**Supplementary Material – Document S1**

This document provides a more detailed description of library preparation methods, targeted sequence capture methods, and a description of how baits used for targeted sequence capture were designed.

1. **BestRAD Library Preparation for Mapping Families**

DNA was digested for 1 hour at 37° C using the restriction enzyme *PstI*-HF (R3140S, New England Biolabs, Ipswich, Massachusetts). Following digestion, the enzyme was heat killed by incubating at 80° C for 20 minutes. Biotinylated bestRAD adapters (Ali et al. 2016) were ligated to *PstI* overhangs using *T4 Ligase* (M0202M, New England Biolabs, Ipswich, Massachusetts) by incubating at 25° C for 12 hours before heat killing the enzyme at 80° C for 20 minutes. Digested and adapter ligated DNA from all individuals within a plate was then pooled and concentrated into 55ul of low EDTA Tris-EDTA buffer (T0230, Teknova, Hollister, California) before shearing on a Covaris E220 Ultrasonicator (Covaris Inc., Woburn, Massachusetts) using the recommended settings for a 300bp mean fragment length. The fragment size distributions and concentrations for sheared and pooled libraries were assessed using a D5000 Tapestation assay (5067-5365, Agilent Technologies, Santa Clara, California) before proceeding.

RAD loci were isolated using M280 streptavidin beads (11205D, Thermo Fisher Scientific, Waltham, Massachusetts) following the exact protocol from Ali et al. (2016) and sequencing adapters were added using the NEB Next Ultra Library Prep Kit for Illumina (E7370S, New England Biolabs, Ipswich, Massachusetts). Adapters were diluted 1:10 prior to ligation and a unique 6 base pair i7 indexing primer was used for each library so they could be pooled. Libraries were amplified for 10 cycles and library quantity and insert size were assessed using Quantit Picogreen (P11496, Thermo Fisher Scientific, Waltham, Massachusetts) and Tapestation D5000 assays (5067-5365, Agilent Technologies, Santa Clara, California), respectively. Finished libraries were then pooled in sets of two and bead cleaned twice using a 0.9:1 bead-to-DNA ratio Ampure XP cleanup (A63881, Beckman Coulter, Brea, California) to remove any residual indexing primers.

1. **Hybridization Capture for Mapping Families**

After preparing RAD libraries for mapping families, the following procedure was used to selectively enrich for 58,889 RAD loci that were previously found to be polymorphic within lake trout (see next section). Target enrichment reactions were carried out using a MyBaits Custom Target Enrichment kit using manufacturer recommendations (MycroArray, Ann Arbor, Michigan; Protocol Version 3). 400 nanograms of DNA were used as input for each reaction. Hybridization reactions were carried out at 65° C for 24 hours and wash reactions were done at 65-67° C in 1.5mL tubes. Following enrichment, pools were PCR amplified for 9 cycles using the KAPA Library Amplification Kit for Illumina (KK2620, KAPA Biosystems, Wilmington, Massachusetts) with universal primers according to the manufacturer recommended protocol. Final enriched pools were quantified using Quantit Picogreen assays (P11496, Thermo Fisher Scientific, Waltham, Massachusetts) run in triplicate and insert sizes were determined using D5000 Tapestation assays (5067-5365, Agilent Technologies, Santa Clara, California). All libraries were pooled in equal amounts before sequencing.

1. **RAD-Capture Bait Selection and Design**

The following procedures were used to discover polymorphic RAD loci within lake trout, select loci for targeted genotyping, and select the final bait panel used for genotyping lake trout families. Variable RAD loci were discovered using *PstI* RAD sequencing carried out on 48 individuals collected from across the lake trout range using the bestRAD protocol (Ali, et al. 2016). These individuals included 3 individuals from the Lewis Lake hatchery strain, 3 individuals from the Seneca Lake hatchery strain, 3 individuals from the Apostle Island hatchery strain, 3 individuals from the Isle Royale hatchery strain, 9 individuals from the Marquette hatchery strain, 3 individuals from the Green Lake hatchery strain, 12 wild born individuals collected from Lake Huron, 2 humpers from Lake Superior, 2 leans from Lake Superior, 2 siscowet from Lake Superior, 2 individuals from Flathead Lake (Montana, USA), one individual from Lake Opeongo (Ontario, Canada), one individual from Lake of the Woods (Ontario, Canada), one individual from Schrader Lake (Alaska, USA), and one individual from Ugashik Lakes (Alaska, USA). Fifty nanograms of double stranded DNA from each sample was used as input for library preparation. Libraries were prepared exactly as described above; however, libraries were sheared using the recommended protocol for 250bp fragments. Prior to sequencing, the concentration and size of the library measured using a combination of Qubit dsDNA high sensitivity (Thermo Fisher, Waltham, Massachusetts), KAPA Illumina Library Quantification qPCR (KAPA Biosystems, Wilmington, Massachusetts), and Caliper LabChipGX HS DNA assays (Caliper Life Sciences, Waltham, Massachusetts). The library was sequenced in two HiSeq 4000 lanes in the 2X150 PE read format, using HiSeq 4000 SBS reagents (Illumina, San Diego, California).

Reads from 3 males and 3 females from the Marquette hatchery strain were used for *de novo* assembly of RAD loci using *Stacks v1.44* (Catchen et al. 2013). These individuals were chosen because they had exceptionally high read depth. We chose individuals from a single hatchery strain in order to minimize diversity and promote the assembly of long contigs. Fastq files were purged of clonal reads using *clone\_filter*. and reads were trimmed whenever the average base quality score across a sliding window of 4 base pairs dropped below q20 using *Trimmomatic v0.36* (Bolger et al. 2014). Reads were re-oriented such that the bestRAD barcodes were always found at the beginning of read 1 using a custom perl script before demultiplexing with *process\_radtags*. Reads less than 140 bp were discarded and reads greater than 140 bp were cropped to 140bp before being used for read clustering and RAD locus assembly.

Loci were identified using *ustacks* (-m 2 -M 4 --model\_type bounded --alpha 0.05 --bound\_low 0 --bound\_high 0.05 -p 32 -d -r –gapped), *cstacks* (-n 2 –gapped), and *sstacks* (--gapped). The paired end reads for all identified loci were deposited in separate fasta files using *sort\_read\_pairs.pl* and were assembled into contigs using the *exec\_velvet.pl* wrapper script. Resulting contigs were concatenated onto 10 pseudo-scaffolds with 500 Ns between each contig. Contig sequences were renamed based on their coordinates within the pseudo-scaffolds and extracted to a new fasta file, which was normalized using *Picard NormalizeFasta v2.8* (http://broadinstitute.github.io/picard/). The process resulted in the discovery and assembly of 1,292,171 RAD loci. Contig lengths ranged from 170 to 669 base pairs, with an average length of 316.71 bp.

In order to discover variable SNP loci, fastq files from two sequencing lanes (including all 48 individuals) were concatenated, and read quality was assessed for read 1 and read 2 files using *Fastqc v0.11.5* (Andrews 2014). In order to avoid genotyping biases associated with clonal reads, duplicates were removed using the *clone\_filter* program. Sequencing adapter contamination was removed from reads using *Trimmomatic v0.36* and reads were truncated whenever the mean Phred score across a window of 4 nucleotides dropped below q15. We further required reads to be greater than 60 bp after applying the trimming steps above. Reads were then re-oriented using a custom script such that the inline index for each individual was always found at the beginning of read 1. Properly oriented fastq files were demultiplexed by individual barcode using *process\_radtags*. Reads were mapped to the *de novo* assembled reference using bwa-mem and resulting SAM files were sorted, converted to bam format, and indexed using *samtools v1.4 (Li et al. 2009)*. BAM files were genotyped using *HaplotypeCaller* and the Genome Analysis Toolkit (GATK) incremental joint genotyping workflow (*v3.7;* McKenna et al. 2010) and the resulting VCF file was filtered using *bcftools* *v1.4.1* (Li 2011).

Data were filtered using a variety of criteria meant to minimize the prevalence of false-positive variants. Before filtering on any variable, we evaluated the distribution of the variable relative to the frequency of the first alternate allele in order to ensure that we were not systematically truncating the distribution of allele frequencies. We also checked that the distributions of z-scores resulting from Wilcoxon Sign Rank Tests output by GATK *genotypeGVCFs* (Mapping Quality Rank Sum, Base Quality Rank Sum, Read Position Rank Sum distributions) were centered on zero and approximately normal. Deviation from normality or a mean different than 0 would indicate systematic biases associated with sequencing, *de novo* assembly, or genotyping.

For SNP loci, we required that QD (Quality standardized by depth) be greater than 2, SOR (strand odds ratio) be less than 3, MQ (Mapping Quality reported by GATK) be greater than 40, FS (Fisher Strand) be less than 60, Base Quality Rank Sum be between -2 and 2, Mapping Quality Rank Sum be between -2 and 2, Read Position Rank Sum be between -2 and 2, and depth across all samples be less than 5000x. We excluded RAD loci containing greater than 10 variants. Additionally, we required that contigs map to a single location in the Atlantic salmon genome and that mapping locations not overlap by more than the length of the cut-site overhang on the Atlantic Salmon genome (Lien et al. 2016).

All remaining SNPs were masked in the *de novo* reference using *Picard FastaAlternateReferenceMaker* v2.8 and RAD locus consensus sequences were extracted using *bedtools getfasta v2.26* (Quinlan and Hall 2010) for RAD loci containing variable SNP loci. RAD loci were re-oriented such that the remainder of the *PstI* cut-site was always found at the beginning of the contig. Loci were removed from the dataset if the remainder of a *PstI* cut-site existed on both ends of a contig. We then used *RepeatMasker v4.0.7* (Smit et al. 1996; <http://www.repeatmasker.org/>) to mask low complexity sequence and repeats using the Atlantic Salmon repeat library, while allowing for up to 10% divergence. Consensus sequences for RAD loci were trimmed whenever a repeat masked or low complexity region was encountered (with the end containing the cut-site being maintained) and loci less than 200bp in length after trimming were excluded. All remaining consensus sequences were cropped to 200bp and aligned to the complete set of 1,292,171 *de novo* assembled RAD loci using *blat* v0.35 (Kent 2002). Loci were removed if they aligned to an off-target locus with greater than 40% similarity (calculated as (matches-mismatches-gaps)/200). The remaining 64,242 loci were submitted to MycroArray for bait design on July 6th 2017 (Ann Arbor, Michigan). Two baits were designed for each locus; one adjacent to the *PstI* cut-site and another offset 80 bp from the cut-site. These baits were input into MycroArray’s complementary bait quality control software. We retained baits only if the following criteria (reported by MycroArray) were met; the top blast hit to the *de novo* assembly was less than 25% soft masked, delta G > -9, zero blast hits to Atlantic salmon or rainbow trout mtDNA, 0 heterodimers with other baits, at most 10 off target blast hits with Tm between 62.5 and 65° C, and fewer than 2 off target blast hits with Tm >= ° C. For loci where both baits passed all filtering criteria, preference was given to the bait closest to the *PstI* cut-site. This resulted in the retention of baits for 58,889 polymorphic RAD loci, which were subsequently used for targeted enrichment of RAD libraries prepared for mapping families (see above).

**Literature Cited:**

Ali, Omar A., Sean M. O’Rourke, Stephen J. Amish, Mariah H. Meek, Gordon Luikart, Carson Jeffres, and Michael R. Miller. "RAD capture (Rapture): flexible and efficient sequence-based genotyping." Genetics 202, no. 2 (2016): 389-400.

Andrews, Simon. "FastQC: a quality control tool for high throughput sequence data. Version 0.11. 5." Babraham Institute, Cambridge, UK http://www. bioinformatics. babraham. ac. uk/projects/fastqc (2014).

Bolger, Anthony M., Marc Lohse, and Bjoern Usadel. "Trimmomatic: a flexible trimmer for Illumina sequence data." Bioinformatics 30, no. 15 (2014): 2114-2120.

Catchen, Julian, Paul A. Hohenlohe, Susan Bassham, Angel Amores, and William A. Cresko. "Stacks: an analysis tool set for population genomics." Molecular Ecology 22, no. 11 (2013): 3124-3140.

Kent, W. James. "BLAT—the BLAST-like alignment tool." *Genome research* 12, no. 4 (2002): 656-664.

Li, Heng, Bob Handsaker, Alec Wysoker, Tim Fennell, Jue Ruan, Nils Homer, Gabor Marth, Goncalo Abecasis, and Richard Durbin. "The sequence alignment/map format and SAMtools." Bioinformatics 25, no. 16 (2009): 2078-2079.

Li, Heng. "A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data." *Bioinformatics* 27, no. 21 (2011): 2987-2993.

Lien, Sigbjørn, Ben F. Koop, Simen R. Sandve, Jason R. Miller, Matthew P. Kent, Torfinn Nome, Torgeir R. Hvidsten et al. "The Atlantic salmon genome provides insights into rediploidization." Nature 533, no. 7602 (2016): 200.

McKenna, Aaron, Matthew Hanna, Eric Banks, Andrey Sivachenko, Kristian Cibulskis, Andrew Kernytsky, Kiran Garimella et al. "The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data." *Genome research* 20, no. 9 (2010): 1297-1303.

Quinlan, Aaron R., and Ira M. Hall. "BEDTools: a flexible suite of utilities for comparing genomic features." *Bioinformatics* 26, no. 6 (2010): 841-842.

Smit, Arian FA, Robert Hubley, and P. Green. "RepeatMasker." (1996): 1996-2010.