Genomics Facilitates Evaluation and Monitoring of McCloud River Redband Trout (*Oncorhynchus mykiss stonei*)

Ensieh Habibi[1*], Michael R. Miller[1], Daphne Gille[2], Leigh Sanders[2], Jeff Rodzen[2], Molly Stephens[3], Amanda J. Finger[1]

Ensieh Habibi[1] ehabibi@ucdavis.edu

Michael R. Miller[1] micmiller@ucdavis.edu

Daphne Gille[2] dagille@ucdavis.edu

Leigh Sanders[2] Leigh.Sanders@wildlife.ca.gov

Jeff Rodzen[2] Jeff.Rodzen@wildlife.ca.gov

Molly Stephens[3] mstephens3@ucmerced.edu

Amanda J. Finger[1] ajfinger@ucdavis.edu

[1]*Department of Animal Science, University of California Davis, Davis, CA 95616*

[2]*California Department of Fish and Wildlife, Genetics Research Laboratory, Sacramento, CA 95834*

[3]*Natural Reserve System, University of California Merced, Merced, CA 95343*

**Corresponding Author*

Abstract

The McCloud River Redband Trout (MRRT; *Oncorhynchus mykiss stonei*) is a unique subspecies of rainbow trout that inhabits the isolated Upper McCloud River of Northern California. A major threat to MRRT is introgressive hybridization with non-native rainbow trout from historical stocking and contemporary unauthorized introductions. To help address this concern, we collected RAD-sequencing data on 308 total individuals from MRRT and other California *O. mykiss* populations and examined population structure using Principal Component and admixture analyses. Our results are consistent with previous studies; we found that populations of MRRT in Sheepheaven, Swamp, Edson, and Moosehead creeks are nonintrogressed. Additionally, we saw no evidence of introgression in Dry Creek, and suggest further investigation to determine if it can be considered a core MRRT conservation population. Sheepheaven Creek was previously thought to be the sole historical lineage of MRRT, but our analysis identified three: Sheepheaven, Edson, and Dry creeks, all of which should be preserved. Finally, we discovered diagnostic and polymorphic SNP markers for monitoring introgression and genetic diversity in MRRT. Collectively, our results provide a valuable resource for the conservation and management of MRRT.

Keywords: conservation genetics, introgression, trout, SNP assay

Introduction

The McCloud River Redband Trout (MRRT; *Oncorhynchus mykiss stonei,* Jordan 1894) is a subspecies of rainbow trout (*O. mykiss ssp.*) that has been isolated in the Upper McCloud River for tens of thousands of years by a series of waterfalls that arose through volcanic activities (Figure 1; Legendre et al. 1972; Miller 1972; Moyle et al. 2008). The long isolation of MRRT is indicated by their unique ancestral characteristics compared to other subspecies of *O. mykiss,* such as the lowest number of gill rakers, greater number of scales along and above the lateral line, and the frequent presence of vestigial basibranchial teeth (Behnke 1992, 2002). These ancestral characteristics have led to the suggestion that MRRT is a distinctive lineage descended from an early invasion of ancestral trout into the McCloud River headwater system (Behnke 1992, 2002).

The Upper McCloud River and MRRT have been heavily impacted by anthropogenic disturbance over the last hundred years, beginning with grazing and logging. Intense cattle grazing in the first half of the 20th century eliminated streamside vegetation, created shallower and wider streams with warmer temperatures, and reduced water quality (Moyle et al. 2008). Logging began in the late 1800s and rapidly expanded with railroad construction through World War II. Logging degraded stream habitat through removal of shade canopy, further increasing in water temperatures, sedimentation, and peak storm flows, while lowering fish habitat diversity (Bolda and Meyers 1997; Moyle et al. 2008). Road construction throughout the Upper McCloud Basin also provided easy access to streams, increasing inputs of sediment and pollutants. However, in the late 20th century, both private and public land managers began to limit logging and grazing in the area (Moyle et al. 2008).

Recurring drought cycles also threaten MRRT, especially given the porous volcanic soil in the region. Even in average or above average water years, many streams in the Upper McCloud region have dry reaches (Pittman 2011). Several times over the past several decades managers have responded to drought conditions and declining MRRT census sizes by translocating fish from drying pools as a conservation measure. Sheepheaven Creek is the most isolated and morphologically distinct population of MRRT (Behnke 2002), and for many years was thought to be the only non-introgressed MRRT population (see below), due to its remote location. In the 1970s, a series of droughts led managers to translocate MRRT from Sheepheaven Creek into presumably fishless Trout and Swamp creeks as a safeguard (Nielsen et al. 1999; Simmons et al. 2010). Sheepheaven Creek experienced another bottleneck during the 1990-1994 drought when the population declined to fewer than 200

individuals (Nielsen et al. 1999). More recently, severe drought conditions from 2013-2015 spurred the California Department of Fish and Wildlife (CDFW) to "rescue" MRRT from three populations (Edson, Swamp, and Moosehead creeks) and place them in the Mount Shasta Hatchery to prevent local extirpation (CDFW 2017). These creeks were prioritized for rescue because genetic analyses (Simmons et al. 2010) indicated that these three populations, along with Sheepheaven Creek had not become introgressed (i.e., were genetically "pure" MRRT) with hatchery rainbow trout (*O. mykiss ssp.*) (see below). The rescued MRRT were spawned across populations in captivity before they could be returned to the wild, founding a new captive MRRT broodstock. Since then, their progeny have been used to stock the Upper McCloud River and McCloud Reservoir with MRRT for recreational angling and overall population reinforcement. As the risk of severe drought increases with climate change, continued management efforts to limit grazing and logging, combined with captive propagation aim to make MRRT more resilient (Moyle et al. 2008).

Another historical and ongoing major threat to MRRT is the introduction of both distantly and closely related nonnative trout species. Brown trout (*Salmo trutta*) and brook trout (*Salvelinus fontinalis*) were introduced as early as late 1800s. Brown and brook trout impact MRRT through a combination of factors such as predation, competition, and disease introductions (Fausch et al. 2006; McGrath and Lewis 2007; Moyle et al. 2008). Yet an even greater threat to MRRT was the introduction of closely related *O. mykiss ssp*., leading to introgressive hybridization (i.e., incorporation of alleles from one species into the gene complex of another (Anderson 1949)), which can result in loss of both genetic identity and locally adapted alleles (Allendorf and Leary 1988; Wolf et al. 2001; White et al. 2018). Following introgressive hybridization with a conspecific, the phenotype and genotypes of introduced trout tend to become pronounced over time, leading to the erosion of the distinctive characteristics of the native taxon (Wilde and Echelle 1992; Moyle et al. 2008; Seehausen et al. 2008).

Due to increasing recognition of the threat of introgressive hybridization – both from authorized and unauthorized introductions, the stocking of hatchery rainbow trout was discontinued by CDFW in 1994, after Berg (1994) reported introgression in extant MRRT populations (Nielsen et al. 1999; M. Dege, CDFW, pers. comm.). Subsequently, several rounds of genetic studies were conducted to discern levels of introgression in the extant MRRT populations (Nielsen 1999; Simmons et al. 2010). To be conservative, the population in Sheepheaven Creek was considered to be the only non-introgressed population of MRRT until genetic analysis

by Simmons et al. (2010). Nielsen et al. (1999) used microsatellite data and found unique genetic characteristics (e.g., number of alleles, heterozygosity, genetic distance) of Sheepheaven MRRT in comparison to other MRRT. The most recent genetic study was by Simmons et al. (2010), who expanded the sampling effort of Nielsen and used ten SNPs (one mitochondrial and nine nuclear SNPs). Simmons found no evidence of introgression in four out of nine sampled MRRT populations (Sheepheaven, Moosehead, Edson, and Swamp creeks) and only low levels of introgression in several other MRRT populations. Based on Simmons et al. (2010), these four non-introgressed populations are currently classified by CDFW as the "core MRRT conservation populations", and due to the complex geology and hydrology of the Upper McCloud Watershed, are almost entirely isolated except during very rare high flow events (Nielsen et al. 1999; Moyle et al. 2008). In addition, CDFW wrote a genetic monitoring plan as part of a strategy to manage and protect MRRT, including preventing, detecting, and monitoring introgression from any unauthorized stocking. For MRRT, the threats of small population sizes, hatchery propagation, and unauthorized stocking make ongoing management actions and genetic monitoring necessary.

This project aims to develop genetic tools to aid in monitoring of introgression and genetic diversity in MRRT populations. Specific goals are to: 1) test previous "core MRRT conservation population" designations with genomic data; 2) evaluate additional putatively non-introgressed populations for possible designation as core conservation populations; and 3) develop two sets of genetic markers for rapid genetic monitoring of MRRT: markers that are diagnostic between MRRT and introduced *O. mykiss ssp*. to monitor introgression, and markers that are polymorphic within MRRT to monitor genetic diversity (e.g., overall diversity and inbreeding) in both the wild and captivity. Managers can use these markers to identify and adapt to changes in introgression levels and genetic diversity to make informed decisions for conservation actions.

Materials and methods

Sample collection and RAD library preparation

We compiled DNA from 308 individuals from a variety of MRRT and other rainbow trout subspecies in California (Table 1). We included every California rainbow trout subspecies for two reasons. First, the placement of MRRT in phylogenetic analysis has not been consistent among previous studies (Berg 1987; Nielsen et al. 1999). Second, the hatchery stocking and translocation records of rainbow trout into the Upper

McCloud River watershed are incomplete and do not include unauthorized introductions (Simmons et al. 2010; M. Dege, CDFW, pers. comm.). The sample collection includes individuals from either archived (some of which were also used in Simmons et al. [2010]) or newly collected samples (Table 1). Our study included three general groups of samples. In addition to the MRRT sample group, we included a group of various rainbow trout samples which we refer to as the "Rainbow Trout Group" (RBTG). The RBTG group includes wild rainbow trout from Surprise Valley, Goose Lake, Pit River, and Warner Valley (also commonly called "redband trout", but here referred to these as "Other Rainbow Trout" because redband trout does not refer to a monophyletic group), hatchery rainbow trout strains (Coleman, Pit, Hot Creek, Mt. Whitney, Mt. Shasta, and Eagle Lake hatcheries), and wild rainbow trout from North Fork American River, Eagle Lake, and Yuba River. The other sample group is called the "Golden Trout Complex" (GTCX) and includes a representative sample of *O. mykiss ssp.* fish from Golden Trout Creek, South Fork Kern River, Little Kern River, and Kern River watersheds (Table 1; Figure 1).

We extracted DNA using the DNeasy extraction kit according to the manufacturer's protocols (Qiagen). After extraction, we prepared libraries for Restriction Site Associated DNA sequencing (RAD-seq) with the SbfI enzyme based on the protocol described in Ali et al. (2016). For sequencing, all libraries were pooled into a single lane for paired end 150 bp sequencing on an Illumina HiSeq 4000 at UC Davis Genome Center.

Alignment, filtration, and population genetic analysis

After sequencing, we de-multiplexed the data into individual samples and aligned them to the rainbow trout reference genome (Pearse et al. 2019) using the MEM algorithm implemented in the software program BWA (Li and Durbin 2009) to generate Sequence Alignment Map (SAM) files for each individual. The SAM files were then converted to Binary Alignment Map (BAM) files using SAMTOOLS (Li and Durbin 2009; Li 2011). We then used SAMTOOLS to sort, filter for proper pairs, remove PCR duplicates, and index the BAM files. At this stage, 19 individuals were removed due to low read numbers, with 289 individuals remaining. We conducted population genetic analyses in ANGSD (Korneliussen et al. 2014), which analyzes BAM files using on a probabilistic framework in the form of genotype likelihoods. For the analyses, we used the SAMTOOLS

genotype likelihood model (-GL 1) with a minimum base quality of 10 (-minQ 10) and minimum mapping quality of 20 (-minMapQ 20).

We performed hierarchical population structure analyses: 1) on all samples to identify potential sources of introgression with MRRT, 2) on MRRT and potential introgression sources to estimate introgression levels as accurately as possible, and 3) on just MRRT to examine within-group population structure. For each analysis, we used PCA plots and admixture analyses. For PCAs, we used PCAngsd (Meisner and Albrechtsen 2018) on all samples that passed initial quality filtering to produce a covariance matrix. Our PCAngsd parameters were: SNP_pval = 1e-6, -doMajorMinor, and -doMaf 1. In addition, sites had to have a -minMaf 0.05 and be present in at least 50% of the individuals. We then used admixture analyses to test various population groupings based on genetic structure and shared ancestry. To do this, we used NgsAdmix (Skotte et al. 2013) to estimate admixture proportions of individuals with various cluster (K) values based on genotyped likelihoods calculated in ANGSD (Korneliussen et al. 2014). We conducted 10 runs for each K value.

To assess genetic diversity of each population, we used theta statistics in ANGSD (-doThetas) and thetaStat (-do_stat) programs (Korneliussen et al. 2013, 2014). To do this, Site Frequency Spectrum (SFS) (-doSaf) was used as a prior to calculate Tajima's Θ (Θ_π) (Tajima 1983). Tajima's Θ estimates theta ($\Theta = 4N\mu$) based on the average number of pairwise nucleotide differences, and when genomic data is used, accurate estimates can be made with even a small sample size (Nelson et al. 2012; Subramanian 2016).

Discovery of candidate diagnostic and polymorphic loci for MRRT populations

Using the results from our admixture and PCA analyses, we designed two sets of candidate SNP loci: one to detect levels of introgression between rainbow trout and MRRT, and a second to monitor within-MRRT genetic diversity. To design the introgression markers (also referred to as MRRT diagnostic markers), we used two overlapping groups of MRRT: $MRRT_A$ and $MRRT_p$. $MRRT_A$ included all MRRT sampled, even those known to have low levels of introgression, and $MRRT_p$ only includes MRRT that are putatively "pure" (i.e., no detectable introgression) by our population structure analyses (the four core populations, Swamp, Sheepheaven, Edson, and Moosehead creeks, plus Dry Creek; see below). We then used these two sample groups to find loci with substantial allele frequency differences between MRRT and a reference rainbow trout (RBT) group (a subset of RBTG) that includes wild and hatchery rainbow trout (Eagle Lake, North Fork American River, Lower Yuba

River, Lower Stanislaus River, Coleman Strain, Eagle Lake Strain, Hot Creek Strain, Mt. Shasta Strain, Pit Strain).

We performed genotype calling in ANGSD to find candidate SNPs with alleles fixed or nearly-fixed in $MRRT_A$ or $MRRT_P$ but not present or present at very low frequency in the RBT group. Since $MRRT_P$ is not introgressed, we expected a higher number of MRRT-unique markers to be captured compared to $MRRT_A$; therefore, for $MRRT_P$, we increased our cutoff to 99% to discover loci with a higher degree of fixation for pure MRRT. After genotype calling, we calculated the minor allele frequency of each locus in each group and selected loci at two levels: with a frequency of ≥90% and ≥99% in the MRRT group, and conversely a frequency of <10% and <1% in the RBT group. The <10% filter was used for both $MRRT_A$ and $MRRT_P$, and the <1% filter was used for $MRRT_P$, specifically. Markers specific to MRRT (both groups) are appropriate for monitoring the entire MRRT population including pure and introgressed, and markers specific to $MRRT_P$ are appropriate for monitoring pure MRRT populations.

In a separate analysis, we investigated variable SNP loci suitable for monitoring the genetic attributes of both $MRRT_A$ and $MRRT_P$ groups. We performed genotype calling using the same pipeline applied for introgression markers, except we used a Hardy Weinberg test filter (-doHWE 1) to remove paralogs. Paralogous loci can cause the misidentification of heterozygous/homozygous genotypes. After removing paralogs, we selected polymorphic loci with minor allele frequencies between 0.25 to 0.45 in each of MRRT groups. We applied this range of allele frequencies to capture loci that are polymorphic with a moderate frequency: alleles with a lower frequency (< 0.25) are not present in all individuals, and alleles with a higher frequency (0.45 -0.5) can be duplicates (i.e., paralogous loci) so this filter further reduces paralogs. We performed the process separately for $MRRT_A$ and $MRRT_P$ with the same allele frequency (between 0.25- 0.45). Furthermore, to avoid linkage disequilibrium (LD) between selected loci in both polymorphic and diagnostic, we chose loci that are at least 5,000,000bp apart.

Validation analysis

We conducted a validation analysis to determine whether the $MRRT_A$ diagnostic markers identified above were truly diagnostic for the MRRT lineage. We designed SNPtype (Fluidigm Corp.) genotyping assays for a set of 44 candidate SNP loci from $MRRT_A$ loci. We obtained dried fin clip samples from fish collected between 2002

and 2007 from three putatively pure MRRT populations (Sheepheaven, Swamp, and Edson creeks) and three populations of hatchery rainbow trout from Crystal Lake (Coleman and Pit strains) and Darrah Springs (Eagle Lake strain) hatcheries. All samples used for validation analyses were taken from the collection of samples used by Simmons et al. (2010) to ensure consistency and allow for comparison across studies. We extracted DNA from the fin clips on the Hamilton Microlab NIMBUS® HD (Hamilton Company) using the Omega Bio-tek Mag-Bind® Blood & Tissue DNA HDQ 96 Kit (Omega Bio-tek, Inc.) according to the manufacturer's instructions. We then amplified the 44 candidate $MRRT_A$ diagnostic SNP loci using Fluidigm® SNP Type Assays and the Juno 96.96 Genotyping Integrated Fluidic Circuit (IFC) on the Juno™ instrument (Fluidigm Corporation) following the manufacturer's protocol. The Fluidigm SNP genotyping method first uses a locus-specific primer (LSP) and a specific target amplification (STA) primer to enrich for DNA sequences containing the SNP of interest. After enrichment, the LSP and two fluorescently labeled allele specific primers (ASP1 and ASP2) amplify the two possible SNP alleles at a particular locus. We collected fluorescent end-point reads on the Biomark™ HD instrument and used the Fluidigm® SNP Genotyping Analysis Software (version 4.5.1) for SNP allele scoring.

Results

Overall population structure and admixture

Three clear clusters were distinguishable in the PCA that corresponded to the Upper McCloud River redband trout (MRRT), rainbow trout (RBTG), and golden trout (GTCX) groups (Fig. 2a; Table 1). The first principal component (PC1) separates MRRT from GTCX and RBTG, and the second principal component (PC2) separated MRRT from GTCX (Fig. 2a). The horizontal distribution of MRRT along PC1 towards RBGT suggests introgression with this group rather than GTCX. The results in the PCA are supported by the admixture analyses with all samples included, using K=2 and K=3 (Fig. 2b). At K=2, MRRT is differentiated from RBTG and GTCX, and at K=3, GTCX, RBTG, and MRRT are three distinct clusters, but there is about 20-25% RBTG ancestry within much of the MRRT group. After confirming that all of our Upper McCloud River samples clustered together and were genetically separated from the other trout subspecies, we next attempted to distinguish putatively pure and introgressed populations of MRRT. From the previous PCA and admixture result (Fig. 2), it is clear that MRRT are an independent cluster that does not group with GTCX or

RBTG. However, the MRRT distribution on PC1 trends towards a subset of RBTG (a small cluster at the right end of the RBTG main cluster), which suggests that introgression in MRRT is with a subset of RBTG (Fig. 2). This RBTG subset includes wild rainbow trout from Eagle Lake and the American River, Steelhead trout from the Yuba and Stanislaus Rivers, hatchery strains (Shasta, Coleman, Eagle Lake, Hot Creek strains) and three creeks from the "Other Rainbow Trout" group: Lincoln, Lost, and Nelson creeks (Table 1). Hence, we performed PCA and admixture analyses on MRRT and this subset of RBTG to most accurately identify the introgression level in MRRT populations (Fig. 3).

The PCA performed with MRRT and the RBTG subset cluster, first suggests that both hatchery and wild Eagle Lake rainbow trout populations are not the likely source of introgression. Even though on the all sample PCA (Fig. 2a) they are clustered separately as a probable introgression source (small circle within RBTG) (Fig. 3A, top right cluster). Second, the MRRT individuals' horizontal and linear distribution along the PC1 towards the RBTG cluster suggests introgression of some MRRT populations (Fig. 3a). An admixture analysis on the same MRRT and RBTG subset sample groups strongly supports the PCA result (Fig. 3b). The correlation coefficient of 99% ($r = 0.995$) between PC1 position and the proportion of rainbow trout ancestry (%RBT) on each individual further supports this concordance.

Our admixture plot showed no evidence of introgression in individuals tested in five populations from the Upper McCloud watershed: Swamp, Edson, Sheepheaven, Dry, and Moosehead creeks (Fig. 3b). However, there is apparent introgression in individuals from Trout, Blue Heron, Raccoon, Cow, Bull, Tate, McKay, Shady Gulch creeks. Furthermore, six sampling locations in the mainstem Upper McCloud River show varying levels of introgression (Fig. 3b). The levels of introgression vary among the locations (Table 3). For example, Bull, Cow, and Trout creeks have the lowest introgression levels (mean %RBT of 0.0205, 0.0547, and 0.0980, respectively), with other locations showing higher levels of introgression (> 0.1910).

Within MRRT population structure

To assist managers in prioritizing MRRT populations for further population genetic investigation, we examined population structure within MRRT. To do this, we conducted admixture analyses with only the MRRT group (Fig. 4). We found three main ancestry groups at K=4: red representing Swamp and Sheepheaven creeks

ancestry group, blue representing Edson and Dry creeks ancestry group, and purple representing Bull Creek ancestry group (Fig. 4, K=4).

We calculated theta statistics (Θ_π) (Table 4) to quantify the overall genetic diversity of the pure MRRT populations. Dry and Edson had the highest (0.000699) and lowest (0.000474) Θ_π values, respectively, although Moosehead Creek's Θ_π value (0.000691) is almost as high as Dry Creek's. Sheepheaven and Swamp creeks had intermediate Θ_π estimates (0.000592 and 0.000561, respectively).

Diagnostic loci for MRRT

We identified putatively non-introgressed MRRT populations based on our admixture and PCA results and identified individuals for two MRRT groups ($MRRT_A$ and $MRRT_P$). We used N=57 individuals from the $MRRT_A$ population group and N=22 for the $MRRT_P$ group (Sheepheaven, Edson, Swamp, Moosehead, and Dry creeks). The RBT group contained 35 wild and hatchery rainbow trout individuals (Eagle Lake, North Fork American River, Lower Yuba River, Lower Stanislaus River, Coleman Strain, Eagle Lake Strain, Hot Creek Strain, Mt. Shasta Strain, Pit Strain).

We discovered 44 differentially fixed SNPs between the broader $MRRT_A$ group and RBT. We found 2,649 diagnostic SNPs at a frequency of ≥90% between the $MRRT_P$ and RBT (Table 2). When we increased the minimum frequency of diagnostic alleles to 99% in the $MRRT_P$, the number of SNPs was reduced to 574. All of these 574 fixed SNPs had the frequency of 100% in $MRRT_P$ and zero percent in RBT before validation (Table 2). Of the $MRRT_A$ loci, 80% were found in $MRRT_P$ with the 90% cutoff, and 50% were found in $MRRT_P$ with the 99% cutoff. For polymorphic markers, we found 6,639 loci in $MRRT_A$ and 7,316 in $MRRT_P$ (Table 2).

Validation results

We successfully amplified and visualized distinct heterozygote and homozygote clusters in 27 of the 44 candidate $MRRT_A$ diagnostic SNP loci using the Fluidigm SNP Type assays (Tables S1 and S2). There were two loci (omy01_36537055 and omy15_57867903) that failed to amplify in more than 90% of individuals in specific $MRRT_P$ and the tested hatchery rainbow trout populations; this data should therefore be interpreted with caution (denoted by asterisks in Table 5). We obtained genotypes for 90 MRRT from putatively non-introgressed locations and 87 hatchery rainbow trout individuals for at least 90% of these two loci; we discarded data from one individual from Edson Creek because it did not meet this threshold. Of the 27 remaining $MRRT_A$

diagnostic loci, 16 were completely fixed between the groups of $MRRT_A$ and hatchery rainbow trout that we tested (Table 5). The frequency of MRRT alleles was greater than 95% for all historically non-introgressed MRRT populations that we tested (Swamp, Sheepheaven, and Edson creeks) in 21 of the candidate diagnostic loci (Table 5); we propose that these loci should be used in assessing introgression between $MRRT_A$ and $MRRT_P$ with hatchery rainbow trout as part of future monitoring efforts.

Discussion

Using population structure analysis, we identified potential MRRT introgression sources from RBTG and we confirmed former findings of four non-introgressed MRRT locations (Edson, Swamp, Sheepheaven, Moosehead), and potentially one more (Dry). In addition, we identified three main historical lineages with MRRT: Sheepheaven and Swamp creeks, Edson and Dry creeks, and Bull Creek. Using these results, we discovered and validated diagnostic and polymorphic SNP loci specific to the $MRRT_P$ (Swamp, Edson, Sheepheaven, Moosehead, and Dry creeks) and $MRRT_A$ (Swamp, Edson, Sheepheaven, Dry, Moosehead, Blue Heron, Bull, Cow, Tate, McCloud River, Raccoon, Trout, McKay, and Shady Gulch Creek) groups to monitor their introgression and genetic diversity both in the wild and the captive breeding program.

Ancestral history of MRRT

Previous research has identified morphological similarities (Schreck and Behnke 1971; Hoopaugh 1974, Gold 1977) and a shared karyotype (2n=58; Thorgaard 1983) between GTCX and MRRT. Behnke (1981) reconciled their morphological similarities and in light of their geographic separation by suggesting that GTCX may have resulted from multiple invasions of a primitive redband trout via the Sacramento River and Tulare Lake, and that all ancient redbands were subsequently extirpated except for those in the Upper McCloud River. However, Gall (1981) found that the trout in the Upper McCloud River differ substantially from golden trout and rainbow trout by meristic, chromosomal, and electrophoretic characteristic traits. Furthermore, subsequent microsatellite data found that the golden trout complex (except Little Kern golden trout) is significantly different from the MRRT (Nielsen et al. 1999), but the authors suggested that inbreeding and introgression could cause significant genetic difference when two stocks with a common ancestry are isolated. Later, Stephens (2007) also reported a more distant common ancestor of MRRT and golden trout complex from AFLP data.

In our analysis we found three main groups (MRRT, RBTG, and GTCX) using PCA and admixture analyses in our samples (Fig. 2a, b). Despite the fact that the level of introgression in the MRRT group found by the admixture analysis was not very high, at K=2, the introgression source in MRRT appears to be from the RBTG and GTCX ancestry group because these two groups are combined at K=2. However, at K=3, when GTCX and RBTG are split, the introgression source is RBTG (Fig. 2b). Our assessment that MRRT is separate from both RBTG and GTCX supports Behnke's (1992) claim that MRRT should have subspecific status. Thus, we did not find evidence of recent shared ancestry between MRRT and the golden trout complex.

Potential sources of introgression

Most hatchery strains of rainbow trout (primarily *O. m. irideus*) used in California were derived from rainbow trout in the upper Sacramento River (Nielsen et al. 1999). The Eagle Lake rainbow trout (*O. m. aquilarum*) was also stocked extensively in California for several decades and is still in use by CDFW. Given the long history of trout stocking (hatchery records indicating stocking of Mt. Shasta hatchery trout in the Upper and Lower McCloud River (M. Dege, CDFW, pers. Comm.)) and the geographic proximity to the McCloud River, we expected the Mt. Shasta Hatchery (Mt. Shasta, California) to be the most probable source of planted fish in the McCloud River. Prior to Mt. Shasta Hatchery, there was also extensive stocking of trout beginning in the late 1800s from Baird Hatchery, which was located near the confluence of the McCloud River and the Pit River but covered by Lake Shasta following the construction of Shasta Dam in the 1940s. Indeed, Berg (1994) reported hatchery introgression of the Upper McCloud fish by assessing several protein loci, ultimately leading to the end of the hatchery stocking in 1994.

Our results show that only a small subset of the rainbow trout that we examined (Fig. 2a, bottom right of top left cluster) can be sources of introgression. This small group includes wild and hatchery rainbow trout, and three populations from the "Other Rainbow Trout" group. Among the hatchery strains used in this study, Mt. Shasta, Coleman, and Hot Creek represent a potential source of introgression, but Pit Strain and possibly Eagle Lake strain are not; Pit Strain is markedly different from other hatchery trout strains used in California; Pit strain rainbow trout were originally sourced from the Pit River, which falls within the geographic range of the general redband trout designation. It is notable that Nielsen et al. (1999) assessed the same hatchery groups as our study but only found significant genetic association (low genetic distance, R_{ST}) between Sheepheaven

Creek MRRT and Eagle Lake hatchery strain rainbow trout, but not other hatchery strains. The conflict may be explained by the power of RAD sequencing with a much greater coverage and thousands of markers in comparison to highly variable but limited number of microsatellite markers for discovering and genotyping polymorphic loci. In addition, mainstem Upper McCloud River has been extensively stocked with a variety of rainbow trout strains and individual hatcheries often stock multiple different strains of domesticated rainbow trout. For example, Mt. Shasta Hatchery has used Shasta, Coleman and Eagle Lake strain trout for various stocking activities (M. Dege, CDFW, pers. comm.). Additionally, incomplete records of stocking from the past century and ad hoc crosses between hatchery strains can explain observing multiple hatchery strains as a source of introgression. Identifying this small group as the source of introgression allows us to more accurately quantify introgression level.

Levels of Introgression

Our results are generally consistent with Nielsen et al. (1999) and Simmons et al. (2010) in supporting the classification of Edson, Swamp, Sheepheaven, and Moosehead creeks as non-introgressed MRRT core conservation streams. As part of marker discovery, we also included samples from five other creeks in the Upper McCloud watershed: Dry, Bull, Cow, Shady Gulch, and Blue Heron. We only used 3-5 samples from each of these creeks for MRRT diagnostic marker discovery, but our introgression analyses suggest that further genetic investigation of Dry Creek as a candidate core conservation stream is warranted because we detected no rainbow trout introgression in our samples. We also found only minor admixture influence in Bull and Cow creeks and recommend that additional genetic sampling be performed to discern whether these creeks should be considered core conservation streams. Furthermore, in future analyses it is important to include samples from different reaches of these creeks; introgression is believed to vary with hydrological connectivity, accessibility, and the potential for unauthorized stocking (M. Dege, CDFW, pers. comm.).

The unique hydrology and connectivity patterns in the Upper McCloud River could explain the introgression patterns we observed. In general, we expect lower stream connectivity and isolation from the public to correlate with lower levels of introgression. The northern streams (Swamp, Edson, Sheepheaven, and Trout creeks) are the least introgressed and the most isolated. These locations are mostly disconnected from the mainstem Upper McCloud except during rare high-flow events (Nielsen et al. 1999) and are largely on private

timber properties (M. Dege, CDFW, pers. comm.). The only non-introgressed population in the south is Moosehead Creek. Although Moosehead Creek is regularly connected by surface flows, a concrete fish barrier near its confluence with the Upper McCloud River prevents fish from ascending into the creek (M. Dege, CDFW, pers. comm.). The tributaries that are mainly south of the Upper McCloud River are expected to be introgressed, because they are connected to the mainstem Upper McCloud River and may have been stocked with hatchery rainbow trout in the past. For example, Raccoon and Tate creeks showed introgression consistent with Simmons et al. (2010). Raccoon and Tate creeks lack fish barriers between the mainstem Upper McCloud River and are easily accessible to the general public, which may explain the observed higher levels introgression.

We expected that Swamp and Trout creeks, which were recipients of MRRT from Sheepheaven Creek in the 1970s, to be genetically similar to Sheepheaven Creek. However, in the admixture analysis, Trout Creek shows low levels of introgression (0.098), but Swamp does not (Fig. 3). Prior to stocking MRRT, there was a rainbow trout eradication action in Trout Creek which may have been incomplete, or rainbow trout may have been unlawfully introduced via a public campground in the Lower Trout Creek. However, Pittman (2011) reported MRRT as the only salmonid species observed in Trout Creek. Although Trout Creek has no discernable barriers to fish passage, the levels of introgression are highly dependent on sampling location, with lower Trout Creek showing patterns of introgression while upper Trout Creek appears to be more “pure” MRRT (Stephens et al. 2013). In addition to Trout Creek, Bull and Cow creeks had low levels of introgression (0.0205, 0.0547, and 0.098 respectively). These creeks are hydrologically isolated from Upper McCloud River most of the year but may be connected during seasonal or higher flow events, especially in the lower section of Bull Creek (S. Plemons and M. Dege, CDFW, pers. comm.) leading to possible introgression with RBT.

Pure MRRT have three ancestral groups

We observed three historical lineages with MRRT using admixture analyses: 1) Sheepheaven and Swamp creeks, 2) Edson and Dry creeks, and 3) Bull Creek (Fig. 4, K=4). Sheepheaven and Swamp creeks’ common ancestry on the admixture plots is consistent with the fact that Swamp Creek was historically fishless prior to a translocation of MRRT from Sheepheaven Creek. Similar to Simmons et al. (2010), our genetic diversity analysis shows that Swamp Creek has a slightly lower genetic diversity (Θ_{π} = 0.000561) than Sheepheaven (Θ_{π}

384 = 0.000592) which is expected for a derived population and consistent with founder effects (Table 4). Similarly,
385 Trout Creek was founded by a translocation from Sheepheaven Creek and shows a higher proportion of
386 common ancestry with Sheepheaven and Swamp creeks than with other introgressed populations.

387 Interestingly, the Edson and Dry creek ancestry seems to be the most common ancestry group (Fig. 4),
388 as this ancestry group was present in most MRRT populations. However, there is less Sheepheaven Creek
389 ancestry in $MRRT_A$ populations than we expected; Sheepheaven was previously thought to be the "sole
390 representative" of MRRT (Behnke 1992, 2002). Simmons et al. (2010) also found private alleles specific to
391 Edson and Moosehead which were not present in Sheepheaven. For the same reason, it is also surprising that
392 Bull Creek, which has the same ancestry group with Edson, Dry, and Moosehead at K=3, shows a unique
393 MRRT signature at K=4.

394 Our results have significant implications for the management of MRRT. For example, our results
395 suggest that Sheepheaven Creek is not the only representative of ancestral MRRT lineages. Edson and Dry
396 creeks may also represent distinct historical genetic lineages; additional genetic testing with a broader sampling
397 distribution within each tributary is warranted to further refine MRRT conservation efforts for each individual
398 creek. In addition, our analyses showed that Edson Creek has the lowest genetic diversity (Θ_π = 0.000474). This
399 indicates that Edson Creek may benefit from targeted conservation management such as supplementation from
400 other pure MRRT populations within the same lineage (e.g. Moosehead). Simmons et al. (2010) reported low
401 allelic richness for Edson and Sheepheaven, and significant inbreeding and a bottleneck in Sheepheaven based
402 on microsatellite data. However, to prioritize the populations for conservation purpose based on their adaptive
403 potential, neutral genetic diversity is not enough and other factors must be considered such as balancing
404 selection on particular segregating sites, effective population size, rate of mutation, and the populations'
405 adaptive diversity (Messer and Petrov 2013; Kardos and Luikart 2021; Teixeira and Huber 2021).

406 **Diagnostic and polymorphic markers**

407 We identified SNP loci that are both diagnostic for identifying non-introgressed and introgressed MRRT and
408 loci that are polymorphic in MRRT. The development of new SNP type assays will facilitate rapid, consistent
409 genetic typing of individual fish from both pure MRRT and introgressed populations to help inform adaptive
410 conservation strategies and actions. We found 44 diagnostic loci (21 after validation) for the entire MRRT

411 population and more than 500 SNP loci for the non-introgressed MRRT, reflective of their unique, divergent
412 lineage from RBTG. Although only $MRRT_A$ diagnostic markers were validated, we would expect a similar
413 validation efficiency for the $MRRT_P$ markers since the marker discovery methods used for $MRRT_P$ and $MRRT_A$
414 were similar.

415 **Management implications and conclusion**

416 The major goal of genetic monitoring in fish and wildlife populations is to help maintain adaptive capacity in
417 unique genetic lineages, providing populations with resiliency against environmental changes into the future.
418 Introgression between threatened native and introduced populations can alter the genetic diversity of native
419 populations and impact their long-term viability (White et al. 2018; Frankham et al. 2009). The historical and
420 ongoing potential for unauthorized introduction of other rainbow trout has exposes the naturally isolated,
421 genetically unique MRRT to introgressive hybridization with nonnative rainbow trout subspecies. Furthermore,
422 MRRT is especially prone to further loss of genetic diversity through processes inherent to small, isolated
423 populations, such as genetic drift and an accumulation of inbreeding effects. Currently, MRRT only exists in a
424 handful of small, isolated populations and also in a conservation hatchery program as a captive broodstock, thus
425 making an adaptive management strategy a necessity.

426 Admixture, population structure, and genetic diversity analyses conducted in this study corroborate
427 previous work that showed that Edson, Swamp, Sheepheaven, and Moosehead creeks are not introgressed with
428 hatchery rainbow trout. However, we suggest further investigation of Dry, Bull, Trout, and Cow creeks with
429 larger sample sizes to further characterize patterns of introgression and to determine which, if any should be
430 considered for core conservation stream status.

431 The monitoring and management of introgressive hybridization in MRRT core conservation
432 populations should be a critical objective of an MRRT management strategy. Currently, most of the recent
433 hatchery-produced MRRT have been stocked in the Upper McCloud River. Hatchery-produced F_1 juveniles,
434 which were the result of the spawning of rescued wild adult MRRT were stocked into both Trout and
435 Moosehead creeks in a one-time event to help restart the populations following the extreme drought conditions.
436 Currently, no authorized stocking is occurring in the creeks with pure MRRT populations, and further stocking
437 activities will likely be limited to the Upper McCloud River between Algoma (Upper McCloud River) and

McCloud Reservoir (M. Dege, CDFW, pers. comm.). Overall, the results from this study will be a valuable resource to assist with future genetic evaluation and monitoring of MRRT.

Acknowledgments

The authors would like to thank California Department of Fish and Wildlife (CDFW) for funding the project under contract P1781006. We like to thank all members at CDFW, particularly, Dave Lentz who obtained the funding for this study, Michael Dege, Andrew Jensen, and Samuel Plemons who provided the samples for this project and their communications and information. The authors would like to thank UC Davis Genome Cnetr for data sequencing. We also gratefully appreciate Marin Rod & Gun Club (MR&GC) for their scholarship and support for E H. We would like to thank Matthew A. Campbell, Visiting Scholar at UC Davis Genomic Variation lab, and Ryan Peek, Postdoctoral Scholar at UC Davis Center of Watershed Science, for their inputs.

Deceleration:

Funding: This study was funded by California Department of Fish and Wildlife (CDFW) under contract P1781006. **Conflicts of interest/Competing interests:** We have no conflicts of interest to disclose.

Availability of data and material:

Author contributions: Resources: E.H., M.R.M., J.R., M.S., and A.J.F .; investigation: E.H., M.R.M., and A.J.F.; formal analysis: E.H., M.R.M., and A.J.F ; validation: D.G., L.S., and J.R.; visualization: E.H., M.R.M., and A.J.F; writing of original draft: E.H., M.R.M., and A.J.F.; writing, review, and editing: E.H., M.R.M., D.G., L.S., J.R., M.S., and A.J.F.; funding acquisition: M.R.M., J.R., and A.J.F ; supervision: M.R.M., and A.J.F

Code availability: Not applicable **Ethics approval:** Not applicable **Consent to participate:** Not applicable

Consent for publication: Not applicable

Figure and Table captions:

Fig. 1. Sampling locations of all samples in this study. The left map shows sampling locations of all the samples by watershed, including wild, hatchery rainbow trout, "Other Rainbow Trout", and Kern River golden trout and rainbow trout. Inset area is the Upper McCloud River (UMCR) watershed. The red box shows sampling locations within the Upper McCloud River watershed, the area above the Middle Falls. Highlighted river in purple is the mainstem McCloud River.

Golden Trout Creek (GTCR), South Fork Kern River (SFKR), Kern River (KRNR), Eagle Lake (EGLK), North Fork American River (NFAR), Lower Stanislaus River (LSTN), Lower Yuba River (LYBA), Coleman Hatchery (COLE), Eagle Lake Hatchery (EGLH), Hot Creek Strain (HTCS), Mt. Shasta Hatchery (MTSH), Pit Strain Hatchery (PITS), Warner Valley (WARV), Goose Lake (GOSL), Surprise Valley (SPRV), North Fork Pit River (NFPR), South Fork Pit River (SFPR), Upper Pit River (UPIT), Lower Pit River (LPIT), Yuba North Fork (YUBA), Upper McCloud River (UMCR): Swamp Creek (SWPC), Edson Creek (EDSN), Sheepheaven Creek (SHPN), Dry Creek (DRYC), Moosehead Creek (MOHD), Bull Creek (BLLC), Cow Creek (COWC), Trout Creek (TRTC), Shady Gulch Creek (SHGU), Raccoon Creek (RCCN), McCloud River (MCLD), McKay Creek (MCKY), , Blue Heron Creek (BLUN), Tate Creek (TATE)

Fig. 2. Population Structure of all samples. **a**. All samples PCA, color represents watershed. Three main groups are distinguishable: Golden Trout Complex (GTCX), Rainbow Trout Group (RBTG), and McCloud River Redband Trout (MRRT). PC1(8.7% variance explained) / PC2(7.65% variance explained). **b**. All samples admixture plots at K=2 (top admixture plot) and K=3 (bottom admixture plot). Blue represents MRRT ancestry group which is different from GTCX (green) and RBTG (red).

Golden Trout Creek (GTCR), South Fork Kern River (SFKR), Kern River (KRNR), Eagle Lake (EGLK), North Fork American River (NFAR), Lower Stainslaus River (LSTN), Lower Yuba River (LYBA), Coleman Hatchery (COLE), Eagle Lake Hatchery (EGLH), Hot Creek Strain (HTCS), Mt. Shasta Hatchery (MTSH), Pit Strain Hatchery (PITS), Warner Valley (WARV), Goose Lake (GOSL), Surprise Valley (SPRV), North Fork Pit River (NFPR), South Fork Pit River (SFPR), Upper Pit River (UPIT), Lower Pit River (LPIT), Yuba North Fork (YUBA), Upper McCloud River (UMCR)

Fig. 3. PCA and admixture analyses of MRRT with a group of a potential source of introgression – a small subset of RBTG. **a**. PCA of MRRT with the RBTG small subset, color represents watershed. The RBTG subset includes: wild rainbow trout from Eagle lake (EGLK) and North Fork American River (NFAR), Steelhead from Lower Yuba River (LYBA) and Stanislaus River (LSTN), hatchery strains from Mt. Shasta (MTSH), Coleman (COLE), Eagle Lake (EGLH), Hot Creek (HTCS), and three from the "Other Rainbow Trout" group: Lincoln and Lost from Lower Pit River watershed (LPIT) and Nelson creeks from Yuba watershed (YUBA), PC1(11.6% variance explained) / PC2(3.44%, variance explained). **b**. admixture plot of MRRT and the RBTG subset cluster. Five pure populations are identified within the MRRT population: Swamp creek (SWPC), Edson Creek (EDSN), Sheepheaven Creek (SHPN), Dry Creek (DRYC), and Moosehead Creek (MOHD)

Fig. 4. Population structure within MRRT group. MRRT group admixture plot. Three major ancestry groups are distinguishable: red represents Swamp and Sheepheaven creeks' ancestry group, blue represents Edson and Dry creeks' ancestry group, and purple represent Bull creek's ancestry group. The population's order is based on increasing rainbow trout ancestry.
Swamp Creek (SWPC), Edson Creek (EDSN), Sheepheaven Creek (SHPN), Dry Creek (DRYC), Moosehead Creek (MOHD), Bull Creek (BLLC), Cow Creek (COWC), Trout Creek (TRTC), Shady Gulch Creek (SHGU), Raccoon Creek (RCCN), McCloud River (MCLD), McKay Creek (MCKY), , Blue Heron Creek (BLUN), Tate Creek (TATE)

Fig. S1. Admixture analysis of all the samples from K=4-6. X-axis shows grouping by watershed. There is no common ancestor between MRRT group (UMCR watershed) and the California Golden Trout Complex (GTCX).
Golden Trout Creek (GTCR), South Fork Kern River (SFKR), Kern River (KRNR), Eagle Lake (EGLK), North Fork American River (NFAR), Lower Stanislaus River (LSTN), Lower Yuba River (LYBA), Coleman Hatchery (COLE), Eagle Lake Hatchery (EGLH), Hot Creek Strain (HTCS), Mt. Shasta Hatchery (MTSH), Pit Strain Hatchery (PITS), Warner Valley (WARV), Goose Lake (GOSL), Surprise Valley (SPRV), North Fork Pit River (NFPR), South Fork Pit River (SFPR), Upper Pit River (UPIT), Lower Pit River (LPIT), Yuba North Fork (YUBA), Upper McCloud River (UMCR)

Table 1. Additional information for all samples included in the analysis and samples removed after the sequencing and alignment qualifying filtering. Three groups of California native trout were used in this study: Rainbow Trout Group (RBTG), Golden Trout Complex (GTCX), Upper McCloud River Redband Trout (MRRT)

Table 2. Number of loci discovered for monitoring introgression (N loci differentiating MRRT from RBT) and genetic diversity monitoring (N Polymorphic Loci)

Table 3. Percentage of Rainbow Trout ancestry in each MRRT population. The third and fourth columns show minimum and maximum percentage of Rainbow Trout (RBT) ancestry and the last column shows the average across the samples of each population.

Table 4. Theta estimate of the five pure MRRT populations. Θ_{π} = Tajima's Θ, Θ_{W} = Watterson's Θ. Dry and Edson, respectively, have the highest and the lowest estimate of the genetic diversity using theta statistic.

Table 5. Frequencies of MRRT alleles by locus for selected MRRT and hatchery rainbow trout populations. Numbers in bold denote MRRT allele frequencies that are less than 95% for a specific population or grouping. * = loci that failed to amplify in more than 90% of individuals tested in a specific population or grouping. *n* = sample size.

Table S1. Diagnostic MRRT SNP marker sequence information. Sequences are shown in the 5' to 3' direction with brackets denoting the two possible SNP alleles. The chromosome name and position refer to the NCBI *Oncorhynchus mykiss* genome assembly accession GCA_002163495.1.

Table S2. Diagnostic MRRT SNP marker Fluidigm assay information. The SNP assay name refers to the *Oncorhynchus mykiss* chromosome number and position from the NCBI genome assembly accession GCA_002163495.1. ASP1 SNP = allele associated with allele-specific primer 1; ASP2 SNP = allele associated

545 with allele-specific primer 2; ASP1 = allele-specific primer 1 sequence; ASP2 = allele-specific primer 2
546 sequence; LSP = locus-specific primer sequence; STA = specific target amplification primer sequence; AMP
547 GC = proportional GC content.

References

Anderson E (1949) Introgressive hybridization. New York, J. Wiley

Ali OA, O'Rourke SM, Amish SJ, Meek MH, Luikart G, Jeffres C, Miller MR (2016) Rad capture (Rapture): Flexible and efficient sequence-based genotyping. Genetics 202(2):389–4

Allendorf FW, Leary RF (1988) Conservation and distribution of genetic variation in a polytypic species, the cut-throat trout. Conservation Biology 2:170-184

Behnke R J (1992) Native trout of western North America. American Fisheries Society Monograph 6, Bethesda, Maryland.

Behnke RJ (2002) Trout and salmon of North America. Free Press, New York.

Berg WJ (1987) Evolutionary genetics of rainbow trout, Parasalmo gairdnerii (Richardson). Doctoral dissertation. University of California, Davis.

Berg WJ (1994) Introgressive hybridization of McCloud River redband trout by hatchery rainbow trout. Report submitted to California Department of Fish and Game, Wild Trout Division. Sacramento, CA.

Bolda KS, Meyers WJ (1997) Conducting a long-term water quality monitoring project: a case study on the McCloud River, California. Journal of Soil and Water Conservation 52(1):49-54

CDFW (2017) McCloud Redband post outplanting monitoring field report; California Department of Fish and Wildlife, Northern Region. 7 pp

Bolda KS, Meyers WJ (1997) Conducting a long-term water quality monitoring project: a case study on the McCloud River, California. *Journal of Soil and Water Conservation*, 52(1): 49+

Gold JR (1977) Systematics of western North American trout (Salmo), with notes on the redband trout of Sheepheaven Creek, California. California Journal of Zoology, 55, 1858 –1873.

Gall G, Bannon M, Smith R, Bentley B (1981) Progress Report: Native Trout of California California Department of Fish and Game, Threatened Trout Committee. 46 pp

Hoopaugh DA (1974) Status of the Redband Trout (Salmo sp.) in California. California Department of Fish and Game, Inland Fisheries Div. Admin. Rpt. No. 74-7. 11 pp

Fausch KD, Rieman B, Dunham JB, Young MK, Peterson DP (2009) Invasion versus isolation: trade-offs in managing native salmonids with barriers to upstream movement. Conservation Biology 23:859–870

Kardos M Luikart G (2021) The genetic architecture of fitness drives population viability during rapid environmental change. American Naturalist, 10.1086/713469

Korneliussen TS, Albrechtsen A, Nielsen R (2014) ANGSD: Analysis of Next Generation Sequencing Data:1–13

Korneliussen TS, Moltke I, Albrechtsen A. et al. (2013). Calculation of Tajima's *D* and other neutrality test statistics from low depth next-generation sequencing data. BMC Bioinformatics **14,** 289

Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25(14):1754–1760

Li H (2011) A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. Bioinformatics 27(21):2987–2993

Legendre P, Schreck CB, Behnke RJ (1972) Taximetric analysis of selected groups of western North American Salmo with respect to phylogenetic divergences. Syst. Zool. 21:292-307

McGrath CC, Lewis WM, Jr (2007) Competition and predation as mechanisms for displacement of greenback cutthroat trout by brook trout. *Transactions of the American Fisheries Society*, 136: 1381–1392

Meisner J, Albrechtsen A (2018) Inferring Population Structure and Admixture Proportions in Low Depth NGS Data. Genetics, 210(2), 719– 731

Messer PW, Petrov DA (2013) Population genomics of rapid adaptation by soft selective sweeps. Trends Ecol. Evol. 28, 659–669 (2013)

Miller RR (1972) Classification of the native trouts of Arizona with the description of a new species, Salmo apache. Copeia 1972:401-422

Moyle P, Israel JA, Purdy SE (2008) Salmon, Steelhead, and Trout in California, Status of an Emblematic Fauna. Center for Watershed Sciences, University of California, Davis

Nelson MR, Wegmann D, Ehm MG, Kessner D, St Jean P, Verzilli C, et al. (2012) An abundance of rare functional variants in 202 drug target genes sequenced in 14,002 people. Science. 337(6090):100–4

Nielsen JL, Crow KD, Fountain MC (1999) Microsatellite diversity and conservation of a relic trout population: McCloud River redband trout. Molecular Ecology 8(Supplement 1):S129–S142

Pittman RG (2011) Minimum Stream Length Requirements for McCloud River Redband Trout (*Oncorhynchus mykiss spp.*) in Trout and Tate creeks in Siskiyou County, California. Dissertation, Humboldt State University

Pearse DE, Barson NJ, Nome T, Gao G, Campbell MA, Abadía-Cardoso A, Anderson EC, Rundio DE, Williams TH, Naish KA, Moen T, Liu S, Kent M, Moser M, Minkley DR, Rondeau EB, Brieuc MSO, Sandve SR, Miller MR, Cedillo L, Baruch K, Hernandez AG, Ben-Zvi G, Shem-Tov D, Barad O, Kuzishchin K, Garza JC, Lindley ST, Koop BF, Thorgaard GH, Palti Y, Lien S (2019) Sex-dependent dominance maintains migration supergene in rainbow trout. Nat Ecol Evol 3:1731–1742

Schreck CB, Behnke RJ (1971) Trouts of the upper Kern River Basin, California with reference to systematics and evolution of western North American Salmo. J. Fish. Res. Board Can. 28:987-998

Skotte L, Korneliussen TS, Albrechtsen A (2013) Estimating individual admixture proportions from next generation sequencing data. Genetics 195(3):693–702

Simmons RE, Lavretsky P, May B (2010) Introgressive Hybridization of Redband Trout in the Upper McCloud River Watershed. Transactions of the American Fisheries Society 139:201-213

Simmons R, Stephens M, May B (2013) Genetic analysis of California redband trout. Final Report to California Department of Fish and Wildlife. CDFW Agreement #P1181006. May 20, 2013. Genomic Variation Laboratory, University of California, Davis. 18Pp

Stephens MR (2007) Contribution of population genetic structure, hybridization, and cultural attitudes to the conservation of threatened native trout. PhD Dissertation. Davis, University of California

Stephens MR, Erickson B, Finger M, Simmons R, May B. (2013) Final report: Genetic analysis of Native California Trout (Phase 6). California Department of Fish and Wildlife Agreement # P1181006

Subramanian S. (2016) The effects of sample size on population genomic analyses implications for the tests of neutrality. BMC Genomics, 17, 123

Tajima F (1983) Evolutionary relationship of DNA sequences in finite populations. Genetics 105:437–460

Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics. 123:585–95

Teixeira JT, Huber CD (2021) The inflated significance of neutral genetic diversity in conservation genetics. Proceedings of the National Academy of Sciences, 118 (10)

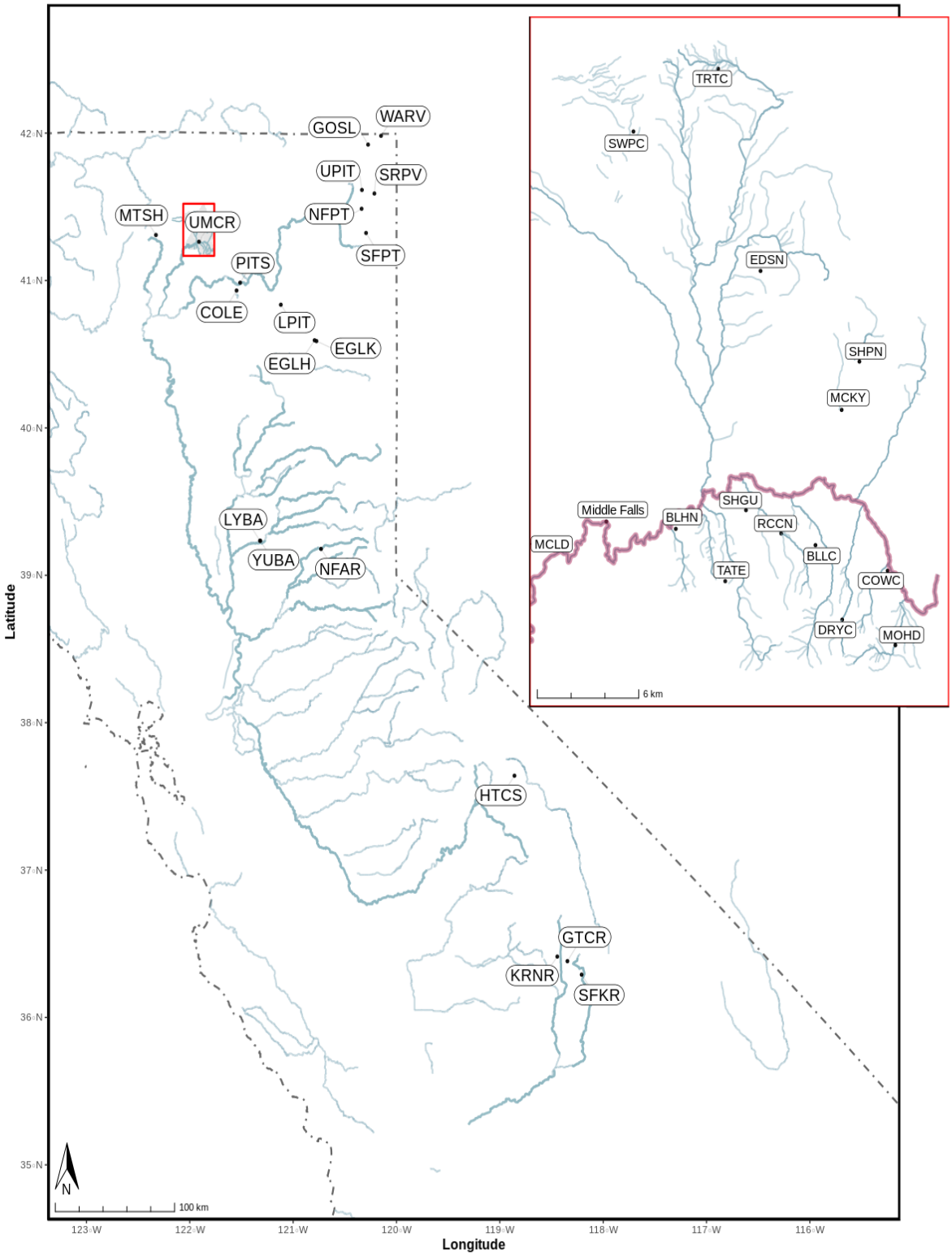
Watterson GA (1975) On the number of segregating sites in genetical models without recombination. Theor. Popul. Biol. 7:256–276

White SL, Miller WL, Dowell SA, Bartron ML, Wagner T (2018) Limited hatchery introgression into wild brook trout (Salvelinus fontinalis) populations despite reoccurring stocking. Evolutionary Applications 11:1567-1581

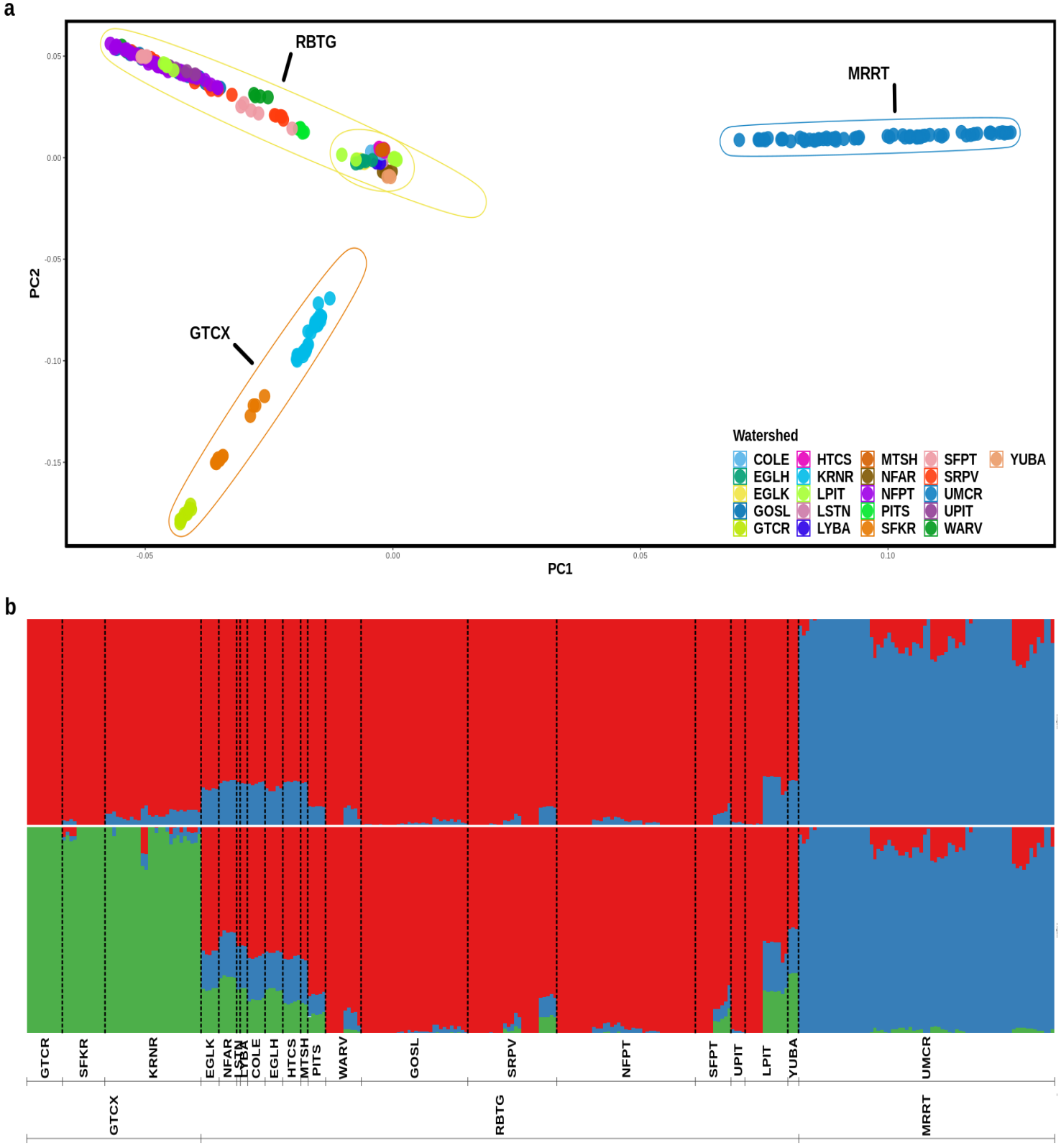
Wolf DE, Takebayashi N, Rieseberg LH (2001) Predicting the risk of extinction through hybridization. Conservation Biology 15(4):1039–1053

636

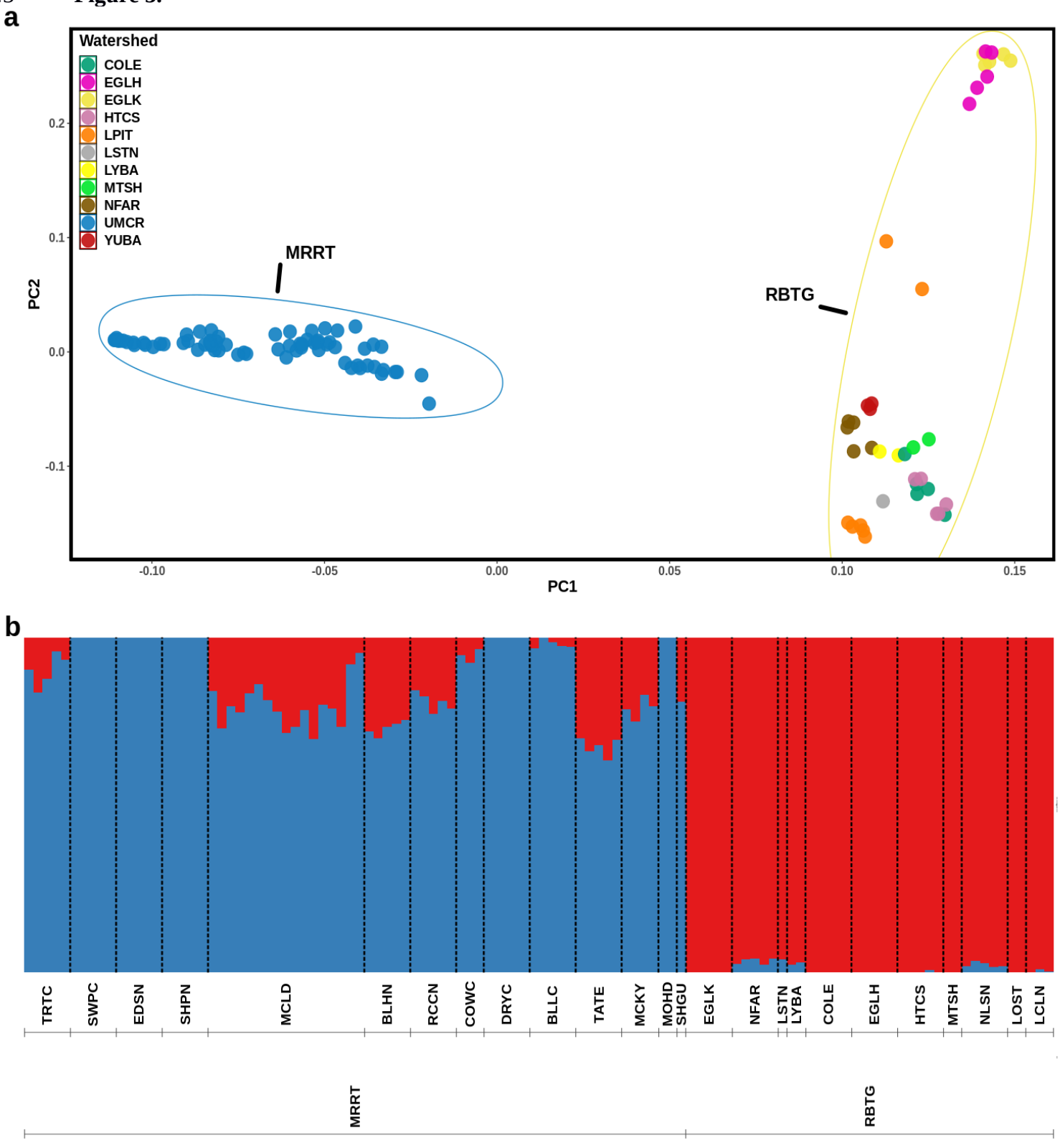
637 **Figure 1.**
638



640
641 **Figure 2.**



643 **Figure 3.**


a
Watershed
COLE
EGLH
EGLK
HTCS
LPIT
LSTN
LYBA
MTSH
NFAR
UMCR
YUBA
MRRT
RBTG
PC2
PC1
0.2
0.1
0.0
-0.1
-0.10
-0.05
0.00
0.05
0.10
0.15
b
TRTC
SWPC
EDSN
SHPN
MCLD
BLHN
RCCN
COWC
DRYC
BLLC
TATE
MCKY
MOHD
SHGU
EGLK
NFAR
LSTN
LYBA
COLE
EGLH
HTCS
MTSH
NLSN
LOST
LCLN
MRRT
RBTG

645
646
647

Figure 4.

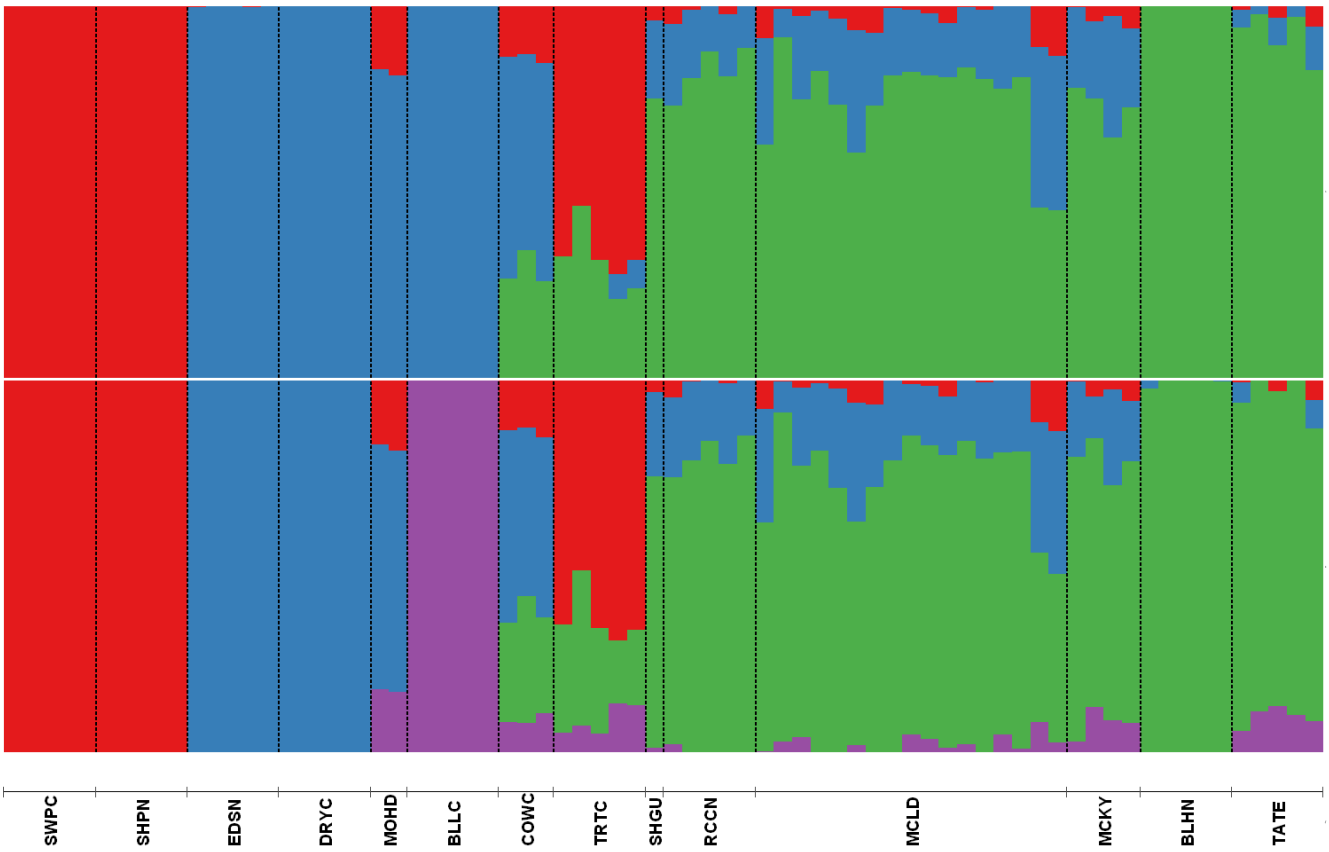


Table 1.

Group	Location	Watershed	Sample Year	Watershed code	N Samples	N Samples removed
Rainbow Trout Group (RBTG)	Lassen Cr (LASN)	Goose Lk	2004	GOLE	5	
	Cottonwood Cr (CTWD)	Goose Lk	2004	GOLE	5	
	Buck Cr (BUCK)	Goose Lk	2009	GOLE	5	
	Davis Cr-Goose Lk (DVGL)	Goose Lk	2009	GOLE	5	
	Willow Cr (WILW)	Goose Lk	2009	GOLE	5	
	Upper Pine Cr (UPNE)	Goose Lk	2003	GOLE	5	
	Lost Cr (LOST)	Lower Pit R	2009	LPIT	5	3
	East Fork Nelson Cr (NLSN)	Lower Pit R	2004	LPIT	5	
	Davis Creek (Pit R)	Lower Pit R	2008	LPIT	5	
	Joseph Cr (JSPH)	NF Pit R	2009	NFPT	5	
	Spring Canyon Cr (SPRC)	NF Pit R	2009	NFPT	5	
	Parker Cr (PARK)	NF Pit R	2003	NFPT	5	
	Shields Cr (SHLD)	NF Pit R	2009	NFPT	5	
	Couch Cr (COUC)	NF Pit R	2003	NFPT	5	1
	N Fork Pit R (NFPT)	NF Pit R	2009	NFPT	5	
	East Cr (EAST)	NF Pit R	2009	NFPT	5	
	Franklin Cr (FRKN)	NF Pit R	2009	NFPT	5	
	Thoms Cr (THMS)	NF Pit R	2009	NFPT	4	
	Parsnip Cr (PSNP)	SF Pit R	2009	SFPT	5	
	Fitzhugh Cr (FTZH)	SF Pit R	2003	SFPT	5	
	Emerson Cr (EMRN)	Surprise V	2009	SRPV	5	
	Mill Cr (MILL)	Surprise V	2009	SRPV	5	
	Cedar Cr (CEDR)	Surprise V	2003	SRPV	5	
	Bidwell Cr (BDWL)	Surprise V	2009	SRPV	5	
	Deep Cr (DEEP)	Surprise V	2009	SRPV	5	
	Dismal Cr (DISM)	Warner V	2002	WARV	5	
	Upper 12 Mile Cr (UTMC)	Warner V	2002	WARV	5	
	Lincoln Cr (LCLN)	Yuba N Fork	2007	YUBA	3	
	Hot Creek Strain (HTCS)	Hatchery	2011	HTCS	5	
	Mt. Whitney Strain (MTWS)	Hatchery	2002	MWTS	2	2
	Mt. Shasta Strain (MTSH)	Hatchery	2002	MTSH	4	2
	Coleman Strain (COLE)	Hatchery	2011	COLE	5	
	Pit Strain (PITS)	Hatchery	2004	PITS	5	
	Eagle Lake (EGLH)	Hatchery	2004	EGLH	5	
	N Fork American R (NFAR)	American R	2006	NFAR	5	
	Eagle Lk (EGLK)	Eagle Lk	2009	EGLK	5	
	Lower Yuba R (LYBA)	Yuba	2009	LYBA	2	
	Battle Creek CNFH	Battle Cr	2009	BATC	1	1
	Lower Stanislaus R (LSTN)	Stanislaus R	2009	LSTN	1	
Total					177	
Upper McCloud River Redband Trout (MRRT)	Sheepheaven Cr (SHPN)	Upper McCloud	2002	UMCR	5	
	Bull Cr (BLLC)	Upper McCloud	2008	UMCR	5	
	Dry Cr (DRYC)	Upper McCloud	2008	UMCR	5	
	McKay Cr (MCKY)	Upper McCloud	2007	UMCR	4	
	Raccoon Cr (RCCN)	Upper McCloud	2007	UMCR	5	

	Swamp Cr (SWPC)	Upper McCloud	2007	UMCR	5	
	Tate Cr (TATE)	Upper McCloud	2008	UMCR	5	
	Blue Heron Cr (BLHN)	Upper McCloud	2008	UMCR	5	
	Trout Cr (TRTC)	Upper McCloud	2007	UMCR	5	
	Shady Gulch Cr (SHGU)	Upper McCloud	2007	UMCR	1	
	Edson Cr (EDSN)	Upper McCloud	2007	UMCR	5	
	Cow Cr (COWC)	Upper McCloud	2007	UMCR	4	1
	Moosehead Cr (MOHD)	Upper McCloud	2016	UMCR	3	1
	McCloud River (MCLD)	Upper McCloud	2007	UMCR	17	
Total					74	2
Golden Trout Group (GTCX)	Volcano Cr Left String (VCLS)	Golden Trout Cr	2005	GTCR	2	
	Salt Lick Cr (STLC)	Golden Trout Cr	2005	GTCR	2	
	Groundhog Cr (GDHG)	Golden Trout Cr	2005	GTCR	2	
	Little Whitney Cr (LTWY)	Golden Trout Cr	2005	GTCR	2	
	Mouth Barrigan (MBNS)	Golden Trout Cr	2005	GTCR	2	
	Upper SFork Kern R (SFKR)	S Fork Kern R	2006	SFKR	2	
	Mulkey Creek (MLKY)	S Fork Kern R	2006	SFKR	2	
	Above Shaefer (ASHF)	S Fork Kern R	2006	SFKR	2	
	Above Ramshaw Barrier (RmsB)	S Fork Kern R	2006	SFKR	2	
	Above Templeton Barrier (TMPB)	SFork Kern R	2006	SFKR	2	
	SFK Below Snake Creek (BSNK)	S Fork Kern R	2006	SFKR	2	
	Wind R (WIND)	Wyoming	2005	WYOM	2	2
	Upper N Fork Clicks Cr (UNFC)	Little Kern R	2005	KRNR	2	2
	Fish Cr (FISH)	Little Kern R	2011	KRNR	2	
	Silver Lake (SILV)	Little Kern R	2007	KRNR	2	
	Rifle Cr (RIFL)	Little Kern R	2006	KRNR	2	
	Sheep Cr (SHPC)	Little Kern R	2001	KRNR	2	1
	Lion Cr (LION)	Little Kern R	2006	KRNR	2	
	Upper Willow Cr (UWLC)	Little Kern R	2011	KRNR	2	
	Tamarack Cr (SMNT)	Little Kern R	2006	KRNR	2	
	S Mountaineer Cr (SMNT)	Little Kern R	2006	KRNR	2	
	Upper Wet Meadow Cr (UWMC)	Little Kern R	2002	KRNR	1	
	Upper Soda Spring Cr (USSC)	Little Kern R	2002	KRNR	2	1
	Nine Lakes North (NLKN)	Kern R	2007	KRNR	2	1
	Chagoopa Cr (CHGC)	Kern R	2007	KRNR	2	
	Chagoopa Lake (CHGL)	Kern R	2006	KRNR	2	
	Kern-Kaweah Cr (KKWC)	Kern R	2007	KRNR	2	
	Upper Chagoopa Cr (UCHG)	Kern R	2007	KRNR	2	
	Picket Cr (PIKT)	Kern R	2007	KRNR	2	1
Total					57	
Total All Samples					308	19

650

Table 2.

Group	N individuals	N loci differentiating MRRT from RBTG		N Polymorphic Loci
		Cutoff 90%	Cutoff 99%	P-value cutoff for HWE 0.1
$MRRT_A$	72	46	-	6639
$MRRT_p$	20	2,649	574	731

Table 3.

Location	N	Minimum RBT %	Maximum RBT %	Mean
Swamp Creek	5	0.0000	0.0000	0.0000
Sheepheaven Creek	5	0.0000	0.0000	0.0000
Edson Creek	5	0.0000	0.0000	0.0000
Dry Creek	5	0.0000	0.0000	0.0000
Moosehead Creek	2	0.0000	0.0000	0.0000
Bull Creek	5	0.0000	0.0331	0.0205
Cow Creek	3	0.0356	0.0758	0.0547
Trout Creek	5	0.0415	0.1640	0.0980
Shady Gulch Creek	1	0.1910	0.1910	0.1910
Raccoon Creek	5	0.1572	0.2283	0.1924
McCloud River	17	0.0456	0.3034	0.2035
McKay Creek	4	0.1726	0.2507	0.2107
Blue Hearon Creek	5	0.2457	0.3016	0.2709
Tate Creek	5	0.3003	0.3678	0.3272

Table 4.

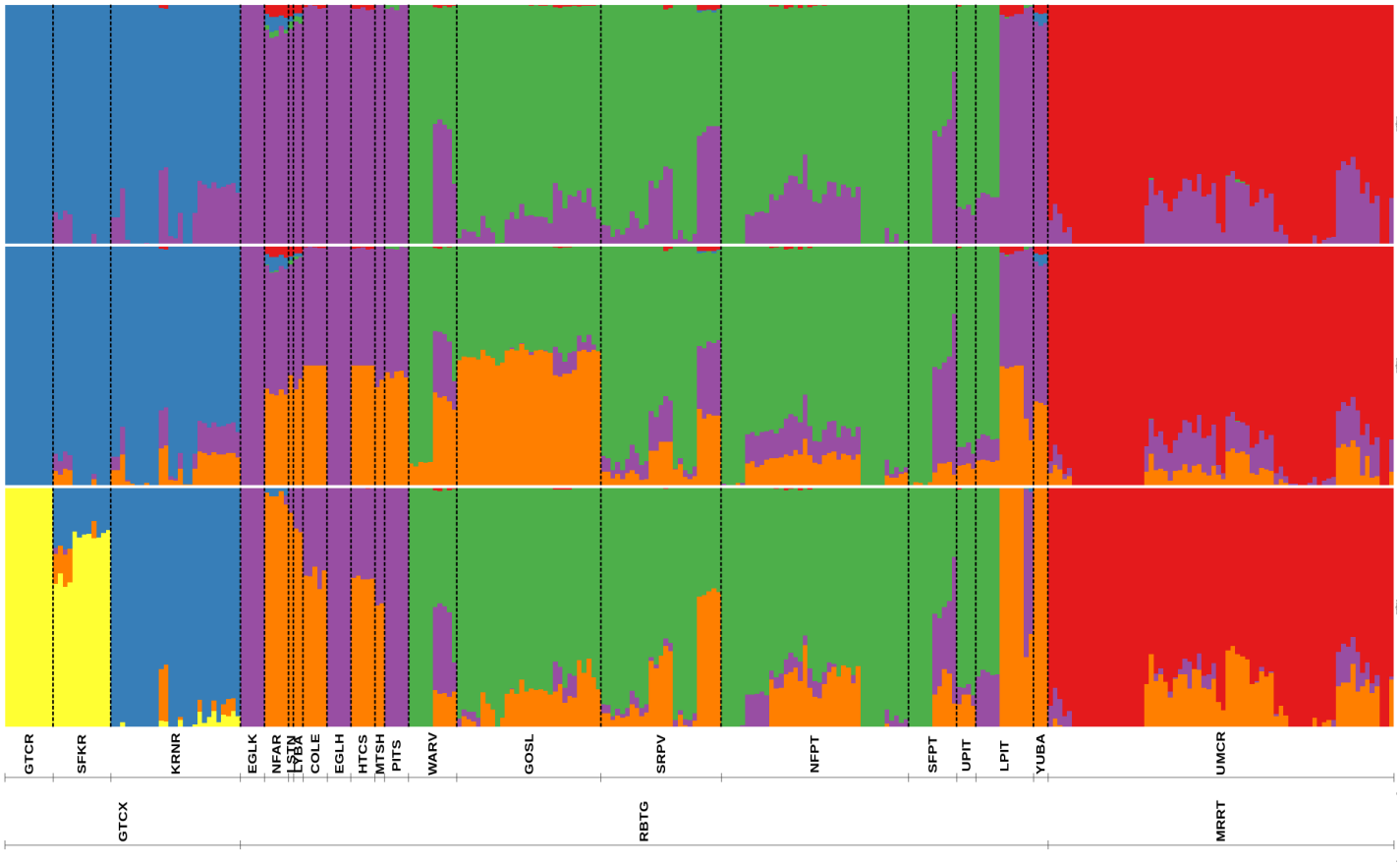
Location	$\Theta\pi$
Dry Creek	0.000699
Edson Creek	0.000474
Moosehead Creek	0.000691
Sheepheaven Creek	0.000592
Swamp Creek	0.000561

667 **Table 5.**
668

	McCloud River Redband Trou				Hatchery Rainbow Trout			
Locus	**Swamp (n = 16)**	**Sheepheaven (n = 20)**	**Edson (n = 54)**	**Overall (n = 90)**	**Crystal: Coleman (n = 20)**	**Crystal: Pit (n = 17)**	**Darrah: Eagle Lake (n = 50)**	**Overall (n = 87)**
omy01_36537055	1.00	1.00	1.00	1.00	0.00*	0.00	0.00	0.00
omy01_36542230	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
omy01_75717390	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
omy04_23890713	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
omy04_57267157	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
omy04_74937939	1.00	1.00	1.00	1.00	0.08	0.00	**0.09**	**0.07**
omy06_1274949	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
omy07_12066898	1.00	1.00	0.98	0.99	0.00	**0.35**	0.02	**0.08**
omy07_12545308	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
omy07_13121873	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
omy07_1624310	1.00	1.00	1.00	1.00	0.05	0.00	**0.15**	**0.10**
omy07_9878739	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
omy11_65719977	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
omy14_22294317	1.00	1.00	1.00	1.00	0.05	0.00	0.00	0.01
omy14_22294535	1.00	1.00	0.98	0.99	0.03	0.00	0.00	0.01
omy15_57867903	1.00	1.00	**0.80***	**0.89***	0.03	0.00	0.00	0.01
omy15_58335356	1.00	1.00	1.00	1.00	**0.15**	0.00	0.00	0.03
omy18_32367886	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
omy18_57370236	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
omy21_13749597	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
omy24_12391438	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
omy24_13095037	1.00	1.00	1.00	1.00	0.00	0.03	0.00	0.01
omy24_16290480	1.00	1.00	1.00	1.00	0.05	0.00	0.00	0.01
omy24_19143772	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
omy24_26165023	1.00	1.00	1.00	1.00	0.05	0.00	0.00	0.01
omy26_12749297	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
omy27_9625126	1.00	1.00	1.00	1.00	**0.08**	0.00	0.05	0

669

Figure S1.


GTCR
SFKR
KRNR
EGLK
NFAR
LSTN
LYBA
COLE
EGLH
HTCS
MTSH
PITS
WARV
GOSL
SRPV
NFPT
SFPT
UPIT
LPIT
YUBA
UMCR
GTCX
RBTG
MRRT

674

675 **Table S1.**

SNP Assay Name	Sequence	Chromosome	Position
Omy01_36537055	GCTATATGCTACCTGCATCTATACTGAACAAAAATATAAACGAGTTCATATAATTTTAATTAATTAATTAGGCCCTAATCTATGAGTTTCACATGACTG GGCATGGATGCAGTCATGGGTGAGCCTGGGAGGGCATAGGCCCATCCACATGGCAGCCAAGCCTACCCACTGGGGAGCCAGGCCCAGCTAATCAGAA TGAG[T/ G]TTTTTTTGCCACAAAAGGGGCTTTATTATAGACATAAATACTTCTCAGCANCCATTAAATTNTCTGGCAATAGCTCTGGTAGACATTACTGCAGTCAT TCCTGCAGGCAGCATGTCAACGCTCTCTCTAAACTTGAGACATATGTTGCTTTGTGTTGTGTGACAAAACTGCACATTTTAAAGTGGCCTTTTATTGTCC CC	Omy01	36537055
Omy01_36542230	GCGCACAATTGGCCCAGCATCGTTCAAGTAAGGGAAGGGTATGGCCGTGAGGGCTTTACCAGGCTCATTGCACTCTAGCGACTCCTTGTTGGGCCGGG CGCCTGCAGGCTGACTTCGGTCATCAGTTGACAGGGTAAGAGCGTCCAGTGTGTGTGCTCTGAACGCCCCCAGGGAGAAACACTCTGAATTTAAAAAT GGAC[A/ G]ATCTGACTCTGGCACTCCAGATTGAATTTAAGAACACACANAAANTCTTACAATGTCTAGCTAGGAATTTGTTATGCTAGCAAGAGGCTGCATAGTA ACAGCACCAACTTCAGGTAGAGAGGCGAAGCGTTAGTACGCTCAACTGAAAGGATACCGTTTACAGCGGGATGGGCAATTCCAGTCCTCGAGGGCCT GACTGG	Omy01	36542230
Omy01_75717390	ATACTTTCCGAAGGCACTTGGCATACATCTCACTTAAGTATTGAAAGTGCAGTGTCACACATCTGATGTACGGATGTTGACATTGCTAGTGATATCAAA TGCATTGAATTTCTCATACATAGCACATGCTGTGTCTCCTCTGTGTGGAGAGATATGCTTGACTTTTAATGATGGCAGTGGCACTGTGACATGACATTG GA[A/ G]AAACAGCTGTTCATATCAACCCTTAATGGGAAACTCAATTAGGTTAGGGGGCTATGATTTCTTAATCCAGCCCTGCAGGTGCTCACATCTGTAGCAT TTTGGAAGGAGAGGGGACATTTGGGTCATAGCCTTGCCTGAAACTTGCAGAGAGTTTCCGTATAGGGGGGGGTNGAGATTTTACAGTTCGAGTTAGTA TCTTT	Omy01	75717390
Omy04_23890713	TGACCGAGTTGGTATGTCTCGTTGTAGGNGTGTGACCGAGTTGGTATGTCTCGTTGTAGGGGTGTGACCGAGTTGGTATGTCTCGTTGTAGGGGTGTGA TCGAGTTGGTATGTCCCTTCCTGCAGGTGTACCCTCTGTGGTTGTGTCCGTTCATGCTCCCCCCGGGCCGGGGGATGGTTCATCCTAAAGGTGANGAGG CA[G/ C]AGCTGTATGTGGACATCGGAGCGTATGGAGAACCCGGAGTTAAACACTTCCAGGCTAAAGCCTCGTGTAGACAGCTGGAGCAGTTTGTCAGAGAGN TTCACGGGTGAGTTTCTTTACACACACACACACACGCACCCCAACAGGGGCTATAGAGAGTAGGGAGTTAGGGTTCAGATTGTGATGATGTGTGTG TGTCTC	Omy04	23890713
Omy04_57267157	AGGAACCCTTTTGTCTTCAATAATCCAAGCCAGATGGGTCATTACATGCCATTTGGGAATATCTAAGTATATCTAAATTGCACGATTGTGTGCATATTT TCCCTGATATTCTAGGAAAAGTATTGGCCTACTTCCTACCTCATGGCTAAATGTAACTTCATGTTTATCTCTTGTTCCCATGGGGTCAAACTAATGCCCG T[T/ A]GGCATATGTAGTTGACTATGTGTGTGGCTCAAATGAACGCCTCTCCTCTACCTTCTTAGGGTCCTACCAAAGTCAGCTCGGTGGCTGATGGCCAACC AGAGGTACGAGGAGGCCATGGATCTGGTCAGGAAAGCAGCTCTGATGAACGGCAAACCCCTGCGGGAGGACATAGAGATATGTCAGGTGTGGTTTGT TGGTCA	Omy04	57267157
Omy04_74937939	ACTCATTTTTTTCATGATGAGGTTGATCATATCCTTCCATTTGTATGATGCCTGGTGGCCAGCCTTATCATAGACTTCCTGCAGGTGTTCCACATAGACT TCACGCTTGCTGTCTTCCAGCCAGAGATCAACTCAGTGAGTGTGTTACTCCCAGTGGATGGGGCTGGCCTGATATTTGGTCAAATAAAATACTTGCAA AT[G/ T]TCAGTGCTGCATTCACTGTGTGTGTGTTTTATTTTCAGCTAAATGGCCTGGACAGCTGTGAGCAGGTGTCTCATGATCTGGGTATCACTGAGACCGAT ATGAATAGGACCACTCCTCTGCTGCTTGAGCTTGTCAAGAGAGGCAGACACAGGGAGAAGGCCTCCATCTTCTCAGCGGTAAAGACACACCAGTACA CAGGG	Omy04	74937939
Omy06_1274949	TGTTGGAAGGTGAACCTTCACCCCAGTCTGAGGTCCTGAGGGCTCTGGAGCAGGGTTTCATCAAGGATCTCTCTGTATTTTGCTCCGTTCATCTTTCTCT CGATCCTGACTAGACTCCCTTTCCCTGCCGCTGTAAAAACATCCCCACAGGATGATGCTGCCACCANCATGCTTTAGCATAAGGATGCTGCCAGGTTTA G[A/ G]CATGTAGCTAGGTATNAACCGGCATAATCCCAACTCATACTACTACCAATACAAACTGATTGTCATAGCTGTAGTATGAACCTGCAGGTAGCTAAA GCTAACCAACTAGGTTCAATGTTAGCTAATTAGTCAACATGAGGCTTTAACTAGCAATGCAAATGGCTCTGAGATATGAATAATATTACTACACAGAT CATACA	Omy06	1274949
Omy07_12066898	AGAGTTAGAATGGAGAGGACTACCACTGATCCCACAAAACCATCTTAACAGAGACTTGCTCATGAATTGATGTTGAAGTCTCGTTAGATCAACATGAT GAGAAAAGCGGTCAATAAATCCTAGAACGCCAACTGTCTTTCTCTGTGCCGGGACCACTATAGGAGGACGGAGTGGTGTTCCGTGTCTTCTGTTCAGT TTGA[C/ G]AAATGAAACAGCATCTCCGGCCTGCAGCAGGTTTCTGTCTCTATCTATTAATAGCCTGGTGTGGGGACAGGATCTCCTGCAGGCTGCAGAACACTGA TGTAAGAGCGGCAGAGAGAGCTTCCCGCCAAAGATTTAATCTCCCCACAAATGCTAACCTTTTGCCATAGGTGTGTATATCCCAGGATCTAACCCTGG AAAGA	Omy07	12066898
Omy07_12545308	AGGCAGTCTTCGACCACCCACCCGAGGGTAGAGAGGCGAGTGAATGTCTTTTCCATCTGAGGCAGGGGACGAGGAGCTCCCAGGAGTTCGCAACCTG GCTGGCGGCGAGNGATGGAACGAAAGGGCCCTGATCGACCACTATCACTGCAGTTTGTTTGAGGACGTCTGCCGAGAGTTGACCTGCAGGGACACCA ACCTCA[C/ T]CTTCGACCAGCTGGTGGACCTGTCCATTCGGTTGGAATACCTGCTGGCTACCCGCGGATGTCCAGAGCGGGGTCTGTCGGTTCCATCCNACAGCACC ACCGCTCTGATGCCCATGGAGCTGGGAGGTGCTGCGCTCAGGGGGACCGGAGCAGGTTCCGTCTCATGCACCATCTGTGGCTGCAGAGGTCACACTGC CGGTC	Omy07	12545308
Omy07_13121873	GTTTCACAAATAGGCAGGGAAACTGACACAAAGAATACAGCTGACACTACAAGGGACAAGGACACTAAGGTATCAACAGAAAGTTTTCCTAAAGAG NACGGTCTTTAGGCTTGTTTTAAATGAGCCTAGAGTCTTTGGGGCCTGCAGGTGGTCCGGGAGGGCATTCCAGAGGCTGGGAGGCTTGTTGCATTTTTT AAAAT[T/ G]ATCTTTTGGGGGTTGACTGCTCTGGGGTTTGATCACTTACATTGTGTGGTCTCGACTCNCCGAATTGCAGAATTGGAATGAAAAGTCTACCTTTTTGG CTTATAACTAGCTAGCTATCTGGCTGTTTGGCTGTTTGGATTTAAAATAATGCAGCCCCAGCTTTTTTCTTTCATCTGTCAGGAAGAATAACACTCAAG GAC	Omy07	13121873
Omy07_9878739	CAAGGGCAGATCGACATATTCTTCACCTAGTCGGCTCAGGGATTCAACCTTTCGGTTGCTGGACCAACTCTTAACTGCTAGGCTACCTGCCGCCCCAAA GCTTTAAATAAGGTCTATAGTCTGTACTTCCTGATGTAAAGTCGAGCGCAGGCCTTTTCATCTTTAACCCAGCTGCAGACTGGCTTCATAGAGCGATAA AT[A/ G]AAGGGGATCCTCCCCACTGGCGTTGTGANACTCAGGACCCCTCCATATCAGGACCACCAGAACCAATGAGAGTTCTCCTGGTGCAGCCTGGCCTGC AGGCCGGGATTCAGGAGCACTGAGGTGTTGTAGCTCCCAGGGGCCTCGCCCCTGACCACGGTGTAGTACAGCCTCTCTTGATCTTATGAAGCACCACT AGTCCA	Omy07	9878739
Omy11_65719977	NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNTTCCTCCCC AACACCTTGGAGGCCTGCAGGAGTATACCACAANTAAAGAATTGCTGTTTTGTGACCCACGCAAAGGGAAAAGGGGCACTGGGGGCAGGCATGGCA GGTGTCTGCTG[G/ A]GCTTATTGAATCTCTCTTCTCTTCNATAGTTCAAACAGTCAGGCCAAGATGGCTGAGTGCCTGGAATCTATCCTCAATCCTCGTAATGCAAAGAAAA CATCCGTCAAGCGACAACCCCAGAATAAAACAGCTTTGGAGTGTCAGCACTTGACGTCTATTCCAAGTAAACAAAGGAAGAAGAAAGATTAACAACA ACAACA	Omy11	65719977
Omy14_22294317	CAACACTTTTGATGAACAGGGCAAAATGGATATATGCCTATAACAGTTTGGATCAGCTTGATCTCCCCCTTTAAATAAAGGACACACCGTGGCTGCCT TCCAAGCAATGGGAACCACGTCAGAGAGGAGAGACAGGTTAAAAAGGTCAGAGATAGGCTTGGAGATGATATGGGCTGCAACCCTAAAGAAGAAAG GGACTT[T/ G]TTTTGTGGTCTAGTTTAAGGAGCTCCTTTAGCACCTCAGACTCAGTGACAGCCTGCAGGGAGAAACTTTGTAGCAGGGGAATAGAGGGAGGAGGAT CGGGGCTAGTCGCATTAGAAGGGGTGGGAGATCAGGAAATGTTGGACGGGCTATGAGGCATGGCTGTGTCAAATAGGAATTCGNACTTAATGAAGTG GTGATTA	Omy14	22294317
Omy14_22294535	AGGAGCTCCTTTAGCACCTCAGACTCAGTGACAGCCTGCAGGGAGAAACTTTGTAGCAGGGGAATAGAGGGAGGAGGATCGGGGCTAGTCGCATTAG AAGGGGTGGGAGATCAGGAAATGTTGGACGGGCTATGAGGCATGGCTGTGTCAAATAGGAATTCGNACTTAATGAAGTGGTGATTAAAGAGCTCAGC	Omy14	22294535

	CATTTG[C/ T]TTCTTGTCAGTAACAACCACATCATCAACATTAAGGGACATGGGCAGCTGTGAGGAGGGTTTATTCTCCAGGTCTTTAACTGTTTTCTAGAACTTCTT GGGGTTAGACCCACAGAGAGAGAACTGCTCCTTAAAGTAACTAACTTTGGCCTTCCGGATAGCCTGAGTGCACGTATTTCTCATTTGCATGAACAAGA GCCT		
Omy15_58335356	TGGCCCTGCCTGCCTCCTGTGCTTATTGTCAGGGTGCAGAGTGCCTGGNTAGAGTAGCCTGACATTGGTTGGCATGGGGAGGCATGCCCTCTCTACCCT GCAGGGGGTCTCTAAAAAGCACATATCCCACACAAATGTCACTGTGAATCTGATCGCCATTCTATAGACACACACACACACACTGCTAGATCCCCAGC CAG[A/ T]NGCCTGACACATCTCAGGGAAAGGCAAATAGAAATGCCATGATCACAACTCTGATGTAACTGCATCACAGTTACACCGTAAAAAATCTAGCTGTTT CAACTAACTATTTTTAAGTAACTGGTTCCACAGCGTTATTTTTGTTCACTCAACTTTTCAGTCCAAGTTTTACCTGAACAACAAAAAAATTACTTGGCCC AAAC	Omy15	58335356
Omy18_32367886	TAAGTTGACTATGCAAACAATTAGGAAGTTCCAACACATTAATGTTTCTTATAAAAACCAGGAAGTCACGCAGTCAGTCGTTCCTCAACTCTCAGCCA AGAGAGACTGGCATGCGTAGTATTTATATTAACCCTCTGATTACAATGAAGAGCAAGACGTTCTGGGCCAGCTGGAGCTTAACTATGTCTTTCTTTGTT GCA[C/ G]TTGACCATATGACTAGACAATAATCACGTTAAGACAAAACTAGAGCCTGCAGGACTTTTTGGAGTGTAGTGTCAAAAAAGCAGAGCACCTCTTAAT TACGGACAGACCTCTCCCCATCTTTACAACTACTGAATCTATATGGTACAGGTGTGGATACAGGAATATGGTTATGCTGTTTCATAACCTATTTTCATA ACTTA	Omy18	32367886
Omy18_57370236	TTTTAAATGAGAGATGTGTTGACCCTTCCTGCAGGTTGACTCCAAGCAGCTGTCTGGCAGAACCCTGTTGTCTGTCCCTGGATACCAGAACAGGGTGG AGTTTGGGACCATGGTCACCTTTAGCTACCCTATAGACGGCCAGGGTGAGTGGTTGGGTTGAGTCCCAAATGGCACCCTATTTAGGGCTCTGGACAAA AGTA[G/ A]TGCACTGTGTAGGGAATAGGGTGCCATTTGAGATGTAATCTCAGTATCTGAGTTGTGATCTTCCTGTATGTCTGCTGTTCTCTGGGATTTTTCCCATC TTTCCTGTTTATACTCTCTCTCTCTCTCTCTTTATTAGATGGAGAGGTGGCAGTGTCGACCACTCGTCCTGAGACCATGTTGGGTGATGTGGCCATCG CT	Omy18	57370236
Omy21_13749597	ACCTGTTCTCCTGCAGGTTTCCCAATTTTTGCCAGGACTCTCAGCGGGCAGAGTGTTTTGGACCGTGCTGTTTGAGGAAATTACAAGGCTGCCAAATCA GCCGGGCTCTAAATCCATTGGAGATGAGAGGGAAAGCTGAAGCCAGGGAAGCCAGAGCGATGAGTATTGCCNTCAATTTCATCAGAGAGGACTCAAA AAGT[T/ G]GACTTTCAAGTGAAGGACATGAAAAAGAATGGGAGGAATATCCAACGACAGTGCACCTTATGTGTAGAATCCATGAGTTAACTAGCCAGTTCATGT TTAATTATGTCCATTTAGCGAATAGTTAAATGTTAGGGCTAGGCCCCTTTTTTCTCAATTTCCACCTGAATGATGTGCCCAAAGTAAACTGCCTGTTACT CAGG	Omy21	13749597
Omy24_12391438	GTTGTCTACAACTTAATTGAGAGAAACTAAATCCAGCAGATGAGGCCCTCCATTAAATGAGTGACACCCCTTGGAATAGACCACAAGTTAGCTTTTAC TAGAAATGCACTGTGATGTAATATAAACTCAGCAAAAAAAATGTCCTCACTGTCAANTGCATTTGTTTTCAGTAAACTTAACATGTTCATACAAATATT TAA[C/ A]TAAGACATGTTACTCCAAGGCACCTGCAAGTTCCCGGATACTTCTGGTGGGAATGGCCCTCGCCCTCTGNTCCAACAGGTCCCAGACGTGCTCAATA GGATTGAGATCCGGGCTCTTCACTGGCCATGGCAGAACACGCACAGAACGAGCAGTATGGCTGGTGGTATTGTCATGCTAAGATGAGCCTGCAGGAA GGGTAC	Omy24	12391438
Omy24_13095037	ATCAAGAAGCTGATTAAACAGCATAATCACTACACATGTGCACCTTGTGCTGGGGACCAGAAAAGGCCAACCTAAAATGTGCAGTTTTGACNTGCTGC CTGCAGGAATGTCCNCCAGAGCTGTTGCCAGAGAATTGAATGTTTTAGAGAATTTGGCAGTATGTCCAACAGGCCTTACAACCGCAGACCACATGTGA CCTC[G/ A]CCAACCCAGGACCTCCACATCTGTTTTCCTCACCTGTGGGATCGTCTGAGACCAGCCAACCAGACAGATGATGAAACTGAGGAGTATTTCTGTCTGT AATAAAGACCTTTTGTGTGGAAAAACTCATTCTGATTGTCTGGGCCTGTGCCTGGCTCCCCACTGGGTGGGCCTGGCTCTCAAGTGGGTGGGCCTATGC CCTC	Omy24	13095037
Omy24_16290480	AGTCCAGCTAAGGGACAATTATTAGAGGAGGGCTTTGCAGTAAGTGTTTTCTACTCACCACCAGTTTTAGCTGGTTTGTAGCCTGCTCCTGCAGGTATC CCACCAGCTCCTGTTGGTAAAACAAGAAAAGCACAAGAAAACAGAAGAGAAGCAGAAAAAAAGAGTTAGATTGAAAGGTTTGACGTTTGTTTTAGCT TTCC[C/ A]CCTGGCTGCTCCCTCCCTCACTTCGGTCTCTGACCTGACCAGATATCTCCACACAGAGACTAGACCACTGTGGTGGGTAGGAGGCTCTAAATCTCTA CATCTTTAAATCTTCTAAGTTTGAGTCCCAAATGGCACCCTATTCCCTATATAGTGTACCACTTTTGACCAGGACCCTAGTCAGAAGCAGTGCACCATG TCTG	Omy24	16290480
Omy24_19143772	TTGATCTCCCCCTTTAAATAAAGGACGAACCGTGGCTTCCTTAAAAGCAATGAGAACTTCCCCAGATTGGAGAGACAGGTTAAAAAGATTGGAGATA GGCTTGGCGATNATAGGGGCAGCAACCAATAAGAAGAAAGGGTCTAAACCAACTGACCCAGATGNTTTTTTGGTGTCAAGTTTAAGGAGCACCTTTA GCACAT[C/ T]GGACTCAGTGACTGCCTGCAGGAAGAAAATTTGACCAGTTTGGCCGGNCAGACAGCTATAGTAAGAGTCTTGGTGGTTCCAAACTTCTTCCATTTAA GAATGATGGAGGACACATGTTCTTGTGAACCTTCAATGCTGCAGAAATGTTTTGGTACACTTCCCCAGATCTGTGCCTCGACACAATTCTGTCTCGGAG CTCT	Omy24	19143772
Omy24_26165023	GCAATTAAAAAGGTTATTTATTGTTAACAAACTATTAACTTTAAACTAAGAAAAATGAATGAGGTGTGGAAGTATCATAATGTAAGGTGTATGTAAAG TGCATGGATGCATGAATCTGTGTGTGTGAATGACTGAGTGAAAACTATGCAAAACTACAACAGAACAAACAAAACAGGTTCATACCTGGAGGAGNGG AGAGA[G/ A]AGAGAGAGAAGTGATTAGTGACACAGTTTAAATACTTAGAGCCCAGGTGACTCCAATCACTAACGACCCTCCTCTGCCTGCAGGAGGAACCGCCCC TGCACTGCAGAGGAGGAGCTGTGACAGCTGTAACATAACAACATTTGGAAAATGTCAAGGNGTCTCAATATTTTCCGAAGGCACTGTATATATTTTCT CACTAT	Omy24	26165023
Omy26_12749297	AGTGCTCACCGGGAGGCTTGGCGGGCAGAGGCCTGCAGGTCGGCTGTCACGTGCAGGAAGTAGCAGCTGAGGAACACTCTCCTCTGCAGCAGGAGGA AGAAGAAGCAGATACTGTCCCAGATGATCCCTGCCTCCTCTACAGGCAGAGTACACGCCTTATCGCTCACACTCTTAGCTGTAGAACACACACACACA AGCAT[A/ T]CAAAGTCNGCTTTGGTTATTCTATAATATCATTGTTTAACCTGCCAATTTGGAGGGTGATGTGTGTTCACTCATAGAGATAAAAGAACTGTATCCATC TGTATGTGTTGAGTACTCACGGTCATAGTAACCCTTCACTGTGCAGACCAGACTGAAGAGTTGGATCACCCAGCAGAAACCTCTCTGCATCTCCACCA CAAA	Omy26	12749297
Omy27_9625126	CACTGTTATTTCCTGTGCCTGTCGCATTGAGTCTGGTCTCTCTTGCACTCCACTGGCCCGGAATGTTCTCCCTTCATTATCAGCCTGACCCAGCCCTGCA GGACACCCCTTCTTCANCTCCACGTAACTCCCCCACCCAGTTCCAGTCCCAGCCCCACAGACGAGGGGGGAGACAGACAGGCCAACAGACTCCCTCGC CA[C/ T]GGTTTATTGACCCAGTTTGCGGCTCAGGTCGCATGGCGGGAGTTGACCAAGAATGGCGCCTAGCGAGGACCACCAACTGAGGGAAGGGATGTGGA GCACCCTGGTGGAGCACCATCCTCTTGCATAATGCTTTATTATTTACCCATATTAATATTGGTCATTGTTATTAATCGAAGGGTAGCCTACCATTTAGAC ATATA	Omy27	9625126

676

677
678 **Table S2.**
679

Fluidigm Assay ID	SNP Assay Name	ASP1 SNP	ASP2 SNP	AMP GC	Primer sequence (5' to 3')	
GTA0251450	omy01_36537055	T	G	0.39	ASP1:	GGCCCAGCTAATCAGAATGAGT
					ASP2:	GGCCCAGCTAATCAGAATGAGG
					LSP:	LSP: TGTCTATAATAAAGCCCCTTTTGTGGCA
					STA:	STA: GCCAAGCCTACCCACTG
GTA0251401	omy01_36542230	A	G	0.48	ASP1:	CTGGAGTGCCAGAGTCAGATT
					ASP2:	CTGGAGTGCCAGAGTCAGATC
					LSP:	ACGCCCCCAGGGAGAAA
					STA:	TGTGTGTTCTTAAATTCAATCTGGAG
GTA0251420	omy01_75717390	A	G	0.44	ASP1:	CATTAAGGGTTGATATGAACAGCTGTTTT
					ASP2:	ATTAAGGGTTGATATGAACAGCTGTTTC
					LSP:	GGCAGTGGCACTGTGACAT
					STA:	CCCCTAACCTAATTGAGTTTCCC
GTA0251458	omy04_23890713	G	C	0.59	ASP1:	TCCGATGTCCACATACAGCTC
					ASP2:	TCCGATGTCCACATACAGCTG
					LSP:	ATGCTCCCCCCGGGC
					STA:	CCGGGTTCTCCATACGCT
GTA0254186	omy04_57267157	T	A	0.46	ASP1:	ACACACATAGTCAACTACATATGCCA
					ASP2:	ACACACATAGTCAACTACATATGCCT
					LSP:	CCATGGGGTCAAACTAATGCCC
					STA:	GAGGCGTTCATTTGAGCCA
GTA0251409	omy04_74937939	G	T	0.42	ASP1:	CACAGTGAATGCAGCACTGAC
					ASP2:	CACAGTGAATGCAGCACTGAA
					LSP:	GGCTGGCCTGATATTTGGTCAA
					STA:	GGCCATTTAGCTGAAAATAAAACACA
GTA0251392	omy06_1274949	A	G	0.40	ASP1:	ATAAGGATGCTGCCAGGTTTAGA
					ASP2:	AGGATGCTGCCAGGTTTAGG
					LSP:	ACAGCTATGACAATCAGTTTGTATTGGTAGT
					STA:	CATGCTTTAGCATAAGGATGCTG
GTA0251403	omy07_12066898	C	G	0.50	ASP1:	CCGTGTCTTCTGTTCAGTTTGAC
					ASP2:	TCCGTGTCTTCTGTTCAGTTTGAG
					LSP:	TGCTGCAGGCCGGAGA
					STA:	ACCACTATAGGAGGACGGAGT
GTA0251422	omy07_12545308	C	T	0.62	ASP1:	GCAGGGACACCAACCTCAC
					ASP2:	GCAGGGACACCAACCTCAT
					LSP:	CCGAATGGACAGGTCCACCA
					STA:	GGACGTCTGCCGAGAGT
GTA0251400	omy07_13121873	T	G	0.44	ASP1:	AGCAGTCAACCCCCAAAAGATA
					ASP2:	GCAGTCAACCCCCAAAAGATC
					LSP:	CCAGAGGCTGGGAGGCTT
					STA:	CAATGTAAGTGATCAAACCCCAGA
GTA0251417	omy07_9878739	A	G	0.51	ASP1:	CAGACTGGCTTCATAGAGCGATAAATA
					ASP2:	AGACTGGCTTCATAGAGCGATAAATG
					LSP:	TCACAACGCCAGTGGGGA
					STA:	GCCTTTTCATCTTTAACCCAGC

GTA0251414	omy11_65719977	G	A	0.47	ASP1:	TGGCAGGTGTCTGCTGG
					ASP2:	ATGGCAGGTGTCTGCTGA
					LSP:	GCCATCTTGGCCTGACTGTTT
					STA:	GCAGGCATGGCAGGT
GTA0251444	omy14_22294317	T	G	0.45	ASP1:	GGAGCTCCTTAAACTAGACCACAAAAA
					ASP2:	GAGCTCCTTAAACTAGACCACAAAAC
					LSP:	TGGGCTGCAACCCTAAAGAAGA
					STA:	TCACTGAGTCTGAGGTGCTAAA
GTA0251452	omy14_22294535	C	T	0.43	ASP1:	TGATTAAAGAGCTCAGCCATTTGC
					ASP2:	GTGATTAAAGAGCTCAGCCATTTGT
					LSP:	ACAGCTGCCCATGTCCCT
					STA:	TGGCTGTGTCAAATAGGAATTCG
GTA0251388	omy15_58335356	A	T	0.56	ASP1:	TGCTAGATCCCCAGCCAGA
					ASP2:	GCTAGATCCCCAGCCAGT
					LSP:	TGCCTTTCCCTGAGATGTGTCA
					STA:	TCTATAGACACACACACACACACT
GTA0251464	omy18_32367886	C	G	0.42	ASP1:	ACGTGATTATTGTCTAGTCATATGGTCAAG
					ASP2:	ACGTGATTATTGTCTAGTCATATGGTCAAC
					LSP:	TGGGCCAGCTGGAGCTTA
					STA:	CTGCAGGCTCTAGTTTTGTCTT
GTA0251441	omy18_57370236	G	A	0.49	ASP1:	CCCTATTCCCTACACAGTGCAC
					ASP2:	CCCTATTCCCTACACAGTGCAT
					LSP:	TGAGTCCCAAATGGCACCCT
					STA:	AGATTACATCTCAAATGGCACCC
GTA0251454	omy21_13749597	T	G	0.41	ASP1:	CTTTTTCATGTCCTTCACTTGAAAGTCA
					ASP2:	TTTTTCATGTCCTTCACTTGAAAGTCC
					LSP:	GCCAGAGCGATGAGTATTGCC
					STA:	GTCGTTGGATATTCCTCCCATTC
GTA0251396	omy24_12391438	C	A	0.31	ASP1:	GGTGCCTTGGAGTAACATGTCTTAG
					ASP2:	GGTGCCTTGGAGTAACATGTCTTAT
					LSP:	TGCATTTGTTTTCAGTAAACTTAACATGTTCAT
					STA:	GAAGTATCCGGGAACTTGCAG
GTA0251462	omy24_13095037	G	A	0.58	ASP1:	GCAGACCACATGTGACCTCG
					ASP2:	GCAGACCACATGTGACCTCA
					LSP:	GACGATCCCACAGGTGAGGA
					STA:	CCAACAGGCCTTACAACCG
GTA0251437	omy24_16290480	C	A	0.54	ASP1:	GTTTGACGTTTGTTTTAGCTTTCCC
					ASP2:	GGTTTGACGTTTGTTTTAGCTTTCCA
					LSP:	TCTGGTCAGGTCAGAGACCGA
					STA:	ACAAGAAAACAGAAGAGAAGCAGAA
GTA0251391	omy24_19143772	C	T	0.51	ASP1:	GTTTAAGGAGCACCTTTAGCACATC
					ASP2:	GTTTAAGGAGCACCTTTAGCACATT
					LSP:	TGCAGGCAGTCACTGAGTCC
					STA:	GTCTAAACCAACTGACCCAGATG
GTA0254169	omy24_26165023	G	A	0.46	ASP1:	CTGTGTCACTAATCACTTCTCTCTCTC
					ASP2:	CTGTGTCACTAATCACTTCTCTCTCTT
					LSP:	ACAAAACAGGTTCATACCTGGAGGA
					STA:	AGTCACCTGGGCTCTAAGTATTTA

GTA0251451	omy26_12749297	A	T	0.38	ASP1:	TGTAGAACACACACACACAAGCATA
					ASP2:	GTAGAACACACACACACAAGCATT
					LSP:	CACATCACCCTCCAAATTGGCA
					STA:	TCGCTCACACTCTTAGCTGT
GTA0251460	omy27_9625126	C	T	0.57	ASP1:	CGCAAACTGGGTCAATAAACCG
					ASP2:	CGCAAACTGGGTCAATAAACCA
					LSP:	GGAGACAGACAGGCCAACAGA
					STA:	CCATGCGACCTGAGCC

680