Application of a method for estimating effective population size and admixture using diagnostic single nucleotide polymorphisms (SNPs): implications for conservation of threatened Paiute cutthroat trout (*Oncorhynchus clarkii seleniris*) in Silver King Creek, California

Amanda J. Finger, Eric C. Anderson, Molly R. Stephens, and Bernard P. May

Abstract: The threatened Paiute cutthroat trout (*Oncorhynchus clarkii seleniris*, PCT) is endemic to Silver King Creek, California, USA, which was stocked with non-native trout beginning in 1930. Single nucleotide polymorphism (SNP) and microsatellite data reveal that the trout population in Silver King Creek is weakly structured and composed of introgressed California golden trout (*Oncorhynchus mykiss aguabonita*, CAGT), hatchery rainbow trout (*Oncorhynchus mykiss*, RT), and some native PCT. Two SNP groups were analyzed: (*i*) one mitochondrial and five autosomal SNPs, diagnostic between Lahontan cutthroat trout (*Oncorhynchus clarkii henshawi*) or PCT and CAGT or RT and (*ii*) one mitochondrial and five autosomal SNPs nearly diagnostic between CAGT and RT. The five autosomal cutthroat–rainbow SNPs were used to jointly estimate the cutthroat trout mixing proportion in Silver King Creek and effective population size (N_e) of the admixed population, using a coalescent-based maximum likelihood method. Given the stocking history of Silver King Creek, there are two different scenarios that bound the range of expected point estimates for N_e. We obtain point estimates of $N_e = 150$ and $N_e = 750$ for Silver King Creek under these two scenarios. This method will be useful in cases with differentiated taxa and in prioritizing conservation and restoration programs where the populations of concern are introgressed.

Résumé : La truite fardée de Paiute (*Oncorhynchus clarkii seleniris*, PCT), une forme menacée, est endémique à Silver King Creek, Californie, É.-U., qui a été empoissonné de truites non indigènes depuis 1930. Des analyses de polymorphisme mononucléotidique (SNP) et des données de microsatellites révèlent que la population de truites de Silver King Creek possède une structure faible et comprend des truites dorées de Californie (*Oncorhynchus mykiss aguabonita*, CAGT) introgressées, des truites arc-en-ciel (*Oncorhynchus mykiss*, RT) de pisciculture et quelques PCT indigènes. Nous avons analysé deux groupes de SNP : (*i*) un SNP mitochondrien et cinq SNP autosomaux diagnostiques entre la truite fardée de Lahontan ou PCT et CAGT ou RT et (*ii*) un SNP mitochondrien et cinq SNP autosomaux presque diagnostiques entre CAGT et RT. Nous avons utilisé les cinq SNP autosomaux fardée–arc-en-ciel pour estimer conjointement la proportion de mélange de la truite fardée dans Silver King Creek et le N_e de la population mixte à l'aide d'une méthode de vraisemblance maximale basée sur la coalescence. Étant donné l'histoire de l'empoissonnement dans Silver King Creek, il y a deux scénarios différents d'établissement des limites des points d'estimation attendus de N_e. Nous obtenons des points d'estimation de $N_e = 150$ et de $N_e = 750$ pour Silver King Creek selon les deux scénarios. Cette méthode pourra s'avérer utile dans les cas où les taxons sont différenciés et dans l'établissement de priorités dans les programmes de conservation et de restauration lorsqu'il y a introgression des populations concernées.

[Traduit par la Rédaction]

Introduction

Hybridization of native populations with introduced species is a major concern for the conservation and management of threatened species (e.g., Rhymer and Simberloff 1996; Epifanio and Nielsen 2000; Allendorf et al. 2004) and can result in negative consequences, including outbreeding depression,

Received 17 June 2010. Accepted 26 March 2011. Published at www.nrcresearchpress.com/cjfas on 10 August 2011.
J21880

Paper handled by Associate Editor Eric Taylor.

A.J. Finger, M.R. Stephens, and B.P. May. Genomic Variation Lab, Department of Animal Science, University of California Davis, One Shields Avenue, Davis, CA 95616, USA.
E.C. Anderson. Fisheries Ecology Division, Southwest Fisheries Science Center, National Marine Fisheries Service, NOAA, 110 Shaffer Road, Santa Cruz, CA 95060, USA.

Corresponding author: A.J. Finger (e-mail: ajfinger@ucdavis.edu).

or reduced population fitness through the loss of co-adapted gene complexes and ecological adaptations (Templeton 1986; Rhymer and Simberloff 1996; Allendorf et al. 2001) and even genomic extinction (Rhymer and Simberloff 1996; Allendorf et al. 2004). To mitigate these consequences, managers should know the genetic identity of populations of interest before undertaking conservation or restoration work on introgressed populations.

Detecting and quantifying introgression is critical to effective management. However, historical records are often incomplete or nonexistent, anecdotal evidence is subject to bias, and morphological data can often be misleading (Allendorf and Leary 1988; Baumsteiger et al. 2005; Metcalf et al. 2007). Genetic markers can provide accurate information, not only of recent introgression, but for additional management-relevant metrics such as effective population size (N_e), evolutionary potential, and genetic variation. N_e is of conservation importance because it allows managers to know how quickly genetic drift and inbreeding are removing genetic diversity from a population. Therefore, it is an important metric to consider when choosing populations for restoration or reintroduction purposes. In addition, knowing N_e changes over time may illuminate how management strategies affect N_e.

Microsatellite loci, with their high variability and mutation rate, have been used extensively to study population structure and hybridization (e.g., Cordes et al. 2006). More recently, SNPs (single nucleotide polymorphisms) have also been used to analyze levels of genetic variation in populations (see Morin et al. 2009 and citations therein) as well to detect hybridization in several taxa, including salmonids (Stephens et al. 2009), plants (Mercure and Bruneau 2008), birds (willow warblers (*Phylloscopus trochilus*), Bensch et al. 2002), and voles (*Microtus* spp.; Belfiore et al. 2003). Relative to microsatellites, SNPs are more readily standardized across platforms, have a better understood mutation model, and have less potential for homoplasy (e.g. Morin et al. 2009). Further, SNP panels can be chosen to be diagnostic between groups of interest, given adequate taxonomic and therefore sequence divergence. Finally, it can be reasonably assumed that an SNP that is fixed in a species today was also fixed for that species within a certain number of generations in the past and future, allowing estimation of N_e and admixture proportions using a temporal method when only current samples can be obtained.

Introgressive hybridization with introduced non-native rainbow trout (*Oncorhynchus mykiss* ssp., herein RT) has played a major role in the decline of all subspecies of native cutthroat trout (*Oncorhynchus clarkii* ssp., herein CT) in western North America (e.g., Allendorf and Leary 1988; Trotter 2008). There are 14 subspecies of CT in North America, two of which went extinct in the 20th century (Behnke 1992). Five are listed as threatened under the Endangered Species Act and seven have been petitioned for listing (Muhlfeld et al. 2009*a*). Indeed, hybridization is the greatest threat to the threatened Paiute cutthroat trout (*Oncorhynchus clarkii seleniris*, herein PCT; US Fish and Wildlife Service 1975). The scenario of the extirpation of the PCT in its limited native range within the Silver King Creek watershed, California, is common among species and subspecies endemic to narrow ranges. The entire native range of the PCT was stocked with other closely related trout species, including California golden trout (*Oncorhynchus mykiss aguabonita*, herein CAGT) and hatchery RT, leading to its extirpation through introgression and competition with non-natives. Out-of-basin native populations persist only because PCT were transferred to fishless waters prior to the putative initial non-native fish stocking events.

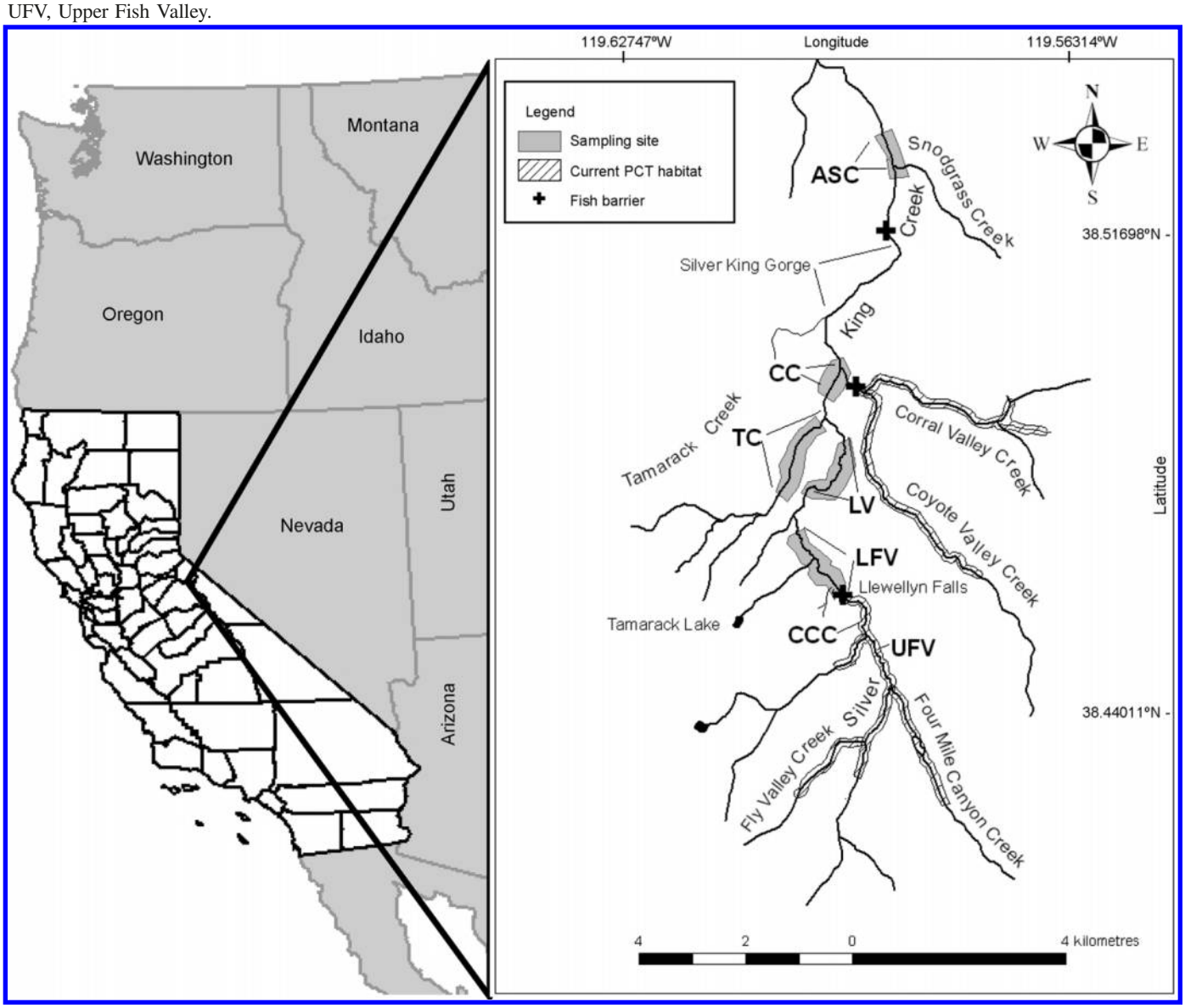
Finger et al. (2009) developed diagnostic SNP markers that distinguish PCT and the closely related Lahontan cutthroat trout (*Oncorhynchus clarkii henshawi*, herein LCT; together referred to as P/LCT) from RT for the purpose of quantifying introgression between the groups. We apply these and other markers to the trout population in Silver King Creek to accomplish three objectives: (*i*) evaluate the relative proportions of CAGT, RT, and P/LCT alleles in the population of fish in Silver King Creek today; (*ii*) describe a novel method that uses fixed SNP markers to simultaneously determine admixture proportions for and estimate N_e of an admixed population over time; and (*iii*) apply this method to the admixed population in Silver King Creek over a range of possible values.

Materials and methods

Study area overview and historical perspective

An approximately 10 km length of Silver King Creek, with the accessible reaches of three small tributaries (Tamarack Creek, Tamarack Lake Creek, and Coyote Valley Creek), constitutes the entire native habitat of the PCT (Fig. 1; Busack and Gall 1981; Cordes et al. 2004). Silver King Creek is a headwater tributary located in a high alpine meadow at 2500 m above sea level in the Carson–Iceberg Wilderness in Alpine County, California, that empties into the east fork of the Carson River. The upstream portion of this section of Silver King Creek is bounded by Llewellyn Falls, and the downstream portion is bounded by Silver King Gorge. These boundaries both act as fish barriers and are thought to have isolated PCT from LCT ~10 000 years ago. PCT generally has no body spots but may have up to nine spots on the body, while LCT is more heavily spotted (Behnke 1965). Between 1930 and 1994, Silver King Creek and Tamarack Lake, a small headwater lake connected to Silver King Creek, were planted with hatchery RT, CAGT, eastern brook trout (*Salvelinus fontinalis*), LCT, and PCT (which may have been introgressed with LCT; Table 1; W. Somer, California Department of Fish and Game, 1 Shields Avenue, Davis, California, personal communication, 2009). All of these non-native fish, except eastern brook trout, are known to hybridize with PCT (Busack and Gall 1981). Multiple reintroductions and efforts to eradicate hybrid fish from portions of Silver King Creek in the 1990s (see Cordes et al. 2004 for overview) culminated in a chemical treatment and restocking of Upper Silver King Creek above Llewellyn Falls in 1991–1993 from populations above fish barriers in the Silver King Creek watershed: Fly Valley Creek and Coyote Valley Creek (Fig. 1). Refuge populations of PCT are now restricted to portions of Silver King Creek in formerly fishless waters above fish barriers (Upper Silver King Creek above Llewellyn Falls, Coyote Valley Creek, Corral Valley Creek, Four Mile Canyon Creek, and Fly Valley Creek) and four out-of-basin locations (Cabin Creek and North Fork Cottonwood Creek in Mono County; Stairway Creek in Madera County; Sharktooth Creek in Fresno County). These refugial PCT populations were determined to

Fig. 1. Location of Silver King Creek in California and map of Silver King Creek between Silver King Gorge and headwaters. Sample locations are shaded in grey, and Paiute cutthroat trout (PCT) populations are shaded in grey with hatch marks. Fish barriers are denoted by a plus (+) symbol. Historic native PCT habitat was from Llewellyn Falls to Silver King Gorge. Abbreviations for hybrid fish in main stem Silver King Creek are as follows: ASC, Silver King Creek above Snodgrass Creek; CC, Canyon sites 1 and 2; TC, Tamarack Creek; LV, Long Valley; and LFV, Lower Fish Valley. Abbreviations for pure PCT above Llewellyn Falls are as follows: CCC, Connell's Cow Camp; and UFV, Upper Fish Valley.



be non-introgressed with RT based on microsatellite data (Cordes et al. 2004). However, morphological evidence suggests that in the main stem portion of Silver King Creek below Llewellyn Falls, remaining P/LCT are introgressed with hatchery RT and CAGT (W. Somer, California Department of Fish and Game, personal communication, 2009). The Paiute Cutthroat Trout Recovery Plan (US Fish and Wildlife Service 2004) suggests the removal of hybrid fish and restocking with PCT in main stem Silver King Creek as part of restoring this species to its native range. Following this current study's genetic inventory of the non-native trout population in Silver King Creek, California Department of Fish and Game plans to chemically treat this reach, restock it with PCT, and adopt a monitoring program (W. Somer, California Department of Fish and Game, personal communication, 2009).

Sample collection and DNA extraction

To genetically characterize the trout population in main stem Silver King Creek, samples of hatchery RT, PCT, and CAGT populations were chosen as references based on presumed genetic similarity to source populations indicated by stocking records (Table 2). Samples from two hatcheries (Hot Creek strain (RT/HCS) and Mount Whitney strain (RT/MWS); collectively referred to as RT references) were selected to represent the RT stocked into Silver King Creek. For CAGT reference populations, fish from Volcano Creek (CAGT/VC), Cottonwood Lakes 2 (CAGT/CL2) and Cottonwood Lakes 4 (CAGT/CL4) were chosen (collectively referred to as CAGT references). CAGT/VC is non-introgressed based on microsatellite (Cordes et al. 2006) and SNP markers (Stephens 2007). Stocking records indicate that

Table 1. Known stocking history of Silver King Creek, with subspecies, years, recorded number of fish (*N*) stocked, and planting location.

Species	Year stocked	*N*	Location
Oncorhynchus mykiss (RT)	1930	5 000	Silver King Creek — below Lewellyn Falls
	1930	5 000	Silver King Creek — below Lewellyn Falls
	1931	10 000	Silver King Creek — below Lewellyn Falls
	1932	10 000	Silver King Creek — below Lewellyn Falls
	1933	10 000	Silver King Creek — below Lewellyn Falls
	1949	8 400	Silver King Creek — below Lewellyn Falls
	1949	5 040[a]	Silver King Creek — below Lewellyn Falls
	1951	6 010	Silver King Creek — below Lewellyn Falls
	1952	5 017	Silver King Creek — below Lewellyn Falls
	1953	4 960	Silver King Creek — below Lewellyn Falls
	1976	960	Silver King Creek — below Lewellyn Falls
	1976	2 900	Silver King Creek — below Lewellyn Falls
Oncorhynchus mykiss aguabonita (CAGT)	1969	1 018	Tamarack Lake
	1972	1 000	Tamarack Lake
	1973	1 141	Tamarack Lake
	1974	2 250	Tamarack Lake
	1976	2 272	Tamarack Lake
Oncorhynchus clarkii seleniris (PCT)	1985	173	Tamarack Lake
	1987	100	Tamarack Lake
	1991	Hundreds	Tamarack lake
Oncorhynchus clarkii henshawi (LCT)	1935	10 000	Silver King Creek — unknown location
	1946	8 700	Silver King Creek — unknown location
	1946	1 740	Silver King Creek – Poison Flat
	1946	1 740	Silver King Creek – Poison Flat
	1947	19 600	Coyote Creek — lower stream
	1947	19 600	Coyote Creek — lower stream
	1947	9 800	Tamarack Creek — lower stream
	1947	4 200	Tamarack Creek — lower stream
	1947	5 600	Silver King Creek – Forks[b]
	1955	1 005	Silver King Creek – Forks[b]
	1957	1 000	Silver King Creek – Forks — mouth
	1959	1 035	Silver King Creek – Forks — mouth
	1962	1 020	Silver King Creek – Forks[b]
	1967	4 000	Silver King Creek – Forks[b]
	1968	5 000	Coyote Creek — mouth to barrier
	1971	4 000	Coyote Creek — mouth to barrier
	1973	3 600	Tamarack Creek — above barrier
	1975	3 600	Tamarack Creek — above barrier
	1976	4 000	Tamarack Lake
	1980	4 200	Tamarack Lake
	1982	4 000	Tamarack Lake
	1987	3 000	Tamarack Lake
Salvelinus fontinalis (eastern brook trout)	1935	5 000	Silver King Creek — unknown location
	1968	500	Tamarack Lake

Note: RT, rainbow trout; CAGT, California golden trout; PCT, Paiute cutthroat trout; LCT, Lahontan cutthroat trout.

[a]Above Llewellyn Falls.

[b]Precise location of Forks is unknown — it is possibly the confluence of Silver King Creek with Coyote Valley Creek.

Cottonwood Lakes were the sources of CAGT planted in Silver King Creek; these populations are thought to be introgressed with hatchery RT (the date of introgression is uncertain, but may have been as early as the 1920s; Cordes et al. 2006). Finally, fish from the chemically treated and restored population in Upper Silver King Creek (from Upper Fish Valley (PCT/UFV) and Connell's Cow Camp (PCT/CCC); collectively referred to as PCT references) were chosen as PCT representatives; these fish were determined to be non-introgressed with RT or LCT, as determined by microsatellite (Cordes et al. 2004) and SNP analysis (Finger et al. 2009).

Fin clips were collected from 250 individual fish in Silver King Creek below Llewellyn Falls in 2006 (50 each from Lower Fish Valley (SKC/LFV), Long Valley (SKC/LV), Tamarack Creek (SKC/TC), Canyon Sites 1 and 2 (SKC/CC), and Silver King Creek above Snodgrass Creek (SKC/ASC); collectively referred to as SKC references). Thirty-

Table 2. Samples included in this study.

Subspecies – sample location	Site name	Abbreviation	Year	*N*
Silver King Creek	Lower Fish Valley	SKC/LFV	2006	50
	Long Valley	SKC/LV	2006	50
	Tamarack Creek	SKC/TC	2006	50
	Canyon Sites 1 and 2	SKC/CC	2006	50
	Silver King Creek above Snodgrass Creek	SKC/ASC	2006	50
Paiute cutthroat trout (*Oncorhynchus clarkii seleniris*)	Upper Fish Valley	PCT/UFV	2006	19
	Connell's Cow Camp	PCT/CCC	2006	19
California golden trout (*Oncorhynchus mykiss aguabonita*)	Volcano Creek	CAGT/VC	2001	34
	Cottonwood Lakes 2 (Lakes 1, 2, 3)	CAGT/CL2	2000	50
	Cottonwood Lakes 4 (Lakes 4, 5)	CAGT/CL4	2000	50
Rainbow trout (*Oncorhynchus mykiss irideus*)	Hot Creek strain, Hot Creek Hatchery	RT/HCS	2002	32
	Mt. Whitney Strain, Mt. Whitney Hatchery	RT/MWS	2002	30
Total				484

Note: Sampled individuals range from 1+ to 4 years old, with most individuals being 2+ years.

eight additional fish were sampled from Upper Silver King Creek (19 from PCT/UFV and 19 from PCT/CCC). To minimize relatedness among individuals sampled, ~10 fish for each sample site were collected at five locations (for up to ~50 individuals per site) within the site (see Fig. 1 for map of sites; shading and labels indicate sample location and total area of each site).

Whole genomic DNA was extracted from each fin clip using the Promega Wizard Extraction Kit and diluted to 5 ng·μL^{-1}. This DNA and DNA previously extracted from individuals from the reference populations were stored at –20 °C (Table 2 lists samples examined in this study).

SNP genotyping

A total of 12 TaqMan assays (Finger et al. 2009; Stephens et al. 2009) were performed on all reference populations listed (Table 2): one mitochondrial and five autosomal loci fixed for differences between P/LCT and *O. mykiss* ssp., (Table 3; B9_288, F5_136, HOXD_287, Rag11_280, URO_302 and Dloop_243; herein P/LCT SNPs); and one mitochondrial and five autosomal loci nearly fixed for differences between CAGT and RT (Table 4; A1A8_94, B9_388, ID1C_77–83, HOXD_170, RAG11_137, Dloop_316; herein CAGT SNPs). These SNPs were determined to be diagnostic or nearly diagnostic by surveying multiple populations of pure PCT, LCT, CAGT, and hatchery RT based on historical records, morphology, and previous genetic studies (see Finger et al. 2009 and Stephens et al. 2009 and citations therein). Though these markers are fixed or nearly fixed for a broad range of populations surveyed, the possibility exists that the subspecies or species of interest in this case may have shared alleles. Data for the six CAGT SNPs for CAGT/CL2, CAGT/CL4, RT/HCS, and RT/MWS were taken from Stephens (2007). TaqMan assays were run in Chromo4 Real-Time PCR Detector (MJ Research – Bio-Rad Laboratories, Inc., Hercules, California) in 96-well reaction plates. On each plate, four wells were reserved for positive and negative controls: one each with template from a known homozygote for each allele, one with DNA template from a known heterozygote, and one well with water as a negative control. MJ Opticon Monitor version 3.1.32 (MJ Research – BioRad Laboratories, Inc.) software was used to determine individual genotypes by creating a scatterplot depicting the endpoint fluorescence of each allele-specific probe in each well. Baseline average fluorescence over the 10- to 30-cycle range was subtracted to reduce noise, and endpoint fluorescence clusters were identified by comparing fluorescence with positive and negative controls.

SNP data analysis

Tests for significant pairwise linkage disequilibrium (LD) were implemented in Arlequin version 3.5 (Excoffier and Lischer 2010). Allele frequencies for each SNP marker were calculated and averaged over each sample location and population. Input files of autosomal SNP data were made for the software Structure version 2.2.3 (Pritchard et al. 2000) using the software Convert (Glaubitz 2004). Structure uses a Bayesian algorithm that, without prior information regarding each individual's membership to a population, computes by Markov chain Monte Carlo (MCMC) the posterior probability for the fraction, q, of each individual's ancestry that originates from each of K different genetic subgroups. The output of the MCMC from multiple runs at different values of K can be used to estimate the most likely K, the number of genetic clusters in the data. Two analyses were done with nuclear SNP data in Structure: (*i*) an analysis with the five autosomal P/LCT SNPs and the five autosomal CAGT SNPs for individuals from PCT, SKC, CAGT, and RT references with three runs each of K values 1–8; and (*ii*) an analysis with the five autosomal CAGT SNPs for SKC, CAGT, and RT references with three runs each of K values 1–5. Both analyses used the admixture model and had a burn-in period of 50 000 and 500 000 iterations. Clumpp version 1.1.2 (Jakobsson and Rosenberg 2007) was used to align multiple runs at each K value. The most likely estimated K value was determined using the method in Evanno et al. (2005). The software Distruct version 1.1 (Rosenberg 2004) was used to create a visual representation of the Structure output at the most likley K value.

Estimation of admixture proportions and N_e

We developed a simple, likelihood-based method for jointly estimating the admixture proportions and N_e of an admixed population with species- or subspecies-diagnostic SNP markers. Intuitively, the fraction of gene copies from each species or subspecies in the admixed population provides in-

Table 3. Paiute or Lahontan cutthroat trout (P/LCT) allele frequencies in Silver King Creek and reference populations at six P/LCT single nucleotide polymorphism loci.

Location	*N*	B9_228[a]	F5_136[a]	Dloop_243[a]	HOXD_287[a]	RAG11_280[a]	URO_302[a]	Avg.
PCT	38	1.0	1.0	1.0	1.0	1.0	1.0	1.0
SKC/LFV	50	0.0	0.0	0.0	0.0	0.01	0.0	<0.01
SKC/LV	50	0.01	0.0	0.0	0.02	0.01	0.0	0.01
SKC/TC	50	0.0	0.0	0.0	0.01	0.0	0.0	<0.01
SKC/CC	50	0.02	0.0	0.0	0.04	0.01	0.0	0.01
SKC/ASC	50	0.05	0.0	0.0	0.06	0.01	0.0	0.02
SKC total	250	0.02	0.0	0.0	0.03	0.01	0.0	0.01
CAGT/VC	34	0.0	0.0	0.0	0.0	0.0	0.0	0.0
CAGT/CL2	46	0.0	0.0	0.0	0.0	0.0	0.0	0.0
CAGT/CL4	46	0.0	0.0	0.0	0.0	0.0	0.0	0.0
RT/HCS	30	0.0	0.0	0.0	0.0	0.0	0.0	0.0
RT/MWS	49	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Note: Refer to Table 2 for location definitions.

[a]Finger et al. 2009.

formation about the admixture proportions, and the variation of that fraction from locus to locus provides information about the degree of genetic drift that has occurred in the population since the admixture event. The latter provides an estimate of N_e of the admixed population, if the time since admixture is known. We formalize this intuitive understanding of our procedure with a likelihood model based on the neutral coalescent (Kingman 1982), which automatically accounts for allele frequency variance due both to genetic drift and finite sampling.

The problem of jointly inferring admixture fractions and N_e of an admixed population was first addressed by Long (1991), who adopted a method-of-moments procedure. More recently several likelihood-based methods have been developed for the joint inference of admixture fractions and N_e. Chikhi et al. (2001) proposed a simple admixture model and provided a likelihood value in terms of the coalescent process. However, using MCMC, their method requires from days to weeks to run. Wang (2003) extended the underlying admixture model of Chikhi et al. (2001) to include an unobserved population from which both populations contributing to the admixture split in the past. He developed a pseudolikelihood for the model using the truncated matrix multiplication machinery introduced in Wang (2001). Wang (2006) described a coalescent-based likelihood suitable for sequence data that is applicable to the model proposed in Wang (2003).

The method we present is based on a simplified version of the admixture model in Chikhi et al. (2001). The likelihood for the model is based on the coalescent process, but, by exploiting the importance-sampling techniques introduced in Anderson (2005), the likelihood can be computed in a matter of seconds. This new method makes assumptions that are suitable only when using species- or subspecies-diagnostic SNPs. In such cases, however, unlike existing methods, our method can rapidly compute the entire joint likelihood surface for the admixture proportion and N_e (details appear in Appendix A). We estimated the N_e of the admixed Silver King Creek population using five autosomal, species-diagnostic P/LCT SNPs (listed in Table 3).

It is likely that the actual history of admixture between P/LCT and RT in Silver King Creek is considerably more complex than the single admixture event of our likelihood model. Given more certainty about the exact history of individual admixture events, we would be inclined to model the entire admixture history for the estimation of N_e. However, though stocking records are available (Table 1), they do not constitute a reliable reconstruction of the history of admixture, and many possible scenarios are possible, depending upon the success of the various introductions. For example, on one extreme, if the only successful introductions of any fish to Silver King Creek were those in 1930, then all the RT/CAGT alleles in Silver King Creek today would have descended from those 10 000 fish introduced in 1930, the PCT alleles would have descended from native fish, and the admixed population would have been undergoing genetic drift for 76 years (1930 to 2006) or roughly 25.3 generations (the average generation time for PCT is thought to be about 3 years; Wong 1975; the average generation time of rainbow trout is 2–3 years; Gall et al. 1988). At the other extreme, it is presumably possible that PCT in Silver King Creek were completely displaced by *O. mykiss* ssp. introductions between 1930 and 1976 and that the P/LCT alleles observed in the creek today originate from the 1991 stocking of PCT planted into Tamarack Lake. Under the latter scenario, drift in the admixed population would have been ongoing for only 15 years (1991 to 2006) or roughly five generations prior to sampling. We provide theory (see Supplemental Materials online[1]) to demonstrate that these two extreme admixture scenarios bound the degree of interlocus allele frequency variance expected in the admixed population and, hence, the possible estimates of N_e. The former scenario leads to an estimate of N_e about five times larger than the latter, since the latter scenario

[1]Supplementary data are available with the article through the journal Web site (http://www.nrcresearchpress.com/cjfas).

Table 4. California golden trout (CAGT) allele frequencies in Silver King Creek and reference populations at six CAGT SNP loci.

Location	A1A8_94[a]	B9_388[a]	ID1C_gap[a]	HOXD_170[b]	RAG11_137[a]	Dloop_316[a]	Avg.
PCT	0.0	1.0	0.0	1.0	—	1.0	—
SKC/LFV	0.10	0.55	0.17	0.41	0.21	0.02	0.24
SKC/LV	0.15	0.47	0.26	0.38	0.18	0.04	0.25
SKC/TC	0.01	0.41	0.04	0.37	0.10	0.0	0.16
SKC/CC	0.08	0.46	0.09	0.37	0.19	0.02	0.20
SKC/ASC	0.06	0.53	0.04	0.42	0.03	0.02	0.18
SKC total	0.08	0.50	0.12	0.39	0.14	0.02	0.20
CAGT/VC	1.0	0.94	1.0	1.0	1.0	1.0	0.99[a]
CAGT/CL2	0.94	0.93	0.92	0.99	0.89	0.87	0.92[c]
CAGT/CL4	0.82	0.94	0.90	0.92	0.93	1.0	0.92[c]
RT/HCS	0.0	0.07	0.0	0.18	0.0	0.0	0.04[c]
RT/MWS	0.0	0.02	0.0	0.02	0.02	0.0	0.01[c]

Note: Refer to Table 2 for location definitions.

[a]A1A8_94, B9_388, RAG11_137, and Dloop_316 and data for CAGT SNPs for the CAGT/VC population are from Stephens et al. (2009).

[b]Sprowles et al. (2006).

[c]Minor allele frequencies for the CAGT SNPs for populations CAGT/CL2, CAGT/CL4, RT/HCS, and RT/MWS are from Stephens (2007).

has about one fifth of the time for the observed degree of interlocus allele frequency variation to occur via drift. Results are given for both of these extreme scenarios.

Microsatellite genotyping

A total of 19 microsatellite loci (Table 5) were amplified in six multiplexed polymerase chain reactions in samples from 11 reference locations (listed in Table 1). We excluded PCT from microsatellite genotyping becazuse of poor amplification. Each reaction had a general master mix with a total volume of 10 μL with 2 ng template DNA, 2 mmol·L^{-1} $MgCl_2$, 125 μmol·L^{-1} each dNTP, and 0.1–0.2 μmol·L^{-1} each forward sequencing primer labeled with a fluorescent dye (either VIC, 6FAM, or NED), 0.1–0.2 μmol·L^{-1} each reverse primer, and 2 U *Taq* polymerase (Table 5). Reactions spent 4 min at 95 °C; 25 cycles of 30 s at 95 °C, 30 s at 58 °C, 45 s at 72 °C, followed by 45 min at 60 °C.

For each individual, a total volume of 10 μL consisting of 1.0 μL of each multiplexed PCR product diluted to 1:5 concentration with water, 0.2 μL LIZ600 size standard, and 8.8 μL formamide was placed in an individual well on a 96-well reaction plate. After being shocked at 95 °C for 3 min, the product was run on an ABI 3130xl Genetic Analyzer (Applied Biosystems, Carlsbad, California). Resulting electropherograms were analyzed with Genemapper version 4.0 software (Applied Biosystems), and allelic sizes were confirmed manually. Individuals were rerun a maximum of two times for each failed locus, with each rerun containing previously genotyped controls. Individuals with <90% genotyping rate were not included in further analysis.

Microsatellite data analysis

Files were composed for the software programs Genepop (Raymond and Rousset 1995), Genetic Data Analysis (GDA; Lewis and Zaykin 2001), and Structure using the software program Convert (Glaubitz 2004). GDA was used to calculate observed (H_o) and expected (H_e) heterozygosities and number of alleles per locus for each population. Genepop was used to detect deviation from Hardy–Weinberg equilibrium for each locus for each population using the probability test. Genepop was also used to detect significant LD for each population and to calculate pairwise θ_{ST} values (Weir and Cockerham 1984). For both Hardy–Weinberg and LD tests, Markov chain parameters were 1000 dememorizations, 100 batches, and 1000 iterations per batch. Sequential Bonferroni corrections were used to calculate significance for multiple comparisons (Rice 1989). Six loci were dropped from further analysis because of low genotyping rates, unscorable alleles, or significant LD. A two-dimensional factorial correspondence analysis (FCA) was performed in Genetix version 4.05 (Belkhir et al. 1996–2004) to depict the relationships between all genotyped individuals. In FCA, two axes are created that are a composite of variables that optimize the differences between individuals. Individuals are then visualized as points along these axes.

As with SNP data, Structure version 2.2.3 was used to estimate the number of genetic clusters (K) among the RT, CAGT, and SKC references. Three runs each were done for K = 1–9 with the admixture model, a burn-in period of 100 000 and 1 000 000 iterations. The estimated K value was obtained as with the SNP data.

Results

SNP genotyping and analysis

The reference PCT fish are fixed for the P/LCT alleles at all six P/LCT SNP loci, and the RT reference fish are fixed for RT alleles at all six CAGT SNP loci. Out of 50 tests for SKC reference locations, none had significant LD ($p > 0.05$). SKC reference fish have very few P/LCT alleles (an average of 0.0 to 0.03; Table 3) at the P/LCT SNP loci. Of the fish with P/LCT alleles, all but one are heterozygous (one individual was homozygous for the P/LCT allele, *A*, at HOXD_287). No SKC reference fish have the P/LCT allele

Table 5. Multiplex primer combinations for microsatellite loci used including multiplex number (Multi), fluorescent label, reference, and primer concentration in multiplex reaction for forward (F) and reverse (R) primers.

Multi	Primer	Label	Primer reference	Primer concentration (μmol·L^{-1})
1	OMM1037	6FAM	Rexroad et al. 2002*a*	F, R = 0.1
	OMM1036	NED	Rexroad et al. 2002*a*	F, R = 0.1
	OMM1089	VIC	Rexroad et al. 2002*b*	F, R = 0.1
2	OtsG85	6FAM	Williamson et al. 2002	F, R = 0.2
	OMM1322	NED	Palti et al. 2002	F, R = 0.1
	OtsG423	VIC	Williamson et al. 2002	F, R = 0.1
3	OMM1082	6FAM	Rexroad et al. 2002*b*	F, R = 0.1
	Omy1009UW	NED	Spies et al. 2005	F, R = 0.1
	OMM1097	VIC	Rexroad et al. 2002*b*	F = 0.1, R = 0.2
4	OMM1046	NED	Rexroad et al. 2002*a*	F, R = 0.1
	OMM1078	VIC	Rexroad et al. 2002*b*	F, R = 0.1
	OMM1051	NED	Rexroad et al. 2002*a*	F, R = 0.1
	OtsG249b	6FAM	Williamson et al. 2002	F, R = 0.1
5	OMM1088	6FAM	Rexroad et al. 2002*b*	F, R = 0.1
	Omy1011UW	NED	Spies et al. 2005	F, R = 0.1
	OMM1058	VIC	Rexroad et al. 2002*b*	F = 0.1, R = 0.2
6	OtsG3	6FAM	Williamson et al. 2002	F, R = 0.1
	OMM1083	NED	Rexroad et al. 2002*b*	F = 0.1, R = 0.2
	OMM1081	VIC	Rexroad et al. 2002*b*	F = 0.1, R = 0.2

at the mitochondrial P/LCT SNP (Dloop_243). The furthest downstream SKC reference location samples have a few more P/LCT alleles (SKC/ASC, between 0.0 and 0.06 across loci; SKC/CC, between 0.0 and 0.04 across loci) than upstream locations (SKC/LFV, directly below Llewellyn Falls, between 0.0 and 0.01 across loci).

Sixteen individuals from main stem Silver King Creek have complete RT genotypes at all of the 12 SNPs, but none have complete CAGT genotypes. The proportion of CAGT alleles varies widely across CAGT SNP loci (between 0.03 and 0.48; Table 4) in Silver King Creek. At the mitochondrial locus (Dloop_316), the proportion of CAGT alleles is 0.02. The CAGT/CL2 and CAGT/CL4 populations each have an average of 0.92 CAGT alleles. In contrast, CAGT/VC has an average of 0.99 CAGT alleles. This supports other evidence of RT introgression in the Cottonwood Lakes populations (Cordes et al. 2006). At the CAGT SNPs, the PCT samples amplified at every locus except Rag11_137, and the PCT reference fish are homozygous for either CAGT alleles (B9_388, HOXD_170, and Dloop_316) or RT alleles (A1A8_94, ID1C_gap). In both sets of SNPs, the B9 and HOXD loci have a higher proportion of the non-RT allele in Silver King Creek fish than in the other SNPs. PCT and CAGT share an allele at B9_388 and HoxD_170 (Finger et al. 2009). It is possible that this inflates the proportion of CAGT alleles in Silver King Creek at these loci, but it is doubtful given the low proportion of P/LCT detected by the P/LCT SNPs.

The most likely K value for the Structure analysis of 10 autosomal SNPs is $K = 3$ genetic subgroups. The first cluster consists of the PCT references (UFV and CCC), the second consists of SKC and RT references, and the third cluster is composed of CAGT references (Fig. 2). In Silver King Creek, several individuals have q values indicating membership to multiple clusters, reflecting the history of hybridization. For the second Structure analysis, with only five autosomal CAGT SNPs, $K = 2$. SKC and RT references form one cluster, and the CAGT references group together to form the second cluster. In this analysis, the RT introgression in CAGT/CL2 and CAGT/CL4 is apparent in some individuals that are inferred to have ancestry from both clusters (Fig. 3*a*).

Microsatellite genotyping and analysis

All individuals from the SKC, RT, and CAGT reference populations were genotyped; those with a genotyping rate less than 90% after two rerun attempts were discarded from further analysis, leaving a total of 421 individual genotypes. Out of 936 pairwise LD comparisons, 88 are significant for LD before a Bonferroni correction and four are significant for LD after the correction. Across all sample locations, the total number of alleles for each locus ranges from 12 (Omm1058) to 28 (OtsG85), and the number of alleles per locus ranges from 5.61 (CAGT/VC) to 11.9 (SKC/LFV) (Table 6). Before a Bonferroni correction, 18 tests were significant for Hardy–Weinberg disequilibrium ($p < 0.05$). After the Bonferroni correction, seven tests were significant ($p < 0.05$). The microsatellite locus Omy1009UW was discarded from further analysis because of significant Hardy–Weinberg disequilibrium in the five collections from Silver King Creek ($p < 0.01$). All other loci are in Hardy–Weinberg equilibrium after the sequential Bonferroni correction for all populations, with two exceptions: Omm1088 in SKC/LFV ($p < 0.01$) and Omm1046 in SKC/CC ($p < 0.01$; Table 6). H_o ranges from 0.58 in CAGT/VC to 0.82 in the SKC/ASC location (Table 7). In general, heterozygosity values in SKC reference locations are greater than those for the CAGT and RT references. At several loci, alleles are present in the SKC reference that are not present in any of the other reference populations, indicating either genetic drift or that the refer-

Fig. 2. Depiction of clustering results from Structure analysis with five autosomal P/LCT SNP markers and five autosomal CAGT SNP markers. The most likely *K* value is 3. The first cluster (black) is the PCT from above Llewellyn Falls (PCT/UFV and PCT/CC). The second cluster (white) consists of hybrid fish from Silver King Creek (SKC/LFV, SKC/LV, SKC/TC, SKC/CC, and SKC/ASC) and the RT reference populations (RT/HCS and RT/MWS). The third cluster (gray) includes CAGT reference populations (CAGT/VC, CAGT/CL2, and CAGT/CL4). Refer to Table 2 for location definitions.

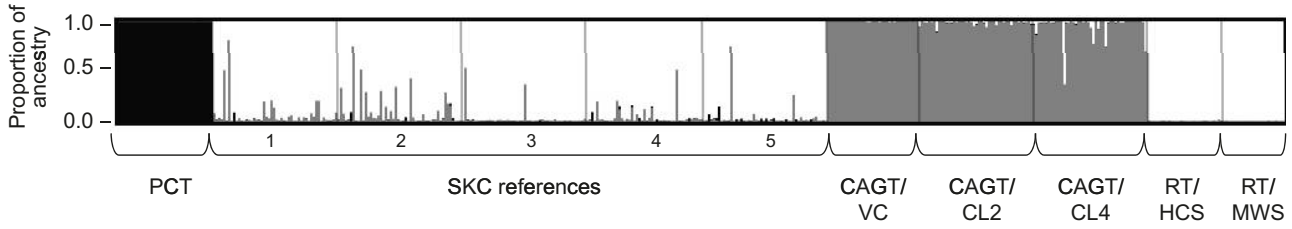
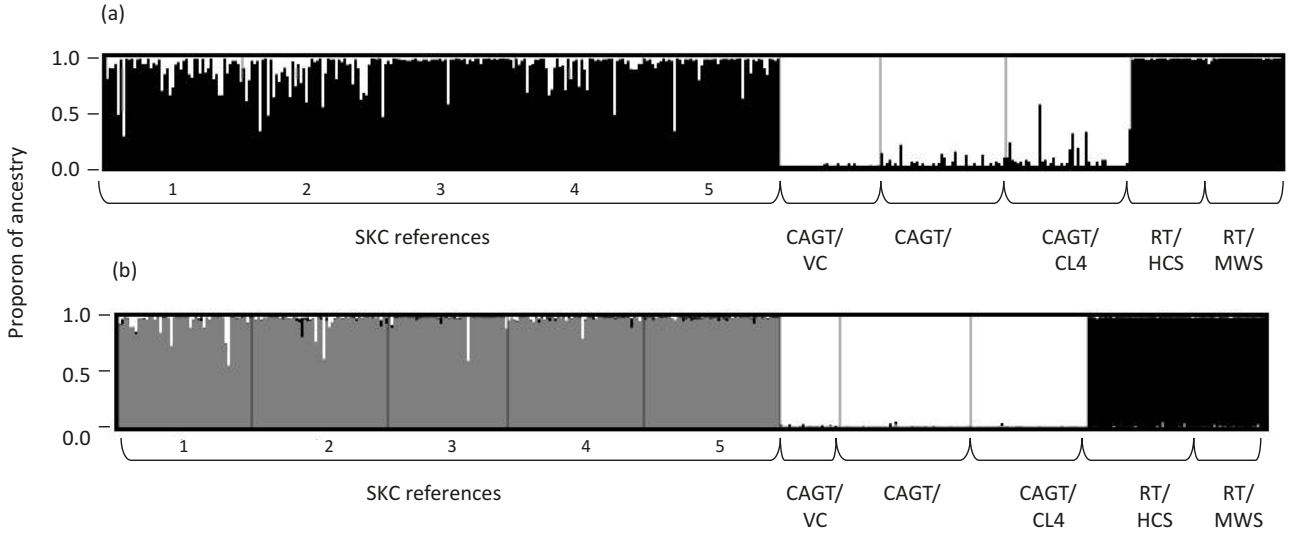


Fig. 3. (*a*, *b*) Depiction of clustering results from two Structure analyses. (*a*) Analysis with five autosomal CAGT SNPs. The most likely *K* value is 2. The first cluster (black) consists of hybrid Silver King Creek fish from five sample locations (1, Lower Fish Valley; 2, Long Valley; 3, Tamarack Creek; 4, Canyon Sites 1 and 2; and 5, Silver King Creek above Snodgrass Creek) and the RT reference populations (Hot Creek Strain (RT/HCS) and Mount Whitney Strain (RT/MWS)). The second cluster (white) includes the CAGT reference populations (Volcano Creek (CAGT/VC) and Cottonwood Lakes 2 and 4 (CAGT/CL2, CAGT/CL4)). (*b*) Depiction of clustering results from Structure analysis with 12 microsatellites including CAGT and RT references. The most likley *K* value is 3. The first cluster (gray) consists of the Silver King Creek fish from five sample locations (1, Lower Fish Valley; 2, Long Valley; 3, Tamarack Creek; 4, Canyon Sites 1 and 2; and 5, Silver King Creek above Snodgrass Creek). The second cluster (white) consists of CAGT reference populations (Volcano Creek (CAGT/VC) and Cottonwood Lakes 2 and 4 (CAGT/CL2, CAGT/CL4)). The third cluster (black) is the RT reference populations (Hot Creek Strain (RT/HCS) and Mount Whitney Strain (RT/MWS)). Refer to Table 2 for location definitions.



ence populations do not fully represent the true parental populations of trout in Silver King Creek.

Pairwise θ_{ST} values (summarized in Table 8) suggest that the population of trout in Silver King Creek is weakly structured (θ_{ST} values range from <0.01 to 0.04). Tamarack Creek (SKC/TC), a tributary to Silver King Creek, is slightly divergent from other SKC reference locations (θ_{ST} = 0.03–0.04), with a slightly higher proportion of RT alleles. Fish from CAGT/CL2 and CAGT/CL4 have a low pairwise θ_{ST} value (0.01) and are each moderately divergent from CAGT/VC (θ_{ST} = 0.20), which in turn is moderately divergent from all other populations (θ_{ST} = 0.20–0.30). This is probably due in part to low genetic diversity in CAGT/VC. In addition, CAGT/VC was the most divergent from RT references (θ_{ST} = 0.30). RT/MWS and RT/HCS are moderately divergent from each other (θ_{ST} = 0.19).

The FCA analysis produces a graph with four distinct clusters; the first consists of the SKC references, located near zero on both axes (Axis 1, 30.8% inertia; Axis 2, 23.1% inertia; Fig. 4). CAGT references group together at the negative end of Axis 2. RT/MWS fish form a cluster between the SKC reference cluster and a final cluster of fish from RT/HCS.

Structure analyses

The optimal *K* value for the Structure analysis is *K* = 3 for the microsatellite analyses. SKC reference fish form one cluster, CAGT references a second cluster, and RT references a third (Fig. 3*b*). In the microsatellite Structure analysis, in contrast with the SNP Structure analysis, SKC references form their own cluster, and the introgression with RT in CAGT/CL2 and CAGT/CL4 is less apparent.

Admixture and N_e

The maximum likelihood estimate (MLE) of the proportion of P/LCT ancestry in the admixed Silver King Creek population is 0.01, and the MLE of the effective size of this

Table 6. Results for 13 unlinked microsatellite loci analyzed in this study.

	OMM1037[a]	OMM1036[a]	OMM1089[b]	OtsG85[c]	OMM1322[d]	Omy1009UW[e]
No. of alleles per locus	15	18	25	28	23	27
Sample location (*N*)						
SKC/LFV (49)	0.67	0.17	0.58	0.21	0.37	**<0.01***
SKC/LV (50)	0.41	0.76	0.42	0.20	**0.02**	**<0.01***
SKC/TC (44)	0.17	0.44	0.10	0.58	**0.04**	**<0.01***
SKC/CC (50)	0.85	0.09	0.07	0.21	0.12	**<0.01***
SKC/ASC (50)	0.09	0.20	0.14	0.96	0.83	**<0.01***
CAGT/VC (22)	1.0	0.26	0.19	0.49	0.07	0.27
CAGT/CL2 (48)	0.53	0.30	0.92	0.20	0.45	0.18
CAGT/CL4 (43)	0.47	0.77	0.70	0.70	0.13	0.13
RT/HCS (31)	0.86	0.51	0.23	0.09	0.55	**0.05**
RT/MWS (34)	0.17	0.59	0.45	0.24	0.14	0.95

Note: Data include citations, total number of alleles per locus, and *p* values for Hardy–Weinberg test. Numbers in bold font are significant before a further analysis because of Hardy–Weinberg disequilibrium in the Silver King Creek references.

[a]Rexroad et al. 2002*a*.

[b]Rexroad et al. 2002*b*.

[c]Williamson et al. 2002.

[d]Palti et al. 2002.

[e]Spies et al. 2005.

Table 7. Results from genotyping 11 sample locations at 12 microsatellites.

Sample location	*N*	H_e	H_o	N_A
SKC/LFV	49	0.85	0.84	12.1
SKC/LV	50	0.84	0.82	12.4
SKC/TC	44	0.81	0.79	10.2
SKC/CC	50	0.84	0.83	12.6
SKC/ASC	50	0.84	0.86	12.1
CAGT/VC	22	0.59	0.58	5.61
CAGT/CL2	48	0.63	0.63	6.77
CAGT/CL4	43	0.62	0.63	6.61
RT/HCS	31	0.77	0.79	7.46
RT/MWS	34	0.74	0.70	7.38

Note: Data shown include sample number (*N*); expected (H_e) and observed (H_o) heterozygosity values; and mean number of alleles per locus per population (N_A).

population since the time of admixture under the $T = 25.3$ generation scenario is 750, and under the $T = 5$ generation scenario the MLE of N_e is 150. Likelihood surfaces with 95% confidence intervals for those two scenarios appear in Figs. 5*a* and 5*b*. The likelihood curves for N_e given each locus, individually, assuming an admixture fraction of 0.01, show that while there is considerable variation across loci, the MLE values (150 and 750, respectively) are within two units of log-likelihood of the maximum for every locus, suggesting that none of the loci are outliers and that the patterns of diversity can be explained by genetic drift alone (Figs. 5*c*, 5*d*).

Discussion

Genetic status of the trout in Silver King Creek

Based on SNP and microsatellite data, Silver King Creek is an admixed population of trout that is composed mostly of hatchery RT and CAGT. Indeed, despite being stocked with PCT and LCT for nearly 20 years after the last planting of *O. mykiss* ssp., the population in Silver King Creek now has very little P/LCT ancestry. Neither PCT nor LCT has successfully coexisted in Silver King Creek with CAGT or hatchery RT. This lack of success of cutthroat trout may be explained by several interacting factors, discussed in more detail below: stocking history in Silver King Creek (propagule pressure), the interactions between species in Silver King Creek, and the physical features of Silver King Creek. PCT were only stocked in Tamarack Lake, and the last time LCT was stocked in main stem Silver King Creek was 1957. Perhaps when P/LCT were stocked in Tamarack Lake, the propagule pressure from the continuous stocking of CAGT and RT from 1930 to 1976 allowed the establishment of a robust population of *O. mykiss* ssp. Subsequent stocking of P/LCT trout in Tamarack Lake was insufficient to establish substantial numbers of P/LCT genes in the main stem. Perhaps the habitat in Tamarack Lake is unsuitable (it is currently fishless) or the P/LCT did not move downstream. However, CAGT were also stocked in Tamarack Lake, and CAGT is genetically well represented in Silver King Creek; CAGT may move downstream more readily than P/LCT. Recent studies have shown that propagule pressure, source connectivity, and habitat quality are major biotic and abiotic factors influencing the spread of hybridization between nonnative RT and native cutthroat trout (e.g., Muhlfeld et al. 2009*c*; Bennett et al. 2010).

Species interactions may also play a role in the genetic composition of trout in Silver King Creek today. Planted P/LCT trout and their hybrids may generally be out-competed by CAGT and RT subspecies and their hybrids. This could be due to preferential mating or greater fitness of hybrids with more RT genes. Metcalf et al. (2008) surveyed hybrid populations of introduced RT and two subspecies of cutthroat trout native to Colorado and found a higher frequency of RT mtDNA than expected in hybrid populations, even in locations where there was more cutthroat trout nuclear DNA than RT nuclear DNA. This suggests asymmetrical pairings and the possibility of, in some cases, RT swamping out cutthroat genes. Our data provides some evidence of asymmetri-

OMM1046[a]	OtsG249b[c]	OMM1088[b]	Omy1011UW[e]	OMM1058[b]	OtsG3[c]	OMM1081[b]
16	17	21	15	12	16	16
1.0	0.45	**<0.01***	0.18	0.76	0.08	0.54
0.08	0.51	0.29	**0.03**	0.05	0.67	0.77
0.02	0.07	**0.02**	0.08	0.19	0.07	0.28
<0.01*	**0.05**	0.05	0.80	0.25	0.37	0.31
0.41	0.58	0.07	0.86	0.24	0.97	0.59
0.16	0.06	0.24	1.0	1.0	—	0.88
0.27	0.75	**0.04**	0.44	1.0	1.0	0.10
0.91	0.94	0.31	0.16	1.0	0.11	0.87
0.08	0.12	**0.02**	**0.04**	0.13	0.41	0.50
0.88	0.14	**0.03**	0.29	0.08	0.71	0.13

Bonferroni correction; numbers with an asterisk (*) are significant after Bonferroni correction. Omy1009UW was discarded before

cal hybridization patterns in Silver King Creek. The P/LCT allele was not present at the Dloop_243 locus, and only 0.02 of the alleles at Dloop_316 were of CAGT ancestry, suggesting lower reproductive success for pairings of CAGT or P/LCT females and male RT than the reciprocal mating. Additionally, behavioral factors such as spawn timing and location or straying rates may make introgression between some taxa more likely than between others.

Physical features of the stream system such as fish barriers can also affect the distribution of alleles (e.g., Rubidge and Taylor 2005, but see Rasmussen et al. 2010). It appears that little gene flow is occurring over Llewellyn Falls into main stem Silver King Creek from the upstream PCT populations. In addition, the two furthest downstream Silver King Creek locations are separated from upstream populations of non-introgressed PCT (Coyote Valley Creek and Corral Valley Creek) by fish barriers of similar size to Llewellyn Falls. This genetic evidence, in addition to anecdotal evidence from 1991–1993 chemical treatments, supports the notion that PCT do not move much (W. Somer, California Department of Fish and Game, personal communication, 2009). It appears that many biotic and abiotic variables can affect the hybridization dynamics between trout species (e.g., Heath et al. 2010; Muhlfeld et al. 2009*c*).

California golden trout × rainbow trout

The hybridization dynamics between CAGT and RT are not as well studied as those between cutthroat trout and RT. In main stem Silver King Creek, though only stocked five times (in Tamarack Lake, between 1969 and 1976), CAGT persist genetically at surprisingly high percentages (0%–48% across SNP loci). Fish from the upstream portion of main stem Silver King Creek (SKC/LFV and SKC/LV) have a slightly higher proportion of CAGT alleles (an average of 0.24 for both locations). CAGT were only planted in Tamarack Lake and must have moved downstream from Tamarack Lake into the main stem, leaving a genetic signature in SKC/LFV and SKC/LV. The high proportion of CAGT in Silver King Creek relative to P/LCT suggests the possibility of hybrid vigor with RT compared with P/LCT × RT hybrids. This may be because CAGT and RT are more closely related to each other than they are to cutthroat trout, so their offspring have greater fitness (e.g., Behnke 1992; McKay et al. 1996). Alternatively, even if the F_2 offspring do not generally have higher fitness, F_1 hybrids may have very high fitness, or a few individuals with a high proportion of non-native admixture may have very high fitness and reproductive success leading to dominance of CAGT and RT.

N_e and admixture analysis

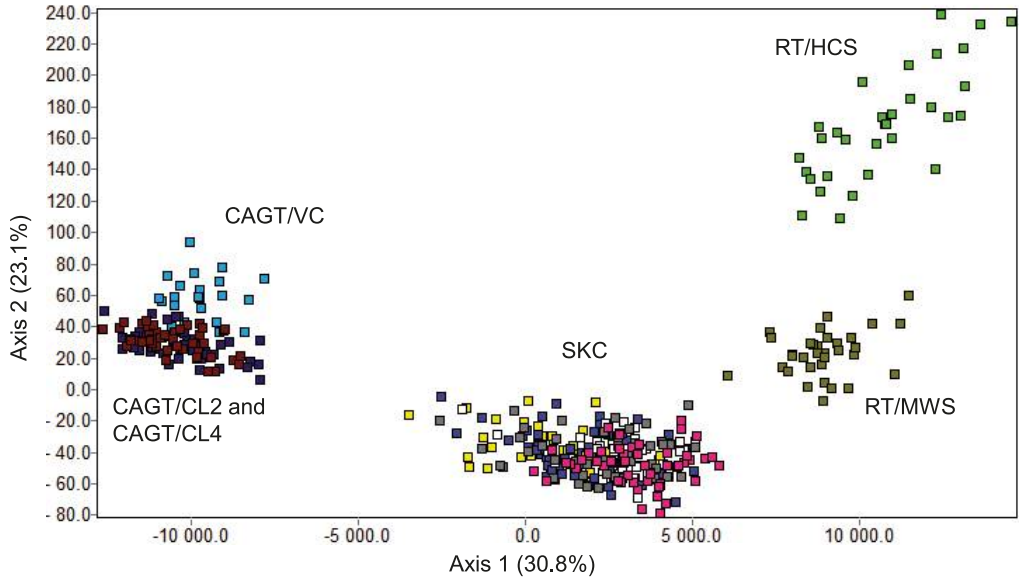
In jointly estimating admixture and N_e of the fish population in Silver King Creek, we used only the autosomal P/LCT SNPs, because the CAGT SNPs are not completely fixed. If species-specific markers were found for CAGT and RT, the method could be used to determine the admixed proportion of CAGT and provide an independent estimate of N_e. Because of uncertainty around the exact history of admixture, we analyzed the data under the two most extreme possible scenarios. The admixture event likely occurred at least three generations ago, but we need more data to be more precise. Given this caveat, the MLE of N_e in Silver King Creek based on the P/LCT SNPs is 150 under a scenario in which all CT alleles descend entirely from the most recent PCT planting (PCT in 1991). Under a different scenario in which all CT alleles descend exclusively from the original native PCT population, and the RT alleles were derived entirely from the 1930 planting of RT, the MLE of N_e is 750.

The census size (N_c) of all Silver King Creek fish below Llewellyn Falls, based on population data from 2000 (W. Somer, California Department of Fish and Game, personal communication, 2009), is estimated to be around 2000 fish. According to this estimation, the ratio of N_e/N_c is 0.08 and 0.38, respectively, under the two scenarios of MLEs of 150 and 750. These values are on the low and high ends of the range for vertebrate, invertebrate, and plant populations (Frankham 1995). In a review of contemporary N_e, all cate-

Table 8. Pairwise θ_{ST} values based on 12 microsatellites.

Location	SKC/LFV	SKC/LV	SKC/TC	SKC/CC	SKC/ASC	CAGT/VC	CAGT/CL2	CAGT/CL4	RT/HCS
SKC/LV	<0.01	—							
SKC/TC	0.04	0.03	—						
SKC/CC	<0.01	<0.01	0.03	—					
SKC/ASC	0.01	0.01	0.03	0.01	—				
CAGT/VC	0.20	0.21	0.24	0.23	0.22	—			
CAGT/CL2	0.16	0.16	0.20	0.18	0.20	0.21	—		
CAGT/CL4	0.17	0.17	0.21	0.19	0.20	0.23	0.01	—	
RT/HCS	0.15	0.15	0.15	0.14	0.15	0.30	0.28	0.29	—
RT/MWS	0.12	0.12	0.14	0.12	0.12	0.30	0.27	0.29	0.19

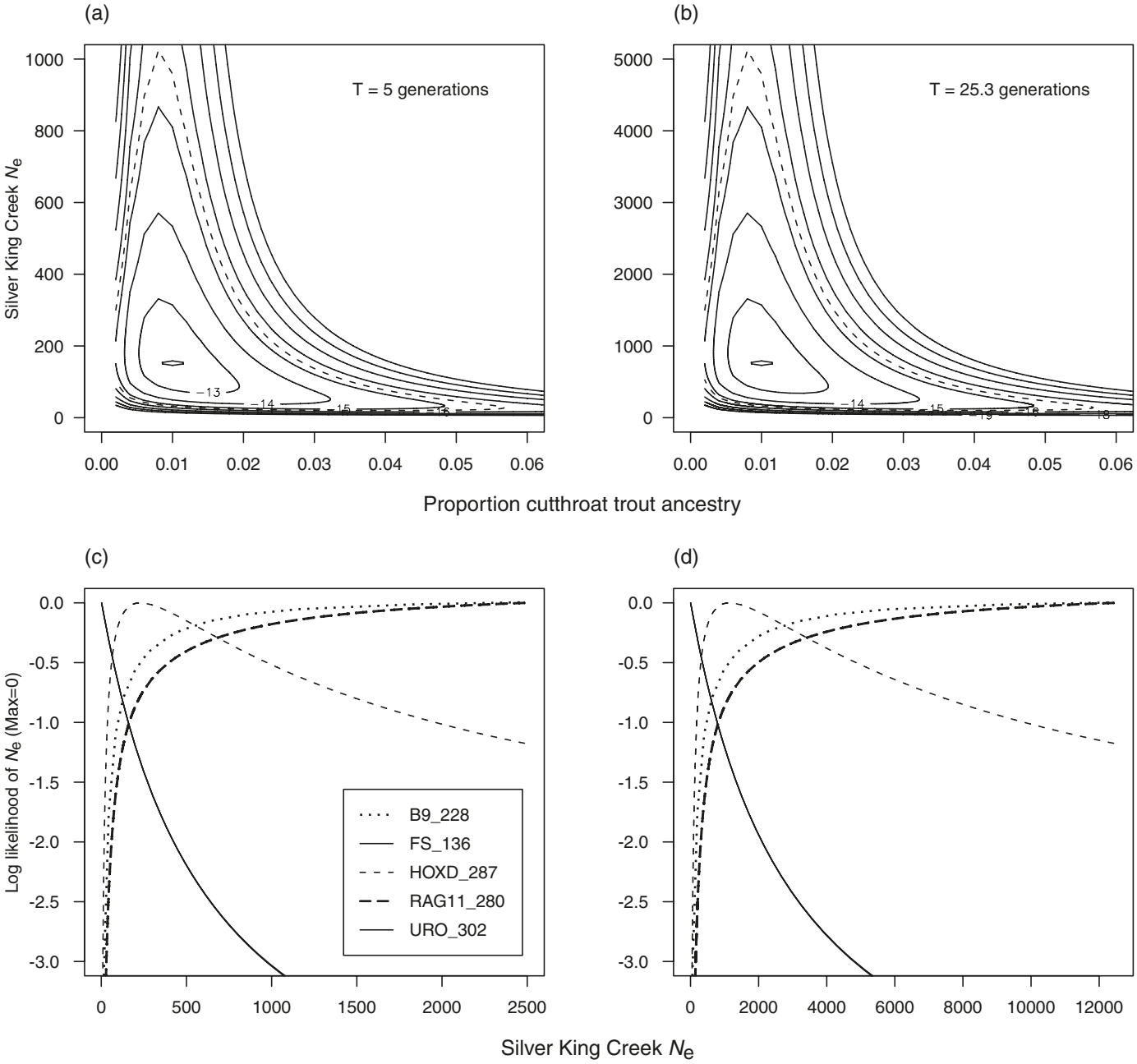
Fig. 4. Factorial correspondence analysis (FCA) analysis of samples genotyped with 12 microsatellites. Four distinct clusters are produced: (*i*) Silver King Creek (SKC) samples, consisting of Lower Fish Valley, Long Valley, Tamarack Creek, Canyon Sites 1 and 2, and SKC above Snodgrass Creek; (*ii*) California golden trout (CAGT) reference populations consisting of Cottonwood Lakes 2 and 4 (CAGT/CL2 and CAGT/CL4) and Volcano Creek (VC); (*iii*) Hot Creek Strain rainbow trout (RT/HCS); and (*iv*) Mount Whitney Strain rainbow trout (RT/MWS).



gories of populations (exploited, stable, and conservation) in all studies surveyed had an average N_e/N_c ratio of 0.19 ± 0.11 (Palstra and Ruzzante 2008). As our two estimates of N_e represent values from two bounding, extreme scenarios for the admixture history, it seems reasonable that they should be just within or slightly above the low and high endpoints of Palstra and Ruzzante's (2008) interval. It should be noted, however, that it is difficult to compare N_e to N_c with an estimate of N_c from only a single point in time. Unrecorded variation over time in N_c would likely drive N_e down, since the rate of inbreeding and allele frequency drift is strongly influenced by small values of population size; our estimate of N_e is an estimate of the harmonic mean of N_e over multiple generations. The harmonic mean of N_c over time would provide a better comparison with our estimated N_e than either a single estimate of N_c in time or the arithmetic mean of N_c over multiple generations (Kalinowski and Waples 2002).

Our method for jointly estimating N_e and the admixture proportion is, like many statistical–genetic methods, based on an idealized model that might be violated in practice. Two departures from the model that one might expect are (*i*) population structure or, equivalently, incomplete admixture of the two species; and (*ii*) non-neutrality of the genetic markers. We briefly consider the effects of each of these on our estimator. If the admixed population is structured so that some locales have more P/LCT than RT ancestry, for example, then the model assumption of independence between loci would be violated. The same would be true if mating between the species is assortative so that the "admixed population" includes a large number of individuals that are close to being pure P/LCT and others that are close to being pure RT. If the population is structured in this way, the first problem is that it is difficult to ensure that a representative sample has been drawn. Assuming, however, that a sample representative of the diversity in the structured population has been drawn, there are still two other effects that the population structure will have: (*i*) the model will overestimate the degree of precision it has in estimating the admixture proportion, because it assumes independence of allelic types between loci within individuals (which is not the case in a structured population); and (*ii*) the model will overestimate N_e because the correlation between alleles at different loci ensures that the interlocus variance in allele frequencies will be smaller than it would be if there were no LD induced by the population structure. Perhaps the simplest way to gain an intuitive under-

Fig. 5. (*a*, *b*) Contour plots of the joint likelihood surface for the admixture proportion of cutthroat trout in Silver King Creek and the N_e of the admixed population under two admixture history scenarios: (*a*) $T = 5$ and (*b*) $T = 25.3$ generations (see text). The dashed line shows the approximate 95% confidence interval. (*c*, *d*) Log-likelihood of N_e for individual loci assuming a cutthroat admixture fraction of 0.01. Values are scaled so the maximum of each curve is 0. (*c*) $T = 5$ generations scenario. (*d*) $T = 25.3$ generations. Of note, none of the curves show a log-likelihood difference greater than 2.0 from the maximum at the maximum likelihood estimate values of 150 and 750, respectively. The log-likelihood curves for FS_136 and URO_302 are entirely overlapping and indistinguishable; both follow the solid black line in the figure.



standing of this is by considering what would happen if every individual from the "admixed" population were either a pure PCT or pure RT. In that case, representative sampling will give an unbiased estimate of the mixing proportion of the two species, but the frequencies of the species-diagnostic alleles will be identical at every locus — the additional loci would provide no extra information for estimating the admixture fractions, and the interlocus variance in allele frequency will be zero, implying the N_e estimate should be infinite. The degree that population structure or assortative mating will bias the estimate of N_e is clearly determined by the correlation of allelic type within individuals. Since we found no evidence for significant LD between the five P/LCT SNPs in the Silver King Creek sample, we suspect that our results are not biased upward by this effect.

The assumptions of our model would also be violated if the SNPs used were linked to regions of the genome that were under selection in the admixed population. In that case, the allele frequencies at every locus would be affected by more than just genetic drift, and the inclusion of some loci

under selection could lead to an increased interlocus allele frequency variance that would bias estimates of N_e downward. In fact, selection's tendency to increase heterogeneity between loci is at the basis of many tests for selection (Lewontin and Krakauer 1973; Beaumont and Balding 2004; Foll and Gaggiotti 2008). The five P/LCT SNPs we used do not show evidence of heterogeneity beyond that expected by genetic drift, which suggests that natural selection has not necessarily affected our estimate of N_e. Had we used more loci, however, it is possible that some could have been linked to regions subject to strong selection. We point out that the likelihood model developed here provides an appropriate null model for detecting outlier loci in admixed populations. As more species- and subspecies-diagnostic SNPs become available and are mapped to the genome, we predict that our method, being fast and computationally efficient, can be incorporated into new tests that use recently admixed populations to identify gene regions involved in hybrid incompatibility, fitness in different environments, and, possibly, speciation.

The relatively high N_e estimates of 150 to 750 fish suggest that genetic drift is only a weak factor in removing genetic diversity over time. These fish have abundant heterozygosity with which to adapt to their environment. Analysis of parental admixture proportions and N_e will be particularly useful in situations of conservation interest when possible source populations must be examined for conservation value and priority (Allendorf et al. 2001), such as when introgression levels of native CT are deemed low enough to be acceptable or when non-introgressed native CT populations no longer exist. In such instances, inexpensive and efficient detection of hybridization is necessary, and knowledge of the demographic and management history of a population is of interest. The application of this method using these or other diagnostic SNPs can facilitate the ranking of populations' conservation utility; those with higher N_e and higher genetic diversity, as well as higher proportion of, for example, native cutthroat trout ancestry, may be better suited for restoration, reintroduction, or translocation. The discovery of fixed diagnostic SNPs for other taxa will enable a wider use of the simultaneous estimation method in such conservation situations.

Management and conservation implications

The biodiversity of native trout in western North America is declining, in large part because of introgression with hatchery fish planted over the last century (e.g., Trotter 2008; Behnke 1992). The result of this widespread stocking is hundreds of hybrid populations and the reduced viability of native fishes. Moreover, managers have limited options when attempting to restore native trout populations in a location where introgression has occurred. Complete restoration involves two steps: (*i*) removal or control of the spread of introgressed fish and (*ii*) restocking with an appropriate source of native fish. For removal or control of introgression, in smaller areas chemical treatment to remove introgressed fish may be the best option, given that once introgression has begun it is nearly impossible to stop. Chemical treatment requires complete removal of all fish present and subsequent monitoring to ensure that treatment success. A second management option, often undertaken when there are legal, political, or practical obstacles to chemical treatment, is to stop stocking fish altogether and hope that the natives recover. For example, in Montana, stocking of non-native RT stopped in 1969; however, there is anecdotal evidence that 70 000 RT were illegally released from a hatchery in 1997 (Muhlfeld et al. 2009*b*), and introgression with native westslope cutthroat trout has continued. There are several factors that may influence this introgression, such as straying rates (Boyer et al. 2008); neighborhood effects (Hitt et al. 2003); and stream temperature, stocking history, and land-use disturbance (Muhlfeld et al. 2009*c*). Regardless of how introgression spreads, it is expected to continue until there are no pure parental types (Hitt et al. 2003).

How managers decide to tackle problems of introgression is context-dependent. In some situations the only native populations of trout may be hybrid populations, and the best option may be to conserve the populations with the lowest proportion of non-native admixture. In contrast, there may be situations where nonhybridized populations remain but are geographically distant. Even if there are hybrid populations in close proximity, managers may wish to suppress or eradicate any hybrid populations, even those with very low non-native admixture.

Diagnostic SNP markers provide an excellent tool for managing introgression and restocking, both of which require considerable time and resources. In the foreseeable future the cost of designing assays will likely decline and throughput will increase, and there will be hundreds of diagnostic markers available. Microsatellites and allozymes have been more standard tools for characterizing hybridization, but it is much more difficult to find diagnostic markers. In addition, in cases where there are only a few individuals left of a population or taxon, SNPs can correctly classify individuals.

After discontinuing the stocking of non-native fish and the removal of introgressed fish, managers must select an appropriate donor source population of native fish for restocking. Initial misidentification of greenback cutthroat trout (*Oncorhynchus clarkii stomias*) sources used for restocking and propagation efforts resulted in the reintroduction of fish introgressed with closely related Colorado River cutthroat trout (*Oncorhynchus clarkii pleuriticus*, a Colorado subspecies of special concern; Metcalf et al. 2007). In addition, those greenback populations that were indeed "pure" had very low genetic diversity, increasing the risk of inbreeding. This example underscores the need for correct genetic characterization of donor sources in conservation situations.

In the case of Silver King Creek, despite the limited stocking of PCT in an interconnected headwater lake and the termination of stocking *O. mykiss* ssp. in 1976, our results suggest that the PCT has not and will not likely recover. For the purpose of reintroducing PCT to Silver King Creek, there is little reason to preserve the current population of trout below Llewellyn Falls, given such few P/LCT alleles and the existence of several refuge populations of non-introgressed PCT. Achieving the goal of successful restoration of PCT will require total removal of the trout in Silver King Creek to prevent future hybridization. Genetic monitoring will play a key role in re-establishing a self-sustaining population of PCT, allowing for early detection of hybridization due to unauthorized stocking events or unsuccessful hybrid eradication.

Acknowledgements

The authors thank William Somer for extensive consultation and the map of Silver King Creek. In addition the authors thank the California Department of Fish and Game and the United States Fish and Wildlife Service for sampling and project support. Thanks also go to the Threatened Trout Crew for samples and Karrigan Bork for valuable comments and editing. Funding for this project was provided through a Section 6 grant for Paiute trout habitat restoration by the US Fish and Wildlife Service, contract No. P0882011 with California Department of Fish and Game.

References

Allendorf, F.W., and Leary, R.F. 1988. Conservation and distribution of genetic variation in a polytypic species, the cutthroat trout. Conserv. Biol. **2**(2): 170–184. doi:10.1111/j.1523-1739.1988.tb00168.x.

Allendorf, F.W., Leary, R.F., Spruell, P., and Wenburg, J.K. 2001. The problems with hybrids: setting conservation guidelines. Trends Ecol. Evol. **16**(11): 613–622. doi:10.1016/S0169-5347(01)02290-X.

Allendorf, F.W., Leary, R.F., Hitt, N.P., Knudsen, K.L., Lundquist, L.L., and Spruell, P. 2004. Intercrosses and the U.S. endangered species act: Should hybridized populations be included as westslope cutthroat trout? Conserv. Biol. **18**(5): 1203–1213. doi:10.1111/j.1523-1739.2004.00305.x.

Anderson, E.C. 2005. An efficient Monte Carlo method for estimating N_e from temporally spaced samples using a coalescent-based likelihood. Genetics, **170**(2): 955–967. doi:10.1534/genetics.104.038349. PMID:15834143.

Beaumont, M.A., and Balding, D.J. 2004. Identifying adaptive genetic divergence among populations from genome scans. Mol. Ecol. **13**(4): 969–980. doi:10.1111/j.1365-294X.2004.02125.x. PMID:15012769.

Baumsteiger, J., Hankin, D., and Loudenslager, E.J. 2005. Genetic analyses of juvenile steelhead, coastal cutthroat trout, and their hybrids differ substantially from field identification. Trans. Am. Fish. Soc. **134**(4): 829–840. doi:10.1577/T04-114.1.

Behnke, R.J. 1965. A systematic study of the family Salmonidae with special reference to the genus *Salmo*. Ph.D. thesis, University of California, Berkeley, Calif.

Behnke, R.J. 1992. Native trout of western North America. American Fisheries Society, Bethesda, Md. AFS Monograph No. 6.

Belfiore, N.M., Hoffman, F.G., Baker, R.G., and Dewoody, J.A. 2003. The use of nuclear and mitochondrial single nucleotide polymorphisms to identify cryptic species. Mol. Ecol. **12**(7): 2011–2017. doi:10.1046/j.1365-294X.2003.01874.x. PMID:12803650.

Belkhir, K., Borsa, P., Chikhi, L., Raufaste, N., and Bonhomme, F. 1996–2004. GENETIX 4.05, logiciel sous Windows pour la génétique des populations. Laboratoire Génome, Populations, Interactions CNRS UMR 5000, Université de Montpellier II, Montpellier, France.

Bennett, S.N., Olson, J.R., Kershner, J.L., and Corbett, P. 2010. Propagule pressure and stream characteristics influence introgression: cutthroat and rainbow trout in British Columbia. Ecol. Appl. **20**(1): 263–277. doi:10.1890/08-0441.1. PMID:20349846.

Bensch, S., Akesson, S., and Irwin, D. 2002. The use of AFLP to find an informative SNP: genetic differences across a migratory divide in willow warblers. Mol. Ecol. **11**(11): 2359–2366. doi:10.1046/j.1365-294X.2002.01629.x. PMID:12406246.

Boyer, M.C., Muhlfeld, C.C., and Allendorf, F.W. 2008. Rainbow trout (*Oncorhynchus mykiss)* invasion and the spread of hybridization with native westslope cutthroat trout (*Oncorhynchus clarkii lewisi*). Can. J. Fish. Aquat. Sci. **65**(4): 658–669. doi:10.1139/f08-001.

Busack, C.A., and Gall, G.A.E. 1981. Introgressive hybridization in populations of Paiute cutthroat trout (*Salmo clarki seleniris*). Can. J. Fish. Aquat. Sci. **38**(8): 939–951. doi:10.1139/f81-127.

Chikhi, L., Bruford, M.W., and Beaumont, M.A. 2001. Estimation of admixture proportions: a likelihood-based approach using Markov chain Monte Carlo. Genetics, **158**(3): 1347–1362. PMID:11454781.

Cordes, J.F., Israel, J.A., and May, B. 2004. Conservation of Paiute cutthroat: the genetic legacy of population transplants in an endemic California salmonid. Calif. Fish Game, **90**(3): 101–118.

Cordes, J.F., Stephens, M.R., Blumberg, M.A., and May, B. 2006. Identifying introgressive hybridization in native populations of California golden trout based on molecular markers. Trans. Am. Fish. Soc. **135**(1): 110–128. doi:10.1577/T05-120.1.

Epifanio, J., and Nielsen, J. 2000. The role of hybridization in the distribution, conservation and management of aquatic species. Rev. Fish Biol. Fish. **10**(3): 245–251. doi:10.1023/A:1016729132297.

Evanno, G., Regnaut, S., and Goudet, J. 2005. Detecting the number of clusters of individuals using the software structure: a simulation study. Mol. Ecol. **14**(8): 2611–2620. doi:10.1111/j.1365-294X.2005.02553.x. PMID:15969739.

Excoffier, L., and Lischer, H.E.L. 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetic analyses under Linux and Windows. Mol. Ecol. Resour. **10**(3): 564–567. doi:10.1111/j.1755-0998.2010.02847.x. PMID:21565059.

Finger, A.J., Stephens, M.R., Clipperton, N.W., and May, B. 2009. Six diagnostic single nucleotide polymorphism markers for detecting introgression between cutthroat and rainbow trouts. Mol. Ecol. Resour. **9**(3): 759–763. doi:10.1111/j.1755-0998.2009.02532.x. PMID:21564737.

Foll, M., and Gaggiotti, O. 2008. A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. Genetics, **180**(2): 977–993. doi:10.1534/genetics.108.092221. PMID:18780740.

Frankham, R. 1995. Effective population size/adult population size ratios in wildlife: a review. Genet. Res. **66**(2): 95–107. doi:10.1017/S0016672300034455. PMID:18976539.

Gall, G.A.E., Baltodano, J., and Huang, N. 1988. Heritability of age at spawning for rainbow trout. Aquaculture, **68**(2): 93–102. doi:10.1016/0044-8486(88)90233-5.

Glaubitz, J.C. 2004. CONVERT: A user-friendly program to reformat diploid genotypic data for commonly used population genetic software packages. Mol. Ecol. Notes, **4**(2): 309–310. doi:10.1111/j.1471-8286.2004.00597.x.

Heath, D., Bettles, C.M., and Roff, D. 2010. Environmental factors associated with reproductive barrier breakdown in sympatric trout populations on Vancouver Island. Evol. Appl. **3**(1): 77–90. doi:10.1111/j.1752-4571.2009.00100.x.

Hitt, N.P., Frissell, C.A., Muhlfeld, C.C., and Allendorf, F.W. 2003. Spread of hybridization between native westslope cutthroat trout, *Oncorhynchus clarki lewisi*, and nonnative rainbow trout, *Oncorhynchus mykiss.* Can. J. Fish. Aquat. Sci. **60**(12): 1440–1451. doi:10.1139/f03-125.

Jakobsson, M., and Rosenberg, N.A. 2007. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics, **23**(14): 1801–1806. doi:10.1093/bioinformatics/btm233. PMID:17485429.

Kalinowski, S.T., and Waples, R.S. 2002. Relationship of effective to census size in fluctuating populations. Conserv. Biol. **16**(1): 129–136. doi:10.1046/j.1523-1739.2002.00134.x.

Kingman, J.F.C. 1982. On the genealogy of large populations. J. Appl. Prob. **19**: 27–43. doi:10.2307/3213548.

Lewis, P.O., and Zaykin, D. 2001. GENETIC DATA ANALYSIS: a computer program for the analysis of allelic data. Version 1.0 (d16c). Free program distributed by the authors, available from http://hydrodictyon.eeb.uconn.edu/people/plewis/software.php.

Lewontin, R.C., and Krakauer, J. 1973. Distribution of gene frequency as a test of the theory of the selective neutrality of polymorphisms. Genetics, **74**(1): 175–195. PMID:4711903.

Long, J.C. 1991. The genetic structure of admixed populations. Genetics, **127**(2): 417–428. PMID:2004712.

McKay, S.J., Devlin, R.H., and Smith, M.J. 1996. Phylogeny of Pacific salmon and trout based on growth hormone type-2 and mitochondrial NADH dehydrogenase subunit 3 DNA sequences. Can. J. Fish. Aquat. Sci. **53**(5): 1165–1176. doi:10.1139/f96-042.

Mercure, M., and Bruneau, A. 2008. Hybridization between the escaped *Rosa rugosa* (Rosacea) and native *R. blanda* in eastern North America. Am. J. Bot. **95**(5): 597–607. doi:10.3732/ajb.2007385. PMID:21632386.

Metcalf, J.L., Pritchard, V.L., Silvestri, S.M., Jenkins, J.B., Wood, J.S., Cowley, D.E., Evans, P.R., Shiozawa, D.K., and Martin, A.P. 2007. Across the great divide: genetic forensics reveals misidentification of endangered cutthroat trout populations. Mol. Ecol. **16**(21): 4445–4454. doi:10.1111/j.1365-294X.2007.03472.x. PMID:17727621.

Metcalf, J.L., Siegle, M.R., and Martin, A.P. 2008. Hybridization dynamics between Colorado's native cutthroat trout and introduced rainbow trout. J. Hered. **99**(2): 149–156. doi:10.1093/jhered/esm118. PMID:18238784.

Morin, P.A., Martien, K.K., and Taylor, B.L. 2009. Assessing statistical power of SNPs for population structure and conservation studies. Mol. Ecol. Resour. **9**(1): 66–73. doi:10.1111/j.1755-0998.2008.02392.x. PMID:21564568.

Muhlfeld, C.C., Kalinowski, S.T., McMahon, T.E., Taper, M.L., Painter, S., Leary, R.F., and Allendorf, F.W. 2009*a*. Hybridization rapidly reduces fitness of a native trout in the wild. Biol. Lett. **5**(3): 328–331. doi:10.1098/rsbl.2009.0033. PMID:19324629.

Muhlfeld, C.C., McMahon, T.E., and Belcer, D. 2009*b*. Spatial and temporal spawning dynamics of native westslope cutthroat trout (*Oncorhynchus clarkii lewisi*), introduced rainbow trout (*Oncorhynchus mykiss*), and their hybrids. Can. J. Fish. Aquat. Sci. **66**: 1153–1168. doi:10.1139/F09-073.

Muhlfeld, C.C., McMahon, T.E., Boyer, M.C., and Gresswell, R.E. 2009*c*. Local-habitat, watershed, and biotic factors in the spread of hybridization between native westslope cutthroat trout and introduced rainbow trout. Trans. Am. Fish. Soc. **138**(5): 1036–1051. doi:10.1577/T08-235.1.

Palstra, F.P., and Ruzzante, D.E. 2008. Genetic estimates of contemporary effective population size: What can they tell us about the importance of genetic stochasticity for wild population persistence? Mol. Ecol. **17**(15): 3428–3447. doi:10.1111/j.1365-294X.2008.03842.x. PMID:19160474.

Palti, Y., Fincham, M.R., and Rexroad, C.E. 2002. Characterization of 38 polymorphic microsatellite markers for rainbow trout (*Oncorhynchus mykiss*). Mol. Ecol. Notes, **2**(4): 449–452. doi:10.1046/j.1471-8286.2002.00274.x.

Pritchard, J.K., Stephens, M., and Donnelly, P. 2000. Inference of population structure using multilocus genotype data. Genetics, **155**(2): 945–959. PMID:10835412.

Rasmussen, J.B., Robinson, M.D., and Heath, D.D. 2010. Ecological consequences of hybridization between native westslope cutthroat (*Oncorhynchus clarkii lewisi*) and introduced rainbow (*Oncorhynchus mykiss*) trout: effects on life history and habitat use. Can. J. Fish. Aquat. Sci. **67**(2): 357–370. doi:10.1139/F09-191.

Raymond, M., and Rousset, F. 1995. GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. J. Hered. **86**(3): 248–249.

Rexroad, C.E., III, Coleman, R.L., Gustafson, A.L., Hershberger, W.K., and Killefer, J. 2002*a*. Development of rainbow trout microsatellite markers from repeat enriched libraries. Mar. Biotechnol. **4**(1): 12–16. doi:10.1007/s10126-001-0058-6. PMID:14961284.

Rexroad, C.E., III, Coleman, R.L., Hershberger, W.K., and Killefer, J. 2002*b*. Rapid communication: thirty-eight polymorphic microsatellite markers for mapping in rainbow trout. J. Anim. Sci. **80**(2): 541–542. PMID:11881939.

Rhymer, J.M., and Simberloff, D. 1996. Extinction by hybridization and introgression. Annu. Rev. Ecol. Syst. **27**(1): 83–109. doi:10.1146/annurev.ecolsys.27.1.83.

Rice, W.R. 1989. Analyzing tables of statistical tests. Evolution, **43**(1): 223–225. doi:10.2307/2409177.

Rosenberg, N.A. 2004. Distruct: a program for the graphical display of population structure. Mol. Ecol. Notes, **4**(1): 137–138. (Also available from http://rosenberglab.bioinformatics.med.umich.edu/distruct.html [modified 19 September 2010].) doi:10.1046/j.1471-8286.2003.00566.x.

Rubidge, E.M., and Taylor, E.B. 2005. An analysis of spatial and environmental factors influencing hybridization between native westslope cutthroat trout (*Oncorhynchus clarki lewisi*) and introduced rainbow trout (*O. mykiss*) in the upper Kootenay River drainage, British Columbia. Conserv. Genet. **6**(3): 369–384. doi:10.1007/s10592-005-4972-4.

Spies, I.B., Brasier, D.J., O'Reilly, P.T.L., Seamons, T.R., and Bentzen, P. 2005. Development and characterization of novel tetra-, tri-, and dinucleotide microsatellite markers in rainbow trout (*Oncorhynchus mykiss*). Mol. Ecol. Notes, **5**(2): 278–281. doi:10.1111/j.1471-8286.2005.00900.x.

Sprowles, A.E., Stephens, M.R., Clipperton, N.W., and May, B.P. 2006. Fishing for SNPs: a targeted locus approach for single nucleotide polymorphisms in rainbow trout. Trans. Am. Fish. Soc. **135**(6): 1698–1721. doi:10.1577/T05-291.1.

Stephens, M.R. 2007. Chapter 2: Comparative analysis of SNP and microsatellite estimates of rainbow trout introgression in native California golden trout. *In* Systematics, genetics and conservation of golden trout. Ph.D. dissertation, Graduate Group in Ecology, University of California, Davis, Calif.

Stephens, M.R., Clipperton, N.W., and May, B.P. 2009. Subspecies-informative SNP assays for evaluating introgression between native golden trout and introduced rainbow trout. Mol. Ecol. Resour. **9**(1): 339–343. doi:10.1111/j.1755-0998.2008.02407.x. PMID:21564644.

Templeton, A.R. 1986. Coadaptation and outbreeding depression. *In* Conservation biology: the science of scarcity and diversity. *Edited by* M.E. Soulé. Sinauer Associates, Sunderland, Mass. pp. 105–116.

Trotter, P. 2008 Cutthroat: native trout of the West. 2nd ed. University of California Press, Berkeley, Calif.

US Fish and Wildlife Service. 1975. Threatened status for three species of trout. Fed. Regist. **40**: 29863–29864.

US Fish and Wildlife Service. 2004. Revised recovery plan for the Paiute cutthroat trout (*Oncorhynchus clarki seleniris*). US Fish and Wildlife Service, Portland, Ore.

Wang, J. 2001. A pseudo-likelihood method for estimating effective population size from temporally spaced samples. Genet. Res. **78**(3): 243–257. doi:10.1017/S0016672301005286. PMID:11865714.

Wang, J. 2003. Maximum-likelihood estimation of admixture proportions from genetic data. Genetics, **164**(2): 747–765. PMID:12807794.

Wang, J. 2006. A coalescent-based estimator of admixture from DNA

sequences. Genetics, **173**(3): 1679–1692. doi:10.1534/genetics.105.054130. PMID:16624918.

Weir, B.S., and Cockerham, C.C. 1984. Estimating *F*-statistics for the analysis of population structure. Evolution, **38**(6): 1358–1370. doi:10.2307/2408641.

Williamson, K.S., Cordes, J.F., and May, B. 2002. Characterization of microsatellite loci in Chinook salmon (*Oncorhynchus tshawytscha*) and cross-species amplification in other salmonids. Mol. Ecol. Notes, **2**(1): 17–19. doi:10.1046/j.1471-8286.2002.00129.x.

Wong, D.M. 1975. Aspects of the life history of the Paiute cutthroat trout, *Salmo clarki seleniris* Snyder, in North Fork Cottonwood Creek, Mono County, California, with notes on behavior in a stream aquarium. M.S. thesis, California State University, Long Beach, Calif.

Appendix A

Figure A1 shows how our method is derived from the model of Chikhi et al. (2001). The conceptual model underlying their method (Fig. A1*a*) includes an admixed population h that is formed in a single admixture event. The allele frequencies at a locus in h at the time of admixture are denoted by p_h^a (the superscript a denotes "at the time of admixture"), which is a weighted average, $\gamma p_1^a + (1-\gamma)p_2^a$, of the allele frequencies in populations 1 and 2 at the time of admixture. γ is the admixture proportion from population 1. Subsequent to the admixture event, T generations elapse before genetic samples are taken from the three populations. At the time of sampling, the allele frequencies in the three populations are no longer p_h^a, p_1^a, and p_2^a; rather, each population has experienced genetic drift to an extent determined by the scaled time $t = \frac{T}{2N_e}$, resulting in allele frequencies p_h^s, p_1^s, and p_2^s (the superscript s denotes, "at the time of sampling"). This model can be represented more formally as a directed acyclic graph (DAG) (Fig. A1*b*), in which each unobserved variable in the model is represented by an unshaded node and each observed variable by a shaded node (for an overview of the use of DAGs to represent joint probability models, see Jordan 2004).

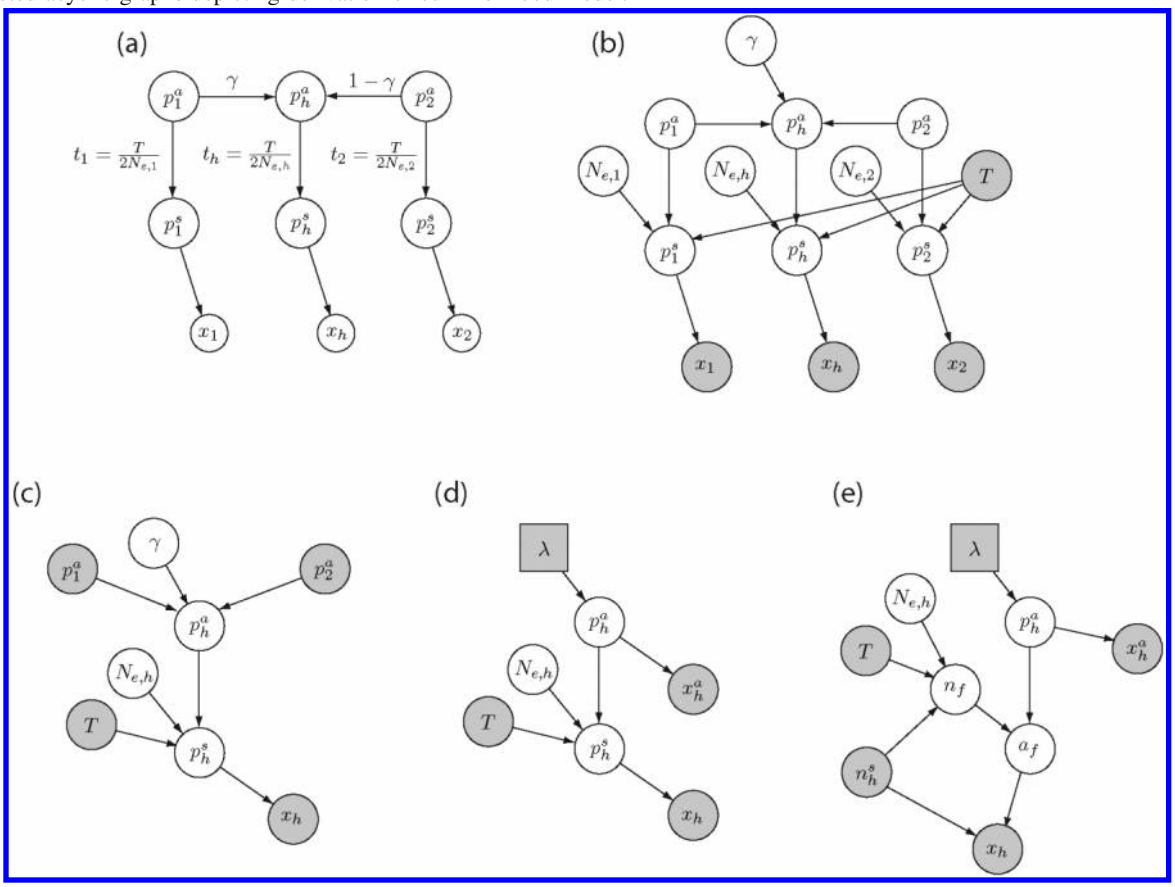
When using species-diagnostic SNP markers, the source populations are assumed fixed for alternate alleles, and hence the genetic drift in those populations is inconsequential. Accordingly, the allele frequencies in those populations at the time of admixture can be taken as observed quantities, simplifying the problem to the DAG shown in Fig. A1*c*. As before, p_h^a is exactly $\gamma p_1^a + (1-\gamma)p_2^a$. It is convenient to express this equality in terms of a probability density for p_h^a, which has all of its mass on the point $\gamma p_1^a + (1-\gamma)p_2^a$. Such a density is precisely the posterior density that would be obtained for p_h^a if an infinitely large sample were taken from the admixed population at the time of admixture and found to contain alleles in the proportions of $\gamma p_1^a + (1-\gamma)p_2^a$. We can therefore represent the model in Fig. A1*c* as the model in Fig. A1*d*, in which is endowed with a prior distribution that is a beta density parameterized by λ (we use $\lambda = (\frac{1}{2}, \frac{1}{2})$ — the unit-information beta prior), and x_h^a is defined to be a very large, imaginary sample of size m gene copies taken from the admixed population at the time of admixture with proportions of alleles observed given by $\gamma p_1^a + (1-\gamma)p_2^a$. As $m \to \infty$, the models in Fig. A1*d* and Fig. A1*c* become identical and, in practice, are indistinguishable for large values of m such as $m = 10\,000$.

The final step in our derivation is to consider the drift in the admixed population from the perspective of the coalescent. The model shown in Fig. A1*d* considers drift in the classical, forward-in-time sense: the allele frequency in the population drifts during the time between the admixture event and the event of sampling . A different way of incorporating the effect of genetic drift in that time interval is to consider its effect on the unobserved coalescent tree connecting the gene copies sampled in x_h, as was introduced by Berthier et al. (2002). Doing so leads to the DAG in Fig. A1*e*. In words, allelic counts x_h from n_h^s gene copies sampled from the admixed population are observed at the present time. These n_h^s gene copies descended from n_f gene copies present at the time of admixture. n_f is unobserved, but its distribution is governed by the coalescent process running back in time for T generations with coalescent rate determined by $N_{e,h}$. The distribution of allelic types in the present-day sample is determined by a_f — the allelic types of the n_f ancestral gene copies (it is assumed there is no mutation occurring within the T generations) — and the properties of the coalescent process. Finally, the allele frequencies at the time of admixture are considered unobserved, but are modeled as if we had a sample x_h^a of size m (with m large) gene copies taken at the time of admixture and we observe alleles in that sample in the proportions of $\gamma p_1^a + (1-\gamma)p_2^a$. It is important to note that x_h^a is a function of γ.

The above model is identical in structure to that in Anderson (2005) in which the likelihood for N_e is computed given two temporally spaced samples — an "historical sample" and a "present-day sample." Complete mathematical details of the Monte Carlo approximation of the likelihood can be found in that paper. We compute the likelihood using the program CoNe described by Anderson (2005). In practice with L diagnostic SNPs, to find the joint likelihood for any pair of values $(\gamma^*, N_{e,h}^*)$, we set $m = 10\,000$ and then create an input data set for CoNe in which the historical sample is of size m gene copies in allelic proportions $\gamma p_1^a + (1-\gamma)p_2^a$ and the present-day sample is just our sample x_h. For diallelic loci, the calculations done by CoNe are exact, so only a single Monte Carlo replicate needs to be done. With repeated calls to CoNe the likelihood over a fine grid of points in the space of γ and $N_{e,h}$ can be calculated in a few seconds, efficiently delivering the joint likelihood for γ and $N_{e,h}$. We construct a confidence interval for the joint estimate of γ and $N_{e,h}$ from this likelihood surface by using the contour that is three units below the likelihood at the maximum likelihood estimate. This approximately corresponds to a 95% confidence interval.

In the supplemental material[1], we present a series of simulations showing that our method gives good maximum likelihood estimates of γ and $N_{e,h}$ and that the confidence intervals on $N_{e,h}$ behave well. The supplement also considers the application of this admixture model to scenarios where there may have been recurrent admixture over time rather than a single admixture event, but great uncertainty about the exact history of admixture. In connection with this, it is worth noting that the degree of drift increases linearly with T (the number of generations assumed since admixture) and inversely with N_e. As a consequence, estimates of N_e made using one value of T can be directly scaled to another value of T (e.g., if one assumes T to be two times longer, than the estimate of N_e will be two times larger).

Fig. A1. Graphical depiction of the derivation of the likelihood model used in this paper. (*a*) The conceptual model of Chikhi et al. (2001). (*b–e*) Directed acyclic graphs depicting derivation of our likelihood model.



References

Anderson, E.C. 2005. An efficient Monte Carlo method for estimating N_e from temporally spaced samples using a coalescent-based likelihood. Genetics, **170**(2): 955–967. doi:10.1534/genetics.104.038349. PMID:15834143.

Berthier, P., Beaumont, M.A., Cornuet, J.M., and Luikart, G. 2002. Likelihood-based estimation of the effective population size using temporal changes in allele frequencies: a genealogical approach. Genetics, **160**(2): 741–751. PMID:11861575.

Chikhi, L., Bruford, M.W., and Beaumont, M.A. 2001. Estimation of admixture proportions: a likelihood-based approach using Markov chain Monte Carlo. Genetics, **158**(3): 1347–1362. PMID:11454781.

Jordan, M.I. 2004. Graphical models. Stat. Sci. **19**(1): 140–155. doi:10.1214/088342304000000026.