



Research

Cite this article: Kovach RP *et al.* 2016
Vive la résistance: genome-wide selection
against introduced alleles in invasive hybrid
zones. *Proc. R. Soc. B* **283**: 20161380.
<http://dx.doi.org/10.1098/rspb.2016.1380>

Received: 17 June 2016

Accepted: 25 October 2016

Subject Areas:

evolution, genomics

Keywords:

hybridization, natural selection, introgression,
invasive species, genomics, temperature

Author for correspondence:

Ryan P. Kovach

e-mail: rkovach@usgs.gov

Electronic supplementary material is available
online at <https://dx.doi.org/10.6084/m9.figshare.c.3573192>.

Vive la résistance: genome-wide selection against introduced alleles in invasive hybrid zones

Ryan P. Kovach¹, Brian K. Hand², Paul A. Hohenlohe³, Ted F. Cosart^{2,4},
Matthew C. Boyer⁵, Helen H. Neville⁶, Clint C. Muhlfeld^{1,2}, Stephen
J. Amish^{2,4}, Kellie Carim⁷, Shawn R. Narum⁸, Winsor H. Lowe⁴,
Fred W. Allendorf⁴ and Gordon Luikart^{2,4}

¹Northern Rocky Mountain Science Center, US Geological Survey, Missoula, MT 59802, USA

²Flathead Biological Station, University of Montana, Polson, MT 59860, USA

³Institute for Bioinformatics and Evolutionary Studies, Department of Biological Sciences, University of Idaho, Moscow, ID 83844, USA

⁴Fish and Wildlife Genomics Group, Division of Biological Sciences, University of Montana, Missoula, MT, USA

⁵Montana Fish, Wildlife, and Parks, Kalispell, MT 59901, USA

⁶Trout Unlimited, Boise, ID 83702, USA

⁷Wildlife Biology Program, University of Montana, Missoula, MT 59812, USA

⁸Hagerman Genetics Laboratory, Columbia River Inter-Tribal Fish Commission, Hagerman, ID 83332, USA

RPK, 0000-0001-5402-2123

Evolutionary and ecological consequences of hybridization between native and invasive species are notoriously complicated because patterns of selection acting on non-native alleles can vary throughout the genome and across environments. Rapid advances in genomics now make it feasible to assess locus-specific and genome-wide patterns of natural selection acting on invasive introgression within and among natural populations occupying diverse environments. We quantified genome-wide patterns of admixture across multiple independent hybrid zones of native westslope cutthroat trout and invasive rainbow trout, the world's most widely introduced fish, by genotyping 339 individuals from 21 populations using 9380 species-diagnostic loci. A significantly greater proportion of the genome appeared to be under selection favouring native cutthroat trout (rather than rainbow trout), and this pattern was pervasive across the genome (detected on most chromosomes). Furthermore, selection against invasive alleles was consistent across populations and environments, even in those where rainbow trout were predicted to have a selective advantage (warm environments). These data corroborate field studies showing that hybrids between these species have lower fitness than the native taxa, and show that these fitness differences are due to selection favouring many native genes distributed widely throughout the genome.

1. Introduction

Biological invasions are increasing due to environmental change and human translocation of species [1], and these invasions can have severe negative impacts on native species through ecological and evolutionary processes. Invasions can lead to hybridization between historically allopatric species, thereby producing complex conservation, evolutionary and legal consequences [2–5]. The negative consequences of hybridization between native and invasive species are numerous, including genomic extinction and loss of biodiversity [6], ecological and genetic homogenization [7], outbreeding depression [8], and increased invasiveness of the non-native taxon [9,10]. However, interspecific hybridization can also benefit native species through heterosis and adaptive introgression [11–13], facilitating colonization of novel environments [14], and adaptation to environmental change [15,16].

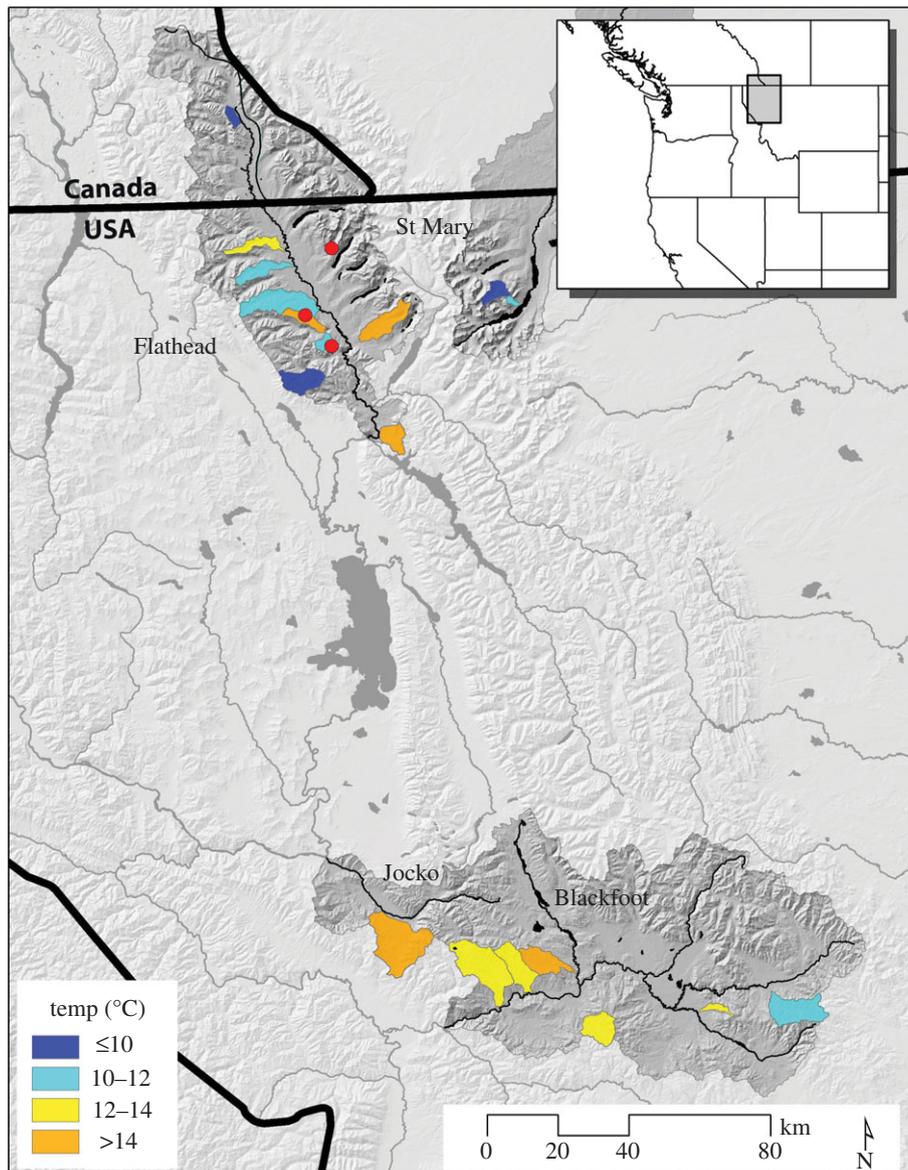


Figure 1. Map of the study areas, including the 21 populations across four major river basins (shaded dark grey) in Montana, USA, and British Columbia, Canada. Stream watersheds occupied by each population are colour-coded (see legend) based on maximum average August stream temperatures at each location. Lake populations are denoted by a red dot.

In practice, the fitness outcomes of invasive hybridization are variable, as the strength and direction of selection frequently vary across environmental gradients [17] and throughout the genome [18]. For example, environmental conditions often influence spatial and temporal dynamics of invasive introgression [19,20], as well as patterns of selection acting on hybrids [21]. Additionally, natural selection can favour introgression of single genes even when there is near-lethal selection against overall genome-wide admixture [11].

Because of the complex relationship between environmental variation, admixture and natural selection, there is limited understanding of the net adaptive consequences of invasive introgression in nature [12,22], and it has been particularly challenging to characterize genome-wide patterns of admixture and natural selection in non-model organisms. However, recent advances in genomics now make it feasible to assess genome-wide patterns of invasive introgression and selection within and between natural populations occupying diverse environments [23,24]. This richer understanding of the architecture and mechanisms influencing invasive introgression is necessary for evaluating the implications of

invasive hybridization and making informed conservation decisions [25].

The ecological, evolutionary and conservation problems associated with invasive hybridization are evident where native westslope cutthroat trout (*Oncorhynchus clarkii lewisi*) and invasive rainbow trout (*O. mykiss*) come into contact due to large-scale human releases of rainbow trout. Despite extensive hybridization where these species come into human-induced sympatry [26], intensive field studies suggest there can be strong selection against rainbow trout admixture [8,27]. However, preliminary genomic data indicate that selection may favour the introgression of several ‘super-invasive’ rainbow trout alleles [28]. Furthermore, other lines of evidence, including measurements of thermal tolerances [29], metabolic traits [30] and spatial patterns in admixture [20], suggest that rainbow trout and their hybrids may have fitness advantages in warmer conditions, potentially leading to adaptive introgression from invasive to native trout in warm environments.

We used genetic samples collected from 21 populations occupying thermally diverse habitats (figure 1; electronic

supplementary material, table S1) and 9380 species-diagnostic loci [31] to assess genome-wide patterns of admixture between rainbow trout and westslope cutthroat trout. We addressed three questions: (i) What are patterns of natural selection acting on invasive admixture? (ii) Do patterns of selection vary depending on temperature? (iii) Is natural selection acting for or against rainbow trout admixture concentrated or dispersed throughout the genome?

2. Material and methods

(a) Sample collection

We obtained tissue samples from juvenile trout in 21 streams and lakes located in Montana, USA and British Columbia, Canada. All sampling locations were chosen to represent local breeding populations. We strategically included locations spanning environmental gradients (i.e. putative selective pressures) thought to influence local adaptation in salmonid fishes [32–34], and spatial patterns of hybridization between rainbow trout and westslope cutthroat trout [20,35–37]. Specifically, we included populations inhabiting both lake and stream environments and also locations with widely varying temperatures (figure 1; electronic supplementary material, table S1).

In total, we obtained genetic samples from four independent river systems—the North Fork Flathead River, Blackfoot River, Jocko River and St Mary River—with the highest number of populations in the North Fork Flathead and Blackfoot River basins. The vast majority of populations (19 of 21) were located near the edge of the hybrid zone between native westslope cutthroat and invasive rainbow trout, where hybridization was present but admixture proportions were relatively low (proportion rainbow trout ancestry less than 0.25; electronic supplementary material, table S1). It is in these populations with low admixture where power was highest to detect rainbow trout alleles at higher than expected frequencies, potentially due to natural selection [38]. The remaining two populations were highly admixed, such that the original westslope cutthroat populations were genomically extinct [39,40]. Throughout this region, admixture began sometime in the early twentieth century (initiation of rainbow trout stocking), but some populations were probably pure westslope cutthroat trout until as late as the early twenty-first century [20,39]. Thus, admixture in any one population has been occurring for approximately 3–25 generations (assuming an approx. 5 year generation time).

(b) RAD genotyping

Detailed methods for identifying diagnostic loci between rainbow trout and westslope cutthroat trout are described elsewhere [31]. In short, we extended previous efforts [28,41] and used RAD sequencing to identify 16 788 fixed species-diagnostic SNP loci that differentiated 66 pure westslope cutthroat trout and 18 pure rainbow trout individuals of diverse origins. Specifically, we used RADseq to genotype 387 individuals from 21 populations containing various amounts of rainbow trout ancestry (electronic supplementary material, table S1). Libraries for RAD sequencing were prepped from genomic DNA according to standard protocols using the restriction enzyme *SbfI* and unique 6 bp barcodes for each sample [42]. We multiplexed RAD libraries (at equal concentration) in six sequencing lanes on an Illumina HiSeq machine. We conducted initial processing of the sequence data from all individuals into RAD loci using several modules from the STACKS software package, v. 1.19 [43]. We used *process_radtags* from STACKS to sort read pairs by barcode and remove any pairs in which the forward read did not contain both a correct barcode and the remaining six bases of the *SbfI* recognition sequence.

Paired-end reads from the same individual were used to identify PCR duplicates. The random shearing step in traditional RAD sequencing produces staggered paired-end reads, so that any set of read pairs with identical sequences across both the forward and reverse reads are likely PCR duplicates from a single original genomic DNA fragment [44], and were removed using the STACKS program *clone_filter*. The remaining read pairs from each individual were aligned to the rainbow trout reference genome [45]. We used the alignment software BOWTIE2 v. 2.1.0 [46] for end-to-end alignment without allowing gaps, and permitted one high-quality nucleotide mismatch per 20 bp.

Given the recent whole-genome duplication in salmonid fishes [47], we used multiple methods to identify diagnostic SNP loci from sequence differences in duplicated genomic regions (homeologues and paralogues) [28]. First, we retained only those loci that aligned uniquely to a single genomic location. We also enabled the Deleveraging and Removal algorithms in STACKS to filter out highly repetitive loci that were also likely paralogous [42]. We assigned diploid genotypes at each nucleotide position in each individual using the bounded maximum-likelihood method, with a minimum Phred quality-score threshold of 10 at each nucleotide, the upper bound of the sequencing error rate set to 0.01, and a likelihood ratio significance level of $\alpha = 0.05$.

For quality control, we further filtered the data to remove all individuals with greater than 75% missing data and all loci that were genotyped in fewer than eight individuals (after individuals with missing data were removed) in each population. Our final dataset included 339 individuals genotyped at 9380 putatively diagnostic loci (electronic supplementary material, table S1 and figure S1). We did not test for linkage disequilibrium or confirmation to Hardy–Weinberg proportions at these loci, because we expected extensive physical linkage and some non-random mating due to recent and ongoing admixture between the species.

(c) Data analysis

To minimize false positives, and the likelihood that false signals of natural selection were due to shared (ancestral or derived) polymorphisms [48], we used a highly conservative strategy for identifying loci that exhibited evidence for selection favouring or resisting RBT admixture. Relative to previous efforts [28], we used an extended baseline (panel of non-hybridized individuals) for identifying fixed diagnostic polymorphisms between rainbow trout and westslope cutthroat trout [31]. Nevertheless, pure westslope cutthroat trout populations are rare in many portions of their historical range; our westslope cutthroat trout individuals were therefore from a limited number of extant populations, largely originating in the South Fork Flathead River and the Clark Fork River. This limited our ability to detect low-frequency polymorphisms that are ancestral in westslope, thus we anticipated that alleles shared with rainbow trout would be present at some markers. Shared polymorphisms at putatively diagnostic loci can be mistaken for loci under selection. Indeed, after increasing the baseline ascertainment panel used previously to identify the larger set of putative diagnostic markers used in this study [31], several alleles that were candidates for selection favouring rainbow trout [29] were flagged as polymorphisms persisting in pure westslope cutthroat trout. Those loci were removed from the dataset.

We used Long's heterogeneity test [49] to identify outlier loci having rainbow trout allele frequencies that significantly deviated from random expectation in each population (see [28] for more details). Long's test accounts for sampling variation as well as genetic drift across loci, making it ideal for identifying loci potentially under selection in hybrid zones [38]. Rainbow trout alleles with higher frequencies relative to selectively neutral loci are candidates for selection favouring rainbow trout, whereas loci that are resistant to rainbow trout admixture have rainbow

trout allele frequencies that are lower than expected by chance (e.g. electronic supplementary material, figure S2). Samples from populations that were nearly pure rainbow trout or westslope cutthroat trout (electronic supplementary material, table S1) provide little information about positive or negative selection, respectively, acting on rainbow trout admixture. Therefore, we used reduced subsets of populations to test for positive and negative selection. Specifically, we grouped populations into three categories: (1) populations where we could detect positive selection favouring rainbow trout admixture (where proportion rainbow trout admixture was less than 0.9, $n = 19$); (2) populations occupying warm habitats (mean August temperature more than 12°C), where we could detect positive selection favouring rainbow trout admixture (a subset of Category 1, $n = 9$); and (3) populations where we could detect negative selection (resistance) acting against rainbow trout admixture (proportion rainbow trout admixture more than 0.1, $n = 8$). Seven of the eight locations where we could detect negative selection against rainbow trout using Long's test were in warm environments. We predicted that there would be more evidence for selection favouring rainbow trout admixture in warm environments.

Within each category, we used a Fisher's combined test to calculate an overall test statistic and associated probability that a particular marker deviated from neutral expectation. p -values were adjusted based on the false discovery rate (a Benjamin and Yekutieli (BY)-adjusted p -value [50,51]) and $\alpha = 0.05$. We further required that each candidate locus with a BY-adjusted p -value of less than 0.05 demonstrate consistent outlier behaviour in multiple populations and river basins. For the entire set of populations where we could detect selection favouring rainbow trout (i.e. category 1, above), we required that outliers demonstrate evidence ($p < 0.05$ without correction for multiple tests) of directional selection favouring rainbow trout in at least one population from the North Fork Flathead River and one from the Blackfoot River, and directional selection for rainbow trout in at least four populations overall (i.e. 21% of populations).

We used similar criteria for directional selection favouring rainbow trout in warm environments (category 2), but reduced the overall number of populations where a locus had to exhibit evidence for selection favouring rainbow trout to two (i.e. 22% of populations in warm environments). Although we required fewer populations for the latter tests, the actual proportion of populations where we required that each locus exhibit a consistent pattern of selection was nearly identical: 21% for all populations, 22% for warm environments only. The same criteria were used for populations where we could detect selection favouring westslope cutthroat trout (category 3): there had to be evidence for resistance to rainbow admixture in at least one population from each of the North Fork Flathead and Blackfoot river systems (i.e. consistent evidence for selection had to be detected in 25% of populations). By requiring evidence for directional selection in both the North Fork Flathead and Blackfoot rivers, we attempted to minimize error associated with shared ancestral polymorphisms.

SNP loci with significant Fisher's combined p -values that were outliers in multiple populations in the North Fork Flathead or the Blackfoot, but not both, were flagged as loci that may have shared polymorphisms and were removed from subsequent genomic cline analyses. Some of these markers may actually be diagnostic (not shared) and subject to natural selection, but we opted to be conservative in the face of various complexities that could produce false signals of selection.

As a complementary method to Long's heterogeneity test, we used genomic clines to identify SNP loci potentially under natural selection [52]. Unlike Long's test, which is unable to detect selection favouring rainbow trout or westslope cutthroat trout in population samples where overall rainbow trout admixture proportions were near 0 or 1, genomic cline analyses use locus-specific and overall

genome-wide admixture proportions across individuals within each population to identify signatures of natural selection. Specifically, a logit-logistic model was used to fit locus-specific rainbow trout allele frequencies against overall genome-wide admixture across all 21 populations. The same analysis was also conducted for the subset of all populations inhabiting warm environments ($n = 11$; electronic supplementary material, table S1). Outliers were identified based on unusual combinations for the estimated parameters defining the slope (u) and cline centre (v) for each locus relative to the observed distribution of slope and cline-centre estimates across all loci. In general, steeper clines indicate loci that are resistant to admixture (i.e. selection acting against rainbow trout admixture), whereas shallower, intermediate clines represent rainbow trout alleles potentially favoured by selection [52]. We also examined locus-specific allele frequencies for each outlier cline to confirm the general pattern of selection. The R package Hiest was used to fit genomic clines and identify statistically significant outlier loci. All p -values were adjusted using the BY-adjusted false discovery rate and $\alpha = 0.05$.

Although Long's tests and genomic cline analyses are similar, there are fundamental differences between the methods, including the unit of analysis (i.e. differences in allele frequencies within a population versus differences in cline shape across populations, respectively) and the underlying data used for each analysis (i.e. the set of populations). We also used more restrictive criteria when evaluating Long's tests. Outlier loci from Long's tests needed to have consistent signatures of selection across all populations and evidence for selection in independent river basins; the same criteria were not applied to the genomic cline analyses. Thus, results from the two approaches are supportive rather than overlapping.

We used multiple tests to assess overarching patterns of selection acting across the genome and across populations. We used χ^2 -tests to compare the total proportion of outlier loci that were candidates for selection favouring or resisting rainbow trout admixture using data from all appropriate populations, and for those populations inhabiting warm environments. Chi-square tests were performed on results from Long's test and the genomic cline analysis individually, and all outlier loci combined (i.e. identified from both tests). For Long's tests, we considered the total proportion of populations where an SNP exhibited evidence for favouring or resisting selection as a measure of strength of evidence (e.g. more populations is stronger evidence).

A nucleotide BLAST using sequence data for each candidate RAD locus and the rainbow trout genome were used to identify whether outlier SNPs were located within or proximate to genes of known function. We compared the RAD loci alignment coordinates with start and end coordinates for the 46 585 predicted gene annotations in the rainbow trout genome [45]. We collected the 177 gene annotations whose start and/or end coordinates either overlapped or were within 50 kbp of the start or end coordinate of the outlier RAD loci. The degree of linkage disequilibrium (LD) between RAD markers and functional loci has not been quantified, and the state of the genome assembly did not allow us to estimate LD extent among our RAD markers. However, we anticipate a relatively large extent of LD across physically linked loci for westslope cutthroat trout versus rainbow trout alleles, given the recent history of admixture [53], so that 50 kb is expected to be relatively conservative for identifying potential genes. We required not only that each locus and gene were on the same chromosome, but that they also shared the same scaffold or contig, because the assembled chromosomes contain sequence gaps of unknown size. To find detailed descriptions for the 177 predicted genes in the genome annotations, we used mRNA sequences corresponding to the coding sequence (CDs) and untranslated regions (UTRs) of the genes as queries in a BLASTn alignment [54]. For each query, we collected the BLAST hit with the highest bitscore. When more than one hit had the highest bitscore, we collected the one that was first in BLAST output order.

Table 1. The number of outlier SNPs where there was evidence for selection favouring or resisting rainbow trout (RBT) admixture. Outlier SNPs are grouped according to statistical methods used to detect selection: Long's heterogeneity tests (LT), genomic cline analysis for all populations (GC), genomic cline analysis for populations inhabiting warm environments (GC warm), all unique SNPs detected as outliers in any analysis (Un) (i.e. at least one analysis including LT, GC or GC warm), and the subset of SNPs that were consistently detected as significant outliers in all three analyses (All).

direction of selection	LT	GC	GC warm	Un	All
favouring RBT	8	4	5	13	0
resisting RBT	49	20	26	67	10

3. Results

(a) Patterns of natural selection acting on invasive admixture

There was consistent evidence for selection against rainbow trout admixture across different populations and hybrid zones (table 1 and figure 2). Rainbow trout alleles at 49 SNP loci (table 1) occurred at lower than expected frequencies—evidence for selection acting against admixture—within multiple independent populations (figure 2; electronic supplementary material, figure S2). Selection against rainbow trout admixture was detected at 22 SNP loci in each of three of the eight populations where we could detect selection favouring westslope cutthroat trout (38% of populations); 15 SNPs were outliers in four of those eight populations (50% of populations) and four SNPs were outliers in five to seven populations (63–88% of populations; figure 2).

Similarly, there were 20 SNPs where rainbow trout allele frequencies were significantly lower than the genome-wide average across populations, as represented by genomic clines (figure 3*a*). Patterns of selection against rainbow admixture generally took two forms. Most outlier loci had lower than expected rainbow trout allele frequencies across a range of genome-wide admixture values (the 'shallow' clines in the lower right quadrant of figure 3). A subset, however, were fixed for the rainbow trout allele in the two highly admixed populations (Gold Creek and Abbott Creek), but the rainbow trout allele was largely absent in other populations—the 'steep' clines centred at genome-wide proportion rainbow trout admixture (pRBT) \sim 0.15 (figure 3). The five loci with this second pattern were all flagged as outliers using Long's test because they had lower than expected frequencies of the rainbow trout allele in 38–63% of populations. There were no populations where the rainbow trout allele was at higher than expected frequencies, suggesting that this pattern reflects, in part, the lack of samples from populations with pRBT between 0.2 and 0.9.

There were considerably fewer SNPs that exhibited patterns suggestive of selection favouring rainbow trout admixture (table 1). Using Long's heterogeneity test and data from all populations where we could detect directional selection favouring rainbow trout admixture ($n = 19$ populations), eight SNPs satisfied our criteria (e.g. at least one population from two drainages showing selection) for selection favouring the rainbow trout admixture (table 1). Four of those SNPs were significant outliers in four populations (21% of populations), two were significant outliers in five populations (26% of

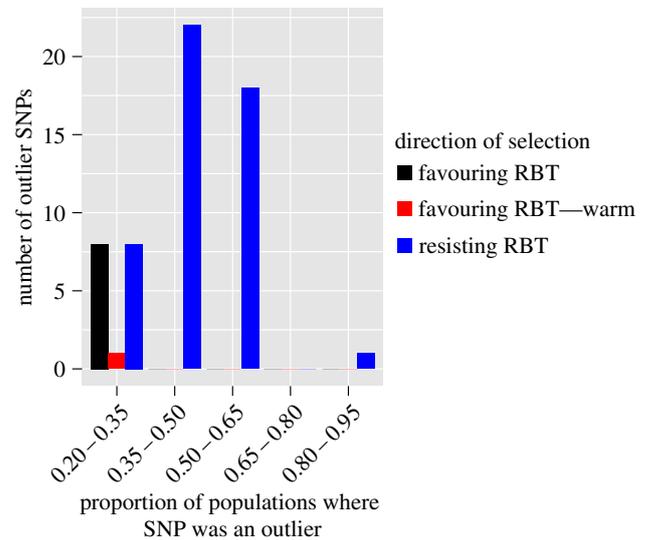


Figure 2. The number of outlier SNPs revealing evidence of selection favouring (red and black bars), or resisting (blue bars) rainbow trout (RBT) admixture. Counts of outliers in each category are binned according to the proportion of populations where each SNP was identified as an outlier (i.e. higher proportions represent SNPs that were identified as a candidate for selection in a larger proportion of populations). Results from outlier tests focused on detecting selection favouring rainbow trout are reported for all populations (black bars), and a subset of nine populations occupying warm environments (more than 12°C mean August temperature; red bars).

populations) and two were significant outliers in six populations (32% of populations; figure 2). When using genomic data from all 21 populations, there were four SNP loci with significant genomic clines suggestive of selection favouring rainbow trout admixture (i.e. the rainbow trout allele was at higher than expected frequencies across genome-wide values of pRBT).

Regardless of method, populations used in the analyses or thermal environment, there were substantially more outlier SNPs where selection appeared to resist rainbow trout admixture (table 1). In all cases, the overall proportion of SNPs where there was evidence for selection opposing rainbow trout alleles was substantially higher than the proportion of SNPs with rainbow trout alleles that were favoured by selection (all pair-wise tests across analyses $p < 0.0001$). SNPs that were candidates for directional selection acting against rainbow trout were identified as outliers in more populations than SNPs that were candidates for selection favouring rainbow trout (figure 2). The subset of SNPs that were consistently identified as outliers in each analysis ($n = 10$) were all candidates for selection acting against rainbow trout (table 1).

(b) Patterns of natural selection in warm habitats

Patterns of selection among populations inhabiting warm environments mirrored those obtained from populations inhabiting the entire thermal spectrum. We only detected one SNP where there was evidence for selection favouring rainbow trout in populations inhabiting warm environments ($n = 9$ populations; figure 2). However, this same SNP was also identified as an outlier using the data from all populations, suggesting that there were no rainbow trout alleles potentially favoured by selection specifically in warm environments. Five

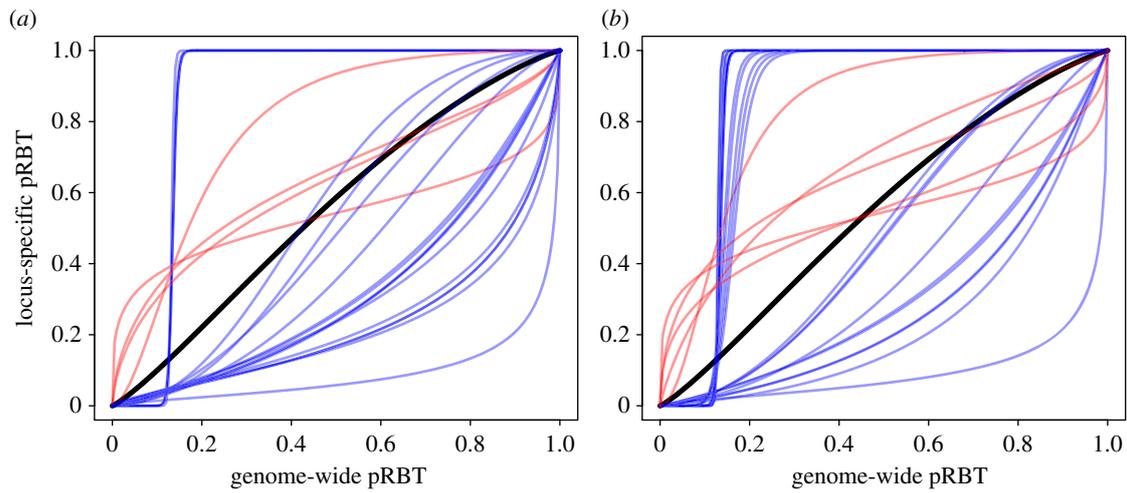


Figure 3. Locus-specific genomic clines that exhibited outlier patterns of introgression representative of selection favouring (red) and resisting (blue) rainbow trout admixture across (a) all populations, and (b) rainbow trout admixture for a subset of the nine populations inhabiting warm environments (more than 12°C mean August temperature). pRBT is the individual-level proportion of rainbow trout admixture. The bold black genomic cline represents genome-wide average across all putatively neutral loci.

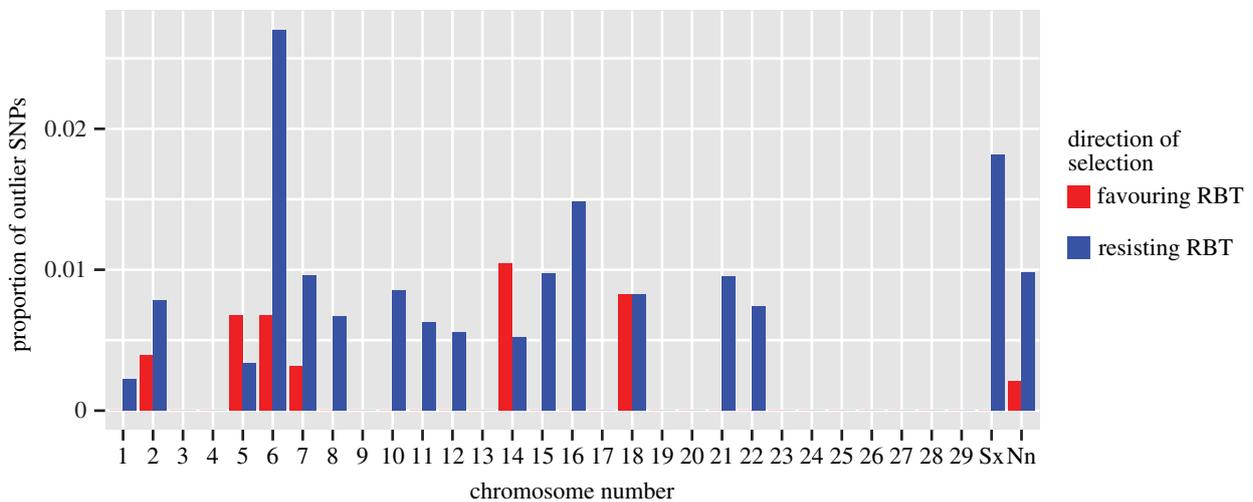


Figure 4. The distribution of outlier SNPs favouring (red) or resisting (blue) rainbow trout (RBT) admixture across the genome (i.e. 30 chromosomes). The proportion of outlier SNPs is relative to the number of SNPs mapped to each chromosome, and includes all unique SNPs identified across all analyses (table 1). The labels on the x-axis represent the 29 autosomal chromosomes, sex chromosomes (Sx) and RAD tags on scaffolds that are not yet mapped to specific chromosomes (Nn). There is no genome assembly for chromosome 25.

genomic clines were suggestive of natural selection favouring rainbow trout admixture in warm environments. By contrast, 26 clines were suggestive of selection opposing rainbow trout admixture (figure 3b), with a subset ($n = 12$) of these 26 outlier clines exhibiting steep slopes centred at pRBT ~ 0.2 , all of which appeared to be under selection resisting rainbow trout admixture similar to the pattern described above (i.e. fixed for the rainbow trout allele in highly admixed populations but absent from other upstream locations). For these loci, Long's tests confirmed that rainbow trout allele frequencies were at lower than expected frequencies within populations, and were never at higher than expected frequencies within any population (i.e. there was evidence for selection opposing rainbow trout admixture).

In summary, there was no evidence for higher rates of adaptive rainbow trout introgression in the warmest streams and lakes. Instead, the proportion of SNPs with rainbow trout alleles potentially favoured by selection in the warmest environments was nearly identical to (genomic cline analyses, figure 3b; $p = 0.999$) or slightly less than (Long's heterogeneity

tests, figure 2; $p = 0.045$) the proportion of SNPs with rainbow trout alleles potentially favoured by selection in any thermal environment.

(c) Genomic distribution and potential functionality of outlier loci

Rainbow trout segments where selection resisted admixture were distributed widely across chromosomes (mapped to 16 of the 28 assembled autosomal chromosomes and to the pair of sex chromosomes), as well as the unmapped portions of the genome (figure 4). There was some evidence for concentrated selection against rainbow trout segments on chromosome 6, but the majority of outlier SNPs mapped to known chromosomes were found on globally ordered but non-anchored portions of the genome [31], making it impossible to identify specific chromosomal regions potentially under selection. Forty of the rainbow trout genome segments potentially under selection (favouring or resisting rainbow trout admixture) were located in or near 126 described gene annotations (electronic supplementary

material, table S2). Specifically, 27 of the RAD loci aligned completely inside gene annotations. The rest ($n = 13$) were within 50 kbp of annotated genes. Highest-scoring BLAST alignments generally yielded *Salmo salar* sequences (121 out of 126). After removing redundant gene annotations, there were 106 unique gene descriptions with potential functions that varied widely from cellular metabolism to gene expression to disease resistance (electronic supplementary material, table S2).

4. Discussion

Genomic data from hybridizing populations of native and invasive trout strongly suggest that natural selection consistently acts against non-native alleles throughout the genome, across populations and in different environments. The consequences of invasive hybridization, particularly fitness outcomes, are notoriously complicated and frequently controversial because it has been challenging to describe patterns of natural selection acting on non-native genetic introgression [4,22]. Here, we show the applied value of genomics for illuminating patterns of selection acting in invasive hybrid zones and further clarify the implications of hybridization for this species of conservation concern. More generally, these data underscore that widespread admixture between native and non-native species can largely reflect large-scale propagation and releases of the non-native species, not positive selection favouring the invasive taxon.

Consistent evidence for selection acting against invasive rainbow trout genomic segments throughout the genome corroborates detailed field studies showing reduced fitness in rainbow trout and their hybrids. In one population, fish with proportion rainbow trout ancestry more than 0.2 suffered from nearly a 50% reduction in reproductive success relative to pure westslope cutthroat trout [8]. Similarly, selection coefficients (s) acting against rainbow trout genotypes were high ($s \sim 0.60$) in two streams with contrasting thermal and hydrologic regimes [27]. Data presented here suggest that selection consistently acts against rainbow trout across the genome in other populations and—more importantly—separate river basins with independent invasion histories.

This pattern of selection against rainbow trout admixture is striking given that we strategically sampled warm environments in multiple independent river basins. Spatial patterns of admixture and laboratory thermal tolerances suggest that rainbow trout may have a selective advantage under warmer temperatures. However, we found no evidence that there were higher rates of adaptive introgression in the warmest environments. The data presented here and in [27] highlight that spatial patterns in admixture should be interpreted carefully with respect to selection. Higher levels of admixture in low-elevation, warmer environments [35–37] do not necessarily imply that rainbow trout and hybrids have higher relative fitness than westslope cutthroat trout. This pattern may simply indicate that selection acting against rainbow trout is not as severe or consistent in these environments, or that human releases of rainbow trout in certain areas overwhelmed selective regimes. Rainbow trout were stocked routinely and abundantly throughout the early part of the twentieth century, and these stocking events occurred predominantly in lower elevation rivers and lakes that often are warmer than those at higher elevation [37,55,56], suggesting that spatial patterns in rainbow trout admixture probably reflect stocking

and subsequent dispersal, more than selection favouring rainbow trout in certain environments [39,57].

We detected positive selection favouring rainbow trout admixture at several loci, but there were no loci where rainbow trout alleles were consistently swept to high frequency across multiple populations. Given the strong selection regimes that appear to be acting against rainbow trout admixture, it may take many generations of recombination before beneficial rainbow trout alleles can successfully overcome genome-wide selection against rainbow trout admixture [58,59]. Furthermore, most adaptive traits are highly polygenic and, thus, most adaptive SNPs are presumably of relatively small effect [18,60], which probably impedes rapid adaptive introgression from rainbow trout to westslope cutthroat trout during early stages of invasive admixture.

Hybridization can act as a valuable source of evolutionary novelty [61], and a stimulus for species invasions, particularly in plants [9]. However, evidence for adaptive introgression between native and invasive vertebrate species remains scarce [12] (but see [38]). Using an extensive baseline of genetically pure individuals from both species during diagnostic SNP discovery and a conservative approach to identify loci potentially under selection, we found that rainbow trout alleles previously inferred to be favoured by selection [28] appear to be shared polymorphisms (alleles) within both species (see Material and methods). Separating signals of selection from shared polymorphisms remains a significant challenge in hybridization genomics, even for model systems (e.g. hominids) [48]. We caution that patterns of adaptive introgression should be interpreted carefully, and recommend that future genomic studies use conservative methods including robust baselines for initial SNP discovery and replication across independent hybrid zones to avoid false positives due to shared polymorphisms between species.

This work has several caveats, which represent opportunities for future research. First, the rainbow trout reference genome remains a work in progress, the majority of which is anchored to chromosomes with unknown order, or is entirely unmapped [45]. A more complete genome assembly is needed to better delineate specific regions of the genome under natural selection, and thereby clarify the genomic basis of fitness differences between species. For example, this would help clarify whether the signal of selection on chromosome 6 is highly concentrated (i.e. one or several genes of large effect resulting in selection acting on a single rainbow trout haplotype block), or dispersed. It also remains unclear whether the signature of selection favouring westslope cutthroat trout represents extrinsic selection acting on portions of the genome associated with local adaptation, or consistent intrinsic selection resulting from hybrid incompatibilities between species [62,63]. Finally, we filtered our data to remove potentially paralogous SNPs during SNP discovery and mapping [31], a necessary strategy, but one with unknown consequences for understanding genomic signatures of selection in organisms like salmonid fishes with residual tetraploidy in their genomes [24,47]. Future progress on the rainbow trout reference genome and emerging methodologies for discovery and genotyping of SNPs in paralogous regions [64] will provide additional insight into selection in this system and many others.

Although adaptive introgression is gaining interest as a process that may have some positive evolutionary and conservation outcomes [12,65], the consequences of human-mediated hybridization are likely to vary widely and

remain challenging to predict [4,22,59]. Our data highlight that native genotypes are often favoured by natural selection [66,67], and thus management strategies focused on minimizing the deleterious impacts of invasive hybridization here and in similar contexts are prudent until an adaptive basis of introgression can be established. This study shows that combining genome-wide data from multiple populations on the leading edge of an invasive hybrid zone can help elucidate overarching patterns of selection across the genome. More specifically, our findings provide strong genome-scale evidence of selection against invasive hybridization across environmentally heterogeneous wild populations.

Data accessibility. The genotypes used in all analyses can be found at Dryad: <http://dx.doi.org/10.5061/dryad.ss035>.

References

1. Simberloff D *et al.* 2013 Impacts of biological invasions: what's what and the way forward. *Trends Ecol. Evol.* **28**, 58–66. (doi:10.1016/j.tree.2012.07.013)
2. Campton DE, Bass AL, Chapman FA, Bowen BW. 2000 Genetic distinction of pallid, shovelnose, and Alabama sturgeon: emerging species and the US Endangered Species Act. *Conserv. Genet.* **1**, 17–32. (doi:10.1023/A:1010121417487)
3. Allendorf FW, Leary RF, Hitt NP, Knudsen KL, Lundquist LL, Spruell P. 2004 Intercrosses and the US Endangered Species Act: should hybridized populations be included as westslope cutthroat trout? *Conserv. Biol.* **18**, 1203–1213. (doi:10.1111/j.1523-1739.2004.00305.x)
4. Allendorf FW, Leary RF, Spruell P, Wenburg JK. 2001 The problems with hybrids: setting conservation guidelines. *Trends Ecol. Evol.* **16**, 613–622. (doi:10.1016/S0169-5347(01)02290-X)
5. Mooney HA, Cleland EE. 2001 The evolutionary impact of invasive species. *Proc. Natl Acad. Sci. USA* **98**, 5446–5451. (doi:10.1073/pnas.091093398)
6. Rhymer JM, Simberloff D. 1996 Extinction by hybridization and introgression. *Annu. Rev. Ecol. Syst.* **27**, 83–109. (doi:10.1146/annurev.ecolsys.27.1.83)
7. Olden JD, LeRoy Poff N, Douglas MR, Douglas ME, Fausch KD. 2004 Ecological and evolutionary consequences of biotic homogenization. *Trends Ecol. Evol.* **19**, 18–24. (doi:10.1016/j.tree.2003.09.010)
8. Muhlfeld CC, Kalinowski ST, McMahon TE, Taper ML, Painter S, Leary RF, Allendorf FW. 2009 Hybridization rapidly reduces fitness of a native trout in the wild. *Biol. Lett.* **5**, 328–331. (doi:10.1126/science.277.5325.494)
9. Ellstrand NC, Schierenbeck KA. 2000 Hybridization as a stimulus for the evolution of invasiveness in plants? *Proc. Natl Acad. Sci. USA* **97**, 7043–7050. (doi:10.1073/pnas.97.13.7043)
10. Schierenbeck KA, Ellstrand NC. 2009 Hybridization and the evolution of invasiveness in plants and other organisms. *Biol. Invasions* **11**, 1093–1105. (doi:10.1007/s10530-008-9388-x)
11. Song Y, Endepols S, Klemann N, Richter D, Matuschka FR, Shih CH, Nachman MW, Kohn MH. 2011 Adaptive introgression of anticoagulant rodent poison resistance by hybridization between old world mice. *Curr. Biol.* **21**, 1296–1301. (doi:10.1016/j.cub.2011.06.043)
12. Hedrick PW. 2013 Adaptive introgression in animals: examples and comparison to new mutation and standing variation as sources of adaptive variation. *Mol. Ecol.* **22**, 4606–4618. (doi:10.1111/mec.12415)
13. The Heliconious Genome Consortium. 2012 Butterfly genome reveals promiscuous exchange of mimicry adaptations among species. *Nature* **487**, 94–98. (doi:10.1038/nature11041)
14. Rieseberg LH *et al.* 2003 Major ecological transitions in wild sunflowers facilitated by hybridization. *Science* **301**, 1211–1216. (doi:10.1126/science.1086949)
15. Grant BR, Grant PR. 1993 Evolution of Darwin's finches caused by a rare climatic event. *Proc. R. Soc. Lond. B* **251**, 111–117. (doi:10.1098/rspb.1993.0016)
16. Becker M *et al.* 2013 Hybridization may facilitate *in situ* survival of endemic species through periods of climate change. *Nat. Clim. Change* **3**, 1039–1043. (doi:10.1038/nclimate2027)
17. Siepielski AM, Gotanda KM, Morrissey MB, Diamond SE, DiBattista JD, Carlson SM. 2013 The spatial patterns of directional phenotypic selection. *Ecol. Lett.* **16**, 1382–1392. (doi:10.1111/ele.12174)
18. Savolainen O, Lascoux M, Merilä J. 2013 Ecological genomics of local adaptation. *Nat. Rev. Genet.* **14**, 807–820. (doi:10.1038/nrg3522)
19. Fitzpatrick BM, Shaffer, HB. 2004 Environment-dependent admixture dynamics in a tiger salamander hybrid zone. *Evolution* **58**, 1282–1293. (doi:10.1111/j.0014-3820.2004.tb01707.x)
20. Muhlfeld CC, Kovach RP, Jones LA, Al-Chokhachy R, Boyer MC, Leary RF, Lowe WH, Luikart G, Allendorf FW. 2014 Invasive hybridization in a threatened species is accelerated by climate change. *Nat. Clim. Change* **4**, 620–624. (doi:10.1038/nclimate2252)
21. Ryan ME, Johnson JR, Fitzpatrick BM, Lowenstine LJ, Picco AM, Shaffer HB. 2013 Lethal effects of water quality on threatened California salamanders but not on co-occurring hybrid salamanders. *Conserv. Biol.* **27**, 95–102. (doi:10.1111/j.1523-1739.2012.01955.x)
22. Crispo E, Moore JS, Lee-Yaw JA, Gray SM, Haller BC. 2011 Broken barriers: human-induced changes to gene flow and introgression in animals. *Bioessays* **33**, 508–518. (doi:10.1002/bies.201000154)
23. Narum SR, Buerkle CA, Davey JW, Miller MR, Hohenlohe PA. 2013 Genotyping-by-sequencing in ecological and conservation genomics. *Mol. Ecol.* **22**, 2841–2847. (doi:10.1111/mec.12350)
24. Andrews KR, Good JM, Miller MR, Luikart G, Hohenlohe PA. 2016 Harnessing the power of RADseq for ecological and evolutionary genomics. *Nat. Rev. Genet.* **17**, 81–92. (doi:10.1038/nrg.2015.28)
25. Allendorf FW, Hohenlohe PA, Luikart G. 2010 Genomics and the future of conservation genetics. *Nat. Rev. Genet.* **11**, 697–709. (doi:10.1038/nrg2844)
26. Allendorf FW, Leary RF. 1988 Conservation and distribution of genetic variation in a polytypic species, the cutthroat trout. *Conserv. Biol.* **2**, 170–184. (doi:10.1111/j.1523-1739.1988.tb00168.x)
27. Kovach RP, Muhlfeld CC, Boyer MC, Lowe WH, Allendorf FW, Luikart G. 2015 Dispersal and selection mediate hybridization between a native and invasive species. *Proc. R. Soc. B* **282**, 20142454. (doi:10.1098/rspb.2014.2454)
28. Hohenlohe PA *et al.* 2013 Genomic patterns of introgression in rainbow and westslope cutthroat trout illuminated by overlapping paired-end RAD sequencing. *Mol. Ecol.* **22**, 3002–3013. (doi:10.1111/mec.12239)
29. Bear EA, McMahon TE, Zale AV. 2007 Comparative thermal requirements of westslope cutthroat trout and rainbow trout: implications for species interactions and development of thermal protection standards. *Trans. Am. Fish. Soc.* **136**, 1113–1121. (doi:10.1577/T06-072.1)
30. Rasmussen JB, Robinson MD, Hontela A, Heath DD. 2012 Metabolic traits of westslope cutthroat trout, introduced rainbow trout and their hybrids in an ecotonal hybrid zone along an elevation gradient.

- Biol. J. Linn. Soc.* **105**, 56–72. (doi:10.1111/j.1095-8312.2011.01768.x)
31. Hand BK *et al.* 2015 Genomics and introgression: discovery and mapping of thousands of species-diagnostic SNPs using RAD sequencing. *Curr. Zool.* **61**, 156–164. (doi:10.1093/czoolo/61.1.146)
 32. Taylor EB. 1991 A review of local adaptation in Salmonidae, with particular reference to Pacific and Atlantic salmon. *Aquaculture* **98**, 185–207. (doi:10.1016/0044-8486(91)90383-1)
 33. Narum SR, Campbell NR, Kozfkay CC, Meyer KA. 2010 Adaptation of redband trout in desert and montane environments. *Mol. Ecol.* **19**, 4622–4637. (doi:10.1111/j.1365-294X.2010.04839.x)
 34. Fraser DJ, Weir LK, Bernatchez L, Hansen MM, Taylor EB. 2011 Extent and scale of local adaptation in salmonid fishes: review and meta-analysis. *Heredity* **106**, 404–420. (doi:10.1038/hdy.2010.167)
 35. Muhlfeld CC, McMahon TE, Boyer MC, Gresswell RE. 2009 Local habitat, watershed, and biotic factors influencing the spread of hybridization between native westslope cutthroat trout and introduced rainbow trout. *Trans. Am. Fish. Soc.* **138**, 1036–1051. (doi:10.1577/T08-235.1)
 36. Rasmussen JB, Robinson MD, Heath DD. 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**, 357–370. (doi:10.1139/F09-191)
 37. Yau MM, Taylor EB. 2013 Environmental and anthropogenic correlates of hybridization between westslope cutthroat trout (*Oncorhynchus clarkii lewisi*) and introduced rainbow trout (*O. mykiss*). *Conserv. Genet.* **14**, 885–900. (doi:10.1007/s10592-013-0485-8)
 38. Fitzpatrick BM, Johnson JR, Kump DK, Smith JJ, Voss SR, Shaffer HB. 2010 Rapid spread of invasive genes into a threatened native species. *Proc. Natl Acad. Sci. USA* **107**, 3606–3610. (doi:10.1073/pnas.0911802107)
 39. Boyer MC, Muhlfeld CC, Allendorf FW. 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**, 658–669. (doi:10.1139/f08-001)
 40. Carim K, Eby L, Pierce R. 2015 Does whirling disease mediate hybridization between a native and nonnative trout? *N. Am. J. Fish. Manage* **35**, 337–351. (doi:10.1080/02755947.2015.1007182)
 41. Amish SJ, Hohenlohe PA, Painter S, Leary RF, Muhlfeld CC, Allendorf FW, Luikart G. 2012 RAD sequencing yields a high success rate for westslope cutthroat and rainbow trout species-diagnostic SNP assays. *Mol. Ecol. Res.* **12**, 653–660. (doi:10.1111/j.1755-0998.2012.03157.x)
 42. Miller MR, Brunelli JP, Wheeler PA, Liu S, Rexroad CE, Palti Y, Doe CQ, Thorgaard GH. 2012 A conserved haplotype controls parallel adaptation in geographically distant salmonid populations. *Mol. Ecol.* **21**, 237–249. (doi:10.1111/j.1365-294X.2011.05305.x)
 43. Catchen J, Hohenlohe PA, Bassham S. 2013 Stacks: an analysis tool set for population genomics. *Mol. Ecol.* **22**, 3124–3140. (doi:10.1111/mec.12354)
 44. Davey JW, Hohenlohe PA, Etter PD, Boone JQ, Catchen JM, Blaxter ML. 2011 Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nat. Rev. Genet.* **12**, 499–510. (doi:10.1038/nrg3012)
 45. Berthelot C *et al.* 2014 The rainbow trout genome provides novel insights into evolution after whole-genome duplication in vertebrates. *Nat. Commun.* **5**, 1–10. (doi:10.1038/ncomms4657)
 46. Langmead B, Salzberg SL. 2012 Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359. (doi:10.1038/nmeth.1923)
 47. Allendorf FW, Bassham S, Cresko WA, Limborg MT, Seeb LW, Seeb JE. 2015 Effects of crossovers between homeologs on inheritance and population genomics in polyploid-derived salmonid fishes. *J. Hered.* **106**, 217–227. (doi:10.1093/jhered/esv015)
 48. Racimo F, Sankararaman S, Nielsen R, Huerta-Sanchez E. 2015 Evidence for archaic adaptive introgression in humans. *Nat. Rev. Genet.* **16**, 359–371. (doi:10.1038/nrg3936)
 49. Long JC. 1991 The genetic structure of admixed populations. *Genetics* **127**, 417–418.
 50. Benjamini Y, Yekutieli D. 2001 The control of the false discovery rate in multiple testing under dependency. *Ann. Stat.* **29**, 1165–1188. (doi:10.1214/aos/1013699998)
 51. Narum SR. 2006 Beyond Bonferroni: less conservative analyses for conservation genetics. *Conserv. Genet.* **7**, 783–787. (doi:10.1007/s10592-005-9056-y)
 52. Fitzpatrick BM. 2013 Alternative forms for genomic clines. *Ecol. Evol.* **3**, 1951–1966. (doi:10.1002/ece3.609)
 53. Buerkle CA, Lexer C. 2008 Admixture as the basis for genetic mapping. *Trends Ecol. Evol.* **23**, 686–694. (doi:10.1016/j.tree.2008.07.008)
 54. Altschul SF, Gish W, Miller W, Myers EW. 1990 Basic local alignment search tool. *J. Molec. Biol.* **215**, 403–410. (doi:10.1016/S0022-2836(05)80360-2)
 55. Hitt NP, Frissell CA, Muhlfeld CC, Allendorf FW. 2003 Spread of hybridization between native westslope cutthroat trout, *Oncorhynchus clarkii lewisi*, and nonnative rainbow trout, *Oncorhynchus mykiss*. *Can. J. Fish. Aquat. Sci.* **60**, 1440–1451. (doi:10.1139/f03-125)
 56. Bennett SN, Olson JR, Kershner JL, Corbett P. 2010 Propagule pressure and stream characteristics influence introgression: cutthroat and rainbow trout in British Columbia. *Ecol. Appl.* **20**, 263–277. (doi:10.1890/08-0441.1)
 57. Lowe WH, Muhlfeld CC, Allendorf FW. 2015 Spatial sorting promotes the spread of maladaptive hybridization. *Trends Ecol. Evol.* **30**, 456–462. (doi:10.1016/j.tree.2015.05.008)
 58. Johnson JR, Fitzpatrick BM, Shaffer HB. 2010 Retention of low-fitness genotypes over six decades of admixture between native and introduced tiger salamanders. *BMC Evol. Biol.* **10**, 147. (doi:10.1186/1471-2148-10-147)
 59. Kovach RP, Luikart G, Lowe WH, Boyer MC, Muhlfeld CC. 2016 Risk and efficacy of human-enabled interspecific hybridization for climate-change adaptation: response to Hamilton and Miller (2016). *Conserv. Biol.* **30**, 428–430. (doi:10.1111/cobi.12678)
 60. Harrison KA, Pavlova A, Scott MT, Sunnucks P. 2014 Using genomics to characterize evolutionary potential for conservation of wild populations. *Evol. Appl.* **7**, 1008–1025. (doi:10.1111/eva.12149)
 61. Anderson E, Stebbins GL. 1954 Hybridization as an evolutionary stimulus. *Evolution* **8**, 378–388. (doi:10.2307/2405784)
 62. Maheshwari S, Barbash DA. 2011 The genetics of hybrid incompatibilities. *Annu. Rev. Genet.* **45**, 331–355. (doi:10.1146/annurev-genet-110410-132514)
 63. Ostberg CO, Hauser L, Pritchard VL, Garza JC, Naish KA. 2013 Chromosome rearrangements, recombination suppression, and limited segregation distortion in hybrids between Yellowstone cutthroat trout (*Oncorhynchus clarkii bouvieri*) and rainbow trout (*O. mykiss*). *BMC Genomics* **14**, 570. (doi:10.1186/1471-2164-14-570)
 64. Limborg MT, Seeb LW, Seeb JE. 2016 Sorting duplicated loci disentangles complexities of polyploid genomes masked by genotyping by sequencing. *Mol. Ecol.* **25**, 2117–2129. (doi:10.1111/mec.13601)
 65. Hamilton JA, Miller JM. 2016 Adaptive introgression as a resource for management and genetic conservation in a changing climate. *Conserv. Biol.* **30**, 33–41. (doi:10.1111/cobi.12574)
 66. Barton NH, Hewitt GM. 1985 Analysis of hybrid zones. *Annu. Rev. Ecol. Syst.* **116**, 113–148. (doi:10.1146/annurev.es.16.110185.000553)
 67. Hereford J. 2009 A quantitative survey of local adaptation and fitness trade-offs. *Am. Nat.* **173**, 579–588. (doi:10.1086/597611)