






Comparing RADseq and microsatellites for estimating genetic diversity and relatedness – Implications for brown trout conservation

Alexandre Lemopoulos^{1,2*}  | Jenni M. Prokkola^{1,3*}  | Silva Uusi-Heikkilä^{2,4}  |
Anti Vasemägi^{2,5,6}  | Ari Huusko⁷ | Pekka Hyvärinen⁷ | Marja-Liisa Koljonen⁸ |
Jarmo Koskiniemi⁹ | Anssi Vainikka¹ 

¹Department of Environmental and Biological Sciences, University of Eastern Finland, Joensuu, Finland

²Department of Biology, University of Turku, Turku, Finland

³Institute of Integrative Biology, University of Liverpool, Liverpool, UK

⁴Department of Biological and Environmental Science, University of Jyväskylä, Jyväskylä, Finland

⁵Department of Aquatic Resources, Institute of Freshwater Research, Swedish University of Agricultural Sciences, Drottningholm, Sweden

⁶Estonian University of Life Sciences, Institute of Veterinary Medicine and Animal Sciences, Tartu, Estonia

⁷Natural Resources Institute Finland (Luke), Kainuu Fisheries Research Station, Paltamo, Finland

⁸Natural Resources Institute Finland (Luke), Helsinki, Finland

⁹Department of Agricultural Sciences, University of Helsinki, Helsinki, Finland

Correspondence

Jenni M. Prokkola, Institute of Integrative Biology, University of Liverpool, Liverpool, UK.

Email: jenni.prokkola@liverpool.ac.uk

Funding information

Suomen Akatemia, Grant/Award Number: 266321 and 286261; Natural Resources Institute Finland (Luke); Finnish Cultural Foundation

Abstract

The conservation and management of endangered species requires information on their genetic diversity, relatedness and population structure. The main genetic markers applied for these questions are microsatellites and single nucleotide polymorphisms (SNPs), the latter of which remain the more resource demanding approach in most cases. Here, we compare the performance of two approaches, SNPs obtained by restriction-site-associated DNA sequencing (RADseq) and 16 DNA microsatellite loci, for estimating genetic diversity, relatedness and genetic differentiation of three, small, geographically close wild brown trout (*Salmo trutta*) populations and a regionally used hatchery strain. The genetic differentiation, quantified as F_{ST} , was similar when measured using 16 microsatellites and 4,876 SNPs. Based on both marker types, each brown trout population represented a distinct gene pool with a low level of interbreeding. Analysis of SNPs identified half- and full-siblings with a higher probability than the analysis based on microsatellites, and SNPs outperformed microsatellites in estimating individual-level multilocus heterozygosity. Overall, the results indicated that moderately polymorphic microsatellites and SNPs from RADseq agreed on estimates of population genetic structure in moderately diverged, small populations, but RADseq outperformed microsatellites for applications that required individual-level genotype information, such as quantifying relatedness and individual-level heterozygosity. The results can be applied to other small populations with low or moderate levels of genetic diversity.

KEYWORDS

ddRADseq, fisheries, population genetics, relatedness, salmonids

*These authors made an equal contribution to this work.

1 | INTRODUCTION

Information on genetic variation within and among populations is paramount in understanding the long-term impacts of human activities. These may include harvesting (Allendorf & Hard, 2009; Henriques et al., 2016), polluting (Paris, King, & Stevens, 2015), habitat destruction and fragmentation (Balkenhol & Waits, 2009; Keyghobadi, 2007; Wofford, Gresswell, & Banks, 2005), and reintroduction and enhancement of wild populations with releases of human-reared individuals (Anderson, Faulds, Atlas, & Quinn, 2013; Cochran-Biederman, Wyman, French, & Loppnow, 2015; Seddon, Armstrong, & Maloney, 2007). Particularly microsatellites, i.e., short repetitive regions in the non-coding DNA, have been used to analyze genetic differentiation and diversity across populations in different taxa for decades (Guichoux et al., 2011). However, modern sequencing methods are being rapidly developed and replacing widely used methods also in conservation genetics (Narum, Buerkle, Davey, Miller, & Hohenlohe, 2013).

Highly polymorphic microsatellite markers, given that they are available for the target species, can typically resolve the genetic structure of populations reliably even among closely related populations, and provide information on population genetic diversity, average kinship and effective population size (N_e). However, microsatellite markers have limitations such as risk of homoplasy for allele size, the presence of null alleles (Putman & Carbone, 2014; Zhang & Hewitt, 2003) or a potentially insufficient number of polymorphic loci in the study species, which may limit their resolution power. Single nucleotide polymorphisms (SNPs) obtained by restriction-site-associated DNA sequencing (RADseq) producing thousands of loci thus provide an appealing alternative, particularly for species without prior genetic information, or for species and populations known to have limited amount of microsatellite variation because of prior population bottlenecks. Likewise, large panmictic populations, for which the differentiation levels are low, are especially challenging for genetic analysis. Consequently, the RADseq approach has rapidly gained popularity (Andrews, Good, Miller, Luikart, & Hohenlohe, 2016; Davey & Blaxter, 2010).

The advantages of SNP markers over microsatellites include their suitability for comparisons of both strongly and weakly diverged populations, and even species, and in revealing ancestral patterns of genetic structuring compared to microsatellites due to the slower mutation rate of SNPs compared to microsatellite regions (Andrews et al., 2016; Zhang & Hewitt, 2003). In addition, the RADseq approach can provide more reliable inferences on population structure (Bruneaux et al., 2013) and improved resolution for data sets with fewer individuals compared to the microsatellite approach (Jeffries et al., 2016). Likewise, Bradbury et al. (2015) demonstrated that SNPs obtained by RADseq were more accurate than microsatellites for characterizing introgression between Atlantic salmon (*Salmo salar*) from the East and West coasts of the Atlantic Ocean. Despite these advances, more information on wider range of species is still needed to compare the performance and cost-efficiency of these

two marker types in determining population structure especially in small populations in need of conservation actions.

In addition to conservation applications focusing on population-level metrics, individual-based metrics, including relatedness, genetic diversity and family structure (full-sib and half-sib information) are valuable for managing hatchery breeding strategies, and understanding demographics (Hauser, Baird, Hilborn, Seeb, & Seeb, 2011; Stadele & Vigilant, 2016) and diversity-fitness-correlations (Hedrick & Kalinowski, 2000) in wild populations. While the fast mutation rate and high polymorphism of microsatellites allow for resolving fine-scale population structuring (Putman & Carbone, 2014), they may be less suitable for inferring genome-wide or individual-level patterns in genetic diversity (Väli, Einarsson, Waits, & Ellegren, 2008), as the number of sampled loci may not be sufficient to represent the total genome of an individual. It has also been demonstrated that SNPs obtained by RADseq produce more precise estimates of relatedness than microsatellites in a range of bird species (Thrasher, Butcher, Campagna, Webster, & Lovette, 2018). Microsatellites may be less efficient for identifying relatives particularly in populations with prior population bottlenecks and lack of gene flow, which limit allelic diversity. Further, the lack of commonly shared loci across species may be a more serious limitation for the use of microsatellite approach compared to RADseq analysis for, e.g., phylogenetic studies (Eaton & Ree, 2013; Near et al., 2018).

Resolving the relationship between individual-level genetic diversity (i.e., heterozygosity) and fitness is a long-standing question in conservation and evolutionary biology: lower diversity is expected to contribute to lower fitness (Hedrick & Kalinowski, 2000). However, the genetic background of populations can influence the observed correlations (e.g., Tiira et al., 2006; Velando, Barros, & Moran, 2015), as can the type of genetic marker used (Miller et al., 2014; Väli et al. 2008). Moreover, published estimates of heterozygosity-fitness correlation (HFC) often have low correlation coefficients (Chapman, Nakagawa, Coltman, Slate, & Sheldon, 2009). In order to reveal a relevant HFC, the diversity of the applied genetic markers needs to represent individual genetic diversity across large areas of the genome, which is often not true for microsatellite panels (Fischer et al., 2017; Väli et al., 2008) as they usually only represent the most variable loci. Consequently, generally lower HFC has been found using microsatellite loci compared to the more numerous SNP loci; there was an almost five-fold increase in HFC when measured by RADseq approach in comparison to 10 microsatellite loci in an endangered species with low genetic diversity, the harbour seal (*Phoca vitulina*; Hoffman et al., 2014). Furthermore, the minimum number of SNPs required for a reliable estimate of individual heterozygosity can vary between populations (Miller et al., 2014), but there is a lack of studies that have included both several populations and a large numerical range of loci in this evaluation (but see Fischer et al. (2017) for a pool-Seq approach).

Many salmonids provide excellent examples of systems where geographically connected (even sympatric) populations can be genetically isolated (e.g., Castric, Bonney, & Bernatchez, 2001; Estoup

et al., 1998; Vähä, Erkinaro, Niemelä, & Primmer, 2007). Due to tremendous changes in their native breeding habitats, including the construction of dams, and overfishing particularly in the feeding areas, a large number of salmonid populations have become extirpated or declined dramatically (e.g., Bradford & Irvine, 2000; Morita & Yamamoto, 2002). As the stocking strategies (restoration or enhancement releases or both) of hatchery fish to maintain some of the impacted populations continue to be optimized, it remains necessary to characterize the best approaches for evaluating differences between hatchery brood stocks and native stocks. Further, the relatively high costs for DNA sequencing and library preparation in RADseq, as well as the potential challenges of obtaining numerous individuals for genotyping from small populations call for assessments whether there is cost-efficient increase in resolution to be gained by using RADseq analysis over microsatellites when using a low number of individuals.

In this study, we analyzed the genetic structure and diversity of three wild and one captive-bred population of brown trout (*Salmo trutta* L.) with both a RADseq approach and a DNA-microsatellite panel commonly used in brown trout population genetic research (e.g., Debes, Gross, & Vasemägi, 2017; Koljonen, Janatuinen, Saura, & Koskiniemi, 2013; Koljonen, Gross, & Koskiniemi, 2014; Swatdipong, Vasemägi, Niva, Koljonen, & Primmer, 2010). Our applied goal was to support making informed management decisions locally. The main methodological goal was to compare the performance of microsatellite and SNP markers especially in the estimation of individual-level genetic diversity and relatedness to provide future reference for studies seeking to tailor the methodology to suit a given purpose.

2 | MATERIAL AND METHODS

2.1 | Sample collection

Fish sampling and breeding were conducted under a license obtained from the national Animal Experiment board in Finland (license number ESAVI/3443/04.10.07/2015). Wild brown trout were caught in the boreal River Oulujoki watershed from three resident brown trout populations in rivers Tuhkajoki (17 September 2013), Pohjajoki (16–17 September 2015) and Vaarainjoki (28–30 September 2010,

15 September–11 October 2011 and 2 October 2012) using electrofishing and transported to the Kainuu Fisheries Research station (www.kfrs.fi), Paltamo, Finland (Table 1, Figure 1). These populations were assumed to be resident due to the absence of observed migrants in these rivers and unpublished experimental data (personal observation P. Hyvärinen, A. Vainikka). River Pohjajoki fish were sampled under anesthesia immediately after the initial capture. Any mortality prior to sampling was negligible in all populations. Samples from the Tuhkajoki, Vaarainjoki and a hatchery brood stock (details below) were collected at Kainuu Fisheries Research Station by snipping a small piece of caudal fin under anesthesia (benzocaine) during artificial breeding between 12 and 22 October 2015. For each river and the hatchery stock, 30 fish were sampled, for a total of 120 individuals.

Samples from the hatchery brood stock included in this study (Ouv) were collected from the second generation of this stock. The stock was founded in October 2000 by using adfluvial brown trout from two previous hatchery brood stocks originating from two connected rivers, River Varisjoki and upstream River Kongasjoki (Figure 1). The former was originally founded in 1960–1970s and fish from other wild populations from the same watercourse were possibly mixed into it without reliable records in the 1960–1970s. The latter was founded in the 1970–1980s.

Total DNA was extracted from fin clips preserved in pure ethanol or dried scales using the Omega bio-tek E.Z.N.A Tissue DNA kit or Macherey-Nagel NucleoSpin Tissue kit. The quality of total DNA was controlled with electrophoresis on a 1% agarose gel and with fluorometric measurements using Qubit 2.0 with Qubit[®] dsDNA HS Assay Kit (ThermoFisher Scientific).

2.2 | Microsatellite analysis

Allelic variation was determined at 16 microsatellite loci (Supporting information Table S1) for all 120 individuals (Table 1). For each sample, two multiplex PCR reactions were performed using the Qiagen Type-it Microsatellite kit in a 10- μ l reaction volume with 3 μ l of extracted DNA, 5 μ l of kit master mix and primers with concentrations and dyes as presented in Supporting information. PCR reactions were carried out in PTC200 Thermal Cyclers (MJ Research), and the temperature profile of the PCR program was suggested in

TABLE 1 Sampling coordinates and summary of individuals used in the analysis with each marker. For RADseq data, *N* before filtering shown in parentheses

| Population | Lat/Lon | Microsatellite data | | RADseq data | |
|----------------|-------------------------------------|---------------------|--------------------|-------------|--------------------|
| | | <i>N</i> | Sex (F/M/immature) | <i>N</i> | Sex (F/M/immature) |
| Pohjajoki | 64° 17' 50.703" N/28° 3' 0.416" E | 30 | 0/4/26 | 11 (12) | 0/4/8 |
| Tuhkajoki | 64° 2' 28.337" N/28° 7' 10.099" E | 30 | 6/12/12 | 9 (11) | 3/3/5 |
| Vaarainjoki | 64° 28' 50.510" N/27° 34' 17.340" E | 30 | 15/15/0 | 29 | 15/14/0 |
| Hatchery stock | Details in text | 30 | 15/15/0 | 26 (28) | 15/13/0 |

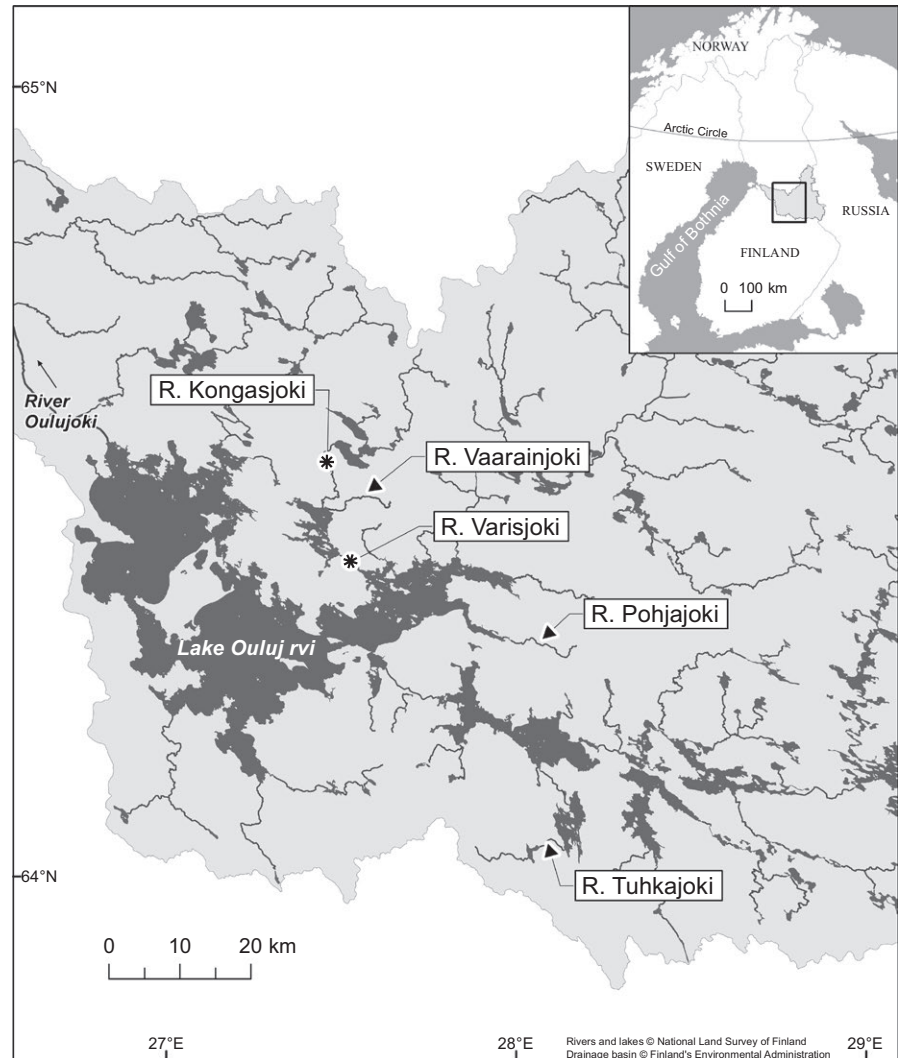


FIGURE 1 A map showing the River Oulujoki watershed. *Salmo trutta* individuals used in the study were collected from rivers Vaarainjoki, Pohjajoki and Tuhkajoki, indicated by arrows. The rivers that the hatchery stock originated from, Kongasjoki and Varisjoki, are shown with asterisks

the Type-it Microsatellite kit manual. The annealing temperature was 56°C. The amplification products were separated by capillary electrophoresis on AB3130 Genetic Analyzers (Applied Biosystems, Foster City, CA). The sizes of the microsatellite alleles were determined using Genemapper v. 4.0 software (Applied Biosystems, Foster City, CA), and manually checked. Deviations from Hardy-Weinberg equilibrium in each population were tested using function *hw.test* from pegas package (Paradis, 2010) in R environment. A few loci showed a significant deviation from equilibrium within populations (2 in Tuhkajoki, 1 in Pohjajoki, 4 in Vaarainjoki and 2 in hatchery stock). However, as all microsatellite loci were in equilibrium in at least two populations, none were excluded from analysis.

2.3 | Sequencing, genotyping and SNP calling

The sequencing libraries were prepared using samples outlined in Table 1 and Supporting information Table S2. From each sample, 100 ng of genomic DNA along with PstI-HF (5'CTGCAG 3') and BamHI-HF (5'GGATCC 3') restriction enzymes was used. The protocol used was the same as in Lemopoulos, Uusi-Heikkilä, Vasemägi et al. (2018a). Briefly, individual barcodes were ligated

to the forward ends before pooling the individuals into a single library. The library was purified with PCR purification and size selected into 280–320 bp fragments on an E-Gel Size select 2% Agarose gel (Invitrogen, CA, USA). Amplification through PCR was then performed and the product was purified using SPRI-Beads, removing fragments <100 bp. The DNA concentration of the libraries was quantified using Qubit 2.0. Size and quantity of contained fragments were assessed with Agilent 2,100 Bioanalyzer (Agilent Technologies, CA, USA). Samples were pooled into one sequencing lane. Four samples were re-sequenced on different sequencing lanes to determine the genotyping error rate. Libraries were sequenced on Illumina HiSeq 2,500 with the rapid run option by a commercial service provider, Turku Centre for Biotechnology (BTK), in Turku, Finland.

A total of 113,885,773 reads were retained after quality filtering (Supporting information Table S2), as well as 6.08 million reads used for estimating sequencing error rate, which was 5% at the SNP level (including the error rate from contig assembly and SNP calling). The average coverage depth per individual after filtering was 42.64x. The obtained RAD-data were analyzed using Stacks, v. 1.40 (Catchen, Hohenlohe, Bassman, Amores, & Cresko, 2013).

Process_Radtags function was used for demultiplexing, quality filtering (q) and cleaning (-c). Orthologous tags were assembled, catalogued and matched using *denovo* pipeline, in which the optimal parameters were obtained following Paris, Stevens, and Catchen (2017). Minimum coverage (-m), maximum mismatches between loci for a single individual (-M) and the maximum mismatches (-n) between loci for catalogue building were all set to 2. All other parameters were set to default. *Population* function was run for the SNP calling. On the first call, only loci that were present in at least 50% of all the individuals were kept while the rest of the parameters were set to defaults. Further filtering was done in R using the *stackr* (Gosselin & Bernatchez, 2016) and *grur* (Gosselin, 2017) packages. To exclude uninformative markers and samples with too much missing data, the number of populations and individuals where a locus had to be present were assessed based on the data (*missing_visualisation* function), after which new parameters were again passed onto STACKS' *population* function. Finally, only loci present in all four populations (-p) and in 60% of the individuals (-r) were retained. Based on this dataset, a total of five individuals with more than 20% of missing data (Table 1), and markers with more than 30% of missing data were discarded (using *stackr*). This was done in order to remove potential sequencing errors and uninformative missing data that could potentially bias the results (see *stackr* package guidelines). In addition, a filter for marker heterozygosity (maximum threshold: 0.5, as in Hohenlohe, Amish, Catchen, Allendorf, and Luikart (2011)) was applied to remove potential sequencing errors. Individual heterozygosity was between 0.14 and 0.22; thus, no individuals were excluded based on heterozygosity. Markers were further filtered for minor allele frequency based on a local (0.02) and a global (0.005) threshold. This dataset was then assessed using *missing_visualisation* function and identity-by-missingness analysis with *grur* to confirm that populations did not cluster based on missing data. Further, only loci that were under Hardy-Weinberg equilibrium as defined by *p*-value threshold >0.05 in at least two populations were retained (HW tests made using *pegas* package; Paradis, 2010 in R).

Because the sex ratio of the samples from two of the studied populations was not known, the SNP dataset was checked for the presence of sex-linked markers to avoid introducing bias into the analysis (Benestan et al., 2017). Genetic variation was compared in all individuals with known sex using BAYESCAN v2.01 (31 females and 32 males; Foll & Gaggiotti, 2008) to detect potential sex-linked outliers. The input file was created using PGD software v2.1.0.3 (Lischer & Excoffier, 2012) and all default parameters were used in Bayescan. No sex-linked outlier loci ($\alpha = 0.1$) were found in the final set of SNPs, likely indicating that these loci were excluded in the other filtering steps or had low coverage. SNPs were not filtered based on neutrality, as the number of loci under selection in the dataset as a whole is expected to be small (see, e.g., outlier analysis on brown trout in Lemopoulos, Uusi-Heikkilä, Huusko, Vasemägi, & Vainikka, 2018b). The final RADseq dataset comprised 4,876 loci and 75 individuals. Trimmed read counts in these samples ranged from 148,214 to 3,116,841 (Supporting information Table S2).

2.4 | Genetic diversity and differentiation between populations

The following analyses were conducted on three datasets: microsatellite data from all individuals, microsatellite data from the same individuals as in the final RADseq data, and the SNPs from the final RADseq data.

Pairwise and global F_{ST} - values (Nei, 1973) were analyzed using package *hierfstat* (Goudet & Jombart, 2015). Total allele counts and allelic richness were measured using package *PopGenReport* v.3.0, and expected heterozygosity (H_e) using package *adegenet* v.2.0.1 (Jombart, 2008; Jombart & Ahmed, 2011) in R. *P*-values for pairwise F_{ST} - values were obtained using 999 MCMC permutations. Effective population size (N_e) was calculated using the method based on the linkage disequilibrium (LD; Hill, 1981; Waples, 2006; Waples & Do, 2010) as implemented in NeEstimator v.2.1 (Do et al., 2014). Minimum allele frequency of 0.02 and non-parametric jackknifed confidence intervals were used for all analyses of N_e (Jones, Ovenden, & Wang, 2016). The estimates of N_e obtained for SNP data are known to suffer from a downward bias and were thus corrected using the equation 1a from Waples, Larson, and Waples (2016) (haploid chromosome number = 40; Leitwein et al., 2017). The corrected N_e was then calculated as: N_e from NeEstimator/0.906.

2.5 | Multivariate analysis and bayesian clustering

The clustering of samples into four populations was visualized using Discriminant Analysis for Principal Components (DAPC) from package *adegenet* v. 2.0.1 (Jombart, 2008; Jombart & Ahmed, 2011) in R. The number of principal components retained for DAPC was obtained by cross-validation, selecting the lowest number of components where the correct assignment probability levelled (13 for SNPs, 4 for both datasets of microsatellites).

The likely number of distinct source populations and admixture between populations were additionally analyzed using STRUCTURE (Pritchard, Stephens, & Donnelly, 2000). STRUCTURE was repeated 20 times using a burn-in of 50,000 followed by 100,000 iterations for microsatellites and a burn-in of 200,000 and 200,000 iterations for the RADseq dataset, for each *K* value 2–7, after which the optimal *K* value was determined using CLUMPAK (Kopelman, Mayzel, Jakobsson, Rosenberg, & Mayrose, 2015), selecting the *K* where the mean ln likelihood converged. Once the optimal *K* (4) was determined, the 20 runs of STRUCTURE using *K* = 4 were combined using LargeKGreedy algorithm in CLUMPP v.1.1.2 (Jakobsson & Rosenberg, 2007), to which the input file was generated using STRUCTURE HARVESTER (Earl & Vonholdt, 2012). The output file of CLUMPP was visualized using DISTRUCT v.1.1 (Rosenberg, 2004). The SNP and STRUCTURE analyses were conducted using the CSC - IT Center for Science Ltd clusters in Finland.

2.6 | Family structure and relatedness

Family structure within populations was assessed in the individuals with both SNP and microsatellite data available using COLONY v.

2.0.6.2 (Jones & Wang, 2010) with random mating model (Wang, 2016). The error rate was set at 0.05 for SNPs (based on repeated genotyping of five individuals) and 0.0001 for microsatellites (based on empirical, unpublished data). Each analysis was run twice using medium run length to confirm that identical results were obtained on both runs. Because the removal of siblings from population genetic datasets can lead to a loss of information on genetic divergence between populations (Waples & Anderson, 2017), all samples were included in the analyses of population structure regardless of family origin. Relatedness was calculated based on the number of alleles shared between individuals (B_{xy} ; Li & Horvitz, 1953). Differences in mean pairwise relatedness based on 1,000 bootstrapped replicates of random subsets of loci (1–16 loci for microsatellites and 1–100 loci for SNPs) were obtained using function *loci.test* from package *demerelate* in R.

2.7 | Individual multilocus heterozygosities

Standardized multilocus heterozygosities (sMLH), i.e., the relationship between the heterozygous loci for an individual and the sum of observed average heterozygosity in the population (Coltman, Pilkington, Smith, & Pemberton, 1999), were measured to compare if microsatellite and SNP markers from the same individuals provide comparable diversity estimates at individual level. sMLH values were obtained using the *inbreedR* package (Stoffel et al., 2016). Pearson correlation was used to evaluate the similarity of the sMLH estimates between microsatellite and SNP markers. In addition, to test how reliably sMLH could be estimated using different numbers of markers, the correlation between two equal-sized, random subsets of markers (i.e., heterozygosity-heterozygosity correlation; Balloux, Amos, & Coulson, 2004) was calculated across all individuals and within populations (repeated for 1,000 times). The strength of the correlation between sMLH values from the same-size subsets of markers indicates the preciseness of the estimates. Subsets of

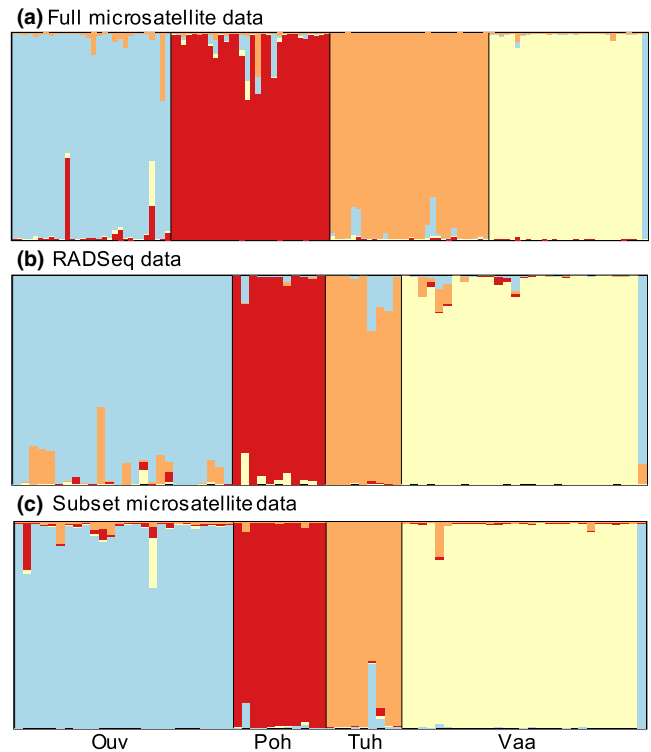


FIGURE 3 DISTRUCT plots showing posterior probabilities of *Salmo trutta* individual genotypes (as bars) assigned to each population based on microsatellite data from all 120 individuals (a), RADseq data from 75 individuals (b), and microsatellite data from the same 75 individuals (c). The expected populations are separated by black lines. Populations from left to right hatchery stock (Ouv), Pohjajoki (Poh), Tuhkajoki (Tuh), Vaarainjoki (Vaa)

eight microsatellites and 100/500/1,000/1,500/2,000 SNPs were evaluated. Violin plots of the results were produced using *ggplot2* (Wickham, 2009) and *easyggplot2* (Kassambara, 2014). R version v. 3.3.2 (R Core Team 2016) was used throughout the analyses.

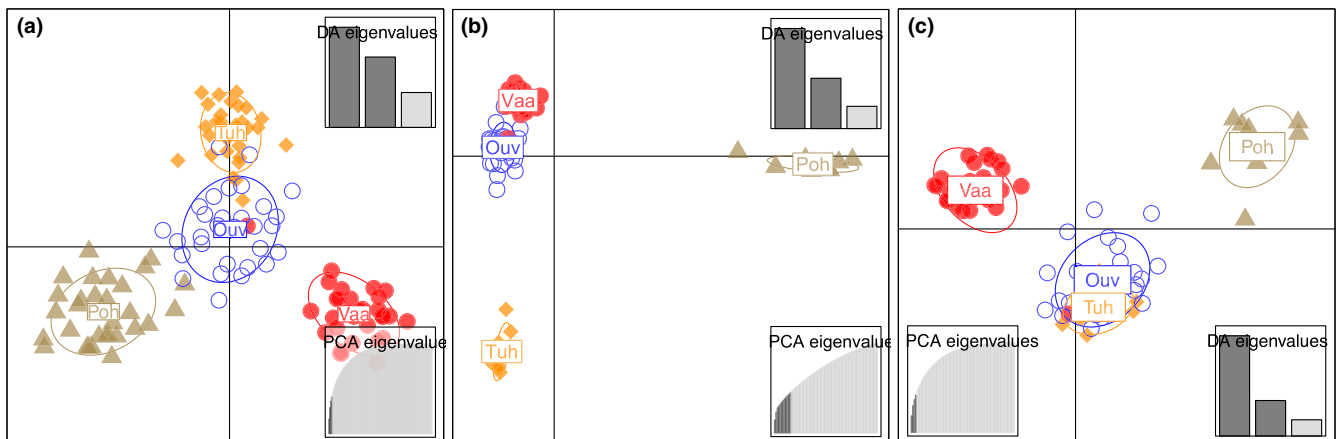


FIGURE 2 Grouping of *Salmo trutta* samples on DAPC based on full microsatellite data (a), RADseq data (b) and microsatellite data (c) from the same individuals. Populations indicated by labelled symbols, hatchery stock (Ouv), Pohjajoki (Poh), Tuhkajoki (Tuh), Vaarainjoki (Vaa)

TABLE 2 Genetic diversity and effective population size (N_e) in all studied populations across the three datasets. Total number of alleles, mean locus-specific allelic richness (Ar mean) and the estimates of expected heterozygosity (H_e) and N_e shown as measured from microsatellites in 120 individuals (A) and RADseq (B) and microsatellites (C) on the same 75 individuals. LD-based estimates of N_e are shown with 95% non-parametric jackknifed confidence intervals. Note the difference in scale for H_e : the theoretical maximum is 1 for multiallelic microsatellites and 0.5 for bi-allelic SNPs

| | A | | | | | B | | | | | C | | | | |
|----------------|--------------|---------|-------|----------------------|--------------|---------|-------|-------------------|--------------|---------|-------|----------------------|--|--|--|
| | Allele count | Ar mean | H_e | N_e | Allele count | Ar mean | H_e | N_e | Allele count | Ar mean | H_e | N_e | | | |
| Pohjajoki | 68 | 4.01 | 0.53 | 14.8 (6.9–34.7) | 6,566 | 1.26 | 0.09 | 13.0 (6.3–28.3) | 46 | 2.75 | 0.48 | 9.7 (2.8–73) | | | |
| Tuhkajoki | 82 | 4.89 | 0.61 | 22.0 (12.4–46.3) | 6,854 | 1.31 | 0.1 | 2.10 (2.9–27.4) | 61 | 3.61 | 0.58 | 10.1 (4.4–28.8) | | | |
| Vaarainjoki | 80 | 4.78 | 0.59 | 32.8 (11.5–infinite) | 7,555 | 1.47 | 0.11 | 24.5 (12.3–123.8) | 80 | 4.75 | 0.59 | 28.8 (10.1–infinite) | | | |
| Hatchery stock | 108 | 6.31 | 0.66 | 53.0 (30.8–133) | 8,380 | 1.61 | 0.14 | 96.5 (60.8–149.0) | 102 | 6.042 | 0.66 | 57.1 (31.2–316.8) | | | |

3 | RESULTS

3.1 | Explorative multivariate analysis

Using DAPC, each population was separated from the others based on both microsatellites (Figure 2a) and SNPs (Figure 2b). The hatchery stock (Ouv) and River Tuhkajoki population clustered together using microsatellite data from individuals included in RADseq (Figure 2c), thus displaying only three groups in total.

3.2 | Bayesian clustering using structure

The average likelihood of 20 independent STRUCTURE runs converged at $K = 4$ with all three datasets. There was low admixture between populations (Figure 3a–c). Both DAPC and STRUCTURE analysis grouped one individual from Vaarainjoki population with the hatchery fish in all datasets (showing that the methods assigned population of origin equally well; Figure 3a–c). In further analyses, this individual was kept in the Vaarainjoki population, as it was originally captured in the river, although it may have been bred in the hatchery.

3.3 | Genetic diversity and differentiation

In total, 3–18 alleles were found in the 16 microsatellite loci (the sum of all alleles was 143), with a mean of 4–6 alleles across loci within populations (Table 2). Nearly half (48%) of the loci had at most four alleles within populations. 4,876 SNPs were included in the final dataset obtained by RADseq. Both microsatellite and SNP markers identified the hatchery population as genetically more diverse than the wild populations based on H_e and allelic richness (Table 2). Two individuals from the Vaarainjoki population shared an identical multilocus genotype based on 16 microsatellite markers, which is highly unusual, but may be explained by the low allelic diversity in most of the loci in the population. These individuals were full-sibs based on SNPs, but in the final set of loci their difference was <2%, which falls within the margin for genotyping error and might indicate that DNA was accidentally collected twice from the same individual, although we consider this unlikely. The average allelic richness was over three times higher for microsatellites than for SNPs (Table 2). The estimates of N_e were overall low (Table 2). The confidence intervals for N_e were overlapping and therefore did not indicate significant differences in the estimates across the datasets.

Based on the complete microsatellite data, the global estimate of F_{ST} was 0.209 across populations, and was very similar to that based on SNPs (0.211). The pairwise F_{ST} values between the three rivers and the hatchery stock were highly significant, and approximately at the same level in all three datasets (Table 3).

3.4 | Family structure and relatedness

According to COLONY output, the estimated number of full-sib families was similar using microsatellite and SNP markers on the

TABLE 3 Pairwise F_{ST} values for three wild brown trout river populations and one hatchery stock obtained using the full microsatellite dataset (A), and RADseq (B) and microsatellite data (C) on the same individuals

| | A | | | B | | | C | | |
|-------------|-----------|-----------|-------------|-----------|-----------|-------------|-----------|-----------|-------------|
| | Pohjajoki | Tuhkajoki | Vaarainjoki | Pohjajoki | Tuhkajoki | Vaarainjoki | Pohjajoki | Tuhkajoki | Vaarainjoki |
| F_{ST} | | | | | | | | | |
| Tuhkajoki | 0.154 | | | 0.202 | | | 0.197 | | |
| Vaarainjoki | 0.150 | 0.138 | | 0.134 | 0.111 | | 0.141 | 0.116 | |
| Hatchery | 0.120 | 0.074 | 0.112 | 0.119 | 0.064 | 0.109 | 0.110 | 0.067 | 0.110 |

TABLE 4 The number of identified full-sib families with average exclusion probabilities and half-sib dyads with average probabilities from microsatellite and SNP data on the same individuals (COLONY software). A comparison of the probabilities of matching half-sib dyads is shown in Figure 4a

| | | Pohjajoki | Tuhkajoki | Vaarainjoki | Hatchery stock |
|-------------------|-----------------|-----------|-----------|-------------|----------------|
| Full-sib families | SNPs | 8/0.98 | 5/0.99 | 25/1.00 | 26/0.97 |
| | Microsatellites | 10/0.44 | 4/0.52 | 26/0.52 | 25/0.66 |
| Half-sib dyads | SNPs | 9/0.56 | 3/0.996 | 34/0.85 | 30/0.72 |
| | Microsatellites | 66/0.18 | 52/0.07 | 334/0.18 | 176/0.26 |

same individuals, when empirically determined error rates (0.0001 and 0.05, respectively) were used for each marker (Table 4). In Tuhkajoki, both markers identified largely the same full-sib families apart from one family identified only with SNPs. In Pohjajoki and Vaarainjoki, none of the identified full-sib pairs matched except for one pair in Vaarainjoki. In the hatchery stock, microsatellites identified one potential pair of full siblings (but with only 41% exclusion probability) that was classified as neither full nor half-sibs based on SNPs. The average exclusion probabilities of full-sib family identification were clearly higher based on SNPs than microsatellites (Table 4), indicating SNPs contained more information to determine sib-ship. The differences between the markers were even more pronounced for half-sib assignment, where microsatellite data yielded four times more half-sib dyads compared to SNP data, with in total 76 based on SNPs and 314 based on microsatellites (Table 4). There

was no correlation between the probabilities of the half-sib dyads that were identified with both markers (Figure 4a, Pearson $r = -0.04$, $N = 39$, $p = 0.806$).

The mean difference in relatedness was low and stable between ca. 80–100 SNP loci (Figure 5b). In contrast, when looking at the microsatellite data (Figure 5a), the mean difference in relatedness compared between 1 and 16 loci was overall higher than that measured by SNPs (Figure 5b).

3.5 | Estimating individual heterozygosity and its accuracy

There was a positive, but relatively moderate correlation, between sMLH measured using 16 microsatellites and 4,876 SNPs (Figure 4b, Pearson $r = 0.45$, $t = 4.29$, $df = 73$, $p < 0.001$). The range of sMLH

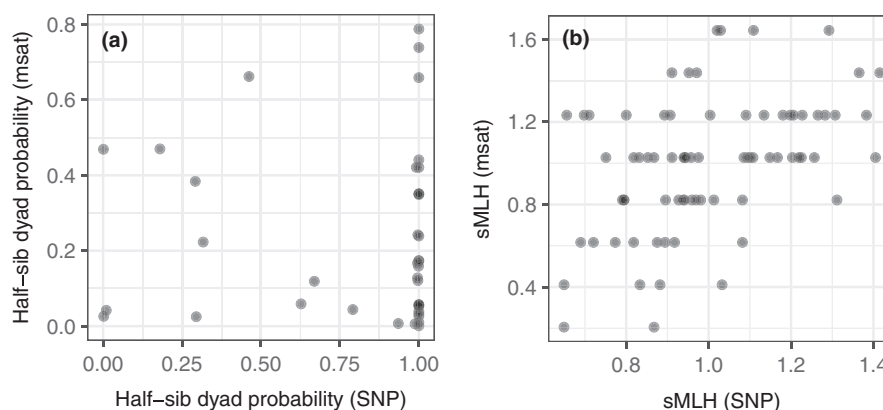


FIGURE 4 Scatter plots showing *Salmo trutta* half-sib assignment probabilities across 39 dyads identified with both SNP and microsatellite markers (a) and individual sMLH values for 75 individuals from both markers (b). Results from the whole dataset of SNPs shown in x-axis and from 16 microsatellites in the y-axis

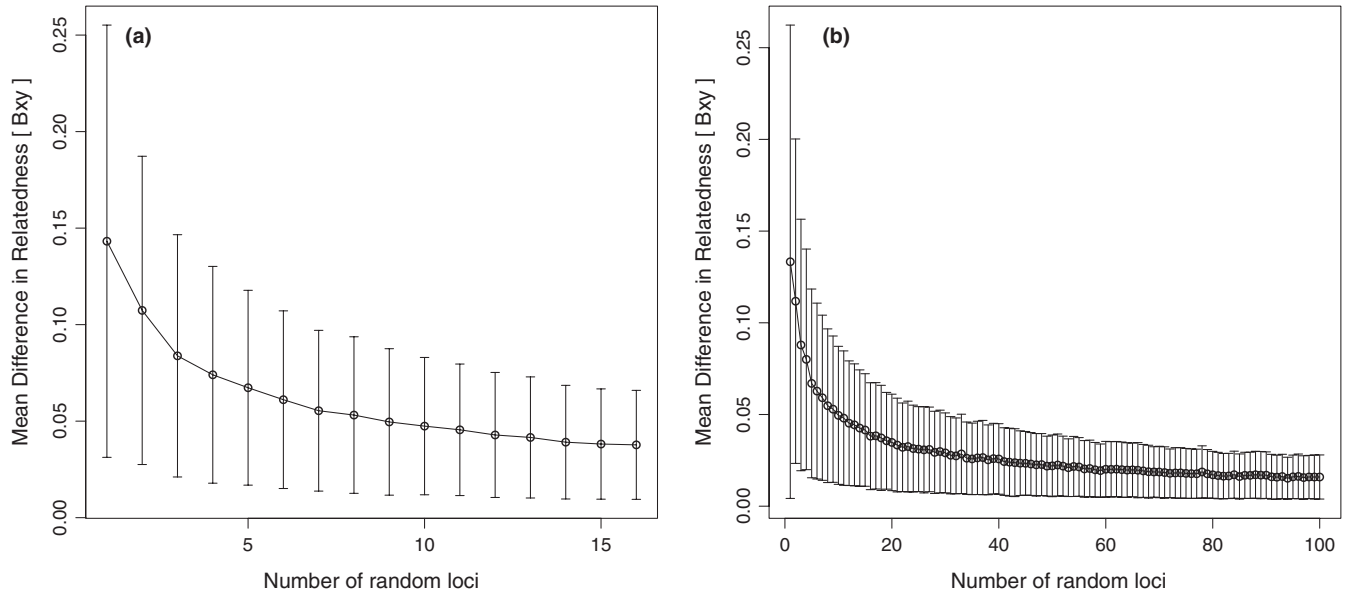


FIGURE 5 Pairwise differences in relatedness using subsets of loci from RADseq (a) or microsatellite data (b) from the same 75 *Salmo trutta* individuals. Pairwise relatedness between individuals was compared between each subset and the maximum number of loci used (100 SNP or 16 microsatellite loci)

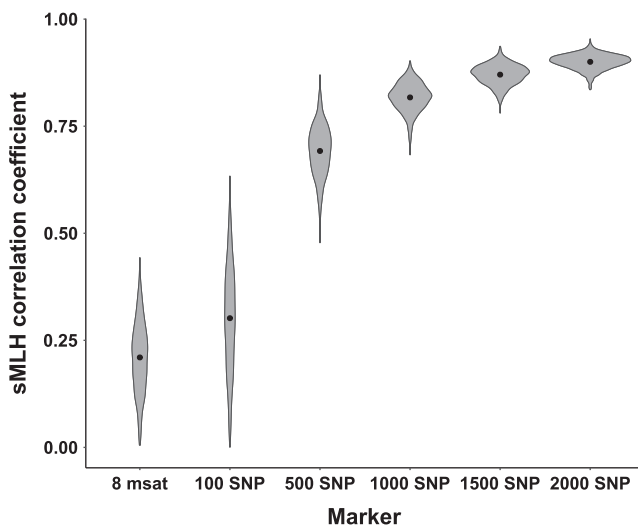


FIGURE 6 Violin plots showing correlations between subsets of sMLH values in *Salmo trutta* according to different markers. For both microsatellites (msats) and SNPs, the correlation between two equal-size randomized subsets was calculated for 1,000 replicated sets of loci. Points showing means within each subset

was much higher for microsatellites (Figure 4b) for which the average within-population standard deviation was 2.5 times higher than for the SNP markers. Further, analysis on subsets of SNP markers revealed that the individual genetic diversity was most precisely measured with 1,500 or more SNPs (mean Pearson $r = 0.87$ for 1,500 SNP subset, $r = 0.90$ for 2,000 SNP subset, Figure 6). There were no clear differences between populations in the number of SNPs required for a high correlation, but Tuhkajoki and Pohjajoki had higher variation in correlations (based on only 9 and 11 samples; Supporting information Figure S1).

4 | DISCUSSION

We conducted a comprehensive evaluation of genetic divergence and diversity, and family structure in small, freshwater, postglacial populations of brown trout using both SNP and microsatellite markers. The results first showed that a moderately diverse microsatellite panel of 16 loci covering a total of 147 alleles can produce similar results as >4,800 SNPs obtained by RADseq for quantifying population divergence, and second, that the resolution of SNPs is higher compared to microsatellites in a multivariate analysis. Third, the results suggest that thousands of SNP markers are needed to reliably estimate individual-level heterozygosity.

4.1 | Performance of markers at population and individual levels

Moderate and high population divergence was equally well reflected by the two marker types, supporting work in other species. For instance, in the Atlantic salmon, as few as nine SNP markers produced F_{ST} values that were correlated to those measured using 14 microsatellites (Ryynänen, Tonteri, Vasemägi, & Primmer, 2007). Other studies have also reported very similar estimates of population divergence between less than a dozen microsatellite loci and >1,000 SNPs in European honeybees (*Apis mellifera mellifera*; Muñoz et al., 2017), *Arabidopsis halleri* (Fischer et al., 2017) and round whitefish (*Prosopium cylindraceum*; Morgan et al., 2017). However, in Crucian carp (*Carassius carassius*), a greater isolation-by-distance was identified by RADseq than by microsatellites across Northern Europe (Jeffries et al., 2016). As in Jeffries et al. (2016), we found stronger divergence between populations using the RADseq data in DAPC when comparing the two marker types

from the same individuals. A recent population genetics study on mud crab (*Rhithropanopeus harrisi*) also concluded that analyzing even a few individuals from each population with RADseq can be highly informative (Forsström, Ahmad, & Vasemägi, 2017). Alternatives to RADseq-based approach include, for instance, reduced sets of SNP loci, which can perform nearly as well as thousands of loci in highly diverged populations (Henriques et al., 2018). In addition, increased resolution can be achieved by sequencing a large number of microsatellites first identified with a whole genome sequence scan (Bradbury et al., 2018).

Our results imply that one of the major advantages of RADseq over microsatellite analysis lies in the power to detect family structure within a population. This was shown by much higher probabilities of identified full-sib and half-sib families. The lower probabilities obtained with the microsatellite loci can be explained by the small N_e in the studied populations. In addition, we found only ≤ 4 alleles in approximately half of the microsatellite loci within populations indicating relatively low diversity dominated by few alleles. Similar allelic richness as in our study have been described also in other brown trout populations from the wild and from hatcheries using partly the same loci (Aho, Rönn, Piironen, & Björklund, 2006; Koljonen et al., 2013; Swatdipong et al., 2010). Low within-population diversity could be a general phenomenon in endangered populations, suggesting that the results can be applied to other species.

Despite the higher probabilities compared to microsatellite markers, half-sib identification using SNPs could be confounded by genotyping errors, but little research has been done to investigate this thus far. However, a previous study comparing a custom-made SNP panel to 10 microsatellite loci for family identification in brown trout found relatively low overlap in full-sib identification between the markers, and more repeatable results when using *ca.* 3,800 SNPs than when using microsatellites (Linlokken, Haugen, Mathew, Johansen, & Lien, 2016). In contrast, 14 microsatellites and 1,728 SNPs agreed on 98% of full-sibs identified from 255 individuals in brown trout from River Altja (Ahmad, Debes, Palomar, & Vasemägi, 2018; Debes et al., 2017). The different results on family identification with the two marker types can therefore be partly explained by sample size; in this study, < 30 individuals from each population were genotyped using both markers, and Linlokken et al. (2016) used 47–48 individuals per population from three populations. In addition, the frequency and diversity of alleles in each population finally determines the ability of microsatellites to correctly identify family structure. Notably, the use of microhaplotype markers, i.e., markers consisting of regions carrying several SNPs, has emerged as an advantageous approach for population genetic studies and particularly relatedness inference (Baetscher, Clemento, Ng, Anderson, & Garza, 2018). Such approach can be performed with RADSeq data and it has already been used in salmonids, e.g., for Chinook salmon stock identification (McKinney, Seeb, & Seeb, 2017).

Accurate quantification of genetic diversity at the individual level is crucial for HFC studies, which are comparing genetic diversity at specific loci or at the genome-wide level to variation in fitness-related traits (Chapman et al., 2009). HFC studies evaluate if the decreased

genetic diversity (i.e., inbreeding depression) can reduce fitness in wild populations or make them more vulnerable to environmental disturbances due to lower adaptability to novel challenges (Hedrick & Kalinowski, 2000; Willi, Buskirk, & Hoffmann, 2006). Before SNP markers became widely available, HFC studies typically evaluated links between fitness and at heterozygosity at single or few highly variable loci (Balloux et al., 2004). With the increasing use of RADseq and other SNP-based genotyping tools, more empirical studies on wider range of species are needed to evaluate the number of markers required to reliably quantify genome-wide patterns of diversity. Recently, Hoffman et al. (2014) measured correlations in sMLH between subsets of SNP loci similar to this study and found that high accuracy was achieved using 2,000 or more loci. On the other hand, Miller et al. (2014) showed differences in the number of markers required for precise sMLH between two populations of the bighorn sheep (*Ovis canadensis*) using 20–412 SNPs and 5–100 microsatellites: in one population, 20 microsatellites performed as well as 75 SNPs, but in a different population more loci would have been needed from both markers. In our study, we did not find clear differences in the number of SNPs required for precise estimates of sMLH between populations. The lack of population-specific differences in sMLH preciseness in our study suggests that the population effects in Miller et al. (2014) might be due to a lower number of SNPs used, or differences in overall genetic diversity between populations. Moreover, in *A. halleri*, increasing the number of SNPs beyond 4,000 up to 300,000 led to an increase in the accuracy of individual genome-wide heterozygosity (Fischer et al., 2017). While the most precise SNP subsets ($> 1,500$ loci) in our study reached ~ 0.9 heterozygosity-heterozygosity correlations, perfectly evaluating the accuracy (i.e., a correlation of 1.0 across all loci) of the sMLH estimates would, however, require whole genome data. Overall, both our results and published work indicates that thousands of SNPs are needed to provide accurate estimates of genome-wide diversity, while estimates based on a few hundred or less SNP markers should be used cautiously.

4.2 | Implications for conservation and management of brown trout

All the rivers from which the study populations originated have undergone some alterations in the environment due to logging, dams, mining or forestry-induced decrease in water quality. These effects combined with historically high fishing pressure have further depressed the number of spawning individuals in the study populations. This explains the low genetic diversity and the N_e estimates that were below 25 according to SNPs. In previous studies on other focal systems, estimates of N_e below 50 have often been observed, which are in line with our observations (Linlokken et al., 2016; Linlokken, Johansen, & Wilson, 2014; Sonstebo, Borgstrom, & Heun, 2007; Vøllestad, 2017) suggesting that many brown trout populations in brooks are naturally small and strongly genetically differentiated. Each of the studied wild populations appeared to be an isolated unit with very limited gene flow from the other rivers or from the hatchery-reared fish, which is in line with their historical and present-day connectivity. Although

Pohjajoki and Vaarainjoki are both connected to Lake Oulujärvi, Vaarainjoki first discharges to Lake Kivesjärvi, which is connected to Oulujärvi through River Varisjoki. It is thus unlikely that resident brown trout would make a journey from Vaarainjoki to Pohjajoki, or *vice versa*. On the other hand, the resident Tuhkajoki population has been separated from the other populations by a natural migration barrier (nowadays also by multiple dams), thus explaining its divergence from the other wild populations. Overall, differences among populations can also be partly explained by founder effects, as they may have been initially formed by few individuals after the last glacial period since *ca.* 10,000 years, as well as by genetic drift, partly due to their resident life-history strategy and small population size even in fully natural conditions. The RADseq approach applied here was also used to assess the population genetic structure of brown trout in River Koutajoki watershed in Finland (Lemopoulos, Uusi-Heikkilä, Vasemägi et al., 2018a), where the pairwise F_{ST} values were in a similar range as in the River Oulujoki watershed. These values are comparable to other brown trout populations with migration barriers (0.099 ± 0.005), albeit higher than those of populations without migration barriers (0.043 ± 0.003 ; Koljonen et al., 2013; values from Vøllestad, 2017).

Both marker types evidenced that stocked fish can disperse to the stream habitats of resident trout, as one individual of identified hatchery origin was caught in Vaarainjoki during spawning time. Hatchery fish are generally stocked directly to Lake Oulujärvi, the close-by Lake Kivesjärvi, or to rivers Varisjoki or Kongasjoki, as 2- or 3-year-old smolts, and can thereafter disperse to Vaarainjoki. Small-scale introductions may have been made also in Tuhkajoki, which could explain the presence of hatchery genotypes in this population. Several studies have shown that introgression from hatchery stocks to wild fish populations occurs especially when the number of stocked fish is high compared to the number of wild individuals (Hansen, 2002; Hindar, Ryman, & Utter, 1991; Ozerov et al., 2016; Salminen, Koljonen, Säisä, & Ruuhijarvi, 2012), as is the case in Lake Oulujärvi area. Further, introgression from hatchery-reared fish can occur at sites distantly located to the stocking location (Vasemägi, Gross, Paaver, Koljonen, & Nilsson, 2005; Finnegan & Stevens 2008). Wild brown trout populations may be locally adapted to their native environments (Jensen et al., 2008; Westley, Ward, & Fleming, 2013) and introductions of genetically differentiated hatchery-reared fish could negatively impact such adaptations (Reed et al., 2015) in case they lack the same adaptative characteristics. Additional genetic studies are required to understand the risks and potential benefits of interbreeding between the resident wild stocks and the migratory stock maintained in captivity and used for stockings and enhancement projects in the region.

Our results suggest that the existing wild populations in the region cannot be supported with the hatchery brood stock without risking the unique genetic composition in the wild populations. Thus, the management decisions on the wild brown trout populations need to balance the low diversity and N_e indicating a high extinction risk with the conservation of seemingly unique genetic composition. Particularly, large enough unique populations should not be mixed

with hatchery stocking in line with the concept of Evolutionarily Significant Units proposed by Fraser and Bernatchez (2001), while the extremely small populations might benefit from the increase of genetic variation from controlled stocking with regional strains.

4.3 | Recommendations for conservation genetics studies

The resolution power of the marker depends on the number and frequency of alleles and loci available for each population. Microsatellites can usually carry a several-fold higher number of alleles than SNPs, but the number of loci can be increased enormously in RADseq analysis compared to microsatellites. Thus, the abundance of SNP loci creates an inevitable advantage in resolution power, especially in cases where microsatellites may not function optimally due to high relatedness within populations or population bottlenecks. Consequently, although microsatellites have been central for characterizing post-glacial phylogenetic relationships in salmonids (e.g., Hansen, Mensberg, & Berg, 1999; Koskinen, Knizhin, Primmer, Schlotterer, & Weiss, 2002; Säisä et al., 2005), and many studies report comparable estimates of genetic divergence between microsatellite and SNP markers, RADseq approach could be preferred over microsatellites for detailed phylogeography studies.

Our study indicates the benefits of RADseq can also be pronounced for small populations with limited genetic diversity due to prior population bottlenecks, which is often the case for endangered populations. Thus, the choice of marker to be used in future studies depends on the goals of the study as well as available resources; compared to microsatellite analysis, RADseq still carries higher per-individual costs, although the overall price is decreasing with decreasing sequencing and reagent costs. There is a relatively high labor and time cost associated with establishing a new microsatellite panel, which RADseq circumvents by having similar preparation costs regardless of prior knowledge on markers. For projects aiming solely at describing population divergence across landscapes using a multiallelic microsatellites panel previously developed for the focal species, the benefits of switching to a RADseq approach can be limited. In contrast, many detailed questions on relatedness, demography, selection, genetic diversity or fine-scale population divergence greatly benefit from an approach based on a large number of SNPs (Andrews et al., 2016). The required number of SNPs for these questions, and therefore sequencing depth and cost, varies by species and populations, but is likely lower for relatedness analysis than for, e.g., HFC studies. In conclusion, SNPs were more informative than microsatellites for describing relatedness and had much higher resolution with low sample sizes from small and isolated populations.

ACKNOWLEDGEMENTS

The study was funded by the Academy of Finland (decision #286261, #266321) and by the Natural Resources Institute Finland (Luke). E. Taskila, P. K. Korhonen, T. Laaksonen, A. Leinonen, A. Toivonen, M.

Hyvönen, K. Ratilainen, J. Aalto, H. Punju, E. Konttinen and H. Laitala are gratefully acknowledged for the help in the sample collection, and O. van der Meer for producing the map of the drainage basin. J.M.P. was also supported by the Finnish Cultural Foundation. The authors would like to thank anonymous reviewers and Dr. Thomas Lilley for their constructive comments on the manuscripts.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

Conceptualization: AVai, SUH, AL, JMP. Investigation: AL, MLK, JK, JMP. Formal analysis: AL, JMP. Writing –Original draft: JMP, AL. Writing –Review and editing: JMP, AL, SUH, AVas, PH, AH, MLK, JK, AVai. Visualization: JMP, AL. Supervision: AVai, SUH, AVas. Funding acquisition: AVai, AH, PH.

DATA ACCESSIBILITY

All sequence data have been deposited in the NCBI Sequence Read Archive (accession number SRP125540). All raw microsatellite data are available in Figshare (<https://doi.org/10.6084/m9.figshare.5314666>).

ORCID

Alexandre Lemopoulos  <https://orcid.org/0000-0002-7997-5526>

Jenni M. Prokkola  <https://orcid.org/0000-0003-2987-4417>

Silva Uusi-Heikkilä  <https://orcid.org/0000-0001-6503-455X>

Anti Vasemägi  <https://orcid.org/0000-0002-2184-5534>

Anssi Vainikka  <https://orcid.org/0000-0002-0172-5615>

REFERENCES

- Ahmad, F., Debes, P. V., Palomar, G., & Vasemägi, A. (2018). Association mapping reveals candidate loci for resistance and anaemic response to an emerging temperature-driven parasitic disease in a wild salmonid fish. *Molecular Ecology*, 27, 1385–1401. <https://doi.org/10.1111/mec.14509>
- Aho, T., Rönn, J., Piironen, J., & Björklund, M. (2006). Impacts of effective population size on genetic diversity in hatchery reared Brown trout (*Salmo trutta* L.) populations. *Aquaculture*, 253, 244–248. <https://doi.org/10.1016/j.aquaculture.2005.09.013>
- Allendorf, F. W., & Hard, J. J. (2009). Human-induced evolution caused by unnatural selection through harvest of wild animals. *Proceedings of the National Academy of Sciences of the United States of America*, 106, 9987–9994. <https://doi.org/10.1073/pnas.0901069106>
- Anderson, J. H., Faulds, P. L., Atlas, W. I., & Quinn, T. P. (2013). Reproductive success of captive bred and naturally spawned Chinook salmon colonizing newly accessible habitat. *Evolutionary Applications*, 6, 165–179. <https://doi.org/10.1111/j.1752-4571.2012.00271.x>
- Andrews, K. R., Good, J. M., Miller, M. R., Luikart, G., & Hohenlohe, P. A. (2016). Harnessing the power of RADseq for ecological and evolutionary genomics. *Nature Reviews Genetics*, 17, 81–92. <https://doi.org/10.1038/nrg.2015.28>
- Baetscher, D. S., Clemento, A. J., Ng, T. C., Anderson, E. C., & Garza, J. C. (2018). Microhaplotypes provide increased power from short-read DNA sequences for relationship inference. *Molecular Ecology Resources*, 18, 296–305. <https://doi.org/10.1111/1755-0998.12737>
- Balkenhol, N., & Waits, L. P. (2009). Molecular road ecology: Exploring the potential of genetics for investigating transportation impacts on wildlife. *Molecular Ecology*, 18, 4151–4164. <https://doi.org/10.1111/j.1365-294X.2009.04322.x>
- Balloux, F., Amos, W., & Coulson, T. (2004). Does heterozygosity estimate inbreeding in real populations? *Molecular Ecology*, 13, 3021–3031.
- Benestan, L., Moore, J.-S., Sutherland, B. J. G., Le Luyer, J., Maaroufi, H., Rougeux, C., ... Bernatchez, L. (2017). Sex matters in massive parallel sequencing: Evidence for biases in genetic parameter estimation and investigation of sex determination systems. *Molecular Ecology*, <https://doi.org/10.1111/mec.14217>
- Bradbury, I. R., Hamilton, L. C., Dempson, B., Robertson, M. J., Bourret, V., Bernatchez, L., & Verspoor, E. (2015). Transatlantic secondary contact in Atlantic Salmon, comparing microsatellites, a single nucleotide polymorphism array and restriction-site associated DNA sequencing for the resolution of complex spatial structure. *Molecular Ecology*, 24, 5130–5144. <https://doi.org/10.1111/mec.13395>
- Bradbury, I. R., Wringe, B. F., Watson, B., Paterson, I., Horne, J., Beiko, R., ... Duffy, S. (2018). Genotyping-by-sequencing of genome-wide microsatellite loci reveals fine-scale harvest composition in a coastal Atlantic salmon fishery. *Evolutionary Applications*, 11, 918–930. <https://doi.org/10.1111/eva.12606>
- Bradford, M. J., & Irvine, J. R. (2000). Land use, fishing, climate change, and the decline of Thompson River, British Columbia, coho salmon. *Canadian Journal of Fisheries and Aquatic Sciences*, 57, 13–16. <https://doi.org/10.1139/f99-283>
- Bruneaux, M., Johnston, S. E., Herczeg, G., Merilä, J., Primmer, C. R., & Vasemägi, A. (2013). Molecular evolutionary and population genomic analysis of the nine-spined stickleback using a modified restriction-site-associated DNA tag approach. *Molecular Ecology*, 22, 565–582. <https://doi.org/10.1111/j.1365-294X.2012.05749.x>
- Cairney, M., Taggart, J. B., & Hoyheim, B. (2000). Characterization of microsatellite and minisatellite loci in Atlantic salmon (*Salmo salar* L.) and cross-species amplification in other salmonids. *Molecular Ecology*, 9, 2175–2178. <https://doi.org/10.1046/j.1365-294X.2000.105312.x>
- Castric, V., Bonney, F., & Bernatchez, L. (2001). Landscape structure and hierarchical genetic diversity in the brook charr, *Salvelinus fontinalis*. *Evolution*, 55, 1016–1028. [https://doi.org/10.1554/0014-3820\(2001\)055\[1016:LSAHGD\]2.0.CO;2](https://doi.org/10.1554/0014-3820(2001)055[1016:LSAHGD]2.0.CO;2)
- Catchen, J., Hohenlohe, P. A., Bassman, S., Amores, A., & Cresko, W. A. (2013). Stacks: An analysis tool set for population genomics. *Molecular Ecology*, 22, 3124–3140. <https://doi.org/10.1111/mec.12354>
- Chapman, J. R., Nakagawa, S., Coltman, D. W., Slate, J., & Sheldon, B. C. (2009). A quantitative review of heterozygosity–fitness correlations in animal populations. *Molecular Ecology*, 18, 2746–2765. <https://doi.org/10.1111/j.1365-294X.2009.04247.x>
- Cochran-Biederman, J. L., Wyman, K. E., French, W. E., & Loppnow, G. L. (2015). Identifying correlates of success and failure of native freshwater fish reintroductions. *Conservation Biology*, 29, 175–186. <https://doi.org/10.1111/cobi.12374>
- Coltman, D. W., Pilkington, J. G., Smith, J. A., & Pemberton, J. M. (1999). Parasite-mediated selection against inbred Soay sheep in a free-living, island population. *Evolution*, 53, 1259–1267. <https://doi.org/10.2307/2640828>
- Davey, J. L., & Blaxter, M. W. (2010). RADseq: Next-generation population genetics. *Briefings in Functional Genomics*, 9, 416–423. <https://doi.org/10.1093/bfpg/elq031>

- Debes, P. V., Gross, R., & Vasemägi, A. (2017). Quantitative genetic variation in, and environmental effects on, pathogen resistance and temperature-dependent disease severity in a wild trout. *American Naturalist*, *190*, 244–265.
- Do, C., Waples, R. S., Peel, D., Macbeth, G. M., Tillett, B. J., & Ovenden, J. R. (2014). NEESTIMATOR v2: Re-implementation of software for the estimation of contemporary effective population size (N_e) from genetic data. *Molecular Ecology Resources*, *14*, 209–214.
- Earl, D. A., & Vonholdt, B. M. (2012). STRUCTURE HARVESTER: A web-site and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, *4*, 359–361. <https://doi.org/10.1007/s12686-011-9548-7>
- Eaton, D. A. R., & Ree, R. H. (2013). Inferring Phylogeny and Introgression using RADseq Data: An example from flowering plants (Pedicularis: Orobanchaceae). *Systematic Biology*, *62*, 689–706. <https://doi.org/10.1093/sysbio/syt032>
- Estoup, A., Presa, P., Krieg, F., Vaiman, D., & Guyomard, R. (1993). (CT) (N) and (GT)(N) microsatellites – a new class of genetic markers for *Salmo trutta* L. (brown trout). *Heredity*, *71*, 488–496. <https://doi.org/10.1038/hdy.1993.167>
- Estoup, A., Rousset, F., Michalakis, Y., Cornuet, J. M., Adriamanga, M., & Guyomard, R. (1998). Comparative analysis of microsatellite and allozyme markers: A case study investigating microgeographic differentiation in brown trout (*Salmo trutta*). *Molecular Ecology*, *7*, 339–353. <https://doi.org/10.1046/j.1365-294X.1998.00362.x>
- Finnegan, A. K., & Stevens, J. R. (2008). Assessing the long-term genetic impact of historical stocking events on contemporary populations of Atlantic salmon, *Salmo salar*. *Fisheries Management and Ecology*, *15*, 315–326. <https://doi.org/10.1111/j.1365-2400.2008.00616.x>
- Fischer, M. C., Rellstab, C., Leuzinger, M., Roumet, M., Gugerli, F., Shimizu, K. K., ... Widmer, A. (2017). Estimating genomic diversity and population differentiation - an empirical comparison of microsatellite and SNP variation in *Arabidopsis halleri*. *BMC Genomics*, *18*. <https://doi.org/10.1186/s12864-016-3459-7>
- Foll, M., & Gaggiotti, O. E. (2008). A genome scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. *Genetics*, *180*, 977–993. <https://doi.org/10.1534/genetics.108.092221>
- Forsström, T., Ahmad, F., & Vasemägi, A. (2017). Invasion genomics: Genotyping-by-sequencing approach reveals regional genetic structure and signatures of temporal selection in an introduced mud crab. *Marine Biology*, *164*, 86. <https://doi.org/10.1007/s00227-017-3210-1>
- Fraser, D. J., & Bernatchez, L. (2001). Adaptive evolutionary conservation: Towards a unified concept for defining conservation units. *Molecular Ecology*, *10*, 2741–2752. <https://doi.org/10.1046/j.1365-294X.2001.t01-1-01411.x>
- Gosselin, T. (2017). grur: An R package tailored for RADseq data imputations. R package version 0.0.1. <https://doi.org/10.5281/zenodo.496176>
- Gosselin, T., & Bernatchez, L. (2016). stackr: GBS/RAD Data Exploration, Manipulation and Visualization using R. R package version 0.4.6. <https://doi.org/10.5281/zenodo.154432>
- Goudet, J., & Jombart, T. (2015). hierfstat: Estimation and Tests of Hierarchical F-Statistics. R package version 0.04-22. <https://CRAN.R-project.org/package=hierfstat>
- Guichoux, E., Lagache, L., Wagner, S., Chaumeil, P., Leger, P., Lepais, O., ... Petit, R. J. (2011). Current trends in microsatellite genotyping. *Molecular Ecology Resources*, *11*, 591–611. <https://doi.org/10.1111/j.1755-0998.2011.03014.x>
- Hansen, M. M. (2002). Estimating the long-term effects of stocking domesticated trout into wild brown trout (*Salmo trutta*) populations: An approach using microsatellite DNA analysis of historical and contemporary samples. *Molecular Ecology*, *11*, 1003–1015. <https://doi.org/10.1046/j.1365-294X.2002.01495.x>
- Hansen, M. M., Mensberg, K. L. D., & Berg, S. (1999). Postglacial recolonization patterns and genetic relationships among whitefish (*Coregonus* sp.) populations in Denmark, inferred from mitochondrial DNA and microsatellite markers. *Molecular Ecology*, *8*, 239–252.
- Hauser, L., Baird, M., Hilborn, R., Seeb, L. W., & Seeb, J. E. (2011). An empirical comparison of SNPs and microsatellites for parentage and kinship assignment in a wild sockeye salmon (*Oncorhynchus nerka*) population. *Molecular Ecology Resources*, *11*, 150–161.
- Hedrick, P. W., & Kalinowski, S. T. (2000). Inbreeding depression in conservation biology. *Annual Review of Ecology and Systematics*, *31*, 139–162. <https://doi.org/10.1146/annurev.ecolsys.31.1.139>
- Henriques, D., Parejo, M., Vignal, A., Wragg, D., Wallberg, A., Webster, M. T., & Pinto, M. A. (2018). Developing reduced SNP assays from whole-genome sequence data to estimate introgression in an organism with complex genetic patterns, the Iberian honeybee (*Apis mellifera iberiensis*). *Evolutionary Applications*, *11*, 1270–1282.
- Henriques, R., von der Heyden, S., Lipinski, M. R., du Toit, N., Kainge, P., Bloomer, P., & Matthee, C. A. (2016). Spatio-temporal genetic structure and the effects of long-term fishing in two partially sympatric offshore demersal fishes. *Molecular Ecology*, *25*, 5843–5861. <https://doi.org/10.1111/mec.13890>
- Hill, W. G. (1981). Estimation of effective population size from data on linkage disequilibrium. *Genetical Research Cambridge*, *38*, 209–216. <https://doi.org/10.1017/S0016672300020553>
- Hindar, K., Ryman, N., & Utter, F. (1991). Genetic effects of cultured fish on natural fish populations. *Canadian Journal of Fisheries and Aquatic Sciences*, *48*, 945–957. <https://doi.org/10.1139/f91-111>
- Hoffman, J. I., Simpson, F., David, P., Rijks, J. M., Kuiken, T., Thorne, M. A. S., ... Dasmahapatra, K. K. (2014). High-throughput sequencing reveals inbreeding depression in a natural population. *Proceedings of the National Academy of Sciences of the United States of America*, *111*, 3775–3780. <https://doi.org/10.1073/pnas.1318945111>
- Hohenlohe, P. A., Amish, S. J., Catchen, J. M., Allendorf, F. W., & Luikart, G. (2011). Next-generation RAD sequencing identifies thousands of SNPs for assessing hybridization between rainbow and westslope cutthroat trout. *Molecular Ecology Resources*, *11*, 117–122.
- Jakobsson, M., & 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*, 1801–1806. <https://doi.org/10.1093/bioinformatics/btm233>
- Jeffries, D. L., Copp, G. H., Handley, L. L., Olsen, K. H., Sayer, C. D., & Hanfling, B. (2016). Comparing RADseq and microsatellites to infer complex phylogeographic patterns, an empirical perspective in the Crucian carp, *Carassius carassius*, L. *Molecular Ecology*, *25*, 2997–3018.
- Jensen, L. F., Hansen, M. M., Pertoldi, C., Holdensgaard, G., Mensberg, K. L. D., & Loeschcke, V. (2008). Local adaptation in brown trout early life-history traits: Implications for climate change adaptability. *Proceedings of the Royal Society B-Biological Sciences*, *275*, 2859–2868. <https://doi.org/10.1098/rspb.2008.0870>
- Jombart, T. (2008). adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics*, *24*, 1403–1405. <https://doi.org/10.1093/bioinformatics/btn129>
- Jombart, T., & Ahmed, I. (2011). adegenet 1.3-1: New tools for the analysis of genome-wide SNP data. *Bioinformatics*, *27*, 3070–3071. <https://doi.org/10.1093/bioinformatics/btr521>
- Jones, A. T., Ovenden, J. R., & Wang, Y. G. (2016). Improved confidence intervals for the linkage disequilibrium method for estimating effective population size. *Heredity*, *117*, 217. <https://doi.org/10.1038/hdy.2016.19>
- Jones, O. R., & Wang, J. L. (2010). COLONY: A program for parentage and sibship inference from multilocus genotype data. *Molecular Ecology Resources*, *10*, 551–555. <https://doi.org/10.1111/j.1755-0998.2009.02787.x>
- Kassambara, A. (2014). easyGgplot2: Perform and customize easily a plot with ggplot2. R package version 1.0.0.9000. <http://www.sthda.com/english/wiki/easyggplot2>
- Keyghobadi, N. (2007). The genetic implications of habitat fragmentation for animals. *Canadian Journal of Zoology*, *85*, 1049–1064.

- Koljonen, M. L., Janatuinen, A., Saura, A., & Koskiniemi, J. (2013) Genetic structure of Finnish and Russian sea trout populations in the Gulf of Finland area. Working papers of the Finnish Game and Fisheries Institute 25/2013. ISBN 978-952-303-067-1.
- Koljonen, M. L., Gross, R., & Koskiniemi, J. (2014). Wild Estonian and Russian sea trout (*Salmo trutta*) in Finnish coastal sea trout catches: Results of genetic mixed-stock analysis. *Hereditas*, *151*, 177–195.
- Kopelman, N. M., Mayzel, J., Jakobsson, M., Rosenberg, N. A., & Mayrose, I. (2015). Clumpak: A program for identifying clustering modes and packaging population structure inferences across K. *Molecular Ecology Resources*, *15*, 1179–1191.
- Koskinen, M. T., Knizhin, I., Primmer, C. R., Schlotterer, C., & Weiss, S. (2002). Mitochondrial and nuclear DNA phylogeography of *Thymallus* spp. (grayling) provides evidence of ice-age mediated environmental perturbations in the world's oldest body of fresh water, Lake Baikal. *Molecular Ecology*, *11*, 2599–2611. <https://doi.org/10.1046/j.1365-294X.2002.01642.x>
- Leitwein, M., Guinand, B., Pouzadoux, J., Desmarais, E., Berrebi, P., & Gagnaire, P. A. (2017). A dense brown trout (*Salmo trutta*) linkage map reveals recent chromosomal rearrangements in the *Salmo* genus and the impact of selection on linked neutral diversity. *G3*, *7*, 1365–1376.
- Lemopoulos, A., Uusi-Heikkilä, S., Huusko, A., Vasemägi, A., & Vainikka, A. (2018b). Comparison of migratory and resident populations of brown trout reveals candidate genes for migration tendency. *Genome Biology and Evolution*, *10*, 1493–1503.
- Lemopoulos, A., Uusi-Heikkilä, S., Vasemägi, A., Huusko, A., Kokko, H., & Vainikka, A. (2018a). Genome-wide divergence patterns support fine-scaled genetic structuring associated with migration tendency in brown trout. *Canadian Journal of Fisheries and Aquatic Sciences*, *75*, 1680–1692.
- Li, C. C., & Horvitz, D. G. (1953). Some methods of estimating the inbreeding coefficient. *American Journal of Human Genetics*, *5*, 107–117.
- Linlokken, A. N., Haugen, T. O., Mathew, P. K., Johansen, W., & Lien, S. (2016). Comparing estimates of number of breeders N-b based on microsatellites and single nucleotide polymorphism of three groups of brown trout (*Salmo trutta* L.). *Fisheries Management and Ecology*, *23*, 152–160.
- Linlokken, A. N., Johansen, W., & Wilson, R. (2014). Genetic structure of brown trout, *Salmo trutta*, populations from differently sized tributaries of Lake Mjøsa in south-east Norway. *Fisheries Management and Ecology*, *21*, 515–525.
- Lischer, H. E. L., & Excoffier, L. (2012). PGDSpider: An automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics*, *28*, 298–299. <https://doi.org/10.1093/bioinformatics/btr642>
- McConnell, S., Hamilton, L., Morris, D., Cook, D., Paquet, D., Bentzen, P., & Wright, J. (1995). Isolation of salmonid microsatellite loci and their application to the population genetics of Canadian east coast stocks of Atlantic salmon. *Aquaculture*, *137*, 19–30. [https://doi.org/10.1016/0044-8486\(95\)01111-0](https://doi.org/10.1016/0044-8486(95)01111-0)
- McKinney, G. J., Seeb, J. E., & Seeb, L. W. (2017). Managing mixed-stock fisheries: Genotyping multi-SNP haplotypes increases power for genetic stock identification. *Canadian Journal of Fisheries and Aquatic Sciences*, *74*, 429–434. <https://doi.org/10.1139/cjfas-2016-0443>
- Miller, J. M., Malenfant, R. M., David, P., Davis, C. S., Poissant, J., Hogg, J. T., ... Coltman, D. W. (2014). Estimating genome-wide heterozygosity: Effects of demographic history and marker type. *Heredity*, *112*, 240–247. <https://doi.org/10.1038/hdy.2013.99>
- Morgan, T. D., Graham, C. F., McArthur, A. G., Raphenya, A. R., Boreham, D. R., Manzon, R. G., ... Somers, C. M. (2017). Genetic population structure of the round whitefish (*Prosopium cylindraceum*) in North America: Multiple markers reveal glacial refugia and regional subdivision. *Canadian Journal of Fisheries and Aquatic Sciences*, *75*, 836–849.
- Morita, K., & Yamamoto, S. (2002). Effects of habitat fragmentation by damming on the persistence of stream-dwelling charr populations. *Conservation Biology*, *16*, 1318–1323. <https://doi.org/10.1046/j.1523-1739.2002.01476.x>
- Muñoz, I., Henriques, D., Jara, L., Johnston, J. S., Chávez-Galarza, J., De La Rúa, P., & Pinto, M. A. (2017). SNPs selected by information content outperform randomly selected microsatellite loci for delineating genetic identification and introgression in the endangered dark European honeybee (*Apis mellifera mellifera*). *Molecular Ecology Resources*, *17*, 783–795.
- Narum, S. R., Buerkle, C. A., Davey, J. W., Miller, M. R., & Hohenlohe, P. A. (2013). Genotyping-by-sequencing in ecological and conservation genomics. *Molecular Ecology*, *22*, 2841–2847. <https://doi.org/10.1111/mec.12350>
- Near, T. J., MacGuigan, D. J., Parker, E., Struthers, C. D., Jones, C. D., & Dornburg, A. (2018). Phylogenetic analysis of Antarctic notothenioids illuminates the utility of RADseq for resolving Cenozoic adaptive radiations. *Molecular Phylogenetics and Evolution*, *129*, 268–279.
- Nei, M. (1973). Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences of the United States of America*, *70*, 3321–3323. <https://doi.org/10.1073/pnas.70.12.3321>
- O'Reilly, P. T., Hamilton, L. C., McConnell, S. K., & Wright, J. M. (1996). Rapid analysis of genetic variation in Atlantic salmon (*Salmo salar*) by PCR multiplexing of dinucleotide and tetranucleotide microsatellites. *Canadian Journal of Fisheries and Aquatic Sciences*, *53*, 2292–2298.
- Ozerov, M. Y., Gross, R., Bruneaux, M., Vähä, J. P., Burimski, O., Pukk, L., & Vasemägi, A. (2016). Genomewide introgressive hybridization patterns in wild Atlantic salmon influenced by inadvertent gene flow from hatchery releases. *Molecular Ecology*, *25*, 1275–1293. <https://doi.org/10.1111/mec.13570>
- Paradis, E. (2010). pegas: An R package for population genetics with an integrated-modular approach. *Bioinformatics*, *26*, 419–420. <https://doi.org/10.1093/bioinformatics/btp696>
- Paris, J. R., King, R. A., & Stevens, J. R. (2015). Human mining activity across the ages determines the genetic structure of modern brown trout (*Salmo trutta* L.) populations. *Evolutionary Applications*, *8*, 573–585.
- Paris, J. R., Stevens, J. R., & Catchen, J. M. (2017). Lost in parameter space: A road map for STACKS. *Methods in Ecology and Evolution*, *8*, 1360–1373
- Paterson, S., Piertney, S. B., Knox, D., Gilbey, J., & Verspoor, E. (2004). Characterization and PCR multiplexing of novel highly variable tetranucleotide Atlantic salmon (*Salmo salar* L.) microsatellites. *Molecular Ecology Notes*, *4*, 160–162. <https://doi.org/10.1111/j.1471-8286.2004.00598.x>
- Poteaux, C., Bonhomme, F., & Berrebi, P. (1999). Microsatellite polymorphism and genetic impact of restocking in Mediterranean brown trout (*Salmo trutta* L.). *Heredity*, *82*, 645–653. <https://doi.org/10.1046/j.1365-2540.1999.00519.x>
- Presca, P., & Guyomard, R. (1996). Conservation of microsatellites in three species of salmonids. *Journal of Fish Biology*, *49*, 1326–1329. <https://doi.org/10.1111/j.1095-8649.1996.tb01800.x>
- Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155*, 945–959.
- Putman, A. I., & Carbone, I. (2014). Challenges in analysis and interpretation of microsatellite data for population genetic studies. *Ecology and Evolution*, *4*, 4399–4428. <https://doi.org/10.1002/ece3.1305>
- Reed, T. E., Prodohl, P., Hynes, R., Cross, T., Ferguson, A., & McGinnity, P. (2015). Quantifying heritable variation in fitness-related traits of wild, farmed and hybrid Atlantic salmon families in a wild river environment. *Heredity*, *115*, 173–184. <https://doi.org/10.1038/hdy.2015.29>
- Rosenberg, N. A. (2004). DISTRUCT: A program for the graphical display of population structure. *Molecular Ecology Notes*, *4*, 137–138. <https://doi.org/10.1046/j.1471-8286.2003.00566.x>
- Ryynänen, H. J., Tonteri, A., Vasemägi, A., & Primmer, C. R. (2007). A comparison of biallelic markers and microsatellites for the estimation

- of population and conservation genetic parameters in Atlantic salmon (*Salmo salar*). *Journal of Heredity*, 98, 692–704. <https://doi.org/10.1093/jhered/esm093>
- Säisä, M., Koljonen, M. L., Gross, R., Nilsson, J., Tahtinen, J., Koskiniemi, J., & Vasemägi, A. (2005). Population genetic structure and postglacial colonization of Atlantic salmon (*Salmo salar*) in the Baltic Sea area based on microsatellite DNA variation. *Canadian Journal of Fisheries and Aquatic Sciences*, 62, 1887–1904.
- Salminen, M., Koljonen, M. L., Säisä, M., & Ruuhijarvi, J. (2012). Genetic effects of supportive stockings on native pikeperch populations in boreal lakes – three cases, three different outcomes. *Hereditas*, 149, 1–15. <https://doi.org/10.1111/j.1601-5223.2011.02230.x>
- Scribner, K. T., Gust, J. R., & Fields, R. L. (1996). Isolation and characterization of novel salmon microsatellite loci: Cross-species amplification and population genetic applications. *Canadian Journal of Fisheries and Aquatic Sciences*, 53, 833–841. <https://doi.org/10.1139/f95-254>
- Seddon, P. J., Armstrong, D. P., & Maloney, R. F. (2007). Developing the science of reintroduction biology. *Conservation Biology*, 21, 303–312. <https://doi.org/10.1111/j.1523-1739.2006.00627.x>
- Slettan, A., Olsaker, I., & Lie, O. (1995). Atlantic salmon, *Salmo salar*, microsatellites at the SSOSL25, SSOSL85, SSOSL311, SSOSL417 loci. *Animal Genetics*, 26, 281–282.
- Slettan, A., Olsaker, I., & Lie, O. (1996). Polymorphic Atlantic salmon, *Salmo salar* L., microsatellites at the SSOSL438, SSOSL439 and SSOSL444 loci. *Animal Genetics*, 27, 57–58. <https://doi.org/10.1111/j.1365-2052.1996.tb01180.x>
- Sonstebo, J. H., Borgstrom, R., & Heun, M. (2007). A comparison of AFLPs and microsatellites to identify the population structure of brown trout (*Salmo trutta* L.) populations from Hardangervidda, Norway. *Molecular Ecology*, 16, 1427–1438. <https://doi.org/10.1111/j.1365-294X.2007.03256.x>
- Stadele, V., & Vigilant, L. (2016). Strategies for determining kinship in wild populations using genetic data. *Ecology and Evolution*, 6, 6107–6120. <https://doi.org/10.1002/ece3.2346>
- Stoffel, M. A., Esser, M., Kardos, M., Humble, E., Nichols, H., David, P., & Hoffman, J. I. (2016). inbreedR: An R package for the analysis of inbreeding based on genetic markers. *Methods in Ecology and Evolution*, 7, 1331–1339.
- Swatdipong, A., Vasemägi, A., Niva, T., Koljonen, M. L., & Primmer, C. R. (2010). High level of population genetic structuring in lake-run brown trout, *Salmo trutta*, of the Inari Basin, northern Finland. *Journal of Fish Biology*, 77, 2048–2071. <https://doi.org/10.1111/j.1095-8649.2010.02784.x>
- Thrasher, D. J., Butcher, B. G., Campagna, L., Webster, M. S., & Lovette, I. J. (2018). Double-digest RAD sequencing outperforms microsatellite loci at assigning paternity and estimating relatedness: A proof of concept in a highly promiscuous bird. *Molecular Ecology Resources*, 00, 1–13. <https://doi.org/10.1111/1755-0998.12771>
- Tiira, K., Laurila, A., Enberg, K., Piironen, J., Aikio, S., Ranta, E., & Primmer, C. R. (2006). Do dominants have higher heterozygosity? Social status and genetic variation in brown trout, *Salmo trutta*. *Behavioral Ecology and Sociobiology*, 59, 657–665. <https://doi.org/10.1007/s00265-005-0094-8>
- Vähä, J. P., Erkinaro, J., Niemelä, E., & Primmer, C. R. (2007). Life-history and habitat features influence the within-river genetic structure of Atlantic salmon. *Molecular Ecology*, 16, 2638–2654. <https://doi.org/10.1111/j.1365-294X.2007.03329.x>
- Väli, U., Einarsson, A., Waits, L., & Ellegren, H. (2008). To what extent do microsatellite markers reflect genome-wide genetic diversity in natural populations? *Molecular Ecology*, 17, 3808–3817.
- Vasemägi, A., Gross, R., Paaver, T., Koljonen, M. L., & Nilsson, J. (2005). Extensive immigration from compensatory hatchery releases into wild Atlantic salmon population in the Baltic sea: Spatio-temporal analysis over 18 years. *Heredity*, 95, 76–83. <https://doi.org/10.1038/sj.hdy.6800693>
- Velando, A., Barros, A., & Moran, P. (2015). Heterozygosity-fitness correlations in a declining seabird population. *Molecular Ecology*, 24, 1007–1018. <https://doi.org/10.1111/mec.13092>
- Vøllestad, L. A. (2017). Chapter 5: Understanding brown trout population genetic structure: A Northern-European perspective. In J. Lobón-Cerviá, & N. Sanz (Eds.), *Brown trout. Biology, Ecology and Management* (pp. 127–144). Hoboken, NJ: Wiley, in press.
- Wang, J. L. (2016). Individual identification from genetic marker data: Developments and accuracy comparisons of methods. *Molecular Ecology Resources*, 16, 163–175. <https://doi.org/10.1111/1755-0998.12452>
- Waples, R. S. (2006). A bias correction for estimates of effective population size based on linkage disequilibrium at unlinked gene loci. *Conservation Genetics*, 7, 167–184. <https://doi.org/10.1007/s10592-005-9100-y>
- Waples, R. S., & Anderson, E. C. (2017). Purging putative siblings from population genetic data sets: A cautionary view. *Molecular Ecology*, 26, 1211–1224. <https://doi.org/10.1111/mec.14022>
- Waples, R. S., & Do, C. (2010). Linkage disequilibrium estimates of contemporary Ne using highly variable genetic markers: A largely untapped resource for applied conservation and evolution. *Evolutionary Applications*, 3, 244–262.
- Waples, R. K., Larson, W. A., & Waples, R. S. (2016). Estimating contemporary effective population size in non-model species using linkage disequilibrium across thousands of loci. *Heredity*, 117, 233. <https://doi.org/10.1038/hdy.2016.60>
- Westley, P. A. H., Ward, E. J., Fleming, I. A. (2013). Fine-scale local adaptation in an invasive freshwater fish has evolved in contemporary time. *Proceedings of the Royal Society Part B - Biological Sciences*, 280, 20122327.
- Wickham, H. (2009). ggplot2: Elegant Graphics for Data Analysis. *Ggplot2: Elegant Graphics for Data Analysis*, 1–212.
- Willi, Y., Van Buskirk, J., & Hoffmann, A. A. (2006). Limits to the adaptive potential of small populations. *Annual Review of Ecology Evolution and Systematics*, 37, 433–458. <https://doi.org/10.1146/annurev.ecolsys.37.091305.110145>
- Wofford, J. E. B., Gresswell, R., & Banks, M. A. (2005). Influence of barriers to movement on within-watershed genetic variation of coastal cutthroat trout. *Ecological Applications*, 15, 628–637. <https://doi.org/10.1890/04-0095>
- Zhang, D. X., & Hewitt, G. M. (2003). Nuclear DNA analyses in genetic studies of populations: Practice, problems and prospects. *Molecular Ecology*, 12, 563–584. <https://doi.org/10.1046/j.1365-294X.2003.01773.x>

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Lemopoulos A, Prokkola JM, Uusi-Heikkilä S, et al. Comparing RADseq and microsatellites for estimating genetic diversity and relatedness – Implications for brown trout conservation. *Ecol Evol*. 2019;9:2106–2120. <https://doi.org/10.1002/ece3.4905>