Genetic profile for Dworshak NFH

Spring Chinook salmon

10/18/2016

Prepared by

Justin Bohling

U.S. Fish and Wildlife Service

Abernathy Fish Technology Center

Applied Program in Conservation Genetics

Longview, WA 98632

Justin_bohling@fws.gov. (360) 425-6072 x 311

National Fish Hatchery Broodstock Genetic Profile

Facility: Dworshak National Fish Hatchery

Stock: Spring Chinook salmon (*Oncorhynchus tshawytscha*)

Parental stock: Hatchery-origin spring Chinook salmon from Carson NFH and Rapid River Fish Hatchery

Year founded: "Carson" stock transferred in 1982, replaced with "Rapid River" stock in 1987

Generation time:

Segregation/Integration history: Segregated. Only hatchery-origin fish trapped at Dworshak are used for broodstock.

Table 1- Broodstock samples analyzed:

Description	Year	*n*	Life stage	Data source	Code
Dworshak NFH	2008	753	Returning adult	FishGen	DWOR2008
Dworshak NFH	2009	766	Returning adult	FishGen	DWOR2009
Dworshak NFH	2010	659	Returning adult	FishGen	DWOR2010
Dworshak NFH	2011	783	Returning adult	FishGen	DWOR2011
Dworshak NFH	2012	1404	Returning adult	FishGen	DWOR2012
Dworshak NFH	2013	1234	Returning adult	FishGen	DWOR2013

Table 2 - Samples analyzed for comparison:

Description	Year	*n*	Life stage	Data source	Code
Kooskia NFH	2008	451	Returning adult	FishGen	KOOS2008
Kooskia NFH	2009	410	Returning adult	FishGen	KOOS2009
Kooskia NFH	2010	568	Returning adult	FishGen	KOOS2010
Kooskia NFH	2011	628	Returning adult	FishGen	KOOS2011
Kooskia NFH	2012	682	Returning adult	FishGen	KOOS2012
Kooskia NFH	2013	624	Returning adult	FishGen	KOOS2013
Nez Perce FH	2008	122	Returning adult	FishGen	NPFH2008
Nez Perce FH	2009	349	Returning adult	FishGen	NPFH2009
Nez Perce FH	2010	465	Returning adult	FishGen	NPFH2010
Nez Perce FH	2011	356	Returning adult	FishGen	NPFH2011
Nez Perce FH	2012	367	Returning adult	FishGen	NPFH2012
Nez Perce FH	2013	391	Returning adult	FishGen	NPFH2013
Powell Satellite Facility	2008	864	Returning adult	FishGen	POWT2008
Powell Satellite Facility	2009	547	Returning adult	FishGen	POWT2009
Powell Satellite Facility	2010	413	Returning adult	FishGen	POWT2010
Powell Satellite Facility	2011	846	Returning adult	FishGen	POWT2011
Powell Satellite Facility	2012	642	Returning adult	FishGen	POWT2012
Powell Satellite Facility	2013	222	Returning adult	FishGen	POWT2013
Rapid River FH	2008	2796	Returning adult	FishGen	RAPH2008
Rapid River FH	2009	2018	Returning adult	FishGen	RAPH2009
Rapid River FH	2010	2291	Returning adult	FishGen	RAPH2010
Rapid River FH	2011	1879	Returning adult	FishGen	RAPH2011
Rapid River FH	2012	1316	Returning adult	FishGen	RAPH2012
Rapid River FH	2013	2004	Returning adult	FishGen	RAPH2013
SF Clearwater Satellite Facility	2008	988	Returning adult	FishGen	SFCW2008
SF Clearwater Satellite Facility	2009	669	Returning adult	FishGen	SFCW2009
SF Clearwater Satellite Facility	2010	741	Returning adult	FishGen	SFCW2010
SF Clearwater Satellite Facility	2011	566	Returning adult	FishGen	SFCW2011
SF Clearwater Satellite Facility	2012	609	Returning adult	FishGen	SFCW2012
SF Clearwater Satellite Facility	2013	435	Returning adult	FishGen	SFCW2013

Genetic markers analyzed:

Ninety-five single nucleotide polymorphism (SNP) markers (Appendix 1).

Table 3: **Diversity within samples**

Expected (H_E) and observed (H_O) heterozygosity, number of loci exhibiting departures from Hardy-Weinberg Equilibrium (HWE), and F_{IS} (a measure of departure from random mating). Values of F_{IS} in bold were bounded by 95% confidence intervals that did not overlap zero.

Number	Population	H_O	H_E	F_{IS}	HWE
1	DWOR2008	0.34	0.35	0.006	7
2	DWOR2009	0.35	0.34	-0.002	14
3	DWOR2010	0.33	0.33	0.003	14
4	DWOR2011	0.34	0.34	-0.008	21
5	DWOR2012	0.35	0.35	0.004	19
6	DWOR2013	0.35	0.35	**-0.014**	12
7	KOOS2008	0.34	0.34	0.006	9
8	KOOS2009	0.34	0.34	0.004	12
9	KOOS2010	0.34	0.34	-0.004	14
10	KOOS2011	0.34	0.34	-0.005	13
11	KOOS2012	0.34	0.34	-0.001	9
12	KOOS2013	0.34	0.34	-0.010	15
13	NPFH2008	0.35	0.35	**0.041**	5
14	NPFH2009	0.35	0.34	-0.011	5
15	NPFH2010	0.35	0.34	**-0.011**	18
16	NPFH2011	0.34	0.35	**0.015**	11
17	NPFH2012	0.35	0.35	-0.008	6
18	NPFH2013	0.34	0.34	0.001	9
19	POWT2008	0.33	0.33	0.002	9
20	POWT2009	0.34	0.34	**0.012**	20
21	POWT2010	0.34	0.34	-0.008	13
22	POWT2011	0.36	0.35	**-0.015**	20
23	POWT2012	0.34	0.33	**-0.010**	13
24	POWT2013	0.35	0.35	-0.006	12
25	RAPH2008	0.34	0.34	-0.002	8
26	RAPH2009	0.34	0.34	-0.004	8
27	RAPH2010	0.34	0.34	**-0.005**	8
28	RAPH2011	0.34	0.34	-0.002	11
29	RAPH2012	0.34	0.34	-0.004	9
30	RAPH2013	0.34	0.34	-0.002	12
31	SFCW2008	0.35	0.34	**-0.015**	14
32	SFCW2009	0.34	0.34	**-0.010**	8
33	SFCW2010	0.35	0.34	**-0.014**	10
34	SFCW2011	0.34	0.35	**0.011**	9
35	SFCW2012	0.35	0.34	-0.009	7
36	SFCW2013	0.34	0.34	0.009	9

Figure 1: Correspondence Analysis (CA) of allele frequencies observed in spring Chinook salmon from Dworshak NFH and adjacent collection sites.

Sample numbers are those listed in Table 3. The primary (horizontal) axis accounted for 58.2% of the inertia in the data and the secondary (vertical) axis accounted for 9.9% of the inertia.

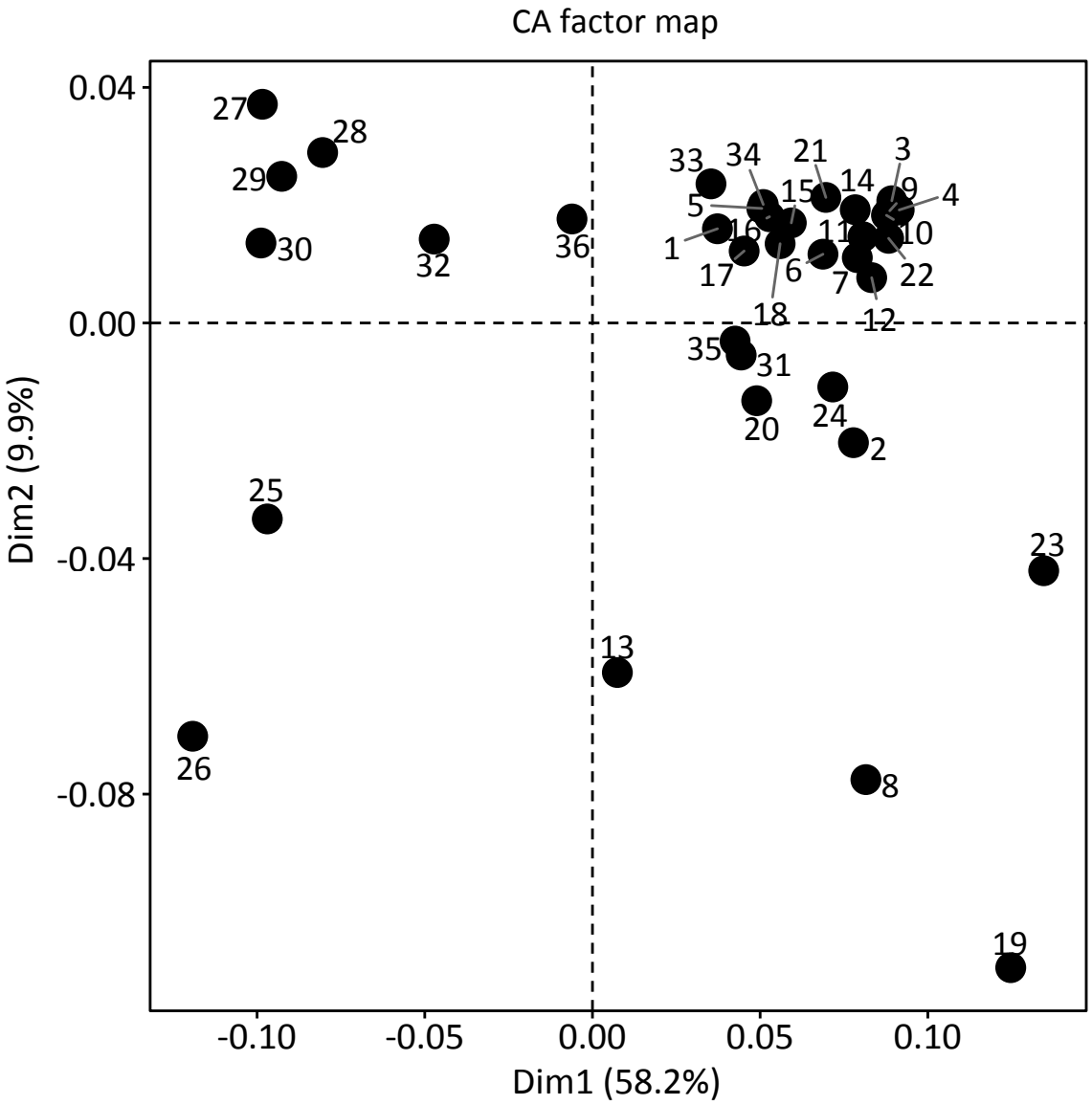


Figure 2: Estimates of effective population size (N_E) for spring Chinook salmon broodstock from Dworshak NFH and adjacent collection sites.

N_E was estimated using the linkage-disequilibrium method. Each point estimate is surrounded by a jackknifed 95% confidence interval.

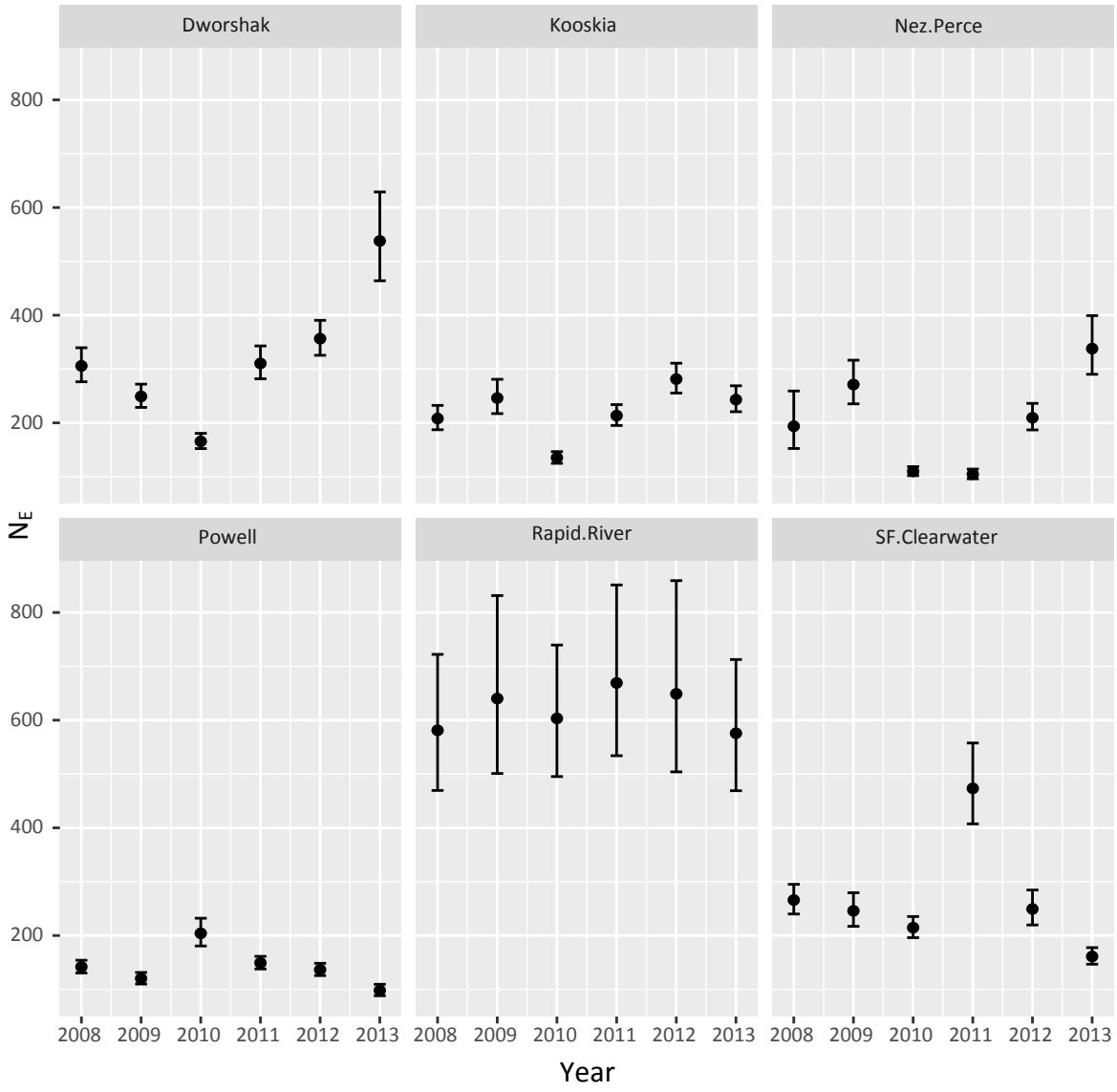
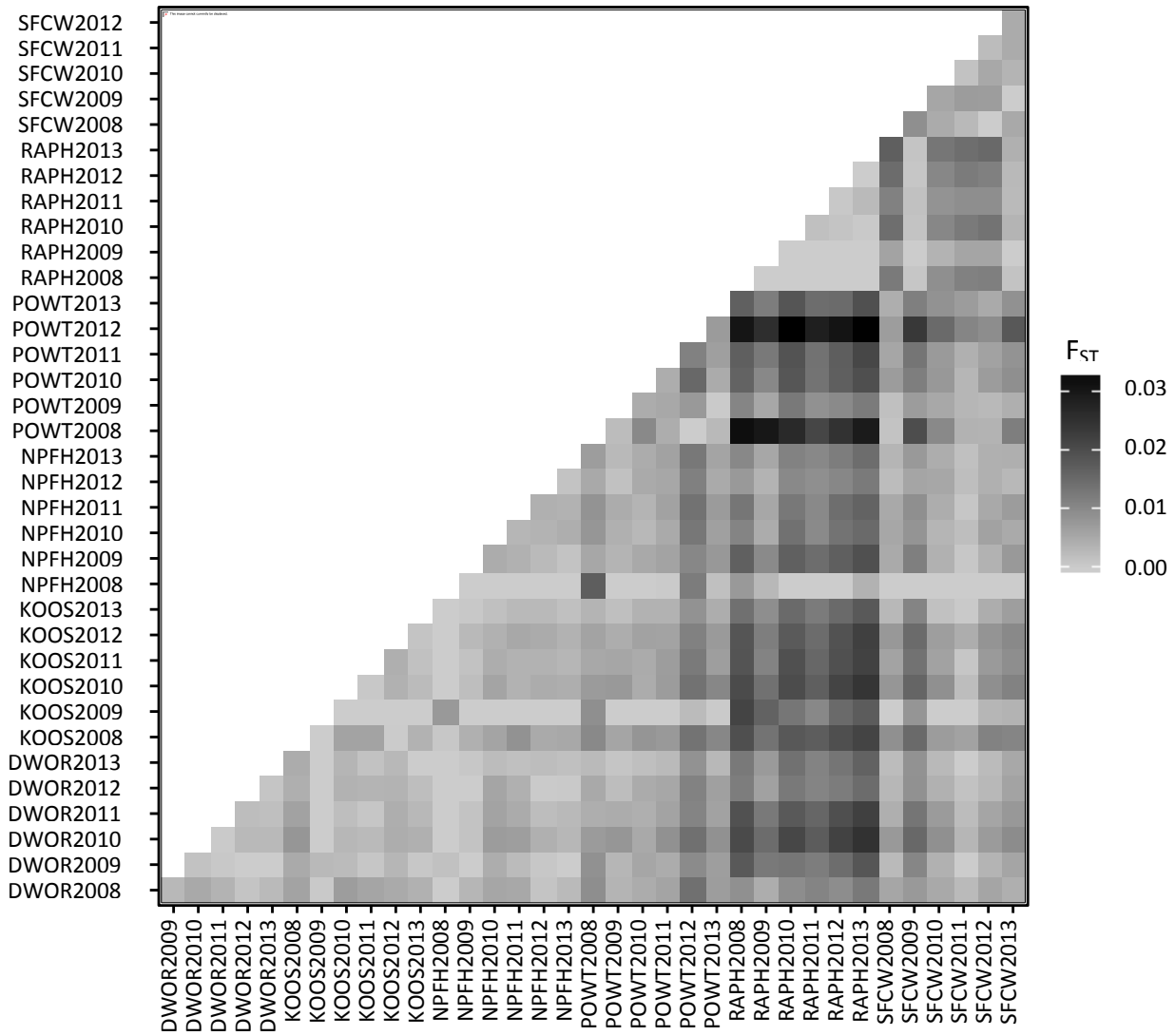


Figure 3: Heatmap of pairwise F_{ST} values between for spring Chinook broodstock from Dworshak NFH and adjacent collection sites.

Estimates were based on using a random sample of 100 fish per broodstock in order to improve computational speed.



Comments

- The goals of this report series are 1) to summarize available genetic information for NFH broodstocks and make that information available to hatchery managers, and 2) to make sure that data for the NFH broodstocks are available for internal hatchery reviews and HET meetings, as well as to our partners.
- Genetic diversity (heterozygosity) was nearly equal across the broodstocks that were analyzed.
- For the number of loci analyzed for this report (95), the number of loci deviating from HWE was high for these broodstocks (between 5-22%). Coupled with the significant F_{IS} values observed from several broodstocks, this suggests deviations from the expectations of random mating. This could be due to the presence of family groups, inbreeding, or genetic drift.
- As revealed by the CA plot, broodstock produced at Dworshak NFH across multiple years clustered high together, suggesting high genetic similarity across cohorts. Broodstock from other hatcheries, especially those from Kooskia NFH and Nez Perce FH, clustered closely to Dworshak stocks.
- The majority of pairwise F_{ST} comparisons were statistically significant, meaning the observed values were greater than expected if there was no differentiation. However, the actual point estimates of F_{ST} (maximum observed value=0.036) were low relative to those observed among wild populations from the Snake River basin. This is a common phenomenon with datasets consisting of a large number of SNPs that provide high statistical power. In practice this suggests low gene flow and high genetic drift in recent generations have fostered allele frequency divergence, but there is low overall differentiation likely due to the common founding history of these stocks.
- Estimates of effective population size at Dworshak NFH were highly variable across years (range 166-538). For four of the six years N_E was between 250-350, suggesting the high (2013) and low (2010) years were outliers. Estimates of N_E were similar to broodstock from other hatcheries aside from Rapid River FH, which produced by far the highest values.

Availability

Genotype data for NFH stocks are available from the FishGen database (www.fishgen.net), and will be provided by AFTC upon request.

Disclaimer

The findings and conclusions in this report are those of the author and do not necessarily represent the views of the U.S. Fish and Wildlife Service.

Appendix 1: Ninety-five single nucleotide polymorphism (SNP) markers used to analyze Dworshak NFH spring Chinook salmon.

Single Nucleotide Polymorphisms		
Ots_100884-287	Ots_94903-99R	Ots_OTALDBINT1-SNP1
Ots_101554-407	Ots_96500-180	Ots_OTDESMIN19-SNP1
Ots_101704-143	Ots_96899-357R	Ots_OTSTF1-SNP1
Ots_102414-395	Ots_ARNT	Ots_P53
Ots_102801-308	Ots_AsnRS-60	Ots_parp3-286
Ots_103122-180	Ots_brp16-64	Ots_pigh-105
Ots_104415-88	Ots_CD59-2	Ots_pop5-96
Ots_105105-613	Ots_CirpA	Ots_ppie-245
Ots_105132-200	Ots_cox1-241	Ots_Prl2
Ots_105385-421	Ots_E2-275	Ots_RAG3
Ots_105407-117	Ots_Est740	Ots_redd1-187
Ots_108820-336	Ots_ETIF1A	Ots_S7-1
Ots_109525-816	Ots_FGF6B_1	Ots_SClkF2R2-135
Ots_110064-383	Ots_GCSH	Ots_SEXY3-1
Ots_110201-363	Ots_GDH-81x	Ots_SWS1op-182
Ots_110495-380	Ots_GPH-318	Ots_TAPBP
Ots_110551-64	Ots_GTH2B-550	Ots_TGFB
Ots_110689-218	Ots_HMGB1-73	Ots_Thio
Ots_112301-43	Ots_hsc71-3'-488	Ots_TLR3
Ots_112419-131	Ots_HSP90B-100	Ots_tpx2-125
Ots_112820-284	Ots_IGF-I.1-76	Ots_txnip-321
Ots_112876-371	Ots_Ikaros-250	Ots_u07-07.161
Ots_113242-216	Ots_IL8R_C8	Ots_u07-17.135
Ots_115987-325	Ots_mapK-3'-309	Ots_u07-18.378
Ots_117432-409	Ots_mapKpr-151	Ots_u07-25.325
Ots_118205-61	Ots_MHC1	Ots_u07-49.290
Ots_118938-325	Ots_MHC2	Ots_u1002-75
Ots_123921-111	Ots_mybp-85	Ots_u211-85
Ots_124774-477	Ots_NFYB-147	Ots_u4-92
Ots_128757-61R	Ots_nkef-192	Ots_u6-75
Ots_129458-451	Ots_NOD1	Ots_unk526
Ots_94857-232R	Ots_ntl-255	Ots_vatf-251

Appendix 2: Pairwise F_{ST} values between spring Chinook salmon broodstock produced at Dworshak NFH and adjacent collection sites.

Sample codes are those listed in Table 2. Bold values indicate comparisons for which the value deviated from zero at the *p*=0.05 level based on 100 permutations. Estimates were based on using a random sample of 100 fish per broodstock in order to improve computational speed.

	DWOR2008	DWOR2009	DWOR2010	DWOR2011	DWOR2012	DWOR2013
DWOR2008	*					
DWOR2009	**0.003**	*				
DWOR2010	**0.005**	**0.002**	*			
DWOR2011	**0.004**	0.001	0.000	*		
DWOR2012	0.001	-0.001	**0.003**	**0.002**	*	
DWOR2013	**0.003**	0.000	**0.003**	**0.002**	0.001	*
KOOS2008	**0.006**	**0.005**	**0.008**	**0.006**	**0.004**	**0.005**
KOOS2009	0.000	**0.003**	-0.001	-0.002	-0.002	-0.005
KOOS2010	**0.007**	**0.003**	**0.003**	**0.002**	**0.004**	**0.003**
KOOS2011	**0.006**	0.001	**0.003**	0.001	**0.004**	**0.001**
KOOS2012	**0.005**	**0.003**	**0.005**	**0.005**	**0.004**	**0.003**
KOOS2013	**0.004**	0.001	**0.004**	**0.003**	**0.002**	-0.001
NPFH2008	-0.004	**0.002**	0.000	-0.003	-0.005	-0.006
NPFH2009	**0.003**	-0.001	**0.001**	**0.001**	-0.001	0.001
NPFH2010	**0.006**	**0.005**	**0.007**	**0.006**	**0.006**	**0.002**
NPFH2011	**0.005**	**0.002**	**0.007**	**0.005**	**0.004**	**0.002**
NPFH2012	0.001	0.001	**0.004**	**0.002**	0.000	**0.002**
NPFH2013	**0.003**	-0.002	**0.003**	**0.003**	0.000	**0.002**
POWT2008	**0.009**	**0.009**	**0.007**	**0.004**	**0.005**	**0.003**
POWT2009	**0.004**	**0.003**	**0.008**	**0.005**	**0.002**	0.001
POWT2010	**0.005**	**0.006**	**0.005**	**0.004**	**0.005**	**0.002**
POWT2011	**0.006**	**0.005**	**0.009**	**0.006**	**0.006**	**0.003**
POWT2012	**0.014**	**0.010**	**0.014**	**0.011**	**0.011**	**0.009**
POWT2013	**0.007**	**0.007**	**0.009**	**0.006**	**0.007**	**0.003**
RAPH2008	**0.009**	**0.018**	**0.021**	**0.020**	**0.012**	**0.013**
RAPH2009	**0.005**	**0.013**	**0.015**	**0.013**	**0.006**	**0.008**
RAPH2010	**0.010**	**0.013**	**0.021**	**0.018**	**0.013**	**0.014**
RAPH2011	**0.011**	**0.012**	**0.018**	**0.016**	**0.011**	**0.011**
RAPH2012	**0.010**	**0.015**	**0.022**	**0.020**	**0.012**	**0.014**
RAPH2013	**0.013**	**0.019**	**0.025**	**0.023**	**0.015**	**0.017**
SFCW2008	**0.006**	**0.003**	**0.008**	**0.004**	**0.003**	**0.002**
SFCW2009	**0.008**	**0.011**	**0.016**	**0.014**	**0.009**	**0.009**
SFCW2010	**0.005**	**0.004**	**0.009**	**0.007**	**0.004**	**0.003**
SFCW2011	**0.003**	0.000	**0.003**	**0.002**	**0.002**	0.000
SFCW2012	**0.006**	**0.004**	**0.007**	**0.006**	**0.003**	**0.003**
SFCW2013	**0.004**	**0.006**	**0.010**	**0.008**	**0.006**	**0.005**

Appendix 3: Glossary

Allele – A unique genetic character state. Each locus has two alleles.

Effective population size (N_E) – The number of individuals in a model population which would lose genetic variation at the same rate as an observed population. Deviations from model behavior in real populations (e.g. unequal sex ratios, some individuals reproducing more than others, etc…) tend to make N_e lower than census size (N).

F_{IS} – Correlation of alleles in an individual relative to the subpopulation in which it occurs. Commonly used as a measure of departure from random mating within a subpopulation.

F_{ST} – Correlation of alleles within the same subpopulation relative to the entire population. Commonly used as a measure of divergence between subpopulations.

Gene flow – Movement of genetic material from one population to another. Implies both physical movement and successful integration into the recipient population.

Genetic Drift – Process of genetic divergence between populations based on random sampling of alleles each generation.

Heterozygosity – Proportion of individuals in a population that are heterozygotes (i.e. do not have two identical alleles at a locus).

Hardy-Weinberg Equilibrium (HWE) – Genotype ratios expected under a random mating model.

Locus – A physical location on the DNA of an organism. The term "locus" is often used synonymously with "marker" or with any type of marker (e.g., "SNP" or "microsatellite").