Deliverable Report

Deliverable No:	4.2	**Delivery Month:**	16
Deliverable Title	Population genetic analysis of wild and comparison with domesticated pikeperch populations to be applied in future breeding programs of the species		
WP No:	4	**WP Lead beneficiary:**	P1. HCMR
WP Title:	Reproduction and Genetics – pikeperch		
Task No:	4.2	**Task Lead beneficiary:**	P1. HCMR
Task Title:	Evaluation of the genetic variation in non-domesticated broodstocks of pikeperch		
Other beneficiaries:	P9. UL		
Status:	Delivered	**Expected month:**	16
…….			

Lead Scientist preparing the Deliverable: Tsigenopoulos, C. (P1. HCMR)
Other Scientists participating: Fontaine, P. (P9. UL), Tsaparis, D. (P1. HCMR)

Objective: The objective of this Deliverable was to assess the genetic variability of wild broodstocks in Europe and compare this variability with that of domesticated pikeperch populations (D4.1 Genetic analysis of domesticated pikeperch broodstocks in Mo12) to be applied in future breeding programs of the species. This information will enable us to define how a future genetic breeding program should be established for sustainable optimal performance of pikeperch through domestication.

Description: After the last ice age (~18.000 years ago) pikeperch or zander (*Sander lucioperca*, Linnaeus, 1758) spread from the Caspian-Black Sea region reaching the Baltic Sea during the Lake Ancylus period approx. 6.000 years ago. In Germany, for example, the Elbe and Danube drainages represented the north-western-most European dispersal limits, until the late 19th century when pikeperch was introduced by humans into rivers beyond these natural frontiers. As a commercially valuable species, pikeperch is still regularly stocked in both its native and non-native ranges. Canals connecting ancestral and novel drainages have further facilitated the dispersal and provided multiple opportunities for secondary contacts of pikeperch populations. Moreover, pikeperch is euryoecious, that is, it is able to cope with a broad range of environmental conditions including low salinity brackish waters, which increases its potential for dispersal and invasion (see references in Eschbach et al., 2014)

Deliverable 4.1 Genetic analysis of domesticated pikeperch broodstocks, provided a first assessment of the genetic diversity of captive pikeperch stocks and because there are only a few (around 10) commercial hatcheries that produce pikeperch in Europe, the genetic diversity was expected to be relatively lower compared to the genetic variability of natural populations (Saisa et al., 2010). In principle, each pikeperch farm uses its own stock, captured either from the wild or supplied by another farmer. Therefore, pikeperch populations differ from one farm to another depending upon the geographical origin of the captured wild populations, which were used as the starting base of the captive stocks.

The primary objective of the present Deliverable was to use the microsatellite multiplex tools developed previously for the species in D4.1 and evaluate the genetic variability of some wild pikeperch populations. Population genetics parameters from wild stocks are compared to those of captive broodstock in commercial RAS farms around Europe (Task 4.2, Deliverable 4.1 Genetic analysis of domesticated pikeperch

broodstocks) with the objective to define how a future genetic breeding program should be established for sustainable optimal performance through domestication of pikeperch.

Material and Methods

Biological material

DNA extractions were completed for all wild and domesticated samples/populations that were obtained using standard protocols (salt precipitation, Miller et al., 1988). In addition to the thirteen cultured populations analyzed in D4.1, eight more populations were genotyped of which one was domesticated (Sweden). Therefore, current results refer to a total of 21 populations and more than 950 fish (**Table 4.2.1**). The Qiagen multiplex PCR kit was used for PCR with the two multiplexes (7-plex and 4-plex).

Table 4.2.1 List of the 21 domesticated and wild pikeperch populations, and number of fish per sample that were genetically analyzed; populations marked in blue were of wild origin.

	Population	Sample Size
1	**Gyori Elore, HTSZ, Hungary**	53
2	Szabolsi, Halaszati Kft, Hungary	49
3	Aquapri A/S Denmark, VanMecklen, Netherlands	54
4	Aquapri A/S Denmark, Czech Rep.	38
5	Aquapri A/S Denmark, Excellence fish, Netherlands	14
6	Aquapri A/S Denmark, Hungary	73
7	Aquapri A/S Denmark, Mosso, Denmark	19
8	IfB Potsdam, Germany	46
9	FGFRI Kainuu fisheries research station, Finland	31
10	FGFRI Laukaa Fish Farm, Finland	20
11	ASIALOR, France	63
12	INAGRO Belgium, German origin	100
13	INAGRO Belgium, Dutch origin	100
14	**Tunisia**	59
15	Svensk Fiskodling AB, Hjalmaren Lake, Sweden	30
16	**Dom de Lindre, France**	51
17	**URAFPA-DAC, Czech Rep.**	70
18	**Sarag L., Poland**	14
19	**Wymoj L., Poland**	11
20	**Oulujarvi L., Finland**	32
21	**Hiidenvesi L., Finland**	31

Microsatellite Loci

Two multiplex PCRs designed with 22 microsatellite loci developed for species phylogenetically close to pikeperch were optimized (see Deliverable 4.1 for loci, primers and PCR conditions). Raw allele sizes were scored using the STR and software (v. 2.4.59 http://www.vgl.ucdavis.edu/STRand). The number of alleles per locus, observed (H_O) and expected heterozygosity (H_E) and linkage disequilibrium (LD) were calculated in GENETIX v. 4.05 (Belkhir et al., 2004), FSTAT 2.9.3 (Goudet 1995) and GenAlEx 6.5 (Peakall and Smouse 2006, 2012) which offers a wide range of population genetic analysis options for the full spectrum of genetic markers within the Microsoft Excel environment on both PC and Macintosh computers.

Deviations from Hardy-Weinberg equilibrium (HWE) across all samples were characterized by F_{IS}. In instances where the observed genotype frequencies deviated significantly from HWE, the Micro-Checker v.2.2.3 program (Van Oosterhout et al., 2004) was used to test for null alleles. The differentiation among locations was also quantified by F_{ST} (using the estimator θ of Weir & Cockerham, 1984).

However, only basic characteristics for loci (PflaL3) were reported and analyses estimating F_{IS} and F_{ST} values were excluded, because the microsatellite loci (PflaL3) showed signs of "null alleles" with significant probability ($P > 0.05$) of "large allele dropout" or "stuttering".

STRUCTURE 2.3.2 (Falush et al., 2003) was used to infer the most likely population structure based on microsatellite data of the 21 pikeperch populations. The calculation was done with an non-admixture model without a priori population information, using a burn-in period of 250,000 and 1,000,000 subsequent MCMC repeats for each k value between one and ten. The most likely number of groups was identified using the Dk criterion (Evanno et al., 2005) and detecting the number of clusters of individuals using the software STRUCTURE. Admixture of populations was calculated such that all individuals were assigned to each of the identified ancestral gene pools. Afterwards, their respective proportions of membership were computed.

Results and Discussion

Basic population genetics parameters & Cross-species microsatellite transferability

Considering a long term breeding program, it is fundamental to ensure sufficient genetic variation within populations, as this determines the potential for selection of desired traits or of adaptation to hostile changes in environmental /rearing conditions. In domesticated stocks, caution is required because the loss of genetic variability within the first generations of breeding practices limits the potential for future genetic improvement from selection practices.

Basic population genetics parameters (allelic richness, heterozygosity indices, inbreeding coefficients) were calculated for both wild and domesticated stocks. The total number of alleles per locus ranged from 8-9 (PflaL3 and PflaL9, respectively) to 23 (Svi4) (**Table 4.2.2**). Therefore, microsatellite loci showed relatively high levels of polymorphism even though some samples were monomorphic (exhibited only one allele) for some loci such as for Za199 and PflaL9 in the "Excellence fish" of Aquapri A/S (population 5), locus Za237 in Kainuu Fisheries Research Station (population 9), Za144 in Laukaa Fish Farm (population 10), Za024 in the Tunisia (population 14) and Pfla3 in Sweden (population 15). The Tunisian population (population 14), was the only wild population that appeared to have one allele at a locus (Za024). Moreover, the microsatellite loci used were developed in other phylogenetically close species and the polymorphism appeared to be related to the species from which the microsatellite originated. The most polymorphic were the loci from *Stizostedion vitreum* (Zvi, 2 loci average 21 alleles), followed by those from *Zingel asper* (Za, 7 loci average 14.85 alleles) and last from *Perca flavescens* (Pfla, 2 loci average 8.5 alleles). This was consistent with recent trends in taxonomy that proposed *Stizostedion* to be grouped into *Sander* genus and that *Zingel* species are the closest to this *Stizostedion-Sander* group, before the perches of the genera *Perca* (*P. fluviatilis*, *P. schrenkii* etc) and *Gymnocephalus* (GenBank taxonomy sequence sources). [*Sander lucioperca* was first described as *Stizostedion lucioperca* by Linnaeus, 1758]

For the 21 populations analyzed, the least number of alleles was encountered in Aquapri's VanMecklen (2.6), Aquapri's "Excellence fish" and Laukaa Fish Farm (2.8), Aquapri's "Mosso fish" (3.1), and the greatest in Hungarian Aquapri's (8.2), Halaszati Kft (7.8) and Inagro's German (7.2) stocks, which were greater than that in all wild stocks (3.7 to 6.2). Likewise, expected heterozygosity (H_E) ranged from 0.3408 in Aquapri's Excellence fish to 0.7194 in Aquapri's Hungarian fish; the later population showed higher values than those in the wild populations (**Table 4.2.3**).

Table 4.2.2 Number of alleles per locus; populations numbers follow those in Table 4.2.1.

Population					Locus						
	PflaL3	Svi18	Za199	Za138	PflaL9	Svi4	Za024	Za038	Za144	Za207	Za237
1	4	6	5	9	3	7	7	6	8	4	7
2	5	12	5	14	4	6	6	8	8	6	9
3	3	3	2	3	2	4	2	3	3	2	2
4	3	4	3	3	3	5	3	4	3	3	2
5	2	4	1	2	1	5	2	4	4	3	2
6	5	13	6	13	3	8	7	9	8	6	9
7	2	4	3	2	4	5	2	3	3	3	2
8	2	4	4	7	5	6	6	4	9	7	5
9	3	4	4	4	3	5	4	4	4	4	1
10	2	3	3	2	5	4	3	3	1	2	2
11	5	7	7	6	4	5	4	4	8	6	3
12	4	9	10	9	6	7	5	5	9	6	6
13	5	5	7	7	5	5	3	3	4	4	4
14	3	5	6	4	2	3	1	3	6	4	3
15	1	3	7	4	4	7	4	6	3	3	3
16	5	6	5	5	4	6	3	4	6	4	3
17	3	5	4	6	3	5	3	4	3	3	2
18	5	6	7	4	4	4	4	5	5	4	3
19	3	4	7	4	4	3	4	4	4	5	3
20	4	5	8	5	4	6	5	4	3	4	4
21	5	5	5	6	5	7	5	4	4	3	3
Total No. Alleles	**9**	**19**	**16**	**20**	**8**	**23**	**14**	**13**	**17**	**11**	**13**

Table 4.2.3 Basic population genetics parameters for all populations analyzed: mean number of alleles per locus, observed (H_O) and expected heterozygosity (H_E), and F_{IS} calculated in GENETIX v. 4.05 (Belkhir et al., 2004). Asterisks indicate significance at p=0.05**.**

	Population	Sample Size	Mean Nb of alleles	H_E	H_O	F_{IS}
1	**Gyori Elore, HTSZ, Hungary**	53	6.2	0.6826	0.7472	-0.08424
2	Szabolsi, Halaszati Kft, Hungary	49	**7.8**	0.7182	0.6759	0.06962*
3	Aquapri A/S Denmark, VanMecklen, Netherlands	54	2.6	0.4675	0.6796	-0.44607
4	Aquapri A/S Denmark, Czech Rep.	38	3.3	0.4616	0.4882	-0.04401*
5	Aquapri A/S Denmark, Excellence fish, Netherlands	14	2.8	0.3408	0.4100	-0.16229
6	Aquapri A/S Denmark, Hungary	73	**8.2**	0.7194	0.7165	0.01110*
7	Aquapri A/S Denmark, Mosso, Denmark	19	3.1	0.4169	0.3985	0.07185*
8	IfB Potsdam, Germany	46	5.7	0.5567	0.5502	0.02343*
9	FGFRI Kainuu fisheries research station, Finland	31	3.7	0.5257	0.5819	-0.09055
10	FGFRI Laukaa Fish Farm, Finland	20	2.8	0.4743	0.6032	-0.24757
11	ASIALOR, France	63	5.4	0.5940	0.5913	0.01261
12	INAGRO Belgium, German origin	100	**7.2**	0.7224	0.8099	-0.11621*
13	INAGRO Belgium, Dutch origin	100	4.7	0.6156	0.6465	-0.04510
14	**Tunisia**	59	3.7	0.4013	0.3585	0.11512*
15	Svensk Fiskodling AB, Hjalmaren Lake, Sweden	30	4.4	0.5250	0.5817	-0.08989
16	**Dom de Lindre, France**	51	4.6	0.5923	0.6706	-0.12237
17	**URAFPA-DAC, Czech Rep.**	70	3.8	0.4692	0.4382	0.07357*
18	**Sarag L., Poland**	14	4.6	0.5763	0.5643	0.05780*
19	**Wymoj L., Poland**	11	4.2	0.6149	0.6764	-0.05217*
20	**Oulujarvi L., Finland**	32	4.8	0.5946	0.5995	0.00787*
21	**Hiidenvesi L., Finland**	31	4.7	0.6034	0.5340	0.13148*

On average, domesticated populations exhibited a slightly higher number of alleles (2.634 versus 2.58, not significantly different with an F-test) (**Fig. 4.2.1**) and amongst the domesticated samples there were populations that were more polymorphic than any wild population [population 2 (from Szabolsi, Halaszati Kft, Hungary), population 6 (Aquapri's Hungarian) and population 12 (INAGRO's from Germany)]. Likewise, unbiased Expected Heterozygosity Estimates were slightly higher in wild population (0.573 versus 0.553, but again not significantly different with an F-test) (**Fig. 4.2.2**); values for wild populations were lower than 0.69 (in the first Hungarian one), whereas the three above mentioned domesticated stocks (populations 2, 6 and 12) showed values above 0.70.

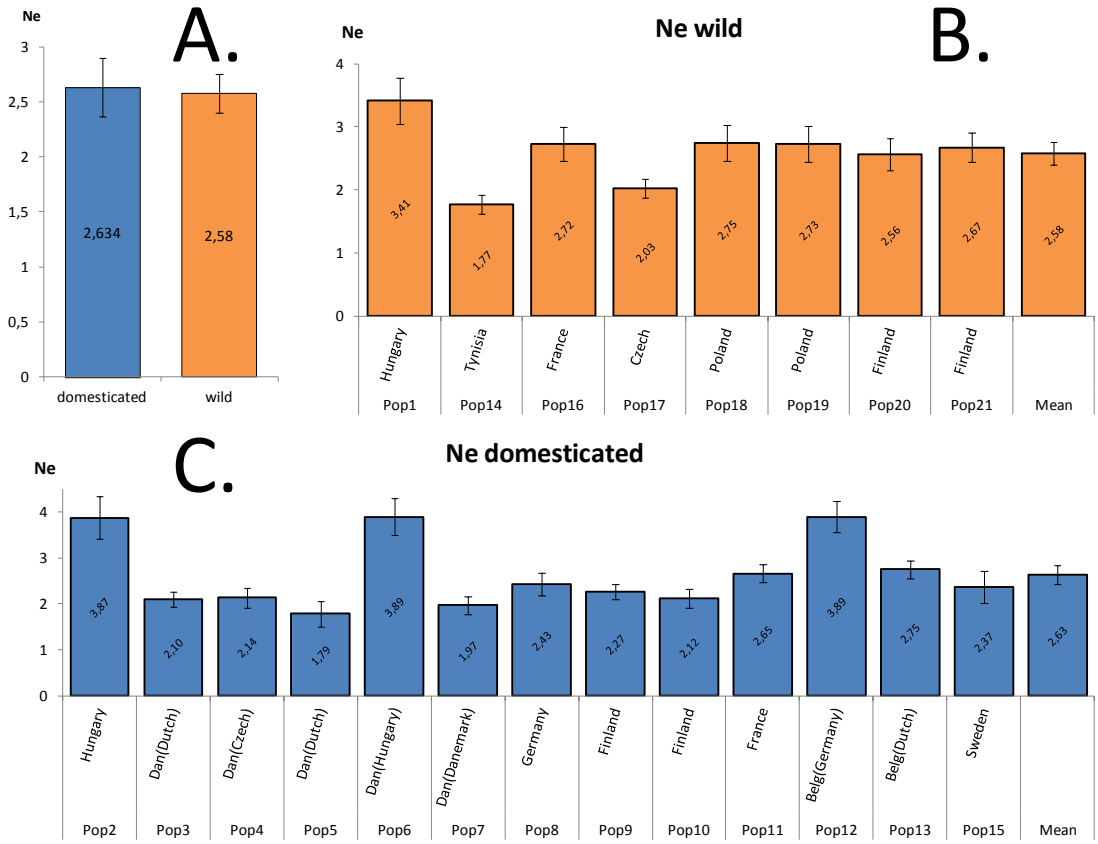

Figure 4.2.1 Mean number of alleles for domesticated and wild populations (A) and for each one separately (B & C).

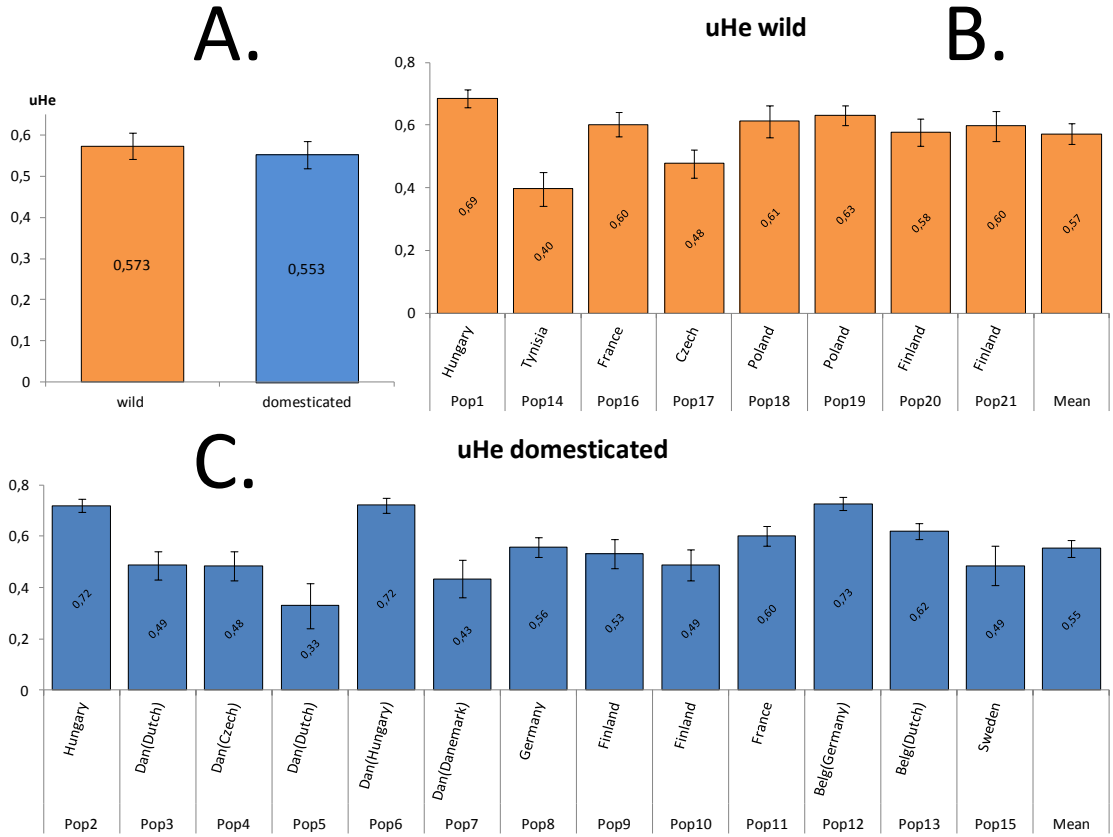

Figure 4.2.2 Estimates of Unbiased Expected Heterozygosity (uHe) for domesticated and wild populations (A) and for each one separately (B & C).

A wide range of F_{IS} values were observed in the 21 populations analyzed (**Tables 4.2.3** and **4.2.4**). In principle, **positive** F_{IS} values indicate that individuals in a population were **more related** than expected under a model of random mating, whereas **negative** F_{IS} values indicate that individuals in a population are **less related** than expected under a model of random mating. The F_{IS} values were high and significant for Hiidenvesi L. in Finland (0.131), in Tunisia (0.115), Halaszati Kft (0.069), Aquapri's Mosso (0.071) and almost all wild samples. Such deviations from Hardy Weinberg equilibrium (HWE) may be due to i) the Wahlund effect, *i.e.*, the reduction in the overall heterozygosity of a population as a result of subpopulation structures (that means if two or more subpopulations have independent allele frequencies, then the overall heterozygosity is reduced, irrespective of whether those subpopulations are in Hardy-Weinberg equilibrium), ii) non-panmixia (inbreeding, groupings of relatives, selection against heterozygotes) or iii) to genotyping errors (null alleles and other scoring errors).

Table 4.2.4 F_{IS} values per locus for the 21 pikeperch populations genotyped for 10 loci (locus PflaL3 is excluded due to null alleles); populations numbers follow those in Table 4.2.1.

Population	Svi18	Za199	Za138	PflaL9	Svi4	Za024	Za038	Za144	Za207	Za237	All
1	-0.409	0.064	-0.032	-0.242	0.003	0.022	0.024	-0.116	-0.161	0.010	-0.084
2	0.199	-0.202	0.330	0.038	-0.061	0.038	-0.057	0.069	0.017	0.230	0.070
3	-0.105	-0.377	-0.615	-0.293	-0.341	0.057	-0.603	-0.640	-1.000	0.000	-0.446
4	0.078	0.641	-0.091	0.395	-0.110	0.193	-0.341	0.037	-0.289	-0.158	-0.044
5	-0.305	NA	0.000	NA	-0.368	1.000	-0.051	-0.300	-0.189	-0.048	-0.162
6	-0.045	-0.013	-0.015	0.021	0.133	0.049	-0.075	0.007	0.088	-0.021	0.011
7	-0.003	-0.094	0.000	0.344	-0.485	0.000	0.402	0.465	-0.138	-0.029	0.072
8	-0.302	-0.001	0.271	0.182	0.150	0.100	-0.005	-0.114	-0.056	-0.161	0.023
9	-0.091	-0.192	-0.211	0.150	-0.122	0.117	-0.291	0.094	-0.230	NA	-0.091
10	-0.137	-0.238	-0.440	-0.080	-0.526	-0.109	-0.302	NA	-0.188	-0.166	-0.248
11	-0.389	0.013	0.303	0.206	0.644	-0.309	0.138	-0.103	0.022	-0.007	0.013
12	0.140	-0.199	-0.060	-0.175	-0.192	0.006	0.109	-0.043	-0.336	-0.471	-0.116
13	-0.217	0.024	-0.177	0.125	-0.004	0.121	-0.109	-0.008	-0.019	-0.186	-0.045
14	-0.169	0.121	0.022	0.269	0.064	NA	-0.056	0.273	0.316	0.172	0.115
15	-0.357	-0.108	-0.102	0.354	-0.306	-0.093	-0.111	-0.211	0.254	0.156	-0.090
16	-0.093	-0.220	-0.042	0.072	-0.130	-0.234	-0.172	-0.196	-0.085	-0.091	-0.122
17	-0.139	0.211	0.186	0.673	0.126	0.096	0.024	0.078	-0.095	0.056	0.074
18	0.227	0.167	-0.111	0.343	-0.324	0.150	-0.167	-0.087	0.122	0.328	0.058
19	-0.119	-0.087	0.209	0.259	-0.500	0.104	-0.240	-0.224	0.080	-0.022	-0.052
20	-0.060	-0.130	0.056	0.243	0.028	-0.133	0.055	0.222	-0.019	-0.221	0.008
21	0.077	0.251	0.005	0.235	0.057	0.221	0.105	0.024	0.128	0.223	0.131

Inbreeding seems an explanation in domesticated and non-random mating is also likely in our case, as deficits were homogeneous among loci (all significant and all non-significant F_{IS} values). Selection against heterozygotes cannot be demonstrated from our results; although microsatellite loci are typically recognized as neutral genetic markers, it is possible that one or more loci are linked to genes or gene groups under selection. The Walhund effect could also explain the deficit of heterozygotes due to the mixing of genetically variable populations to form a new domesticated stock, which might be the case in some aquaculture companies' practices.

Finally, F_{ST} values are frequently used as a summary of genetic differentiation among groups. It depends on the allele frequencies at a locus, showing specific properties linked to genetic diversity. Population differentiation was estimated across samples using the F_{ST} estimate by Weir & Cockerham's (1984) (**Table 4.2.5**). The smallest F_{ST} estimate values were between the two wild Finnish samples (0.021) and the Finnoscandinavian (wild and domesticated) samples in general (F_{ST} <0.18). Next, a close relationship was observed (F_{ST} <0.11) between the two Hungarian populations with the Aquapri's Hungarian one (population 6) (**Table 4.2.5**). Also a close relationship was observed between the two Czech populations (F_{ST} =0.03 between Aquapri's population and the wild one) and that the two German ones (F_{ST} =0.16 between IfB Potsdam and Inagro's). Lastly, a close relationship was observed between the wild sample of Domaine de Lindre and INAGRO's German samples (F_{ST} =0.11) and Aquapri's Mosso sample with the wild from Wymoj L. in Poland (F_{ST} =0.087).

Table 4.2.5 Pairwise Population Theta (θ) calculations [Weir & Cockerham's (1984) F_{ST}'s] between the twenty-one populations; populations numbers follow those in Tables 4.2.1.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
1		0.09290	0.27379	0.25377	0.30706	0.11096	0.36453	0.23820	0.35349	0.37247	0.21651	0.17514	0.29317	0.32680	0.33166	0.17908	0.25729	0.25825	0.27075	0.31423	0.31237
2			0.26190	0.25382	0.28725	0.01107	0.33350	0.22514	0.32040	0.34246	0.21432	0.13854	0.25806	0.31393	0.29288	0.18892	0.24719	0.23192	0.22608	0.28015	0.28441
3				0.12311	0.27969	0.26942	0.30152	0.15416	0.36933	0.44913	0.14936	0.19025	0.29213	0.25978	0.27234	0.10994	0.13251	0.20304	0.25359	0.32925	0.31654
4					0.20315	0.26953	0.34269	0.10511	0.39394	0.47000	0.18356	0.16520	0.30121	0.22369	0.30959	0.10657	0.03661	0.25044	0.28528	0.34892	0.35242
5						0.30601	0.41686	0.21680	0.43994	0.50969	0.25966	0.22664	0.30216	0.35744	0.34764	0.25005	0.17858	0.29582	0.31873	0.38035	0.39222
6							0.33530	0.23424	0.31672	0.33659	0.21339	0.14520	0.26399	0.32174	0.29141	0.20289	0.26336	0.24767	0.23668	0.27974	0.28503
7								0.25244	0.23673	0.36220	0.23900	0.20825	0.15052	0.28816	0.16212	0.24509	0.37533	0.15072	0.08743	0.18048	0.18721
8									0.32053	0.38605	0.16109	0.16478	0.25167	0.20189	0.25419	0.12060	0.14054	0.19165	0.19687	0.29028	0.29146
9										0.20971	0.29800	0.20454	0.21880	0.38667	0.18268	0.30641	0.42287	0.22947	0.19845	0.05904	0.07506
10											0.34287	0.25771	0.23901	0.47150	0.30535	0.36628	0.47521	0.33185	0.25694	0.14888	0.14894
11												0.13184	0.24508	0.22387	0.22857	0.04616	0.18916	0.19943	0.18863	0.25896	0.25345
12													0.19147	0.22641	0.18097	0.11796	0.17591	0.16062	0.14126	0.16699	0.16678
13														0.24927	0.18392	0.25241	0.33406	0.16212	0.09797	0.15177	0.18713
14															0.33817	0.21290	0.29536	0.29825	0.27279	0.35719	0.36237
15																0.23454	0.34206	0.15058	0.16687	0.14029	0.10804
16																	0.12742	0.18707	0.19218	0.26504	0.26355
17																		0.27227	0.29958	0.37687	0.38079
18																			0.03594	0.16009	0.14526
19																				0.11963	0.12153
20																					0.02123
21																					

All results mentioned above based on F_{ST} values can also be visualized based on a Factorial Correspondence Analysis graph using the GENETIX v. 4.05 (Belkhir et al., 2004) software (**Fig. 4.2.3**).

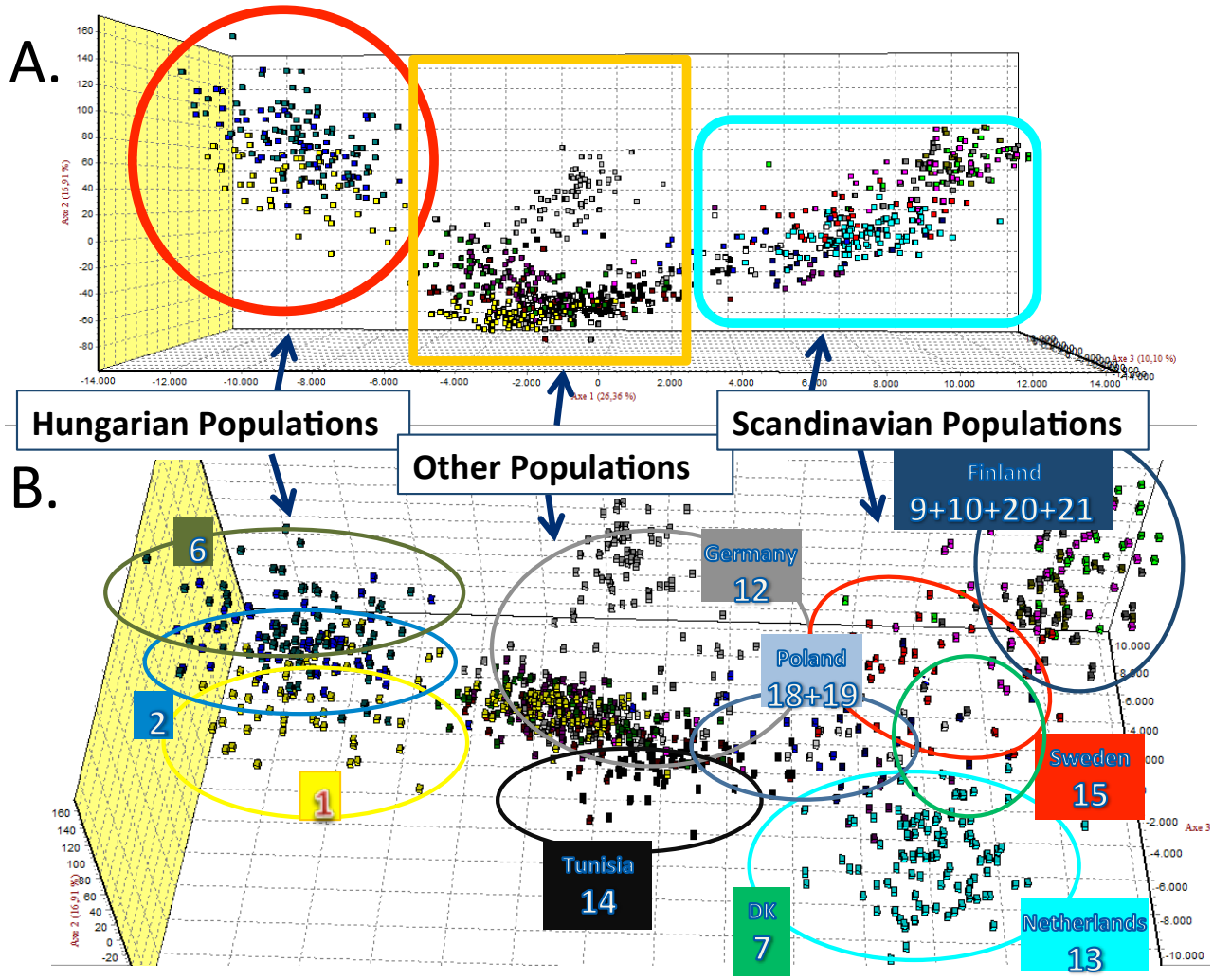

Figure 4.2.3 Factorial Correspondence Analysis (FCA) for all 21 populations and ten loci using the GENETIX v. 4.05 (Belkhir et al., 2004) software; populations numbers follow those in **Tables 4.2.1**.

Structure Analysis

Admixture analysis with STRUCTURE suggested a k value of two as the most likely number of existing clusters based on the ΔK criterion (**Fig. 4.2.4**). The first cluster comprises all the populations from Netherlands/Denmark/Poland and northwards to Sweden/Finland (light blue in **Fig. 4.2.5** for K=2) and the second one all the remaining populations (red in 4.2.5 for K=2). If we progressively take into account three and four clusters groupings in **Fig. 4.2.5**, the two above mentioned clusters are further divided to a Scandinavian Sweden/Finland cluster (dark blue for K=4) and a Hungarian one (red for K=4), respectively.

Interestingly, the populations of the same areas showed low to medium levels of admixture. For the 'Scandinavian' group, the Aquapri's "Mosso fish" (population 7) contained some 19% of 'northern' genotypes and inversely the Swedish sample (population 15) had some 5% of the 'southern' genotypes of the

first (blue) cluster. The situation in Polish lakes is transitory with the first one (population 18) containing 25% of cluster two (red-orange) genotypes. For the second (red) pikeperch cluster, there is a clear grouping of the Hungarian samples (populations 1, 2 and Aquapri's Hungarian population 6) with the Inagro's German origin fish (73% of the population 12); the later seem to be differentiated from those held in IfB Potsdam (population 8, German fish too) that are 'orange' type at 96%. Finally, Asialor's sample (population 11) showed 82.5% of cluster two (orange) and 17.3% of cluster one (light blue) genotypes.

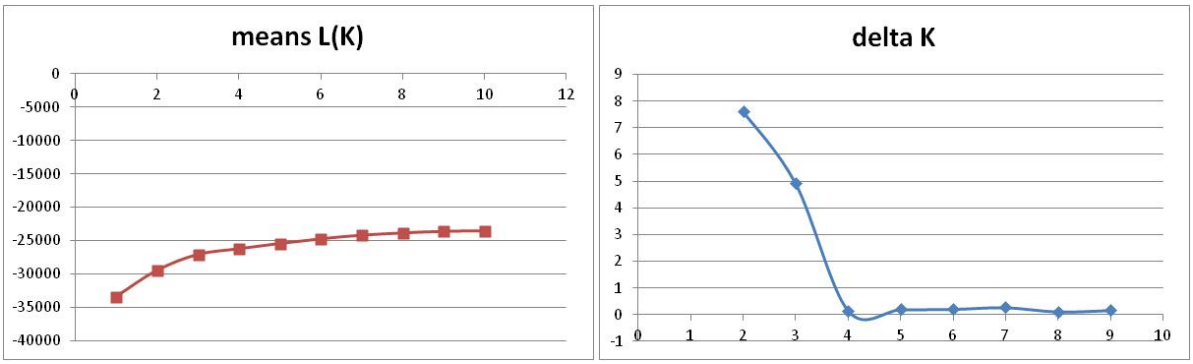

Figure 4.2.4 Admixture analysis revealed two genetic clusters as the most likely number, as indicated by a decrease in likelihood (A) and an increase in variance of calculated probabilities Δ(K) (B) Determination of the number of clusters (*K*) including all 10 repetitions for each *K* without geographical area as a prior. The highest peak denotes the most likely number of clusters according to the Pritchard Bayes formula. PD: probability of data.

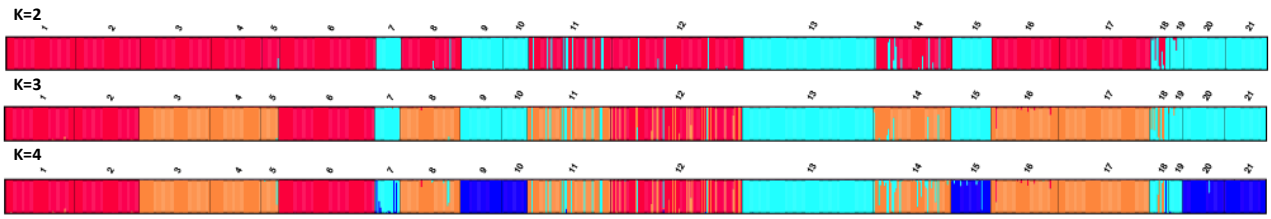

Figure 4.2.5 Bayesian individual assignment implemented in STRUCTURE for *K* = 2, 3 and 4 clusters without using geographical area as a prior. The *y*-axis represents the probability of assignment of an individual to each cluster and each color. Population numbers follow those in **Tables 4.2.1**.

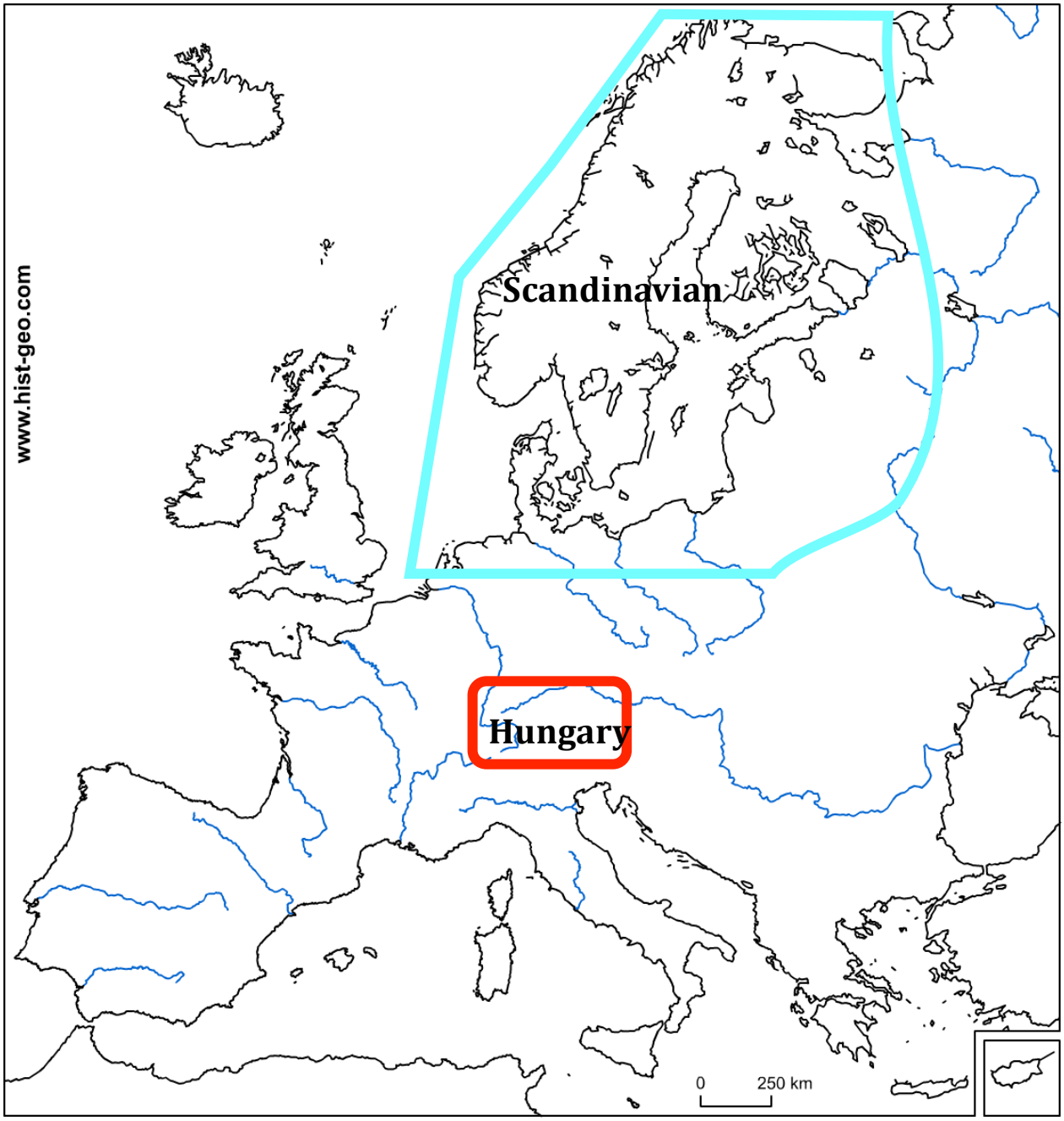

Figure 4.2.6 Map of Europe showing the major pikeperch genetic groups; colors follow those in **Fig. 4.2.5**.

Conclusions

Genetic studies in this species were until today very scarce and information was lacking on the genetic structure of wild populations, which in turn is a prerequisite for its successful conservation, and in the case of DIVERSIFY it is necessary to monitor the changes that may result from culture practices.

In this report we provide evidence that pikeperch populations in Europe are part of at least two genetically differentiated groups. The first is found in northern Europe from Netherlands/Denmark to the West and Poland (at least) to the East to the North of Finland (see **Fig. 4.2.6**). This is the group probably referred also

as "Baltic Sea" stock by Björklund et al. (2007) and Poulet et al. (2009). The second group comprises all remaining populations in Central Europe to as south as Tunisia (and probably Spain, Italy and Northern Greece).

In the second stock, the Hungarian populations are having a key-position being different from those found geographically near, *e.g.*, from Czech Rep. and Germany. It might be another stock associated with Hungarian lakes, as opposed to all other populations that probably dispersed through the Danube River west- and southwards.

Most populations analyzed seemed to contain fish of a single origin. The exception was 1) Aquapri's "Mosso fish" (population 7), which had some 19% of 'northern' genotypes, 2) the Swedish sample (population 15), which had ~ 5% of 'southern' genotypes of the same stock, 3) Sarag L. in Poland, which had 25% of cluster two genotypes, 4) the Inagro's German origin sample, which had 73% of Hungarian origin and 24% of other stock two fish, and 5) Asialor's sample, which had 82.5% of cluster two and 17.3% of cluster one genotypes.

It seems that there is no single "German stock"; the Inagro's German origin had some 73% of Hungarian origin, whereas those bred in IfB Potsdam are 'orange' type at 96%. Of course, one must always think of trade practices, which have absolutely to be taken into account for future breeding programmes in the species, *i.e.*, correct stock allocation.

In general, the mean heterozygosity estimates and the count of the number of alleles per population indicate that domesticated samples do not suffer from inbreeding. There are few domesticated populations that either due to their small sample size (Aqaupri's "Excellence" and "Mosso", and Laukaa in Finland) or their a priori known use as 'selected' fish, which indicates the notion of some level of inbreeding (*e.g.*, "Excellence fish" in Aquapri). Interestingly, the number of alleles in domesticated samples is slightly higher than that in the wild (2.63 versus 2.58, **Fig. 4.2.1**), whereas the unbiased heterozygosity is slightly lower (0.553 versus 0.573). At this point, we should also make a special reference to the 'light red line' used to distinguish populations considered as 'wild' in this study but may be a case intentionally (re)introduced stocks (population 14 from Tunisia and population 17 from Czech Rep.).

For the wild populations, the studies already performed show different mean heterozygosity (or gene diversity) values (>0.65 in Eschbach et al., 2014, >0.49 in Bjorklumnd et al., 2007, 0.39 in Saisa et al., 2010 and 0.51 in Salminen et al., 2012) and allelic richness (>4.8 in Eschbach et al., 2014, >3.38 in Bjorklumnd et al., 2007, 5.3 in Saisa et al., 2010 and 4.1 in Salminen et al., 2012). However, these results have to be considered with caution since other genetic studies were performed either with different microsatellite loci (Eschbach et al., 2014) or different population genetics parameters were measured (e.g. allelic richness and gene diversity instead of unbiased heterozygosity in Bjorklumnd et al., 2007, Saisa et al., 2010 and Salminen et al., 2012).

Last, we should bear in mind that besides inbreeding that reduces genetic diversity and the effective population size, outbreeding is also a major concern for future breeding programmes. Outbreeding is simply the crossing of different stocks, *i.e.*, locally adapted populations/strains with others that are significantly different genetically. Fish in a given (wild) population/strain possess a particular arrangement of alleles at different loci (co-adapted gene complexes). Crossing (hybridization) between the reared strains may potentially lead to a breakdown of these complexes resulting in reduced fitness. Often, the first generation is highly variable, but the difficulty to demonstrate outbreeding depression is that its severity is evident after the second and subsequent generations. For aquaculture purposes, these are the two potential strategies to consider the population closed for breeding purposes as the variability is sufficient and similar to wild populations or to bring in new strains to the population (crossing or hybridization). The scientists involved should decide whether the benefits from crossing different strains outweigh any later detrimental effects on fitness coming from outbreeding depression. For the domesticated samples analyzed, this is the case for 1) Aquapri's "Mosso fish" that had some 19% of Swedish/Finnish genotypes, 2) the Swedish sample that had ~ 5% of 'southern' genotypes of the same stock, 3) the Inagro's German origin sample that had 73% of Hungarian origin fish, and 5) the Asialor's sample that had 82.5% of cluster two and 17.3% of cluster one genotypes.

Deviations: There were no major deviations from the general outline in the DOW. As previously mentioned, one minor deviation was the development of two instead of one single microsatellite multiplex for genotyping, and the final exclusion of one locus (Pfla3) due to null alleles that lead to the use of ten loci for the population genetics analyses in the species. The adequate collaboration between the research teams implicated in WP4 and the commercial farms permitted us to analyze more fish and populations than initially planned: we report results on 21 populations and more than 950 fish, whereas we had proposed a minimum of 13 populations and 650 fish.

References

Belkhir, K., Borsa, P., Goudet, J., Chicki, L., Bonhomme, F., 1996-2004. GENETIX 4.05, logiciel sous WindowsTM pour la génetique des populations. Laboratoire Génome, Populations, Interactions, CNRS UMR 5000, Université de Montpellier II, Montpellier (France). Available on http://www.univ-montp2.fr/genetix/genetix/ genetix.htm.

Björklund, M., Aho, T., and Larsson, L.C. (2007). Genetic differentiation in pikeperch (*Sander lucioperca*): the relative importance of gene flow, drift and common history. Journal of Fish Biology *71*, 264-278.

Eschbach, E., Nolte, A.W., Kohlmann, K., Kersten, P., Kail, J., and Arlinghaus, R. (2014). Population differentiation of zander (*Sander lucioperca*) across native and newly colonized ranges suggests increasing admixture in the course of an invasion. Evolutionary Applications *7*, 555-568.

Evanno, G., Regnaut, S., and Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular Ecology *14*, 2611-2620.

Falush, D., Stephens, M., and Pritchard, J.K. (2003). Inference of Population Structure Using Multilocus Genotype Data:Linked Loci and Correlated Allele Frequencies. Genetics *164*, 1567–1587.

Goudet, J. (1995). Fstat version 1.2: a computer program to calculate Fstatistics. Journal of Heredity *86*, 485-486.

Miller, S.A., Dykes, D.D., and Polesky, H.F. (1988). A simple salting out procedure for extracting DNA from human nucleated cells. Nucl Acids Res *16*, 1215-.

Peakall, R., and Smouse, P.E. (2012). GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research - an update. Bioinformatics *28*, 2537-2539.

Pritchard, J.K., Stephens, M., and Donnelly, P. (2000). Inference of Population Structure Using Multilocus Genotype Data. Genetics *155*, 945–959.

Säisä, M., Salminen, M., Koljonen, M.-L., and Ruuhijärvi, J. (2010). Coastal and freshwater pikeperch (*Sander lucioperca*) populations differ genetically in the Baltic Sea basin. Hereditas *147*, 205-214.

Salminen, M., Koljonen, M.-L., Säisä, M., and Ruuhijärvi, J. (2012). Genetic effects of supportive stockings on native pikeperch populations in boreal lakes – three cases, three different outcomes. Hereditas *149*, 1-15.

Weir, B.S., and Cockerham, C.C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution,* 38, 1358-1370.