Appendix 2:

RAD Genotyping

Population Samples:

Tissue samples from 64 juvenile red drum sampled from West Matagorda Bay, TX, in the spring of 2008 were randomly chosen from a larger set of samples collected for a prior study; sample collection procedures were described in Carson et al. (2014). DNA was extracted using a phenol-chloroform-isoamyl (PCI) extraction protocol and double-digest restriction-site associated DNA (ddRAD) libraries were prepared following procedures outlined in Peterson et al. (2012). RAD libraries were sequenced on an Illumina HiSeq 2000 DNA sequencer.

Sequence data were first demultiplexed by individual barcode sequence, using the process_radtags program from the STACKS software package (Catchen et al. 2011). Read mapping and SNP calling were performed with the *dDocent* pipeline (Puritz et al. 2014). Reads were mapped to a reduced-representation genome sequence, developed as part of an ongoing study and consisting of red drum RAD fragments sequenced on the Illumina MiSeq platform and assembled to recover the entire sequence of each fragment. The raw SNP dataset (consisting of SNP, indel, and complex polymorphisms, but hereafter referred to as SNPs, unless otherwise specified) consisted of 524,657 SNP genotypes for 64 individuals. Genotypes and individuals were filtered based on numerous criteria. First, initial filtering, to remove loci with more than 50% missing genotypes and individuals with less than 10x mean coverage across loci, removed 73,123 SNPs and 6 individuals from the analysis. Genotypes with a depth of < 10 reads and a Phred quality score of less than 20 were then filtered, followed by removal of loci with more than 5% missing data across individuals and a minor allele frequency of less than 0.05. A total of 6748 SNPs across 3224 RAD-tags remained after these filtering steps. These genotypes were subject to another round of filtering for quality control: filters for allele balance across forward and reverse reads, improper pairing of reads, excessively high depth and low quality score (genotypes), and maximum mean depth (loci). This resulted in 6048 SNPs across 3048 RAD-tags. SNP loci were then filtered further if they did not conform to expectations of Hardy-Weinberg equilibrium (HWE) (*P* value < 0.001); 212 SNPs (66 RAD-tags) were removed. The remaining complex polymorphisms were decomposed to allelic primitives with the vcfallelicprimitives program in the vcflib package (https://github.com/ekg/vcflib), resulting in 6,189 SNPs and indel polymorphisms. These were collapsed into haplotypes for each RAD-tag, using a custom Perl script. The script filtered indel polymorphisms (unless they were the only polymorphism on the RAD-tag) and RAD-tags for which less than 95% of individuals could be successfully haplotyped across the RAD-tag. A GENEPOP file, consisting of (haplotype) genotypes at 2101 multi-allelic RAD-tags was produced using the script and implemented for NE2 and LINKNE analyses.

Mapping Panel:

DNA from tissue samples for two parents and 108 full-sib progeny was extracted and ddRAD libraries prepared as above. Samples were obtained as part of prior studies in which a microsatellite-based linkage map was constructed for red drum; details regarding the mapping cross and sample collection can be found in Portnoy et al. (2010) and references therein.

Bioinformatic processing of SNP data was performed as above, with the exception that filters for HWE were not applied to the data. SNPs were collapsed into haplotypes for each RAD-tag and the resulting genotypes output into JOINMAP (van Ooijen 2012) format. The pairwise recombination rate for each pair of loci was calculated for each parent, using a custom Perl script. When a recombination rate could be estimated for both parents (both loci were segregating in an informative manner), the rate was averaged between parents. When the recombination rate could be estimated for only one parent, the sex-specific rate was used. Recombination rates were output as a square matrix of pairwise values.

References

Carson EW, Bumguardner BW, Fisher M, Saillant E, Gold JR (2014). Spatial and temporal variation in recovery of hatchery-released red drum (Sciaenops ocellatus) in stock-enhancement of Texas bays and estuaries. *Fish Res* **151**: 191–198.

Catchen JM, Amores A, Hohenlohe P, Cresko W, Postlethwait JH (2011). Stacks: building and genotyping Loci de novo from short-read sequences. *G3 (Bethesda)* **1**: 171–82.

Van Ooijen JW (2012). JoinMap 4.1, Software for the calculation of genetic linkage maps in experimental populations of diploid species. *Softw Calc Genet Link maps Exp Popul Kyazma BV, Wageningen, Netherlands*.

Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE (2012). Double digest RADseq: An inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS One* **7**: e37135.

Portnoy DS, Renshaw M a., Hollenbeck CM, Gold JR (2010). A genetic linkage map of red drum, Sciaenops ocellatus. *Anim Genet* **41**: 630–641.

Puritz JB, Hollenbeck CM, Gold JR (2014). dDocent: a RADseq, variant-calling pipeline designed for population genomics of non-model organisms. *PeerJ* **2**: e431.