Supplementary material:

Publication title: A century of genetic homogenization in Baltic salmon – evidence from archival DNA

Johan Östergren[1], Stefan Palm[1], John Gilbey[2], Göran Spong[3], Johan Dannewitz[1], Helena Königsson[3], John Persson[1], and Anti Vasemägi[1,4]

[1] Swedish University of Agricultural Sciences, Department of Aquatic Resources, Institute of Freshwater Research, Stångholmsvägen 2, SE-178 93 Drottningholm, Sweden

[2] Marine Scotland Science, Freshwater Fisheries Laboratory, Faskally, Pitlochry, PH16 5LB, Scotland, U.K.

[3] Swedish University of Agricultural Sciences, Department of Wildlife, Fish, and Environmental Studies, SE-901 83 Umeå, Sweden

[4]Chair of Aquaculture, Institute of Veterinary Medicine and Animal Sciences, Estonian University of Life Sciences, Tartu, Estonia

Keywords: *Salmo salar*, historical DNA, human-induced, genetic change, conservation

Corresponding author: Johan Östergren, fax: +46 10 478 42 62, email: johan.ostergren@slu.se

Running title: Genetic homogenization in Baltic salmon

DOI: 10.1098/rspb.2020-3147

Table S1. Historical and contemporary samples included in the study. River status is based on ICES (2018). The years of ended wild production is the time span when natural production of salmon in rivers developed by hydropower ceased to exist.

Nr	Sampled river	Sample time period	Abbreviation	River status	Catchment size (km^2)	Sampled life stage	Years of ended wild production	Sampling Years	*n*
1	Torneälven	Contemporary	To_C	Wild	40157	Parr	-	2012	54
2	Torneälven	Historical	To_H	Wild	40157	Adult	-	1928	44
3	Kalixälven	Contemporary	Ka_C	Wild	18130	Parr	-	2012	56
4	Kalixälven	Historical	Ka_H	Wild	18130	Adult	-	1928, 1929, 1930	66
5	Luleälven	Contemporary	Lu_C	Reared	25241	Parr	1961-71	2014	60
6	Luleälven	Historical	Lu_H	Wild	25241	Adult	-	1928, 1934, 1938	122
7	Byskeälven	Contemporary	By_C	Wild	3662	Parr	-	2015	60
8	Byskeälven	Historical	By_H	Wild	3662	Adult	-	1930, 1932, 1934	65
9	Rickleån	Contemporary	Ri_C	Wild	1649	Smolt	-	2014	51
10	Rickleån	Historical	Ri_H	Wild	1649	Adult	-	1961, 1962, 1963	72
11	Vindelälven	Contemporary	Vi_C	Wild	14227	Smolt	-	2015	56
12	Vindelälven	Historical	Vi_H	Wild	14227	Adult	-	1938	42
13	Öreälven	Contemporary	Or_C	Wild	3029	Parr	-	2015	59
14	Öreälven	Historical	Or_H	Wild	3029	Adult	-	1932, 1933, 1934	31
15	Ångermanälven	Contemporary	An_C	Reared	31864	Parr	1959	2014	78
16	Ångermanälven	Historical	An_H	Wild	31864	Adult	-	1926, 1927, 1928	109
17	Indalsälven	Contemporary	In_C	Reared	26727	Adult	1949-55	2013	59
18	Indalsälven	Historical	In_H	Wild	26727	Adult	-	1927, 1928, 1929	65
19	Ljungan	Contemporary	Ln_C	Wild	12851	Parr	-	2014	48
20	Ljungan	Historical	Ln_H	Wild	12851	Adult	-	1926, 1927, 1929	56
21	Ljusnan	Contemporary	Ls_C	Reared	19828	Adult	1930-45	2013	58
22	Ljusnan	Historical	Ls_H	Wild	19828	Adult	-	1926, 1927, 1929	87
23	Dalälven	Contemporary	Da_C	Reared	28954	Parr	1915, 1986	2014	57
24	Dalälven	Historical	Da_H	Wild	28954	Adult	-	1946	49
25	Mörrumsån	Contemporary	Mo_C	Wild	3369	Smolt	-	2010, 2013	91
26	Mörrumsån	Historical	Mo_H	Wild	3369	Adult	-	1920, 1928, 1929	85

Table S2. Statistics of SNPs selected from Karlsson et al. 2011. MAF=Minor Allele Frequency, HO= Observed heterozygosity level, HE= Expected heterozygosity level, HWE= p-value for HWE equilibrium deviation, Amp success= Amplification success.

SNP #	Marker reference	Chromosome	MAF	HO	HE	HWE[2]	Amp. success
1	Ss_mt_SNP1	mt	0,08	N/A	N/A	N/A	1,00
2	Ss_mt_SNP2	mt	0,17	N/A	N/A	N/A	1,00
3	St_mt_SNP7	mt	N/A	N/A	N/A	N/A	1,00
4	St_qsdy03_SNP353		N/A	N/A	N/A	N/A	0,51
5	St_qsdy05_SNP250		N/A	N/A	N/A	N/A	0,53
6	St_qsdy06_SNP113b		N/A	N/A	N/A	N/A	0,48
7	Ss_qsdy09_SNP80		N/A	N/A	N/A	N/A	0,50
8	Ss_qsdY13_SNP200		N/A	N/A	N/A	N/A	0,51
9	rs159401191	ssa27	0,12	0,18	0,21	0,00	0,98
10	rs159401228	ssa13	0,29	0,37	0,41	0,00	1,00
11	rs159401232	ssa16	0,49	0,48	0,50	0,00	0,97
12	rs159401336	ssa18	0,21	0,33	0,33	0,39	0,99
13	rs159401382	ssa18	0,18	0,28	0,29	0,18	0,99
14	rs159401422	ssa24	0,47	0,46	0,50	0,00	1,00
15	rs159401453	ssa04	0,09	0,16	0,17	0,00	0,99
16	rs159403189	ssa05	0,37	0,46	0,47	0,08	0,99
17	rs159403215	ssa06	0,24	0,35	0,37	0,04	0,99
18	rs159403335	ssa18	0,06	0,12	0,12	0,26	1,00
19	rs159403339	ssa14	0,06	0,11	0,11	0,01	0,96
20	rs159403491	ssa09	0,49	0,44	0,50	0,00	0,99
21	rs159403500	ssa04	0,40	0,45	0,48	0,02	0,98
22	rs159403514	ssa27	0,17	0,28	0,28	0,45	0,98
23	rs159401490	ssa24	0,28	0,38	0,40	0,00	0,97
24	rs159403556	ssa28	0,42	0,44	0,49	0,00	1,00
25	rs159403646	ssa26	0,09	0,15	0,17	0,00	0,99
26	rs159403753	ssa21	0,12	0,19	0,21	0,00	0,99
27	rs159403791	ssa01	0,36	0,41	0,46	0,00	0,98
28	rs159403950	ssa03	0,43	0,44	0,49	0,00	0,99
29	rs159406144	ssa19	0,28	0,39	0,40	0,19	0,99
30	rs159406168	ssa01	0,09	0,15	0,16	0,00	1,00
31	rs159407643	ssa09	0,44	0,50	0,49	0,32	0,99
32	rs159407670	ssa23	0,47	0,46	0,50	0,00	0,99
33	rs159401148	ssa16	0,40	0,48	0,48	0,72	0,99
34	rs159401155	ssa17	0,29	0,39	0,41	0,00	0,95
35	rs159401158	ssa04	0,47	0,50	0,50	0,22	0,99
36	rs159401159	ssa21	0,48	0,48	0,50	0,00	0,99
37	rs159401164	ssa13	0,49	0,43	0,50	0,00	0,98
38	rs159401165	ssa14	0,37	0,36	0,47	0,00	0,99
39	rs159401171	ssa03	0,08	0,15	0,15	0,65	1,00
40	rs159401176	ssa20	0,15	0,24	0,25	0,00	1,00
41	rs159401504	ssa22	0,37	0,42	0,46	0,00	0,99
42	rs159401181	ssa28	0,19	0,27	0,31	0,00	0,97
43	rs159401182	ssa10	0,45	0,49	0,50	0,06	1,00
44	rs159401206	ssa24	0,13	0,22	0,22	0,21	0,96
45	rs159401226	ssa18	0,12	0,18	0,21	0,00	0,92
46	rs159401252	ssa19	0,29	0,38	0,41	0,00	0,95
47	rs159401259	ssa29	0,42	0,46	0,49	0,00	0,99

48	rs159401514	ssa25	0,37	0,44	0,47	0,00	0,99
49	rs159401300	ssa08	0,25	0,36	0,37	0,05	0,99
50	rs159401541	ssa13	0,15	0,21	0,25	0,00	0,99
51	rs159401547	ssa24	0,34	0,43	0,45	0,00	0,99
52	rs159401589	ssa27	0,45	0,45	0,49	0,00	0,99
53	rs159401655	ssa20	0,29	0,40	0,41	0,52	0,97
54	rs159401695	ssa27	0,12	0,21	0,21	0,52	1,00
55	rs159401720	ssa07	0,32	0,43	0,44	0,17	0,92
56	rs159401792	ssa04	0,25	0,35	0,38	0,00	0,99
57	rs159401834	ssa22	0,45	0,49	0,50	0,37	1,00
58	rs159401859	ssa22	0,30	0,39	0,42	0,00	1,00
59	rs159401876	ssa26	0,23	0,31	0,35	0,00	0,97
60	rs159401881	ssa21	0,23	0,25	0,35	0,00	0,99
61	rs159402012	ssa24	0,14	0,21	0,24	0,00	0,95
62	rs159402033	ssa01	0,13	0,20	0,22	0,00	0,99
63	rs159402044	ssa28	0,39	0,43	0,48	0,00	0,99
64	rs159402119	ssa26	0,24	0,27	0,37	0,00	1,00
65	rs159402192	ssa14	0,21	0,31	0,33	0,01	0,99
66	rs159402311	ssa03	0,12	0,19	0,21	0,00	1,00
67	rs159402319	ssa21	0,47	0,50	0,50	0,89	0,99
68	rs159402369	ssa17	0,28	0,41	0,40	0,84	0,99
69	rs159402473	ssa13	0,43	0,49	0,49	0,93	0,99
70	rs159402518	ssa17	0,45	0,47	0,50	0,03	1,00
71	rs159402533	ssa18	0,32	0,40	0,43	0,00	0,99
72	rs159402534	ssa19	0,17	0,27	0,28	0,19	0,98
73	rs159402577	ssa15	0,29	0,41	0,41	0,41	0,99
74	rs159402594	ssa10	0,29	0,40	0,41	0,26	1,00
75	rs159402631	ssa04	0,40	0,46	0,48	0,00	0,98
76	rs159402634	ssa10	0,18	0,24	0,29	0,00	0,96
77	rs159402660	ssa25	0,44	0,46	0,49	0,00	1,00
78	rs159402683	ssa07	0,27	0,38	0,39	0,00	0,99
79	rs159402726	ssa12	0,28	0,41	0,40	0,74	0,96
80	rs159402794	ssa28	0,31	0,38	0,43	0,00	0,99
81	rs159402799	ssa13	0,04	0,07	0,08	0,00	1,00
82	rs159402800	ssa29	0,22	0,33	0,35	0,00	0,99
83	rs159402815	ssa05	0,23	0,34	0,36	0,00	1,00
84	rs159402944	ssa27	0,15	0,25	0,25	0,27	1,00
85	rs159402977	ssa25	0,11	0,19	0,20	0,00	0,96
86	rs159402986	ssa17	0,36	0,43	0,46	0,02	0,99
87	rs159403001	ssa27	0,46	0,49	0,50	0,42	0,99
88	rs159403005	ssa09	0,25	0,34	0,38	0,00	0,99
89	rs159403008	ssa09	0,36	0,45	0,46	0,60	1,00
90	rs159403009	ssa28	0,20	0,30	0,32	0,00	0,99
91	rs159403054	ssa14	0,35	0,45	0,46	0,28	1,00
92	rs159403070	ssa04	0,10	0,18	0,19	0,23	1,00
93	rs159401575	ssa02	0,32	0,41	0,44	0,01	0,86
94	rs159401722	ssa03	0,49	0,47	0,50	0,00	0,93
95	rs159401921	ssa20	0,41	0,45	0,49	0,00	0,93
96	rs159402062	ssa23	0,22	0,27	0,34	0,00	0,94
Means			0,28	0,35	0,37		0,96

[1)] Assays designed to amplify in presence of a Y-chromosome, monomorphic. [2)] with over 3000 individuals, even small deviations from expected H levels are significant (SEMs are very narrow).

Table S3. Summary of statistics for historical and contemporary samples included in the study showing sample abbreviations as in table S1, expected heterozygosity (H_e), F_{IS} values, and results from significant tests of F_{IS} values in FSTAT (F_{IS} sign.) based on 42640 randomizations (* = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$), and significance levels after correction for multiple tests ($p < 0.0006$) indicated in bold text.

Sample	Historical			Contemporary		
	H_e	F_{IS}	F_{IS} sign.	H_e	F_{IS}	F_{IS} sign.
To	0.344	-0.018	n.s.	0.342	0.002	n.s.
Ka	0.339	-0.032	*	0.346	0.018	n.s.
Lu	0.348	**-0.043**	***	0.361	-0.021	n.s.
By	0.352	-0.011	n.s.	0.353	-0.008	n.s.
Ri	0.356	0.004	n.s.	0.355	-0.016	n.s.
Vi	0.334	0.029	n.s.	0.340	-0.012	n.s.
Or	0.349	0.036	*	0.365	0.017	n.s.
An	0.360	0.007	n.s.	0.362	0.009	n.s.
In	0.342	-0.008	n.s.	0.341	-0.019	n.s.
Ln	0.359	0.016	n.s.	0.358	0.042	*
Ls	0.363	0.009	n.s.	0.371	0	n.s.
Da	0.350	-0.018	n.s.	0.356	0	n.s.
Mo	0.322	-0.012	n.s.	0.352	-0.037	**

Table S4. Pairwise F_{ST} values of historical and contemporary samples included in the study. Abbreviations as in table S1. Significance levels (* = $p < 0.05$, NS = non-significant) based on test by FSTAT between pairs of samples (6500 permutations).

	To_H	To_C	Ka_H	Ka_C	Lu_H	Lu_C	By_H	By_C	Ri_H	Ri_C	Vi_H	Vi_C	Or_H	Or_C	An_H	An_C	In_H	In_C	Ln_H	Ln_C	Ls_H	Ls_C	Da_H	Da_C	Mo_H	Mo_C
To_H		NS	NS	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
To_C	**0,003**		NS	NS	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Ka_H	**0,004**	**0,001**		NS	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Ka_C	0,011	**0,002**	**0,003**		*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Lu_H	0,036	0,028	0,025	0,034		*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Lu_C	0,030	0,027	0,023	0,031	0,014		*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
By_H	0,043	0,042	0,040	0,047	0,048	0,043		*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
By_C	0,037	0,036	0,035	0,039	0,036	0,033	0,017		*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Ri_H	0,060	0,059	0,055	0,061	0,043	0,023	0,065	0,059		*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Ri_C	0,048	0,044	0,040	0,050	0,030	0,018	0,053	0,052	0,028		*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Vi_H	0,084	0,075	0,075	0,078	0,061	0,044	0,089	0,081	0,028	0,043		NS	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Vi_C	0,079	0,072	0,072	0,078	0,054	0,035	0,079	0,072	0,024	0,035	**0,006**		*	*	*	*	*	*	*	*	*	*	*	*	*	*
Or_H	0,084	0,078	0,081	0,083	0,062	0,046	0,077	0,082	0,035	0,058	0,070	0,064		*	*	*	*	*	*	*	*	*	*	*	*	*
Or_C	0,064	0,061	0,059	0,063	0,043	0,029	0,058	0,055	0,031	0,035	0,060	0,054	0,021		*	*	*	*	*	*	*	*	*	*	*	*
An_H	0,053	0,050	0,049	0,055	0,042	0,020	0,058	0,049	0,010	0,026	0,029	0,025	0,046	0,035		*	*	*	*	*	*	*	*	*	*	*
An_C	0,066	0,062	0,065	0,068	0,043	0,027	0,057	0,053	0,018	0,035	0,043	0,035	0,051	0,042	0,016		*	*	*	*	*	*	*	*	*	*
In_H	0,067	0,065	0,064	0,070	0,056	0,038	0,071	0,069	0,013	0,044	0,039	0,038	0,050	0,053	0,016	0,020		*	*	*	*	*	*	*	*	*
In_C	0,066	0,055	0,058	0,062	0,047	0,034	0,058	0,061	0,018	0,037	0,044	0,039	0,051	0,051	0,018	0,023	0,011		*	*	*	*	*	*	*	*
Ln_H	0,066	0,065	0,068	0,074	0,049	0,026	0,062	0,056	0,013	0,031	0,048	0,037	0,045	0,041	0,019	0,023	0,026	0,026		NS	*	*	*	*	*	*
Ln_C	0,061	0,054	0,056	0,058	0,037	0,022	0,061	0,053	0,011	0,027	0,043	0,034	0,037	0,034	0,015	0,017	0,018	0,014	**0,005**		*	*	*	*	*	*
Ls_H	0,071	0,068	0,070	0,073	0,054	0,033	0,073	0,070	0,023	0,032	0,046	0,046	0,052	0,046	0,027	0,035	0,031	0,028	0,013	0,013		*	*	*	*	*
Ls_C	0,060	0,056	0,061	0,067	0,046	0,026	0,063	0,058	0,022	0,031	0,039	0,040	0,052	0,039	0,016	0,025	0,028	0,028	0,011	0,018	0,011		*	*	*	*
Da_H	0,114	0,109	0,114	0,116	0,101	0,074	0,125	0,112	0,050	0,073	0,065	0,075	0,091	0,082	0,056	0,072	0,057	0,063	0,042	0,048	0,024	0,033		*	*	*
Da_C	0,079	0,079	0,078	0,087	0,072	0,044	0,085	0,079	0,031	0,048	0,049	0,051	0,066	0,062	0,034	0,046	0,040	0,041	0,022	0,028	0,020	0,019	0,030		*	*
Mo_H	0,194	0,195	0,192	0,200	0,191	0,167	0,207	0,205	0,156	0,167	0,182	0,189	0,197	0,180	0,164	0,166	0,170	0,175	0,154	0,155	0,132	0,144	0,132	0,112		*
Mo_C	0,129	0,132	0,131	0,139	0,132	0,106	0,140	0,139	0,089	0,105	0,115	0,121	0,128	0,118	0,093	0,096	0,096	0,102	0,087	0,091	0,071	0,076	0,079	0,061	0,025	

Table S4. Global F_{ST} in historical and contemporary samples in pre-defined population groups calculated using FSTAT. Confidence intervals (95% CI) based on bootstrapping over loci is given. Despite reduction in genetic divergence in all studied groups, permutation analysis did not reveal significant temporal change in any of the studied groups (1000 permutations, two-sided $p > 0.05$).

Population group	Global F_{ST}	
	Historical (95% CI)	Contemporary (95% CI)
All (A)	0.074 (0.063–0.087)	0.059 (0.051–0.067)
Reared (B)	0.046 (0.036–0.057)	0.031(0.026 – 0.037)
Wild (C)	0.101 (0.084–0.122)	0.076 (0.065–0.088)
Wild without Mörrum (D)	0.054 (0.047–0.062)	0.049 (0.040–0.058)

Table S6. Average variance effective population size (*Ne*) and immigration rate (*m*), estimated from temporal genetic change observed within 13 Baltic salmon populations (MLNe software; Wang and Whitlock 2003). G is length (years) of assumed generation intervals. See text for details.

		Time between samples [1]			
River population	*G*	Years	Generations	*Ne* (95% CI)	*m* (95% CI)
Torneälven	7	89	13	692 (291 − 5278)	0.01 (0.00 − 0.03)
Kalixälven	7	88	13	1066 (420 − 20000)	0.00 (0.00 − 0.02)
Luleälven	6	84	14	289 (148 − 589)	0.08 (0.06 − 0.13)
Byskeälven	6	87	15	219 (132 − 372)	0.03 (0.02 − 0.05)
Rickleån	6	52	9	41 (21 − 73)	0.25 (0.13 − 0.67)
Vindeälven	6	79	13	472 (229 − 1547)	0.02 (0.01 − 0.04)
Öreälven	6	86	14	120 (70 − 203)	0.06 (0.04 − 0.09)
Ångermanälven	6	90	15	161 (83 − 279)	0.04 (0.01 − 0.09)
Indalsälven	6	85	14	252 (136 − 469)	0.04 (0.02 − 0.07)
Ljungan	6	91	15	489 (223 − 1535)	0.04 (0.03 − 0.07)
Ljusnan	6	86	14	270 (147 − 517)	0.05 (0.03 − 0.08)
Dalälven	6	71	12	88 (55 − 137)	0.06 (0.04 − 0.09)
Mörrumsån	5	88	18	301 (192 − 475)	0.02 (0.02 − 0.02)

[1] rounded to full integers

Figures.

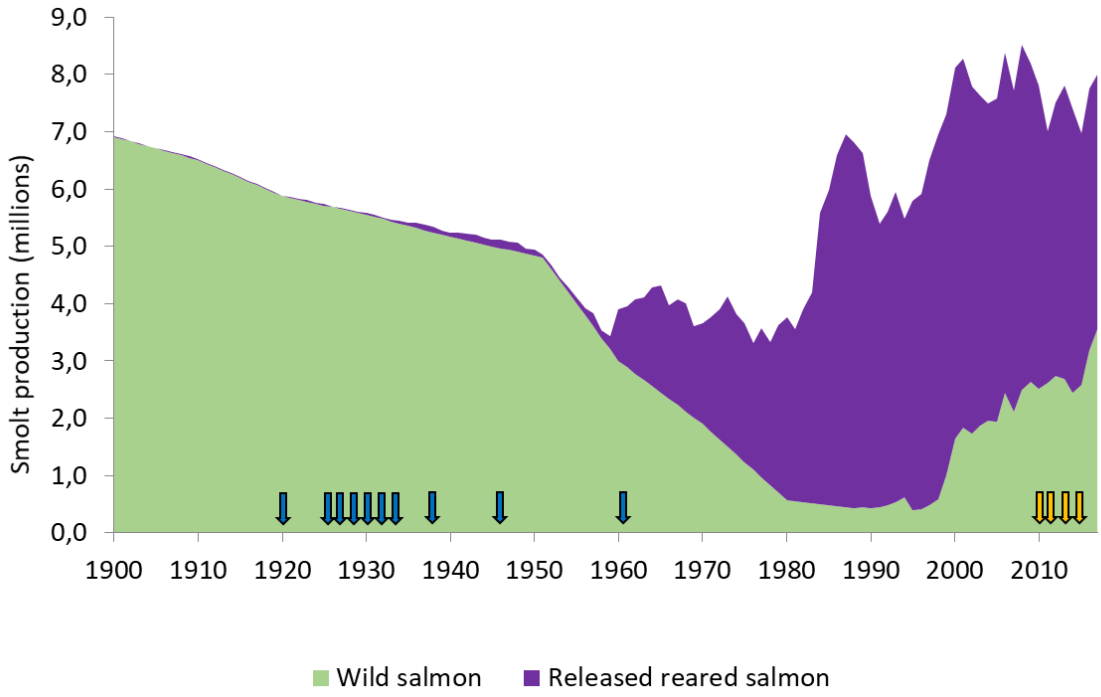

Figure S1. Wild (green) and hatchery reared (purple) Atlantic salmon smolt production expressed as smolt equivalents in the Baltic Sea between the years 1900–2017. Arrows indicate years of sampling for historical (blue) and contemporary (orange) samples in the study.

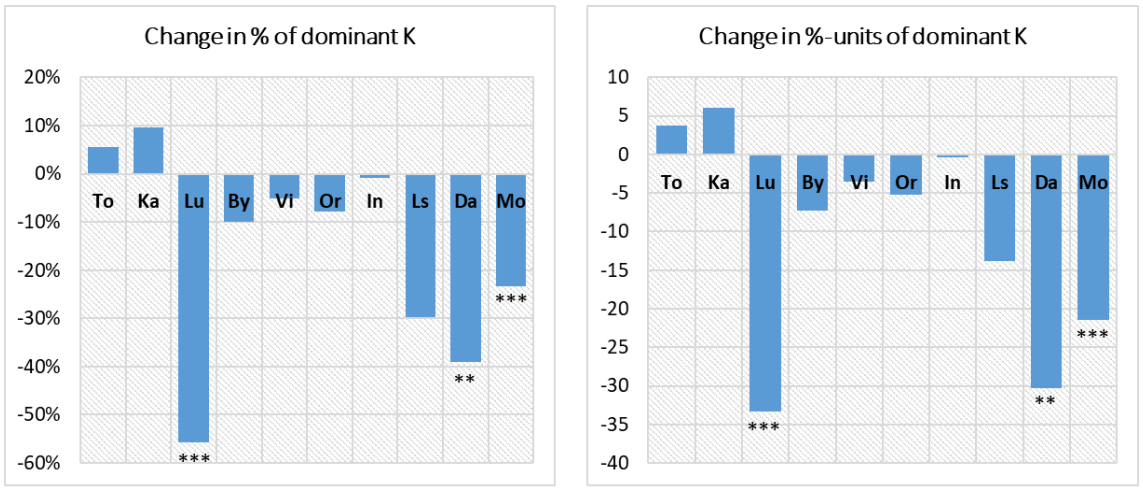

Figure S2. Direction and amount of temporal change in proportion of the dominant cluster (*K*) in 10 salmon populations. Change in percent (contemporary vs. historical, left panel) and in percent-units (contemporary – historical, right panel) is shown. Stars indicate significance level (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$). Sample abbreviations as in table S1.

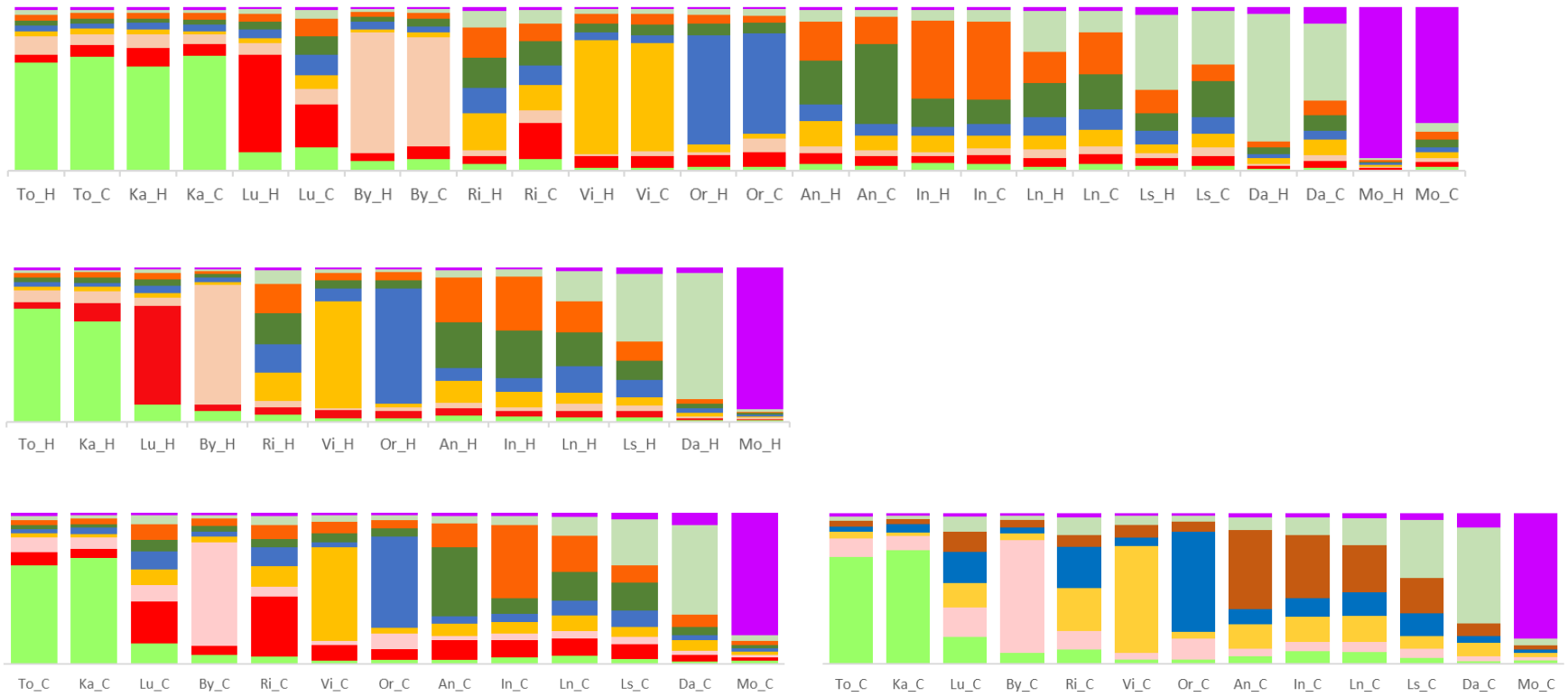

Figure S3. Results from STRUCTURE showing clusters K = 9 for all samples analyzed together (historical and contemporary; upper panel), and for historical (mid panel) and contemporary (lower panel) samples separately. The lower right panel shows K = 7 for the contemporary samples. Sample abbreviations as in Table S1. The clusters are very similar independent of analyses done, however for K = 7 (most likely K for contemporary samples) the red cluster for Lu_H no longer exist, and the dark green cluster (An_C) is joint with the orange cluster (In_C) and creates a new cluster (brown in figure).

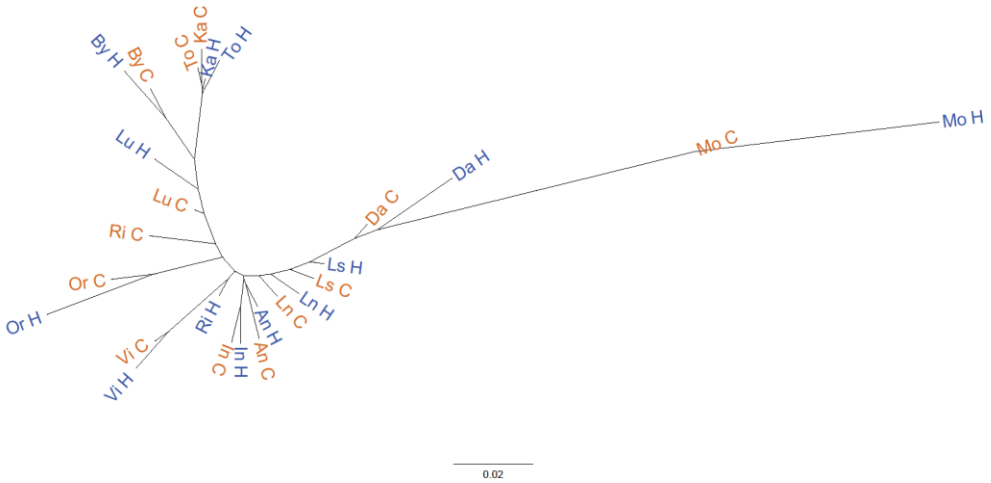

Figure S4. Dendrogram (neighbor-joining, unrooted, Cavalli-Sforza and Edwards chord distance, without bootstrap values). Historical samples (blue) and contemporary samples (orange) are indicated. Sample names as in Table S1.

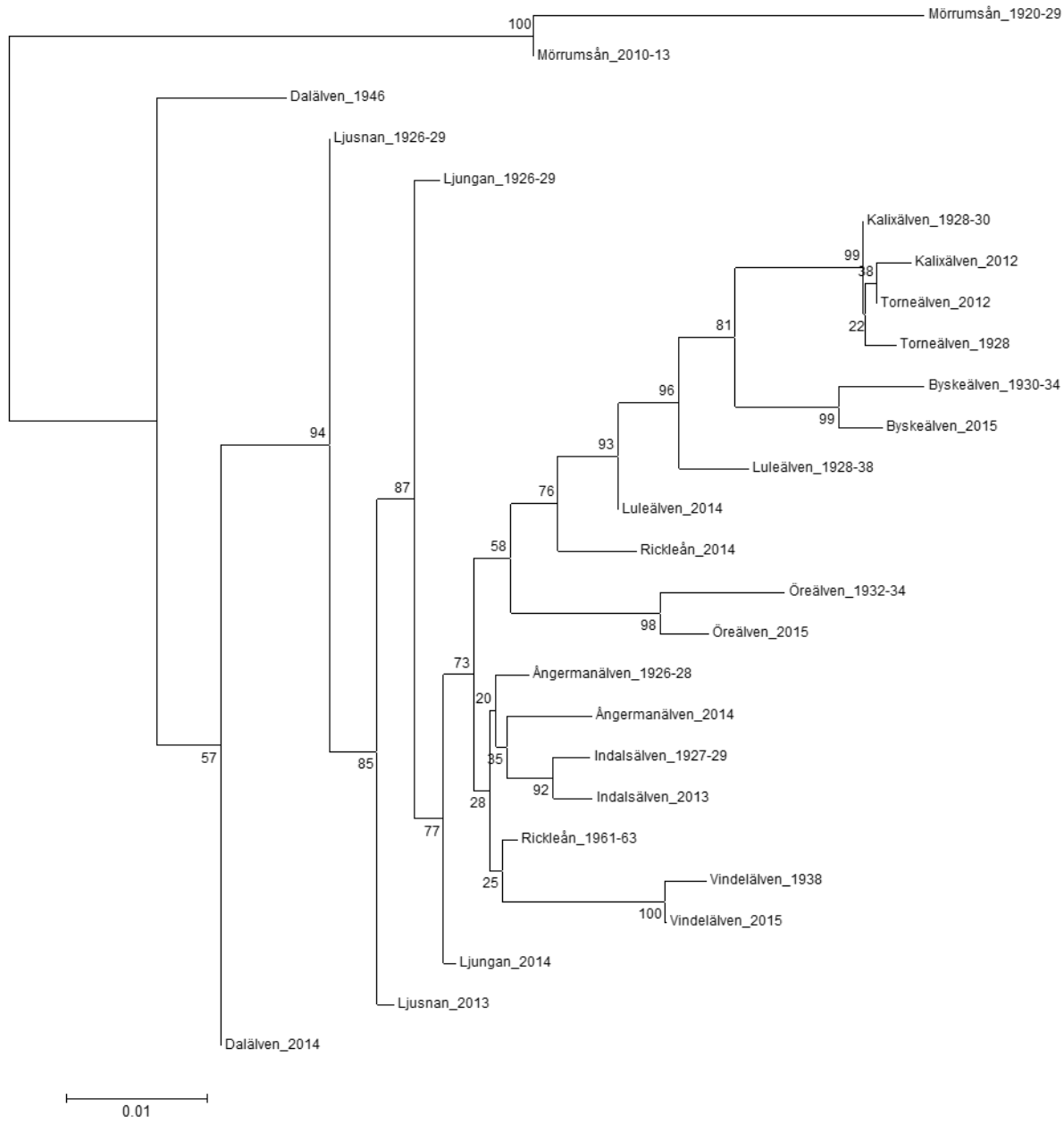

Figure S5. Dendrogram (neighbor-joining, unrooted, chord distance) with bootstrap values. Sample names as in Table S1.

Appendix A1, Supplemental material.

Testing for non-neutrality

A series of tests were carried out aimed at detecting putative "outliers" among the 82 SNPs, i.e. loci showing higher or lower genetic differentiation than expected under selective neutrality. Three different statistical approaches were employed, viz. ARLEQUIN (Excoffier *et al*. 2005), BAYESCAN (Foll & Gaggiotti 2008) and OUTFLANK (Whitlock & Lotterhos 2015). Historic and contemporary samples were always analysed separately. Since presence of hierarchical genetic structuring may yield spurious significances (Excoffier *et al*. 2009) analyses were run both including (s=13) and excluding (s=12) the genetically and geographically most deviating population (R. Mörrumsån; Figure 4). When the Mörrumsån sample was included, both a non-hierarchal and a hierarchical analysis was performed with ARLEQUIN, as this option is possible with this software.

Tests with ARLEQUIN were based on 50 000 coalescent simulations for an Island model with 100 demes per group (one or two groups for non-hierarchical and hierarchical tests, respectively) using pairwise distances (Gamma a = 0.81, Min. DAF frequency = 0.01) and with expected heterozygosity ranging from 0 to 0.5. Default settings were used when running BAYESCAN, except for varying prior parameters FDR (false discovery rate; 0.05 or 0.01) and PO (Prior Odds; 10, 50 or 100); all in all yielding six different combinations/tests per analysed data-set. The tests with OutFLANK were performed using the default settings suggested (i.e. min Heterozygosity 0.1, left and right F_{ST} trim 0.05).

In total, 34 separate tests were carried out using different combinations of data and statistical method. Results are listed in Table A1. Overall, the historic data set displayed a higher proportion of putative outlier loci than the contemporary one. In line with Excoffier *et al*. 2009, more significant outliers were also identified when the divergent Mörrumsån samples were included. Results from BAYESCAN were clearly sensitive to prior settings for parameters FDR and PO, where the latter indicates how more likely a model with selection is compared to a neutral one. As expected, reducing FDR from 0.05 to 0.01 resulted in lower numbers of significant outliers. Likewise, increasing PO from 10 to 50 and 100 resulted in higher numbers (Table A1). OutFLANK, which is suggested to have "much lower false positive rates and comparable power, as shown by simulation" (Whitlock & Lotterhos 2015), yielded only a single significant locus (*Ss_SNP40*) for the historic data set which included Mörrumsån (Table A1).

In summary, a single locus was identified as an outlier by the most conservative of the outlier tests, OutFLANK, and this locus was only identified in one of the datasets analysed in this way. Further, as there was little consistency in which loci were identified as significant outliers by all the different analyses performed (Table A2) it was concluded that the loci identified were likely to be false positives (Type I errors). If loci linked to chromosome regions under strong selection were present it would be expected that such loci would appear in most/all of the analysis. As this was not the case, we decided to retain all 82 SNPs for the downstream analysis and assume none were strongly influenced by selection.

References

Excoffier L, Laval G, Schneider S (2005) Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* 1:47-50.

Excoffier L, Hofer T, Foll M (2009) Detecting loci under selection in a hierarchically structured population. *Heredity* 103:285–298.

Foll M, Gaggiotti O (2008) A Genome-Scan Method to Identify Selected Loci Appropriate for Both Dominant and Codominant Markers: A Bayesian Perspective. *Genetics* 180:977–993.

Whitlock MC, Lotterhos KE (2015) Reliable detection of loci responsible for local adaptation: inference of a null model through trimming the distribution of F_{ST}. *The American Naturalist* 186: S24-36.

Table A1:1. Results from 34 outlier tests based on different combinations of data sets, method/software and settings (see text for details). Loci listed when using Arlequin have P<= 1%, OutFLANK q-threshold is 0.05.

			Putative outlier loci identified	
Data set	**Test No.**	**Program (settings)**	**Balancing selection**	**Directional selection**
Historic (s=13)	1	Arlequin (non-hierarchical)	*Ss_SNP154*	*Ss_SNP146, Ss_SNP40*
	2	Arlequin (hierarchical, s=12+1)	-	-
	3	BayeScan (Prior Odds = 10, FDR=0.05)	*Ss_SNP154, Ss_SNP69*	*Ss_SNP145, Ss_SNP146,*
	4	BayeScan (Prior Odds = 50, FDR=0.05)	*Ss_SNP154*	*Ss_SNP146*
	5	BayeScan (Prior Odds = 100, FDR=0.05)	-	-
	6	BayeScan (Prior Odds = 10, FDR=0.01)	*Ss_SNP154*	*Ss_SNP145, Ss_SNP146*
	7	BayeScan (Prior Odds = 50, FDR=0.01)	-	-
	8	BayeScan (Prior Odds = 100, FDR=0.01)	-	-
	9	OutFLANK (Min Heterozygosity 0.1, left and right F_{ST} trim	-	*Ss_SNP40*
Historic (s=12)	10	Arlequin (non-hierarchical)	-	*Ss_SNP145, Ss_SNP146,*
	11	BayeScan (Prior Odds = 10, FDR=0.05)	*Ss_SNP154*	*Ss_SNP145, Ss_SNP73*
	12	BayeScan (Prior Odds = 50, FDR=0.05)	-	*Ss_SNP145*
	13	BayeScan (Prior Odds = 100, FDR=0.05)	-	-
	14	BayeScan (Prior Odds = 10, FDR=0.01)	-	*Ss_SNP145*
	15	BayeScan (Prior Odds = 50, FDR=0.01)	-	-
	16	BayeScan (Prior Odds = 100, FDR=0.01)	-	-
	17	OutFLANK (Min Heterozygosity 0.1, left and right F_{ST} trim	-	-
Contemporary (s=13)	18	Arlequin (non-hierarchical)	-	*Ss_SNP146, Ss_SNP39,*
	19	Arlequin (hierarchical, s=12+1)	-	*Ss_SNP40*
	20	BayeScan (Prior Odds = 10, FDR=0.05)	-	-
	21	BayeScan (Prior Odds = 50, FDR=0.05)	-	-
	22	BayeScan (Prior Odds = 100, FDR=0.05)	-	-
	23	BayeScan (Prior Odds = 10, FDR=0.01)	-	-
	24	BayeScan (Prior Odds = 50, FDR=0.01)	-	-
	25	BayeScan (Prior Odds = 100, FDR=0.01)	-	-
	26	OutFLANK (Min Heterozygosity 0.1, left and right F_{ST} trim	-	-
Contemporary (s=12)	27	Arlequin (non-hierarchical)	-	*Ss_SNP73, Ss_SNP80*
	28	BayeScan (Prior Odds = 10, FDR=0.05)	-	-
	29	BayeScan (Prior Odds = 50, FDR=0.05)	-	-
	30	BayeScan (Prior Odds = 100, FDR=0.05)	-	-
	31	BayeScan (Prior Odds = 10, FDR=0.01)	-	-

	32	BayeScan (Prior Odds = 50, FDR=0.01)	-	-
	33	BayeScan (Prior Odds = 100, FDR=0.01)	-	-
	34	OutFLANK (Min Heterozygosity 0.1, left and right F_{ST} trim	-	-

Table A1:2. Eight SNPs (out of 82) identified as outliers in various analyses performed (see Table Sx for test number details).

Locus	Test number where SNP identified as significant outlier	
	Balancing selection	Directional selection
Ss_SNP145		3, 6, 10, 11, 12, 14
Ss_SNP146		1, 3, 4, 6, 10, 18
Ss_SNP154	1, 3, 4, 6, 11	
Ss_SNP39		18
Ss_SNP40		1, 9, 18, 19
Ss_SNP69	3	
Ss_SNP73		3, 10, 11, 27
Ss_SNP80		27

Appendix A2, Supplemental material.

Statistical power

Power to detect genetic heterogeneity for the present 82 SNP markers was evaluated using POWSIM 4.1 (Ryman and Palm, 2006). POWSIM simulates sampling of genes from a specified number of populations which have diverged due to random drift to a predefined expected level of genetic differentiation quantified as F_{ST}. Samples from the simulated populations are used for testing for genetic homogeneity at each locus separately using traditional χ^2-test and Fisher's exact test, and across loci by means of sums of χ^2 and "Fisher's method", respectively. The proportion of significance results (here $P<0.05$) obtained after having repeated the above simulation procedure a large number of times (here 1000) yields an estimate of power (or α-error, when $F_{ST}=0$).

As initial allele frequencies for the simulations, we used the average frequencies in the total material (all 26 historic and contemporary samples combined). Two basic scenarios were evaluated, with two (s=2) and 13 (s=13) samples being compared, mimicking the present temporal (within rivers) and spatial (between rivers) comparisons. Sample sizes were always set to n=50 to approximate the 26 empirical samples (range of 31-122, mean 65, median 59; Table S1).

Estimates of power to detect genetic heterogeneity for the present 82 SNP markers and average allele frequencies are presented in Figure A2:1. Regardless of number of samples and statistical method, power was 100 % at F_{ST}=0.015 (and above). At lower levels of divergence, as expected, power was always higher when comparing 13 than two samples. Further, as shown by Ryman *et al*. (2006) for di-allelic markers, summation of χ^2 yielded higher power than Fisher's method, especially when making pairwise sample comparisons.

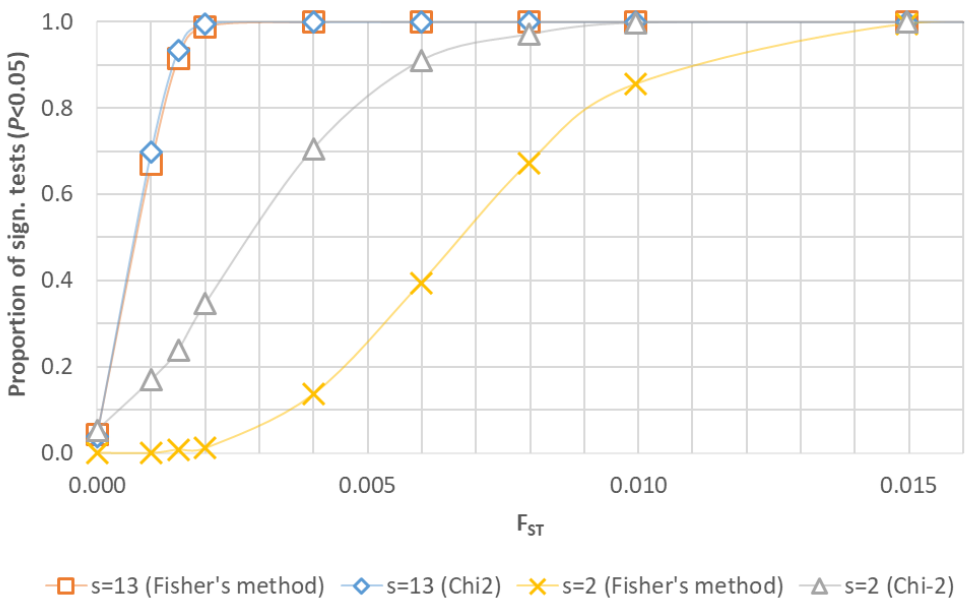

Figure A2:1. Simulated estimates (average of 1000 simulations) of power and α-error for the present 82 SNP loci and average allele frequencies at various levels of genetic divergence (F_{ST}). Simulations were carried out using 13 and 2 population split models (s=13 and s=2). See text for details.

References

Ryman N, Palm S (2006). POWSIM: a computer program for assessing statistical power when testing for genetic differentiation. *Molecular Ecology Notes* 6: 600-602.

Ryman N, Palm S, André C, Carvalho GR, Dahlgren TG, Jorde PE, Laikre L, Larsson LC, Palmé A, Ruzzante, DE (2006). Power for detecting genetic divergence: differences between statistical methods and marker loci. *Molecular Ecology* 15:2031-2045.