6-2017

MOLECULAR EVOLUTION AND PHYLOGEOGRAPHY OF MITOCHONDRIAL DNA CYTOCHROME B GENE IN SOUTHERN CALIFORNIA SANTA ANA SPECKLED DACE (Rhinichthys osculus)

Pia Marie VanMeter
California State University - San Bernardino, pvbionut@gmail.com

Recommended Citation

VanMeter, Pia Marie, "MOLECULAR EVOLUTION AND PHYLOGEOGRAPHY OF MITOCHONDRIAL DNA CYTOCHROME B GENE IN SOUTHERN CALIFORNIA SANTA ANA SPECKLED DACE (Rhinichthys osculus)" (2017). *Electronic Theses, Projects, and Dissertations*. 561.
https://scholarworks.lib.csusb.edu/etd/561

MOLECULAR EVOLUTION AND PHYLOGEOGRAPHY OF MITOCHONDRIAL DNA CYTOCHROME B GENE IN SOUTHERN CALIFORNIA SANTA ANA SPECKLED DACE (*Rhinichthys osculus*)

A Thesis

Presented to the

Faculty of

California State University,

San Bernardino

In Partial Fulfillment

of the Requirements for the Degree

Master of Science

in

Biology

---

by

Pia Marie VanMeter

June 2017

MOLECULAR EVOLUTION AND PHYLOGEOGRAPHY OF MITOCHONDRIAL DNA CYTOCHROME B GENE IN SOUTHERN CALIFORNIA SANTA ANA SPECKLED DACE (*Rhinichthys osculus*)

A Thesis
Presented to the
Faculty of
California State University,
San Bernardino

by
Pia Marie VanMeter
June 2017

Approved by:

Anthony Metcalf, Ph. D., Committee Chair, Biology   Date

James Ferrari, Ph. D., Committee Member

David Polcyn, Ph. D., Committee Member

ABSTRACT

In this study, I sequenced the mitochondrial cytochrome b gene to elucidate the extent and pattern of genetic variations among and within populations of *Rhinichthys osculus* (Santa Ana Speckled Dace) found in the different watersheds in Southern California, Central California Coast and Eastern California Desert. I described and analyzed the structural characteristics and pattern of base sequence substitutions in the cytochrome b gene to understand the molecular evolution of the gene. The Phylogenetic analyses showed that the Southern California Santa Ana Speckled dace is a distinct population from the Central California Coast dace population and Eastern California Desert dace population, and is more closely related to the Colorado River speckled dace population. There is a high degree of genetic variation among all populations including a significant genetic structure associated with watersheds, mountain ranges, and geographic grouping based on locations. The SWISS-Model automated protein structure homology elucidated the conserved and invariant residues within the cytochrome b gene where the amino acid substitutions are located in the trans-membrane of the protein sequence. The implication for conservation and management of the Southern California Santa Ana Speckled dace is high because of habitat lost for this distinct dace population. The data from this study will contribute to preserving the genetic variability of the Santa Ana Speckled Dace as a separate taxa and species, as well as to help maintain intact the population in the different Southern California creeks.

ACKNOWLEDGEMENTS

To the love of my life, James Jay VanMeter who is the “Wind Beneath My Wings.” You gave me the strength to pursue my dreams and always championed my ability even when I doubted myself. Thank you for all your love, support, and friendship. You are truly my soul mate and my best friend. I will always be grateful that you decided to pursue graduate school at the same time. You made our late nights and CSUSB weekends one of my most cherished moments. We are stronger together because of our experience.

I would like to thank my advisor, Tony Metcalf who took a chance in taking a graduate student who worked as a full time biology teacher. I know that my students continue to benefit from what I learned in graduate school. I also would like to thank my committee members Jim Ferrari and David Polcyn for your patience with the slow progress of this thesis study. It is finally completed! The three of you have always been my inspiration since my undergraduate days at CSUSB. I always aspire to motivate and inspire in my high school students the love for science the way that you inspired me.

I dedicate this work to my parents Crispin L. Oqueriza Jr. and Yolanda M. Oqueriza who always instilled in me the importance of education and hard work. I know that the last six months I was persona non grata. Now, I can finally call home every weekend.

And finally I thank my family and friends for all of their love and support during my late nights and curtailed weekend visits.

TABLE OF CONTENTS

CHAPTER TWO: MATERIALS AND METHODS

CHAPTER ONE

INTRODUCTION

There are 250 native species of cyprinidae and the most widespread species is known locally as the Speckled Dace, *Rhinichthys osculus,* a Western North American cyprinid that has been described extensively (Jordan & Evermann, 1896; Hubbs and Miller, 1948; Cornelius, 1969; Hubbs et al., 1974; Woodman, 1992; Moyle, 2002; Pfrender et al., 2004; Oakey et al., 2004). *R. osculus* is native to all major western drainages from the Columbia Basin, Canada in the North to Sonora Mexico in the south (Lee et al., 1980). *Rhinichthys osculus* populations are highly endemic and show large differences in morphological traits (Scott and Crossman, 1998). Because of the differences in morphological traits, the species *R. osculus* is considered to be a widely distributed species with multiple subspecies (Hubbs et al., 1974), rather than the 12 different species as originally described (Jordan and Evermann, 1896). There is a limited amount of information regarding the understanding of the relationship among the different speckled dace populations. Researchers investigating each population of the species from different regions of North America treat *R. osculus* differently depending on what region the fish is found. For example, Deacon and Williams (1984) list 10 distinct subspecies and five undescribed *R. osculus* taxa from Nevada State, while in the state of Oregon *R. osculus* is treated as a habitat generalist because of its abundance (Bond et al., 1988; Pfrender, 2004). To more fully understand the distribution of Speckled Dace, I will review the systematics and the geographic distribution defining these taxa.

Distribution of Cypriniformes

The cypriniformes are an order of ray-finned fish that contains six families, with 321 genera and approximately 3,268 species (Nelson, 1994). The cyprinid family, which includes the minnows and carp, is one of the largest families of fish within this order and is distributed across Africa, Eurasia, and North America (Banarescu and Coad, 1991). Three families of the cypriniformes family comprise the native primary freshwater fishes in California, the Cyprinidae (minnows), Catostomidae (suckers) and Salmonidae (trout). The largest family of freshwater fish found in California is the Cyprinidae and they are used to reconstruct the biogeographic history of the different regions on the western North America (Hubbs and Miller, 1948; Minckley et al., 1986; Moyle, 2002). Researchers employ cyprinids in evolutionary studies because of their wide distribution and high genus and/or species diversity (Nelson, 1994). Researchers have used the geological history of the region, the current distribution of the cyprinidae species, and extensive fossil records documenting their historical distribution to map the evolutionary and phylogeographic pattern of the species around the world (Pfrender, 2004).

There are three zoogeographic provinces in California (Hickman, J.C., 1993) including the California Floristic province (CA-FP), the Great Basin province, and the Desert province. The Southwestern California region of the CA-FP consists of four sub regions and smaller districts from Baja, California north to Point Conception (Figure 1, Appendix A) (Moyle, 2002). The major watersheds included in the Southwestern California region are the San Gabriel, the Santa Ana, and Los Angeles rivers and their tributaries (Figure 1, Appendix A) (SAWPA, 2005). The only streams that have endemic groups of fresh water fishes such as the Santa Ana sucker (family Catostomidae), arroyo chub, and speckled dace (both family Cyprinidae), are found in the Los Angeles basin which consists of the Santa Ana, San Gabriel and Los Angeles Rivers (Moyle, 2002). The origins of the endemic fresh water fish groups have puzzled zoogeographers because their closest relatives known are in the Colorado River drainage for the Santa Ana Sucker and speckled dace, and in Mexico for the arroyo chub (Cornelius, 1969; Minckley et al., 1986; Moyle, 2002). Minckley (1986) argued that ancestors of the arroyo chub rode into the South Coast region on a shifting continental plate that split from the continent farther south during the Pleistocene period due to tectonic activity, and are descended from related fishes of the Mexican plateau. Both the speckled dace and catostomid suckers such as the Santa Ana Sucker are fish that are capable of living in small, swift streams and have distribution patterns throughout the West that suggest dispersal through streams, most likely entering the region by the way of stream connections to the ancient Colorado River drainage (Miller, 1946; Cornelius, 1969; Spencer & Dowling, 2008).

Rhinichthys osculus (Speckled Dace) Systematics

The speckled dace *R. osculus* is a small, highly variable species identifiable by a thick caudal peduncle, a small sub-terminal mouth, a pointed snout, and small scales (Figure 2, Appendix A). Speckled dace have elongated bodies and large fins. The origin of their dorsal fin is well behind their pelvic fins with 6-8 rays, while their anal fin normally has 7 rays. Their pharyngeal teeth are strongly hooked with a slight grinding surface and they have a tiny barbell at the end of each maxilla and a small frenum that attaches the snout to the middle of the upper lip (Cornelius, 1969; Moyle, 2002). Their color is highly variable, but most fish over 3 cm in length have a darkly speckled pattern on their back and sides. The dark speckles on the sides often overlap together to resemble a dark lateral band. Speckled dace also possess a spot at the base of the caudal peduncle, with a stripe that runs on the head through the snout. Dace colors range from dusky yellow to dark olive with a belly yellowish to whitish in color (Cornelius, 1969). The speckled dace have adapted a downward facing mouth in order to feed on small invertebrates such as snails and insect larvae living on the stream bottom (Figure 2, Appendix A) (Moyle, 2002).

The genus *Rhinichthys* is found in almost every drainage of Western North America, but with only eight recognized species in abundance (Moyle, 2002). The abundant distribution of the speckled dace on both sides of the Sierra Nevada and the Transverse Ranges in California may be attributed to headwater stream capture. Headwater stream fish capture occurs when a graded stream erode backward towards two separate divided drainage system to form one drainage system (Jordan, 1896), or by separation of population because of a geographical barrier to gene flow such as tectonic change in elevation or basalt flow, or development of falls, or death of intervening habitat (Hubbs and Miller, 1948). Another explanation for the abundant distribution of the speckled dace is lake spillover where water is released with a population of fish included and a barrier function is reestablished (Hubbs and Miller, 1948). Finally geographical separation through isolation by distance (Wright, 1931) may play a role in speckled dace biogeographical patterns. The wide distribution of dace shows their ability to colonize new areas through headwater capture (Moyle, 2002) and their ability to adapt to new environments. Their adaptability is reflected in the variability of their body shapes (Hubbs et al., 1974). Small and chunky forms are observed to live in springs and slow-moving streams while large, streamlined forms live in fast-moving turbulent streams where the bottom is rocky. The distinctive morphology is not fixed genetically, but is capable of changing with the developmental environment (Jordan and Evermann, 1896; Hubbs and Miller, 1948; Cornelius, 1969; Hubbs et al., 1974; Woodman 1992; Pfrender et al., 2004; Oakey et al., 2004). Modern molecular and morphometric techniques need to be used to resolve the relationships among the different forms of speckled dace as well as their specific regions (Sada, 1989).

Morphological analyses conducted by Cornelius (1969) on *Rhinichthys osculus* found in Southern California, known as the Santa Ana speckled dace, although limited by small sample size, led him to conclude that the Santa Ana Speckled Dace warranted subspecies status. Cornelius also suggested that the Santa Ana Speckled Dace is more closely related to dace populations found in the Colorado River drainage rather than Northern California populations.

The Santa Ana Speckled Dace is also of particular importance because it was reported completely extirpated from the Los Angeles River drainage system (SAWPA, 1995). The number of speckled dace in other watersheds has declined rapidly due to habitat fragmentation. The primary cause of dace habitat reduction is due to human activity including introduction of non-native species, water diversions, urbanization of watersheds, and other human-caused factors (Swift et al., 1993). In 1995 the California Department of Fish and Game listed the Santa Ana speckled dace as a species of special concern, and in 1998 the U.S. Forest Service listed the Santa Ana speckled dace as a sensitive species. (Santa Ana Speckled Dace Recovery Project, 2005). This proposal is part of a larger study being conducted regarding the speckled dace and its phylogeographic distribution and evolutionary lineage in Southern California.

Formation of the Geographic Features of Southern California

The Southern California region consists of diverse topographic terrain. Southern California mountain ranges include the San Bernardino Mountains, the San Gabriel Mountains, The Santa Susana Mountains and the Santa Monica Mountains (Hickman, 1993). The Cajon Pass separates the San Gabriel and San

Bernardino Mountains. The San Fernando Pass separates the San Gabriel and the Santa Susana mountains (Schoenherr, 1992). The Southern California region contains over half of the human population of the state.

The Transverse and Peninsular ranges are both fault-block ranges (Figure 3, Appendix A). The Transverse Ranges lie on the “big bend” of the San Andreas Fault and the mountains of this region of Southern California have been rotated about 90° clockwise due to the northward motion of the Pacific plate along the San Andreas Fault (Seeber et al., 2000). The Peninsular Range includes mountains in the Los Angeles Basin, and extends from the northernmost ranges of the series of mountains that make up the Baja California peninsula in Mexico while the San Jacinto Mountains, the Santa Rosa Mountains, and the Laguna Mountains which run from north to south (Kimbrough et al., 2001). The Los Angeles Basin is a flood plain and surrounded by the previously described mountains, crossed by a series of north-south trending faults. The Los Angeles, San Gabriel, and Santa Ana rivers have deposited alluvial outwash from the mountains to the coast for millions of years (Warrick et al, 2003).

The climate of Southern California is due to a combination of its latitude and its proximity to the cold waters of the Pacific Ocean. This gives the area a combination of maritime and Mediterranean climates. (Schoenherr, 1992).

Southern California Watersheds

A watershed is a region where water that travels both underground and on the surface moves through a network of drainage pathways that include streams, tributaries and rivers (SAWPA, 2004). Watersheds follow major ridgelines and can normally be discerned using topographic maps (Figure 1, Appendix A).

The Santa Ana River watershed is the largest watershed in Southern California. It is about 2800 square miles and located geographically between Los Angeles and San Diego (SAWPA, 2004). The Santa Ana River watershed is a group of interconnected inland basins and coastal basins that flow southwestward to the Pacific Ocean (Mitchell, 2006). The east-west crest alignment of the San Gabriel and San Bernardino Mountains separates the Santa Ana River basin from the Mojave Desert. In the south, the watershed boundary divides the Santa Margarita river drainage area from the San Jacinto River which terminates in Lake Elsinore and feeds into the Santa Ana River via Temescal wash (Mitchell, 2006). The tributaries of the Santa Ana River Watershed upstream of Seven Oaks Dam include Mill Creek, City Creek, Plunge Creek, Twin/Strawberry Creek, Cajon Creek, and Lytle Creek. The Santa Ana River proper extends to the U.S. Army Corps of Engineers' Santa Ana River Project located at Anaheim Fore bay recharge area (Mitchell, 2006).

The Los Angeles River Watershed covers 834 square miles, extending from the eastern portion of the Santa Monica Mountains to the San Gabriel Mountains in the west. The watershed is shaped by the Los Angeles River which flows from its headwaters in the mountains eastward to the northern corner of Griffith Park, then flows across the coastal plain of the Los Angeles Basin into San Pedro Bay near Long Beach (Los Angeles County Watershed Management 2008). There are eight major L.A. River tributaries which include the Burbank Western Channel, Los Angeles River drainage, the Pacoima Wash, the Tujunga Wash, the Vertugo Wash, the Arroyo Seco, Compton Creek, and the Rio Hondo (Los Angeles County Watershed Management, 2008).

The San Gabriel River watershed is located in Los Angeles County. It is the largest watershed in the San Gabriel mountain drainage system with a total area of 648 square miles and a length of 75 miles that begins northeast of Los Angeles in the Los Angeles National Forest (San Gabriel Mountains Regional Conservancy, 2005). The San Gabriel River watershed consists of three major upper forks (North, West, and East). The uppermost portion of the San Gabriel watershed located in the San Gabriel Mountains and extending to the Morris Dam is the least altered habitat within the watershed. Above Morris dam, the area is mountainous and dissected by deep narrow ravines that contain creek and stream tributaries to the San Gabriel River (San Gabriel Mountains Regional Conservancy, 2005).

Phylogeography

The two types of general approaches used in the phylogeographic analysis of molecular data are phylogeny estimation by method of genetic distance, and estimation by method of character state. Both methods seek to infer the true phylogeny of a particular taxa using gene trees, which are constructed according to specific criteria.

Distance-based phylogenetic methods typically use genetic distance data matrices to construct phylogenetic trees by either the minimum evolution or neighbor-joining methods. In the minimum evolution approach, multiple gene trees are constructed based on the genetic distance between related individuals, and the tree(s) with the overall shortest total distance between all individuals is selected as the most probable gene tree representing the phylogeny of the taxa (Hillis et al, 1996). In the neighbor-joining approach, haplotype pairs separated by the smallest corrected genetic distance are joined to form a tree branch, and subsequent joining of additional haplotypes by the criteria of minimum distance result in construction of a neighbor-joining (NJ) tree (Hall, 2004).

Character state-based phylogenetic methods assign a character state at each nucleotide position, or character, of a particular DNA sequence for each individual being compared. The four character states possible consist of the four nucleotides adenine (A), guanine (G), cytosine (C), and thymine (T). The numbers of observed character-state differences between individuals are then used to infer the genetic relationship among the taxa being compared and construct a phylogenetic tree (or trees). There are three commonly used character-state based approaches to phylogenetic estimation. The method of maximum parsimony (MP) searches for the tree (or trees) containing the lowest overall number of character state changes, or evolutionary steps (Swofford et al, 1996). The method of maximum likelihood (ML) searches for the tree maximizing the probability of observing the genetic data given a particular model of evolution (Hillis et al, 1996). The evolutionary model used in ML analysis, as well as in other approaches, is selected based on the observed characteristics of the DNA sequence data used in the analysis (Posada and Crandall, 1998). Bayesian analysis uses empirical data of the molecular variation to develop an a posteriori hypothesis (Hall, 2004), wherein the empirical data will suggest a type of model of evolution using Modeltest 3.7 (Posada and Crandall 1998) and the probability that the data will create a particular phylogenetic tree. The Bayesian approach to analysis of phylogenies will use the information to create a tree based on a specific evolutionary model and empirical data.

Phylogeography of Freshwater Fish

Phylogeography is the field of study concerned with geographical distribution of organisms and their genealogical lineages specifically at the intraspecific level (Avise, 2000). In the early years up to the present time molecular phylogenetic studies were conducted using allele-based differences in the net charges of proteins, such as allozyme assays (Kocher et al., 1989). Mitochondrial DNA (mtDNA) is widely used in phylogeographic studies (Avise, 2000) because the mtDNA genome possesses the properties of maternal or uniparental transmission (Dawid et al., 1972), intraspecific variation that can be undetectable in allozyme assays (Moritz et al., 1992), and absence of genetic recombination (Hayashi et al., 1985). Mitochondrial DNA cytochrome *b* gene and the control region are used population genetics and phylogenetic studies because both have evolutionary rates that can be analyze to resolve population structure. Protein-coding cytochrome *b* gene evolves slower than the non-coding control region, when genetic distances between taxa are very close. When taxa are separated by less than 5% genetic distance cytochrome b evolves faster than the control region (Tang et. al, 2006). The field of phylogeography studies the evolutionary genetics of species-level relationships (phylogenetics) and the genetic relationships among individuals (genealogy) (Avise, 2000). Phylogeography attempts to explain how animal and plant distributions are influenced by historical events. Another aspect of phylogeography is landscape genetics, which is concerned with the emergence of recent variation and differentiation over short geographical distances, and has valuable application in conservation biology (Avise, 1998). Freshwater fish are ideal organisms to study biogeographic patterns in landscape genetics because they are restricted to particular lakes and river drainages. The evolutionary history of a fresh water fish species population is highly linked to the geological history of the landscape they occupy (Bernatchez and Wilson, 1998). An example of this is the variation found in the endangered cyprinid fish *Anaecypris hispanica* (Alves et al., 2001). Phylogenetic analysis using the mitochondrial DNA cytochrome b gene as a molecular marker was used to reveal the presence of geographical population structure. The data obtained in the study inferred limited or no gene flow between populations, suggesting the need for revision of conservation strategies for *A. hispanica*. Morphological studies of cyprinid fish have failed to reach a consensus regarding the number of monophyletic subfamilies (Cunha et al, 2002) and created disagreements over their taxonomic alignment. Cunha et al. (2002) tried to elucidate the phylogenetic relationships of Eurasian and American cyprinids using cytochrome b sequences and they found that the sequence data were not consistent with the previously proposed relationships, especially with the Asian taxa. Further studies are needed to improve the understanding of the historical and taxonomic relationship of the cyprinid families as well as elucidating the relationships of Asian and African cyprinid taxa (Cunha et al., 2002).

The coalescent theory is a conceptual framework within which plylogeographic data is analyzed. Coalescent theory states that any two non-recombining haplotypes sampled from a population will 'coalesce' at their most recent common ancestor (Griffiths and Tavare, 1997; Herbots, 1997; Taib, 1997). It interprets the distribution of times to common ancestry in a gene tree accounting for the effects of the various evolutionary forces including genetic drift, migration, population size changes, and selection (Zardoya et al., 1998). Haplotypes and their distributions in single populations can provide evidence of historical changes in population size, and of past separations of clades that have subsequently come into secondary contact (Zardoya et al., 1998). In a study focusing on the European cyprinid *Leuciscus leuciscus*, Costedoat et al. (2006)

used cytochrome b and other mtDNA loci to investigate historical cyprinid population dynamics including expansion, migration, bottlenecks, and divergence over time. The authors concluded that the relationships between lineages showed evidence of complex biogeographical patterns due to Pliocene and Pleistocene paleoclimatic events, based on the observed divergence of mtDNA nucleotide sequences.

Phylogeography of Cyprinidae

Geographically separated populations may arise from the splitting of a previously continuous population (vicariance), or through creation of a new population by dispersal (Avise, 2000). For example, a Greek study of the chub species *Leuciscus cephalus* found the difference in the distributional range between both sides of the Balkan Peninsula directly linked to the differences in the dispersion rate and extension of their distribution range (Zardoya et al., 1999). The river confluence due to sea level lowering and river capture in Western Greece was seen to be the main contributor to the observed chub dispersion during the middle Miocene period, resulting in genetically distinct populations (Zardoya et al., 1999).

Genetic differentiation among populations is promoted by drift and limited by gene flow. Perdices et al. (2005) analyzed the phylogenetic structure of twenty-eight populations of the Asian cyprinid *Opsariichthys bidens* across its range in China. Five mtDNA cytochrome *b* lineages were detected overall within the Yangtze, Pearl, and Hai He drainages. The authors detected cytochrome *b* sequence divergence as large as twenty percent between lineages, suggesting taxonomic status based on morphology underestimated the differences among populations. Drainage-restricted haplotypes with high frequency and moderate nucleotide diversity were inferred by the authors as evidence of historic long-term interruption of gene flow leading to genetic drift, although the observed distribution of some common haplotypes suggest a recent connection allowing gene flow across some drainages.

Phylogeography can help resolve questions regarding cryptic species as well, and demonstrates the importance of sampling across entire biogeographical ranges in making taxonomic decisions (Ketmaier et al., 2004). The Ketmaier et al. study showed that distinction among clades using the mtDNA cytochrome b gene did not support previous taxonomic assumptions based on morphology. The three genera *Telestes, Leuciscus, and Phoxinellus* were found not to have a monophyletic origin, and some *Leuciscus* and *Phoxinellus* taxa were discovered that should be included in the genus *Telestes* based on molecular comparison*.* In a related study (Ketmaier et al., 2004), cytochrome b data was used to study two lineages of *Leuciscinae* cyprinids in Iceland and taxonomic assumptions based on morphology were analyzed, with clades showing discordance. The authors used nested clade analysis (Templeton, 1980) to retrace the radiations of the two lineages and the time period over which it occurred during the Miocene period (Ketmaier et al., 2004).

Phylogeographic analysis has proven useful in investigations of whether populations are native to an area or have recently been introduced, and to determine the source area of the introduction (Mesquita et al., 2001). In a study of *Squalius aradensis*, the researchers were able to determine that the Arade drainage in southern Portugal represented the evolutionary center of the species. This type of information has importance in conservation biology, because in revealing the place of origin for the particular species being studied it provides a point of reference regarding what type of conservation will be most effective for that particular organism (Mesquita et al. 2001).

Phylogeography of Speckled Dace

The diversification of *Rhinichthys osculus* in western North America indicates that two large clades (Colorado and Snake River) are represented due to regional aridity and tectonic movements (Oakey et al., 2004). Pfrender et al. (2004) studied the biogeographic patterns and current distribution of molecular genetic variation among populations of *Rhinichthys osculus* in five major drainage systems in Oregon State. Their analysis used a 670 base pair segment of the mitochondrial cytochrome b gene. The study revealed a deep divergence among the major drainages and a high genetic diversity within the major basins. They applied a molecular clock to find the divergence times among the drainages and concluded the observed pattern of genetic variation is due to vicariant events during the late Miocene to early Pliocene period.

Smith and Dowling (2008) studied *Rhinichthys osculus* populations in the states of Nevada, Utah, and Arizona using the mitochondrial cytochrome b gene and incorporated geological data to estimate branch length and to calculate the divergence of the *Rhinichthys osculus* populations found in the different drainages. They concluded that the observed heterogeneity in the mitochondrial DNA rate of evolution was explained in part to differences in body size, environmental temperature and correlated life history traits (latitude and elevation of habitat). Smith and Dowling (2008) used difference in body size, environmental temperature, and life-history traits with geological data to calculate rates of nucleotide substitution and ages of nodes on the constructed phylogenetic tree. They estimated that the speckled dace species diverged from its sister species *Rhinichthys cataractae* and *Rhinichthys obtusus* at 6.3 Ma with an average divergence rate of 1.8% per million years and that the speckled dace have been in the Colorado drainage system for 3.6 million years (Smith and Dowling, 2008). Speckled dace were described as having dispersed through the Colorado River drainage and former drainage connections to the Los Angeles Basin in the past 1.9 million years. The estimated divergence among speckled dace lineages of the upper and lower Colorado River drainages occurred between 1.3-1.9 million years. The Colorado River speckled dace genetic divergence of allopatric lineages was accompanied by morphological adaptations to different stream substrates, and the small genetic distances observed among these populations of speckled dace show that there has been recent gene flow and a lack of reproductive isolation (Smith and Dowling, 2008).

Population Genetics

The goal of population genetics is to assess the degree and extent of genetic variation within and among subpopulations of particular taxa. The range a particular species occupies is dependent on such factors as the species' biological and ecological requirements for growth, development, and reproduction, the potential for the species to migrate to suitable surrounding geographical regions, and geographic barriers which may reduce or prohibit the ability to disperse (Lomolino et al, 2006). For example, Phillipsen and Metcalf (2009) found that there was significant genetic structure and the distribution of genetic diversity in *P. cadaverina* was due to the watersheds and mountain ranges separating the populations.

Sewall Wright (1931) developed a mathematical framework, known as *F*-statistics, by which genetic variation and population structure within and among subpopulations may be estimated. *F*-statistics may be used to estimate the degree of genetic differentiation at three population levels: 1) at the level of individuals relative to a particular subpopulation (F_{IS}); 2) at the level of individuals relative to the total populations (F_{IT}), and; 3) at the level of a subpopulation relative to the total population (F_{ST}). Hoekzema and Sidlauskas (2014) used population level genetic analysis of microsatellites to illustrate how different populations of *R. osculus* in the Oregon's Great Basin interact through geographic distance. The study concluded that the Foskett Spring speckled dace was a distinct population with no recent gene flow from the surrounding Warner Basin population.

The value of F_{ST}, known as the fixation index, provides a measure of the reduction in variation of the total population of the total population due to the differentiation among subpopulations, relative to the total population variation expected with no population subdivision (Conner and Hartl, 2004). A value of 0 for F_{ST} indicates no genetic divergence exists among subpopulations being compared, while a value of 1 for F_{ST} indicates complete population differentiation among subpopulations, resulting from fixation of alternate alleles of the same gene, and loss of heterozygosity within each respective subpopulation. Wright (1931; 1951; 1969) has described how F_{ST} may be determined from the variance in allele frequencies among subpopulations. The mathematics behind the fixation index F_{ST} is based on specific assumptions, namely that allelic variation is due to neutral processes, inheritance of alleles follows a diploid pattern, and the total population is infinite in size and panmictic (Wright, 1931). Since natural populations are generally small for the most part, there exists some controversy regarding the application of Wright's statistics to real populations (Nei, 2005).

Analogs of the fixation index *F* statistics have been developed specifically for haploid allele inheritance (such as occurs in mtDNA). These include the coefficient of gene differentiation (G_{ST}; Nei, 2005), Theta (θ; Weir and Cockerham, 1984), K_{ST} (Hudson et al., 1992), and θ_{ST} (Excoffier et al., 1992).

Tajima's D and Fu's F_s are both tests of molecular neutrality. The Tajima's D statistics compares estimates of the number of segregating sites to the mean pairwise differences between DNA sequences (Tajima, 1989). Fu's F_s statistic takes into account the number of different haplotypes in the sample based on Ewens' sampling distribution (Fu, 1997). Significant departure from the null hypothesis of neutral evolution may be caused by selection, migration, demographic changes such as bottlenecks or population expansion.

Past demographic history may be inferred using mismatch distribution analysis (Schneider and Excoffier, 1999). This technique tests the probability that a population fits a model of demographic expansion by comparing the observed distribution of differences between DNA sequences to a simulated distribution assuming a finite-sites model of substitution with heterogeneous mutation rates (Schneider and Excoffier, 1999).

Molecular Evolution of Mitochondrial DNA (mtDNA)

The mitochondrion is believed to have evolved via an endosymbiont origin. This is supported by the discovery of DNA in the mitochondria and from the realization that all mitochondria in all eukaryotes contain DNA and replicate independently (Scheffler, 1999). Mitochondrial DNA codes specific and unique genes for cellular respiration. The mitochondrion also possesses its own machinery of protein synthesis (Cecchini, 2003). Mitochondrial DNA (mtDNA) is a circular, double stranded DNA and a particular arrangement and order of 37 genes in most multicellular animals and some protozoans, including 22 transfer RNAs (tRNA) with some exceptions (Figure 4, Appendix A) (Cantatore et al. 1987). There are 13 messenger RNAs (mRNA) that code for polypeptide chains involved in ATP synthesis and the electron transport chain, 2 ribosomal RNAs (rRNA) responsible for translating mitochondrial mRNAs to protein, and a non-coding sequence called the control region which contains the origin of replication and regulates the transcription of the heavy (H) and light (L) strand transcription and the H-strand replication of the mitochondrial genome (Wilson et al. 1985). The animal mtDNA molecule lacks introns and intergenic space because of its small size, while plant mtDNA is larger and includes both introns and intergenic space between the genes (Scheffler 1999).

Complete mitochondrial DNA sequences have been determined for some vertebrates including several mammals, frogs, amphibians, and reptiles (Brown et al., 1982; Johansen et al., 1990; Belle et al., 2005; Akimoto et al., 2006, Boykin et al., 2007, Krause et al., 2010). Animal mtDNA is a 15-20 kb circular molecule containing 37 coding regions or genes involved in the electron transport chain and the process of oxidative phosphorylation in the mitochondria (Scheffler, 1999).

Mitochondrial DNA is often used in evolutionary studies because of its relatively high rate of divergence (Wilson et al., 1985). Mitochondrial DNA is used as a molecular marker because it evolves faster than nuclear DNA and each region has a different rate of evolution (Avise, 2004). For example, Brown, W.M. et al. (1979) showed that primate and rodent mtDNA nucleotide substitution occurs from 5 to 10 times faster than in single copy nuclear DNA by comparing the extent of sequence divergence in mitochondrial and single copy nuclear DNA. The authors determined percent minimum and estimated sequence difference in mtDNA with thermo stability of heteroduplexes in single copy nuclear DNA, with an assumption that percent difference equals Δt_m. Brown, W.M. et al.(1979) also stated that the lack of or inefficiency of the mtDNA replication complex editing function and high turnover rate cause the high rate of mutation, while in nuclear DNA the replication editing complex is very efficient and able to correct errors during rounds of replication. Studies that compare the mtDNA cytochrome b coding region and non-coding control region rate of mutation showed that the non-coding control region has a higher rate of mutation than the coding region, but for closely related taxa of less than ten percent sequence divergence the control region evolved at a slightly slower rate (Tang et al, 2006). However, in fish the rate of evolution of the mtDNA control region has been estimated to be up to 43 times more rapid than cytochrome b coding region with a strong transition/transversion bias (McMillen and Palumbi, 1997).

Animal mitochondrial DNA (mtDNA) is a marker of choice for phylogeographic studies because it is maternally inherited and does not undergo significant recombination, so that each molecular lineage has a distinct genealogical history (Dawid et al., 1972). Currently, mitochondrial DNA sequence data is the main vehicle used to determine genetic variation among individuals across a species' range. Mitochondrial DNA has become the most popular genetic marker for this purpose and has been employed in more than 80% of published studies (Avise, 2004). The advantages of mtDNA include versatile PCR primers which amplify mtDNA sequences without purification (Graur and Li, 2000), a high mtDNA copy number that can be extracted from museum materials and archaeological remains like bones and teeth (Futuyma, 2005), and a high mutation rate compared with the nuclear genome which results in relatively rapid genetic variation (Wilson et al., 1985).

The regions of the mtDNA genome most often used in genetic studies are the cytochrome b gene, the control region, transfer RNAs, and ribosomal RNAs (Figure 4, Appendix A) (Kvist et al., 2003). Studies have shown that DNA sequences coding for certain select amino acids are conserved across different vertebrate classes, however other coding DNA sequences and amino acids are observed to be more variable (Futuyma and Edwards, 2005). Differences in the DNA sequences are largely due to base substitutions in the protein-coding genes. The cytochrome b (cyt b) gene has become a universal metric to compare rates of evolution (Avise, 2004). The cyt b gene is variable enough for population level questions but conserved enough to clarify deeper phylogenetic relationships (Tang et al., 2006).

The average effective population size of mtDNA is ¼ that of diploid nuclear genes because there is a dual effect resulting from the haploid inheritance of mtDNA and mtDNA's normal pattern of single parent inheritance (Avise et al. 1988), so haplotype frequencies can drift rapidly, creating genetic differences among populations in relatively short time (Birky et al., 1983). Haplotype distributions and their relative frequencies can then be used to infer historical relationships among populations (Avise, 2000).

Structure and Molecular Evolution of Cytochrome b gene

Cytochrome b is trans-membrane protein with a central catalytic subunit of ubiquinol, namely cytochrome c reductase, an enzyme present in the respiratory chain of the mitochondria (Figure 5, Appendix A) (Lydeard, C. and Roe, K.V. 1997). The cytochrome b gene codes for part of the cytochrome bc1 complex in the electron transport chain (ETS) of the cellular respiration process. The main function of the cytochrome bc1 complex in the electron transport chain is the transport and the generation of adenosine tri-phosphate (ATP) which is the primary energy carrying molecule of cells. The cytochrome b protein contains two heme groups, the b-562 high potential heme (b_H) and b-566 low potential heme (b_L) with different spectroscopic and redox properties, eight trans-membrane helices with extra and intermembrane loops, and two quinone reacting centers (Esposti, M.D. et. al. 1993). Cytochrome b extra and intermembrane loops are located in two regions: inside the mitochondria (positive side), and in the intra-membrane space of the mitochondria (negative side). The N-terminus and C-terminus is also located in the intra-membrane space of the mitochondria (Figure 6A, Appendix A). The cytochrome b protein is translated from the cytochrome b gene located in the mitochondrial DNA. There is natural variation in the structure of cytochrome b, and certain regions/residues are conserved due to the importance of their position within the three-dimensional protein structure. Esposti et al. (1993) compared 800 sequences of mitochondrial cytochrome b from different organisms and found 9 invariant amino acids. Esposti et. al. (1993) also found that there are 27 highly conserved amino acids in cytochrome b that are found to be the same in most organisms. In the eight trans-membrane, alpha helix structures in the matrix, helices B, C, and A are the most conserved, helices D and F are less conserved, and helices E and H are the least conserved. There is also an uneven distribution of conservation in the regions of the cytochrome b protein that protrudes at the two sides of the membrane (Figure 6A, Appendix A).

Irwin et al. (1991) also found that comparing the cytochrome b sequences of different mammals support the current structure-function models for membrane-spanning protein (Figure 5, Appendix A). The variable positions of the amino acid residues were found within the trans-membrane segments or at the amino and carboxyl end of the protein while the outer surface of the protein appears to be evolving more slowly than either the trans-membrane region or the inner surface (Figure 5, Appendix A) (Irwin, D.M. et al. 1991). The variable changes in the trans-membrane portions are changes between hydrophobic (leucine, isoleucine, and valine) residues (Figure 6B, Appendix A) (Irwin, D.M. et al. 1991). For example, Yang and Yoder (1999) found that the relative rates of substitution at 1st, 2nd, and third positions within codons of cytochrome b are in the proportion of 1:0.26:10.23 estimated by a maximum likelihood model. Conserved and variable sites in cytochrome b can also be estimated by codon.

The complete mtDNA cytochrome b gene has been used to look at genetic distances in vertebrates since 1998, and is the most extensively sequenced gene in vertebrates (Johns and Avise, 1998). The evolutionary dynamics and the biochemistry of the cytochrome b gene protein product are better characterized than most other molecular systems because of the amount of research and studies using the cytochrome b gene (Howell, 1989; Esposti et al., 1993). The understanding of mtDNA cytochrome b has expanded through DNA sequencing and analysis of its protein sequence, which assist in explaining how the cytochrome b protein has evolved through time while still maintaining its specific function in the oxidative phosphorylation process (Figure 4, Appendix A) (Esposti et al., 1993). Changes in the cytochrome b protein sequence can greatly affect the function of the protein and its conformation, which may lead to changes in a species (Figure 5B, Appendix A). The evolution of mitochondrial proteins occurs predominantly by single nucleotide substitution, while the typical lack of recombination makes the interpretation of any observed differences in molecular evolution within or between groups more concrete (Brown et al., 1996). For example Ma et al. (1993) studied the structure and evolution of opossum, guinea pig and porcupine cytochrome b by calculating the number of nucleotide substitutions per nonsynonymous and synonymous sites. The study showed that in the opossum, guinea pig and porcupine cytochrome b gene, most of the amino acid residues involved in the two reaction centers were conserved and the eight trans-membrane domains possessed nonsynonymous substitutions that created the most divergent regions in the cytochrome b sequences among the cutherian lineages. McClellan et al. (2001) estimated the influence of selection on the variable amino acid sites of the functional domain found in the cytochrome b protein. The researchers used the rate of synonymous and non-synonymous substitutions to calculate the differences among the three functional cyt b domains: the inter-membrane, matrix, and trans-membrane regions found in the cytochrome b protein. The study showed that the inter-membrane domain evolved more slowly than the functional domains found within the mitochondrial matrix and the trans-membrane region because evolutionary constraints are imposed by the function of the Q redox center. The 105 cyt b amino acid residues important for the Q redox center function are found within the mitochondrial inter-membrane region in mammals, and 29% of the amino acid residues are completely or mostly conserved in metazoans (Zhang et al. 1998). The matrix domain of cyt b is located entirely within the inner surface of the mitochondrial membrane and includes 65 amino acid residues that have a high proportion of polar and basic R-groups, and is believed to have no part in the proton-input function of the Q redox center (Griffith and Tavare 1997). The comparison of the relative proportions of amino acid residues in the cytochrome b protein were seen through the relative patterns of non-synonymous substitution found in different mammals such as pocket gophers and cetariodactyls. Another study done by Prusak and Grzybowski (2004) showed that there is a strong non-random distribution of variable nucleotides in the cytochrome b sequences of vertebrates, with the highest differences found at the third codon position, which are often silent substitutions. The smallest observed nucleotide substitution difference occurred in the second codon position, and changes at the first codon position were less frequent than in third position because this can result in amino acid substitution that could be detrimental to protein function. For example, the quail, frog, python, and elk exhibit C over A nucleotide bias, and artiodactyl species contain fewer pyrimidines than purines in their cyt b nucleotide distribution. The amino acid sequences of all 14 species examined showed that the overall rate of amino acid identity is 61%, which was a similar observation seen by Irvin et al. (1991). Prusak and Grzybowski (2004) used the structural model of cytochrome b to illustrate the region being studied and showed conserved regions because of proximity to Q redox center.

PROPOSED STUDY

This study has two broad goals. The first goal will be to estimate the phylogeography of the *Rhinichthys osculus* (Santa Ana Speckled dace), and the second goal will be to characterize the molecular evolution of the cytochrome b gene. The broad extent and pattern of genetic variation among populations of *Rhinichthys osculus* (Santa Ana Speckled dace) found in California (Figure 1, Appendix A) will be determined using the entire mtDNA cytochrome b gene as a genetic marker (Figure 7, Appendix A). This study will determine whether the subpopulations of Santa Ana speckled dace have significant population structure due to restricted gene flow among the different tributaries, watersheds, and mountain ranges. Recent studies (Cunha et al., 2002; Pfrender et al., 2004; Oakey et al., 2004; Smith and Dowling, 2008) have shown that climate, geology, and hydrology have played major roles in the initiation and divergence of genetic lineages of *Rhinichthys osculus* in Western North America. Other studies have also identified the Transverse Range of California as a biogeographical barrier that significantly influenced the range and population structures of numerous Southern California flora and fauna (reviewed in Calsbeek et al., 2003; Chatzimanolis et al, 2007; Phillipsen and Metcalf, 2009). The prevailing notion is that populations of Northern and Southern California speckled dace separated by the Transverse Range were considered to be one and the same (*R. osculus carringtonii*) (Culver and Hubbs, 1917; LaRivers, 1952; Hubbs et al., 1974), even though Cornelius (1969) questioned this grouping on the basis of morphology. This study will determine whether the Santa Ana speckled dace is genetically distinct from the northern central coast and eastern desert populations of *R. osculus*.

The second goal of this study is to describe and analyze the structural characteristics and pattern of base sequence substitutions in the cytochrome b gene to better understand the molecular evolution of the gene. The cytochrome b gene is a protein-coding gene with a specific function within the cellular respiration of organisms (Figure 5, Appendix A). Because of the importance of the cytochrome b protein in cellular respiration, one of the main goals of this study is to map the locations of the conserved and non-conserved regions within the cytochrome b gene sequence and analyze the types of base substitutions. By translating the DNA sequence of cytochrome b to amino acid sequences the study will attempt to predict the gene’s molecular protein conformation and likelihood of deleting substitutions based on domain found in the trans-membrane region of the mitochondria (Figure 6A & 6B, Appendix A).

This information will contribute to answering the question of taxonomic identity of the Santa Ana Speckled Dace and the designation of this species as an evolutionary significant unit. The Santa Ana Speckled Dace is currently listed as a species of special concern by the California Department of Fish and Game.

CHAPTER TWO

MATERIALS AND METHODS

Data Collection

The specimens of *Rhinichthys osculus* used in this study were sampled by the California Department of Fish and Wildlife and the United States Forest Service and with Dr. Anthony Metcalf, Department of Biology, and California State University San Bernardino under auspicious of the United States Forest Service and California Department of Fish and Wildlife. Specimens of *osculus* were sampled from different watersheds and mountain ranges in Southern California. The four watersheds sampled include the Santa Ana R., the San Gabriel R., the Santa Maria R. and San Luis Obispo R. The populations used form a representative sampling of *R. osculus* within Southern California as specimen availability allows (Figure 1, Appendix A). Samples of *R. osculus* from populations inhabiting the Owens and Colorado Rivers located at the Eastern Sierra Nevada and the Rocky Mountain ranges respectively, was used as an out group. Sampling locations were varied to allow analysis of genetic variation within and among the four major Southern California watersheds (Table 1, Appendix B). A minimum of seven specimens was collected from each watershed if allowed by availability.

Sampling location within watersheds included different tributaries. The Santa Ana River watershed R. osculus samples were sampled from Plunge

Creek, City Creek, Twin/Strawberry Creek, Cajon Creek, Lytle Creek, and Indian Creek (Figure 8A, Appendix A). The San Gabriel watershed sample sites include the North and West forks of the San Gabriel River, and Cattle Canyon Creek (Figure 8A, Appendix A). The sampling locations within the Santa Maria watershed include the Cuyama River, Manzana Creek, and Davey Brown Creek (Figure 8B, Appendix A). The San Luis Obispo watershed includes sampling sites Stenner Creek, Brizziolari Creek, and the San Luis Obispo Creek tributaries (Figure 8B, Appendix A). The specimens sampled from the Owens River were taken from the tributaries Marvin's Marsh and Big Pine Creek (Figure 8C, Appendix A). The Colorado River sequences and Oregon sequence via GenBank (Accession number DQ990216.1-DQ990232.1) came from Arizona State University courtesy of Dr. Thomas Dowling's research laboratory. The *R. osculus* samples obtained from the Owens River and the sequences from the Colorado River were utilized to further examine the phylogenetic relationship and origins of the Santa Ana Speckled Dace populations within California, while individuals of the closely related cyprinid taxa *Rhinichthys obtusus and Rhinichthys cataractae* were used to root all native Southern California populations of *R. osculus* in the phylogenetic analysis (Table 1, Appendix B).

Molecular Methods

The techniques used for the genetic analysis of the speckled dace study are as follows. For each specimen, approximately 30 milligrams of wet muscle tissue obtained close to the tail were lyophilized using vacuum desiccation for approximately 1 hour and ground to a fine powder in liquid nitrogen. Total genomic DNA was extracted using a DNEasy kit and visualized using agarose gel electrophoresis, and DNA concentration and purity determined at 280 nm using UV spectrophotometry with dilution factor of 1:20. An approximately 1260 base pair fragment of the mitochondrial cytochrome b gene including the tRNA glutamine and tRNA threonine partials were amplified by polymerase chain reaction (PCR) using external primers H16526/L15267 (Figure 7, Appendix A) (Briolay et al, Kotlik et al 1998). PCR products were detected by inspecting ethidium bromide stained electrophoresis gels under UV light. To achieve the highest purity of the PCR amplication products, a Qiagen QIAquick Gel Extraction kit was used. Direct DNA sequencing of the PCR amplication product and DNA sequencing of cloned PCR product were used as necessary. For clone sequencing, the Invitrogen TOPO TA Cloning Kit was used to incorporate the PCR product amplicon containing the cytochrome b gene (1260 bp) to a plasmid vector. The cloning reaction was as follows: 3 µL of amplified PCR product, 1 µL salt solution (200 mM NaCl, 10 mM $MgCl_2$), 1 µL sterile dH_2O, and 1 µL of TOPO vector incubated at room temperature for 30 minutes. Bacterial transformation of Mach1-T1 Competent Cells were performed by addition of 2 µL of TOPO cloning reaction to a vial of freeze-dried Chemically Competent *E. coli* cells, and incubated on ice for 5 minutes. The cell solution were heat-shocked for 30 seconds at 42˚C, followed by addition of 250 µL of S.O.C. medium and incubated on a shaking water bath at 37˚C for 1 hour. The 35 µL of transformed cells were plated on LB+kanamycin (LB+kan) plates (50 mg/mL) and incubated overnight at 37˚C. Ten isolated transformed cell bacterial colonies were selected, transferred to LB+kan nutrient broth culture tubes, and incubated overnight at 37˚C. A QIAprep Spin Miniprep Kit was used to extract and isolate transformed Plasmid DNA from the broth cultures. EcoR1 restriction digestion analysis and PCR amplification of the inserted region were used to test the 10 plasmid DNA samples for the presence of the cytochrome b gene. Bi-directional cycle sequencing reactions were performed on at least four positive plasmid samples using M13F and M13R primers complimentary to the M13 primer sites within the TOPO vector DNA flanking the cytochrome b gene insertion site. Specific internal primers designed for this study completed the double stranded cytochrome b gene sequence (Figure 7, Appendix A). Both strands of each amplicon were sequenced. The TOPO cytochrome b gene sequencing products were run on a LI-COR 4300 model sequencer (LI-COR, Lincoln, NE). Resulting DNA sequences were proofread and assembled using the programs AlignIR (LI-COR), LASERGENE (DNASTAR, Inc.) and MEGALIGN (LI-COR), and identical sequences/consensus sequences collapsed using MacClade (Maddison and Maddison 1989), aligned using ClustalX (Thompson et al. 1997) and ambiguous bases resolved using LASERGENE (DNASTAR, Inc.), Mega 6.0 (Tamura et al. 2013), and MEGALIGN (LI-COR).

Sequence Analysis

The aligned mtDNA cytochrome b consensus sequences were used to estimate the genetic variation within and among Southern California *Rhinichthys osculus* populations in the different watersheds and construct the cytochrome b protein structure. This analysis illustrate whether population structure exists in Southern California and also elucidate the historical biogeographical processes that may have played a role in shaping the variation within and among populations in the different watersheds.

Prior to the analysis the sequences were edited to include just the cytochrome b sequence, to maintain consistent codon positioning with cytochrome b. The pattern of nucleotide substitutions, insertions, and deletions within the cytochrome b region may reveal information relating to the molecular evolution of the gene including codon patterns of substitution within codons and among the different parts of the gene such as trans-membrane, inter-membrane, and extra-membrane (Figure 5, Appendix A). Multiple alignment of the cytochrome b region was obtained using the software package Clustal X (Thompson et al, 1997, Larkin et al, 2007), Mega 6.0 and Mega 7.0 (Tamura et al, 2013) to locate the presence and location of genetic variation in the gene. Genetic variation was described using descriptive statistics including nucleotide frequency, number of haplotypes, number of polymorphic sites, number and type of substitutions, number of transitions (Ti) and transversions (Tv), Ti/Tv ratio, and uncorrected and/or corrected genetic distance among haplotypes. To evaluate sequence dissimilarity, calculation of pair-wise genetic distances on haplotypes will be conducted. Pair-wise genetic distance is an estimation of the number of nucleotide substitutions per nucleotide site between two DNA sequences. Based on factors such as differing substitution rates for Ti and Tv, or multiple substitutions at one site, genetic distance may be uncorrected or corrected. Descriptive statistics are determined using Paup* 4.10b (Swofford 2003), Mega 6.0 (Kumar et al. 2001, Tamura et al. 2013), and Modeltest 3.7 (Posada and Crandall, 1998).

Phylogeography

In this study, the phylogenetic relationships among related Speckled Dace cytochrome b haplotypes was inferred from gene trees and estimated using the methods of: 1) genetic distance; 2) maximum parsimony (MP), and; 3) Bayesian analysis using PAUP, Mega 7.01 (Tamura et. al, 2013) and MrBayes (Ronquist et al. 2012) software for phylogeny testing. Relative time divergence was calculated for each region using Mega 7.01 (Tamura et. al, 2013). Modeltest 3.7 is used to determine the type of evolutionary model that the speckled dace cytochrome b sequence best fits (Posada and Crandall, 1998). To visualize intraspecific genetic variation within *Rhinichthys osculus*, a haplotype minimum spanning network was created using the software program TCS v.1.21 (Clement et. al. 2000).

Population Genetics

Hierachical analysis of 1) region, 2) watershed, and 3) drainages will be calculated using AMOVA to analyze phylogenetic structure and breaks. Tajima's D (Tajima, 1989) and Fu's Fs (Fu, 1997) will be used to detect population decline, stability, or exponential growth using genetic variation within populations at the DNA level. AMOVA was performed using the software programs Arlequin 3.5 (Excoffier and Lischer , 2010) and GenAlEx 6.501 (Peakall and Smouse, 2006) to test for significance of any detected differences in the proportion of variable nucleotides present, which may reveal relative variability of the cytochrome b gene. Also, demographic understanding of historical population patterns will be analyzed using the Tajima 1D (Tajima 1989) and Fu's F (Fu, 1997) methods using GenAlEx 6.501 (Peakall and Smouse, 2006) and Arlequin 3.5 (Escoffier and Lischer, 2010).

Migration rates (N_m) were estimated to visualize gene flow between regions, watersheds, mountain ranges using GenAlEx 6.501 (Peakall and Smouse, 2006). GenAlEx 6.501 was used to calculate the Principle coordinates analysis (PCoA) for the different regions, watersheds, mountain ranges, and individual tributaries. Mismatch distribution of pairwise nucleotide differences among *R. osculus* was also calculated using Arlequin 3.4 (Excoffier and Lischer, 2010).

Protein Conformation Analysis

The goal of the protein conformation analysis is to further examine the molecular evolution of cytochrome b gene and the molecular level changes to determine the pattern of base and amino acid mutation across the portions of the gene. Base composition of cytochrome b genes will be analyzed by using the method of Prager and Wilson (1988). To determine the spatial pattern of substitutions within the cytochrome b sequence, the number of inferred nucleotide substitutions at each codon position will be compared to the number expected according to Poisson and negative binomial distributions (Holmquist 1983). A computer molecular model of the protein will also be analyzed using proteomic (Chou and Fasman 1978; Barker et al. 1996) and bioinformatic protocols (Bairoch 1991; Hofmann et al. 1999). Model sequence was submitted to raptorX server to predict the protein model and calculate P-value, alignment score, secondary structure prediction (3 class and 8 class), solvent accessibility and disorder prediction (Kallberg et al., 2012). Protein model summary results include best template for the submitted sequence and calculation of this analysis will further strengthen the understanding of how the amino acid sequence of cytochrome b is susceptible to mutational change, and to compare the rates of intra-, extra-, or trans-membrane evolution of the gene and the respective effect on the protein conformation (Figure 6A & 6B, Appendix A). This information will be used in the Bayesian analysis as a way to partition the cytochrome b sequence and to calculate the models of evolution. The Chi-square (X^2) test was utilize to compare observed nucleotide and amino acid substitutions found in the cytochrome b (nucleotide and amino acid) sequences and domain partitions with the expected calculated null hypothesis where there is no difference found in the cytochrome b (nucleotide and amino acid) sequences and domain partitions.

CHAPTER THREE

RESULTS

Sequence Analysis

The complete 1155 base pair cytochrome *b* gene from 92 specimens of *Rhinicthys osculus* was sequenced from 3 distinct geographic locations and yielded 34 haplotypes, representing 7 watershed locations (Table 2, Appendix B). The sequences were trimmed to 1140 bp for comparison with 11 *R. osculus* specimens from the Colorado River (AZ), 1 specimen from Lake County (OR), 1 specimen from Columbia River (B.C., Canada) and 2 different species as out-groups. The average cytochrome *b* nucleotide composition among all haplotypes and regions was T=0.290, C=0.276, A=0.259, and G=0.176. There were 965 conserved sites and 190 variable sites. There were 135 3rd- position base substitutions, 17 were 2nd-position base substitutions, and 38 were 1st-position base substitutions. There were 175 parsimony informative sites. The average number of nucleotide substitutions among all haplotypes was 54 third-position codon substitutions, 1 second-position codon substitution, and 8 first-position codon substitutions. The overall Transition/Transversion (ti/tv) ratio was 6.75. There were 147 synonymous nucleotide substitutions and 28 non-synonymous nucleotide substitutions, with 11 substitutions occurring in the intra-membrane or extra-membrane region, and 17 substitutions occurring in the trans-membrane region. The cytochrome *b* base frequency and substitution information is illustrated in Table 4, Appendix B.

Phylogeography

The methods of distance, maximum parsimony, maximum likelihood and Bayesian method were used to analyze the phylogeny of *R. osculus.* Best-fit evolutionary models were generated through Modeltest 3.7 (Posada and Crandall 1998) and MEGA 6.0 and are listed in Table 6, Appendix B. MEGA 6.0 generated consensus trees of distance (NJ), parsimony (MP), and maximum likelihood (ML) which were consistent and contained similar topologies (Figure 9, Appendix A). Three types of Bayesian trees were created, using data partition by nucleotide bases, by codon positions, and by intra- versus extra-membrane domain separation. The three tree topologies produced using Bayesian analysis (Figure 10, Appendix A) were congruent to the first three methods of phylogeny estimation. There were four distinct well-supported clades: a southern clade consisting of the haplotypes found within the Santa Ana and San Gabriel watersheds (colored Green in Figures 10, Appendix A); a Colorado River clade consisting of GenBank sequences (colored yellow in Figure 10, Appendix A); a Central Coast clade consisting of haplotypes found within the Santa Maria and Cuyama River watershed, and in San Luis Obispo Creek (colored Pink in Figures 10, Appendix A); and an Eastern desert clade consisting of haplotypes found within the Owens River watershed (colored Lavender in Figures 10, Appendix A).

The uncorrected p-distance between Santa Ana Speckled Dace and Colorado, Central Coast, Eastern Desert dace is 6.0%, 8.6%, and 9.3%. The estimated divergence time between the Santa Ana Speckled Dace and Colorado River dace is 3 million years. Relative Divergence Time for each clade was 0.07 for Colorado River dace population, 0.11 for the Northern California dace population and the Southern California Santa Ana speckled dace population was 0.06 (Figure 11, Appendix A).

Population Genetics

The samples were analyzed using Analysis of Molecular Variance (AMOVA) to infer the amount of genetic variation among the geographic regions of *R. osculus* specimens and within the Santa Ana Speckled Dace in Southern California. There was a high degree of genetic variation among all *R. osculus* populations including a significant genetic structure associated with watersheds, mountain ranges, and geographic groupings based on locations within California from different specimen regions as well as the Colorado specimens obtained from GenBank. Intra-population haplotype diversity (*h*), percentage of variable sites, nucleotide diversity (Π), Tajima's D, and Fu's F_S were determined using GenAlEx 6.501 and ARLEQUIN 3.5 and are summarized in Table 7A-7B, Appendix B. The genetic variance of populations among regions (Φ_{PT}), sequences among populations (Φ_{RT}), and among regions (Φ_{PR}) was tested by grouping populations by watersheds mountain ranges, and geographic locations.

Φ_{PT} values among the five *R. osculus* populations sampled ranged from 0.9526 to 0.9699 ($p<0.000001$) in the three geographic models tested, indicating high levels of population structure among all populations of *R. osculus* (Table 7A-7B, Appendix B). When grouped by watersheds, mountain ranges, or general geographic location within California, the degree of total genetic variation attributed to differences among groups was in excess of 89% (Φ_{RT} = 0.92190, 0.89789, and 0.93430 respectively). Populations also possessed genetic structuring within geographic groups, with Φ_{PR} = 0.61453, 0.53595, and 0.53226 for watersheds, mountains, and geographic areas respectively. Table 8, Appendix B summarizes the results of AMOVA and Φ-statistic calculations.

AMOVA analyses of cytochrome *b* genetic variation within populations partitioned by watershed, mountain ranges, and geographic regions showed Φ_{PT} values of 0.97, 0.95, and 0.97 respectively. Among population groups the Φ_{RT} values were 0.92, 0.90, and 0.93, and among populations within groups the Φ_{PR} values were 0.61, 0.54, and 0.53 (Table 8, Appendix B). The model with the highest percentage of molecular variance of all models tested was the data partition by 3 regions (Southern California, Central California Coast, and Eastern Desert). In the 3 Region model, the molecular variance attributed to differences in geographic regions was 92%, the molecular variance among populations within regions was 5%, and within-population molecular variance was 3% with the following Φ-statistics: Φ_{RT} =0.934, Φ_{PR} =0.532, Φ_{PT} = 0.969. The average among region migration rate Nm was low, Nm = 0.015. A fourth AMOVA was calculated with the addition of dace from a 4th geographic region (Colorado River), and the results were similar to the 3 geographic region partition described above. The Φ-statistics for the four region AMOVA were Φ_{RT} =0.905, Φ_{PR} =0.414, and Φ_{PT} =0.944 with an overall migration rate Nm = 0.037 respectively between the 4 regions (Figure 13A-13B, Appendix A). Significant population structure exists among populations in all groups (Figure 13A-13B, Appendix A), which is congruent with the Bayesian analysis results and tree that revealed 4 distinct clades (Figure 10, Appendix A).

Additional population-level analysis within the Southern California clade was performed using AMOVA. Southern California populations were partitioned by nearest mountain range, by watershed, and by the individual tributary the population occupied. The percentage of molecular variance present when partitioning Southern Speckled Dace populations by nearest mountain range showed that there was greater molecular variance among dace populations within (55%) rather than among (45%) mountain ranges (Figure 14, Appendix A), with Φ_{PT} =0.451 and a migration rate N_m of 0.609 among mountain ranges. Partitioning Southern Dace populations by watersheds showed a percentage of molecular variance that was larger among watersheds (72%), and smaller within watersheds (28%). The value of Φ_{PT} =0.724, with a migration rate Nm of 0.191 (Figure 15, Appendix A). Partitioning populations by tributaries also showed greater molecular variance present among tributaries (66%) rather than within tributaries (34%), with Φ_{PT} =0.664 and an among tributaries migration rate N_m =

0.253 (Figure 16, Appendix A). Mantel test analysis showed isolation by distance of Santa Ana Speckled Dace from the Central California Coast dace and Eastern California Desert dace populations with $r^2 = 0.510$ that was similar to a second Mantel test that included the Colorado dace population with $r^2 = 0.347$ (Figure 17-18, Appendix A).

Principle coordinates analysis (PCoA) identified three distinct speckled dace populations; a Southern California population, a Central Coast California dace population, and an Eastern California Desert dace population (Figure 19, Appendix A). This is also consistent with a second PCoA analysis, but with the addition of the Colorado dace population creating a fourth distinct population, and still preserving the 3 original population groups within California. The Colorado dace population is more closely related to the Southern California population than either the Central Coast or the Eastern Desert populations (Figure 20, Appendix A). The cytochrome b average percent genetic distance between each population was as follows: Southern California to Central California uncorrected p-distance was 8.620%, Southern California to Eastern Desert p-distance was 9.343%, Southern California to Colorado p-distance was 6.012%, Central California to Eastern Desert p-distance was 6.920%, Central California to Colorado p-distance was 7.379%, and the Eastern Desert to Colorado average p-distance was 7.491% (Table 9, Appendix B).

Mismatch distribution of pairwise nucleotide differences among *R. osculus* individuals in the Southern California region indicates a population structure that follows a model of spatial expansion assuming constant deme size. However, the Southern California region populations do not fit a model of sudden demographic expansion. Mismatch distribution indicates the Central Coast and Eastern Desert regions each fit both the spatial expansion assuming constant deme size, and the sudden expansion models, within their respective population structures (Figure 21A-21C, Appendix A).

Molecular Evolution and Protein Conformation Analysis of Cytochrome b gene

The cytochrome *b* gene contains 384 amino acids, and consists of 8 trans-membrane segments which span amino acids 33-53, 86-106, 115-135, 178-198, 224-244, 288-308, 323-343, and 349-369 (Figure 22A-22B, Appendix A). There were 4 heme group locations; at amino acid 83, 97, 182, and 196. The intra-membrane and extra-membrane segments were located at amino acids 1-32, 54-85, 107-114, 136-177, 199-223, 245-287, 309-322, 344-348 and 370-384. The N-Terminal and C-terminal sections of the cytochrome *b* gene are located in the cytoplasmic membrane. There were a total of 28 amino acid substitutions in cytochrome *b*. Eleven amino acid substitutions occurred within the intra-membrane and extra-membrane domains, while 17 amino acid substitutions occurred within the trans-membrane domains. Cytochrome *b* contained 13 hydrophobic to hydrophobic amino acid substitutions, 6 hydrophobic to hydrophilic amino acid substitutions, 4 charged to polar amino acid substitutions, and 2 glycine to charged amino acid substitutions (Table 10, Appendix B). The speckled dace cytochrome b gene contains 356 conserved amino acids and 28 variable amino acid sites, of which 12 are parsimony informative sites, resulting in 23 amino acid haplotypes (Table 10, Appendix B). Chi-squared (χ^2) analysis indicated that there were significant differences in the frequency of single nucleotide polymorphisms in the intra- and extra-membrane cytochrome *b* domains compared to the trans-membrane domains (χ^2= 6.598; p-value= 0.0102). There was also a significant difference in base substitution frequency by codon position (χ^2 = 125.192; p-value= 6.5×10^{-28}), and in the frequency of hydrophobic to hydrophobic, or hydrophilic to hydrophilic amino acid substitutions (χ^2= 4.804; p-value= 0.0284). There were no significant differences in single nucleotide polymorphisms in amino acid substitutions by domain (χ^2= 1.226; p-value= 0.2122), by hydrophobic to hydrophilic substitutions (χ^2= 0.8076; p-value= 0.3688), or by uncharged/charged substitutions (χ^2 = 2.4093; p-value= 0.1206) (Table 5, Appendix B).

The cytochrome *b* sequences were analyzed using the SWISS-MODEL automated protein structure homology-modeling server and the Swiss Pdb-Viewer Software, to elucidate the conserved and invariant residues within the cytochrome b gene as well as the protein conformation of the protein. There are 356 conserved residues and 28 variable residues out of a total of 384 amino acids present (Table 11, Appendix B). The amino acids located in the trans-membrane domain were all hydrophobic. Trans-membrane amino acids were highly conserved. The symmetric motif of the trans-membrane glycines and histidines were characteristic of other cytochrome *b* amino acid sequences from other vertebrate animals. The most conserved regions in cytochrome *b* are the trans-membrane helices, and the least variable regions are the extra-membrane amino acid sequences and helices. There was periodicity in the 8 trans-membrane helices. The dominant trans-membrane conformation observed was an alpha helix containing 21 amino acid residues. The overall shape of cytochrome *b* gene is illustrated in Figure 22A-22B, Appendix A. The shape illustrates both the form and function of the cytochrome *b* protein. RaptorX data analysis revealed that protein model illustrated on Figure 22A-22B, Appendix A have a p-value = 5.27×10^{-10} with an overall uGD(GDT) score of 340(88) and 14(3%) positions predicted as disordered in the N-terminus/C-terminus of the amino acid sequence (Figure 22A-22B, Appendix A). Figure 23A-23B, Appendix A illustrates the placements of the 3-class secondary structures, 8-class secondary structures, solvent accessibility, and disorder prediction of the protein model. Sixty-seven (67%) of the secondary structure are helices and 31% are coiled. Solvent exposure of the protein sequences is as follows: 38% fully exposed, 42% medium exposure, and 18% buried (Figure 24, Appendix A).

CHAPTER FOUR

DISCUSSION

This study had two broad goals. The first goal was to describe the phylogeny of the Santa Ana Speckled Dace (*R. osculus*) in relation to populations inhabiting other geographic regions within California. I found that the Santa Ana Speckled Dace is significantly genetically distinct from *R. osculus* populations in other geographic regions, a pattern consistent with species level status.

The second goal was to describe the structure and molecular evolution of the mitochondrial DNA cytochrome *b* gene of *R. osculus*. I found that the majority of the nucleotide substitutions in cytochrome *b* are synonymous mutations which do not affect the primary, secondary, or tertiary structure of the gene.

Phylogeography

This study showed evidence of significant population structure for *Rhinichthys osculus* across geographic regions. Thirty-four unique haplotypes representing 5 watersheds within California were identified (Table 2, Appendix B). Phylogeographic analysis revealed three reciprocally monophyletic clades of *R. osculus* in California; (1) a Southern California clade consisting of Santa Ana Speckled Dace populations inhabiting the Santa Ana River and San Gabriel River watersheds; (2) a Central California Coast clade consisting of *R. osculus* populations inhabiting San Luis Obispo Creek and the Santa Maria River watershed; and (3) an Eastern California Desert clade comprised of *R. osculus* populations inhabiting the Owens River watershed (Figure 9-10, Appendix A). A fourth population consisting of *R. osculus* from the Colorado River was included to generate a 4 region analysis (Figure 9-10, Appendix A). Southern California to Central California uncorrected p-distance was 8.7%, Southern California to Eastern Desert p-distance was 9.5%, Southern California to Colorado p-distance was 6.1% (Table 9, Appendix B). Partitioning nucleotide base sequences by domain or codon position in the Bayesian analysis yielded congruent trees. There is distinct separation between the Southern California clade (the Santa Ana Speckled Dace) from the Central Coast and Eastern Desert clades. This study confirms that the Santa Ana Speckled Dace is most closely related to the Colorado River speckled dace (Smith and Dowling, 2008), and genetically distinct from dace inhabiting other regions of California.

Early studies of *Rhinichthys osculus* elucidated the distribution pattern of the organism throughout the western United States by way of stream connections to the ancient Colorado River drainage (Miller, 1946; Cornelius, 1969). Spencer et al. (2008) stated the most plausible dispersal route of *R. osculus* to Southern California was from the lower Colorado River drainage via the Mojave River (Figure 10, Appendix A). Cornelius (1969) suggested that the speckled dace inhabiting the Central Coast of California originated through migration from Southern California. This study does not corroborate that conclusion. Inclusion of dace specimens from Oregon and British Columbia in the phylogenetic analysis suggests that the California Central Coast clade is most closely related to speckled dace found in the Pacific Northwest region. This finding is similar to the findings of Oakey et al. (2004) using restriction enzyme analysis. Both the California Central Coast and Eastern Desert populations are separated from the Southern California population by geographic distance and by the Transverse Range, leading to restricted gene flow among populations. Between the Central Coast clade and Eastern Desert clade, there are two mountain ranges (the Coast Range and the Sierra-Nevada Mountains), and the presence of these considerable physical barriers create isolated populations within each region. These geographic barriers and isolation by distance account for the absence of gene flow among the 3 geographic regions (Figure 14A, Appendix A; $N_m = 0.015$). The Colorado River dace population is partitioned from the California populations by even greater geographic distance, the Mojave Desert, the Central & Eastern Transverse Ranges, and the Peninsular Range. The expectation would be for gene flow among the 4 regions, including the Colorado River, to be even lower than the flow among the 3 California regions. However, in the population analysis the gene flow for the 4 region model including the Colorado River was observed to be higher (Figure 14B, Appendix A; $N_m = 0.037$). This supports the hypothesis that the Santa Ana Speckled Dace was derived from the Colorado River population, which is in agreement with the conclusion of Dowling et al. (2008) and Oakey et al. (2004).

Principal coordinates analysis (PCoA) supports that the Southern California Santa Ana speckled dace is genetically distinct from the Central Coast and Eastern Desert dace populations. This further verifies that the Southern California Santa Ana Speckled Dace population is unique from dace populations found in the California Central Coast, Eastern Desert, and the Colorado River. The results of a Mantel test (Figure 17-18, Appendix A) indicates that the Santa Ana Speckled Dace differs genetically from the other dace populations due to isolation by distance from the other populations ($R^2 = 0.347$, $p = 0.002$). This has implications relating to conservation because the Southern California dace habitat has been affected by drought, fire, flood, and increasing human activity within Southern California through conversion of wilderness areas to residential and industrial use (Los Angeles County Watershed Management, 2008).

The estimated divergence time between the Santa Ana Speckled Dace and Colorado River dace is three million years. The relative divergence time among each region ranged from 3.0% to 7.0% (Figure 11, Appendix A), using the molecular clock calibrated for cytochrome b gene in cyprinid fishes (Pfrender et. al. 2004 and Dowling et. al. 2008).

Population Genetics

The AMOVA analysis of cytochrome *b* showed significant genetic structure among populations of *Rhinichthys osculus* when partitioned based on a three-region geographic model (Southern, West Coast, Eastern Desert

California). When partitioned by 3 geographic regions, the Φ_{RT} value among groups was .92 with percent total genetic variation of 92.22% attributable to differences among regions (Table 8, Appendix B). This further verifies the results found in the phylogeographic analysis. The results among populations within groups and within populations showed that there was gene flow within each geographic region, but little to no gene flow between populations separated by large geographic distances. This concurs with the results of Dowling et al (2008) and Oakey et al (2004) which found the populations of *Rhinichthys osculus* in California were distinct from the populations found in the Colorado River and Oregon, respectively. Data partitioning of California populations based on mountain ranges and watersheds showed similar results, which further verifies that genetic variation within the speckled dace correlates to geographic location, with little to no gene flow among geographic regions (Figure17, Appendix A; N_m = 0.015). This further illustrates that the Southern California Santa Ana Speckled Dace population constitutes a genetically distinct and unique population of *R. osculus*.

The minimum spanning haplotype network (Figure 12, Appendix A) identified an ancestral haplotype H5 that contained 18 speckled dace individuals from City Creek, Cajon Creek and Lytle Creek that are all contained within the Santa Ana River watershed. Other haplotypes surrounding the ancestral haplotype contain dace individuals from Plunge Creek, Twin Creek and Lytle Creek and form a star-shaped pattern, which is one indication of an expanding population (Cooke et al., 2012). The three Indian creek haplotypes H1-H3 located in the San Jacinto River differ from the ancestral lineage H5 by at least 11 mutational steps, creating a distinct population from the ancestral haplotype and other closely related haplotypes of the Santa Ana River. Haplotype H9 contains five San Gabriel River individuals and two Twin creek individuals. Haplotype H9 differs from the ancestral lineage by 11 mutational steps and represents a semi-heterogeneous population of individuals from the Santa Ana and San Gabriel watersheds, presently disconnected but with apparent historical gene flow between the Santa Ana and San Gabriel River populations (Figure 12, Appendix A).

Mismatch distribution analysis (Figure 21A, Appendix A) indicates that the Southern California population does not fit a model of expectation of sudden population expansion (Figure 21A, Appendix A), but does fit a model of spatial expansion with constant deme size. This correlates with the minimum haplotype spanning network (Figure 12, Appendix A), where the star-like patterns observed for the Southern California haplotypes are consistent with an ancestral population spreading to other areas (Cooke et al., 2012).

Neutrality tests (Table 7A and 7B, Appendix B) indicate that the Santa Ana Speckled Dace do not significantly differ from a model of neutral evolution, with populations evolving according to the standard coalescent model (Tajima, 1997). A large negative value for Fu’s F_s (Table 7B, Appendix B) suggests historical population growth for the Southern California clade (Fu, 1997), consistent with the star-like haplotype network (Figure 12, Appendix A) and the results of the mismatch analysis (Table 7B, Appendix B). Population genetic data correlates with the phylogenetic analysis.

This illustrates how populations of the Santa Ana Speckled Dace have spread through different watersheds, and subsequently become isolated due to habitat fragmentation that resulted in gene flow restrictions (Figure 12, Appendix A). Once isolated, speckled dace populations began to diverge through neutral evolution through drift. These results are consistent with the analysis described for the California tree frog, *P. cadaverina*, which shares some of the same habitat locations as the Santa Ana Speckled Dace (Phillipsen and Metcalf, 2009). Phillipsen and Metcalf (2009) suggested that landscape habitats are important factors that affected the population structure of stream-dwelling vertebrates such as *P. cadaverina*.

Molecular Evolution of Cytochrome b gene

Analysis using the protein-modeling software showed that Cytochrome b gene is supporting the current structure-function model for membrane spanning protein (Irwin et al 1991). The Chi-square (X^2) analysis compared the observed nucleotide and amino acid substitutions found in the cytochrome b (nucleotide and amino acid) sequences and domain partitions with the expected calculated nucleotide and amino acid substitutions. The null hypothesis is that there is no difference found between the observed and expected nucleotide and amino acid substitutions and domain partitions in the cytochrome *b* gene sequences. The average nucleotide substitution is occurring in the 3rd codon position ($X^2 = 125.2$, $P<6.53E^{-28}$) and are synonymous substitutions where the mutations will not change the amino acid drastically that the protein conformation will also change (Broughton, et al., 2006) (Table 5, Appendix B). Because of these findings, the structure-function model illustration is similar to the illustration created by Irwin et al 1991 that showed hypervariable residues occur in the transmembrane segments rather than the outer surface of the membrane ($X^2 = 6.598$, $P<0.0102$). The Hydrophobic-Hydrophobic Amino Acid Substitutions was higher in the transmembrane than the inner-membrane ($X^2 = 4.804$, $P<0.0284$). The results showed that hydrophobic-hydrophobic amino acid substitutions are better suited for the transmembrane than the inner membrane where there is more interaction with solution molecules found in the cytoplasm or in the inner membrane space. Illustration of cytochrome *b* gene model showed that amino acid changes within the trans-membrane did not change the conformation of the protein. The predicted model showed that the predominant 3-class secondary structure was a helix (Figure 23A, Appendix A), which was congruent to Irwin et al., 1991 findings. Even though parts of the protein sequence was highly exposed in intra/inter membrane, the changes in amino acid did not change the conformation of the protein (Figure 22, Appendix A). This illustrates that the evolution of cytochrome *b* can be used in phylogenetic studies because observed synonymous and non-synonymous substitutions did not affect protein conformation to the extent that the protein is non-functional. Next steps will be to simulate mutations or substitutions that will render the protein non-functional and analyze the constraint of the protein before rendered non-functional. With the use of the evolutionary rate of cytochrome *b* gene, a molecular clock can provide an estimate of the date of divergence from the ancestral lineage. With the help of a protein-modeling program, a protein model of the cytochrome *b* was created to illustrate map locations of the conserved and non-conserved regions based on amino acid substitution locations. The results of this study contribute to describing the taxonomic identity of the Santa Ana Speckled dace as a distinct taxon (perhaps species level) and its evolutionary significance.

Implications for Conservation and Management

Due to habitat fragmentation primarily attributed to human activity including introduction of non-native species, water diversions, urbanization of watersheds, and other human factors, the number of speckled dace found in Southern California have declined (Swift et al. 1993). The Santa Ana Speckled dace was listed as a species of special concern by the California Department of Fish and Game in 1995 and as a sensitive species by the U.S. Forest Service in 1998. From 2004-2005, California suffered wildfires and flooding which resulted in further compromising the speckled dace habitat. A recent Forest Service survey (SAWPA 2004) showed that in the Santa Ana watershed the population of speckled dace is now limited to three tributaries of the Santa Ana River that include Lytle Creek, Cajon Creek, and Plunge Creek (SAWPA 2004; Santa Ana Speckled Dace Recovery Project 2005). Twin and City Creek surveys indicated that the speckled dace populations in those locations were extirpated (Santa Ana Speckled Dace Recovery Project 2005). The extirpation of Santa Ana Speckled Dace populations would possibly move the Santa Ana Speckled Dace to federal listing especially if it can be shown that there is significant genetic differentiation existing between the populations inhabiting the Santa Ana Watershed (Waples 1995) and those of the San Gabriel, Santa Maria, San Luis Obispo, and the Owens River watersheds. Specific management strategies for the remaining watershed populations and drainage areas, with the goal of conservation of the species, may be suggested once there exists certain knowledge regarding the degree of genetic variation among the remaining fish populations (Moritz 1995). The management strategies might include captive breeding programs or species reintroductions into locations historically populated by dace. This will then contribute to preserving the genetic variability of the Santa Ana Speckled Dace as a separate taxa and species, as well as help maintain intact Santa Ana Speckled Dace populations in different creeks. Mitochondrial DNA itself is not sufficient to suggest within region reintroduction strategies.

APPENDIX A

FIGURES

Figure 1. *Rhinichthys osculus* California Range Map. Areas bordered in dark are known to contain *Rhinichthys osculus* within river drainages of that region. Watershed sampling locations are denoted in red numbered circles. 1=Colorado River; 2=Lytle Creek/Santa Ana River; 3=San Gabriel River; 4=Santa Maria River; 5=San Luis Obispo River; 6=Owens River.

Figure 2. *Rhinichthys osculus*, the Santa Ana Speckled Dace (Barrett, P. 2003. USFWS).

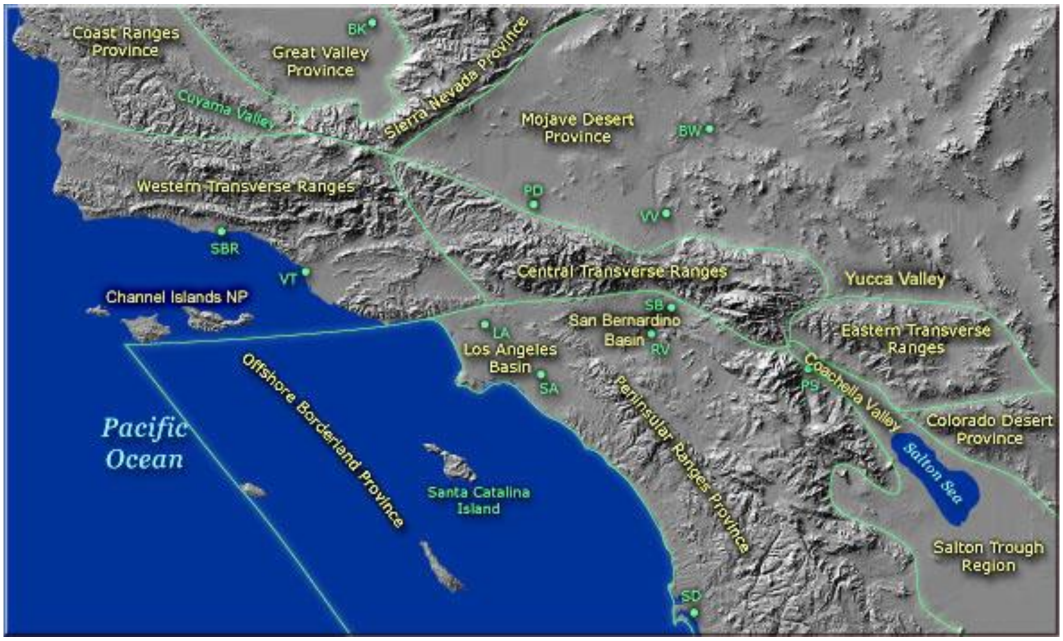


Figure 3. Spatial location of the Transverse and Peninsular Ranges.

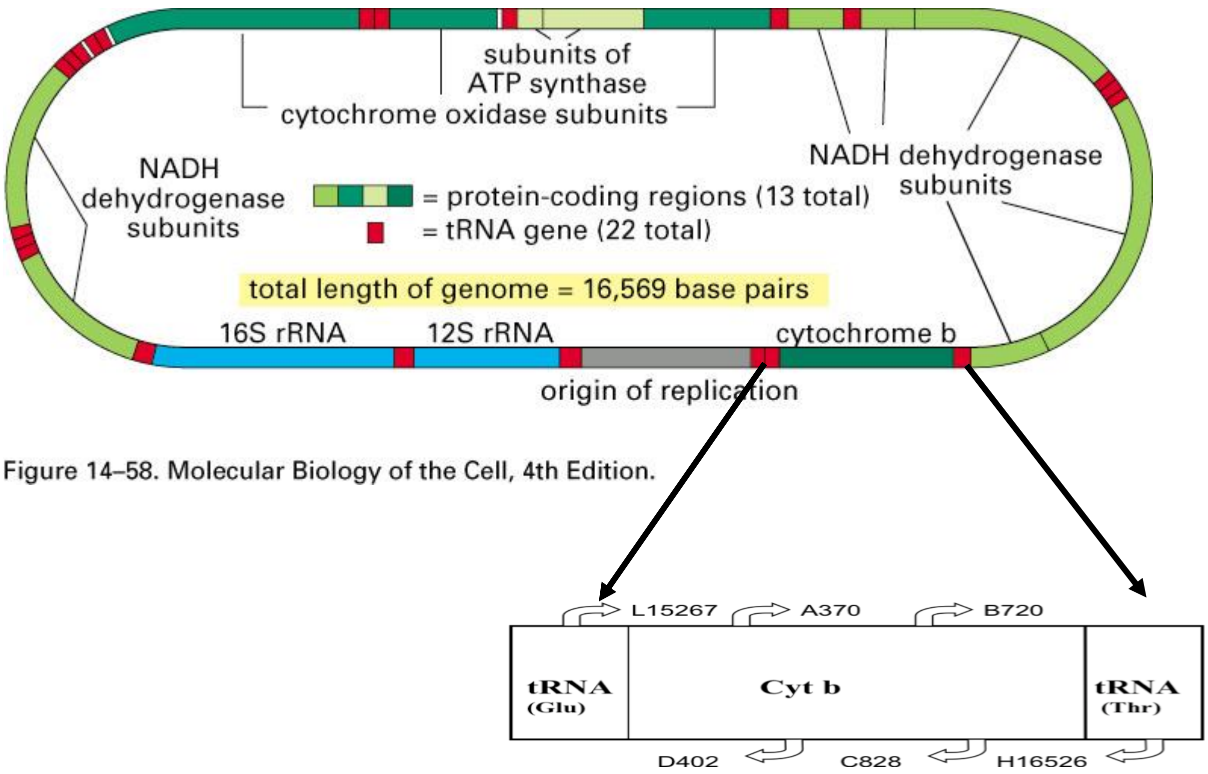


Figure 14–58. Molecular Biology of the Cell, 4th Edition.

Figure 4. Structure of Vertebrate Mitochondrial DNA. The coding mtDNA cytochrome b gene is shaded in dark green including the tRNA glutamine (Glu) and tRNA threonine (Thr) ends shaded in red. (from Molecular Biology of the Cell, 4th ed.)

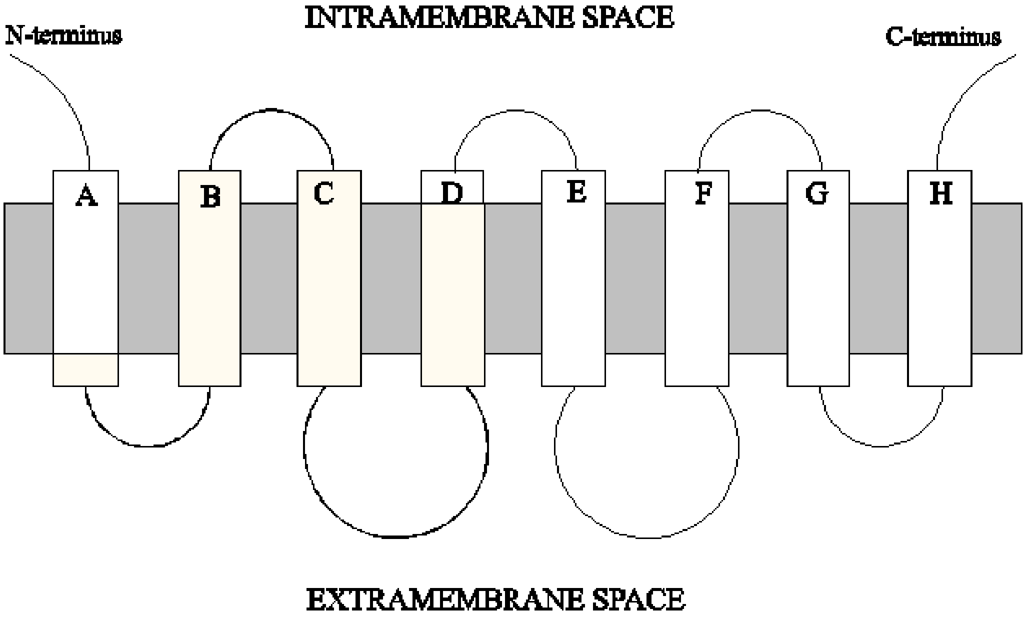


Figure 5: Structure of the cytochrome b protein. The location of 8 trans-membrane helices diagram by letters A-H with extra- and inter-membrane loops including the N and C terminus located in the intra-membrane space (Kvist, L. 2000).

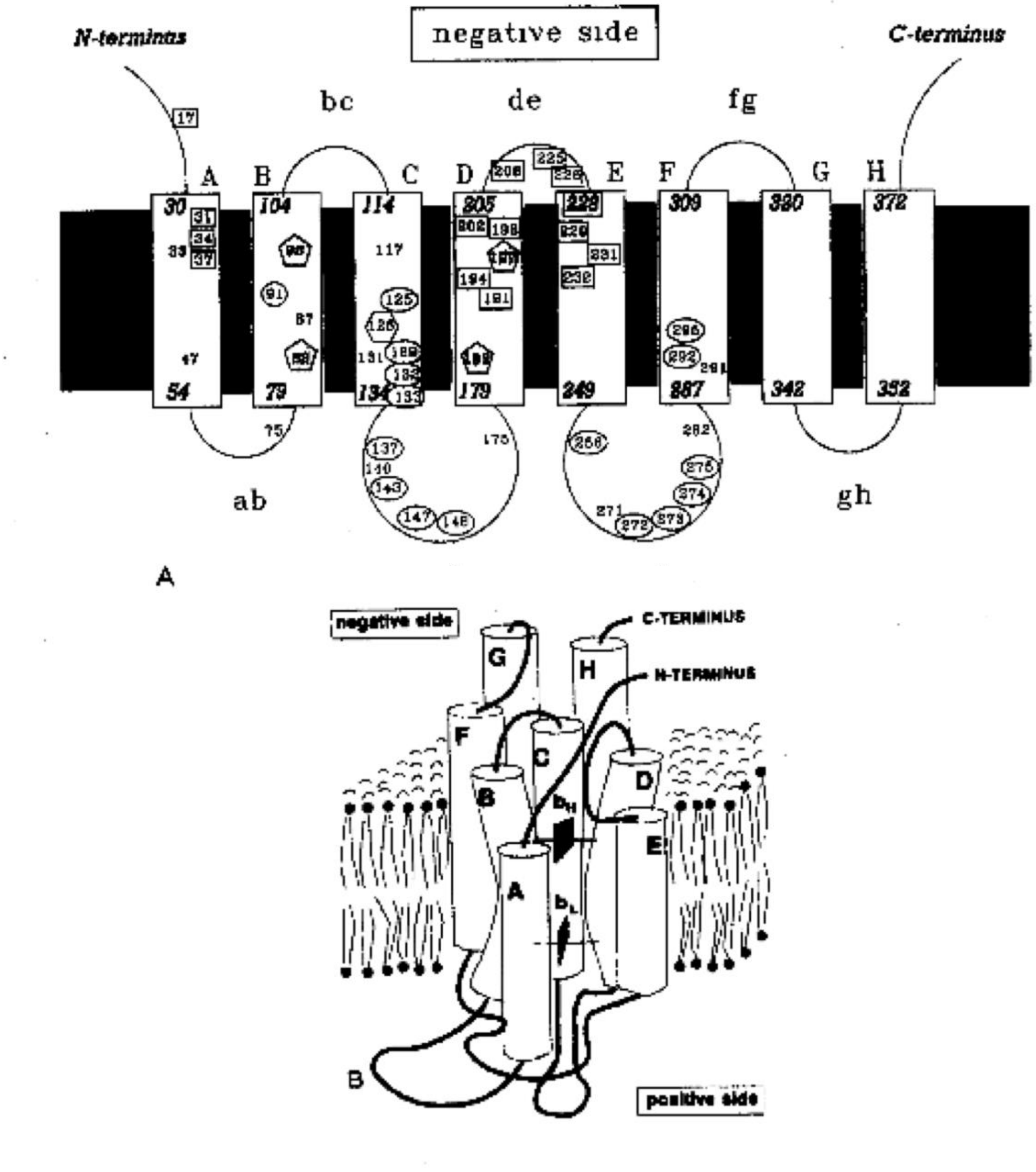
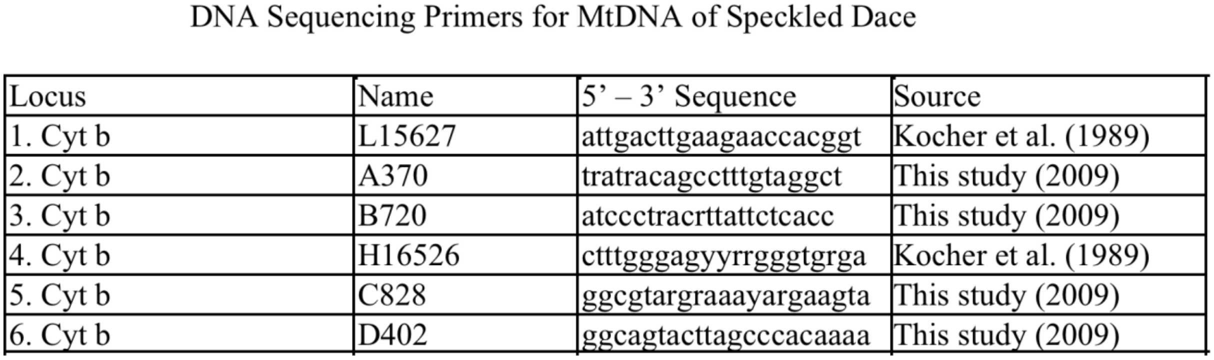


Figure 6A & 6B: Structure of the cytochrome b protein. 6A – Overall view of the positions of key residues in the eight helices model of cytochrome b. 6B- Tentative representation of the overall folding of cytochrome b (Esposti et. al 1993)



DNA Sequencing Primers for MtDNA of Speckled Dace

Locus	Name	5’ – 3’ Sequence	Source
1. Cyt b	L15627	attgacttgaagaaccacggt	Kocher et al. (1989)
2. Cyt b	A370	tratracagcctttgtaggct	This study (2009)
3. Cyt b	B720	atccctracrttattctcacc	This study (2009)
4. Cyt b	H16526	ctttgggagyyrrgggtgrga	Kocher et al. (1989)
5. Cyt b	C828	ggcgtargraaayargaagta	This study (2009)
6. Cyt b	D402	ggcagtacttagcccacaaaa	This study (2009)

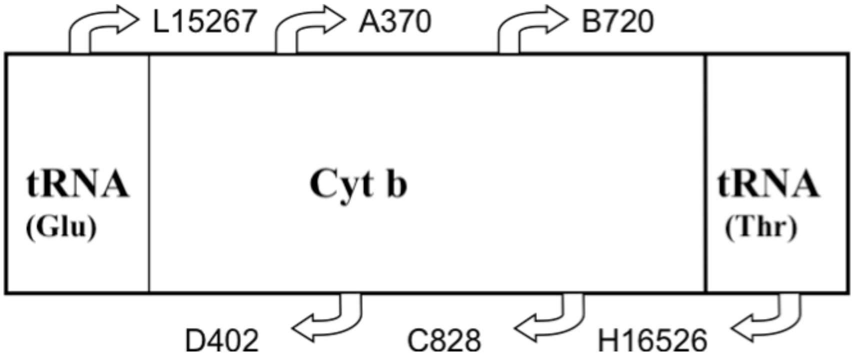


Figure 7. Map of Primer Locations on the Mitochondrial Genome of *Rhinichthys osculus including DNA sequence Primers*. Approximate locations of primers to be used in PCR and sequencing are shown. Forward primers are depicted above and reverse primers below the diagram. Figure is not shown to scale.

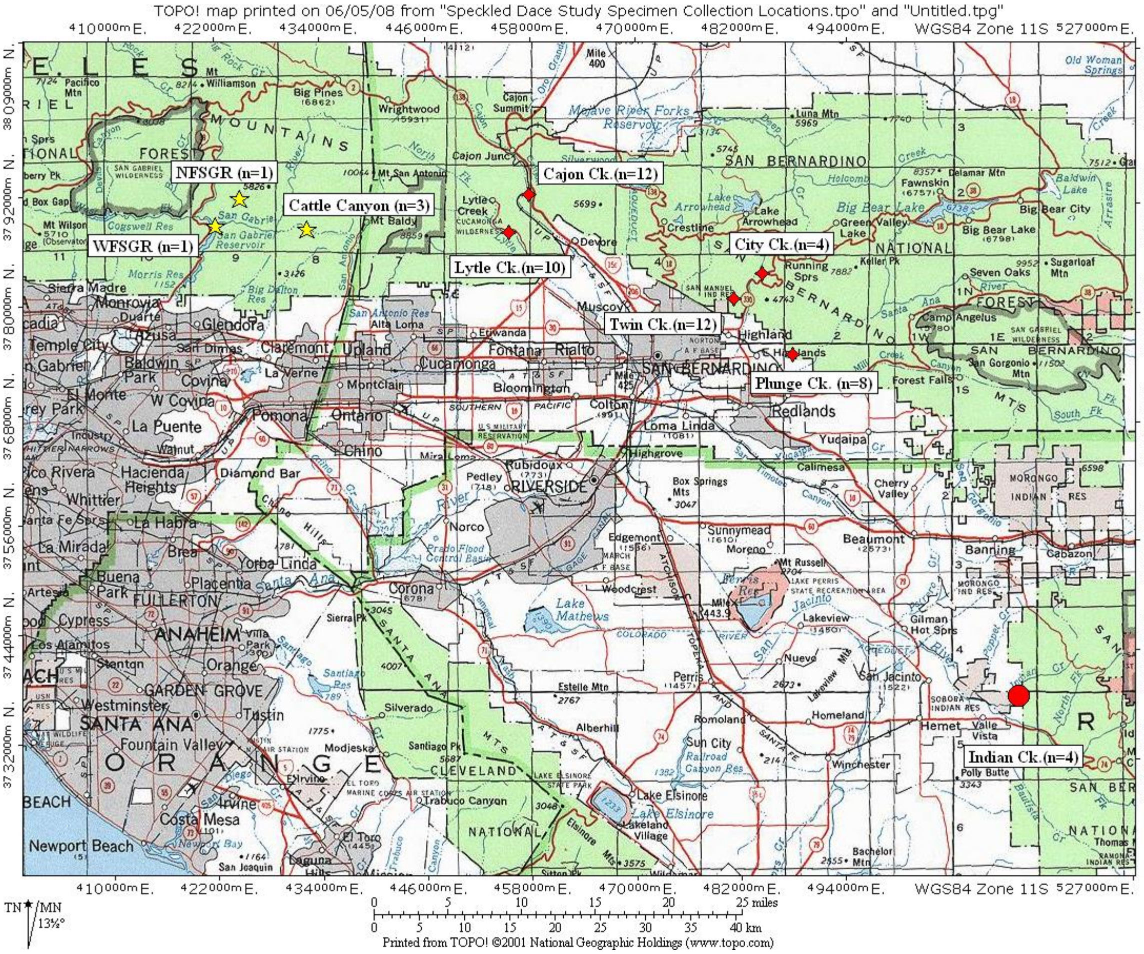


Figure 8A. *Rhinichthys osculus* sampling sites in the Santa Ana River (SAR) and San Gabriel River (SGR) Watersheds. SAR sites = Lytle Creek, Cajon Creek, Twin Creek, City Creek, Plunge Creek, and Indian Creek; SGR sites = North Fork SGR, West Fork SGR, Cattle Canyon Creek.

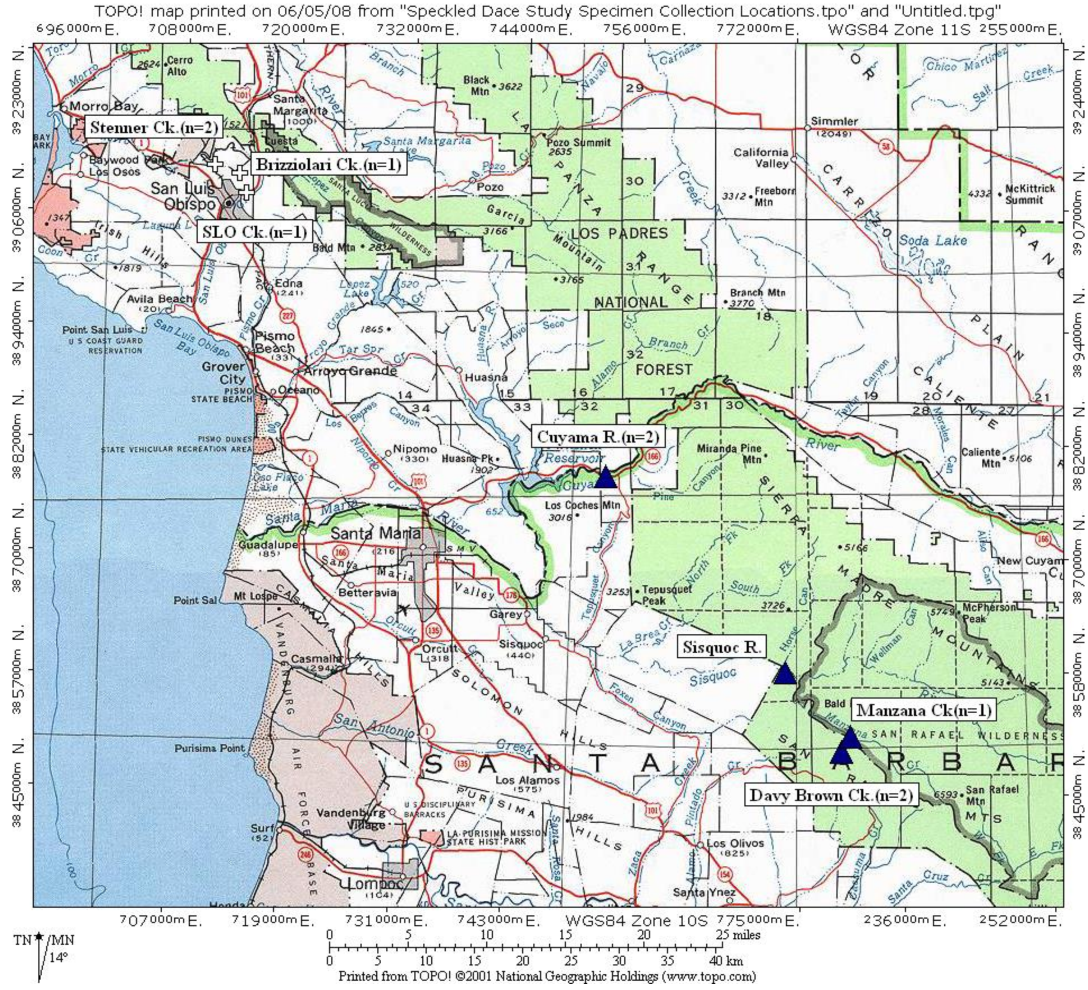


Figure 8B. *Rhinichthys osculus* sampling sites in the San Luis Obispo River (SLO) and Santa Maria River (SMR) Watersheds. SLO sites (white crosses) = San Luis Obispo Creek, Stenner Creek, and Brizziolari Creek; SMR sites (blue triangles) = Cuyama River, Sisquoc River, Manzana Creek, and Davy Brown Creek.

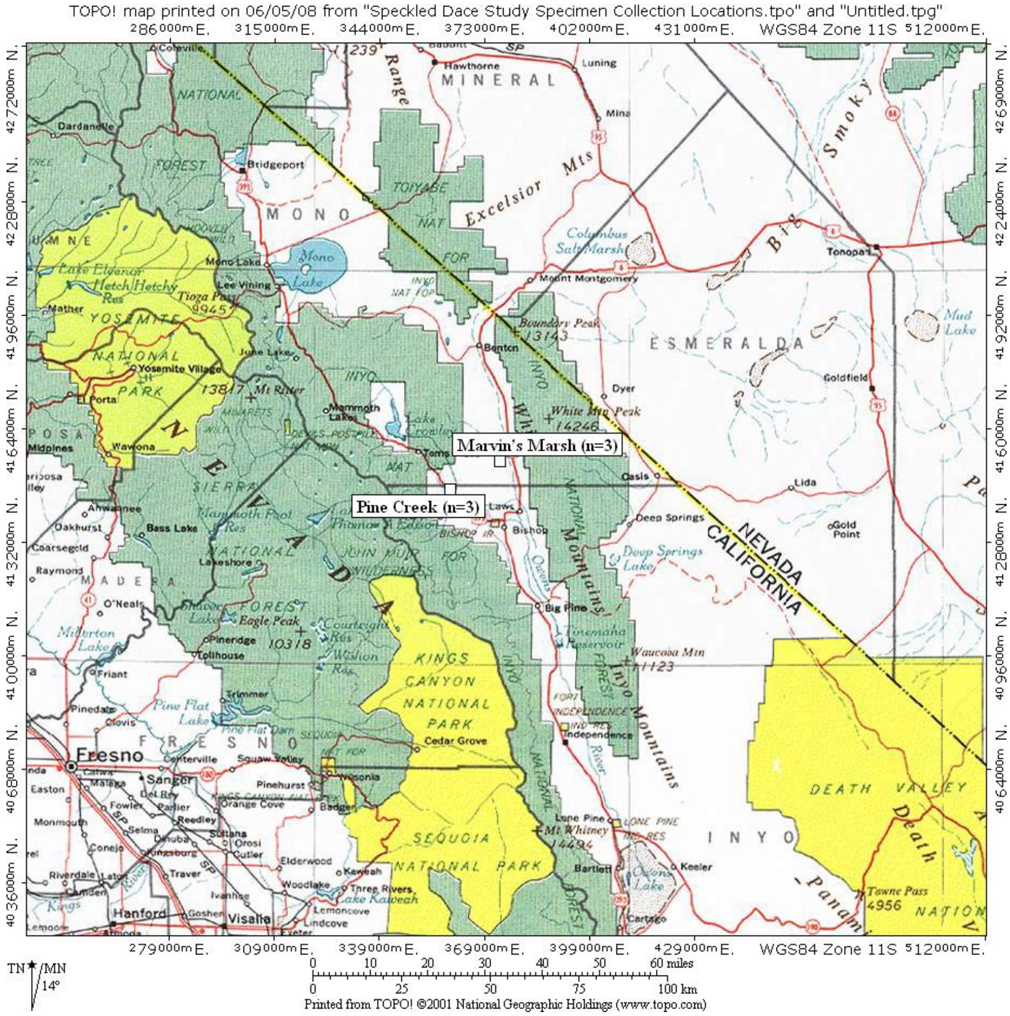


Figure 8C. *Rhinichthys osculus* sampling sites in the Owens River Watershed. Owens River sites (white squares) = Marvin's Marsh and Pine Creek.

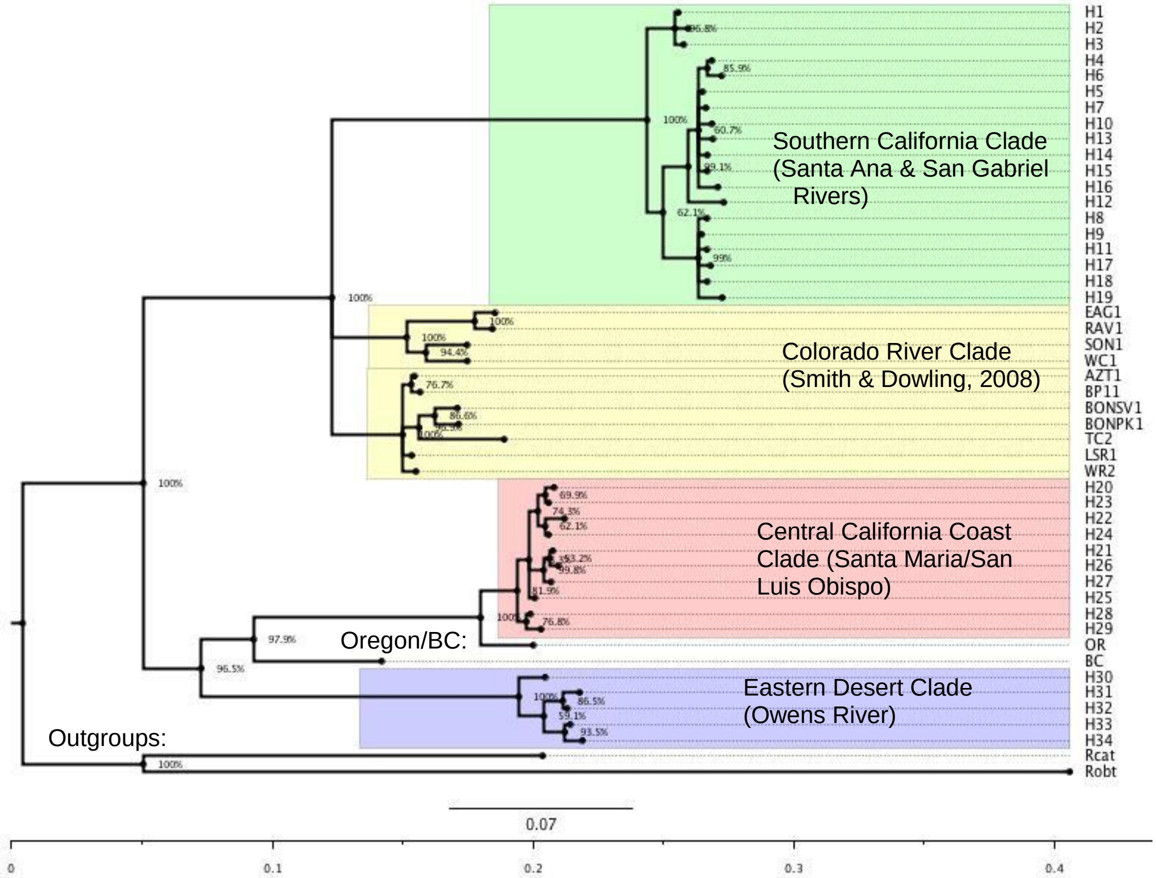


Figure 9: Bootstrapped consensus Maximum Likelihood-based Tree. Like the distance and maximum parsimony analysis, the consensus maximum likelihood tree shows strong bootstrap support for the four clades (Southern California, Colorado River, Central California Coast, and Eastern Desert) with *R. osculus* representatives from Oregon and British Columbia rooted with two outgroups: *R. cataractae* and *R. obtusus*.

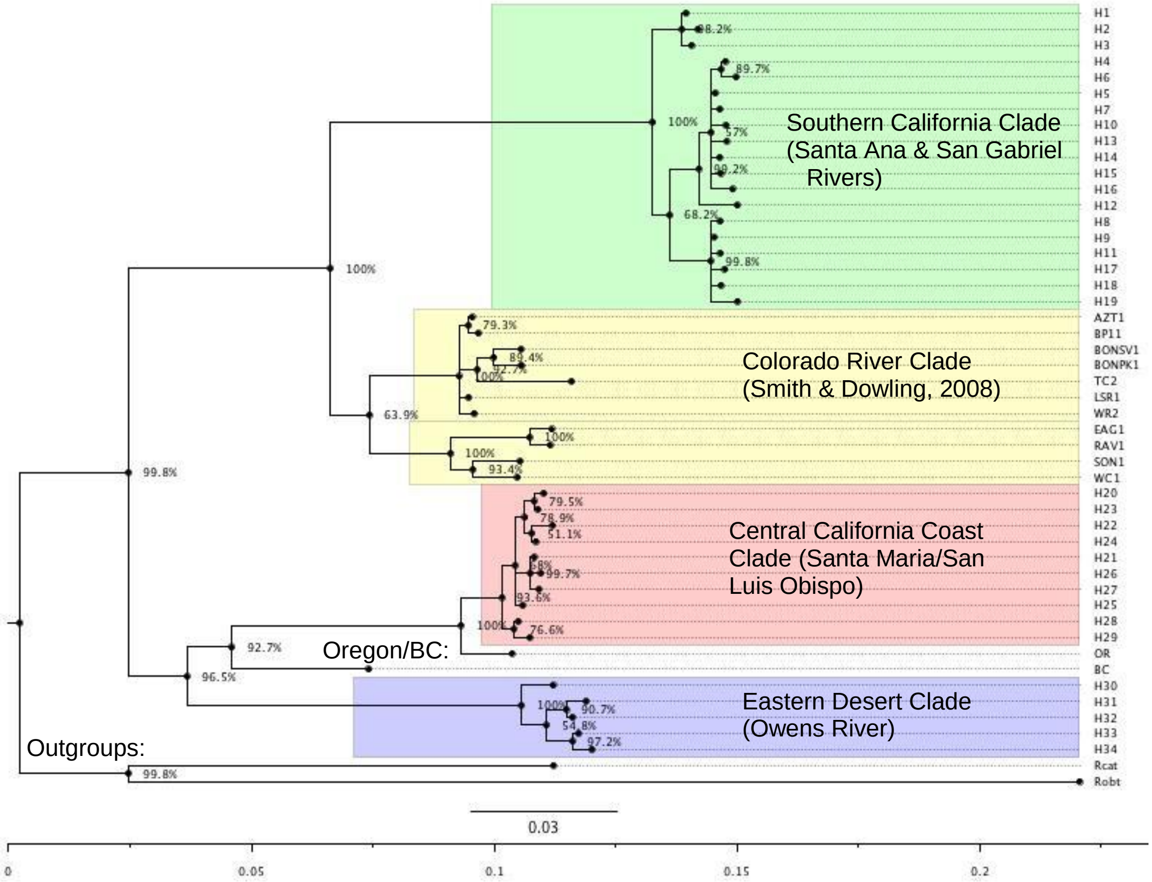
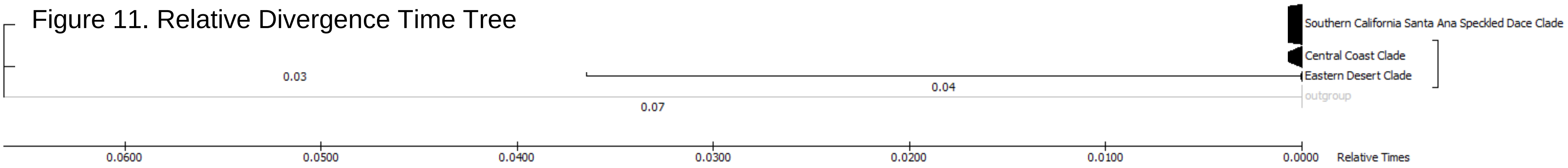


Figure 10: Bootstrapped consensus Bayesian Tree (Nucleotide base, Domain, and codon position). Like the maximum likelihood tree, the consensus Bayesian tree illustrated strong bootstrap support for the four clades (Southern California, Colorado River, Central California Coast, and Eastern Desert) with *R. osculus* representatives from Oregon and British Columbia rooted with two outgroups: *R. cataractae* and *R. obtusus*.

Figure 11. Relative Divergence Time Tree



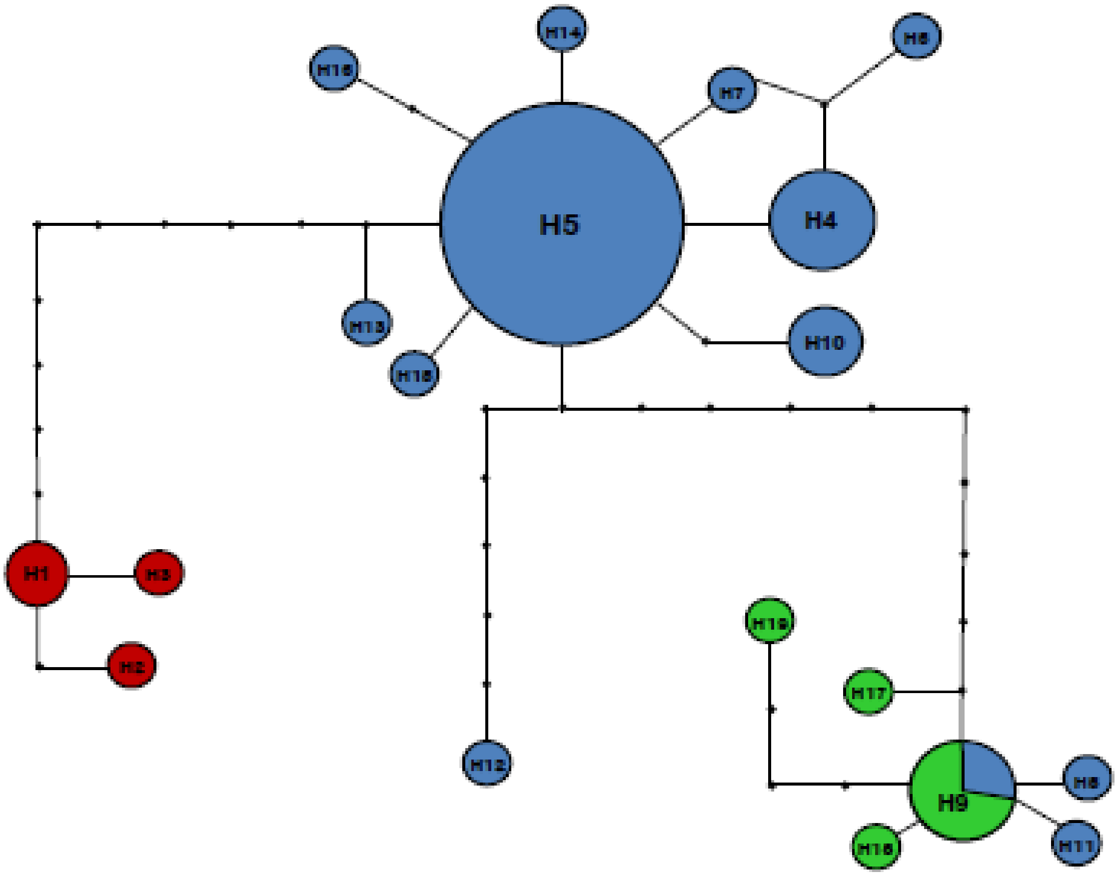
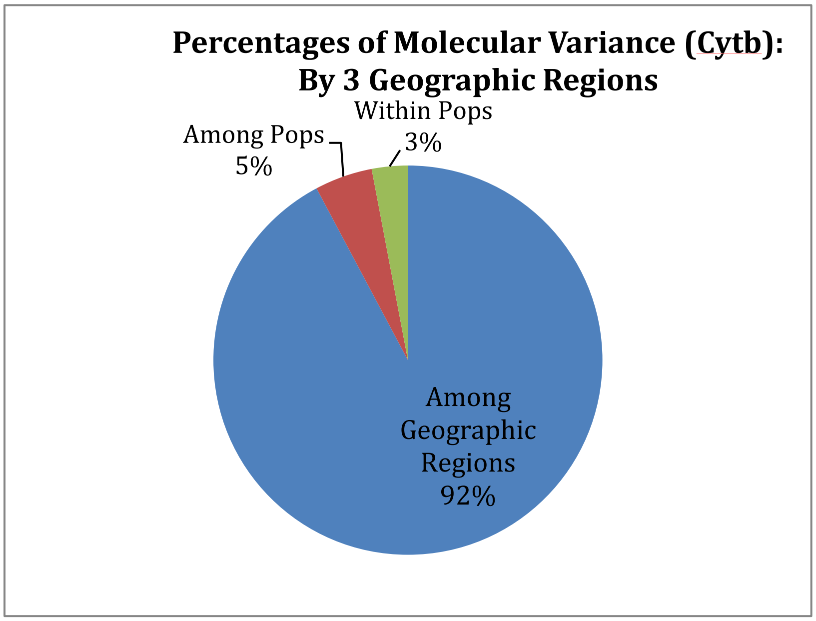


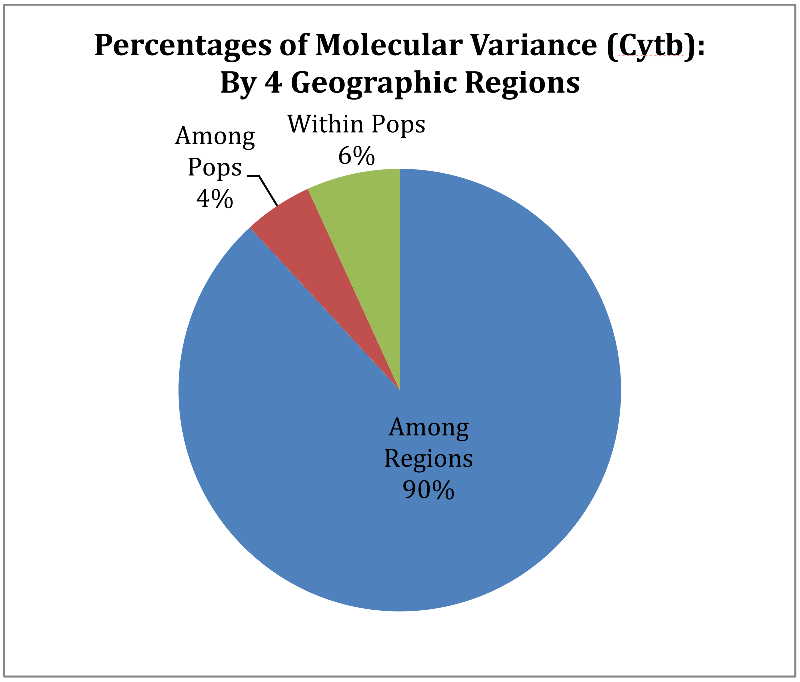
Figure 12: Minimum Spanning Haplotype Network.
Each line represents one mutational step, circle size is relative to the number of individuals with the same haplotype with the largest circle represents inferred ancestral haplotypes, and dots represent unsampled haplotypes. (Red=San Jacinto R., Blue=Santa Ana R., Green=San Gabriel R.)



A. 3 Geographic Regions

Stat	Value	P (rand>=data)
Φ_{RT}	0.922	0.000
Φ_{PR}	0.616	0.000
Φ_{PT}	0.970	0.000

Nm (Haploid) = 0.015

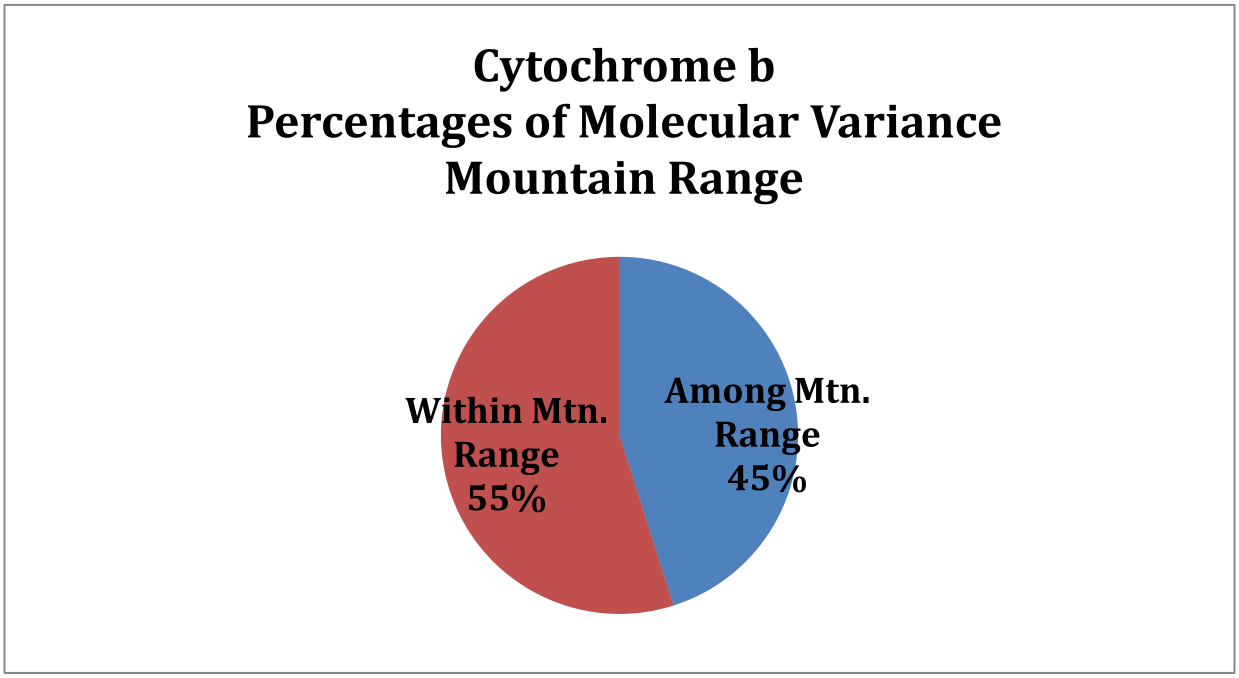


B. 4 Geographic Regions

Stat	Value	P (rand>=data)
Φ_{RT}	0.905	0.000
Φ_{PR}	0.414	0.000
Φ_{PT}	0.944	0.000

N_m (Haploid) = 0.037

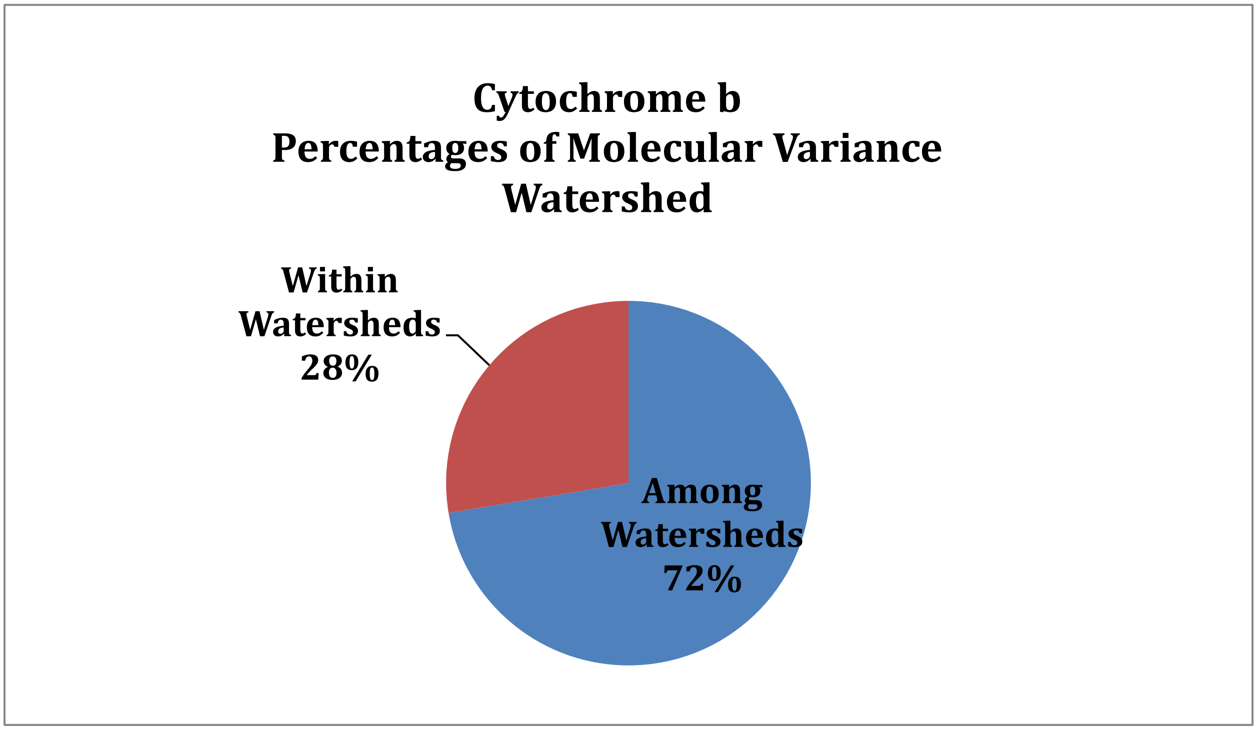
Figure 13A-13B. Cytochrome b Percentages of Molecular Variance Summary. The Model showing the highest percentage of total variation attributable to differences among groups was partition by three & four Geographic Regions (Southern, Central Coast, E. Desert California and including Colorado R.)



San Jacinto Mtns:
Indian Creek
San Bernardino Mtns:
City, Plunge, Twin, and
Cajon Creek
San Gabriel Mtns:
Lytle, Cattle Canyon,
EF, WF, and NF of
San Gabriel River

$\Phi_{PT} = 0.451$
Nm = 0.609
$p = 0.0001$

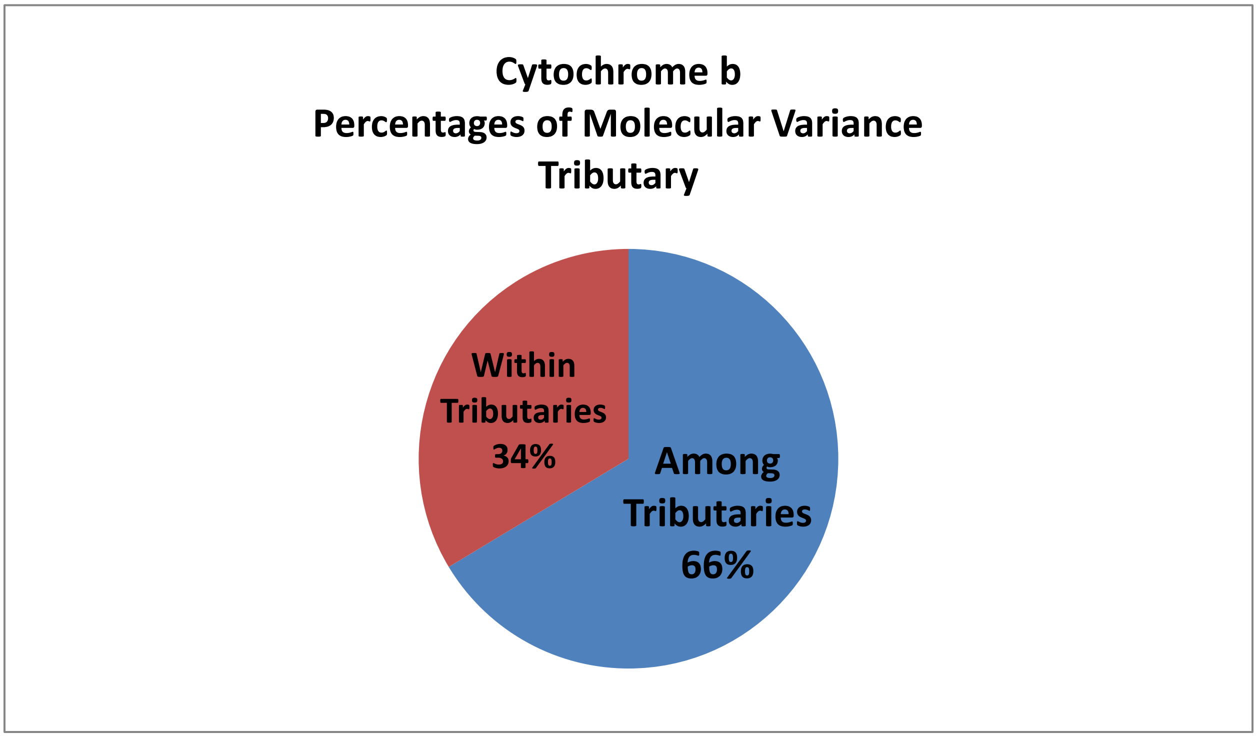
Figure 14. Cytochrome b Percentages of Molecular Variance Summary.
The Model showing the highest percentage of total variation attributable to differences among groups was partition by Mountain Ranges (San Jacinto Mountains, San Bernardino Mountains, and San Gabriel Mountains)



San Jacinto R.:
Indian Creek
San Ana R.:
City, Plunge, Twin, and
Cajon Creek
San Gabriel R.:
Lytle, Cattle Canyon,
EF, WF, and NF of
San Gabriel River

$\Phi_{PT} = 0.724$
Nm = 0.191
p = 0.0001

Figure 15. Cytochrome b Percentages of Molecular Variance Summary. The Model showing the highest percentage of total variation attributable to differences among groups was partition by Watershed (San Jacinto River, San Ana River, and San Gabriel River)



Tributaries:
Indian Creek
City Creek
Plunge Creek
Twin Creek
Cajon Creek
Lytle Creek
San Gabriel River

$\Phi_{PT} = 0.664$
$N_m = 0.253$
$p = 0.0001$

Figure 16. Cytochrome b Percentages of Molecular Variance Summary. The Model showing the highest percentage of total variation attributable to differences among groups was partition by Tributaries.

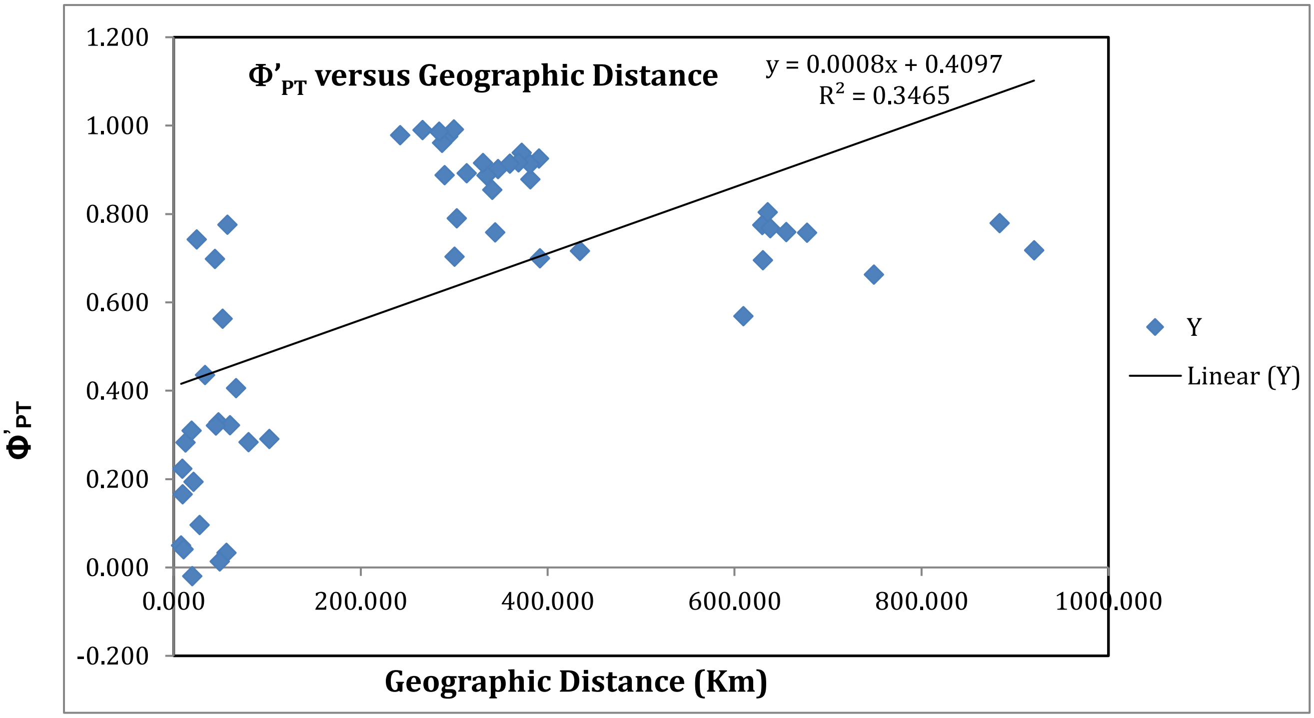


Figure 17. Four Geographic Regions Mantel test cytochrome b gene. Isolation by distance, Φ'_{PT} versus geographic distance of 4 geographic regions (Southern California, Central Coast, Eastern Desert, and Colorado River).

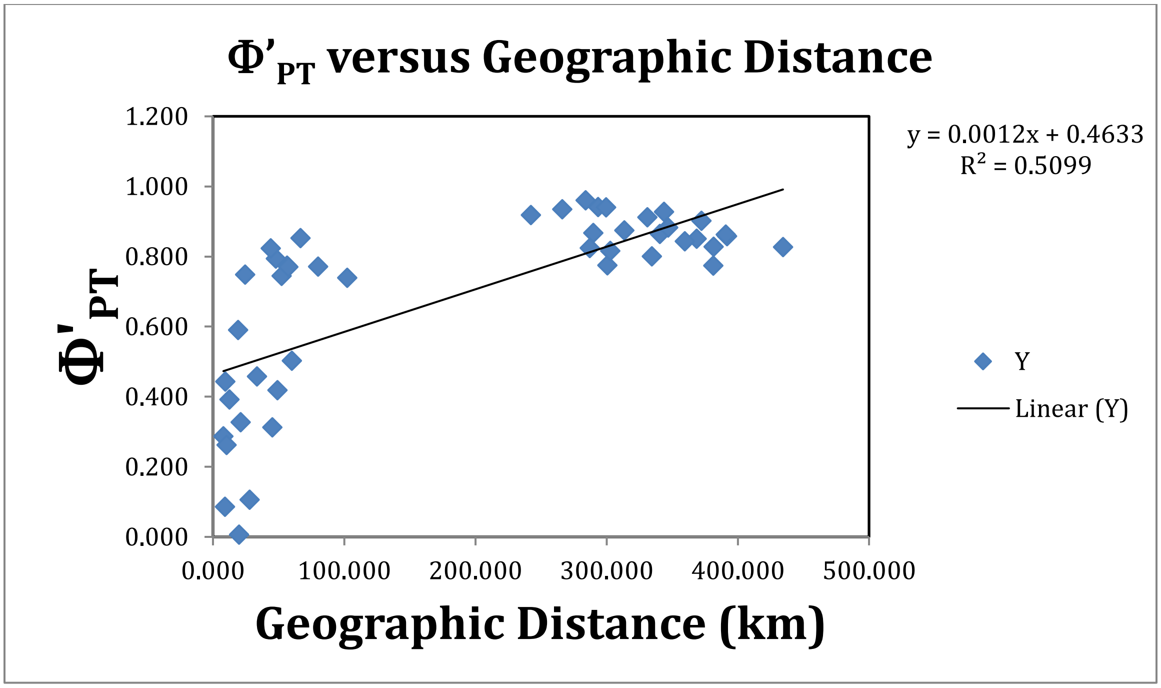
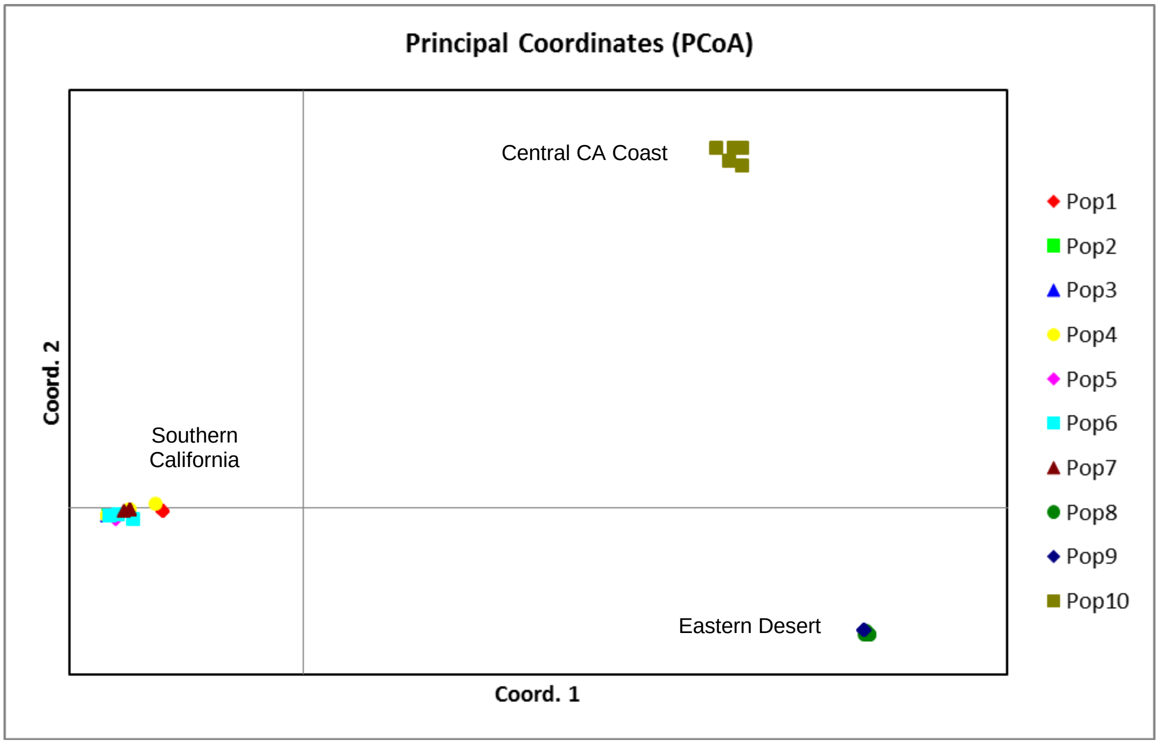


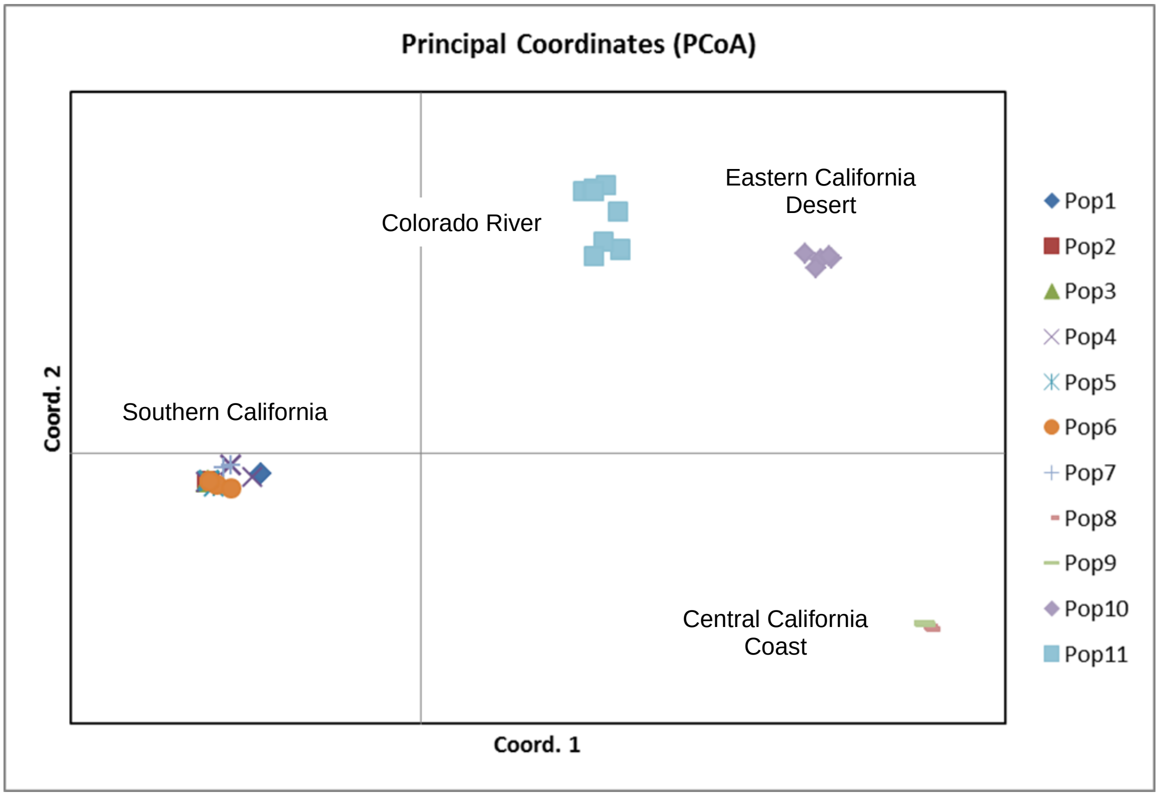
Figure 18. Three Geographic Regions Mantel test cytochrome b gene. Isolation by distance, Φ'$_{PT}$ versus geographic distance of 3 geographic regions (Southern California, Central Coast, and Eastern Desert).



Axis No.	**1**	**2**	**3**
Eigen Value	135.673	34.932	12.553

Figure 19. Principal Coordinates (PCoA) Cytb Analysis Summary.
PCoA analysis is calculated by using 3 regions and geographic distance as coordinates. 3 regions = Southern California, Central California Coast, Eastern Desert; 10 populations = 10 creeks/rivers with total N = 77 individuals.

Population Key: Pop1=IC, Pop2=CC, Pop3=PC, Pop4=TWIN, Pop5=CAJ, Pop6=LC, Pop 7= SGR, Pop8=SLO, Pop9=SMR, and Pop10=OWR



Axis No.	1	2	3
EigenValue	132.394	40.619	27.173

Figure 20. Principal Coordinates (PCoA) Cytb Analysis Summary. PCoA analysis is calculated by using 4 regions and geographic distance as coordinates. 4 regions = Southern California, Central California Coast, Eastern Desert and Colorado R.; 10 populations = 11 creeks/rivers with total N = 88 individuals.

Population Key: Pop1=IC, Pop2=CC, Pop3=PC, Pop4=TWIN, Pop5=CAJ, Pop6=LC, Pop 7= SGR, Pop8=SLO, Pop9=SMR, Pop10=OWR, and Pop11=COR

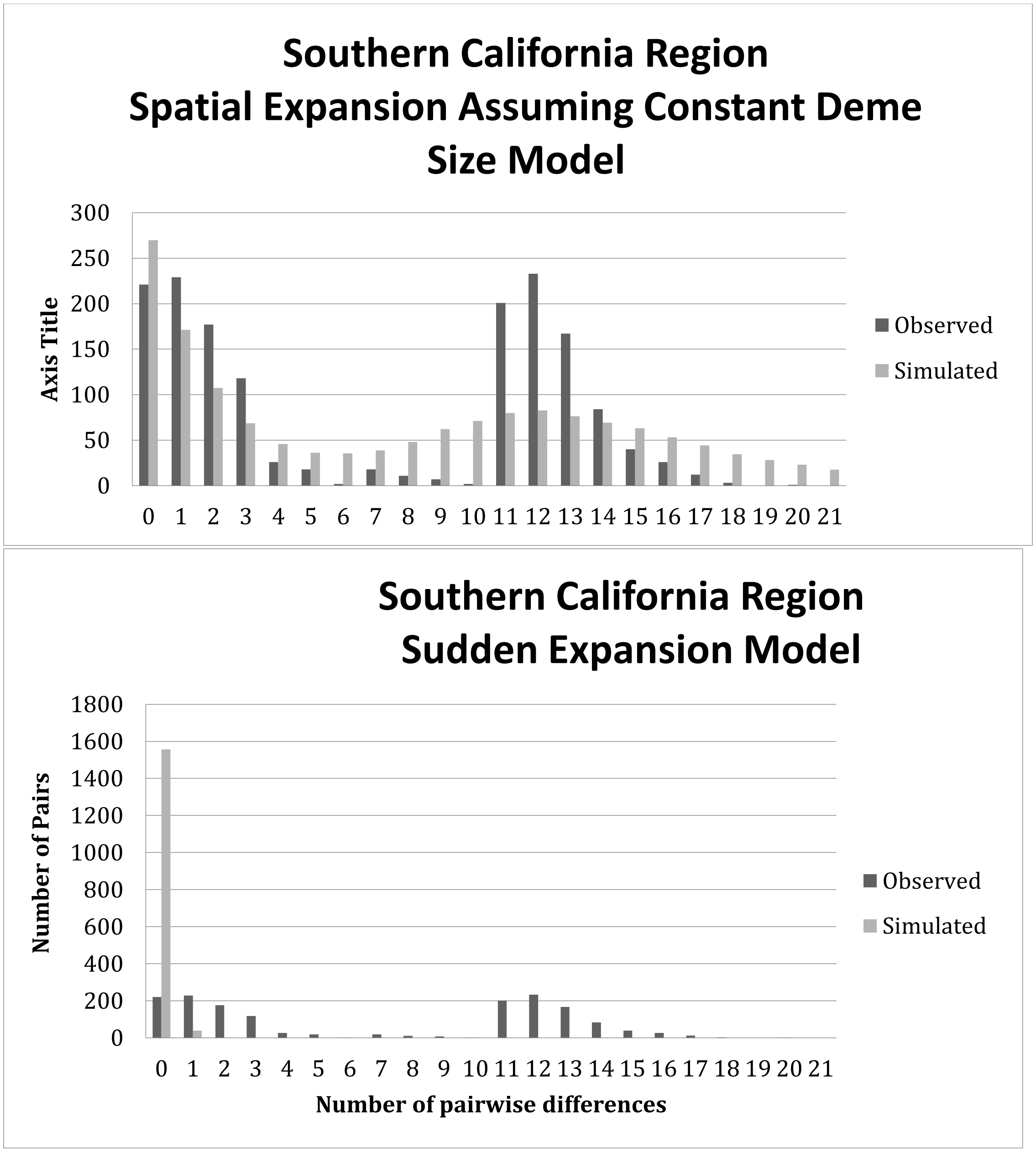
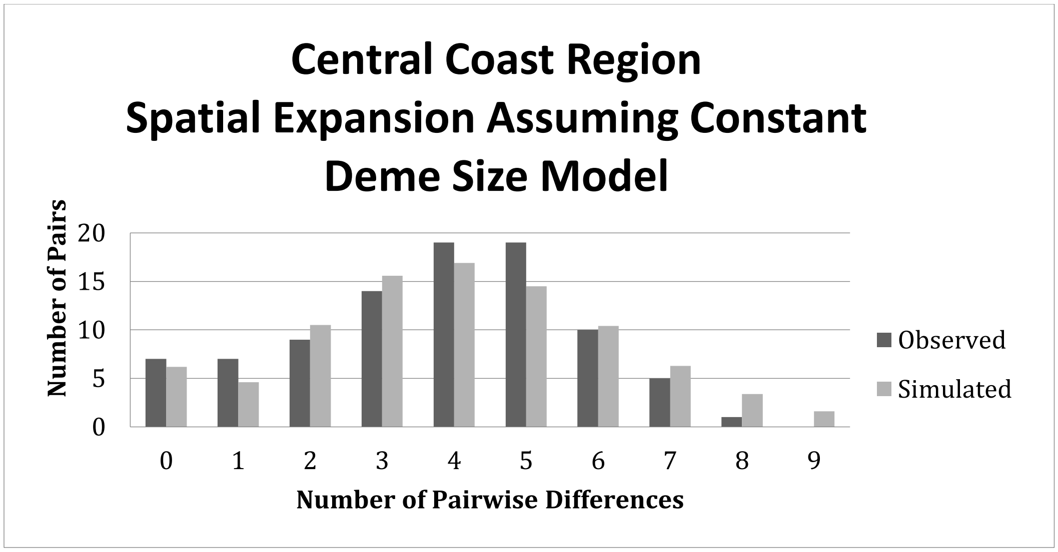


Figure 21A. Southern California Region Mismatch distributions models of pairwise nucleotide differences among *R. osculus* mtDNA haplotypes.



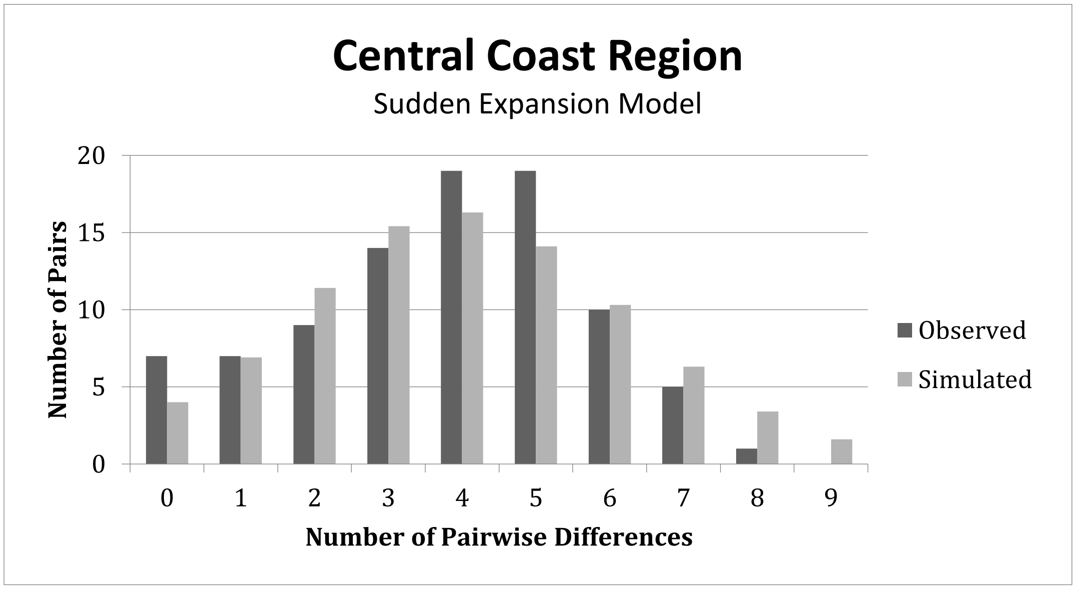


Figure 21B. Central Coast Region Mismatch distributions models of pairwise nucleotide differences among *R. osculus* mtDNA haplotypes.

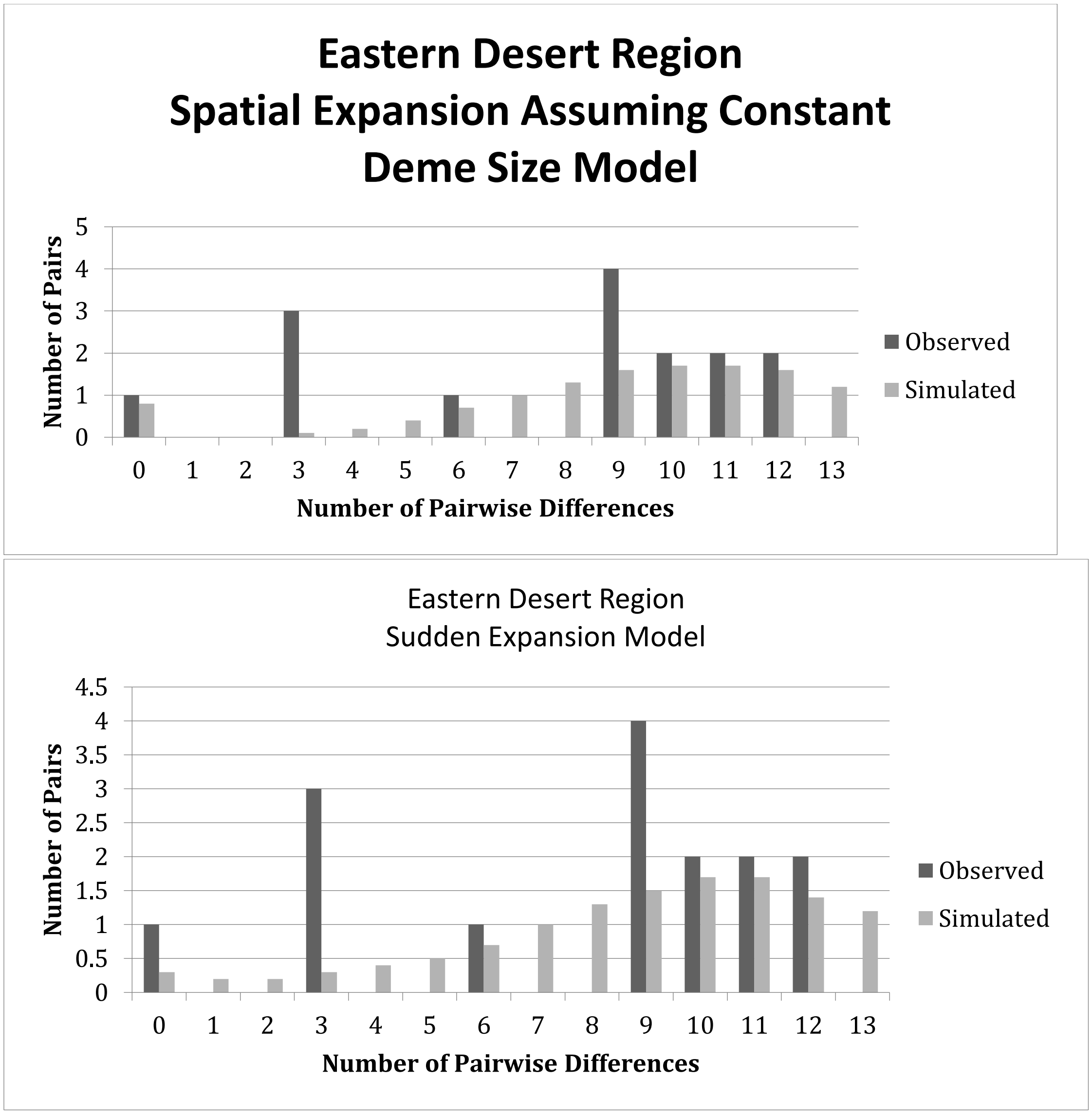
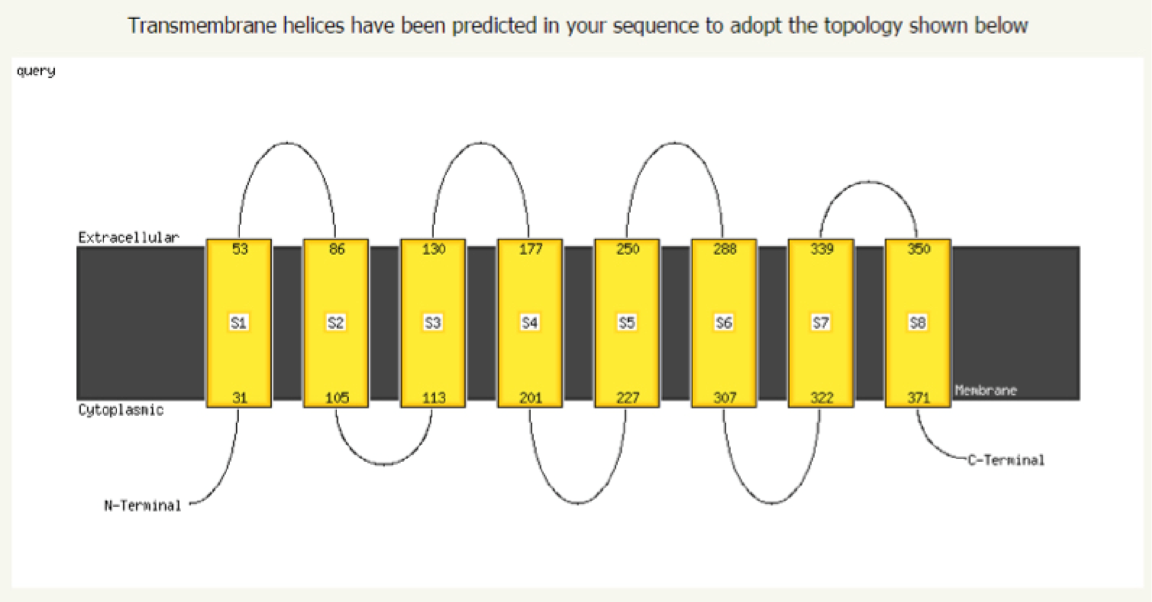
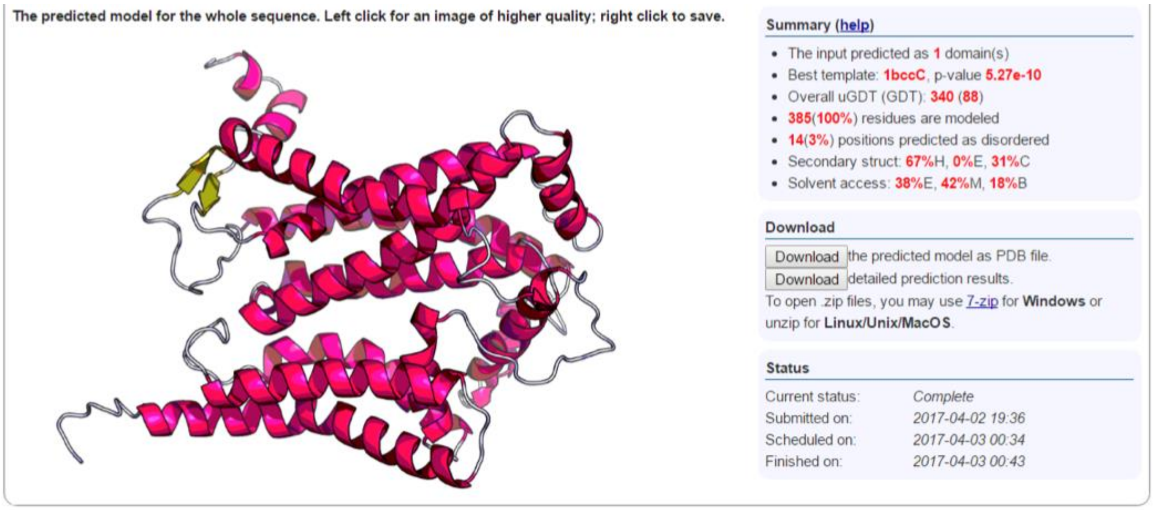


Figure 21C. Eastern Desert Region Mismatch distributions models of pairwise nucleotide differences among *R. osculus* mtDNA haplotypes.



A. Cytochrome b gene trans-membrane helices topology



The predicted model for the whole sequence. Left click for an image of higher quality; right click to save.

Summary (help)

- The input predicted as **1** domain(s)
- Best template: **1bccC**, p-value **5.27e-10**
- Overall uGDT (GDT): **340** (**88**)
- **385**(**100%**) residues are modeled
- **14**(**3%**) positions predicted as disordered
- Secondary struct: **67%**H, **0%**E, **31%**C
- Solvent access: **38%**E, **42%**M, **18%**B

Download

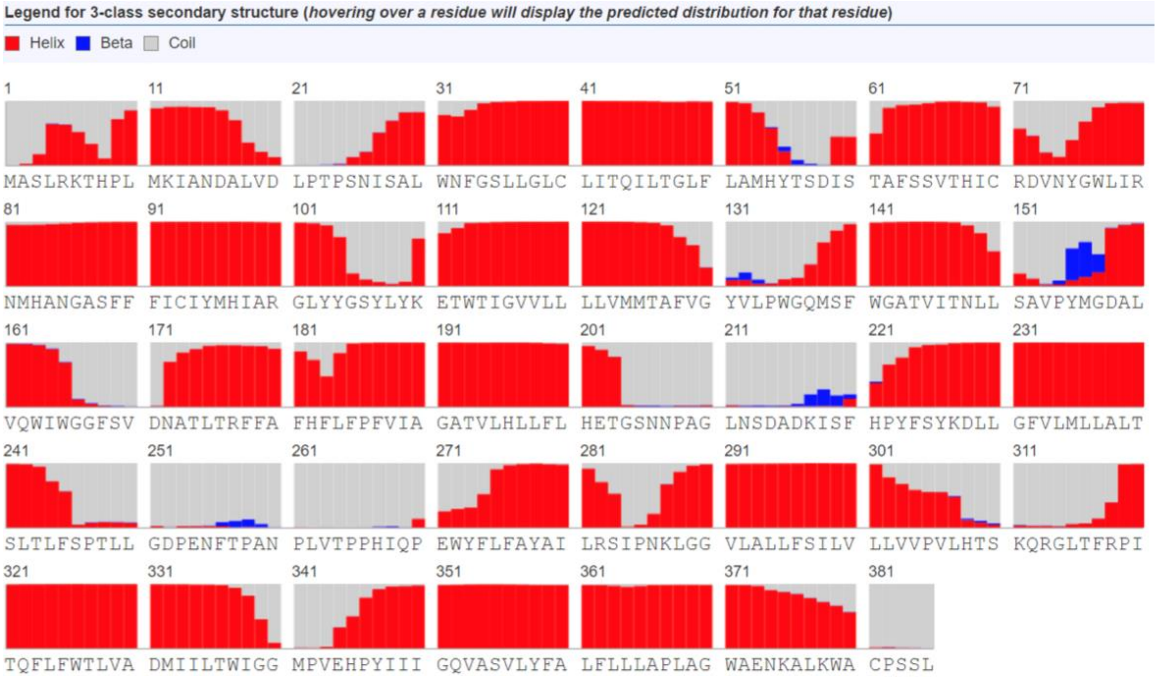
Download the predicted model as PDB file.
Download detailed prediction results.
To open .zip files, you may use 7-zip for **Windows** or unzip for **Linux/Unix/MacOS**.

Status

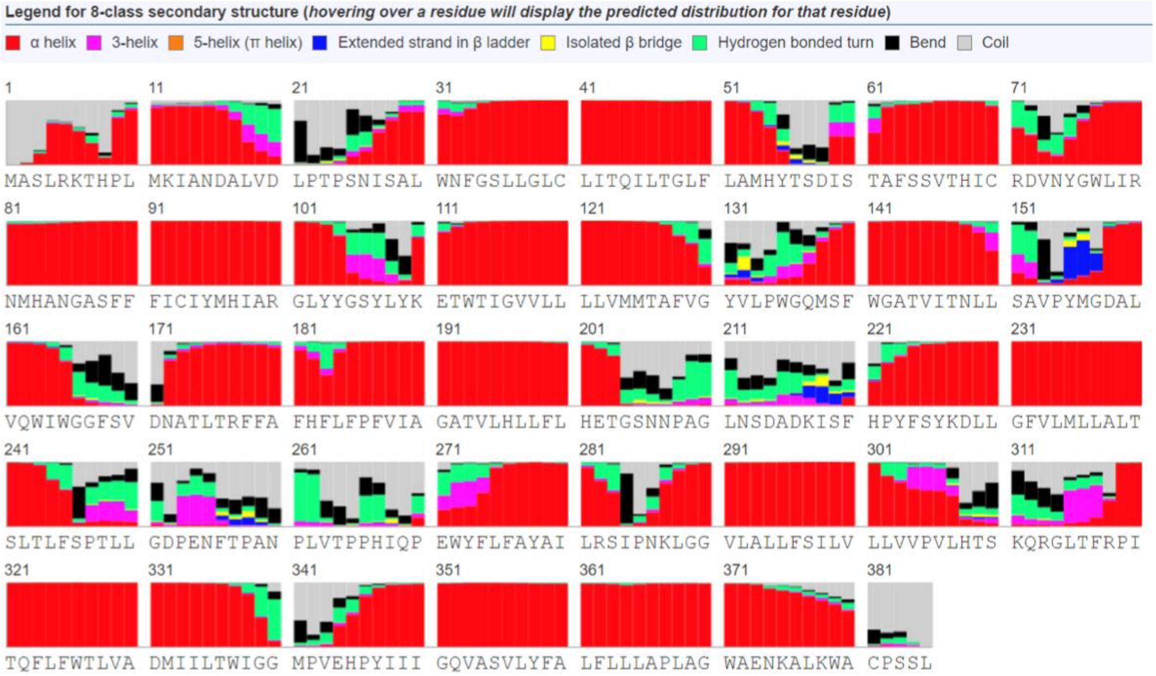
Current status:	*Complete*
Submitted on:	*2017-04-02 19:36*
Scheduled on:	*2017-04-03 00:34*
Finished on:	*2017-04-03 00:43*

B. Cytochrome b gene protein model & summary prediction results

Figure 22A-22B. Cytochrome b gene protein trans-membrane helices topology (A) and protein predicted model & summary results (B).

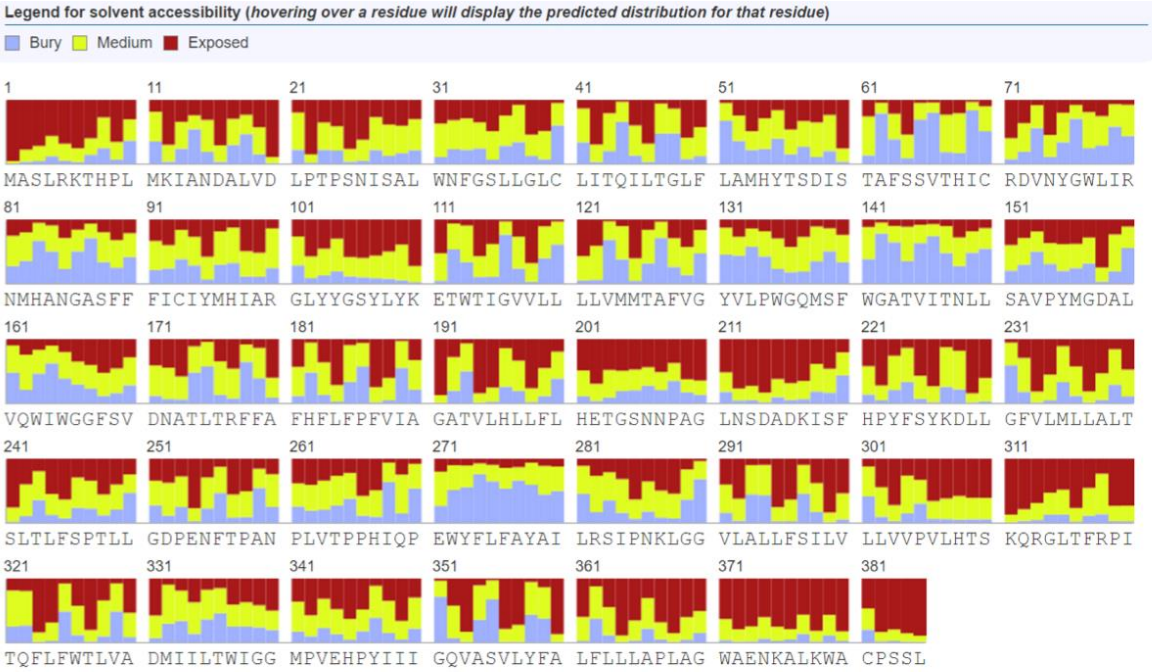


A. Cytochrome b model 3-class secondary structure



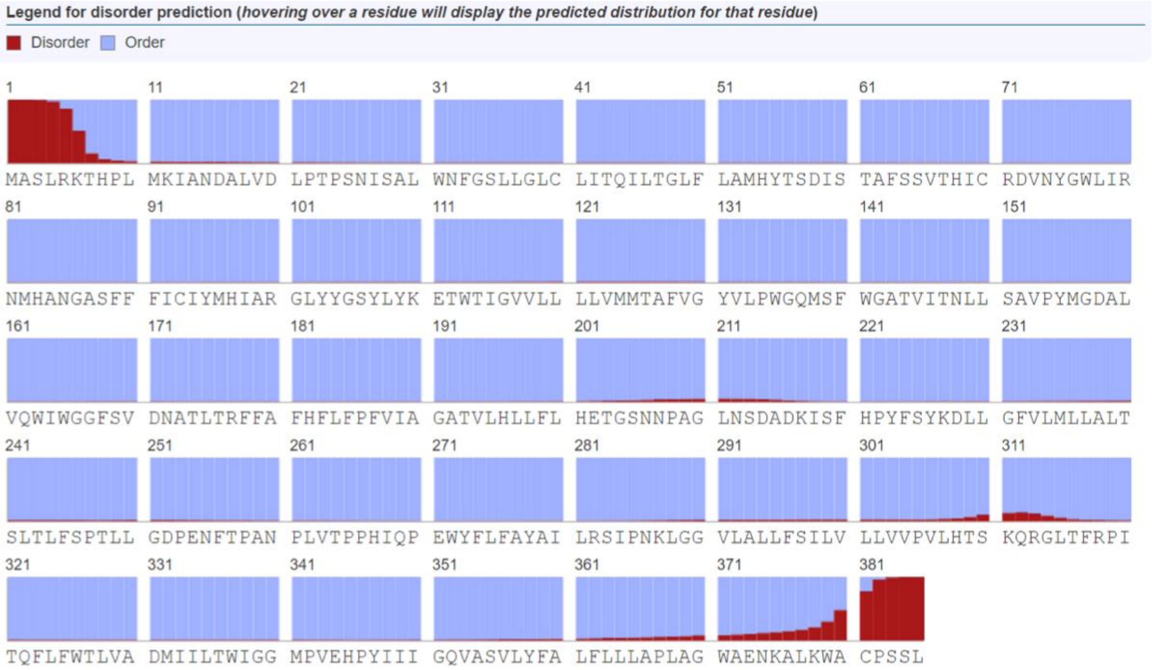
B. Cytochrome b model 8-class secondary structure

Figure 23A-23B. Cytochrome b gene secondary (3-class & 8-class) structure.



Legend for solvent accessibility (*hovering over a residue will display the predicted distribution for that residue*)

A. Solvent Accessibility model



Legend for disorder prediction (*hovering over a residue will display the predicted distribution for that residue*)

B. Disorder Prediction model

Figure 24A-24B. Cytochrome b gene solvent accessibility model (A) & disorder prediction model (B).

APPENDIX B

TABLES

Table 1. *Rhinichthys osculus* sampling locations with sample sizes in Southern California. Mountain Range codes are as follows: San Bernardino (SB), San Gabriel (SG), Coast Ranges (CR), Eastern Sierra Nevada (SN), and Rocky Mountains (RM).

Watershed Name	***Tributaries***	***Mountain Ranges***	***Total Number of Specimen (N)***	***Specimens Sequence to Date (n)***
San Jacinto	Indian Creek	SB	5	4
Santa Ana	Plunge Creek	SB	16	8
	City Creek	SB	11	4
	Twin/Strawberry Creek	SB	12	12
	Lytle Creek	SB	10	10
	Cajon Creek	SB	15	12
San Gabriel (SGR)	Cattle Canyon	SG	4	4
	SGR, East Fork	SG	3	3
	SGR, North Fork	SG	2	2
	SGR, West Fork	SG	3	3
Santa Maria	Manzana Creek	CR	2	2
	Davy Brown Ck.	CR	2	2
	Cuyama River	CR	2	2
San Luis Obispo	San Luis Obispo Ck	CR	3	3
	Brizziolari Creek	CR	3	3
	Stenner Creek	CR	3	3
Owens River	Pine Creek	SN	3	3
	Marvin’s Marsh	SN	3	3
Colorado River	Colorado River	RM	11	11

Table 2A. Rhinichthys osculus Haplotype table with 92 specimens from 3 distinct geographic locations with 6 watershed locations.

Geographic Location	Watershed	Haplo #	Specimen ID																	
Southern California	San Jacinto R.	H1	IC1	IC2	IC5															
		H2	IC3																	
		H3	IC4																	
		H4	CC1	CC9	PC1	PC4	PC7	PC11	PC13	PC14										
	Santa Ana R.	H5	CC2	CC6	CAJ1	CAJ2	CAJ3	CAJ4	CAJ5	CAJ8	CAJ9	CAJ10	CAJ11	CAJ12	CAJ15	LCK3	LCK5	LCK6	LCK8	LCK10
		H6	PC10																	
		H7	PC12																	
		H8	TWIN1	TWIN12																
	Santa Ana R./ San Gabriel R.	H9	TWIN2	TWIN7	CATC2	EFSGR1	NFSGR1	NFSGR2	WFSGR1											
	Santa Ana R.	H10	TWIN3	TWIN5	TWIN6	TWIN8	TWIN9	TWIN11												
		H11	TWIN4																	
		H12	TWIN10																	
		H13	CAJ6																	
		H14	CAJ14																	
		H15	LCK4																	
		H16	LCK9																	
	San Gabriel R.	H17	CATC1																	
		H18	CATC3																	
		H19	EFSGR2																	
Central CA Coast	San Luis Obispo R.	H20	STC1																	
		H21	STC2	SLO3																
		H22	STC3																	
		H23	BC1																	
		H24	BC2																	
		H25	BC3																	
		H26	SLO1																	
		H27	SLO2																	
Eastern CA Desert	Santa Maria R.	H28	MNC1	MNC2	CUYR2	DBC2														
		H29	CUYR1																	
	Owens R.	H30	PNC1																	
		H31	PNC2	PNC3																
		H32	MM1																	
		H33	MM2																	
		H34	MM3																	

Table 2B. *Rhinichthys osculus* outgroup GenBank downloaded specimens.

Geographic Location	Watershed Locations	GENBANK Specimens Outgroup	Authors
Colorado River AZ	Animas River, San Juan River, CO	AZT1	Smith and Dowling (2008)
	Bonneville Park Canyon Ck., UT	BONpk1	
	Bonneville, Sevier River, UT	BONsv1	
	Animas River, San Juan River, CO	BP11	
	Eagle Creek, Gila River, AZ	EAG1	
	Little Snake River, Green River, CO	LSR1	
	Aravaipa Creek, Gila River, AZ	RAV1	
	Sonoita Creek, Santa Cruz River, AZ	SON1	
	Trout Creek, Bill Williams River, AZ	TC2	
	West Clear Creek, Verde River, AZ	WCC1	
	White River, Green River, CO	WR1	
	White River, Green River, CO	WR2	
Lake County, OR	Cotton Creek, OR	RO11652	Shonhuth, Shiozawa, Dowling, and Mayden (2012)
Columbia R., BC, CAN	Kettle River, B,C., Canada	ROSC	McPhail and Taylor (2009)
	Rhinichthys obtusus	Robt	Dowling, Chow, and Smith (Unpublished)
	Rhinichthys cataractae	Rcat	McPhail and Taylor (2009)

Table 3. Variable Sites within 1155-Base Pair Rhinichthys osculus Haplotypes.

	Domain: Exmem-1													Transmem-1							Exmem-2										Transmem-2									Exmem-3			Transmem-3													Exmem-4										Transmem-4											
	2	2	2	3	3	4	4	5	6	8	8	9	9	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	5
	1	4	7	6	9	6	8	7	9	1	6	0	3	0	0	2	2	3	3	5	7	7	9	0	0	1	1	2	3	4	6	8	8	9	0	0	1	1	1	3	4	4	5	5	5	6	6	6	7	7	8	9	9	0	0	0	0	1	2	2	2	2	3	4	4	4	5	6	6	6	6	7	7	7	8	9	0
H1	C	T	A	G	T	G	T	C	T	T	C	G	G	2	8	1	3	6	8	3	1	7	2	4	6	0	9	2	4	0	4	5	8	0	3	6	2	5	6	3	1	2	1	4	5	0	6	9	2	5	7	7	9	2	5	7	8	4	0	3	6	9	5	1	5	8	4	2	4	5	8	1	7	8	6	8	7
H2	.	.	.	.	.	.	.	.	.	.	.	.	.	A	C	T	G	C	G	A	C	C	G	C	T	T	T	C	C	A	A	T	A	A	A	G	T	G	T	G	C	T	G	C	C	C	G	C	A	A	A	C	G	G	A	G	T	G	C	A	C	C	C	G	C	T	G	C	A	C	A	A	T	C	G	T	A
H3	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
H4	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
H5	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
H6	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
H7	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	G	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
H8	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
H9	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	T	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	A	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.
H10	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	A	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.
H11	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.
H12	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	A	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.
H13	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	A	G	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.
H14	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
H15	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
H16	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
H17	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	A	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
H18	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	T	.	G	.	.	.	.	.	.	.	A	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.
H19	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	A	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.
H20	.	.	G	A	.	A	.	T	C	.	.	A	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	A	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.
H21	.	.	G	A	.	A	.	T	C	.	.	A	.	G	.	.	A	.	A	G	T	.	A	T	.	C	A	.	.	.	G	C	.	.	.	.	C	.	.	.	G	.	A	T	.	.	A	A	G	G	G	.	.	C	.	.	A	A	T	.	T	.	G	A	.	C	.	.	.	T	G	C	C	T	A	C	C
H22	.	.	G	A	.	A	.	T	C	.	.	A	.	G	.	.	A	.	A	C	T	.	A	T	.	C	A	.	.	.	G	C	.	.	.	.	C	.	.	.	G	.	A	T	.	.	A	A	G	G	G	.	.	C	.	.	A	A	T	.	T	.	G	A	.	C	.	.	.	T	G	C	C	T	A	C	C
H23	.	.	G	A	.	A	.	T	C	.	.	A	.	G	.	.	A	.	A	G	T	.	A	T	.	C	A	.	.	.	G	C	.	.	.	.	C	.	.	.	G	.	A	T	.	.	A	A	G	G	G	.	.	C	.	.	A	A	T	.	T	.	G	A	.	C	.	.	.	T	G	G	C	T	A	C	C
H24	.	.	G	A	.	A	.	T	C	.	.	A	.	G	.	.	A	.	A	G	T	.	A	T	.	C	A	.	.	.	G	C	.	.	.	.	C	.	.	.	G	.	A	T	.	.	A	A	G	G	G	.	.	C	.	.	A	A	T	.	T	.	G	A	.	C	.	.	.	T	G	C	C	T	A	C	C
H25	.	.	G	A	.	A	.	T	C	.	.	A	.	G	.	.	A	.	A	G	T	.	A	T	.	C	A	.	.	.	G	C	.	.	.	.	C	.	.	.	G	.	A	T	.	.	A	A	G	G	G	.	.	C	.	.	A	A	T	.	T	.	G	A	.	C	.	.	.	T	G	C	C	T	A	C	C
H26	.	.	G	A	.	A	.	T	C	.	.	A	.	G	.	.	A	.	A	G	T	.	A	T	.	C	A	.	.	.	G	C	.	.	.	.	C	.	.	.	G	.	A	T	.	.	A	A	G	G	G	.	.	C	.	.	A	A	T	.	T	.	G	A	.	C	.	.	.	T	G	C	C	T	A	C	C
H27	.	.	G	A	.	A	.	T	C	.	.	A	.	G	.	.	A	.	A	C	T	.	A	T	.	C	A	.	.	.	G	C	.	.	.	.	C	.	.	.	G	.	A	T	.	.	A	A	G	G	G	.	.	C	.	.	A	A	T	.	T	.	G	A	.	C	.	.	.	T	G	C	C	T	A	C	C
H28	.	.	G	A	.	A	.	T	C	.	.	A	.	G	.	.	A	.	A	C	T	.	A	T	.	C	A	.	.	.	G	C	.	.	.	.	C	.	.	.	G	.	A	T	.	.	A	A	G	G	G	.	.	C	.	.	A	A	T	.	T	.	G	A	.	C	.	.	.	T	G	C	C	T	A	C	C
H29	.	.	G	A	.	A	.	T	C	.	.	A	.	G	.	.	A	.	A	G	T	.	A	T	.	C	A	.	.	.	.	C	G	.	.	.	C	.	.	.	G	.	A	T	.	.	A	A	G	G	G	.	.	C	.	.	A	A	T	.	T	.	G	A	.	C	.	.	.	T	G	C	C	T	A	C	C
H30	T	C	.	A	C	.	C	T	C	.	.	A	A	G	.	.	A	.	A	G	T	.	A	T	.	C	A	.	.	.	.	C	G	G	.	.	C	.	.	.	G	.	A	T	.	.	A	A	G	G	G	.	.	C	.	.	A	A	T	.	T	.	G	A	.	C	.	.	T	T	G	C	C	T	A	C	C
H31	T	C	.	A	C	.	C	T	C	.	.	A	A	.	.	.	A	T	A	.	.	.	A	.	.	C	.	.	T	.	.	.	.	.	G	C	.	A	.	A	A	C	A	T	.	T	A	T	G	G	.	T	A	A	.	.	A	A	T	G	T	T	G	A	.	C	.	T	.	T	.	.	C	.	A	.	C
H32	T	C	.	A	C	.	C	T	C	.	.	A	A	.	T	.	A	T	A	.	.	.	A	.	C	C	.	.	T	.	.	C	.	.	G	C	.	A	.	A	A	C	A	T	.	T	A	T	G	G	.	T	A	A	.	.	A	A	T	G	T	T	G	A	.	C	.	T	.	T	.	C	C	.	.	.	C
H33	T	C	.	A	C	.	C	T	C	.	.	A	A	.	T	.	A	T	A	.	.	.	A	.	.	C	.	.	T	.	.	C	.	.	G	C	.	A	.	A	A	C	A	T	.	T	A	T	G	G	.	T	A	A	.	.	A	A	T	G	T	T	G	A	.	C	.	T	.	T	.	C	C	T	.	.	C
H34	T	C	.	A	C	.	C	T	C	.	.	A	A	.	.	.	A	T	A	.	.	.	A	.	.	C	.	.	T	.	.	.	.	.	G	C	.	A	.	A	A	C	A	T	.	T	A	T	G	G	.	T	A	A	.	.	A	A	T	G	.	T	G	A	.	C	.	T	.	T	.	C	C	T	A	.	C
														.	.	.	A	T	A	.	.	.	A	.	.	C	.	.	T	.	.	.	.	.	G	C	.	A	.	A	A	C	A	T	.	T	A	T	G	G	.	T	A	A	.	.	A	A	T	G	.	T	G	A	.	C	.	T	.	T	.	C	C	T	A	.	C

Table 3. Variable Sites within 1155-Base Pair Rhinichthys osculus Haplotypes continued.

	Exmem-5	Transmem-5																		Exmem-6											Transmem-6															Exmem-7																		Transmem-7											
	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	9	9	9	9
	2	3	4	5	5	5	6	6	6	6	6	6	7	7	7	8	8	8	9	9	2	3	3	4	4	4	5	5	6	6	7	7	8	8	9	9	0	0	0	1	1	2	2	2	3	3	3	4	5	5	6	6	7	8	9	9	1	2	2	2	3	4	5	6	7	7	8	8	8	9	9	0	0	1	2
	2	4	2	0	2	5	1	2	3	5	6	7	2	3	9	0	2	5	2	8	1	0	1	0	5	6	1	7	3	6	1	5	1	4	0	9	0	5	8	4	5	1	6	9	0	5	8	4	2	3	5	8	5	9	5	8	6	3	5	9	1	6	2	4	0	9	0	2	5	4	5	3	9	6	4
H1	G	T	T	C	G	T	T	G	T	A	T	C	G	C	A	G	C	A	T	T	C	G	T	A	A	G	G	C	C	A	T	A	A	T	T	A	C	A	G	T	C	T	G	G	C	C	T	T	G	T	T	C	G	T	A	A	G	T	A	G	T	C	C	G	G	A	C	A	A	T	C	A	A	G	T
H2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
H3	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
H4	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	C	.	A	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	G	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.
H5	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	C	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	G	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.
H6	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	C	.	A	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	G	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.
H7	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	C	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	G	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.
H8	.	.	.	.	.	.	.	.	.	.	.	T	.	.	G	.	.	.	C	.	.	.	.	G	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
H9	.	.	.	.	.	.	.	.	.	.	.	T	.	.	G	.	.	.	C	.	.	.	.	G	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
H10	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	C	.	.	.	C	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	G	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.
H11	.	.	.	.	.	.	.	.	.	.	.	T	.	.	G	.	.	.	C	.	.	.	.	G	.	A	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
H12	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	G	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.
H13	.	.	.	.	.	.	.	.	.	.	.	T	A	.	.	.	.	.	C	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.
H14	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	C	.	.	.	.	G	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	G	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.
H15	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	C	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	T	.	.	G	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.
H16	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	C	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	G	.	A	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.
H17	.	.	.	.	.	.	.	.	.	.	.	T	.	.	G	.	.	.	C	.	.	.	.	G	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
H18	.	.	.	.	.	.	.	.	.	.	.	T	.	.	G	.	.	.	C	.	.	.	.	G	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
H19	.	.	.	.	.	.	.	T	G	T	A	T	.	.	G	.	.	.	C	.	.	.	.	G	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
H20	A	C	A	.	A	C	C	.	.	.	.	.	.	T	G	.	.	.	C	.	.	A	.	G	.	.	A	T	.	.	.	.	G	C	.	.	.	G	A	.	.	.	A	A	T	.	C	A	C	C	.	.	C	C	.	G	A	C	G	.	C	A	T	A	.	G	.	.	G	C	.	.	.	.	C
H21	A	C	.	.	A	C	C	.	.	.	.	.	.	T	G	.	.	.	C	.	.	A	.	G	.	.	A	T	.	.	.	.	G	C	.	.	.	G	A	.	.	.	A	A	T	.	C	A	.	C	.	.	.	C	.	G	A	C	G	.	C	A	T	A	.	G	.	.	G	C	.	.	.	.	C
H22	A	C	.	.	A	C	C	.	.	.	.	.	.	T	G	.	.	.	C	.	.	A	.	G	.	.	A	T	.	.	C	.	G	C	.	.	.	G	A	.	T	.	A	A	T	.	C	A	C	C	.	.	.	C	.	G	A	C	G	.	C	A	T	A	.	G	.	.	G	C	.	.	.	.	C
H23	A	C	.	.	A	C	C	.	.	.	.	.	.	T	G	.	.	.	C	.	.	A	.	G	.	.	A	T	.	.	.	.	G	C	.	.	.	G	A	.	.	.	A	A	T	.	C	A	C	C	.	.	C	C	.	G	A	C	G	.	C	A	T	A	.	G	.	.	G	C	.	.	.	.	C
H24	A	C	.	.	A	C	C	.	.	.	.	.	.	T	G	.	.	.	C	.	.	A	.	G	.	.	A	T	.	.	.	.	G	C	.	.	.	G	A	.	.	.	A	A	T	.	C	A	C	C	.	.	.	C	.	G	A	C	G	.	C	A	T	A	.	G	.	.	G	C	.	.	.	.	C
H25	A	C	.	.	A	C	C	.	.	.	.	.	.	T	G	.	.	.	C	.	.	A	.	G	.	.	A	T	.	.	.	.	G	C	.	.	.	G	A	.	.	.	A	A	T	.	C	A	.	C	.	.	C	C	.	G	A	C	G	.	C	A	T	A	.	G	.	.	G	C	.	.	.	.	C
H26	A	C	.	.	A	C	C	.	.	.	.	.	.	T	G	.	.	.	C	.	.	A	.	G	.	.	A	T	.	.	.	.	G	C	.	.	.	G	A	.	.	.	A	A	T	.	C	A	.	C	.	.	.	C	.	G	A	C	G	.	C	A	T	A	.	G	.	.	G	C	.	.	.	.	C
H27	A	C	.	.	A	C	C	.	.	.	.	.	.	T	G	.	.	.	C	.	.	A	.	G	.	.	A	T	.	.	.	.	G	C	.	.	.	G	A	.	.	.	A	A	T	.	C	A	.	C	.	.	C	C	.	G	A	C	G	.	C	A	T	A	.	G	.	.	G	C	.	.	.	.	C
H28	A	C	.	.	A	C	C	.	.	.	.	.	.	T	G	.	.	.	C	.	.	A	.	G	.	.	A	T	.	.	.	.	G	C	.	.	.	G	A	.	.	.	A	A	T	.	C	A	.	C	.	.	.	C	.	G	A	C	G	.	C	A	T	A	.	G	.	.	G	C	.	.	.	.	C
H29	A	C	.	.	A	C	C	.	.	.	.	.	.	T	G	.	.	.	C	.	.	A	.	G	.	.	A	T	.	.	.	.	G	C	.	.	.	G	A	.	.	.	A	A	T	.	C	A	.	C	.	.	.	C	.	G	A	C	G	.	C	A	T	A	.	G	.	.	G	C	.	.	.	.	C
H30	.	C	.	T	A	.	C	.	.	.	.	.	.	T	G	A	T	G	C	C	.	A	.	G	G	.	A	T	.	.	.	.	G	C	A	G	.	.	A	C	.	.	A	A	T	.	A	A	.	C	C	.	.	.	.	.	.	.	G	.	C	G	T	A	A	G	T	G	.	.	T	T	G	A	.
H31	.	C	.	T	A	.	C	.	.	.	.	.	.	T	.	A	T	.	C	C	.	A	.	G	G	.	A	T	.	G	.	.	G	C	A	G	.	.	A	C	.	.	A	A	T	.	A	A	.	C	C	.	.	.	G	.	.	.	G	.	C	G	T	A	A	.	T	G	.	.	T	T	.	A	.
H32	.	C	.	T	A	.	C	.	.	.	.	.	.	T	.	A	T	G	C	C	.	A	.	G	G	.	A	T	.	G	.	.	G	C	A	G	.	.	A	C	.	.	A	A	T	.	A	A	.	C	C	.	.	.	G	.	.	.	G	.	C	G	T	A	A	.	T	G	.	.	T	T	.	A	.
H33	.	C	.	T	.	.	C	.	.	.	.	.	.	T	.	A	T	G	C	C	.	A	.	G	G	.	A	T	.	G	.	.	G	C	A	G	.	.	A	C	.	.	A	.	T	.	A	A	.	C	C	.	.	.	G	.	.	.	G	.	C	G	T	A	A	.	T	G	.	.	T	T	.	A	.
H34	.	C	.	T	.	.	C	.	.	.	.	.	.	T	.	A	T	G	C	C	.	A	.	G	G	.	A	T	.	.	.	.	G	C	A	G	.	.	A	C	.	.	A	.	T	.	A	A	.	C	C	.	.	.	.	G	.	.	G	.	C	G	T	A	A	.	T	G	.	.	T	T	.	A	.

Table 3. Variable Sites within 1155-Base Pair Rhinichthys osculus Haplotypes continued.

Exmem-8 Transmem-8 Exmem-9

								1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	9	9	9	9	9	9	9	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1
	3	5	6	8	8	8	9	0	0	1	1	2	2	4	5	6	7	8	9	0	0	1	1
	6	7	9	4	5	7	9	3	5	4	7	3	6	1	7	2	1	9	0	1	7	0	9
H1	A	A	T	A	G	G	C	T	A	T	A	G	T	C	G	G	G	C	C	C	G	C	A
H2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
H3	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
H4	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
H5	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
H6	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
H7	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
H8	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
H9	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
H10	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
H11	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
H12	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
H13	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
H14	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
H15	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
H16	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
H17	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
H18	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
H19	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
H20	.	G	C	.	.	C	.	.	.	.	.	A	C	T	A	C	A	T	G	.	A	G	G
H21	.	G	C	.	A	C	.	.	.	.	.	A	C	T	A	C	A	T	G	.	A	G	G
H22	.	G	C	.	.	C	.	.	.	.	.	A	C	T	A	C	A	T	G	.	A	G	G
H23	.	G	C	.	.	C	.	.	.	.	.	A	C	T	A	C	A	T	G	.	A	G	G
H24	.	G	C	.	.	C	.	.	.	.	.	A	C	T	A	C	A	T	G	.	A	G	G
H25	.	G	C	.	.	C	.	.	.	.	.	A	C	T	A	C	A	T	G	.	A	G	G
H26	.	G	C	.	A	C	.	C	.	.	.	A	C	T	A	C	A	T	G	.	A	G	G
H27	.	G	C	.	A	C	.	.	.	.	.	A	C	T	A	C	A	T	G	.	A	G	G
H28	.	G	C	.	.	C	.	.	.	.	.	A	C	T	A	C	A	T	G	.	A	G	G
H29	.	G	C	.	.	C	.	.	.	.	.	A	C	T	A	C	A	T	G	.	A	G	G
H30	G	.	C	G	.	.	.	.	G	C	G	A	C	T	A	.	.	T	G	G	A	A	G
H31	G	.	C	G	.	.	.	.	G	C	G	A	C	T	A	.	.	T	G	G	A	A	G
H32	G	.	C	G	.	.	.	.	G	C	G	A	C	T	A	.	.	T	G	G	A	A	G
H33	G	.	C	G	.	.	.	.	G	C	G	A	C	T	A	.	.	T	G	G	A	A	G
H34	G	.	C	G	.	.	.	.	G	C	G	A	C	T	A	.	.	T	G	G	A	A	G

R. osculus haplotype frequency and variable site data for 1155-bp sequences. Numbers above the columns correspond to nucleotide position relative to 1155-bp sequence; dots indicate identity with haplotype H1 and location with the cytochrome b gene.

Table 4. Cytochrome b Gene Statistics for 1155-Base Pair Haplotypes.

		Overall (1155 bp)	Extra-membrane (588 bp)	Trans-membrane (567 bp)
Base Frequencies	A	0.259	0.307	0.208
	T	0.290	0.236	0.347
	C	0.276	0.298	0.254
	G	0.176	0.159	0.191
Average Nucleotide Substitutions	1st Pos	8	1	7
	2nd Pos	1	1	0
	3rd Pos	53	23	30
	Total	62	25	37
	Ti	54	22	32
	Tv	8	3	5
	Ti/Tv	6.75	7.33	6.4
Nucleotide Substitutions	Synonymous	147	59	88
	Nonsynonymous	28	11	17
	Total	175	70	105

Base Frequencies and Nucleotide substitution data were illustrated for the overall cytochrome b sequence and divided by extra-membrane domain and trans-membrane domain regions.

Table 5: Chi-Square Analysis of Rhinicthys osculus

Chi-Square Analysis of Dace Cytb SNPs by Domain/Codon Position		Chi-Square Analysis of Dace Cytb Amino Acid Substitution Type by Domain		
Location	Chi-square P-value	Substititution Type	Chi-square P-value	
Domain	0.0102	Domain	0.2122	
Codon Position	6.53E-28	Hydrophobic/Hydrophilic	0.3688	
		Hydrophobic/Hydrophobic	0.0284	
		Uncharged/Charged	0.1206	Total Variable Sites = 28

Table 6. Summary of Best-fit Evolutionary Models for Cytochrome *b* gene by mtDNA Region.

mtDNA Nucleotide Sequence		Best-fit Evolutionary Model
	Overall	Tamura-Nei (TN93) with Gamma Distribution and Evolutionary Invariable site (G+I)
Domain Partition	Transmembrane	Kimura2-parameter (K2) with Gamma Distribution (G)
	Intra/Extramembrane	Hasegawa-Kishino-Yano (HKY) with Gamma Distribution (G)
Codon Partition	1st Position	Kimura 2-parameter (K2)
	2nd Position	Hasegawa-Kishino-Yano (HKY)
	3rd Position	Tamura-Nei (TN93) with Gamma Distribution (G)
Amino Acid	Overall	General Reversible Mitochondrial (mtREV24)
	Transmembrane	Jones-Taylor-Thornton (JTT)
	Intra/Extramembrane	Jones-Taylor-Thornton (JTT)

Table 7A. Statistics of Neurtality tests and mismatch distribution analysis by location partition

	Neutrality tests					Mismatch Distribution			
Location Partition	Tajima's D	(P-value)	Fu's Fs	(P-value)	Location Partition 3-Regions	SSD	(P-value)	Raggedness Inde	(P-value)
4-Regions	-0.3945	0.4034	-0.0321	0.4001	Santa Ana	0.0305	0.4225	0.0273	0.869
3-Regions	-0.4309	0.3873	-0.1374	0.4286	Central Coast	0.0054	0.8065	0.0214	0.934
Mountains	-0.4309	0.3812	-0.1374	0.432	Eastern Desert	0.0794	0.229	0.2	0.5135
Watersheds	-0.4309	0.383	-0.1374	0.4257					

Table 7B. Results of Population statistics for Santa Ana Speckled Dace, All Populations.

		Diversity Indices		***Neutrality Tests***		***Mismatch Distribution***		
		Haplotype Diversity h	***Nucleotide Diversity P***	***Tajima's D***	***Fu's Fs***	***P***		
Region	***Number of localities***	***(+/- 95% CI)***	***(+/- 95% CI)***	***(P-value)***	***(P-value)***	***(SSDobs)***	***Raggedness***	***P (Ragobs)***
Southern		0.8475	0.001761	-0.8369	-5.30392	0.0955	0.054	0.2835
California	7	(+/- 0.0322)	(+/- 0.001116)	(0.2145)	(0.0125)			
Central		0.7333	0.000974	-0.8223	0.16615	0.018	0.918	0.005
Coast	2	(+/- 0.1199)	(+/- 0.000782)	(0.2345)	(0.5155)			
Eastern		1	0.002637	-0.1247	-2.69605	0.3915	0.1333	0.509
Desert	1	(+/- 0.0962)	(+/- 0.001843)	(0.4765)	(0.0300)			

Table 8. Results of AMOVA analysis of genetic variation within and among populations of Rhinichthys osculus partitioned by geographic models based on watersheds, mountain ranges, geographic regions. All Phi statistics were found significant (P<0.01)

	Model							
	Watersheds (5 groups)		Mountains (5 groups)		Regions (3 groups)		Regions (4 groups)	
	% total variation	F-statistics	% total variation	F-statistics	% total variation	F-statistics	% total variation	F-statistics
Among groups	91.65	Φ-CT = .92	89.83	Φ-CT = .90	92.22	Φ-CT = .92	76	Φ-CT = .765
Among populations within groups	4.46	Φ-SC = .53	5.47	Φ-SC = .54	4.80	Φ-SC = .62	4	Φ-SC = .161
Within populations	3.90	Φ-PT = .96	4.70	Φ-PT = .95	2.99	Φ-PT = .97	20	Φ-PT = .803

Table 9. Cytochrome *b* gene average percent genetic distance between populations.

Population Comparison	***Cytochrome b Distance (%)***
Southern CA/Central CA Coast	8.620242
Southern CA/Eastern CA Desert	9.342561
Southern CA/Colorado River	6.012111
Central CA Coast/Eastern CA Desert	6.920415
Central CA Coast/Colorado River	7.378893
Eastern Ca Desert/Colorado River	7.491349

Table 10. Amino Acid Substitutions with domain and amino acid location number (A = Charged, B=Hydrophilic, C = Special Case,D = Hydrophobic, / = change to and +/- =charge)

Amino Acid Number	**Substitutions**	**Domain**
16	A-/B	exmembrane
29	D/D	exmembrane
69	D/B	exmembrane
97	A+/A+	transmembrane
106	B/D	transmembrane
114	B/B/B-	exmembrane
136	C/A-	exmembrane
152	D/B	transmembrane
155	D/D	transmembrane
160	D/D	transmembrane
181	D/D	transmembrane
188	D/C	transmembrane
189	D/D	transmembrane
191	C/A-	transmembrane
194	D/D	transmembrane
207	B/A+	exmembrane
214	B/A-	exmembrane
216	A-/B	exmembrane
224	D/B	transmembrane
239	D/D	transmembrane
241	B/C	transmembrane
251	C/D	exmembrane
259	D/C	exmembrane
277	D/B	exmembrane
306	D/D	transmembrane
329	D/D	transmembrane
353	D/D	transmembrane

364	D/D	transmembrane

Table 11. Amino Acid Haplotype Table with 28 amino acid variable sites.

						1	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	3	3	3	3
		1	2	6	9	0	1	3	5	5	6	8	8	8	9	9	0	1	1	2	3	4	5	5	7	0	2	5	6
		6	9	9	7	6	4	6	2	5	0	1	8	9	1	4	7	4	6	4	9	1	1	9	7	6	9	3	4
H1		D	A	I	H	S	T	G	A	Y	L	F	V	I	G	V	N	N	D	F	L	S	G	A	A	V	V	V	L
H2		.	.	.	.	.	.	D	.	.	.	.	.	.	.	.	.	.	.	.	.	P	.	.	.	.	.	.	.
H3		.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	S	.	.	.	.
H4		.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	K	D	.	.	.	.	.	.	.	.	.	.	.
H5		.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	D	.	.	.	.	.	.	.	.	.	.	.
H6		.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	K	D	.	.	.	.	.	.	.	.	.	.	.
H7		.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	D	.	.	.	.	.	.	.	.	.	.	.
H8		.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	D	.	.	.	.	.	.	.	.	.	.	.
H9		.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	D	.	.	.	.	.	.	.	.	.	.	.
H10		.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	D	.	.	.	.	.	.	.	.	.	.	.
H11		.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	D	N	.	.	.	.	.	.	.	.	.	.
H12		.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	D	.	.	.	.	.	.	.	.	.	.	.
H13		.	.	.	.	.	.	.	.	.	.	.	.	.	D	.	.	D	.	.	.	.	.	.	.	.	.	.	.
H14		.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	D	.	.	.	.	.	.	.	.	.	.	.
H15		.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	D	.	.	.	.	.	.	.	.	.	.	.
H16		.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	D	.	.	.	.	.	.	.	.	.	.	.
H17		.	V	.	.	.	.	.	.	.	.	.	.	.	.	.	.	D	.	.	.	.	.	.	.	.	.	.	.
H18		.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	D	.	.	.	.	.	.	.	.	.	.	.
H19		.	.	.	.	.	.	.	.	.	.	.	C	Y	.	.	.	D	.	.	.	.	.	.	.	.	.	.	.
H20		N	.	.	.	.	S	.	.	.	F	Y	.	.	.	.	.	D	.	.	.	.	A	P	.	.	.	I	V
H21		N	.	.	.	.	S	.	.	.	F	.	.	.	.	.	.	D	.	.	.	.	.	.	.	.	I	I	V
H22		N	.	.	.	.	S	.	.	.	F	.	.	.	.	.	.	D	.	S	F	.	A	.	.	.	.	I	V
H23		N	.	.	.	.	S	.	.	.	F	.	.	.	.	.	.	D	.	.	.	.	A	P	.	.	.	I	V
H24		N	.	.	.	.	S	.	.	.	F	.	.	.	.	.	.	D	.	.	.	.	A	.	.	.	.	I	V
H25		N	.	.	.	.	S	.	.	.	F	.	.	.	.	.	.	D	.	.	.	.	.	P	.	.	.	I	V
H26		N	.	.	.	.	S	.	.	.	F	.	.	.	.	.	.	D	.	.	.	.	.	.	.	.	I	I	V
H27		N	.	.	.	.	S	.	.	.	F	.	.	.	.	.	.	D	.	.	.	.	.	P	.	.	I	I	V
H28		N	.	.	.	.	S	.	.	.	F	.	.	.	.	.	.	D	.	.	.	.	.	.	.	.	.	I	V
H29		N	.	.	R	.	S	.	.	F	F	.	.	.	.	.	.	D	.	.	.	.	.	.	.	.	.	I	V
H30		.	.	.	.	.	N	.	.	.	.	.	.	.	.	I	.	D	.	.	.	.	.	.	.	I	.	I	V
H31		.	.	T	.	.	N	.	.	.	.	.	.	.	.	I	.	D	.	.	.	.	.	.	.	I	.	I	V
H32		.	.	.	.	.	N	.	.	.	F	.	.	.	.	I	.	D	.	.	.	.	.	.	.	I	.	I	V
H33		.	.	.	.	.	N	.	.	.	F	.	.	.	.	I	.	D	.	.	.	.	.	.	.	I	.	I	V
H34		.	.	.	.	.	N	.	.	.	F	.	.	.	.	I	.	D	.	.	.	.	.	.	.	I	.	I	V

APPENDIX C
INPUT FILES

Mr. Bayes Input File

```
#NEXUS

Begin data;
        Dimensions ntax=49 nchar=1140;
        Format datatype= DNA gap=-;
        Matrix
H1
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGAC
CTTCCAACTCCATCTAATATTTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGC
TTGATTACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCA
ACTGCCTTTTCGTCAGTAACACACATCTGTCGAGACGTTAACTATGGCTGACTCATCCGA
AACATACATGCTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGC
GGACTGTACTATGGGTCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCCTACTC
CTTTTGGTCATAATAACAGCCTTTGTAGGCTATGTACTGCCGTGAGGTCAAATGTCTTTC
TGAGGCGCCACCGTCATTACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTT
GTCCAGTGGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCC
TTCCACTTCCTGTTTCCGTTTGTTATCGCCGGCGCAACAGTCCTACATTTGTTATTTTTA
CACGAAACGGGATCAAACAACCCTGCCGGGTTAAACTCCAACGCAGATAAGATCTCCTTC
CACCCATACTTCTCATACAAAGATCTCCTTGGCTTTGTACTAATATTGTTAGCTCTCACA
TCCCTGACGCTATTCTCTCCAACTCTGCTCGGTGACCCAGAAAATTTCACCCCGGCAAAC
CCCCTGGTTACCCCACCACACATCCAACCGGAGTGGTACTTCTTATTTGCTTACGCCATT
CTACGCTCTATCCCAAACAAGCTGGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTA
CTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATC
ACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATCATTTTAACATGAATTGGAGGC
ATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCA
CTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGAGCCGAAAATAAAGCATTGAAATGAGCT

H2
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGAC
CTTCCAACTCCATCTAATATTTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGC
TTGATTACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCA
ACTGCCTTTTCGTCAGTAACACACATCTGTCGAGACGTTAACTATGGCTGACTCATCCGA
AACATACATGCTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGC
GGACTGTACTATGGGTCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCCTACTC
CTTTTGGTCATAATAACAGCCTTTGTAGGCTATGTACTGCCGTGAGATCAAATGTCTTTC
TGAGGCGCCACCGTCATTACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTT
GTCCAGTGGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCC
TTCCACTTCCTGTTTCCGTTTGTTATCGCCGGCGCAACAGTCCTACATTTGTTATTTTTA
CACGAAACGGGATCAAACAACCCTGCCGGGTTAAACTCCAACGCAGATAAGATCTCCTTC
CACCCATACTTCTCATACAAAGATCTCCTTGGCTTTGTACTAATATTGTTAGCTCTCACA
CCCCTGACGCTATTCTCTCCAACTCTGCTCGGTGACCCAGAAAATTTCACCCCGGCAAAC
CCCCTGGTTACCCCACCACACATCCAACCGGAGTGGTACTTCTTATTTGCTTACGCCATT
CTACGCTCTATCCCAAACAAGCTGGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTA
CTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATC
ACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATCATTTTAACATGAATTGGAGGC
ATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCA
CTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGAGCCGAAAATAAAGCATTGAAATGAGCT

H3
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGAC
```

CTTCCAACTCCATCTAATATTTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGC
TTGATTACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCA
ACTGCCTTTTCGTCAGTAACACACATCTGTCGAGACGTTAACTATGGCTGACTCATCCGA
AACATACATGCTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGC
GGACTGTACTATGGGTCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCCTACTC
CTTTTGGTCATAATAACAGCCTTTGTAGGCTATGTACTGCCGTGAGGTCAAATGTCTTTC
TGAGGCGCCACCGTCATTACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTT
GTCCAGTGGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCC
TTCCACTTCCTGTTTCCGTTTGTTATCGCCGGCGCAACAGTCCTACATTTGTTATTTTTA
CACGAAACGGGATCAAACAACCCTGCCGGGTTAAACTCCAACGCAGATAAGATCTCCTTC
CACCCATACTTCTCATACAAAGATCTCCTTGGCTTTGTACTAATATTGTTAGCTCTCACA
TCCCTGACGCTATTCTCTCCAACTCTGCTCGGTGACCCAGAAAATTTCACCCCGGCAAAC
CCCCTGGTTACCCCACCACACATCCAACCGGAGTGGTACTTCTTATTTTCTTACGCCATT
CTACGCTCTATCCCAAACAAGCTGGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTA
CTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATC
ACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATCATTTTAACATGAATTGGAGGC
ATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCA
CTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGAGCCGAAAATAAAGCATTGAAATGAGCT

H4
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGAC
CTTCCAACTCCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGC
TTGATTACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCA
ACTGCCTTTTCGTCAGTAACACACATCTGTCGAGACGTTAATTATGGCTGACTCATCCGG
AACATACATGCTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGC
GGACTGTACTATGGGTCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCCTACTC
CTTTTGGTCATAATAACAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTC
TGAGGCGCCACCGTCATTACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTT
GTCCAGTGGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCC
TTCCACTTCCTGTTTCCGTTTGTTATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTA
CACGAAACGGGATCAAACAAACCTGCCGGGTTAAACTCCGACGCAGATAAGATCTCCTTC
CACCCATACTTCTCATACAAAGATCTCCTTGGCTTTGTACTAATATTGTTAGCTCTCACA
TCCCTGACGCTATTCTCTCCAACTCTGCTCGGTGACCCAGAAAATTTTACCCCGGCAAAC
CCCCTGGTTACCCCGCCACACATCCAACCGGAGTGGTACTTCTTATTTGCTTACGCCATT
CTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTA
CTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATC
ACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATTATTTTAACATGAATTGGAGGC
ATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCA
CTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGAGCCGAAAATAAAGCATTGAAATGAGCT

H5
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGAC
CTTCCAACTCCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGC
TTGATTACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCA
ACTGCCTTTTCGTCAGTAACACACATCTGTCGAGACGTTAATTATGGCTGACTCATCCGG
AACATACATGCTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGC
GGACTGTACTATGGGTCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCCTACTC
CTTTTGGTCATAATAACAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTC
TGAGGCGCCACCGTCATTACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTT
GTCCAGTGGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCC
TTCCACTTCCTGTTTCCGTTTGTTATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTA
CACGAAACGGGATCAAACAACCCTGCCGGGTTAAACTCCGACGCAGATAAGATCTCCTTC
CACCCATACTTCTCATACAAAGATCTCCTTGGCTTTGTACTAATATTGTTAGCTCTCACA

TCCCTGACGCTATTCTCTCCAACTCTGCTCGGTGACCCAGAAAATTTTACCCCGGCAAAC
CCCCTGGTTACCCCGCCACACATCCAACCGGAGTGGTACTTCTTATTTGCTTACGCCATT
CTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTA
CTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATC
ACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATTATTTTAACATGAATTGGAGGC
ATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCA
CTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGAGCCGAAAATAAAGCATTGAAATGAGCT

H6
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGAC
CTTCCAACTCCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGC
TTGATTACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCA
ACTGCCTTTTCGTCAGTAACACACATCTGTCGAGACGTTAATTATGGCTGACTCATCCGG
AACATACATGCTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGC
GGACTGTACTATGGGGCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCCTACTC
CTTTTGGTCATAATAACAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTC
TGAGGCGCCACCGTCATTACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTT
GTCCAGTGGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCC
TTCCACTTCCTGTTTCCGTTTGTTATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTA
CACGAAACGGGATCAAACAAACCTGCCGGGTTAAACTCCGACGCAGATAAGATCTCCTTC
CACCCATACTTCTCATACAAAGATCTCCTTGGCTTTGTATTAATATTGTTAGCTCTCACA
TCCCTGACGCTATTCTCTCCAACTCTGCTCGGTGACCCAGAAAATTTTACCCCGGCAAAC
CCCCTGGTTACCCCGCCACACATCCAACCGGAGTGGTACTTCTTATTTGCTTACGCCATT
CTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTA
CTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATC
ACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATTATTTTAACATGAATTGGAGGC
ATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCA
CTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGAGCCGAAAATAAAGCATTGAAATGAGCT

H7
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGAC
CTTCCAACTCCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGC
TTGATTACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCA
ACTGCCTTTTCGTCAGTAACACACATCTGTCGAGACGTTAATTATGGCTGACTCATCCGG
AACATACATGCTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGC
GGACTGTACTATGGGTCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCCTACTC
CTTTTGGTCATAATAACAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTC
TGAGGCGCCACCGTCATTACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTT
GTCCAGTGGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCC
TTCCACTTCCTGTTTCCGTTTGTTATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTA
CACGAAACGGGATCAAACAACCCTGCCGGGTTAAACTCCGACGCAGATAAGATCTCCTTC
CACCCATACTTCTCATACAAAGATCTCCTTGGCTTTGTATTAATATTGTTAGCTCTCACA
TCCCTGACGCTATTCTCTCCAACTCTGCTCGGTGACCCAGAAAATTTTACCCCGGCAAAC
CCCCTGGTTACCCCGCCACACATCCAACCGGAGTGGTACTTCTTATTTGCTTACGCCATT
CTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTA
CTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATC
ACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATTATTTTAACATGAATTGGAGGC
ATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCA
CTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGAGCCGAAAATAAAGCATTGAAATGAGCT

H8
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGAC
CTTCCAACTCCATCTAATATTTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGC

CTGATTACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATTTCA
ACTGCCTTTTCGTCAGTAACACACATCTGTCGAGACGTTAACTATGGCTGACTCATCCGG
AACATACATGCTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGC
GGACTGTACTATGGATCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCTTACTC
CTTTTGGTCATAATAACAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTC
TGAGGCGCCACCGTCATTACGAATTTATTATCAGCAGTCCCCTACATAGGAGACGCTCTT
GTCCAGTGGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCC
TTCCACTTCCTGTTTCCGTTTGTTATTGCCGGCGCAACGGTCCTACATTTGCTATTTTTA
CACGAAACGGGATCAAACAACCCTGCCGGGTTAAACTCCGACGCAGATAAGATCTCCTTC
CACCCATACTTCTCGTACAAAGATCTCCTTGGCTTTGTACTAATATTGTTAGCTCTCACA
TCCCTGACGCTATTCTCTCCAACTCTGCTCGGTGACCCAGAAAATTTCACCCCGGCAAAC
CCCCTGGTTACCCCGCCACACATCCAACCGGAGTGGTACTTCTTATTTGCTTACGCCATT
CTACGCTCTATCCCAAACAAGCTGGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTA
CTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATC
ACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATCATTTTAACATGAATTGGAGGC
ATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCA
CTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGAGCCGAAAATAAAGCATTGAAATGAGCT

H9
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGAC
CTTCCAACTCCATCTAATATTTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGC
TTGATTACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATTTCA
ACTGCCTTTTCGTCAGTAACACACATCTGTCGAGACGTTAACTATGGCTGACTCATCCGG
AACATACATGCTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGC
GGACTGTACTATGGATCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCTTACTC
CTTTTGGTCATAATAACAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTC
TGAGGCGCCACCGTCATTACGAATTTATTATCAGCAGTCCCCTACATAGGAGACGCTCTT
GTCCAGTGGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCC
TTCCACTTCCTGTTTCCGTTTGTTATTGCCGGCGCAACGGTCCTACATTTGCTATTTTTA
CACGAAACGGGATCAAACAACCCTGCCGGGTTAAACTCCGACGCAGATAAGATCTCCTTC
CACCCATACTTCTCGTACAAAGATCTCCTTGGCTTTGTACTAATATTGTTAGCTCTCACA
TCCCTGACGCTATTCTCTCCAACTCTGCTCGGTGACCCAGAAAATTTCACCCCGGCAAAC
CCCCTGGTTACCCCGCCACACATCCAACCGGAGTGGTACTTCTTATTTGCTTACGCCATT
CTACGCTCTATCCCAAACAAGCTGGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTA
CTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATC
ACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATCATTTTAACATGAATTGGAGGC
ATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCA
CTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGAGCCGAAAATAAAGCATTGAAATGAGCT

H10
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGAC
CTTCCAACTCCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGC
TTGATTACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCA
ACTGCCTTTTCGTCAGTAACACACATCTGTCGAGACGTTAATTATGGCTGACTCATCCGG
AACATACATGCTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGC
GGACTGTACTATGGGTCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCCTACTC
CTTTTGGTCATAATAACAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTC
TGAGGCGCCACCGTCATTACGAATCTATTATCAACAGTCCCCTACATAGGAGACGCTCTT
GTCCAGTGGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCC
TTCCACTTCCTGTTTCCGTTTGTTATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTA
CACGAAACGGGATCAAACAACCCTGCCGGGCTAAACTCCGACGCAGATAAGATCTCCTTC
CACCCATACTTCTCATACAAAGATCTCCTTGGCTTTGTACTAATATTGTTAGCTCTCACA
TCCCTGACGCTATTCTCTCCAACTCTGCTCGGTGACCCAGAAAATTTTACCCCGGCAAAC

CCCCTGGTTACCCCGCCACACATCCAACCGGAGTGGTACTTCTTATTTGCTTACGCCATT
CTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTA
CTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATC
ACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATTATTTTAACATGAATTGGAGGC
ATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCA
CTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGAGCCGAAAATAAAGCATTGAAATGAGCT

H11
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGAC
CTTCCAACTCCATCTAATATTTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGC
TTGATTACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATTTCA
ACTGCCTTTTCGTCAGTAACACACATCTGTCGAGACGTTAACTATGGCTGACTCATCCGG
AACATACATGCTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGC
GGACTGTACTATGGATCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCTTACTC
CTTTTGGTCATAATAACAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTC
TGAGGCGCCACCGTCATTACGAATTTATTATCAGCAGTCCCCTACATAGGAGACGCTCTT
GTCCAGTGGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCC
TTCCACTTCCTGTTTCCGTTTGTTATTGCCGGCGCAACGGTCCTACATTTGCTATTTTTA
CACGAAACGGGATCAAACAACCCTGCCGGGTTAAACTCCGACGCAAATAAGATCTCCTTC
CACCCATACTTCTCGTACAAAGATCTCCTTGGCTTTGTACTAATATTGTTAGCTCTCACA
TCCCTGACGCTATTCTCTCCAACTCTGCTCGGTGACCCAGAAAATTTCACCCCGGCAAAC
CCCCTGGTTACCCCGCCACACATCCAACCGGAGTGGTACTTCTTATTTGCTTACGCCATT
CTACGCTCTATCCCAAACAAGCTGGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTA
CTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATC
ACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATCATTTTAACATGAATTGGAGGC
ATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCA
CTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGAGCCGAAAATAAAGCATTGAAATGAGCT

H12
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGAC
CTTCCAACTCCATCTAATATTTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGC
TTGATTACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCA
ACTGCCTTTTCGTCAGTAACACACATCTGTCGAGACGTTAATTATGGCTGACTCATCCGG
AACATACATGCTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGC
GGACTGTACTATGGGTCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCCTACTC
CTTTTGGTCATAATAACAGCCTTTGTAGGCTATGTACTACCATGGGGTCAAATATCTTTC
TGAGGCGCCACCGTCATTACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTT
GTCCAATGGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCC
TTCCACTTCCTATTTCCGTTTGTTATCGCCGGCGCAACAGTCCTACATTTGCTATTTTTA
CACGAAACGGGATCAAACAACCCTGCCGGGTTAAACTCCGACGCAGATAAGATCTCCTTC
CACCCATACTTCTCATACAAAGATCTCCTTGGCTTTGTACTAATATTGTTAGCTCTCACA
TCCCTGACGCTATTCTCTCCAACTCTGCTCGGTGACCCAGAAAATTTTACCCCGGCAAAC
CCCCTGGTTACCCCGCCACACATCCAACCGGAGTGGTACTTCTTATTTGCTTACGCCATT
CTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTA
CTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATC
ACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATTATTTTAACATGAATTGGAGGC
ATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCA
CTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGAGCCGAAAATAAAGCATTGAAATGAGCT

H13
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGAC
CTTCCAACTCCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGC
TTGATTACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCA

ACTGCCTTTTCGTCAGTAACACACATCTGTCGAGACGTTAATTATGGCTGACTCATCCGG
AACATACATGCTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGC
GGACTGTACTATGGGTCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCCTACTC
CTTTTGGTCATAATAACAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTC
TGAGGCGCCACCGTCATTACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTT
GTCCAGTGGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCC
TTCCACTTCCTGTTTCCGTTTGTTATTGCCGACGCAACAGTCCTACATTTGCTATTTTTA
CACGAAACGGGATCAAACAACCCTGCCGGGTTAAACTCCGACGCAGATAAGATCTCCTTC
CACCCATACTTCTCATACAAAGATCTCCTTGGCTTTGTACTAATATTGTTAGCTCTCACA
TCCCTGACGCTATTCTCTCCAACTCTGCTCGGTGACCCAGAAAATTTTACCCCGGCAAAC
CCCCTGGTTACCCCACCACACATCCAACCGGAGTGGTACTTCTTATTTGCTTACGCCATT
CTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTA
CTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATC
ACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATTATTTTAACATGAATTGGAGGC
ATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCA
CTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGAGCCGAAAATAAAGCATTGAAATGAGCT

H14
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGAC
CTTCCAACTCCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGC
TTGATTACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCA
ACTGCCTTTTCGTCAGTAACACACATCTGTCGAGACGTTAATTATGGCTGACTCATCCGG
AACATACATGCTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGC
GGACTGTACTATGGGTCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCCTACTC
CTTTTGGTCATAATAACAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTC
TGAGGCGCCACCGTCATTACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTT
GTCCAGTGGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCC
TTCCACTTCCTGTTTCCGTTTGTTATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTA
CACGAAACGGGATCAAACAACCCTGCCGGGTTAAACTCCGACGCAGATAAGATCTCCTTC
CATCCATACTTCTCATACAAAGATCTCCTTGGCTTTGTACTAATATTGTTAGCTCTCACA
TCCCTGACGCTATTCTCTCCAACTCTGCTCGGTGACCCAGAAAATTTTACCCCGGCAAAC
CCCCTGGTTACCCCGCCACACATCCAACCGGAGTGGTACTTCTTATTTGCTTACGCCATT
CTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTA
CTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATC
ACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATTATTTTAACATGAATTGGAGGC
ATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCA
CTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGAGCCGAAAATAAAGCATTGAAATGAGCT

H15
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGAC
CTTCCAACTCCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGC
TTGATTACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCA
ACTGCCTTTTCGTCAGTAACACACATCTGTCGAGACGTTAATTATGGCTGACTCATCCGG
AACATACATGCTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGC
GGACTGTACTATGGGTCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCCTACTC
CTTTTGGTCATAATAACAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTC
TGAGGCGCCACCGTCATTACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTT
GTCCAGTGGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCC
TTCCACTTCCTGTTTCCGTTTGTTATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTA
CACGAAACGGGATCAAACAACCCTGCCGGGTTAAACTCCGACGCAGATAAGATCTCCTTC
CACCCATACTTCTCATACAAAGATCTCCTTGGCTTTGTACTAATATTGTTAGCTCTCACA
TCCCTGACGCTATTCTCTCCAACACTGCTCGGTGACCCAGAAAATTTTACCCCGGCAAAC
CCCCTGGTTACCCCGCCACACATCCAACCGGAGTGGTACTTCTTATTTGCTTACGCCATT

CTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTA
CTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATC
ACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATTATTTTAACATGAATTGGAGGC
ATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCA
CTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGAGCCGAAAATAAAGCATTGAAATGAGCT

H16
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGAC
CTTCCAACTCCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGC
TTGATTACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCA
ACTGCCTTTTCGTCAGTAACACACATCTGTCGAGACGTTAATTATGGCTGACTCATCCGG
AACATACATGCTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGC
GGACTGTACTATGGGTCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTTCTACTC
CTTTTAGTCATAATAACAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTC
TGAGGCGCCACCGTCATTACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTT
GTCCAGTGGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCC
TTCCACTTCCTGTTTCCGTTTGTTATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTA
CACGAAACGGGATCAAACAACCCTGCCGGGTTAAACTCCGACGCAGATAAGATCTCCTTC
CACCCATACTTCTCATACAAAGATCTCCTTGGCTTTGTACTAATATTGTTAGCTCTCACA
TCCCTGACGCTATTCTCTCCAACTCTGCTCGGTGACCCAGAAAATTTTACCCCGGCAAAC
CCCCTGGTTACCCCGCCACACATCCAACCGGAGTGATACTTCTTATTTGCTTACGCCATT
CTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTA
CTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATC
ACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATTATTTTAACATGAATTGGAGGC
ATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCA
CTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGAGCCGAAAATAAAGCATTGAAATGAGCT

H17
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGAC
CTTCCAACTCCATCTAATATTTCAGTACTGTGGAACTTCGGATCCCTCCTAGGATTATGC
TTGATTACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATTTCA
ACTGCCTTTTCGTCAGTAACACACATCTGTCGAGACGTTAATTATGGCTGACTCATCCGG
AACATACATGCTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGC
GGACTGTACTATGGATCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCTTACTC
CTTTTGGTCATAATAACAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTC
TGAGGCGCCACCGTCATTACGAATTTATTATCAGCAGTCCCCTACATAGGAGACGCTCTT
GTCCAGTGGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCC
TTCCACTTCCTGTTTCCGTTTGTTATTGCCGGCGCAACGGTCCTACATTTGCTATTTTTA
CACGAAACGGGATCAAACAACCCTGCCGGGTTAAACTCCGACGCAGATAAGATCTCCTTC
CACCCATACTTCTCGTACAAAGATCTCCTTGGCTTTGTACTAATATTGTTAGCTCTCACA
TCCCTGACGCTATTCTCTCCAACTCTGCTCGGTGACCCAGAAAATTTCACCCCGGCAAAC
CCCCTGGTTACCCCGCCACACATCCAACCGGAGTGGTACTTCTTATTTGCTTACGCCATT
CTACGCTCTATCCCAAACAAGCTGGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTA
CTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATC
ACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATCATTTTAACATGAATTGGAGGC
ATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCA
CTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGAGCCGAAAATAAAGCATTGAAATGAGCT

H18
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGAC
CTTCCAACTCCATCTAATATTTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGC
TTGATTACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATTTCA

ACTGCCTTTTCGTCAGTAACACACATCTGTCGAGACGTTAACTATGGCTGACTCATCCGG
AACATACATGCTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGC
GGACTGTACTATGGATCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCTTACTC
CTTTTGGTCATAATAACAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTC
TGAGGCGCCACCGTCATTACGAATTTATTATCAGCAGTCCCCTACATAGGAGACGCTCTT
GTCCAGTGGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCC
TTCCACTTCCTGTTTCCGTTTGTTATTGCCGGCGCAACGGTCCTACATTTGCTATTTTTA
CACGAAACGGGATCAAACAACCCTGCCGGGTTAAACTCCGACGCAGATAAGATCTCCTTC
CACCCATACTTCTCGTACAAAGATCTCCTTGGCTTTGTACTAATATTGTTAGCTCTCACA
TCCCTGACGCTATTTTCTCCAACTCTGCTCGGTGACCCAGAAAATTTCACCCCGGCAAAC
CCCCTGGTTACCCCGCCACACATCCAACCGGAGTGGTACTTCTTATTTGCTTACGCCATT
CTACGCTCTATCCCAAACAAGCTGGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTA
CTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATC
ACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATCATTTTAACATGAATTGGAGGC
ATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCA
CTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGAGCCGAAAATAAAGCATTGAAATGAGCT

H19
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGAC
CTTCCAACTCCATCTAATATTTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGC
TTGATTACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATTTCA
ACTGCCTTTTCGTCAGTAACACACATCTGTCGAGACGTTAACTATGGCTGACTCATCCGG
AACATACATGCTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGC
GGACTGTACTATGGATCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCTTACTC
CTTTTGGTCATAATAACAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTC
TGAGGCGCCACCGTCATTACGAATTTATTATCAGCAGTCCCCTACATAGGAGACGCTCTT
GTCCAGTGGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCC
TTCCACTTCCTGTTTCCGTTTTGTTATGCCGGCGCAACGGTCCTACATTTGCTATTTTTA
CACGAAACGGGATCAAACAACCCTGCCGGGTTAAACTCCGACGCAGATAAGATCTCCTTC
CACCCATACTTCTCGTACAAAGATCTCCTTGGCTTTGTACTAATATTGTTAGCTCTCACA
TCCCTGACGCTATTCTCTCCAACTCTGCTCGGTGACCCAGAAAATTTCACCCCGGCAAAC
CCCCTGGTTACCCCGCCACACATCCAACCGGAGTGGTACTTCTTATTTGCTTACGCCATT
CTACGCTCTATCCCAAACAAGCTGGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTA
CTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATC
ACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATCATTTTAACATGAATTGGAGGC
ATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCA
CTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGAGCCGAAAATAAAGCATTGAAATGAGCT

H20
ATGGCAAGCCTACGAAAAACCCATCCGCTAATAAAAATTGCTAACAATGCACTAGTTGAC
CTTCCAACCCCATCTAATATTTCAGCACTATGGAACTTCGGGTCCCTCCTAGGATTATGC
TTAATTACTCAAATCCTAACAGGACTATTCCTGGCCATACATTATACTTCTGATATCTCA
ACTGCCTTTTCATCAGTAACACATATCTGCCGAGACGTAAACTATGGCTGACTCATCCGA
AACATACATGCTAACGGAGCATCGTTCTTCTTTATCTGTATTTACATACACATTGCCCGC
GGACTGTACTACGGGTCATACCTCTATAAAGAGACCTGAAGTATTGGCGTAGTTCTACTC
CTTTTAGTAATGATGACAGCCTTTGTGGGCTATGTACTGCCCTGAGGACAAATATCTTTT
TGAGGTGCCACCGTGATTACAAATCTACTATCAGCAGTCCCCTATATGGGCGACGCCTTT
GTCCAATGGATTTGAGGCGGCTTTTCCGTAGATAACGCAACATTAACACGATTCTTCGCC
TACCACTTCCTATTCCCGTTCGTTATCGCCGGTGCAACGGTCCTACATTTGCTATTTTTA
CACGAAACGGGATCAAACAACCCTGCCGGATTAAACTCCGACGCAGATAAAATCTCTTTC
CACCCATACTTCTCATACAAGGACCTCCTTGGCTTTGTACTAATGTTATTAGCTCTCACA
TCCCTAACATTATTCTCCCCAACACTGCTCGCCGACCCAGAAAATTTCACCCCGCCAAAC
CCCCTGGTCACCCCACCGCACATCCAACCGGAGTGATACTTCCTGTTTGCCTACGCCATT

CTACGATCTATTCCAAACAAGCTAGGAGGGGTCCTAGCGCTATTGTTTAGCATCCTAGTA
CTATTAGTAGTGCCAGTTTTACACACCTCGAAGCAACGAGGACTAACTTTCCGTCCGATC
ACCCAGTTCTTATTCTGAACCTTAGTCGCAGACATAATCATTTTAACATGAATTGGAGGC
ATACCCGTAGAACACCCATATATCATCATTGGCCAAATCGCCTCGGTGCTATACTTCGCA
CTCTTCCTTGTTCTCGCCCCCCTCGCAGGGTGAGCCGAGAATAAAGCATTGAAATGAGCT

H21
ATGGCAAGCCTACGAAAAACCCATCCGCTAATAAAAATTGCTAACAATGCACTAGTTGAC
CTTCCAACCCCATCTAATATTTCAGCACTATGGAACTTCGGGTCCCTCCTAGGATTATGC
TTAATTACTCAAATCCTAACAGGACTATTCCTCGCCATACATTATACTTCTGATATCTCA
ACTGCCTTTTCATCAGTAACACATATCTGCCGAGACGTAAACTATGGCTGACTCATCCGA
AACATACATGCTAACGGAGCATCGTTCTTCTTTATCTGTATTTACATACACATTGCCCGC
GGACTGTACTACGGGTCATACCTCTATAAAGAGACCTGAAGTATTGGCGTAGTTCTACTC
CTTTTAGTAATGATGACAGCCTTTGTGGGCTATGTACTGCCCTGAGGACAAATATCTTTT
TGAGGTGCCACCGTGATTACAAATCTACTATCAGCAGTCCCCTATATGGGCGACGCCTTT
GTCCAATGGATTTGAGGCGGCTTTTCCGTAGATAACGCAACATTAACACGATTCTTCGCC
TTCCACTTCCTATTCCCGTTCGTTATCGCCGGTGCAACGGTCCTACATTTGCTATTTTTA
CACGAAACGGGATCAAACAACCCTGCCGGATTAAACTCCGACGCAGATAAAATCTCTTTC
CACCCATACTTCTCATACAAGGACCTCCTTGGCTTTGTACTAATGTTATTAGCTCTCACA
TCCCTAACATTATTCTCCCCAACACTGCTCGGCGACCCAGAAAATTTCACCCCGGCAAAC
CCCCTGGTCACCCCACCGCACATCCAACCGGAGTGATACTTCCTGTTTGCCTACGCCATT
CTACGATCTATTCCAAACAAGCTAGGAGGGGTCCTAGCGCTATTGTTTAGCATCCTAGTA
CTATTAGTAGTGCCAGTTTTACACACCTCGAAGCAACGAGGACTAACTTTCCGTCCGATC
ACCCAGTTCTTATTCTGAACCTTAATCGCAGACATAATCATTTTAACATGAATTGGAGGC
ATACCCGTAGAACACCCATATATCATCATTGGCCAAATCGCCTCGGTGCTATACTTCGCA
CTCTTCCTTGTTCTCGCCCCCCTCGCAGGGTGAGCCGAGAATAAAGCATTGAAATGAGCT

H22
ATGGCAAGCCTACGAAAAACCCATCCGCTAATAAAAATTGCTAACAATGCACTAGTTGAC
CTTCCAACCCCATCTAATATTTCAGCACTATGGAACTTCGGGTCCCTCCTAGGATTATGC
TTAATTACTCAAATCCTAACAGGACTATTCCTGGCCATACATTATACTTCTGATATCTCA
ACTGCCTTTTCATCAGTAACACATATCTGCCGAGACGTAAACTATGGCTGACTCATCCGA
AACATACATGCTAACGGAGCATCGTTCTTCTTTATCTGTATTTACATACACATTGCCCGC
GGACTGTACTACGGGTCATACCTCTATAAAGAGACCTGAAGTATTGGCGTAGTTCTACTC
CTTTTAGTAATGATGACAGCCTTTGTGGGCTATGTACTGCCCTGAGGACAAATATCTTTT
TGAGGTGCCACCGTGATTACAAATCTACTATCAGCAGTCCCCTATATGGGGGACGCCTTT
GTCCAATGGATTTGAGGCGGCTTTTCCGTAGATAACGCAACATTAACACGATTCTTCGCC
TTCCACTTCCTATTCCCGTTCGTTATCGCCGGTGCAACGGTCCTACATTTGCTATTTTTA
CACGAAACGGGATCAAACAACCCTGCCGGATTAAACTCCGACGCAGATAAAATCTCTTTC
CACCCATACTCCTCATACAAGGACCTCCTTGGCTTTGTACTAATGTTATTAGCTTTCACA
TCCCTAACATTATTCTCCCCAACACTGCTCGCCGACCCAGAAAATTTCACCCCGGCAAAC
CCCCTGGTCACCCCACCGCACATCCAACCGGAGTGATACTTCCTGTTTGCCTACGCCATT
CTACGATCTATTCCAAACAAGCTAGGAGGGGTCCTAGCGCTATTGTTTAGCATCCTAGTA
CTATTAGTAGTGCCAGTTTTACACACCTCGAAGCAACGAGGACTAACTTTCCGTCCGATC
ACCCAGTTCTTATTCTGAACCTTAGTCGCAGACATAATCATTTTAACATGAATTGGAGGC
ATACCCGTAGAACACCCATATATCATCATTGGCCAAATCGCCTCGGTGCTATACTTCGCA
CTCTTCCTTGTTCTCGCCCCCCTCGCAGGGTGAGCCGAGAATAAAGCATTGAAATGAGCT

H23
ATGGCAAGCCTACGAAAAACCCATCCGCTAATAAAAATTGCTAACAATGCACTAGTTGAC
CTTCCAACCCCATCTAATATTTCAGCACTATGGAACTTCGGGTCCCTCCTAGGATTATGC
TTAATTACTCAAATCCTAACAGGACTATTCCTGGCCATACATTATACTTCTGATATCTCA
ACTGCCTTTTCATCAGTAACACATATCTGCCGAGACGTAAACTATGGCTGACTCATCCGA

AACATACATGCTAACGGAGCATCGTTCTTCTTTATCTGTATTTACATACACATTGCCCGC
GGACTGTACTACGGGTCATACCTCTATAAAGAGACCTGAAGTATTGGCGTAGTTCTACTC
CTTTTAGTAATGATGACAGCCTTTGTGGGCTATGTACTGCCCTGAGGACAAATATCTTTT
TGAGGTGCCACCGTGATTACAAATCTACTATCAGCAGTCCCCTATATGGGCGACGCCTTT
GTCCAATGGATTTGAGGCGGCTTTTCCGTAGATAACGCAACATTAACACGATTCTTCGCC
TTCCACTTCCTATTCCCGTTCGTTATCGCCGGTGCAACGGTCCTACATTTGCTATTTTTA
CACGAAACGGGATCAAACAACCCTGCCGGATTAAACTCCGACGCAGATAAAATCTCTTTC
CACCCATACTTCTCATACAAGGACCTCCTTGGCTTTGTACTAATGTTATTAGCTCTCACA
TCCCTAACATTATTCTCCCCAACACTGCTCGCCGACCCAGAAAATTTCACCCCGCCAAAC
CCCCTGGTCACCCCACCGCACATCCAACCGGAGTGATACTTCCTGTTTGCCTACGCCATT
CTACGATCTATTCCAAACAAGCTAGGAGGGGTCCTAGCGCTATTGTTTAGCATCCTAGTA
CTATTAGTAGTGCCAGTTTTACACACCTCGAAGCAACGAGGACTAACTTTCCGTCCGATC
ACCCAGTTCTTATTCTGAACCTTAGTCGCAGACATAATCATTTTAACATGAATTGGAGGC
ATACCCGTAGAACACCCATATATCATCATTGGCCAAATCGCCTCGGTGCTATACTTCGCA
CTCTTCCTTGTTCTCGCCCCCCTCGCAGGGTGAGCCGAGAATAAAGCATTGAAATGAGCT

H24
ATGGCAAGCCTACGAAAAACCCATCCGCTAATAAAAATTGCTAACAATGCACTAGTTGAC
CTTCCAACCCCATCTAATATTTCAGCACTATGGAACTTCGGGTCCCTCCTAGGATTATGC
TTAATTACTCAAATCCTAACAGGACTATTCCTGGCCATACATTATACTTCTGATATCTCA
ACTGCCTTTTCATCAGTAACACATATCTGCCGAGACGTAAACTATGGCTGACTCATCCGA
AACATACATGCTAACGGAGCATCGTTCTTCTTTATCTGTATTTACATACACATTGCCCGC
GGACTGTACTACGGGTCATACCTCTATAAAGAGACCTGAAGTATTGGCGTAGTTCTACTC
CTTTTAGTAATGATGACAGCCTTTGTGGGCTATGTACTGCCCTGAGGACAAATATCTTTT
TGAGGTGCCACCGTGATTACAAATCTACTATCAGCAGTCCCCTATATGGGCGACGCCTTT
GTCCAATGGATTTGAGGCGGCTTTTCCGTAGATAACGCAACATTAACACGATTCTTCGCC
TTCCACTTCCTATTCCCGTTCGTTATCGCCGGTGCAACGGTCCTACATTTGCTATTTTTA
CACGAAACGGGATCAAACAACCCTGCCGGATTAAACTCCGACGCAGATAAAATCTCTTTC
CACCCATACTTCTCATACAAGGACCTCCTTGGCTTTGTACTAATGTTATTAGCTCTCACA
TCCCTAACATTATTCTCCCCAACACTGCTCGCCGACCCAGAAAATTTCACCCCGGCAAAC
CCCCTGGTCACCCCACCGCACATCCAACCGGAGTGATACTTCCTGTTTGCCTACGCCATT
CTACGATCTATTCCAAACAAGCTAGGAGGGGTCCTAGCGCTATTGTTTAGCATCCTAGTA
CTATTAGTAGTGCCAGTTTTACACACCTCGAAGCAACGAGGACTAACTTTCCGTCCGATC
ACCCAGTTCTTATTCTGAACCTTAGTCGCAGACATAATCATTTTAACATGAATTGGAGGC
ATACCCGTAGAACACCCATATATCATCATTGGCCAAATCGCCTCGGTGCTATACTTCGCA
CTCTTCCTTGTTCTCGCCCCCCTCGCAGGGTGAGCCGAGAATAAAGCATTGAAATGAGCT

H25
ATGGCAAGCCTACGAAAAACCCATCCGCTAATAAAAATTGCTAACAATGCACTAGTTGAC
CTTCCAACCCCATCTAATATTTCAGCACTATGGAACTTCGGGTCCCTCCTAGGATTATGC
TTAATTACTCAAATCCTAACAGGACTATTCCTGGCCATACATTATACTTCTGATATCTCA
ACTGCCTTTTCATCAGTAACACATATCTGCCGAGACGTAAACTATGGCTGACTCATCCGA
AACATACATGCTAACGGAGCATCGTTCTTCTTTATCTGTATTTACATACACATTGCCCGC
GGACTGTACTACGGGTCATACCTCTATAAAGAGACCTGAAGTATTGGCGTAGTTCTACTC
CTTTTAGTAATGATGACAGCCTTTGTGGGCTATGTACTGCCCTGAGGACAAATATCTTTT
TGAGGTGCCACCGTGATTACAAATCTACTATCAGCAGTCCCCTATATGGGCGACGCCTTT
GTCCAATGGATTTGAGGCGGCTTTTCCGTAGATAACGCAACATTAACACGATTCTTCGCC
TTCCACTTCCTATTCCCGTTCGTTATCGCCGGTGCAACGGTCCTACATTTGCTATTTTTA
CACGAAACGGGATCAAACAACCCTGCCGGATTAAACTCCGACGCAGATAAAATCTCTTTC
CACCCATACTTCTCATACAAGGACCTCCTTGGCTTTGTACTAATGTTATTAGCTCTCACA
TCCCTAACATTATTCTCCCCAACACTGCTCGGCGACCCAGAAAATTTCACCCCGCCAAAC
CCCCTGGTCACCCCACCGCACATCCAACCGGAGTGATACTTCCTGTTTGCCTACGCCATT
CTACGATCTATTCCAAACAAGCTAGGAGGGGTCCTAGCGCTATTGTTTAGCATCCTAGTA

CTATTAGTAGTGCCAGTTTTACACACCTCGAAGCAACGAGGACTAACTTTCCGTCCGATC
ACCCAGTTCTTATTCTGAACCTTAGTCGCAGACATAATCATTTTAACATGAATTGGAGGC
ATACCCGTAGAACACCCATATATCATCATTGGCCAAATCGCCTCGGTGCTATACTTCGCA
CTCTTCCTTGTTCTCGCCCCCCTCGCAGGGTGAGCCGAGAATAAAGCATTGAAATGAGCT

H26
ATGGCAAGCCTACGAAAAACCCATCCGCTAATAAAAATTGCTAACAATGCACTAGTTGAC
CTTCCAACCCCATCTAATATTTCAGCACTATGGAACTTCGGGTCCCTCCTAGGATTATGC
TTAATTACTCAAATCCTAACAGGACTATTCCTCGCCATACATTATACTTCTGATATCTCA
ACTGCCTTTTCATCAGTAACACATATCTGCCGAGACGTAAACTATGGCTGACTCATCCGA
AACATACATGCTAACGGAGCATCGTTCTTCTTTATCTGTATTTACATACACATTGCCCGC
GGACTGTACTACGGGTCATACCTCTATAAAGAGACCTGAAGTATTGGCGTAGTTCTACTC
CTTTTAGTAATGATGACAGCCTTTGTGGGCTATGTACTGCCCTGAGGACAAATATCTTTT
TGAGGTGCCACCGTGATTACAAATCTACTATCAGCAGTCCCCTATATGGGCGACGCCTTT
GTCCAATGGATTTGAGGCGGCTTTTCCGTAGATAACGCAACATTAACACGATTCTTCGCC
TTCCACTTCCTATTCCCGTTCGTTATCGCCGGTGCAACGGTCCTACATTTGCTATTTTTA
CACGAAACGGGATCAAACAACCCTGCCGGATTAAACTCCGACGCAGATAAAATCTCTTTC
CACCCATACTTCTCATACAAGGACCTCCTTGGCTTTGTACTAATGTTATTAGCTCTCACA
TCCCTAACATTATTCTCCCCAACACTGCTCGGCGACCCAGAAAATTTCACCCCGGCAAAC
CCCCTGGTCACCCCACCGCACATCCAACCGGAGTGATACTTCCTGTTTGCCTACGCCATT
CTACGATCTATTCCAAACAAGCTAGGAGGGGTCCTAGCGCTATTGTTTAGCATCCTAGTA
CTATTAGTAGTGCCAGTTTTACACACCTCGAAGCAACGAGGACTAACTTTCCGTCCGATC
ACCCAGTTCTTATTCTGAACCTTAATCGCAGACATAATCATTCTAACATGAATTGGAGGC
ATACCCGTAGAACACCCATATATCATCATTGGCCAAATCGCCTCGGTGCTATACTTCGCA
CTCTTCCTTGTTCTCGCCCCCCTCGCAGGGTGAGCCGAGAATAAAGCATTGAAATGAGCT

H27
ATGGCAAGCCTACGAAAAACCCATCCGCTAATAAAAATTGCTAACAATGCACTAGTTGAC
CTTCCAACCCCATCTAATATTTCAGCACTATGGAACTTCGGGTCCCTCCTAGGATTATGC
TTAATTACTCAAATCCTAACAGGACTATTCCTCGCCATACATTATACTTCTGATATCTCA
ACTGCCTTTTCATCAGTAACACATATCTGCCGAGACGTAAACTATGGCTGACTCATCCGA
AACATACATGCTAACGGAGCATCGTTCTTCTTTATCTGTATTTACATACACATTGCCCGC
GGACTGTACTACGGGTCATACCTCTATAAAGAGACCTGAAGTATTGGCGTAGTTCTACTC
CTTTTAGTAATGATGACAGCCTTTGTGGGCTATGTACTGCCCTGAGGACAAATATCTTTT
TGAGGTGCCACCGTGATTACAAATCTACTATCAGCAGTCCCCTATATGGGCGACGCCTTT
GTCCAATGGATTTGAGGCGGCTTTTCCGTAGATAACGCAACATTAACACGATTCTTCGCC
TTCCACTTCCTATTCCCGTTCGTTATCGCCGGTGCAACGGTCCTACATTTGCTATTTTTA
CACGAAACGGGATCAAACAACCCTGCCGGATTAAACTCCGACGCAGATAAAATCTCTTTC
CACCCATACTTCTCATACAAGGACCTCCTTGGCTTTGTACTAATGTTATTAGCTCTCACA
TCCCTAACATTATTCTCCCCAACACTGCTCGGCGACCCAGAAAATTTCACCCCGCCAAAC
CCCCTGGTCACCCCACCGCACATCCAACCGGAGTGATACTTCCTGTTTGCCTACGCCATT
CTACGATCTATTCCAAACAAGCTAGGAGGGGTCCTAGCGCTATTGTTTAGCATCCTAGTA
CTATTAGTAGTGCCAGTTTTACACACCTCGAAGCAACGAGGACTAACTTTCCGTCCGATC
ACCCAGTTCTTATTCTGAACCTTAATCGCAGACATAATCATTTTAACATGAATTGGAGGC
ATACCCGTAGAACACCCATATATCATCATTGGCCAAATCGCCTCGGTGCTATACTTCGCA
CTCTTCCTTGTTCTCGCCCCCCTCGCAGGGTGAGCCGAGAATAAAGCATTGAAATGAGCT

H28
ATGGCAAGCCTACGAAAAACCCATCCGCTAATAAAAATTGCTAACAATGCACTAGTTGAC
CTTCCAACCCCATCTAATATTTCAGCACTATGGAACTTCGGGTCCCTCCTAGGATTATGC
TTAATTACTCAAATCCTAACAGGACTATTCCTGGCCATACATTATACTTCTGATATCTCA
ACTGCCTTTTCATCAGTAACACATATCTGCCGAGACGTAAACTATGGCTGACTCATCCGA
AACATACATGCTAACGGAGCATCATTCTTCTTTATCTGTATTTACATGCACATTGCCCGC

GGACTGTACTACGGGTCATACCTCTATAAAGAGACCTGAAGTATTGGCGTAGTTCTACTC
CTTTTAGTAATGATGACAGCCTTTGTGGGCTATGTACTGCCCTGAGGACAAATATCTTTT
TGAGGTGCCACCGTGATTACAAATCTACTATCAGCAGTCCCCTATATGGGCGACGCCTTT
GTCCAATGGATTTGAGGCGGCTTTTCCGTAGATAACGCAACATTAACACGATTCTTCGCC
TTCCACTTCCTATTCCCGTTCGTTATCGCCGGTGCAACGGTCCTACATTTGCTATTTTTA
CACGAAACGGGATCAAACAACCCTGCCGGATTAAACTCCGACGCAGATAAAATCTCTTTC
CACCCATACTTCTCATACAAGGACCTCCTTGGCTTTGTACTAATGTTATTAGCTCTCACA
TCCCTAACATTATTCTCCCCAACACTGCTCGGCGACCCAGAAAATTTCACCCCGGCAAAC
CCCCTGGTCACCCCACCGCACATCCAACCGGAGTGATACTTCCTGTTTGCCTACGCCATT
CTACGATCTATTCCAAACAAGCTAGGAGGGGTCCTAGCGCTATTGTTTAGCATCCTAGTA
CTATTAGTAGTGCCAGTTTTACACACCTCGAAGCAACGAGGACTAACTTTCCGTCCGATC
ACCCAGTTCTTATTCTGAACCTTAGTCGCAGACATAATCATTTTAACATGAATTGGAGGC
ATACCCGTAGAACACCCATATATCATCATTGGCCAAATCGCCTCGGTGCTATACTTCGCA
CTCTTCCTTGTTCTCGCCCCCCTCGCAGGGTGAGCCGAGAATAAAGCATTGAAATGAGCT

H29
ATGGCAAGCCTACGAAAAACCCATCCGCTAATAAAAATTGCTAACAATGCACTAGTTGAC
CTTCCAACCCCATCTAATATTTCAGCACTATGGAACTTCGGGTCCCTCCTAGGATTATGC
TTAATTACTCAAATCCTAACAGGACTATTCCTGGCCATACATTATACTTCTGATATCTCA
ACTGCCTTTTCATCAGTAACACATATCTGCCGAGACGTAAACTATGGCTGACTCATCCGA
AACATACATGCTAACGGAGCATCATTCTTCTTTATCTGTATTTACATGCGCATTGCCCGC
GGACTGTACTACGGGTCATACCTCTATAAAGAGACCTGAAGTATTGGCGTAGTTCTACTC
CTTTTAGTAATGATGACAGCCTTTGTGGGCTATGTACTGCCCTGAGGACAAATATCTTTT
TGAGGTGCCACCGTGATTACAAATCTACTATCAGCAGTCCCCTTTATGGGCGACGCCTTT
GTCCAATGGATTTGAGGCGGCTTTTCCGTAGATAACGCAACATTAACACGATTCTTCGCC
TTCCACTTCCTATTCCCGTTCGTTATCGCCGGTGCAACGGTCCTACATTTGCTATTTTTA
CACGAAACGGGATCAAACAACCCTGCCGGATTAAACTCCGACGCAGATAAAATCTCTTTC
CACCCATACTTCTCATACAAGGACCTCCTTGGCTTTGTACTAATGTTATTAGCTCTCACA
TCCCTAACATTATTCTCCCCAACACTGCTCGGCGACCCAGAAAATTTCACCCCGGCAAAC
CCCCTGGTCACCCCACCGCACATCCAACCGGAGTGATACTTCCTGTTTGCCTACGCCATT
CTACGATCTATTCCAAACAAGCTAGGAGGGGTCCTAGCGCTATTGTTTAGCATCCTAGTA
CTATTAGTAGTGCCAGTTTTACACACCTCGAAGCAACGAGGACTAACTTTCCGTCCGATC
ACCCAGTTCTTATTCTGAACCTTAGTCGCAGACATAATCATTTTAACATGAATTGGAGGC
ATACCCGTAGAACACCCATATATCATCATTGGCCAAATCGCCTCGGTGCTATACTTCGCA
CTCTTCCTTGTTCTCGCCCCCCTCGCAGGGTGAGCCGAGAATAAAGCATTGAAATGAGCT

H30
ATGGCAAGCCTACGAAAAACTCACCCACTAATAAAAATCGCTAACGACGCACTAGTTGAC
CTTCCAACCCCATCTAATATTTCAGCACTATGAAACTTCGGATCCCTCCTAGGATTATGC
TTAATTACTCAAATCTTAACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCA
ACTGCCTTTTCATCAGTAACACACATCTGCCGAGACGTTAACTATGGCTGACTTATCCGA
AACATACATGCTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGC
GGGCTCTACTATGGATCATACCTCTATAAAGAAACCTGAAACATTGGCGTAGTTCTACTT
CTTTTAGTTATGATGACAGCCTTTGTAGGCTATGTATTACCATGAGGACAAATATCTTTT
TGGGGTGCTACCGTGATTACAAATCTACTATCAGCAGTCCCTTATATAGGAGACGCCCTT
GTCCAATGGATTTGAGGTGGCTTTTCCGTAGATAACGCAACGTTAACACGATTCTTCGCC
TTCCACTTCTTATTTCCGTTCGTTATCGCCGGTGCAACGATTCTGCATTTGCTATTTCTA
CACGAAACGGGATCAAACAACCCTGCCGGATTAAACTCCGACGCGGATAAAATCTCTTTC
CACCCATACTTCTCATACAAGGACCTCCTAGGCTTTGTGCTAATATTATTAGCCCTCACA
TCCCTAACATTATTCTCACCAACACTGCTCGGCGACCCAGAAAACTTCACCCCGGCAAAC
CCCCTGGTTACCCCACCACACATCCAACCGGAGTGGTACTTCTTGTTTGCCTACGCCATT
CTACGGTCTATTCCAAACAAGCTAGGAGGAGTCCTAGCGTTGTTATTTAGCATTTTAGTA
CTTTTAGTGGTGCCAATTTTACATACCTCGAAGCAGCGAGGACTAACTTTCCGTCCAATC

ACCCAGTTCTTATTCTGAACCTTGGTGGCAGACATAATCATTTTGACATGAATCGGGGGC
ATACCCGTAGAACACCCATATATCATCATTGGCCAAATCGCGTCGGTGCTGTACTTCGCA
CTCTTCCTTGTTCTCGCCCCGCTCGCAGGATGAGCCGAGAATAAAGCATTGAAATGAGCT

H31
ATGGCAAGCCTACGAAAAACTCACCCACTAATAAAAATCGCTAACGACGCACTAGTTGAC
CTTCCAACCCCATCTAATATTTCAGCACTATGAAACTTCGGATCCCTTCTAGGATTATGC
TTAATTACTCAAATCTTAACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCA
ACTGCCTTTTCATCAGTAACACACACCTGCCGAGACGTTAACTATGGCTGACTTATCCGA
AACATACATGCTAACGGAGCATCATTCTTCTTTATCTGTATTTACATACACATTGCCCGC
GGGCTCTACTATGGATCATACCTCTATAAAGAAACCTGAAACATTGGCGTAGTTCTACTT
CTTTTAGTTATGATGACAGCCTTTGTAGGCTATGTATTACCATGAGGACAAATATCTTTT
TGGGGTGCTACCGTGATTACAAATCTACTATCAGCAGTCCCTTATATAGGCGACGCCCTT
GTCCAGTGGATTTGAGGTGGCTTTTCCGTAGATAACGCAACGTTAACACGATTCTTCGCC
TTCCACTTCTTATTTCCGTTCGTTATCGCCGGTGCAACAATTCTACATTTGCTATTTCTA
CACGAAACGGGATCAAACAACCCTGCCGGATTAAACTCCGACGCGGATAAAATCTCTTTC
CACCCGTACTTCTCATACAAGGACCTCCTAGGCTTTGTGCTAATATTATTAGCCCTCACA
TCCCTAACATTATTCTCACCAACACTGCTCGGCGACCCAGAAAACTTCACCCCGGCAAAC
CCCCTGGTTACCCCGCCACACATCCAACCGGAGTGGTACTTCTTGTTTGCCTACGCCATT
CTACGGTCTATTCCAAACAAGCTAGGAGGAGTCCTAGCATTGTTATTTAGCATTTTAGTA
CTTTTAGTAGTGCCAATTTTACATACCTCGAAGCAGCGAGGACTAACTTTCCGTCCAATC
ACCCAGTTCTTATTCTGAACCTTGGTGGCAGACATAATCATTTTGACATGAATCGGGGGC
ATACCCGTAGAACACCCATATATCATCATTGGCCAAATCGCGTCGGTGCTGTACTTCGCA
CTCTTCCTTGTTCTCGCCCCGCTCGCAGGATGAGCCGAGAATAAAGCATTGAAATGAGCT

H32
ATGGCAAGCCTACGAAAAACTCACCCACTAATAAAAATCGCTAACGACGCACTAGTTGAC
CTTCCAACCCCATCTAATATTTCAGCACTATGAAACTTCGGATCCCTTCTAGGATTATGC
TTAATTACTCAAATCTTAACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCA
ACTGCCTTTTCATCAGTAACACACATCTGCCGAGACGTTAACTATGGCTGACTTATCCGA
AACATACATGCTAACGGAGCATCATTCTTCTTTATCTGTATTTACATACACATTGCCCGC
GGGCTCTACTATGGATCATACCTCTATAAAGAAACCTGAAACATTGGCGTAGTTCTACTT
CTTTTAGTTATGATGACAGCCTTTGTAGGCTATGTATTACCATGAGGACAAATATCTTTT
TGGGGTGCTACCGTGATTACAAATCTACTATCAGCAGTCCCTTATATAGGCGACGCCTTT
GTCCAGTGGATTTGAGGTGGCTTTTCCGTAGATAACGCAACGTTAACACGATTCTTCGCC
TTCCACTTCTTATTTCCGTTCGTTATCGCCGGTGCAACAATTCTGCATTTGCTATTTCTA
CACGAAACGGGATCAAACAACCCTGCCGGATTAAACTCCGACGCGGATAAAATCTCTTTC
CACCCGTACTTCTCATACAAGGACCTCCTAGGCTTTGTGCTAATATTATTAGCCCTCACA
TCCCTAACATTATTCTCACCAACACTGCTCGGCGACCCAGAAAACTTCACCCCGGCAAAC
CCCCTGGTTACCCCGCCACACATCCAACCGGAGTGGTACTTCTTGTTTGCCTACGCCATT
CTACGGTCTATTCCAAACAAGCTAGGAGGAGTCCTAGCATTGTTATTTAGCATTTTAGTA
CTTTTAGTAGTGCCAATTTTACATACCTCGAAGCAGCGAGGACTAACTTTCCGTCCAATC
ACCCAGTTCTTATTCTGAACCTTGGTGGCAGACATAATCATTTTGACATGAATCGGGGGC
ATACCCGTAGAACACCCATATATCATCATTGGCCAAATCGCGTCGGTGCTGTACTTCGCA
CTCTTCCTTGTTCTCGCCCCGCTCGCAGGATGAGCCGAGAATAAAGCATTGAAATGAGCT

H33
ATGGCAAGCCTACGAAAAACTCACCCACTAATAAAAATCGCTAACGACGCACTAGTTGAC
CTTCCAACCCCATCTAATATTTCAGCACTATGAAACTTCGGATCCCTCCTAGGATTATGC
TTAATTACTCAAATCTTAACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCA
ACTGCCTTTTCATCAGTAACACACATCTGCCGAGACGTTAACTATGGCTGACTTATCCGA
AACATACATGCTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGC
GGGCTCTACTATGGATCATACCTCTATAAAGAAACCTGAAACATTGGCGTAGTTCTACTT

CTTTTAGTTATGATGACAGCCTTTGTAGGCTATGTATTACCATGAGGACAAATATCTTTT
TGGGGCGCTACCGTGATTACAAATCTACTATCAGCAGTCCCTTATATAGGCGACGCCTTT
GTCCAATGGATTTGAGGTGGCTTTTCCGTAGATAACGCAACGTTAACACGATTCTTCGCC
TTCCACTTCTTGTTTCCGTTCGTTATCGCCGGTGCAACAATTCTGCATTTGCTATTTCTA
CACGAAACGGGATCAAACAACCCTGCCGGATTAAACTCCGACGCGGATAAAATCTCTTTC
CACCCGTACTTCTCATACAAGGACCTCCTAGGCTTTGTGCTAATATTATTAGCCCTCACA
TCCCTAACGTTATTCTCACCAACACTGCTCGGCGACCCAGAAAACTTCACCCCGGCAAAC
CCCCTGGTTACCCCGCCACACATCCAACCGGAGTGGTACTTCTTGTTTGCCTACGCCATT
CTACGGTCTATTCCAAACAAGCTAGGAGGAGTCCTAGCATTGTTATTTAGCATTTTAGTA
CTTTTAGTAGTGCCAATTTTACATACCTCGAAGCAGCGAGGACTAACTTTCCGTCCAATC
ACCCAGTTCTTATTCTGAACCTTGGTGGCAGACATAATCATTTTGACATGAATCGGGGGC
ATACCCGTAGAACACCCATATATCATCATTGGCCAAATCGCGTCGGTGCTGTACTTCGCA
CTCTTCCTTGTTCTCGCCCCGCTCGCAGGATGAGCCGAGAATAAAGCATTGAAATGAGCT

H34
ATGGCAAGCCTACGAAAAACTCACCCACTAATAAAAATCGCTAACGACGCACTAGTTGAC
CTTCCAACCCCATCTAATATTTCAGCACTATGAAACTTCGGATCCCTCCTAGGATTATGC
TTAATTACTCAAATCTTAACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCA
ACTGCCTTTTCATCAGTAACACACATCTGCCGAGACGTTAACTATGGCTGACTTATCCGA
AACATACATGCTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGC
GGGCTCTACTATGGATCATACCTCTATAAAGAAACCTGAAACATTGGCGTAGTTCTACTT
CTTTTAGTTATGATGACAGCCTTTGTAGGCTATGTATTACCATGAGGACAAATATCTTTT
TGGGGCGCTACCGTGATTACAAATCTACTATCAGCAGTCCCTTATATAGGCGACGCCTTT
GTCCAATGGATTTGAGGTGGCTTTTCCGTAGATAACGCAACGTTAACACGATTCTTCGCC
TTCCACTTCTTGTTTCCGTTCGTTATCGCCGGTGCAACAATTCTGCATTTGCTATTTCTA
CACGAAACGGGATCAAACAACCCTGCCGGATTAAACTCCGACGCGGATAAAATCTCTTTC
CACCCATACTTCTCATACAAGGACCTCCTAGGCTTTGTGCTAATATTATTAGCCCTCACA
TCCCTAACGTTATTCTCACCAACACTGCTCGGCGACCCAGAAAACTTCACCCCGGCAAAC
CCCCTGGTTACCCCACCGCACATCCAACCGGAGTGGTACTTCTTGTTTGCCTACGCCATT
CTACGGTCTATTCCAAACAAGCTAGGAGGAGTCCTAGCATTGTTATTTAGCATTTTAGTA
CTTTTAGTAGTGCCAATTTTACATACCTCGAAGCAGCGAGGACTAACTTTCCGTCCAATC
ACCCAGTTCTTATTCTGAACCTTGGTGGCAGACATAATCATTTTGACATGAATCGGGGGC
ATACCCGTAGAACACCCATATATCATCATTGGCCAAATCGCGTCGGTGCTGTACTTCGCA
CTCTTCCTTGTTCTCGCCCCGCTCGCAGGATGAGCCGAGAATAAAGCATTGAAATGAGCT

AZT1
ATGGCAAGCCTACGAAAAACCCACCCACTAATAAAAATTGCTAACGATGCACTAGTCGAC
CTTCCAACTCCATCTAATATCTCAGCACTATGGAACTTCGGATCCCTCCTAGGATTATGC
TTGATTACGCAAATCCTAACAGGACTGTTCCTGGCCATGCATTATACTTCCGATATCTCA
ACTGCCTTTTCGTCAGTAACACACATCTGCCGAGACGTTAACTATGGCTGACTCATCCGA
AACATACATGCTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGC
GGACTGTACTATGGGTCTTACCTCTATAAAGAGACCTGAACCATTGGCGTGGTCCTACTC
CTTTTGGTCATAATAACGGCCTTCGTAGGCTATGTGCTACCGTGGGGTCAAATATCTTTT
TGGGGCGCCACCGTCATTACAAACCTACTATCAGCAGTCCCTTATATGGGAGACGCCCTT
GTCCAGTGGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCC
TTCCACTTCCTGTTTCCGTTCGTTATCGCCGGTGCGACGGTCCTGCATTTGCTATTTTTA
CACGAAACGGGATCAAACAACCCTGCCGGATTAAACTCCGACGCGGATAAAATCTCTTTC
CACCCATACTTCTCGTACAAAGATCTCCTTGGCTTTGTACTAATGTTATTAGCTCTCACA
TCCCTGACATTATTCTCCCCAACTCTGCTCGGCGACCCAGAAAATTTCACCCCAGCAAAC
CCCCTGGTTACCCCGCCACACATCCAACCGGAGTGATACTTCTTATTTGCCTACGCCATT
CTACGCTCTATTCCAAACAAGCTGGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTA
CTATTAGTGGTGCCAATTTTACATACCTCAAAGCAGCGAGGACTAACTTTCCGTCCAATC
ACCCAGTTCTTATTCTGAACCTTGGTGGCAGACATAATCATTTTGACATGAATTGGAGGC

ATGCCCGTAGAACACCCATACATCATCATCGGCCAAATCGCGTCGGTGCTGTACTTCGCA
CTCTTCCTTCTTCTCGCCCCCCTCGCAGGGTGAGCCGAGAATAAAGCATTGAAATGAGCT

BP11
ATGGCAAGCCTACGAAAAACCCACCCACTAATAAAAATTGCTAACGATGCACTAGTCGAC
CTTCCAACTCCATCTAATATCTCAGCACTATGGAACTTCGGATCCCTCCTAGGATTATGC
TTGATTACGCAAATCCTAACAGGACTGTTCCTGGCCATGCATTATACTTCCGATATCTCA
ACTGCCTTTTCGTCAGTAACACACATCTGCCGAGACGTTAACTATGGCTGACTCATCCGA
AACATACATGCTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGC
GGACTGTACTATGGGTCTTACCTCTATAAAGAGACCTGAACCATTGGCGTGGTCCTACTC
CTTTTGGTCATAATAACGGCCTTCGTAGGCTATGTGCTACCGTGGGGTCAAATATCTTTT
TGGGGCGCCACCGTCATTACAAACCTACTATCAGCAGTCCCTTATATGGGAGACGCCCTT
GTCCAGTGGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCC
TTCCACTTCCTGTTTCCGTTCGTTATCGCCGGTGCGACGGTCCTGCACTTGCTATTTTTA
CACGAAACGGGATCAAACAACCCTGCCGGATTAAACTCCGACGCGGATAAAATCTCTTTC
CACCCATACTTCTCGTACAAAGATCTCCTTGGCTTTGTACTAATGTTATTAGCTCTCACA
TCCCTGACATTATTCTCCCCAACTCTGCTCGGCGACCCAGAAAATTTCACCCCAGCAAAC
CCCCTGGTTACCCCGCCACACATCCAACCGGAGTGATACTTCTTATTTGCCTACGCCATT
CTACGCTCTATTCCAAACAAGCTGGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTA
CTATTAGTGGTGCCAATTTTACATACCTCAAAGCAGCGAGGACTAACTTTCCGTCCAATC
ACCCAGTTCTTATTCTGAACCTTGGTGGCAGACATAATCATTTTGACATGAATTGGAGGC
ATGCCCGTAGAACACCCATACATCATCATCGGCCAAATCGCGTCGGTGCTGTACTTCGCA
CTCTTCCTTCTTCTCGCCCCCCTCGCAGGGTGAGCCGAGAATAAAGCATTGAAATGAGCT

BONSV1
ATGGCAAGCCTACGAAAAACCCACCCACTAATAAAAATTGCTAACGATGCACTAGTCGAC
CTTCCAACTCCATCTAATATCTCAGCACTATGGAACTTCGGATCCCTCCTAGGATTATGC
TTGATTACGCAAATCCTAACAGGACTGTTCCTAGCCATGCATTATACTTCCGATATCTCA
ACTGCCTTTTCGTCAGTAACACACATCTGCCGAGACGTTAACTATGGCTGACTCATCCGA
AACATACATGCTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGC
GGACTGTACTATGGGTCTTACCTCTATAAAGAGACCTGAACCATTGGTGTGGTCCTACTC
CTTTTGGTCATAATAACGGCCTTCGTAGGCTATGTGCTACCGTGGGGTCAAATATCTTTT
TGGGGCGCCACCGTCATTACAAACCTATTATCAGCAGTCCCTTATATGGGAGACGCCCTT
GTCCAGTGGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCC
TTCCACTTCCTGTTTCCGTTCGTTATCGCCGGTGCGACGGTCCTGCATTTGCTATTTTTA
CACGAAACCGGATCAAACAACCCTGCCGGATTAAACTCCGACGCGGATAAAATCTCTTTC
CACCCATACTTCTCGTACAAAGATCTCCTTGGCTTTGTACTAATGTTATTAGCTCTCACA
TCCCTGACATTATTCTCCCCAACTCTGCTTGGCGACCCAGAAAATTTCACCCCGGCAAAC
CCCCTGGTTACCCCGCCACACATCCAACCGGAGTGATACTTCTTATTTGCCTACGCCATT
CTACGTTCTATTCCAAACAAGCTGGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTA
CTATTAGTCGTGCCAATTTTACATACCTCAAAGCAGCGAGGACTAACTTTCCGTCCAATC
ACCCAGTTCTTATTCTGGACCTTGGTGGCAGACATAATCATTTTGACATGAATTGGAGGC
ATGCCCGTAGAACACCCATACATCATCATCGGCCAAATCGCGTCGGTGCTGTACTTCGCA
CTCTTCCTTCTTCTCGCCCCCCTCGCAGGGTGAGCCGAGAATAAAGCATTGAAATGAGCT

BONPK1
ATGGCAAGCCTACGAAAAACCCACCCACTAATAAAAATTGCTAACGATGCACTAGTCGAC
CTTCCAACCCCATCTAATATCTCAGCACTATGGAACTTCGGATCCCTCCTAGGATTATGC
TTGATTACGCAAATCCTAACAGGACTGTTCCTGGCCATGCATTATACTTCCGATATCTCA
ACTGCCTTTTCGTCAGTAACACACATCTGCCGAGACGTTAACTATGGCTGACTCATCCGA
AACATACATGCTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGC
GGACTGTACTATGGGTCTTACCTCTATAAAGAGACCTGAACCATTGGTGTGGTCCTACTC
CTTTTGGTCATAATAACGGCCTTCGTAGGCTATGTGCTACCGTGGGGCCAAATATCTTTT

TGGGGCGCCACCGTCATTACAAACCTACTATCAGCAGTCCCTTATATGGGAGACGCCCTT
GTCCAGTGGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCC
TTCCACTTCCTGTTTCCGTTCGTTATCGCCGGTGCGACAGTCCTGCATTTGCTATTTTTA
CACGAAACGGGATCAAACAACCCTGCCGGATTAAACTCCGACGCGGATAAAATCTCTTTC
CACCCATACTTCTCGTACAAAGATCTCCTTGGCTTTGTACTAATGTTATTAGCTCTTACA
TCCCTGACATTATTCTCCCCAACTCTGCTCGGCGACCCAGAAAATTTCACCCCGGCAAAC
CCCCTGGTTACCCCGCCACACATCCAACCGGAGTGATACTTCTTATTTGCCTACGCCATT
CTACGTTCTATTCCAAACAAGCTGGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTA
CTATTAGTCGTGCCAATTTTACATACCTCAAAGCAGCGAGGACTAACTTTCCGTCCAATC
ACCCAGTTCTTATTCTGGACCTTGGTGGCAGACATAATCATTTTGACATGAATTGGAGGC
ATGCCCGTAGAACACCCATACATCATCATCGGCCAAATCGCGTCGGTGCTGTACTTCGCA
CTCTTCCTTCTTCTCGCCCCCCTCGCAGGGTGAGCCGAGAATAAAGCATTGAAATGAGCT

EAG1
ATGGCAAGCCTACGAAAAACCCACCCACTAATAAAAATTGCTAACGATGCACTAGTTGAC
CTTCCGACCCCATCTAATATCTCAGCACTATGGAACTTCGGATCCCTCCTAGGATTATGC
TTGATTACTCAAATCCTAACAGGACTATTCCTCGCCATGCATTATACTTCCGATATCTCA
ACTGCCTTTTCGTCGGTAACACACATCTGCCGAGACGTTAACTATGGTTGACTCATCCGA
AACATACATGCTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGC
GGACTGTACTATGGGTCATACCTCTACAAAGCGACCTGAACTATTGGCGTGGTCCTACTC
CTTTTGGTCATAATAACTGCCTTTGTAGGCTATGTGCTGCCCTGAGGCCAAATATCTTTC
TGGGGTGCCACCGTCATTACAAATCTACTATCAGCAGTCCCTTATATAGGAGACGCCCTT
GTTCAGTGGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCC
TTCCACTTCCTGTTTCCGTTCGTTATCGCCGGTGCGACGGTCCTGCATTTGCTATTTCTA
CACGAAACGGGGTCAAACAACCCTGCCGGATTAAACTCCGACGCGGATAAAATCTCTTTC
CACCCATACTTCTCATACAAAGATCTCCTTGGCTTTGTACTAATGTTATTAGCTCTTACA
TCCCTGACATTATTCTCCCCAACTCTGCTCGGTGACCCAGAAAATTTCACCCCAGCAAAC
CCCCTGGTTACCCCGCCACACATCCAACCGGAGTGATACTTCTTATTTGCCTACGCCATT
CTGCGCTCTATTCCAAACAAGCTAGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTA
CTGTTAGTCGTGCCAATTTTACATACCTCAAAGCAGCGAGGACTAACTTTCCGCCCAATC
ACCCAGTTCCTATTCTGAACCTTAGTGGCAGACATAATCATTTTAACATGAATTGGAGGC
ATACCCGTAGAACACCCATATATCATCATTGGCCAAATCGCGTCGGTGCTGTACTTCGCA
CTCTTCCTTCTTCTCGCCCCCCTCGCAGGGTGAGCCGAGAATAAAGCATTGAAATGAGCT

LSR1
ATGGCAAGCCTACGAAAAACCCACCCACTAATAAAAATTGCTAACGATGCACTAGTCGAC
CTTCCAACTCCATCTAATATCTCAGCACTATGGAACTTCGGATCCCTCCTAGGATTATGC
TTGATTACGCAAATCCTAACAGGACTGTTCCTGGCCATGCATTATACTTCCGATATCTCA
ACTGCCTTTTCGTCAGTAACACACATCTGCCGAGACGTTAACTATGGCTGACTCATCCGA
AACATACATGCTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGC
GGACTGTACTATGGGTCTTACCTCTATAAAGAGACCTGAACCATTGGCGTGGTCCTACTC
CTTTTGGTCATAATAACGGCCTTCGTAGGCTATGTGCTACCGTGGGGTCAAATATCTTTT
TGGGGCGCCACCGTCATTACAAACCTACTATCAGCAGTCCCTTATATGGGAGACGCCCTT
GTCCAGTGGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCC
TTCCACTTCCTGTTTCCGTTCGTTATCGCCGGTGCGACGGTCCTGCATTTGCTATTTTTA
CACGAAACGGGATCAAACAACCCTGCCGGATTAAACTCCGACGCGGATAAAATCTCTTTC
CACCCATACTTCTCGTACAAAGATCTCCTTGGCTTTGTACTAATGTTATTAGCTCTCACA
TCCCTGACATTATTCTCCCCAACTCTGCTCGGCGACCCAGAAAATTTCACCCCGGCAAAC
CCCCTGGTTACCCCGCCACACATCCAACCGGAGTGATACTTCTTATTTGCCTACGCCATT
CTACGCTCTATTCCAAACAAGCTGGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTA
CTATTAGTGGTGCCAATTTTACATACCTCAAAGCAGCGAGGACTAACTTTCCGTCCAATC
ACCCAGTTCTTATTCTGAACCTTGGTGGCAGATATAATCATTTTGACATGAATTGGAGGC
ATGCCCGTAGAACACCCATACATCATCATCGGCCAAATCGCGTCGGTGCTGTACTTCGCA

CTCTTCCTTCTTCTCGCCCCCCTCGCAGGGTGAGCCGAGAATAAAGCATTGAAATGAGCT

RAV1
ATGGCAAGCCTACGAAAAACCCACCCACTAATAAAAATTGCTAACGATGCACTAGTTGAC
CTTCCGACCCCATCTAATATCTCAGCACTATGGAACTTCGGATCCCTCCTAGGATTATGC
TTGATTACTCAAATCCTAACAGGACTATTCCTGGCCATGCATTATACTTCCGATATCTCA
ACTGCCTTTTCGTCGGTAACACACATCTGCCGAGACGTTAACTATGGTTGACTCATCCGG
AACATACATGCTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGC
GGACTGTACTATGGGTCATACCTCTACAAAGCGACCTGAACTATTGGCGTGGTCCTACTC
CTTTTGGTTATAATAACTGCCTTTGTAGGCTATGTGCTGCCCTGAGGCCAAATATCTTTC
TGGGGTGCCACCGTCATTACAAATCTACTATCAGCAGTCCCTTATATAGGAGACGCCCTT
GTTCAGTGGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCC
TTCCACTTCCTGTTTCCGTTCGTTATCGCCGGTGCGACGGTCCTGCATTTGCTATTTCTA
CACGAAACGGGATCAAACAACCCTGCCGGATTAAACTCCGACGCGGATAAAATCTCTTTC
CACCCATACTTCTCATACAAAGATCTCCTTGGCTTTGTACTAATGTTATTAGCTCTTACA
TCCCTGACATTATTCTCCCCAACTCTGCTCGGTGACCCAGAAAATTTCACCCCAGCAAAC
CCCCTGGTTACCCCGCCACACATCCAACCGGAGTGATACTTCTTATTTGCCTACGCCATT
CTGCGCTCTATTCCAAACAAGCTGGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTA
CTGTTAGTCGTGCCAATTTTACATACCTCAAAGCAGCGAGGACTAACTTTCCGCCCAATC
ACCCAGTTCCTATTCTGAACCTTAGTGGCAGACATAATCATTTTAACATGAATTGGAGGC
ATACCCGTAGAACACCCATATATCATCATTGGCCAAATCGCATCGGTGCTGTACTTCGCA
CTCTTCCTTCTTCTCGCCCCCCTCGCAGGGTGAGCCGAGAATAAAGCATTGAAATGAGCT

SON1
ATGGCAAGCCTACGAAAAACCCACCCACTAATAAAAATTGCTAACGATGCACTAGTCGAC
CTTCCGACCCCATCCAATATCTCAGCACTATGGAACTTCGGATCCCTCCTAGGATTATGC
TTGATTACTCAAATCCTAACAGGACTATTCCTGGCCATGCATTATACTTCCGATATCTCA
ACTGCCTTTTCGTCGGTAACACACATCTGCCGAGACGTTAACTATGGCTGGCTCATCCGA
AACATACATGCTAACGGAGCATCATTCTTCTTCATCTGTATTTATATACACATTGCCCGC
GGACTGTACTATGGGTCATACCTCTATAAAGCGACCTGAACTATTGGCGTGGTCCTACTC
CTTTTGGTCATAGTAACGGCCTTTGTAGGCTATGTGCTGCCCTGAGGCCAAATATCTTTC
TGGGGCGCCACCGTCATTACAAATCTACTATCAGCAGTCCCTTATATGGGAGACGCCCTT
GTCCAGTGGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCC
TTCCACTTCCTATTTCCGTTCGTTATCGCCGGTGCGACGGTCCTGCATTTGCTATTTTTA
CACGAAACGGGATCAAGCAACCCTGCCGGATTAAACTCCGACGCGGATAAAATCTCTTTC
CACCCATACTTCTCATACAAAGATCTCCTTGGCTTTGTACTAATGTTATTAGCTCTTACA
TCCCTGACATTATTCTCCCCAACTCTGCTCGGTGACCCAGAAAATTTTACCCCAGCAAAC
CCCCTGGTTACCCCGCCACATATCCAACCGGAGTGATACTTCTTATTTGCCTACGCCATT
CTACGCTCTATTCCAAACAAGCTGGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTA
CTGTTAGTCGTGCCAATTTTACATACCTCAAAGCAGCGAGGACTAACTTTCCGGCCAATC
ACCCAGTTCTTATTCTGAACCTTAGTGGCAGACATAATCATTTTAACATGAATTGGAGGC
ATACCCGTAGAACACCCATACATCATCATCGGCCAAATCGCGTCGGTGCTGTACTTCGCA
CTCTTCCTTCTTCTCGCCCCCCTCGCAGGGTGAGCCGAGAATAAAGCATTGAAATGAGCT

TC2
ATGGCAAGCCTACGAAAAACCCACCCGCTAATAAAAATTGCTAACGATGCACTAGTCGAC
CTTCCAACTCCATCTAATATCTCAGCACTATGGAACTTCGGGTCCCTCCTAGGATTATGC
TTGATTACGCAAATCCTAACAGGGCTGTTCCTAGCCATGCATTATACTTCCGATATCTCA
ACTGCCTTTTCGTCGGTGACACACATCTGCCGAGACGTTAACTATGGCTGACTCATCCGA
AACATACATGCTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGC
GGACTGTACTATGGGTCTTACCTCTATAAAGAGACCTGAACCATTGGTGTAGTCCTACTC
CTTTTGGTCATAATAACAGCCTTCGTAGGCTATGTGCTACCCTGGGGCCAAATATCTTTT
TGGGGCGCCACCGTCATTACAAACCTACTATCAGCAGTCCCTTATATAGGAGACGCCCTT

GTCCAGTGGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCC
TTCCACTTCCTGTTTCCGTTCGTTATCGCCGGTGCGACGGTCCTGCATTTGCTATTTTTA
CACGAAACGGGATCAAACAACCCTGCCGGATTAAACTCCGACGCGGATAAAATCTCTTTC
CACCCATACTTCTCGTACAAAGATCTCCTTGGCTTTGTACTAATGTTATTAGCTCTCACA
TCCCTGACATTATTCTCCCCAACTCTGCTCGGCGACCCAGAAAATTTCACCCCGGCAAAC
CCCCTGGTTACCCCGCCACACATCCAACCGGAGTGATACTTCTTATTTGCCTACGCCATT
CTACGCTCTATTCCAAACAAGCTGGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTA
CTATTAATTGTGCCAATTTTACACACCTCAAAGCAGCGAGGACTAACTTTCCGTCCAATC
ACCCAGTTCTTATTCTGAACCTTGGTGGCAGACATAATCATTTTGACATGAATTGGAGGC
ATGCCCGTAGAACACCCATACATCATCATCGGCCAAATCGCGTCGGTGCTGTACTTCGCG
CTCTTCCTTCTTCTCGCCCCCCTCGCANGGTGAGCCGAGAATAAAGCATTGAAATGAGCT

WC1
ATGGCAAGCCTACGAAAAACCCACCCACTAATAAAAATTGCTAACGATGCACTAGTCGAC
CTTCCGACCCCATCTAATATCTCAGCACTATGGAACTTCGGATCCCTCCTAGGATTATGC
TTGATTACTCAAATCCTAACAGGACTATTCCTGGCCATGCATTATACTTCCGATATCTCA
ACTGCCTTTTCGTCGGTAACACACATCTGCCGAGACGTTAATTATGGCTGGCTCATCCGA
AACATACATGCTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGC
GGACTGTACTATGGATCATACCTCTATAAAGCGACCTGAACTATTGGCGTGGTCCTACTC
CTTTTGGTCATAATAACGGCCTTTGTAGGCTATGTGCTGCCCTGAGGCCAAATATCTTTC
TGGGGCGCCACCGTCATTACAAATTTACTATCAGCAGTCCCTTATATGGGAGACGCCCTT
GTCCAGTGGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCC
TTCCACTTCCTATTTCCGTTCGTTATCGCCGGTGCGACGGTCCTGCATTTGCTATTTTTA
CACGAAACGGGATCAAACAACCCTGCCGGATTAAACTCCGACGCAGATAAAATCTCTTTC
CACCCATACTTCTCATACAAAGATCTCCTCGGCTTTGTACTAATATTATTAGCTCTTACA
TCCCTGACATTATTCTCCCCAACTCTGCTCGGTGACCCAGAAAATTTCACCCCAGCAAAC
CCCCTGGTTACCCCGCCACACATCCAACCGGAGTGATACTTCTTATTTGCCTACGCCATT
CTACGCTCTATTCCAAACAAGCTGGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTA
CTGTTAGTCGTGCCAATTTTACATACCTCAAAGCAGCGAGGACTAACTTTCCGGCCAATC
ACCCAGTTCTTATTCTGAACCTTAGTGGCAGACATAATCATTTTAACATGAATTGGAGGC
ATACCCGTAGAACACCCATACATCATCATTGGCCAAATCGCATCGGTGCTGTACTTCGCA
CTCTTCCTTCTTCTCGCCCCCCTCGCAGGGTGAGCCGAGAATAAAGCATTGAAATGAGCT

WR2
ATGGCAAGCCTACGAAAAACCCACCCACTAATAAAAATTGCTAACGATGCACTAGTCGAC
CTTCCAACTCCATCTAATATCTCAGCGCTATGGAACTTCGGATCTCTCCTAGGATTATGC
TTGATTACGCAAATCCTAACAGGACTGTTCCTGGCCATGCATTATACTTCCGATATCTCA
ACTGCCTTTTCGTCAGTAACACACATCTGCCGAGACGTTAACTATGGCTGACTCATCCGA
AACATACATGCTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGC
GGACTGTACTATGGGTCTTACCTCTATAAAGAGACCTGAACCATTGGCGTGGTCCTACTC
CTTTTGGTCATAATAACGGCCTTCGTAGGCTATGTGCTACCGTGGGGTCAAATATCTTTT
TGGGGCGCCACCGTCATTACAAACCTACTATCAGCAGTCCCTTATATGGGAGACGCCCTT
GTCCAGTGGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCC
TTCCACTTCCTGTTTCCGTTCGTTATCGCCGGTGCGACGGTCCTGCATTTGCTATTTTTA
CACGAAACGGGATCAAACAACCCTGCCGGATTAAACTCCGACGCGGATAAAATCTCTTTC
CACCCATACTTCTCGTACAAAGATCTCCTTGGCTTTGTACTAATGTTATTAGCTCTCACA
TCCCTGACATTATTCTCCCCAACTCTGCTCGGCGACCCAGAAAATTTCACCCCGGCAAAC
CCCCTGGTTACCCCGCCACACATCCAACCGGAGTGATACTTCTTATTTGCCTACGCCATT
CTACGCTCTATTCCAAACAAGCTGGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTA
CTATTAGTGGTGCCAATTTTACATACCTCAAAGCAGCGAGGACTAACTTTCCGTCCAATC
ACCCAGTTCTTATTCTGAACCTTGGTGGCAGACATAATCATTTTGACATGAATTGGAGGC
ATGCCCGTAGAACACCCATACATCATCATCGGCCAAATCGCGTCGGTGCTGTACTTCGCA
CTCTTCCTTCTTCTCGCCCCCCTCGCAGGGTGAGCCGAGAATAAAGCATTGAAATGAGCT

OR
ATGGCAAGCCTACGAAAAACCCATCCGCTAATAAAAATTGCTAACAATGCACTAGTTGAC
CTTCCAACCCCATCTAATATTTCAGCACTATGGAGCTTCGGATCCCTCCTAGGATTATGC
TTAATTACTCAAATCCTAACAGGACTATTCCTGGCCATACATTATACTTCCGATATCTCA
ACTGCCTTTTCATCAGTAACACATATCTGCCGAGACGTAAACTATGGCTGACTCATCCGA
AACATACATGCTAACGGAGCATCGTTCTTCTTTATCTGTATTTACATGCACATTGCCCGC
GGACTGTACTACGGGTCATACCTCTATAAAGAGACCTGAAGTGTTGGCGTAGTTCTACTC
CTTTTAGTAATGATGACGGCCTTTGTGGGCTATGTACTGCCCTGAGGACAAATATCTTTT
TGAGGTGCCACCGTAATTACAAATCTACTATCAGCAGTCCCCTATATAGGCGACGCCCTT
GTCCAGTGGATTTGAGGCGGCTTTTCCGTAGATAACGCAACATTAACACGATTCTTCGCC
TTCCACTTCCTATTCCCGTTCGTTATCGCCGGTGCAACGGTCCTACATTTGCTATTTTTA
CACGAAACGGGATCAAACAACCCTGCCGGATTAAACTCCGACGCAGATAAAATCTCTTTC
CACCCATACTTCTCATACAAGGACCTCCTTGGCTTTGTACTAATGTTATTAGCTCTCACA
TCCCTAACATTATTCTCCCCAACACTGCTCGGCGACCCAGAAAATTTCACCCCGGCAAAC
CCCCTGGTCACCCCACCGCACATCCAGCCGGAGTGATACTTCCTGTTTGCCTATGCCATT
CTACGATCTATTCCAAACAAGCTAGGAGGGGTCCTAGCGCTATTGTTTAGCATCCTAGTA
CTATTAGTAGTGCCAATTTTACACACCTCAAAGCAACGAGGACTAACTTTCCGTCCGATC
ACCCAGTTCTTATTCTGAACCTTAGTCGCAGACATAATCATTTTAACATGAATTGGAGGC
ATACCCGTAGAACACCCATATATCATCATTGGCCAAATCGCCTCGGTGCTATACTTCGCA
CTCTTCCTTGTTCTCGCCCCCCTCGCAGGGTGAGCCGAGAATAAAGCATTGAAATGAGCT

BC
ATGGCAAGCCTACGAAAAACCCACCCGCTAATAAAAATTGCTAACGATGCACTAGTTGAC
CTTCCAACCCCATCTAATATTTCAGCACTATGGAACTTCGGATCCCTCCTAGGATTATGC
TTAATTACTCAAATCCTAACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCA
ACTGCCTTTTCATCAGTAACACACATCTGCCGGGACGTAAACTATGGCTGACTCATCCGA
AACATACATGCTAATGGAGCATCATTCTTCTTTATCTGTATTTACATACACATTGCCCGC
GGACTGTACTATGGGTCATACCTCTATAAAGAGACCTGAAATATTGGCGTAGTCCTACTT
CTTTTAGTCATAATGACAGCCTTTGTAGGCTATGTGCTACCATGAGGGCAAATGTCTTTT
TGGGGTGCCACCGTGATTACAAATCTACTATCAGCAGTCCCCTATATAGGCGACGCCCTT
GTCCAGTGGATTTGAGGTGGCTTTTCCGTAGATAACGCAACGTTAACACGATTCTTCGCC
TTCCACTTCCTATTTCCGTTCGTTATCGCCGGTGCGACGGTCC-GCATTTGCTATTTTTA
CACGAAACCGGTTCAAACAACCCTGCCGGGTTAAACTCCGACGCAGATAAGATCTC-TTC
CACCCATACTTCTCATACAAGGACCTCCTTGGCTTTGTCCTAATACTATTAGCTCTCACA
TCCCTGACATT-TTCTCCCCAACACTGCTCGGCGACCCAGAAAATTTCACCCCTGCAAAC
CCCCTGGTTACCCCACCACACATCCAACCGGAGTGATACTTCCTGTTTGCCTACGCCATT
CTACGGTCTATTCCAAACAAGCTAGGAGGGGTCCTAGCACTATTATTTAGCATCCTGGTA
CTATTAGTGGTGCCAATTTTACACACCTCAAAGCAACGAGGACTAACTTT-CGTCCAATC
ACCCAGTTCTTATTCTGAACCTTGGTGGCAGACATAATCATTTTAACATGAATTGGAGGC
ATACC-GTAGAACACCCATATATCATCATTGGCCAAATCGGGTTGGTGCTATACTTCGCA
CTCTTCCTTATTCTCGCCCC-CTCGCAGGGTGAGCCGAGAATAAAGCATTGAAATGAGCT

Robt
ATGGCAAGCCTACGAAAAACCCACCCACTGATAAAAATCGCCAACGGTGCACTAGTCGAC
CTCCCAACACCATCTAACATTTCAGCGCTATGGAACTTCGGGTCCCTCCTGGGATTATGC
TTAATTACTCAGATCCTAACAGGACTATTCCTAGCTATACATTATACCTCTGATATCTCA
ACTGCATTTTCATCCGTAACACACATCTGTCGTGATGTAAATTATGGCTGACTCATTCGG
AACATGCATGCCAACGGCGCATCATTCTTCTTTATCTGTATTTACATGCACATTGCCCGC
GGCCTCTACTACGGCTCATACCTTTATAAGGAAACCTGAAACATTGGCGTAGTTCTACTT
CTTCTGGTAATAATGACAGCCTTCGTGGGCTATGTGCTCCCATGAGGTCAAATGTCTTTT
TGGGGGGCCACCGTAATCACAAATCTATTATCAGCAGTCCCCTATATAGGAGACACCCTT
GTCCAGTGGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGGTTCTTCGCC

```
TTTCACTTCTTATTCCCGTTCGTTATCGCCGGTGCAACAGTCCTGCATTTATTATTTCTT
CACGAGACCGGGTCAAACAACCCTGCCGGGCTAAACTCCGACGCGGATAAAATCTCTTTC
CACCCGTACTTCTCATATAAGGACCTCCTTGGCTTTGTACTTATACTACTAGCTCTCACA
TCCCTAACATTGTTCTCCCCCACCCTTCTAGGCGACCCAGAAAATTTCACCCCAGCAAAT
CCCCTGGTTACTCCACCACACATCCAACCGGAGTGATACTTCTTGTTTGCCTACGCCATT
CTACGATCCATTCCAAACAAGCTAGGAGGGGTCCTGGCATTATTATTTAGCATTCTGGTG
TTACTGGTAGTCCCCATCTTACACACCTCAAAACAACGTGGATTAACCTTCCGCCCAATC
ACTCAGTTCTTATTCTGAACCCTGGTGGCGGATATAGCCATCTTAACATGAATTGGGGGC
ATGCCCGTAGAACACCCATATATTATCATCGGCCAAATCGCGTCAGTGCTATACTTCGCA
CTCTTCCTTCTTCTCGCCCCGCTCGCAGGGTGGGCAGAGAATAAAGCATTGAAATGAGCT

Rcat
ATGGCAAGCCTACGAAAAACCCACCCACTAATAAAAATCGCTAACGATGCACTAGTCGAC
CTTCCAACCCCGTCTAATATTTCAGCGCTATGAAACTTCGGATCCCTCCTAGGATTATGC
TTAATTACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACCTCCGATATCTCA
ACTGCATTTTCATCCGTAACACACATCTGTCGAGACGTTAACTATGGCTGACTCATCCGG
AATATACATGCTAACGGGGCATCATTCTTCTTTATCTGTATTTACATACACATTGCCCGC
GGCCTATACTACGGGTCGTACCTTTATAAGGAGACCTGAAATATCGGCGTTGTTTTACTT
CTCCTAGTCATAATAACAGCCTTCGTGGGCTATGTGCTCCCATGGGGACAAATATCTTTT
TGAGGCGCTACCGTTATTACGAACCTACTATCAGCAGTGCCTTATATGGGTGACGCCCTC
GTCCAGTGGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTCTTCGCC
TTCCACTTCCTATTTCCGTTCGTTATCGCCGGCGCGACAGTGCTACACTTGCTATTTCTA
CACGAGACGGGATCAAACAACCCTGCCGGACTAAATTCCGACGCAGATAAAATCTCTTTC
CACCCATACTTCTCATACAAGGACCTCCTTGGCTTTGTACTAATGTTATTAGCTCTCACA
TCCCTCACATTATTTTCCCCAACCCTGCTCGGGGACCCAGAAAATTTCACCCCAGCAAAT
CCGCTGGTTACCCCGCCACACATCCAACCGGAGTGATACTTCTTATTTGCCTACGCCATT
CTACGATCTATTCCAAACAAACTGGGAGGGGTCCTAGCACTATTATTTAGCATTTTAGTA
CTATTGGTAGTGCCGATTTTACACACCTCGAAGCAACGAGGACTGACTTTCCGCCCGATC
ACTCAGTTTTTATTCTGAACCTTGGTGGCAGATATGATCATTTTAACATGAATTGGAGGT
ATACCCGTAGAGCACCCATACATCATCATCGGCCAAATCGCATCGCTGCTATATTTCGCA
CTCTTCCTTGTTCTCGCCCCTCTCGCAGGGTGAGCGGAGAATAAAGCATTGAAATGAGCT
;
end;

begin trees;
[this tree is a ML tree from the same alignment]
tree dacecytbhaplomodtree =
((((((((((((H5:0.00000001,H16:0.00276030):0.00000001,H15:0.00091501):0.00000001,H14:0.000
91284):0.00000001,H13:0.00183362):0.00000001,H10:0.00183170):0.00000001,H7:0.00091286
):0.00000001,(H4:0.00000001,H6:0.00183218):0.00091409):0.00089405,H12:0.00469237):0.003
89743,(H9:0.00000001,(H11:0.00090989,(H17:0.00182493,(H18:0.00090966,(H8:0.00090960,H
19:0.00000001):0.00000001):0.00000001):0.00000001):0.00000001):0.00556906):0.00024897,(
H2:0.00184023,(H1:0.00000001,H3:0.00091877):0.00000001):0.00439885):0.04773399,(((EAG1
:0.00289690,RAV1:0.00270682):0.01184196,(SON1:0.00668360,WC1:0.00658214):0.00231629)
:0.01198269,(TC2:0.01457100,((BONSV1:0.00397895,BONPK1:0.00341140):0.00257118,((AZT
1:0.00000001,BP11:0.00090266):0.00090177,(LSR1:0.00090102,WR2:0.00180586):0.00000001
):0.00142078):0.00000001):0.01333313):0.00459398):0.03170739,((((H33:0.00000001,H34:0.00
276639):0.00286071,H30:0.00661689):0.00180165,(H31:0.00273182,H32:0.00003864):0.00101
850):0.05294794,(BC:0.02090450,(OR:0.00654739,((H28:0.00000001,H29:0.00183574):0.00092
155,(((H21:0.00000001,H27:0.00092131):0.00000001,H26:0.00091796):0.00185578,((H22:0.002
77257,H24:0.00000001):0.00000001,(H23:0.00000001,(H20:0.00091913,H25:0.00091923):0.000
00001):0.00091922):0.00092784):0.00090280):0.00607815):0.03493494):0.00625870):0.008372
05):0.03683255,Rcat:0.06298758,Robt:0.14500151);
```

```
end;

begin mrbayes;
        log start replace;
        set autoclose = no nowarn=no;
        charset exmem = 1-96 160-255 319-342 406-447 511-531 595-669 733-861 925-966
1030-1044 1108-1140;
        charset transmem = 97-159 256-318 343-405 448-510 532-594 670-732 862-924 967-
1029 1045-1107;
        partition dacecytb = 2:exmem, transmem;
        set partition = dacecytb;
        lset applyto = (all) nst=2 rates = gamma;
        prset applyto = (all); [For JC add option statefreqpr = fixed(equal]
        unlink revmat=(all) shape=(all) pinvar=(all) statefreq=(all) tratio= (all);
        showmodel;
        startvals tau = dacecytbhaplomodtree V = dacecytbhaplomodtree;
        mcmc ngen=10000000 printfreq=1000 samplefreq=100 nchains=4 temp=0.2 checkfreq =
50000 diagnfreq = 1000 stopval = 0.01 stoprule = yes nperts = 1;
        sumt relburnin = yes burninfrac = 0.25 contype = halfcompat conformat = simple;
        sump relburnin = yes burninfrac =0.25;
        log stop;
end;
```

GenAlEx FASTA Input File

```
>IC1
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCTTCCAACT
CCATCTAATATTTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGATTACTCAAATCCTGA
CAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTTTTCGTCAGTAACACACATCTG
TCGAGACGTTAACTATGGCTGACTCATCCGAAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGT
ATTTATATACACATTGCCCGCGGACTGTACTATGGGTCATACCTCTATAAAGAGACCTGAACTATTGGC
GTGGTCCTACTCCTTTTGGTCATAATAACAGCCTTTGTAGGCTATGTACTGCCGTGAGGTCAAATGTCTT
TCTGAGGCGCCACCGTCATTACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGT
GGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCC
GTTTGTTATCGCCGGCGCAACAGTCCTACATTTGTTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGTTAAACTCCAACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCTTGGCTTT
GTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGGTGACCCAGAAAATT
TCACCCCGGCAAACCCCCTGGTTACCCCACCACACATCCAACCGGAGTGGTACTTCTTATTTGCTTACG
CCATTCTACGCTCTATCCCAAACAAGCTGGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTACTATT
AGTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATCACCCAGTTTTTATTC
TGAACCTTAGTGGCAGACATAATCATTTTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCA
TCATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGA
GCCGAAAATAAAGCATTGAAATGAGCT----------------
>IC2
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCTTCCAACT
CCATCTAATATTTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGATTACTCAAATCCTGA
CAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTTTTCGTCAGTAACACACATCTG
TCGAGACGTTAACTATGGCTGACTCATCCGAAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGT
ATTTATATACACATTGCCCGCGGACTGTACTATGGGTCATACCTCTATAAAGAGACCTGAACTATTGGC
GTGGTCCTACTCCTTTTGGTCATAATAACAGCCTTTGTAGGCTATGTACTGCCGTGAGGTCAAATGTCTT
TCTGAGGCGCCACCGTCATTACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGT
GGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCC
GTTTGTTATCGCCGGCGCAACAGTCCTACATTTGTTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGTTAAACTCCAACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCTTGGCTTT
```

GTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGGTGACCCAGAAAATT
TCACCCCGGCAAACCCCCTGGTTACCCCACCACACATCCAACCGGAGTGGTACTTCTTATTTGCTTACG
CCATTCTACGCTCTATCCCAAACAAGCTGGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTACTATT
AGTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATCACCCAGTTTTTATTC
TGAACCTTAGTGGCAGACATAATCATTTTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCA
TCATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGA
GCCGAAAATAAAGCATTGAAATGAGCT----------------
>IC3
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCTTCCAACT
CCATCTAATATTTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGATTACTCAAATCCTGA
CAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTTTTCGTCAGTAACACACATCTG
TCGAGACGTTAACTATGGCTGACTCATCCGAAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGT
ATTTATATACACATTGCCCGCGGACTGTACTATGGGTCATACCTCTATAAAGAGACCTGAACTATTGGC
GTGGTCCTACTCCTTTTGGTCATAATAACAGCCTTTGTAGGCTATGTACTGCCGTGAGATCAAATGTCTT
TCTGAGGCGCCACCGTCATTACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGT
GGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCC
GTTTGTTATCGCCGGCGCAACAGTCCTACATTTGTTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGTTAAACTCCAACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCTTGGCTTT
GTACTAATATTGTTAGCTCTCACACCCCTGACGCTATTCTCTCCAACTCTGCTCGGTGACCCAGAAAATT
TCACCCCGGCAAACCCCCTGGTTACCCCACCACACATCCAACCGGAGTGGTACTTCTTATTTGCTTACG
CCATTCTACGCTCTATCCCAAACAAGCTGGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTACTATT
AGTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATCACCCAGTTTTTATTC
TGAACCTTAGTGGCAGACATAATCATTTTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCA
TCATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGA
GCCGAAAATAAAGCATTGAAATGAGCT----------------
>IC4
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCTTCCAACT
CCATCTAATATTTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGATTACTCAAATCCTGA
CAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTTTTCGTCAGTAACACACATCTG
TCGAGACGTTAACTATGGCTGACTCATCCGAAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGT
ATTTATATACACATTGCCCGCGGACTGTACTATGGGTCATACCTCTATAAAGAGACCTGAACTATTGGC
GTGGTCCTACTCCTTTTGGTCATAATAACAGCCTTTGTAGGCTATGTACTGCCGTGAGGTCAAATGTCTT
TCTGAGGCGCCACCGTCATTACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGT
GGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCC
GTTTGTTATCGCCGGCGCAACAGTCCTACATTTGTTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGTTAAACTCCAACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCTTGGCTTT
GTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGGTGACCCAGAAAATT
TCACCCCGGCAAACCCCCTGGTTACCCCACCACACATCCAACCGGAGTGGTACTTCTTATTTTCTTACG
CCATTCTACGCTCTATCCCAAACAAGCTGGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTACTATT
AGTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATCACCCAGTTTTTATTC
TGAACCTTAGTGGCAGACATAATCATTTTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCA
TCATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGA
GCCGAAAATAAAGCATTGAAATGAGCT----------------
>IC5
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCTTCCAACT
CCATCTAATATTTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGATTACTCAAATCCTGA
CAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTTTTCGTCAGTAACACACATCTG
TCGAGACGTTAACTATGGCTGACTCATCCGAAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGT
ATTTATATACACATTGCCCGCGGACTGTACTATGGGTCATACCTCTATAAAGAGACCTGAACTATTGGC
GTGGTCCTACTCCTTTTGGTCATAATAACAGCCTTTGTAGGCTATGTACTGCCGTGAGGTCAAATGTCTT
TCTGAGGCGCCACCGTCATTACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGT
GGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCC
GTTTGTTATCGCCGGCGCAACAGTCCTACATTTGTTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGTTAAACTCCAACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCTTGGCTTT
GTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGGTGACCCAGAAAATT
TCACCCCGGCAAACCCCCTGGTTACCCCACCACACATCCAACCGGAGTGGTACTTCTTATTTGCTTACG
CCATTCTACGCTCTATCCCAAACAAGCTGGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTACTATT
AGTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATCACCCAGTTTTTATTC
TGAACCTTAGTGGCAGACATAATCATTTTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCA

```
TCATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGA
GCCGAAAATAAAGCATTGAAATGAGCT----------------
>CC1
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCTTCCAACT
CCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGATTACTCAAATCCTGA
CAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTTTTCGTCAGTAACACACATCTG
TCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGT
ATTTATATACACATTGCCCGCGGACTGTACTATGGGTCATACCTCTATAAAGAGACCTGAACTATTGGC
GTGGTCCTACTCCTTTTGGTCATAATAACAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTT
TCTGAGGCGCCACCGTCATTACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGT
GGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCC
GTTTGTTATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAAACCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCTTGGCTTT
GTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGGTGACCCAGAAAATT
TTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGTGGTACTTCTTATTTGCTTACG
CCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTACTATT
AGTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATCACCCAGTTTTTATTC
TGAACCTTAGTGGCAGACATAATTATTTTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCA
TCATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGA
GCCGAAAATAAAGCATTGAAATGAGCT----------------
>CC2
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCTTCCAACT
CCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGATTACTCAAATCCTGA
CAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTTTTCGTCAGTAACACACATCTG
TCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGT
ATTTATATACACATTGCCCGCGGACTGTACTATGGGTCATACCTCTATAAAGAGACCTGAACTATTGGC
GTGGTCCTACTCCTTTTGGTCATAATAACAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTT
TCTGAGGCGCCACCGTCATTACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGT
GGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCC
GTTTGTTATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCTTGGCTTT
GTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGGTGACCCAGAAAATT
TTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGTGGTACTTCTTATTTGCTTACG
CCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTACTATT
AGTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATCACCCAGTTTTTATTC
TGAACCTTAGTGGCAGACATAATTATTTTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCA
TCATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGA
GCCGAAAATAAAGCATTGAAATGAGCT----------------
>CC6
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCTTCCAACT
CCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGATTACTCAAATCCTGA
CAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTTTTCGTCAGTAACACACATCTG
TCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGT
ATTTATATACACATTGCCCGCGGACTGTACTATGGGTCATACCTCTATAAAGAGACCTGAACTATTGGC
GTGGTCCTACTCCTTTTGGTCATAATAACAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTT
TCTGAGGCGCCACCGTCATTACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGT
GGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCC
GTTTGTTATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCTTGGCTTT
GTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGGTGACCCAGAAAATT
TTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGTGGTACTTCTTATTTGCTTACG
CCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTACTATT
AGTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATCACCCAGTTTTTATTC
TGAACCTTAGTGGCAGACATAATTATTTTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCA
TCATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGA
GCCGAAAATAAAGCATTGAAATGAGCT----------------
>CC9
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCTTCCAACT
CCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGATTACTCAAATCCTGA
```

CAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTTTTCGTCAGTAACACACATCTG
TCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGT
ATTTATATACACATTGCCCGCGGACTGTACTATGGGTCATACCTCTATAAAGAGACCTGAACTATTGGC
GTGGTCCTACTCCTTTTGGTCATAATAACAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTT
TCTGAGGCGCCACCGTCATTACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGT
GGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCC
GTTTGTTATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAAACCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCTTGGCTTT
GTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGGTGACCCAGAAAATT
TTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGTGGTACTTCTTATTTGCTTACG
CCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTACTATT
AGTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATCACCCAGTTTTTATTC
TGAACCTTAGTGGCAGACATAATTATTTTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCA
TCATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGA
GCCGAAAATAAAGCATTGAAATGAGCT----------------
>PC1
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCTTCCAACT
CCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGATTACTCAAATCCTGA
CAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTTTTCGTCAGTAACACACATCTG
TCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGT
ATTTATATACACATTGCCCGCGGACTGTACTATGGGTCATACCTCTATAAAGAGACCTGAACTATTGGC
GTGGTCCTACTCCTTTTGGTCATAATAACAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTT
TCTGAGGCGCCACCGTCATTACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGT
GGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCC
GTTTGTTATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAAACCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCTTGGCTTT
GTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGGTGACCCAGAAAATT
TTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGTGGTACTTCTTATTTGCTTACG
CCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTACTATT
AGTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATCACCCAGTTTTTATTC
TGAACCTTAGTGGCAGACATAATTATTTTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCA
TCATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGA
GCCGAAAATAAAGCATTGAAATGAGCT----------------
>PC4
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCTTCCAACT
CCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGATTACTCAAATCCTGA
CAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTTTTCGTCAGTAACACACATCTG
TCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGT
ATTTATATACACATTGCCCGCGGACTGTACTATGGGTCATACCTCTATAAAGAGACCTGAACTATTGGC
GTGGTCCTACTCCTTTTGGTCATAATAACAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTT
TCTGAGGCGCCACCGTCATTACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGT
GGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCC
GTTTGTTATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAAACCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCTTGGCTTT
GTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGGTGACCCAGAAAATT
TTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGTGGTACTTCTTATTTGCTTACG
CCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTACTATT
AGTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATCACCCAGTTTTTATTC
TGAACCTTAGTGGCAGACATAATTATTTTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCA
TCATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGA
GCCGAAAATAAAGCATTGAAATGAGCT----------------
>PC7
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCTTCCAACT
CCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGATTACTCAAATCCTGA
CAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTTTTCGTCAGTAACACACATCTG
TCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGT
ATTTATATACACATTGCCCGCGGACTGTACTATGGGTCATACCTCTATAAAGAGACCTGAACTATTGGC
GTGGTCCTACTCCTTTTGGTCATAATAACAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTT
TCTGAGGCGCCACCGTCATTACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGT

```
GGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCC
GTTTGTTATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAAACCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCTTGGCTTT
GTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGGTGACCCAGAAAATT
TTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGTGGTACTTCTTATTTGCTTACG
CCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTACTATT
AGTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATCACCCAGTTTTTATTC
TGAACCTTAGTGGCAGACATAATTATTTTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCA
TCATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGA
GCCGAAAATAAAGCATTGAAATGAGCT----------------
>PC10
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCTTCCAACT
CCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGATTACTCAAATCCTGA
CAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTTTTCGTCAGTAACACACATCTG
TCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGT
ATTTATATACACATTGCCCGCGGACTGTACTATGGGGCATACCTCTATAAAGAGACCTGAACTATTGGC
GTGGTCCTACTCCTTTTGGTCATAATAACAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTT
TCTGAGGCGCCACCGTCATTACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGT
GGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCC
GTTTGTTATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAAACCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCTTGGCTTT
GTATTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGGTGACCCAGAAAATT
TTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGTGGTACTTCTTATTTGCTTACG
CCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTACTATT
AGTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATCACCCAGTTTTTATTC
TGAACCTTAGTGGCAGACATAATTATTTTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCA
TCATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGA
GCCGAAAATAAAGCATTGAAATGAGCT----------------
>PC11
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCTTCCAACT
CCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGATTACTCAAATCCTGA
CAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTTTTCGTCAGTAACACACATCTG
TCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGT
ATTTATATACACATTGCCCGCGGACTGTACTATGGGTCATACCTCTATAAAGAGACCTGAACTATTGGC
GTGGTCCTACTCCTTTTGGTCATAATAACAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTT
TCTGAGGCGCCACCGTCATTACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGT
GGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCC
GTTTGTTATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAAACCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCTTGGCTTT
GTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGGTGACCCAGAAAATT
TTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGTGGTACTTCTTATTTGCTTACG
CCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTACTATT
AGTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATCACCCAGTTTTTATTC
TGAACCTTAGTGGCAGACATAATTATTTTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCA
TCATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGA
GCCGAAAATAAAGCATTGAAATGAGCT----------------
>PC12
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCTTCCAACT
CCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGATTACTCAAATCCTGA
CAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTTTTCGTCAGTAACACACATCTG
TCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGT
ATTTATATACACATTGCCCGCGGACTGTACTATGGGTCATACCTCTATAAAGAGACCTGAACTATTGGC
GTGGTCCTACTCCTTTTGGTCATAATAACAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTT
TCTGAGGCGCCACCGTCATTACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGT
GGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCC
GTTTGTTATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCTTGGCTTT
GTATTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGGTGACCCAGAAAATT
TTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGTGGTACTTCTTATTTGCTTACG
```

```
CCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTACTATT
AGTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATCACCCAGTTTTTATTC
TGAACCTTAGTGGCAGACATAATTATTTTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCA
TCATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGA
GCCGAAAATAAAGCATTGAAATGAGCT----------------
>PC13
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCTTCCAACT
CCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGATTACTCAAATCCTGA
CAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTTTTCGTCAGTAACACACATCTG
TCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGT
ATTTATATACACATTGCCCGCGGACTGTACTATGGGTCATACCTCTATAAAGAGACCTGAACTATTGGC
GTGGTCCTACTCCTTTTGGTCATAATAACAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTT
TCTGAGGCGCCACCGTCATTACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGT
GGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCC
GTTTGTTATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAAACCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCTTGGCTTT
GTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGGTGACCCAGAAAATT
TTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGTGGTACTTCTTATTTGCTTACG
CCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTACTATT
AGTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATCACCCAGTTTTTATTC
TGAACCTTAGTGGCAGACATAATTATTTTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCA
TCATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGA
GCCGAAAATAAAGCATTGAAATGAGCT----------------
>PC14
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCTTCCAACT
CCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGATTACTCAAATCCTGA
CAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTTTTCGTCAGTAACACACATCTG
TCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGT
ATTTATATACACATTGCCCGCGGACTGTACTATGGGTCATACCTCTATAAAGAGACCTGAACTATTGGC
GTGGTCCTACTCCTTTTGGTCATAATAACAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTT
TCTGAGGCGCCACCGTCATTACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGT
GGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCC
GTTTGTTATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAAACCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCTTGGCTTT
GTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGGTGACCCAGAAAATT
TTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGTGGTACTTCTTATTTGCTTACG
CCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTACTATT
AGTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATCACCCAGTTTTTATTC
TGAACCTTAGTGGCAGACATAATTATTTTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCA
TCATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGA
GCCGAAAATAAAGCATTGAAATGAGCT----------------
>TWIN1
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCTTCCAACT
CCATCTAATATTTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCCTGATTACTCAAATCCTGA
CAGGACTATTCCTAGCCATACATTATACTTCCGATATTTCAACTGCCTTTTCGTCAGTAACACACATCTGT
CGAGACGTTAACTATGGCTGACTCATCCGGAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGTA
TTTATATACACATTGCCCGCGGACTGTACTATGGATCATACCTCTATAAAGAGACCTGAACTATTGGCGT
GGTCTTACTCCTTTTGGTCATAATAACAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTC
TGAGGCGCCACCGTCATTACGAATTTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGG
ATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGT
TTGTTATTGCCGGCGCAACGGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGCCG
GGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCGTACAAAGATCTCCTTGGCTTTGT
ACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGGTGACCCAGAAAATTTC
ACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGTGGTACTTCTTATTTGCTTACGC
CATTCTACGCTCTATCCCAAACAAGCTGGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTACTATTA
GTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCT
GAACCTTAGTGGCAGACATAATCATTTTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCAT
CATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGAG
CCGAAAATAAAGCATTGAAATGAGCT----------------
```

```
>TWIN2
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCTTCCAACT
CCATCTAATATTTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGATTACTCAAATCCTGA
CAGGACTATTCCTAGCCATACATTATACTTCCGATATTTCAACTGCCTTTTCGTCAGTAACACACATCTGT
CGAGACGTTAACTATGGCTGACTCATCCGGAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGTA
TTTATATACACATTGCCCGCGGACTGTACTATGGATCATACCTCTATAAAGAGACCTGAACTATTGGCGT
GGTCTTACTCCTTTTGGTCATAATAACAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTC
TGAGGCGCCACCGTCATTACGAATTTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGG
ATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGT
TTGTTATTGCCGGCGCAACGGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGCCG
GGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCGTACAAAGATCTCCTTGGCTTTGT
ACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGGTGACCCAGAAAATTTC
ACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGTGGTACTTCTTATTTGCTTACGC
CATTCTACGCTCTATCCCAAACAAGCTGGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTACTATTA
GTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCT
GAACCTTAGTGGCAGACATAATCATTTTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCAT
CATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGAG
CCGAAAATAAAGCATTGAAATGAGCT----------------
>TWIN3
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCTTCCAACT
CCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGATTACTCAAATCCTGA
CAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTTTTCGTCAGTAACACACATCTG
TCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGT
ATTTATATACACATTGCCCGCGGACTGTACTATGGGTCATACCTCTATAAAGAGACCTGAACTATTGGC
GTGGTCCTACTCCTTTTGGTCATAATAACAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTT
TCTGAGGCGCCACCGTCATTACGAATCTATTATCAACAGTCCCCTACATAGGAGACGCTCTTGTCCAGT
GGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCC
GTTTGTTATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGCTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCTTGGCTTT
GTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGGTGACCCAGAAAATT
TTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGTGGTACTTCTTATTTGCTTACG
CCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTACTATT
AGTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATCACCCAGTTTTTATTC
TGAACCTTAGTGGCAGACATAATTATTTTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCA
TCATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGA
GCCGAAAATAAAGCATTGAAATGAGCT----------------
>TWIN4
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCTTCCAACT
CCATCTAATATTTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGATTACTCAAATCCTGA
CAGGACTATTCCTAGCCATACATTATACTTCCGATATTTCAACTGCCTTTTCGTCAGTAACACACATCTGT
CGAGACGTTAACTATGGCTGACTCATCCGGAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGTA
TTTATATACACATTGCCCGCGGACTGTACTATGGATCATACCTCTATAAAGAGACCTGAACTATTGGCGT
GGTCTTACTCCTTTTGGTCATAATAACAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTC
TGAGGCGCCACCGTCATTACGAATTTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGG
ATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGT
TTGTTATTGCCGGCGCAACGGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGCCG
GGTTAAACTCCGACGCAAATAAGATCTCCTTCCACCCATACTTCTCGTACAAAGATCTCCTTGGCTTTGT
ACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGGTGACCCAGAAAATTTC
ACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGTGGTACTTCTTATTTGCTTACGC
CATTCTACGCTCTATCCCAAACAAGCTGGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTACTATTA
GTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCT
GAACCTTAGTGGCAGACATAATCATTTTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCAT
CATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGAG
CCGAAAATAAAGCATTGAAATGAGCT----------------
>TWIN5
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCTTCCAACT
CCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGATTACTCAAATCCTGA
CAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTTTTCGTCAGTAACACACATCTG
TCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGT
```

```
ATTTATATACACATTGCCCGCGGACTGTACTATGGGTCATACCTCTATAAAGAGACCTGAACTATTGGC
GTGGTCCTACTCCTTTTGGTCATAATAACAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTT
TCTGAGGCGCCACCGTCATTACGAATCTATTATCAACAGTCCCCTACATAGGAGACGCTCTTGTCCAGT
GGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCC
GTTTGTTATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGCTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCTTGGCTTT
GTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGGTGACCCAGAAAATT
TTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGTGGTACTTCTTATTTGCTTACG
CCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTACTATT
AGTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATCACCCAGTTTTTATTC
TGAACCTTAGTGGCAGACATAATTATTTTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCA
TCATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGA
GCCGAAAATAAAGCATTGAAATGAGCT----------------
>TWIN6
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCTTCCAACT
CCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGATTACTCAAATCCTGA
CAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTTTTCGTCAGTAACACACATCTG
TCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGT
ATTTATATACACATTGCCCGCGGACTGTACTATGGGTCATACCTCTATAAAGAGACCTGAACTATTGGC
GTGGTCCTACTCCTTTTGGTCATAATAACAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTT
TCTGAGGCGCCACCGTCATTACGAATCTATTATCAACAGTCCCCTACATAGGAGACGCTCTTGTCCAGT
GGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCC
GTTTGTTATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGCTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCTTGGCTTT
GTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGGTGACCCAGAAAATT
TTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGTGGTACTTCTTATTTGCTTACG
CCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTACTATT
AGTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATCACCCAGTTTTTATTC
TGAACCTTAGTGGCAGACATAATTATTTTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCA
TCATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGA
GCCGAAAATAAAGCATTGAAATGAGCT----------------
>TWIN7
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCTTCCAACT
CCATCTAATATTTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGATTACTCAAATCCTGA
CAGGACTATTCCTAGCCATACATTATACTTCCGATATTTCAACTGCCTTTTCGTCAGTAACACACATCTGT
CGAGACGTTAACTATGGCTGACTCATCCGGAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGTA
TTTATATACACATTGCCCGCGGACTGTACTATGGATCATACCTCTATAAAGAGACCTGAACTATTGGCGT
GGTCTTACTCCTTTTGGTCATAATAACAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTC
TGAGGCGCCACCGTCATTACGAATTTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGG
ATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGT
TTGTTATTGCCGGCGCAACGGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGCCG
GGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCGTACAAAGATCTCCTTGGCTTTGT
ACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGGTGACCCAGAAAATTTC
ACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGTGGTACTTCTTATTTGCTTACGC
CATTCTACGCTCTATCCCAAACAAGCTGGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTACTATTA
GTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCT
GAACCTTAGTGGCAGACATAATCATTTTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCAT
CATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGAG
CCGAAAATAAAGCATTGAAATGAGCT----------------
>TWIN8
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCTTCCAACT
CCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGATTACTCAAATCCTGA
CAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTTTTCGTCAGTAACACACATCTG
TCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGT
ATTTATATACACATTGCCCGCGGACTGTACTATGGGTCATACCTCTATAAAGAGACCTGAACTATTGGC
GTGGTCCTACTCCTTTTGGTCATAATAACAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTT
TCTGAGGCGCCACCGTCATTACGAATCTATTATCAACAGTCCCCTACATAGGAGACGCTCTTGTCCAGT
GGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCC
GTTTGTTATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
```

```
CGGGCTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCTTGGCTTT
GTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGGTGACCCAGAAAATT
TTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGTGGTACTTCTTATTTGCTTACG
CCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTACTATT
AGTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATCACCCAGTTTTTATTC
TGAACCTTAGTGGCAGACATAATTATTTTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCA
TCATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGA
GCCGAAAATAAAGCATTGAAATGAGCT----------------
>TWIN9
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCTTCCAACT
CCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGATTACTCAAATCCTGA
CAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTTTTCGTCAGTAACACACATCTG
TCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGT
ATTTATATACACATTGCCCGCGGACTGTACTATGGGTCATACCTCTATAAAGAGACCTGAACTATTGGC
GTGGTCCTACTCCTTTTGGTCATAATAACAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTT
TCTGAGGCGCCACCGTCATTACGAATCTATTATCAACAGTCCCCTACATAGGAGACGCTCTTGTCCAGT
GGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCC
GTTTGTTATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGCTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCTTGGCTTT
GTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGGTGACCCAGAAAATT
TTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGTGGTACTTCTTATTTGCTTACG
CCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTACTATT
AGTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATCACCCAGTTTTTATTC
TGAACCTTAGTGGCAGACATAATTATTTTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCA
TCATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGA
GCCGAAAATAAAGCATTGAAATGAGCT----------------
>TWIN10
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCTTCCAACT
CCATCTAATATTTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGATTACTCAAATCCTGA
CAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTTTTCGTCAGTAACACACATCTG
TCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGT
ATTTATATACACATTGCCCGCGGACTGTACTATGGGTCATACCTCTATAAAGAGACCTGAACTATTGGC
GTGGTCCTACTCCTTTTGGTCATAATAACAGCCTTTGTAGGCTATGTACTACCATGGGGTCAAATATCTT
TCTGAGGCGCCACCGTCATTACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAAT
GGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTATTTCC
GTTTGTTATCGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCTTGGCTTT
GTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGGTGACCCAGAAAATT
TTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGTGGTACTTCTTATTTGCTTACG
CCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTACTATT
AGTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATCACCCAGTTTTTATTC
TGAACCTTAGTGGCAGACATAATTATTTTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCA
TCATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGA
GCCGAAAATAAAGCATTGAAATGAGCT----------------
>TWIN11
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCTTCCAACT
CCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGATTACTCAAATCCTGA
CAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTTTTCGTCAGTAACACACATCTG
TCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGT
ATTTATATACACATTGCCCGCGGACTGTACTATGGGTCATACCTCTATAAAGAGACCTGAACTATTGGC
GTGGTCCTACTCCTTTTGGTCATAATAACAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTT
TCTGAGGCGCCACCGTCATTACGAATCTATTATCAACAGTCCCCTACATAGGAGACGCTCTTGTCCAGT
GGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCC
GTTTGTTATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGCTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCTTGGCTTT
GTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGGTGACCCAGAAAATT
TTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGTGGTACTTCTTATTTGCTTACG
CCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTACTATT
AGTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATCACCCAGTTTTTATTC
```

```
TGAACCTTAGTGGCAGACATAATTATTTTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCA
TCATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGA
GCCGAAAATAAAGCATTGAAATGAGCT----------------
>TWIN12
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCTTCCAACT
CCATCTAATATTTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGATTACTCAAATCCTGA
CAGGACTATTCCTAGCCATACATTATACTTCCGATATTTCAACTGCCTTTTCGTCAGTAACACACATCTGT
CGAGACGTTAACTATGGCTGACTCATCCGGAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGTA
TTTATATACACATTGCCCGCGGACTGTACTATGGATCATACCTCTATAAAGAGACCTGAACTATTGGCGT
GGTCTTACTCCTTTTGGTCATAATAACAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTC
TGAGGCGCCACCGTCATTACGAATTTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGG
ATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGT
TTGTTATTGCCGGCGCAACGGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGCCG
GGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCGTACAAAGATCTCCTTGGCTTTGT
ACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGGTGACCCAGAAAATTTC
ACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGTGGTACTTCTTATTTGCTTACGC
CATTCTACGCTCTATCCCAAACAAGCTGGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTACTATTA
GTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCT
GAACCTTAGTGGCAGACATAATCATTTTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCAT
CATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGAG
CCGAAAATAAAGCATTGAAATGAGCT----------------
>CAJ1
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCTTCCAACT
CCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGATTACTCAAATCCTGA
CAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTTTTCGTCAGTAACACACATCTG
TCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGT
ATTTATATACACATTGCCCGCGGACTGTACTATGGGTCATACCTCTATAAAGAGACCTGAACTATTGGC
GTGGTCCTACTCCTTTTGGTCATAATAACAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTT
TCTGAGGCGCCACCGTCATTACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGT
GGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCC
GTTTGTTATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCTTGGCTTT
GTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGGTGACCCAGAAAATT
TTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGTGGTACTTCTTATTTGCTTACG
CCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTACTATT
AGTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATCACCCAGTTTTTATTC
TGAACCTTAGTGGCAGACATAATTATTTTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCA
TCATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGA
GCCGAAAATAAAGCATTGAAATGAGCT----------------
>CAJ2
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCTTCCAACT
CCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGATTACTCAAATCCTGA
CAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTTTTCGTCAGTAACACACATCTG
TCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGT
ATTTATATACACATTGCCCGCGGACTGTACTATGGGTCATACCTCTATAAAGAGACCTGAACTATTGGC
GTGGTCCTACTCCTTTTGGTCATAATAACAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTT
TCTGAGGCGCCACCGTCATTACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGT
GGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCC
GTTTGTTATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCTTGGCTTT
GTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGGTGACCCAGAAAATT
TTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGTGGTACTTCTTATTTGCTTACG
CCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTACTATT
AGTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATCACCCAGTTTTTATTC
TGAACCTTAGTGGCAGACATAATTATTTTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCA
TCATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGA
GCCGAAAATAAAGCATTGAAATGAGCT----------------
>CAJ3
```

ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCTTCCAACT
CCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGATTACTCAAATCCTGA
CAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTTTTCGTCAGTAACACACATCTG
TCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGT
ATTTATATACACATTGCCCGCGGACTGTACTATGGGTCATACCTCTATAAAGAGACCTGAACTATTGGC
GTGGTCCTACTCCTTTTGGTCATAATAACAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTT
TCTGAGGCGCCACCGTCATTACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGT
GGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCC
GTTTGTTATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCTTGGCTTT
GTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGGTGACCCAGAAAATT
TTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGTGGTACTTCTTATTTGCTTACG
CCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTACTATT
AGTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATCACCCAGTTTTTATTC
TGAACCTTAGTGGCAGACATAATTATTTTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCA
TCATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGA
GCCGAAAATAAAGCATTGAAATGAGCT----------------
>CAJ4
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCTTCCAACT
CCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGATTACTCAAATCCTGA
CAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTTTTCGTCAGTAACACACATCTG
TCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGT
ATTTATATACACATTGCCCGCGGACTGTACTATGGGTCATACCTCTATAAAGAGACCTGAACTATTGGC
GTGGTCCTACTCCTTTTGGTCATAATAACAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTT
TCTGAGGCGCCACCGTCATTACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGT
GGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCC
GTTTGTTATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCTTGGCTTT
GTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGGTGACCCAGAAAATT
TTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGTGGTACTTCTTATTTGCTTACG
CCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTACTATT
AGTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATCACCCAGTTTTTATTC
TGAACCTTAGTGGCAGACATAATTATTTTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCA
TCATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGA
GCCGAAAATAAAGCATTGAAATGAGCT----------------
>CAJ5
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCTTCCAACT
CCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGATTACTCAAATCCTGA
CAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTTTTCGTCAGTAACACACATCTG
TCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGT
ATTTATATACACATTGCCCGCGGACTGTACTATGGGTCATACCTCTATAAAGAGACCTGAACTATTGGC
GTGGTCCTACTCCTTTTGGTCATAATAACAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTT
TCTGAGGCGCCACCGTCATTACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGT
GGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCC
GTTTGTTATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCTTGGCTTT
GTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGGTGACCCAGAAAATT
TTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGTGGTACTTCTTATTTGCTTACG
CCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTACTATT
AGTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATCACCCAGTTTTTATTC
TGAACCTTAGTGGCAGACATAATTATTTTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCA
TCATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGA
GCCGAAAATAAAGCATTGAAATGAGCT----------------
>CAJ6
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCTTCCAACT
CCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGATTACTCAAATCCTGA
CAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTTTTCGTCAGTAACACACATCTG
TCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGT
ATTTATATACACATTGCCCGCGGACTGTACTATGGGTCATACCTCTATAAAGAGACCTGAACTATTGGC

```
GTGGTCCTACTCCTTTTGGTCATAATAACAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTT
TCTGAGGCGCCACCGTCATTACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGT
GGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCC
GTTTGTTATTGCCGACGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGCC
GGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCTTGGCTTTG
TACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGGTGACCCAGAAAATTT
TACCCCGGCAAACCCCCTGGTTACCCCACCACACATCCAACCGGAGTGGTACTTCTTATTTGCTTACGC
CATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTACTATTA
GTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCT
GAACCTTAGTGGCAGACATAATTATTTTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCAT
CATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGAG
CCGAAAATAAAGCATTGAAATGAGCT----------------
>CAJ8
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCTTCCAACT
CCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGATTACTCAAATCCTGA
CAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTTTTCGTCAGTAACACACATCTG
TCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGT
ATTTATATACACATTGCCCGCGGACTGTACTATGGGTCATACCTCTATAAAGAGACCTGAACTATTGGC
GTGGTCCTACTCCTTTTGGTCATAATAACAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTT
TCTGAGGCGCCACCGTCATTACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGT
GGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCC
GTTTGTTATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCTTGGCTTT
GTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGGTGACCCAGAAAATT
TTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGTGGTACTTCTTATTTGCTTACG
CCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTACTATT
AGTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATCACCCAGTTTTTATTC
TGAACCTTAGTGGCAGACATAATTATTTTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCA
TCATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGA
GCCGAAAATAAAGCATTGAAATGAGCT----------------
>CAJ9
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCTTCCAACT
CCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGATTACTCAAATCCTGA
CAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTTTTCGTCAGTAACACACATCTG
TCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGT
ATTTATATACACATTGCCCGCGGACTGTACTATGGGTCATACCTCTATAAAGAGACCTGAACTATTGGC
GTGGTCCTACTCCTTTTGGTCATAATAACAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTT
TCTGAGGCGCCACCGTCATTACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGT
GGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCC
GTTTGTTATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCTTGGCTTT
GTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGGTGACCCAGAAAATT
TTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGTGGTACTTCTTATTTGCTTACG
CCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTACTATT
AGTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATCACCCAGTTTTTATTC
TGAACCTTAGTGGCAGACATAATTATTTTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCA
TCATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGA
GCCGAAAATAAAGCATTGAAATGAGCT----------------
>CAJ10
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCTTCCAACT
CCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGATTACTCAAATCCTGA
CAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTTTTCGTCAGTAACACACATCTG
TCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGT
ATTTATATACACATTGCCCGCGGACTGTACTATGGGTCATACCTCTATAAAGAGACCTGAACTATTGGC
GTGGTCCTACTCCTTTTGGTCATAATAACAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTT
TCTGAGGCGCCACCGTCATTACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGT
GGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCC
GTTTGTTATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCTTGGCTTT
```

```
GTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGGTGACCCAGAAAATT
TTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGTGGTACTTCTTATTTGCTTACG
CCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTACTATT
AGTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATCACCCAGTTTTTATTC
TGAACCTTAGTGGCAGACATAATTATTTTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCA
TCATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGA
GCCGAAAATAAAGCATTGAAATGAGCT----------------
>CAJ11
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCTTCCAACT
CCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGATTACTCAAATCCTGA
CAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTTTTCGTCAGTAACACACATCTG
TCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGT
ATTTATATACACATTGCCCGCGGACTGTACTATGGGTCATACCTCTATAAAGAGACCTGAACTATTGGC
GTGGTCCTACTCCTTTTGGTCATAATAACAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTT
TCTGAGGCGCCACCGTCATTACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGT
GGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCC
GTTTGTTATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCTTGGCTTT
GTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGGTGACCCAGAAAATT
TTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGTGGTACTTCTTATTTGCTTACG
CCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTACTATT
AGTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATCACCCAGTTTTTATTC
TGAACCTTAGTGGCAGACATAATTATTTTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCA
TCATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGA
GCCGAAAATAAAGCATTGAAATGAGCT----------------
>CAJ12
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCTTCCAACT
CCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGATTACTCAAATCCTGA
CAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTTTTCGTCAGTAACACACATCTG
TCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGT
ATTTATATACACATTGCCCGCGGACTGTACTATGGGTCATACCTCTATAAAGAGACCTGAACTATTGGC
GTGGTCCTACTCCTTTTGGTCATAATAACAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTT
TCTGAGGCGCCACCGTCATTACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGT
GGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCC
GTTTGTTATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCTTGGCTTT
GTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGGTGACCCAGAAAATT
TTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGTGGTACTTCTTATTTGCTTACG
CCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTACTATT
AGTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATCACCCAGTTTTTATTC
TGAACCTTAGTGGCAGACATAATTATTTTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCA
TCATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGA
GCCGAAAATAAAGCATTGAAATGAGCT----------------
>CAJ14
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCTTCCAACT
CCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGATTACTCAAATCCTGA
CAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTTTTCGTCAGTAACACACATCTG
TCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGT
ATTTATATACACATTGCCCGCGGACTGTACTATGGGTCATACCTCTATAAAGAGACCTGAACTATTGGC
GTGGTCCTACTCCTTTTGGTCATAATAACAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTT
TCTGAGGCGCCACCGTCATTACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGT
GGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCC
GTTTGTTATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCATCCATACTTCTCATACAAAGATCTCCTTGGCTTT
GTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGGTGACCCAGAAAATT
TTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGTGGTACTTCTTATTTGCTTACG
CCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTACTATT
AGTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATCACCCAGTTTTTATTC
TGAACCTTAGTGGCAGACATAATTATTTTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCA
```

```
TCATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGA
GCCGAAAATAAAGCATTGAAATGAGCT----------------
>CAJ15
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCTTCCAACT
CCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGATTACTCAAATCCTGA
CAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTTTTCGTCAGTAACACACATCTG
TCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGT
ATTTATATACACATTGCCCGCGGACTGTACTATGGGTCATACCTCTATAAAGAGACCTGAACTATTGGC
GTGGTCCTACTCCTTTTGGTCATAATAACAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTT
TCTGAGGCGCCACCGTCATTACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGT
GGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCC
GTTTGTTATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCTTGGCTTT
GTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGGTGACCCAGAAAATT
TTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGTGGTACTTCTTATTTGCTTACG
CCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTACTATT
AGTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATCACCCAGTTTTTATTC
TGAACCTTAGTGGCAGACATAATTATTTTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCA
TCATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGA
GCCGAAAATAAAGCATTGAAATGAGCT----------------
>LCK1
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAAATTGCTAACGATGCACTAGTCGACCTTCCAACT
CCATCTAATATCTCAGCACTATGGAACTTCGGGTCCCTCCTAGGATTATGCTTAATTACTCAAATCCTAA
CAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTTTTCGTCAGTAACACACATCTG
CCGAGACGTTAATTATGGCTGACTCATCCGGAATATACATGCTAACGGAGCATCATTCTTCTTTATCTGT
ATTTACATGCACATTGCCCGCGGACTGTACTATGGGTCATACCTCTATAAAGAGACCTGAACTATTGGC
GTGGTTCTACTCCTTTTAGTCATGATGACAGCCTTTGTAGGCTACGTGCTACCATGAGGTCAAATATCTT
TTTGAGGTGCCACCGTGATTACAAATCTACTATCAGCAGTCCCCTATATAGGCGACGCTCTTGTCCAAT
GGATTTGAGGTGGCTTTTCCGTAGATAACGCAACGTTAACACGATTCTTCGCCTTCCACTTCCTATTTCC
GTTTGTTATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCTTGGCTTT
GTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGGTGACCCAGAAAATT
TTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGTGATACTTCTTATTTGCTTACG
CCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTACTATT
AGTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATCACCCAGTTTTTATTC
TGAACCTTAGTGGCAGACATAATTATTTTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCA
TCATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGA
GCCGAAAATAAAGCATTGAAATGAGCT----------------
>LCK3
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCTTCCAACT
CCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGATTACTCAAATCCTGA
CAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTTTTCGTCAGTAACACACATCTG
TCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGT
ATTTATATACACATTGCCCGCGGACTGTACTATGGGTCATACCTCTATAAAGAGACCTGAACTATTGGC
GTGGTCCTACTCCTTTTGGTCATAATAACAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTT
TCTGAGGCGCCACCGTCATTACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGT
GGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCC
GTTTGTTATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCTTGGCTTT
GTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGGTGACCCAGAAAATT
TTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGTGGTACTTCTTATTTGCTTACG
CCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTACTATT
AGTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATCACCCAGTTTTTATTC
TGAACCTTAGTGGCAGACATAATTATTTTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCA
TCATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGA
GCCGAAAATAAAGCATTGAAATGAGCT----------------
>LCK4
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCTTCCAACT
CCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGATTACTCAAATCCTGA
```

```
CAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTTTTCGTCAGTAACACACATCTG
TCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGT
ATTTATATACACATTGCCCGCGGACTGTACTATGGGTCATACCTCTATAAAGAGACCTGAACTATTGGC
GTGGTCCTACTCCTTTTGGTCATAATAACAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTT
TCTGAGGCGCCACCGTCATTACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGT
GGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCC
GTTTGTTATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCTTGGCTTT
GTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACACTGCTCGGTGACCCAGAAAATT
TTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGTGGTACTTCTTATTTGCTTACG
CCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTACTATT
AGTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATCACCCAGTTTTTATTC
TGAACCTTAGTGGCAGACATAATTATTTTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCA
TCATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGA
GCCGAAAATAAAGCATTGAAATGAGCT----------------
>LCK5
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCTTCCAACT
CCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGATTACTCAAATCCTGA
CAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTTTTCGTCAGTAACACACATCTG
TCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGT
ATTTATATACACATTGCCCGCGGACTGTACTATGGGTCATACCTCTATAAAGAGACCTGAACTATTGGC
GTGGTCCTACTCCTTTTGGTCATAATAACAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTT
TCTGAGGCGCCACCGTCATTACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGT
GGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCC
GTTTGTTATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCTTGGCTTT
GTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGGTGACCCAGAAAATT
TTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGTGGTACTTCTTATTTGCTTACG
CCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTACTATT
AGTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATCACCCAGTTTTTATTC
TGAACCTTAGTGGCAGACATAATTATTTTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCA
TCATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGA
GCCGAAAATAAAGCATTGAAATGAGCT----------------
>LCK6
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCTTCCAACT
CCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGATTACTCAAATCCTGA
CAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTTTTCGTCAGTAACACACATCTG
TCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGT
ATTTATATACACATTGCCCGCGGACTGTACTATGGGTCATACCTCTATAAAGAGACCTGAACTATTGGC
GTGGTCCTACTCCTTTTGGTCATAATAACAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTT
TCTGAGGCGCCACCGTCATTACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGT
GGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCC
GTTTGTTATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCTTGGCTTT
GTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGGTGACCCAGAAAATT
TTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGTGGTACTTCTTATTTGCTTACG
CCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTACTATT
AGTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATCACCCAGTTTTTATTC
TGAACCTTAGTGGCAGACATAATTATTTTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCA
TCATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGA
GCCGAAAATAAAGCATTGAAATGAGCT----------------
>LCK7
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAAATTGCTAACGATGCACTAGTCGACCTTCCAACT
CCATCTAATATCTCAGCACTATGGAACTTCGGGTCCCTCCTAGGATTATGCTTAATTACTCAAATCCTAA
CAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTTTTCGTCAGTAACACACATCTG
CCGAGACGTTAATTATGGCTGACTCATCCGGAATATACATGCTAACGGAGCATCATTCTTCTTTATCTGT
ATTTACATGCACATTGCCCGCGGACTGTACTATGGGTCATACCTCTATAAAGAGACCTGAACTATTGGC
GTGGTTCTACTCCTTTTAGTCATGATAACAGCCTTTGTAGGCTACGTGCTACCATGAGGTCAAATATCTT
TCTGAGGCGCCACCGTGATTACAAATCTACTATCAGCAGTCCCCTATATAGGAGACGCTCTTGTCCAAT
```

```
GGATTTGAGGTGGCTTTTCCGTAGATAACGCAACGTTAACACGATTCTTCGCCTTCCACTTCCTATTTCC
GTTTGTTATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCTTGGCTTT
GTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGGTGACCCAGAAAATT
TTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGTGATACTTCTTATTTGCTTACG
CCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTACTATT
AGTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATCACCCAGTTTTTATTC
TGAACCTTAGTGGCAGACATAATTATTTTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCA
TCATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGA
GCCGAAAATAAAGCATTGAAATGAGCT----------------
>LCK8
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCTTCCAACT
CCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGATTACTCAAATCCTGA
CAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTTTTCGTCAGTAACACACATCTG
TCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGT
ATTTATATACACATTGCCCGCGGACTGTACTATGGGTCATACCTCTATAAAGAGACCTGAACTATTGGC
GTGGTCCTACTCCTTTTGGTCATAATAACAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTT
TCTGAGGCGCCACCGTCATTACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGT
GGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCC
GTTTGTTATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCTTGGCTTT
GTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGGTGACCCAGAAAATT
TTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGTGGTACTTCTTATTTGCTTACG
CCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTACTATT
AGTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATCACCCAGTTTTTATTC
TGAACCTTAGTGGCAGACATAATTATTTTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCA
TCATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGA
GCCGAAAATAAAGCATTGAAATGAGCT----------------
>LCK9
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCTTCCAACT
CCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGATTACTCAAATCCTGA
CAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTTTTCGTCAGTAACACACATCTG
TCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGT
ATTTATATACACATTGCCCGCGGACTGTACTATGGGTCATACCTCTATAAAGAGACCTGAACTATTGGC
GTGGTTCTACTCCTTTTAGTCATAATAACAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTT
TCTGAGGCGCCACCGTCATTACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGT
GGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCC
GTTTGTTATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCTTGGCTTT
GTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGGTGACCCAGAAAATT
TTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGTGATACTTCTTATTTGCTTACG
CCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTACTATT
AGTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATCACCCAGTTTTTATTC
TGAACCTTAGTGGCAGACATAATTATTTTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCA
TCATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGA
GCCGAAAATAAAGCATTGAAATGAGCT----------------
>LCK10
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCTTCCAACT
CCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGATTACTCAAATCCTGA
CAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTTTTCGTCAGTAACACACATCTG
TCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGT
ATTTATATACACATTGCCCGCGGACTGTACTATGGGTCATACCTCTATAAAGAGACCTGAACTATTGGC
GTGGTCCTACTCCTTTTGGTCATAATAACAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTT
TCTGAGGCGCCACCGTCATTACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGT
GGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCC
GTTTGTTATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCTTGGCTTT
GTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGGTGACCCAGAAAATT
TTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGTGGTACTTCTTATTTGCTTACG
```

```
CCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTACTATT
AGTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATCACCCAGTTTTTATTC
TGAACCTTAGTGGCAGACATAATTATTTTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCA
TCATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGA
GCCGAAAATAAAGCATTGAAATGAGCT----------------
>CATC1
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCTTCCAACT
CCATCTAATATTTCAGTACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGATTACTCAAATCCTGA
CAGGACTATTCCTAGCCATACATTATACTTCCGATATTTCAACTGCCTTTTCGTCAGTAACACACATCTGT
CGAGACGTTAATTATGGCTGACTCATCCGGAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGTA
TTTATATACACATTGCCCGCGGACTGTACTATGGATCATACCTCTATAAAGAGACCTGAACTATTGGCGT
GGTCTTACTCCTTTTGGTCATAATAACAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTC
TGAGGCGCCACCGTCATTACGAATTTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGG
ATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGT
TTGTTATTGCCGGCGCAACGGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGCCG
GGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCGTACAAAGATCTCCTTGGCTTTGT
ACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGGTGACCCAGAAAATTTC
ACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGTGGTACTTCTTATTTGCTTACGC
CATTCTACGCTCTATCCCAAACAAGCTGGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTACTATTA
GTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCT
GAACCTTAGTGGCAGACATAATCATTTTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCAT
CATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGAG
CCGAAAATAAAGCATTGAAATGAGCT----------------
>CATC2
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCTTCCAACT
CCATCTAATATTTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGATTACTCAAATCCTGA
CAGGACTATTCCTAGCCATACATTATACTTCCGATATTTCAACTGCCTTTTCGTCAGTAACACACATCTGT
CGAGACGTTAACTATGGCTGACTCATCCGGAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGTA
TTTATATACACATTGCCCGCGGACTGTACTATGGATCATACCTCTATAAAGAGACCTGAACTATTGGCGT
GGTCTTACTCCTTTTGGTCATAATAACAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTC
TGAGGCGCCACCGTCATTACGAATTTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGG
ATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGT
TTGTTATTGCCGGCGCAACGGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGCCG
GGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCGTACAAAGATCTCCTTGGCTTTGT
ACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGGTGACCCAGAAAATTTC
ACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGTGGTACTTCTTATTTGCTTACGC
CATTCTACGCTCTATCCCAAACAAGCTGGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTACTATTA
GTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCT
GAACCTTAGTGGCAGACATAATCATTTTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCAT
CATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGAG
CCGAAAATAAAGCATTGAAATGAGCT----------------
>CATC3
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCTTCCAACT
CCATCTAATATTTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGATTACTCAAATCCTGA
CAGGACTATTCCTAGCCATACATTATACTTCCGATATTTCAACTGCCTTTTCGTCAGTAACACACATCTGT
CGAGACGTTAACTATGGCTGACTCATCCGGAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGTA
TTTATATACACATTGCCCGCGGACTGTACTATGGATCATACCTCTATAAAGAGACCTGAACTATTGGCGT
GGTCTTACTCCTTTTGGTCATAATAACAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTC
TGAGGCGCCACCGTCATTACGAATTTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGG
ATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGT
TTGTTATTGCCGGCGCAACGGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGCCG
GGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCGTACAAAGATCTCCTTGGCTTTGT
ACTAATATTGTTAGCTCTCACATCCCTGACGCTATTTTCTCCAACTCTGCTCGGTGACCCAGAAAATTTC
ACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGTGGTACTTCTTATTTGCTTACGC
CATTCTACGCTCTATCCCAAACAAGCTGGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTACTATTA
GTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCT
GAACCTTAGTGGCAGACATAATCATTTTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCAT
CATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGAG
CCGAAAATAAAGCATTGAAATGAGCT----------------
```

```
>EFSGR1
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCTTCCAACT
CCATCTAATATTTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGATTACTCAAATCCTGA
CAGGACTATTCCTAGCCATACATTATACTTCCGATATTTCAACTGCCTTTTCGTCAGTAACACACATCTGT
CGAGACGTTAACTATGGCTGACTCATCCGGAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGTA
TTTATATACACATTGCCCGCGGACTGTACTATGGATCATACCTCTATAAAGAGACCTGAACTATTGGCGT
GGTCTTACTCCTTTTGGTCATAATAACAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTC
TGAGGCGCCACCGTCATTACGAATTTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGG
ATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGT
TTGTTATTGCCGGCGCAACGGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGCCG
GGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCGTACAAAGATCTCCTTGGCTTTGT
ACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGGTGACCCAGAAAATTTC
ACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGTGGTACTTCTTATTTGCTTACGC
CATTCTACGCTCTATCCCAAACAAGCTGGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTACTATTA
GTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCT
GAACCTTAGTGGCAGACATAATCATTTTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCAT
CATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGAG
CCGAAAATAAAGCATTGAAATGAGCT----------------
>EFSGR2
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCTTCCAACT
CCATCTAATATTTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGATTACTCAAATCCTGA
CAGGACTATTCCTAGCCATACATTATACTTCCGATATTTCAACTGCCTTTTCGTCAGTAACACACATCTGT
CGAGACGTTAACTATGGCTGACTCATCCGGAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGTA
TTTATATACACATTGCCCGCGGACTGTACTATGGATCATACCTCTATAAAGAGACCTGAACTATTGGCGT
GGTCTTACTCCTTTTGGTCATAATAACAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTC
TGAGGCGCCACCGTCATTACGAATTTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGG
ATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGT
TTTGTTATGCCGGCGCAACGGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGCCG
GGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCGTACAAAGATCTCCTTGGCTTTGT
ACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGGTGACCCAGAAAATTTC
ACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGTGGTACTTCTTATTTGCTTACGC
CATTCTACGCTCTATCCCAAACAAGCTGGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTACTATTA
GTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCT
GAACCTTAGTGGCAGACATAATCATTTTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCAT
CATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGAG
CCGAAAATAAAGCATTGAAATGAGCT----------------
>NFSGR1
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCTTCCAACT
CCATCTAATATTTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGATTACTCAAATCCTGA
CAGGACTATTCCTAGCCATACATTATACTTCCGATATTTCAACTGCCTTTTCGTCAGTAACACACATCTGT
CGAGACGTTAACTATGGCTGACTCATCCGGAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGTA
TTTATATACACATTGCCCGCGGACTGTACTATGGATCATACCTCTATAAAGAGACCTGAACTATTGGCGT
GGTCTTACTCCTTTTGGTCATAATAACAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTC
TGAGGCGCCACCGTCATTACGAATTTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGG
ATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGT
TTGTTATTGCCGGCGCAACGGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGCCG
GGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCGTACAAAGATCTCCTTGGCTTTGT
ACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGGTGACCCAGAAAATTTC
ACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGTGGTACTTCTTATTTGCTTACGC
CATTCTACGCTCTATCCCAAACAAGCTGGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTACTATTA
GTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCT
GAACCTTAGTGGCAGACATAATCATTTTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCAT
CATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGAG
CCGAAAATAAAGCATTGAAATGAGCT----------------
>NFSGR2
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCTTCCAACT
CCATCTAATATTTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGATTACTCAAATCCTGA
CAGGACTATTCCTAGCCATACATTATACTTCCGATATTTCAACTGCCTTTTCGTCAGTAACACACATCTGT
CGAGACGTTAACTATGGCTGACTCATCCGGAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGTA
```

```
TTTATATACACATTGCCCGCGGACTGTACTATGGATCATACCTCTATAAAGAGACCTGAACTATTGGCGT
GGTCTTACTCCTTTTGGTCATAATAACAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTC
TGAGGCGCCACCGTCATTACGAATTTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGG
ATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGT
TTGTTATTGCCGGCGCAACGGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGCCG
GGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCGTACAAAGATCTCCTTGGCTTTGT
ACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGGTGACCCAGAAAATTTC
ACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGTGGTACTTCTTATTTGCTTACGC
CATTCTACGCTCTATCCCAAACAAGCTGGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTACTATTA
GTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCT
GAACCTTAGTGGCAGACATAATCATTTTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCAT
CATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGAG
CCGAAAATAAAGCATTGAAATGAGCT----------------
>WFSGR1
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCTTCCAACT
CCATCTAATATTTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGATTACTCAAATCCTGA
CAGGACTATTCCTAGCCATACATTATACTTCCGATATTTCAACTGCCTTTTCGTCAGTAACACACATCTGT
CGAGACGTTAACTATGGCTGACTCATCCGGAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGTA
TTTATATACACATTGCCCGCGGACTGTACTATGGATCATACCTCTATAAAGAGACCTGAACTATTGGCGT
GGTCTTACTCCTTTTGGTCATAATAACAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTC
TGAGGCGCCACCGTCATTACGAATTTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGG
ATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGT
TTGTTATTGCCGGCGCAACGGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGCCG
GGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCGTACAAAGATCTCCTTGGCTTTGT
ACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGGTGACCCAGAAAATTTC
ACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGTGGTACTTCTTATTTGCTTACGC
CATTCTACGCTCTATCCCAAACAAGCTGGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTACTATTA
GTAGTGCCAGTTTTACATACCTCGAAGCAACGAGGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCT
GAACCTTAGTGGCAGACATAATCATTTTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCAT
CATTGGCCAAGTCGCGTCGGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGAG
CCGAAAATAAAGCATTGAAATGAGCT----------------
>STC1
ATGGCAAGCCTACGAAAAACCCATCCGCTAATAAAAATTGCTAACAATGCACTAGTTGACCTTCCAACC
CCATCTAATATTTCAGCACTATGGAACTTCGGGTCCCTCCTAGGATTATGCTTAATTACTCAAATCCTAA
CAGGACTATTCCTGGCCATACATTATACTTCTGATATCTCAACTGCCTTTTCATCAGTAACACATATCTGC
CGAGACGTAAACTATGGCTGACTCATCCGAAACATACATGCTAACGGAGCATCGTTCTTCTTTATCTGTA
TTTACATACACATTGCCCGCGGACTGTACTACGGGTCATACCTCTATAAAGAGACCTGAAGTATTGGCG
TAGTTCTACTCCTTTTAGTAATGATGACAGCCTTTGTGGGCTATGTACTGCCCTGAGGACAAATATCTTT
TTGAGGTGCCACCGTGATTACAAATCTACTATCAGCAGTCCCCTATATGGGCGACGCCTTTGTCCAATG
GATTTGAGGCGGCTTTTCCGTAGATAACGCAACATTAACACGATTCTTCGCCTACCACTTCCTATTCCCG
TTCGTTATCGCCGGTGCAACGGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGCC
GGATTAAACTCCGACGCAGATAAAATCTCTTTCCACCCATACTTCTCATACAAGGACCTCCTTGGCTTTG
TACTAATGTTATTAGCTCTCACATCCCTAACATTATTCTCCCCAACACTGCTCGCCGACCCAGAAAATTT
CACCCCGCCAAACCCCCTGGTCACCCCACCGCACATCCAACCGGAGTGATACTTCCTGTTTGCCTACG
CCATTCTACGATCTATTCCAAACAAGCTAGGAGGGGTCCTAGCGCTATTGTTTAGCATCCTAGTACTATT
AGTAGTGCCAGTTTTACACACCTCGAAGCAACGAGGACTAACTTTCCGTCCGATCACCCAGTTCTTATT
CTGAACCTTAGTCGCAGACATAATCATTTTAACATGAATTGGAGGCATACCCGTAGAACACCCATATATC
ATCATTGGCCAAATCGCCTCGGTGCTATACTTCGCACTCTTCCTTGTTCTCGCCCCCCTCGCAGGGTGA
GCCGAGAATAAAGCATTGAAATGAGCT----------------
>STC2
ATGGCAAGCCTACGAAAAACCCATCCGCTAATAAAAATTGCTAACAATGCACTAGTTGACCTTCCAACC
CCATCTAATATTTCAGCACTATGGAACTTCGGGTCCCTCCTAGGATTATGCTTAATTACTCAAATCCTAA
CAGGACTATTCCTCGCCATACATTATACTTCTGATATCTCAACTGCCTTTTCATCAGTAACACATATCTGC
CGAGACGTAAACTATGGCTGACTCATCCGAAACATACATGCTAACGGAGCATCGTTCTTCTTTATCTGTA
TTTACATACACATTGCCCGCGGACTGTACTACGGGTCATACCTCTATAAAGAGACCTGAAGTATTGGCG
TAGTTCTACTCCTTTTAGTAATGATGACAGCCTTTGTGGGCTATGTACTGCCCTGAGGACAAATATCTTT
TTGAGGTGCCACCGTGATTACAAATCTACTATCAGCAGTCCCCTATATGGGCGACGCCTTTGTCCAATG
GATTTGAGGCGGCTTTTCCGTAGATAACGCAACATTAACACGATTCTTCGCCTTCCACTTCCTATTCCCG
TTCGTTATCGCCGGTGCAACGGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGCC
```

```
GGATTAAACTCCGACGCAGATAAAATCTCTTTCCACCCATACTTCTCATACAAGGACCTCCTTGGCTTTG
TACTAATGTTATTAGCTCTCACATCCCTAACATTATTCTCCCCAACACTGCTCGGCGACCCAGAAAATTT
CACCCCGGCAAACCCCCTGGTCACCCCACCGCACATCCAACCGGAGTGATACTTCCTGTTTGCCTACG
CCATTCTACGATCTATTCCAAACAAGCTAGGAGGGGTCCTAGCGCTATTGTTTAGCATCCTAGTACTATT
AGTAGTGCCAGTTTTACACACCTCGAAGCAACGAGGACTAACTTTCCGTCCGATCACCCAGTTCTTATT
CTGAACCTTAATCGCAGACATAATCATTTTAACATGAATTGGAGGCATACCCGTAGAACACCCATATATC
ATCATTGGCCAAATCGCCTCGGTGCTATACTTCGCACTCTTCCTTGTTCTCGCCCCCCTCGCAGGGTGA
GCCGAGAATAAAGCATTGAAATGAGCT----------------
>STC3
ATGGCAAGCCTACGAAAAACCCATCCGCTAATAAAAATTGCTAACAATGCACTAGTTGACCTTCCAACC
CCATCTAATATTTCAGCACTATGGAACTTCGGGTCCCTCCTAGGATTATGCTTAATTACTCAAATCCTAA
CAGGACTATTCCTGGCCATACATTATACTTCTGATATCTCAACTGCCTTTTCATCAGTAACACATATCTGC
CGAGACGTAAACTATGGCTGACTCATCCGAAACATACATGCTAACGGAGCATCGTTCTTCTTTATCTGTA
TTTACATACACATTGCCCGCGGACTGTACTACGGGTCATACCTCTATAAAGAGACCTGAAGTATTGGCG
TAGTTCTACTCCTTTTAGTAATGATGACAGCCTTTGTGGGCTATGTACTGCCCTGAGGACAAATATCTTT
TTGAGGTGCCACCGTGATTACAAATCTACTATCAGCAGTCCCCTATATGGGGGACGCCTTTGTCCAATG
GATTTGAGGCGGCTTTTCCGTAGATAACGCAACATTAACACGATTCTTCGCCTTCCACTTCCTATTCCCG
TTCGTTATCGCCGGTGCAACGGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGCC
GGATTAAACTCCGACGCAGATAAAATCTCTTTCCACCCATACTCCTCATACAAGGACCTCCTTGGCTTTG
TACTAATGTTATTAGCTTTCACATCCCTAACATTATTCTCCCCAACACTGCTCGCCGACCCAGAAAATTT
CACCCCGGCAAACCCCCTGGTCACCCCACCGCACATCCAACCGGAGTGATACTTCCTGTTTGCCTACG
CCATTCTACGATCTATTCCAAACAAGCTAGGAGGGGTCCTAGCGCTATTGTTTAGCATCCTAGTACTATT
AGTAGTGCCAGTTTTACACACCTCGAAGCAACGAGGACTAACTTTCCGTCCGATCACCCAGTTCTTATT
CTGAACCTTAGTCGCAGACATAATCATTTTAACATGAATTGGAGGCATACCCGTAGAACACCCATATATC
ATCATTGGCCAAATCGCCTCGGTGCTATACTTCGCACTCTTCCTTGTTCTCGCCCCCCTCGCAGGGTGA
GCCGAGAATAAAGCATTGAAATGAGCT----------------
>BC1
ATGGCAAGCCTACGAAAAACCCATCCGCTAATAAAAATTGCTAACAATGCACTAGTTGACCTTCCAACC
CCATCTAATATTTCAGCACTATGGAACTTCGGGTCCCTCCTAGGATTATGCTTAATTACTCAAATCCTAA
CAGGACTATTCCTGGCCATACATTATACTTCTGATATCTCAACTGCCTTTTCATCAGTAACACATATCTGC
CGAGACGTAAACTATGGCTGACTCATCCGAAACATACATGCTAACGGAGCATCGTTCTTCTTTATCTGTA
TTTACATACACATTGCCCGCGGACTGTACTACGGGTCATACCTCTATAAAGAGACCTGAAGTATTGGCG
TAGTTCTACTCCTTTTAGTAATGATGACAGCCTTTGTGGGCTATGTACTGCCCTGAGGACAAATATCTTT
TTGAGGTGCCACCGTGATTACAAATCTACTATCAGCAGTCCCCTATATGGGCGACGCCTTTGTCCAATG
GATTTGAGGCGGCTTTTCCGTAGATAACGCAACATTAACACGATTCTTCGCCTTCCACTTCCTATTCCCG
TTCGTTATCGCCGGTGCAACGGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGCC
GGATTAAACTCCGACGCAGATAAAATCTCTTTCCACCCATACTTCTCATACAAGGACCTCCTTGGCTTTG
TACTAATGTTATTAGCTCTCACATCCCTAACATTATTCTCCCCAACACTGCTCGCCGACCCAGAAAATTT
CACCCCGCCAAACCCCCTGGTCACCCCACCGCACATCCAACCGGAGTGATACTTCCTGTTTGCCTACG
CCATTCTACGATCTATTCCAAACAAGCTAGGAGGGGTCCTAGCGCTATTGTTTAGCATCCTAGTACTATT
AGTAGTGCCAGTTTTACACACCTCGAAGCAACGAGGACTAACTTTCCGTCCGATCACCCAGTTCTTATT
CTGAACCTTAGTCGCAGACATAATCATTTTAACATGAATTGGAGGCATACCCGTAGAACACCCATATATC
ATCATTGGCCAAATCGCCTCGGTGCTATACTTCGCACTCTTCCTTGTTCTCGCCCCCCTCGCAGGGTGA
GCCGAGAATAAAGCATTGAAATGAGCT----------------
>BC2
ATGGCAAGCCTACGAAAAACCCATCCGCTAATAAAAATTGCTAACAATGCACTAGTTGACCTTCCAACC
CCATCTAATATTTCAGCACTATGGAACTTCGGGTCCCTCCTAGGATTATGCTTAATTACTCAAATCCTAA
CAGGACTATTCCTGGCCATACATTATACTTCTGATATCTCAACTGCCTTTTCATCAGTAACACATATCTGC
CGAGACGTAAACTATGGCTGACTCATCCGAAACATACATGCTAACGGAGCATCGTTCTTCTTTATCTGTA
TTTACATACACATTGCCCGCGGACTGTACTACGGGTCATACCTCTATAAAGAGACCTGAAGTATTGGCG
TAGTTCTACTCCTTTTAGTAATGATGACAGCCTTTGTGGGCTATGTACTGCCCTGAGGACAAATATCTTT
TTGAGGTGCCACCGTGATTACAAATCTACTATCAGCAGTCCCCTATATGGGCGACGCCTTTGTCCAATG
GATTTGAGGCGGCTTTTCCGTAGATAACGCAACATTAACACGATTCTTCGCCTTCCACTTCCTATTCCCG
TTCGTTATCGCCGGTGCAACGGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGCC
GGATTAAACTCCGACGCAGATAAAATCTCTTTCCACCCATACTTCTCATACAAGGACCTCCTTGGCTTTG
TACTAATGTTATTAGCTCTCACATCCCTAACATTATTCTCCCCAACACTGCTCGCCGACCCAGAAAATTT
CACCCCGGCAAACCCCCTGGTCACCCCACCGCACATCCAACCGGAGTGATACTTCCTGTTTGCCTACG
CCATTCTACGATCTATTCCAAACAAGCTAGGAGGGGTCCTAGCGCTATTGTTTAGCATCCTAGTACTATT
AGTAGTGCCAGTTTTACACACCTCGAAGCAACGAGGACTAACTTTCCGTCCGATCACCCAGTTCTTATT
```

```
CTGAACCTTAGTCGCAGACATAATCATTTTAACATGAATTGGAGGCATACCCGTAGAACACCCATATATC
ATCATTGGCCAAATCGCCTCGGTGCTATACTTCGCACTCTTCCTTGTTCTCGCCCCCCTCGCAGGGTGA
GCCGAGAATAAAGCATTGAAATGAGCT----------------
>BC3
ATGGCAAGCCTACGAAAAACCCATCCGCTAATAAAAATTGCTAACAATGCACTAGTTGACCTTCCAACC
CCATCTAATATTTCAGCACTATGGAACTTCGGGTCCCTCCTAGGATTATGCTTAATTACTCAAATCCTAA
CAGGACTATTCCTGGCCATACATTATACTTCTGATATCTCAACTGCCTTTTCATCAGTAACACATATCTGC
CGAGACGTAAACTATGGCTGACTCATCCGAAACATACATGCTAACGGAGCATCGTTCTTCTTTATCTGTA
TTTACATACACATTGCCCGCGGACTGTACTACGGGTCATACCTCTATAAAGAGACCTGAAGTATTGGCG
TAGTTCTACTCCTTTTAGTAATGATGACAGCCTTTGTGGGCTATGTACTGCCCTGAGGACAAATATCTTT
TTGAGGTGCCACCGTGATTACAAATCTACTATCAGCAGTCCCCTATATGGGCGACGCCTTTGTCCAATG
GATTTGAGGCGGCTTTTCCGTAGATAACGCAACATTAACACGATTCTTCGCCTTCCACTTCCTATTCCCG
TTCGTTATCGCCGGTGCAACGGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGCC
GGATTAAACTCCGACGCAGATAAAATCTCTTTCCACCCATACTTCTCATACAAGGACCTCCTTGGCTTTG
TACTAATGTTATTAGCTCTCACATCCCTAACATTATTCTCCCCAACACTGCTCGGCGACCCAGAAAATTT
CACCCCGCCAAACCCCCTGGTCACCCCACCGCACATCCAACCGGAGTGATACTTCCTGTTTGCCTACG
CCATTCTACGATCTATTCCAAACAAGCTAGGAGGGGTCCTAGCGCTATTGTTTAGCATCCTAGTACTATT
AGTAGTGCCAGTTTTACACACCTCGAAGCAACGAGGACTAACTTTCCGTCCGATCACCCAGTTCTTATT
CTGAACCTTAGTCGCAGACATAATCATTTTAACATGAATTGGAGGCATACCCGTAGAACACCCATATATC
ATCATTGGCCAAATCGCCTCGGTGCTATACTTCGCACTCTTCCTTGTTCTCGCCCCCCTCGCAGGGTGA
GCCGAGAATAAAGCATTGAAATGAGCT----------------
>SLO1
ATGGCAAGCCTACGAAAAACCCATCCGCTAATAAAAATTGCTAACAATGCACTAGTTGACCTTCCAACC
CCATCTAATATTTCAGCACTATGGAACTTCGGGTCCCTCCTAGGATTATGCTTAATTACTCAAATCCTAA
CAGGACTATTCCTCGCCATACATTATACTTCTGATATCTCAACTGCCTTTTCATCAGTAACACATATCTGC
CGAGACGTAAACTATGGCTGACTCATCCGAAACATACATGCTAACGGAGCATCGTTCTTCTTTATCTGTA
TTTACATACACATTGCCCGCGGACTGTACTACGGGTCATACCTCTATAAAGAGACCTGAAGTATTGGCG
TAGTTCTACTCCTTTTAGTAATGATGACAGCCTTTGTGGGCTATGTACTGCCCTGAGGACAAATATCTTT
TTGAGGTGCCACCGTGATTACAAATCTACTATCAGCAGTCCCCTATATGGGCGACGCCTTTGTCCAATG
GATTTGAGGCGGCTTTTCCGTAGATAACGCAACATTAACACGATTCTTCGCCTTCCACTTCCTATTCCCG
TTCGTTATCGCCGGTGCAACGGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGCC
GGATTAAACTCCGACGCAGATAAAATCTCTTTCCACCCATACTTCTCATACAAGGACCTCCTTGGCTTTG
TACTAATGTTATTAGCTCTCACATCCCTAACATTATTCTCCCCAACACTGCTCGGCGACCCAGAAAATTT
CACCCCGGCAAACCCCCTGGTCACCCCACCGCACATCCAACCGGAGTGATACTTCCTGTTTGCCTACG
CCATTCTACGATCTATTCCAAACAAGCTAGGAGGGGTCCTAGCGCTATTGTTTAGCATCCTAGTACTATT
AGTAGTGCCAGTTTTACACACCTCGAAGCAACGAGGACTAACTTTCCGTCCGATCACCCAGTTCTTATT
CTGAACCTTAATCGCAGACATAATCATTCTAACATGAATTGGAGGCATACCCGTAGAACACCCATATATC
ATCATTGGCCAAATCGCCTCGGTGCTATACTTCGCACTCTTCCTTGTTCTCGCCCCCCTCGCAGGGTGA
GCCGAGAATAAAGCATTGAAATGAGCT----------------
>SLO2
ATGGCAAGCCTACGAAAAACCCATCCGCTAATAAAAATTGCTAACAATGCACTAGTTGACCTTCCAACC
CCATCTAATATTTCAGCACTATGGAACTTCGGGTCCCTCCTAGGATTATGCTTAATTACTCAAATCCTAA
CAGGACTATTCCTCGCCATACATTATACTTCTGATATCTCAACTGCCTTTTCATCAGTAACACATATCTGC
CGAGACGTAAACTATGGCTGACTCATCCGAAACATACATGCTAACGGAGCATCGTTCTTCTTTATCTGTA
TTTACATACACATTGCCCGCGGACTGTACTACGGGTCATACCTCTATAAAGAGACCTGAAGTATTGGCG
TAGTTCTACTCCTTTTAGTAATGATGACAGCCTTTGTGGGCTATGTACTGCCCTGAGGACAAATATCTTT
TTGAGGTGCCACCGTGATTACAAATCTACTATCAGCAGTCCCCTATATGGGCGACGCCTTTGTCCAATG
GATTTGAGGCGGCTTTTCCGTAGATAACGCAACATTAACACGATTCTTCGCCTTCCACTTCCTATTCCCG
TTCGTTATCGCCGGTGCAACGGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGCC
GGATTAAACTCCGACGCAGATAAAATCTCTTTCCACCCATACTTCTCATACAAGGACCTCCTTGGCTTTG
TACTAATGTTATTAGCTCTCACATCCCTAACATTATTCTCCCCAACACTGCTCGGCGACCCAGAAAATTT
CACCCCGCCAAACCCCCTGGTCACCCCACCGCACATCCAACCGGAGTGATACTTCCTGTTTGCCTACG
CCATTCTACGATCTATTCCAAACAAGCTAGGAGGGGTCCTAGCGCTATTGTTTAGCATCCTAGTACTATT
AGTAGTGCCAGTTTTACACACCTCGAAGCAACGAGGACTAACTTTCCGTCCGATCACCCAGTTCTTATT
CTGAACCTTAATCGCAGACATAATCATTTTAACATGAATTGGAGGCATACCCGTAGAACACCCATATATC
ATCATTGGCCAAATCGCCTCGGTGCTATACTTCGCACTCTTCCTTGTTCTCGCCCCCCTCGCAGGGTGA
GCCGAGAATAAAGCATTGAAATGAGCT----------------
>SLO3
```

```
ATGGCAAGCCTACGAAAAACCCATCCGCTAATAAAAATTGCTAACAATGCACTAGTTGACCTTCCAACC
CCATCTAATATTTCAGCACTATGGAACTTCGGGTCCCTCCTAGGATTATGCTTAATTACTCAAATCCTAA
CAGGACTATTCCTCGCCATACATTATACTTCTGATATCTCAACTGCCTTTTCATCAGTAACACATATCTGC
CGAGACGTAAACTATGGCTGACTCATCCGAAACATACATGCTAACGGAGCATCGTTCTTCTTTATCTGTA
TTTACATACACATTGCCCGCGGACTGTACTACGGGTCATACCTCTATAAAGAGACCTGAAGTATTGGCG
TAGTTCTACTCCTTTTAGTAATGATGACAGCCTTTGTGGGCTATGTACTGCCCTGAGGACAAATATCTTT
TTGAGGTGCCACCGTGATTACAAATCTACTATCAGCAGTCCCCTATATGGGCGACGCCTTTGTCCAATG
GATTTGAGGCGGCTTTTCCGTAGATAACGCAACATTAACACGATTCTTCGCCTTCCACTTCCTATTCCCG
TTCGTTATCGCCGGTGCAACGGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGCC
GGATTAAACTCCGACGCAGATAAAATCTCTTTCCACCCATACTTCTCATACAAGGACCTCCTTGGCTTTG
TACTAATGTTATTAGCTCTCACATCCCTAACATTATTCTCCCCAACACTGCTCGGCGACCCAGAAAATTT
CACCCCGGCAAACCCCCTGGTCACCCCACCGCACATCCAACCGGAGTGATACTTCCTGTTTGCCTACG
CCATTCTACGATCTATTCCAAACAAGCTAGGAGGGGTCCTAGCGCTATTGTTTAGCATCCTAGTACTATT
AGTAGTGCCAGTTTTACACACCTCGAAGCAACGAGGACTAACTTTCCGTCCGATCACCCAGTTCTTATT
CTGAACCTTAATCGCAGACATAATCATTTTAACATGAATTGGAGGCATACCCGTAGAACACCCATATATC
ATCATTGGCCAAATCGCCTCGGTGCTATACTTCGCACTCTTCCTTGTTCTCGCCCCCCTCGCAGGGTGA
GCCGAGAATAAAGCATTGAAATGAGCT----------------
>MNCR1
ATGGCAAGCCTACGAAAAACCCATCCGCTAATAAAAATTGCTAACAATGCACTAGTTGACCTTCCAACC
CCATCTAATATTTCAGCACTATGGAACTTCGGGTCCCTCCTAGGATTATGCTTAATTACTCAAATCCTAA
CAGGACTATTCCTGGCCATACATTATACTTCTGATATCTCAACTGCCTTTTCATCAGTAACACATATCTGC
CGAGACGTAAACTATGGCTGACTCATCCGAAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGTA
TTTACATGCACATTGCCCGCGGACTGTACTACGGGTCATACCTCTATAAAGAGACCTGAAGTATTGGCG
TAGTTCTACTCCTTTTAGTAATGATGACAGCCTTTGTGGGCTATGTACTGCCCTGAGGACAAATATCTTT
TTGAGGTGCCACCGTGATTACAAATCTACTATCAGCAGTCCCCTATATGGGCGACGCCTTTGTCCAATG
GATTTGAGGCGGCTTTTCCGTAGATAACGCAACATTAACACGATTCTTCGCCTTCCACTTCCTATTCCCG
TTCGTTATCGCCGGTGCAACGGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGCC
GGATTAAACTCCGACGCAGATAAAATCTCTTTCCACCCATACTTCTCATACAAGGACCTCCTTGGCTTTG
TACTAATGTTATTAGCTCTCACATCCCTAACATTATTCTCCCCAACACTGCTCGGCGACCCAGAAAATTT
CACCCCGGCAAACCCCCTGGTCACCCCACCGCACATCCAACCGGAGTGATACTTCCTGTTTGCCTACG
CCATTCTACGATCTATTCCAAACAAGCTAGGAGGGGTCCTAGCGCTATTGTTTAGCATCCTAGTACTATT
AGTAGTGCCAGTTTTACACACCTCGAAGCAACGAGGACTAACTTTCCGTCCGATCACCCAGTTCTTATT
CTGAACCTTAGTCGCAGACATAATCATTTTAACATGAATTGGAGGCATACCCGTAGAACACCCATATATC
ATCATTGGCCAAATCGCCTCGGTGCTATACTTCGCACTCTTCCTTGTTCTCGCCCCCCTCGCAGGGTGA
GCCGAGAATAAAGCATTGAAATGAGCT----------------
>MNCR2
ATGGCAAGCCTACGAAAAACCCATCCGCTAATAAAAATTGCTAACAATGCACTAGTTGACCTTCCAACC
CCATCTAATATTTCAGCACTATGGAACTTCGGGTCCCTCCTAGGATTATGCTTAATTACTCAAATCCTAA
CAGGACTATTCCTGGCCATACATTATACTTCTGATATCTCAACTGCCTTTTCATCAGTAACACATATCTGC
CGAGACGTAAACTATGGCTGACTCATCCGAAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGTA
TTTACATGCACATTGCCCGCGGACTGTACTACGGGTCATACCTCTATAAAGAGACCTGAAGTATTGGCG
TAGTTCTACTCCTTTTAGTAATGATGACAGCCTTTGTGGGCTATGTACTGCCCTGAGGACAAATATCTTT
TTGAGGTGCCACCGTGATTACAAATCTACTATCAGCAGTCCCCTATATGGGCGACGCCTTTGTCCAATG
GATTTGAGGCGGCTTTTCCGTAGATAACGCAACATTAACACGATTCTTCGCCTTCCACTTCCTATTCCCG
TTCGTTATCGCCGGTGCAACGGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGCC
GGATTAAACTCCGACGCAGATAAAATCTCTTTCCACCCATACTTCTCATACAAGGACCTCCTTGGCTTTG
TACTAATGTTATTAGCTCTCACATCCCTAACATTATTCTCCCCAACACTGCTCGGCGACCCAGAAAATTT
CACCCCGGCAAACCCCCTGGTCACCCCACCGCACATCCAACCGGAGTGATACTTCCTGTTTGCCTACG
CCATTCTACGATCTATTCCAAACAAGCTAGGAGGGGTCCTAGCGCTATTGTTTAGCATCCTAGTACTATT
AGTAGTGCCAGTTTTACACACCTCGAAGCAACGAGGACTAACTTTCCGTCCGATCACCCAGTTCTTATT
CTGAACCTTAGTCGCAGACATAATCATTTTAACATGAATTGGAGGCATACCCGTAGAACACCCATATATC
ATCATTGGCCAAATCGCCTCGGTGCTATACTTCGCACTCTTCCTTGTTCTCGCCCCCCTCGCAGGGTGA
GCCGAGAATAAAGCATTGAAATGAGCT----------------
>CUYR1
ATGGCAAGCCTACGAAAAACCCATCCGCTAATAAAAATTGCTAACAATGCACTAGTTGACCTTCCAACC
CCATCTAATATTTCAGCACTATGGAACTTCGGGTCCCTCCTAGGATTATGCTTAATTACTCAAATCCTAA
CAGGACTATTCCTGGCCATACATTATACTTCTGATATCTCAACTGCCTTTTCATCAGTAACACATATCTGC
CGAGACGTAAACTATGGCTGACTCATCCGAAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGTA
TTTACATGCGCATTGCCCGCGGACTGTACTACGGGTCATACCTCTATAAAGAGACCTGAAGTATTGGCG
```

```
TAGTTCTACTCCTTTTAGTAATGATGACAGCCTTTGTGGGCTATGTACTGCCCTGAGGACAAATATCTTT
TTGAGGTGCCACCGTGATTACAAATCTACTATCAGCAGTCCCCTTTATGGGCGACGCCTTTGTCCAATG
GATTTGAGGCGGCTTTTCCGTAGATAACGCAACATTAACACGATTCTTCGCCTTCCACTTCCTATTCCCG
TTCGTTATCGCCGGTGCAACGGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGCC
GGATTAAACTCCGACGCAGATAAAATCTCTTTCCACCCATACTTCTCATACAAGGACCTCCTTGGCTTTG
TACTAATGTTATTAGCTCTCACATCCCTAACATTATTCTCCCCAACACTGCTCGGCGACCCAGAAAATTT
CACCCCGGCAAACCCCCTGGTCACCCCACCGCACATCCAACCGGAGTGATACTTCCTGTTTGCCTACG
CCATTCTACGATCTATTCCAAACAAGCTAGGAGGGGTCCTAGCGCTATTGTTTAGCATCCTAGTACTATT
AGTAGTGCCAGTTTTACACACCTCGAAGCAACGAGGACTAACTTTCCGTCCGATCACCCAGTTCTTATT
CTGAACCTTAGTCGCAGACATAATCATTTTAACATGAATTGGAGGCATACCCGTAGAACACCCATATATC
ATCATTGGCCAAATCGCCTCGGTGCTATACTTCGCACTCTTCCTTGTTCTCGCCCCCCTCGCAGGGTGA
GCCGAGAATAAAGCATTGAAATGAGCT----------------
>CUYR2
ATGGCAAGCCTACGAAAAACCCATCCGCTAATAAAAATTGCTAACAATGCACTAGTTGACCTTCCAACC
CCATCTAATATTTCAGCACTATGGAACTTCGGGTCCCTCCTAGGATTATGCTTAATTACTCAAATCCTAA
CAGGACTATTCCTGGCCATACATTATACTTCTGATATCTCAACTGCCTTTTCATCAGTAACACATATCTGC
CGAGACGTAAACTATGGCTGACTCATCCGAAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGTA
TTTACATGCACATTGCCCGCGGACTGTACTACGGGTCATACCTCTATAAAGAGACCTGAAGTATTGGCG
TAGTTCTACTCCTTTTAGTAATGATGACAGCCTTTGTGGGCTATGTACTGCCCTGAGGACAAATATCTTT
TTGAGGTGCCACCGTGATTACAAATCTACTATCAGCAGTCCCCTATATGGGCGACGCCTTTGTCCAATG
GATTTGAGGCGGCTTTTCCGTAGATAACGCAACATTAACACGATTCTTCGCCTTCCACTTCCTATTCCCG
TTCGTTATCGCCGGTGCAACGGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGCC
GGATTAAACTCCGACGCAGATAAAATCTCTTTCCACCCATACTTCTCATACAAGGACCTCCTTGGCTTTG
TACTAATGTTATTAGCTCTCACATCCCTAACATTATTCTCCCCAACACTGCTCGGCGACCCAGAAAATTT
CACCCCGGCAAACCCCCTGGTCACCCCACCGCACATCCAACCGGAGTGATACTTCCTGTTTGCCTACG
CCATTCTACGATCTATTCCAAACAAGCTAGGAGGGGTCCTAGCGCTATTGTTTAGCATCCTAGTACTATT
AGTAGTGCCAGTTTTACACACCTCGAAGCAACGAGGACTAACTTTCCGTCCGATCACCCAGTTCTTATT
CTGAACCTTAGTCGCAGACATAATCATTTTAACATGAATTGGAGGCATACCCGTAGAACACCCATATATC
ATCATTGGCCAAATCGCCTCGGTGCTATACTTCGCACTCTTCCTTGTTCTCGCCCCCCTCGCAGGGTGA
GCCGAGAATAAAGCATTGAAATGAGCT----------------
>DBC2
ATGGCAAGCCTACGAAAAACCCATCCGCTAATAAAAATTGCTAACAATGCACTAGTTGACCTTCCAACC
CCATCTAATATTTCAGCACTATGGAACTTCGGGTCCCTCCTAGGATTATGCTTAATTACTCAAATCCTAA
CAGGACTATTCCTGGCCATACATTATACTTCTGATATCTCAACTGCCTTTTCATCAGTAACACATATCTGC
CGAGACGTAAACTATGGCTGACTCATCCGAAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGTA
TTTACATGCACATTGCCCGCGGACTGTACTACGGGTCATACCTCTATAAAGAGACCTGAAGTATTGGCG
TAGTTCTACTCCTTTTAGTAATGATGACAGCCTTTGTGGGCTATGTACTGCCCTGAGGACAAATATCTTT
TTGAGGTGCCACCGTGATTACAAATCTACTATCAGCAGTCCCCTATATGGGCGACGCCTTTGTCCAATG
GATTTGAGGCGGCTTTTCCGTAGATAACGCAACATTAACACGATTCTTCGCCTTCCACTTCCTATTCCCG
TTCGTTATCGCCGGTGCAACGGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGCC
GGATTAAACTCCGACGCAGATAAAATCTCTTTCCACCCATACTTCTCATACAAGGACCTCCTTGGCTTTG
TACTAATGTTATTAGCTCTCACATCCCTAACATTATTCTCCCCAACACTGCTCGGCGACCCAGAAAATTT
CACCCCGGCAAACCCCCTGGTCACCCCACCGCACATCCAACCGGAGTGATACTTCCTGTTTGCCTACG
CCATTCTACGATCTATTCCAAACAAGCTAGGAGGGGTCCTAGCGCTATTGTTTAGCATCCTAGTACTATT
AGTAGTGCCAGTTTTACACACCTCGAAGCAACGAGGACTAACTTTCCGTCCGATCACCCAGTTCTTATT
CTGAACCTTAGTCGCAGACATAATCATTTTAACATGAATTGGAGGCATACCCGTAGAACACCCATATATC
ATCATTGGCCAAATCGCCTCGGTGCTATACTTCGCACTCTTCCTTGTTCTCGCCCCCCTCGCAGGGTGA
GCCGAGAATAAAGCATTGAAATGAGCT----------------
>PNC1
ATGGCAAGCCTACGAAAAACTCACCCACTAATAAAAATCGCTAACGACGCACTAGTTGACCTTCCAACC
CCATCTAATATTTCAGCACTATGAAACTTCGGATCCCTCCTAGGATTATGCTTAATTACTCAAATCTTAAC
AGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTTTTCATCAGTAACACACATCTGC
CGAGACGTTAACTATGGCTGACTTATCCGAAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGTA
TTTATATACACATTGCCCGCGGGCTCTACTATGGATCATACCTCTATAAAGAAACCTGAAACATTGGCGT
AGTTCTACTTCTTTTAGTTATGATGACAGCCTTTGTAGGCTATGTATTACCATGAGGACAAATATCTTTTT
GGGGTGCTACCGTGATTACAAATCTACTATCAGCAGTCCCTTATATAGGAGACGCCCTTGTCCAATGGA
TTTGAGGTGGCTTTTCCGTAGATAACGCAACGTTAACACGATTCTTCGCCTTCCACTTCTTATTTCCGTT
CGTTATCGCCGGTGCAACGATTCTGCATTTGCTATTTCTACACGAAACGGGATCAAACAACCCTGCCGG
ATTAAACTCCGACGCGGATAAAATCTCTTTCCACCCATACTTCTCATACAAGGACCTCCTAGGCTTTGTG
```

```
CTAATATTATTAGCCCTCACATCCCTAACATTATTCTCACCAACACTGCTCGGCGACCCAGAAAACTTCA
CCCCGGCAAACCCCCTGGTTACCCCACCACACATCCAACCGGAGTGGTACTTCTTGTTTGCCTACGCC
ATTCTACGGTCTATTCCAAACAAGCTAGGAGGAGTCCTAGCGTTGTTATTTAGCATTTTAGTACTTTTAG
TGGTGCCAATTTTACATACCTCGAAGCAGCGAGGACTAACTTTCCGTCCAATCACCCAGTTCTTATTCTG
AACCTTGGTGGCAGACATAATCATTTTGACATGAATCGGGGGCATACCCGTAGAACACCCATATATCAT
CATTGGCCAAATCGCGTCGGTGCTGTACTTCGCACTCTTCCTTGTTCTCGCCCCGCTCGCAGGATGAG
CCGAGAATAAAGCATTGAAATGAGCT----------------
>PNC2
ATGGCAAGCCTACGAAAAACTCACCCACTAATAAAAATCGCTAACGACGCACTAGTTGACCTTCCAACC
CCATCTAATATTTCAGCACTATGAAACTTCGGATCCCTTCTAGGATTATGCTTAATTACTCAAATCTTAAC
AGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTTTTCATCAGTAACACACACCTGC
CGAGACGTTAACTATGGCTGACTTATCCGAAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGTA
TTTACATACACATTGCCCGCGGGCTCTACTATGGATCATACCTCTATAAAGAAACCTGAAACATTGGCGT
AGTTCTACTTCTTTTAGTTATGATGACAGCCTTTGTAGGCTATGTATTACCATGAGGACAAATATCTTTTT
GGGGTGCTACCGTGATTACAAATCTACTATCAGCAGTCCCTTATATAGGCGACGCCCTTGTCCAGTGGA
TTTGAGGTGGCTTTTCCGTAGATAACGCAACGTTAACACGATTCTTCGCCTTCCACTTCTTATTTCCGTT
CGTTATCGCCGGTGCAACAATTCTACATTTGCTATTTCTACACGAAACGGGATCAAACAACCCTGCCGG
ATTAAACTCCGACGCGGATAAAATCTCTTTCCACCCGTACTTCTCATACAAGGACCTCCTAGGCTTTGTG
CTAATATTATTAGCCCTCACATCCCTAACATTATTCTCACCAACACTGCTCGGCGACCCAGAAAACTTCA
CCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGTGGTACTTCTTGTTTGCCTACGCC
ATTCTACGGTCTATTCCAAACAAGCTAGGAGGAGTCCTAGCATTGTTATTTAGCATTTTAGTACTTTTAGT
AGTGCCAATTTTACATACCTCGAAGCAGCGAGGACTAACTTTCCGTCCAATCACCCAGTTCTTATTCTGA
ACCTTGGTGGCAGACATAATCATTTTGACATGAATCGGGGGCATACCCGTAGAACACCCATATATCATC
ATTGGCCAAATCGCGTCGGTGCTGTACTTCGCACTCTTCCTTGTTCTCGCCCCGCTCGCAGGATGAGC
CGAGAATAAAGCATTGAAATGAGCT----------------
>PNC3
ATGGCAAGCCTACGAAAAACTCACCCACTAATAAAAATCGCTAACGACGCACTAGTTGACCTTCCAACC
CCATCTAATATTTCAGCACTATGAAACTTCGGATCCCTTCTAGGATTATGCTTAATTACTCAAATCTTAAC
AGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTTTTCATCAGTAACACACACCTGC
CGAGACGTTAACTATGGCTGACTTATCCGAAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGTA
TTTACATACACATTGCCCGCGGGCTCTACTATGGATCATACCTCTATAAAGAAACCTGAAACATTGGCGT
AGTTCTACTTCTTTTAGTTATGATGACAGCCTTTGTAGGCTATGTATTACCATGAGGACAAATATCTTTTT
GGGGTGCTACCGTGATTACAAATCTACTATCAGCAGTCCCTTATATAGGCGACGCCCTTGTCCAGTGGA
TTTGAGGTGGCTTTTCCGTAGATAACGCAACGTTAACACGATTCTTCGCCTTCCACTTCTTATTTCCGTT
CGTTATCGCCGGTGCAACAATTCTACATTTGCTATTTCTACACGAAACGGGATCAAACAACCCTGCCGG
ATTAAACTCCGACGCGGATAAAATCTCTTTCCACCCGTACTTCTCATACAAGGACCTCCTAGGCTTTGTG
CTAATATTATTAGCCCTCACATCCCTAACATTATTCTCACCAACACTGCTCGGCGACCCAGAAAACTTCA
CCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGTGGTACTTCTTGTTTGCCTACGCC
ATTCTACGGTCTATTCCAAACAAGCTAGGAGGAGTCCTAGCATTGTTATTTAGCATTTTAGTACTTTTAGT
AGTGCCAATTTTACATACCTCGAAGCAGCGAGGACTAACTTTCCGTCCAATCACCCAGTTCTTATTCTGA
ACCTTGGTGGCAGACATAATCATTTTGACATGAATCGGGGGCATACCCGTAGAACACCCATATATCATC
ATTGGCCAAATCGCGTCGGTGCTGTACTTCGCACTCTTCCTTGTTCTCGCCCCGCTCGCAGGATGAGC
CGAGAATAAAGCATTGAAATGAGCT----------------
>MM1
ATGGCAAGCCTACGAAAAACTCACCCACTAATAAAAATCGCTAACGACGCACTAGTTGACCTTCCAACC
CCATCTAATATTTCAGCACTATGAAACTTCGGATCCCTTCTAGGATTATGCTTAATTACTCAAATCTTAAC
AGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTTTTCATCAGTAACACACATCTGC
CGAGACGTTAACTATGGCTGACTTATCCGAAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGTA
TTTACATACACATTGCCCGCGGGCTCTACTATGGATCATACCTCTATAAAGAAACCTGAAACATTGGCGT
AGTTCTACTTCTTTTAGTTATGATGACAGCCTTTGTAGGCTATGTATTACCATGAGGACAAATATCTTTTT
GGGGTGCTACCGTGATTACAAATCTACTATCAGCAGTCCCTTATATAGGCGACGCCTTTGTCCAGTGGA
TTTGAGGTGGCTTTTCCGTAGATAACGCAACGTTAACACGATTCTTCGCCTTCCACTTCTTATTTCCGTT
CGTTATCGCCGGTGCAACAATTCTGCATTTGCTATTTCTACACGAAACGGGATCAAACAACCCTGCCGG
ATTAAACTCCGACGCGGATAAAATCTCTTTCCACCCGTACTTCTCATACAAGGACCTCCTAGGCTTTGTG
CTAATATTATTAGCCCTCACATCCCTAACATTATTCTCACCAACACTGCTCGGCGACCCAGAAAACTTCA
CCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGTGGTACTTCTTGTTTGCCTACGCC
ATTCTACGGTCTATTCCAAACAAGCTAGGAGGAGTCCTAGCATTGTTATTTAGCATTTTAGTACTTTTAGT
AGTGCCAATTTTACATACCTCGAAGCAGCGAGGACTAACTTTCCGTCCAATCACCCAGTTCTTATTCTGA
ACCTTGGTGGCAGACATAATCATTTTGACATGAATCGGGGGCATACCCGTAGAACACCCATATATCATC
```

```
ATTGGCCAAATCGCGTCGGTGCTGTACTTCGCACTCTTCCTTGTTCTCGCCCCGCTCGCAGGATGAGC
CGAGAATAAAGCATTGAAATGAGCT----------------
>MM2
ATGGCAAGCCTACGAAAAACTCACCCACTAATAAAAATCGCTAACGACGCACTAGTTGACCTTCCAACC
CCATCTAATATTTCAGCACTATGAAACTTCGGATCCCTCCTAGGATTATGCTTAATTACTCAAATCTTAAC
AGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTTTTCATCAGTAACACACATCTGC
CGAGACGTTAACTATGGCTGACTTATCCGAAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGTA
TTTATATACACATTGCCCGCGGGCTCTACTATGGATCATACCTCTATAAAGAAACCTGAAACATTGGCGT
AGTTCTACTTCTTTTAGTTATGATGACAGCCTTTGTAGGCTATGTATTACCATGAGGACAAATATCTTTTT
GGGGCGCTACCGTGATTACAAATCTACTATCAGCAGTCCCTTATATAGGCGACGCCTTTGTCCAATGGA
TTTGAGGTGGCTTTTCCGTAGATAACGCAACGTTAACACGATTCTTCGCCTTCCACTTCTTGTTTCCGTT
CGTTATCGCCGGTGCAACAATTCTGCATTTGCTATTTCTACACGAAACGGGATCAAACAACCCTGCCGG
ATTAAACTCCGACGCGGATAAAATCTCTTTCCACCCGTACTTCTCATACAAGGACCTCCTAGGCTTTGTG
CTAATATTATTAGCCCTCACATCCCTAACGTTATTCTCACCAACACTGCTCGGCGACCCAGAAAACTTCA
CCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGTGGTACTTCTTGTTTGCCTACGCC
ATTCTACGGTCTATTCCAAACAAGCTAGGAGGAGTCCTAGCATTGTTATTTAGCATTTTAGTACTTTTAGT
AGTGCCAATTTTACATACCTCGAAGCAGCGAGGACTAACTTTCCGTCCAATCACCCAGTTCTTATTCTGA
ACCTTGGTGGCAGACATAATCATTTTGACATGAATCGGGGGCATACCCGTAGAACACCCATATATCATC
ATTGGCCAAATCGCGTCGGTGCTGTACTTCGCACTCTTCCTTGTTCTCGCCCCGCTCGCAGGATGAGC
CGAGAATAAAGCATTGAAATGAGCT----------------
>MM3
ATGGCAAGCCTACGAAAAACTCACCCACTAATAAAAATCGCTAACGACGCACTAGTTGACCTTCCAACC
CCATCTAATATTTCAGCACTATGAAACTTCGGATCCCTCCTAGGATTATGCTTAATTACTCAAATCTTAAC
AGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTTTTCATCAGTAACACACATCTGC
CGAGACGTTAACTATGGCTGACTTATCCGAAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGTA
TTTATATACACATTGCCCGCGGGCTCTACTATGGATCATACCTCTATAAAGAAACCTGAAACATTGGCGT
AGTTCTACTTCTTTTAGTTATGATGACAGCCTTTGTAGGCTATGTATTACCATGAGGACAAATATCTTTTT
GGGGCGCTACCGTGATTACAAATCTACTATCAGCAGTCCCTTATATAGGCGACGCCTTTGTCCAATGGA
TTTGAGGTGGCTTTTCCGTAGATAACGCAACGTTAACACGATTCTTCGCCTTCCACTTCTTGTTTCCGTT
CGTTATCGCCGGTGCAACAATTCTGCATTTGCTATTTCTACACGAAACGGGATCAAACAACCCTGCCGG
ATTAAACTCCGACGCGGATAAAATCTCTTTCCACCCATACTTCTCATACAAGGACCTCCTAGGCTTTGTG
CTAATATTATTAGCCCTCACATCCCTAACGTTATTCTCACCAACACTGCTCGGCGACCCAGAAAACTTCA
CCCCGGCAAACCCCCTGGTTACCCCACCGCACATCCAACCGGAGTGGTACTTCTTGTTTGCCTACGCC
ATTCTACGGTCTATTCCAAACAAGCTAGGAGGAGTCCTAGCATTGTTATTTAGCATTTTAGTACTTTTAGT
AGTGCCAATTTTACATACCTCGAAGCAGCGAGGACTAACTTTCCGTCCAATCACCCAGTTCTTATTCTGA
ACCTTGGTGGCAGACATAATCATTTTGACATGAATCGGGGGCATACCCGTAGAACACCCATATATCATC
ATTGGCCAAATCGCGTCGGTGCTGTACTTCGCACTCTTCCTTGTTCTCGCCCCGCTCGCAGGATGAGC
CGAGAATAAAGCATTGAAATGAGCT----------------
>AZT1
ATGGCAAGCCTACGAAAAACCCACCCACTAATAAAAATTGCTAACGATGCACTAGTCGACCTTCCAACT
CCATCTAATATCTCAGCACTATGGAACTTCGGATCCCTCCTAGGATTATGCTTGATTACGCAAATCCTAA
CAGGACTGTTCCTGGCCATGCATTATACTTCCGATATCTCAACTGCCTTTTCGTCAGTAACACACATCTG
CCGAGACGTTAACTATGGCTGACTCATCCGAAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGT
ATTTATATACACATTGCCCGCGGACTGTACTATGGGTCTTACCTCTATAAAGAGACCTGAACCATTGGC
GTGGTCCTACTCCTTTTGGTCATAATAACGGCCTTCGTAGGCTATGTGCTACCGTGGGGTCAAATATCT
TTTTGGGGCGCCACCGTCATTACAAACCTACTATCAGCAGTCCCTTATATGGGAGACGCCCTTGTCCAG
TGGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTC
CGTTCGTTATCGCCGGTGCGACGGTCCTGCATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTG
CCGGATTAAACTCCGACGCGGATAAAATCTCTTTCCACCCATACTTCTCGTACAAAGATCTCCTTGGCTT
TGTACTAATGTTATTAGCTCTCACATCCCTGACATTATTCTCCCCAACTCTGCTCGGCGACCCAGAAAAT
TTCACCCCAGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGTGATACTTCTTATTTGCCTAC
GCCATTCTACGCTCTATTCCAAACAAGCTGGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTACTAT
TAGTGGTGCCAATTTTACATACCTCAAAGCAGCGAGGACTAACTTTCCGTCCAATCACCCAGTTCTTATT
CTGAACCTTGGTGGCAGACATAATCATTTTGACATGAATTGGAGGCATGCCCGTAGAACACCCATACAT
CATCATCGGCCAAATCGCGTCGGTGCTGTACTTCGCACTCTTCCTTCTTCTCGCCCCCCTCGCAGGGT
GAGCCGAGAATAAAGCATTGAAATGAGCT----------------
>BP11
ATGGCAAGCCTACGAAAAACCCACCCACTAATAAAAATTGCTAACGATGCACTAGTCGACCTTCCAACT
CCATCTAATATCTCAGCACTATGGAACTTCGGATCCCTCCTAGGATTATGCTTGATTACGCAAATCCTAA
```

```
CAGGACTGTTCCTGGCCATGCATTATACTTCCGATATCTCAACTGCCTTTTCGTCAGTAACACACATCTG
CCGAGACGTTAACTATGGCTGACTCATCCGAAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGT
ATTTATATACACATTGCCCGCGGACTGTACTATGGGTCTTACCTCTATAAAGAGACCTGAACCATTGGC
GTGGTCCTACTCCTTTTGGTCATAATAACGGCCTTCGTAGGCTATGTGCTACCGTGGGGTCAAATATCT
TTTTGGGGCGCCACCGTCATTACAAACCTACTATCAGCAGTCCCTTATATGGGAGACGCCCTTGTCCAG
TGGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTC
CGTTCGTTATCGCCGGTGCGACGGTCCTGCACTTGCTATTTTTACACGAAACGGGATCAAACAACCCTG
CCGGATTAAACTCCGACGCGGATAAAATCTCTTTCCACCCATACTTCTCGTACAAAGATCTCCTTGGCTT
TGTACTAATGTTATTAGCTCTCACATCCCTGACATTATTCTCCCCAACTCTGCTCGGCGACCCAGAAAAT
TTCACCCCAGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGTGATACTTCTTATTTGCCTAC
GCCATTCTACGCTCTATTCCAAACAAGCTGGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTACTAT
TAGTGGTGCCAATTTTACATACCTCAAAGCAGCGAGGACTAACTTTCCGTCCAATCACCCAGTTCTTATT
CTGAACCTTGGTGGCAGACATAATCATTTTGACATGAATTGGAGGCATGCCCGTAGAACACCCATACAT
CATCATCGGCCAAATCGCGTCGGTGCTGTACTTCGCACTCTTCCTTCTTCTCGCCCCCCTCGCAGGGT
GAGCCGAGAATAAAGCATTGAAATGAGCT----------------
>BONSV1
ATGGCAAGCCTACGAAAAACCCACCCACTAATAAAAATTGCTAACGATGCACTAGTCGACCTTCCAACT
CCATCTAATATCTCAGCACTATGGAACTTCGGATCCCTCCTAGGATTATGCTTGATTACGCAAATCCTAA
CAGGACTGTTCCTAGCCATGCATTATACTTCCGATATCTCAACTGCCTTTTCGTCAGTAACACACATCTG
CCGAGACGTTAACTATGGCTGACTCATCCGAAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGT
ATTTATATACACATTGCCCGCGGACTGTACTATGGGTCTTACCTCTATAAAGAGACCTGAACCATTGGTG
TGGTCCTACTCCTTTTGGTCATAATAACGGCCTTCGTAGGCTATGTGCTACCGTGGGGTCAAATATCTTT
TTGGGGCGCCACCGTCATTACAAACCTATTATCAGCAGTCCCTTATATGGGAGACGCCCTTGTCCAGTG
GATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCG
TTCGTTATCGCCGGTGCGACGGTCCTGCATTTGCTATTTTTACACGAAACCGGATCAAACAACCCTGCC
GGATTAAACTCCGACGCGGATAAAATCTCTTTCCACCCATACTTCTCGTACAAAGATCTCCTTGGCTTTG
TACTAATGTTATTAGCTCTCACATCCCTGACATTATTCTCCCCAACTCTGCTTGGCGACCCAGAAAATTT
CACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGTGATACTTCTTATTTGCCTACG
CCATTCTACGTTCTATTCCAAACAAGCTGGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTACTATT
AGTCGTGCCAATTTTACATACCTCAAAGCAGCGAGGACTAACTTTCCGTCCAATCACCCAGTTCTTATTC
TGGACCTTGGTGGCAGACATAATCATTTTGACATGAATTGGAGGCATGCCCGTAGAACACCCATACATC
ATCATCGGCCAAATCGCGTCGGTGCTGTACTTCGCACTCTTCCTTCTTCTCGCCCCCCTCGCAGGGTG
AGCCGAGAATAAAGCATTGAAATGAGCT----------------
>BONPK1
ATGGCAAGCCTACGAAAAACCCACCCACTAATAAAAATTGCTAACGATGCACTAGTCGACCTTCCAACC
CCATCTAATATCTCAGCACTATGGAACTTCGGATCCCTCCTAGGATTATGCTTGATTACGCAAATCCTAA
CAGGACTGTTCCTGGCCATGCATTATACTTCCGATATCTCAACTGCCTTTTCGTCAGTAACACACATCTG
CCGAGACGTTAACTATGGCTGACTCATCCGAAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGT
ATTTATATACACATTGCCCGCGGACTGTACTATGGGTCTTACCTCTATAAAGAGACCTGAACCATTGGTG
TGGTCCTACTCCTTTTGGTCATAATAACGGCCTTCGTAGGCTATGTGCTACCGTGGGGCCAAATATCTT
TTTGGGGCGCCACCGTCATTACAAACCTACTATCAGCAGTCCCTTATATGGGAGACGCCCTTGTCCAGT
GGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCC
GTTCGTTATCGCCGGTGCGACAGTCCTGCATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGATTAAACTCCGACGCGGATAAAATCTCTTTCCACCCATACTTCTCGTACAAAGATCTCCTTGGCTTT
GTACTAATGTTATTAGCTCTTACATCCCTGACATTATTCTCCCCAACTCTGCTCGGCGACCCAGAAAATT
TCACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGTGATACTTCTTATTTGCCTACG
CCATTCTACGTTCTATTCCAAACAAGCTGGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTACTATT
AGTCGTGCCAATTTTACATACCTCAAAGCAGCGAGGACTAACTTTCCGTCCAATCACCCAGTTCTTATTC
TGGACCTTGGTGGCAGACATAATCATTTTGACATGAATTGGAGGCATGCCCGTAGAACACCCATACATC
ATCATCGGCCAAATCGCGTCGGTGCTGTACTTCGCACTCTTCCTTCTTCTCGCCCCCCTCGCAGGGTG
AGCCGAGAATAAAGCATTGAAATGAGCT----------------
>EAG1
ATGGCAAGCCTACGAAAAACCCACCCACTAATAAAAATTGCTAACGATGCACTAGTTGACCTTCCGACC
CCATCTAATATCTCAGCACTATGGAACTTCGGATCCCTCCTAGGATTATGCTTGATTACTCAAATCCTAA
CAGGACTATTCCTCGCCATGCATTATACTTCCGATATCTCAACTGCCTTTTCGTCGGTAACACACATCTG
CCGAGACGTTAACTATGGTTGACTCATCCGAAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGT
ATTTATATACACATTGCCCGCGGACTGTACTATGGGTCATACCTCTACAAAGCGACCTGAACTATTGGC
GTGGTCCTACTCCTTTTGGTCATAATAACTGCCTTTGTAGGCTATGTGCTGCCCTGAGGCCAAATATCTT
TCTGGGGTGCCACCGTCATTACAAATCTACTATCAGCAGTCCCTTATATAGGAGACGCCCTTGTTCAGT
```

```
GGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCC
GTTCGTTATCGCCGGTGCGACGGTCCTGCATTTGCTATTTCTACACGAAACGGGGTCAAACAACCCTGC
CGGATTAAACTCCGACGCGGATAAAATCTCTTTCCACCCATACTTCTCATACAAAGATCTCCTTGGCTTT
GTACTAATGTTATTAGCTCTTACATCCCTGACATTATTCTCCCCAACTCTGCTCGGTGACCCAGAAAATT
TCACCCCAGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGTGATACTTCTTATTTGCCTACG
CCATTCTGCGCTCTATTCCAAACAAGCTAGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTACTGTT
AGTCGTGCCAATTTTACATACCTCAAAGCAGCGAGGACTAACTTTCCGCCCAATCACCCAGTTCCTATT
CTGAACCTTAGTGGCAGACATAATCATTTTAACATGAATTGGAGGCATACCCGTAGAACACCCATATATC
ATCATTGGCCAAATCGCGTCGGTGCTGTACTTCGCACTCTTCCTTCTTCTCGCCCCCCTCGCAGGGTGA
GCCGAGAATAAAGCATTGAAATGAGCT----------------
>LSR1
ATGGCAAGCCTACGAAAAACCCACCCACTAATAAAAATTGCTAACGATGCACTAGTCGACCTTCCAACT
CCATCTAATATCTCAGCACTATGGAACTTCGGATCCCTCCTAGGATTATGCTTGATTACGCAAATCCTAA
CAGGACTGTTCCTGGCCATGCATTATACTTCCGATATCTCAACTGCCTTTTCGTCAGTAACACACATCTG
CCGAGACGTTAACTATGGCTGACTCATCCGAAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGT
ATTTATATACACATTGCCCGCGGACTGTACTATGGGTCTTACCTCTATAAAGAGACCTGAACCATTGGC
GTGGTCCTACTCCTTTTGGTCATAATAACGGCCTTCGTAGGCTATGTGCTACCGTGGGGTCAAATATCT
TTTTGGGGCGCCACCGTCATTACAAACCTACTATCAGCAGTCCCTTATATGGGAGACGCCCTTGTCCAG
TGGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTC
CGTTCGTTATCGCCGGTGCGACGGTCCTGCATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTG
CCGGATTAAACTCCGACGCGGATAAAATCTCTTTCCACCCATACTTCTCGTACAAAGATCTCCTTGGCTT
TGTACTAATGTTATTAGCTCTCACATCCCTGACATTATTCTCCCCAACTCTGCTCGGCGACCCAGAAAAT
TTCACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGTGATACTTCTTATTTGCCTAC
GCCATTCTACGCTCTATTCCAAACAAGCTGGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTACTAT
TAGTGGTGCCAATTTTACATACCTCAAAGCAGCGAGGACTAACTTTCCGTCCAATCACCCAGTTCTTATT
CTGAACCTTGGTGGCAGATATAATCATTTTGACATGAATTGGAGGCATGCCCGTAGAACACCCATACAT
CATCATCGGCCAAATCGCGTCGGTGCTGTACTTCGCACTCTTCCTTCTTCTCGCCCCCCTCGCAGGGT
GAGCCGAGAATAAAGCATTGAAATGAGCT----------------
>RAV1
ATGGCAAGCCTACGAAAAACCCACCCACTAATAAAAATTGCTAACGATGCACTAGTTGACCTTCCGACC
CCATCTAATATCTCAGCACTATGGAACTTCGGATCCCTCCTAGGATTATGCTTGATTACTCAAATCCTAA
CAGGACTATTCCTGGCCATGCATTATACTTCCGATATCTCAACTGCCTTTTCGTCGGTAACACACATCTG
CCGAGACGTTAACTATGGTTGACTCATCCGGAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGT
ATTTATATACACATTGCCCGCGGACTGTACTATGGGTCATACCTCTACAAAGCGACCTGAACTATTGGC
GTGGTCCTACTCCTTTTGGTTATAATAACTGCCTTTGTAGGCTATGTGCTGCCCTGAGGCCAAATATCTT
TCTGGGGTGCCACCGTCATTACAAATCTACTATCAGCAGTCCCTTATATAGGAGACGCCCTTGTTCAGT
GGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCC
GTTCGTTATCGCCGGTGCGACGGTCCTGCATTTGCTATTTCTACACGAAACGGGATCAAACAACCCTGC
CGGATTAAACTCCGACGCGGATAAAATCTCTTTCCACCCATACTTCTCATACAAAGATCTCCTTGGCTTT
GTACTAATGTTATTAGCTCTTACATCCCTGACATTATTCTCCCCAACTCTGCTCGGTGACCCAGAAAATT
TCACCCCAGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGTGATACTTCTTATTTGCCTACG
CCATTCTGCGCTCTATTCCAAACAAGCTGGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTACTGTT
AGTCGTGCCAATTTTACATACCTCAAAGCAGCGAGGACTAACTTTCCGCCCAATCACCCAGTTCCTATT
CTGAACCTTAGTGGCAGACATAATCATTTTAACATGAATTGGAGGCATACCCGTAGAACACCCATATATC
ATCATTGGCCAAATCGCATCGGTGCTGTACTTCGCACTCTTCCTTCTTCTCGCCCCCCTCGCAGGGTGA
GCCGAGAATAAAGCATTGAAATGAGCT----------------
>SON1
ATGGCAAGCCTACGAAAAACCCACCCACTAATAAAAATTGCTAACGATGCACTAGTCGACCTTCCGACC
CCATCCAATATCTCAGCACTATGGAACTTCGGATCCCTCCTAGGATTATGCTTGATTACTCAAATCCTAA
CAGGACTATTCCTGGCCATGCATTATACTTCCGATATCTCAACTGCCTTTTCGTCGGTAACACACATCTG
CCGAGACGTTAACTATGGCTGGCTCATCCGAAACATACATGCTAACGGAGCATCATTCTTCTTCATCTG
TATTTATATACACATTGCCCGCGGACTGTACTATGGGTCATACCTCTATAAAGCGACCTGAACTATTGGC
GTGGTCCTACTCCTTTTGGTCATAGTAACGGCCTTTGTAGGCTATGTGCTGCCCTGAGGCCAAATATCT
TTCTGGGGCGCCACCGTCATTACAAATCTACTATCAGCAGTCCCTTATATGGGAGACGCCCTTGTCCAG
TGGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTATTTC
CGTTCGTTATCGCCGGTGCGACGGTCCTGCATTTGCTATTTTTACACGAAACGGGATCAAGCAACCCTG
CCGGATTAAACTCCGACGCGGATAAAATCTCTTTCCACCCATACTTCTCATACAAAGATCTCCTTGGCTT
TGTACTAATGTTATTAGCTCTTACATCCCTGACATTATTCTCCCCAACTCTGCTCGGTGACCCAGAAAAT
TTTACCCCAGCAAACCCCCTGGTTACCCCGCCACATATCCAACCGGAGTGATACTTCTTATTTGCCTAC
```

```
GCCATTCTACGCTCTATTCCAAACAAGCTGGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTACTGT
TAGTCGTGCCAATTTTACATACCTCAAAGCAGCGAGGACTAACTTTCCGGCCAATCACCCAGTTCTTATT
CTGAACCTTAGTGGCAGACATAATCATTTTAACATGAATTGGAGGCATACCCGTAGAACACCCATACAT
CATCATCGGCCAAATCGCGTCGGTGCTGTACTTCGCACTCTTCCTTCTTCTCGCCCCCCTCGCAGGGT
GAGCCGAGAATAAAGCATTGAAATGAGCT----------------
>TC2
ATGGCAAGCCTACGAAAAACCCACCCGCTAATAAAAATTGCTAACGATGCACTAGTCGACCTTCCAACT
CCATCTAATATCTCAGCACTATGGAACTTCGGGTCCCTCCTAGGATTATGCTTGATTACGCAAATCCTAA
CAGGGCTGTTCCTAGCCATGCATTATACTTCCGATATCTCAACTGCCTTTTCGTCGGTGACACACATCT
GCCGAGACGTTAACTATGGCTGACTCATCCGAAACATACATGCTAACGGAGCATCATTCTTCTTTATCT
GTATTTATATACACATTGCCCGCGGACTGTACTATGGGTCTTACCTCTATAAAGAGACCTGAACCATTGG
TGTAGTCCTACTCCTTTTGGTCATAATAACAGCCTTCGTAGGCTATGTGCTACCCTGGGGCCAAATATCT
TTTTGGGGCGCCACCGTCATTACAAACCTACTATCAGCAGTCCCTTATATAGGAGACGCCCTTGTCCAG
TGGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTC
CGTTCGTTATCGCCGGTGCGACGGTCCTGCATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTG
CCGGATTAAACTCCGACGCGGATAAAATCTCTTTCCACCCATACTTCTCGTACAAAGATCTCCTTGGCTT
TGTACTAATGTTATTAGCTCTCACATCCCTGACATTATTCTCCCCAACTCTGCTCGGCGACCCAGAAAAT
TTCACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGTGATACTTCTTATTTGCCTAC
GCCATTCTACGCTCTATTCCAAACAAGCTGGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTACTAT
TAATTGTGCCAATTTTACACACCTCAAAGCAGCGAGGACTAACTTTCCGTCCAATCACCCAGTTCTTATT
CTGAACCTTGGTGGCAGACATAATCATTTTGACATGAATTGGAGGCATGCCCGTAGAACACCCATACAT
CATCATCGGCCAAATCGCGTCGGTGCTGTACTTCGCGCTCTTCCTTCTTCTCGCCCCCCTCGCAGGGT
GAGCCGAGAATAAAGCATTGAAATGAGCT----------------
>WC1
ATGGCAAGCCTACGAAAAACCCACCCACTAATAAAAATTGCTAACGATGCACTAGTCGACCTTCCGACC
CCATCTAATATCTCAGCACTATGGAACTTCGGATCCCTCCTAGGATTATGCTTGATTACTCAAATCCTAA
CAGGACTATTCCTGGCCATGCATTATACTTCCGATATCTCAACTGCCTTTTCGTCGGTAACACACATCTG
CCGAGACGTTAATTATGGCTGGCTCATCCGAAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGT
ATTTATATACACATTGCCCGCGGACTGTACTATGGATCATACCTCTATAAAGCGACCTGAACTATTGGCG
TGGTCCTACTCCTTTTGGTCATAATAACGGCCTTTGTAGGCTATGTGCTGCCCTGAGGCCAAATATCTTT
CTGGGGCGCCACCGTCATTACAAATTTACTATCAGCAGTCCCTTATATGGGAGACGCCCTTGTCCAGTG
GATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTATTTCCG
TTCGTTATCGCCGGTGCGACGGTCCTGCATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGCC
GGATTAAACTCCGACGCAGATAAAATCTCTTTCCACCCATACTTCTCATACAAAGATCTCCTCGGCTTTG
TACTAATATTATTAGCTCTTACATCCCTGACATTATTCTCCCCAACTCTGCTCGGTGACCCAGAAAATTTC
ACCCCAGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGTGATACTTCTTATTTGCCTACGCC
ATTCTACGCTCTATTCCAAACAAGCTGGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTACTGTTAG
TCGTGCCAATTTTACATACCTCAAAGCAGCGAGGACTAACTTTCCGGCCAATCACCCAGTTCTTATTCTG
AACCTTAGTGGCAGACATAATCATTTTAACATGAATTGGAGGCATACCCGTAGAACACCCATACATCATC
ATTGGCCAAATCGCATCGGTGCTGTACTTCGCACTCTTCCTTCTTCTCGCCCCCCTCGCAGGGTGAGC
CGAGAATAAAGCATTGAAATGAGCT----------------
>WR2
ATGGCAAGCCTACGAAAAACCCACCCACTAATAAAAATTGCTAACGATGCACTAGTCGACCTTCCAACT
CCATCTAATATCTCAGCGCTATGGAACTTCGGATCTCTCCTAGGATTATGCTTGATTACGCAAATCCTAA
CAGGACTGTTCCTGGCCATGCATTATACTTCCGATATCTCAACTGCCTTTTCGTCAGTAACACACATCTG
CCGAGACGTTAACTATGGCTGACTCATCCGAAACATACATGCTAACGGAGCATCATTCTTCTTTATCTGT
ATTTATATACACATTGCCCGCGGACTGTACTATGGGTCTTACCTCTATAAAGAGACCTGAACCATTGGC
GTGGTCCTACTCCTTTTGGTCATAATAACGGCCTTCGTAGGCTATGTGCTACCGTGGGGTCAAATATCT
TTTTGGGGCGCCACCGTCATTACAAACCTACTATCAGCAGTCCCTTATATGGGAGACGCCCTTGTCCAG
TGGATTTGAGGTGGCTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTC
CGTTCGTTATCGCCGGTGCGACGGTCCTGCATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTG
CCGGATTAAACTCCGACGCGGATAAAATCTCTTTCCACCCATACTTCTCGTACAAAGATCTCCTTGGCTT
TGTACTAATGTTATTAGCTCTCACATCCCTGACATTATTCTCCCCAACTCTGCTCGGCGACCCAGAAAAT
TTCACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGTGATACTTCTTATTTGCCTAC
GCCATTCTACGCTCTATTCCAAACAAGCTGGGAGGGGTCCTAGCACTATTATTTAGCATTCTAGTACTAT
TAGTGGTGCCAATTTTACATACCTCAAAGCAGCGAGGACTAACTTTCCGTCCAATCACCCAGTTCTTATT
CTGAACCTTGGTGGCAGACATAATCATTTTGACATGAATTGGAGGCATGCCCGTAGAACACCCATACAT
CATCATCGGCCAAATCGCGTCGGTGCTGTACTTCGCACTCTTCCTTCTTCTCGCCCCCCTCGCAGGGT
GAGCCGAGAATAAAGCATTGAAATGAGCT----------------
```

Arlequin Input File

```
#================================
#Project file created by Arlequin
#================================

#    (symbol # is used to comment what follows on a line)

[Profile]
    Title="Dace MtDNA Cytb Region Analysis"
    NbSamples=10
    DataType=DNA  # - {DNA, RFLP, MICROSAT, STANDARD, FREQUENCY}
    GenotypicData=0  # - {0, 1}
    GameticPhase=0  # - {0, 1}
    LocusSeparator=NONE  # - {TAB, WHITESPACE, NONE}
    RecessiveData=0  # - {0, 1}
    MissingData='?'  # A single character specifying missing data
#    Some advanced settings the experienced user can uncomment
#    Frequency= ABS            # - {ABS, REL}
#    FrequencyThreshold= 1.0e-5 # - (Any real number, usually between 1.0e-7 and 1.e-3)
#    EpsilonValue= 1.0e-7       # - (Any real number, usually between 1.0e-12 and 1.0e-5)

[Data]

[[HaplotypeDefinition]]
HaplListName="Dace MtDNA Cytb Haplotype List"
HaplList={
H1
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT
TCCAACTCCATCTAATATTTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGAT
TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTT
TTCGTCAGTAACACACATCTGTCGAGACGTTAACTATGGCTGACTCATCCGAAACATACATG
CTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATG
GGTCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCCTACTCCTTTTGGTCATAATAA
CAGCCTTTGTAGGCTATGTACTGCCGTGAGGTCAAATGTCTTTCTGAGGCGCCACCGTCATT
ACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGGATTTGAGGTGG
CTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGTTTGTT
ATCGCCGGCGCAACAGTCCTACATTTGTTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGTTAAACTCCAACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCT
TGGCTTTGTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGG
TGACCCAGAAAATTTCACCCCGGCAAACCCCCTGGTTACCCCACCACACATCCAACCGGAG
TGGTACTTCTTATTTGCTTACGCCATTCTACGCTCTATCCCAAACAAGCTGGGAGGGGTCCT
AGCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACG
AGGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATCAT
TTTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGT
CGGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGAGCCGAAAATAA
AGCATTGAAATGAGCTTGCCCTAGTAGCTTAG
H2
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT
TCCAACTCCATCTAATATTTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGAT
TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTT
TTCGTCAGTAACACACATCTGTCGAGACGTTAACTATGGCTGACTCATCCGAAACATACATG
```

CTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATG
GGTCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCCTACTCCTTTTGGTCATAATAA
CAGCCTTTGTAGGCTATGTACTGCCGTGAGATCAAATGTCTTTCTGAGGCGCCACCGTCATT
ACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGGATTTGAGGTGG
CTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGTTTGTT
ATCGCCGGCGCAACAGTCCTACATTTGTTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGTTAAACTCCAACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCT
TGGCTTTGTACTAATATTGTTAGCTCTCACACCCCTGACGCTATTCTCTCCAACTCTGCTCGG
TGACCCAGAAAATTTCACCCCGGCAAACCCCCTGGTTACCCCACCACACATCCAACCGGAG
TGGTACTTCTTATTTGCTTACGCCATTCTACGCTCTATCCCAAACAAGCTGGGAGGGGTCCT
AGCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACG
AGGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATCAT
TTTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGT
CGGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGAGCCGAAAATAA
AGCATTGAAATGAGCTTGCCCTAGTAGCTTAG
H3
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT
TCCAACTCCATCTAATATTTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGAT
TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTT
TTCGTCAGTAACACACATCTGTCGAGACGTTAACTATGGCTGACTCATCCGAAACATACATG
CTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATG
GGTCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCCTACTCCTTTTGGTCATAATAA
CAGCCTTTGTAGGCTATGTACTGCCGTGAGGTCAAATGTCTTTCTGAGGCGCCACCGTCATT
ACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGGATTTGAGGTGG
CTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGTTTGTT
ATCGCCGGCGCAACAGTCCTACATTTGTTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGTTAAACTCCAACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCT
TGGCTTTGTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGG
TGACCCAGAAAATTTCACCCCGGCAAACCCCCTGGTTACCCCACCACACATCCAACCGGAG
TGGTACTTCTTATTTTCTTACGCCATTCTACGCTCTATCCCAAACAAGCTGGGAGGGGTCCTA
GCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGA
GGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATCATT
TTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTC
GGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGAGCCGAAAATAAA
GCATTGAAATGAGCTTGCCCTAGTAGCTTAG
H4
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT
TCCAACTCCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGAT
TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTT
TTCGTCAGTAACACACATCTGTCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATG
CTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATG
GGTCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCCTACTCCTTTTGGTCATAATAA
CAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTCTGAGGCGCCACCGTCATT
ACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGGATTTGAGGTGG
CTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGTTTGTT
ATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAAACCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCT
TGGCTTTGTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGG
TGACCCAGAAAATTTTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAG
TGGTACTTCTTATTTGCTTACGCCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTA
GCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGA
GGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATTATT
TTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTC

GGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGAGCCGAAAATAAA
GCATTGAAATGAGCTTGCCCTAGTAGCTTAG
H5
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT
TCCAACTCCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGAT
TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTT
TTCGTCAGTAACACACATCTGTCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATG
CTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATG
GGTCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCCTACTCCTTTTGGTCATAATAA
CAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTCTGAGGCGCCACCGTCATT
ACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGGATTTGAGGTGG
CTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGTTTGTT
ATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCT
TGGCTTTGTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGG
TGACCCAGAAAATTTTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAG
TGGTACTTCTTATTTGCTTACGCCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTA
GCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGA
GGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATTATT
TTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTC
GGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGAGCCGAAAATAAA
GCATTGAAATGAGCTTGCCCTAGTAGCTTAG
H6
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT
TCCAACTCCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGAT
TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTT
TTCGTCAGTAACACACATCTGTCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATG
CTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATG
GGGCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCCTACTCCTTTTGGTCATAATAA
CAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTCTGAGGCGCCACCGTCATT
ACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGGATTTGAGGTGG
CTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGTTTGTT
ATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAAACCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCT
TGGCTTTGTATTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGG
TGACCCAGAAAATTTTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAG
TGGTACTTCTTATTTGCTTACGCCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTA
GCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGA
GGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATTATT
TTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTC
GGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGAGCCGAAAATAAA
GCATTGAAATGAGCTTGCCCTAGTAGCTTAG
H7
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT
TCCAACTCCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGAT
TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTT
TTCGTCAGTAACACACATCTGTCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATG
CTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATG
GGTCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCCTACTCCTTTTGGTCATAATAA
CAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTCTGAGGCGCCACCGTCATT
ACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGGATTTGAGGTGG
CTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGTTTGTT
ATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC

CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCT
TGGCTTTGTATTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGG
TGACCCAGAAAATTTTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAG
TGGTACTTCTTATTTGCTTACGCCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTA
GCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGA
GGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATTATT
TTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTC
GGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGAGCCGAAAATAAA
GCATTGAAATGAGCTTGCCCTAGTAGCTTAG
H8
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT
TCCAACTCCATCTAATATTTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCCTGAT
TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATTTCAACTGCCTTT
TCGTCAGTAACACACATCTGTCGAGACGTTAACTATGGCTGACTCATCCGGAACATACATGC
TAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATGG
ATCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCTTACTCCTTTTGGTCATAATAAC
AGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTCTGAGGCGCCACCGTCATTA
CGAATTTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGGATTTGAGGTGGC
TTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGTTTGTTA
TTGCCGGCGCAACGGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGCC
GGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCGTACAAAGATCTCCTT
GGCTTTGTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGGT
GACCCAGAAAATTTCACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGT
GGTACTTCTTATTTGCTTACGCCATTCTACGCTCTATCCCAAACAAGCTGGGAGGGGTCCTA
GCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGA
GGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATCATT
TTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTC
GGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGAGCCGAAAATAAA
GCATTGAAATGAGCTTGCCCTAGTAGCTTAG
H9
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT
TCCAACTCCATCTAATATTTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGAT
TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATTTCAACTGCCTTT
TCGTCAGTAACACACATCTGTCGAGACGTTAACTATGGCTGACTCATCCGGAACATACATGC
TAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATGG
ATCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCTTACTCCTTTTGGTCATAATAAC
AGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTCTGAGGCGCCACCGTCATTA
CGAATTTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGGATTTGAGGTGGC
TTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGTTTGTTA
TTGCCGGCGCAACGGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGCC
GGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCGTACAAAGATCTCCTT
GGCTTTGTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGGT
GACCCAGAAAATTTCACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGT
GGTACTTCTTATTTGCTTACGCCATTCTACGCTCTATCCCAAACAAGCTGGGAGGGGTCCTA
GCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGA
GGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATCATT
TTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTC
GGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGAGCCGAAAATAAA
GCATTGAAATGAGCTTGCCCTAGTAGCTTAG
H10
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT
TCCAACTCCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGAT
TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTT

TTCGTCAGTAACACACATCTGTCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATG
CTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATG
GGTCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCCTACTCCTTTTGGTCATAATAA
CAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTCTGAGGCGCCACCGTCATT
ACGAATCTATTATCAACAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGGATTTGAGGTGG
CTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGTTTGTT
ATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGCTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCC
TTGGCTTTGTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCG
GTGACCCAGAAAATTTTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGA
GTGGTACTTCTTATTTGCTTACGCCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCC
TAGCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAAC
GAGGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATTA
TTTTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCG
TCGGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGAGCCGAAAATA
AAGCATTGAAATGAGCTTGCCCTAGTAGCTTAG
H11
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT
TCCAACTCCATCTAATATTTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGAT
TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATTTCAACTGCCTTT
TCGTCAGTAACACACATCTGTCGAGACGTTAACTATGGCTGACTCATCCGGAACATACATGC
TAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATGG
ATCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCTTACTCCTTTTGGTCATAATAAC
AGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTCTGAGGCGCCACCGTCATTA
CGAATTTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGGATTTGAGGTGGC
TTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGTTTGTTA
TTGCCGGCGCAACGGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGCC
GGGTTAAACTCCGACGCAAATAAGATCTCCTTCCACCCATACTTCTCGTACAAAGATCTCCTT
GGCTTTGTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGGT
GACCCAGAAAATTTCACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGT
GGTACTTCTTATTTGCTTACGCCATTCTACGCTCTATCCCAAACAAGCTGGGAGGGGTCCTA
GCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGA
GGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATCATT
TTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTC
GGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGAGCCGAAAATAAA
GCATTGAAATGAGCTTGCCCTAGTAGCTTAG
H12
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT
TCCAACTCCATCTAATATTTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGAT
TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTT
TTCGTCAGTAACACACATCTGTCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATG
CTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATG
GGTCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCCTACTCCTTTTGGTCATAATAA
CAGCCTTTGTAGGCTATGTACTACCATGGGGTCAAATATCTTTCTGAGGCGCCACCGTCATT
ACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAATGGATTTGAGGTGG
CTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTATTTCCGTTTGTT
ATCGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCT
TGGCTTTGTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGG
TGACCCAGAAAATTTTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAG
TGGTACTTCTTATTTGCTTACGCCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTA
GCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGA
GGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATTATT

TTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTC
GGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGAGCCGAAAATAAA
GCATTGAAATGAGCTTGCCCTAGTAGCTTAG
H13
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT
TCCAACTCCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGAT
TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTT
TTCGTCAGTAACACACATCTGTCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATG
CTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATG
GGTCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCCTACTCCTTTTGGTCATAATAA
CAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTCTGAGGCGCCACCGTCATT
ACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGGATTTGAGGTGG
CTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGTTTGTT
ATTGCCGACGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCT
TGGCTTTGTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGG
TGACCCAGAAAATTTTACCCCGGCAAACCCCCTGGTTACCCCACCACACATCCAACCGGAGT
GGTACTTCTTATTTGCTTACGCCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTA
GCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGA
GGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATTATT
TTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTC
GGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGAGCCGAAAATAAA
GCATTGAAATGAGCTTGCCCTAGTAGCTTAG
H14
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT
TCCAACTCCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGAT
TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTT
TTCGTCAGTAACACACATCTGTCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATG
CTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATG
GGTCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCCTACTCCTTTTGGTCATAATAA
CAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTCTGAGGCGCCACCGTCATT
ACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGGATTTGAGGTGG
CTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGTTTGTT
ATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCATCCATACTTCTCATACAAAGATCTCCT
TGGCTTTGTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGG
TGACCCAGAAAATTTTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAG
TGGTACTTCTTATTTGCTTACGCCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTA
GCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGA
GGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATTATT
TTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTC
GGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGAGCCGAAAATAAA
GCATTGAAATGAGCTTGCCCTAGTAGCTTAG
H15
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT
TCCAACTCCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGAT
TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTT
TTCGTCAGTAACACACATCTGTCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATG
CTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATG
GGTCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCCTACTCCTTTTGGTCATAATAA
CAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTCTGAGGCGCCACCGTCATT
ACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGGATTTGAGGTGG
CTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGTTTGTT

ATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCT
TGGCTTTGTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACACTGCTCGG
TGACCCAGAAAATTTTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAG
TGGTACTTCTTATTTGCTTACGCCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTA
GCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGA
GGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATTATT
TTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTC
GGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGAGCCGAAAATAAA
GCATTGAAATGAGCTTGCCCTAGTAGCTTAG
H16
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT
TCCAACTCCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGAT
TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTT
TTCGTCAGTAACACACATCTGTCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATG
CTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATG
GGTCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTTCTACTCCTTTTAGTCATAATAA
CAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTCTGAGGCGCCACCGTCATT
ACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGGATTTGAGGTGG
CTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGTTTGTT
ATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCT
TGGCTTTGTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGG
TGACCCAGAAAATTTTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAG
TGATACTTCTTATTTGCTTACGCCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTA
GCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGA
GGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATTATT
TTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTC
GGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGAGCCGAAAATAAA
GCATTGAAATGAGCTTGCCCTAGTAGCTTAG
H17
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT
TCCAACTCCATCTAATATTTCAGTACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGAT
TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATTTCAACTGCCTTT
TCGTCAGTAACACACATCTGTCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATGC
TAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATGG
ATCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCTTACTCCTTTTGGTCATAATAAC
AGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTCTGAGGCGCCACCGTCATTA
CGAATTTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGGATTTGAGGTGGC
TTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGTTTGTTA
TTGCCGGCGCAACGGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGCC
GGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCGTACAAAGATCTCCTT
GGCTTTGTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGGT
GACCCAGAAAATTTCACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGT
GGTACTTCTTATTTGCTTACGCCATTCTACGCTCTATCCCAAACAAGCTGGGAGGGGTCCTA
GCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGA
GGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATCATT
TTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTC
GGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGAGCCGAAAATAAA
GCATTGAAATGAGCTTGCCCTAGTAGCTTAG
H18
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT
TCCAACTCCATCTAATATTTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGAT

TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATTTCAACTGCCTTT
TCGTCAGTAACACACATCTGTCGAGACGTTAACTATGGCTGACTCATCCGGAACATACATGC
TAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATGG
ATCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCTTACTCCTTTTGGTCATAATAAC
AGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTCTGAGGCGCCACCGTCATTA
CGAATTTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGGATTTGAGGTGGC
TTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGTTTGTTA
TTGCCGGCGCAACGGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGCC
GGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCGTACAAAGATCTCCTT
GGCTTTGTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTTTCTCCAACTCTGCTCGGT
GACCCAGAAAATTTCACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGT
GGTACTTCTTATTTGCTTACGCCATTCTACGCTCTATCCCAAACAAGCTGGGAGGGGTCCTA
GCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGA
GGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATCATT
TTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTC
GGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGAGCCGAAAATAAA
GCATTGAAATGAGCTTGCCCTAGTAGCTTAG
H19
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT
TCCAACTCCATCTAATATTTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGAT
TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATTTCAACTGCCTTT
TCGTCAGTAACACACATCTGTCGAGACGTTAACTATGGCTGACTCATCCGGAACATACATGC
TAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATGG
ATCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCTTACTCCTTTTGGTCATAATAAC
AGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTCTGAGGCGCCACCGTCATTA
CGAATTTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGGATTTGAGGTGGC
TTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGTTTTGTT
ATGCCGGCGCAACGGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGCC
GGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCGTACAAAGATCTCCTT
GGCTTTGTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGGT
GACCCAGAAAATTTCACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGT
GGTACTTCTTATTTGCTTACGCCATTCTACGCTCTATCCCAAACAAGCTGGGAGGGGTCCTA
GCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGA
GGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATCATT
TTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTC
GGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCCTCGCGGGCTGAGCCGAAAATAAA
GCATTGAAATGAGCTTGCCCTAGTAGCTTAG
H20
ATGGCAAGCCTACGAAAAACCCATCCGCTAATAAAAATTGCTAACAATGCACTAGTTGACCTT
CCAACCCCATCTAATATTTCAGCACTATGGAACTTCGGGTCCCTCCTAGGATTATGCTTAATT
ACTCAAATCCTAACAGGACTATTCCTGGCCATACATTATACTTCTGATATCTCAACTGCCTTTT
CATCAGTAACACATATCTGCCGAGACGTAAACTATGGCTGACTCATCCGAAACATACATGCT
AACGGAGCATCGTTCTTCTTTATCTGTATTTACATACACATTGCCCGCGGACTGTACTACGG
GTCATACCTCTATAAAGAGACCTGAAGTATTGGCGTAGTTCTACTCCTTTTAGTAATGATGAC
AGCCTTTGTGGGCTATGTACTGCCCTGAGGACAAATATCTTTTTGAGGTGCCACCGTGATTA
CAAATCTACTATCAGCAGTCCCCTATATGGGCGACGCCTTTGTCCAATGGATTTGAGGCGGC
TTTTCCGTAGATAACGCAACATTAACACGATTCTTCGCCTACCACTTCCTATTCCCGTTCGTT
ATCGCCGGTGCAACGGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGATTAAACTCCGACGCAGATAAAATCTCTTTCCACCCATACTTCTCATACAAGGACCTCCT
TGGCTTTGTACTAATGTTATTAGCTCTCACATCCCTAACATTATTCTCCCCAACACTGCTCGC
CGACCCAGAAAATTTCACCCCGCCAAACCCCCTGGTCACCCCACCGCACATCCAACCGGAG
TGATACTTCCTGTTTGCCTACGCCATTCTACGATCTATTCCAAACAAGCTAGGAGGGGTCCT
AGCGCTATTGTTTAGCATCCTAGTACTATTAGTAGTGCCAGTTTTACACACCTCGAAGCAACG

AGGACTAACTTTCCGTCCGATCACCCAGTTCTTATTCTGAACCTTAGTCGCAGACATAATCAT
TTTAACATGAATTGGAGGCATACCCGTAGAACACCCATATATCATCATTGGCCAAATCGCCTC
GGTGCTATACTTCGCACTCTTCCTTGTTCTCGCCCCCCTCGCAGGGTGAGCCGAGAATAAA
GCATTGAAATGAGCTTGCCCTAGTAGCTTAG
H21
ATGGCAAGCCTACGAAAAACCCATCCGCTAATAAAAATTGCTAACAATGCACTAGTTGACCTT
CCAACCCCATCTAATATTTCAGCACTATGGAACTTCGGGTCCCTCCTAGGATTATGCTTAATT
ACTCAAATCCTAACAGGACTATTCCTCGCCATACATTATACTTCTGATATCTCAACTGCCTTTT
CATCAGTAACACATATCTGCCGAGACGTAAACTATGGCTGACTCATCCGAAACATACATGCT
AACGGAGCATCGTTCTTCTTTATCTGTATTTACATACACATTGCCCGCGGACTGTACTACGG
GTCATACCTCTATAAAGAGACCTGAAGTATTGGCGTAGTTCTACTCCTTTTAGTAATGATGAC
AGCCTTTGTGGGCTATGTACTGCCCTGAGGACAAATATCTTTTTGAGGTGCCACCGTGATTA
CAAATCTACTATCAGCAGTCCCCTATATGGGCGACGCCTTTGTCCAATGGATTTGAGGCGGC
TTTTCCGTAGATAACGCAACATTAACACGATTCTTCGCCTTCCACTTCCTATTCCCGTTCGTT
ATCGCCGGTGCAACGGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGATTAAACTCCGACGCAGATAAAATCTCTTTCCACCCATACTTCTCATACAAGGACCTCCT
TGGCTTTGTACTAATGTTATTAGCTCTCACATCCCTAACATTATTCTCCCCAACACTGCTCGG
CGACCCAGAAAATTTCACCCCGGCAAACCCCCTGGTCACCCCACCGCACATCCAACCGGAG
TGATACTTCCTGTTTGCCTACGCCATTCTACGATCTATTCCAAACAAGCTAGGAGGGGTCCT
AGCGCTATTGTTTAGCATCCTAGTACTATTAGTAGTGCCAGTTTTACACACCTCGAAGCAACG
AGGACTAACTTTCCGTCCGATCACCCAGTTCTTATTCTGAACCTTAATCGCAGACATAATCAT
TTTAACATGAATTGGAGGCATACCCGTAGAACACCCATATATCATCATTGGCCAAATCGCCTC
GGTGCTATACTTCGCACTCTTCCTTGTTCTCGCCCCCCTCGCAGGGTGAGCCGAGAATAAA
GCATTGAAATGAGCTTGCCCTAGTAGCTTAG
H22
ATGGCAAGCCTACGAAAAACCCATCCGCTAATAAAAATTGCTAACAATGCACTAGTTGACCTT
CCAACCCCATCTAATATTTCAGCACTATGGAACTTCGGGTCCCTCCTAGGATTATGCTTAATT
ACTCAAATCCTAACAGGACTATTCCTGGCCATACATTATACTTCTGATATCTCAACTGCCTTTT
CATCAGTAACACATATCTGCCGAGACGTAAACTATGGCTGACTCATCCGAAACATACATGCT
AACGGAGCATCGTTCTTCTTTATCTGTATTTACATACACATTGCCCGCGGACTGTACTACGG
GTCATACCTCTATAAAGAGACCTGAAGTATTGGCGTAGTTCTACTCCTTTTAGTAATGATGAC
AGCCTTTGTGGGCTATGTACTGCCCTGAGGACAAATATCTTTTTGAGGTGCCACCGTGATTA
CAAATCTACTATCAGCAGTCCCCTATATGGGGGACGCCTTTGTCCAATGGATTTGAGGCGGC
TTTTCCGTAGATAACGCAACATTAACACGATTCTTCGCCTTCCACTTCCTATTCCCGTTCGTT
ATCGCCGGTGCAACGGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGATTAAACTCCGACGCAGATAAAATCTCTTTCCACCCATACTCCTCATACAAGGACCTCCT
TGGCTTTGTACTAATGTTATTAGCTTTCACATCCCTAACATTATTCTCCCCAACACTGCTCGC
CGACCCAGAAAATTTCACCCCGGCAAACCCCCTGGTCACCCCACCGCACATCCAACCGGAG
TGATACTTCCTGTTTGCCTACGCCATTCTACGATCTATTCCAAACAAGCTAGGAGGGGTCCT
AGCGCTATTGTTTAGCATCCTAGTACTATTAGTAGTGCCAGTTTTACACACCTCGAAGCAACG
AGGACTAACTTTCCGTCCGATCACCCAGTTCTTATTCTGAACCTTAGTCGCAGACATAATCAT
TTTAACATGAATTGGAGGCATACCCGTAGAACACCCATATATCATCATTGGCCAAATCGCCTC
GGTGCTATACTTCGCACTCTTCCTTGTTCTCGCCCCCCTCGCAGGGTGAGCCGAGAATAAA
GCATTGAAATGAGCTTGCCCTAGTAGCTTAG
H23
ATGGCAAGCCTACGAAAAACCCATCCGCTAATAAAAATTGCTAACAATGCACTAGTTGACCTT
CCAACCCCATCTAATATTTCAGCACTATGGAACTTCGGGTCCCTCCTAGGATTATGCTTAATT
ACTCAAATCCTAACAGGACTATTCCTGGCCATACATTATACTTCTGATATCTCAACTGCCTTTT
CATCAGTAACACATATCTGCCGAGACGTAAACTATGGCTGACTCATCCGAAACATACATGCT
AACGGAGCATCGTTCTTCTTTATCTGTATTTACATACACATTGCCCGCGGACTGTACTACGG
GTCATACCTCTATAAAGAGACCTGAAGTATTGGCGTAGTTCTACTCCTTTTAGTAATGATGAC
AGCCTTTGTGGGCTATGTACTGCCCTGAGGACAAATATCTTTTTGAGGTGCCACCGTGATTA
CAAATCTACTATCAGCAGTCCCCTATATGGGCGACGCCTTTGTCCAATGGATTTGAGGCGGC

TTTTCCGTAGATAACGCAACATTAACACGATTCTTCGCCTTCCACTTCCTATTCCCGTTCGTT
ATCGCCGGTGCAACGGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGATTAAACTCCGACGCAGATAAAATCTCTTTCCACCCATACTTCTCATACAAGGACCTCCT
TGGCTTTGTACTAATGTTATTAGCTCTCACATCCCTAACATTATTCTCCCCAACACTGCTCGC
CGACCCAGAAAATTTCACCCCGCCAAACCCCCTGGTCACCCCACCGCACATCCAACCGGAG
TGATACTTCCTGTTTGCCTACGCCATTCTACGATCTATTCCAAACAAGCTAGGAGGGGTCCT
AGCGCTATTGTTTAGCATCCTAGTACTATTAGTAGTGCCAGTTTTACACACCTCGAAGCAACG
AGGACTAACTTTCCGTCCGATCACCCAGTTCTTATTCTGAACCTTAGTCGCAGACATAATCAT
TTTAACATGAATTGGAGGCATACCCGTAGAACACCCATATATCATCATTGGCCAAATCGCCTC
GGTGCTATACTTCGCACTCTTCCTTGTTCTCGCCCCCCTCGCAGGGTGAGCCGAGAATAAA
GCATTGAAATGAGCTTGCCCTAGTAGCTTAG
H24
ATGGCAAGCCTACGAAAAACCCATCCGCTAATAAAAATTGCTAACAATGCACTAGTTGACCTT
CCAACCCCATCTAATATTTCAGCACTATGGAACTTCGGGTCCCTCCTAGGATTATGCTTAATT
ACTCAAATCCTAACAGGACTATTCCTGGCCATACATTATACTTCTGATATCTCAACTGCCTTTT
CATCAGTAACACATATCTGCCGAGACGTAAACTATGGCTGACTCATCCGAAACATACATGCT
AACGGAGCATCGTTCTTCTTTATCTGTATTTACATACACATTGCCCGCGGACTGTACTACGG
GTCATACCTCTATAAAGAGACCTGAAGTATTGGCGTAGTTCTACTCCTTTTAGTAATGATGAC
AGCCTTTGTGGGCTATGTACTGCCCTGAGGACAAATATCTTTTTGAGGTGCCACCGTGATTA
CAAATCTACTATCAGCAGTCCCCTATATGGGCGACGCCTTTGTCCAATGGATTTGAGGCGGC
TTTTCCGTAGATAACGCAACATTAACACGATTCTTCGCCTTCCACTTCCTATTCCCGTTCGTT
ATCGCCGGTGCAACGGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGATTAAACTCCGACGCAGATAAAATCTCTTTCCACCCATACTTCTCATACAAGGACCTCCT
TGGCTTTGTACTAATGTTATTAGCTCTCACATCCCTAACATTATTCTCCCCAACACTGCTCGC
CGACCCAGAAAATTTCACCCCGGCAAACCCCCTGGTCACCCCACCGCACATCCAACCGGAG
TGATACTTCCTGTTTGCCTACGCCATTCTACGATCTATTCCAAACAAGCTAGGAGGGGTCCT
AGCGCTATTGTTTAGCATCCTAGTACTATTAGTAGTGCCAGTTTTACACACCTCGAAGCAACG
AGGACTAACTTTCCGTCCGATCACCCAGTTCTTATTCTGAACCTTAGTCGCAGACATAATCAT
TTTAACATGAATTGGAGGCATACCCGTAGAACACCCATATATCATCATTGGCCAAATCGCCTC
GGTGCTATACTTCGCACTCTTCCTTGTTCTCGCCCCCCTCGCAGGGTGAGCCGAGAATAAA
GCATTGAAATGAGCTTGCCCTAGTAGCTTAG
H25
ATGGCAAGCCTACGAAAAACCCATCCGCTAATAAAAATTGCTAACAATGCACTAGTTGACCTT
CCAACCCCATCTAATATTTCAGCACTATGGAACTTCGGGTCCCTCCTAGGATTATGCTTAATT
ACTCAAATCCTAACAGGACTATTCCTGGCCATACATTATACTTCTGATATCTCAACTGCCTTTT
CATCAGTAACACATATCTGCCGAGACGTAAACTATGGCTGACTCATCCGAAACATACATGCT
AACGGAGCATCGTTCTTCTTTATCTGTATTTACATACACATTGCCCGCGGACTGTACTACGG
GTCATACCTCTATAAAGAGACCTGAAGTATTGGCGTAGTTCTACTCCTTTTAGTAATGATGAC
AGCCTTTGTGGGCTATGTACTGCCCTGAGGACAAATATCTTTTTGAGGTGCCACCGTGATTA
CAAATCTACTATCAGCAGTCCCCTATATGGGCGACGCCTTTGTCCAATGGATTTGAGGCGGC
TTTTCCGTAGATAACGCAACATTAACACGATTCTTCGCCTTCCACTTCCTATTCCCGTTCGTT
ATCGCCGGTGCAACGGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGATTAAACTCCGACGCAGATAAAATCTCTTTCCACCCATACTTCTCATACAAGGACCTCCT
TGGCTTTGTACTAATGTTATTAGCTCTCACATCCCTAACATTATTCTCCCCAACACTGCTCGG
CGACCCAGAAAATTTCACCCCGCCAAACCCCCTGGTCACCCCACCGCACATCCAACCGGAG
TGATACTTCCTGTTTGCCTACGCCATTCTACGATCTATTCCAAACAAGCTAGGAGGGGTCCT
AGCGCTATTGTTTAGCATCCTAGTACTATTAGTAGTGCCAGTTTTACACACCTCGAAGCAACG
AGGACTAACTTTCCGTCCGATCACCCAGTTCTTATTCTGAACCTTAGTCGCAGACATAATCAT
TTTAACATGAATTGGAGGCATACCCGTAGAACACCCATATATCATCATTGGCCAAATCGCCTC
GGTGCTATACTTCGCACTCTTCCTTGTTCTCGCCCCCCTCGCAGGGTGAGCCGAGAATAAA
GCATTGAAATGAGCTTGCCCTAGTAGCTTAG
H26
ATGGCAAGCCTACGAAAAACCCATCCGCTAATAAAAATTGCTAACAATGCACTAGTTGACCTT

CCAACCCCATCTAATATTTCAGCACTATGGAACTTCGGGTCCCTCCTAGGATTATGCTTAATT
ACTCAAATCCTAACAGGACTATTCCTCGCCATACATTATACTTCTGATATCTCAACTGCCTTTT
CATCAGTAACACATATCTGCCGAGACGTAAACTATGGCTGACTCATCCGAAACATACATGCT
AACGGAGCATCGTTCTTCTTTATCTGTATTTACATACACATTGCCCGCGGACTGTACTACGG
GTCATACCTCTATAAAGAGACCTGAAGTATTGGCGTAGTTCTACTCCTTTTAGTAATGATGAC
AGCCTTTGTGGGCTATGTACTGCCCTGAGGACAAATATCTTTTTGAGGTGCCACCGTGATTA
CAAATCTACTATCAGCAGTCCCCTATATGGGCGACGCCTTTGTCCAATGGATTTGAGGCGGC
TTTTCCGTAGATAACGCAACATTAACACGATTCTTCGCCTTCCACTTCCTATTCCCGTTCGTT
ATCGCCGGTGCAACGGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGATTAAACTCCGACGCAGATAAAATCTCTTTCCACCCATACTTCTCATACAAGGACCTCCT
TGGCTTTGTACTAATGTTATTAGCTCTCACATCCCTAACATTATTCTCCCCAACACTGCTCGG
CGACCCAGAAAATTTCACCCCGGCAAACCCCCTGGTCACCCCACCGCACATCCAACCGGAG
TGATACTTCCTGTTTGCCTACGCCATTCTACGATCTATTCCAAACAAGCTAGGAGGGGTCCT
AGCGCTATTGTTTAGCATCCTAGTACTATTAGTAGTGCCAGTTTTACACACCTCGAAGCAACG
AGGACTAACTTTCCGTCCGATCACCCAGTTCTTATTCTGAACCTTAATCGCAGACATAATCAT
TCTAACATGAATTGGAGGCATACCCGTAGAACACCCATATATCATCATTGGCCAAATCGCCT
CGGTGCTATACTTCGCACTCTTCCTTGTTCTCGCCCCCCTCGCAGGGTGAGCCGAGAATAAA
GCATTGAAATGAGCTTGCCCTAGTAGCTTAG
H27
ATGGCAAGCCTACGAAAAACCCATCCGCTAATAAAAATTGCTAACAATGCACTAGTTGACCTT
CCAACCCCATCTAATATTTCAGCACTATGGAACTTCGGGTCCCTCCTAGGATTATGCTTAATT
ACTCAAATCCTAACAGGACTATTCCTCGCCATACATTATACTTCTGATATCTCAACTGCCTTTT
CATCAGTAACACATATCTGCCGAGACGTAAACTATGGCTGACTCATCCGAAACATACATGCT
AACGGAGCATCGTTCTTCTTTATCTGTATTTACATACACATTGCCCGCGGACTGTACTACGG
GTCATACCTCTATAAAGAGACCTGAAGTATTGGCGTAGTTCTACTCCTTTTAGTAATGATGAC
AGCCTTTGTGGGCTATGTACTGCCCTGAGGACAAATATCTTTTTGAGGTGCCACCGTGATTA
CAAATCTACTATCAGCAGTCCCCTATATGGGCGACGCCTTTGTCCAATGGATTTGAGGCGGC
TTTTCCGTAGATAACGCAACATTAACACGATTCTTCGCCTTCCACTTCCTATTCCCGTTCGTT
ATCGCCGGTGCAACGGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGATTAAACTCCGACGCAGATAAAATCTCTTTCCACCCATACTTCTCATACAAGGACCTCCT
TGGCTTTGTACTAATGTTATTAGCTCTCACATCCCTAACATTATTCTCCCCAACACTGCTCGG
CGACCCAGAAAATTTCACCCCGCCAAACCCCCTGGTCACCCCACCGCACATCCAACCGGAG
TGATACTTCCTGTTTGCCTACGCCATTCTACGATCTATTCCAAACAAGCTAGGAGGGGTCCT
AGCGCTATTGTTTAGCATCCTAGTACTATTAGTAGTGCCAGTTTTACACACCTCGAAGCAACG
AGGACTAACTTTCCGTCCGATCACCCAGTTCTTATTCTGAACCTTAATCGCAGACATAATCAT
TTTAACATGAATTGGAGGCATACCCGTAGAACACCCATATATCATCATTGGCCAAATCGCCTC
GGTGCTATACTTCGCACTCTTCCTTGTTCTCGCCCCCCTCGCAGGGTGAGCCGAGAATAAA
GCATTGAAATGAGCTTGCCCTAGTAGCTTAG
H28
ATGGCAAGCCTACGAAAAACCCATCCGCTAATAAAAATTGCTAACAATGCACTAGTTGACCTT
CCAACCCCATCTAATATTTCAGCACTATGGAACTTCGGGTCCCTCCTAGGATTATGCTTAATT
ACTCAAATCCTAACAGGACTATTCCTGGCCATACATTATACTTCTGATATCTCAACTGCCTTTT
CATCAGTAACACATATCTGCCGAGACGTAAACTATGGCTGACTCATCCGAAACATACATGCT
AACGGAGCATCATTCTTCTTTATCTGTATTTACATGCACATTGCCCGCGGACTGTACTACGG
GTCATACCTCTATAAAGAGACCTGAAGTATTGGCGTAGTTCTACTCCTTTTAGTAATGATGAC
AGCCTTTGTGGGCTATGTACTGCCCTGAGGACAAATATCTTTTTGAGGTGCCACCGTGATTA
CAAATCTACTATCAGCAGTCCCCTATATGGGCGACGCCTTTGTCCAATGGATTTGAGGCGGC
TTTTCCGTAGATAACGCAACATTAACACGATTCTTCGCCTTCCACTTCCTATTCCCGTTCGTT
ATCGCCGGTGCAACGGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGATTAAACTCCGACGCAGATAAAATCTCTTTCCACCCATACTTCTCATACAAGGACCTCCT
TGGCTTTGTACTAATGTTATTAGCTCTCACATCCCTAACATTATTCTCCCCAACACTGCTCGG
CGACCCAGAAAATTTCACCCCGGCAAACCCCCTGGTCACCCCACCGCACATCCAACCGGAG
TGATACTTCCTGTTTGCCTACGCCATTCTACGATCTATTCCAAACAAGCTAGGAGGGGTCCT

AGCGCTATTGTTTAGCATCCTAGTACTATTAGTAGTGCCAGTTTTACACACCTCGAAGCAACG
AGGACTAACTTTCCGTCCGATCACCCAGTTCTTATTCTGAACCTTAGTCGCAGACATAATCAT
TTTAACATGAATTGGAGGCATACCCGTAGAACACCCATATATCATCATTGGCCAAATCGCCTC
GGTGCTATACTTCGCACTCTTCCTTGTTCTCGCCCCCCTCGCAGGGTGAGCCGAGAATAAA
GCATTGAAATGAGCTTGCCCTAGTAGCTTAG
H29
ATGGCAAGCCTACGAAAAACCCATCCGCTAATAAAAATTGCTAACAATGCACTAGTTGACCTT
CCAACCCCATCTAATATTTCAGCACTATGGAACTTCGGGTCCCTCCTAGGATTATGCTTAATT
ACTCAAATCCTAACAGGACTATTCCTGGCCATACATTATACTTCTGATATCTCAACTGCCTTTT
CATCAGTAACACATATCTGCCGAGACGTAAACTATGGCTGACTCATCCGAAACATACATGCT
AACGGAGCATCATTCTTCTTTATCTGTATTTACATGCGCATTGCCCGCGGACTGTACTACGG
GTCATACCTCTATAAAGAGACCTGAAGTATTGGCGTAGTTCTACTCCTTTTAGTAATGATGAC
AGCCTTTGTGGGCTATGTACTGCCCTGAGGACAAATATCTTTTTGAGGTGCCACCGTGATTA
CAAATCTACTATCAGCAGTCCCCTTTATGGGCGACGCCTTTGTCCAATGGATTTGAGGCGGC
TTTTCCGTAGATAACGCAACATTAACACGATTCTTCGCCTTCCACTTCCTATTCCCGTTCGTT
ATCGCCGGTGCAACGGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGATTAAACTCCGACGCAGATAAAATCTCTTTCCACCCATACTTCTCATACAAGGACCTCCT
TGGCTTTGTACTAATGTTATTAGCTCTCACATCCCTAACATTATTCTCCCCAACACTGCTCGG
CGACCCAGAAAATTTCACCCCGGCAAACCCCCTGGTCACCCCACCGCACATCCAACCGGAG
TGATACTTCCTGTTTGCCTACGCCATTCTACGATCTATTCCAAACAAGCTAGGAGGGGTCCT
AGCGCTATTGTTTAGCATCCTAGTACTATTAGTAGTGCCAGTTTTACACACCTCGAAGCAACG
AGGACTAACTTTCCGTCCGATCACCCAGTTCTTATTCTGAACCTTAGTCGCAGACATAATCAT
TTTAACATGAATTGGAGGCATACCCGTAGAACACCCATATATCATCATTGGCCAAATCGCCTC
GGTGCTATACTTCGCACTCTTCCTTGTTCTCGCCCCCCTCGCAGGGTGAGCCGAGAATAAA
GCATTGAAATGAGCTTGCCCTAGTAGCTTAG
H30
ATGGCAAGCCTACGAAAAACTCACCCACTAATAAAAATCGCTAACGACGCACTAGTTGACCT
TCCAACCCCATCTAATATTTCAGCACTATGAAACTTCGGATCCCTCCTAGGATTATGCTTAAT
TACTCAAATCTTAACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTTT
TCATCAGTAACACACATCTGCCGAGACGTTAACTATGGCTGACTTATCCGAAACATACATGCT
AACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGGCTCTACTATGGA
TCATACCTCTATAAAGAAACCTGAAACATTGGCGTAGTTCTACTTCTTTTAGTTATGATGACA
GCCTTTGTAGGCTATGTATTACCATGAGGACAAATATCTTTTTGGGGTGCTACCGTGATTACA
AATCTACTATCAGCAGTCCCTTATATAGGAGACGCCCTTGTCCAATGGATTTGAGGTGGCTT
TTCCGTAGATAACGCAACGTTAACACGATTCTTCGCCTTCCACTTCTTATTTCCGTTCGTTAT
CGCCGGTGCAACGATTCTGCATTTGCTATTTCTACACGAAACGGGATCAAACAACCCTGCCG
GATTAAACTCCGACGCGGATAAAATCTCTTTCCACCCATACTTCTCATACAAGGACCTCCTAG
GCTTTGTGCTAATATTATTAGCCCTCACATCCCTAACATTATTCTCACCAACACTGCTCGGCG
ACCCAGAAAACTTCACCCCGGCAAACCCCCTGGTTACCCCACCACACATCCAACCGGAGTG
GTACTTCTTGTTTGCCTACGCCATTCTACGGTCTATTCCAAACAAGCTAGGAGGAGTCCTAG
CGTTGTTATTTAGCATTTTAGTACTTTTAGTGGTGCCAATTTTACATACCTCGAAGCAGCGAG
GACTAACTTTCCGTCCAATCACCCAGTTCTTATTCTGAACCTTGGTGGCAGACATAATCATTT
TGACATGAATCGGGGGCATACCCGTAGAACACCCATATATCATCATTGGCCAAATCGCGTCG
GTGCTGTACTTCGCACTCTTCCTTGTTCTCGCCCCGCTCGCAGGATGAGCCGAGAATAAAG
CATTGAAATGAGCTTGCCCTAGTAGCTTAG
H31
ATGGCAAGCCTACGAAAAACTCACCCACTAATAAAAATCGCTAACGACGCACTAGTTGACCT
TCCAACCCCATCTAATATTTCAGCACTATGAAACTTCGGATCCCTTCTAGGATTATGCTTAATT
ACTCAAATCTTAACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTTTT
CATCAGTAACACACACCTGCCGAGACGTTAACTATGGCTGACTTATCCGAAACATACATGCT
AACGGAGCATCATTCTTCTTTATCTGTATTTACATACACATTGCCCGCGGGCTCTACTATGGA
TCATACCTCTATAAAGAAACCTGAAACATTGGCGTAGTTCTACTTCTTTTAGTTATGATGACA
GCCTTTGTAGGCTATGTATTACCATGAGGACAAATATCTTTTTGGGGTGCTACCGTGATTACA

AATCTACTATCAGCAGTCCCTTATATAGGCGACGCCCTTGTCCAGTGGATTTGAGGTGGCTT
TTCCGTAGATAACGCAACGTTAACACGATTCTTCGCCTTCCACTTCTTATTTCCGTTCGTTAT
CGCCGGTGCAACAATTCTACATTTGCTATTTCTACACGAAACGGGATCAAACAACCCTGCCG
GATTAAACTCCGACGCGGATAAAATCTCTTTCCACCCGTACTTCTCATACAAGGACCTCCTA
GGCTTTGTGCTAATATTATTAGCCCTCACATCCCTAACATTATTCTCACCAACACTGCTCGGC
GACCCAGAAAACTTCACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGT
GGTACTTCTTGTTTGCCTACGCCATTCTACGGTCTATTCCAAACAAGCTAGGAGGAGTCCTA
GCATTGTTATTTAGCATTTTAGTACTTTTAGTAGTGCCAATTTTACATACCTCGAAGCAGCGA
GGACTAACTTTCCGTCCAATCACCCAGTTCTTATTCTGAACCTTGGTGGCAGACATAATCATT
TTGACATGAATCGGGGGCATACCCGTAGAACACCCATATATCATCATTGGCCAAATCGCGTC
GGTGCTGTACTTCGCACTCTTCCTTGTTCTCGCCCCGCTCGCAGGATGAGCCGAGAATAAA
GCATTGAAATGAGCTTGCCCTAGTAGCTTAG
H32
ATGGCAAGCCTACGAAAAACTCACCCACTAATAAAAATCGCTAACGACGCACTAGTTGACCT
TCCAACCCCATCTAATATTTCAGCACTATGAAACTTCGGATCCCTTCTAGGATTATGCTTAATT
ACTCAAATCTTAACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTTTT
CATCAGTAACACACATCTGCCGAGACGTTAACTATGGCTGACTTATCCGAAACATACATGCT
AACGGAGCATCATTCTTCTTTATCTGTATTTACATACACATTGCCCGCGGGCTCTACTATGGA
TCATACCTCTATAAAGAAACCTGAAACATTGGCGTAGTTCTACTTCTTTTAGTTATGATGACA
GCCTTTGTAGGCTATGTATTACCATGAGGACAAATATCTTTTTGGGGTGCTACCGTGATTACA
AATCTACTATCAGCAGTCCCTTATATAGGCGACGCCTTTGTCCAGTGGATTTGAGGTGGCTT
TTCCGTAGATAACGCAACGTTAACACGATTCTTCGCCTTCCACTTCTTATTTCCGTTCGTTAT
CGCCGGTGCAACAATTCTGCATTTGCTATTTCTACACGAAACGGGATCAAACAACCCTGCCG
GATTAAACTCCGACGCGGATAAAATCTCTTTCCACCCGTACTTCTCATACAAGGACCTCCTA
GGCTTTGTGCTAATATTATTAGCCCTCACATCCCTAACATTATTCTCACCAACACTGCTCGGC
GACCCAGAAAACTTCACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGT
GGTACTTCTTGTTTGCCTACGCCATTCTACGGTCTATTCCAAACAAGCTAGGAGGAGTCCTA
GCATTGTTATTTAGCATTTTAGTACTTTTAGTAGTGCCAATTTTACATACCTCGAAGCAGCGA
GGACTAACTTTCCGTCCAATCACCCAGTTCTTATTCTGAACCTTGGTGGCAGACATAATCATT
TTGACATGAATCGGGGGCATACCCGTAGAACACCCATATATCATCATTGGCCAAATCGCGTC
GGTGCTGTACTTCGCACTCTTCCTTGTTCTCGCCCCGCTCGCAGGATGAGCCGAGAATAAA
GCATTGAAATGAGCTTGCCCTAGTAGCTTAG
H33
ATGGCAAGCCTACGAAAAACTCACCCACTAATAAAAATCGCTAACGACGCACTAGTTGACCT
TCCAACCCCATCTAATATTTCAGCACTATGAAACTTCGGATCCCTCCTAGGATTATGCTTAAT
TACTCAAATCTTAACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTTT
TCATCAGTAACACACATCTGCCGAGACGTTAACTATGGCTGACTTATCCGAAACATACATGCT
AACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGGCTCTACTATGGA
TCATACCTCTATAAAGAAACCTGAAACATTGGCGTAGTTCTACTTCTTTTAGTTATGATGACA
GCCTTTGTAGGCTATGTATTACCATGAGGACAAATATCTTTTTGGGGCGCTACCGTGATTACA
AATCTACTATCAGCAGTCCCTTATATAGGCGACGCCTTTGTCCAATGGATTTGAGGTGGCTTT
TCCGTAGATAACGCAACGTTAACACGATTCTTCGCCTTCCACTTCTTGTTTCCGTTCGTTATC
GCCGGTGCAACAATTCTGCATTTGCTATTTCTACACGAAACGGGATCAAACAACCCTGCCGG
ATTAAACTCCGACGCGGATAAAATCTCTTTCCACCCGTACTTCTCATACAAGGACCTCCTAG
GCTTTGTGCTAATATTATTAGCCCTCACATCCCTAACGTTATTCTCACCAACACTGCTCGGCG
ACCCAGAAAACTTCACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGTG
GTACTTCTTGTTTGCCTACGCCATTCTACGGTCTATTCCAAACAAGCTAGGAGGAGTCCTAG
CATTGTTATTTAGCATTTTAGTACTTTTAGTAGTGCCAATTTTACATACCTCGAAGCAGCGAG
GACTAACTTTCCGTCCAATCACCCAGTTCTTATTCTGAACCTTGGTGGCAGACATAATCATTT
TGACATGAATCGGGGGCATACCCGTAGAACACCCATATATCATCATTGGCCAAATCGCGTCG
GTGCTGTACTTCGCACTCTTCCTTGTTCTCGCCCCGCTCGCAGGATGAGCCGAGAATAAAG
CATTGAAATGAGCTTGCCCTAGTAGCTTAG

```
H34
ATGGCAAGCCTACGAAAAACTCACCCACTAATAAAAATCGCTAACGACGCACTAGTTGACCT
TCCAACCCCATCTAATATTTCAGCACTATGAAACTTCGGATCCCTCCTAGGATTATGCTTAAT
TACTCAAATCTTAACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTTT
TCATCAGTAACACACATCTGCCGAGACGTTAACTATGGCTGACTTATCCGAAACATACATGCT
AACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGGCTCTACTATGGA
TCATACCTCTATAAAGAAACCTGAAACATTGGCGTAGTTCTACTTCTTTTAGTTATGATGACA
GCCTTTGTAGGCTATGTATTACCATGAGGACAAATATCTTTTTGGGGCGCTACCGTGATTACA
AATCTACTATCAGCAGTCCCTTATATAGGCGACGCCTTTGTCCAATGGATTTGAGGTGGCTTT
TCCGTAGATAACGCAACGTTAACACGATTCTTCGCCTTCCACTTCTTGTTTCCGTTCGTTATC
GCCGGTGCAACAATTCTGCATTTGCTATTTCTACACGAAACGGGATCAAACAACCCTGCCGG
ATTAAACTCCGACGCGGATAAAATCTCTTTCCACCCATACTTCTCATACAAGGACCTCCTAGG
CTTTGTGCTAATATTATTAGCCCTCACATCCCTAACGTTATTCTCACCAACACTGCTCGGCGA
CCCAGAAAACTTCACCCCGGCAAACCCCCTGGTTACCCCACCGCACATCCAACCGGAGTGG
TACTTCTTGTTTGCCTACGCCATTCTACGGTCTATTCCAAACAAGCTAGGAGGAGTCCTAGC
ATTGTTATTTAGCATTTTAGTACTTTTAGTAGTGCCAATTTTACATACCTCGAAGCAGCGAGG
ACTAACTTTCCGTCCAATCACCCAGTTCTTATTCTGAACCTTGGTGGCAGACATAATCATTTT
GACATGAATCGGGGGCATACCCGTAGAACACCCATATATCATCATTGGCCAAATCGCGTCG
GTGCTGTACTTCGCACTCTTCCTTGTTCTCGCCCCGCTCGCAGGATGAGCCGAGAATAAAG
CATTGAAATGAGCTTGCCCTAGTAGCTTAG

}

[[Samples]]

   SampleName="IC"
   SampleSize= 5  #Fictive number, but must match the sume of haplotype frequencies given
below
          SampleData= {
     #Example of a sample consisiting of haplotypic data (2 haplotypes, 2 loci):
     H1 3
     H2 1
        H3 1
   }

   SampleName="CC"
   SampleSize= 4  #Fictive number, but must match the sume of haplotype frequencies given
below
          SampleData= {
     #Example of a sample consisiting of haplotypic data (2 haplotypes, 2 loci):
     H4 2
     H5 2
   }

   SampleName="PC"
   SampleSize= 8  #Fictive number, but must match the sume of haplotype frequencies given
below
          SampleData= {
     #Example of a sample consisiting of haplotypic data (2 haplotypes, 2 loci):
     H4 6
     H6 1
        H7 1
   }
```

```
    SampleName="TWIN"
    SampleSize= 12  #Fictive number, but must match the sume of haplotype frequencies given below
            SampleData= {
      #Example of a sample consisiting of haplotypic data (2 haplotypes, 2 loci):
      H8 2
      H9 2
          H10 6
          H11 1
          H12 1
    }

    SampleName="CAJ"
    SampleSize= 13  #Fictive number, but must match the sume of haplotype frequencies given below
            SampleData= {
      #Example of a sample consisiting of haplotypic data (2 haplotypes, 2 loci):
      H5 11
      H13 1
          H14 1
    }

    SampleName="LCK"
    SampleSize= 7  #Fictive number, but must match the sume of haplotype frequencies given below
            SampleData= {
      #Example of a sample consisiting of haplotypic data (2 haplotypes, 2 loci):
      H5 5
      H15 1
          H16 1
    }

    SampleName="SGR"
    SampleSize= 8  #Fictive number, but must match the sume of haplotype frequencies given below
            SampleData= {
      #Example of a sample consisiting of haplotypic data (2 haplotypes, 2 loci):
      H9 5
      H17 1
          H18 1
          H19 1
    }

    SampleName="SLO"
    SampleSize= 9  #Fictive number, but must match the sume of haplotype frequencies given below
            SampleData= {
      #Example of a sample consisiting of haplotypic data (2 haplotypes, 2 loci):
      H20 1
      H21 2
          H22 1
          H23 1
```

```
        H24 1
        H25 1
        H26 1
        H27 1
   }

   SampleName="SMR"
   SampleSize= 5  #Fictive number, but must match the sume of haplotype frequencies given below
          SampleData= {
     #Example of a sample consisiting of haplotypic data (2 haplotypes, 2 loci):
     H28 4
        H29 1
   }

   SampleName="OWR"
   SampleSize= 6  #Fictive number, but must match the sume of haplotype frequencies given below
          SampleData= {
     #Example of a sample consisiting of haplotypic data (2 haplotypes, 2 loci):
     H30 1
     H31 2
        H32 1
        H33 1
        H34 1
   }



        [[Structure]]

                StructureName="New Edited Structure"
                NbGroups=5

                Group={
                        "IC"
                }

                Group={
                        "CC"
                        "PC"
                        "TWIN"
                        "CAJ"
                }

                Group={
                        "LCK"
                        "SGR"
                }

                Group={
                        "SMR"
                        "SLO"
```

```
	}

	Group={
		"OWR"
	}
```

TCS Input File

```
#NEXUS

Begin data;
	Dimensions ntax=57 nchar=1155;
	Format datatype=nucleotide gap=- missing=? matchchar=.;
	Matrix
IC1
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT
TCCAACTCCATCTAATATTTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGAT
TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTT
TTCGTCAGTAACACACATCTGTCGAGACGTTAACTATGGCTGACTCATCCGAAACATACATG
CTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATG
GGTCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCCTACTCCTTTTGGTCATAATAA
CAGCCTTTGTAGGCTATGTACTGCCGTGAGGTCAAATGTCTTTCTGAGGCGCCACCGTCATT
ACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGGATTTGAGGTGG
CTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGTTTGTT
ATCGCCGGCGCAACAGTCCTACATTTGTTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGTTAAACTCCAACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCT
TGGCTTTGTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGG
TGACCCAGAAAATTTCACCCCGGCAAACCCCCTGGTTACCCCACCACACATCCAACCGGAG
TGGTACTTCTTATTTGCTTACGCCATTCTACGCTCTATCCCAAACAAGCTGGGAGGGGTCCT
AGCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACG
AGGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATCAT
TTTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGT
CGGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCTCGCGGGCTGAGCCGAAAATAAA
GCATTGAAATGAGCTTGCCCTAGTAGCTTAG
IC2
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT
TCCAACTCCATCTAATATTTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGAT
TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTT
TTCGTCAGTAACACACATCTGTCGAGACGTTAACTATGGCTGACTCATCCGAAACATACATG
CTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATG
GGTCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCCTACTCCTTTTGGTCATAATAA
CAGCCTTTGTAGGCTATGTACTGCCGTGAGGTCAAATGTCTTTCTGAGGCGCCACCGTCATT
ACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGGATTTGAGGTGG
CTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGTTTGTT
ATCGCCGGCGCAACAGTCCTACATTTGTTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGTTAAACTCCAACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCT
TGGCTTTGTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGG
TGACCCAGAAAATTTCACCCCGGCAAACCCCCTGGTTACCCCACCACACATCCAACCGGAG
TGGTACTTCTTATTTGCTTACGCCATTCTACGCTCTATCCCAAACAAGCTGGGAGGGGTCCT
AGCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACG
```

AGGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATCAT
TTTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGT
CGGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCTCGCGGGCTGAGCCGAAAATAAA
GCATTGAAATGAGCTTGCCCTAGTAGCTTAG

IC3

ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT
TCCAACTCCATCTAATATTTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGAT
TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTT
TTCGTCAGTAACACACATCTGTCGAGACGTTAACTATGGCTGACTCATCCGAAACATACATG
CTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATG
GGTCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCCTACTCCTTTTGGTCATAATAA
CAGCCTTTGTAGGCTATGTACTGCCGTGAGATCAAATGTCTTTCTGAGGCGCCACCGTCATT
ACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGGATTTGAGGTGG
CTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGTTTGTT
ATCGCCGGCGCAACAGTCCTACATTTGTTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGTTAAACTCCAACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCT
TGGCTTTGTACTAATATTGTTAGCTCTCACACCCCTGACGCTATTCTCTCCAACTCTGCTCGG
TGACCCAGAAAATTTCACCCCGGCAAACCCCCTGGTTACCCCACCACACATCCAACCGGAG
TGGTACTTCTTATTTGCTTACGCCATTCTACGCTCTATCCCAAACAAGCTGGGAGGGGTCCT
AGCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACG
AGGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATCAT
TTTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGT
CGGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCTCGCGGGCTGAGCCGAAAATAAA
GCATTGAAATGAGCTTGCCCTAGTAGCTTAG

IC4

ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT
TCCAACTCCATCTAATATTTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGAT
TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTT
TTCGTCAGTAACACACATCTGTCGAGACGTTAACTATGGCTGACTCATCCGAAACATACATG
CTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATG
GGTCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCCTACTCCTTTTGGTCATAATAA
CAGCCTTTGTAGGCTATGTACTGCCGTGAGGTCAAATGTCTTTCTGAGGCGCCACCGTCATT
ACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGGATTTGAGGTGG
CTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGTTTGTT
ATCGCCGGCGCAACAGTCCTACATTTGTTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGTTAAACTCCAACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCT
TGGCTTTGTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGG
TGACCCAGAAAATTTCACCCCGGCAAACCCCCTGGTTACCCCACCACACATCCAACCGGAG
TGGTACTTCTTATTTTCTTACGCCATTCTACGCTCTATCCCAAACAAGCTGGGAGGGGTCCTA
GCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGA
GGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATCATT
TTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTC
GGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCTCGCGGGCTGAGCCGAAAATAAAG
CATTGAAATGAGCTTGCCCTAGTAGCTTAG

IC5

ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT
TCCAACTCCATCTAATATTTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGAT
TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTT
TTCGTCAGTAACACACATCTGTCGAGACGTTAACTATGGCTGACTCATCCGAAACATACATG
CTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATG
GGTCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCCTACTCCTTTTGGTCATAATAA
CAGCCTTTGTAGGCTATGTACTGCCGTGAGGTCAAATGTCTTTCTGAGGCGCCACCGTCATT
ACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGGATTTGAGGTGG

CTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGTTTGTT
ATCGCCGGCGCAACAGTCCTACATTTGTTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGTTAAACTCCAACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCT
TGGCTTTGTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGG
TGACCCAGAAAATTTCACCCCGGCAAACCCCCTGGTTACCCCACCACACATCCAACCGGAG
TGGTACTTCTTATTTGCTTACGCCATTCTACGCTCTATCCCAAACAAGCTGGGAGGGGTCCT
AGCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACG
AGGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATCAT
TTTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGT
CGGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCTCGCGGGCTGAGCCGAAAATAAA
GCATTGAAATGAGCTTGCCCTAGTAGCTTAG
CC1
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT
TCCAACTCCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGAT
TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTT
TTCGTCAGTAACACACATCTGTCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATG
CTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATG
GGTCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCCTACTCCTTTTGGTCATAATAA
CAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTCTGAGGCGCCACCGTCATT
ACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGGATTTGAGGTGG
CTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGTTTGTT
ATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAAACCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCT
TGGCTTTGTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGG
TGACCCAGAAAATTTTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAG
TGGTACTTCTTATTTGCTTACGCCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTA
GCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGA
GGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATTATT
TTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTC
GGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCTCGCGGGCTGAGCCGAAAATAAAG
CATTGAAATGAGCTTGCCCTAGTAGCTTAG
CC2
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT
TCCAACTCCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGAT
TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTT
TTCGTCAGTAACACACATCTGTCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATG
CTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATG
GGTCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCCTACTCCTTTTGGTCATAATAA
CAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTCTGAGGCGCCACCGTCATT
ACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGGATTTGAGGTGG
CTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGTTTGTT
ATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCT
TGGCTTTGTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGG
TGACCCAGAAAATTTTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAG
TGGTACTTCTTATTTGCTTACGCCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTA
GCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGA
GGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATTATT
TTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTC
GGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCTCGCGGGCTGAGCCGAAAATAAAG
CATTGAAATGAGCTTGCCCTAGTAGCTTAG
CC6
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT

TCCAACTCCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGAT
TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTT
TTCGTCAGTAACACACATCTGTCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATG
CTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATG
GGTCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCCTACTCCTTTTGGTCATAATAA
CAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTCTGAGGCGCCACCGTCATT
ACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGGATTTGAGGTGG
CTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGTTTGTT
ATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCT
TGGCTTTGTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGG
TGACCCAGAAAATTTTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAG
TGGTACTTCTTATTTGCTTACGCCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTA
GCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGA
GGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATTATT
TTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTC
GGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCTCGCGGGCTGAGCCGAAAATAAAG
CATTGAAATGAGCTTGCCCTAGTAGCTTAG
CC9
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT
TCCAACTCCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGAT
TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTT
TTCGTCAGTAACACACATCTGTCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATG
CTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATG
GGTCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCCTACTCCTTTTGGTCATAATAA
CAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTCTGAGGCGCCACCGTCATT
ACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGGATTTGAGGTGG
CTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGTTTGTT
ATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAAACCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCT
TGGCTTTGTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGG
TGACCCAGAAAATTTTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAG
TGGTACTTCTTATTTGCTTACGCCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTA
GCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGA
GGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATTATT
TTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTC
GGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCTCGCGGGCTGAGCCGAAAATAAAG
CATTGAAATGAGCTTGCCCTAGTAGCTTAG
PC1
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT
TCCAACTCCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGAT
TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTT
TTCGTCAGTAACACACATCTGTCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATG
CTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATG
GGTCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCCTACTCCTTTTGGTCATAATAA
CAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTCTGAGGCGCCACCGTCATT
ACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGGATTTGAGGTGG
CTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGTTTGTT
ATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAAACCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCT
TGGCTTTGTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGG
TGACCCAGAAAATTTTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAG
TGGTACTTCTTATTTGCTTACGCCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTA

GCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGA
GGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATTATT
TTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTC
GGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCTCGCGGGCTGAGCCGAAAATAAAG
CATTGAAATGAGCTTGCCCTAGTAGCTTAG
PC4
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT
TCCAACTCCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGAT
TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTT
TTCGTCAGTAACACACATCTGTCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATG
CTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATG
GGTCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCCTACTCCTTTTGGTCATAATAA
CAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTCTGAGGCGCCACCGTCATT
ACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGGATTTGAGGTGG
CTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGTTTGTT
ATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAAACCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCT
TGGCTTTGTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGG
TGACCCAGAAAATTTTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAG
TGGTACTTCTTATTTGCTTACGCCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTA
GCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGA
GGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATTATT
TTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTC
GGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCTCGCGGGCTGAGCCGAAAATAAAG
CATTGAAATGAGCTTGCCCTAGTAGCTTAG
PC7
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT
TCCAACTCCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGAT
TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTT
TTCGTCAGTAACACACATCTGTCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATG
CTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATG
GGTCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCCTACTCCTTTTGGTCATAATAA
CAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTCTGAGGCGCCACCGTCATT
ACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGGATTTGAGGTGG
CTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGTTTGTT
ATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAAACCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCT
TGGCTTTGTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGG
TGACCCAGAAAATTTTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAG
TGGTACTTCTTATTTGCTTACGCCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTA
GCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGA
GGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATTATT
TTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTC
GGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCTCGCGGGCTGAGCCGAAAATAAAG
CATTGAAATGAGCTTGCCCTAGTAGCTTAG
PC10
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT
TCCAACTCCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGAT
TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTT
TTCGTCAGTAACACACATCTGTCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATG
CTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATG
GGGCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCCTACTCCTTTTGGTCATAATAA
CAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTCTGAGGCGCCACCGTCATT

ACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGGATTTGAGGTGG
CTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGTTTGTT
ATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAAACCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCT
TGGCTTTGTATTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGG
TGACCCAGAAAATTTTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAG
TGGTACTTCTTATTTGCTTACGCCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTA
GCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGA
GGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATTATT
TTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTC
GGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCTCGCGGGCTGAGCCGAAAATAAAG
CATTGAAATGAGCTTGCCCTAGTAGCTTAG
PC11
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT
TCCAACTCCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGAT
TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTT
TTCGTCAGTAACACACATCTGTCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATG
CTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATG
GGTCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCCTACTCCTTTTGGTCATAATAA
CAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTCTGAGGCGCCACCGTCATT
ACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGGATTTGAGGTGG
CTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGTTTGTT
ATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAAACCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCT
TGGCTTTGTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGG
TGACCCAGAAAATTTTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAG
TGGTACTTCTTATTTGCTTACGCCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTA
GCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGA
GGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATTATT
TTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTC
GGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCTCGCGGGCTGAGCCGAAAATAAAG
CATTGAAATGAGCTTGCCCTAGTAGCTTAG
PC12
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT
TCCAACTCCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGAT
TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTT
TTCGTCAGTAACACACATCTGTCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATG
CTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATG
GGTCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCCTACTCCTTTTGGTCATAATAA
CAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTCTGAGGCGCCACCGTCATT
ACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGGATTTGAGGTGG
CTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGTTTGTT
ATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCT
TGGCTTTGTATTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGG
TGACCCAGAAAATTTTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAG
TGGTACTTCTTATTTGCTTACGCCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTA
GCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGA
GGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATTATT
TTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTC
GGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCTCGCGGGCTGAGCCGAAAATAAAG
CATTGAAATGAGCTTGCCCTAGTAGCTTAG

PC13
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT
TCCAACTCCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGAT
TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTT
TTCGTCAGTAACACACATCTGTCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATG
CTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATG
GGTCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCCTACTCCTTTTGGTCATAATAA
CAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTCTGAGGCGCCACCGTCATT
ACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGGATTTGAGGTGG
CTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGTTTGTT
ATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAAACCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCT
TGGCTTTGTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGG
TGACCCAGAAAATTTTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAG
TGGTACTTCTTATTTGCTTACGCCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTA
GCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGA
GGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATTATT
TTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTC
GGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCTCGCGGGCTGAGCCGAAAATAAAG
CATTGAAATGAGCTTGCCCTAGTAGCTTAG
PC14
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT
TCCAACTCCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGAT
TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTT
TTCGTCAGTAACACACATCTGTCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATG
CTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATG
GGTCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCCTACTCCTTTTGGTCATAATAA
CAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTCTGAGGCGCCACCGTCATT
ACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGGATTTGAGGTGG
CTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGTTTGTT
ATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAAACCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCT
TGGCTTTGTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGG
TGACCCAGAAAATTTTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAG
TGGTACTTCTTATTTGCTTACGCCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTA
GCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGA
GGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATTATT
TTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTC
GGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCTCGCGGGCTGAGCCGAAAATAAAG
CATTGAAATGAGCTTGCCCTAGTAGCTTAG
TWIN1
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT
TCCAACTCCATCTAATATTTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCCTGAT
TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATTTCAACTGCCTTT
TCGTCAGTAACACACATCTGTCGAGACGTTAACTATGGCTGACTCATCCGGAACATACATGC
TAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATGG
ATCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCTTACTCCTTTTGGTCATAATAAC
AGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTCTGAGGCGCCACCGTCATTA
CGAATTTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGGATTTGAGGTGGC
TTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGTTTGTTA
TTGCCGGCGCAACGGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGCC
GGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCGTACAAAGATCTCCTT
GGCTTTGTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGGT

GACCCAGAAAATTTCACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGT
GGTACTTCTTATTTGCTTACGCCATTCTACGCTCTATCCCAAACAAGCTGGGAGGGGTCCTA
GCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGA
GGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATCATT
TTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTC
GGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCTCGCGGGCTGAGCCGAAAATAAAG
CATTGAAATGAGCTTGCCCTAGTAGCTTAG
TWIN2
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT
TCCAACTCCATCTAATATTTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGAT
TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATTTCAACTGCCTTT
TCGTCAGTAACACACATCTGTCGAGACGTTAACTATGGCTGACTCATCCGGAACATACATGC
TAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATGG
ATCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCTTACTCCTTTTGGTCATAATAAC
AGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTCTGAGGCGCCACCGTCATTA
CGAATTTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGGATTTGAGGTGGC
TTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGTTTGTTA
TTGCCGGCGCAACGGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGCC
GGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCGTACAAAGATCTCCTT
GGCTTTGTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGGT
GACCCAGAAAATTTCACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGT
GGTACTTCTTATTTGCTTACGCCATTCTACGCTCTATCCCAAACAAGCTGGGAGGGGTCCTA
GCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGA
GGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATCATT
TTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTC
GGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCTCGCGGGCTGAGCCGAAAATAAAG
CATTGAAATGAGCTTGCCCTAGTAGCTTAG
TWIN3
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT
TCCAACTCCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGAT
TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTT
TTCGTCAGTAACACACATCTGTCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATG
CTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATG
GGTCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCCTACTCCTTTTGGTCATAATAA
CAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTCTGAGGCGCCACCGTCATT
ACGAATCTATTATCAACAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGGATTTGAGGTGG
CTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGTTTGTT
ATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGCTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCC
TTGGCTTTGTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCG
GTGACCCAGAAAATTTTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGA
GTGGTACTTCTTATTTGCTTACGCCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCC
TAGCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAAC
GAGGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATTA
TTTTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCG
TCGGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCTCGCGGGCTGAGCCGAAAATAAA
GCATTGAAATGAGCTTGCCCTAGTAGCTTAG
TWIN4
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT
TCCAACTCCATCTAATATTTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGAT
TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATTTCAACTGCCTTT
TCGTCAGTAACACACATCTGTCGAGACGTTAACTATGGCTGACTCATCCGGAACATACATGC
TAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATGG

ATCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCTTACTCCTTTTGGTCATAATAAC
AGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTCTGAGGCGCCACCGTCATTA
CGAATTTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGGATTTGAGGTGGC
TTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGTTTGTTA
TTGCCGGCGCAACGGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGCC
GGGTTAAACTCCGACGCAAATAAGATCTCCTTCCACCCATACTTCTCGTACAAAGATCTCCTT
GGCTTTGTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGGT
GACCCAGAAAATTTCACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGT
GGTACTTCTTATTTGCTTACGCCATTCTACGCTCTATCCCAAACAAGCTGGGAGGGGTCCTA
GCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGA
GGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATCATT
TTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTC
GGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCTCGCGGGCTGAGCCGAAAATAAAG
CATTGAAATGAGCTTGCCCTAGTAGCTTAG
TWIN5
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT
TCCAACTCCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGAT
TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTT
TTCGTCAGTAACACACATCTGTCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATG
CTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATG
GGTCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCCTACTCCTTTTGGTCATAATAA
CAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTCTGAGGCGCCACCGTCATT
ACGAATCTATTATCAACAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGGATTTGAGGTGG
CTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGTTTGTT
ATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGCTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCC
TTGGCTTTGTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCG
GTGACCCAGAAAATTTTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGA
GTGGTACTTCTTATTTGCTTACGCCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCC
TAGCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAAC
GAGGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATTA
TTTTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCG
TCGGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCTCGCGGGCTGAGCCGAAAATAAA
GCATTGAAATGAGCTTGCCCTAGTAGCTTAG
TWIN6
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT
TCCAACTCCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGAT
TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTT
TTCGTCAGTAACACACATCTGTCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATG
CTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATG
GGTCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCCTACTCCTTTTGGTCATAATAA
CAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTCTGAGGCGCCACCGTCATT
ACGAATCTATTATCAACAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGGATTTGAGGTGG
CTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGTTTGTT
ATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGCTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCC
TTGGCTTTGTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCG
GTGACCCAGAAAATTTTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGA
GTGGTACTTCTTATTTGCTTACGCCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCC
TAGCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAAC
GAGGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATTA
TTTTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCG

TCGGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCTCGCGGGCTGAGCCGAAAATAAA
GCATTGAAATGAGCTTGCCCTAGTAGCTTAG
TWIN7
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT
TCCAACTCCATCTAATATTTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGAT
TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATTTCAACTGCCTTT
TCGTCAGTAACACACATCTGTCGAGACGTTAACTATGGCTGACTCATCCGGAACATACATGC
TAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATGG
ATCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCTTACTCCTTTTGGTCATAATAAC
AGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTCTGAGGCGCCACCGTCATTA
CGAATTTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGGATTTGAGGTGGC
TTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGTTTGTTA
TTGCCGGCGCAACGGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGCC
GGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCGTACAAAGATCTCCTT
GGCTTTGTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGGT
GACCCAGAAAATTTCACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGT
GGTACTTCTTATTTGCTTACGCCATTCTACGCTCTATCCCAAACAAGCTGGGAGGGGTCCTA
GCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGA
GGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATCATT
TTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTC
GGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCTCGCGGGCTGAGCCGAAAATAAAG
CATTGAAATGAGCTTGCCCTAGTAGCTTAG
TWIN8
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT
TCCAACTCCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGAT
TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTT
TTCGTCAGTAACACACATCTGTCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATG
CTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATG
GGTCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCCTACTCCTTTTGGTCATAATAA
CAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTCTGAGGCGCCACCGTCATT
ACGAATCTATTATCAACAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGGATTTGAGGTGG
CTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGTTTGTT
ATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGCTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCC
TTGGCTTTGTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCG
GTGACCCAGAAAATTTTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGA
GTGGTACTTCTTATTTGCTTACGCCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCC
TAGCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAAC
GAGGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATTA
TTTTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCG
TCGGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCTCGCGGGCTGAGCCGAAAATAAA
GCATTGAAATGAGCTTGCCCTAGTAGCTTAG
TWIN9
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT
TCCAACTCCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGAT
TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTT
TTCGTCAGTAACACACATCTGTCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATG
CTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATG
GGTCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCCTACTCCTTTTGGTCATAATAA
CAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTCTGAGGCGCCACCGTCATT
ACGAATCTATTATCAACAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGGATTTGAGGTGG
CTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGTTTGTT
ATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC

CGGGCTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCC
TTGGCTTTGTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCG
GTGACCCAGAAAATTTTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGA
GTGGTACTTCTTATTTGCTTACGCCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCC
TAGCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAAC
GAGGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATTA
TTTTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCG
TCGGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCTCGCGGGCTGAGCCGAAAATAAA
GCATTGAAATGAGCTTGCCCTAGTAGCTTAG
TWIN10
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT
TCCAACTCCATCTAATATTTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGAT
TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTT
TTCGTCAGTAACACACATCTGTCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATG
CTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATG
GGTCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCCTACTCCTTTTGGTCATAATAA
CAGCCTTTGTAGGCTATGTACTACCATGGGGTCAAATATCTTTCTGAGGCGCCACCGTCATT
ACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAATGGATTTGAGGTGG
CTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTATTTCCGTTTGTT
ATCGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCT
TGGCTTTGTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGG
TGACCCAGAAAATTTTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAG
TGGTACTTCTTATTTGCTTACGCCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTA
GCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGA
GGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATTATT
TTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTC
GGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCTCGCGGGCTGAGCCGAAAATAAAG
CATTGAAATGAGCTTGCCCTAGTAGCTTAG
TWIN11
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT
TCCAACTCCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGAT
TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTT
TTCGTCAGTAACACACATCTGTCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATG
CTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATG
GGTCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCCTACTCCTTTTGGTCATAATAA
CAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTCTGAGGCGCCACCGTCATT
ACGAATCTATTATCAACAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGGATTTGAGGTGG
CTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGTTTGTT
ATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGCTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCC
TTGGCTTTGTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCG
GTGACCCAGAAAATTTTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGA
GTGGTACTTCTTATTTGCTTACGCCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCC
TAGCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAAC
GAGGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATTA
TTTTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCG
TCGGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCTCGCGGGCTGAGCCGAAAATAAA
GCATTGAAATGAGCTTGCCCTAGTAGCTTAG
TWIN12
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT
TCCAACTCCATCTAATATTTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGAT
TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATTTCAACTGCCTTT

TCGTCAGTAACACACATCTGTCGAGACGTTAACTATGGCTGACTCATCCGGAACATACATGC
TAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATGG
ATCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCTTACTCCTTTTGGTCATAATAAC
AGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTCTGAGGCGCCACCGTCATTA
CGAATTTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGGATTTGAGGTGGC
TTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGTTTGTTA
TTGCCGGCGCAACGGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGCC
GGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCGTACAAAGATCTCCTT
GGCTTTGTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGGT
GACCCAGAAAATTTCACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGT
GGTACTTCTTATTTGCTTACGCCATTCTACGCTCTATCCCAAACAAGCTGGGAGGGGTCCTA
GCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGA
GGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATCATT
TTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTC
GGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCTCGCGGGCTGAGCCGAAAATAAAG
CATTGAAATGAGCTTGCCCTAGTAGCTTAG
CAJ1
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT
TCCAACTCCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGAT
TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTT
TTCGTCAGTAACACACATCTGTCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATG
CTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATG
GGTCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCCTACTCCTTTTGGTCATAATAA
CAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTCTGAGGCGCCACCGTCATT
ACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGGATTTGAGGTGG
CTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGTTTGTT
ATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCT
TGGCTTTGTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGG
TGACCCAGAAAATTTTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAG
TGGTACTTCTTATTTGCTTACGCCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTA
GCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGA
GGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATTATT
TTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTC
GGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCTCGCGGGCTGAGCCGAAAATAAAG
CATTGAAATGAGCTTGCCCTAGTAGCTTAG
CAJ2
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT
TCCAACTCCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGAT
TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTT
TTCGTCAGTAACACACATCTGTCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATG
CTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATG
GGTCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCCTACTCCTTTTGGTCATAATAA
CAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTCTGAGGCGCCACCGTCATT
ACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGGATTTGAGGTGG
CTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGTTTGTT
ATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCT
TGGCTTTGTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGG
TGACCCAGAAAATTTTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAG
TGGTACTTCTTATTTGCTTACGCCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTA
GCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGA
GGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATTATT

TTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTC
GGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCTCGCGGGCTGAGCCGAAAATAAAG
CATTGAAATGAGCTTGCCCTAGTAGCTTAG
CAJ3
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT
TCCAACTCCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGAT
TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTT
TTCGTCAGTAACACACATCTGTCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATG
CTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATG
GGTCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCCTACTCCTTTTGGTCATAATAA
CAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTCTGAGGCGCCACCGTCATT
ACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGGATTTGAGGTGG
CTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGTTTGTT
ATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCT
TGGCTTTGTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGG
TGACCCAGAAAATTTTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAG
TGGTACTTCTTATTTGCTTACGCCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTA
GCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGA
GGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATTATT
TTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTC
GGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCTCGCGGGCTGAGCCGAAAATAAAG
CATTGAAATGAGCTTGCCCTAGTAGCTTAG
CAJ4
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT
TCCAACTCCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGAT
TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTT
TTCGTCAGTAACACACATCTGTCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATG
CTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATG
GGTCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCCTACTCCTTTTGGTCATAATAA
CAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTCTGAGGCGCCACCGTCATT
ACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGGATTTGAGGTGG
CTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGTTTGTT
ATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCT
TGGCTTTGTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGG
TGACCCAGAAAATTTTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAG
TGGTACTTCTTATTTGCTTACGCCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTA
GCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGA
GGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATTATT
TTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTC
GGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCTCGCGGGCTGAGCCGAAAATAAAG
CATTGAAATGAGCTTGCCCTAGTAGCTTAG
CAJ5
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT
TCCAACTCCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGAT
TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTT
TTCGTCAGTAACACACATCTGTCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATG
CTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATG
GGTCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCCTACTCCTTTTGGTCATAATAA
CAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTCTGAGGCGCCACCGTCATT
ACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGGATTTGAGGTGG
CTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGTTTGTT

ATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCT
TGGCTTTGTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGG
TGACCCAGAAAATTTTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAG
TGGTACTTCTTATTTGCTTACGCCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTA
GCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGA
GGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATTATT
TTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTC
GGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCTCGCGGGCTGAGCCGAAAATAAAG
CATTGAAATGAGCTTGCCCTAGTAGCTTAG
CAJ6
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT
TCCAACTCCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGAT
TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTT
TTCGTCAGTAACACACATCTGTCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATG
CTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATG
GGTCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCCTACTCCTTTTGGTCATAATAA
CAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTCTGAGGCGCCACCGTCATT
ACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGGATTTGAGGTGG
CTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGTTTGTT
ATTGCCGACGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCT
TGGCTTTGTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGG
TGACCCAGAAAATTTTACCCCGGCAAACCCCCTGGTTACCCCACCACACATCCAACCGGAGT
GGTACTTCTTATTTGCTTACGCCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTA
GCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGA
GGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATTATT
TTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTC
GGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCTCGCGGGCTGAGCCGAAAATAAAG
CATTGAAATGAGCTTGCCCTAGTAGCTTAG
CAJ8
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT
TCCAACTCCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGAT
TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTT
TTCGTCAGTAACACACATCTGTCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATG
CTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATG
GGTCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCCTACTCCTTTTGGTCATAATAA
CAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTCTGAGGCGCCACCGTCATT
ACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGGATTTGAGGTGG
CTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGTTTGTT
ATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCT
TGGCTTTGTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGG
TGACCCAGAAAATTTTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAG
TGGTACTTCTTATTTGCTTACGCCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTA
GCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGA
GGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATTATT
TTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTC
GGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCTCGCGGGCTGAGCCGAAAATAAAG
CATTGAAATGAGCTTGCCCTAGTAGCTTAG
CAJ9
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT
TCCAACTCCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGAT

TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTT
TTCGTCAGTAACACACATCTGTCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATG
CTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATG
GGTCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCCTACTCCTTTTGGTCATAATAA
CAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTCTGAGGCGCCACCGTCATT
ACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGGATTTGAGGTGG
CTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGTTTGTT
ATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCT
TGGCTTTGTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGG
TGACCCAGAAAATTTTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAG
TGGTACTTCTTATTTGCTTACGCCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTA
GCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGA
GGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATTATT
TTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTC
GGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCTCGCGGGCTGAGCCGAAAATAAAG
CATTGAAATGAGCTTGCCCTAGTAGCTTAG
CAJ10
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT
TCCAACTCCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGAT
TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTT
TTCGTCAGTAACACACATCTGTCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATG
CTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATG
GGTCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCCTACTCCTTTTGGTCATAATAA
CAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTCTGAGGCGCCACCGTCATT
ACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGGATTTGAGGTGG
CTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGTTTGTT
ATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCT
TGGCTTTGTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGG
TGACCCAGAAAATTTTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAG
TGGTACTTCTTATTTGCTTACGCCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTA
GCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGA
GGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATTATT
TTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTC
GGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCTCGCGGGCTGAGCCGAAAATAAAG
CATTGAAATGAGCTTGCCCTAGTAGCTTAG
CAJ11
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT
TCCAACTCCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGAT
TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTT
TTCGTCAGTAACACACATCTGTCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATG
CTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATG
GGTCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCCTACTCCTTTTGGTCATAATAA
CAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTCTGAGGCGCCACCGTCATT
ACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGGATTTGAGGTGG
CTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGTTTGTT
ATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCT
TGGCTTTGTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGG
TGACCCAGAAAATTTTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAG
TGGTACTTCTTATTTGCTTACGCCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTA
GCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGA

GGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATTATT
TTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTC
GGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCTCGCGGGCTGAGCCGAAAATAAAG
CATTGAAATGAGCTTGCCCTAGTAGCTTAG
CAJ12
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT
TCCAACTCCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGAT
TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTT
TTCGTCAGTAACACACATCTGTCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATG
CTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATG
GGTCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCCTACTCCTTTTGGTCATAATAA
CAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTCTGAGGCGCCACCGTCATT
ACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGGATTTGAGGTGG
CTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGTTTGTT
ATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCT
TGGCTTTGTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGG
TGACCCAGAAAATTTTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAG
TGGTACTTCTTATTTGCTTACGCCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTA
GCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGA
GGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATTATT
TTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTC
GGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCTCGCGGGCTGAGCCGAAAATAAAG
CATTGAAATGAGCTTGCCCTAGTAGCTTAG
CAJ14
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT
TCCAACTCCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGAT
TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTT
TTCGTCAGTAACACACATCTGTCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATG
CTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATG
GGTCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCCTACTCCTTTTGGTCATAATAA
CAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTCTGAGGCGCCACCGTCATT
ACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGGATTTGAGGTGG
CTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGTTTGTT
ATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCATCCATACTTCTCATACAAAGATCTCCT
TGGCTTTGTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGG
TGACCCAGAAAATTTTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAG
TGGTACTTCTTATTTGCTTACGCCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTA
GCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGA
GGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATTATT
TTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTC
GGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCTCGCGGGCTGAGCCGAAAATAAAG
CATTGAAATGAGCTTGCCCTAGTAGCTTAG
CAJ15
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT
TCCAACTCCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGAT
TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTT
TTCGTCAGTAACACACATCTGTCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATG
CTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATG
GGTCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCCTACTCCTTTTGGTCATAATAA
CAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTCTGAGGCGCCACCGTCATT
ACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGGATTTGAGGTGG

CTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGTTTGTT
ATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCT
TGGCTTTGTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGG
TGACCCAGAAAATTTTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAG
TGGTACTTCTTATTTGCTTACGCCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTA
GCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGA
GGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATTATT
TTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTC
GGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCTCGCGGGCTGAGCCGAAAATAAAG
CATTGAAATGAGCTTGCCCTAGTAGCTTAG
LCK3
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT
TCCAACTCCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGAT
TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTT
TTCGTCAGTAACACACATCTGTCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATG
CTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATG
GGTCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCCTACTCCTTTTGGTCATAATAA
CAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTCTGAGGCGCCACCGTCATT
ACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGGATTTGAGGTGG
CTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGTTTGTT
ATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCT
TGGCTTTGTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGG
TGACCCAGAAAATTTTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAG
TGGTACTTCTTATTTGCTTACGCCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTA
GCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGA
GGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATTATT
TTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTC
GGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCTCGCGGGCTGAGCCGAAAATAAAG
CATTGAAATGAGCTTGCCCTAGTAGCTTAG
LCK4
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT
TCCAACTCCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGAT
TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTT
TTCGTCAGTAACACACATCTGTCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATG
CTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATG
GGTCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCCTACTCCTTTTGGTCATAATAA
CAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTCTGAGGCGCCACCGTCATT
ACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGGATTTGAGGTGG
CTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGTTTGTT
ATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCT
TGGCTTTGTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACACTGCTCGG
TGACCCAGAAAATTTTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAG
TGGTACTTCTTATTTGCTTACGCCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTA
GCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGA
GGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATTATT
TTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTC
GGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCTCGCGGGCTGAGCCGAAAATAAAG
CATTGAAATGAGCTTGCCCTAGTAGCTTAG
LCK5
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT

TCCAACTCCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGAT
TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTT
TTCGTCAGTAACACACATCTGTCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATG
CTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATG
GGTCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCCTACTCCTTTTGGTCATAATAA
CAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTCTGAGGCGCCACCGTCATT
ACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGGATTTGAGGTGG
CTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGTTTGTT
ATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCT
TGGCTTTGTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGG
TGACCCAGAAAATTTTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAG
TGGTACTTCTTATTTGCTTACGCCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTA
GCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGA
GGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATTATT
TTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTC
GGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCTCGCGGGCTGAGCCGAAAATAAAG
CATTGAAATGAGCTTGCCCTAGTAGCTTAG
LCK6
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT
TCCAACTCCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGAT
TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTT
TTCGTCAGTAACACACATCTGTCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATG
CTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATG
GGTCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCCTACTCCTTTTGGTCATAATAA
CAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTCTGAGGCGCCACCGTCATT
ACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGGATTTGAGGTGG
CTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGTTTGTT
ATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCT
TGGCTTTGTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGG
TGACCCAGAAAATTTTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAG
TGGTACTTCTTATTTGCTTACGCCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTA
GCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGA
GGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATTATT
TTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTC
GGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCTCGCGGGCTGAGCCGAAAATAAAG
CATTGAAATGAGCTTGCCCTAGTAGCTTAG
LCK8
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT
TCCAACTCCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGAT
TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTT
TTCGTCAGTAACACACATCTGTCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATG
CTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATG
GGTCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCCTACTCCTTTTGGTCATAATAA
CAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTCTGAGGCGCCACCGTCATT
ACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGGATTTGAGGTGG
CTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGTTTGTT
ATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCT
TGGCTTTGTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGG
TGACCCAGAAAATTTTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAG
TGGTACTTCTTATTTGCTTACGCCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTA

GCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGA
GGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATTATT
TTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTC
GGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCTCGCGGGCTGAGCCGAAAATAAAG
CATTGAAATGAGCTTGCCCTAGTAGCTTAG
LCK9
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT
TCCAACTCCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGAT
TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTT
TTCGTCAGTAACACACATCTGTCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATG
CTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATG
GGTCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTTCTACTCCTTTTAGTCATAATAA
CAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTCTGAGGCGCCACCGTCATT
ACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGGATTTGAGGTGG
CTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGTTTGTT
ATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCT
TGGCTTTGTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGG
TGACCCAGAAAATTTTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAG
TGATACTTCTTATTTGCTTACGCCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTA
GCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGA
GGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATTATT
TTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTC
GGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCTCGCGGGCTGAGCCGAAAATAAAG
CATTGAAATGAGCTTGCCCTAGTAGCTTAG
LCK10
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT
TCCAACTCCATCTAATATCTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGAT
TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATCTCAACTGCCTT
TTCGTCAGTAACACACATCTGTCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATG
CTAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATG
GGTCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCCTACTCCTTTTGGTCATAATAA
CAGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTCTGAGGCGCCACCGTCATT
ACGAATCTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGGATTTGAGGTGG
CTTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGTTTGTT
ATTGCCGGCGCAACAGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGC
CGGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCATACAAAGATCTCCT
TGGCTTTGTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGG
TGACCCAGAAAATTTTACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAG
TGGTACTTCTTATTTGCTTACGCCATTCTACGCTCTATCCCAAACAAGCTAGGAGGGGTCCTA
GCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGA
GGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATTATT
TTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTC
GGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCTCGCGGGCTGAGCCGAAAATAAAG
CATTGAAATGAGCTTGCCCTAGTAGCTTAG
CATC1
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT
TCCAACTCCATCTAATATTTCAGTACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGAT
TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATTTCAACTGCCTTT
TCGTCAGTAACACACATCTGTCGAGACGTTAATTATGGCTGACTCATCCGGAACATACATGC
TAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATGG
ATCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCTTACTCCTTTTGGTCATAATAAC
AGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTCTGAGGCGCCACCGTCATTA

CGAATTTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGGATTTGAGGTGGC
TTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGTTTGTTA
TTGCCGGCGCAACGGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGCC
GGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCGTACAAAGATCTCCTT
GGCTTTGTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGGT
GACCCAGAAAATTTCACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGT
GGTACTTCTTATTTGCTTACGCCATTCTACGCTCTATCCCAAACAAGCTGGGAGGGGTCCTA
GCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGA
GGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATCATT
TTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTC
GGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCTCGCGGGCTGAGCCGAAAATAAAG
CATTGAAATGAGCTTGCCCTAGTAGCTTAG
CATC2
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT
TCCAACTCCATCTAATATTTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGAT
TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATTTCAACTGCCTTT
TCGTCAGTAACACACATCTGTCGAGACGTTAACTATGGCTGACTCATCCGGAACATACATGC
TAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATGG
ATCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCTTACTCCTTTTGGTCATAATAAC
AGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTCTGAGGCGCCACCGTCATTA
CGAATTTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGGATTTGAGGTGGC
TTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGTTTGTTA
TTGCCGGCGCAACGGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGCC
GGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCGTACAAAGATCTCCTT
GGCTTTGTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGGT
GACCCAGAAAATTTCACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGT
GGTACTTCTTATTTGCTTACGCCATTCTACGCTCTATCCCAAACAAGCTGGGAGGGGTCCTA
GCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGA
GGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATCATT
TTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTC
GGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCTCGCGGGCTGAGCCGAAAATAAAG
CATTGAAATGAGCTTGCCCTAGTAGCTTAG
CATC3
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT
TCCAACTCCATCTAATATTTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGAT
TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATTTCAACTGCCTTT
TCGTCAGTAACACACATCTGTCGAGACGTTAACTATGGCTGACTCATCCGGAACATACATGC
TAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATGG
ATCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCTTACTCCTTTTGGTCATAATAAC
AGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTCTGAGGCGCCACCGTCATTA
CGAATTTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGGATTTGAGGTGGC
TTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGTTTGTTA
TTGCCGGCGCAACGGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGCC
GGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCGTACAAAGATCTCCTT
GGCTTTGTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTTTCTCCAACTCTGCTCGGT
GACCCAGAAAATTTCACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGT
GGTACTTCTTATTTGCTTACGCCATTCTACGCTCTATCCCAAACAAGCTGGGAGGGGTCCTA
GCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGA
GGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATCATT
TTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTC
GGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCTCGCGGGCTGAGCCGAAAATAAAG
CATTGAAATGAGCTTGCCCTAGTAGCTTAG

EFSGR1
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT
TCCAACTCCATCTAATATTTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGAT
TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATTTCAACTGCCTTT
TCGTCAGTAACACACATCTGTCGAGACGTTAACTATGGCTGACTCATCCGGAACATACATGC
TAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATGG
ATCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCTTACTCCTTTTGGTCATAATAAC
AGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTCTGAGGCGCCACCGTCATTA
CGAATTTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGGATTTGAGGTGGC
TTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGTTTGTTA
TTGCCGGCGCAACGGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGCC
GGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCGTACAAAGATCTCCTT
GGCTTTGTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGGT
GACCCAGAAAATTTCACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGT
GGTACTTCTTATTTGCTTACGCCATTCTACGCTCTATCCCAAACAAGCTGGGAGGGGTCCTA
GCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGA
GGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATCATT
TTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTC
GGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCTCGCGGGCTGAGCCGAAAATAAAG
CATTGAAATGAGCTTGCCCTAGTAGCTTAG
EFSGR2
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT
TCCAACTCCATCTAATATTTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGAT
TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATTTCAACTGCCTTT
TCGTCAGTAACACACATCTGTCGAGACGTTAACTATGGCTGACTCATCCGGAACATACATGC
TAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATGG
ATCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCTTACTCCTTTTGGTCATAATAAC
AGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTCTGAGGCGCCACCGTCATTA
CGAATTTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGGATTTGAGGTGGC
TTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGTTTTGTT
ATGCCGGCGCAACGGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGCC
GGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCGTACAAAGATCTCCTT
GGCTTTGTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGGT
GACCCAGAAAATTTCACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGT
GGTACTTCTTATTTGCTTACGCCATTCTACGCTCTATCCCAAACAAGCTGGGAGGGGTCCTA
GCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGA
GGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATCATT
TTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTC
GGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCTCGCGGGCTGAGCCGAAAATAAAG
CATTGAAATGAGCTTGCCCTAGTAGCTTAG
NFSGR1
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT
TCCAACTCCATCTAATATTTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGAT
TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATTTCAACTGCCTTT
TCGTCAGTAACACACATCTGTCGAGACGTTAACTATGGCTGACTCATCCGGAACATACATGC
TAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATGG
ATCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCTTACTCCTTTTGGTCATAATAAC
AGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTCTGAGGCGCCACCGTCATTA
CGAATTTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGGATTTGAGGTGGC
TTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGTTTGTTA
TTGCCGGCGCAACGGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGCC
GGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCGTACAAAGATCTCCTT
GGCTTTGTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGGT

```
GACCCAGAAAATTTCACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGT
GGTACTTCTTATTTGCTTACGCCATTCTACGCTCTATCCCAAACAAGCTGGGAGGGGTCCTA
GCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGA
GGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATCATT
TTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTC
GGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCTCGCGGGCTGAGCCGAAAATAAAG
CATTGAAATGAGCTTGCCCTAGTAGCTTAG
NFSGR2
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT
TCCAACTCCATCTAATATTTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGAT
TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATTTCAACTGCCTTT
TCGTCAGTAACACACATCTGTCGAGACGTTAACTATGGCTGACTCATCCGGAACATACATGC
TAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATGG
ATCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCTTACTCCTTTTGGTCATAATAAC
AGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTCTGAGGCGCCACCGTCATTA
CGAATTTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGGATTTGAGGTGGC
TTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGTTTGTTA
TTGCCGGCGCAACGGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGCC
GGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCGTACAAAGATCTCCTT
GGCTTTGTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGGT
GACCCAGAAAATTTCACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGT
GGTACTTCTTATTTGCTTACGCCATTCTACGCTCTATCCCAAACAAGCTGGGAGGGGTCCTA
GCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGA
GGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATCATT
TTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTC
GGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCTCGCGGGCTGAGCCGAAAATAAAG
CATTGAAATGAGCTTGCCCTAGTAGCTTAG
WFSGR1
ATGGCAAGCCTACGAAAAACCCATCCACTAATAAAGATTGCTAACGATGCACTAGTCGACCT
TCCAACTCCATCTAATATTTCAGCACTGTGGAACTTCGGATCCCTCCTAGGATTATGCTTGAT
TACTCAAATCCTGACAGGACTATTCCTAGCCATACATTATACTTCCGATATTTCAACTGCCTTT
TCGTCAGTAACACACATCTGTCGAGACGTTAACTATGGCTGACTCATCCGGAACATACATGC
TAACGGAGCATCATTCTTCTTTATCTGTATTTATATACACATTGCCCGCGGACTGTACTATGG
ATCATACCTCTATAAAGAGACCTGAACTATTGGCGTGGTCTTACTCCTTTTGGTCATAATAAC
AGCCTTTGTAGGCTATGTACTGCCGTGGGGTCAAATGTCTTTCTGAGGCGCCACCGTCATTA
CGAATTTATTATCAGCAGTCCCCTACATAGGAGACGCTCTTGTCCAGTGGATTTGAGGTGGC
TTTTCAGTAGATAACGCAACGTTAACACGATTTTTCGCCTTCCACTTCCTGTTTCCGTTTGTTA
TTGCCGGCGCAACGGTCCTACATTTGCTATTTTTACACGAAACGGGATCAAACAACCCTGCC
GGGTTAAACTCCGACGCAGATAAGATCTCCTTCCACCCATACTTCTCGTACAAAGATCTCCTT
GGCTTTGTACTAATATTGTTAGCTCTCACATCCCTGACGCTATTCTCTCCAACTCTGCTCGGT
GACCCAGAAAATTTCACCCCGGCAAACCCCCTGGTTACCCCGCCACACATCCAACCGGAGT
GGTACTTCTTATTTGCTTACGCCATTCTACGCTCTATCCCAAACAAGCTGGGAGGGGTCCTA
GCACTATTATTTAGCATTCTAGTACTATTAGTAGTGCCAGTTTTACATACCTCGAAGCAACGA
GGACTAACTTTCCGTCCAATCACCCAGTTTTTATTCTGAACCTTAGTGGCAGACATAATCATT
TTAACATGAATTGGAGGCATGCCTGTAGAACACCCATACATCATCATTGGCCAAGTCGCGTC
GGTGCTGTACTTCGCACTCTTCCTCCTTCTCGCCCCCTCGCGGGCTGAGCCGAAAATAAAG
CATTGAAATGAGCTTGCCCTAGTAGCTTAG
;
End;
```

References

Akimoto, S., Itoi, S., Sezaki, K., Borsa, P., and Watabe, S. (2006) Identification of alfonsino, *Beryx mollis* and *B. splendens* collected in Japan, based on the mitochondrial cyctochrome b gene and their comparison with those collected in New Caledonia. *Fisheries Science* 72:202-207.

Alves, M.J., Coelho, H., Collares-Pereira, M.J., and Coelho, M.M. (2001). Mitochondrial DNA Variation in the Highly Endangered cyprinid fish *Anaecypris hispanica*: importance for conservation. *Heredity* 87:463-473.

Avise, J.C. (2004). Molecular Markers, Natural History and Evolution. 2nd ed. Sinauer Associates. Sunderland, Ma.

Avise, J.C. (20000. Phylogeography: The History and Formation of Species. Harvard University Press. Cambridge, MA.

Avise, John C. (1998). The history and perview of phylogeography: a personal reflection. *Molecular Ecology* 7:371-379.

Avise, J.C., Ball, R.M., Arnold, J. (1988). Current versus historical population sizes in vertebrate species with high gene flow: A comparison based on mitochondrial DNA lineages and inbreeding theory for neutral mutations. *Mol. Biol. Evol.* 5:331-344.

Banarescu, P. and Coad, B.W. (1991). Cyprinids of Eurasia. *Fish and Fish*. 3:667. Chapman and Hall, London.

Barrett, Paul. (2003) *Rhinichthys osculus* (Santa Ana Speckled dace) photo. San Gabriel River, Los Angeles Basin, California: USFWS.

Barroch, A. (1991). Prosite: A Dictionary of Sites and Patterns in Proteins. *Nucleic Acids Res.* 19:2241-2245.

Belle, E.M.S., Piganeau G., Gardner, M., and Eyre-Walker, A. (2005). An Investigation of the Variation in the Transition bias among various animal mitochondrial DNA. *Gene* 355:58-66.

Bernatchez, L., and C.C. Wilson. (1998). Comparative phylogeography of Nearctic and Palearctic fishes. *Mol. Ecol*. 7: 431 – 452.

Birky, C.W., Jr., T. Maruyama, and P. Fuerst. (1983). An approach to population and evolutionary theory for genes in mitochondria and chloroplasts, and some results. *Genetics* 103: 513 – 527.

Bond, C.E, Rexstad, E., Hughes, R.M. (1988). Habitat use of 25 common species of Oregon freshwater fishes. *Northwest Sci*. 62:223-232.

Boykin, L.M., Shatters, R.G., Rosell, R.C. McKenzie, C.L., Bagnal, R.A., DeBarro, P. and Frohlich, D.R. (2007). Global relationships *Bemisia tabaci* (Hemiptera: Aleyrodidae) revealed using Bayesian analysis of mitochondrial COI DNA sequences. *Mol. Phylogenet. Evol.* 44(3):1306-1319.

Briolay, J., Galtier, N., Brito, R.M., and Bouvet, Y. (1998). Molecular Phylogeny of Cyprinidae Inferred from cytochrome b DNA Sequences. *Molecular Phylogenetics and Evolution* 9(1):100-108.

Britten, Roy J. (1986). Rates of DNA Sequence Evolution Differ between Taxonomic Groups. *Science, New Series* 231(4744):1393-1398.

Broughton, R.E., and Reneau, P.C. (2006). Spatial Covaration of Mutation and Nonsynonymous Substitution Rates in Vertebrate Mitochondrial Genomes. *Mol. Biol. Evol.* 23(8):1516-1524.

Brown, J.R., Beckenbach, K., Beckenbach, A.T., and Smith, M.J. (1996). Length Variation, Heteroplasmy and Sequence Divergence in Mitochondrial DNA of Four Species of Sturgeon (*Acipenser)*. *Genetics* 142:525-535.

Brown, W.M., George, M., Wilson A. C. (1979). Rapid Evolution of Animal Mitochondrial DNA. *Proc. Natl. Acad. Sci.* 76:4:1967-1971.

Brown, W.M., Prager, E.M., Wang, A., Wilson, A. C. (1982). Mitochondrial DNA sequences of primates: Tempo and mode of evolution. *J. Mol. Evol.* 18:225-39.

Calsbeek, R., Thompson, J.N., and Richardson, J.E.. (2003). Patterns of molecular evolution and diversification in a biodiversity hotspot: the California floristic province. *Molecular Ecology* 12: 1021 – 1029.

Cantatore, P., M. N. Gadaleta, M. Roberti, C. Saccone, and A. C. Wilson. (1987). Duplication and remoulding of tRNA genes during the evolutionary rearrangement of mitochondrial genomes. *Nature* 329: 853 – 855.

Cecchini, G. (2003). Function and Structure of Complex II of the Respiration Chain*. Annu. Rev. Biochem*. 72:77-109.

Chatzimanolis, S. and M.S. Caterino. (2007). Toward a better understanding of the “Transverse Range Break;” lineage diversification in southern California. *Evolution* 61 (9): 2127 – 2147.

Chou, P. Y., & Fasman, G. D. (1978). Empirical predictions of protein conformation. *Annual review of biochemistry*, *47*(1), 251-276.

Clement, M., Posada, D. C. K. A., & Crandall, K. A. (2000). TCS: a computer program to estimate gene genealogies. *Molecular ecology*, *9*(10), 1657-1659.

Conner, J.K. and Hartl, D.L. (2004). A Primer of Ecological Genetics. Sinauer Associates, Sutherland, MA:304.

Cooke, G.M., N.L. Chao, and L.B. Beheragaray. (2012). Divergent natural selection with gene flow along major environmental gradients in Amazonia: insights from genome scans, population genetics, and phylogeography of the the characin fish Triportheus albus. *Molec. Ecol.* 21: 2410 – 2427.

Cornelius, R.H. (19690. The Systematics and Zoogeography of *Rhinichthys Osculus* (Girard) in Southern California. Unpublished masters thesis, California State University, Fullerton.

Costedoat, C., Chappaz, R., Barascud, B., Guillard, O., and Giles, A. (2006). Heterogeneous Colonization pattern of European Cyprinids, as highlighted by the dace complex (Teleostei: Cyprinidae). *Molecular Phylogenetics and Evolution* 41:127-148.

Culver, G.B., and C.L. Hubbs. (1917). The fishes of the Santa Ana system of streams in Southern California. *Lorquinia* 1 (11): 82 – 83.

Cunha, C., Mesquita, N., Dowling, T.E., Gilles, A., and Coelho, M. M. (2002). Phylogenetic Relationships of Eurasian and American Cyprinids using Cytochrome b Sequences. *Journal of Fish Biology* 61:929-944.

Dawid, I.B. and Blackler, A.W. (1972). Maternal and cytoplasmic inheritance of mitochondrial DNA in *Xenopus*. *Dev. Biol.* 29:152-161.

Deacon, J.E., and Williams, J.E. (1984). Annotated list of the fishes of Nevada. *Proc. Biol. Soc. Wash*. 97:103-118.

Esposti, D., De Vries, M.S., Crimi, M., Ghelli, A., Paternello, T., and Meyer, A. (1993). Mitochondrial cytochrome b: evolution and structure of the protein. *Biochimica et Biophysica Acta* 1143:243-271.

Excoffier, L., & Lischer, H. E. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular ecology resources*, *10*(3), 564-567.

Excoffier, L., P.E. Smouse, and J.M. Quattro. (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* 131: 479 – 491.

Fu, Y.X. (1997). Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* 147, 915-925.

Futuyma, D.J. 2005. Evolution. Sinauer Associates, Sunderland, MA.

Graur, Dan and Li, Wen-Hsiung (2000). Fundamentals of Molecular Evolution 2nd ed. Sinauer Associates, Inc. Sunderland, MA.

Griffith, R. and Tavare, S. (1997). Computational methods for the coalescent. Pp. 165-182 in Progress in Population genetics and Human Evolution, P. Donnelly and S. Tavare (eds.) Springer-Verlag, New York.

Hall, B.G. (20040. Phylogenetic trees made easy: a how-to manual. Sunderland, MA: Sinauer Associates.

Hayashi, J.I., Tagashira, Y., and Yoshida, M.C. (1985). Absence of extensive recombination between inter- and intraspacies mitochondrial DNA in mammalian cells. *Exp. Cell. Res.* 160:387-395.

Herbots, H.M. (1997). The structured coalescent. Pp. 231-235 in Progress in Human Population Genetics and Human Evolution, P. Donnelly and S. Tavare (eds.). Springer-Verlag, New York.

Hickman, James C. (1993). The Jepson Manual Higher Plants of California. University of California Press.

Hillis, D.M., Mable, B.K., Larson, Davis, S.K., and Zimmer, E.A. (1996). Nucleic Acids IV: Sequencing and Cloning. In Molecular Systematics. D.M. Hillis, C. Moritz, and B.K. Mable (Eds). Sinauer Associates, Inc. Sunderland, MA. pp. 321-381.

Hoekzema, K. and Sidlauskas, B.L. (2014). Molecular phylogenetics and microsatellite analysis reveal cryptic species of speckled dace (Cyprinidae: *Rhinichthys osculus*) in Oregon’s Great Basin. *Mol. Biol. Evol.* 77:238-250.

Hofmann, K., Bucher, P., Falquet, L., & Bairoch, A. (1999). The PROSITE database, its status in 1999. *Nucleic Acids Research*, *27*(1), 215-219.

Holmquist, R. (1983). Transitions and transversions in evolutionary descent: an approach to understanding. *Journal of Molecular Evolution*, *19*(2), 134-144.

Hubbs, C.L., Miller, R.R., and Hubbs. L.C. (1974). Hydrographic history and relict fishes of the north-central great basin. *Memoirs of the California Academy of Sciences*, Vol. 7. San Francisco, CA.

Hubbs, C.L., and R.R. Miller. (1948). Correlation between fish distribution and the hydrographic history in the desert basins of western United States. In: The Great Basin, with emphasis on Glacial and post-glacial times. Pp 17 – 144. *Bull. Univ. Utah 38, Biol. Ser.* 10(7).

Hudson, R.R., M. Slatkin, and W.P. Maddison. (1992). Estimation of levels of gene flow from DNA sequence data. *Genetics* 132: 583 – 589.

Irwin, D.M., Kocher, T.D., and Wilson, A.C. (1991). Evolution of the Cytochrome b Gene of Mammals. *J Mol Evol* 32:128-144.

Johansen, S., Guddal, P.H., and Johansen, T. (1990). Organization of the mitochondrial genome of Atlantic cod, *Gadus morhua*. *Nucleic Acids Research* 18(3):411-419.

Johns, G.C.,and Avise, J.C. (1998). A Comparative Summary of Genetic Distances in the Vertebrates from the Mitochondrial Cytochrome b Gene. *Mol. Biol. Evol.* 15(11);1481-1490.

Jordan, D.S., and Evermann B.W. (1896). The Fishes of North and Middle America: a descriptive catalogue of the species of fish-like vertebrates found in the waters of North America, North of the Isthmus of Panama. *Bull. U.S. Natl. Mus.* 47:1-1240.

Källberg, M., Wang, H., Wang, S., Peng, J., Wang, Z., Lu, H., & Xu, J. (2012). Template-based protein structure modeling using the RaptorX web server. *Nature protocols*, *7*(8), 1511-1522.

Ketmaier, Valerio, Bianco, Pier Giorgio, Cabolli, Marina, Krivokapic, Marijan, Caniglia, Romolo, and De Matthaeis, Elivira. (2004). Molecular Phylogeny of Two Lineages of *Leuciscinae* Cyprinids (Telestes and Scardinius) from the peri-mediterranean area based on cytochrome b data. *Molecular Phylogenetics and Evolution.* 32:1061-1071.

Kimbrough, D.L., Smith, D.P., Mahoney, J.B., Moore, T.E., Grove, M., Gastril, R.G., Ortega-Rivera, A., and Fanning, C.M., (2001). Forearc-basin sedimentary response to rapid Late Cretaceous batholith emplacement in the Peninsular Ranges of Southern and Baja California. *Geology* 29(6):491-494.

Kocher, T.D., Thomas, W.K., Meyer, A., Edwards, S.V., Paabo, S., and Villablanca, F.X., (1989). Dynamics of mitochondrial DNA evolution in animals: Amplication and Sequencing with conserved primers. *Proc. Natl. Acad. Sci. USA* 86:6196-6200.

Kotlik, P. and Berrebi, P. (2001). Phylogeography of the barbell (*Barbus barbus*) assessed by mitochondrial DNA variation. *Molecular Ecology* 10:2177-2185.

Krause, J., Fu, Q., Good, J.M., Viola, B., Shunkov, M.V., Derevianko, A.P., and Paabo, S. (2010). The complete mitochondrial DNA genome of an unknown hominin from Southern Siberia. *Nature* 464:894-897.

Kvist, L., Martens, J., Nazarenko, A.A., and Orell, M. (2003). Paternal leakage of mitochondrial DNA in the great tit (*Parus major)*. *Mol. Biol. Evol.* 20:243-247.

LaRivers, I. (1952). A key to Nevada Fishes. *S. Calif. Acad. Sci. Bull.* 51(3): 86 – 102.

Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., & Thompson, J. D. (2007). Clustal W and Clustal X version 2.0. *bioinformatics*, *23*(21), 2947-2948.

Lee, D.S., C.R. Gilbert, C.H. Hocutt, and R.E. Jenkins, D.E. McCallister, J.R. Stauffer Jr. (1980). Atlas of North American freshwater fishes. North Carolina State Museum of Nat. History, Raleigh NC.

Lomolino, M.V., Riddle, B.R., and Brown, J.H. (2006) Biogeography, 3rd ed. Sinauer Associates. Sunderland, MA.

Los Angeles County watershed Management (2008). Watershed Management Plan for the Los Angeles County Watershed.

Lydeard, C., & Roe, K. J. (1997). The phylogenetic utility of the mitochondrial cytochrome b gene for inferring relationships among actinopterygian fishes. *Molecular systematics of fishes*, 285-303.

Ma, Din Pow, Zharkikh, A., Graur, D., VandeBerg, J.L., and Li, Wen-Hsiung Li, (1993). Structure and Evolution of Opossum, Guinea Pig, and Porcupine Cytochrome b Genes. *Journal of Molecular Evolution* 36:327-334.

Maddison, W.P., and D.R. Maddison. (1989). Interactive analysis of phylogeny and character evolution using the computer program MacClade. *Folia Primatol (Basel)* 53: 190 – 202.

McClellan, David. A., and McCracken, Kevin G. (2001). Estimating the Influence of Selection on the Variable Amino Acid Sites of the Cytochrome b *Protein Functional Domains*. 18(6):917-925.

McMillan, W.O., and S.R. Palumbi. (1997). Rapid rate of control-region evolution in pacific butterflyfishes (Chaetodontidae). *J. Mol. Evol.* 45: 473 – 484.

Mesquita, N., Carvalho, G., Shaw, P., Crespo, E., and Coelho, M.M., (2001). River basin-related genetic structuring in an endangered species, *Chondrostoma lusitanicum*, based on mtDNA sequencing and RFLP analysis. *Heredity* 86:253-264.

Miller, R.R. (1946). On relations between fish distribution and Pleistocene hydrography in eastern California and southwestern Nevada, with a map of the Pleistocene waters. *J. Geology. LIV* 1: 43 – 53.

Minckley, W.L., Hendrickson, D.A.,and Bond, C.E. (1986). Geography of Western North American Freshwater Fishes: Description and Relationships to Intracontinental Tectonism. In: Hocutt, C.H., Wiley, E.O. (Eds), The Zoogeography of North American Freshwater Fishes. Wiley, New York, 519-613.

Mitchell, Patrick (2006). Santa Ana River guide: From crest to coast—110 miles along Southern California's largest river system. *Wilderness Press* 1-246.

Moritz, C. (1995). Uses of Molecular Phylogenies for Conservation. Phil. Trans. R. Soc. Lond. B. 349: 113-118.

Moritz, C., C.J. Schneider, and D.B. Wake. (1992). Evolutionary relationships within the *Ensatina eschscholtzii* complex confirm the ring species interpretation. *Systematic Biology* 41: 273 – 291.

Moyle, P.B (2002). Inland Fishes of California. University of California Press. Berkeley, CA.

Nei, M., 2005. Selectionism and Neutralism in Molecular Evolution. *Mol. Biol. Evol.* 22(12):2318-2342.

Nelson, J., (1994). Fishes of the World. Third Edition. New York, NY: John Wiley and Sons.

Oakey, D.D., Douglas, M.E., and Douglas, M.R. (2004). Small Fish in a Large Landscape: Diversification of *Rhinichthys osculus* (Cyprinidae) in Western North America. *Copeia* 2: 207-221.

Peakall, R. O. D., & Smouse, P. E. (2006). GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular ecology notes*, *6*(1), 288-295.

Perdices, A., D. Sayanda, and M.M. Coelho. (20050. Mitochondrial diversity of *Opsariichthys bidens* (Teleostei, Cyprinidae) in three Chinese drainages. *Mol. Phyl. Evol.* 37: 920 – 927.

Pfrender, M.E., Hicks, J., Lynch, M. (2004). Biogeographic patterns and current distribution of molecular-genetic variation among populations of speckled dace, *Rhinichthys osculus* (Girard). *Molecular Phylogenetics and Evolution* 30: 490-502.

Phillipsen, I.C., and A.E. Metcalf. (2009). Phylogeography of a stream-dwelling frog (*Pseudacris cadaverina*) in Southern California. *Molecular Phylogenetics and Evolution* 53(1): 152 – 170.

Posada, D., and K.A. Crandall. (1998). MODELTEST: testing the model of DNA substitution. *Bioinformatics* 14: 817 – 818.

Posada, David and Crandall, Keith A. (1998). Modeltest: Testing the model of DNA Substitution. *Bioinformatics Applications Note* 14(9):817-818.

Prager, E. M., & Wilson, A. C. (1988). Ancient origin of lactalbumin from lysozyme: analysis of DNA and amino acid sequences. *Journal of Molecular Evolution*, *27*(4), 326-335.

Prusak, Beata and Grzybowski, Tomasz. (2004). Non-random base composition in codons of mitochondrial cytochrome b gene in vertebrates. *Acta Biochimica Polonica* 51(4):897-096.

Ronquist, F., Teslenko, M., Van Der Mark, P., Ayres, D. L., Darling, A., Höhna, S. & Huelsenbeck, J. P. (2012). MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic biology*, *61*(3), 539-542.

San Gabriel Mountains Regional Conservancy. (2005). Watershed Management Plan for the San Gabriel River above Whittier Narrows.

Santa Ana Speckled Dace Recovery Project. (2005). United States Forest Service and California Department of Fish and Game Bi-Annual Report. Kerwin Russell, editor. April – October 2005.

Santa Ana River Watershed Project Authority (SAWPA). (2004-2005). Old, Grand Prix, and Padua Fires (October, 2003): Burn Impacts to Water Systems and Resources. Santa Ana River Watershed Area. San Bernardino National Forest, California. Santa Ana Watershed Project Authority. Summary Report. Riverside, California.

Scheffler, Immo E. (1999). “Biogenesis” Mitochondria. Wiley-Liss, Inc,. Publication, 48-92.

Scott, W.B., and Crossman, E.J. (1998). Freshwater Fishes of Canada. Galt House Publications LTD, Oakville, ON, Canada.

Schneider, S. and Excoffier, L. (1999). Estimation of Past Demographic Parameters From the Distribution of Pairwise Differences When the Mutation Rates Vary Among Sites: Application to Human Mitochondrial DNA. *Genetics* 152:1079-1089.

Schoenherr, Allan (1992). A Natural History of California. University of California Press. 1-500.

Seeber, Leonardo, and Sorlien, Christopher (2000). Listric thrusts in the Western Transverse Ranges, California. *GSA Bulletin*, 112(7):1067-1079.

Smith, G.R. and Dowling, T.E. (2008). Correlating hydrographic events and divergence times of Speckled dace (Rhinichthys: Teleostei : Cyprinidae) in the Colorado River drainage. *The Geological Society of America – Special Paper*, 301-315.

Spencer, J.E., Smith, G.R., and Dowling, T.E.. (2008). Middle to late Cenozoic geology, hydrography, and fish evolution in the American southwest. In: Late Cenozoic drainage history of the southwestern Great Basin and lower Colorado River region: geologic and biotic perspectives. (Reheis, M.C, R. Hershler, and D.M. Miller, eds). *The Geological Society of America*, special paper 439. Boulder, Colo. Pp 279 – 299.

Swift, C.C., Hugland, T.R., Ruiz, M., and Fisher, R.N. (1993). The status and distribution of the freshwater fishes of Southern California. Bull. *S. Calif. Acad. Sci.* 92:101-167.

Swofford, D.L., G.J.Olson, P.J. Waddell, and D.M. Hillis. (1996). Phylogenetic Inference. Pp. 407 – 514 in Molecular Systematics (D.M. Hillis, C. Moritz, and B.K. Mable, eds.). Sinauer Assoc., Sunderland, MA.

Swofford, D.L. (2003). PAUP*: Phylogenetic analysis using parsimony (*and other methods). Sinauer Associates, Inc. Sunderland, MA.

Taib, Z. (1997). Branching processes and evolution. Pp. 321-329 in Progress in Human Population Genetics and Human Evolution, P. Donnelly and S. Tavare (eds.). Springer-Verlag, New York.

Tajima, Fumio (1989). Statistical Method for Testing the Neutral Mutation Hypothesis by DNA Polymorphism. *Genetics* 123:585-595.

Tamura K, Stecher G, Peterson D, Filipski, A, and Kumar S. (2013). MEGA6: Molecular Evolutionary Genetics Analysi version 6.0. *Molecular Biology and Evolution* 30:2725-2729.

Tang, Q., H. Liu, R. Mayden, and B. Xiong. (2006). Comparison of evolutionary rates in the mitochondrial DNA cytochrome *b* gene and control region and their implications for phylogeny of the Cobitoidea (Teleostei: Cypriniformes). *Mol. Phylogen. Evol.* 39: 347 – 357.

Templeton, A.R. (1980). The theory of speciation via the founder principle. *Genetics* 94: 1011 – 1038.

Thompson, J.D., T.J. Gibson, F. Plewniak, F. Jeanmougin, and D.G. Higgins. (1997). The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* 25: 4876 – 4882.

Waples, R. S. (1995). Evolutionary significant units and the conservation of biological diversity under the Endangered Species Act. In: Evolution and the Aquatic Ecosystem: Defining Unique Units in Population Conservation. American Fisheries Society Symposium 17. Bethesda, Maryland.

Warrick J.A. and Milliman, J.D. (2003)(. Hyperpycnal sediment discharge from semiarid Couthern California Rivers: Implication for coastal sediment budgets. *Geology* 31(9):781-784.

Weir, B.S. and Cockerham, C.C. (1984). Estimating F-Statistics for Analysis of Population Structure. *Evolution* 38:1358-1370.

Wilson, A.C., R.L. Cann, S.M. Carr, M. George, U. B. Gyllensten, K. M. Helm-Bychowski, R. G. Higuchi, S. R. Palumbi, E. M. Prager, R. D. Sage, and M. Stoneking. (1985). Mitochondrial DNA and two perspectives on evolutionary genetics. *Biol. J. Linn. Soc*. 26: 375 – 400.

Woodman, D.A. (1992). Systematic relationships within the cyprinid genus *Rhinichthys*. In: Mayden, R.L. (Ed.), Systematics, Historical Ecology, and North American Freshwater Fishes. *Stanford University Press, Stanford, CA.* pp. 374 – 391.

Wright, S. 1931. Evolution in Mendelian populations. Genetics. 16: 97 – 159.

Zardoya, R., and Doadrio, I. (1999). Molecular Evidences on the Evolutionary and Biogeographical Patterns of European Cyprinids. *J Mol Evol* 49:227-237.

Zardoya, R., Economidis, P.S. and Doadrio, I., (1999). Phylogenetic Relationships of Greek Cyprinidae: Molecular Evidence for at Least Two Origins of the Greek Cyprinid Fauna. *Molecular Phylogenetics and Evolution* 13(1):121-131.

Zardoya, R., Doadrio, I. (1998). Phylogenetic Relationships of Iberian cyprinids: Systematic and Biogeographical Implications. *Proc. R. Soc. Lond. B* 265:1365-1372.

Zhang, A., Huang, L., Shulmeister, Y., Chi, I., Kim, K., Huang, L.W., Croftss, A.R., Berry, E.A., and Kim, S.H. (1998). Electron transfer by domain movement in cytochrome bc_1. *Nature* 392:677-684.