CALIFORNIA STATE UNIVERSITY, NORTHRIDGE

The Genetic Diversity and Population Genetic Structure of the Shovelnose Guitarfish (*Pseudobatos productus*) from Southern California to Baja California Sur

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Biology

By

Alexandra Elyse Meyer

December 2019

The thesis of Alexandra Elyse Meyer is approved:

_____	_____
Dr. Christopher L. Chabot	Date
_____	_____
Dr. Jeanne M. Robertson	Date
_____	_____
Dr. Larry G. Allen, Chair	Date

California State University, Northridge

Acknowledgements

This project would not have been possible without the contributions from a number of amazing scientists and supporters. I would like to thank Dr. Larry Allen for taking me on even though he is not an "elasmo guy" and believes that doing genetic research is akin to practicing hoodoo and black magic. His passion for knowledge and love of fishes has been truly inspirational. I would like to thank Dr. Chris Chabot, who has definitely been in the trenches with me on this project. As my "Lab Dad" he has helped me through countless hours of questions and panic and many years of frustrating lab work. Without him, my project would have never gotten off the ground. I would like to thank Dr. Jeanne Robertson, for giving me my first introduction to molecular ecology and being incredibly patient with someone who had never held a micropipette or done any sort of molecular benchwork. She has always been there for me whenever I needed help or encouragement. I would like to thank Dr. Mike Franklin for his many years of patience and support as both my boss and mentor. Without him I would have likely not been employed as a GA or instructional staff, and my many years supporting him in lab courses has helped shape my desire and ability to teach.

My project would not have been completed without the help of numerous individuals who either provided me with my samples or helped me collect them. Dr. Andy Nosal and Dr. Dan Cartamil provided me with my samples from Santa Catalina Island, La Jolla, and Guererro Negro. Dr. Felipe Galván-Magaña provided me with my samples from Bahía Tortugas. Dennis Dunn, Denis Mahaffy, and the crew of the R/V Yellowfin were integral to my collection of specimens from San Pedro.

To Calvin Won, my older brother from an Asian mother, thank you for always being there for me as well as teaching me how to fish. I'm sorry I'm not better at it. To Justin Cordova, thank you for helping me make the connections I need in the shark world and for making me laugh when I need it most. Your ability to get yourself into and out of trouble is awe inspiring. To Emily Ladin, thank you for sharing your boisterous personality and trying to keep our lab from falling apart. Special thanks go to Nathaniel Chard, for helping with DNA sequencing and extraction as well as being my first guinea pig when it came to teaching bench skills.

Without numerous funding sources, my project would never have been possible. Thank you PADI Foundation, the International Women's Fishing Association, and CSUN Thesis support for your contributions. I would also like to thank all my family and friends who came out in droves to donate to my crowd source funding. Without your love and support I would have been both miserable and unable to buy my reagents.

To Jenni Henley, my best friend, thank you for all the years you've put up with me. You've been with me through thick and thin, through frustrating years of high school and exciting years of college. We've laughed a lot, cried a little, and managed to hold each other together despite it all. Maybe one day we will finally finish our massive backlog of comic books.

Lastly, I must thank my parents for letting me move back home and helping make my graduate career as least financially painful as possible. You have always encouraged me to follow my passions, even when it was something insane like wanting to become an Orca trainer or chasing sharks. Your love of science and natural curiosity has inspired my own. I promise that one day I will clean my room.

Dedication

To my parents, who taught me to always ask questions and never give up.

Table of Contents

List of Figures

List of Tables

Abstract

The Genetic Diversity and Population Genetic Structure of the Shovelnose Guitarfish (*Pseudobatos productus*) from Southern California to Baja California Sur

By

Alexandra Elyse Meyer

Master of Science in Biology

The Shovelnose Guitarfish, *Pseudobatos productus* (Ayres 1854), is a shark-like batoid ray that lives in shallow soft-bottom habitats such as bays and estuaries. This species ranges from San Francisco Bay, California to the Gulf of California, Mexico, and is targeted in the artisanal elasmobranch fisheries throughout Baja California and the Gulf of California. In many cases, localized fishing pressure can lead to a loss of genetic diversity, a reduction in effective population size, and local extinction. Genetic assessments of fish stocks are seldom applied to management strategies, though there is strong evidence that such assessments will better inform policy. This study sought to assess the genetic diversity and connectivity of *P. productus* from southern California to Baja California Sur, Mexico. Genetic diversity and connectivity were assessed using the complete mitochondrial control region as well as genomic SNP data produced using the 2b-RAD library preparation and sequencing protocol. Both markers show evidence of high levels of connectivity from southern California to northern Baja California Sur, indicating admixture across international boundaries. Mitochondrial DNA evidence

supports higher levels of population structure while genomic SNP data supports admixture between mainland localities. This may be indicative of sex-specific movement patterns, with higher levels of male-mediated gene flow between localities. Based on these data, *P. productus* should be managed as a bi-national species with sex-based restrictions as males likely contribute more to genetic admixture than females.

Introduction

Genetic Diversity and Exploitation

Similar to the effect of species diversity on the long-term stability of communities, intraspecific genetic diversity is often integral to the temporal stability of a population as well as influence its system function (Agashe, 2009; Prieto *et al*., 2015). Populations lacking genetic diversity are often at risk of collapse or localized extinction due to limited adaptive potential (Willi *et al*., 2006) or the effects of increased inbreeding (Frankham, 1995). In the case of species harvested for human use, this risk is often related to the level of exploitation. In heavily exploited populations, the removal of individuals either by targeted or non-selective means typically results in a loss of genetic diversity. This in turn can lead to a loss of effective population size as observed in species such as the Giant Sea Bass (*Stereolepis gigas*) and Orange Roughy (*Hoplostethus atlanticus*), particularly when breeding aggregations are targeted (Smith *et al*., 1991; Gaffney *et al*., 2007; Chabot *et al*., 2015b).

Despite the relationship between genetic diversity and population stability, seemingly stable fisheries stocks are commonly assessed using generalized models that frequently do not factor in specific life history characteristics and these stocks are often reassessed only when there are significant changes in stock abundance or distribution (Begg *et al*., 1999). Numerous studies have supported the use of genetic stock identification to inform fisheries management policies, however, this is rarely implemented (Ryman, 1991; Shaklee and Bentzen, 1998; Shaklee *et al*., 1999). Additionally, many species span political boundaries, either internationally or at the state level, complicating management policies even further. These “trans-boundary” species

are often managed by unique political entities with different policies despite possibly belonging to the same genetic stock (Chabot *et al*., 2015b; Reiss *et al*., 2009).

Elasmobranchs in general are particularly vulnerable to the effects of this mismatch between genetic stocks and management due to their often-similar life history characteristics. Many elasmobranch life histories are characterized by long lifespans, slow growth rates, late age-of-maturity, long gestation periods that produce large, well-formed young, low fecundity, and aggregating behaviors for the purposes of feeding or breeding (Hoenig and Gruber, 1990; Musick *et al*., 2000). In addition, species that reach a total maximum length over 100 cm in combination with these life-history characteristics appear to be particularly vulnerable to overfishing and population collapse (Frisk *et al*., 2001). The historic over-exploitation of species such as the Soupfin Shark (*Galeorhinus galeus*), Porbeagle (*Lamna nasus*), Basking Shark (*Cetorhinus maximus*), Spiny Dogfish (*Squalus acanthias*), and Barndoor Skate (*Raja laevis*) have led to population collapse and near extinction in some cases (Casey and Myers, 1998; Stevens *et al*., 2000). Of the 1,050 known species of elasmobranchs, the International Union for the Conservation of Nature classifies 204 species in one of three threatened categories (Vulnerable, Endangered, Critically Endangered) and another 115 as Near-Threatened ("IUCN Red List of Threatened Species," 2019). Due to its life-history characteristics, distribution, and fisheries status, population genetic assessment is critical for the conservation and management of the Shovelnose Guitarfish (*Pseudobatos productus*, Ayres 1854).

Distribution and Putative Biogeographic Barriers

Pseudobatos productus ranges from San Francisco Bay, California through the Gulf of California, Mexico, though it is most common south of Point Conception, California (Timmons and Bray, 1997; Love, 2011; Figure 1.). Shovelnose Guitarfish are a coastal benthos-associated species, commonly inhabiting shallow sandy or muddy embayments and estuaries shallower than 13 m, though they have been observed at depths as great as 91 m (Love, 2011). *Pseudobatos productus* mate and pup during the late spring and summer months, with females arriving earlier in the spring and remaining there until fall and males arriving in the summer as mating aggregations form (Villaviencio-Garayzar, 1993; Márquez-Farías, 2007; Love, 2011). Shovelnose Guitarfish are thought to move offshore outside of breeding season, though there is evidence of site fidelity to proposed nursery sites and other nearshore sites from year to year (Dubois, 1981; Villaviencio-Garayzar, 1993; Sandoval-Castillo *et al*., 2004).

Populations of *P. productus* span several predominant phylogeographic barriers along the coast of California and Baja California, including Point Conception, the San Pedro Channel, the San Quintín Upwelling-Zone, Punta Eugenia, and Cabo San Lucas, all which have been shown to have varying degrees of impact on population connectivity of coastal fishes. For example, Point Conception often provides a biogeographic barrier, but not a strong phylogeographic break, between central and southern California, and is believed to be the transition zone between the northern Oregonian Province and southern San Diegan Province (Briggs, 1974; Horn and Allen, 1978; Horn *et al*., 2006). By contrast, Cabo San Lucas often provides both a phylogeographic and biogeographic break between the Gulf of California and the Pacific Coast of Baja California, limiting gene flow between the two regions as seen in species such as the Golden Cownose Ray

(*Rhinoptera steindachneri*) and the Pacific Angel Shark (*Squatina californica*) (Dawson *et al*., 2006; Ramírez-Amaro *et al*., 2017; Sandoval-Castillo and Rocha-Olivares, 2011). Populations of demersal species such as the Round Stingray (*Urobatis halleri*) show restricted gene flow across the San Pedro Channel, separating the southern California Channel Islands from mainland southern California (Plank *et al*., 2010), while numerous other demersal and shallow-water species show varying levels of phylogeographic structure across Punta Eugenia (Dawson, 2001).

Life History

Maximum total length estimates for *Pseudobatos productus* range from 162 cm in females to 114 cm in males (Downton-Hoffmann, 2007; Juaristi Videgaray, 2016) and the species has an estimated longevity of 11 years and a generation time of 9 years (Timmons and Bray, 1997). Location-based estimates of size at sexual maturity range from 57–111 cm in females and 63–110 cm in males, with both sexes maturing at approximately 7 years (Downton-Hoffmann, 2007; Dubois, 1981; Juaristi Videgaray, 2016; Márquez-Farías, 2007; Timmons and Bray, 1997; Villaviencio-Garayzar, 1993). Total gestation time in *P. productus* is approximately 12 months with a 5-month embryonic growth phase. The reproductive cycle of *P. productus* is annual, producing a single litter of pups per year with ovulation soon after parturition (Márquez-Farías, 2007; Romo-Curiel *et al*., 2016). Additionally, there is no evidence supporting sperm storage, indicating that fertilization occurs during the reproductive event (Juaristi Videgaray, 2016). *Pseudobatos productus* are a lecithotrophic viviparous species in which pups are nourished exclusively by the yolk sac with no additional maternal contribution (Romo-Curiel *et al*., 2016). Litter sizes range from 2–18 pups, however the average litter size is 5

pups with larger females producing larger litters (Downton-Hoffmann, 2007; Márquez-Farías, 2007; Villaviencio-Garayzar, 1993). Similar to size at sexual maturity, size at parturition varies by location ranging from 17–23 cm (Márquez-Farías, 2007; Villaviencio-Garayzar, 1993). As with all elasmobranchs, pups are born fully formed and independent of parental care.

Fisheries

Pseudobatos productus is classified as "Near-Threatened" with a declining population trend, likely as the result of fisheries harvest and human impact with the primary threat being targeted elasmobranch fisheries (Farrugia *et al*., 2016). In Californian waters there is no commercial fishery for *P. productus* and individuals are landed as the result of incidental trawling by-catch or from limited harvest by recreational anglers (Farrugia *et al*., 2016). In Mexican waters, both within the Gulf of California and along the Pacific coast, *P. productus* is landed as a primary target of the artisanal ray fishery as well as bycatch in demersal trawls from other fisheries (Farrugia *et al*., 2016). In some locations, gillnets used to catch demersal species are modified with additional nylon lines to increase the entanglement and retention rate of guitarfish (Salazar-Hermoso and Villavicencio-Garayzar, 1999; Márquez-Farías, 2011).
Pseudobatos productus is the most commonly landed species in the artisanal elasmobranch fishery, comprising at least 20% of the total reported catch of any state yearly with recorded landings often numbering in the thousands (Bizzarro *et al*., 2009; Cartamil *et al*., 2011; Ramirez-Amaro *et al*., 2013; Smith *et al*., 2009a). In one extreme example, nearly 46,000 individuals were landed in 1999 as part of the artisanal fishery in Sonora (Bizzarro *et al*., 2009). Although the catch of *P. productus* has a strong seasonal

bias towards the spring and summer months, individuals are landed throughout the year at most locations (Márquez-Farías, 2007).

Unlike, the artisanal shark fishery in the Mexican Pacific which has been documented since the 1930s, the targeted ray fishery is a relatively recent development, with the first documentation of landings in 1986, developing alongside the use of bottom-set gillnets in commercial teleost fisheries (Márquez-Farías, 2002). Despite being a relatively young documented fishery, catches in some locations such as Bahía Almejas, Baja California Sur have steeply declined from the late 1990s likely as a result of intense fishing pressure that frequently removes mature individuals (Farrugia *et al*., 2016). Gillnets are often set at the mouths of bays, channels, and estuaries frequently used by *P. productus* to move inshore during the breeding season, increasing the chance of catching reproductive individuals (Márquez-Farías, 2011; Salazar-Hermoso and Villavicencio-Garayzar, 1999). Large, gravid females are frequently caught during the spring and summer months due to the targeted fishing of these breeding aggregations (Bizzarro *et al*., 2009; Márquez-Farías, 2011; Smith *et al*., 2009a). Catches of *P. productus* often deviate from a 1:1 sex ratio, with higher numbers of females being landed when compared to males, sometimes doubling or tripling the male catch (Bizzarro *et al*., 2009; Cartamil *et al*., 2011; Márquez-Farías, 2011; Ramirez-Amaro *et al*., 2013; Salazar Hermoso and Villavicencio-Garayzar, 1999; Smith *et al*., 2009a). This deviation is often attributed to the difference in size between males and females, with larger individuals being more susceptible to entanglement, as well as the difference in the duration of time spent in shallow-water habitats vulnerable to fishing (Márquez-Farías, 2005; Márquez-Farías, 2011). This removal of mature females is of particular concern as

prior assessments of population connectivity reveal genetic isolation of populations, indicating low female migration between some sites (Sandoval-Castillo *et al*., 2004).

Assessment of Population Genetic Structure

Over the last several decades, the use of mitochondrial DNA, particularly the non-coding mitochondrial control region (mtCR) or "D-loop", has been the marker of choice for phylogeographic studies due to its relatively rapid mutation rate with no genetic recombination between generations (Figure 2). Frequently, mtCR analyses have been combined with neutral nuclear based approaches (e.g., microsatellites) in order to provide more complete assessments of population structure, as nuclear approaches remove the female-specific bias inherent to mitochondrial markers due to bi-parental inheritance (Freeland *et al*., 2011). More recently, with the advent of next generation sequencing technology (NGS), studies now typically incorporate the use of single nucleotide polymorphisms (SNPs) discovered through genome-wide screening for the assessment of population differentiation. These SNPs are often produced through various restriction-site associated DNA sequencing methods (RADseq) as they reliably generate large numbers of small fragments that are distributed throughout the genome and are independent of each other (Andrews and Luikart, 2014; Puritz *et al*., 2014; Rowe *et al*., 2011).

The population genetic structure of *Pseudobatos productus* was previously assessed using Polymerase Chain Reaction-restriction fragment length polymorphism (PCR-RFLP) analysis of the mitochondrial control region (mtCR) (Sandoval-Castillo *et al*., 2004). The study assessed the degree of differentiation between populations located on the Pacific coast of Baja California Sur and those located in the northern Gulf of California, finding indications of a deep phylogenetic separation across the Baja

California Peninsula and the possible existence of a new cryptic species (Sandoval-Castillo *et al*., 2004). However, there has been no assessment of the population connectivity of *P. productus* along the Pacific coast including the waters within the United States. Therefore, the aim of this study was to assess the population structure of *P. productus* from southern California to Baja California Sur using the analysis of SNPs generated via the 2b-RAD protocol (Wang *et al*., 2012) as well as complete mitochondrial control region sequences.

Materials and Methods

Sample Collection

Tissue samples were collected from the fins and muscle tissues of *P. productus* caught from southern California to Baja California Sur. Tissues were collected using sterile techniques and stored in 95% ethanol in a -20 ºC freezer for long term storage. In southern California, specimens were collected using hook and line, beach seines, and shallow water otter trawls. In Baja California Sur, muscle tissue was collected from individuals landed as part of the artisanal elasmobranch fishery. Tissue was collected from sub-adult and adult individuals of both sexes. Samples were collected from 6 localities from Los Angeles to Baja California Sur from 2012–2018, with the majority of samples collected in the spring and summer. At least 6 individuals were collected from each locality, with the exception of Dockweiler Beach, California where only 1 individual was collected (Santa Catalina Island, CA, n=13; Long Beach, CA, n=6; La Jolla, CA, n=10; Guerrero Negro, BCS, n=15; Bahía Tortugas, BCS, n=27) (Table 1; Figure 1). The sample from Dockweiler beach was grouped with individuals from Long Beach due to geographic proximity to form a general locality hereafter referred to as San Pedro that is comprised of samples from the Los Angeles/Long Beach region.

Total genomic DNA was extracted from each sample using the DNEasy Blood and Tissue Kit® (Qiagen, Valencia, CA) following the manufacturer's protocol. DNA concentration (ng/µl) and 260/280 nm absorbance ratio were quantified using a NanoDrop™ 2000 Spectrophotometer (Thermo Scientific) and overall quality was assessed visually using 1% agarose gels.

Mitochondrial DNA Amplification and Sequencing

The currently unpublished mitochondrial control region (mtCR) of *P. productus* was amplified using species-specific primers designed by Sandoval-Castillo *et al.* (2004) for use in PCR-RFLP analysis: FWD: CbRhino1163 5'- CYT ACT TCT CAT TAT TCC TNA TCC TCC TAC C -3' and REV: 12Srev326 5'- ACT CGT ATA ACC GCG GTG GCT -3'. Oligonucleotides were ordered from Integrated DNA Technologies (San Diego, CA) and reconstituted to a concentration of 100µM using Buffer AE (10mM Tris-Cl and 0.5mM EDTA; Qiagen, Valencia CA). Amplification was performed in a 10µl volume using 5µl HotStarTaq Master Mix® (Qiagen, Valencia, CA), 1µl each of 2µM forward and reverse primers, 2.1µl nuclease free water (NFW), 0.4µl of 10µM Bovine Serum Albumin (BSA, New England Biolabs, Ipswich, MA), and 1.5µl DNA template (10nM-140nM). Polymerase Chain Reaction (PCR) was performed in an Applied Biosystems GeneAmp PCR System 9700® thermocycler with an initial denaturation step of 94 °C for 15 minutes, followed by 35 cycles of 94 °C for 15 seconds, 63 °C for 120 seconds, 72 °C for 45 seconds, and a final extension of 72 °C for 10 minutes. A subsample of each PCR product was visualized using a 1% TAE agarose gel and electrophoresed at 90-95 volts for 45 minutes to confirm the presence of the expected PCR product with a length of ~1500bp.

Successfully amplified PCR products were sequenced in two directions using the PCR primer CbRhino1163 as the forward primer and two nested sequencing primers (REV 12SrRNA: 12Srev195 5'- AGG CTA GGA CCA AAC CTG CT -3' and REV Internal-mtCR: 12sDLoop101 5'- GGG TTT TTC GAG GAT ACC GTG -3'). Both nested sequencing primers were designed using Geneious 10.2.3 (https://www.geneious.com) and Primer3 (Koressaar and Remm, 2007; Untergasser *et*

al., 2012). These two primers were designed and optimized from preliminary samples to reduce sequence coverage within the flanking conserved 12SrRNA region in order to increase read coverage within the mtCR as well as provide sequence overlap for the forward and reverse reads. PCR products were purified using ExoSap-IT Express (Applied Biosystems, Thermo Fisher Scientific, Waltham, MA) and sent to Laragen, Inc. (Culver City, CA) for Sanger Sequencing.

Mitochondrial DNA Sequence Verification

All sequences for each individual were aligned to each other and to publicly available mitochondrial sequences from 7 closely related species (Rhinopristiformes: *Rhinobatos hynnicephalus* GenBank Accession # (GB#)KF534708.1; *Rhinobatos schlegellii* GB#KJ140136.1; *Rhychobatus australiae* GB#KU746824.1; *Pristis clavata* GB#KF381507.1; *Pristis pectinata* GB#KP400584.1; *Anoxypristis cuspidata* GB#KP233202.1; *Rhina ancylostoma* GB#KU721837.1; Figure 2) using the De Novo Assemble tool in Geneious 10.2.3 under the High Sensetivity setting and validated by eye. A multiple sequence alignment (MSA) of all validated mitochondrial control region sequences from *P. productus* was generated in Geneious using the Multiple Alignment tool (Geneious 10.2.3) under the default settings. From the MSA produced by Geneious, single nucleotide polymorphisms (SNPs) were verified by eye and each variant sequence was given a haplotype identifier based on location (e.g., CAT 1–13, DW 1, LB 1–6, etc.) and were uploaded to GenBank (NCBI).

Mitochondrial DNA Analysis

jModelTest 2 (Darriba *et al*., 2012) was used to determine the best fit evolutionary model for the mtCR. It was found that the generalized time-reversal model (GTR+1+G) was the best fit model, however, this model is not available in many of the programs used for subsequent analyses. As a result, the Tamura-Nei model was used as it was also identified as good model and is available in all of the software used for downstream analyses.

Arlequin 3.5 (Excoffier and Lischer, 2010) was used to calculate standard molecular diversity indices including relative nucleotide composition, number of polymorphic loci, haplotype diversity (h), nucleotide diversity (π), mean pairwise nucleotide differences among populations, and the population mutation parameter θ_S (which is based on the number of segregating sites, sample size, and θ for a sample of non-recombining DNA). Female effective population size, N_{ef}, was estimated for each population using the mutation rates of 0.0067 mutations per million years (Schultz *et al*., 2008) and 0.024 mutations per million years (Hoelzel *et al*., 2006) based on minimum and maximum mutation rates observed in elasmobranchs. N_{ef} was calculated using the equation $\theta_S = 2N_{ef}u$ where $u = 2\mu k$, with μ being the mutation rate and k being the number of nucleotides in the sequence. Arlequin 3.5 (Excoffier and Lischer, 2010) was also used to calculate two estimators of recent population expansion, Tajima's D (Tajima, 1989) and Fu's FS statistic (Fu, 1997).

The estimation of genetic divergence between localities was calculated using multiple pairwise and hierarchical AMOVAs (Analysis of MOlecular VAriance; Excoffier *et al*., 1992) to generate the diversity metrics Φ_{ST} and Φ_{RT}, which are analogous to Wright's F-statistics, using GenAlEx v6.503 (Peakall and Smouse, 2006, 2012). To

determine the significance of hierarchical and pairwise AMOVAs after multiple tests, the false discovery rate of Benjamini and Yekutieli (Benjamini and Yekutieli, 2001; Narum, 2006) was used to determine the significant value of alpha based on the total number of pairwise comparisons.

A Median-Joining haplotype network was generated using PopART to visualize the relatedness of haplotypes (Leigh and Bryant, 2015). The genetic population assignment of individuals was assessed using STRUCTURE 2.3.4 (Falush *et al*., 2003, 2007; Hubisz *et al*., 2009; Pritchard *et al*., 2000). STRUCTURE uses posterior probabilities generated from estimated allele frequencies to assign individual membership to the most likely cluster, "*K*". Population assignment was tested for *K* values of 1–6 with 20 independent iterations of 1,000,000 Markov-chain Monte Carlo repetitions per *K* and a burn-in of 10% using the admixture model. CLUMPAK (Kopelman *et al*., 2015) was used to determine the most likely value of *K* according to the Evanno method (Evanno *et al*., 2005) and the natural-log probability of *K* (Falush *et al*., 2003, 2007; Hubisz *et al*., 2009; Pritchard *et al*., 2000).

Genomic DNA Preparation

Given that high-quality genomic DNA (gDNA) is required for successful amplification using the 2b-RAD genotyping technique (Wang *et al*., 2012), total gDNA concentration was assessed using a Qubit 3.0 Flourometer (Invitrogen, Thermo Fisher Scientific, Waltham, MA) prior to library preparation. Since 2b-RAD genotyping works best at low volumes of high concentration gDNA, all samples were concentrated using ethanol precipitation based on methods described by Green and Sambrook (2016). gDNA was then reconstituted using Tris-HCl (pH 8.5) to a final concentration of ~100ng/µl.

Genomic DNA Library Generation and Sequencing

Genomic sequencing libraries were generated using the 2b-RAD genotyping technique (Wang *et al.*, 2012) with minor modifications based on updated protocols provided by Eli Meyer (2018). A complete list of oligonucleotides used in the generation of 2b-RAD libraries for sequencing can be found in Appendix C. The 2b-RAD technique consists of four steps: digestion of gDNA, ligation of double stranded library adaptors, PCR attachment of dual-indexing barcodes and amplification of gDNA, and fragment size selection to purify generated libraries (Figure 3). The digestion mix consisted of 1.2µl of 10X Buffer R (Thermo Scientific cat#:ER1801 2U/µl *AlfI* restriction enzyme kit, Waltham, MA) 0.8µl of 150µM SAM (cat#:ER1801), 0.5µl of *AlfI* restriction enzyme (cat#:ER1801), and 1.5µl of nuclease free water (NFW) per sample. A total of 8µl of concentrated gDNA was added to the 4µl digestion master mix, creating a total reaction volume of 12µl that was incubated overnight at 37 ºC.

Due to *P. productus* having a large genome of ~4Gb (Hinegardner, 1976; Schwartz and Maddock, 1986), the feasibility of generating libraries capable of sequencing the entire genome of each sample with adequate depth was low. As a result, a 1/4 reduction scheme approach based on adapters from Meyer (2018) was used to generate libraries with adequate coverage of the genome, albeit reduced, with increased sequencing depth for each fragment of the genome. Immediately prior to ligation, two double stranded ligation adaptors were assembled. 2µL of 100µM 5ILL-NR oligo (Appendix C) were combined with 2µL of 100µM anti-ILL oligo (Appendix C) and 96µL NFW to create Adaptor 1. The process was repeated to create Adaptor 2, in which 2µL of 100µM 3ILL-NR oligo (Appendix C) were combined with 2µL of 100µM anti-ILL oligo (Appendix C) and 96µL NFW. Both adaptor mixes were incubated at room

temperature for 10 minutes. Following adapter assembly, a ligation mix consisting of 1µl of 10mM ATP (New England Biolabs, Ipswich, MA, #P0756L), 4µl of 10X T4 ligase buffer, 1µl of T4 DNA ligase (New England Biolabs, Ipswich, MA, #M0202L), 0.5µl of 20mg/ml Bovine Serum Albumin (BSA, New England Biolabs, Ipswich, MA), 5µl of Adaptor 1, 5µl of Adaptor 2, and 23.5µl of NFW was combined with 12µl of digested gDNA and incubated overnight at 16 ºC.

Amplification of the gDNA/adapter complex and the attachment of sequencing barcodes was performed in a single PCR consisting of two reaction mixes. The first, a PCR mix consisting of 22µl of NFW, 2µl of 10mM dNTP (New England Biolabs, Ipswich, MA, #N0446S), 2µl of 10µM ILL-Lib1 (Appendix C), 2µl of 10µM ILL-Lib2 (Appendix C), 1µl of Q5 DNA Polymerase, 20µl of 5X Q5 Buffer (New England Biolabs, Ipswich, MA, #M0491L), and 1µl of 20mg/ml Bovine Serum Albumin (BSA, New England Biolabs, Ipswich, MA). The second, a gDNA/adapter complex-dual-indexing mix consisting of 5µl of an Illumina i7 index (Appendix C) and 5µl of an Illumina i5 index (Appendix C) that were combined with 40µl of each gDNA/adapter complex in unique sample combinations (Table 2). For the final PCR mix, 50µl of the PCR master mix was combined with 50µl of the gDNA/adapter complex-dual-indexing mix for a total reaction volume of 100µl. PCR was performed in an Applied Biosystems GeneAmp PCR System 9700® thermocycler with an initial denaturation step of 98 °C for 30 seconds, followed by 25 cycles of 98 °C for 5 seconds, 60 °C for 20 seconds, 72 °C for 10 seconds, and a final extension of 72 °C for 2 minutes.

After PCR, 2% Tris-Borate-EDTA low melt agarose gels were used to visualize libraries followed by their extraction from the gels. For each library, 20µl of gel loading dye was combined with 100µl of the PCR product and electrophoresed on low voltage

(95 V) for 2 hours until bands were adequately separated. Gels were visualized on a blue-light table (Clare Chemical Research, Dark Reader DR-89X) to minimize damage to libraries and a target fragment size of ~166bp was cut from the gel. Libraries were then extracted from gels using a MinElute Gel Extraction Kit (Qiagen, Valencia, CA) according to the manufacturer's protocols.

Prior to the assembly of a pooled sequencing library, the quantity of DNA for each extracted library was determined by Qubit. Based on the quantity of DNA in each sample, concentrations were adjusted to 10nM and then individual libraries were pooled together at equimolar concentrations into a single sequencing library. Pooled libraries were then sequenced on an Illumina HiSeq 4000 Next Generation Sequencer using a 50-cycle single read protocol at the University of Chicago Genomics Facility. Sequences were demultiplexed and sequencing adaptors were trimmed following run completion by the University of Chicago prior to any data analysis.

Genomic DNA Read Processing

Raw sequence reads were processed using a custom Perl script pipeline from the Meyer Lab (Meyer, 2018) with minor modification to optimize the number of reads and their quality as well as reduce artifacts within the data inherent to high-throughput sequencing. A complete list of all scripts and their code can be found in Appendix D. Given that *AlfI* restriction endonucleases generate predictable 36bp-long fragments, the raw 50-character reads were trimmed to 36 characters to remove the additional 14 characters produced as an artifact of the sequencing protocol (TruncateFastq.pl). Low quality reads were removed using bbduk.sh (Bushnell, 2019) with parameters set to remove reads with an average $Q<30$. As *P. productus* does not have a known reference

genome, a *de novo* reference genome was generated by creating a cluster-derived reference (CDR) based on representative sequences from the raw reads (BuildRef.pl). Reads were aligned to the CDR using the SHRiMP software package (David *et al*., 2011). Poor reads that did not properly map to the CDR and ambiguous reads that mapped to multiple locations within the CDR were filtered using SAMFilter.pl. Genotypes were determined from nucleotide frequencies using CallGenotypes.pl. Monomorphic loci were filtered using PolyFilter.pl, leaving only loci at which 2 or more genotypes were observed through all samples. MDFilter.pl was used to filter low-coverage loci and retain loci that were genotyped in all 72 individuals to minimize missing data in downstream analyses. Repetitive sequences were filtered using RepTagFilter.pl to remove bias inherent to high-throughput genotyping-by-sequencing methods due to the inability to correctly map repetitive sequences to corresponding unique loci in the *de novo* CDR genome. This filtering process removes loci that produce an unusually high number of SNPs, leaving only loci in which 2 SNPs are present. Putative linked SNPs that occur on the same *AlfI* fragment were filtered using OneSNPPerTag.pl to remove downstream bias generated by the analyses of linked traits.

Genomic DNA Analysis

The estimation of genetic divergence between localities was calculated using multiple pairwise and hierarchical AMOVAs (Analysis of MOlecular VAriance; Excoffier *et al*., 1992) to generate the diversity metrics F_{ST} and F_{RT} in Arlequin 3.5 (Excoffier and Lischer, 2010). To determine the significance of hierarchical and pairwise AMOVAs after multiple tests, the false discovery rate of Benjamini and Yekutieli (Benjamini and Yekutieli, 2001; Narum, 2006) was used to determine the significant

value of alpha based on the total number of pairwise comparisons. The genetic population assignment of individuals was assessed using STRUCTURE 2.3.4 (Falush *et al*., 2003, 2007; Hubisz *et al*., 2009; Pritchard *et al*., 2000). Population assignment was tested for *K* values of 1–6 with 20 independent iterations of 1,000,000 Markov-chain Monte Carlo repetitions per *K* and a burn-in of 10% using the admixture model. CLUMPAK (Kopelman *et al*., 2015) was used to determine the most likely value of *K* populations according the Evanno method (Evanno *et al*., 2005) and the natural-log probability of *K* (Falush *et al*., 2003, 2007; Hubisz *et al*., 2009; Pritchard *et al*., 2000). Genetic clustering was visualized using a principal-components analysis (PCA) using *adegenet* in R (Jombart, 2008; Jombart and Ahmed, 2011).

Results

Mitochondrial DNA Analyses

Complete mitochondrial control regions were amplified from 70 individuals with each individual possessing a unique haplotype (Table 3; GenBank Accession #MN783979 – MN784048). Due to overall poor DNA quality, complete mitochondrial control regions were successfully amplified and sequenced from only 13 individuals from Guerrero Negro. The complete control region of *P. productus* is between 1,316bp – 1,320bp in length with an average composition of 22.51% cytosine, 32.63% thymine, 32.26% adenine, and 12.60% guanine. The complete control region of *P. productus* is roughly 200bp longer than the known complete control regions of other members of the family Rhinobatidae. Haplotypes of *P. productus* contained a total of 95 polymorphic sites, which is approximately 7.2% of the entire control region (Table 3). Of the 95 polymorphic sites, there were 40 transitions, 41 transversions, and 20 indels across all samples. The majority of polymorphic sites occurred between 700bp – 1,150bp in the second hypervariable region towards the $tRNA^{Phe}$ and 12sRNA genes. Overall haplotype diversity was high ($h = 1.00$) while total nucleotide diversity was low ($\pi = 0.011 \pm 0.006$) (Table 4). Nucleotide diversity across all localities was consistently low ($\pi = 0.007 \pm 0.004 – 0.017 \pm 0.009$) with the lowest being observed at Santa Catalina Island, San Pedro, and Bahía Tortugas and the highest at Guerrero Negro (Table 4). Tests of population expansion via Tajima's D statistic were non-significant across all populations, however both Santa Catalina Island and Bahía Tortugas had significantly negative Fu's F_S statistics (Table 4). Maximum effective female population sizes of *P. productus* ranged between 157,665 (based on a mutation rate of 0.024 mutations per

million years) and 564,770 (based on a mutation rate of 0.0067 mutations per million years) (Table 4).

Significant mitochondrial DNA genetic structure was observed with 23.18% ($p = 0.0001$) of global variation occurring among sampled localities (Table 5). Significant pairwise assessments of Φ_{ST} ranged in value between 0.1225 and 0.4859 with Santa Catalina Island and Bahía Tortugas significantly diverging from all other sampling localities ($p \leq 0.017$, Table 6). Hierarchical AMOVAs testing divergence across putative phylogeographic barriers revealed three putative mtDNA regions with two putative barriers to geneflow ($\Phi_{RT} = 0.1626$, $p = 0.0001$, Table 7). The first barrier separates Santa Catalina Island from Long Beach and the second barrier separates Guerrero Negro from Bahía Tortugas while Long Beach, La Jolla, and Guerrero Negro experience geneflow between localities. There is no evidence for isolation by distance between any of the sampling localities ($r^2 = 0.013$, $p = 0.25$).

Haplotype network analyses, while heavily reticulated, revealed two distinct clusters with one grouping comprised of individuals sampled in the north and the other of individuals sampled in the south with mixing of haplotypes between the two (Figure 4). Population assignment via STRUCTURE revealed a minimum of $K = 2$, with Santa Catalina Island and Bahía Tortugas being assigned to different clusters. Increasing to $K = 3$ retains these same clusters but begins to segregate Guerrero Negro from other localities (Figure 5). The Evanno method estimated $K = 2$ as the maximum value of K while the natural-log probability of K estimated $K = 3$ as the most likely value of K (Figure 6).

Genomic DNA Analyses

A total of 573 SNP loci were detected across all individuals. The average genetic diversity across loci was high at 0.254 ± 0.013. Significant genomic DNA structure was observed with 0.54% ($p < 0.0001$) of the variation occurring among populations (Table 8). Significant pairwise assessments of F_{ST} ranged in value from 0.004 to 0.0443 with Santa Catalina Island and San Pedro significantly diverging from all other sampling localities ($p \leq 0.017$, Table 6). Hierarchical AMOVAs testing divergence across putative phylogeographic barriers revealed no significant structure among regions based on location (Table 9). Population assignment via STRUCTURE revealed a minimum of $K = 2$ with Santa Catalina Island slightly separating from all other localities (Figure 7). Both the Evanno method and the natural-log probability of K method estimated $K = 2$ as the most likely number of clusters. The principal-components analysis (PCA) revealed an overall lack of distinct clustering supporting very little differentiation among sampled localities (Figure 8).

Discussion

Depending on the molecular marker used, data from the present study were able to detect either two or three genetically isolated populations of *P. productus*. Mitochondrial DNA evidence supports three divergent populations, as shown by both pairwise Φ_{ST} and hierarchical AMOVA analyses. This is further supported by both STRUCTURE analyses and the haplotype network that demonstrate two distinct clusters (Santa Catalina Island and Bahía Tortugas) with a high degree of mixing occurring between the two (Figure 4, 5). Congruent observations from both markers describe a single population of *P. productus* along the mainland from San Pedro, California to Guerrero Negro, Baja California Sur with an isolated population existing on Santa Catalina Island.

Putative Mechanisms Separating Populations of Pseudobatos productus

San Pedro Channel

Pseudobatos productus appears to have restricted gene flow across the San Pedro Channel, a span of water that separates the mainland from Santa Catalina Island. Of those studied at Santa Catalina Island, relatively few species show evidence of genetic divergence between the island and the mainland, likely due to the proximity of Santa Catalina Island to the mainland (approximately 45 km) and local current patterns that do not appear to inhibit larval dispersal (Dawson, 2001). Species that do show evidence of structure across the San Pedro Channel similar to that exhibited by *P. productus* include Round Stingray, Leopard Shark (*Triakis semifasciata*), and Black Surfperch (*Embiotoca jacksoni*), all viviparous species that prefer shallow coastal waters (Bernardi, 2000; Lewallen *et al*., 2007; Plank *et al*., 2010). The deepest part of the channel is the San

Pedro Basin, which is approximately 905 m and 20 km wide, that likely produces a sizeable barrier to viviparous species regularly inhabiting shallower waters with low dispersal potential for offspring due to reproductive characteristics (Dartnell and Gardner, 2009). Congruent molecular evidence that *P. productus* exhibits genetic structure between Santa Catalina Island and the mainland (although strong with mitochondrial markers and weak with nuclear markers) may suggest that movement varies between males and females across the channel with males potentially crossing more frequently than females.

Population Connectivity Across the United States-Mexican Border

The coastline of northern Baja California is characterized by discontinuous areas of seasonal upwelling events, particularly along the south-facing stretches of coasts of prominent headlands, such as Cabo San Quintín (Emerson, 1956). These upwelling zones can produce extreme variations in water temperature, with temperature fluctuations being as much as 10 °C cooler on the southern sides of headlands (Emerson, 1956). The upwelling zone at San Quintín is most active during the spring and summer months, characterized by strong northwest winds and offshore currents that have been shown to affect the abundance of marine organisms (Emerson, 1956). Given the hydrology of the area, the San Quintín Upwelling Zone appears to have a far greater impact on the population structure of summer-spawning fishes or species that use warm-water temperatures as settlement cues, such as the White Seabass (*Atractoscion nobilis*), Spotted Sand Bass (*Paralabrax maculatofasciatus*), and to a lesser degree, Barred Sand Bass (*Paralabrax nebulifer*) (Franklin *et al*., 2016; Paterson *et al*., 2015; Salomon, 2005). Species of teleost fishes that are not exclusively summer-spawners such as the California

Halibut (*Paralichthys californicus*), or species with a long pelagic larval phase such as the Opaleye (*Girella nigricans*), appear to be unaffected by the San Quintín Upwelling Zone as high levels of geneflow have been observed across it (Craig *et al*., 2011; Terry *et al*., 2000). Among broadcast spawning invertebrates, both the Tidepool Copepod (*Tigriopus californicus*) and the subtidal California Spiny Lobster (*Panulirus interruptus*) show connectivity across the upwelling zone despite having vastly different larval life-cycles and habitat usage (García-Rodríguez and Perez-Enriquez, 2006; Peterson *et al*., 2013).

Viviparous teleost species such as the Black Surfperch appear to be unaffected by the San Quintín Upwelling Zone, likely due to their mode of reproduction that produces well-formed young that do not rely on current patterns for dispersal (Bernardi, 2000). Elasmobranchs appear to be similarly unaffected by the upwelling zone, as the reproductive biology of all elasmobranch species, both viviparous and oviparous, produces well-formed nektonic young that are capable of dispersal independent of local current patterns. Although relatively few studies have assessed the genetic structure of elasmobranch species in both Californian and Mexican waters, the majority show evidence of high levels of genetic connectivity. For example, assessment of the Brown Smoothhound Shark (*Mustelus henlei*) provided evidence of a single genetic population from southern California into the Gulf of California using both nuclear and mitochondrial DNA markers (Chabot *et al*., 2015). Likewise, the Round Stingray was shown to have similar degrees of population connectivity when assessed with microsatellite markers (Plank *et al*., 2010). *Pseudobatos productus* appears to exhibit the same degree of connectivity across the San Quintín Upwelling Zone, with a high degree of geneflow

observed between locations in southern California and those along the northern Baja California coast.

Punta Eugenia

Restricted geneflow across Punta Eugenia has been observed among several species including Opaleye, Black Surfperch, and Tidepool Copepod, however, this has been attributed to a lack of useable habitat in the form of rocky reefs and tidepools just north of Punta Eugenia in the shallow, sandy Laguna Ojo de Liebre where Guerrero Negro is located (Bernardi, 2000; Peterson *et al*., 2013; Terry *et al*., 2000). In addition, genetic structuring across Punta Eugenia is found in species that are often restricted to estuaries such as the California Killifish (*Fundulus parvipinnis*) and Longjaw Mudsucker (*Gillichthys mirabilis*) or in species with pelagic larval phases such as the Spotted Sand Bass that may be affected by both current patterns and a scarcity of habitat space used by breeding adults across Point Eugenia (Bernardi and Talley, 2000; Huang and Bernardi, 2001; Salomon, 2005).

Similar to these species, *P. productus* appears to exhibit mitochondrial structuring across Punta Eugenia, a finding that is unique among previously studied rays including the confamilial Banded Guitarfish (*Zapteryx exasperata*) (Castillo-Páez *et al*., 2014). *Pseudobatos productus* and the Banded Guitarfish have similar life history characteristics and are known to frequently share habitat space, yet, the Banded Guitarfish exhibits genetic connectively along the Pacific coast of Baja California while *P. productus* does not (Castillo-Páez *et al*., 2014). Relatively few studies have examined the population connectivity of elasmobranchs along the Pacific coast of southern California and Baja California and even fewer have included samples collected from northern Baja

California. Those that do include samples from northern Baja California frequently indicate a high degree of genetic connectivity along the Pacific coast of Baja California, even across Punta Eugenia (Castillo-Páez *et al*., 2014; Ramírez-Amaro *et al*., 2017). The observed mitochondrial structuring across Punta Eugenia may be the result of a physical barrier represented by the geography of the area which females are less likely to cross while males are apparently more vagile.

Peninsula of Baja California

Patterns of genetic divergence across Cabo San Lucas, dividing Pacific coast populations from those in the Gulf of California, have been observed in a number of marine species including elasmobranchs and teleost fishes. Structuring that has been observed in teleost species with a pelagic larval phase is largely attributed to differences in water temperature and current patterns that may affect the distributions of adults and settlement of larvae such as the White Seabass (Franklin *et al*., 2016). Strong signals of genetic divergence between the Pacific coast and the Gulf of California observed in teleost species such as the Longjaw Mudsucker, Opaleye, and Spotted Sand Bass have been attributed to population structuring caused by the formation of the Baja California Peninsula, glaciation events, as well as the possibility of incipient speciation resulting from vicariance (Bernardi *et al*., 2003). Similar patterns of strong divergence across Cabo San Lucas have been observed in numerous demersal and coastal-associated elasmobranch species such as the Banded Guitarfish, Golden Cownose Ray, Butterfly Ray (*Gymnura marmorata*), and the Pacific Angel Shark (Castillo-Páez *et al*., 2014; Ramírez-Amaro *et al*., 2017; Jonathan Sandoval-Castillo and Rocha-Olivares, 2011; Smith *et al*., 2009b). Although not directly tested in the present study, population

subdivision has been observed in *P. productus* across Cabo San Lucas resulting in an isolated population within the Gulf of California (Sandoval-Castillo *et al*., 2004). Based on the phylogeographic breaks described above and the lack of population connectivity observed in previous studies along with those of the present study, there would appear to be at least four populations of *P. productus* within the northeastern Pacific based on mitochondrial data: 1) Santa Catalina Island, 2) southern California down to Punta Eugenia, 3) below Punta Eugenia to Cabo San Lucas, and 4) within the Gulf of California.

Putative Mechanisms Affecting Population Connectivity

The mitochondrial structuring exhibited by *P. productus* may potentially be a result of some degree of reproductive philopatry that may have produced geographically isolated reproductive units as seen in sawfishes of the genus *Pristis* (Phillips *et al*., 2011). *Pseudobatos productus* has been documented to exhibit seasonal residency to particular nursery sites in California (Farrugia *et al*., 2011) as well as some degree of site fidelity to inshore habitats between years (Dubois, 1981; Villaviencio-Garayzar, 1993; Sandoval-Castillo *et al*., 2004). Taken in combination with the relative lack of structure suggested by genomic DNA evidence, this mitochondrial structuring may also be sex dependent. Similar patterns of discordance between genomic or nuclear markers and mitochondrial DNA have been reported in numerous coastal shark species including the Grey Reef Shark (*Carcharhinus amblyrhynchos*) (Momigliano *et al*., 2017) and the Bonnethead Shark (*Sphyrna tiburo*) (Portnoy *et al*., 2015). For both, discordance has been attributed to higher levels of male movement and some degree of female philopatry (Momigliano *et al*., 2017; Portnoy *et al*., 2015). In addition to female philopatry to nurseries or sites that

produce conditions favorable for reproduction, prey availability may influence where individuals are found throughout the year as well as their fine-scale movement (Carlisle and Starr, 2010; Nosal *et al*., 2013).

Other causes of population structuring are usually attributed to long-term genetic patterns caused by historic changes in both the shape of the coastline and habitat availability either due to tectonic activity or glaciation periods (Dawson *et al*., 2006). Southern California and the Baja California peninsula have been sites of tectonic upheaval for the last several million years that may have influenced historic population structure in numerous species (Bernardi *et al*., 2003). The formation of the Gulf of California as well as the positioning of Baja California likely began during the late Miocene approximately 12 million to 3.5 million years ago (Holt *et al*., 2000). Studies of population structure and long-term phylogeographic separation in terrestrial species such as desert mice and side-blotched lizards as well as marine fauna including the Opaleye and Spotted Sand Bass support the opening and closure of numerous trans-peninsula seaways connecting the Pacific coast with the Gulf of California (Lindell *et al*., 2006; Riddle *et al*., 2000; Stepien *et al*., 2001; Terry *et al*., 2000; Upton and Murphy, 1997). During the late Pliocene, approximately 3 million years ago, two seaways were proposed that connected the northern and southern Gulf of California with the Pacific Ocean (Bernardi *et al*., 2003). Further genetic structuring in both marine and terrestrial fauna has been attributed to the opening of a mid-peninsular seaway connecting the Pacific Ocean to the Gulf of California south of Punta Eugenia approximately 1 million years ago during the mid-Pleistocene (Bernardi *et al*., 2003; Lindell *et al*., 2006; Upton and Murphy, 1997). While these Pliocene and Pleistocene seaways may have had an impact on the structure of populations of *P. productus* between the Pacific coast and the Gulf of

California, they are unlikely to have contributed to the structure observed in Baja California Sur as the proposed Pleistocene mid-peninsular seaway crossed the peninsula at Punta Abreojos south of Punta Eugenia.

Climate change and variation in sea levels associated with the Last Glacial Maximum (LGM) approximately 20 thousand years ago have been cited as potential causes of population structure in species such as the California Killifish and the Tidewater Goby (*Eucyclogobius newberryi*) in both southern California and the Baja peninsula (Dawson *et al*., 2001; Dolby *et al*., 2018). Reconstructions of coastal habitats and tidal estuaries have hypothesized that sea level recession caused by the LGM likely reduced the amount of usable habitat and decreases in water temperatures impacted the genetic structure of the California Killifish, Longjaw Mudsucker, and Shadow Goby (*Quietula y-cauda*) (Dolby *et al*., 2018). These species were likely restricted to shallow water refugia during periods of low sea level followed by colonization into suitable habitat as sea level rose. This loss of shared habitat space could account for structuring seen at the mitochondrial marker in *P. productus*. Similarly, habitat space in southern California was likely heavily affected by the sea level decrease associated with the LGM.

Historic geological data suggests that the northern California Channel Islands may have been connected to the mainland during the LGM, allowing for expansion and contraction of numerous species based on habitat availability (Dawson, 2001; Dawson *et al*., 2006). It is unlikely, however, that populations of *P. productus* between the mainland and the southern Channel Islands would experience increased connectivity during this period as the land bridge generated between the mainland and California Channel Islands likely connected farther north near Ventura rather than San Pedro (Bernardi, 2000; Dawson, 2001). In addition, due to the shape of the San Pedro Shelf and San Pedro Basin,

it is likely that shallow coastal habitat useable by *P. productus* would be further restricted across the deeper waters.

The low levels of mitochondrial nucleotide diversity along with the high haplotype diversity observed in the present study of *P. productus* are often attributed to a demographic expansion from a small effective population (Grant and Bowen, 1998). Evidence of rapid population expansion was detected in two sites (Santa Catalina Island and Bahía Tortugas) using the Fu's F_S statistic (Table 4), however, this test was found to be non-significant in all other sample localities. Similar to the Barred Sand Bass, the haplotype network generated in this study for *P. productus* lacks the typical starburst pattern often associated with haplotypic expansion from a bottleneck or small population indicating that such population expansion may have occurred in the distant past (Paterson *et al*., 2015). Estimates of effective population size for *P. productus* are fairly high, indicating the possibility of large, genetically stable populations throughout the range similar to patterns observed in the Brown Smoothhound Shark (Chabot *et al*., 2015a). Despite genetic patterns of expansion and long-term stability, populations of *P. productus* appear to be declining within Baja California based on observational studies (Farrugia *et al*., 2016). Due to this incongruence, future monitoring of the genetic diversity of *P. productus* may be essential for the successful management of the species, particularly in areas where harvesting occurs.

Cryptic Species

Within the Gulf of California, putative cryptic species have been suggested for *P. productus*, the Golden Cownose Ray, and the Butterfly Ray due to high levels of mitochondrial genetic divergence observed between/among sampled populations

(Sandoval-Castillo *et al*., 2004; Sandoval-Castillo and Rocha-Olivares, 2011; Smith *et al*., 2009b). Recently, a new species of *Pseudobatos* was described from 82 specimens obtained from the Gulf of California in the 1940s and 1950s (Rutledge, 2019). *Pseudobatos buthi* is described as being morphologically most similar to *P. productus* in most aspects including coloration but differs in the size and shape of the rostrum as well as in the scale and thorn patterns found on the head (Rutledge, 2019). *Pseudobatos buthi* was found to have a smaller size at maturity in both sexes (48 cm) and shorter observed total length (68.5 cm) than *P. productus* (Rutledge, 2019). This difference in size at maturity and maximum total length observed in *P. buthi* may explain differences in life history characteristics reported for *P. productus* between the Gulf of California and along the Pacific coast (Downton-Hoffmann, 2007; Dubois, 1981; Juaristi Videgaray, 2016; Márquez-Farías, 2007; Timmons and Bray, 1997; Villaviencio-Garayzar, 1993). An assessment of the genetic structure of *P. buthi* and phylogenetic comparisons of *P. productus* and *P. buthi* would be helpful in order to further resolve the evolutionary relationships between the two species.

In addition to *P. buthi*, there are two other described congeneric species of guitarfish found within the Gulf of California that share habitat space with *P. productus*: the Speckled Guitarfish (*Pseudobatos glaucostigmus*) and the Whitesnout Guitarfish (*Pseudobatos leucorhynchus*). All four species of guitarfish are morphologically very similar with slight variations in color and size being the immediately distinguishable characteristics. These species inhabit coastal waters in a clinal pattern that includes overlapping habitat space within the Gulf of California. While *P. productus* is found from California waters and into the Gulf of California, the Speckled Guitarfish ranges from southern Baja California Sur to Ecuador, the Whitesnout Guitarfish occurs within the

Gulf of California and south to Ecuador, and *P. buthi* has been described only within the Gulf of California (Bizzarro 2016a; Bizzarro 2016b; Farrugia *et al*. 2016; Rutledge 2019). Of the four species, population trends of only *P. productus* have been assessed by the IUCN within the last decade with the species being classified as Near-Threatened with a declining trend, while the Whitesnout Guitarfish was assessed in 2006 and classified similarly (Bizzarro, 2016b; Farrugia *et al*., 2016). Both *P. buthi* and the Speckled Guitarfish are considered Data Deficient by the IUCN despite likely being impacted by both intentional harvest in artisanal fisheries and accidental bycatch in trawl-based fisheries (Bizzarro, 2016a; Rutledge, 2019). In order to fully understand the impact of removal of individuals on population structure and diversity of Pacific *Pseudobatos* spp., further studies and stock assessments are required as there is a scarcity of information on the Whitesnout Guitarfish, the Speckled Guitarfish, and *P. buthi*.

Fisheries Management

While *P. productus* is a heavily fished, bi-national species, relatively little management strategies have been enacted in order to protect the species from overexploitation. In California, there are no specific regulations to the recreational harvest of guitarfish, just a general catch limit of 10 individuals per day regardless of size and sex ("California ocean sport fishing regulations," 2019). In Mexico, there has been an increase in fishing regulations applied to the artisanal fishery, however, none specifically have been applied for *P. productus*. In 2007, a regulation was implemented which required the use of a logbook for accurate catch reporting as well as defined restricted fishing areas (Farrugia *et al*., 2016). As a result of the 2007 regulations, *P. productus* as well as other guitarfish species are recorded in the fishery as "pez guitarra", however,

differentiation between species is not required (Farrugia *et al*., 2016). In 2012, a regulation was implemented to close the fishery from May 31 to July 31, a period that coincides with the peak of breeding season for many coastal elasmobranch species (Farrugia *et al*., 2016). Despite this ban, *P. productus* remains vulnerable to fishing pressures due to their extended use of shallow coastal areas throughout the year.

The implementation of management policies specific to the genus *Pseudobatos* along with species-specific management strategies based on observed population structuring in *P. productus* will likely promote both the health of the species as well as that of the fishery. Within the Mexican fishery, region-based catch limits and size restrictions should be set, particularly during the spring months when large, gravid female *P. productus* remain unprotected by the fishing closure. Under the current management system, only late-arriving individuals to aggregation sites are protected by policy while those who arrive earlier in the year are vulnerable (Márquez-Farías, 2007). Due to the observed mitochondrial genetic structure described in the present study, populations of *P. productus* north of Punta Eugenia should be assessed and managed independently of those to the south. Current assessments of catch by the artisanal fisheries are typically conducted within broad ranges of the states of Baja California and Baja California Sur, so an assessment of catches within regions of genetic similarity are necessary for a more accurate understanding of the impact of harvest on populations. Given the evidence of strong genetic divergence of *P. productus* across Cabo San Lucas, populations should be managed as separate entities between southern Baja California Sur and the Gulf of California. Additionally, current fisheries along the Pacific coast and within the Gulf of California should be assessed for the presence of *P. productus* and *P. buthi* as both species are very similar morphologically and incorrect identifications may result in

erroneous catch counts (Rutledge, 2019). The inclusion of accurate species identifications of *Pseudobatos* spp. within the fisheries (along with bycatch) would be an ideal addition to current management policies, particularly in the Gulf of California where species' ranges overlap.

Collaborative management between Californian and Baja Californian fisheries authorities is necessary to protect the area of genetic continuity that stretches from southern California to Guerrero Negro. Currently, management policy is divided between the two entities with different levels of harvest regulations ("California ocean sport fishing regulations," 2019; Farrugia *et al*., 2016). As previously mentioned, there are no specific regulations to the recreational harvest of guitarfish within Californian waters. This policy should be revised to include guitarfish with other federally regulated "groundfish" species and to have catch limits set similarly to those species with shared levels of population structuring, life history characteristics, and that inhabit similar environments as *P. productus*. Additionally, the implementation of regionally based policies that manage stocks at the California Channel Islands separately from those on the mainland will help maintain genetic diversity within isolated populations. Though the protection of *P. productus* was not the aim of restrictions to trawling fisheries and the implementation of harvest restrictions within marine protected areas, the species has likely benefitted from them as mortality due to bycatch has been greatly reduced (Farrugia *et al*., 2016).

The combination of population genetic structure and accurate stock assessments are essential for the creation of management strategies used to promote the health and longevity of targeted species as well as the fisheries that harvest them. In addition, studies of genetic diversity can provide a genetic baseline for populations that can be used to

detect localized declines in genetic diversity. Trans-international assessments of population structure are also essential in species which may experience different harvest rates and policy applications throughout its range, to further prevent depletion within isolated populations. As shown with *P. productus*, observed patterns of population connectivity seldom align with international or state boundaries that are often used to delineate management units.

Conclusion

This study sought to assess the genetic diversity and population genetic structure of *P. productus* from southern California to the southern Pacific coast of Baja California. Both mitochondrial and genomic DNA evidence suggests a high degree of geneflow and connectivity between San Pedro, California and Guerrero Negro, Baja California Sur with an isolated population being detected at Santa Catalina Island. Mitochondrial evidence supports further structuring of *P. productus* with an additionally isolated population being observed at Bahía Tortugas, BCS. Based on these findings, populations of *P. productus* should be managed based on areas of genetic relatedness instead of by political boundaries. Furthermore, international cooperation may be necessary to protect areas of high geneflow, such as observed across the United States-Mexico border.

Literature Cited

Agashe, D. (2009). The stabilizing effect of intraspecific genetic variation on population dynamics in novel and ancestral habitats. *The American Naturalist*, 174(2), 255–267.

Andrews, K. R., and Luikart, G. (2014). Recent novel approaches for population genomics data analysis. *Molecular Ecology*, 23(7), 1661–1667.

Begg, G. A., Friedland, K. D., and Pearce, J. B. (1999). Stock identification and its role in stock assessment and fisheries management: An overview. *Fisheries Research*, 43(1–3), 1–8.

Benjamini, Y., and Yekutieli, D. (2001). The control of the false discovery rate in multiple testing under dependency. *The Annals of Statistics*, 29(4), 1165–1188.

Bernardi, G. (2000). Barriers to gene flow in *Embiotoca jacksoni*, a marine fish lacking a pelagic larval stage. *Evolution*, 54(1), 226–237.

Bernardi, G., Findley, L., and Rocha-Olivares, A. (2003). Vicariance and dispersal across Baja California in disjunct marine fish populations. *Evolution*, 57(7), 1599–1609.

Bernardi, G., and Talley, D. (2000). Genetic evidence for limited dispersal in the coastal California killifish, *Fundulus parvipinnis*. *Journal of Experimental Marine Biology and Ecology*, 255(2), 187–199.

Bizzarro, J. J. (2016a). *Pseudobatos glaucostigmus*. *The IUCN Red List of Threatened Species*, 8235. Retrieved from https://www.iucnredlist.org/species/161457/103931285

Bizzarro, J. J. (2016b). *Pseudobatos leucorhynchus*. *The IUCN Red List of Threatened Species*, 8235. Retrieved from https://www.iucnredlist.org/species/60168/103935082

Bizzarro, J. J., Smith, W. D., Márquez-Farías, J. F., Tyminski, J., and Hueter, R. E. (2009). Temporal variation in the artisanal elasmobranch fishery of Sonora, Mexico. *Fisheries Research*, 97(1–2), 103–117.

Briggs, J. C. (1974). Marine zoogeography. New York: McGraw-Hill.

Bushnell, B. (2019). BBMap. Retrieved from https://sourceforge.net/projects/bbmap/

California ocean sport fishing regulations. (2019). California Department of Fish and Wildlife. Retrieved from https://www.wildlife.ca.gov/Regulations

Carlisle, A. B., and Starr, R. M. (2010). Tidal movements of female leopard sharks (*Triakis semifasciata*) in Elkhorn Slough, California. *Environmental Biology of Fishes*, 89(1), 31–45.

Cartamil, D., Santana-Morales, O., Escobedo-Olvera, M., Kacev, D., Castillo-Geniz, L., Graham, J. B., Rubin, R. D., Sosa-Nishizaki, O. (2011). The artisanal elasmobranch fishery of the Pacific coast of Baja California, Mexico. *Fisheries Research*, 108(2–3), 393–403.

Casey, J. M., and Myers, R. A. (1998). Near extinction of a large, widely distributed fish. *Science*, 281(5377), 690–692.

Castillo-Páez, A., Sosa-Nishizaki, O., Sandoval-Castillo, J., Galván-Magaña, F., Blanco-Parra, M. D. P., and Rocha-Olivares, A. (2014). Strong population structure and shallow mitochondrial phylogeny in the banded guitarfish, *Zapteryx exasperata* (Jordan y Gilbert, 1880), from the northern Mexican Pacific. *Journal of Heredity*, 105(1), 91–100.

Chabot, C. L., Espinoza, M., Mascareñas-Osorio, I., and Rocha-Olivares, A. (2015a). The effect of biogeographic and phylogeographic barriers on gene flow in the brown smoothhound shark, *Mustelus henlei*, in the northeastern Pacific. *Ecology and Evolution*, 5(8), 1585–1600.

Chabot, C. L., Hawk, H. A., and Allen, L. G. (2015b). Low contemporary effective population size detected in the Critically Endangered giant sea bass, *Stereolepis gigas*, due to fisheries overexploitation. *Fisheries Research*, 172, 71–78.

Craig, M. T., Fodrie, F. J., Allen, L. G., Chartier, L. A., and Toonen, R. J. (2011). Discordant phylogeographic and biogeographic breaks in California halibut. *Bulletin, Southern California Academy of Sciences*, 110(3), 141–151

Darriba, D., Taboada, G. L., Doallo, R., and Posada, D. (2012). jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods*, 9(8), 772.

Dartnell, P., and Gardner, J. V. (2009). Seafloor terrain analysis and geomorphology of the greater Los Angeles margin and San Pedro Basin, Southern California. In H. J. Lee and W. R. Normark (Eds) Earth Science in the Urban Ocean: The Southern California Continental Borderland., 454, 9–28.

David, M., Dzamba, M., Lister, D., Ilie, L., and Brudno, M. (2011). SHRiMP2: Sensitive yet practical short read mapping. *Bioinformatics*, 27(7), 1011–1012.

Dawson, M. N. (2001). Phylogeography in coastal marine animals: A solution from California?. *Journal of Biogeography*, 28(6), 723–736

Dawson, M. N., Staton, J. L., and Jacobs, D. K. (2001). Phylogeography of the tidewater goby, *Eucyclogobius newberryi* (Teleostei, Gobiidae), in coastal California. *Evolution*; 55(6), 1167–1179.

Dawson, M. N., Waples, R. S., and Bernardi, G. (2006). Phylogeography. In L. G. Allen, D. J. Pondella, and M. H. Horn (Eds.), The Ecology of Marine Fishes, California and Adjacent Waters (pp. 26–54). Los Angeles, CA: University of California Press, Berkeley.

Dolby, G. A., Ellingson, R. A., Findley, L. T., and Jacobs, D. K. (2018). How sea-level change mediates genetic divergence in coastal species across regions with varying tectonic and sediment processes. *Molecular Ecology*, 27(4), 994–1011.

Downton-Hoffmann, C. A. (2007). Biología del Pez Guitarra *Rhinobatos productus* (Ayres, 1856), en Baja California Sur, México. (Doctoral dissertation). Instituto Politecnico Nacional. Centro Interdisciplnario de Ciencias Marinas.

Dubois, A. J. (1981). Studies on fishes in Mugu Lagoon, California. (Master's Thesis). University of California, Santa Barbara.

Emerson, W. K. (1956). Upwelling and associated marine life along Pacific Baja California, Mexico. *Journal of Paleontology*, 30(2), 393–397.

Evanno, G., Regnaut, S., and Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Molecular Ecology*, 14(8), 2611–2620.

Excoffier, L., and Lischer, H. E. L. (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., Smouse, P. E., and Quattro, J. M. (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. *Genetics*, 131(2), 479–491.

Falush, D., Stephens, M., and Pritchard, J. K. (2003). Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies. *Genetics*, 164(4), 1557–1587.

Falush, D., Stephens, M., and Pritchard, J. K. (2007). Inference of population structure using multilocus genotype data: Dominant markers and null alleles. *Molecular Ecology Notes*, 7(4), 574–578.

Farrugia, T. J., Márquez-Farías, F., Freedman, R. M., Lowe, C. G., Smith, W. D., and Bizzarro, J. J. (2016). *Pseudobatos productus*. *The IUCN Red List of Threatened Species*, 8235. Retrieved from https://www.iucnredlist.org/species/60171/104004394

Farrugia, T. J., Espinoza, M., and Lowe, C. G. (2011). Abundance, habitat use and movement patterns of the shovelnose guitarfish (*Rhinobatos productus*) in a restored southern California estuary. *Marine and Freshwater Research*, 62(6), 648–657.

Frankham, R. (1995). Inbreeding and extinction: Island populations. *Conservation Biology*, 9(4), 792–799.

Franklin, M. P., Chabot, C. L., and Allen, L. G. (2016). A baseline investigation into the population structure of white seabass, *Atractoscion nobilis*, in California and Mexican waters using microsatellite DNA analysis. *Bulletin, Southern California Academy of Sciences*, 115(2), 126–135.

Freeland, J., Kirk, H., and Petersen, S. (2011). Molecular ecology. Chichester, West Sussex, UK: Wiley-Blackwell.

Frisk, M. G., Miller, T. J., and Fogarty, M. J. (2001). Estimation and analysis of biological parameters in elasmobranch fishes: A comparative life history study. *Canadian Journal of Fisheries and Aquatic Sciences*, 58(5), 969–981.

Fu, Y. X. (1997). Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*, 147(2), 915–925.

Gaffney, P. M., Rupnow, J., and Domeier, M. L. (2007). Genetic similarity of disjunct populations of the giant sea bass *Stereolepis gigas*. *Journal of Fish Biology*, 70, 111–124.

García-Rodríguez, F. J., and Perez-Enriquez, R. (2006). Genetic differentiation of the California spiny lobster *Panulirus interruptus* (Randall, 1840) along the west coast of the Baja California Peninsula, Mexico. *Marine Biology*, 148(3), 621–629.

Grant, W. S., and Bowen, B. W. (1998). Shallow population histories in deep evolutionary lineages of marine fishes: Insights from sardines and anchovies and lessons for coservation. *Journal of Heredity*, 89, 415–426.

Green, M. R., and Sambrook, J. (2016). Precipitation of DNA with ethanol. Cold Spring Harbor Protocols, 2016(12).

Hinegardner, R. (1976). The cellular DNA content of sharks, rays and some other fishes. *Comparative Biochemistry and Physiology Part B: Comparative Biochemistry*, 55(3), 367–370.

Hoelzel, A. R., Shivji, M. S., Magnussen, J., and Francis, M. P. (2006). Low worldwide genetic diversity in the basking shark (*Cetorhinus maximus*). *Biology Letters*, 2(4), 639–642.

Hoenig, J. M., and Gruber, S. H. (1990). Life-history patterns in the elasmobranchs: Implications for fisheries management. *Life-History Patterns in the Elasmobranchs: Implications for Fisheries Management. NOAA Technical Report NMFS*, 90, 1–16.

Holt, J. W., Holt, E. W., and Stock, J. M. (2000). An age constraint of Gulf of California rifting from the Santa Rosalía basin, Baja California Sur, Mexico. *Bulletin of the Geological Society of America*, 112(4), 540–549.

Horn, M. H., Allen, L. G., and Lea, R. N. (2006). Biogeography. In L. G. Allen, D. J. Pondella, and M. H. Horn (Eds.), The Ecology of Marine Fishes, California and Adjacent Waters (pp. 3–16). Los Angeles, CA: University of California Press, Berkeley.

Horn, M. H., and Allen, L. G. (1978). A distributional analysis of California coastal marine fishes. *Journal of Biogeography*, 5(1), 23–42.

Huang, D., and Bernardi, G. (2001). Disjunct Sea of Cortez-Pacific ocean *Gillichthys mirabilis* populations and the evolutionary origin of their Sea of Cortez endemic relative, *Gillichthys seta*. *Marine Biology*, 138(2), 421–428.

Hubisz, M. J., Falush, D., Stephens, M., and Pritchard, J. K. (2009). Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources*, 9(5), 1322–1332.

IUCN Red List of Threatened Species (2019) Retrieved from https://www.iucnredlist.org/search?taxonomies=100043&searchType=species

Jombart, T. (2008). Adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics*, 24(11), 1403–1405.

Jombart, T., and Ahmed, I. (2011). adegenet 1.3-1: New tools for the analysis of genome-wide SNP data. *Bioinformatics*, 27(21), 3070–3071.

Juaristi Videgaray, D. (2016). Biología reproductiva del Pez Guitarra *Pseudobatos productus* (Ayres, 1856), en Bahía Tortugas, Baja California Sur, México. (Master's Thesis). Instituto Politecnico Nacional. Centro Interdisciplnario de Ciencias Marinas.

Kopelman, N. M., Mayzel, J., Jakobsson, M., Rosenberg, N. A., and Mayrose, I. (2015). CLUMPAK: a program for identifying clustering modes and packaging population structure inferences across K. *Molecular Ecology Resources*, 15(5), 1179–1191.

Koressaar, T., and Remm, M. (2007). Enhancements and modifications of primer design program Primer3. *Bioinformatics*, 23(10), 1289–1291.

Leigh, J. W., and Bryant, D. (2015). POPART: Full-feature software for haplotype network construction. *Methods in Ecology and Evolution*, 6(9), 1110–1116.

Lewallen, E. A., Anderson, T. W., and Bohonak, A. J. (2007). Genetic structure of leopard shark (*Triakis semifasciata*) populations in California waters. *Marine Biology*, 152(3) 599–609.

Lindell, J., Ngo, A., and Murphy, R. W. (2006). Deep genealogies and the mid-peninsular seaway of Baja California. *Journal of Biogeography*, 33(8), 1327–1331.

Love, M. (2011). Certainly more than you want to know about the fishes of the Pacific Coast. (pp. 66–67). Santa Barbara, CA: Really Big Press.

Márquez-Farías, J. F. (2002). The artisanal ray fishery in the Gulf of California: development, fisheries research and management issues. *IUCN Shark Specialist Group. Shark News*, 14, 1–5.

Márquez-Farías, J. F. (2005). Gillnet mesh selectivity for the shovelnose guitarfish (*Rhinobatos productus*) from fishery-dependent data in the artisanal ray fishery of the Gulf of California, Mexico. *Journal of Northwest Atlantic Fishery Science*, 35, 443–452.

Márquez-Farías, J. F. (2007). Reproductive biology of shovelnose guitarfish *Rhinobatos productus* from the eastern Gulf of California México. *Marine Biology*, 151(4), 1445–1454.

Márquez-Farías, J. F. (2011). Assessment of the impact of gillnets on the population structure of the shovelnose guitarfish *Rhinobatos productus* from the Gulf of California, Mexico. *Ciencias Marinas*, 37(3), 293–304.

Meyer, E. (2018). 2b-RAD Tools and Protocols. Retrieved from http://people.oregonstate.edu/~meyere/tools.html

Momigliano, P., Harcourt, R., Robbins, W. D., Jaiteh, V., Mahardika, G. N., Sembiring, A., and Stow, A. (2017). Genetic structure and signatures of selection in grey reef sharks (*Carcharhinus amblyrhynchos*). *Heredity* 119(3), 142–153

Musick, J. A., Burgess, G., Cailliet, G., Camhi, M., and Fordham, S. (2000). Management of sharks and their relatives (Elasmobranchii). *Fisheries*, 25(3), 9–13.

Narum, S. R. (2006). Beyond Bonferroni: Less conservative analyses for conservation genetics. *Conservation Genetics*, 7(5), 783–787.

Nosal, A. P., Cartamil, D. C., Long, J. W., Lührmann, M., Wegner, N. C., and Graham, J. B. (2013). Demography and movement patterns of leopard sharks (*Triakis semifasciata*) aggregating near the head of a submarine canyon along the open coast of southern California, USA. *Environmental Biology of Fishes*, 96(7), 865–878.

Paterson, C. N., Chabot, C. L., Robertson, J. M., Erisman, B., Cota-Nieto, J. J., and Allen, L. G. (2015). The genetic diversity and population structure of barred sand bass, *Paralabrax nebulifer*: A historically important fisheries species off southern and Baja California. *California Cooperative Oceanic Fisheries Investigations Reports*, 56, 1–13.

Peakall, R., and Smouse, P. E. (2006). GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, 6(1), 288–295.

Peakall, R., and Smouse, P. E. (2012). GenALEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research-an update. *Bioinformatics*, 28(19), 2537–2539.

Peterson, D. L., Kubow, K. B., Connolly, M. J., Kaplan, L. R., Wetkowski, M. M., Leong, W., Phillips, B. C., and Edmands, S. (2013). Reproductive and phylogenetic divergence of tidepool copepod populations across a narrow geographical boundary in Baja California. *Journal of Biogeography*, 40(9), 1664–1675.

Phillips, N. M., Chaplin, J. A., Morgan, D. L., and Peverell, S. C. (2011). Population genetic structure and genetic diversity of three critically endangered *Pristis* sawfishes in Australian waters. *Marine Biology*, 158(4), 903–915.

Plank, S. M., Lowe, C. G., Feldheim, K. A., Wilson, R. R., and Brusslan, J. A. (2010). Population genetic structure of the round stingray *Urobatis halleri* (Elasmobranchii: Rajiformes) in southern California and the Gulf of California. *Journal of Fish Biology*, 77(2), 329–340.

Portnoy, D. S., Puritz, J. B., Hollenbeck, C. M., Gelsleichter, J., Chapman, D., and Gold, J. R. (2015). Selection and sex-biased dispersal in a coastal shark: The influence of philopatry on adaptive variation. *Molecular Ecology*, 24(23), 5877–5885.

Prieto, I., Violle, C., Barre, P., Durand, J. L., Ghesquiere, M., and Litrico, I. (2015). Complementary effects of species and genetic diversity on productivity and stability of sown grasslands. *Nature Plants*, 1(4), 1–5.

Pritchard, J. K., Stephens, M., and Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, 155(2), 945–959.

Puritz, J. B., Matz, M. V., Toonen, R. J., Weber, J. N., Bolnick, D. I., and Bird, C. E. (2014). Demystifying the RAD fad. *Molecular Ecology*, 23, 5937–5942.

Ramírez-Amaro, S. R., Cartamil, D., Galván-Magaña, F., Gonzalez-Barba, G., Graham, J. B., Carrera-Fernandez, M., Escobar-Sanchez, O., Sosa-Nishizaki, O., and Rochin-Alamillo, A. (2013). The artisanal elasmobranch fishery of the Pacific coast of Baja California Sur, Mexico, management implications. *Scientia Marina*, 77(3), 473–487.

Ramírez-Amaro, S., Ramírez-Macías, D., Vázquez-Juárez, R., Flores-Ramírez, S., Galván-Magaña, F., and Gutiérrez-Rivera, J. N. (2017). Population structure of the Pacific angel shark (*Squatina californica*) along the northwestern coast of Mexico based on the mitochondrial DNA control region. *Ciencias Marinas*, 43(1), 69–80.

Reiss, H., Hoarau, G., Dickey-Collas, M., and Wolff, W. J. (2009). Genetic population structure of marine fish: Mismatch between biological and fisheries management units. *Fish and Fisheries*, 10(4), 361–395.

Riddle, B. R., Hafner, D. J., Alexander, L. F., and Jaeger, J. R. (2000). Cryptic vicariance in the historical assembly of a Baja California Peninsular Desert biota. *Proceedings of the National Academy of Sciences of the United States of America*, 97(26), 14438–14443.

Romo-Curiel, A. E., Sosa-Nishizaki, O., Pérez-Jiménez, J. C., and Rodríguez-Medrano, M. C. (2016). Reproductive cycle and maternal–embryonic nutritional relationship of shovelnose guitarfish *Pseudobatos productus* in the Gulf of California. *Journal of Fish Biology*, 90(3), 889–905.

Rowe, H. C., Renaut, S., and Guggisberg, A. (2011). RAD in the realm of next-generation sequencing technologies. *Molecular Ecology*, 20(17), 3499–3502.

Rutledge, K. M. (2019). A new guitarfish of the genus *Pseudobatos* (Batoidea: Rhinobatidae) with key to the guitarfishes of the Gulf of California. *Copeia*, 107(3), 451–463.

Ryman, N. (1991). Conservation genetics considerations in fishery management. *Journal of Fish Biology*, 39, 211–224.

Salazar-Hermoso, F., and Villavicencio-Garayzar, C. (1999). Relative abundance of the shovelnose guitarfish *Rhinobatos productus* (Ayres, 1856) (Pisces: Rhinobatidae) in Bahía Almejas, Baja California Sur, from 1991 to 1995. *Ciencias Marinas*, 25(3), 401–422.

Salomon, M. P. (2005). Gene flow among populations of the spotted sand bass, *Paralabrax maculatofasciatus*, from southern California and two populations along the Pacific coast of Baja California. (Master's Thesis). California State University, Northridge

Sandoval-Castillo, J., Rocha-Olivares, A., Villavicencio-Garayzar, C., and Balart, E. (2004). Cryptic isolation of Gulf of California shovelnose guitarfish evidenced by mitochondrial DNA. *Marine Biology*, 145(5), 983–988.

Sandoval-Castillo, J., and Rocha-Olivares, A. (2011). Deep mitochondrial divergence in Baja California populations of an aquilopelagic elasmobranch: The golden cownose ray. *Journal of Heredity*, 102(3), 269–274.

Schultz, J. K., Feldheim, K. A., Gruber, S. H., Ashley, M. V., McGovern, T. M., and Bowen, B. W. (2008). Global phylogeography and seascape genetics of the lemon sharks (genus *Negaprion*). *Molecular Ecology*, 17(24), 5336–5348.

Schwartz, F. J., and Maddock, M. B. (1986). Comparisons of karyotypes and cellular DNA contents within and between major lines of elasmobranchs. In T. Uyeno, R. Arai, T. Taniuchi, and K. Matsuura (Eds.), Indo-Pacific Fish Biology (pp. 148–157). Ichthyological Society Japan, Tokyo.

Shaklee, J. B., and Bentzen, P. (1998). Genetic Identification of Stocks of Marine Fish and Shellfish. *Bulletin of Marine Science*, 62(2), 589–621.

Shaklee, J. B., Beacham, T. D., Seeb, L., and White, B. A. (1999). Managing fisheries using genetic data: Case studies from four species of Pacific salmon. *Fisheries Research*, 43, 45–78.

Smith, P. J., Francis, R. I. C. C., and McVeagh, M. (1991). Loss of genetic diversity due to fishing pressure. *Fisheries Research*, 10(3-4), 309–316.

Smith, W. D., Bizzarro, J. J., and Cailliet, G. M. (2009a). The artisanal elasmobranch fishery on the east coast of Baja California, Mexico: Characteristics and management considerations. *Ciencias Marinas*, 35(2), 209–236.

Smith, W. D., Bizzarro, J. J., Richards, V. P., Nielsen, J., Márquez-Farías, F., and Shivji, M. S. (2009b). Morphometric convergence and molecular divergence: The taxonomic status and evolutionary history of *Gymnura crebripunctata* and *Gymnura marmorata* in the eastern Pacific Ocean. *Journal of Fish Biology*, 75(4), 761–783.

Stepien, C. A., Rosenblatt, R. H., and Bargmeyer, B. A. (2001). Phylogeography of the spotted sand bass, *Paralabrax maculatofasciatus*: Divergence of Gulf of California and Pacific Coast populations. *Evolution*, 55(9), 1852–1862.

Stevens, J. D., Bonfil, R., Dulvy, N. K., and Walker, P. A. (2000). The effects of fishing on sharks, rays, and chimaeras (chondrichthyans), and the implications for marine ecosystems. *ICES Journal of Marine Science*, 57(3), 476–494.

Tajima, F. (1989). Statistical Method for Testing the Neutral Mutation Hypothesis by DNA Polymorphism. *Genetics*, 123(3), 585–595.

Terry, A., Bucciarelli, G., and Bernardi, G. (2000). Restricted gene flow and incipient speciation in disjunct Pacific Ocean and Sea of Cortez populations of a reef fish species, *Girella nigricans*. *Evolution*, 54(2), 652–659.

Timmons, M., and Bray, R. N. (1997). Age, growth, and sexual maturity of shovelnose guitarfish, *Rhinobatos productus* (Ayres). *Fishery Bulletin*, 95(2), 349–359.

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B. C., Remm, M., and Rozen, S. G. (2012). Primer3-new capabilities and interfaces. *Nucleic Acids Research*, 40(15), 1–12.

Upton, D. E., and Murphy, R. W. (1997). Phylogeny of the side-blotched lizards (Phrynosomatidae: *Uta*) based on mtDNA sequences: support for a midpeninsular seaway in Baja California. *Molecular Phylogenetics and Evolution*, 8(1), 104–113.

Villaviencio-Garayzar, C. J. (1993). Biología reproductiva de *Rhinobatos productus* (Pisces: Rhinobatidae), en Bahía Almejas, Baja California Sur, México. *Revista de Biología Tropical*, 41(3B), 777–782.

Wang, S., Meyer, E., McKay, J. K., and Matz, M. V. (2012). 2b-RAD: a simple and flexible method for genome-wide genotyping. *Nature Methods*, 9(8), 808–810.

Willi, Y., van Buskirk, J., and Hoffmann, A. A. (2006). Limits to the adaptive potential of small populations. *Annual Review of Ecology, Evolution, and Systematics*, 37, 433–458.

Appendix A: Figures

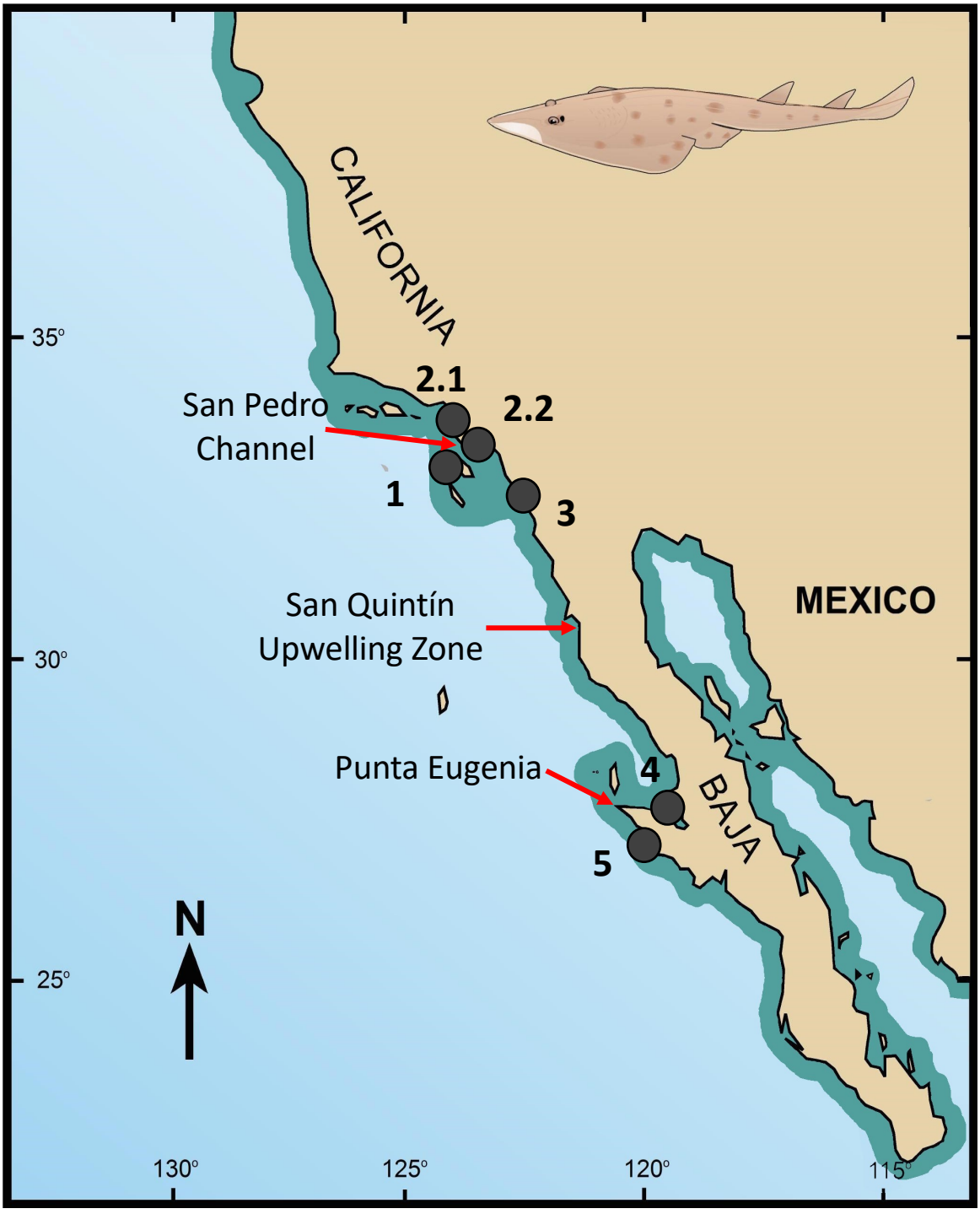


Figure 1. Distribution of *Pseudobatos productus* in shaded area with marked sampling locations. 1) Santa Catalina Island, CA, 2.1) Dockweiler Beach, CA, 2.2) Long Beach, CA (2.1 and 2.2 later combined as the sampling region San Pedro), 3) La Jolla, CA, 4) Guerrero Negro, BCS, 5) Bahía Tortugas, BCS. Major known phylogeographic or biogeographic barriers indicated by red arrows.

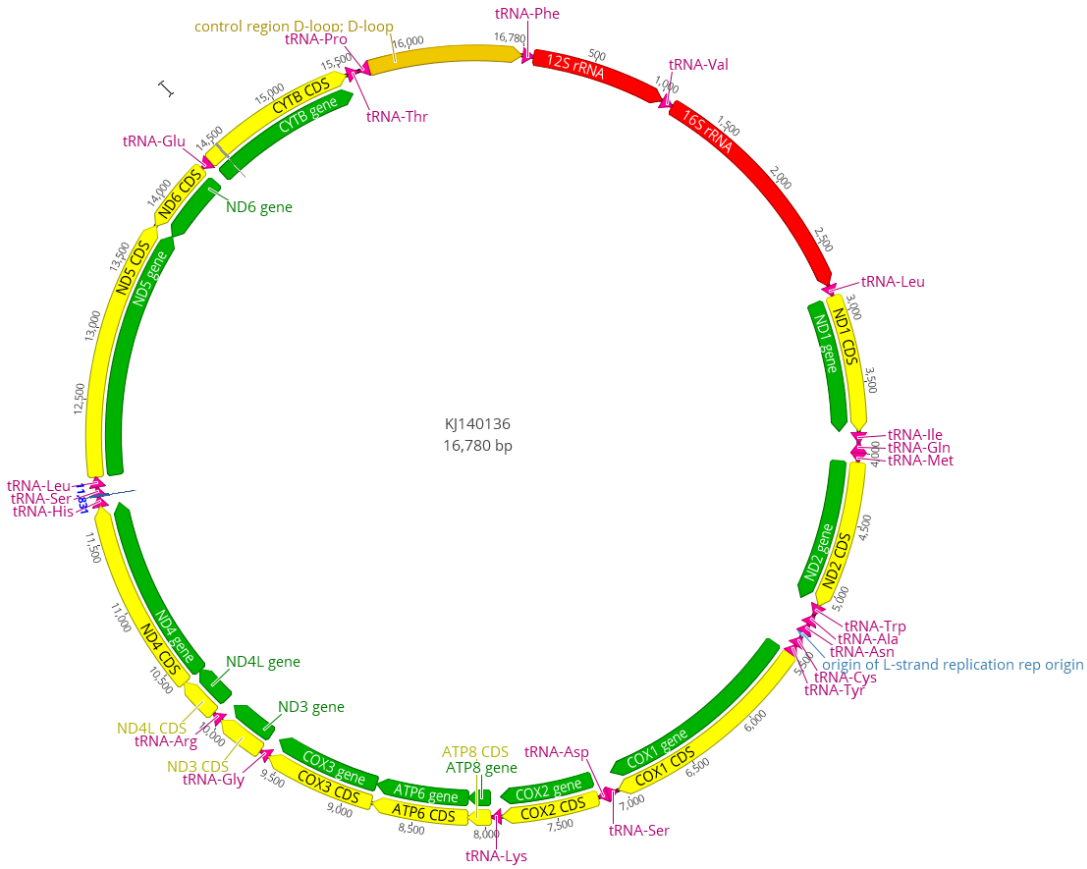


Figure 2. Complete mitochondrial genome of *Rhinobatos schlegelii* (GenBank Accession #KJ140136.1) indicating the position of the control region (D-Loop) and flanking tRNAs.

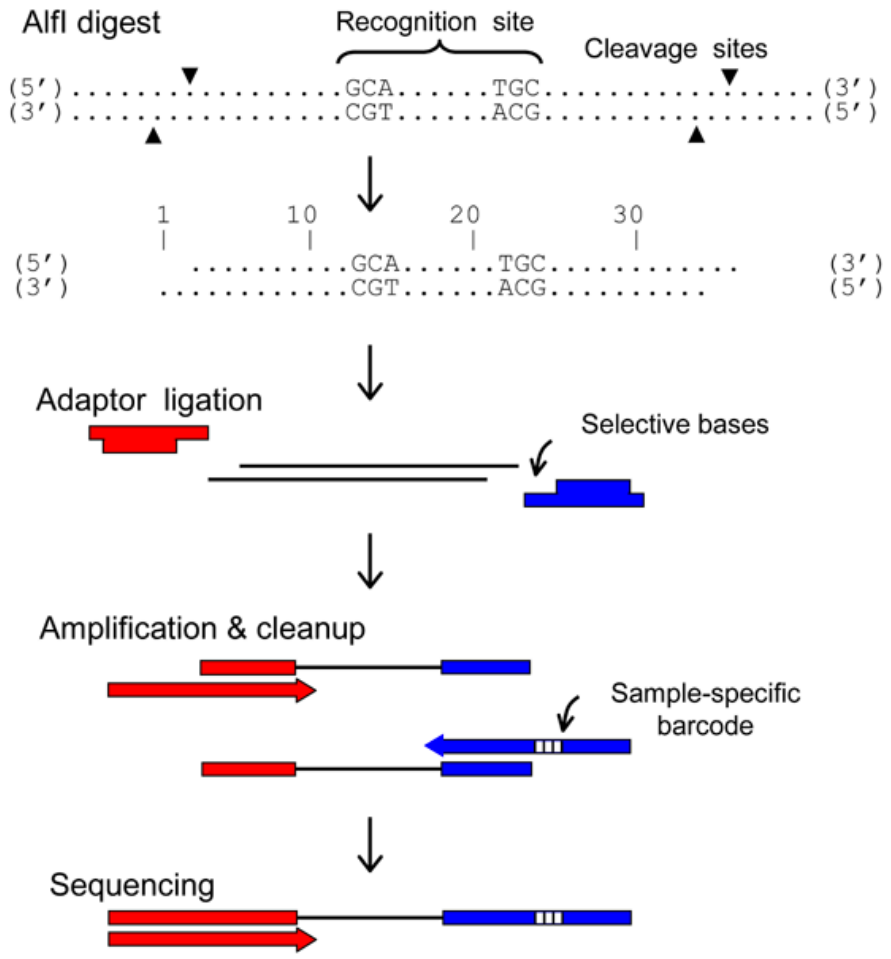


Figure 3. The steps of 2b-RAD library generation. From Eli Meyer (2018).

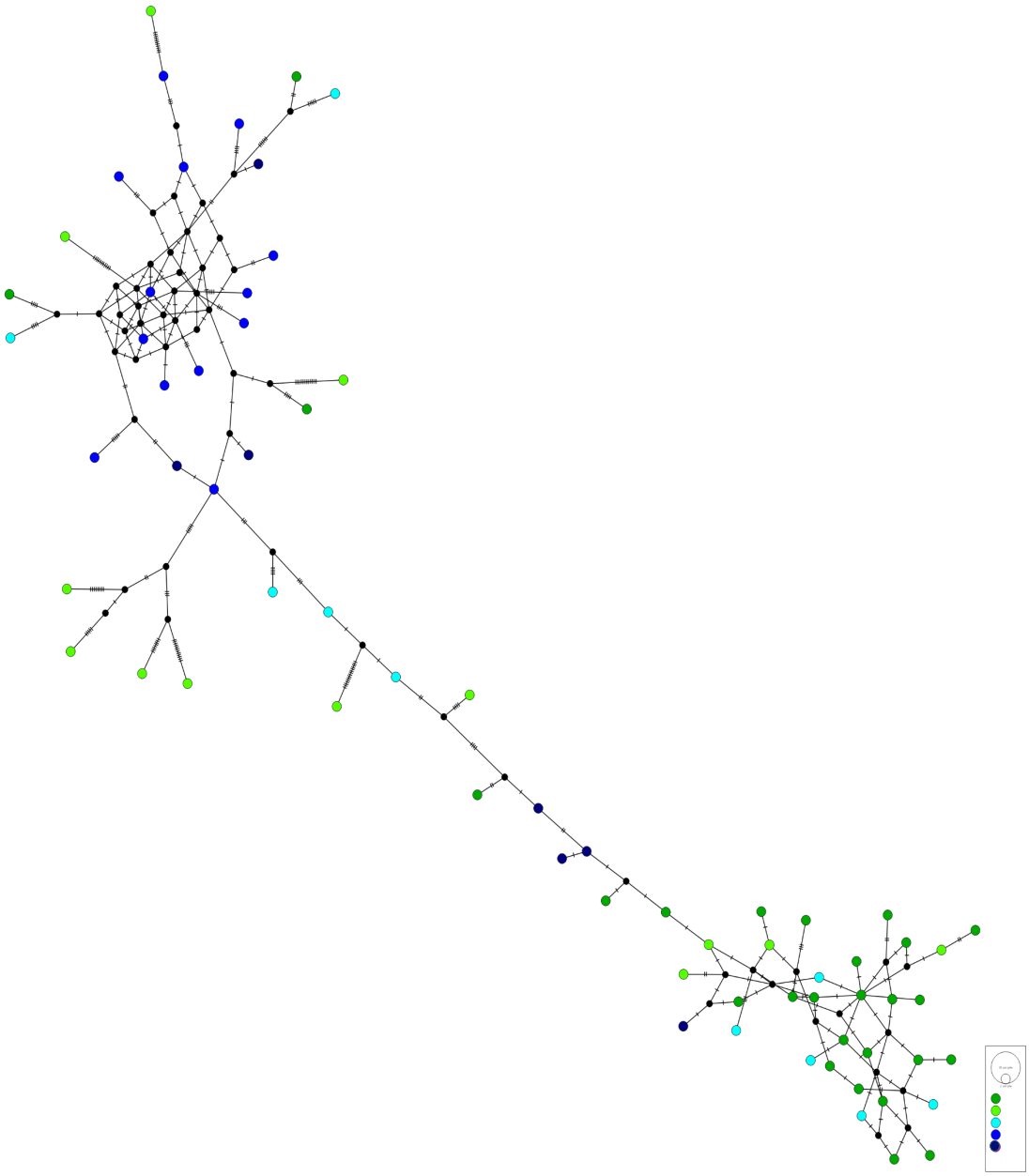
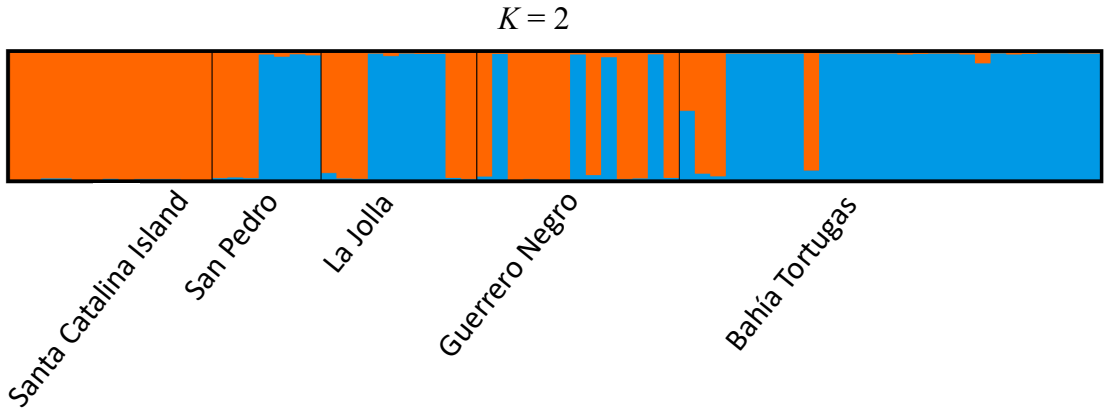


Figure 4. Haplotype network for *Pseudobatos productus*. Each colored circle represents a unique haplotype and each color corresponds to a sampling location. Bahía Tortugas (dark green), Guerrero Negro (light green), La Jolla (bright blue), Catalina (blue), and San Pedro (dark blue). Circle size corresponds to the number of individuals per haplotype. Each black circle represents a hypothetical "ancestor" or missed haplotype. The hashmarks on joining lines represent the number of mutational steps between haplotypes. A total of 70 unique haplotypes were found in 70 individuals.

a)



b)

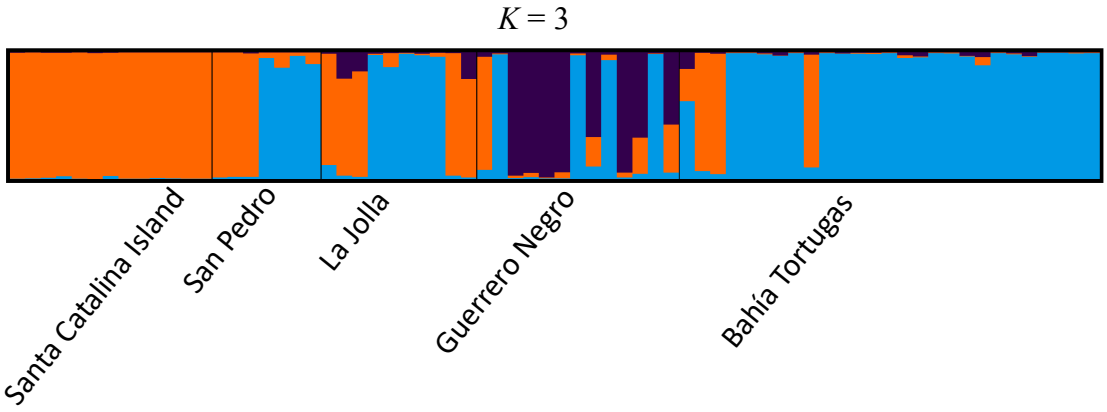
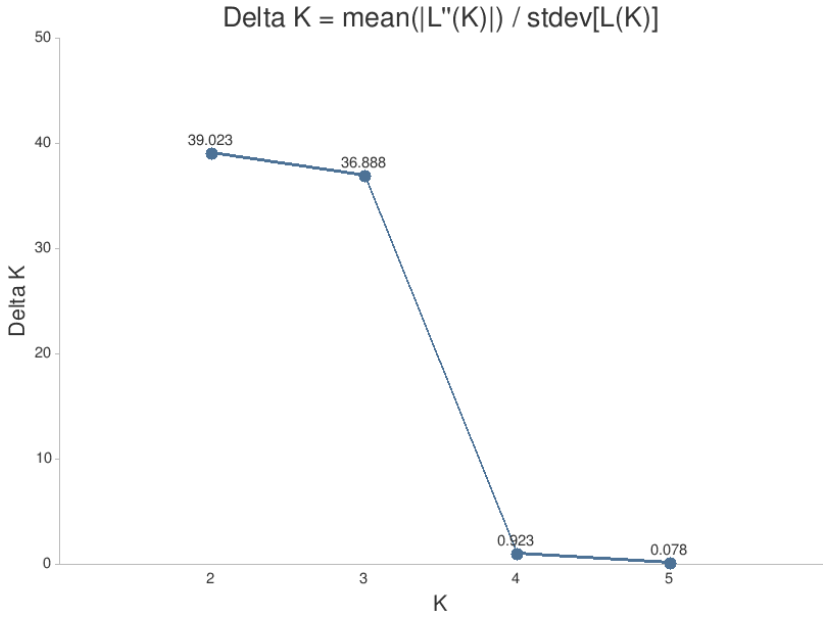


Figure 5. Population assignment of *Pseudobatos productus* estimated by STRUCTURE using mitochondrial DNA. Each bar represents a single individual by population with percentage ancestry to *K* populations represented vertically. Californian sampling locations: Santa Catalina Island, San Pedro, and La Jolla. Mexican sampling locations: Guerrero Negro, BCS and Bahía Tortugas, BCS. **a)** Assignment to $K = 2$. **b)** Assignment to $K = 3$.

a)



b)

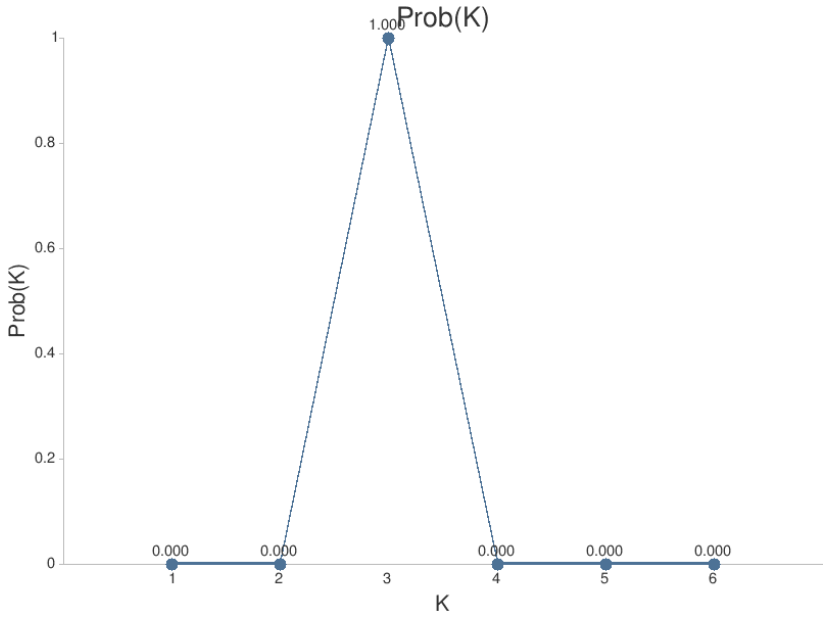


Figure 6. Determination of the most likely value of *K* estimated by CLUMPAK using mitochondrial DNA. **a)** Evanno DeltaK method estimates $K = 2$. **b)** Natural-log probability of *K* estimates $K = 3$.

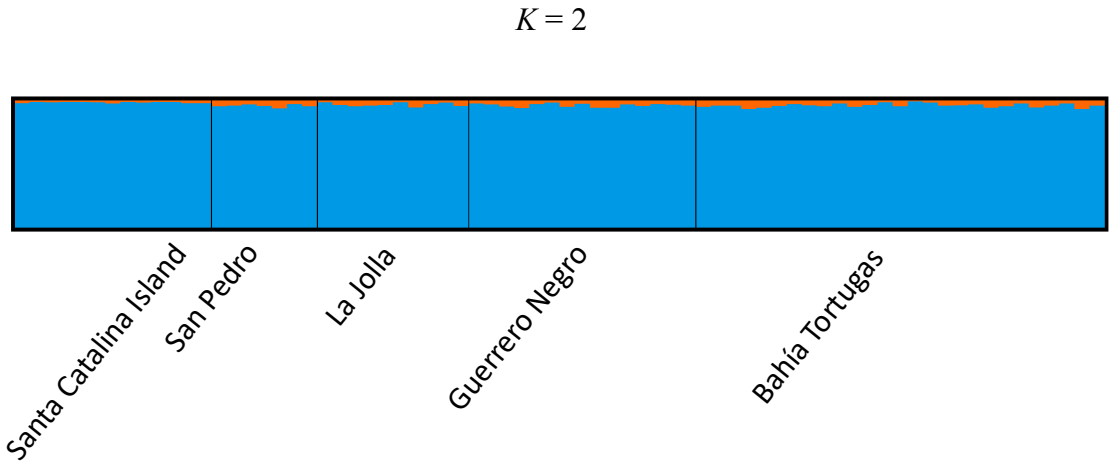


Figure 7. Population assignment of *Pseudobatos productus* estimated by STRUCTURE using genomic DNA. Each bar represents a single individual by population with percentage ancestry to *K* populations represented vertically. Californian sampling locations: Santa Catalina Island, San Pedro, and La Jolla. Mexican sampling locations: Guerrero Negro, BCS and Bahía Tortugas, BCS.

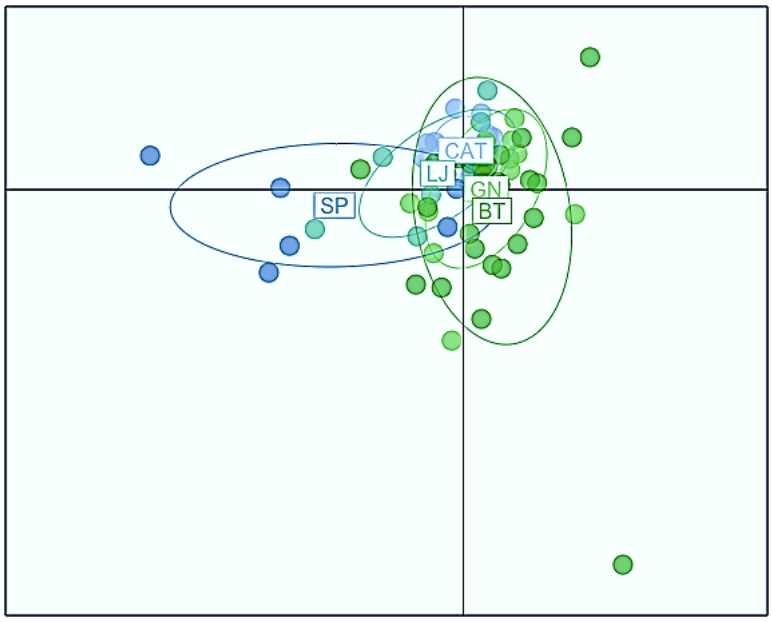


Figure 8. Genetic clustering via principal-components analysis (PCA) using *adegenet* for *Pseudobatos productus* using genomic DNA.

Appendix B: Tables

Table 1. Samples of *Pseudobatos productus*.

Sample ID	Location	Sex	TL (mm)	Analysis
RPCAT1	Santa Catalina Island, CA	F	1230	mtCR/gDNA
RPCAT2	Santa Catalina Island, CA	F	1270	mtCR/gDNA
RPCAT3	Santa Catalina Island, CA	F	1320	mtCR/gDNA
RPCAT4	Santa Catalina Island, CA	F	1235	mtCR/gDNA
RPCAT5	Santa Catalina Island, CA	F	1230	mtCR/gDNA
RPCAT6	Santa Catalina Island, CA	F	1310	mtCR/gDNA
RPCAT7	Santa Catalina Island, CA	F	1120	mtCR/gDNA
RPCAT8	Santa Catalina Island, CA	F	1265	mtCR/gDNA
TS263	Santa Catalina Island, CA	F	880	mtCR/gDNA
TS264	Santa Catalina Island, CA	M	1130	mtCR/gDNA
TS270	Santa Catalina Island, CA	M	1030	mtCR/gDNA
TS272	Santa Catalina Island, CA	M	960	mtCR/gDNA
TS273	Santa Catalina Island, CA	M	1090	mtCR/gDNA
RPDW1	Dockweiler Beach, CA	M	890	mtCR/gDNA
RPLBH1	Long Beach, CA	F	583	mtCR/gDNA
RPLBH2	Long Beach, CA	M	320	mtCR/gDNA
RPLBH3	Long Beach, CA	F	760	mtCR/gDNA
RPLBH4	Long Beach, CA	M	650	mtCR/gDNA
RPLBH5	Long Beach, CA	F	620	mtCR/gDNA
RPLBH6	Long Beach, CA	F	670	mtCR/gDNA
RP001	La Jolla, CA	F	1320	mtCR/gDNA
RP002	La Jolla, CA	F	1420	mtCR/gDNA
RP003	La Jolla, CA	F	1230	mtCR/gDNA
RP116	La Jolla, CA	F	1360	mtCR/gDNA
RP216	La Jolla, CA	M	1240	mtCR/gDNA
RP316	La Jolla, CA	F	1470	mtCR/gDNA
RP416	La Jolla, CA	F	1400	mtCR/gDNA
RP516	La Jolla, CA	F	1360	mtCR/gDNA
RP616	La Jolla, CA	M	1350	mtCR/gDNA
RP716	La Jolla, CA	F	1460	mtCR/gDNA
LCRP3	Guerrero Negro, BCS	M	810	mtCR/gDNA
LCRP4	Guerrero Negro, BCS	M	770	mtCR/gDNA
LCRP5	Guerrero Negro, BCS	M	860	mtCR/gDNA
LCRP11	Guerrero Negro, BCS	M	880	mtCR/gDNA
LCRP13	Guerrero Negro, BCS	M	740	mtCR/gDNA
LCRP15	Guerrero Negro, BCS	M	850	gDNA

Table 1. Samples of *Pseudobatos productus* (cont.).

Sample ID	Location	Sex	TL (mm)	Analysis
LCRP22	Guerrero Negro, BCS	M	580	mtCR/gDNA
LCRP25	Guerrero Negro, BCS	F	1000	gDNA
LCRP48	Guerrero Negro, BCS	F	—	mtCR/gDNA
LCRP49	Guerrero Negro, BCS	F	—	mtCR/gDNA
LCRP50	Guerrero Negro, BCS	M	—	mtCR/gDNA
LCRP51	Guerrero Negro, BCS	M	—	mtCR/gDNA
LCRP54	Guerrero Negro, BCS	F	—	mtCR/gDNA
LCRP56	Guerrero Negro, BCS	F	1415	mtCR/gDNA
LCRP57	Guerrero Negro, BCS	M	1030	mtCR/gDNA
BTRP1	Bahía Tortugas, BCS	F	786	mtCR/gDNA
BTRP2	Bahía Tortugas, BCS	F	922	mtCR/gDNA
BTRP3	Bahía Tortugas, BCS	M	1024	mtCR/gDNA
BTRP4	Bahía Tortugas, BCS	F	904	mtCR/gDNA
BTRP5	Bahía Tortugas, BCS	M	996	mtCR/gDNA
BTRP6	Bahía Tortugas, BCS	F	1054	mtCR/gDNA
BTRP7	Bahía Tortugas, BCS	F	1008	mtCR/gDNA
BTRP8	Bahía Tortugas, BCS	M	928	mtCR/gDNA
BTRP9	Bahía Tortugas, BCS	H	756	mtCR/gDNA
BTRP10	Bahía Tortugas, BCS	M	1146	mtCR/gDNA
BTRP11	Bahía Tortugas, BCS	F	842	mtCR/gDNA
BTRP12	Bahía Tortugas, BCS	F	922	mtCR/gDNA
BTRP13	Bahía Tortugas, BCS	F	926	mtCR/gDNA
BTRP14	Bahía Tortugas, BCS	F	936	mtCR/gDNA
BTRP15	Bahía Tortugas, BCS	F	1064	mtCR/gDNA
BTRP16	Bahía Tortugas, BCS	F	924	mtCR/gDNA
BTRP17	Bahía Tortugas, BCS	M	842	mtCR/gDNA
BTRP18	Bahía Tortugas, BCS	F	842	mtCR/gDNA
BTRP19	Bahía Tortugas, BCS	M	1014	mtCR/gDNA
BTRP20	Bahía Tortugas, BCS	F	980	mtCR/gDNA
BTRP21	Bahía Tortugas, BCS	F	1122	mtCR/gDNA
BTRP22	Bahía Tortugas, BCS	F	910	mtCR/gDNA
BTRP23	Bahía Tortugas, BCS	F	970	mtCR/gDNA
BTRP24	Bahía Tortugas, BCS	F	1036	mtCR/gDNA
BTRP25	Bahía Tortugas, BCS	F	1160	mtCR/gDNA
BTRP26	Bahía Tortugas, BCS	F	940	mtCR/gDNA
BTRP27	Bahía Tortugas, BCS	M	762	mtCR/gDNA

Table 2. Unique dual-indices applied to samples of *Pseudobatos productus* for genomic library sequencing.

Sample	i7 Index Name	i7 Index Sequence	i5 Index Name	i5 Index Sequence
RPCAT1	UTBC74	ACGATT	HT28	CAGCA
RPCAT2	UTBC74	ACGATT	HT29	GGCGA
RPCAT3	UTBC74	ACGATT	HT33	TGTTG
RPCAT4	UTBC74	ACGATT	HT38	AAAAT
RPCAT5	UTBC74	ACGATT	HT5	AACCC
RPCAT6	UTBC91	TGCAAA	HT13	ATTAC
RPCAT7	UTBC91	TGCAAA	HT17	CCAAT
RPCAT8	UTBC91	TGCAAA	HT20	CACCG
TS263	UTBC91	TGCAAA	HT28	CAGCA
TS264	UTBC91	TGCAAA	HT29	GGCGA
TS270	UTBC91	TGCAAA	HT33	TGTTG
TS272	UTBC91	TGCAAA	HT38	AAAAT
TS273	TSBC06	GCCAAT	HT7	GGAAC
RPDW1	TSBC06	GCCAAT	HT9	CGTAG
RPLBH1	TSBC06	GCCAAT	HT12	TATCT
RPLBH2	TSBC06	GCCAAT	HT19	TGGTC
RPLBH3	TSBC15	ATGTCA	HT7	GGAAC
RPLBH4	TSBC15	ATGTCA	HT9	CGTAG
RPLBH5	TSBC15	ATGTCA	HT12	TATCT
RPLBH6	TSBC15	ATGTCA	HT19	TGGTC
RP001	UTBC74	ACGATT	HT13	ATTAC
RP002	UTBC74	ACGATT	HT17	CCAAT
RP003	UTBC74	ACGATT	HT20	CACCG
RP116	UTBC72	AAGTAT	HT17	CCAAT
RP216	UTBC72	AAGTAT	HT20	CACCG
RP316	UTBC72	AAGTAT	HT28	CAGCA
RP416	UTBC72	AAGTAT	HT29	GGCGA
RP516	UTBC72	AAGTAT	HT33	TGTTG
RP616	UTBC72	AAGTAT	HT38	AAAAT
RP716	UTBC91	TGCAAA	HT5	AACCC
LCRP3	TSBC42	TAATCG	HT20	CACCG
LCRP4	TSBC42	TAATCG	HT28	CAGCA
LCRP5	UTBC68	GGAACT	HT29	GGCGA
LCRP11	TSBC42	TAATCG	HT33	TGTTG
LCRP13	TSBC42	TAATCG	HT38	AAAAT
LCRP15	UTBC68	GGAACT	HT5	AACCC

Table 2. Unique dual-indices applied to samples of *Pseudobatos productus* for genomic library sequencing (cont.).

Sample	i7 Index Name	i7 Index Sequence	i5 Index Name	i5 Index Sequence
LCRP22	UTBC68	GGAACT	HT13	ATTAC
LCRP25	UTBC68	GGAACT	HT17	CCAAT
LCRP48	UTBC68	GGAACT	HT20	CACCG
LCRP49	UTBC68	GGAACT	HT28	CAGCA
LCRP50	TSBC42	TAATCG	HT29	GGCGA
LCRP51	UTBC68	GGAACT	HT33	TGTTG
LCRP54	UTBC68	GGAACT	HT38	AAAAT
LCRP56	UTBC72	AAGTAT	HT5	AACCC
LCRP57	UTBC72	AAGTAT	HT13	ATTAC
BTRP1	TSBC12	CTTGTA	HT5	AACCC
BTRP2	TSBC12	CTTGTA	HT13	ATTAC
BTRP3	TSBC12	CTTGTA	HT17	CCAAT
BTRP4	TSBC12	CTTGTA	HT20	CACCG
BTRP5	TSBC12	CTTGTA	HT28	CAGCA
BTRP6	TSBC12	CTTGTA	HT29	GGCGA
BTRP7	TSBC12	CTTGTA	HT33	TGTTG
BTRP8	TSBC12	CTTGTA	HT38	AAAAT
BTRP9	TSBC13	AGTCAA	HT5	AACCC
BTRP10	TSBC13	AGTCAA	HT13	ATTAC
BTRP11	TSBC13	AGTCAA	HT17	CCAAT
BTRP12	TSBC13	AGTCAA	HT20	CACCG
BTRP13	TSBC13	AGTCAA	HT28	CAGCA
BTRP14	TSBC13	AGTCAA	HT29	GGCGA
BTRP15	TSBC13	AGTCAA	HT33	TGTTG
BTRP16	TSBC13	AGTCAA	HT38	AAAAT
BTRP17	TSBC34	CATGGC	HT5	AACCC
BTRP18	TSBC34	CATGGC	HT13	ATTAC
BTRP19	TSBC34	CATGGC	HT17	CCAAT
BTRP20	TSBC34	CATGGC	HT20	CACCG
BTRP21	TSBC34	CATGGC	HT28	CAGCA
BTRP22	TSBC34	CATGGC	HT29	GGCGA
BTRP23	TSBC34	CATGGC	HT33	TGTTG
BTRP24	TSBC34	CATGGC	HT38	AAAAT
BTRP25	TSBC42	TAATCG	HT5	AACCC
BTRP26	TSBC42	TAATCG	HT13	ATTAC
BTRP27	TSBC42	TAATCG	HT17	CCAAT

Table 3. Nucleotide changes representative of haplotypes of *Pseudobatos productus* relative to the consensus sequence.

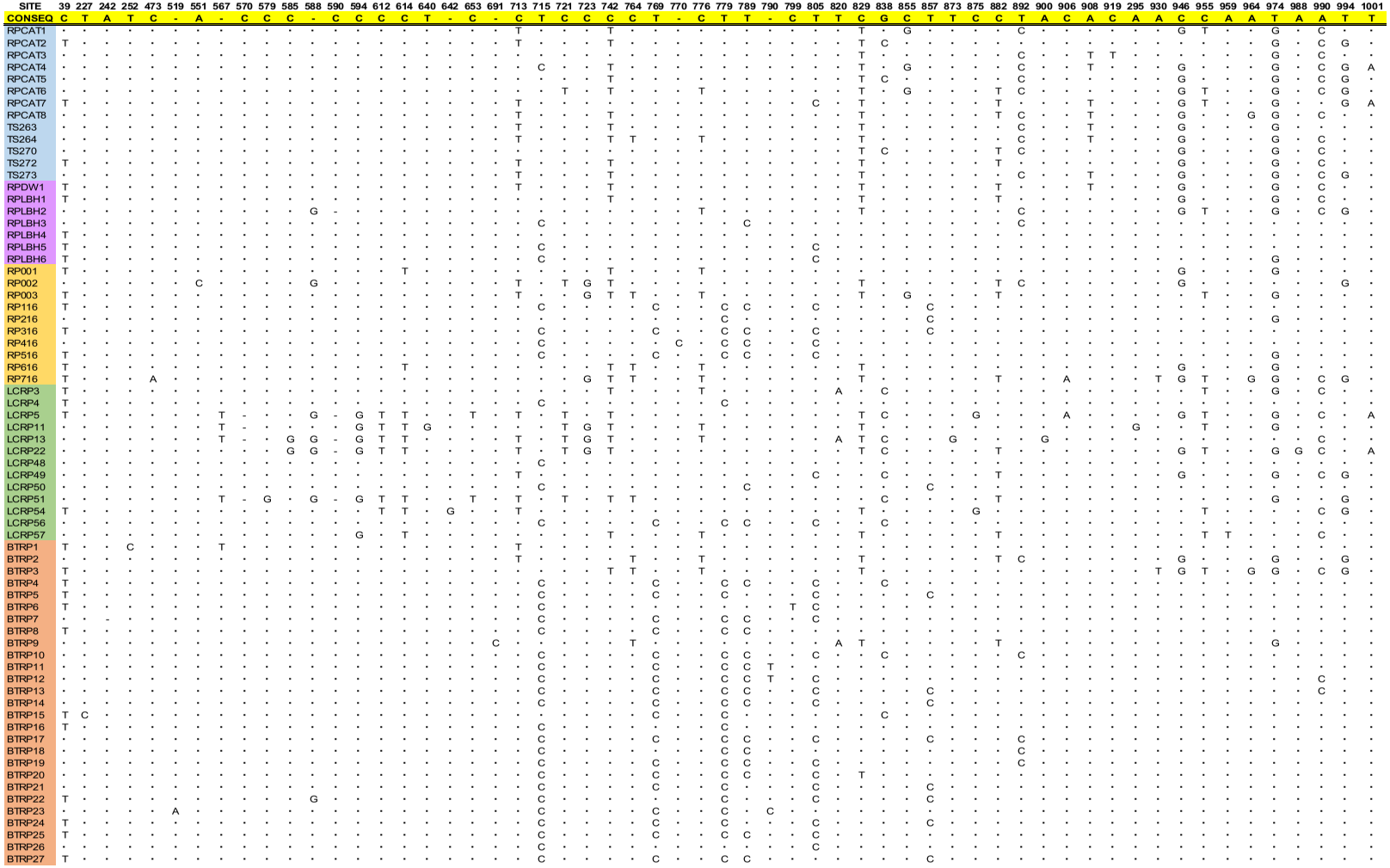


Table 3. Nucleotide changes representative of haplotypes of *Pseudobatos productus* relative to the consensus sequence (cont.).

SITE	1029	1040	1044	1046	1049	1060	1067	1069	1086	1089	1097	1103	1107	1112	1125	1127	1129	1143	1154	1188	1189	1211	1276	1280	1294	1312	1313	1315	1316	1317	1318	1319	1320	1321	1322	1323
CONSEQ	A	A	C	A	A	T	C	A	A	A	A	T	T	C	T	G	C	T	C	T	T	A	-	-	-	T	G	-	C	A	-	C	A	C	-	T
RPCAT1	.	.	T	.	.	.	.	.	.	.	.	.	.	.	A	.	.	G	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
RPCAT2	.	.	T	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	T	.	.	.	.	C	.	.	.	.	.	.
RPCAT3	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
RPCAT4	.	.	T	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
RPCAT5	.	.	T	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
RPCAT6	.	.	T	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
RPCAT7	.	.	T	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	T	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.
RPCAT8	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
TS263	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.
TS264	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
TS270	.	.	T	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
TS272	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.
TS273	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
RPDW1	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.
RPLBH1	.	.	T	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	T	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.
RPLBH2	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
RPLBH3	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
RPLBH4	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.
RPLBH5	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.
RPLBH6	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.
RP001	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.
RP002	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.
RP003	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	T	.	.	.	-	-	.	.	.	.	G	.
RP116	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	A
RP216	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
RP316	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	A	.	.	.	.	.	.	.	.	.
RP416	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
RP516	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.
RP616	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.
RP716	.	.	T	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	T	.	.	.	.	.	T	.	A	.	.	.	.	.	.	.	G	.
LCRP3	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.
LCRP4	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.
LCRP5	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	T	C	.	T	.	.	T	T	.	-	.	.
LCRP11	.	.	.	.	.	.	.	.	.	.	C	.	.	T	A	.	.	G	T	.	.	.	.	.	.	.	.	C	.	-	.	.	.	.	.	.
LCRP13	.	.	T	G	.	.	.	.	.	.	.	G	.	T	.	.	A	G	T	.	.	.	.	.	.	.	.	C	.	-	.	.	.	A	.	.
LCRP22	G	.	T	.	C	.	.	.	.	.	.	.	.	.	.	.	.	G	G	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.
LCRP48	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
LCRP49	.	.	.	G	.	G	.	.	.	.	.	.	A	.	.	A	.	.	T	.	.	.	G	.	.	.	.	.	.	T	.	.	T	.	.	.
LCRP50	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
LCRP51	G	G	T	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
LCRP54	.	.	T	G	.	.	.	.	.	.	.	.	.	T	.	.	.	.	T	.	.	.	.	G	T	.	.	.	.	.	.	.	G	.	.	.
LCRP56	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.
LCRP57	.	G	T	.	.	.	A	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.
BTRP1	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.
BTRP2	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
BTRP3	.	.	T	.	.	.	.	.	.	.	.	.	.	T	.	.	.	G	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.
BTRP4	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	A	.	.	.	.	.	.	.	.	.
BTRP5	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	A	.	.	.	.	.	.	.	.	.
BTRP6	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.
BTRP7	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.
BTRP8	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.
BTRP9	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.
BTRP10	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
BTRP11	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	-	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
BTRP12	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
BTRP13	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
BTRP14	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
BTRP15	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.
BTRP16	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	G	A	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.
BTRP17	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
BTRP18	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
BTRP19	.	.	.	G	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.
BTRP20	.	.	.	.	.	.	.	.	G	C	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
BTRP21	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
BTRP22	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.
BTRP23	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
BTRP24	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.
BTRP25	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.
BTRP26	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
BTRP27	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.

Table 4. Summary of mtDNA genetic diversity of *Pseudobatos productus*. Haplotype number (*N*), haplotype diversity (*h*), mean nucleotide diversity (π), mean pairwise difference (MPD), population mutation parameter (θ_S), female effective population size (N_{ef}) based on mutation rates of 0.024 and 0.0067 mutation per million years respectively, Tajima's D, and Fu's F_S.

Locality	*N*	*h*	π	MPD	θ_S	N_{ef}	Tajima's *D*	Fu's F_S
Santa Catalina Island	13	1.00	0.007 ± 0.004	8.551 ± 4.229	8.3784	66,117 - 236,839	-0.256	-6.451*
San Pedro	7	1.00	0.007 ± 0.004	9.429 ± 4.934	7.7551	61,199 - 219,219	0.452	-1.668
La Jolla	10	1.00	0.012 ± 0.006	15.200 ± 7.433	12.372	97,633 - 349,729	0.428	-2.398
Guerrero Negro	13	1.00	0.017 ± 0.009	22.923 ± 10.801	19.979	157,665 - 564,770	-0.180	-2.856
Bahía Tortugas	27	1.00	0.007 ± 0.004	8.749 ± 4.167	9.599	75,753 - 271,352	-0.841	-22.509*

Table 5. Fixation index for *Pseudobatos productus* based on mitochondrial DNA.

Source of Variation	Degrees of Freedom	Sum of Squares	Variance Component	Percentage of Variation
Among Population	4	122.108	1.853	23.18%
Within Populations	65	399.207	6.142	76.82%
Total	69	521.314	7.995	100.00%
Fixation Index (Φ_{ST}) 0.2318				
$p < 0.0001$				

Table 6. Pairwise Φ_{ST} and F_{ST} values for *Pseudobatos productus*. Φ_{ST} values are given below the diagonal and F_{ST} values are given above. * indicates p-values ≤ 0.017 after Benjamini and Yekutieli modified false discovery rate correction. ** indicates p-values ≤ 0.0001.

	Santa Catalina Island	San Pedro	La Jolla	Guerrero Negro	Bahía Tortugas
Santa Catalina Island	—	0.0443**	0.0198**	0.0062*	0.0069*
San Pedro	0.2095*	—	0.0040*	0.0128*	0.0149**
La Jolla	0.2498**	0.0000	—	0.0000	0.0000
Guerrero Negro	0.1847**	0.0567	0.0673	—	0.0000
Bahía Tortugas	0.4859**	0.1993*	0.1225*	0.2506**	—

Table 7. Results of hierarchical AMOVA using mitochondrial DNA testing putative biogeographic barriers, San Pedro Channel and Punta Eugenia. Catalina was separated from San Pedro and Guerrero Negro was separated from Bahía Tortugas.

Source of Variation	Degrees of Freedom	Sum of Squares	Variance Component	Percentage of Variation
Among Regions	2	95.666	1.334	16.26%
Among Populations	2	26.442	0.730	8.89%
Within Populations	65	399.207	6.142	74.84%
Total	69	521.314	8.206	100.00%
Fixation Index (Φ_{RT}) 0.1626				
p = 0.0001				

Table 8. Fixation index for *Pseudobatos productus* based on genomic DNA.

Source of Variation	Degrees of Freedom	Sum of Squares	Variance Component	Percentage of Variation
Among Population	4	33.267	0.03923	.54%
Within Populations	139	1007.893	7.25103	99.46%
Total	143	1041.16	7.29026	100.00%
Fixation Index (F_{ST}) 0.0054				
$p < 0.0001$				

Table 9. Results of hierarchical AMOVAs using genomic DNA testing putative biogeographic barriers.

a) Separating Santa Catalina Island from all other localities

Source of Variation	Degrees of Freedom	Sum of Squares	Variance Component	Percentage of Variation
Among Regions	1	8.223	-0.00407	0.18%
Among Populations	3	25.044	0.04082	0.44%
Within Populations	139	1007.893	7.25103	99.38%
Total	143	1041.160	7.28778	100.00%
Fixation Index (F_{RT}) 0.0000				
p = 0.4003				

b) Separating Santa Catalina Island and San Pedro from all other localities

Source of Variation	Degrees of Freedom	Sum of Squares	Variance Component	Percentage of Variation
Among Regions	1	8.873	0.01296	0.18%
Among Populations	3	24.394	0.03233	0.44%
Within Populations	139	1007.893	7.25103	99.38%
Total	143	1041.160	7.29633	100.00%
Fixation Index (F_{RT}) 0.0018				
p = 0.2997				

c) Separating Santa Catalina Island and Bahía Tortugas from all other localities

Source of Variation	Degrees of Freedom	Sum of Squares	Variance Component	Percentage of Variation
Among Regions	2	16.190	-0.02896	0.00%
Among Populations	2	17.077	0.06337	0.87%
Within Populations	139	1007.893	7.25103	99.53%
Total	143	1041.160	7.28544	100.00%
Fixation Index (F_{RT}) 0.0000				
p = 0.5028				

Appendix C

a) Oligonucleotides used in the generation of ligation adaptors. These oligonucleotides are attached during the ligation process.

Name	Sequence
anti-lLL	AGATCGGAAGAGC(InvdT)
5lLL-NR	CTACACGACGCTCTTCCGATCTNR
3lLL-NR	CAGACGTGTGCTCTTCCGATCTNR

b) Oligonucleotides used to bind gDNA constructs to the Illumina flow-cell for sequencing. These oligonucleotides are attached and amplified during the PCR process.

Name	Construct End	Sequence
lLL-Lib1	5'	AATGATACGGCGACCACCGA
lLL-Lib2	3'	CAAGCAGAAGACGGCATACGA

c) Complete oligonucleotides used to attach index sequences to the Adaptor 1 end of the gDNA/adaptor complex. These oligonucleotides are attached and amplified during the PCR process.

Name	Sequence
HT5	AATGATACGGCGACCACCGAGATCTACACAACCCACACTCTTTCCCTACACGACGCTCTTCCGATCT
HT07	AATGATACGGCGACCACCGAGATCTACACGGAACACACTCTTTCCCTACACGACGCTCTTCCGATCT
HT09	AATGATACGGCGACCACCGAGATCTACACCGTAGACACTCTTTCCCTACACGACGCTCTTCCGATCT
HT12	AATGATACGGCGACCACCGAGATCTACACTATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT
HT13	AATGATACGGCGACCACCGAGATCTACACATTACACACTCTTTCCCTACACGACGCTCTTCCGATCT
HT17	AATGATACGGCGACCACCGAGATCTACACCCAATACACTCTTTCCCTACACGACGCTCTTCCGATCT
HT19	AATGATACGGCGACCACCGAGATCTACACTGGTCACACTCTTTCCCTACACGACGCTCTTCCGATCT
HT20	AATGATACGGCGACCACCGAGATCTACACCACCGACACTCTTTCCCTACACGACGCTCTTCCGATCT
HT28	AATGATACGGCGACCACCGAGATCTACACCAGCAACACTCTTTCCCTACACGACGCTCTTCCGATCT
HT29	AATGATACGGCGACCACCGAGATCTACACGGCGAACACTCTTTCCCTACACGACGCTCTTCCGATCT
HT33	AATGATACGGCGACCACCGAGATCTACACTGTTGACACTCTTTCCCTACACGACGCTCTTCCGATCT
HT38	AATGATACGGCGACCACCGAGATCTACACAAAATACACTCTTTCCCTACACGACGCTCTTCCGATCT

d) Complete oligonucleotides used to attach index sequences to the Adaptor 2 end of the gDNA/adaptor complex. These oligonucleotides are attached and amplified during the PCR process.

Name	Sequence
UTBC68	CAAGCAGAAGACGGCATACGAGATAGTTCCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC
UTBC72	CAAGCAGAAGACGGCATACGAGATATACTTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC
UTBC74	CAAGCAGAAGACGGCATACGAGATAATCGTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC
UTBC91	CAAGCAGAAGACGGCATACGAGATTTTGCAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC
TSBC06	CAAGCAGAAGACGGCATACGAGATATTGGCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC
TSBC12	CAAGCAGAAGACGGCATACGAGATTACAAGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC
TSBC13	CAAGCAGAAGACGGCATACGAGATTTGACTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC
TSBC15	CAAGCAGAAGACGGCATACGAGATTGACATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC
TSBC34	CAAGCAGAAGACGGCATACGAGATGCCATGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC
TSBC42	CAAGCAGAAGACGGCATACGAGATCGATTAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC

Appendix D

All perl scripts were written by Eli Meyer (Meyer, 2018)

Custom perl-script pipeline used to prepare genomic data for analyses:

1. Truncate reads using TruncateFastq.pl
2. Quality filter reads using bbduk.sh
3. Generate a cluster derived reference genome using BuildRef.pl
4. Align reads to reference genome using gmapper function within SHRiMP
5. Filter weak and ambiguous reads using SAMFilter.pl
6. Parse alignments and call base pairs using SAMBaseCounts.pl
7. Combine all base calls into a single file prior to calling genotypes with CombineBaseCounts.pl
8. Determine genotypes from nucleotide frequencies using CallGenotypes.pl
9. Filter to include only polymorphic loci using PolyFilter.pl
10. Filter reads to exclude low coverage loci that were genotyped in too few samples to be informative using MDFilter.pl
11. Filter samples to exclude repetitive sites using RepTagFilter.pl
12. Filter samples to exclude samples with multiple SNPs per tag to remove linkage using OneSNPPerTag.pl

TruncateFastq.pl

```
#!/usr/bin/perl
# written by E Meyer, eli.meyer@science.oregonstate.edu
# distributed without any guarantees or restrictions

# -- check arguments and print usage statement
$scriptname=$0; $scriptname =~ s/.+\///g;
$usage = <<USAGE;
Truncates a set of short reads in FASTQ format to keep the region specified
Usage:   $scriptname -i input -s start -e end -o output
Required arguments:
       -i input      file of short reads to be filtered, fastq format
       -s start      beginning of the region to keep, nucleotide position
       -e end        end of the region to keep, nucleotide position
       -o output     a name for the output file (fastq format)
USAGE
if ($#ARGV < 3 || $ARGV[0] eq "-h") {print "\n", "-"x60, "\n", $scriptname, "\n",
$usage, "-"x60, "\n\n"; exit;}

# -- module and executable dependencies
$mod1="Getopt::Std";
unless(eval("require $mod1")) {print "$mod1 not found. Exiting\n"; exit;}
use Getopt::Std;

# get variables from input
getopts('i:s:e:o:h');# in this example a is required, b is optional, h is help
if (!$opt_i ||!$opt_s ||!$opt_e ||!$opt_o || $opt_h) {print "\n", "-"x60, "\n",
$scriptname, "\n", $usage, "-"x60, "\n\n"; exit;}

my $seqfile = $opt_i;       # raw reads, fastq format
my $startloc = $opt_s-1;    # beginning of region to keep, base 1
my $endloc = $opt_e-1;             # end of region to keep, base 1
my $outfile = $opt_o;       # name for output file, fastq format

# loop through fastq file and truncate sequences and quality scores
open (IN, $seqfile);
open (OUT, ">$outfile");
my $switch = 0;
while(<IN>)
       {
       chomp;
       $count++;
       if ($count==1) {$ss = substr($_, 0, 4);}
       if ($_ =~ /^$ss/)
              {
              print OUT $_, "\n";
              next;
              }
       if ($_ =~ /^\+$/)
              {
              print OUT "$_", "\n";
              next;
              }
       else
              {
              $ssi = $_;
              $slen = length($ssi);
              $subi = substr ($ssi, $startloc, $endloc-$startloc+1);
              print OUT $subi, "\n";
              }
       }
close(IN);
```

BuildRef.pl

```
#!/usr/bin/env perl
# written by E Meyer, eli.meyer@science.oregonstate.edu
# distributed without any guarantees or restrictions

# -- check arguments and print usage statement
$scriptname=$0; $scriptname =~ s/.+\///g;
$usage = <<USAGE;
Builds a reference for de novo analysis of 2bRAD sequences from samples
without a sequenced genome. The script filters, clusters, and compares
similar sequences to infer the set of loci present in the species of
interest, using a subset of reads from the samples themselves.
Usage: $scriptname -i input -o output <OPTIONS>
Required arguments:
       -i input      The set of processed (truncated, HQ) reads (FASTQ) to be used
for reference
                     development. Ideally this should include 10-20 million reads
spanning the range
                     of known diversity (e.g. from all populations in your study).
Prepare this
                     ahead of time by concatenating together a subset of reads from
your samples.
       -o output     a name for the output file, to be used as a reference in
mapping and genotyping
Options:
       -v overwrite  0=do not overwrite existing files; use them for analysis.
(default)
                     1=do not use existing files; overwrite them with new files.
       -n mincov     Minimum depth to qualify as a valid allele. (default=2)
       -q threshold  Quality scores below this threshold are low quality
(default=30)
       -x number     Maximum number of low quality bases allowed for reference
construction (default=0)
       -m mismatches Maximum number of mismatches allowed in clustering of related
alleles (default=2)
       -d distance   Minimum number of bases required to resolve sub-clusters
(default=1)
       -a haplotypes For very large clusters containing more than this number of
unique sequences,
                     do not attempt to resolve sub clusters. These indicate
repetitive sequences
                     which are not useful for genotyping anyway, and resolving these
large clusters
                     is computationally intensive. (default=32)
USAGE
if ($#ARGV < 2 || $ARGV[0] eq "-h") {print "\n", "-"x60, "\n", $scriptname, "\n",
$usage, "-"x60, "\n\n"; exit;}

# -- module and executable dependencies
# use this block if checking for executable dependencies
# copy the block and edit to check for additional Perl modules required by the
script
$mod1="File::Which";
unless(eval("require $mod1")) {print "$mod1 not found. Exiting\n"; exit;}
use File::Which;
$mod2="Bio::TreeIO";
unless(eval("require $mod2")) {print "$mod2 not found. Exiting\n"; exit;}
use Bio::TreeIO;
$mod1="Getopt::Std";
unless(eval("require $mod1")) {print "$mod1 not found. Exiting\n"; exit;}
use Getopt::Std;
$mod1="Bio::SeqIO";
unless(eval("require $mod1")) {print "$mod1 not found. Exiting\n"; exit;}
use Bio::SeqIO;
```

```
$mod2="Bio::Seq::Quality";
unless(eval("require $mod2")) {print "$mod2 not found. Exiting\n"; exit;}
use Bio::Seq::Quality;

# use this block and edit to check for executables required by the script
$dep1 = "raxmlHPC";
unless (defined(which($dep1))) {print $dep1, " not found. Exiting.\n"; exit;}
$dep2 = "cd-hit-est";
unless (defined(which($dep2))) {print $dep2, " not found. Exiting.\n"; exit;}

# get variables from input
getopts('i:o:v:n:q:x:m:d:a:h');
if (!$opt_i || !$opt_o || $opt_h) {print "\n", "-"x60, "\n", $scriptname, "\n",
$usage, "-"x60, "\n\n"; exit;}
if ($opt_v) {$ow_opt=$opt_v;} else {$ow_opt = 0;}# overwrite options: 1=overwrite
if ($ow_opt ne 0 andand $ow_opt ne 1) {print "\n", "-"x60, "\n", $scriptname, "\n",
$usage, "-"x60, "\n\n"; exit;}
if ($opt_n) {$mincov=$opt_n;} else {$mincov = 2;}# min coverage for valid alleles
if ($opt_q) {$qthd=$opt_q;} else {$qthd = 30;}          # quality score threshold
if ($opt_x) {$maxlq=$opt_x;} else {$maxlq = 0;}         # max LQ positions threshold
if ($opt_a) {$maxmems=$opt_a;} else {$maxmems = 32;}    # maximum cluster size for
sub cluster analysis
$infile = $opt_i;           # name of the input sequence file (fasta format)
$outfile = $opt_o;          # a name for the final output file (reference to be used
for mapping)
if ($opt_d) {$depthi = $opt_d;} else {$depthi = 1;}
$mindepth = 0.028 * $depthi;# min branch length and depth to select clades. 0.028
corresponds to 1 bp distance
if ($opt_m) {$maxmm = $opt_m;} else {$maxmm = 2;}
$initc = (36 - $maxmm)/36;  # 0.944 = 2 bp distance for initial clustering of
alleles and paralogs

# define additional variables. edit these if you understand the consequences.
$site = "\\w{12}GCA\\w{6}TGC\\w{12}";     # recognition site in Perl regexp
$minmems = 4; # minimum number of sequences in a clade to consider tree building.
Don't set below 4.

# convert fastq input to fasta
my $wrap = 1000;            # only needs to be longer than maximum read length
my $fastqfile = $infile;
$nameroot = $fastqfile;
$nameroot =~ s/\..+//;
my $osfile = $nameroot.".fasta";
my $oqfile = $nameroot.".qual";

if ($ow_opt eq 0)
       {
       if (-e $osfile andand -e $oqfile)
              {
              print "Found $osfile and $oqfile . Using these ...\n";
              goto CONV;
              }
       }
print "Converting FASTQ input into FASTA...\n";
my $inseqs = new Bio::SeqIO(-file=>$fastqfile, -format=>"fastq");
my $outseqs = new Bio::SeqIO(-file=>">$osfile", -format=>"fasta", -width=>$wrap);
my $outquals = new Bio::SeqIO(-file=>">$oqfile", -format=>"qual", -width=>$wrap);
my %sh; my $scount = 0;
while ($seq = $inseqs->next_seq)
       {$sh{$seq->display_id} = $seq->seq;
       $outseqs->write_seq($seq);
       $outquals->write_seq($seq);
       $scount++;
       }

print "Converted ", $scount, " reads from FASTQ to FASTA and QUAL.\n";
```

```
CONV:                          # end of file conversion section
$seqfile = $osfile;
$qualfile = $oqfile;

# fix sequence name characters because raxmlHP is picky
print "Checking input sequence names for \":\" characters...\n";
$checkseq = `head -n 1 $seqfile | grep \"\:\" -m 1`;
$checkqual = `head -n 1 $qualfile | grep \"\:\" -m 1`;
print "Finished checking input files.\n";
if (length($checkseq)>0)
        {
        print "Found \":\" in $seqfile. Correcting...\n";
        system("perl -pi -e \"s/\:/-/g\" $seqfile");
        print "Finished.\n";
        }
if (length($checkqual)>0)
        {
        print "Found \":\" in $qualfile. Correcting...\n";
        system("perl -pi -e \"s/\:/-/g\" $qualfile");
        print "Finished.\n";
        }

# -- counting input
$inseq = `grep \">\" -c  $seqfile`;
$inqual = `grep \">\" -c  $qualfile`;
chomp($inseq); chomp($inqual);

# -- screen for Ns in input reads
print "Checking for Ns in input...\n";
open(SF, $seqfile);
while(<SF>)
        {chomp;
        if ($_ =~ />/) {$name = $_; $name =~ s/>//; $name =~ s/\s+.*//; next;}
        else    {
                if ($_ =~ /N/i)
                        {
                        $nhash{$name}++;
                        $withn++;
                        }
                }
        }
print "Finished.\n";

# -- designate positions to exclude from quality filtering
# low diversity at these positions leads to falsely low quality scores,
# and we'll be checking for accuracy at those positions more directly anyway
@excvec = qw{13 14 15 16 22 23 24};
foreach $e (@excvec) {$exch{$e}++;}

# -- stringently filter and truncate raw reads and scores
if ($ow_opt eq 0)
        {
        if (-e "VHQ.qual" andand -e "VHQ.fasta")
                {
                print "VHQ.fasta and VHQ.qual found. Using existing files...\n";
                $outqual = `grep -c \">\" VHQ.qual`;
                $outseq = `grep -c \">\" VHQ.fasta`;
                chomp($outseq); chomp($outqual);
                $nlq = $inseq - $outseq - $withn;
                goto VHQ;
                }
        }
print "Conducting stringent quality filtering...\n";
my $nqual = $lqn = 0;
open(QF, $qualfile);
open(OQ, ">VHQ.qual");
```

```
while(<QF>)
	{chomp;
	if ($_ =~ />/) {$name = $_; $name =~ s/>//; $name =~ s/\s+.*//; next;}
	else	{@qa = split(" ", $_);
		$lqn = 0; $poscount = 0;
		foreach $q (@qa)
			{
			$poscount++;
			if (exists($exch{$poscount})) {next;}
			if ($q < $qthd) {$lqn++;}
			}
		if (exists($nhash{$name})) {next;}
		if ($lqn>$maxlq) {$nlq++; next;}
		elsif ($lqn <= $maxlq)
			{print OQ ">", $name, "\n";
			print OQ join(" ", @qa), "\n";
			$gh{$name}++;
			$outqual++;
			}
		}
	}
close(QF);

open(SF, $seqfile);
open(OS, ">VHQ.fasta");
while(<SF>)
	{chomp;
	if ($_ =~ />/) {$name = $_; $name =~ s/>//; $name =~ s/\s+.*//; next;}
	else	{
		if (exists($gh{$name}))
			{print OS ">", $name, "\n";
			print OS $_, "\n";
			$outseq++;
			}
		}
	}
close(SF);
close(OS);
print "Finished quality filtering.\n";
VHQ:                                # end of quality filtering section.

print $inseq, " sequences in.\n";
print $inqual, " scores in.\n";
print $withn, " sequences with Ns excluded.\n";
print $nlq, " low quality sequences excluded.\n";
print $outseq, " sequences out.\n";
print $outqual, " scores out.\n";

# -- Filter for perfect match to restriction site
print "Filtering for perfect matches to restriction site...\n";
if ($ow_opt eq 0)
	{
	if (-e "matches.fasta")
		{
		print "matches.fasta found. Using existing file...\n";
		goto MATCHES;
		}
	}
system("grep -P \"$site\" VHQ.fasta -B 1 > matches.fasta");
system("perl -pi -e \"s/^--\n//g\" matches.fasta");
MATCHES:                            # end of restriction site filtering section

$nmatches = `grep \">\" matches.fasta -c`;
chomp($nmatches);
print "Done filtering for site matches. $nmatches matches found.\n";
```

```
# -- if VHQ sites coverage is high (> 10 million) enforce a minimum coverage setting
of at least 1/5M reads
$newmincov = int($nmatches/5000000+0.5);
if ($newmincov > $mincov) {$mincov = $newmincov;}

# -- this step identifies sequences observed repeatedly in the VHQ dataset
# -- we call these "valid alleles"
# -- cluster reads at 1.0 and extract valid alleles based on depth
print "Beginning initial clustering step...\n";
if ($ow_opt eq 0)
        {
        if (-e "va.clstr")
                {
                print "va.clstr found. Using existing file...\n";
                goto CLSTR1;
                }
        }
system("date");
system("cd-hit-est -i matches.fasta -M 0 -d 0 -c 1 -o va >va.log");
print "Finished with initial clustering.\n";
CLSTR1:                             # end of initial clustering section

print "Parsing initial clusters...\n";
open(IN, "va.clstr");
while(<IN>)
        {
        chomp;
        if ($_ =~ /^>/)
                {
                $cname = $_;
                $cname =~ s/>//;
                $cname =~ s/ /_/;
                $cnum++;
                }
        else
                {
                $ch{$cname}++;
                $allreads++;
                }
        }
close(IN);

# -- apply coverage filter to identify sequences identified more than $mincov times
in the VHQ data
foreach $c (sort(keys(%ch)))
        {
        if ($ch{$c}<$mincov) {next;}
        $pass++;
        $passh{$c}++;
        }

# -- identify representative sequences from valid clusters
open(IN, "va.clstr");
$status = 0;
while(<IN>)
        {
        chomp;
        if ($_ =~ /^>/)
                {
                $cname = $_;
                $cname =~ s/ /_/;
                $cname =~ s/>//;
                if (exists($passh{$cname})) {$status = 1;}
                }
        elsif ($status > 0 andand $_ =~ /\*/)
                {
```

```
                $repname = $_;
                $repname =~ s/.*>//;
                $repname =~ s/\..+//;
                $reph{$repname}++;
                $status = 0;
                }
        }
close(IN);

# -- extract representative sequences from valid clusters
open(IN,"va");
open(OUT, ">filtered_va.fasta");
$status = 0;
while(<IN>)
        {
        chomp;
        if ($_ =~ />/)
                {
                $rname = $_;
                $rname =~ s/>//;
                $rname =~ s/ .+//;
                if(exists($reph{$rname})) {print OUT $_, "\n"; $status++;}
                }
        elsif ($status > 0)
                {
                print OUT $_, "\n";
                $status = 0;
                }
        }
close(IN);
close(OUT);
print "Finished parsing initial clusters. \n";
print "Identified $pass putative alleles (sequences observed at least $mincov
times).\n";
system("date");

# cluster valid alleles to identify groups of related alleles that may represent
paralogs or alleles
print "Beginning second clustering step...\n";
if ($ow_opt eq 0)
        {
        if (-e "gra.clstr")
                {
                print "gra.clstr found. Using existing file...\n";
                goto CLSTR2;
                }
        }
system("cd-hit-est -i filtered_va.fasta -M 0 -d 0 -c $initc -o gra >gra.log");
print "Finished second clustering step.\n";
CLSTR2:                             # end of second clustering step
print "Parsing cluster output...\n";

# parse clusters to record cluster membership of each sequence
open(IN, "gra.clstr");
$status = 0; $gra_c = 0;
while(<IN>)
        {
        chomp;
        if ($_ =~ /^>/)
                {
                $cname = $_;
                $cname =~ s/ /_/;
                $cname =~ s/>//;
                $gra_c++;
                }
        else
```

```
                {
                $memname = $_;
                $memname =~ s/.*>//;
                $memname =~ s/\..+//;
                if ($_ =~ /\*/)
                        {
                        $grareps{$cname} = $memname;
                        }
                $grah{$cname}{$memname}++;
                }
        }
print "Finished parsing.\n";
print $gra_c, " clusters (groups of related alleles) identified.\n";
system("date");

# -- build trees for each cluster in an effort to distinguish between paralogs and
alleles
print "Testing for clusters containing multiple loci...\n";
$ctestno = 0;
@allcs = keys(%grah);
$nallcs = @allcs;
foreach $c (sort(keys(%grah)))
        {                       # count members in each cluster
        %cih = %{$grah{$c}};
        @cmems = sort(keys(%cih));
        $ncmems = @cmems;
        $ctestno++;
# If fewer than 4 alleles were observed, it's not possible to use sequence
relationships
# among alleles to infer the number of loci in the cluster.
        if ($ncmems < $minmems)
                {
# pick representative seq from this cluster and add it to reference as a single
locus
                $outh{$grareps{$c}}++;
                if ($ncmems < 2) {$mono_loc++;}
                else   {$poly_one++;}
                next;
                }
# If more than the maximum number of sequences were observed, this indicates a
highly
# repetitive cluster that we'd like to exclude because such loci will be error
prone.
        if ($ncmems > $maxmems)
                {
# pick representative seq from this cluster and add it to reference as a single
locus
                $outh{$grareps{$c}}++;
                $poly_max++;
                next;
                }
# for clusters of the right size for tree building
        print STDERR "Resolving cluster number $ctestno of $nallcs, with $ncmems
members ... ";
# extract sequences of all cluster members
        $c4tree++;
        $found = 0;
        open(SEQ, "filtered_va.fasta");
        open(TMP, ">tmp.fasta");
        while(<SEQ>)
                {
                chomp;
                if ($_ =~ />/)
                        {
                        $status = 0;
                        $_ =~ s/>//;
```

```
                        $_ =~ s/\s.*//;
                        if (exists($cih{$_}))
                                {
                                $outname = $_;
                                $outname =~ s/\:/-/g;
                                print TMP ">", $outname, "\n";
                                $status++;
                                $found++;
                                next;
                                }
                        }
                elsif ($status > 0)
                        {
                        print TMP $_, "\n";
                        $status = 0;
                        }
                if ($found >= $ncmems) {last;}
                }
        close(SEQ); close(TMP);

# tree building code begins here
# prepare for tree building
        if(glob("*.tree"))    {system("rm *.tree");}
        %cladeh = %deph = %blenh = ();

# build tree describing relationships among members of the cluster
        $nseqs = `grep -c \">\" tmp.fasta`;
        chomp($nseqs);
        if ($nseqs ne 0)
        {
        system("raxmlHPC -s tmp.fasta -p 123 -m GTRCAT -n out.tree > ml.log");

# parse tree to identify clades that differ by at least the critical
# minimum difference ($mindepth variable)
        if (-e "RAxML_bestTree.out.tree")
                {
                $treeobj = new Bio::TreeIO(-file=>"RAxML_bestTree.out.tree", -
format=>"newick");
                $tree = $treeobj->next_tree;
                $rootnode = $tree->get_root_node;
                }

# store id, depth, and branch length for each node
        foreach $node ($rootnode->get_all_Descendents)
                {
                if (defined($node->id))
                        {
                        @lineage = $tree->get_lineage_nodes($node);
                        $cladeh{$lineage[-1]->internal_id}{$node->id}++;
                        $deph{$lineage[-1]->internal_id} = $lineage[-1]->depth;
                        $blenh{$lineage[-1]->internal_id} = $lineage[-1]-
>branch_length;
                        }
                }

# identify number of loci in each cluster and representative sequence for each locus
        %tmpfateh = ();
        foreach $tclade (sort(keys(%cladeh)))
                {
                $cladetest++; $cci = 0;
                %th = %{$cladeh{$tclade}};
                @tha = sort(keys(%th));
                $ntha = @tha;
# store a representative of clade at root
                if ($deph{$tclade}==0)
                        {
```

```
                        $outh{$tha[0]}++;
                        $tmpfateh{$tclade} = $tha[0];
                        }
# store a representative of each clade judged to be sufficiently different to
# represent a seperate locus
                elsif ($deph{$tclade}>=$mindepth andand $blenh{$tclade}>=$mindepth)
                        {
                        $outh{$tha[0]}++;
                        $tmpfateh{$tclade} = $tha[0];
                        }
                }
# end tree building code
        @tmpkeys = keys(%tmpfateh);
        $nkeys = @tmpkeys;
        if ($nkeys > 1)
                {
                $splitclust++;
                $newclust += $nkeys;
                print STDERR " split into $nkeys loci.\n";
                }
        else
                {
                $remainsingle++;
                print STDERR " left as a single locus.\n";
                }
        }
        }


print "Finished tree analysis.\n";
system("date");
print $mono_loc, " singleton clusters\n";
print $poly_one, " clusters made of too few sequences to build useful trees\n";
print $c4tree, " cluster with sufficient diversity to build trees\n";
print $remainsingle, " clusters lacked sub-clusters and remained together\n";
print $splitclust, " clusters were split into two or more sub-clusters\n";
print "Those clusters were split into $newclust sub-clusters.\n";

@outlocs = sort(keys(%outh));
$nout = @outlocs;
print $nout, " loci identified altogether.\n";

# write out valid sites to a de novo reference file
open(IN, "filtered_va.fasta");
open(OUT, ">$outfile");
$status = 0; $found = 0;
while(<IN>)
        {
        chomp;
        if ($_ =~ /^>/)
                {
                $sid = $_;
                $sid =~ s/>//;
                $sid =~ s/ .*//g;
                $altsid = $sid;
                $altsid =~ s/\:/-/g;
                if (exists($outh{$sid}) || exists($outh{$altsid}))
                        {
                        $status = 1;
                        $found++;
                        print OUT ">denovoLocus", $found, "\n";
                        next;
                        }
                else
                        {
                        next;
```

```
                        }
                }
        else
                {
                if ($status == 1)
                        {
                        print OUT $_, "\n";
                        }
                $status = 0;
                }
        }
close(IN);
close(OUT);
print "Wrote $found sequences to output file $outfile.\n";
print "This is now ready to use as a reference for mapping and genotyping.\n";
system("date");
```

SAMFilter.pl

```
#!/usr/bin/env perl
# written by E Meyer, eli.meyer@science.oregonstate.edu
# distributed without any guarantees or restrictions

# -- check arguments and print usage statement
$scriptname=$0; $scriptname =~ s/.+\///g;
$usage = <<USAGE;
Filters the alignments produced by mapping short reads against a reference,
excluding ambiguous, short, and weak matches.
NOTE: make sure that when a read matches multiple reference sequences (ambigous)
your mapper reports at least two alignments in the output. This is NOT the default
behavior for some mappers, but is required to exclude ambiguous matches before
genotyping.

Usage:  $scriptname -i input -m matches -o output <options>
Required arguments:
       -i input      Output from any short read mapper, in SAM format.
       -m matches    Minimum number of matching bases required to consider an
alignment valid.
       -o output     A name for the filtered output (SAM format).
Options:
       -c option     1: Report the number of reads matching each reference sequence
                     in a separate output files "counts.tab". 0: Don't produce this
file (default).
       -l length     Minimum length of aligned region (match, mismatch, + gaps)
required to consider
                     an alignment valid. Only relevant if your mapper uses local
alignment. For global
                     alignments, this is set equal to -m.
USAGE
if ($#ARGV < 3 || $ARGV[0] eq "-h") {print "\n", "-"x60, "\n", $scriptname, "\n",
$usage, "-"x60, "\n\n"; exit;}

# -- module and executable dependencies
$mod1="Getopt::Std";
unless(eval("require $mod1")) {print "$mod1 not found. Exiting\n"; exit;}
use Getopt::Std;

# get variables from input
getopts('i:m:o:c:l:h');      # in this example a is required, b is optional, h is
help
if (!$opt_i || !$opt_m || !$opt_o || $opt_h) {print "\n", "-"x60, "\n", $scriptname,
"\n", $usage, "-"x60, "\n\n"; exit;}
if ($opt_c) {$cprint = 1;} else {$cprint = 0;}
if ($opt_l) {$athd = $opt_l;} else {$a_thd = $opt_m;}
my $infile = $opt_i;
my $mthd = $opt_m;
my $outfile = $opt_o;
$ambig = $tooshort = 0;

# read in sam file output and build a hash, counting raw mappings
open(IN, $infile);
my %maph;
while(<IN>)
       {
       if ($_ =~ /^@/) {next;}
       chomp;
       $rowi = $_;
       $rowi =~ s/^>//;
       @cols = split("\t", $rowi);
       $ncols = @cols;

       if ($cols[2] eq "*") {next;}
```

```
        $rawmap++;

# -- extract alignment length and apply alignment threshold
         $numstr = $cols[5];
         @chars = split("M", $numstr);
        $aligni = $mismatchi = 0;
        foreach $c (@chars)
                {
                $c =~ s/.+\D//g;
                if ($c > 0) {$aligni+=$c;}
                }
        if ($aligni < $athd) {$tooshort++; next;}


# -- extract mismatches and apply mismatch threshold
        foreach $flag (@cols[11..$ncols])
                {
                if ($flag =~ /NM\:i/)
                        {
                        $flag =~ s/NM\:i\://;
                        $mismatchi = $flag;
                        }
                }
        $matchi = $aligni - $mismatchi;
        if ($matchi < $mthd) {$tooweak++; next;}

# -- add extracted data to hash
        $maph{$cols[0]}{$cols[2]}{"count"}++;
        $maph{$cols[0]}{$cols[2]}{"string"} = $_;
        $maph{$cols[0]}{$cols[2]}{"align"} = $aligni;
        $maph{$cols[0]}{$cols[2]}{"match"} = $matchi;
        }

# count number of reads with one or more matches passing thresholds
@mra = keys(%maph);
$nmr = @mra;

# select final set of unique mappings
open(OUT, ">$outfile");
open(IN, $infile);
while(<IN>)
        {
        chomp;
        if ($_ =~ /^@/) {print OUT $_, "\n";}
        else {last;}
        }

my %refch;
foreach $r (@mra)
        {
        %rh = %{$maph{$r}};
        @ma = keys(%rh);
        $nma = @ma;
        if ($nma>1)
                {
                @moa = sort{$rh{$b}{"match"}<=>$rh{$a}{"match"}}(keys(%rh));
                if ($rh{$moa[0]}{"match"}==$rh{$moa[1]}{"match"})
                        {
                        $ambig++;
                        next;
                        }
                }
        if ($nma == 1) {@moa = @ma;}
        $unimap++;
        $refch{$moa[0]}++;
        print OUT $rh{$moa[0]}{"string"}, "\n";
```

```
        }

if ($cprint eq 1)
{
open(CTS, ">counts.tab");
@sref = sort{$refch{$b}<=>$refch{$a}}(keys(%refch));
foreach $s (@sref)
        {
        print CTS $s, "\t", $refch{$s}, "\n";
        }
}

print $rawmap, " raw mappings altogether.\n";
print $nmr, " reads had one or more matches\n";
print $tooshort, " excluded for short alignments\n";
print $tooweak, " excluded for weak matches\n";
print $ambig, " excluded for ambiguous matches\n";
print $unimap, " unique mappings remained\n";
```

SAMBaseCounts.pl

```
#!/usr/bin/perl
# written by E Meyer, eli.meyer@science.oregonstate.edu
# distributed without any guarantees or restrictions

# -- check arguments and print usage statement
$scriptname=$0; $scriptname =~ s/.+\///g;
$usage = <<USAGE;
Counts nucleotide frequencies at each locus in a 2bRAD sequence data set.
Usage: $scriptname -i input -r reference -o <OPTIONS>
Required arguments:
       -i input             input alignments, SAM format
       -r reference         reference used to generate the input alignments, FASTA
format
       -o output            a name for the output file (tab delimited text)
Options:
       -c coverage          loci with lower coverage are discarded (default: 3)
USAGE
if ($#ARGV < 2 || $ARGV[0] eq "-h") {print "\n", "-"x60, "\n", $scriptname, "\n",
$usage, "-"x60, "\n\n"; exit;}

# -- module and executable dependencies
$mod1="Getopt::Std";
unless(eval("require $mod1")) {print "$mod1 not found. Exiting\n"; exit;}
use Getopt::Std;

# get variables from input
getopts('i:r:o:c:h');# in this example a is required, b is optional, h is help
if (!$opt_i || !$opt_r || !$opt_o || $opt_h) {print "\n", "-"x60, "\n", $scriptname,
"\n", $usage, "-"x60, "\n\n"; exit;}
if ($opt_c) {$thd = $opt_c;} else {$thd = 3;}
$infile = $opt_i;
$reffile = $opt_r;
$outfile = $opt_o;
# load reference into memory
system("date");
print "beginning analysis of nucleotide frequencies in $infile...\n";
print "Reading reference...\n";
open(REF,$reffile);
while(<REF>)
       {
       chomp;
       if ($_ =~ /\>/)
              {
              $_ =~ s/\>//;
              $refi = $_;
              }
       elsif ($_ =~ /[ACGT]/)
              {
              @bits = split("", $_);
              $bcount = 0;
              foreach $b (@bits)
                     {
                     $bcount++;
                     $rrh{$refi}{$bcount} = $b;
                     }
              }
       }
close(REF);
print "Finished loading reference. into memory.\n";

# define positions to be ignored (recognition sites, which are invariant)
foreach $p (13, 14, 15, 22, 23, 24) {$exch{$p}++;}             # for AlfI
```

```
# read in sam file output line by line
print "Reading alignments and counting nucleotides at each position...\n";
open(IN, $infile);
my %nfh;
while(<IN>)
	{
	if ($_ =~ /^@/) {next;}
	chomp;
	$rowi = $_;
	$rowi =~ s/^>//;
	@cols = split("\t", $rowi);
	$ncols = @cols;

	if ($cols[2] eq "*") {next;}
	$rawmap++;

# -- build a key for each read matching each position to a reference position
	$cig = $cols[5];
	@ciga = split(/(?<=\D)/, $cig);
	@codea = (); $refloc = $cols[3];
	foreach $c (@ciga)
		{
		if ($c =~ /M/)
			{
			$c =~ s/M//;
			for ($d=1; $d<=$c; $d++) {push @codea, $refloc; $refloc++;}
			}
		elsif ($c =~ /D/)
			{
			$c =~ s/D//;
			for ($d=1; $d<=$c; $d++) {$refloc++;}
			}
		elsif ($c =~ /I/)
			{
			$c =~ s/I//;
			for ($d=1; $d<=$c; $d++) {push @codea, 0;}
			}
		}

# -- count each base toward nucleotide frequencies at the corresponding reference
position
	@reada = split("",$cols[9]);

	if (@reada ne @codea) {print "Something is wrong.\n @reada \n @codea\n";
exit;}
	$readpos = -1;
	foreach $b (@reada)
		{
		$readpos++;
		if ($codea[$readpos] eq 0) {next;}
		if ($codea[$readpos] eq "") {next;}
		$nfh{$cols[2]}{$codea[$readpos]}{$b}++;
		}
	}
print "Finished reading and counting alignments.\n";

# -- write output, applying coverage threshold per locus
print "Writing output...\n";
open(OUT, ">$outfile");
print OUT "Tag\tLocus\tRef\tA\tC\tG\tT\n";
@nbins = qw{A C G T};
$rownum = 0;
foreach $ref (sort(keys(%nfh)))
	{
	%refh = %{$nfh{$ref}};
	foreach $loc (sort{$a<=>$b}(keys(%refh)))
```

```
                {
                %loch = %{$refh{$loc}};
                $rownum++;
                $covi = 0;
                foreach $n (@nbins)
                        {
                        if (exists($loch{$n})) {$covi += $loch{$n};}
                        }
                if(exists($exch{$loc})) {next;}
                if ($covi < $thd) {next;}
                print OUT $ref, "\t", $loc, "\t", $rrh{$ref}{$loc}, "\t";
                foreach $n (@nbins)
                        {
                        if (exists($loch{$n})) {print OUT $loch{$n}, "\t";}
                        else   {print OUT "0\t";}
                        }
                print OUT "\n";
                }
        }
print "Finished with analysis of nucleotide frequencies.\n";
system("date");
print "\n";
```

CombineBaseCounts.pl

```
#!/usr/bin/perl
# written by E Meyer, eli.meyer@science.oregonstate.edu
# distributed without any guarantees or restrictions

# -- check arguments and print usage statement
$scriptname=$0; $scriptname =~ s/.+\///g;
$usage = <<USAGE;
Counts the number of times each allele was observed, for each locus, in a collection
of
2bRAD data describing nucleotide frequencies for each locus and sample (the output
from
SAMBaseCounts.pl).

Output format: columns 1=tag, 2=position, 3=reference allele,
5=allele counts for sample 1 (A/C/G/T), 6=for sample 2, etc..
Missing data are shown as "NA" for all alleles.

Usage: $scriptname file_1 file_2 ... file_n > output_file
Where:
       files 1-n:    nucleotide frequencies (output from SAMBasecaller.pl) for each
sample
       output_file:  a name for the output; tab-delimited text
USAGE
if ($#ARGV < 1 || $ARGV[0] eq "-h") {print "\n", "-"x60, "\n", $scriptname, "\n",
$usage, "-"x60, "\n\n"; exit;}

my %bigh;
my $maxthird = 2;

print "Tag\tLocus\tRef\t";
$countrow = 0;
foreach $argi (0..$#ARGV)
       {
       $countrow++;
       if ($countround==1) {next;}
       $tmpname = $ARGV[$argi];
       #$tmpname =~ s/(.+).tab/;
       ($tmpname, $foo) = split (/\.tab/, $tmpname);
       #$tmpname =~ s/.+\///g;
       print $tmpname, "\t";
       print STDERR "reading ", $ARGV[$argi], "\n";
       open(TAB, $ARGV[$argi]);
       $rownum = 0;
       while(<TAB>)
              {
              chomp;
              $rownum++; if ($rownum==1) {next;}
              @cols = split("\t", $_);
              $bigh{$cols[0]}{$cols[1]}{"ref"} = $cols[2];
              $bigh{$cols[0]}{$cols[1]}{$argi}{"A"} = $cols[3];
              $bigh{$cols[0]}{$cols[1]}{$argi}{"C"} = $cols[4];
              $bigh{$cols[0]}{$cols[1]}{$argi}{"G"} = $cols[5];
              $bigh{$cols[0]}{$cols[1]}{$argi}{"T"} = $cols[6];
              }
       }
print "\n";

foreach $r (sort(keys(%bigh)))
       {
       %ch = %{$bigh{$r}};
       foreach $l (sort{$a<=>$b}(keys(%ch)))
              {
              print $r, "\t", $l, "\t", $bigh{$r}{$l}{"ref"}, "\t";
```

```
                foreach $argi (0..$#ARGV)
                        {
                        if (exists($bigh{$r}{$l}{$argi}))
                                {
                                foreach $base (qw{A C G T})
                                        {
                                        print $bigh{$r}{$l}{$argi}{$base};
                                        unless ($base eq "T") {print "/";}
                                        }
                                }
                        else {print "0/0/0/0";}
                        print "\t";
                        }
                print "\n";
                }
        }
exit;
```

*Lines 33-34 modified to include sample names of any format:

```
#$tmpname =~ s/(.+).tab/;
($tmpname, $foo) = split (/\.tab/, $tmpname);
#$tmpname =~ s/.+\///g;
```

CallGenotypes.pl

```
#!/usr/bin/perl
# written by E Meyer, eli.meyer@oregonstate.edu
# distributed without any guarantees or restrictions

# -- check arguments and print usage statement
$scriptname=$0; $scriptname =~ s/.+\///g;
$usage = <<USAGE;

Determines SNP genotypes from nucleotide frequencies. Input file contains nucleotide
frequencies
from multiple samples. Output file lists the genotypes called from those
frequencies.
Usage: $scriptname -i input -o output <OPTIONS>
Required arguments:
       -i input      Input file, tab delimited text file of nucleotide frequencies
(output from NtFrequences.pl)
                     column = tag, column 2 = locus, column 3 = reference genotype
                     subsequent columns = nucleotide frequences in each sample, as
A/C/G/T (e.g. 0/0/10/12)
       -o output     A name for the output file (tab delimited text)
Options:
       -c coverage   Minimum coverage required to determine genotypes. Lower
coverage loci wil be discarded.
                     Default: 10
       -e ends       y: exclude terminal positions in alignments where errors may
arise during ligation. (default)
                     n: do not exclude terminal positions.
       -m method     "nf": nucleotide frequencies (classic method; the default).
                       This method determines genotypes directly from nucleotide
frequencies, using thresholds
                       defined by the user. If minor allele frequency (MAF) <= x at
a locus, the genotype is called
                       homozygous for the major allele at that locus. If MAF >= n,
genotype is called heterozygous.
                       Genotypes are not called at intermediate MAF (if x > MAF > n)
where errors are likely.
                     "pgf": NF informed by population genotype frequencies. (an
update on the classic method)
                       This method first identifies valid alleles at each locus
based on their frequency in the
                       population (the two most common alleles with frequencies >=y
times in >= q individuals),
                       then applies relaxed nucleotide frequency thresholds for
those alleles (using y instead
                       of n for valid alleles).
                     "bgc" = Bayesian Genotype Caller
                       This method calls the BGC software, which implements a
maximum-likelihood (ML) method for
                       calling genotypes that incorporates prior population-level
information on genotype
                       frequencies and error rates from a genotype-frequency
estimator. For more details see
                       (Maruki and Lynch, [doi: 10.1534/g3.117.039008], and cite
that paper if using this option.
       Options for method "nf" or "pgf":
       -x max_MAF    Maximum frequency of the minor allele you're willing to ignore
and call the position
                     homozygous for the major allele (0-1). Default: 0.01
       -n min_MAF    Minimum frequency of the minor allele you're willing to accept
as evidence of
                     heterozygosity, and call the locus heterozygous (0-1). Default:
0.25
```

```
       -r min_reads  Because low frequencies translate into 1 or fewer reads at low
coverage, the script
                     also imposes a minimum read number for detection of
heterozygotes. (default: 2)
       Options for method "pgf":
       -y frequency  Minimum frequency a second allele must be detected to be
considered valid.
                     (default: 0.05)
       -q samples    Each allele must present in at least q samples to be considered
valid.
                     (default: 2)
       Options for method "bgc":
       -p p-value    Critical p-value for the chi-square polymorphism test (BGC)
                     (default: 0.05)
       -v maxcov     Coverage at which the pipeline switches from BGC (for low
coverage data) to
                     HGC (for high coverage data). Default=80 (i.e. HGC above 80).
Examples:
  CallGenotypes.pl -i allele_counts.tab -o genotypes.tab       # basic usage
  CallGenotypes.pl -i allele_counts.tab -o genotypes.tab -c 20        # increase
coverage threshold
  CallGenotypes.pl -i allele_counts.tab -o genotypes.tab -m bgc       # use Bayesian
Genotype Caller
  CallGenotypes.pl -i allele_counts.tab -o genotypes.tab -m pgf -y 0.05      # use
population method
USAGE
if ($#ARGV < 1 || $ARGV[0] eq "-h") {print "\n", "-"x60, "\n", $scriptname, "\n",
$usage, "-"x60, "\n\n"; exit;}

# -- module and executable dependencies
$mod1="Getopt::Std";
unless(eval("require $mod1")) {print "$mod1 not found. Exiting\n"; exit;}
use Getopt::Std;
$mod1="Bio::SeqIO";
unless(eval("require $mod1")) {print "$mod1 not found. Exiting\n"; exit;}
use Bio::SeqIO;
$mod1="File::Which";
unless(eval("require $mod1")) {print "$mod1 not found. Exiting\n"; exit;}
use File::Which;

# get variables from input
getopts('i:o:e:c:m:x:n:y:q:p:r:v:h');     # in this example a is required, b is
optional, h is help
if (!$opt_i || !$opt_o || $opt_h) {print "\n", "-"x60, "\n", $scriptname, "\n",
$usage, "-"x60, "\n\n"; exit;}
$ftabfile = $opt_i;  # input filename
$outfile = $opt_o;   # output filename
if ($opt_m)   {$method = $opt_m;} else    {$method = "nf";}
if ($method ne "bgc" andand $method ne "nf" andand $method ne "pgf") {print "\n", "-
"x60, "\n", $scriptname, "\n", $usage, "-"x60, "\n\n"; exit;}
if ($method eq "bgc")
       {
       $dep1 = "GFE";
       unless (defined(which($dep1))) {print $dep1, " not found. Exiting.\n"; exit;}
       $dep1 = "BGC";
       unless (defined(which($dep1))) {print $dep1, " not found. Exiting.\n"; exit;}
       $dep1 = "HGC";
       unless (defined(which($dep1))) {print $dep1, " not found. Exiting.\n"; exit;}
       }
if ($opt_c) {$mincov = $opt_c;} else {$mincov = 10;}           # default min
coverage to be genotyped
if ($opt_x) {$maxhom = $opt_x;} else {$maxhom = 0.01;}         # default maximum MAF
ignored and called homozygous for major allele
if ($opt_n) {$minhet = $opt_n;} else {$minhet = 0.25;}         # default minimum MAF
considered heterozygous
```

```
if ($opt_r) {$minreads = $opt_r;} else {$minreads = 2;}        # default minimum MA
read depth for heterozygous calls
if ($opt_y) {$minall = $opt_y;} else {$minall = 0.05;}         # default minimum
coverage per allele to define valid alleles
if ($opt_q) {$minsamp = $opt_q;} else {$minsamp = 2;}          # default minimum
number of samples to define valid alleles
if ($opt_p) {$bgcp = $opt_p;} else {$bgcp = 0.05;}             # default critical p
value for polymorphism test (BGC)
if ($opt_e) {$eopt = $opt_e;} else {$eopt = "y";}       # exclude terminal positions
from alignments
if ($opt_v) {$maxcov = $opt_v;} else {$maxcov = 80;}           # max coverage for
BGC

system("date");

# Nucleotide Frequency Method - the classic approach
if ($method eq "nf")
{
# make a big hash of all base calls
print "Reading $ftabfile ...\n";
my %bh;
@boa = qw{A C G T};
open(FTAB, $ftabfile);
$rowno = 0;
while(<FTAB>)
        {chomp;
        $rowno++;
        @cols = split("\t", $_);
        if ($rowno == 1)
                {
                @header = @cols;
                next;
                }
        for ($i=3; $i<@cols; $i++)
                {
                @bases = split("/", $cols[$i]);
                $jno = 0;
                foreach $j (@bases)
                        {
                        $bh{$cols[0]}{$cols[1]}{$header[$i]}{$boa[$jno]} = $j;
                        $jno++;
                        }
                }
        }
print "Finished. \n";
print $rowno-1, " loci with nucleotide frequency data. \n";
system("date");
close(FTAB);

@samlabs = @header[3..@cols-1];

# loop through all loci and samples, applying genotyping rules using threshold
defined above
# in this method, genotyping is based directly on nucleotide frequency thresholds
print "Determining genotypes from nucleotide frequencies ...\n";
open(OUT,">$outfile");
print OUT "scaffold", "\t", "position", "\t", join ("\t", @samlabs), "\n";
@sites = sort(keys(%bh));
$countgt = $gtlocno = $sampno = 0;
foreach $s (@sites)
        {
        %siteh = %{$bh{$s}};
        @loca = sort{$a<=>$b}(keys(%siteh));
        foreach $l (@loca)
                {
                $sgt = 0;
```

```
                if ($l == 1 | $l == 36)
                        {
                        if($eopt eq "y")
                                {
                                next;
                                }
                        }
                %loch = %{$siteh{$l}};
                @gta = ();
                foreach $s (@samlabs)
                        {
                        %samh = %{$loch{$s}};
                        %tmph = ();
                        foreach $b (@boa) {$tmph{$b} = $samh{$b};}
                        @saa = sort{$tmph{$b}<=>$tmph{$a}}(keys(%tmph));
                        $majmin = $tmph{$saa[0]}+$tmph{$saa[1]};
                        $minreject = $majmin * $maxhom;
                        $gti = ""; $rati = "";
#                       print $saa[0], ":", $tmph{$saa[0]}, ",", $saa[1], ":",
$tmph{$saa[1]}, "[", $majmin, "]";
# first conditions where we arent interested in or cannot compare nucleotide
frequencies
# exclude loci with >2 alleles observed, and loci with low coverage
                        if (($tmph{$saa[2]}>=$minreject)||($majmin<$mincov))
                                {$gti = 0;}
# then conditions where we can compare nucleotide frequencies
                        else
                                {
                                $rati = $tmph{$saa[1]}/($majmin);
#                               print "(", $rati, ")";
                                if ($rati<$maxhom) {$gti = $saa[0];}
                                elsif ($rati>=$minhet) {@rsa = sort(@saa[0..1]); $gti =
"@rsa";}
                                elsif (($rati >= $maxhom andand $rati <
$minhet)||($tmph{$saa[1]} < $minreads)) {$gti = 0;}
                                }
#                       print "\t";
                        push @gta, $gti;
                        if ($gti =~ /[ACGT]/) {$sgt++;}
                        }
#               print join("\t", @gta), "\n";
                $countgt += $sgt;
                if ($sgt > 0)
                        {
                        $gtlocno++;
                        print OUT $s, "\t", $l, "\t", join("\t", @gta), "\n";
                        }
                }
        }
print "Finished. \n";
$sampno = @samlabs;
print "Called genotypes at $gtlocno loci in $sampno samples ($countgt genotypes
altogether)\n";
system("date");
}

# Nucleotide Frequencies Informed by Population Genotype Frequencies
# in this method, valid alleles are defined at each locus by counting the number of
times each
# allele occurs at high enough coverage in enough samples. Each sample is then
genotyped only for those
# alleles based on observed numbers and user defined nucleotide frequency
thresholds. The increased
# confidence in the valid alleles provided by population analysis makes it possible
to relax MAF
```

```
# threshold for heterozygous calls, determining genotypes at loci with intermediate
MAFs that could
# otherwise not be determined directly from nucleotide frequencies.
if ($method eq "pgf")
{
# make a big hash of all base calls
print "Reading $ftabfile ...\n";
my %bh;
@boa = qw{A C G T};
open(FTAB, $ftabfile);
$rowno = 0;
while(<FTAB>)
	{chomp;
	$rowno++;
	@cols = split("\t", $_);
	if ($rowno == 1)
		{
		@header = @cols;
		next;
		}
	for ($i=3; $i<@cols; $i++)
		{
		@bases = split("/", $cols[$i]);
		$jno = 0;
		foreach $j (@bases)
			{
			$bh{$cols[0]}{$cols[1]}{$header[$i]}{$boa[$jno]} = $j;
			$jno++;
			}
		}
	}
print "Finished. \n";
print $rowno-1, " loci with nucleotide frequency data. \n";
system("date");
close(FTAB);

@samlabs = @header[3..@cols-1];

# loop through all loci and samples, applying genotyping rules using thresholds
defined above.
print "Determining genotypes from population genotype frequencies and sample
nucleotide frequencies ...\n";
open(OUT,">$outfile");
print OUT "scaffold", "\t", "position", "\t", join ("\t", @samlabs), "\n";
@sites = sort(keys(%bh));
foreach $s (@sites)
	{
	%siteh = %{$bh{$s}};
	@loca = sort{$a<=>$b}(keys(%siteh));
# filter out terminal bases
	foreach $l (@loca)
		{
		$sgt = 0;
		if ($l == 1 | $l == 36)
			{
			if($eopt eq "y")
				{
				next;
				}
			}
		%loch = %{$siteh{$l}};
		@gta = (); %vah = ();
# first identify valid alleles
		foreach $s (@samlabs)
			{
			%samh = %{$loch{$s}};
```

```
                        %tmph = (); $covi = 0;
                        foreach $b (@boa) {$tmph{$b} = $samh{$b}; $covi += $samh{$b};}
                        @saa = sort{$tmph{$b}<=>$tmph{$a}}(keys(%tmph));
                        if ($covi > 0)
                                {
                                if ((($tmph{$saa[0]}/$covi) >=
$minall)andand($tmph{$saa[0]} >= $minreads)) {$vah{$saa[0]}++;}
                                if ((($tmph{$saa[1]}/$covi) >=
$minall)andand($tmph{$saa[1]} >= $minreads)) {$vah{$saa[1]}++;}
                                }
                        }
                @tmp = sort{$vah{$b}<=>$vah{$a}}(keys(%vah));
                %sva = (); $nva = 0;
                foreach $t (@tmp)
                        {
                        if ($vah{$t} >= $minsamp) {$sva{$t}++; $nva++;}
                        }
#               print $s, "\t", $l, "\t", $nva, " valid alleles\n";
# filter out positions with more than two alleles
# then genotype samples based on frequency of those alleles in each sample
# taking into account the population genotype frequencies determined above
                foreach $s (@samlabs)
                        {
                        %samh = %{$loch{$s}};
                        %tmph = ();
                        foreach $b (@boa) {$tmph{$b} = $samh{$b};}
                        @saa = sort{$tmph{$b}<=>$tmph{$a}}(keys(%tmph));
                        $majmin = $tmph{$saa[0]}+$tmph{$saa[1]};
                        $minreject = $majmin * $maxhom;
                        $gti = ""; $rati = "";
#                       print $saa[0], ":", $tmph{$saa[0]}, ",", $saa[1], ":",
$tmph{$saa[1]}, "[", $majmin, "]";
# first conditions where we arent interested in or cannot compare nucleotide
frequencies
# exclude loci with >2 alleles observed, and loci with low coverage
                        if (($tmph{$saa[2]}>=$minreject)||($majmin<$mincov))
                                {
                                $gti = 0;
#                               print "r";
                                }
# then conditions where we can compare nucleotide frequencies
                        else
                                {
                                $rati = $tmph{$saa[1]}/($majmin);
#                               print "(", $rati, ")";
                                if ($rati<$maxhom) {$gti = $saa[0];}
                                elsif ($rati>=$minhet) {@rsa = sort(@saa[0..1]); $gti =
"@rsa";}
                                elsif (($rati >= $maxhom andand $rati <
$minhet)||($tmph{$saa[1]} < $minreads))
                                        {
                                        if(exists($sva{$saa[1]}))   # if minor allele in
this sample is one
                                                {                   # of top 2 in pgf,
handle this way
                                                if (($rati >= $minall) andand
($tmph{$saa[1]} >= $minreads))
                                                        {@rsa = sort(@saa[0..1]); $gti =
"@rsa";}
                                                else
                                                        {$gti = 0;}
                                                }
                                        else                        # if minor allele is
not one of the top
                                                {                   # 2 in pgf handle it
this way
```

```
                                        $gti = 0;
                                        }
                                }
                        }
#                       print "\t";
                        push @gta, $gti;
                        if ($gti ne 0) {$sgt++;}
                        }
#               print join("\t", @gta), "\t", $sgt, "\n";
                $countgt += $sgt;
                if ($sgt > 0)
                        {
                        $gtlocno++;
                        print OUT $s, "\t", $l, "\t", join("\t", @gta), "\n";
                        }
                }
        }
print "Finished. \n";
$sampno = @samlabs;
print "Called genotypes at $gtlocno loci in $sampno samples ($countgt genotypes
altogether)\n";
system("date");
}

# Bayesian Genotyper Caller Method
if ($method eq "bgc")
{
# estimate genotype frequencies in population
$gfe_out = $ftabfile.".gfe.out";
print "Running GFE ...\n";
system "GFE -cv $bgcp -mode c -as 1 -min_cov $mincov -max_cov $maxcov -in $ftabfile
-out $gfe_out > gfe_log.txt";
print "Finished.\n";
system("date");

# call genotypes at low coverage sites using ML approach informed by prior
population estimates
$bgc_out = $ftabfile.".bgc.out";
print "Running BGC ...\n";
system "BGC -as 1 -min_cov $mincov -max_cov $maxcov -in $gfe_out -out $bgc_out >
bgc_log.txt";
print "Finished.\n";
system("date");

# call genotypes at high coverage sites using ML approach informed by prior
population estimates
$hgc_out = $ftabfile.".hgc.out";
print "Running HGC ...\n";
HGC:
system "HGC -in $ftabfile -out $hgc_out -min_cov $maxcov > hgc_log.txt";
$numrows=`wc -l $hgc_out`;
$numrows =~ s/\s.+//;
chomp($numrows);
if ($numrows eq 1) {goto HGC;}            # because HGC sometimes exits without
throwing an error. This catches it.
print "Finished.\n";
system("date");

# integrate output from BGC and HGC
$rownum=0;
open(IN, $hgc_out);
while(<IN>)
        {
        chomp;
        $rownum++;
        @cols = split("\t", $_);
```

```
        if ($rownum==1) {@header=@cols; next;}
        $len=@cols; $end=$len-8;
        for ($a=11; $a<$end; $a++)
                {
                unless ($cols[$a] eq "NA")
                        {
                        $fh{$cols[0]}{$cols[1]}{$header[$a]} = $cols[$a];
                        }
                }
        }
close(IN);

open(IN, $bgc_out);
open(OUT, ">integrated_bgc.tab");
$rownum=0;
while(<IN>)
        {
        chomp;
        $rownum++;
        @cols = split("\t", $_);
        if ($rownum==1) {print OUT $_, "\n"; @header=@cols; next;}
        $end=@cols-1;
        for ($a=5; $a<$end; $a++)
                {
                if(exists($fh{$cols[0]}{$cols[1]}{$header[$a]}))
                        {
                        $cols[$a] = $fh{$cols[0]}{$cols[1]}{$header[$a]};
                        }
                }
        print OUT join("\t", @cols), "\n";
        }
close(IN);

# convert output to same format as the other methods
print "Converting output file...\n";
open(OUT, ">$outfile");
foreach $a (qw{0 1 2 3 4})
        {$exc_col{$a}++;}
foreach $a (qw{2 3 4})
        {$exa_col{$a}++;}
open(IN, "integrated_bgc.tab");
$rownum = 0;
while(<IN>)
        {
        $rownum++;
        chomp;
        @cols = split("\t", $_);
        $cno = 0;
        if ($rownum==1)
                {
                foreach $c (@cols)
                        {
                        unless (exists($exa_col{$cno}))
                                {
                                print OUT $c, "\t";
                                }
                        $cno++;
                        }
                print OUT "\n";
                next;
                }

        $nonmd = 0; $cno = 0;
        foreach $c (@cols)
                {
                unless (exists($exc_col{$cno}))
```

```
                        {
                        if ($c ne "NA")       {$nonmd++;}
                        }
                $cno++;
                }
        if ($nonmd < 1) {next;}

        if ($eopt eq "y")
                {
                if ($cols[1] == 1 || $cols[1] == 36)
                        {
                        next;
                        }
                }

        $cno = 0; $gtino = 0;
        foreach $c (@cols)
                {
                if ($cno < 2)
                        {
                        print OUT $c, "\t";
                        }
                unless (exists($exc_col{$cno}))
                        {
                        @alls = split("", $c);
                        if ($c eq "NA")
                                {
                                print OUT "0\t";
                                }
                        elsif ($alls[0] eq $alls[1])
                                {
                                $countgt++; $gtino++;
                                print OUT $alls[0], "\t";
                                }
                        else
                                {
                                $countgt++; $gtino++;
                                print OUT $alls[0], " " , $alls[1], "\t";
                                }
                        }
                $cno++;
                }
        if ($gtino>0) {$gtloc++;}
        print OUT "\n";
        }
print "Finished.\n";
$sampno = @cols - 5;
print "Genotyped $gtloc loci in $sampno samples ($countgt genotypes altogether)\n";
system("date");
}
```

PolyFilter.pl

```
#!/usr/bin/env perl
# written by E Meyer, eli.meyer@science.oregonstate.edu
# distributed without any guarantees or restrictions

# -- check arguments and print usage statement
$scriptname=$0; $scriptname =~ s/.+\///g;
$usage = <<USAGE;
Excludes loci containing too few classes of genotypes or numbers of alleles (keeps
polymorphic loci).
Usage: $scriptname -i input <OPTIONS>
Required arguments:
       -i input      name of the input file, a matrix of genotypes or allele counts
                     (see -m for format)
Options:
       -n genotypes  minimum number of unique genotypes (for -m g) or alleles (for -
m a)
                     required to consider a locus polymorphic (default=2)
       -m mode       g=genotypes (default).
                       Input file contains genotypes from individuals.
                       Input format: rows=loci and columns=samples.
                          row 1 = column label, column 1 = tag, column 2 = position
                          subsequent columns contain genotypes for each sample
                     a=allele counts.
                       Input file contains allele counts from pooled samples.
                       Input format: rows=loci and columns=samples.
                          row 1 = column label, column 1 = tag, column 2 = position
                          column 3 = major allele, column 4 = minor allele
                          subsequent pairs of columns contain allele counts
                       (major then minor) for each sample
       -v min_cov    (for -m a) minimum coverage required to consider an allele
present
                     (default=2)
       -s min_samp   minimum number of samples in which an allele must be present to
be counted.
                     (default=1)
       -p option     y=print filtered loci and summary; n=only print summary
                     (default=n)
       -o output     a name for the output file (loci passing this filter) (required
if -p y)

USAGE

# -- module and executable dependencies
$mod1="Getopt::Std";
unless(eval("require $mod1")) {print "$mod1 not found. Exiting\n"; exit;}
use Getopt::Std;

# get variables from input
getopts('i:o:n:p:s:v:m:h'); # in this example a is required, b is optional, h is
help
if (!$opt_i || $opt_h) {print "\n", "-"x60, "\n", $scriptname, "\n", $usage, "-"x60,
"\n\n"; exit;}
if ($opt_n) {$minall = $opt_n;} else {$minall = 2;}     # min genotypes or alleles
to consider polymorphic
if ($opt_p) {$pvar = $opt_p;} else {$pvar = "n";}# whether to print loci passing
filter
if ($pvar eq "y" andand !$opt_o) {print "\n", "-"x60, "\n", $scriptname, "\n",
$usage, "-"x60, "\n\n"; exit;}
if ($opt_o) {$ovar = $opt_o;} else {$ovar = "null";}    # name for output file
if ($opt_m) {$mode = $opt_m;} else {$mode = "g";}# mode
if ($mode ne "g" andand $mode ne "a") {print "\n", "-"x60, "\n", $scriptname, "\n",
$usage, "-"x60, "\n\n"; exit;}
if ($opt_v) {$mincov = $opt_v;} else {$mincov = 2;}     # min cov, for -m a
```

```
if ($opt_s) {$minfreq = $opt_s;} else {$minfreq = 1;}    # min samps, for -m a
my $opt = $pvar;

# mode g
if ($mode eq "g")
{
open(IN, $opt_i);
open(OUT, ">$ovar");
while(<IN>)
        {
        chomp;
        $rowcount++;
        if($rowcount==1){if($opt eq "y") {print OUT $_, "\n";} next;}
        @cols = split("\t", $_);
        $ncols = @cols;
        %ahi = ();
        for ($a=2; $a<$ncols; $a++)
                {
                if ($cols[$a] eq "0") {next;}
                $ahi{$cols[$a]}++;

                }
        @alla = sort(keys(%ahi));
        $passno = 0;
        foreach $z (@alla)
                {
                if ($ahi{$z} >= $minfreq) {$passno++;}
                }
        $nalla = @alla;
        if ($passno < $minall) {$nopoly++; next;}
        if ($opt eq "y") {print OUT $_, "\n";}
        }
close(IN); close(OUT);
print STDERR $rowcount-1, " loci in input file\n";
print STDERR $nopoly, " loci with fewer than ", $minall, " genotypes\n";
print STDERR $rowcount-1-$nopoly, " polymorphic loci (at least ", $minall, "
genotypes in $minfreq or more samples)\n";
}

# mode a
elsif ($mode eq "a")
{
open(IN, $opt_i);
open(OUT, ">$ovar");
while(<IN>)
        {
        chomp;
        $rowcount++;
        if($rowcount==1){if($opt eq "y") {print OUT $_, "\n";} next;}
        @cols = split("\t", $_);
        $ncols = @cols;
        $gt_no = $alt_no = $alt_freq = 0;
        for ($a=4; $a<$ncols; $a+=2)
                {
                if ($cols[$a] eq "NA") {next;}
                $gt_no++;
                if ($cols[$a+1] >= $mincov)
                        {
                        $alt_no++;
                        }
                }
        $alt_freq = $alt_no / $gt_no;
        if ($alt_freq eq 0) {$mono++; next;}
        if ($alt_freq < $minfreq) {$toolow++; next;}
        if ($opt eq "y") {print OUT $_, "\n";}
        }
```

```
close(IN); close(OUT);
print STDERR $rowcount-1, " loci in input file\n";
print STDERR $toolow, " loci with minor alleles present in fewer than $minfreq of
sample\n";
print STDERR $mono, " monomorphic loci\n";
print STDERR $rowcount-1-$toolow-$mono, " polymorphic loci\n";
}

if (-e "null")
        {
        system("rm null");
        }
```

MDFilter.pl

```
#!/usr/bin/env perl
# written by E Meyer, eli.meyer@science.oregonstate.edu
# distributed without any guarantees or restrictions

# -- check arguments and print usage statement
$scriptname=$0; $scriptname =~ s/.+\///g;
$usage = <<USAGE;

Excludes loci containing too many missing data (genotyped in too few samples)
Usage: $scriptname -i input -n min_data <OPTIONS>
Required arguments:
       -i input      name of the input file, a matrix of genotypes or allele counts
                     (see -m for format)
       -n min_data   loci that were genotyped in fewer samples than this will be
excluded
Options:
       -m mode       g=genotypes (default).
                       Input file contains genotypes from individuals.
                       Input format: rows=loci and columns=samples.
                          row 1 = column label, column 1 = tag, column 2 = position
                          subsequent columns contain genotypes for each sample
                     a=allele counts.
                       Input file contains allele counts from pooled samples.
                       Input format: rows=loci and columns=samples.
                          row 1 = column label, column 1 = tag, column 2 = position
                          column 3 = major allele, column 4 = minor allele
                          subsequent pairs of columns contain allele counts
                       (major then minor) for each sample
       -p option     y=print filtered loci and summary; n=only print summary
                     (default=n)
       -o output     a name for the output file (loci passing this filter) (required
if -p y)

USAGE

# -- module and executable dependencies
$mod1="Getopt::Std";
unless(eval("require $mod1")) {print "$mod1 not found. Exiting\n"; exit;}
use Getopt::Std;

# get variables from input
getopts('i:o:n:p:m:h');     # in this example a is required, b is optional, h is
help
if (!$opt_i || !$opt_n || $opt_h) {print "\n", "-"x60, "\n", $scriptname, "\n",
$usage, "-"x60, "\n\n"; exit;}
if ($opt_p) {$pvar = $opt_p;} else {$pvar = "n";}# whether to print loci passing
filter
if ($pvar eq "y" andand !$opt_o) {print "\n", "-"x60, "\n", $scriptname, "\n",
$usage, "-"x60, "\n\n"; exit;}
if ($opt_o) {$ovar = $opt_o;} else {$ovar = "null";}    # name for output file
if ($opt_m) {$mode = $opt_m;} else {$mode = "g";}# name for output file
if ($mode ne "g" andand $mode ne "a") {print "\n", "-"x60, "\n", $scriptname, "\n",
$usage, "-"x60, "\n\n"; exit;}

my $mindat = $opt_n;
my $opt = $pvar;

# mode g
if ($mode eq "g")
{
open(IN, $opt_i);
open(OUT, ">$ovar");
$toofew = 0;
```

```
while(<IN>)
	{
	chomp;
	$rowcount++;
	if($rowcount==1){if($opt eq "y") {print OUT $_, "\n";} next;}
	@cols = split("\t", $_);
	$ncols = @cols;
	if ($rowcount==2)
		{
		$maxmiss = $ncols - 2 - $mindat;
		}
	$missi = 0;
	for ($a=2; $a<$ncols; $a++)
		{
		if ($cols[$a]eq "0") {$missi++;}
		}
	if ($missi>$maxmiss) {$toofew++; next;}
	if ($opt eq "y") {print OUT $_, "\n";}
	}
close(IN); close(OUT);

print STDERR $rowcount-1, " loci\n";
print STDERR $toofew, " loci with more than ", $maxmiss, " missing data (i.e.
genotyped in fewer than $mindat samples)\n";
print STDERR $rowcount-1-$toofew, " loci with no more than more than ", $maxmiss, "
missing data\n";
}

# mode a
elsif ($mode eq "a")
{
open(IN, $opt_i);
open(OUT, ">$ovar");
$toofew = 0;
while(<IN>)
	{
	chomp;
	$rowcount++;
	if($rowcount==1){if($opt eq "y") {print OUT $_, "\n";} next;}
	@cols = split("\t", $_);
	$ncols = @cols;
	if ($rowcount==2)
		{
		$maxmiss = ($ncols - 4)/2 - $mindat;
		}
	$missi = 0;
	for ($a=4; $a<$ncols; $a+=2)
		{
		if (($cols[$a] eq 0 andand $cols[$a+1] eq 0)||($cols[$a] eq "NA"
andand $cols[$a+1] eq "NA")) {$missi++;}
		}
	if ($missi>$maxmiss) {$toofew++; next;}
	if ($opt eq "y") {print OUT $_, "\n";}
	}
close(IN); close(OUT);
print STDERR $rowcount-1, " loci\n";
print STDERR $toofew, " loci with more than ", $maxmiss, " missing data (i.e.
genotyped in fewer than $mindat samples)\n";
print STDERR $rowcount-1-$toofew, " loci with no more than more than ", $maxmiss, "
missing data\n";
}
if (-e "null")
	{
	system("rm null");
	}
```

RepTagFilter.pl

```
#!/usr/bin/env perl
# written by E Meyer, eli.meyer@science.oregonstate.edu
# distributed without any guarantees or restrictions

# -- check arguments and print usage statement
$scriptname=$0; $scriptname =~ s/.+\///g;
$usage = <<USAGE;

Excludes tags containing too many SNPs, suggesting repetive regions of the genome
Usage: $scriptname -i input -n max_snps <OPTIONS>
Required arguments:
       -i input      name of the SNP input file, a matrix of genotypes or allele
counts
                     (see -m for format)
                     note: the script assumes the input only includes polymorphic
loci
       -n max_snps   all SNPs from tags containing more than this number of SNPs
will be excluded
Options:
       -m mode       g=genotypes (default).
                       Input file contains genotypes from individuals.
                       Input format: rows=loci and columns=samples.
                          row 1 = column label, column 1 = tag, column 2 = position
                          subsequent columns contain genotypes for each sample
                     a=allele counts.
                       Input file contains allele counts from pooled samples.
                       Input format: rows=loci and columns=samples.
                          row 1 = column label, column 1 = tag, column 2 = position
                          column 3 = major allele, column 4 = minor allele
                          subsequent pairs of columns contain allele counts
                       (major then minor) for each sample
       -p option     y=print filtered loci and summary; n=only print summary
                     (default=n)
       -o output     a name for the output file (loci passing this filter) (required
if -p y)

USAGE

# -- module and executable dependencies
$mod1="Getopt::Std";
unless(eval("require $mod1")) {print "$mod1 not found. Exiting\n"; exit;}
use Getopt::Std;

# get variables from input
getopts('i:o:n:p:m:h');     # in this example a is required, b is optional, h is
help
if (!$opt_i || !$opt_n || $opt_h) {print "\n", "-"x60, "\n", $scriptname, "\n",
$usage, "-"x60, "\n\n"; exit;}
if ($opt_p) {$pvar = $opt_p;} else {$pvar = "n";}# whether to print loci passing
filter
if ($pvar eq "y" andand !$opt_o) {print "\n", "-"x60, "\n", $scriptname, "\n",
$usage, "-"x60, "\n\n"; exit;}
if ($opt_o) {$ovar = $opt_o;} else {$ovar = "null";}    # name for output file
if ($opt_m) {$mode = $opt_m;} else {$mode = "g";}# name for output file
if ($mode ne "g" andand $mode ne "a") {print "\n", "-"x60, "\n", $scriptname, "\n",
$usage, "-"x60, "\n\n"; exit;}
# note that choice of mode does nothing in this script
# its only included for consistency with other filtering scripts

my $maxsnps = $opt_n;
my $opt = $pvar;

open(IN, $opt_i);
```

```
open(OUT, ">$ovar");

while(<IN>)
	{
	chomp;
	$rowcount++;
	if($rowcount==1){next;}
	@cols = split("\t", $_);
	$ncols = @cols;
	$ch{$cols[0]}++;
	}
close(IN);

$rowcount = 0;
open(IN, $opt_i);
while(<IN>)
	{
	chomp;
	$rowcount++;
	if($rowcount==1){if($opt eq "y") {print OUT $_, "\n";} next;}
	@cols = split("\t", $_);
	$ncols = @cols;
	if ($ch{$cols[0]}>$maxsnps) {$toomany++; next;}
	if ($opt eq "y") {print OUT $_, "\n";}
	}
close(IN); close(OUT);
if (-e "null")
	{
	system("rm null");
	}
print STDERR $rowcount-1, " loci\n";
print STDERR $toomany, " loci with more than ", $maxsnps, " SNPs per tag\n";
print STDERR $rowcount-1-$toomany, " loci with no more than more than ", $maxsnps, "
SNPS per tag\n";
```

OneSNPPerTag.pl

```
#!/usr/bin/env perl
# written by E Meyer, eli.meyer@science.oregonstate.edu
# distributed without any guarantees or restrictions

# -- check arguments and print usage statement
$scriptname=$0; $scriptname =~ s/.+\///g;
$usage = <<USAGE;

Selects a single SNP from each tag in a matrix or genotypes or allele counts.
Chooses the locus with the least missing data.
Usage: $scriptname -i input <OPTIONS>
Required arguments:
       -i input      name of the input file, a matrix of genotypes or allele counts
                     (see -m for format)
Options:
       -m mode       g=genotypes (default).
                       Input file contains genotypes from individuals.
                       Input format: rows=loci and columns=samples.
                          row 1 = column label, column 1 = tag, column 2 = position
                          subsequent columns contain genotypes for each sample
                     a=allele counts.
                       Input file contains allele counts from pooled samples.
                       Input format: rows=loci and columns=samples.
                          row 1 = column label, column 1 = tag, column 2 = position
                          column 3 = major allele, column 4 = minor allele
                          subsequent pairs of columns contain allele counts
                       (major then minor) for each sample
       -p option     y=print filtered loci and summary; n=only print summary
                     (default=n)
       -o output     a name for the output file (loci passing this filter) (required
if -p y)

USAGE

# -- module and executable dependencies
$mod1="Getopt::Std";
unless(eval("require $mod1")) {print "$mod1 not found. Exiting\n"; exit;}
use Getopt::Std;

# get variables from input
getopts('i:o:p:m:h');# in this example a is required, b is optional, h is help
if (!$opt_i || $opt_h) {print "\n", "-"x60, "\n", $scriptname, "\n", $usage, "-"x60,
"\n\n"; exit;}
if ($opt_p) {$pvar = $opt_p;} else {$pvar = "n";}# whether to print loci passing
filter
if ($pvar eq "y" andand !$opt_o) {print "\n", "-"x60, "\n", $scriptname, "\n",
$usage, "-"x60, "\n\n"; exit;}
if ($opt_o) {$ovar = $opt_o;} else {$ovar = "null";}    # name for output file
if ($opt_m) {$mode = $opt_m;} else {$mode = "g";}# name for output file
if ($mode ne "g" andand $mode ne "a") {print "\n", "-"x60, "\n", $scriptname, "\n",
$usage, "-"x60, "\n\n"; exit;}

my $opt = $pvar;

if ($mode eq "g")
{
open(IN, $opt_i);
open(OUT, ">$ovar");
while(<IN>)
       {
       chomp;
       $rowcount++;
       if($rowcount==1){next;}
```

```
        @cols = split("\t", $_);
        $ncols = @cols;
        $mdcount = 0;
        foreach $a (@cols[2..$ncols]) {if ($a eq "0") {$mdcount++;}}
        $ch{$cols[0]}{$cols[1]} = $mdcount;
        }
close(IN);
foreach $tag (sort(keys(%ch)))
        {
        %tagh = %{$ch{$tag}};
        @sorted = sort{$tagh{$a}<=>$tagh{$b}}(keys(%tagh));
        $useh{$tag} = $sorted[0];
        }
open(IN, $opt_i);
$rowcount = 0;
while(<IN>)
        {
        chomp;
        $rowcount++;
        if($rowcount==1){if($opt eq "y") {print OUT $_, "\n";} next;}
        @cols = split("\t", $_);
        if($useh{$cols[0]} ne $cols[1]) {next;}
        if ($opt eq "y") {print OUT $_, "\n";}
        $kept++;
        }
close(IN); close(OUT);
print STDERR $rowcount-1, " loci\n";
print STDERR $kept, " SNPs selected.\n";
}

elsif ($mode eq "a")
{
open(IN, $opt_i);
open(OUT, ">$ovar");
while(<IN>)
        {
        chomp;
        $rowcount++;
        if($rowcount==1){next;}
        @cols = split("\t", $_);
        $ncols = @cols;
        $mdcount = 0;
        for ($a=4; $a<$ncols; $a+=2)
                {
                if (($cols[$a] eq "0" andand $cols[$a+1] eq "0")||($cols[$a] eq "NA"
andand $cols[$a+1] eq "NA"))
                        {$mdcount++;}
                }
        $ch{$cols[0]}{$cols[1]} = $mdcount;
        }
close(IN);

foreach $tag (sort(keys(%ch)))
        {
        %tagh = %{$ch{$tag}};
        @sorted = sort{$tagh{$a}<=>$tagh{$b}}(keys(%tagh));
        $useh{$tag} = $sorted[0];
        }

$rowcount = 0;
open(IN, $opt_i);
while(<IN>)
        {
        chomp;
        $rowcount++;
        if($rowcount==1){if($opt eq "y") {print OUT $_, "\n";} next;}
```

```
        @cols = split("\t", $_);
        if($useh{$cols[0]} ne $cols[1]) {next;}
        if ($opt eq "y") {print OUT $_, "\n";}
        $kept++;
        }
close(IN);
print STDERR $rowcount-1, " loci\n";
print STDERR $kept, " SNPs selected.\n";
}

if (-e "null")
        {
        system("rm null");
        }
```

Appendix E

>Haplotype: CAT1 GB#MN784029 (Sample: RPCAT1)
AAAGATCAACACATGACCAACCTGCATAAACGCATGTACTTGCGTGAACGCAAATGCATTTGTGAAAACAC
AAATGCAATTGCAAAATTATGCAGGTTGGTCAGTCTGACATATAAACTATTTAGCCCCATACTATGTATAT
CCTCATTAAATATACATACTATGCTTAATCCTCATTAATAGACTGTCCCCATATACATATTATGCTTAATC
CACATCATACATATTATGCTTAACCCCCACATTACATATTATGCTTAATCCACATCATACATATTATGCTT
AATCCACATCATACATACGTCTATCTCACTCCCCATACACACATATATGTGTGCTTAATCCCCATTAATCC
ATACTCCCCTACTCCATTACATGCTATGCTTAATCCTCATTAATAGACTGTCCCCATAGTACATACTATGC
TTAATCCTCATTAACATACTATGCTTAATCCCCATTAATAGACATTCCCCTATTTCATTACATCTTACGTA
TCATTAATCCTCATTACTAGAAAATCCTCTACTCCATTACAATCATTCACTCACCCTCATTAATTTACTTC
TAATACTGCTTATGCGGGCTGGCGAGAAATAACCAATACTCTCTTATAAGACCCCTTTGCACGGTTTGTGG
TACTGTAATCGAATAATCCCCATCAATTGTTCAAATCCCTCATTTGGCACCCTTAAAAAGGTATACGTTTT
TTAATCGCGTCAGAACTCCTTTTCCTTTAGTTCCCTTAATTGTCATTTCGACTTTTAATCGTTTCAAGATT
TTTCAGTCCTCCCAGTTTTTTTTTTAGGGGAATGAAGCAATTGTTAACCCGCCAGGGTTGGATACGGGTTTA
TCATTTAAACCTGTGCTACCCTCGACATAATACTTGTATGACCTAACTACTTTAACATTCATAGGACATTA
TTTGTCAAGATTGAGGTAATTAATTGTTACGGGAGAATAATTGGGAAGTCATAATTGACCCAGTGGGTAAT
GAATAATCAGATATACATTTAAGAAAGACATTTCATGGACTTTGACCATTAAGTGTAGTGAACAACTATTT
AATTAATGTAAGGAGCATTGAACTACATGGTACATATATCGATATTTTGGCAAAGACTTATTTATATGGGG
ACCCCTGGGTGGAAAAGAAAAAAAAAAACAATATACATAAATTTTTGGAAAAACCCCCCCTCCCCCCAAAT
ATACACGGTATCCTCGAAAAACCCCTAAAACGAGGGCCGATTATATATTTTTCTTTTAATGACATGTAGAA
TTAATTATCATCATTATATATAGTGTCACACTATGGCAT

>Haplotype: CAT2 GB#MN784030 (Sample: RPCAT2)
AAAGATCAACACATGACCAACCTGCATAAACGCATGTATTTGCGTGAACGCAAATGCATTTGTGAAAACAC
AAATGCAATTGCAAAATTATGCAGGTTGGTCAGTCTGACATATAAACTATTTAGCCCCATACTATGTATAT
CCTCATTAAATATACATACTATGCTTAATCCTCATTAATAGACTGTCCCCATATACATATTATGCTTAATC
CACATCATACATATTATGCTTAACCCCCACATTACATATTATGCTTAATCCACATCATACATATTATGCTT
AATCCACATCATACATACGTCTATCTCACTCCCCATACACACATATATGTGTGCTTAATCCCCATTAATCC
ATACTCCCCTACTCCATTACATGCTATGCTTAATCCTCATTAATAGACTGTCCCCATAGTACATACTATGC
TTAATCCTCATTAACATACTATGCTTAATCCCCATTAATAGACATTCCCCTATTTCATTACATCTTACGTA
TCATTAATCCTCATTACTAGAAAATCCTCTACTCCATTACAATCATTCACTCACCCTCATTAATTTACTTC
TAATACTGCTTATGCGGGCTGGCGAGAAATAACCAATACTCTCTTATAAGACCCCTTTGCACGGTTTGTGG
TACTGTAATCGAATAATCCCCATCAATTGTTCAAATCCCTCATTTGGCACCCTTAAAAAGGTATACGTTTT
TTAATCGCGTCAGAACTCCTTTTCCTTTAGTTCCCTTAATTGTCATTTCGACTTTTAATCGTTTCAAGATT
TTTCAGTCCTCCCAGTTTTTTTTTTAGGGGAATGAAGCAATTGTTAACCCCCCAGGGTTGGATACGGCTTTA
TCATTTAAACCTGTGCTACCCTCGACATAATATTTGTATGACCTAACTACTTTAACATTCATAGGACATTA
TTTGTCAAGATTGAGCTAATTAATCGTTACGGGAGAATAATTGGGAAGTCATAATTGACCCAGGGGGTAAT
GAATAATCAGATATACATTTAAGAAAGACATTTCATGGACTTTGACCATTAAGTGTAGTGAACAAATATTT
AATTAATGTAAGGAGCATTGAACTACATGGTACATATATCGATATTTTGGCATAGACTTATTTATATGGGT
ACCCCTGGGTTGAAAAGAAAAAAAAAAACAATATACATAAATTTTTGGAAAAACCCCCCCTCCCCCCAAAT
ATACACGGTATCCTCGAAAAACCCCTAAAACGAGGGCCGATTATATATTTTTCTTTTAATGACATGTAGAA
TTAATTTATCATCATTATATATAGTGTCACCACTATGGCAT

>Haplotype: CAT3 GB#MN784031 (Sample: RPCAT3)
AAAGATCAACACATGACCAACCTGCATAAACGCATGTACTTGCGTGAACGCAAATGCATTTGTGAAAACAC
AAATGCAATTGCAAAATTATGCAGGTTGGTCAGTCTGACATATAAACTATTTAGCCCCATACTATGTATAT
CCTCATTAAATATACATACTATGCTTAATCCTCATTAATAGACTGTCCCCATATACATATTATGCTTAATC
CACATCATACATATTATGCTTAACCCCCACATTACATATTATGCTTAATCCACATCATACATATTATGCTT
AATCCACATCATACATACGTCTATCTCACTCCCCATACACACATATATGTGTGCTTAATCCCCATTAATCC
ATACTCCCCTACTCCATTACATGCTATGCTTAATCCTCATTAATAGACTGTCCCCATAGTACATACTATGC
TTAATCCTCATTAACATACTATGCTTAATCCCCATTAATAGACATTCCCCTATTTCATTACATCTTACGTA
TCATTAATCCTCATTACTAGAAAATCCTCTACTCCATTACAATCATTCACTCACCCTCATTAATTTACTTC
TAATACTGCTTATGCGGGCTGGCGAGAAATAACCAATACTCTCTTATAAGACCCCTTTGCACGGTTTGTGG
TACTGTAATCGAATAATCCCCATCAATTGTTCAAATCCCTCATTTGGCACCCTTAAAAAGGTATACGTCTT
TTAATCGCGTCAGAACTCCTTTTCCTCTAGTTCCCTTAATTGTCATTTCGACTTTTAATCGTTTCAAGATT
TTTCAGTCCTCCCAGTTTTTTTTTTAGGGGAATGAAGCAATTGTTAACCCGCCAGGGTTGGATACGGCTTTA
TCATTTAAACCTGTGCTACCCTCGACATAATACTTGTATGACCTAACTTCTTTAACATTTATAGGACATTA
TTTGTCAAGATTGAGCTAATTAATCGTTACGGGAGAATAATTGGGAAGTCATAATTGACCCAGTGGGTAAT

GAATAATCAGATATACATTTAAGAAAGACATTTCATGGACTTTGACCATTAAGTGTAGTGAACAACTATTT
AATTAATGTAAGGAGCATTGAACTACATGGTACATATATCGATATTTTGGCATAGACTTATTTATATGGGT
ACCCCTGGGTTGAAAAGAAAAAAAAAAAACAATATACATAAATTTTTGGAAAAACCCCCCCTCCCCCCAAAT
ATACACGGTATCCTCGAAAAACCCCTAAAACGAGGGCCGATTATATATTTTTCTTTTAATGACATGTAGAA
TTAATTATCATCATTATATATAGTGTCACACTATGGCAT

>Haplotype: CAT4 GB#MN784032 (Sample: RPCAT4)
AAAGATCAACACATGACCAACCTGCATAAACGCATGTACTTGCGTGAACGCAAATGCATTTGTGAAAACAC
AAATGCAATTGCAAAATTATGCAGGTTGGTCAGTCTGACATATAAACTATTTAGCCCCATACTATGTATAT
CCTCATTAAATATACATACTATGCTTAATCCTCATTAATAGACTGTCCCCATATACATATTATGCTTAATC
CACATCATACATATTATGCTTAACCCCCACATTACATATTATGCTTAATCCACATCATACATATTATGCTT
AATCCACATCATACATACGTCTATCTCACTCCCCATACACACATATATGTGTGCTTAATCCCCATTAATCC
ATACTCCCCTACTCCATTACATGCTATGCTTAATCCTCATTAATAGACTGTCCCCATAGTACATACTATGC
TTAATCCTCATTAACATACTATGCTTAATCCCCATTAATAGACATTCCCCTATTTCATTACATCTTACGTA
TCATTAATCCTCATTACTAGAAAATCCTCTACTCCATTACAATCATTCACTCACCCTCATTAATTTACTTC
TAATACTGCTTATGCGGGCTGGCGAGAAATAACCAATACTCTCTTATAAGACCCCTTTGCACGGTTTGTGG
TACTGTAATCGAATAATCCCCATCAATTGTTCAAATCCCTCATTTGGCACCCTTAAAAAGGTATACGTCTC
TTAATCGCGTCAGAACTCCTTTTCCTTTAGTTCCCTTAATTGTCATTTCGACTTTTAATCGTTTCAAGATT
TTTCAGTCCTCCCAGTTTTTTTTTAGGGGAATGAAGCAATTGTTAACCCGCCAGGGTTGGATACGGGTTTA
TCATTTAAACCTGTGCTACCCTCGACATAATACTTGTATGACCTAACTTCTTTAACATTCATAGGACATTA
TTTGTCAAGATTGAGGTAATTAATCGTTACGGGAGAATAATTGGGAAGTCATAATTGACCCAGGGGGTAAA
GAATAATCAGATATACATTTAAGAAAGACATTTCATGGACTTTGACCATTAAGTGTAGTGAACAAATATTT
AATTAATGTAAGGAGCATTGAACTACATGGTACATATATCGATATTTTGGCATAGACTTATTTATATGGGT
ACCCCTGGGTTGAAAAGAAAAAAAAAAAACAATATACATAAATTTTTGGAAAAACCCCCCCTCCCCCCAAAT
ATACACGGTATCCTCGAAAAACCCCTAAAACGAGGGCCGATTATATATTTTTCTTTTAATGACATGTAGAA
TTAATTATCATCATTATATATAGTGTCACACTATGGCAT

>Haplotype: CAT5 GB#MN784033 (Sample: RPCAT5)
AAAGATCAACACATGACCAACCTGCATAAACGCATGTACTTGCGTGAACGCAAATGCATTTGTGAAAACAC
AAATGCAATTGCAAAATTATGCAGGTTGGTCAGTCTGACATATAAACTATTTAGCCCCATACTATGTATAT
CCTCATTAAATATACATACTATGCTTAATCCTCATTAATAGACTGTCCCCATATACATATTATGCTTAATC
CACATCATACATATTATGCTTAACCCCCACATTACATATTATGCTTAATCCACATCATACATATTATGCTT
AATCCACATCATACATACGTCTATCTCACTCCCCATACACACATATATGTGTGCTTAATCCCCATTAATCC
ATACTCCCCTACTCCATTACATGCTATGCTTAATCCTCATTAATAGACTGTCCCCATAGTACATACTATGC
TTAATCCTCATTAACATACTATGCTTAATCCCCATTAATAGACATTCCCCTATTTCATTACATCTTACGTA
TCATTAATCCTCATTACTAGAAAATCCTCTACTCCATTACAATCATTCACTCACCCTCATTAATTTACTTC
TAATACTGCTTATGCGGGCTGGCGAGAAATAACCAATACTCTCTTATAAGACCCCTTTGCACGGTTTGTGG
TACTGTAATCGAATAATCCCCATCAATTGTTCAAATCCCTCATTTGGCACCCTTAAAAAGGTATACGTCTT
TTAATCGCGTCAGAACTCCTTTTCCTTTAGTTCCCTTAATTGTCATTTCGACTTTTAATCGTTTCAAGATT
TTTCAGTCCTCCCAGTTTTTTTTTAGGGGAATGAAGCAATTGTTAACCCCCCAGGGTTGGATACGGCTTTA
TCATTTAAACCTGTGCTACCCTCGACATAATACTTGTATGACCTAACTACTTTAACATTCATAGGACATTA
TTTGTCAAGATTGAGGTAATTAATCGTTACGGGAGAATAATTGGGAAGTCATAATTGACCCAGGGGGTAAT
GAATAATCAGATATACATTTAAGAAAGACATTTCATGGACTTTGACCATTAAGTGTAGTGAACAAATATTT
AATTAATGTAAGGAGCATTGAACTACATGGTACATATATCGATATTTTGGCATAGACTTATTTATATGGGT
ACCCCTGGGTTGAAAAGAAAAAAAAAAAACAATATACATAAATTTTTGGAAAAACCCCCCCTCCCCCCAAAT
ATACACGGTATCCTCGAAAAACCCCTAAAACGAGGGCCGATTATATATTTTTCTTTTAATGACATGTAGAA
TTAATTATCATCATTATATATAGTGTCACACTATGGCAT

>Haplotype: CAT6 GB#MN784034 (Sample: RPCAT6)
AAAGATCAACACATGACCAACCTGCATAAACGCATGTACTTGCGTGAACGCAAATGCATTTGTGAAAACAC
AAATGCAATTGCAAAATTATGCAGGTTGGTCAGTCTGACATATAAACTATTTAGCCCCATACTATGTATAT
CCTCATTAAATATACATACTATGCTTAATCCTCATTAATAGACTGTCCCCATATACATATTATGCTTAATC
CACATCATACATATTATGCTTAACCCCCACATTACATATTATGCTTAATCCACATCATACATATTATGCTT
AATCCACATCATACATACGTCTATCTCACTCCCCATACACACATATATGTGTGCTTAATCCCCATTAATCC
ATACTCCCCTACTCCATTACATGCTATGCTTAATCCTCATTAATAGACTGTCCCCATAGTACATACTATGC
TTAATCCTCATTAACATACTATGCTTAATCCCCATTAATAGACATTCCCCTATTTCATTACATCTTACGTA
TCATTAATCCTCATTACTAGAAAATCCTCTACTCCATTACAATCATTCACTCACCCTCATTAATTTACTTC
TAATACTGCTTATGCGGGCTGGCGAGAAATAACCAATACTCTCTTATAAGACCCCTTTGCACGGTTTGTGG
TACTGTAATCGAATAATCCCCATCAATTGTTCAAATCCCTCATTTGGCACCCTTAAAAAGGTATACGTCTT

TTAATTGCGTCAGAACTCCTTTTCCTTTAGTTCCCTTAATTGTCATTTCGACTTTTAATTGTTTCAAGATT
TTTCAGTCCTCCCAGTTTTTTTTTTAGGGGAATGAAGCAATTGTTAACCCGCCAGGGTTGGATACGGGTTTA
TCATTTAAACCTGTGCTACCCTTGACATAATACTTGTATGACCTAACTACTTTAACATTCATAGGACATTA
TTTGTCAAGATTGAGGTAATTAATTGTTACGGGAGAATAATTGGGAAGTCATAATTGACCCAGGGGGTAAT
GAATAATCAGATATACATTTAAGAAAGACATTTCATGGACTTTGACCATTAAGTGTAGTGAACAACTTTTT
AATTAATGTAAGGAGCATTGAACTACATGGTACATATATCGATATTTTGGCATAGACTTATTTATATGGGT
ACCCCTGGGTTGAAAAGAAAAAAAAAAACAATATACATAAATTTTTGGAAAAACCCCCCCTCCCCCCAAAT
ATACACGGTATCCTCGAAAAACCCCTAAAACGAGGGCCGATTATATATTTTTCTTTTAATGACATGTAGAA
TTAATTATCATCATTATATATAGTGTCACACTATGGCAT

>Haplotype: CAT7 GB#MN784035 (Sample: RPCAT7)
AAAGATCAACACATGACCAACCTGCATAAACGCATGTATTTGCGTGAACGCAAATGCATTTGTGAAAACAC
AAATGCAATTGCAAAATTATGCAGGTTGGTCAGTCTGACATATAAACTATTTAGCCCCATACTATGTATAT
CCTCATTAAATATACATACTATGCTTAATCCTCATTAATAGACTGTCCCCATATACATATTATGCTTAATC
CACATCATACATATTATGCTTAACCCCCACATTACATATTATGCTTAATCCACATCATACATATTATGCTT
AATCCACATCATACATACGTCTATCTCACTCCCCATACACACATATATGTGTGCTTAATCCCCATTAATCC
ATACTCCCCTACTCCATTACATGCTATGCTTAATCCTCATTAATAGACTGTCCCCATAGTACATACTATGC
TTAATCCTCATTAACATACTATGCTTAATCCCCATTAATAGACATTCCCCTATTTCATTACATCTTACGTA
TCATTAATCCTCATTACTAGAAAATCCTCTACTCCATTACAATCATTCACTCACCCTCATTAATTTACTTC
TAATACTGCTTATGCGGGCTGGCGAGAAATAACCAATACTCTCTTATAAGACCCCTTTGCACGGTTTGTGG
TACTGTAATCGAATAATCCCCATCAATTGTTCAAATCCCTCATTTGGCACCCTTAAAAAGGTATACGTTTT
TTAATCGCGTCAGAACTCCTTTTCCTCTAGTTCCCTTAATTGTCATTTCGACTTTTAATCGTTTCAAGATT
TTTCAGTCCTCCCAGTCTTTTTTTTAGGGGAATGAAGCAATTGTTAACCCGCCAGGGTTGGATACGGCTTTA
TCATTTAAACCTGTGCTACCCTTGACATAATATTTGTATGACCTAACTTCTTTAACATTCATAGGACATTA
TTTGTCAAGATTGAGGTAATTAATTGTTACGGGAGAATAATTGGGAAGTCATAATTGACACAGGGGGTAAA
GAATAATCAGATATACATTTAAGAAAGACATTTCATGGACTTTGACCATTAAGTGTAGTGAACAACTATTT
AATTAATGTAAGGAGCATTGAACTACATGGTACATATATTGATATTTTGGCATAGACTTATTTATATGGGT
ACCCCTGGGTTGAAAAGAAAAAAAAAAACAATATACATAAATTTTTGGAAAAACCCCCCCTCCCCCCAAAT
ATACACGGTATCCTCGAAAAACCCCTAAAACGAGGGCCGATTATATATTTTTCTTTTAATGACATGTAGAA
TTAATTTATCATCATTATATATAGTGTCACACTATGGCAT

>Haplotype: CAT8 GB#MN784036 (Sample: RPCAT8)
AAAGATCAACACATGACCAACCTGCATAAACGCATGTACTTGCGTGAACGCAAATGCATTTGTGAAAACAC
AAATGCAATTGCAAAATTATGCAGGTTGGTCAGTCTGACATATAAACTATTTAGCCCCATACTATGTATAT
CCTCATTAAATATACATACTATGCTTAATCCTCATTAATAGACTGTCCCCATATACATATTATGCTTAATC
CACATCATACATATTATGCTTAACCCCCACATTACATATTATGCTTAATCCACATCATACATATTATGCTT
AATCCACATCATACATACGTCTATCTCACTCCCCATACACACATATATGTGTGCTTAATCCCCATTAATCC
ATACTCCCCTACTCCATTACATGCTATGCTTAATCCTCATTAATAGACTGTCCCCATAGTACATACTATGC
TTAATCCTCATTAACATACTATGCTTAATCCCCATTAATAGACATTCCCCTATTTCATTACATCTTACGTA
TCATTAATCCTCATTACTAGAAAATCCTCTACTCCATTACAATCATTCACTCACCCTCATTAATTTACTTC
TAATACTGCTTATGCGGGCTGGCGAGAAATAACCAATACTCTCTTATAAGACCCCTTTGCACGGTTTGTGG
TACTGTAATCGAATAATCCCCATCAATTGTTCAAATCCCTCATTTGGCACCCTTAAAAAGGTATACGTTTT
TTAATCGCGTCAGAACTCCTTTTCCTTTAGTTCCCTTAATTGTCATTTCGACTTTTAATCGTTTCAAGATT
TTTCAGTCCTCCCAGTTTTTTTTTTAGGGGAATGAAGCAATTGTTAACCCGCCAGGGTTGGATACGGCTTTA
TCATTTAAACCTGTGCTACCCTTGACATAATACTTGTATGACCTAACTTCTTTAACATTCATAGGACATTA
TTTGTCAAGATTGAGGTAATTAATCGTTACGGGGGAATAATTGGGAAGTCATAATTGACCCAGTGGGTAAT
GAATAATCAGATATACATTTAAGAAAGACATTTCATGGACTTTGACCATTAAGTGTAGTGAACAACTATTT
AATTAATGTAAGGAGCATTGAACTACATGGTACATATATCGATATTTTGGCATAGACTTATTTATATGGGT
ACCCCTGGGTTGAAAAGAAAAAAAAAAACAATATACATAAATTTTTGGAAAAACCCCCCCTCCCCCCAAAT
ATACACGGTATCCTCGAAAAACCCCTAAAACGAGGGCCGATTATATATTTTTCTTTTAATGACATGTAGAA
TTAATTATCATCATTATATATAGTGTCACACTATGGCAT

>Haplotype: CAT9 GB#MN784044 (Sample: TS263)
AAAGATCAACACATGACCAACCTGCATAAACGCATGTACTTGCGTGAACGCAAATGCATTTGTGAAAACAC
AAATGCAATTGCAAAATTATGCAGGTTGGTCAGTCTGACATATAAACTATTTAGCCCCATACTATGTATAT
CCTCATTAAATATACATACTATGCTTAATCCTCATTAATAGACTGTCCCCATATACATATTATGCTTAATC
CACATCATACATATTATGCTTAACCCCCACATTACATATTATGCTTAATCCACATCATACATATTATGCTT
AATCCACATCATACATACGTCTATCTCACTCCCCATACACACATATATGTGTGCTTAATCCCCATTAATCC
ATACTCCCCTACTCCATTACATGCTATGCTTAATCCTCATTAATAGACTGTCCCCATAGTACATACTATGC

TTAATCCTCATTAACATACTATGCTTAATCCCCATTAATAGACATTCCCCTATTTCATTACATCTTACGTA
TCATTAATCCTCATTACTAGAAAATCCTCTACTCCATTACAATCATTCACTCACCCTCATTAATTTACTTC
TAATACTGCTTATGCGGGCTGGCGAGAAATAACCAATACTCTCTTATAAGACCCCTTTGCACGGTTTGTGG
TACTGTAATCGAATAATCCCCATCAATTGTTCAAATCCCTCATTTGGCACCCTTAAAAAGGTATACGTTTT
TTAATCGCGTCAGAACTCCTTTTCCTTTAGTTCCCTTAATTGTCATTTCGACTTTTAATCGTTTCAAGATT
TTTCAGTCCTCCCAGTTTTTTTTTAGGGGAATGAAGCAATTGTTAACCCGCCAGGGTTGGATACGGCTTTA
TCATTTAAACCTGTGCTACCCTCGACATAATACTTGTATGACCTAACTTCTTTAACATTCATAGGACATTA
TTTGTCAAGATTGAGGTAATTAATCGTTACGGGAGAATAATTGGGAAGTCATAATTGACACAGTGGGTAAT
GAATAATCAGATATACATTTAAGAAAGACATTTCATGGACTTTGACCATTAAGTGTAGTGAACAACTATTT
AATTAATGTAAGGAGCATTGAACTACATGGTACATATATCGATATTTTGGCATAGACTTATTTATATGGGT
ACCCCTGGGTTGAAAAGAAAAAAAAAAAACAATATACATAAATTTTTGGAAAAACCCCCCCTCCCCCCAAAT
ATACACGGTATCCTCGAAAAACCCCTAAAACGAGGGCCGATTATATATTTTTCTTTTAATGACATGTAGAA
TTAATTATCATCATTATATATAGTGTCACCACTATGGCAT

>Haplotype: CAT10 GB#MN784045 (Sample: TS264)
AAAGATCAACACATGACCAACCTGCATAAACGCATGTACTTGCGTGAACGCAAATGCATTTGTGAAAACAC
AAATGCAATTGCAAAATTATGCAGGTTGGTCAGTCTGACATATAAACTATTTAGCCCCATACTATGTATAT
CCTCATTAAATATACATACTATGCTTAATCCTCATTAATAGACTGTCCCCATATACATATTATGCTTAATC
CACATCATACATATTATGCTTAACCCCCACATTACATATTATGCTTAATCCACATCATACATATTATGCTT
AATCCACATCATACATACGTCTATCTCACTCCCCATACACACATATATGTGTGCTTAATCCCCATTAATCC
ATACTCCCCTACTCCATTACATGCTATGCTTAATCCTCATTAATAGACTGTCCCCATAGTACATACTATGC
TTAATCCTCATTAACATACTATGCTTAATCCCCATTAATAGACATTCCCCTATTTCATTACATCTTACGTA
TCATTAATCCTCATTACTAGAAAATCCTCTACTCCATTACAATCATTCACTCACCCTCATTAATTTACTTC
TAATACTGCTTATGCGGGCTGGCGAGAAATAACCAATACTCTCTTATAAGACCCCTTTGCACGGTTTGTGG
TACTGTAATCGAATAATCCCCATCAATTGTTCAAATCCCTCATTTGGCACCCTTAAAAAGGTATACGTTTT
TTAATCGCGTCAGAACTCCTTTTCCTTTAGTTCCCTTAATTGTCATTTTGACTTTTAATTGTTTCAAGATT
TTTCAGTCCTCCCAGTTTTTTTTTAGGGGAATGAAGCAATTGTTAACCCGCCAGGGTTGGATACGGCTTTA
TCATTTAAACCTGTGCTACCCTCGACATAATACTTGTATGACCTAACTTCTTTAACATTCATAGGACATTA
TTTGTCAAGATTGAGGTAATTAATCGTTACGGGAGAATAATTGGGAAGTCATAATTGACCCAGTGGGTAAT
GAATAATCAGATATACATTTAAGAAAGACATTTCATGGACTTTGACCATTAAGTGTAGTGAACAACTATTT
AATTAATGTAAGGAGCATTGAACTACATGGTACATATATCGATATTTTGGCATAGACTTATTTATATGGGT
ACCCCTGGGTCGAAAAGAAAAAAAAAAAACAATATACATAAATTTTTGGAAAAACCCCCCCTCCCCCCAAAT
ATACACGGTATCCTCGAAAAACCCCTAAAACGAGGGCCGATTATATATTTTTCTTTTAATGACATGTAGAA
TTAATTATCATCATTATATATAGTGTCACACTATGGCAT

>Haplotype: CAT11 GB#MN784046 (Sample: TS270)
AAAGATCAACACATGACCAACCTGCATAAACGCATGTACTTGCGTGAACGCAAATGCATTTGTGAAAACAC
AAATGCAATTGCAAAATTATGCAGGTTGGTCAGTCTGACATATAAACTATTTAGCCCCATACTATGTATAT
CCTCATTAAATATACATACTATGCTTAATCCTCATTAATAGACTGTCCCCATATACATATTATGCTTAATC
CACATCATACATATTATGCTTAACCCCCACATTACATATTATGCTTAATCCACATCATACATATTATGCTT
AATCCACATCATACATACGTCTATCTCACTCCCCATACACACATATATGTGTGCTTAATCCCCATTAATCC
ATACTCCCCTACTCCATTACATGCTATGCTTAATCCTCATTAATAGACTGTCCCCATAGTACATACTATGC
TTAATCCTCATTAACATACTATGCTTAATCCCCATTAATAGACATTCCCCTATTTCATTACATCTTACGTA
TCATTAATCCTCATTACTAGAAAATCCTCTACTCCATTACAATCATTCACTCACCCTCATTAATTTACTTC
TAATACTGCTTATGCGGGCTGGCGAGAAATAACCAATACTCTCTTATAAGACCCCTTTGCACGGTTTGTGG
TACTGTAATCGAATAATCCCCATCAATTGTTCAAATCCCTCATTTGGCACCCTTAAAAAGGTATACGTCTT
TTAATCGCGTCAGAACTCCTTTTCCTCTAGTTCCCTTAATTGTCATTTCGACTTTTAATCGTTTCAAGATT
TTTCAGTCCTCCCAGTTTTTTTTTAGGGGAATGAAGCAATTGTTAACCCCCCAGGGTTGGATACGGCTTTA
TCATTTAAACCTGTGCTACCCTTGACATAATACTTGTATGACCTAACTACTTTAACATTCATAGGACATTA
TTTGTCAAGATTGAGGTAATTAATCGTTACGGGAGAATAATTGGGAAGTCATAATTGACCCAGTGGGTAAT
GAATAATCAGATATACATTTAAGAAAGACATTTCATGGACTTTGACCATTAAGTGTAGTGAACAACTATTT
AATTAATGTAAGGAGCATTGAACTACATGGTACATATATCGATATTTTGGCAAAGACTTATTTATATGGGT
ACCCCTGGGTTGAAAAGAAAAAAAAAAACAATATACATAAATTTTTGGAAAAACCCCCCCTCCCCCCAAAT
ATACACGGTATCCTCGAAAAACCCCTAAAACGAGGGCCGATTATATATTTTTCTTTTAATGACATGTAGAA
TTAATTATCATCATTATATATAGTGTCACACTATGGCAT

>Haplotype: CAT12 GB#MN784047 (Sample: TS272)
AAAGATCAACACATGACCAACCTGCATAAACGCATGTATTTGCGTGAACGCAAATGCATTTGTGAAAACAC
AAATGCAATTGCAAAATTATGCAGGTTGGTCAGTCTGACATATAAACTATTTAGCCCCATACTATGTATAT

CCTCATTAAATATACATACTATGCTTAATCCTCATTAATAGACTGTCCCCATATACATATTATGCTTAATC
CACATCATACATATTATGCTTAACCCCCACATTACATATTATGCTTAATCCACATCATACATATTATGCTT
AATCCACATCATACATACGTCTATCTCACTCCCCATACACACATATATGTGTGCTTAATCCCCATTAATCC
ATACTCCCCTACTCCATTACATGCTATGCTTAATCCTCATTAATAGACTGTCCCCATAGTACATACTATGC
TTAATCCTCATTAACATACTATGCTTAATCCCCATTAATAGACATTCCCCTATTTCATTACATCTTACGTA
TCATTAATCCTCATTACTAGAAAATCCTCTACTCCATTACAATCATTCACTCACCCTCATTAATTTACTTC
TAATACTGCTTATGCGGGCTGGCGAGAAATAACCAATACTCTCTTATAAGACCCCTTTGCACGGTTTGTGG
TACTGTAATCGAATAATCCCCATCAATTGTTCAAATCCCTCATTTGGCACCCTTAAAAAGGTATACGTTTT
TTAATCGCGTCAGAACTCCTTTTCCTTTAGTTCCCTTAATTGTCATTTCGACTTTTAATCGTTTCAAGATT
TTTCAGTCCTCCCAGTTTTTTTTTAGGGGAATGAAGCAATTGTTAACCCGCCAGGGTTGGATACGGCTTTA
TCATTTAAACCTGTGCTACCCTTGACATAATATTTGTATGACCTAACTACTTTAACATTCATAGGACATTA
TTTGTCAAGATTGAGGTAATTAATCGTTACGGGAGAATAATTGGGAAGTCATAATTGACCCAGTGGGTAAT
GAATAATCAGATATACATTTAAGAAAGACATTTCATGGACTTTGACCATTAAGTGTAGTGAACAACTATTT
AATTAATGTAAGGAGCATTGAACTACATGGTACATATATCGATATTTTGGCATAGACTTATTTATATGGGT
ACCCCTGGGTTGAAAAGAAAAAAAAAAACAATATACATAAATTTTTGGAAAAACCCCCCCTCCCCCCAAAT
ATACACGGTATCCTCGAAAAACCCCTAAAACGAGGGCCGATTATATATTTTTCTTTTAATGACATGTAGAA
TTAATTTATCATCATTATATATAGTGTCACACTATGGCAT

>Haplotype: CAT13 GB#MN784048 (Sample: TS273)
AAAGATCAACACATGACCAACCTGCATAAACGCATGTACTTGCGTGAACGCAAATGCATTTGTGAAAACAC
AAATGCAATTGCAAAATTATGCAGGTTGGTCAGTCTGACATATAAACTATTTAGCCCCATACTATGTATAT
CCTCATTAAATATACATACTATGCTTAATCCTCATTAATAGACTGTCCCCATATACATATTATGCTTAATC
CACATCATACATATTATGCTTAACCCCCACATTACATATTATGCTTAATCCACATCATACATATTATGCTT
AATCCACATCATACATACGTCTATCTCACTCCCCATACACACATATATGTGTGCTTAATCCCCATTAATCC
ATACTCCCCTACTCCATTACATGCTATGCTTAATCCTCATTAATAGACTGTCCCCATAGTACATACTATGC
TTAATCCTCATTAACATACTATGCTTAATCCCCATTAATAGACATTCCCCTATTTCATTACATCTTACGTA
TCATTAATCCTCATTACTAGAAAATCCTCTACTCCATTACAATCATTCACTCACCCTCATTAATTTACTTC
TAATACTGCTTATGCGGGCTGGCGAGAAATAACCAATACTCTCTTATAAGACCCCTTTGCACGGTTTGTGG
TACTGTAATCGAATAATCCCCATCAATTGTTCAAATCCCTCATTTGGCACCCTTAAAAAGGTATACGTTTT
TTAATCGCGTCAGAACTCCTTTTCCTTTAGTTCCCTTAATTGTCATTTCGACTTTTAATCGTTTCAAGATT
TTTCAGTCCTCCCAGTTTTTTTTTAGGGGAATGAAGCAATTGTTAACCCGCCAGGGTTGGATACGGCTTTA
TCATTTAAACCTGTGCTACCCTCGACATAATACTTGTATGACCTAACTTCTTTAACATTCATAGGACATTA
TTTGTCAAGATTGAGGTAATTAATCGTTACGGGAGAATAATTGGGAAGTCATAATTGACCCAGGGGGTAAT
GAATAATCAGATATACATTTAAGAAAGACATTTCATGGACTTTGACCATTAAGTGTAGTGAACAACTATTT
AATTAATGTAAGGAGCATTGAACTACATGGTACATATATCGATATTTTGGCATAGACTTATTTATATGGGT
ACCCCTGGGTTGAAAAGAAAAAAAAAAACAATATACATAAATTTTTGGAAAAACCCCCCCTCCCCCCAAAT
ATACACGGTATCCTCGAAAAACCCCTAAAACGAGGGCCGATTATATATTTTTCTTTTAATGACATGTAGAA
TTAATTATCATCATTATATATAGTGTCACACTATGGCAT

>Haplotype: DW1 GB#MN784037 (Sample: RPDW1)
AAAGATCAACACATGACCAACCTGCATAAACGCATGTATTTGCGTGAACGCAAATGCATTTGTGAAAACAC
AAATGCAATTGCAAAATTATGCAGGTTGGTCAGTCTGACATATAAACTATTTAGCCCCATACTATGTATAT
CCTCATTAAATATACATACTATGCTTAATCCTCATTAATAGACTGTCCCCATATACATATTATGCTTAATC
CACATCATACATATTATGCTTAACCCCCACATTACATATTATGCTTAATCCACATCATACATATTATGCTT
AATCCACATCATACATACGTCTATCTCACTCCCCATACACACATATATGTGTGCTTAATCCCCATTAATCC
ATACTCCCCTACTCCATTACATGCTATGCTTAATCCTCATTAATAGACTGTCCCCATAGTACATACTATGC
TTAATCCTCATTAACATACTATGCTTAATCCCCATTAATAGACATTCCCCTATTTCATTACATCTTACGTA
TCATTAATCCTCATTACTAGAAAATCCTCTACTCCATTACAATCATTCACTCACCCTCATTAATTTACTTC
TAATACTGCTTATGCGGGCTGGCGAGAAATAACCAATACTCTCTTATAAGACCCCTTTGCACGGTTTGTGG
TACTGTAATCGAATAATCCCCATCAATTGTTCAAATCCCTCATTTGGCACCCTTAAAAAGGTATACGTTTT
TTAATCGCGTCAGAACTCCTTTTCCTTTAGTTCCCTTAATTGTCATTTCGACTTTTAATCGTTTCAAGATT
TTTCAGTCCTCCCAGTTTTTTTTTAGGGGAATGAAGCAATTGTTAACCCGCCAGGGTTGGATACGGCTTTA
TCATTTAAACCTGTGCTACCCTTGACATAATATTTGTATGACCTAACTTCTTTAACATTCATAGGACATTA
TTTGTCAAGATTGAGGTAATTAATCGTTACGGGAGAATAATTGGGAAGTCATAATTGACCCAGTGGGTAAT
GAATAATCAGATATACATTTAAGAAAGACATTTCATGGACTTTGACCATTAAGTGTAGTGAACAACTATTT
AATTAATGTAAGGAGCATTGAACTACATGGTACATATATCGATATTTTGGCATAGACTTATTTATATGGGT
ACCCCTGGGTTGAAAAGAAAAAAAAAAACAATATACATAAATTTTTGGAAAAACCCCCCCTCCCCCCAAAT
ATACACGGTATCCTCGAAAAACCCCTAAAACGAGGGCCGATTATATATTTTTCTTTTAATGACATGTAGAA
TTAATTTATCATCATTATATATAGTGTCACACTATGGCAT

>Haplotype: LB1 GB#MN784038 (Sample: RPLBH1)
AAAGATCAACACATGACCAACCTGCATAAACGCATGTATTTGCGTGAACGCAAATGCATTTGTGAAAACAC
AAATGCAATTGCAAAATTATGCAGGTTGGTCAGTCTGACATATAAACTATTTAGCCCCATACTATGTATAT
CCTCATTAAATATACATACTATGCTTAATCCTCATTAATAGACTGTCCCCATATACATATTATGCTTAATC
CACATCATACATATTATGCTTAACCCCCACATTACATATTATGCTTAATCCACATCATACATATTATGCTT
AATCCACATCATACATACGTCTATCTCACTCCCCATACACACATATATGTGTGCTTAATCCCCATTAATCC
ATACTCCCCTACTCCATTACATGCTATGCTTAATCCTCATTAATAGACTGTCCCCATAGTACATACTATGC
TTAATCCTCATTAACATACTATGCTTAATCCCCATTAATAGACATTCCCCTATTTCATTACATCTTACGTA
TCATTAATCCTCATTACTAGAAAATCCTCTACTCCATTACAATCATTCACTCACCCTCATTAATTTACTTC
TAATACTGCTTATGCGGGCTGGCGAGAAATAACCAATACTCTCTTATAAGACCCCTTTGCACGGTTTGTGG
TACTGTAATCGAATAATCCCCATCAATTGTTCAAATCCCTCATTTGGCACCCTTAAAAAGGTATACGTCTT
TTAATCGCGTCAGAACTCCTTTTCCTTTAGTTCCCTTAATTGTCATTTCGACTTTTAATCGTTTCAAGATT
TTTCAGTCCTCCCAGTTTTTTTTTTAGGGGAATGAAGCAATTGTTAACCCGCCAGGGTTGGATACGGCTTTA
TCATTTAAACCTGTGCTACCCTTGACATAATATTTGTATGACCTAACTACTTTAACATTCATAGGACATTA
TTTGTCAAGATTGAGGTAATTAATCGTTACGGGAGAATAATTGGGAAGTCATAATTGACCCAGTGGGTAAT
GAATAATCAGATATACATTTAAGAAAGACATTTCATGGACTTTGACCATTAAGTGTAGTGAACAACTATTT
AATTAATGTAAGGAGCATTGAACTACATGGTACATATATTGATATTTTGGCATAGACTTATTTATATGGGT
ACCCCTGGGTTGAAAAGAAAAAAAAAAACAATATACATAAATTTTTGGAAAAACCCCCCCTCCCCCCAAAT
ATACACGGTATCCTCGAAAAACCCCTAAAACGAGGGCCGATTATATATTTTTCTTTTAATGACATGTAGAA
TTAATTTATCATCATTATATATAGTGTCACACTATGGCAT

>Haplotype: LB2 GB#MN784039 (Sample: RPLBH2)
AAAGATCAACACATGACCAACCTGCATAAACGCATGTACTTGCGTGAACGCAAATGCATTTGTGAAAACAC
AAATGCAATTGCAAAATTATGCAGGTTGGTCAGTCTGACATATAAACTATTTAGCCCCATACTATGTATAT
CCTCATTAAATATACATACTATGCTTAATCCTCATTAATAGACTGTCCCCATATACATATTATGCTTAATC
CACATCATACATATTATGCTTAACCCCCACATTACATATTATGCTTAATCCACATCATACATATTATGCTT
AATCCACATCATACATACGTCTATCTCACTCCCCATACACACATATATGTGTGCTTAATCCCCATTAATCC
ATACTCCCCTACTCCATTACATGCTATGCTTAATCCTCATTAATAGACTGTCCCCATAGTACATACTATGC
TTAATCCTCATTAACATACTATGCTTAATCCCCATTAATAGACATTCCCCTATTTCATTACATCTTACGTA
TCATTAATCCTCATTACTAGAAAATCCTCTACTCCATTACAATCATTCACTCACCCTCATTAATTTACTTC
TAATACTGCTTATGCGGGGTGGCGAGAAATAACCAATACTCTCTTATAAGACCCCTTTGCACGGTTTGTGG
TACTGTAATCGAATAATCCCCATCAATTGTTCAAATCCCTCATTTGGCACCCTTAAAAAGGTATACGTCTT
TTAATCGCGTCAGAACTCCTTTTCCTCTAGTTCCCTTAATTGTCATTTCGACTTTTAATTGTTTCAAGATT
TTTCAGTCCTCCCAGTTTTTTTTTTAGGGGAATGAAGCAATTGTTAACCCGCCAGGGTTGGATACGGCTTTA
TCATTTAAACCTGTGCTACCCTCGACATAATACTTGTATGACCTAACTACTTTAACATTCATAGGACATTA
TTTGTCAAGATTGAGGTAATTAATTGTTACGGGAGAATAATTGGGAAGTCATAATTGACCCAGGGGGTAAT
GAATAATCAGATATACATTTAAGAAAGACATTTCATGGACTTTGACCATTAAGTGTAGTGAACAACTATTT
AATTAATGTAAGGAGCATTGAACTACATGGTACATATATCGATATTTTGGCATAGACTTATTTATATGGGT
ACCCCTGGGTTGAAAAGAAAAAAAAAAACAATATACATAAATTTTTGGAAAAACCCCCCCTCCCCCCAAAT
ATACACGGTATCCTCGAAAAACCCCTAAAACGAGGGCCGATTATATATTTTTCTTTTAATGACATGTAGAA
TTAATTATCATCATTATATATAGTGTCACACTATGGCAT

>Haplotype: LB3 GB#MN784040 (Sample: RPLBH3)
AAAGATCAACACATGACCAACCTGCATAAACGCATGTACTTGCGTGAACGCAAATGCATTTGTGAAAACAC
AAATGCAATTGCAAAATTATGCAGGTTGGTCAGTCTGACATATAAACTATTTAGCCCCATACTATGTATAT
CCTCATTAAATATACATACTATGCTTAATCCTCATTAATAGACTGTCCCCATATACATATTATGCTTAATC
CACATCATACATATTATGCTTAACCCCCACATTACATATTATGCTTAATCCACATCATACATATTATGCTT
AATCCACATCATACATACGTCTATCTCACTCCCCATACACACATATATGTGTGCTTAATCCCCATTAATCC
ATACTCCCCTACTCCATTACATGCTATGCTTAATCCTCATTAATAGACTGTCCCCATAGTACATACTATGC
TTAATCCTCATTAACATACTATGCTTAATCCCCATTAATAGACATTCCCCTATTTCATTACATCTTACGTA
TCATTAATCCTCATTACTAGAAAATCCTCTACTCCATTACAATCATTCACTCACCCTCATTAATTTACTTC
TAATACTGCTTATGCGGGCTGGCGAGAAATAACCAATACTCTCTTATAAGACCCCTTTGCACGGTTTGTGG
TACTGTAATCGAATAATCCCCATCAATTGTTCAAATCCCTCATTTGGCACCCTTAAAAAGGTATACGTCTC
TTAATCGCGTCAGAACTCCTTTTCCTCTAGTTCCCTTAATTGTCATTTCGACTTTTAATCGTTTCAAGATT
TCTCAGTCCTCCCAGTTTTTTTTTTAGGGGAATGAAGCAATCGTTAACCCGCCAGGGTTGGATACGGCTTTA
TCATTTAAACCTGTGCTACCCTCGACATAATACTTGTATGACCTAACTACTTTAACATTCATAGGACATTA
TTTGTCAAGATTGAGCTAATTAATCGTTACGGGAGAATAATTGTGAAGTCATAATTGACACAGTGGGTAAT
GAATAATCAGATATACATTTAAGAAAGACATTTCATGGACTTCGACCATTAAGTGTAGTGAACAACTATTT
AATTAATGTAAGGAGCATTGAACTACATGGTACATATATCGATATTTTGGCATAGACTTATTTATATGGGT

ACCCCTGGGTCGAAAAGAAAAAAAAAAACAATATACATAAATTTTTGGAAAAACCCCCCCTCCCCCCAAAT
ATACACGGTATCCTCGAAAAACCCCTAAAACGAGGGCCGATTATATATTTTTCTTTTAATGACATGTAGAA
TTAATTATCATCATTATATATAGTGTCACACTATGGCAT

>Haplotype: LB4 GB#MN784041 (Sample: RPLBH4)
AAAGATCAACACATGACCAACCTGCATAAACGCATGTATTTGCGTGAACGCAAATGCATTTGTGAAAACAC
AAATGCAATTGCAAAATTATGCAGGTTGGTCAGTCTGACATATAAACTATTTAGCCCCATACTATGTATAT
CCTCATTAAATATACATACTATGCTTAATCCTCATTAATAGACTGTCCCCATATACATATTATGCTTAATC
CACATCATACATATTATGCTTAACCCCCACATTACATATTATGCTTAATCCACATCATACATATTATGCTT
AATCCACATCATACATACGTCTATCTCACTCCCCATACACACATATATGTGTGCTTAATCCCCATTAATCC
ATACTCCCCTACTCCATTACATGCTATGCTTAATCCTCATTAATAGACTGTCCCCATAGTACATACTATGC
TTAATCCTCATTAACATACTATGCTTAATCCCCATTAATAGACATTCCCCTATTTCATTACATCTTACGTA
TCATTAATCCTCATTACTAGAAAATCCTCTACTCCATTACAATCATTCACTCACCCTCATTAATTTACTTC
TAATACTGCTTATGCGGGCTGGCGAGAAATAACCAATACTCTCTTATAAGACCCCTTTGCACGGTTTGTGG
TACTGTAATCGAATAATCCCCATCAATTGTTCAAATCCCTCATTTGGCACCCTTAAAAAGGTATACGTCTT
TTAATCGCGTCAGAACTCCTTTTCCTCTAGTTCCCTTAATTGTCATTTCGACTTTTAATCGTTTCAAGATT
TTTCAGTCCTCCCAGTTTTTTTTTAGGGGAATGAAGCAATCGTTAACCCGCCAGGGTTGGATACGGCTTTA
TCATTTAAACCTGTGCTACCCTCGACATAATATTTGTATGACCTAACTACTTTAACATTCATAGGACATTA
TTTGTCAAGATTGAGCTAATTAATCGTTACGGGAGAATAATTGTGAAGTCATAATTGACACAGTGGGTAAT
GAATAATCAGATATACATTTAAGAAAGACATTTCATGGACTTCGACCATTAAGTGTAGTGAACAACTATTT
AATTAATGTAAGGAGCATTGAACTACATGGTACATATATCGATATTTTGGCATAGACTTATTTATATGGGT
ACCCCTGGGTCGAAAAGAAAAAAAAAAAACAATATACATAAATTTTTGGAAAAACCCCCCCTCCCCCCAAAT
ATACACGGTATCCTCGAAAAACCCCTAAAACGAGGGCCGATTATATATTTTTCTTTTAATGACATGTAGAA
TTAATTTATCATCATTATATATAGTGTCACACTATGGCAT

>Haplotype: LB5 GB#MN784042 (Sample: RPLBH5)
AAAGATCAACACATGACCAACCTGCATAAACGCATGTATTTGCGTGAACGCAAATGCATTTGTGAAAACAC
AAATGCAATTGCAAAATTATGCAGGTTGGTCAGTCTGACATATAAACTATTTAGCCCCATACTATGTATAT
CCTCATTAAATATACATACTATGCTTAATCCTCATTAATAGACTGTCCCCATATACATATTATGCTTAATC
CACATCATACATATTATGCTTAACCCCCACATTACATATTATGCTTAATCCACATCATACATATTATGCTT
AATCCACATCATACATACGTCTATCTCACTCCCCATACACACATATATGTGTGCTTAATCCCCATTAATCC
ATACTCCCCTACTCCATTACATGCTATGCTTAATCCTCATTAATAGACTGTCCCCATAGTACATACTATGC
TTAATCCTCATTAACATACTATGCTTAATCCCCATTAATAGACATTCCCCTATTTCATTACATCTTACGTA
TCATTAATCCTCATTACTAGAAAATCCTCTACTCCATTACAATCATTCACTCACCCTCATTAATTTACTTC
TAATACTGCTTATGCGGGCTGGCGAGAAATAACCAATACTCTCTTATAAGACCCCTTTGCACGGTTTGTGG
TACTGTAATCGAATAATCCCCATCAATTGTTCAAATCCCTCATTTGGCACCCTTAAAAAGGTATACGTCTC
TTAATCGCGTCAGAACTCCTTTTCCTCTAGTTCCCTTAATTGTCATTTCGACTTTTAATCGTTTCAAGATT
TTTCAGTCCTCCCAGTCTTTTTTTTAGGGGAATGAAGCAATCGTTAACCCGCCAGGGTTGGATACGGCTTTA
TCATTTAAACCTGTGCTACCCTCGACATAATATTTGTATGACCTAACTACTTTAACATTCATAGGACATTA
TTTGTCAAGATTGAGCTAATTAATCGTTACGGGAGAATAATTGTGAAGTCATAATTGACACAGTGGGTAAT
GAATAATCAGATATACATTTAAGAAAGACATTTCATGGACTTCGACCATTAAGTGTAGTGAACAACTATTT
AATTAATGTAAGGAGCATTGAACTACATGGTACATATATCGATATTTTGGCATAGACTTATTTATATGGGT
ACCCCTGGGTCGAAAAGAAAAAAAAAAAACAATATACATAAATTTTTGGAAAAACCCCCCCTCCCCCCAAAT
ATACACGGTATCCTCGAAAAACCCCTAAAACGAGGGCCGATTATATATTTTTCTTTTAATGACATGTAGAA
TTAATTTATCATCATTATATATAGTGTCACACTATGGCAT

>Haplotype: LB6 GB#MN784043 (Sample: RPLBH6)
AAAGATCAACACATGACCAACCTGCATAAACGCATGTATTTGCGTGAACGCAAATGCATTTGTGAAAACAC
AAATGCAATTGCAAAATTATGCAGGTTGGTCAGTCTGACATATAAACTATTTAGCCCCATACTATGTATAT
CCTCATTAAATATACATACTATGCTTAATCCTCATTAATAGACTGTCCCCATATACATATTATGCTTAATC
CACATCATACATATTATGCTTAACCCCCACATTACATATTATGCTTAATCCACATCATACATATTATGCTT
AATCCACATCATACATACGTCTATCTCACTCCCCATACACACATATATGTGTGCTTAATCCCCATTAATCC
ATACTCCCCTACTCCATTACATGCTATGCTTAATCCTCATTAATAGACTGTCCCCATAGTACATACTATGC
TTAATCCTCATTAACATACTATGCTTAATCCCCATTAATAGACATTCCCCTATTTCATTACATCTTACGTA
TCATTAATCCTCATTACTAGAAAATCCTCTACTCCATTACAATCATTCACTCACCCTCATTAATTTACTTC
TAATACTGCTTATGCGGGCTGGCGAGAAATAACCAATACTCTCTTATAAGACCCCTTTGCACGGTTTGTGG
TACTGTAATCGAATAATCCCCATCAATTGTTCAAATCCCTCATTTGGCACCCTTAAAAAGGTATACGTCTC
TTAATCGCGTCAGAACTCCTTTTCCTCTAGTTCCCTTAATTGTCATTTCGACTTTTAATCGTTTCAAGATT
TTTCAGTCCTCCCAGTCTTTTTTTTAGGGGAATGAAGCAATCGTTAACCCGCCAGGGTTGGATACGGCTTTA

TCATTTAAACCTGTGCTACCCTCGACATAATATTTGTATGACCTAACTACTTTAACATTCATAGGACATTA
TTTGTCAAGATTGAGCTAATTAATCGTTACGGGAGAATAATTGGGAAGTCATAATTGACACAGTGGGTAAT
GAATAATCAGATATACATTTAAGAAAGACATTTCATGGACTTCGACCATTAAGTGTAGTGAACAACTATTT
AATTAATGTAAGGAGCATTGAACTACATGGTACATATATCGATATTTTGGCATAGACTTATTTATATGGGT
ACCCCTGGGTCGAAAAGAAAAAAAAAAACAATATACATAAATTTTTGGAAAAACCCCCCCTCCCCCCAAAT
ATACACGGTATCCTCGAAAAACCCCTAAAACGAGGGCCGATTATATATTTTTCTTTTAATGACATGTAGAA
TTAATTTATCATCATTATATATAGTGTCACACTATGGCAT

>Haplotype: LJ1 GB#MN784019 (Sample: RP001)
AAAGATCAACACATGACCAACCTGCATAAACGCATGTATTTGCGTGAACGCAAATGCATTTGTGAAAACAC
AAATGCAATTGCAAAATTATGCAGGTTGGTCAGTCTGACATATAAACTATTTAGCCCCATACTATGTATAT
CCTCATTAAATATACATACTATGCTTAATCCTCATTAATAGACTGTCCCCATATACATATTATGCTTAATC
CACATCATACATATTATGCTTAACCCCCACATTACATATTATGCTTAATCCACATCATACATATTATGCTT
AATCCACATCATACATACGTCTATCTCACTCCCCATACACACATATATGTGTGCTTAATCCCCATTAATCC
ATACTCCCCTACTCCATTACATGCTATGCTTAATCCTCATTAATAGACTGTCCCCATAGTACATACTATGC
TTAATCCTCATTAACATACTATGCTTAATCCCCATTAATAGACATTCCCCTATTTCATTACATCTTACGTA
TCATTAATCCTCATTACTAGAAAATCCTCTACTCCATTACAATCATTCACTCACCCTCATTAATTTACTTC
TAATACTGCTTATGCGGGCTGGCGAGAAATAACCAATACTCTTTTATAAGACCCCTTTGCACGGTTTGTGG
TACTGTAATCGAATAATCCCCATCAATTGTTCAAATCCCTCATTTGGCACCCTTAAAAAGGTATACGTCTT
TTAATCGCGTCAGAACTCCTTTTCCTTTAGTTCCCTTAATTGTCATTTCGACTTTTAATTGTTTCAAGATT
TTTCAGTCCTCCCAGTTTTTTTTTAGGGGAATGAAGCAATCGTTAACCCGCCAGGGTTGGATACGGCTTTA
TCATTTAAACCTGTGCTACCCTCGACATAATATTTGTATGACCTAACTACTTTAACATTCATAGGACATTA
TTTGTCAAGATTGAGGTAATTAATCGTTACGGGAGAATAATTGGGAAGTCATAATTGACACAGTGGGTAAT
GAATAATCAGATATACATTTAAGAAAGACATTTCATGGACTTTGACCATTAAGTGTAGTGAACAACTATTT
AATTAATGTAAGGAGCATTGAACTACATGGTACATATATCGATATTTTGGCATAGACTTATTTATATGGGT
ACCCCTGGGTTGAAAAGAAAAAAAAAAACAATATACATAAATTTTTGGAAAAACCCCCCCTCCCCCCAAAT
ATACACGGTATCCTCGAAAAACCCCTAAAACGAGGGCCGATTATATATTTTTCTTTTAATGACATGTAGAA
TTAATTTATCATCATTATATATAGTGTCACACTATGGCAT

>Haplotype: LJ2 GB#MN784020 (Sample: RP002)
AAAGATCAACACATGACCAACCTGCATAAACGCATGTACTTGCGTGAACGCAAATGCATTTGTGAAAACAC
AAATGCAATTGCAAAATTATGCAGGTTGGTCAGTCTGACATATAAACTATTTAGCCCCATACTATGTATAT
CCTCATTAAATATACATACTATGCTTAATCCTCATTAATAGACTGTCCCCATATACATATTATGCTTAATC
CACATCATACATATTATGCTTAACCCCCACATTACATATTATGCTTAATCCACATCATACATATTATGCTT
AATCCACATCATACATACGTCTATCTCACTCCCCATACACACATATATGTGTGCTTAATCCCCATTAATCC
ATACTCCCCTACTCCATTACATGCTATGCTTAATCCTCATTAATAGACTGTCCCCATAGTACATACTATGC
TTAATCCTCATTAACATACTATGCTTAATCCCCATTAATAGACATTCCCCTATTTCATTACATCTTACGTA
TCATTAATCCTCATTACTAGAAAATCCTCTACTCCATTACAATCATTCACTCCCCCTCATTAATTTACTTC
TAATACTGCTTATGCGGGGCTGGCGAGAAATAACCAATACTCTCTTATAAGACCCCTTTGCACGGTTTGTG
GTACTGTAATCGAATAATCCCCATCAATTGTTCAAATCCCTCATTTGGCACCCTTAAAAAGGTATACGTTT
TTTAATTGGGTCAGAACTCCTTTTCCTTTAGTTCCCTTAATTGTCATTTCGACTTTTAATCGTTTCAAGAT
TTTTCAGTCCTCCCAGTTTTTTTTTAGGGGAATGAAGCAATTGTTAACCCGCCAGGGTTGGATACGGCTTT
ATCATTTAAACCTGTGCTACCCTTGACATAATACTTGTATGACCTAACTACTTTAACATTCATAGGACATT
ATTTGTCAAGATTGAGGTAATTAATCGTTACGGGAGAATAATTGTGAAGTCATAATTGACACAGGGGGTAA
TGAATAATCAGATATACATTTAAGAAAGACATTTCATGGACTTTGACCATTAAGTGTAGTGAACAACTATT
TAATTAATGTAAGGAGCATTGAACTACATGGTACATATATCGATATTTTGGCATAGACTTATTTATATGGG
GACCCCTGGGTTGAAAAGAAAAAAAAAAACAATATACATAAATTTTTGGAAAAACCCCCCCTCCCCCCAAA
TATACACGGTATCCTCGAAAAACCCCTAAAACGAGGGCCGATTATATATTTTTCTTTTAATGACATGTAGA
ATTAATTATCATCATTATATATAGTGTCACACGTATGGCAT

>Haplotype: LJ3 GB#MN784021 (Sample: RP003)
AAAGATCAACACATGACCAACCTGCATAAACGCATGTATTTGCGTGAACGCAAATGCATTTGTGAAAACAC
AAATGCAATTGCAAAATTATGCAGGTTGGTCAGTCTGACATATAAACTATTTAGCCCCATACTATGTATAT
CCTCATTAAATATACATACTATGCTTAATCCTCATTAATAGACTGTCCCCATATACATATTATGCTTAATC
CACATCATACATATTATGCTTAACCCCCACATTACATATTATGCTTAATCCACATCATACATATTATGCTT
AATCCACATCATACATACGTCTATCTCACTCCCCATACACACATATATGTGTGCTTAATCCCCATTAATCC
ATACTCCCCTACTCCATTACATGCTATGCTTAATCCTCATTAATAGACTGTCCCCATAGTACATACTATGC
TTAATCCTCATTAACATACTATGCTTAATCCCCATTAATAGACATTCCCCTATTTCATTACATCTTACGTA
TCATTAATCCTCATTACTAGAAAATCCTCTACTCCATTACAATCATTCACTCACCCTCATTAATTTACTTC

TAATACTGCTTATGCGGGCTGGCGAGAAATAACCAATACTCTCTTATAAGACCCCTTTGCACGGTTTGTGG
TACTGTAATCGAATAATCCCCATCAATTGTTCAAATCCCTCATTTGGCACCCTTAAAAAGGTATACGTTTT
TTAATCGGGTCAGAACTCCTTTTCCTTTAGTTCCCTTAATTGTCATTTTGACTTTTAATTGTTTCAAGATT
TTTCAGTCCTCCCAGTTTTTTTTTTAGGGGAATGAAGCAATTGTTAACCCGCCAGGGTTGGATACGGGTTTA
TCATTTAAACCTGTGCTACCCTTGACATAATATTTGTATGACCTAACTACTTTAACATTCATAGGACATTA
TTTGTCAAGATTGAGCTAATTAATTGTTACGGGAGAATAATTGGGAAGTCATAATTGACACAGTGGGTAAT
GAATAATCAGATATACATTTAAGAAAGACATTTCATGGACTTTGACCATTAAGTGTAGTGAACAACTATTT
AATTAATGTAAGGAGCATTGAACTACATGGTACATATATCGATATTTTGGCATAGACTTATTTATATGGGT
ACCCCTGGGTTGAAAAGAAAAAAAAAAAACAATATACATAAATTTTTGGAAAAACCCCCCCTCCCCCCAAAT
ATACACGGTATCCTCGAAAAACCCCTAAAACGAGGGCCGATTATATATTTTTCTTTTAATGACATGTAGAA
TTAATTTATCATCATTATATATAGTGTCACACGTATGGCAT

>Haplotype: LJ4 GB#MN784022 (Sample: RP116)
AAAGATCAACACATGACCAACCTGCATAAACGCATGTATTTGCGTGAACGCAAATGCATTTGTGAAAACAC
AAATGCAATTGCAAAATTATGCAGGTTGGTCAGTCTGACATATAAACTATTTAGCCCCATACTATGTATAT
CCTCATTAAATATACATACTATGCTTAATCCTCATTAATAGACTGTCCCCATATACATATTATGCTTAATC
CACATCATACATATTATGCTTAACCCCCACATTACATATTATGCTTAATCCACATCATACATATTATGCTT
AATCCACATCATACATACGTCTATCTCACTCCCCATACACACATATATGTGTGCTTAATCCCCATTAATCC
ATACTCCCCTACTCCATTACATGCTATGCTTAATCCTCATTAATAGACTGTCCCCATAGTACATACTATGC
TTAATCCTCATTAACATACTATGCTTAATCCCCATTAATAGACATTCCCCTATTTCATTACATCTTACGTA
TCATTAATCCTCATTACTAGAAAATCCTCTACTCCATTACAATCATTCACTCACCCTCATTAATTTACTTC
TAATACTGCTTATGCGGGCTGGCGAGAAATAACCAATACTCTCTTATAAGACCCCTTTGCACGGTTTGTGG
TACTGTAATCGAATAATCCCCATCAATTGTTCAAATCCCTCATTTGGCACCCTTAAAAAGGTATACGTCTC
TTAATCGCGTCAGAACTCCTTTTCCTCTAGTTCCCTTAATTGTCATTTCGACTCTTAATCGTCTCAAGATT
TCTCAGTCCTCCCAGTCTTTTTTTTAGGGGAATGAAGCAATCGTTAACCCGCCAGGGTTGGATACGGCTCTA
TCATTTAAACCTGTGCTACCCTCGACATAATATTTGTATGACCTAACTACTTTAACATTCATAGGACATTA
TTTGTCAAGATTGAGCTAATTAATCGTTACGGGAGAATAATTGTGAAGTCATAATTGACACAGTGGGTAAT
GAATAATCAGATATACATTTAAGAAAGACATTTCATGGACTTCGACCATTAAGTGTAGTGAACAACTATTT
AATTAATGTAAGGAGCATTGAACTACATGGTACATATATCGATATTTTGGCATAGACTTATTTATATGGGT
ACCCCTGGGTCGAAAAGAAAAAAAAAAAACAATATACATAAATTTTTGGAAAAACCCCCCCTCCCCCCAAAT
ATACACGGTATCCTCGAAAAACCCCTAAAACGAGGGCCGATTATATATTTTTCTTTTAATGACATGTAGAA
TTAATTTATCATCATTATATATAGTGTCACAATGGCAT

>Haplotype: LJ5 GB#MN784023 (Sample: RP216)
AAAGATCAACACATGACCAACCTGCATAAACGCATGTACTTGCGTGAACGCAAATGCATTTGTGAAAACAC
AAATGCAATTGCAAAATTATGCAGGTTGGTCAGTCTGACATATAAACTATTTAGCCCCATACTATGTATAT
CCTCATTAAATATACATACTATGCTTAATCCTCATTAATAGACTGTCCCCATATACATATTATGCTTAATC
CACATCATACATATTATGCTTAACCCCCACATTACATATTATGCTTAATCCACATCATACATATTATGCTT
AATCCACATCATACATACGTCTATCTCACTCCCCATACACACATATATGTGTGCTTAATCCCCATTAATCC
ATACTCCCCTACTCCATTACATGCTATGCTTAATCCTCATTAATAGACTGTCCCCATAGTACATACTATGC
TTAATCCTCATTAACATACTATGCTTAATCCCCATTAATAGACATTCCCCTATTTCATTACATCTTACGTA
TCATTAATCCTCATTACTAGAAAATCCTCTACTCCATTACAATCATTCACTCACCCTCATTAATTTACTTC
TAATACTGCTTATGCGGGCTGGCGAGAAATAACCAATACTCTCTTATAAGACCCCTTTGCACGGTTTGTGG
TACTGTAATCGAATAATCCCCATCAATTGTTCAAATCCCTCATTTGGCACCCTTAAAAAGGTATACGTCTT
TTAATCGCGTCAGAACTCCTTTTCCTCTAGTTCCCTTAATTGTCATTTCGACTTTTAATCGTCTCAAGATT
TTTCAGTCCTCCCAGTTTTTTTTTTAGGGGAATGAAGCAATCGTTAACCCGCCAGGGTTGGATACGGCTCTA
TCATTTAAACCTGTGCTACCCTCGACATAATATTTGTATGACCTAACTACTTTAACATTCATAGGACATTA
TTTGTCAAGATTGAGCTAATTAATCGTTACGGGAGAATAATTGGGAAGTCATAATTGACACAGTGGGTAAT
GAATAATCAGATATACATTTAAGAAAGACATTTCATGGACTTCGACCATTAAGTGTAGTGAACAACTATTT
AATTAATGTAAGGGGCATTGAACTACATGGTACATATATCGATATTTTGGCATAGACTTATTTATATGGGT
ACCCCTGGGTCGAAAAGAAAAAAAAAAAACAATATACATAAATTTTTGGAAAAACCCCCCCTCCCCCCAAAT
ATACACGGTATCCTCGAAAAACCCCTAAAACGAGGGCCGATTATATATTTTTCTTTTAATGACATGTAGAA
TTAATTATCATCATTATATATAGTGTCACACTATGGCAT

>Haplotype: LJ6 GB#MN784024 (Sample: RP316)
AAAGATCAACACATGACCAACCTGCATAAACGCATGTATTTGCGTGAACGCAAATGCATTTGTGAAAACAC
AAATGCAATTGCAAAATTATGCAGGTTGGTCAGTCTGACATATAAACTATTTAGCCCCATACTATGTATAT
CCTCATTAAATATACATACTATGCTTAATCCTCATTAATAGACTGTCCCCATATACATATTATGCTTAATC
CACATCATACATATTATGCTTAACCCCCACATTACATATTATGCTTAATCCACATCATACATATTATGCTT

AATCCACATCATACATACGTCTATCTCACTCCCCATACACACATATATGTGTGCTTAATCCCCATTAATCC
ATACTCCCCTACTCCATTACATGCTATGCTTAATCCTCATTAATAGACTGTCCCCATAGTACATACTATGC
TTAATCCTCATTAACATACTATGCTTAATCCCCATTAATAGACATTCCCCTATTTCATTACATCTTACGTA
TCATTAATCCTCATTACTAGAAAATCCTCTACTCCATTACAATCATTCACTCACCCTCATTAATTTACTTC
TAATACTGCTTATGCGGGCTGGCGAGAAATAACCAATACTCTCTTATAAGACCCCTTTGCACGGTTTGTGG
TACTGTAATCGAATAATCCCCATCAATTGTTCAAATCCCTCATTTGGCACCCTTAAAAAGGTATACGTCTC
TTAATCGCGTCAGAACTCCTTTTCCTCTAGTTCCCTTAATTGTCATTTCGACTCTTAATCGTCTCAAGATT
TCTCAGTCCTCCCAGTCTTTTTTTAGGGGAATGAAGCAATCGTTAACCCGCCAGGGTTGGATACGGCTCTA
TCATTTAAACCTGTGCTACCCTCGACATAATATTTGTATGACCTAACTACTTTAACATTCATAGGACATTA
TTTGTCAAGATTGAGCTAATTAATCGTTACGGGAGAATAATTGTGAAGTCATAATTGACACAGTGGGTAAT
GAATAATCAGATATACATTTAAGAAAGACATTTCATGGACTTCGACCATTAAGTGTAGTGAACAACTATTT
AATTAATGTAAGGGGCATTGAACTACATGGTACATATATCGATATTTTGGCATAGACTTATTTATATGGGT
ACCCCTGGGTCGAAAAGAAAAAAAAAAAACAATATACATAAATTTTTGGAAAAACCCCCCCTCCCCCCAAAT
ATACACGGTATCCTCGAAAAACCCCTAAAACGAGGGCCGATTATATATTTTTCTTTTAATGACATGTAGAA
TTAATTTATCATCATTATATATAGTATCACACTATGGCAT

>Haplotype: LJ7 GB#MN784025 (Sample: RP416)
AAAGATCAACACATGACCAACCTGCATAAACGCATGTACTTGCGTGAACGCAAATGCATTTGTGAAAACAC
AAATGCAATTGCAAAATTATGCAGGTTGGTCAGTCTGACATATAAACTATTTAGCCCCATACTATGTATAT
CCTCATTAAATATACATACTATGCTTAATCCTCATTAATAGACTGTCCCCATATACATATTATGCTTAATC
CACATCATACATATTATGCTTAACCCCCACATTACATATTATGCTTAATCCACATCATACATATTATGCTT
AATCCACATCATACATACGTCTATCTCACTCCCCATACACACATATATGTGTGCTTAATCCCCATTAATCC
ATACTCCCCTACTCCATTACATGCTATGCTTAATCCTCATTAATAGACTGTCCCCATAGTACATACTATGC
TTAATCCTCATTAACATACTATGCTTAATCCCCATTAATAGACATTCCCCTATTTCATTACATCTTACGTA
TCATTAATCCTCATTACTAGAAAATCCTCTACTCCATTACAATCATTCACTCACCCTCATTAATTTACTTC
TAATACTGCTTATGCGGGCTGGCGAGAAATAACCAATACTCTCTTATAAGACCCCTTTGCACGGTTTGTGG
TACTGTAATCGAATAATCCCCATCAATTGTTCAAATCCCTCATTTGGCACCCTTAAAAAGGTATACGTCTC
TTAATCGCGTCAGAACTCCTTTTCCTCTAGTTCCCTTAATTGTCATTTCGACTTCTTAATCGTCTCAAGAT
TTCTCAGTCCTCCCAGTCTTTTTTTAGGGGAATGAAGCAATCGTTAACCCGCCAGGGTTGGATACGGCTTT
ATCATTTAAACCTGTGCTACCCTCGACATAATATTTGTATGACCTAACTACTTTAACATTCATAGGACATT
ATTTGTCAAGATTGAGCTAATTAATCGTTACGGGAGAATAATTGTGAAGTCATAATTGACACAGTGGGTAA
TGAATAATCAGATATACATTTAAGAAAGACATTTCATGGACTTCGACCATTAAGTGTAGTGAACAACTATT
TAATTAATGTAAGGGGCATTGAACTACATGGTACATATATCGATATTTTGGCATAGACTTATTTATATGGG
TACCCCTGGGTCGAAAAGAAAAAAAAAAACAATATACATAAATTTTTGGAAAAACCCCCCCTCCCCCCAAA
TATACACGGTATCCTCGAAAAACCCCTAAAACGAGGGCCGATTATATATTTTTCTTTTAATGACATGTAGA
ATTAATTATCATCATTATATATAGTGTCACACTATGGCAT

>Haplotype: LJ8 GB#MN784026 (Sample: RP516)
AAAGATCAACACATGACCAACCTGCATAAACGCATGTATTTGCGTGAACGCAAATGCATTTGTGAAAACAC
AAATGCAATTGCAAAATTATGCAGGTTGGTCAGTCTGACATATAAACTATTTAGCCCCATACTATGTATAT
CCTCATTAAATATACATACTATGCTTAATCCTCATTAATAGACTGTCCCCATATACATATTATGCTTAATC
CACATCATACATATTATGCTTAACCCCCACATTACATATTATGCTTAATCCACATCATACATATTATGCTT
AATCCACATCATACATACGTCTATCTCACTCCCCATACACACATATATGTGTGCTTAATCCCCATTAATCC
ATACTCCCCTACTCCATTACATGCTATGCTTAATCCTCATTAATAGACTGTCCCCATAGTACATACTATGC
TTAATCCTCATTAACATACTATGCTTAATCCCCATTAATAGACATTCCCCTATTTCATTACATCTTACGTA
TCATTAATCCTCATTACTAGAAAATCCTCTACTCCATTACAATCATTCACTCACCCTCATTAATTTACTTC
TAATACTGCTTATGCGGGCTGGCGAGAAATAACCAATACTCTCTTATAAGACCCCTTTGCACGGTTTGTGG
TACTGTAATCGAATAATCCCCATCAATTGTTCAAATCCCTCATTTGGCACCCTTAAAAAGGTATACGTCTC
TTAATCGCGTCAGAACTCCTTTTCCTCTAGTTCCCTTAATTGTCATTTCGACTCTTAATCGTCTCAAGATT
TCTCAGTCCTCCCAGTCTTTTTTTAGGGGAATGAAGCAATCGTTAACCCGCCAGGGTTGGATACGGCTTTA
TCATTTAAACCTGTGCTACCCTCGACATAATATTTGTATGACCTAACTACTTTAACATTCATAGGACATTA
TTTGTCAAGATTGAGCTAATTAATCGTTACGGGAGAATAATTGGGAAGTCATAATTGACACAGTGGGTAAT
GAATAATCAGATATACATTTAAGAAAGACATTTCATGGACTTCGACCATTAAGTGTAGTGAACAACTATTT
AATTAATGTAAGGGGCATTGAACTACATGGTACATATATCGATATTTTGGCATAGACTTATTTATATGGGT
ACCCCTGGGTCGAAAAGAAAAAAAAAAAACAATATACATAAATTTTTGGAAAAACCCCCCCTCCCCCCAAAT
ATACACGGTATCCTCGAAAAACCCCTAAAACGAGGGCCGATTATATATTTTTCTTTTAATGACATGTAGAA
TTAATTTATCATCATTATATATAGTGTCACACTATGGCAT

>Haplotype: LJ9 GB#MN784027 (Sample: RP616)
AAAGATCAACACATGACCAACCTGCATAAACGCATGTATTTGCGTGAACGCAAATGCATTTGTGAAAACAC
AAATGCAATTGCAAAATTATGCAGGTTGGTCAGTCTGACATATAAACTATTTAGCCCCATACTATGTATAT
CCTCATTAAATATACATACTATGCTTAATCCTCATTAATAGACTGTCCCCATATACATATTATGCTTAATC
CACATCATACATATTATGCTTAACCCCCACATTACATATTATGCTTAATCCACATCATACATATTATGCTT
AATCCACATCATACATACGTCTATCTCACTCCCCATACACACATATATGTGTGCTTAATCCCCATTAATCC
ATACTCCCCTACTCCATTACATGCTATGCTTAATCCTCATTAATAGACTGTCCCCATAGTACATACTATGC
TTAATCCTCATTAACATACTATGCTTAATCCCCATTAATAGACATTCCCCTATTTCATTACATCTTACGTA
TCATTAATCCTCATTACTAGAAAATCCTCTACTCCATTACAATCATTCACTCACCCTCATTAATTTACTTC
TAATACTGCTTATGCGGGCTGGCGAGAAATAACCAATACTCTTTTATAAGACCCCTTTGCACGGTTTGTGG
TACTGTAATCGAATAATCCCCATCAATTGTTCAAATCCCTCATTTGGCACCCTTAAAAAGGTATACGTCTT
TTAATCGCGTCAGAACTCCTTTTCCTTTAGTTCCCTTAATTGTCATTTTGACTTTTAATTGTTTCAAGATT
TTTCAGTCCTCCCAGTTTTTTTTTTAGGGGAATGAAGCAATTGTTAACCCGCCAGGGTTGGATACGGCTTTA
TCATTTAAACCTGTGCTACCCTCGACATAATATTTGTATGACCTAACTACTTTAACATTCATAGGACATTA
TTTGTCAAGATTGAGGTAATTAATCGTTACGGGAGAATAATTGGGAAGTCATAATTGACACAGTGGGTAAT
GAATAATCAGATATACATTTAAGAAAGACATTTCATGGACTTTGACCATTAAGTGTAGTGAACAACTATTT
AATTAATGTAAGGAGCATTGAACTACATGGTACATATATCGATATTTTGGCATAGACTTATTTATATGGGT
ACCCCTGGGTTGAAAAGAAAAAAAAAAACAATATACATAAATTTTTGGAAAAACCCCCCCTCCCCCCAAAT
ATACACGGTATCCTCGAAAAACCCCTAAAACGAGGGCCGATTATATATTTTTCTTTTAATGACATGTAGAA
TTAATTTATCATCATTATATATAGTGTCACACTATGGCAT

>Haplotype: LJ10 GB#MN784028 (Sample: RP716)
AAAGATCAACACATGACCAACCTGCATAAACGCATGTATTTGCGTGAACGCAAATGCATTTGTGAAAACAC
AAATGCAATTGCAAAATTATGCAGGTTGGTCAGTCTGACATATAAACTATTTAGCCCCATACTATGTATAT
CCTCATTAAATATACATACTATGCTTAATCCTCATTAATAGACTGTCCCCATATACATATTATGCTTAATC
CACATCATACATATTATGCTTAACCCCCACATTACATATTATGCTTAATCCACATCATACATATTATGCTT
AATCCACATCATACATACGTCTATCTCACTCCCCATACACACATATATGTGTGCTTAATCCCCATTAATCC
ATACTCCCCTACTCCATTACATGCTATGCTTAATCCTCATTAATAGACTGTCCCCATAGTACATACTATGC
TTAATCCTCATTAACATACTATGCTTAATCCCCATTAATAGACATTACCCTATTTCATTACATCTTACGTA
TCATTAATCCTCATTACTAGAAAATCCTCTACTCCATTACAATCATTCACTCACCCTCATTAATTTACTTC
TAATACTGCTTATGCGGGCTGGCGAGAAATAACCAATACTCTCTTATAAGACCCCTTTGCACGGTTTGTGG
TACTGTAATCGAATAATCCCCATCAATTGTTCAAATCCCTCATTTGGCACCCTTAAAAAGGTATACGTCTT
TTAATCGGGTCAGAACTCCTTTTCCTTTAGTTCCCTTAATTGTCATTTTGACTTTTAATTGTTTCAAGATT
TTTCAGTCCTCCCAGTTTTTTTTTTAGGGGAATGAAGCAATTGTTAACCCGCCAGGGTTGGATACGGCTTTA
TCATTTAAACCTGTGCTACCCTTGACATAATATTTGTATGACCTAAATACTTTAACATTCATAGGACATTT
TTTGTCAAGATTGAGGTAATTAATTGTTACGGGGGAATAATTGGGAAGTCATAATTGACCCAGGGGGTAAT
GAATAATCAGATATACATTTAAGAAAGACATTTCATGGACTTTGACCATTAAGTGTAGTGAACAACTATTT
AATTAATGTAAGGAGCATTGAACTACATGGTACATATATTGATATTTTGGCATAGACTTATTTATATGGGT
ACCCCTGGGTTGAAAAGAAAAAAAAAAACAATATACATAAATTTTTGGAAAAACCCCCCCTCCCCCCAAAT
ATACACGGTATCCTCGAAAAACCCCTAAAACGAGGGCCGATTATATATTTTTCTTTTAATGACATGTAGAA
TTAATTTATCATCATTATATATAGTATCACACGTATGGCAT

>Haplotype: GN1 GB#MN784006 (Sample: LCRP3)
AAAGATCAACACATGACCAACCTGCATAAACGCATGTATTTGCGTGAACGCAAATGCATTTGTGAAAACAC
AAATGCAATTGCAAAATTATGCAGGTTGGTCAGTCTGACATATAAACTATTTAGCCCCATACTATGTATAT
CCTCATTAAATATACATACTATGCTTAATCCTCATTAATAGACTGTCCCCATATACATATTATGCTTAATC
CACATCATACATATTATGCTTAACCCCCACATTACATATTATGCTTAATCCACATCATACATATTATGCTT
AATCCACATCATACATACGTCTATCTCACTCCCCATACACACATATATGTGTGCTTAATCCCCATTAATCC
ATACTCCCCTACTCCATTACATGCTATGCTTAATCCTCATTAATAGACTGTCCCCATAGTACATACTATGC
TTAATCCTCATTAACATACTATGCTTAATCCCCATTAATAGACATTCCCCTATTTCATTACATCTTACGTA
TCATTAATCCTCATTACTAGAAAATCCTCTACTCCATTACAATCATTCACTCACCCTCATTAATTTACTTC
TAATACTGCTTATGCGGGCTGGCGAGAAATAACCAATACTCTCTTATAAGACCCCTTTGCACGGTTTGTGG
TACTGTAATCGAATAATCCCCATCAATTGTTCAAATCCCTCATTTGGCACCCTTAAAAAGGTATACGTCTT
TTAATCGCGTCAGAACTCCTTTTCCTTTAGTTCCCTTAATTGTCATTTCGACTTTTAATTGTTTCAAGATT
TTTCAGTCCTCCCAGTTTTTTTTTTAGGGGAAAGAAGCAATCGTTAACCCCCCAGGGTTGGATACGGCTTTA
TCATTTAAACCTGTGCTACCCTCGACATAATATTTGTATGACCTAACTACTTTAACATTCATAGGACATTA
TTTGTCAAGATTGAGCTAATTAATTGTTACGGGAGAATAATTGGGAAGTCATAATTGACCCAGTGGGTAAT
GAATAATCAGATATACATTTAAGAAAGACATTTCATGGACTTTGACCATTAAGTGTAGTGAACAACTATTT
AATTAATGTAAGGAGCATTGAACTACATGGTACATATATCGATATTTTGGCATAGACTTATTTATATGGGT

ACCCCTGGGTTGAAAAGAAAAAAAAAAACAATATACATAAATTTTTGGAAAAACCCCCCCTCCCCCCAAAT
ATACACGGTATCCTCGAAAAACCCCTAAAACGAGGGCCGATTATATATTTTTCTTTTAATGACATGTAGAA
TTAATTTATCATCATTATATATAGTGTCACACTATGGCAT

>Haplotype: GN2 GB#MN784007 (Sample: LCRP4)
AAAGATCAACACATGACCAACCTGCATAAACGCATGTATTTGCGTGAACGCAAATGCATTTGTGAAAACAC
AAATGCAATTGCAAAATTATGCAGGTTGGTCAGTCTGACATATAAACTATTTAGCCCCATACTATGTATAT
CCTCATTAAATATACATACTATGCTTAATCCTCATTAATAGACTGTCCCCATATACATATTATGCTTAATC
CACATCATACATATTATGCTTAACCCCCACATTACATATTATGCTTAATCCACATCATACATATTATGCTT
AATCCACATCATACATACGTCTATCTCACTCCCCATACACACATATATGTGTGCTTAATCCCCATTAATCC
ATACTCCCCTACTCCATTACATGCTATGCTTAATCCTCATTAATAGACTGTCCCCATAGTACATACTATGC
TTAATCCTCATTAACATACTATGCTTAATCCCCATTAATAGACATTCCCCTATTTCATTACATCTTACGTA
TCATTAATCCTCATTACTAGAAAATCCTCTACTCCATTACAATCATTCACTCACCCTCATTAATTTACTTC
TAATACTGCTTATGCGGGCTGGCGAGAAATAACCAATACTCTCTTATAAGACCCCTTTGCACGGTTTGTGG
TACTGTAATCGAATAATCCCCATCAATTGTTCAAATCCCTCATTTGGCACCCTTAAAAAGGTATACGTCTC
TTAATCGCGTCAGAACTCCTTTTCCTCTAGTTCCCTTAATTGTCATTTCGACTTTTAATCGTCTCAAGATT
TTTCAGTCCTCCCAGTTTTTTTTTAGGGGAATGAAGCAATCGTTAACCCGCCAGGGTTGGATACGGCTTTA
TCATTTAAACCTGTGCTACCCTCGACATAATATTTGTATGACCTAACTACTTTAACATTCATAGGACATTA
TTTGTCAAGATTGAGCTAATTAATCGTTACGGGAGAATAATTGTGAAGTCATAATTGACACAGTGGGTAAT
GAATAATCAGATATACATTTAAGAAAGACATTTCATGGACTTCGACCATTAAGTGTAGTGAACAACTATTT
AATTAATGTAAGGGGCATTGAACTACATGGTACATATATCGATATTTTGGCATAGACTTATTTATATGGGT
ACCCCTGGGTCGAAAAGAAAAAAAAAAAACAATATACATAAATTTTTGGAAAAACCCCCCCTCCCCCCAAAT
ATACACGGTATCCTCGAAAAACCCCTAAAACGAGGGCCGATTATATATTTTTCTTTTAATGACATGTAGAA
TTAATTTATCATCATTATATATAGTGTCACACTATGGCAT

>Haplotype: GN3 GB#MN784008 (Sample: LCRP5)
AAAGATCAACACATGACCAACCTGCATAAACGCATGTATTTGCGTGAACGCAAATGCATTTGTGAAAACAC
AAATGCAATTGCAAAATTATGCAGGTTGGTCAGTCTGACATATAAACTATTTAGCCCCATACTATGTATAT
CCTCATTAAATATACATACTATGCTTAATCCTCATTAATAGACTGTCCCCATATACATATTATGCTTAATC
CACATCATACATATTATGCTTAACCCCCACATTACATATTATGCTTAATCCACATCATACATATTATGCTT
AATCCACATCATACATACGTCTATCTCACTCCCCATACACACATATATGTGTGCTTAATCCCCATTAATCC
ATACTCCCCTACTCCATTACATGCTATGCTTAATCCTCATTAATAGACTGTCCCCATAGTACATACTATGC
TTAATCCTCATTAACATACTATGCTTAATCCCCATTAATAGACATTCCCCTATTTCATTACATCTTACGTA
TCATTAATCCTCATTACTAGAAAATCCTCTACTCCATTACAATCATTCACTCACCCTCATTAATTTACTTT
TAATACTGCTTATGCGGGGTGGGGAGAAATAACCAATACTTTTTTATAAGACCCCTTTGCACGGTTTGTGG
TACTGTAATTGAATAATCCCCATCAATTGTTCAAATCCCTCATTTGGCACCCTTAAAAAGGTATACGTTTT
TTAATTGCGTCAGAACTCCTTTTCCTTTAGTTCCCTTAATTGTCATTTCGACTTTTAATCGTTTCAAGATT
TTTCAGTCCTCCCAGTTTTTTTTTAGGGGAATGAAGCAATTGTTAACCCCCCAGGGTTGGATACGGCTTTA
TCATTTAAACCTGTGGTACCCTCGACATAATATTTGTATGACCTAAATACTTTAACATTCATAGGACATTA
TTTGTCAAGATTGAGGTAATTAATTGTTACGGGAGAATAATTGGGAAGTCATAATTGACCCAGTGGGTAAA
GAATAATCAGATATACATTTAAGAAAGACATTTCATGGACTTCGACCATTAAGTGTAGTGAACAACTATTT
AATTAATGTAAGGAGCATTGAACTACATGGTACATATATCGATATTTTGGCATAGACTTATTTATATGGGT
ACCCCTGGGTTGAAAAGAAAAAAAAAAAACAATATACATAAATTTTTGGAAAAACCCCCCCTCCCCCCAAAT
ATACACGGTATCCTCGAAAAACCCCTAAAACGAGGGCCGATTATATATTTTTCTTTTAATGACATGTAGAA
TTAATTTATCATCATTATATATAGTGTTCATTATATGGCAT

>Haplotype: GN4 GB#MN784009 (Sample: LCRP11)
AAAGATCAACACATGACCAACCTGCATAAACGCATGTACTTGCGTGAACGCAAATGCATTTGTGAAAACAC
AAATGCAATTGCAAAATTATGCAGGTTGGTCAGTCTGACATATAAACTATTTAGCCCCATACTATGTATAT
CCTCATTAAATATACATACTATGCTTAATCCTCATTAATAGACTGTCCCCATATACATATTATGCTTAATC
CACATCATACATATTATGCTTAACCCCCACATTACATATTATGCTTAATCCACATCATACATATTATGCTT
AATCCACATCATACATACGTCTATCTCACTCCCCATACACACATATATGTGTGCTTAATCCCCATTAATCC
ATACTCCCCTACTCCATTACATGCTATGCTTAATCCTCATTAATAGACTGTCCCCATAGTACATACTATGC
TTAATCCTCATTAACATACTATGCTTAATCCCCATTAATAGACATTCCCCTATTTCATTACATCTTACGTA
TCATTAATCCTCATTACTAGAAAATCCTCTACTCCATTACAATCATTCACTCACCCTCATTAATTTACTTT
TAATACTGCTTATGCGGGCTGGGGAGAAATAACCAATACTTTTTTATAAGACCCCTTTGCACGGTTTGGGG
TACTGTAATCGAATAATCCCCATCAATTGTTCAAATCCCTCATTTGGCACCCTTAAAAAGGTATACGTCTT
TTAATTGGGTCAGAACTCCTTTTCCTTTAGTTCCCTTAATTGTCATTTCGACTTTTAATTGTTTCAAGATT
TTTCAGTCCTCCCAGTTTTTTTTTAGGGGAATGAAGCAATTGTTAACCCGCCAGGGTTGGATACGGCTTTA

TCATTTAAACCTGTGCTACCCTCGACATAATATTTGTATGACCTAACTACTTTAACATTCATAGGGCATTA
TTTGTCAAGATTGAGCTAATTAATTGTTACGGGAGAATAATTGGGAAGTCATAATTGACACAGTGGGTAAT
GAATAATCAGATATACATTTAAGAAAGACATTTCATGGACTTCGACCATTAAGTGTAGTGAACAACTATTT
AATTAATGTAAGGAGCATTGAACTCCATGGTACATATATTGATATTTTGGCAAAGACTTATTTATATGGGG
ACCCCTGGGTTGAAAAGAAAAAAAAAAACAATATACATAAATTTTTGGAAAAACCCCCCCTCCCCCCAAAT
ATACACGGTATCCTCGAAAAACCCCTAAAACGAGGGCCGATTATATATTTTTCTTTTAATGACATGTAGAA
TTAATTATCATCATTATATATAGCGTCCCACTATGGCAT

>Haplotype: GN5 GB#MN784010 (Sample: LCRP13)
AAAGATCAACACATGACCAACCTGCATAAACGCATGTACTTGCGTGAACGCAAATGCATTTGTGAAAACAC
AAATGCAATTGCAAAATTATGCAGGTTGGTCAGTCTGACATATAAACTATTTAGCCCCATACTATGTATAT
CCTCATTAAATATACATACTATGCTTAATCCTCATTAATAGACTGTCCCCATATACATATTATGCTTAATC
CACATCATACATATTATGCTTAACCCCCACATTACATATTATGCTTAATCCACATCATACATATTATGCTT
AATCCACATCATACATACGTCTATCTCACTCCCCATACACACATATATGTGTGCTTAATCCCCATTAATCC
ATACTCCCCTACTCCATTACATGCTATGCTTAATCCTCATTAATAGACTGTCCCCATAGTACATACTATGC
TTAATCCTCATTAACATACTATGCTTAATCCCCATTAATAGACATTCCCCTATTTCATTACATCTTACGTA
TCATTAATCCTCATTACTAGAAAATCCTCTACTCCATTACAATCATTCACTCACCCTCATTAATTTACTTT
TAATACTGCTTATGGGGGGTGGGGAGAAATAACCAATACTTTTTTATAAGACCCCTTTGCACGGTTTGTGG
TACTGTAATCGAATAATCCCCATCAATTGTTCAAATCCCTCATTTGGCACCCTTAAAAAGGTATACGTTTT
TTAATTGGGTCAGAACTCCTTTTCCTTTAGTTCCCTTAATTGTCATTTCGACTTTTAATTGTTTCAAGATT
TTTCAGTCCTCCCAGTTTTTTTTTAGGGGAAAGAAGCAATTGTTAACCCCCCAGGGTTGGATACGGCTTTA
TCATTTAAACCTGGGCTACCCTCGACATAATATTTGTATGGCCTAACTACTTTAACATTCATAGGACATTA
TTTGTCAAGATTGAGCTAATTAATCGTTACGGGAGAATAATTGTGAAGTCATAATTGACCCAGTGGGTAAT
GAATAATCAGATATACATTTAAGAAAGACATTTCATGGACTTTGGCCATTAAGTGTAGTGAACAACTATTT
AATTAATGTAAGGAGCATTGAACTACATGGGACATATATTGATATTTTGGCATAGAATTATTTATATGGGG
ACCCCTGGGTTGAAAAGAAAAAAAAAAACAATATACATAAATTTTTGGAAAAACCCCCCCTCCCCCCAAAT
ATACACGGTATCCTCGAAAAACCCCTAAAACGAGGGCCGATTATATATTTTTCTTTTAATGACATGTAGAA
TTAATTATCATCATTATATATAGTGTCCCAATATGGCAT

>Haplotype: GN6 GB#MN784011 (Sample: LCRP22)
AAAGATCAACACATGACCAACCTGCATAAACGCATGTACTTGCGTGAACGCAAATGCATTTGTGAAAACAC
AAATGCAATTGCAAAATTATGCAGGTTGGTCAGTCTGACATATAAACTATTTAGCCCCATACTATGTATAT
CCTCATTAAATATACATACTATGCTTAATCCTCATTAATAGACTGTCCCCATATACATATTATGCTTAATC
CACATCATACATATTATGCTTAACCCCCACATTACATATTATGCTTAATCCACATCATACATATTATGCTT
AATCCACATCATACATACGTCTATCTCACTCCCCATACACACATATATGTGTGCTTAATCCCCATTAATCC
ATACTCCCCTACTCCATTACATGCTATGCTTAATCCTCATTAATAGACTGTCCCCATAGTACATACTATGC
TTAATCCTCATTAACATACTATGCTTAATCCCCATTAATAGACATTCCCCTATTTCATTACATCTTACGTA
TCATTAATCCTCATTACTAGAAAATCCTCTACTCCATTACAATCATTCACTCACCCTCATTAATTTACTTC
TAATACTGCTTATGGGGGGTGGGGAGAAATAACCAATACTTTTTTATAAGACCCCTTTGCACGGTTTGTGG
TACTGTAATCGAATAATCCCCATCAATTGTTCAAATCCCTCATTTGGCACCCTTAAAAAGGTATACGTTTT
TTAATTGGGTCAGAACTCCTTTTCCTTTAGTTCCCTTAATTGTCATTTCGACTTTTAATCGTTTCAAGATT
TTTCAGTCCTCCCAGTTTTTTTTTTAGGGGAATGAAGCAATTGTTAACCCCCCAGGGTTGGATACGGCTTTA
TCATTTAAACCTGTGCTACCCTTGACATAATATTTGTATGACCTAACTACTTTAACATTCATAGGACATTA
TTTGTCAAGATTGAGGTAATTAATTGTTACGGGAGAATAATTGGGAAGTCATAATTGGCCCAGTGGGTAAA
GAATAATCAGATATACATTTAAGAAAGGCATTTCATGGACTTTGACCCTTAAGTGTAGTGAACAACTATTT
AATTAATGTAAGGAGCATTGAACTACATGGTACATATATCGATATTTTGGCATAGACTTATTTATATGGGG
ACCCCTGGGTGGAAAAGAAAAAAAAAAACAATATACATAAATTTTTGGAAAAACCCCCCCTCCCCCCAAAT
ATACACGGTATCCTCGAAAAACCCCTAAAACGAGGGCCGATTATATATTTTTCTTTTAATGACATGTAGAA
TTAATTATCATCATTATATATAGTGTCTCACTATGGCATGCTAGTGTAGCTT

>Haplotype: GN7 GB#MN784012 (Sample: LCRP48)
AAAGATCAACACATGACCAACCTGCATAAACGCATGTACTTGCGTGAACGCAAATGCATTTGTGAAAACAC
AAATGCAATTGCAAAATTATGCAGGTTGGTCAGTCTGACATATAAACTATTTAGCCCCATACTATGTATAT
CCTCATTAAATATACATACTATGCTTAATCCTCATTAATAGACTGTCCCCATATACATATTATGCTTAATC
CACATCATACATATTATGCTTAACCCCCACATTACATATTATGCTTAATCCACATCATACATATTATGCTT
AATCCACATCATACATACGTCTATCTCACTCCCCATACACACATATATGTGTGCTTAATCCCCATTAATCC
ATACTCCCCTACTCCATTACATGCTATGCTTAATCCTCATTAATAGACTGTCCCCATAGTACATACTATGC
TTAATCCTCATTAACATACTATGCTTAATCCCCATTAATAGACATTCCCCTATTTCATTACATCTTACGTA
TCATTAATCCTCATTACTAGAAAATCCTCTACTCCATTACAATCATTCACTCACCCTCATTAATTTACTTC

TAATACTGCTTATGCGGGCTGGCGAGAAATAACCAATACTCTCTTATAAGACCCCTTTGCACGGTTTGTGG
TACTGTAATCGAATAATCCCCATCAATTGTTCAAATCCCTCATTTGGCACCCTTAAAAAGGTATACGTCTC
TTAATCGCGTCAGAACTCCTTTTCCTCTAGTTCCCTTAATTGTCATTTCGACTTTTAATCGTTTCAAGATT
TTTCAGTCCTCCCAGTTTTTTTTTAGGGGAATGAAGCAATCGTTAACCCGCCAGGGTTGGATACGGCTTTA
TCATTTAAACCTGTGCTACCCTCGACATAATATTTGTATGACCTAACTACTTTAACATTCATAGGACATTA
TTTGTCAAGATTGAGCTAATTAATCGTTACGGGAGAATAATTGTGAAGTCATAATTGACACAGTGGGTAAT
GAATAATCAGATATACATTTAAGAAAGACATTTCATGGACTTCGACCATTAAGTGTAGTGAACAACTATTT
AATTAATGTAAGGGGCATTGAACTACATGGTACATATATCGATATTTTGGCATAGACTTATTTATATGGGT
ACCCCTGGGTCGAAAAGAAAAAAAAAAACAATATACATAAATTTTTGGAAAAACCCCCCCTCCCCCCAAAT
ATACACGGTATCCTCGAAAAACCCCTAAAACGAGGGCCGATTATATATTTTTCTTTTAATGACATGTAGAA
TTAATTATCATCATTATATATAGTGTCACACTATGGCAT

>Haplotype: GN8 GB#MN784013 (Sample: LCRP49)
AAGCTACACTAGCATGCCATAGAGAGACACTATATATAATGATGATAATTAATTCTACATGTCCATTAAAA
GAAAAATATATAATCGGCCCTCGTTTTAGGGGTTTTTCGAGGATACCGTGTATATTTGGGGGGAGGGGGGG
TTTTTCCAAAAATTTATGTATATTGTTTTTTTTTTTCTTTTCAACCCAGGGGTACCCATATAAATAAGTTT
ATGCCAAAATATCGATATTTGTACCATGTAGTTCAATGCTCCTTACATTAATTAAATAGTTGTTCCCTACA
CTTAATGGCCGAAGTCCATGAAATGTCTTTCTTAAATGTATATCTGATTATTCATTACCCCCTGGGTCAAT
TATGACTTCCCAATTATTCTCCCGTAACGATTAATTACCTCAATCTTGACAAATAATGTCCTATGAATGTT
AAAGTAGTTAGGTCATACAAATATTATGTCAAGGGTAGCACAGGTTTAAATGATAAAGCCGTATCCAACCC
TGGGGGGTTAACGATTGCTTCATTCCCCTAAAAAAAGACTGGGAGGACTGAAAAATCTTGAAACGATTAAA
AGTCGAAATGACAATTAAGGGAACTAGAGGAAAAGGAGTTCTGACGCGATTAAAAAACGTATACCTTTTTA
AGGGTGCCAAATGAGGGATTTGAACAATTGATGGGGATTATTCGATTACAGTACCACAAACCGTGCAAAGG
GGTCTTATAAGAGAGTATTGGTTATTTCTCGCCAGCCCGCATAAGCAGTATTAGAAGTAAATTAATGAGGG
TGAGTGAATGATTGTAATGGAGTAGAGGATTTTCTAGTAATGAGGATTAATGATACGTAAGATGTAATGAA
ATAGGGGAATGTCTATTAATGGGGATTAAGCATAGTATGTTAATGAGGATTAAGCATAGTATGTACTATGG
GGACAGTCTATTAATGAGGATTAAGCATAGCATGTAATGGAGTAGGGGAGTATGGATTAATGGGGATTAAG
CACACATATATGTGTGTATGGGGAGTGAGATAGACGTATGTATGATGTGGATTAAGCATAATATGTATGAT
GTGGATTAAGCATAATATGTAATGTGGGGGTTAAGCATAATATGTATGATGTGGATTAAGCATAATATGTA
TATGGGGACAGTCTATTAATGAGGATTAAGCATAGTATGTATATTTAATGAGGATATACATAGTATGGGGC
TAAATAGTTTATATGTCAGACTGACCAACCTGCATAATTTTGCAATTGCATTTGTGTTTTCACAAATGCAT
TTGCGTTCACGCAAGTACATGCGTTTATGCAGGTTGGT

>Haplotype: GN9 GB#MN784014 (Sample: LCRP50)
AAAGATCAACACATGACCAACCTGCATAAACGCATGTACTTGCGTGAACGCAAATGCATTTGTGAAAACAC
AAATGCAATTGCAAAATTATGCAGGTTGGTCAGTCTGACATATAAACTATTTAGCCCCATACTATGTATAT
CCTCATTAAATATACATACTATGCTTAATCCTCATTAATAGACTGTCCCCATATACATATTATGCTTAATC
CACATCATACATATTATGCTTAACCCCCACATTACATATTATGCTTAATCCACATCATACATATTATGCTT
AATCCACATCATACATACGTCTATCTCACTCCCCATACACACATATATGTGTGCTTAATCCCCATTAATCC
ATACTCCCCTACTCCATTACATGCTATGCTTAATCCTCATTAATAGACTGTCCCCATAGTACATACTATGC
TTAATCCTCATTAACATACTATGCTTAATCCCCATTAATAGACATTCCCCTATTTCATTACATCTTACGTA
TCATTAATCCTCATTACTAGAAAATCCTCTACTCCATTACAATCATTCACTCACCCTCATTAATTTACTTC
TAATACTGCTTATGCGGGCTGGCGAGAAATAACCAATACTCTCTTATAAGACCCCTTTGCACGGTTTGTGG
TACTGTAATCGAATAATCCCCATCAATTGTTCAAATCCCTCATTTGGCACCCTTAAAAAGGTATACGTCTC
TTAATCGCGTCAGAACTCCTTTTCCTCTAGTTCCCTTAATTGTCATTTCGACTTTTAATCGTTTCAAGATT
TCTCAGTCCTCCCAGTTTTTTTTTAGGGGAATGAAGCAATCGTTAACCCGCCAGGGTTGGATACGGCTCTA
TCATTTAAACCTGTGCTACCCTCGACATAATATTTGTATGACCTAACTACTTTAACATTCATAGGACATTA
TTTGTCAAGATTGAGCTAATTAATCGTTACGGGAGAATAATTGTGAAGTCATAATTGACACAGTGGGTAAT
GAATAATCAGATATACATTTAAGAAAGACATTTCATGGACTTCGACCATTAAGTGTAGTGAACAACTATTT
AATTAATGTAAGGGGCATTGAACTACATGGTACATATATCGATATTTTGGCATAGACTTATTTATATGGGT
ACCCCTGGGTTGAAAAGAAAAAAAAAAAACAATATACATAAATTTTTGGAAAAACCCCCCCTCCCCCCAAAT
ATACACGGTATCCTCGAAAAACCCCTAAAACGAGGGCCGATTATATATTTTTCTTTTAATGACATGTAGAA
TTAATTATCATCATTATATATAGTGTCACACTATGGCAT

>Haplotype: GN10 GB#MN784015 (Sample: LCRP51)
AAAGATCAACACATGACCAACCTGCATAAACGCATGTACTTGCGTGAACGCAAATGCATTTGTGAAAACAC
AAATGCAATTGCAAAATTATGCAGGTTGGTCAGTCTGACATATAAACTATTTAGCCCCATACTATGTATAT
CCTCATTAAATATACATACTATGCTTAATCCTCATTAATAGACTGTCCCCATATACATATTATGCTTAATC
CACATCATACATATTATGCTTAACCCCCACATTACATATTATGCTTAATCCACATCATACATATTATGCTT

AATCCACATCATACATACGTCTATCTCACTCCCCATACACACATATATGTGTGCTTAATCCCCATTAATCC
ATACTCCCCTACTCCATTACATGCTATGCTTAATCCTCATTAATAGACTGTCCCCATAGTACATACTATGC
TTAATCCTCATTAACATACTATGCTTAATCCCCATTAATAGACATTCCCCTATTTCATTACATCTTACGTA
TCATTAATCCTCATTACTAGAAAATCCTCTACTCCATTACAATCATTCACTCACCCTCATTAATTTACTTT
TAATACTGGTTATGCGGGGTGGGGAGAAATAACCAATACTTTTTTATAAGACCCCTTTGCACGGTTTGTGG
TACTGTAATTGAATAATCCCCATCAATTGTTCAAATCCCTCATTTGGCACCCTTAAAAAGGTATACGTTTT
TTAATTGCGTCAGAACTCCTTTTCCTTTAGTTCCCTTAATTGTCATTTTGACTTTTAATCGTTTCAAGATT
TTTCAGTCCTCCCAGTTTTTTTTTAGGGGAATGAAGCAATCGTTAACCCCCCAGGGTTGGATACGGCTTTA
TCATTTAAACCTGTGCTACCCTTGACATAATATTTGTATGACCTAACTACTTTAACATTCATAGGACATTA
TTTGTCAAGATTGAGCTAATTAATCGTTACGGGAGAATAATTGGGAAGTCATAATTGACACAGGGGGTAAT
GAATAATCAGATATACATTTAAGAAAGGCATTTCATGGGCTTTGGCCATTAAGTGTAGTGAACAACTATTT
AATTAATGTAAGGAGCATTGAACTACATGGTACATATATCGATATTTTGGCATAGACTTATTTATATGGGT
ACCCCTGGGTTGAAAAGAAAAAAAAAAAACAATATACATAAATTTTTGGAAAAACCCCCCCTCCCCCCAAAT
ATACACGGTATCCTCGAAAAACCCCTAAAACGAGGGCCGATTATATATTTTTCTTTTAATGACATGTAGAA
TTAATTATCATCATTATATATAGTGTCACACTATGGCAT

>Haplotype: GN11 GB#MN784016 (Sample: LCRP54)
AAAGATCAACACATGACCAACCTGCATAAACGCATGTATTTGCGTGAACGCAAATGCATTTGTGAAAACAC
AAATGCAATTGCAAAATTATGCAGGTTGGTCAGTCTGACATATAAACTATTTAGCCCCATACTATGTATAT
CCTCATTAAATATACATACTATGCTTAATCCTCATTAATAGACTGTCCCCATATACATATTATGCTTAATC
CACATCATACATATTATGCTTAACCCCCACATTACATATTATGCTTAATCCACATCATACATATTATGCTT
AATCCACATCATACATACGTCTATCTCACTCCCCATACACACATATATGTGTGCTTAATCCCCATTAATCC
ATACTCCCCTACTCCATTACATGCTATGCTTAATCCTCATTAATAGACTGTCCCCATAGTACATACTATGC
TTAATCCTCATTAACATACTATGCTTAATCCCCATTAATAGACATTCCCCTATTTCATTACATCTTACGTA
TCATTAATCCTCATTACTAGAAAATCCTCTACTCCATTACAATCATTCACTCACCCTCATTAATTTACTTC
TAATACTGCTTATGCGGGCTGGCGAGAAATAACCAATACTTTTTTATAAGACCCCTTTGCACGGTTTGTGG
GTACTGTAATCGAATAATCCCCATCAATTGTTCAAATCCCTCATTTGGCACCCTTAAAAAGGTATACGTTT
TTTAATCGCGTCAGAACTCCTTTTCCTCTAGTTCCCTTAATTGTCATTTCGACTTTTAATCGTTTCAAGAT
TTTTCAGTCCTCCCAGTTTTTTTTTAGGGGAATGAAGCAATTGTTAACCCGCCAGGGTTGGATACGGCTTT
ATCATTTAAACCTGTGGTACCCTCGACATAATATTTGTATGACCTAACTACTTTAACATTCATAGGACATT
ATTTGTCAAGATTGAGCTAATTAATTGTTACGGGAGAATAATTGTGAAGTCATAATTGACCCAGGGGGTAA
TGAATAATCAGATATACATTTAAGAAAGACATTTCATGGACTTTGGCCATTAAGTGTAGTGAACAACTATT
TAATTAATGTAAGGAGCATTGAACTACATGGTACATATATTGATATTTTGGCATAGACTTATTTATATGGG
TACCCCTGGGTTGAAAAGAAAAAAAAAAACAATATACATAAATTTTTGGAAAAACCCCCCCTCCCCCCAAA
TATACACGGTATCCTCGAAAAACCCCTAAAACGAGGGCCGATTATATATTTTTCTTTTAATGACAGTGTAG
AATTAATTTATCATCATTATATATAGTGTCACGCTATGGCAT

>Haplotype: GN12 GB#MN784018 (Sample: LCRP56)
AAAGATCAACACATGACCAACCTGCATAAACGCATGTACTTGCGTGAACGCAAATGCATTTGTGAAAACAC
AAATGCAATTGCAAAATTATGCAGGTTGGTCAGTCTGACATATAAACTATTTAGCCCCATACTATGTATAT
CCTCATTAAATATACATACTATGCTTAATCCTCATTAATAGACTGTCCCCATATACATATTATGCTTAATC
CACATCATACATATTATGCTTAACCCCCACATTACATATTATGCTTAATCCACATCATACATATTATGCTT
AATCCACATCATACATACGTCTATCTCACTCCCCATACACACATATATGTGTGCTTAATCCCCATTAATCC
ATACTCCCCTACTCCATTACATGCTATGCTTAATCCTCATTAATAGACTGTCCCCATAGTACATACTATGC
TTAATCCTCATTAACATACTATGCTTAATCCCCATTAATAGACATTCCCCTATTTCATTACATCTTACGTA
TCATTAATCCTCATTACTAGAAAATCCTCTACTCCATTACAATCATTCACTCACCCTCATTAATTTACTTC
TAATACTGCTTATGCGGGCTGGCGAGAAATAACCAATACTCTCTTATAAGACCCCTTTGCACGGTTTGTGG
TACTGTAATCGAATAATCCCCATCAATTGTTCAAATCCCTCATTTGGCACCCTTAAAAAGGTATACGTCTC
TTAATCGCGTCAGAACTCCTTTTCCTCTAGTTCCCTTAATTGTCATTTCGACTCTTAATCGTCTCAAGATT
TCTCAGTCCTCCCAGTCTTTTTTTAGGGGAATGAAGCAATCGTTAACCCCCCAGGGTTGGATACGGCTTTA
TCATTTAAACCTGTGCTACCCTCGACATAATATTTGTATGACCTAACTACTTTAACATTCATAGGACATTA
TTTGTCAAGATTGAGCTAATTAATCGTTACGGGAGAATAATTGTGAAGTCATAATTGACACAGTGGGTAAT
GAATAATCAGATATACATTTAAGAAAGACATTTCATGGACTTCGACCATTAAGTGTAGTGAACAACTATTT
AATTAATGTAAGGGGCATTGAACTACATGGTACATATATCGATATTTTGGCATAGACTTATTTATATGGGT
ACCCCTGGGTCGAAAAGAAAAAAAAAAAACAATATACATAAATTTTTGGAAAAACCCCCCCTCCCCCCAAAT
ATACACGGTATCCTCGAAAAACCCCTAAAACGAGGGCCGATTATATATTTTTCTTTTAATGACATGTAGAA
TTAATTATCATCATTATATATAGTATCACACTATGGCAT

>Haplotype: GN13 GB#MN784017 (Sample: LCRP57)
AAAGATCAACACATGACCAACCTGCATAAACGCATGTACTTGCGTGAACGCAAATGCATTTGTGAAAACAC
AAATGCAATTGCAAAATTATGCAGGTTGGTCAGTCTGACATATAAACTATTTAGCCCCATACTATGTATAT
CCTCATTAAATATACATACTATGCTTAATCCTCATTAATAGACTGTCCCCATATACATATTATGCTTAATC
CACATCATACATATTATGCTTAACCCCCACATTACATATTATGCTTAATCCACATCATACATATTATGCTT
AATCCACATCATACATACGTCTATCTCACTCCCCATACACACATATATGTGTGCTTAATCCCCATTAATCC
ATACTCCCCTACTCCATTACATGCTATGCTTAATCCTCATTAATAGACTGTCCCCATAGTACATACTATGC
TTAATCCTCATTAACATACTATGCTTAATCCCCATTAATAGACATTCCCCTATTTCATTACATCTTACGTA
TCATTAATCCTCATTACTAGAAAATCCTCTACTCCATTACAATCATTCACTCACCCTCATTAATTTACTTC
TAATACTGCTTATGCGGGCTGGGGAGAAATAACCAATACTCTTTTATAAGACCCCTTTGCACGGTTTGTGG
TACTGTAATCGAATAATCCCCATCAATTGTTCAAATCCCTCATTTGGCACCCTTAAAAAGGTATACGTCTT
TTAATCGCGTCAGAACTCCTTTTCCTTTAGTTCCCTTAATTGTCATTTCGACTTTTAATTGTTTCAAGATT
TTTCAGTCCTCCCAGTTTTTTTTTTAGGGGAATGAAGCAATTGTTAACCCGCCAGGGTTGGATACGGCTTTA
TCATTTAAACCTGTGCTACCCTTGACATAATATTTGTATGACCTAACTACTTTAACATTCATAGGACATTA
TTTGTCAAGATTGAGCTAATTAATTGTTTCGGGAGAATAATTGTGAAGTCATAATTGACCCAGTGGGTAAT
GAATAATCAGATATACATTTAAGAAAGACATTTCATGGGCTTTGACCATTAAGTGTAGTGAACAAATATTT
AATTAATGTAAGGAGCCTTGAACTACATGGTACATATATCGATATTTTGGCATAGACTTTTTTATATGGGT
ACCCCTGGGTCGAAAAGAAAAAAAAAAACAATATACATAAATTTTTGGAAAAACCCCCCCTCCCCCCAAAT
ATACACGGTATCCTCGAAAAACCCCTAAAACGAGGGCCGATTATATATTTTTCTTTTAATGACATGTAGAA
TTAATTATCATCATTATATATAGTGTCACCCTATGGCAT

>Haplotype: BT1 GB# MN783979 (Sample: BTRP1)
AAAGATCAACACATGACCAACCTGCATAAACGCATGTATTTGCGTGAACGCAAATGCATTTGTGAAAACAC
AAATGCAATTGCAAAATTATGCAGGTTGGTCAGTCTGACATATAAACTATTTAGCCCCATACTATGTATAT
CCTCATTAAATATACATACTATGCTTAATCCTCATTAATAGACTGTCCCCATATACATATTATGCTTAATC
CACATCATACATATTATGCTTAACCCCCACATTACATACTATGCTTAATCCACATCATACATATTATGCTT
AATCCACATCATACATACGTCTATCTCACTCCCCATACACACATATATGTGTGCTTAATCCCCATTAATCC
ATACTCCCCTACTCCATTACATGCTATGCTTAATCCTCATTAATAGACTGTCCCCATAGTACATACTATGC
TTAATCCTCATTAACATACTATGCTTAATCCCCATTAATAGACATTCCCCTATTTCATTACATCTTACGTA
TCATTAATCCTCATTACTAGAAAATCCTCTACTCCATTACAATCATTCACTCACCCTCATTAATTTACTTT
CTAATACTGCTTATGCGGGCTGGCGAGAAATAACCAATACTCTCTTATAAGACCCCTTTGCACGGTTTGTG
GTACTGTAATCGAATAATCCCCATCAATTGTTCAAATCCCTCATTTGGCACCCTTAAAAAGGTATACGTTT
TTTAATCGCGTCAGAACTCCTTTTCCTCTAGTTCCCTTAATTGTCATTTCGACTTTTAATCGTTTCAAGAT
TTTTCAGTCCTCCCAGTTTTTTTTTAGGGGAATGAAGCAATCGTTAACCCGCCAGGGTTGGATACGGCTTT
ATCATTTAAACCTGTGCTACCCTCGACATAATATTTGTATGACCTAACTACTTTAACATTCATAGGACATT
ATTTGTCAAGATTGAGCTAATTAATCGTTACGGGAGAATAATTGTGAAGTCATAATTGACACAGTGGGTAA
TGAATAATCAGATATACATTTAAGAAAGACATTTCATGGACTTTGACCATTAAGTGTAGTGAACAACTATT
TAATTAATGTAAGGAGCATTGAACTACATGGTACATATATCGATATTTTGGCATAGACTTATTTATATGGG
TACCCCTGGGTCGAAAAGAAAAAAAAAAACAATATACATAAATTTTTGGAAAAACCCCCCCTCCCCCCAAA
TATACACGGTATCCTCGAAAAACCCCTAAAACGAGGGCCGATTATATATTTTTCTTTTAATGACATGTAGA
ATTAATTTATCATCATTATATATAGTGTCACACTATGGCAT

>Haplotype: BT2 GB# MN783980 (Sample: BTRP2)
AAAGATCAACACATGACCAACCTGCATAAACGCATGTACTTGCGTGAACGCAAATGCATTTGTGAAAACAC
AAATGCAATTGCAAAATTATGCAGGTTGGTCAGTCTGACATATAAACTATTTAGCCCCATACTATGTATAT
CCTCATTAAATATACATACTATGCTTAATCCTCATTAATAGACTGTCCCCATATACATATTATGCTTAATC
CACATCATACATATTATGCTTAACCCCCACATTACATATTATGCTTAATCCACATCATACATATTATGCTT
AATCCACATCATACATACGTCTATCTCACTCCCCATACACACATATATGTGTGCTTAATCCCCATTAATCC
ATACTCCCCTACTCCATTACATGCTATGCTTAATCCTCATTAATAGACTGTCCCCATAGTACATACTATGC
TTAATCCTCATTAACATACTATGCTTAATCCCCATTAATAGACATTCCCCTATTTCATTACATCTTACGTA
TCATTAATCCTCATTACTAGAAAATCCTCTACTCCATTACAATCATTCACTCACCCTCATTAATTTACTTC
TAATACTGCTTATGCGGGCTGGCGAGAAATAACCAATACTCTCTTATAAGACCCCTTTGCACGGTTTGTGG
TACTGTAATCGAATAATCCCCATCAATTGTTCAAATCCCTCATTTGGCACCCTTAAAAAGGTATACGTTTT
TTAATCGCGTCAGAACTCCTTTTCCTCTAGTTCCCTTAATTGTCATTTTGACTTTTAATTGTTTCAAGATT
TTTCAGTCCTCCCAGTTTTTTTTTTAGGGGAATGAAGCAATTGTTAACCCGCCAGGGTTGGATACGGCTTTA
TCATTTAAACCTGTGCTACCCTTGACATAATACTTGTATGACCTAACTACTTTAACATTCATAGGACATTA
TTTGTCAAGATTGAGGTAATTAATCGTTACGGGAGAATAATTGGGAAGTCATAATTGACACAGGGGGTAAT
GAATAATCAGATATACATTTAAGAAAGACATTTCATGGACTTTGACCATTAAGTGTAGTGAACAACTATTT
AATTAATGTAAGGAGCATTGAACTACATGGTACATATATCGATATTTTGGCATAGACTTATTTATATGGGG

ACCCCTGGGTCGAAAAGAAAAAAAAAAACAATATACATAAATTTTTGGAAAAACCCCCCCTCCCCCCAAAT
ATACACGGTATCCTCGAAAAACCCCTAAAACGAGGGCCGATTATATATTTTTCTTTTAATGACATGTAGAA
TTAATTATCATCATTATATATAGTGTCACACTATGGCAT

>Haplotype: BT3 GB# MN783981 (Sample: BTRP3)
AAAGATCAACACATGACCAACCTGCATAAACGCATGTATTTGCGTGAACGCAAATGCATTTGTGAAAACAC
AAATGCAATTGCAAAATTATGCAGGTTGGTCAGTCTGACATATAAACTATTTAGCCCCATACTATGTATAT
CCTCATTAAATATACATACTATGCTTAATCCTCATTAATAGACTGTCCCCATATACATATTATGCTTAATC
CACATCATACATATTATGCTTAACCCCCACATTACATATTATGCTTAATCCACATCATACATATTATGCTT
AATCCACATCATACATACGTCTATCTCACTCCCCATACACACATATATGTGTGCTTAATCCCCATTAATCC
ATACTCCCCTACTCCATTACATGCTATGCTTAATCCTCATTAATAGACTGTCCCCATAGTACATACTATGC
TTAATCCTCATTAACATACTATGCTTAATCCCCATTAATAGACATTCCCCTATTTCATTACATCTTACGTA
TCATTAATCCTCATTACTAGAAAATCCTCTACTCCATTACAATCATTCACTCACCCTCATTAATTTACTTC
TAATACTGCTTATGCGGGCTGGCGAGAAATAACCAATACTCTCTTATAAGACCCCTTTGCACGGTTTGTGG
TACTGTAATCGAATAATCCCCATCAATTGTTCAAATCCCTCATTTGGCACCCTTAAAAAGGTATACGTCTT
TTAATCGCGTCAGAACTCCTTTTCCTTTAGTTCCCTTAATTGTCATTTTGACTTTTAATTGTTTCAAGATT
TTTCAGTCCTCCCAGTTTTTTTTTAGGGGAATGAAGCAATTGTTAACCCGCCAGGGTTGGATACGGCTTTA
TCATTTAAACCTGTGCTACCCTCGACATAATATTTGTATGACCTAACTACTTTAACATTCATAGGACATTT
TTTGTCAAGATTGAGGTAATTAATTGTTACGGGGGAATAATTGGGAAGTCATAATTGACCCAGGGGGTAAT
GAATAATCAGATATACATTTAAGAAAGACATTTCATGGACTTTGACCATTAAGTGTAGTGAACAACTATTT
AATTAATGTAAGGAGCATTGAACTACATGGTACATATATTGATATTTTGGCATAGACTTATTTATATGGGG
ACCCCTGGGTCGAAAAGAAAAAAAAAAAACAATATACATAAATTTTTGGAAAAACCCCCCCTCCCCCCAAAT
ATACACGGTATCCTCGAAAAACCCCTAAAACGAGGGCCGATTATATATTTTTCTTTTAATGACATGTAGAA
TTAATTTATCATCATTATATATAGTGTCACACTATGGCAT

>Haplotype: BT4 GB# MN783982 (Sample: BTRP4)
AAAGATCAACACATGACCAACCTGCATAAACGCATGTATTTGCGTGAACGCAAATGCATTTGTGAAAACAC
AAATGCAATTGCAAAATTATGCAGGTTGGTCAGTCTGACATATAAACTATTTAGCCCCATACTATGTATAT
CCTCATTAAATATACATACTATGCTTAATCCTCATTAATAGACTGTCCCCATATACATATTATGCTTAATC
CACATCATACATATTATGCTTAACCCCCACATTACATATTATGCTTAATCCACATCATACATATTATGCTT
AATCCACATCATACATACGTCTATCTCACTCCCCATACACACATATATGTGTGCTTAATCCCCATTAATCC
ATACTCCCCTACTCCATTACATGCTATGCTTAATCCTCATTAATAGACTGTCCCCATAGTACATACTATGC
TTAATCCTCATTAACATACTATGCTTAATCCCCATTAATAGACATTCCCCTATTTCATTACATCTTACGTA
TCATTAATCCTCATTACTAGAAAATCCTCTACTCCATTACAATCATTCACTCACCCTCATTAATTTACTTC
TAATACTGCTTATGCGGGCTGGCGAGAAATAACCAATACTCTCTTATAAGACCCCTTTGCACGGTTTGTGG
TACTGTAATCGAATAATCCCCATCAATTGTTCAAATCCCTCATTTGGCACCCTTAAAAAGGTATACGTCTC
TTAATCGCGTCAGAACTCCTTTTCCTCTAGTTCCCTTAATTGTCATTTCGACTCTTAATCGTCTCAAGATT
TCTCAGTCCTCCCAGTCTTTTTTTAGGGGAATGAAGCAATCGTTAACCCCCCAGGGTTGGATACGGCTTTA
TCATTTAAACCTGTGCTACCCTCGACATAATATTTGTATGACCTAACTACTTTAACATTCATAGGACATTA
TTTGTCAAGATTGAGCTAATTAATCGTTACGGGAGAATAATTGTGAAGTCATAATTGACACAGTGGGTAAT
GAATAATCAGATATACATTTAAGAAAGACATTTCATGGACTTCGACCATTAAGTGTAGTGAACAACTATTT
AATTAATGTAAGGAGCATTGAACTACATGGTACATATATCGATATTTTGGCATAGACTTATTTATATGGGT
ACCCCTGGGTCGAAAAGAAAAAAAAAAAACAATATACATAAATTTTTGGAAAAACCCCCCCTCCCCCCAAAT
ATACACGGTATCCTCGAAAAACCCCTAAAACGAGGGCCGATTATATATTTTTCTTTTAATGACATGTAGAA
TTAATTTATCATCATTATATATAGTATCACACTATGGCAT

>Haplotype: BT5 GB# MN783983 (Sample: BTRP5)
AAAGATCAACACATGACCAACCTGCATAAACGCATGTATTTGCGTGAACGCAAATGCATTTGTGAAAACAC
AAATGCAATTGCAAAATTATGCAGGTTGGTCAGTCTGACATATAAACTATTTAGCCCCATACTATGTATAT
CCTCATTAAATATACATACTATGCTTAATCCTCATTAATAGACTGTCCCCATATACATATTATGCTTAATC
CACATCATACATATTATGCTTAACCCCCACATTACATATTATGCTTAATCCACATCATACATATTATGCTT
AATCCACATCATACATACGTCTATCTCACTCCCCATACACACATATATGTGTGCTTAATCCCCATTAATCC
ATACTCCCCTACTCCATTACATGCTATGCTTAATCCTCATTAATAGACTGTCCCCATAGTACATACTATGC
TTAATCCTCATTAACATACTATGCTTAATCCCCATTAATAGACATTCCCCTATTTCATTACATCTTACGTA
TCATTAATCCTCATTACTAGAAAATCCTCTACTCCATTACAATCATTCACTCACCCTCATTAATTTACTTC
TAATACTGCTTATGCGGGCTGGCGAGAAATAACCAATACTCTCTTATAAGACCCCTTTGCACGGTTTGTGG
TACTGTAATCGAATAATCCCCATCAATTGTTCAAATCCCTCATTTGGCACCCTTAAAAAGGTATACGTCTC
TTAATCGCGTCAGAACTCCTTTTCCTCTAGTTCCCTTAATTGTCATTTCGACTCTTAATCGTCTCAAGATT
TTTCAGTCCTCCCAGTCTTTTTTTAGGGGAATGAAGCAATCGTTAACCCGCCAGGGTTGGATACGGCTCTA

TCATTTAAACCTGTGCTACCCTCGACATAATATTTGTATGACCTAACTACTTTAACATTCATAGGACATTA
TTTGTCAAGATTGAGCTAATTAATCGTTACGGGAGAATAATTGTGAAGTCATAATTGACACAGTGGGTAAT
GAATAATCAGATATACATTTAAGAAAGACATTTCATGGACTTCGACCATTAAGTGTAGTGAACAACTATTT
AATTAATGTAAGGAGCATTGAACTACATGGTACATATATCGATATTTTGGCATAGACTTATTTATATGGGT
ACCCCTGGGTCGAAAAGAAAAAAAAAAACAATATACATAAATTTTTGGAAAAACCCCCCCTCCCCCCAAAT
ATACACGGTATCCTCGAAAAACCCCTAAAACGAGGGCCGATTATATATTTTTCTTTTAATGACATGTAGAA
TTAATTTATCATCATTATATATAGTATCACACTATGGCAT

>Haplotype: BT6 GB#MN783984 (Sample: BTRP6)
AAAGATCAACACATGACCAACCTGCATAAACGCATGTATTTGCGTGAACGCAAATGCATTTGTGAAAACAC
AAATGCAATTGCAAAATTATGCAGGTTGGTCAGTCTGACATATAAACTATTTAGCCCCATACTATGTATAT
CCTCATTAAATATACATACTATGCTTAATCCTCATTAATAGACTGTCCCCATATACATATTATGCTTAATC
CACATCATACATATTATGCTTAACCCCCACATTACATATTATGCTTAATCCACATCATACATATTATGCTT
AATCCACATCATACATACGTCTATCTCACTCCCCATACACACATATATGTGTGCTTAATCCCCATTAATCC
ATACTCCCCTACTCCATTACATGCTATGCTTAATCCTCATTAATAGACTGTCCCCATAGTACATACTATGC
TTAATCCTCATTAACATACTATGCTTAATCCCCATTAATAGACATTCCCCTATTTCATTACATCTTACGTA
TCATTAATCCTCATTACTAGAAAATCCTCTACTCCATTACAATCATTCACTCACCCTCATTAATTTACTTC
TAATACTGCTTATGCGGGCTGGCGAGAAATAACCAATACTCTCTTATAAGACCCCTTTGCACGGTTTGTGG
TACTGTAATCGAATAATCCCCATCAATTGTTCAAATCCCTCATTTGGCACCCTTAAAAAGGTATACGTCTC
TTAATCGCGTCAGAACTCCTTTTCCTCTAGTTCCCTTAATTGTCATTTCGACTTTTAATCGTTTCAAGATT
TTTCAGTCCTTCCAGTCTTTTTTTAGGGGAATGAAGCAATCGTTAACCCGCCAGGGTTGGATACGGCTTTA
TCATTTAAACCTGTGCTACCCTCGACATAATATTTGTATGACCTAACTACTTTAACATTCATAGGACATTA
TTTGTCAAGATTGAGCTAATTAATCGTTACGGGAGAATAATTGTGAAGTCATAATTGACACAGTGGGTAAT
GAATAATCAGATATACATTTAAGAAAGACATTTCATGGACTTCGACCATTAAGTGTAGTGAACAACTATTT
AATTAATGTAAGGGGCATTGAACTACATGGTACATATATCGATATTTTGGCATAGACTTATTTATATGGGT
ACCCCTGGGTCGAAAAGAAAAAAAAAAACAATATACATAAATTTTTGGAAAAACCCCCCCTCCCCCCAAAT
ATACACGGTATCCTCGAAAAACCCCTAAAACGAGGGCCGATTATATATTTTTCTTTTAATGACATGTAGAA
TTAATTTATCATCATTATATATAGTGTCACACTATGGCAT

>Haplotype: BT7 GB#MN783985 (Sample: BTRP7)
AAAGATCAACACATGACCAACCTGCATAAACGCATGTACTTGCGTGAACGCAAATGCATTTGTGAAAACAC
AAATGCAATTGCAAAATTATGCAGGTTGGTCAGTCTGACATATAAACTATTTAGCCCCATACTATGTATAT
CCTCATTAAATATACATACTATGCTTAATCCTCATTAATAGACTGTCCCCATATACATATTATGCTTAATC
CACATCATACATATTATGCTTAACCCCCCATTACATATTATGCTTAATCCACATCATACATATTATGCTTA
ATCCACATCATACATACGTCTATCTCACTCCCCATACACACATATATGTGTGCTTAATCCCCATTAATCCA
TACTCCCCTACTCCATTACATGCTATGCTTAATCCTCATTAATAGACTGTCCCCATAGTACATACTATGCT
TAATCCTCATTAACATACTATGCTTAATCCCCATTAATAGACATTCCCCTATTTCATTACATCTTACGTAT
CATTAATCCTCATTACTAGAAAATCCTCTACTCCATTACAATCATTCACTCACCCTCATTAATTTACTTCT
AATACTGCTTATGCGGGCTGGCGAGAAATAACCAATACTCTCTTATAAGACCCCTTTGCACGGTTTGTGGT
ACTGTAATCGAATAATCCCCATCAATTGTTCAAATCCCTCATTTGGCACCCTTAAAAAGGTATACGTCTCT
TAATCGCGTCAGAACTCCTTTTCCTCTAGTTCCCTTAATTGTCATTTCGACTCTTAATCGTCTCAAGATTT
CTCAGTCCTCCCAGTCTTTTTTTAGGGGAATGAAGCAATCGTTAACCCGCCAGGGTTGGATACGGCTTTAT
CATTTAAACCTGTGCTACCCTCGACATAATATTTGTATGACCTAACTACTTTAACATTCATAGGACATTAT
TTGTCAAGATTGAGCTAATTAATCGTTACGGGAGAATAATTGTGAAGTCATAATTGACACAGTGGGTAATG
AATAATCAGATATACATTTAAGAAAGACATTTCATGGACTTCGACCATTAAGTGTAGTGAACAACTATTTA
ATTAATGTAAGGGGCATTGAACTACATGGTACATATATCGATATTTTGGCATAGACTTATTTATATGGGTA
CCCCTGGGTCGAAAAGAAAAAAAAAAACAATATACATAAATTTTTGGAAAAACCCCCCCTCCCCCCAAATA
TACACGGTATCCTCGAAAAACCCCTAAAACGAGGGCCGATTATATATTTTTCTTTTAATGACATGTAGAAT
TAATTATCATCATTATATATAGTGTGCACACTATGGCAT

>Haplotype: BT8 GB#MN783986 (Sample: BTRP8)
AAAGATCAACACATGACCAACCTGCATAAACGCATGTATTTGCGTGAACGCAAATGCATTTGTGAAAACAC
AAATGCAATTGCAAAATTATGCAGGTTGGTCAGTCTGACATATAAACTATTTAGCCCCATACTATGTATAT
CCTCATTAAATATACATACTATGCTTAATCCTCATTAATAGACTGTCCCCATATACATATTATGCTTAATC
CACATCATACATATTATGCTTAACCCCCACATTACATATTATGCTTAATCCACATCATACATATTATGCTT
AATCCACATCATACATACGTCTATCTCACTCCCCATACACACATATATGTGTGCTTAATCCCCATTAATCC
ATACTCCCCTACTCCATTACATGCTATGCTTAATCCTCATTAATAGACTGTCCCCATAGTACATACTATGC
TTAATCCTCATTAACATACTATGCTTAATCCCCATTAATAGACATTCCCCTATTTCATTACATCTTACGTA
TCATTAATCCTCATTACTAGAAAATCCTCTACTCCATTACAATCATTCACTCACCCTCATTAATTTACTTC

TAATACTGCTTATGCGGGCTGGCGAGAAATAACCAATACTCTCTTATAAGACCCCTTTGCACGGTTTGTGG
TACTGTAATCGAATAATCCCCATCAATTGTTCAAATCCCTCATTTGGCACCCTTAAAAAGGTATACGTCTC
TTAATCGCGTCAGAACTCCTTTTCCTCTAGTTCCCTTAATTGTCATTTCGACTCTTAATCGTCTCAAGATT
TCTCAGTCCTCCCAGTTTTTTTTTAGGGGAATGAAGCAATCGTTAACCCGCCAGGGTTGGATACGGCTTTA
TCATTTAAACCTGTGCTACCCTCGACATAATATTTGTATGACCTAACTACTTTAACATTCATAGGACATTA
TTTGTCAAGATTGAGCTAATTAATCGTTACGGGAGAATAATTGTGAAGTCATAATTGACACAGTGGGTAAT
GAATAATCAGATATACATTTAAGAAAGACATTTCATGGACTTCGACCATTAAGTGTAGTGAACAACTATTT
AATTAATGTAAGGAGCATTGAACTACATGGTACATATATCGATATTTTGGCATAGACTTATTTATATGGGT
ACCCCTGGGTCGAAAAGAAAAAAAAAAAACAATATACATAAATTTTTGGAAAAACCCCCCCTCCCCCCAAAT
ATACACGGTATCCTCGAAAAACCCCTAAAACGAGGGCCGATTATATATTTTTCTTTTAATGACATGTAGAA
TTAATTTATCATCATTATATATAGTGTCACACTATGGCAT

>Haplotype: BT9 GB#MN783987 (Sample: BTRP9)
AAAGATCAACACATGACCAACCTGCATAAACGCATGTACTTGCGTGAACGCAAATGCATTTGTGAAAACAC
AAATGCAATTGCAAAATTATGCAGGTTGGTCAGTCTGACATATAAACTATTTAGCCCCATACTATGTATAT
CCTCATTAAATATACATACTATGCTTAATCCTCATTAATAGACTGTCCCCATATACATATTATGCTTAATC
CACATCATACATATTATGCTTAACCCCCACATTACATATTATGCTTAATCCACATCATACATATTATGCTT
AATCCACATCATACATACGTCTATCTCACTCCCCATACACACATATATGTGTGCTTAATCCCCATTAATCC
ATACTCCCCTACTCCATTACATGCTATGCTTAATCCTCATTAATAGACTGTCCCCATAGTACATACTATGC
TTAATCCTCATTAACATACTATGCTTAATCCCCATTAATAGACATTCCCCTATTTCATTACATCTTACGTA
TCATTAATCCTCATTACTAGAAAATCCTCTACTCCATTACAATCATTCACTCACCCTCATTAATTTACTTC
TAATACTGCTTATGCGGGCTGGCGAGAAATAACCAATACTCTCTTATAAGACCCCTTTGCACGGTTTGTGG
TACTGTAATCGAATAATCCCCATCAATTGTTCAAATCCCTCATTTGGCACCCTTAAAAAGGTATACGTCT
TTTAATCGCGTCAGAACTCCTTTTCCTCTAGTTCCCTTAATTGTCATTTTGACTTTTAATCGTTTCAAGAT
TTTTCAGTCCTCCCAGTTTTTTTTTAGGGGAAAGAAGCAATTGTTAACCCGCCAGGGTTGGATACGGCTTT
ATCATTTAAACCTGTGCTACCCTTGACATAATATTTGTATGACCTAACTACTTTAACATTCATAGGACATT
ATTTGTCAAGATTGAGCTAATTAATCGTTACGGGAGAATAATTGGGAAGTCATAATTGACACAGTGGGTAA
TGAATAATCAGATATACATTTAAGAAAGACATTTCATGGACTTTGACCATTAAGTGTAGTGAACAACTATT
TAATTAATGTAAGGAGCATTGAACTACATGGTACATATATCGATATTTTGGCATAGACTTATTTATATGGG
TACCCCTGGGTTGAAAAGAAAAAAAAAAAACAATATACATAAATTTTTGGAAAAACCCCCCCTCCCCCCAAA
TATACACGGTATCCTCGAAAAACCCCTAAAACGAGGGCCGATTATATATTTTTCTTTTAATGACATGTAGA
ATTAATTATCATCATTATATATAGTGTCACACTTATGGCAT

>Haplotype: BT10 GB#MN783988 (Sample: BTRP10)
AAAGATCAACACATGACCAACCTGCATAAACGCATGTACTTGCGTGAACGCAAATGCATTTGTGAAAACAC
AAATGCAATTGCAAAATTATGCAGGTTGGTCAGTCTGACATATAAACTATTTAGCCCCATACTATGTATAT
CCTCATTAAATATACATACTATGCTTAATCCTCATTAATAGACTGTCCCCATATACATATTATGCTTAATC
CACATCATACATATTATGCTTAACCCCCACATTACATATTATGCTTAATCCACATCATACATATTATGCTT
AATCCACATCATACATACGTCTATCTCACTCCCCATACACACATATATGTGTGCTTAATCCCCATTAATCC
ATACTCCCCTACTCCATTACATGCTATGCTTAATCCTCATTAATAGACTGTCCCCATAGTACATACTATGC
TTAATCCTCATTAACATACTATGCTTAATCCCCATTAATAGACATTCCCCTATTTCATTACATCTTACGTA
TCATTAATCCTCATTACTAGAAAATCCTCTACTCCATTACAATCATTCACTCACCCTCATTAATTTACTTC
TAATACTGCTTATGCGGGCTGGCGAGAAATAACCAATACTCTCTTATAAGACCCCTTTGCACGGTTTGTGG
TACTGTAATCGAATAATCCCCATCAATTGTTCAAATCCCTCATTTGGCACCCTTAAAAAGGTATACGTCTC
TTAATCGCGTCAGAACTCCTTTTCCTCTAGTTCCCTTAATTGTCATTTCGACTCTTAATCGTCTCAAGATT
TCTCAGTCCTCCCAGTCTTTTTTTTAGGGGAATGAAGCAATCGTTAACCCCCCAGGGTTGGATACGGCTTTA
TCATTTAAACCTGTGCTACCCTCGACATAATACTTGTATGACCTAACTACTTTAACATTCATAGGACATTA
TTTGTCAAGATTGAGCTAATTAATCGTTACGGGAGAATAATTGTGAAGTCATAATTGACACAGTGGGTAAT
GAATAATCAGATATACATTTAAGAAAGACATTTCATGGACTTCGACCATTAAGTGTAGTGAACAACTATTT
AATTAATGTAAGGGGCATTGAACTACATGGTACATATATCGATATTTTGGCATAGACTTATTTATATGGGT
ACCCCTGGGTCGAAAAGAAAAAAAAAAAACAATATACATAAATTTTTGGAAAAACCCCCCCTCCCCCCAAAT
ATACACGGTATCCTCGAAAAACCCCTAAAACGAGGGCCGATTATATATTTTTCTTTTAATGACATGTAGAA
TTAATTATCATCATTATATATAGTGTCACACTATGGCAT

>Haplotype: BT11 GB#MN783989 (Sample: BTRP11)
AAAGATCAACACATGACCAACCTGCATAAACGCATGTACTTGCGTGAACGCAAATGCATTTGTGAAAACAC
AAATGCAATTGCAAAATTATGCAGGTTGGTCAGTCTGACATATAAACTATTTAGCCCCATACTATGTATAT
CCTCATTAAATATACATACTATGCTTAATCCTCATTAATAGACTGTCCCCATATACATATTATGCTTAATC
CACATCATACATATTATGCTTAACCCCCACATTACATATTATGCTTAATCCACATCATACATATTATGCTT

AATCCACATCATACATACGTCTATCTCACTCCCCATACACACATATATGTGTGCTTAATCCCCATTAATCC
ATACTCCCCTACTCCATTACATGCTATGCTTAATCCTCATTAATAGACTGTCCCCATAGTACATACTATGC
TTAATCCTCATTAACATACTATGCTTAATCCCCATTAATAGACATTCCCCTATTTCATTACATCTTACGTA
TCATTAATCCTCATTACTAGAAAATCCTCTACTCCATTACAATCATTCACTCACCCTCATTAATTTACTTC
TAATACTGCTTATGCGGGCTGGCGAGAAATAACCAATACTCTCTTATAAGACCCCTTTGCACGGTTTGTGG
TACTGTAATCGAATAATCCCCATCAATTGTTCAAATCCCTCATTTGGCACCCTTAAAAAGGTATACGTCTC
TTAATCGCGTCAGAACTCCTTTTCCTCTAGTTCCCTTAATTGTCATTTCGACTCTTAATCGTCTCAAGATT
TCTTCAGTCCTCCCAGTTTTTTTTTAGGGGAATGAAGCAATCGTTAACCCGCCAGGGTTGGATACGGCTTT
ATCATTTAAACCTGTGCTACCCTCGACATAATATTTGTATGACCTAACTACTTTAACATTCATAGGACATT
ATTTGTCAAGATTGAGCTAATTAATCGTTACGGGAGAATAATTGTGAAGTCATAATTGACACAGTGGGTAA
TGAATAATCAGATATACATTTAAGAAAGACATTTCATGGACTTCGACCATTAAGTGTAGTGAACAACTATT
TAATTAATGTAAGGGGCATTGAACTACATGGTACATATATCGATATTTTGGCATAGACTTATTTATATGGG
TACCCCTGGGTCGAAAAGAAAAAAAAAAACAATATACATAAATTTTGGAAAAACCCCCCCTCCCCCCAAAT
ATACACGGTATCCTCGAAAAACCCCTAAAACGAGGGCCGATTATATATTTTTCTTTTAATGACATGTAGAA
TTAATTATCATCATTATATATAGTGTCACACTATGGCAT

>Haplotype: BT12 GB#MN783990 (Sample: BTRP12)
AAAGATCAACACATGACCAACCTGCATAAACGCATGTACTTGCGTGAACGCAAATGCATTTGTGAAAACAC
AAATGCAATTGCAAAATTATGCAGGTTGGTCAGTCTGACATATAAACTATTTAGCCCCATACTATGTATAT
CCTCATTAAATATACATACTATGCTTAATCCTCATTAATAGACTGTCCCCATATACATATTATGCTTAATC
CACATCATACATATTATGCTTAACCCCCACATTACATATTATGCTTAATCCACATCATACATATTATGCTT
AATCCACATCATACATACGTCTATCTCACTCCCCATACACACATATATGTGTGCTTAATCCCCATTAATCC
ATACTCCCCTACTCCATTACATGCTATGCTTAATCCTCATTAATAGACTGTCCCCATAGTACATACTATGC
TTAATCCTCATTAACATACTATGCTTAATCCCCATTAATAGACATTCCCCTATTTCATTACATCTTACGTA
TCATTAATCCTCATTACTAGAAAATCCTCTACTCCATTACAATCATTCACTCACCCTCATTAATTTACTTC
TAATACTGCTTATGCGGGCTGGCGAGAAATAACCAATACTCTCTTATAAGACCCCTTTGCACGGTTTGTGG
TACTGTAATCGAATAATCCCCATCAATTGTTCAAATCCCTCATTTGGCACCCTTAAAAAGGTATACGTCTC
TTAATCGCGTCAGAACTCCTTTTCCTCTAGTTCCCTTAATTGTCATTTCGACTCTTAATCGTCTCAAGATT
TCTTCAGTCCTCCCAGTCTTTTTTTTAGGGGAATGAAGCAATCGTTAACCCGCCAGGGTTGGATACGGCTTT
ATCATTTAAACCTGTGCTACCCTCGACATAATATTTGTATGACCTAACTACTTTAACATTCATAGGACATT
ATTTGTCAAGATTGAGCTAATTAATCGTTACGGGAGAATAATTGTGAAGTCATAATTGACCCAGTGGGTAA
TGAATAATCAGATATACATTTAAGAAAGACATTTCATGGACTTCGACCATTAAGTGTAGTGAACAACTATT
TAATTAATGTAAGGGGCATTGAACTACATGGTACATATATCGATATTTTGGCATAGACTTATTTATATGGG
TACCCCTGGGTCGAAAAGAAAAAAAAAAACAATATACATAAATTTTTGGAAAAACCCCCCCTCCCCCCAAA
TATACACGGTATCCTCGAAAAACCCCTAAAACGAGGGCCGATTATATATTTTTCTTTTAATGACATGTAGA
ATTAATTATCATCATTATATATAGTGTCACACTATGGCAT

>Haplotype: BT13 GB#MN783991 (Sample: BTRP13)
AAAGATCAACACATGACCAACCTGCATAAACGCATGTACTTGCGTGAACGCAAATGCATTTGTGAAAACAC
AAATGCAATTGCAAAATTATGCAGGTTGGTCAGTCTGACATATAAACTATTTAGCCCCATACTATGTATAT
CCTCATTAAATATACATACTATGCTTAATCCTCATTAATAGACTGTCCCCATATACATATTATGCTTAATC
CACATCATACATATTATGCTTAACCCCCACATTACATATTATGCTTAATCCACATCATACATATTATGCTT
AATCCACATCATACATACGTCTATCTCACTCCCCATACACACATATATGTGTGCTTAATCCCCATTAATCC
ATACTCCCCTACTCCATTACATGCTATGCTTAATCCTCATTAATAGACTGTCCCCATAGTACATACTATGC
TTAATCCTCATTAACATACTATGCTTAATCCCCATTAATAGACATTCCCCTATTTCATTACATCTTACGTA
TCATTAATCCTCATTACTAGAAAATCCTCTACTCCATTACAATCATTCACTCACCCTCATTAATTTACTTC
TAATACTGCTTATGCGGGCTGGCGAGAAATAACCAATACTCTCTTATAAGACCCCTTTGCACGGTTTGTGG
TACTGTAATCGAATAATCCCCATCAATTGTTCAAATCCCTCATTTGGCACCCTTAAAAAGGTATACGTCTC
TTAATCGCGTCAGAACTCCTTTTCCTCTAGTTCCCTTAATTGTCATTTCGACTCTTAATCGTCTCAAGATT
TCTCAGTCCTCCCAGTCTTTTTTTTAGGGGAATGAAGCAATCGTTAACCCGCCAGGGTTGGATACGGCTCTA
TCATTTAAACCTGTGCTACCCTCGACATAATATTTGTATGACCTAACTACTTTAACATTCATAGGACATTA
TTTGTCAAGATTGAGCTAATTAATCGTTACGGGAGAATAATTGTGAAGTCATAATTGACCCAGTGGGTAAT
GAATAATCAGATATACATTTAAGAAAGACATTTCATGGACTTCGACCATTAAGTGTAGTGAACAACTATTT
AATTAATGTAAGGAGCATTGAACTACATGGTACATATATCGATATTTTGGCATAGACTTATTTATATGGGT
ACCCCTGGGTCGAAAAGAAAAAAAAAAACAATATACATAAATTTTTGGAAAAACCCCCCCTCCCCCCAAAT
ATACACGGTATCCTCGAAAAACCCCTAAAACGAGGGCCGATTATATATTTTTCTTTTAATGACATGTAGAA
TTAATTATCATCATTATATATAGTGTCACACTATGGCAT

>Haplotype: BT14 GB#MN783992 (Sample: BTRP14)
AAAGATCAACACATGACCAACCTGCATAAACGCATGTACTTGCGTGAACGCAAATGCATTTGTGAAAACAC
AAATGCAATTGCAAAATTATGCAGGTTGGTCAGTCTGACATATAAACTATTTAGCCCCATACTATGTATAT
CCTCATTAAATATACATACTATGCTTAATCCTCATTAATAGACTGTCCCCATATACATATTATGCTTAATC
CACATCATACATATTATGCTTAACCCCCACATTACATATTATGCTTAATCCACATCATACATATTATGCTT
AATCCACATCATACATACGTCTATCTCACTCCCCATACACACATATATGTGTGCTTAATCCCCATTAATCC
ATACTCCCCTACTCCATTACATGCTATGCTTAATCCTCATTAATAGACTGTCCCCATAGTACATACTATGC
TTAATCCTCATTAACATACTATGCTTAATCCCCATTAATAGACATTCCCCTATTTCATTACATCTTACGTA
TCATTAATCCTCATTACTAGAAAATCCTCTACTCCATTACAATCATTCACTCACCCTCATTAATTTACTTC
TAATACTGCTTATGCGGGCTGGCGAGAAATAACCAATACTCTCTTATAAGACCCCTTTGCACGGTTTGTGG
TACTGTAATCGAATAATCCCCATCAATTGTTCAAATCCCTCATTTGGCACCCTTAAAAAGGTATACGTCTC
TTAATCGCGTCAGAACTCCTTTTCCTCTAGTTCCCTTAATTGTCATTTCGACTCTTAATCGTCTCAAGATT
TCTCAGTCCTCCCAGTCTTTTTTTAGGGGAATGAAGCAATCGTTAACCCGCCAGGGTTGGATACGGCTCTA
TCATTTAAACCTGTGCTACCCTCGACATAATATTTGTATGACCTAACTACTTTAACATTCATAGGACATTA
TTTGTCAAGATTGAGCTAATTAATCGTTACGGGAGAATAATTGTGAAGTCATAATTGACACAGTGGGTAAT
GAATAATCAGATATACATTTAAGAAAGACATTTCATGGACTTCGACCATTAAGTGTAGTGAACAACTATTT
AATTAATGTAAGGAGCATTGAACTACATGGTACATATATCGATATTTTGGCATAGACTTATTTATATGGGT
ACCCCTGGGTCGAAAAGAAAAAAAAAAACAATATACATAAATTTTTGGAAAAACCCCCCCTCCCCCCAAAT
ATACACGGTATCCTCGAAAAACCCCTAAAACGAGGGCCGATTATATATTTTTCTTTTAATGACATGTAGAA
TTAATTATCATCATTATATATAGTGTCACACTATGGCAT

>Haplotype: BT15 GB#MN783993 (Sample: BTRP15)
AAAGATCAACACATGACCAACCTGCATAAACGCATGTATTTGCGTGAACGCAAATGCATTTGTGAAAACAC
AAATGCAATTGCAAAATTATGCAGGTTGGTCAGTCTGACATATAAACTATTTAGCCCCATACTATGTATAT
CCTCATTAAATATACATACTATGCTTAATCCTCATTAATAGACTGTCCCCATATACATATTATGCTTAATC
CACATCATACATACTATGCTTAACCCCCACATTACATATTATGCTTAATCCACATCATACATATTATGCTT
AATCCACATCATACATACGTCTATCTCACTCCCCATACACACATATATGTGTGCTTAATCCCCATTAATCC
ATACTCCCCTACTCCATTACATGCTATGCTTAATCCTCATTAATAGACTGTCCCCATAGTACATACTATGC
TTAATCCTCATTAACATACTATGCTTAATCCCCATTAATAGACATTCCCCTATTTCATTACATCTTACGTA
TCATTAATCCTCATTACTAGAAAATCCTCTACTCCATTACAATCATTCACTCACCCTCATTAATTTACTTC
TAATACTGCTTATGCGGGCTGGCGAGAAATAACCAATACTCTCTTATAAGACCCCTTTGCACGGTTTGTGG
TACTGTAATCGAATAATCCCCATCAATTGTTCAAATCCCTCATTTGGCACCCTTAAAAAGGTATACGTCTT
TTAATCGCGTCAGAACTCCTTTTCCTCTAGTTCCCTTAATTGTCATTTCGACTCTTAATCGTCTCAAGATT
TTTCAGTCCTCCCAGTTTTTTTTTTAGGGGAATGAAGCAATCGTTAACCCCCCAGGGTTGGATACGGCTTTA
TCATTTAAACCTGTGCTACCCTCGACATAATATTTGTATGACCTAACTACTTTAACATTCATAGGACATTA
TTTGTCAAGATTGAGCTAATTAATCGTTACGGGAGAATAATTGTGAAGTCATAATTGACACAGTGGGTAAT
GAATAATCAGATATACATTTAAGAAAGACATTTCATGGACTTCGACCATTAAGTGTAGTGAACAACTATTT
AATTAATGTAAGGGGCATTGAACTACATGGTACATATATCGATATTTTGGCATAGACTTATTTATATGGGT
ACCCCTGGGTCGAAAAGAAAAAAAAAAACAATATACATAAATTTTTGGAAAAACCCCCCCTCCCCCCAAAT
ATACACGGTATCCTCGAAAAACCCCTAAAACGAGGGCCGATTATATATTTTTCTTTTAATGACATGTAGAA
TTAATTTATCATCATTATATATAGTGTCACACTATGGCAT

>Haplotype: BT16 GB#MN783994 (Sample: BTRP16)
AAAGATCAACACATGACCAACCTGCATAAACGCATGTATTTGCGTGAACGCAAATGCATTTGTGAAAACAC
AAATGCAATTGCAAAATTATGCAGGTTGGTCAGTCTGACATATAAACTATTTAGCCCCATACTATGTATAT
CCTCATTAAATATACATACTATGCTTAATCCTCATTAATAGACTGTCCCCATATACATATTATGCTTAATC
CACATCATACATATTATGCTTAACCCCCACATTACATATTATGCTTAATCCACATCATACATATTATGCTT
AATCCACATCATACATACGTCTATCTCACTCCCCATACACACATATATGTGTGCTTAATCCCCATTAATCC
ATACTCCCCTACTCCATTACATGCTATGCTTAATCCTCATTAATAGACTGTCCCCATAGTACATACTATGC
TTAATCCTCATTAACATACTATGCTTAATCCCCATTAATAGACATTCCCCTATTTCATTACATCTTACGTA
TCATTAATCCTCATTACTAGAAAATCCTCTACTCCATTACAATCATTCACTCACCCTCATTAATTTACTTC
TAATACTGCTTATGCGGGCTGGCGAGAAATAACCAATACTCTCTTATAAGACCCCTTTGCACGGTTTGTGG
TACTGTAATCGAATAATCCCCATCAATTGTTCAAATCCCTCATTTGGCACCCTTAAAAAGGTATACGTCTC
TTAATCGCGTCAGAACTCCTTTTCCTCTAGTTCCCTTAATTGTCATTTCGACTTTTAATCGTCTCAAGATT
TTTCAGTCCTCCCAGTTTTTTTTTTAGGGGAATGAAGCAATCGTTAACCCGCCAGGGTTGGATACGGCTTTA
TCATTTAAACCTGTGCTACCCTCGACATAATATTTGTATGACCTAACTACTTTAACATTCATAGGACATTA
TTTGTCAAGATTGAGCTAATTAATCGTTACGGGAGAATAATTGTGAAGTCATAATTGACACAGTGGGTAAT
GAATAATCAGATATACATTTAAGAAAGACATTTCATGGACTTCGACCATTAAGTGTAGTGAACAACTATTT
AATTAATGTAAGGGGCATTGAACTACATGGTACATATATCGATATTTTGGCATAGACTTATTTATATGGGT

ACCCCTGGGTCGAAAAGAAAAAAAAAAACAATATACATAAATTTGAGGAAAAACCCCCCCTCCCCCCAAAT
ATACACGGTATCCTCGAAAAACCCCTAAAACGAGGGCCGATTATATATTTTTCTTTTAATGACATGTAGAA
TTAATTTATCATCATTATATATAGTGTCACACTATGGCAT

>Haplotype: BT17 GB#MN783995 (Sample: BTRP17)
AAAGATCAACACATGACCAACCTGCATAAACGCATGTACTTGCGTGAACGCAAATGCATTTGTGAAAACAC
AAATGCAATTGCAAAATTATGCAGGTTGGTCAGTCTGACATATAAACTATTTAGCCCCATACTATGTATAT
CCTCATTAAATATACATACTATGCTTAATCCTCATTAATAGACTGTCCCCATATACATATTATGCTTAATC
CACATCATACATATTATGCTTAACCCCCACATTACATATTATGCTTAATCCACATCATACATATTATGCTT
AATCCACATCATACATACGTCTATCTCACTCCCCATACACACATATATGTGTGCTTAATCCCCATTAATCC
ATACTCCCCTACTCCATTACATGCTATGCTTAATCCTCATTAATAGACTGTCCCCATAGTACATACTATGC
TTAATCCTCATTAACATACTATGCTTAATCCCCATTAATAGACATTCCCCTATTTCATTACATCTTACGTA
TCATTAATCCTCATTACTAGAAAATCCTCTACTCCATTACAATCATTCACTCACCCTCATTAATTTACTTC
TAATACTGCTTATGCGGGCTGGCGAGAAATAACCAATACTCTCTTATAAGACCCCTTTGCACGGTTTGTGG
TACTGTAATCGAATAATCCCCATCAATTGTTCAAATCCCTCATTTGGCACCCTTAAAAAGGTATACGTCTC
TTAATCGCGTCAGAACTCCTTTTCCTCTAGTTCCCTTAATTGTCATTTCGACTCTTAATCGTCTCAAGATT
TCTCAGTCCTCCCAGTCTTTTTTTAGGGGAATGAAGCAATCGTTAACCCGCCAGGGTTGGATACGGCTCTA
TCATTTAAACCTGTGCTACCCTCGACATAATACTTGTATGACCTAACTACTTTAACATTCATAGGACATTA
TTTGTCAAGATTGAGCTAATTAATCGTTACGGGAGAATAATTGTGAAGTCATAATTGACACAGTGGGTAAT
GAATAATCAGATATACATTTAAGAAAGACATTTCATGGACTTCGACCATTAAGTGTAGTGAACAACTATTT
AATTAATGTAAGGGGCATTGAACTACATGGTACATATATCGATATTTTGGCATAGACTTATTTATATGGGT
ACCCCTGGGTCGAAAAGAAAAAAAAAAAACAATATACATAAATTTTTGGAAAAACCCCCCCTCCCCCCAAAT
ATACACGGTATCCTCGAAAAACCCCTAAAACGAGGGCCGATTATATATTTTTCTTTTAATGACATGTAGAA
TTAATTATCATCATTATATATAGTGTCACACTATGGCAT

>Haplotype: BT18 GB#MN783996 (Sample: BTRP18)
AAAGATCAACACATGACCAACCTGCATAAACGCATGTACTTGCGTGAACGCAAATGCATTTGTGAAAACAC
AAATGCAATTGCAAAATTATGCAGGTTGGTCAGTCTGACATATAAACTATTTAGCCCCATACTATGTATAT
CCTCATTAAATATACATACTATGCTTAATCCTCATTAATAGACTGTCCCCATATACATATTATGCTTAATC
CACATCATACATATTATGCTTAACCCCCACATTACATATTATGCTTAATCCACATCATACATATTATGCTT
AATCCACATCATACATACGTCTATCTCACTCCCCATACACACATATATGTGTGCTTAATCCCCATTAATCC
ATACTCCCCTACTCCATTACATGCTATGCTTAATCCTCATTAATAGACTGTCCCCATAGTACATACTATGC
TTAATCCTCATTAACATACTATGCTTAATCCCCATTAATAGACATTCCCCTATTTCATTACATCTTACGTA
TCATTAATCCTCATTACTAGAAAATCCTCTACTCCATTACAATCATTCACTCACCCTCATTAATTTACTTC
TAATACTGCTTATGCGGGCTGGCGAGAAATAACCAATACTCTCTTATAAGACCCCTTTGCACGGTTTGTGG
TACTGTAATCGAATAATCCCCATCAATTGTTCAAATCCCTCATTTGGCACCCTTAAAAAGGTATACGTCTC
TTAATCGCGTCAGAACTCCTTTTCCTCTAGTTCCCTTAATTGTCATTTCGACTTTTAATCGTCTCAAGATT
TCTCAGTCCTCCCAGTTTTTTTTTAGGGGAATGAAGCAATCGTTAACCCGCCAGGGTTGGATACGGCTTTA
TCATTTAAACCTGTGCTACCCTCGACATAATACTTGTATGACCTAACTACTTTAACATTCATAGGACATTA
TTTGTCAAGATTGAGCTAATTAATCGTTACGGGAGAATAATTGTGAAGTCATAATTGACACAGTGGGTAAT
GAATAATCAGATATACATTTAAGAAAGACATTTCATGGACTTCGACCATTAAGTGTAGTGAACAACTATTT
AATTAATGTAAGGGGCATTGAACTACATGGTACATATATCGATATTTTGGCATAGACTTATTTATATGGGT
ACCCCTGGGTCGAAAAGAAAAAAAAAAAACAATATACATAAATTTTTGGAAAAACCCCCCCTCCCCCCAAAT
ATACACGGTATCCTCGAAAAACCCCTAAAACGAGGGCCGATTATATATTTTTCTTTTAATGACATGTAGAA
TTAATTATCATCATTATATATAGTGTCACACTATGGCAT

>Haplotype: BT19 GB#MN783997 (Sample: BTRP19)
AAAGATCAACACATGACCAACCTGCATAAACGCATGTACTTGCGTGAACGCAAATGCATTTGTGAAAACAC
AAATGCAATTGCAAAATTATGCAGGTTGGTCAGTCTGACATATAAACTATTTAGCCCCATACTATGTATAT
CCTCATTAAATATACATACTATGCTTAATCCTCATTAATAGACTGTCCCCATATACATATTATGCTTAATC
CACATCATACATATTATGCTTAACCCCCACATTACATATTATGCTTAATCCACATCATACATATTATGCTT
AATCCACATCATACATACGTCTATCTCACTCCCCATACACACATATATGTGTGCTTAATCCCCATTAATCC
ATACTCCCCTACTCCATTACATGCTATGCTTAATCCTCATTAATAGACTGTCCCCATAGTACATACTATGC
TTAATCCTCATTAACATACTATGCTTAATCCCCATTAATAGACATTCCCCTATTTCATTACATCTTACGTA
TCATTAATCCTCATTACTAGAAAATCCTCTACTCCATTACAATCATTCACTCACCCTCATTAATTTACTTC
TAATACTGCTTATGCGGGCTGGCGAGAAATAACCAATACTCTCTTATAAGACCCCTTTGCACGGTTTGTGG
TACTGTAATCGAATAATCCCCATCAATTGTTCAAATCCCTCATTTGGCACCCTTAAAAAGGTATACGTCTC
TTAATCGCGTCAGAACTCCTTTTCCTCTAGTTCCCTTAATTGTCATTTCGACTCTTAATCGTCTCAAGATT
TCTCAGTCCTCCCAGTCTTTTTTTAGGGGAATGAAGCAATCGTTAACCCGCCAGGGTTGGATACGGCTTTA

TCATTTAAACCTGTGCTACCCTCGACATAATACTTGTATGACCTAACTACTTTAACATTCATAGGACATTA
TTTGTCAAGATTGAGCTAATTAATCGTTACGGGAGAATAATTGTGAAGTCATAATTGACACAGTGGGTAAT
GAATAATCAGATATACATTTAAGAAAGACATTTCATGGACTTCGGCCATTAAGTGTAGTGAACAACTATTT
AATTAATGTAAGGGGCATTGAACTACATGGTACATATATCGATATTTTGGCATAGACTTATTTATATGGGT
ACCCCTGGGTCGAAAAGAAAAAAAAAAACAATATACATAAATTTTTGGAAAAACCCCCCCTCCCCCCCAAT
ATACACGGTATCCTCGAAAAACCCCTAAAACGAGGGCCGATTATATATTTTTCTTTTAATGACATGTAGAA
TTAATTATCATCATTATATATAGTGTCACACTATGGCAT

>Haplotype: BT20 GB#MN783998 (Sample: BTRP20)
AAAGATCAACACATGACCAACCTGCATAAACGCATGTACTTGCGTGAACGCAAATGCATTTGTGAAAACAC
AAATGCAATTGCAAAATTATGCAGGTTGGTCAGTCTGACATATAAACTATTTAGCCCCATACTATGTATAT
CCTCATTAAATATACATACTATGCTTAATCCTCATTAATAGACTGTCCCCATATACATATTATGCTTAATC
CACATCATACATATTATGCTTAACCCCCACATTACATATTATGCTTAATCCACATCATACATATTATGCTT
AATCCACATCATACATACGTCTATCTCACTCCCCATACACACATATATGTGTGCTTAATCCCCATTAATCC
ATACTCCCCTACTCCATTACATGCTATGCTTAATCCTCATTAATAGACTGTCCCCATAGTACATACTATGC
TTAATCCTCATTAACATACTATGCTTAATCCCCATTAATAGACATTCCCCTATTTCATTACATCTTACGTA
TCATTAATCCTCATTACTAGAAAATCCTCTACTCCATTACAATCATTCACTCACCCTCATTAATTTACTTC
TAATACTGCTTATGCGGGCTGGCGAGAAATAACCAATACTCTCTTATAAGACCCCTTTGCACGGTTTGTGG
TACTGTAATCGAATAATCCCCATCAATTGTTCAAATCCCTCATTTGGCACCCTTAAAAAGGTATACGTCTC
TTAATCGCGTCAGAACTCCTTTTCCTCTAGTTCCCTTAATTGTCATTTCGACTCTTAATCGTCTCAAGATT
TCTCAGTCCTCCCAGTCTTTTTTTAGGGGAATGAAGCAATTGTTAACCCGCCAGGGTTGGATACGGCTTTA
TCATTTAAACCTGTGCTACCCTCGACATAATATTTGTATGACCTAACTACTTTAACATTCATAGGACATTA
TTTGTCAAGATTGAGCTAATTAATCGTTACGGGAGAATAATTGTGAAGTCATAATTGACACAGTGGGTAAT
GAATAATCAGATATACATTTAAGAAAGACATTTCATGGACTTCGACCATTAAGTGTAGTGAACAACTATTT
AATTAATGTAAGGGGCCTTGAACTACATGGTACATATATCGATATTTTGGCATAGACTTATTTATATGGGT
ACCCCTGGGTTGAAAAGAAAAAAAAAAACAATATACATAAATTTTTGGAAAAACCCCCCCTCCCCCCAAAT
ATACACGGTATCCTCGAAAAACCCCTAAAACGAGGGCCGATTATATATTTTTCTTTTAATGACATGTAGAA
TTAATTATCATCATTATATATAGTGTCACACTATGGCAT

>Haplotype: BT21 GB#MN783999 (Sample: BTRP21)
AAAGATCAACACATGACCAACCTGCATAAACGCATGTACTTGCGTGAACGCAAATGCATTTGTGAAAACAC
AAATGCAATTGCAAAATTATGCAGGTTGGTCAGTCTGACATATAAACTATTTAGCCCCATACTATGTATAT
CCTCATTAAATATACATACTATGCTTAATCCTCATTAATAGACTGTCCCCATATACATATTATGCTTAATC
CACATCATACATATTATGCTTAACCCCCACATTACATATTATGCTTAATCCACATCATACATATTATGCTT
AATCCACATCATACATACGTCTATCTCACTCCCCATACACACATATATGTGTGCTTAATCCCCATTAATCC
ATACTCCCCTACTCCATTACATGCTATGCTTAATCCTCATTAATAGACTGTCCCCATAGTACATACTATGC
TTAATCCTCATTAACATACTATGCTTAATCCCCATTAATAGACATTCCCCTATTTCATTACATCTTACGTA
TCATTAATCCTCATTACTAGAAAATCCTCTACTCCATTACAATCATTCACTCACCCTCATTAATTTACTTC
TAATACTGCTTATGCGGGCTGGCGAGAAATAACCAATACTCTCTTATAAGACCCCTTTGCACGGTTTGTGG
TACTGTAATCGAATAATCCCCATCAATTGTTCAAATCCCTCATTTGGCACCCTTAAAAAGGTATACGTCTC
TTAATCGCGTCAGAACTCCTTTTCCTCTAGTTCCCTTAATTGTCATTTCGACTCTTAATCGTCTCAAGATT
TTTCAGTCCTCCCAGTCTTTTTTTAGGGGAATGAAGCAATCGTTAACCCGCCAGGGTTGGATACGGCTCTA
TCATTTAAACCTGTGCTACCCTCGACATAATATTTGTATGACCTAACTACTTTAACATTCATAGGACATTA
TTTGTCAAGATTGAGCTAATTAATCGTTACGGGAGAATAATTGTGAAGTCATAATTGACACAGTGGGTAAT
GAATAATCAGATATACATTTAAGAAAGACATTTCATGGACTTCGACCATTAAGTGTAGTGAACAACTATTT
AATTAATGTAAGGGGCATTGAACTACATGGTACATATATCGATATTTTGGCATAGACTTATTTATATGGGT
ACCCCTGGGTCGAAAAGAAAAAAAAAAACAATATACATAAATTTTTGGAAAAACCCCCCCTCCCCCCAAAT
ATACACGGTATCCTCGAAAAACCCCTAAAACGAGGGCCGATTATATATTTTTCTTTTAATGACATGTAGAA
TTAATTATCATCATTATATATAGTGTCACACTATGGCAT

>Haplotype: BT22 GB#MN784000 (Sample: BTRP22)
AAAGATCAACACATGACCAACCTGCATAAACGCATGTATTTGCGTGAACGCAAATGCATTTGTGAAAACAC
AAATGCAATTGCAAAATTATGCAGGTTGGTCAGTCTGACATATAAACTATTTAGCCCCATACTATGTATAT
CCTCATTAAATATACATACTATGCTTAATCCTCATTAATAGACTGTCCCCATATACATATTATGCTTAATC
CACATCATACATATTATGCTTAACCCCCACATTACATATTATGCTTAATCCACATCATACATATTATGCTT
AATCCACATCATACATACGTCTATCTCACTCCCCATACACACATATATGTGTGCTTAATCCCCATTAATCC
ATACTCCCCTACTCCATTACATGCTATGCTTAATCCTCATTAATAGACTGTCCCCATAGTACATACTATGC
TTAATCCTCATTAACATACTATGCTTAATCCCCATTAATAGACATTCCCCTATTTCATTACATCTTACGTA
TCATTAATCCTCATTACTAGAAAATCCTCTACTCCATTACAATCATTCACTCACCCTCATTAATTTACTTC

TAATACTGCTTATGCGGGGCTGGCGAGAAATAACCAATACTCTCTTATAAGACCCCTTTGCACGGTTTGTG
GTACTGTAATCGAATAATCCCCATCAATTGTTCAAATCCCTCATTTGGCACCCTTAAAAAGGTATACGTCT
CTTAATCGCGTCAGAACTCCTTTTCCTCTAGTTCCCTTAATTGTCATTTCGACTTTTAATCGTCTCAAGAT
TTTTCAGTCCTCCCAGTCTTTTTTTAGGGGAATGAAGCAATCGTTAACCCGCCAGGGTTGGATACGGCTCT
ATCATTTAAACCTGTGCTACCCTCGACATAATATTTGTATGACCTAACTACTTTAACATTCATAGGACATT
ATTTGTCAAGATTGAGCTAATTAATCGTTACGGGAGAATAATTGTGAAGTCATAATTGACACAGTGGGTAA
TGAATAATCAGATATACATTTAAGAAAGACATTTCATGGACTTCGACCATTAAGTGTAGTGAACAACTATT
TAATTAATGTAAGGAGCATTGAACTACATGGTACATATATCGATATTTTGGCATAGACTTATTTATATGGG
TACCCCTGGGTCGAAAAGAAAAAAAAAAACAATATACATAAATTTTTGGAAAAACCCCCCCTCCCCCCAAA
TATACACGGTATCCTCGAAAAACCCCTAAAACGAGGGCCGATTATATATTTTTCTTTTAATGACATGTAGA
ATTAATTTATCATCATTATATATAGTGTCACACTATGGCAT

>Haplotype: BT23 GB#MN784001 (Sample: BTRP23)
AAAGATCAACACATGACCAACCTGCATAAACGCATGTACTTGCGTGAACGCAAATGCATTTGTGAAAACAC
AAATGCAATTGCAAAATTATGCAGGTTGGTCAGTCTGACATATAAACTATTTAGCCCCATACTATGTATAT
CCTCATTAAATATACATACTATGCTTAATCCTCATTAATAGACTGTCCCCATATACATATTATGCTTAATC
CACATCATACATATTATGCTTAACCCCCACATTACATATTATGCTTAATCCACATCATACATATTATGCTT
AATCCACATCATACATACGTCTATCTCACTCCCCATACACACATATATGTGTGCTTAATCCCCATTAATCC
ATACTCCCCTACTCCATTACATGCTATGCTTAATCCTCATTAATAGACTGTCCCCATAGTACATACTATGC
TTAATCCTCATTAACATACTATGCTTAATCCCCATTAATAGACATTCCCCTATTTCATTACATCTTACGTA
TCATTAATCCTCATTACTAGAAAAATCCTCTACTCCATTACAATCATTCACTCACCCTCATTAATTTACTT
CTAATACTGCTTATGCGGGCTGGCGAGAAATAACCAATACTCTCTTATAAGACCCCTTTGCACGGTTTGTG
GTACTGTAATCGAATAATCCCCATCAATTGTTCAAATCCCTCATTTGGCACCCTTAAAAAGGTATACGTCT
CTTAATCGCGTCAGAACTCCTTTTCCTCTAGTTCCCTTAATTGTCATTTCGACTCTTAATCGTCTCAAGAT
TTTCTCAGTCCTCCCAGTTTTTTTTTAGGGGAATGAAGCAATCGTTAACCCGCCAGGGTTGGATACGGCTT
TATCATTTAAACCTGTGCTACCCTCGACATAATATTTGTATGACCTAACTACTTTAACATTCATAGGACAT
TATTTGTCAAGATTGAGCTAATTAATCGTTACGGGAGAATAATTGTGAAGTCATAATTGACACAGTGGGTA
ATGAATAATCAGATATACATTTAAGAAAGACATTTCATGGACTTCGACCATTAAGTGTAGTGAACAACTAT
TTAATTAATGTAAGGGGCATTGAACTACATGGTACATATATCGATATTTTGGCATAGACTTATTTATATGG
GTACCCCTGGGTCGAAAAGAAAAAAAAAAACAATATACATAAATTTTTGGAAAAACCCCCCCTCCCCCCAA
ATATACACGGTATCCTCGAAAAACCCCTAAAACGAGGGCCGATTATATATTTTTCTTTTAATGACATGTAG
AATTAATTATCATCATTATATATAGTGTCACACTATGGCAT

>Haplotype: BT24 GB#MN784002 (Sample: BTRP24)
AAAGATCAACACATGACCAACCTGCATAAACGCATGTATTTGCGTGAACGCAAATGCATTTGTGAAAACAC
AAATGCAATTGCAAAATTATGCAGGTTGGTCAGTCTGACATATAAACTATTTAGCCCCATACTATGTATAT
CCTCATTAAATATACATACTATGCTTAATCCTCATTAATAGACTGTCCCCATATACATATTATGCTTAATC
CACATCATACATATTATGCTTAACCCCCACATTACATATTATGCTTAATCCACATCATACATATTATGCTT
AATCCACATCATACATACGTCTATCTCACTCCCCATACACACATATATGTGTGCTTAATCCCCATTAATCC
ATACTCCCCTACTCCATTACATGCTATGCTTAATCCTCATTAATAGACTGTCCCCATAGTACATACTATGC
TTAATCCTCATTAACATACTATGCTTAATCCCCATTAATAGACATTCCCCTATTTCATTACATCTTACGTA
TCATTAATCCTCATTACTAGAAAATCCTCTACTCCATTACAATCATTCACTCACCCTCATTAATTTACTTC
TAATACTGCTTATGCGGGCTGGCGAGAAATAACCAATACTCTCTTATAAGACCCCTTTGCACGGTTTGTGG
TACTGTAATCGAATAATCCCCATCAATTGTTCAAATCCCTCATTTGGCACCCTTAAAAAGGTATACGTCTC
TTAATCGCGTCAGAACTCCTTTTCCTCTAGTTCCCTTAATTGTCATTTCGACTCTTAATCGTCTCAAGATT
TTTCAGTCCTCCCAGTCTTTTTTTAGGGGAATGAAGCAATCGTTAACCCGCCAGGGTTGGATACGGCTCTA
TCATTTAAACCTGTGCTACCCTCGACATAATATTTGTATGACCTAACTACTTTAACATTCATAGGACATTA
TTTGTCAAGATTGAGCTAATTAATCGTTACGGGAGAATAATTGTGAAGTCATAATTGACACAGTGGGTAAT
GAATAATCAGATATACATTTAAGAAAGACATTTCATGGACTTCGACCATTAAGTGTAGTGAACAACTATTT
AATTAATGTAAGGGGCATTGAACTACATGGTACATATATCGATATTTTGGCATAGACTTATTTATATGGGT
ACCCCTGGGTCGAAAAGAAAAAAAAAAAACAATATACATAAATTTTTGGAAAAACCCCCCCTCCCCCCAAAT
ATACACGGTATCCTCGAAAAACCCCTAAAACGAGGGCCGATTATATATTTTTCTTTTAATGACATGTAGAA
TTAATTTATCATCATTATATATAGTGTCACACTATGGCAT

>Haplotype: BT25 GB#MN784003 (Sample: BTRP25)
AAAGATCAACACATGACCAACCTGCATAAACGCATGTATTTGCGTGAACGCAAATGCATTTGTGAAAACAC
AAATGCAATTGCAAAATTATGCAGGTTGGTCAGTCTGACATATAAACTATTTAGCCCCATACTATGTATAT
CCTCATTAAATATACATACTATGCTTAATCCTCATTAATAGACTGTCCCCATATACATATTATGCTTAATC
CACATCATACATATTATGCTTAACCCCCACATTACATATTATGCTTAATCCACATCATACATATTATGCTT

AATCCACATCATACATACGTCTATCTCACTCCCCATACACACATATATGTGTGCTTAATCCCCATTAATCC
ATACTCCCCTACTCCATTACATGCTATGCTTAATCCTCATTAATAGACTGTCCCCATAGTACATACTATGC
TTAATCCTCATTAACATACTATGCTTAATCCCCATTAATAGACATTCCCCTATTTCATTACATCTTACGTA
TCATTAATCCTCATTACTAGAAAATCCTCTACTCCATTACAATCATTCACTCACCCTCATTAATTTACT
AAAGATCAACACATGACCAACCTGCATAAACGCATGTATTTGCGTGAACGCAAATGCATTTGTGAAAACAC
AAATGCAATTGCAAAATTATGCAGGTTGGTCAGTCTGACATATAAACTATTTAGCCCCATACTATGTATAT
CCTCATTAAATATACATACTATGCTTAATCCTCATTAATAGACTGTCCCCATATACATATTATGCTTAATC
CACATCATACATATTATGCTTAACCCCCACATTACATATTATGCTTAATCCACATCATACATATTATGCTT
AATCCACATCATACATACGTCTATCTCACTCCCCATACACACATATATGTGTGCTTAATCCCCATTAATCC
ATACTCCCCTACTCCATTACATGCTATGCTTAATCCTCATTAATAGACTGTCCCCATAGTACATACTATGC
TTAATCCTCATTAACATACTATGCTTAATCCCCATTAATAGACATTCCCCTATTTCATTACATCTTACGTA
TCATTAATCCTCATTACTAGAAAATCCTCTACTCCATTACAATCATTCACTCACCCTCATTAATTTACTTC
TAATACTGCTTATGCGGGCTGGCGAGAAATAACCAATACTCTCTTATAAGACCCCTTTGCACGGTTTGTGG
TACTGTAATCGAATAATCCCCATCAATTGTTCAAATCCCTCATTTGGCACCCTTAAAAAGGTATACGTCTC
TTAATCGCGTCAGAACTCCTTTTCCTCTAGTTCCCTTAATTGTCATTTCGACTCTTAATCGTCTCAAGATT
TCTCAGTCCTCCCAGTCTTTTTTTAGGGGAATGAAGCAATCGTTAACCCGCCAGGGTTGGATACGGCTTTA
TCATTTAAACCTGTGCTACCCTCGACATAATATTTGTATGACCTAACTACTTTAACATTCATAGGACATTA
TTTGTCAAGATTGAGCTAATTAATCGTTACGGGAGAATAATTGTGAAGTCATAATTGACACAGTGGGTAAT
GAATAATCAGATATACATTTAAGAAAGACATTTCATGGACTTCGACCATTAAGTGTAGTGAACAACTATTT
AATTAATGTAAGGGGCATTGAACTACATGGTACATATATCGATATTTTGGCATAGACTTATTTATATGGGT
ACCCCTGGGTCGAAAAGAAAAAAAAAAAACAATATACATAAATTTTTGGAAAAACCCCCCCTCCCCCCAAAT
ATACACGGTATCCTCGAAAAACCCCTAAAACGAGGGCCGATTATATATTTTTCTTTTAATGACATGTAGAA
TTAATTTATCATCATTATATATAGTGTCACACTATGGCAT

>Haplotype: BT26 GB#MN784004 (Sample: BTRP26)
AAAGATCAACACATGACCAACCTGCATAAACGCATGTACTTGCGTGAACGCAAATGCATTTGTGAAAACAC
AAATGCAATTGCAAAATTATGCAGGTTGGTCAGTCTGACATATAAACTATTTAGCCCCATACTATGTATAT
CCTCATTAAATATACATACTATGCTTAATCCTCATTAATAGACTGTCCCCATATACATATTATGCTTAATC
CACATCATACATATTATGCTTAACCCCCACATTACATATTATGCTTAATCCACATCATACATATTATGCTT
AATCCACATCATACATACGTCTATCTCACTCCCCATACACACATATATGTGTGCTTAATCCCCATTAATCC
ATACTCCCCTACTCCATTACATGCTATGCTTAATCCTCATTAATAGACTGTCCCCATAGTACATACTATGC
TTAATCCTCATTAACATACTATGCTTAATCCCCATTAATAGACATTCCCCTATTTCATTACATCTTACGTA
TCATTAATCCTCATTACTAGAAAATCCTCTACTCCATTACAATCATTCACTCACCCTCATTAATTTACTTC
TAATACTGCTTATGCGGGCTGGCGAGAAATAACCAATACTCTCTTATAAGACCCCTTTGCACGGTTTGTGG
TACTGTAATCGAATAATCCCCATCAATTGTTCAAATCCCTCATTTGGCACCCTTAAAAAGGTATACGTCTC
TTAATCGCGTCAGAACTCCTTTTCCTCTAGTTCCCTTAATTGTCATTTCGACTTTTAATCGTTTCAAGATT
TTTCAGTCCTCCCAGTCTTTTTTTAGGGGAATGAAGCAATCGTTAACCCGCCAGGGTTGGATACGGCTTTA
TCATTTAAACCTGTGCTACCCTCGACATAATATTTGTATGACCTAACTACTTTAACATTCATAGGACATTA
TTTGTCAAGATTGAGCTAATTAATCGTTACGGGAGAATAATTGTGAAGTCATAATTGACACAGTGGGTAAT
GAATAATCAGATATACATTTAAGAAAGACATTTCATGGACTTCGACCATTAAGTGTAGTGAACAACTATTT
AATTAATGTAAGGGGCATTGAACTACATGGTACATATATCGATATTTTGGCATAGACTTATTTATATGGGT
ACCCCTGGGTCGAAAAGAAAAAAAAAAAACAATATACATAAATTTTTGGAAAAACCCCCCCTCCCCCCAAAT
ATACACGGTATCCTCGAAAAACCCCTAAAACGAGGGCCGATTATATATTTTTCTTTTAATGACATGTAGAA
TTAATTATCATCATTATATATAGTGTCACACTATGGCAT

>Haplotype: BT27 GB#MN784005 (Sample: BTRP27)
AAAGATCAACACATGACCAACCTGCATAAACGCATGTATTTGCGTGAACGCAAATGCATTTGTGAAAACAC
AAATGCAATTGCAAAATTATGCAGGTTGGTCAGTCTGACATATAAACTATTTAGCCCCATACTATGTATAT
CCTCATTAAATATACATACTATGCTTAATCCTCATTAATAGACTGTCCCCATATACATATTATGCTTAATC
CACATCATACATATTATGCTTAACCCCCACATTACATATTATGCTTAATCCACATCATACATATTATGCTT
AATCCACATCATACATACGTCTATCTCACTCCCCATACACACATATATGTGTGCTTAATCCCCATTAATCC
ATACTCCCCTACTCCATTACATGCTATGCTTAATCCTCATTAATAGACTGTCCCCATAGTACATACTATGC
TTAATCCTCATTAACATACTATGCTTAATCCCCATTAATAGACATTCCCCTATTTCATTACATCTTACGTA
TCATTAATCCTCATTACTAGAAAATCCTCTACTCCATTACAATCATTCACTCACCCTCATTAATTTACTTC
TAATACTGCTTATGCGGGCTGGCGAGAAATAACCAATACTCTCTTATAAGACCCCTTTGCACGGTTTGTGG
TACTGTAATCGAATAATCCCCATCAATTGTTCAAATCCCTCATTTGGCACCCTTAAAAAGGTATACGTCTC
TTAATCGCGTCAGAACTCCTTTTCCTCTAGTTCCCTTAATTGTCATTTCGACTCTTAATCGTCTCAAGATT
TCTCAGTCCTCCCAGTTTTTTTTTAGGGGAATGAAGCAATCGTTAACCCGCCAGGGTTGGATACGGCTCTA
TCATTTAAACCTGTGCTACCCTCGACATAATATTTGTATGACCTAACTACTTTAACATTCATAGGACATTA

TTTGTCAAGATTGAGCTAATTAATCGTTACGGGAGAATAATTGTGAAGTCATAATTGACACAGTGGGTAAT
GAATAATCAGATATACATTTAAGAAAGACATTTCATGGACTTCGACCATTAAGTGTAGTGAACAACTATTT
AATTAATGTAAGGAGCATTGAACTACATGGTACATATATCGATATTTTGGCATAGACTTATTTATATGGGT
ACCCCTGGGTCGAAAAGAAAAAAAAAAAACAATATACATAAATTTTTGGAAAAACCCCCCCTCCCCCCAAAT
ATACACGGTATCCTCGAAAAACCCCTAAAACGAGGGCCGATTATATATTTTTCTTTTAATGACATGTAGAA
TTAATTTATCATCATTATATATAGTGTCACACTATGGCAT