmentation and/or Long Distance Colonization
total cladogram	13	32: all populations	Restricted Gene Flow with Isolation by Distance

Clades without genetic or geographic variation: 1-4, 1-5

TABLE 13. Pairwise F_{ST} values for 32 populations of *Gila* in the lower Colorado River from ND2 sequences. Significant values ($P < 0.05$) are marked by asterisks. (see Appendix 1. and Fig. 2. for localities)

	EVR	*LSALT*	*SPRVE*	*VDP*	*WAK*	*WVW*
EVR						
LSALT	0.800 *					
SPRVE	0.753 *	0.080 *				
VDP	0.758 *	0.017 NS	0.072 NS			
WAK	1.000 *	0.567 *	0.333 *	0.475 *		
WVW	0.782 *	0.437 *	0.419 *	0.321 *	0.698 *	
WCL	0.911 *	0.417 *	0.217 *	0.351 *	0.061 NS	0.650 *
FOS	1.000 *	0.763 *	0.736 *	0.742 *	1.000 *	0.806 *
MAR	0.971 *	0.622 *	0.603 *	0.570 *	0.958 *	0.654 *
SPRSA	0.935 *	0.425 *	0.437 *	0.362 *	0.897 *	0.507 *
TON	0.863 *	0.246 *	0.249 *	0.183 *	0.715 *	0.416 *
BLK	0.683 *	0.096 *	0.145 *	0.122 *	0.524 *	0.354 *
CHR	0.650 *	0.199 *	0.234 *	0.154 *	0.530 *	0.288 *
TURAZ	1.000 *	0.792 *	0.742 *	0.747 *	1.000 *	0.772 *
BAS	0.951 *	0.251 *	0.378 *	0.309 *	0.926 *	0.588 *
ODN	0.808 *	0.207 *	0.295 *	0.222 *	0.726 *	0.459 *
RDF	0.860 *	0.266 *	0.322 *	0.213 *	0.782 *	0.399 *
ARA	0.867 *	0.097 *	0.268 *	0.134 *	0.800 *	0.457 *
BLU	0.987 *	0.600 *	0.568 *	0.536 *	0.981 *	0.623 *
CCR	1.000 *	0.616 *	0.585 *	0.555 *	1.000 *	0.640 *
SHY	1.000 *	0.763 *	0.708 *	0.720 *	1.000 *	0.763 *
SAB	1.000 *	0.257 *	0.324 *	0.194 *	1.000 *	0.419 *
SIL	1.000 *	0.359 *	0.489 *	0.429 *	1.000 *	0.689 *
BON	0.819 *	0.096 NS	0.025 NS	0.085 NS	0.306 *	0.482 *
EFE	0.878 *	0.109 *	0.268 *	0.151 *	0.812 *	0.447 *
LEG	0.806 *	-0.011 NS	0.059 NS	0.038 NS	0.540 *	0.442 *
UEG	0.863 *	0.127 *	0.263 *	0.153 *	0.790 *	0.421 *
HCN	1.000 *	0.304 *	0.385 *	0.248 *	1.000 *	0.482 *
TURNM	0.923 *	0.213 *	0.326 *	0.266 *	0.876 *	0.558 *
NMGILA	0.928 *	0.110 NS	0.271 *	0.181 *	0.880 *	0.519 *
TRT	0.959 *	0.791 *	0.765 *	0.770 *	0.947 *	0.814 *
BOL	0.938 *	0.652 *	0.630 *	0.605 *	0.913 *	0.671 *

TABLE 13. (continued), *ND2 F_{ST} values*

	WCL	*FOS*	*MAR*	*SPRSA*	*TON*	*BLK*
EVR						
LSALT						
SPRVE						
VDP						
WAK						
WVW						
WCL						
FOS	0.915 *					
MAR	0.839 *	0.975 *				
SPRSA	0.750 *	0.943 *	0.799 *			
TON	0.560 *	0.882 *	0.351 *	0.469 *		
BLK	0.462 *	0.611 *	0.517 *	0.356 *	0.261 *	
CHR	0.493 *	0.650 *	0.368 *	0.287 *	0.152 *	0.126 *
TURAZ	0.907 *	1.000 *	0.970 *	0.931 *	0.856 *	0.638 *
BAS	0.780 *	0.921 *	0.869 *	0.749 *	0.614 *	0.119 *
ODN	0.637 *	0.753 *	0.667 *	0.500 *	0.397 *	0.103 *
RDF	0.652 *	0.877 *	0.670 *	0.404 *	0.259 *	0.230 *
ARA	0.666 *	0.834 *	0.703 *	0.450 *	0.360 *	0.101 *
BLU	0.835 *	0.989 *	0.924 *	0.826 *	0.684 *	0.479 *
CCR	0.850 *	1.000 *	0.944 *	0.855 *	0.712 *	0.492 *
SHY	0.905 *	1.000 *	0.971 *	0.934 *	0.858 *	0.601 *
SAB	0.745 *	1.000 *	0.880 *	0.588 *	0.304 *	0.235 *
SIL	0.857 *	1.000 *	0.953 *	0.880 *	0.759 *	0.193 *
BON	0.130 *	0.806 *	0.690 *	0.535 *	0.315 *	0.213 *
EFE	0.670 *	0.853 *	0.720 *	0.500 *	0.357 *	0.073 *
LEG	0.385 *	0.762 *	0.654 *	0.474 *	0.287 *	0.096 *
UEG	0.654 *	0.843 *	0.699 *	0.473 *	0.331 *	0.075 *
HCN	0.779 *	1.000 *	0.899 *	0.644 *	0.373 *	0.286 *
TURNM	0.731 *	0.879 *	0.829 *	0.697 *	0.559 *	0.119 *
NMGILA	0.711 *	0.896 *	0.819 *	0.655 *	0.492 *	0.077 *
TRT	0.897 *	0.950 *	0.934 *	0.903 *	0.859 *	0.682 *
BOL	0.826 *	0.945 *	0.873 *	0.790 *	0.701 *	0.558 *

TABLE 13. (continued), *ND2 F_{ST} values*

	CHR	*TURAZ*	*BAS*	*ODN*	*RDF*	*ARA*
EVR						
LSALT						
SPRVE						
VDP						
WAK						
WVW						
WCL						
FOS						
MAR						
SPRSA						
TON						
BLK						
CHR						
TURAZ	0.616 *					
BAS	0.296 *	0.945 *				
ODN	0.211 *	0.672 *	0.153 NS			
RDF	0.166 *	0.780 *	0.570 *	0.244 *		
ARA	0.175 *	0.861 *	0.188 *	0.161 *	0.286 *	
BLU	0.422 *	0.987 *	0.881 *	0.641 *	0.657 *	0.696 *
CCR	0.435 *	1.000 *	0.905 *	0.660 *	0.684 *	0.719 *
SHY	0.617 *	1.000 *	0.919 *	0.739 *	0.860 *	0.833 *
SAB	0.150 *	1.000 *	0.790 *	0.398 *	0.122 NS	0.319 *
SIL	0.394 *	1.000 *	0.019 NS	0.316 *	0.730 *	0.336 *
BON	0.289 *	0.810 *	0.484 *	0.366 *	0.409 *	0.357 *
EFE	0.173 *	0.871 *	0.266 *	0.185 *	0.285 *	0.039 NS
LEG	0.213 *	0.797 *	0.244 *	0.208 *	0.331 *	0.150 *
UEG	0.167 *	0.856 *	0.315 *	0.204 *	0.263 *	0.083 *
HCN	0.197 *	1.000 *	0.825 *	0.460 *	0.168 *	0.374 *
TURNM	0.283 *	0.918 *	0.060 NS	0.202 *	0.542 *	0.190 *
NMGILA	0.225 *	0.924 *	0.035 NS	0.153 *	0.459 *	0.048 NS
TRT	0.703 *	0.957 *	0.873 *	0.778 *	0.859 *	0.832 *
BOL	0.513 *	0.936 *	0.837 *	0.682 *	0.684 *	0.714 *

TABLE 13. (continued), *ND2 F_{ST} values*

	BLU	*CCR*	*SHY*	*SAB*	*SIL*	*BON*
EVR						
LSALT						
SPRVE						
VDP						
WAK						
WVW						
WCL						
FOS						
MAR						
SPRSA						
TON						
BLK						
CHR						
TURAZ						
BAS						
ODN						
RDF						
ARA						
BLU						
CCR	0.975 *					
SHY	0.987 *	1.000 *				
SAB	0.942 *	1.000 *	1.000 *			
SIL	0.980 *	1.000 *	1.000 *	1.000 *		
BON	0.666 *	0.684 *	0.782 *	0.449 *	0.606 *	
EFE	0.711 *	0.737 *	0.844 *	0.338 *	0.443 *	0.359 *
LEG	0.627 *	0.646 *	0.747 *	0.354 *	0.362 *	0.045 NS
UEG	0.685 *	0.711 *	0.827 *	0.303 *	0.491 *	0.352 *
HCN	0.954 *	1.000 *	1.000 *	0.000 NS	1.000 *	0.512 *
TURNM	0.831 *	0.854 *	0.873 *	0.699 *	0.150 *	0.406 *
NMGILA	0.824 *	0.852 *	0.891 *	0.639 *	0.170 *	0.367 *
TRT	0.936 *	0.944 *	0.942 *	0.922 *	0.925 *	0.813 *
BOL	0.877 *	0.892 *	0.938 *	0.807 *	0.908 *	0.700 *

TABLE 13. (continued), *ND2 F_{ST} values*

	EFE	*LEG*	*UEG*	*HCN*	*TURNM*	*NMGILA*	*TRT*
EVR							
LSALT							
SPRVE							
VDP							
WAK							
WVW							
WCL							
FOS							
MAR							
SPRSA							
TON							
BLK							
CHR							
TURAZ							
BAS							
ODN							
RDF							
ARA							
BLU							
CCR							
SHY							
SAB							
SIL							
BON							
EFE							
LEG	0.171 *						
UEG	0.000 NS	0.187 *					
HCN	0.400 *	0.416 *	0.367 *				
TURNM	0.250 *	0.157 *	0.290 *	0.748 *			
NMGILA	0.121 *	0.121 *	0.179 *	0.694 *	0.060 NS		
TRT	0.839 *	0.784 *	0.832 *	0.934 *	0.848 *	0.858 *	
BOL	0.721 *	0.671 *	0.703 *	0.833 *	0.806 *	0.793 *	0.919 *

TABLE 14. Pairwise F_{ST} values for 32 populations of *Gila* in the lower Colorado River from Tpi-B sequences. Significant values ($P < 0.05$) are marked by asterisks. (see Appendix 1. and Fig. 2. for localities)

	EVR	*LSALT*	*SPRVE*	*VDP*	*WAK*	*WVW*
EVR						
LSALT	0.368 *					
SPRVE	0.143 *	0.114 *				
VDP	0.274 *	0.113 *	0.127 *			
WAK	0.088 NS	0.249 *	0.030 NS	0.206 *		
WVW	0.501 *	0.343 *	0.331 *	0.076 *	0.436 *	
WCL	0.251 *	0.027 NS	0.056 *	0.034 NS	0.164 *	0.184 *
FOS	0.418 *	0.325 *	0.280 *	0.075 *	0.369 *	-0.014 NS
MAR	0.680 *	0.492 *	0.534 *	0.199 *	0.628 *	0.070 *
SPRSA	0.410 *	0.305 *	0.251 *	0.063 *	0.350 *	-0.005 NS
TON	0.851 *	0.593 *	0.678 *	0.290 *	0.793 *	0.217 *
BLK	0.481 *	0.237 *	0.309 *	0.051 *	0.412 *	0.058 *
CHR	0.401 *	-0.021 NS	0.118 *	0.098 *	0.267 *	0.331 *
TURAZ	0.502 *	0.275 *	0.310 *	0.043 NS	0.424 *	-0.001 NS
BAS	0.410 *	0.305 *	0.251 *	0.063 *	0.350 *	-0.005 NS
ODN	0.143 *	0.114 *	-0.026 NS	0.127 *	0.030 NS	0.331 *
RDF	0.531 *	0.307 *	0.350 *	0.064 *	0.459 *	0.010 NS
ARA	0.311 *	0.000 NS	0.048 *	0.120 *	0.177 *	0.355 *
BLU	0.000 NS	0.363 *	0.139 *	0.268 *	0.085 NS	0.494 *
CCR	1.000 *	0.265 *	0.653 *	0.434 *	0.910 *	0.735 *
SHY	0.616 *	0.023 NS	0.134 *	0.180 *	0.363 *	0.474 *
SAB	0.606 *	0.028 NS	0.162 *	0.200 *	0.381 *	0.495 *
SIL	0.326 *	0.290 *	0.209 *	0.070 *	0.287 *	0.021 NS
BON	0.316 *	0.015 NS	0.030 NS	0.126 *	0.160 *	0.369 *
EFE	0.020 NS	0.298 *	0.064 *	0.207 *	0.019 NS	0.398 *
LEG	0.321 *	0.018 NS	0.081 *	0.025 NS	0.211 *	0.183 *
UEG	0.070 *	0.218 *	0.004 NS	0.182 *	-0.024 NS	0.396 *
HCN	0.328 *	0.226 *	0.174 *	0.032 NS	0.269 *	0.015 NS
TURNM	0.000 NS	0.358 *	0.135 *	0.262 *	0.081 NS	0.487 *
NMGILA	0.345 *	-0.006 NS	0.080 *	0.045 *	0.221 *	0.236 *
TRT	0.008 NS	0.370 *	0.120 *	0.300 *	0.014 NS	0.532 *
BOL	0.582 *	0.050 *	0.236 *	0.273 *	0.421 *	0.561 *

TABLE 14. (continued), *Tpi-B F_{ST} values*

	WCL	*FOS*	*MAR*	*SPRSA*	*TON*	*BLK*
EVR						
LSALT						
SPRVE						
VDP						
WAK						
WVW						
WCL						
FOS	0.164 *					
MAR	0.355 *	0.129 *				
SPRSA	0.142 *	-0.020 NS	0.163 *			
TON	0.471 *	0.287 *	0.028 NS	0.331 *		
BLK	0.131 *	0.090 *	0.090 *	0.098 *	0.168 *	
CHR	0.021 NS	0.315 *	0.483 *	0.295 *	0.590 *	0.219 *
TURAZ	0.138 *	0.024 NS	0.056 *	0.034 NS	0.168 *	-0.007 NS
BAS	0.142 *	-0.020 NS	0.163 *	-0.026 NS	0.331 *	0.098 *
ODN	0.056 *	0.280 *	0.534 *	0.251 *	0.678 *	0.309 *
RDF	0.173 *	0.044 NS	0.034 NS	0.059 NS	0.129 *	-0.006 NS
ARA	0.025 NS	0.325 *	0.523 *	0.302 *	0.639 *	0.275 *
BLU	0.247 *	0.413 *	0.675 *	0.404 *	0.847 *	0.475 *
CCR	0.401 *	0.711 *	0.794 *	0.729 *	0.897 *	0.566 *
SHY	0.098 *	0.445 *	0.626 *	0.431 *	0.756 *	0.360 *
SAB	0.114 *	0.466 *	0.639 *	0.454 *	0.762 *	0.378 *
SIL	0.130 *	-0.005 NS	0.211 *	-0.017 NS	0.375 *	0.128 *
BON	0.033 NS	0.335 *	0.543 *	0.311 *	0.667 *	0.293 *
EFE	0.179 *	0.323 *	0.604 *	0.299 *	0.773 *	0.406 *
LEG	-0.020 NS	0.169 *	0.349 *	0.149 *	0.469 *	0.109 *
UEG	0.132 *	0.331 *	0.597 *	0.307 *	0.757 *	0.381 *
HCN	0.079 *	-0.003 NS	0.188 *	-0.014 NS	0.342 *	0.077 *
TURNM	0.242 *	0.407 *	0.670 *	0.397 *	0.844 *	0.469 *
NMGILA	-0.015 NS	0.221 *	0.402 *	0.200 *	0.520 *	0.147 *
TRT	0.257 *	0.448 *	0.701 *	0.441 *	0.856 *	0.508 *
BOL	0.166 *	0.530 *	0.680 *	0.523 *	0.782 *	0.441 *

TABLE 14. (continued), *Tpi-B F_{ST} values*

	CHR	*TURAZ*	*BAS*	*ODN*	*RDF*	*ARA*
EVR						
LSALT						
SPRVE						
VDP						
WAK						
WVW						
WCL						
FOS						
MAR						
SPRSA						
TON						
BLK						
CHR						
TURAZ	0.258 *					
BAS	0.295 *	0.034 NS				
ODN	0.118 *	0.310 *	0.251 *			
RDF	0.290 *	-0.024 NS	0.059 NS	0.350 *		
ARA	0.000 NS	0.301 *	0.302 *	0.048 *	0.337 *	
BLU	0.395 *	0.495 *	0.404 *	0.139 *	0.524 *	0.306 *
CCR	0.292 *	0.664 *	0.729 *	0.653 *	0.672 *	0.382 *
SHY	0.028 NS	0.406 *	0.431 *	0.134 *	0.437 *	0.003 NS
SAB	0.035 NS	0.428 *	0.454 *	0.162 *	0.457 *	0.016 NS
SIL	0.283 *	0.064 *	-0.017 NS	0.209 *	0.094 *	0.276 *
BON	0.016 NS	0.318 *	0.311 *	0.030 NS	0.354 *	-0.020 NS
EFE	0.316 *	0.405 *	0.299 *	0.064 *	0.441 *	0.233 *
LEG	0.010 NS	0.124 *	0.149 *	0.081 *	0.157 *	0.028 NS
UEG	0.232 *	0.386 *	0.307 *	0.004 NS	0.423 *	0.144 *
HCN	0.214 *	0.029 NS	-0.014 NS	0.174 *	0.057 NS	0.219 *
TURNM	0.388 *	0.487 *	0.397 *	0.135 *	0.517 *	0.300 *
NMGILA	-0.012 NS	0.171 *	0.200 *	0.080 *	0.204 *	0.006 NS
TRT	0.406 *	0.533 *	0.441 *	0.120 *	0.561 *	0.304 *
BOL	0.059 *	0.500 *	0.523 *	0.236 *	0.525 *	0.055 *

TABLE 14. (continued), *Tpi-B F_{ST} values*

	BLU	*CCR*	*SHY*	*SAB*	*SIL*	*BON*
EVR						
LSALT						
SPRVE						
VDP						
WAK						
WVW						
WCL						
FOS						
MAR						
SPRSA						
TON						
BLK						
CHR						
TURAZ						
BAS						
ODN						
RDF						
ARA						
BLU						
CCR	1.000 *					
SHY	0.607 *	0.524 *				
SAB	0.599 *	0.461 *	-0.041 NS			
SIL	0.321 *	0.692 *	0.400 *	0.422 *		
BON	0.310 *	0.458 *	0.002 NS	0.018 NS	0.281 *	
EFE	0.019 NS	0.886 *	0.436 *	0.454 *	0.233 *	0.223 *
LEG	0.315 *	0.420 *	0.104 *	0.120 *	0.142 *	0.041 NS
UEG	0.067 *	0.836 *	0.295 *	0.321 *	0.249 *	0.125 *
HCN	0.322 *	0.650 *	0.336 *	0.360 *	-0.012 NS	0.226 *
TURNM	0.000 NS	1.000 *	0.599 *	0.591 *	0.316 *	0.304 *
NMGILA	0.339 *	0.385 *	0.067 NS	0.080 *	0.190 *	0.019 NS
TRT	0.007 NS	0.959 *	0.574 *	0.575 *	0.353 *	0.302 *
BOL	0.576 *	0.311 *	-0.015 NS	-0.019 NS	0.488 *	0.066 *

TABLE 14. (continued), *Tpi-B F_{ST} values*

	EFE	*LEG*	*UEG*	*HCN*	*TURNM*	*NMGILA*	*TRT*
EVR							
LSALT							
SPRVE							
VDP							
WAK							
WVW							
WCL							
FOS							
MAR							
SPRSA							
TON							
BLK							
CHR							
TURAZ							
BAS							
ODN							
RDF							
ARA							
BLU							
CCR							
SHY							
SAB							
SIL							
BON							
EFE							
LEG	0.231 *						
UEG	0.006 NS	0.174 *					
HCN	0.230 *	0.083 *	0.230 *				
TURNM	0.017 NS	0.308 *	0.064 *	0.316 *			
NMGILA	0.253 *	-0.023 NS	0.183 *	0.126 *	0.332 *		
TRT	0.015 NS	0.333 *	0.028 NS	0.355 *	0.006 NS	0.354 *	
BOL	0.490 *	0.174 *	0.380 *	0.435 *	0.571 *	0.126 *	0.564 *

TABLE 15. Pairwise F_{ST} values for 32 populations of *Gila* in the lower Colorado River from S7 sequences. Significant values ($P < 0.05$) are marked by asterisks. (see Appendix 1. and Fig. 2. for localities)

	EVR	*LSALT*	*SPRVE*	*VDP*	*WAK*	*WVW*
EVR						
LSALT	0.368 *					
SPRVE	0.143 *	0.114 *				
VDP	0.274 *	0.113 *	0.127 *			
WAK	0.088 NS	0.249 *	0.030 NS	0.206 *		
WVW	0.501 *	0.343 *	0.331 *	0.076 *	0.436 *	
WCL	0.251 *	0.027 NS	0.056 *	0.034 NS	0.164 *	0.184 *
FOS	0.418 *	0.325 *	0.280 *	0.075 *	0.369 *	-0.014 NS
MAR	0.680 *	0.492 *	0.534 *	0.199 *	0.628 *	0.070 *
SPRSA	0.410 *	0.305 *	0.251 *	0.063 *	0.350 *	-0.005 NS
TON	0.851 *	0.593 *	0.678 *	0.290 *	0.793 *	0.217 *
BLK	0.481 *	0.237 *	0.309 *	0.051 *	0.412 *	0.058 *
CHR	0.401 *	-0.021 NS	0.118 *	0.098 *	0.267 *	0.331 *
TURAZ	0.502 *	0.275 *	0.310 *	0.043 NS	0.424 *	-0.001 NS
BAS	0.410 *	0.305 *	0.251 *	0.063 *	0.350 *	-0.005 NS
ODN	0.143 *	0.114 *	-0.026 NS	0.127 *	0.030 NS	0.331 *
RDF	0.531 *	0.307 *	0.350 *	0.064 *	0.459 *	0.010 NS
ARA	0.311 *	0.000 NS	0.048 *	0.120 *	0.177 *	0.355 *
BLU	0.000 NS	0.363 *	0.139 *	0.268 *	0.085 NS	0.494 *
CCR	1.000 *	0.265 *	0.653 *	0.434 *	0.910 *	0.735 *
SHY	0.616 *	0.023 NS	0.134 *	0.180 *	0.363 *	0.474 *
SAB	0.606 *	0.028 NS	0.162 *	0.200 *	0.381 *	0.495 *
SIL	0.326 *	0.290 *	0.209 *	0.070 *	0.287 *	0.021 NS
BON	0.316 *	0.015 NS	0.030 NS	0.126 *	0.160 *	0.369 *
EFE	0.020 NS	0.298 *	0.064 *	0.207 *	0.019 NS	0.398 *
LEG	0.321 *	0.018 NS	0.081 *	0.025 NS	0.211 *	0.183 *
UEG	0.070 *	0.218 *	0.004 NS	0.182 *	-0.024 NS	0.396 *
HCN	0.328 *	0.226 *	0.174 *	0.032 NS	0.269 *	0.015 NS
TURNM	0.000 NS	0.358 *	0.135 *	0.262 *	0.081 NS	0.487 *
NMGILA	0.345 *	-0.006 NS	0.080 *	0.045 *	0.221 *	0.236 *
TRT	0.008 NS	0.370 *	0.120 *	0.300 *	0.014 NS	0.532 *
BOL	0.582 *	0.050 *	0.236 *	0.273 *	0.421 *	0.561 *

TABLE 15. (continued), *S7 F_{ST} values*

	WCL	*FOS*	*MAR*	*SPRSA*	*TON*	*BLK*
EVR						
LSALT						
SPRVE						
VDP						
WAK						
WVW						
WCL						
FOS	0.399 *					
MAR	0.404 *	0.000 NS				
SPRSA	-0.022 NS	0.447 *	0.453 *			
TON	0.341 *	0.000 NS	0.000 NS	0.382 *		
BLK	0.020 NS	0.585 *	0.590 *	0.027 NS	0.526 *	
CHR	-0.021 NS	0.444 *	0.450 *	-0.025 NS	0.381 *	0.025 NS
TURAZ	0.016 NS	0.622 *	0.628 *	0.023 NS	0.556 *	-0.025 NS
BAS	-0.014 NS	0.395 *	0.401 *	-0.020 NS	0.332 *	0.062 NS
ODN	-0.008 NS	0.550 *	0.555 *	-0.005 NS	0.484 *	-0.016 NS
RDF	0.110 *	0.738 *	0.743 *	0.131 *	0.684 *	0.011 NS
ARA	0.082 *	0.164 *	0.167 *	0.077 *	0.127 *	0.220 *
BLU	0.335 *	0.000 NS	0.000 NS	0.375 *	0.000 NS	0.520 *
CCR	0.703 *	0.864 *	0.867 *	0.715 *	0.831 *	0.705 *
SHY	0.306 *	0.000 NS	0.000 NS	0.342 *	0.000 NS	0.489 *
SAB	0.694 *	0.837 *	0.840 *	0.708 *	0.792 *	0.734 *
SIL	0.397 *	-0.002 NS	-0.001 NS	0.440 *	-0.011 NS	0.580 *
BON	0.095 *	0.755 *	0.760 *	0.116 *	0.697 *	0.002 NS
EFE	0.086 *	0.300 *	0.305 *	0.083 *	0.244 *	0.129 *
LEG	0.067 *	0.181 *	0.185 *	0.062 *	0.138 *	0.173 *
UEG	0.166 *	0.281 *	0.286 *	0.162 *	0.222 *	0.237 *
HCN	0.049 NS	0.251 *	0.255 *	0.044 NS	0.202 *	0.187 *
TURNM	-0.016 NS	0.539 *	0.545 *	-0.016 NS	0.470 *	-0.011 NS
NMGILA	0.089 *	0.185 *	0.189 *	0.087 *	0.140 *	0.241 *
TRT	0.187 *	0.477 *	0.482 *	0.202 *	0.424 *	0.175 *
BOL	0.419 *	0.000 NS	0.000 NS	0.469 *	0.000 NS	0.604 *

TABLE 15. (continued), *S7 F_{ST} values*

	CHR	*TURAZ*	*BAS*	*ODN*	*RDF*	*ARA*
EVR						
LSALT						
SPRVE						
VDP						
WAK						
WVW						
WCL						
FOS						
MAR						
SPRSA						
TON						
BLK						
CHR						
TURAZ	0.022 ^{NS}					
BAS	-0.019 ^{NS}	0.059 ^{NS}				
ODN	-0.006 ^{NS}	-0.020 ^{NS}	0.020 ^{NS}			
RDF	0.128 *	0.008 ^{NS}	0.182 *	0.046 ^{NS}		
ARA	0.081 *	0.219 *	0.043 ^{NS}	0.162 *	0.358 *	
BLU	0.374 *	0.548 *	0.325 *	0.477 *	0.678 *	0.123 *
CCR	0.715 *	0.698 *	0.724 *	0.703 *	0.704 *	0.748 *
SHY	0.342 *	0.513 *	0.293 *	0.443 *	0.648 *	0.104 *
SAB	0.709 *	0.728 *	0.706 *	0.719 *	0.764 *	0.701 *
SIL	0.438 *	0.614 *	0.387 *	0.543 *	0.728 *	0.161 *
BON	0.113 *	-0.002 ^{NS}	0.167 *	0.034 ^{NS}	-0.030 ^{NS}	0.343 *
EFE	0.085 *	0.118 *	0.082 *	0.102 *	0.186 *	0.088 *
LEG	0.065 *	0.165 *	0.040 ^{NS}	0.124 *	0.282 *	-0.006 ^{NS}
UEG	0.166 *	0.222 *	0.150 *	0.199 *	0.302 *	0.114 *
HCN	0.047 ^{NS}	0.187 *	0.012 ^{NS}	0.128 *	0.335 *	-0.008 ^{NS}
TURNM	-0.016 ^{NS}	-0.014 ^{NS}	0.005 ^{NS}	-0.026 ^{NS}	0.064 ^{NS}	0.139 *
NMGILA	0.090 *	0.242 *	0.048 ^{NS}	0.180 *	0.390 *	-0.021 ^{NS}
TRT	0.203 *	0.163 *	0.220 *	0.175 *	0.165 *	0.283 *
BOL	0.466 *	0.644 *	0.417 *	0.571 *	0.755 *	0.177 *

TABLE 15. (continued), *S7 F_{ST} values*

	BLU	*CCR*	*SHY*	*SAB*	*SIL*	*BON*
EVR						
LSALT						
SPRVE						
VDP						
WAK						
WVW						
WCL						
FOS						
MAR						
SPRSA						
TON						
BLK						
CHR						
TURAZ						
BAS						
ODN						
RDF						
ARA						
BLU						
CCR	0.827 *					
SHY	0.000 NS	0.808 *				
SAB	0.786 *	0.820 *	0.758 *			
SIL	-0.012 NS	0.864 *	-0.020 NS	0.833 *		
BON	0.690 *	0.695 *	0.656 *	0.754 *	0.742 *	
EFE	0.237 *	0.535 *	0.210 *	0.593 *	0.306 *	0.168 *
LEG	0.133 *	0.669 *	0.111 *	0.640 *	0.183 *	0.262 *
UEG	0.215 *	0.522 *	0.186 *	0.584 *	0.287 *	0.280 *
HCN	0.197 *	0.757 *	0.172 *	0.719 *	0.243 *	0.321 *
TURNM	0.462 *	0.702 *	0.426 *	0.713 *	0.531 *	0.050 NS
NMGILA	0.135 *	0.756 *	0.112 *	0.710 *	0.176 *	0.377 *
TRT	0.417 *	0.554 *	0.391 *	0.623 *	0.485 *	0.154 *
BOL	0.000 NS	0.874 *	0.000 NS	0.850 *	0.001 NS	0.773 *

TABLE 15. (continued), *S7 F_{ST} values*

	EFE	*LEG*	*UEG*	*HCN*	*TURNM*	*NMGILA*	*TRT*
EVR							
LSALT							
SPRVE							
VDP							
WAK							
WVW							
WCL							
FOS							
MAR							
SPRSA							
TON							
BLK							
CHR							
TURAZ							
BAS							
ODN							
RDF							
ARA							
BLU							
CCR							
SHY							
SAB							
SIL							
BON							
EFE							
LEG	0.020 ^{NS}						
UEG	-0.002 ^{NS}	0.034 ^{NS}					
HCN	0.107 *	0.015 ^{NS}	0.147 *				
TURNM	0.092 *	0.104 *	0.183 *	0.104 *			
NMGILA	0.103 *	0.003 ^{NS}	0.125 *	-0.015 ^{NS}	0.155 *		
TRT	0.134 *	0.220 *	0.206 *	0.285 *	0.176 *	0.291 *	
BOL	0.320 *	0.197 *	0.301 *	0.268 *	0.562 *	0.201 *	0.496 *

TABLE 16. Management units for *Gila* populations in the lower Colorado River basin. Units are designated by morphological species (R, N and I) and organized in numbered groups by genetic similarity at ND2, Tpi-B and S7 loci.

robusta	*nigra*	*intermedia*
R1) Cherry Cr.	N1) Tonto Cr.	I1) Williamson Val. Wsh
Black R.	Marsh Cr.	
Low. Salt R.		I2) Walker Cr.
Verde R. (Perkinville)	N2) Spring Cr.(Salt)	
		I3) Spring Cr. (Verde)
R2) Aravaipa Cr.	N3) West Clear Cr.	
Low. Eagle Cr.		I4) Blue R.
	N4) East Verde R.	
R3) Boulder Cr.		I5) Cienega Cr.
	N5) Fossil Cr.	
R4) Trout Cr.		I6) Sabino Cr.
	N6) East, Middle and West Forks	
	Gila R., NM	I7) Sheehy Spr.
		I8) Turkey Cr., AZ
		Bass Can.
		Redfield Can.
		Odonnell Can.
		I9) East Fork Eagle Cr.
		upper Eagle Cr.
		Bonita Cr.
		Harden Cienega Cr.
		Turkey Cr., NM
		I10) Silver Cr.

APPENDIX 1

LOCALITIES, LOCALITY ABBREVIATIONS, TAXONOMIC IDENTITY AND NUMBER OF INDIVIDUALS SAMPLED PER LOCALITY

Localities are arranged by drainage sub-basin of the Gila and Bill Williams Rivers. All sub-basins are within the Gila River basin except for the Bill Williams River, which is a distinct tributary of the lower Colorado River. Locality numbers correspond to the localities map, Figure 1. Taxonomic identity of sampled individuals follows Minckley and DeMarais (2000).

Verde River sub-basin
1. East Verde River (EVR), Gila Co., AZ [*G. nigra*; 20 individuals sampled]
2. Verde River at Salt River confluence (LSALT), Maricopa Co., AZ [*G. robusta*; 29]
3. Spring Creek (SPRVE), Yavapai Co, AZ [*G. intermedia*; 20]
4. Verde River at Perkinsville (VDP), Yavapai Co, AZ [*G. robusta*; 20]
5. Walker Creek (WAK), Yavapai Co, AZ [*G. intermedia*; 15]
6. West Clear Creek (WCL), Yavapai Co, AZ [*G. nigra*; 29]
7. Williamson Valley Wash (WVW), Yavapai Co, AZ [*G. intermedia*; 20]
8. Fossil Creek (FOS), Yavapai Co, AZ [*G. nigra*; 26]

Salt River sub-basin
9. Marsh Creek (MAR), Gila Co., AZ [*G. nigra*; 27]
10. Spring Creek (SPRSA), Gila Co., AZ [*G. nigra*; 20]
11. Tonto Creek (TON), Gila Co., AZ [*G. nigra*; 16]
12. Cherry Creek, CHR, Gila Co., AZ [*G. robusta*; 21]
13. Black River (BLK), Apache Co., AZ [*G. robusta*; 24]

San Pedro River sub-basin
14. Aravaipa Creek (ARA), Pinal Co., AZ [*G. robusta*; 24]
15. Bass Canyon (BAS), Cochise Co., AZ [*G. intermedia*; 20]
16. O'Donnell Canyon (ODN), Santa Cruz Co., AZ [*G. intermedia*; 20]
17. Redfield Canyon (RDF), Pima Co., AZ [*G. intermedia*; 20]
19. Turkey Creek (TURAZ), Santa Cruz Co., AZ [*G. intermedia*; 18]

San Carlos River sub-basin
18. Blue River (BLU), Gila, Co., AZ [*G. intermedia*; 19]

Santa Cruz River sub-basin
20. Sheehy Spring (SHY), Santa Cruz Co., AZ [*G. intermedia*; 11]
21. Cienega Creek (CCR), Pima Co., AZ [*G. intermedia*; 20]
22. Sabino Creek (SAB), Pima Co., AZ [*G. intermedia*; 14]

Gila River sub-basin, middle section
23. East Fork Eagle (EFE), Greenlee Co., AZ [*G. intermedia*; 20]
24. Eagle Creek - upper (UEG), Greenlee Co., AZ [*G. intermedia*; 18]
25. Eagle Creek- lower (LEG), Greenlee Co., AZ [*G. robusta*; 20]
26. Bonita Creek (BON), Graham Co., AZ [*G. intermedia*; 20]

Gila River sub-basin, upper section
27. Harden-Cienega Creek (HCN), Greenlee, Co., AZ [*G. intermedia*; 22]
28. Turkey Creek (TURNM), Grant Co., NM [*G. intermedia*; 18]
29. East, Middle and West Forks Gila River (NMGILA), Catron Co., NM [*G. nigra*; 18]

Agua Fria River basin
30. Silver Creek (SIL), Yavapai Co., AZ [*G. intermedia*; 29]

Bill Williams River basin
31. Trout Creek (TRT), Mohave Co., AZ [*G. robusta*; 30]
32. Boulder Creek (BOL), Yavapai Co., AZ [*G. robusta*; 30]

APPENDIX 2

LIST OF HAPLOTYPE AND ALLELE SEQUENCES FOR ND2, TPI-B, AND S7

ND2 (788bp)

AA
CACCACCACCCCCGCGCAGTGGAAGCGGCCACAAAATACTTTCTAACCCAGG
CCACCGCAGCAGCTATAATTCTGTTTGCAAGCATAACAAATGCCTGAATCAC
AGGGGGGTGAGACATAAGTAATATATCGGATCCCATTGCCAGCACAATAGTC
ATCGCCGCCCTGGCACTCAAAATTGGACTAGCACCTATACACTTCTGAATACC
TGAAGTTTTACAGGGGCTAGATCTTTTAACCGGACTAATTTTATCCACCTGAC
AAAAACTTGCTCCACTTGCTCTCATTATTCAAACAGCCCAAGCTATCGACCCC
CTACTGCTGACGGCCCTGGGCCTATTGTCCACACTAATCGGTGGTTGAGGAG
GGCTAAACCAGACCCAGCTCCGGAAAATTTTAGCCCTTGCAATTACCTGATC
GAAAAGCCCTGTCTTGACAGCAACCACTGCCTTAGTATTACTGTCATTGGGCG
GTTTGCCTCCTCTTACAGGGTTTATGCCAAAATGACTTATTTTACAAGAGTTG
GCAAAACAGAGCCTTCCCCTTGCCGCTACAGTTATGGCTCTCGCAGCCCTCCT
CAGCCTGTACTTCTACTTACGACTCTGCTATGCAATGACACTCACCATTTCTC
CTAACACTGTTAATTCAGCCACGCCCTGACGGGTTCAAACAACTCAAGCCTCT
CTCCCCGTGGCCCTAGCTACTACGCTAGCACTGGGCCTTCTTCCCATAACTCC
AGCCATTGTAATGCTAGTCACCTAGGGGCTTAGGATAACATTAGACCAAGA

AB
CACCACCACCCCCGCGCAGTGGAAGCGGCCACAAAATACTTTCTAACCCAGG
CCACCGCAGCAGCTATAATTCTGTTTGCAAGCATAACAAATGCCTGAATCAC
AGGGGGGTGAGACATAAGTAATATATCGGATCCCATTGCCAGCACAATAGTC
ATCGCCGCCCTGGCACTCAAAATTGGACTAGCACCTATACACTTCTGAATACC
TGAAGTTTTACAGGGGCTAGATCTTTTAACCGGACTAATTTTATCCACCTGAC
AAAAACTTGCTCCACTTGCTCTCATTATTCAAACAGCCCAAGCTATCGACCCC
CTACTGCTGACGGCCCTGGGCCTATTGTCCACACTAATCGGTGGTTGAGGAG
GGCTAAACCAGACCCAGCTCCGGAAAATTTTAGCCCTTGCAATTACCTGATC
GAAAAGCCCTGTCTTGACAGCAACCACTGCCTTAGTATTACTGTCATTGGGCG
GTTTGCCTCCTCTTACAGGGTTTATGCCAAAATGACTTATTTTACAAGAGTTG
GCAAAACAGAGCCTTCCCCTTGCCGCTACAGTTATGGCTCTCGCAGCCCTCCT
CAGCCTGTACTTCTACTTACGACTCTGCTATGCAATGACACTCACCATTTCTC
CTAACACTGTTAATTCAGCCACACCCTGACGGGTTCAAACAACTCAAGCCTCT
CTCCCCGTGGCCCTAGCTACTACGCTAGCACTGGGCCTTCTTCCCATAACTCC
AGCCATTGTAATGCTAGTCACCTAGGGGCTTAGGATAACATTAGACCAAGA

AD
CACCACCACCCCCGCGCAGTGGAAGCGGCCACAAAATACTTTCTAACCCAGG
CCACCGCAGCAGCTATAATTCTGTTTGCAAGCATAACAAATGCCTGAATCAC
AGGGGGGTGAGACATAAGTAATATATCGGATCCCATTGCCAGCACAATAGTC
ATCGCCGCCCTGGCACTCAAAATTGGACTAGCACCTATACACTTCTGAATACC
TGAAGTTTTACAGGGGCTAGATCTTTTAACCGGACTAATTTTATCCACCTGAC
AAAAACTTGCTCCACTTGCTCTCATTATTCAAACAGCCCAAGCTATCGACCCC
CTACTGCTGACGGCCCTGGGCCTATTGTCCACACTAATCGGTGGTTGAGGAG

GGCTAAACCAGACCCAGCTCCGGAAAATTTTAGCCCTTGCAATTACCTGATC
GAAAAGCCCTGTCTTGACAGCAACCACTGCCTTAGTATTACTGTCATTGGGCG
GTTTGCCTCCTCTTACAGGGTTTATGCCAAAATGACTTATTTTACAAGAGTTG
GCAAAGCAGAGCCTTCCCCTTGCCGCTACAGTTATGGCTCTCGCAGCCCTCCT
CAGCCTGTACTTCTACTTACGACTCTGCTATGCAATGACACTCACCATTTCTC
CTAACACTGTTAATTCAGCCACACCCTGACGGGTTCAAACAACTCAAGCCTCT
CTCCCCGTGGCCCTAGCTACTACGCTAGCACTGGGCCTTCTTCCCATAACTCC
AGCCATTGTAATGCTAGTCACCTAGGGGCTTAGGATAACATTAGACCAAGA

AE
CACCACCACCCCCGCGCAGTGGAAGCGGCCACAAAATACTTTCTAACCCAGG
CCACCGCAGCAGCTATAATTCTGTTTGCAAGCATAACAAATGCCTGAATCAC
AGGGGGGTGAGACATAAGTAATATATCGGATCCCATTGCCAGCACAATAGTC
ATCGCCGCCCTGGCACTCAAAATTGGACTAGCACCTATACACTTCTGAATACC
TGAAGTTTTACAGGGGCTAGATCTTTTAACCGGACTAATTTTATCCACCTGAC
AAAAACTTGCTCCACTTGCTCTCATTATTCAAACAGCCCAAGCTATCGACCCC
CTACTGCTGACGGCCCTGGGCCTATTGTCCACACTAATCGGTGGTTGAGGAG
GGCTAAACCAGACCCAGCTCCGGAAAATTTTAGCCCTTGCAATTACCTGATC
GAAAAGCCCTGTCTTGACAGCAACCACTGCCTTAGTATTACTCTCATTGGGCG
GCTTGCCTCCTCTTACAGGGTTTATGCCAAAATGACTTATTTTACAAGAGTTG
GCAAAACAGAGCCTTCCCCTTGCCGCTACAGTTATGGCTCTCGCAGCCCTCCT
CAGCCTGTACTTCTACTTACGACTCTGCTATGCAATGACACTCACCATTTCTC
CTAACACTGTTAATTCAGCCACACCCTGACGGGTTCAAACAACTCAAGCCTCT
CTCCCCGTGGCCCTAGCTACTACGCTAGCACTGGGCCTTCTTCCCATAACTCC
AGCCATTGTAATGCTAGTCACCTAGGGGCTTAGGATAACATTAGACCAAGA

AF
CACCACCACCCCCGCGCAGTGGAAGCGGCCACAAAATACTTTCTAACCCAGG
CCACCGCAGCAGCTATAATTCTGTTTGCAAGCATAACAAATGCCTGAATCAC
AGGGGGGTGAGACATAAGTAATATATCGGATCCCATTGCCAGCACAATAGTC
ATCGCCGCCCTGGCACTCAAAATTGGACTAGCACCTATACACTTCTGAATACC
TGAAGTTTTACAGGGGCTAGATCTTTTAACCGGACTAATTTTATCCACCTGAC
AAAAACTTGCTCCACTTGCTCTCATTATTCAAACAGCCCAAGCTATCGACCCC
CTACTGCTGACGGCCCTGGGCCTATTGTCCACACTAATCGGTGGTTGAGGAG
GGCTAAACCAGACCCAGCTCCGGAAAATTTTAGCCCTTGCAATTACCTGATC
GAAAAGCCCTGTCTTAACAGCAACCACTGCCTTAGTATTACTGTCATTGGGCG
GTTTGCCTCCTCTTACAGGGTTTATGCCAAAATGACTTATTTTACAAGAGTTG
GCAAAACAGAGCCTTCCCCTTGCCGCTACAGTTATGGCTCTCGCAGCCCTCCT
CAGCCTGTACTTCTACTTACGACTCTGCTATGCAATGACACTCACCATTTCTC
CTAACACTGTTAATTCAGCCACACCCTGACGGGTTCAAACAACTCAAGCCTCT
CTCCCCGTGGCCCTAGCTACTACGCTAGCACTGGGCCTTCTTCCCATAACTCC
AGCCATTGTAATGCTAGTCACCTAGGGGCTTAGGATAACATTAGACCAAGA

AG
CACCACCACCCCCGCGCAGTGGAAGCGGCCACAAAATACTTTCTAACCCAGG
CCACCGCAGCAGCTATAATTCTGTTTGCAAGCATAACAAATGCCTGAATCAC
AGGGGGGTGAGACATAAGTAATATATCGGATCCCATTGCCAGCACAATAGTC
ATCGCCGCCCTGGCACTCAAAATTGGACTAGCACCTATACACTTCTGAATACC
TGAAGTTTTACAGGGGCTAGATCTTTTAACCGGACTAATTTTATCCACCTGAC
AAAAACTTGCTCCACTTGCTCTCATTATTCAAACAGCCCAAGCTATCGACCCC
CTACTGCTGACGGCCCTGGGCCTATTGTCCACACTAATCGGTGGTTGAGGAG
GGCTAAACCAGACCCAGCTCCGGAAAATTTTAGCCCTTGCAATTACCTGATC
GAAAAGCCCTGTCTTGACAGCAACCACTGCCTTAGTATTACTGTCATTGGGCG
GTTTGCCTCCTCTTACAGGGTTTATGCCAAAATGACTTATTTTACAAGAGTTG
ACAAAACAGAGCCTTCCCCTTGCCGCTACAGTTATGGCTCTCGCAGCCCTCCT
CAGCCTGTACTTCTACTTACGACTCTGCTATGCAATGACACTCACCATTTCTC
CTAACACTGTTAATTCAGCCACACCCTGGCGGGTTCAAACAACTCAAGCCTCT
CTCCCCGTGGCCCTAGCTACTACGCTAGCACTGGGCCTTCTTCCCATAACTCC
AGCCATTGTAATGCTAGTCACCTAGGGGCTTAGGATAACATTAGACCAAGA

AH
CACCACCACCCCCGCGCAGTGGAAGCGGCCACAAAATACTTTCTAACCCAGG
CCACCGCAGCAGCTATAATTCTGTTTGCAAGCATAACAAATGCCTGAATCAC
AGGGGGGTGAGACATAAGTAATATATCGGATCCCATTGCCAGCACAATAGTC
ATCGCCGCCCTGGCACTCAAAATTGGACTAGCACCTATACACTTCTGAATACC
TGAAGTTTTACAGGGGCTAGATCTTTTAACCGGACTAATTTTATCCACCTGAC
AAAAACTTGCTCCACTTGCTCTCATTATTCAAACAGCCCAAGCTATCGACCCC
CTACTGCTGACGGCCCTGGGCCTATTGTCCACACTAATCGGTGGTTGAGGAG
GGCTAAACCAGACCCAGCTCCGGAAAATTTTAGCCCTTGCAATTACCTGATC
GAAAAGCCCTGTCTTGACAGCAACCACTGCCTTAGTATTACTGTCATTGGGCG
GTTTGCCTCCTCTTACAGGGTTTATGCCAAAATGACTTATTTTACAAGAGTTG
GCAAAACAGGGCCTTCCCCTTGCCGCTACAGTTATGGCTCTCGCAGCCCTCCT
CAGCCTGTACTTCTACTTACGACTCTGCTATGCAATGACACTCACCATTTCTC
CTAACACTGTTAATTCAGCCACGCCCTGACGGGTTCAAACAACTCAAGCCTCT
CTCCCCGTGGCCCTAGCTACTACGCTAGCACTGGGCCTTCTTCCCATAACTCC
AGCCATTGTAATGCTAGTCACCTAGGGGCTTAGGATAACATTAGACCAAGA

AJ
CACCACCACCCCCGCGCAGTGGAAGCGGCCACAAAATACTTTCTAACCCAGG
CCACCGCAGCAGCTATAATTCTGTTTGCAAGCATAACAAATGCCTGAATCAC
AGGGGGGTGAGACATAAGTAATATATCGGATCCCATTGCCAGCACAATAGTC
ATCGCCGCCCTGGCACTCAAAATTGGACTAGCACCTATACACTTCTGAATACC
TGAAGTTTTACAGGGGCTAGATCTTTTAACCGGACTAATTTTATCCACCTGAC
AAAAACTTGCTCCACTTGCTCTCATTATTCAAACAGCCCAAGCTATCGACCCC
CTACTGCTGACGGCCCTGGGCCTATTGTCCACACTAATCGGTGGTTGAGGAG
GGCTAAACCAGACCCAGCTCCGGAAAATTTTAGCCCTTGCAATTACCTGATC
GAAAAGCCCTGTCTTGACAGCAACCACTGCCTTAGTATTACTGTCATTGGGCG

GTTTGCCTCCTCTTACAGGGTTTATGCCAAAATGACTTATTTTACAAGAGTTG
GCAAAACAGGGCCTTCCCCTTGCCGCTACAGTTATGGCTCTCGCAGCCCTCCT
CAGCCTGTACTTCTACTTACGACTTTGCTATGCAATGACACTCACCATTTCTCC
TAACACTGTTAATTCAGCCACGCCCTGACGGGTTCAAACAACTCAAGCCTCTC
TCCCCGTGGCCCTAGCTACTACGCTAGCACTGGGCCTTCTTCCCATAACTCCA
GCCATTGTAATGCTAGTCACCTAGNNNNNNNNNNNNNNNNNNNNNNNNN

AK
CACCACCACCCCCGCGCAGTGGAAGCGGCCACAAAATACTTTCTAACCCAGG
CCACCGCAGCAGCTATAATTCTGTTTGCAAGCATAACAAATGCCTGAATCAC
AGGGGGGTGAGACATAAGTAATATATCGGATCCCATTGCCAGCACAATAGTC
ATCGCCGCCCTGGCACTCAAAATTGGACTAGCACCTATACACTTCTGAATACC
TGAAGTTTTACAGGGGCTAGATCTTTTAACCGGACTAATTTTATCCACCTGAC
AAAAACTTGCTCCACTTGCTCTCATTATTCAAACAGCCCAAGCTATCGACCCC
CTACTGCTGACGGCCCTGGGCCTATTGTCCACACTAATCGGTGGTTGAGGAG
GGCTAAACCAGACCCAGCTCCGGAAAATTTTAGCCCTTGCAATTACCTGATC
GAAAAGCCCTGTCTTGACAGCAACCACTGCCTTAGTATTACTGTCATTGGGCG
GTTTGCCTCCTCTTACAGGGTTTATGCCGAAATGACTTATTTTACAAGAGTTG
GCAAAACAGAGCCTTCCCCTTGCCGCTACAGTTATGGCTCTCGCAGCCCTCCT
CAGCCTGTACTTCTACTTACGACTCTGCTATGCAATGACACTCACCATTTCTC
CTAACACTGTTAATTCAGCCACACCCTGACGGGTTCAAACAACTCAAGCCTCT
CTCCCCGTGGCCCTAGCTACTACGCTAGCACTGGGCCTTCTTCCCATAACTCC
AGCCATTGTAATGCTAGTCACCTAGGGGCTTAGGATAACATTAGACCAAGA

AN
CACCACCACCCCCGCGCAGTGGAAGCGGCCACAAAATACTTTCTAACCCAGG
CCACCGCAGCAGCTATAATTCTGTTTGCAAGCATAACAAATGCCTGAATCAC
AGGGGGGTGAGACATAAGTAATATATCGGATCCCATTGCCAGCACAATAGTC
ATCGCCGCCCTGGCACTCAAAATTGGACTAGCACCTATACACTTCTGAATACC
TGAAGTTTTACAGGGGCTAGATCTTTTAACCGGACTAATTTTATCCACCTGAC
AAAAACTTGCTCCACTTGCTCTCATTATTCAAACAGCCCAAGCTATCGACCCC
CTACTGCTGACGGCCCTGGGCCTATTGTCCACACTAATCGGTGGTTGAGGAG
GGCTAAACCAGACCCAGCTCCGGAAAATTTTAGCCCTTGCAATTACCTGATC
GAAAAGCCCTGTCTTGACAGCAACCACTGCCTTAGTATTACTGTCATTGGGCG
GTTTGCCTCCTCTTACAGGGTTTATGCCAAAATGACTTATTTTACAAGAGTTG
GCAAAACAGAGCCTTCCCCTTGCCGCTGCAGTTATGGCTCTCGCAGCCCTCCT
CAGCCTGTACTTCTACTTACGACTCTGCTATGCAATGACACTCACCATTTCTC
CTAACACTGTTAATTCAACCACGCCCTGACGGGTTCAAACAACTCAAGCCTCT
CTCCCCGTGGCCCTAGCTACTACGCTAGCACTGGGCCTTCTTCCCATAACTCC
AGCCATTGTAATGCTAGTCACCTAGGGGCTTAGGATAACATTAGACCAAGA

AP
CACCACCACCCCCGCGCAGTGGAAGCGGCCACAAAATACTTTCTAACCCAGG
CCACCGCAGCAGCTATAATTCTGTTTGCAAGCATAACAAATGCCTGAATCAC

AGGGGGGTGAGACATAAGTAATATATCGGATCCCATTGCCAGCACAATAGTC
ATCGCCGCCCTGGCACTCAAAATTGGACTAGCACCTATACACTTCTGAATACC
TGAAGTTTTACAGGGGCTAGATCTTTTAACCGGACTAATTTTATCCACCTGAC
AAAAACTTGCTCCACTTGCTCTCATTATTCAAACAGCCCAAGCTATCGACCCC
CTACTGCTGACGGCCCTGGGCCTATTGTCCACACTAATCGGTGGTTGAGGAG
GGCTAAACCAGACCCAGCTCCGGAAAATTTTAGCCCTTGCAATTACCTGATC
GAAAAGCCCTGTCTTGACAGCAACCACTGCCTTAGTATTACTGTCATTGGGCG
GTTTGCCTCCTCTTACAGGGTTTATGCCAAAATGACTTATTTTACAAGAGTTG
GCAAAACAGAGCCTTCCCCTTGCCGCTACAGTTATGGCTCTCGCAGCCCTCCT
CAGCCTGTACTTCTACTTACGACTCTGCTATGCAATGACACTCACCATTTCTC
CTAACACTGTTAATTCAGCCACACCCTGACGGGTTCAAGCAACTCAAGCCTCT
CTCCCCGTGGCCCTAGCTACTACGCTAGCACTGGGCCTTCTTCCCATAACTCC
AGCCATTGTAATGCTAGTCACCTAGGNGNNNNNNNNNNNNNNNNNNNNNNN

AR
CACCACCACCCCCGCGCAGTGGAAGCGGCCACAAAATACTTTCTAACCCAGG
CCACCGCAGCAGCTATAATTCTGTTTGCAAGCATAACAAATGCCTGAATCAC
AGGGGGGTGAGACATAAGTAATATATCGGATCCCATTGCCAGCACAATAGTC
ATCGCCGCCCTGGCACTCAAAATTGGACTAGCACCTATACACTTCTGAATACC
TGAAGTTTTACAGGGGCTAGATCTTTTAACCGGACTAATTTTATCCACCTGAC
AAAAACTTGCTCCACTTGCTCTCATTATTCAAACAGCCCAAGCTATCGACCCC
CTACTGCTGACGGCCCTGGGCCTATTGTCCACACTAATCGGTGGTTGAGGAG
GGCTAAACCAGACCCAGCTCCGGAAAATTTTAGCCCTTGCAATTACCTGATC
GAAAAGCCCTGTCTTGACAGCAACCACTGCCTTAGTATTACTGTCATTGGGTG
GTTTGCCTCCTCTTACAGGGTTTATGCCAAAATGACTTATTTTACAAGAGTTG
GCAAAACAGAGCCTTCCCCTTGCCGCTACAGTTATGGCTCTCGCAGCCCTCCT
CAGCCTGTACTTCTACTTACGACTCTGCTATGCAATGACACTCACCATTTCTC
CTAACACTGTTAATTCAGCCACACCCTGACGGGTTCAAACAACTCAAGCCTCT
CTCCCCGTGGCCCTAGCTACTACGCTAGCACTGGGCCTTCTTCCCATAACTCC
AGCCATTGTAATGCTAGTCACCTAGGGGCTTAGGATAACATTAGACCAAGA

AS
CACCACCACCCCCGCGCAGTGGAAGCGGCCACAAAATACTTTCTAACCCAGG
CCACCGCAGCAGCTATAATTCTGTTTGCAAGCATAACAAATGCCTGAATCAC
AGGGGGGTGAGACATAAGTAATATATCGGATCCCATTGCCAGCACAATAGTC
ATCGCCGCCCTGGCACTCAAAATTGGACTAGCACCTATACACTTCTGAATACC
TGAAGTTTTACAGGGGCTAGATCTTTTAACCGGACTAATTTTATCCACCTGAC
AAAAACTTGCTCCACTTGCTCTCATTATTCAAACAGCCCAAGCTATCGACCCC
CTACTGCTGACGGCCCTGGGCCTATTGTCCACACTAATCGGTGGTTGAGGAG
GGCTAAACCAGACCCAGCTCCGGAAAATTTTAGCCCTTGCAATTACCTGATC
GAAAAGCCCTGTCTTGACAGCAACCACTGCCTTAGTATTACTGTCATTGGGCG
GTTTGCCTCCTCTTACAGGGTTTATGCCAAAATGACTTATTTTACAAGAGTTG
GCAAAACAGAGCCTTCCCCTTGCCGCTACAGTTATGGCTCTCGCAGCCCTCCT
CAGCCTGTACTTCTACTTACGACTCTGCTATGCAATGACACTCACCATTTCTC

CTAACACTGTTAATTCAGCCACACCCTGACGGGTTCAAGCAACTCAAGCCTCT
CTCCCCGTGGCCCTAACTACTACGCTAGCACTGGGCCTTCTTCCCATAACTCC
AGCCATTGTAATGCTAGTCACCTANNNNNNNNNNNNNNNNNNNNNNNNNN

AT
CACCACCACCCCCGCGCAGTGGAAGCGGCCACAAAATACTTTCTAACCCAGG
CCACCGCAGCAGCTATAATTCTGTTTGCAAGCATAACAAATGCCTGAATCAC
AGGGGGGTGAGACATAAGTAATATATCGGATCCCATTGCCAGCACAATAGTC
ATCGCCGCCCTGGCACTCAAAATTGGACTAGCACCTATACACTTCTGAATACC
TGAAGTTTTACAGGGGCTAGATCTTTTAACCGGACTAATTTTATCCACCTGAC
AAAAACTTGCTCCACTTGCTCTCATTATTCAAACAGCCCAAGCTATCGACCCC
CTACTGCTGACGGCCCTGGGCCTATTGTCCACACTAATCGGTGGTTGAGGAG
GGCTAAACCAGACCCAGCTCCGGAAAATTTTAGCCCTTGCAATTACCTGATC
GAAAAGCCCTGTCTTGACAGCAACCACTGCCTTAGTATTACTGTCATTGGGCG
GCTTGCCTCCTCTTACAGGGTTTATGCCAAAATGACTTATTTTACAAGAGTTG
GCAAAACAGAGCCTTCCCCTTGCCGCTACAGTTATGGCTCTCGCAGCCCTCCT
CAGCCTGTACTTCTACTTACGACTCTGCTATGCAATGACACTCACCATTTCTC
CTAACACTGTTAATTCAGCCACACCCTGACGGGTTCAAACAACTCAAGCCTCT
CTCCCCGTGGCCCTAGCTACTACGCTAGCACTGGGCCTTCTTCCCATAACTCC
AGCCATTGTAATGCTAGTCACCTAGGGGNNNNNNNNNNNNNNNNNNNNNNN

BA
CACCACCACCCCCGCGCAGTGGAAGCGGCCACAAAATACTTTCTAACCCAGG
CCACCGCAGCAGCTATAATTCTGTTTGCAAGCATAACAAATGCCTGAATCAC
AGGGGGGTGAGACATAAATAATATATCGGATCCCATTGCCAGCACAATAGTC
ATCGCCGCCCTGGCACTCAAAATTGGACTAGCACCTATACACTTCTGAATACC
TGAAGTTTTACAGGGGCTAGATCTTTTAACCGGACTAATTTTATCCACCTGAC
AAAAACTTGCTCCACTTGCTCTCATTATTCAAACAGCCCAAGCTATCGACCCC
CTACTGCTGACGGCCCTGGGCCTATTGTCCACACTAATCGGTGGTTGAGGAG
GGCTAAACCAGACCCAGCTCCGGAAAATTTTAGCCCTTGCAATTACCTGATC
GAAAAGCCCTGTCTTGACAGCAACCACTGCCTTAGTATTACTGTCATTGGGCG
GTTTGCCTCCTCTTACAGGGTTTATGCCAAAATGACTTATTTTACAAGAGTTG
GCAAAACAGAGCCTTCCCCTTGCCGCTACAGTTATGGCTCTCGCAGCCCTCCT
CAGCCTGTACTTCTACTTACGACTCTGCTATGCAATGACACTCACCATTTCTC
CTAACACTGTTAATTCAGCCACGCCCTGACGGGTTCAAACAACTCAAGCCTCT
CTCCCCGTGGCCCTAGCTACTACGCTAGCACTGGGCCTTCTTCCCATAACTCC
AGCCATTGTAATGCTAGTCACCTAGGGGCTTAGGATAACATTAGACCAAGA

BG
CACCACCACCCCCGCGCAGTGGAAGCGGCCACAAAATACTTTCTAACCCAGG
CCACCGCAGCAGCTATAATTCTGTTTGCAAGCATAACAAATGCCTGAATCAC
AGGGGGGTGAGACATAAATAATATATCGGATCCCATTGCCAGCACAATAGTC
ATCGCCGCCCTGGCACTCAAAATTGGACTAGCACCTATACACTTCTGAATACC
TGAAGTTTTACAGGGGCTAGATCTTTTAACCGGACTAATTTTATCCACCTGAC

AAAAACTTGCTCCACTTGCTCTCATTATTCAAACAGCCCAAGCTATCGACCCC
CTACTGCTGACGGCCCTGGGCCTATTGTCCACACTAATCGGTGGTTGAGGAG
GGCTAAACCAGACCCAGCTCCGGAAAATTTTAGCCCTTGCAATTACCTGATC
GAAAAGCCCTGTCTTGACAGCAACCACTGCCTTAGTATTACTGTCATTGGGCG
GTTTGCCTCCTCTTACAGGGTTTATGCCAAAATGACTTATTTTACAAGAGTTG
ACAAAACAGAGCCTTCCCCTTGCCGCTACAGTTATGGCTCTCGCAGCCCTCCT
CAGCCTGTACTTCTACTTACGACTCTGCTATGCAATGACACTCACCATTTCTC
CTAACACTGTTAATTCAGCCACACCCTGGCGGGTTCAAACAACTCAAGCCTCT
CTCCCCGTGGCCCTAGCTACTACGCTAGCACTGGGCCTTCTTCCCATAACTCC
AGCCATTGTAATGCTAGTCACCTAGGGGCTTAGGATAACATTAGACCAAGA

CA
CACCACCACCCCCGCGCAGTGGAAGCGGCCACAAAATACTTTCTAACCCAGG
CCGCCGCAGCAGCTATAATTCTGTTTGCAAGCATAACAAATGCCTGAATCAC
AGGGGGGTGAGACATAAGTAATATATCGGATCCCATTGCCAGCACAATAGTC
ATCGCCGCCCTGGCACTCAAAATTGGACTAGCACCTATACACTTCTGAATACC
TGAAGTTTTACAGGGGCTAGATCTTTTAACCGGACTAATTTTATCCACCTGAC
AAAAACTTGCTCCACTTGCTCTCATTATTCAAACAGCCCAAGCTATCGACCCC
CTACTGCTGACGGCCCTGGGCCTATTGTCCACACTAATCGGTGGTTGAGGAG
GGCTAAACCAGACCCAGCTCCGGAAAATTTTAGCCCTTGCAATTACCTGATC
GAAAAGCCCTGTCTTGACAGCAACCACTGCCTTAGTATTACTGTCATTGGGCG
GTTTGCCTCCTCTTACAGGGTTTATGCCAAAATGACTTATTTTACAAGAGTTG
GCAAAACAGAGCCTTCCCCTTGCCGCTACAGTTATGGCTCTCGCAGCCCTCCT
CAGCCTGTACTTCTACTTACGACTCTGCTATGCAATGACACTCACCATTTCTC
CTAACACTGTTAATTCAGCCACGCCCTGACGGGTTCAAACAACTCAAGCCTCT
CTCCCCGTGGCCCTAGCTACTACGCTAGCACTGGGCCTTCTTCCCATAACTCC
AGCCATTGTAATGCTAGTCACCTAGGGGCTTAGGATAACATTAGACCAAGA

CB
CACCACCACCCCCGCGCAGTGGAAGCGGCCACAAAATACTTTCTAACCCAGG
CCGCCGCAGCAGCTATAATTCTGTTTGCAAGCATAACAAATGCCTGAATCAC
AGGGGGGTGAGACATAAGTAATATATCGGATCCCATTGCCAGCACAATAGTC
ATCGCCGCCCTGGCACTCAAAATTGGACTAGCACCTATACACTTCTGAATACC
TGAAGTTTTACAGGGGCTAGATCTTTTAACCGGACTAATTTTATCCACCTGAC
AAAAACTTGCTCCACTTGCTCTCATTATTCAAACAGCCCAAGCTATCGACCCC
CTACTGCTGACGGCCCTGGGCCTATTGTCCACACTAATCGGTGGTTGAGGAG
GGCTAAACCAGACCCAGCTCCGGAAAATTTTAGCCCTTGCAATTACCTGATC
GAAAAGCCCTGTCTTGACAGCAACCACTGCCTTAGTATTACTGTCATTGGGCG
GTTTGCCTCCTCTTACAGGGTTTATGCCAAAATGACTTATTTTACAAGAGTTG
GCAAAACAGAGCCTTCCCCTTGCCGCTACAGTTATGGCTCTCGCAGCCCTCCT
CAGCCTGTACTTCTACTTACGACTCTGCTATGCAATGACACTCACCATTTCTC
CTAACACTGTTAATTCAGCCACACCCTGACGGGTTCAAACAACTCAAGCCTCT
CTCCCCGTGGCCCTAGCTACTACGCTAGCACTGGGCCTTCTTCCCATAACTCC
AGCCATTGTAATGCTAGTCACCTAGGGGCTTAGGATAACATTAGACCAAGA

CK
CACCACCACCCCCGCGCAGTGGAAGCGGCCACAAAATACTTTCTAACCCAGG
CCGCCGCAGCAGCTATAATTCTGTTTGCAAGCATAACAAATGCCTGAATCAC
AGGGGGGTGAGACATAAGTAATATATCGGATCCCATTGCCAGCACAATAGTC
ATCGCCGCCCTGGCACTCAAAATTGGACTAGCACCTATACACTTCTGAATACC
TGAAGTTTTACAGGGGCTAGATCTTTTAACCGGACTAATTTTATCCACCTGAC
AAAAACTTGCTCCACTTGCTCTCATTATTCAAACAGCCCAAGCTATCGACCCC
CTACTGCTGACGGCCCTGGGCCTATTGTCCACACTAATCGGTGGTTGAGGAG
GGCTAAACCAGACCCAGCTCCGGAAAATTTTAGCCCTTGCAATTACCTGATC
GAAAAGCCCTGTCTTGACAGCAACCACTGCCTTAGTATTACTGTCATTGGGCG
GTTTGCCTCCTCTTACAGGGTTTATGCCGAAATGACTTATTTTACAAGAGTTG
GCAAAACAGAGCCTTCCCCTTGCCGCTACAGTTATGGCTCTCGCAGCCCTCCT
CAGCCTGTACTTCTACTTACGACTCTGCTATGCAATGACACTCACCATTTCTC
CTAACACTGTTAATTCAGCCACACCCTGACGGGTTCAAACAACTCAAGCCTCT
CTCCCCGTGGCCCTAGCTACTACGCTAGCACTGGGCCTTCTTCCCATAACTCC
AGCCATTGTAATGCTAGTCACCTAGGGGCTTAGGATAACATTAGACCAAGA

DB
CACCACCACCCCCGCGCAGTGGAAGCGGCCACAAAATACTTTCTAACCCAGG
CCACCGCAGCAGCTATAATTCTGTTTGCAAGCATAACAAATGCCTGAATCAC
AGGGGAGTGAGACATAAGTAATATATCGGATCCCATTGCCAGCACAATAGTC
ATCGCCGCCCTGGCACTCAAAATTGGACTAGCACCTATACACTTCTGAATGCC
TGAAGTTTTACAAGGGCTAGATCTTTTAACCGGACTAATTTTATCCACCTGAC
AAAAACTTGCTCCACTTGCTCTCATTATTCAAACAGCCCAAGCTATCGACCCC
CTACTGCTGACGGCCCTGGGCCTATTGTCCACACTAATCGGTGGTTGAGGAG
GGCTAAACCAGACCCAGCTCCGGAAAATTTTAGCCCTTGCAATTACCTGATC
GAAAAGCCCTGTCTTGACAGCAACCACTGCCTTAGTATTACTGTCATTGGGCG
GTTTGCCTCCTCTTACAGGGTTTATGCCAAAATGACTTATTTTACAAGAGTTG
GCAAAACAGAGCCTTCCCCTTGCCGCTACAGTTATGGCTCTCGCAGCCCTCCT
CAGCCTGTACTTCTACTTACGACTCTGCTATGCAATGACACTCACCATTTCTC
CTAACACTGTTAATTCAGCCACACCCTGACGGGTTCAAACAACTCAAGCCTCT
CTCCCCGTGGCCCTAGCTACTACGCTAGCACTGGGCCTTCTTCCCATAACTCC
AGCCATTGTAATGCTAGTCACCTAGGGGCTTAGGATAACATTAGACCAAGA

EB
CACCACCACCCCCGCGCAGAGGAAGCGGCCACAAAATACTTTCTAACCCAGG
CCACCGCAGCAGCTATAATTCTGTTTGCAAGCATAACAAATGCCTGAATCAC
AGGGGGGTGAGACATAAGTAATATGTCGGATCCCATTGCCAGCACAATAGTC
ATCGCCGCCCTGGCACTCAAAATTGGACTAGCACCTATACACTTCTGAATACC
TGAAGTTTTACAGGGGCTAGATCTTTTAACCGGACTAATTTTATCCACCTGAC
AAAAGCTTGCTCCACTTGCTCTCATTATTCAAACAGCCCAAGCTATCGACCCC
CTACTGCTGACGGCCCTGGGCCTATTGTCCACACTAATCGGtGGTTGAGGAGG
GCTAAACCAGACCCAGCTCCGGAAAATTTTAGCCCTTGCAATTACCTGATCG
AAAAGCCCTGTCTTGACAGCAACCACTGCCTTAGTATTACTGTCATTGGGCGG

TTTGCCTCCTCTTACAGGGTTTATGCCAAAATGACTTATTTTACAAGAGTTGG
CAAAACAGAGCCTTCCCCTTGCCGCTACAGTTATGGCTCTCGCAGCCCTCCTC
AGCCTGTACTTCTACTTACGACTCTGCTATGCAATGACACTCACCATTTCTCCT
AACACTGTTAATTCAGCCACACCCTGACGGGTTCAAACAACTCAAGCCTCTCT
CCCCGTGGCCCTAGCTACTACGCTAGCACTGGGCCTTCTTCCCATAACTCCAG
CCATTGTAATGCTAGTCACCTAGGGGCTTAGGATAACATTAGACCAAGA

FA
CACCACCACCCCCGCGCAGTGGAAGCGGCCACAAAATACTTTCTAACCCAGG
CCACCGCAGCAGCTATAATTCTGTTTGCAAGCATAACAAATGCATGAATCAC
AGGGGGGTGAGACATAAGTAATATATCGGATCCCATTGCCAGCACAATAGTC
ATCGCCGCCCTGGCACTCAAAATTGGACTAGCACCTATACACTTCTGAATACC
TGAAGTTTTACAGGGGCTANATCTTTTAACCGGACTAATTTTATCCACCTGAC
AAAAACTTGCTCCACTTGCTCTCATTATTCAAACAGCCCAAGCTATCGACCCC
CTACTGCTGACGGCCCTGGGCCTATTGTCCACACTAATCGGTGGTTGAGGAG
GGCTAAACCAAACCCAGNTCCGGAAAATTTTAGCCCTTGCAATTACCTGATC
GAAAAGCCCTGTCTTGACAGCAACCACTGCCTTAGTATTACTGTCATTGGGCG
GTTTGCCTCCTCTTACAGGGTTTATGCCAAAATGACTTATTTTACAAGAGTTG
GCAAAACAGAGCCTTCCCCTTGCCGCTACAGTTATGGCTCTCGCAGCCCTCCT
CAGCCTGTACTTCTACTTACGACTCTGCTATGCAATGACACTCACCATTTCTC
CTAACACTGTTAATTCAGCCACGCCCTGACGGGTTCAAACAACTCAAGCCTCT
CTCCCCGTGGCCCTAGCTACTACGCTAGCACTGGGCCTTCTTCCCATAACTCC
AGCCATTGTAATGCTAGTCACCTAGGGGCTTAGGATAACATTAGACCAAGA

FC
CACCACCACCCCCGCGCAGTGGAAGCGGCCACAAAATACTTTCTAACCCAGG
CCACCGCAGCAGCTATAATTCTGTTTGCAAGCATAACAAATGCATGAATCAC
AGGGGGGTGAGACATAAGTAATATATCGGATCCCATTGCCAGCACAATAGTC
ATCGCCGCCCTGGCACTCAAAATTGGACTAGCACCTATACACTTCTGAATACC
TGAAGTTTTACAGGGGCTANATCTTTTAACCGGACTAATTTTATCCACCTGAC
AAAAACTTGCTCCACTTGCTCTCATTATTCAAACAGCCCAAGCTATCGACCCC
CTACTGCTGACGGCCCTGGGCCTATTGTCCACACTAATCGGTGGTTGAGGAG
GGCTAAACCAAACCCAGNTCCGGAAAATTTTAGCCCTTGCAATTACCTGATC
GAAAAGCCCTGTCTTGACAGCAACCACTGCCTTAGTATTACTGTCATTGGGCG
GTTTGCCTCCTCTTACAGGGTTTATGCCAAAATGACTTATTTTACAAGAGTTG
GCAAAACAGAGCCTTCCCCTTGCCGCCACAGTTATGGCTCTCGCAGCCCTCCT
CAGCCTGTACTTCTACTTACGACTCTGCTATGCAATGACACTCACCATTTCTC
CTAACACTGTTAATTCAGCCACACCCTGACGGGTTCAAACAACTCAAGCCTCT
CTCCCCGTGGCCCTAGCTACTACGCTAGCACTGGGCCTTCTTCCCATAACTCC
AGCCATTGTAATGCTAGTCACCTAGGGGCTTAGGATAACATTAGACCAAGA

GI
CACCACCACCCCCGCGCAGTGGAAGCGGCCACAAAATACTTTCTAACCCAGG
CCACCGCAGCAGCTATAATTCTGTTTGCAAGCATAACAAATGCCTGAATCAC

AGGGGGGTGAGACATAAGTAATATATCGGATCCCATTGCCAGCACAATAGTC
ATCGCCGCCCTGGCACTCAAAATTGGACTAGCACCTATACACTTCTGAATACC
TGAAGTTTTACAGGGGCTAGATCTTTTAACCGGACTAATTTTATCCACCTGAC
AAAAGCTTGCTCCACTTGCTCTCATTATTCAAACAGCCCAAGCTATCGACCCC
CTACTGCTGACGGCCCTGGGCCTATTGTCCACACTAATCGGTGGTTGAGGGG
GGCTAAACCAGACCCAGCTCCGGAAAATTTTAGCCCTTGCAATTACCTGATC
GAAAAGCCCTGTCTTGACAGCAACCACTGCCTTAGTATTACTGTCATTGGGCG
GTTTGCCTCCTCTTACAGGGTTTATGCCAAAATGACTTATTTTACAAGAGTTG
GCAAAACAGAGCCTTCCCCTTGCCGCTACAGTTATGGCTCTCGCAGCCCTCCT
CAGCCTGTACTTCTACTTACGACTCTGCTATGCAATGACACTCACCATTTCTC
CTAACACTGTTAATTCAGCCACACCCTGACGGGCTCAAACAACTCAAGCCTC
TCTCCCCGTGGCCCTAGCTACTACGCTAGCACTGGGCCTTCTTCCCATAACTC
CAGCCATTGTAATGCTAGTCACCTAGGGGCTTAGGATAACATTAGACCAAGA

GK
CACCACCACCCCCGCGCAGTGGAAGCGGCCACAAAATACTTTCTAACCCAGG
CCACCGCAGCAGCTATAATTCTGTTTGCAAGCATAACAAATGCCTGAATCAC
AGGGGGGTGAGACATAAGTAATATATCGGATCCCATTGCCAGCACAATAGTC
ATCGCCGCCCTGGCACTCAAAATTGGACTAGCACCTATACACTTCTGAATACC
TGAAGTTTTACAGGGGCTAGATCTTTTAACCGGACTAATTTTATCCACCTGAC
AAAAGCTTGCTCCACTTGCTCTCATTATTCAAACAGCCCAAGCTATCGACCCC
CTACTGCTGACGGCCCTGGGCCTATTGTCCACACTAATCGGTGGTTGAGGGG
GGCTAAACCAGACCCAGCTCCGGAAAATTTTAGCCCTTGCAATTACCTGATC
GAAAAGCCCTGTCTTGACAGCAACCACTGCCTTAGTATTACTGTCATTGGGCG
GTTTGCCTCCTCTTACAGGGTTTATGCCGAAATGACTTATTTTACAAGAGTTG
GCAAAACAGAGCCTTCCCCTTGCCGCTACAGTTATGGCTCTCGCAGCCCTCCT
CAGCCTGTACTTCTACTTACGACTCTGCTATGCAATGACACTCACCATTTCTC
CTAACACTGTTAATTCAGCCACACCCTGACGGGTTCAAACAACTCAAGCCTCT
CTCCCCGTGGCCCTAGCTACTACGCTAGCACTGGGCCTTCTTCCCATAACTCC
AGCCATTGTAATGCTAGTCACCTAGGGGCTTAGGATAACATTAGACCAAGA

HB
CACCACCACCCCCGCGCAGTGGAAGCGGCCACAAAATACTTTCTAACCCAGG
CCACCGCAGCAGCTATAATTCTGTTTGCAAGCATAACAAATGCCTGAATCAC
AGGGGGGTGAGACATAAGTAATATATCGGATCCCATTGCCAGCACAATAGTC
ATCGCCGCCCTGGCACTCAAAATTGGACTAGCACCTATACACTTCTGAATACC
TGAAGTTTTACAGGGGCTAGATCTTTTAACCGGACTAATTTTATCCACCTGAC
AAAAGCTTGCTCCACTTGCTCTCATTATTCAAACAGCCCAAGCTATCGACCCC
CTACTGCTGACGGCCCTGGGCCTATTGTCCACACTAATCGGTGGTTGAGGAG
GGCTAAACCAGACCCAGCTCCGGAAAATTTTAGCCCTTGCAATTACCTGATC
GAAAAGCCCTGTCTTGACAGCAACCACTGCCTTAGTATTACTGTCATTGGGCG
GTTTGCCTCCTCTTACAGGGTTTATGCCAAAATGACTTATTTTACAAGAGTTG
GCAAAACAGAGCCTTCCCCTTGCCGCTACAGTTATGGCTCTCGCAGCCCTCCT
CAGCCTGTACTTCTACTTACGACTCTGCTATGCAATGACACTCACCATTTCTC

CTAACACTGTTAATTCAGCCACACCCTGACGGGTTCAAACAACTCAAGCCTCT
CTCCCCGTGGCCCTAGCTACTACGCTAGCACTGGGCCTTCTTCCCATAACTCC
AGCCATTGTAATGCTAGTCACCTAGGGGCTTAGGATAACATTAGACCAAGA

HO
CACCACCACCCCCGCGCAGTGGAAGCGGCCACAAAATACTTTCTAACCCAGG
CCACCGCAGCAGCTATAATTCTGTTTGCAAGCATAACAAATGCCTGAATCAC
AGGGGGGTGAGACATAAGTAATATATCGGATCCCATTGCCAGCACAATAGTC
ATCGCCGCCCTGGCACTCAAAATTGGACTAGCACCTATACACTTCTGAATACC
TGAAGTTTTACAGGGGCTAGATCTTTTAACCGGACTAATTTTATCCACCTGAC
AAAAGCTTGCTCCACTTGCTCTCATTATTCAAACAGCCCAAGCTATCGACCCC
CTACTGCTGACGGCCCTGGGCCTATTGTCCACACTAATCGGTGGTTGAGGAG
GGCTAAACCAGACCCAGCTCCGGAAAATTTTAGCCCTTGCAATTACCTGATC
GAAAAGCCCTGTCTTAACAGCAACCACTGCCTTAGTATTACTGTCATTGGGCG
GTTTGCCTCCTCTTACAGGGTTTATGCCAAAATGACTTATTTTACAAGAGTTG
GCAAAACAGAGCCTTCCCCTTGCCGCCACAGTTATGGCTCTCGCAGCCCTCCT
CAGCCTGTACTTCTACTTACGACTCTGCTATGCGATGACACTCACCATTTCTC
CTAACACTGTTAATTCAGCCACACCCTGACGGGTTCAAACAACTCAAGCCTCT
CTCCCCGTGGCCCTAGCTACTACGCTAGCACTGGGCCTTCTTCCCATAACTCC
AGCCATTGTAATGCTAGTCACCTAGGGGCTTAGGATAACATTAGACCAAGA

IB
CACCACCACCCCCGCGCAGTGGAAGCGGCCACAAAATACTTTCTAACCCAGG
CCACCGCAGCAGCTATAATTCTGTTTGCAAGCATAACAAATGCCTGAATCAC
AGGGGGGTGAGACATAAGTAATATATCGGATCCCATTGCCAGCACAATAGTC
ATCGCCGCCCTGGCACTCAAAATTGGACTAGCACCTATACACTTCTGAATACC
TGAAGTTTTACAGGGGCTAGATCTTTTAACCGGACTAATTTTATCCACCTGAC
AAAAACTTGCTCCACTTGCTCTCATTATTCAAACAGCCCAAGCTATCGACCCC
CTACTGCTGACGGCCCTAGGCCTATTGTCCACACTAATCGGTGGCTGAGGAG
GGCTAAACCAGACCCAGCTCCGGAAAATTTTAGCCCTTGCAATTACCTGATC
GAAAAGCCCTGTCTTGACAGCAACCACTGCCTTAGTATTACTGTCATTGGGCG
GTTTGCCTCCTCTTACAGGGTTTATGCCAAAATGACTTATTTTACAAGAGTTG
GCAAAACAGAGCCTTCCCCTTGCCGCTACAGTTATGGCTCTCGCAGCCCTCCT
CAGCCTGTACTTCTACTTACGACTCTGCTATGCAATGACACTCACCATTTCTC
CTAACACTGTTAATTCAGCCACACCCTGACGGGTTCAAACAACTCAAGCCTCT
CTCCCCGTGGCCCTAGCTACTACGCTAGCACTGGGCCTTCTTCCCATAACTCC
AGCCATTGTAATGCTAGTCACCTAGGGGCTTAGGATAACATTAGACCAAGA

IK
CACCACCACCCCCGCGCAGTGGAAGCGGCCACAAAATACTTTCTAACCCAGG
CCACCGCAGCAGCTATAATTCTGTTTGCAAGCATAACAAATGCCTGAATCAC
AGGGGGGTGAGACATAAGTAATATATCGGATCCCATTGCCAGCACAATAGTC
ATCGCCGCCCTGGCACTCAAAATTGGACTAGCACCTATACACTTCTGAATACC
TGAAGTTTTACAGGGGCTAGATCTTTTAACCGGACTAATTTTATCCACCTGAC

AAAAACTTGCTCCACTTGCTCTCATTATTCAAACAGCCCAAGCTATCGACCCC
CTACTGCTGACGGCCCTAGGCCTATTGTCCACACTAATCGGTGGCTGAGGAG
GGCTAAACCAGACCCAGCTCCGGAAAATTTTAGCCCTTGCAATTACCTGATC
GAAAAGCCCTGTCTTGACAGCAACCACTGCCTTAGTATTACTGTCATTGGGCG
GTTTGCCTCCTCTTACAGGGTTTATGCCGAAATGACTTATTTTACAAGAGTTG
GCAAAACAGAGCCTTCCCCTTGCCGCTACAGTTATGGCTCTCGCAGCCCTCCT
CAGCCTGTACTTCTACTTACGACTCTGCTATGCAATGACACTCACCATTTCTC
CTAACACTGTTAATTCAGCCACACCCTGACGGGTTCAAACAACTCAAGCCTCT
CTCCCCGTGGCCCTAGCTACTACGCTAGCACTGGGCCTTCTTCCCATAACTCC
AGCCATTGTAATGCTAGTCACCTAGGGGCTTAGGATAACATTAGACCAAGA

JB
CACCACCACCCCCGCGCAGTGGAAGCGGCCACAAAATACTTTCTAACCCAGG
CCACCGCAGCAGCTATAATTCTGTTTGCAAGCATAACAAATGCCTGAATCAC
AGGGGGGTGAGACATAAGTAATATATCGGATCCCATTGCCAGCACAATAGTC
ATCGCCGCCCTGGCACTCAAAATTGGACTAGCACCTATACACTTCTGAATACC
TGAAGTTTTACAGGGGCTAGATCTTTTAACCGGACTAATTTTATCCACCTGAC
AAAAACTTGCTCCACTTGCTCTCATTATTCAAACAGCCCAAGCTATCGACCCC
CTACTGCTGACGGCCCTGGGCCTATTGTCCGCACTAATCnGTGGTTGAgGAGG
GCTAAACCAGACCCAGCTCCGGAAAATTTTAGCCCTTGCAATTACCTGATCG
AAAAGCCCTGTCTTGACAGCAACCACTGCCTTAGTATTACTGTCATTGGGCGG
TTTGCCTCCTCTTACAGGGTTTATGCCAAAATGACTTATTTTACAAGAGTTGG
CAAAACAGAGCCTTCCCCTTGCCGCTACAGTTATGGCTCTCGCAGCCCTCCTC
AGCCTGTACTTCTACTTACGACTCTGCTATGCAATGACACTCACCATTTCTCCT
AACACTGTTAATTCAGCCACACCCTGACGGGTTCAAACAACTCAAGCCTCTCT
CCCCGTGGCCCTAGCTACTACGCTAGCACTGGGCCTTCTTCCCATAACTCCAG
CCATTGTAATGCTAGTCACCTAGGGGCTTAGGATAACATTAGACCAAGA

KM
CACCACCACCCCCGCGCAGNGGAAGCGGCCACAAAATACTTTCTAACCCAGG
CCACCGCAGCAGCTATAATTCTGTTTGCAAGCATAACAAATGCCTGAATCAC
AGGGGGGTGAGACATAAGNAATATATCGGATCCCATTGCCAGCACAATAGTC
ATCGCCGCCCTGGCACTCAAAATTGGACTAGCACCTATACACTTATGAATAC
CTGAAGTTTTACAGGGGCTAGATCTTTTAACCGGACTAATTTTATCCACCTGA
CAAAAACTTGCTCCACTTGCTCTCATTATTCAAACAGCCCAAGCTATCGACCC
CCTACTGCTGACGGTCCTGGGCCTATTGTCCACACTAATCGGTGGTTGAGGAG
GGCTAAACCAGACCCAGCTCCGGAAAATTTTAGCCCTTGCAATTACCTGATC
GAAAAGCCCTGTCTTGACAGCAACCACTGCCTTAGTATTACTGTCATTGGGCG
GTTTGCCTCCTCTTACAGGGTTTATGCCAAAATGACTTATTTTACAAGAGTTG
GCAAAACAGAGCCTTCCCCTTGCCGCTACAGTTATGGCTCTCGCAGCCCTCCT
TAGCCTGTACTTCTACTTACGACTCTGCTATGCAATGACACTCACCATTTCTCC
TAACACTGTTAATTCAGCCACGCCCTGACGGGTTCAAACAACTCAAGCCTCTC
TCCCCGTGGCCCTAGCTACTACGCTAGCACTGGGCCTTCTTCCCATAACTCCA
GCCATTGTAATGCTAGTCACCTAGGGGCTTAGGATAACATTAGACCAAGA

LA
CACCACCACCCCCGCGCAGTGGAAGCGGCCACAAAATACTTTCTAACCCAGG
CCACCGCAACAGCTATAATTCTGTTTGCAAGCATAACAAATGCCTGAATCAC
AGGGGGGTGAGACATAAGTAATATATCGGATCCCATTGCCAGCACAATAGTC
ATCGCCGCCCTGGCACTCAAAATTGGACTAGCACCTATACACTTCTGAATACC
TGAAGTTTTACAGGGGCTAGATCTTTTAACCGGACTAATTTTATCCACCTGAC
AAAAACTTGCTCCACTTGCTCTCATTATTCAAACAGCCCAAGCTATCGACCCC
CTACTGCTGACGGCCCTGGGCCTATTGTCCACACTAATCGGTGGTTGAGGAG
GGCTAAACCAGACCCAGCTCCGGAAAATTTTAGCCCTTGCAATTACCTGATC
GAAAAGCCCTGTCTTGACAGCAACCACTGCCTTAGTATTACTGTCATTGGGCG
GTTTGCCTCCTCTTACAGGGTTTATGCCAAAATGACTTATTTTACAAGAGTTG
GCAAAACAGAGCCTTCCCCTTGCCGCTACAGTTATGGCTCTCGCAGCCCTCCT
CAGCCTGTACTTCTACTTACGACTCTGCTATGCAATGACACTCACCATTTCTC
CTAACACTGTTAATTCAGCCACGCCCTGACGGGTTCAAACAACTCAAGCCTCT
CTCCCCGTGGCCCTAGCTACTACGCTAGCACTGGGCCTTCTTCCCATAACTCC
AGCCATTGTAATGCTAGTCACCTAGGGGCTTAGGATAACATTAGACCAAGA

MB
CACCACCACCCCCGCGCAGAGGAAGCGGCCACAAAATACTTTCTAACCCAGG
CCACCGCAGCAGCTATAATTCTGTTTGCAAGCATAACAAATGCCTGAATCAC
AGGGGGGTGAGACATAAGTAATATATCGGATCCCATTGCCAGCACAATAGTC
ATCGCCGCCCTGGCACTCAAAATTGGACTAGCACCTATACACTTCTGAATACC
TGAAGTTTTACAGGGGCTAGATCTTTTAACCGGACTAATTTTATCCACCTGAC
AAAAACTTGCTCCACTTGCTCTCATTATTCAAACAGCCCAAGCTATCGACCCC
CTACTGCTGACGGCCCTGGGCCTATTGTCCACACTAATCGGTGGTTGAGGAG
GGCTAAACCAGACCCAGCTCCGGAAAATTTTAGCCCTTGCAATTACCTGATC
GAAAAGCCCTGTCTTGACAGCAACCACTGCCTTAGTATTACTGTCATTGGGCG
GTTTGCCTCCTCTTACAGGGTTTATGCCAAAATGACTTATTTTACAAGAGTTG
GCAAAACAGAGCCTTCCCCTTGCCGCTACAGTTATGGCTCTCGCAGCCCTCCT
CAGCCTGTACTTCTACTTACGACTCTGCTATGCAATGACACTCACCATTTCTC
CTAACACTGTTAATTCAGCCACACCCTGACGGGTTCAAACAACTCAAGCCTCT
CTCCCCGTGGCCCTAGCTACTACGCTAGCACTGGGCCTTCTTCCCATAACTCC
AGCCATTGTAATGCTAGTCACCTAGGGGCTTAGGATAACATTAGACCAAGA

NA
CACCACCACCCCCGCGCAGTGGAAGCGGCCACAAAATACTTTCTAACCCAGG
CCACCGCAGCAGCTATGATTCTGTTTGCAAGCATAACAAATGCCTGAATCAC
AGGGGGATGAGACATAAGTAATATATCGGATCCCATTGCCAGCACAATAGTC
ATCGCCGCCCTGGCACTCAAAATTGGACTAGCACCTATACACTTCTGAATACC
TGAAGTTTTACAGGGACTAGATCTTTTAACCGGACTAATTTTATCCACCTGAC
AAAAACTTGCTCCACTTGCTCTCATTATTCAAACAGCCCAAGCTATCGACCCC
CTACTGCTGACGGCCCTGGGCCTATTGTCCACACTAATCGGTGGTTGAGGAG
GGCTAAACCAGACCCAGCTCCGGAAAATTTTAGCCCTTGCAATTACCTGATC
GAAAAGCCCTGTCTTGACAGCAACCACTGCCTTAGTATTACTGTCATTGGGCG

GTTTGCCTCCTCTTACAGGGTTTATGCCAAAATGACTTATTTTACAAGAGTTG
GCAAAACAGAGCCTTCCCCTTGCCGCTACAGTTATGGCTCTCGCAGCCCTCCT
CAGCCTGTACTTCTACTTACGACTCTGCTATGCAATGACACTCACCATTTCTC
CTAACACTGTTAATTCAGCCACGCCCTGACGGGTTCAAACAACTCAAGCCTCT
CTCCCCGTGGCCCTAGCTACTACGCTAGCACTGGGCCTTCTTCCCATAACTCC
AGCCATTGTAATGCTAGTCACCTAGGGGCTTAGGATAACATTAGACCAAGA

NL
CACCACCACCCCCGCGCAGTGGAAGCGGCCACAAAATACTTTCTAACCCAGG
CCACCGCAGCAGCTATGATTCTGTTTGCAAGCATAACAAATGCCTGAATCAC
AGGGGGATGAGACATAAGTAATATATCGGATCCCATTGCCAGCACAATAGTC
ATCGCCGCCCTGGCACTCAAAATTGGACTAGCACCTATACACTTCTGAATACC
TGAAGTTTTACAGGGACTAGATCTTTTAACCGGACTAATTTTATCCACCTGAC
AAAAACTTGCTCCACTTGCTCTCATTATTCAAACAGCCCAAGCTATCGACCCC
CTACTGCTGACGGCCCTGGGCCTATTGTCCACACTAATCGGTGGTTGAGGAG
GGCTAAACCAGACCCAGCTCCGGAAAATTTTAGCCCTTGCAATTACCTGATC
GAAAAGCCCTGTCTTGACAGCAACCACTGCCTTAGTATTACTGTCATTGGGCG
GTTTGCCTCCTCTTACAGGGTTTATGCCAAAATGACTTATTTTACAAGAGTTG
GCAAAACAGAGCCTTCCCCTTGCCGCTACAGTTATGGCTCTCGCAGCCCTCCT
CAGCCTGTACTTCTACTTACGACTCTGCTATGCAATGACACTCACCATTTCTC
CTAACACTGTTAATTCAGCCACGCCCTGACGGGTTCAAACAACTCAAGCCTCT
CTCCCCGTGGCCCTAGCTACTACACTAGCACTGGGCCTTCTTCCCATAACTCC
AGCCATTGTAATGCTAGTCACCTAGGGGCTTAGGATAACATTAGACCAAGA

OB
CACCACCACCCCCGCGCAGNGGAAGCGGCCACAAAATACTTTCTAACCCAGG
CCACCGCAGCAGCTATAATTCTGTTTGCAAGCATAACAAATGCCTGAATCAC
AGGGGGGTGAGACATAAGTAATATATCGGACCCCATTGCCAGCACAATAGTC
ATCGCCGCCCTGGCACTCAAAATTGGACTAGCACCTATACACTTCTGAATACC
TGAAGTTTTACAGGGGCTAGATCTTTTAACCGGACTAATTTTATCCACCTGAC
AAAAGCTTGCTCCACTTGCTCTCATTATTCAAACAGCCCAAGCTATCGACCCC
CTACTGCTGACGGCCCTGGGCCTATTGTCCACACTAATCGGTGGTTGAGGGG
GGCTAAACCAGACCCAGCTCCGGAAAATTTTAGCCCTTGCAATTACCTGATC
GAAAAGCCCTGTCTTGACAGCAACCACTGCCTTAGTATTACTGTCATTGGGCG
GTTTGCCTCCTCTTACAGGGTTTATGCCAAAATGACTTATTTTACAAGAGTTG
GCAAAACAGAGCCTTCCCCTTGCCGCTACAGTTATGGCTCTCGCAGCCCTCCT
CAGCCTGTACTTCTACTTACGACTCTGCTATGCAATGACACTCACCATTTCTC
CTAACACTGTTAATTCAGCCACACCCTGACGGGTTCAAACAACTCAAGCCTCT
CTCCCCGTGGCCCTAGCTACTACGCTAGCACTGGGCCTTCTTCCCATAACTCC
AGCCATTGTAATGCTAGTCACCTAGGGGCTTAGGATAACATTAGACCAAGA

PD
CACCACCACCCCCGCGCAGTGGAAGCGGCCACAAAATACTTTCTAACCCAGG
CCACCGCAGCAGCTATAATTCTGTTTGCAAGCATAACAAATGCCTGAATCAC

AGGGGGGTGAGACATAAGTAATATATCGGATCCCATTGCCAGCACAATAATC
ATCGCCGCCCTGGCACTCAAAATTGGACTAGCACCTATACACTTCTGAATACC
TGAAGTTTTACAGGGGCTAGATCTTTTAACCGGACTAATTTTATCCACCTGAC
AAAAACTTGCTCCACTTGCTCTCATTATTCAAACAGCCCAAGCTATCGACCCC
CTACTGCTGACGGCCCTGGGCCTATTGTCCACACTAATCGGTGGTTGAGGAG
GGCTAAACCAGACCCAGCTCCGGAAAATTTTAGCCCTTGCAATTACCTGATC
GAAAAGCCCTGTCTTGACAGCAACCACTGCCTTAGTATTACTGTCATTGGGCG
GTTTGCCTCCTCTTACAGGGTTTATGCCAAAATGACTTATTTTACAAGAGTTG
GCAAAGCAGAGCCTTCCCCTTGCCGCTACAGTTATGGCTCTCGCAGCCCTCCT
CAGCCTGTACTTCTACTTACGACTCTGCTATGCAATGACACTCACCATTTCTC
CTAACACTGTTAATTCAGCCACACCCTGACGGGTTCAAACAACTCAAGCCTCT
CTCCCCGTGGCCCTAGCTACTACGCTAGCACTGGGCCTTCTTCCCATAACTCC
AGCCATTGTAATGCTAGTCACCTAGGGGCTTAGGATAACATTAGACCAAGA

QG
CACCACCACCCCCGCGCAGTGGAAGCGGCCACAAAATACTTTCTAACCCAGG
CCACCGCAGCAGCTATAATTCTGTTTGCAAGCATAACAAATGCCTGAATCAC
AGGGGGGTGAGACATAAGTAATATATCGGACCCCATTGCCAGCACAATAGTC
ATCGCCGCCCTGGCACTCAAAATTGGACTAGCACCTATACACTTCTGAATACC
TGAAGTTTTACAGGGGCTAGATCTTTTAACCGGACTAATTTTATCCACCTGAC
AAAAACTTGCTCCACTTGCTCTCATTATTCAAACAGCCCAAGCTATCGACCCC
CTACTGCTGACGGCCCTGGGCCTATTGTCCACACTAATCGGTGGTTGAGGAG
GGCTAAACCAGACCCAGCTCCGGAAAATTTTAGCCCTTGCAATTACCTGATC
GAAAAGCCCTGTCTTGACAGCAACCACTGCCTTAGTATTACTGTCATTGGGCG
GTTTGCCTCCTCTTACAGGGTTTATGCCAAAATGACTTATTTTACAAGAGTTG
ACAAAACAGAGCCTTCCCCTTGCCGCTACAGTTATGGCTCTCGCAGCCCTCCT
CAGCCTGTACTTCTACTTACGACTCTGCTATGCAATGACACTCACCATTTCTC
CTAACACTGTTAATTCAGCCACACCCTGGCGGGTTCAAACAACTCAAGCCTCT
CTCCCCGTGGCCCTAGCTACTACGCTAGCACTGGGCCTTCTTCCCATAACTCC
AGCCATTGTAATGCTAGTCACCTAGGGGCTTAGGATAACATTAGACCAAGA

RA
CACCACCACCCCCGCGCAGTGGAAGCGGCCACAAAATACTTTCTAACCCAGG
CCACCGCAGCAGCTATAATTCTGTTTGCAAGCATAACAAATGCCTGAATCAC
AGGGGAGTGAGACATAAGTAATATATCGGATCCCATTGCCAGCACAATAGTC
ATCGCCGCCCTGGCACTCAAAATTGGACTAGCACCTATACACTTCTGAATACC
TGAAGTTTTACAAGGGCTAGATCTTTTAACCGGACTAATTTTATCCACCTGAC
AAAAACTTGCTCCACTTGCTCTCATTATTCAAACAGCCCAAGCTATCGACCCC
CTACTGCTGACGGCCCTGGGCCTATTGTCCACACTAATCGGTGGTTGAGGAG
GGCTAAACCAGACCCAGCTCCGGAAAATTTTAGCCCTTGCAATTACCTGATC
GAAAAGCCCTGTCTTGACAGCAACCACTGCCTTAGTATTACTGTCATTGGGCG
GTTTGCCTCCTCTTACAGGGTTTATGCCAAAATGACTTATTTTACAAGAGTTG
GCAAAACAGAGCCTTCCCCTTGCCGCTACAGTTATGGCTCTCGCAGCCCTCCT
CAGCCTGTACTTCTACTTACGACTCTGCTATGCAATGACACTCACCATTTCTC

CTAACACTGTTAATTCAGCCACGCCCTGACGGGTTCAAACAACTCAAGCCTCT
CTCCCCGTGGCCCTAGCTACTACGCTAGCACTGGGCCTTCTTCCCATAACTCC
AGCCATTGTAATGCTAGTCACCTAGGGGCTTAGGATAACATTAGACCAAGA

TPI (328bp)

A
GAGAGGGAGGCTGGAATCACGGAGAAAGTTGTTTTTGCACAGACCAAGTTCA
TTGCAGGTATTTTTCAATTTAGTGGTGTATATA--TTATAAAACATTCCTTGTAT
TTAAAATAAAACAATCATCTCTTTTAGATAATGTGAAGGATTGGAGCAAAGT
GGTCCTTGCTTATGAGCCTGTTTGGGCCATTGGCACTGGGAAAACTGCATCAC
CCCAACAGGTACAGACGAGACATTTGGATTAGAACATGGGGGTTTTTCAAAC
ATGTTCTCTGGTTTTATTGAACTCTGTGACGAGTCTCTCCCTGCTGTTCTTACA
GGCTCAGGAG

B
GAGAGGGAGGCTGGAATCACGGAGAAAGTTGTTTTTGCACAGACCAAGTTCA
TTGCAGGTATTTTTCAATTTAGTGGTGTATATA--TTATAAAACATTCCTTGTAT
TTAAAATAAAACAATCATCTCTTTTAGATAATGTGAAGGATTGGAGCAAAGT
GGTCCTTGCTTATGAGCCTGTTTGGGCCATTGGCACTGGGAAAACTGCATCAC
CCCAACAGGTACAGACGAGACATTTGGATTAGAACATGGGGGGTATTCAAAC
ATGTTCTCTGGTTTTATTGAACTCTGTGACGAGTCTCTCCCTGCTGTTCTTACA
GGCTCAGGAG

C
GAGAGGGAGGCTGGAATCACGGAGAAAGTTGTTTTTGCACAGACCAAGTTCA
TTGCAGGTATTTTTCAATTTAGTGGTGTATATA--TTATAAAACATTCCTTGTAT
TTAAAATAAAACAATCATCTCTTTTAGATAATGTGAAGGATTGGAGCAAAGT
GGTCCTTGCTTATGAGCCTGTTTGGGCCATTGGCACTGGGAAAACTGCATCAC
CCCAACAGGTACAGACGAGACATTTGGATTAGAACATGGGGGTTTTTCAAAC
ATGTTCTCTGGTTTTATTGAACTCTGTGATGAGTCTCTCCCTGCTGTTCTTACA
GGCTCAGGAG

D
NNNNNNNNNNNNNNNNNNNNNNNNNNNTTGTTTTTGCACAGACCAAGTTC
ATTGCAGGTATTTTTCAATTTAGTGGTGTATATA--TTATAAAA CATTCCTTGT
ATTTAAAATAAAACAATCATCTCTTTTAGATAATGTGAAGGATTGGAGCAAA
GTGGTCCTTGCTTATGAGCCTGTTTGGGCCATTGGCACTGGGAAAACTGCATC
ACCCCAACAGGTACAGACGAGACATTTGGATTAGAACATGGGGGGTATTCAA
ACATGTTCTCTGGTTTTATTGAACTCTGTGATGAGTCTCTCCCTGCTGTTCTTA
CAGGCTCAGGAG

E
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNTGTTTTTGCACAGACCAAGTTC
ATTGCAGGTATTTTTCAATTTAGTGGTGTGTATATATTATAAAACATTCCTTGT
ATTTAAAATAAAACAATCATCTCTTTTAGATAATGTGAAGGATTGGAGCAAA
GTGGTCCTTGCTTATGAGCCTGTTTGGGCCATTGGCAGTGGGAAAACTGCATC
ACCCCAACAGGTACAGACGAGACATTTGGATTAGAACATGGGGGTTTTTCAA
ACATGTTCTCTGGTTTTATTGAACTCTGTGACGAGTCTCTCCCTGCTGTTCGTA
CAGGCTCAGGAG

S7 (358bp)

A
GTTGAAAGAG GTGAGTCTAC TGAATGTGCC CAGCTATAGG GATATCAAGT
TAAAATGTAG AAAATAAGTA ATATTTACCT CCACGCATGA GCTTCATATG
ATATTTTCAC TGAGATGCGT TAGATGCGAG TATTTAAGAA AGTGGATGGT
TATTTTAGAA ACAGGAATTT ATAAATTTAA GATGTGCTAG CAGACGTGTC
TGTACTACAG TACTACTGCG GCCCATGTGG TGTTCTAATA TGCCTGAAAA
TGCCTCTATT AAGTAAAGTA CTATATCTCG AAGACTTATA TTTGTGATAG
ATTAAAACTT GTAATGTAAA TGTAGTGATT CTGTGCTAGC TAATCAATGC
TAACGGCA

B
GTTGAAAGAG GTGAGTCTAC TGAATGTGCC CAGCTATAGG GATATCAAGT
TAAAATGTAG AAAATAAGTA ATATTTACCT CCACGCATGA GCTTCATATG
ATATTTTCAC TGAGATGCGT TAGATGCGAG TATTTAAGAA AGTGGATGGT
TATTTTAGAA ACAGGAATTT ATAAATTTAA GATGTGCTAG CAGACGTGTC
TGTACTACAG TACTACTGCG GCCCATGTGG TGTTCTAATA TGCCTGAAAA
TGCCTCTATT AAGTAAAGTA CTATATCTCG GAGACTTATA TTTGTGATAG
ATTAAAACTT GTAATGTAAA TGTAGTGATT CTGTGCTAGC TAATCAATGC
TAACGGCA

C
GTTGAAAGAG GTGAGTCTAC TGAATGTGCC CAGCTATAGG GATATCAAGT
TAAAATGTAG AAAATAAGTA ATATTTACCT CCACGCATGA GCTTCATATA
ATATTTTCAC TGAGATGCGT TAGATGCGAG TATTTAAGAA AGTGGATGGT
TATTTTAGAA ACAGGAATTT ATAAATTTAA GATGTGCTAG CAGACGTGTC
TGTACTACAG TGCTACTGCG GCCCATGTGG TGTTCTAATA TGCCTGAAAA
TGCCTCTATT AAGTAAAGTA CTATATCTCG AAGACTTATA TTTGTGATAG
ATTAAAACTT GTAATGTAAA TGTAGTGATT CTGTGCTAGC TAATCAATGC
TAACGGCA

D
GTTGAAAGAG GTGAGTCTAC TGAATGTGCC CAGCTATAGG GATATCAAGT
TAAAATGTAG AAAATAAGTA ATATTTACCT CCACGCATGA GCTTCATATT

ATATTTTTAC TGAGATGCGT TAGATGCGAG TATTTAAGAA AGTGGATGGT
TATTTTAGAA ACAGGAATTT ATAAATTTAA GATGTGCTAG CAGACGTGTC
TGTACTACAG TACTACTGCG GCCCATGTGG TGTTCTAATA TGCCTGAAAA
TGCCTCTATT AAGTAAAGTA CTATATCTCG GAGACTTATA TTTGTGATAG
ATTAAAACTT GTAATGTAAA TGTAGTGATT CTGTGCTAGC TAATCAATGC
TAACGGCA

E
GTTGAAAGAG GTGAGTCTAC TGAATGTGCC CAGCTATAGG GATATCAAGT
TAAAATGTAG AAAATAAGTA ATATTTACCT CCACGCATGA GCTTCATATG
ATATTTTCAC TGAGATGCGT TAGATGCGAG TATTTAAGAA AGTGGATGGT
TATTTTAGAA ACAGGAATTT ATAAATTTAA GATGTGCTAG CAGACGTGTC
T--ACTACAG TACTACTGCG GCCCATGTGG TGTTCTAATA TGCCTGAAAA
TGCCTCTATT AAGTAAAGTA CTATATCTCG AAGACTTATA TTTGTGATAG
ATTAAAACTT GTAATGTAAA TGTAGTGATT CTGTGCTAGC TAATCAATGC
TAACGGCA

F
GTTGAAAGAG GTGAGTCTAC TGAATGTGCC CAGCTATAGG GATATCAAGT
TAAAATGTAG AAAATAAGTA ATATTTACCT CCACGCATGA GCTTCATATG
ATATTTTCAC TGAGATGCGT TAGATGCGAG TATTTAAGAA AGTGGATGGT
TATTTTAGAA ACAGGAATTT ATAAATTTAA GATGTGCTAG CAGACGTGTC
TGTACTACAG TACTACTGCG GCCCATGTGG TGTTCTAATA TGCCTGAAAA
TGCCTCTATT AAGTAAAGCA CTATATCTCG AAGACTTATA TTTGTGATAG
ATTAAAACTT GTAATGTAAA TGTAGTGATT CTGTGCTAGC TAATCAATGC
TAACGGCA

G
GTTGAAAGAG GTGAGTCTAC TGAATGTGCC CAGCTATAGG GATATCAAGT
TAAAATGTAG AAAATAAGTA ATATTTACCT CCACGCATGA GCTTCATATG
ATATTTTCAC TGAGATGCGT TAGATGCGAG TATTTAAGAA AGTGGATGGT
TATTTTAGAA ACAGGAATTT ATAAATTTAA GATGTGCTAG CAGACGTGTC
TGTACTACAG TACTACTGCG GCCCATGTGG TGTTCTAATA TGCCTGAAAA
TGCCTCTATT AAGTAAAGCA CTACATCTCG AAGACTTATA TTTGTGATAG
ATTAAAACTT GTAATGTAAA TGTAGTGATT CTGTGCTAGC TAATCAATGC
TAACGGCA

H
GTTGAAAGAG GTGAGTCTAC TGAATGTGCC CAGCTATAGG GATATCAAGT
TAAAATGTAG AAAATAAGTA ATATTTACCT CCACGCATGA GCTTCATATG
ATATTTTCAC TGAGATGCGT TAGATGCGAG TATTTAAGAA AGTGGATGGT
TATTTTAGAA ACAGGAATTT ATAAATTTAA GATGTGCTAG CAGACGTGTC
TGTACTACAG TACTACTGCG GCCCATGCGG TGTTCTAATA TGCCTGAAAA
TGCCTCTATT AAGTAAAGTA CTATATCTCG AAGACTTATA TTTGTGATAG

ATTAAAACTT GTAATGTAAA TGTAGTGATT CTGTGCTAGC TAATCAATGC
TAACGGCA

I
GTTGAAAGAG GTGAGTCTAC TGAATGTGCC CAGCTATAGG GATATCAAGT
TAAAATGTAG AAAATAAGTA ATATTTACCT TTACGCATGA GCTTCATATG
ATATTTTCAC TGAGATGCGT TAGATGCGAG TATTTAAGAA AGTGGATGGT
TATTTTAGAA ACAGGAATTT ATAAATTTAA GATGTGCTAG CAGACGTGTC
TGTACTACAG TACTACTGCG GCCCATGTGG TGTTCTAATA TGCCTGAAAA
TGCCTCTATT AAGTAAAGTA CTATATCTCG AAGACTTATA TTTGTGATAG
ATTAAAACTT GTAATGTAAA TGTAGTGATT CTGTGCTAGC TAATCAATGC
TAACGGCA

J
GTTGAAAGAG GTGAGTCTAC TGAATGTGCC CAGCTATAGG GATATCAAGT
TAAAATGTAG AAAATAAGTA ATATTTACCT CCACGCATGA GCTTCATATA
ATATTTTCAC TGAGATGCGT TAGATGCGAG TATTTAAGAA AGTGGATGGT
TATTTTAGAA ACAGGAATTT ATAAATTTAA GATGTGCTAG CAGACGTGTC
TGTACTACAG TACTACTGCG GCCCATGTGG TGTTCTAATA TGCCTGAAAA
TGCCTCTATT AAGTAAAGTA CTATATCTCG AAGACTTATA TTTGTGATAG
ATTAAAACTT GTAATGTAAA TGTAGTGATT CTGTGCTAGC TAATCAATGC
TAACGGCA

K
GTTGAAAGAG GTGAGTCTAC TGAATGTGCC CAGCTATAGG GATATCAAGT
TAAAATGTAG AAAATAAGTA ATATTTACCT CCACGCATGA GCTTCATATT
ATATTTTTAC TGAGATGCGT TAGATGCGAG TATTTAAGAA AGTGGATGGT
TATTTTAGAA ACAGGAATTT ATAAATTTAA GATGTGCTAG CAGACGTGTC
TGTACTACAG TACTACTGCG GCCCATGTGG TGTTCTAATA TGCCTGAAAA
TGCCTCTATT AAGTAAAGTA CTATATCTCG AAGACTTATA TTTGTGATAG
ATTAAAACTT GTAATGTAAA TGTAGTGATT CTGTGCTAGC TAATCAATGC
TAACGGCA

L
GTTGAAAGAG GTGAGTCTAC TGAATGTGCC CAGCTATAGG GATATCAAGT
TAAAATGTAG AAAATAAGTA ATATTTACCT CCACGCGTGA GCTTCATATG
ATATTTTCAC TGAGATGCGT TAGATGCGAG TATTTAAGAA AGTGGATGGT
TATTTTAGAA ACAGGAATTT ATAAATTTAA GATGTGCTAG CAGACGTGTC
TGTACTACAG TACTACTGCG GCCCATGTGG TGTTCTAATA TGCCTGAAAA
TGCCTCTATT AAGTAAAGTA CTATATCTCG AAGACTTATA TTTGTGATAG
ATTAAAACTT GTAATGTAAA TGTAGTGATT CTGTGCTAGC TAATCAATGC
TAACGGCA

M
GTTGAAAGAG GTGAGTCTAC TGAATGTGCC CAGCTATAGG GATATCAAGT
TAAAATGTAG AAAATAAGTA ATATTTACCT CCACGCATGA GCTTCATATG
ATATTTTCAC TGAGATGCGT TAGATGCGAG TATTTAAGAA AGTGGATGGT
TATTTTAGAA ACAGGAATTT ATAAATTTAA GATGTGCTAG CAGACGTGTC
TGTACTAC-- ------TGCG GCCCATGTGG TGTTCTAATA TGCCTGAAAA TGCC
TCTATT AAGTAAAGTA CTATATCTCG AAGACTTATA TTTGTGATAG ATTA
AAACTT GTAATGTAAA TGTAGTGATT CTGTGCTAGC TAATCAATGC TAAC
GGCA