Deliverable Report

Deliverable No:	4.1	**Delivery Month:**	12
Deliverable Title	Genetic analysis of domesticated pikeperch broodstocks		
WP No:	4	**WP Lead beneficiary:**	P1. HCMR
WP Title:	Reproduction and Genetics – pikeperch		
Task No:	4.1	**Task Lead beneficiary:**	P9. UL
Task Title:	Evaluation of the genetic variation in available domesticated broodstocks of pikeperch		
Other beneficiaries:	P1. HCMR		
Status:	Delivered	**Expected month:**	12
…….			

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

Objective: The objective of this Deliverable was to assess the genetic variability of captive broodstocks in commercial farms in Europe operating in recirculating aquaculture systems (RAS), and later compare this variability with that of wild populations (D4.2 Population genetics analysis of wild and comparison with domesticated pikeperch populations 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: The pikeperch (*Sander lucioperca*) is a temperate Eurasian freshwater fish species, which tolerates brackish water, with growing interest for the European aquaculture. Currently, wild populations show signs of decline in many areas of its natural range of distribution (from Finland to the Aegean Sea, and East to the Aral and Caspian basins) due to human activities, such as the destruction of natural habitats and/or overfishing. The species has also been introduced in northern Russia, Italy, Spain, Turkey, the North African countries (from Morocco to Tunisia) and many other regions; it is a fish that may live for 17 years and can reach 100 cm in length and 20 kg in weight (www.fishbase.org).

Currently, there are no assessments of the genetic diversity of captive pikeperch stocks and because there are only a few commercial hatcheries that produce pikeperch (around 10) in Europe, the genetic diversity is 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 at the base of the captive stocks.

Genetic studies carried out so far in this species are scarce and there is still a lack of information 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. From a population genetics point of view, Björklund et al. (2007) analyzed genetically with six microsatellites two Fennoscandian regions (North and South of Finland and Sweden) and found that the allelic richness and the degree of differentiation were significantly higher in the North (F_{ST} =0.20) than in the South (F_{ST} =0.064). They also reported that when comparing these two areas, the percentage of correct assignments of individual genotypes to their population of origin was much greater in northern than in southern samples, *i.e.* the

northern samples were easier to differentiate genetically. Björklund et al. (2007) also suggested that there has been more gene flow between populations in southern areas than in northern areas, where the importance of genetic drift has been greater. Later, Poulet et al. (2009) used the same set of six microsatellite loci as above to show that values of allelic richness and unbiased expected heterozygosity observed in the Rhône drainage introduced populations in France were similar, or even higher, compared to native populations from the Baltic Sea drainage studied in Björklund et al. (2007). This might be explained by multiple introductions in the Rhône drainage, but also by demographic strategy that would have facilitated population persistence in this fragmented habitat. Similarly, heterozygote deficits (revealed by F_{IS} values) have been detected, but were also found in native populations suggesting that mating among relatives could also result from the mating behavior of the species, maybe reinforced in the Rhône drainage by the reduced carrying capacity of the artificial canals and their respective isolation. Despite harsh environmental conditions and suspected inbreeding, the pikeperch has successfully maintained viable populations in the Rhône delta. This study suggested that one of the factors in this invasive success, apart from the species ecology, could be the maintenance of a good level of genetic diversity in introduced pikeperch populations. This genetic diversity probably stems from i) its popularity as game fish and food resource, which led to numerous stockings and an increasing propagation pressure, and ii) the reproductive strategy of the species.

Furthermore, Saisa et al. (2010) performed a microsatellite DNA based analysis (12 loci) among three coastal and five freshwater populations of pikeperch in the northern part of the Baltic Sea drainage basin and reported marked genetic differentiation between the coastal and lake populations with high F_{ST} values between these population groups (as high as 0.25). The three coastal populations grouped tightly together, whereas the freshwater samples formed a looser group with the lake populations showing, in general, higher genetic diversity than the coastal ones (similar to Björklund et al., 2007). Recently, Salminen et al. (2012) assessed the genetic consequences and gene flow of pikeperch in three boreal lakes based on admixture model analysis and comparison of the pre- and post-release patterns of genetic variability at nine microsatellite loci in the recipient populations. They reported that the genetic structure of populations were disrupted by the releases of fish; in two out of the three cases, the release of fish from foreign populations caused significant changes in the genetic structure of the recipient population. The genetic analyses showed that the current stocking practices create an effective artificial gene flow that may strongly shape and reduce the genetic differentiation among the remaining native pikeperch populations. A common feature of all three cases was the lack of prior appraisal of the potential genetic and ecological risks in relation to the expected benefits of the release programmes.

The primary objective of the present Deliverable was to develop a highly informative and efficient microsatellite multiplex for the species, which ideally should consist of more than 10 markers to allow the adequate genotyping of all pikeperch populations sampled. This microsatellite multiplex is expected to be used for genotyping purposes on a capillary sequencer and evaluate the genetic variability of captive broodstock in commercial RAS farms around Europe and finally compare this variability with the variability of wild stocks/populations (Task 4.2, Deliverable 4.2 Population genetics analysis of wild and comparison with domesticated pikeperch populations to be applied in future breeding programs of the species). The overall objective of these two Deliverables is to define how a future genetic breeding program should be established for sustainable optimal performance through domestication of pikeperch.

Material and Methods

Microsatellite Loci

Microsatellite loci were selected that had been successfully used for pikeperch. Multiplex were optimized for 22 loci (grouped into two multiplexes; Table 4.1.1) that were reported in: Leclerc et al. (2000) for the yellow perch (*Perca flavescens* – Code: Pfla), Borer et al. (1999) and Wirth et al. (1999) in walleye (*Stizostedion vitreum* – Code: Zvi), and Dubut et al (2010) in the Rhone streber (*Zingel asper* – Code: Za) (see Table 4.1.1). Those reported in pikeperch by Kohlmann & Kersten (2008) have shown low number of alleles (2-6 alleles in a population of 25 fish) and low expected heterozygosity (0.334 – 0.777) and have not been previously used in a multiplex to genotype populations. Microsatellite loci were first ordered by increasing size in base pairs (bp) and the size range (reported in the species described), and in each range one of the

primers for each microsatellite locus the reverse (code: R) was fluorescently labelled with dyes that conformed to P1. HCMR's sequencing technology (ABI 3730).

Table 4.1.1 Characteristics of the 22 microsatellite loci used in pikeperch. Loci were grouped into two multiplexes (upper and lower part of the table). For each locus, we report the accession number in NCBI, the repeat motif (sequence), the size range, the number of alleles (Na) and the observed and expected heterozygosity (Ho and He, respectively) in the species first described (Pfla for *Perca flavescens*, Zvi for *Stizostedion vitreum* and Za for *Zingel asper*), the fluorescent dye we used in the automatic sequencer for the reverse primer and the primers' sequences.

Locus	Acc.Number	Repeat Sequence	Size Range	Na	Ho	He	Dye	Forward sequence	Reverse sequence
Svi33	G36966	(AC)14	075–083	3	0.25	0.21	6-FAM	CAGGACTGCTGTGTATAGACTTG	GATATAGCTTTCTGCTGGGGTC
PflaL3	AF211828	(TG)18	101–119	8	0.34	0.29	6-FAM	GCCGAATGTGATTGAATG	CGCTAAAGCCAACTTAATG
SviL8	AF144741	(TG)22	107–145	8	0.34	0.20	Atto-565	GCTTATACGTCGTTCTTATG	ATGGAGAAGCAAGTTGAG
SviL11	AF144744	(TG)26G(TG)8	115–121	3	0.12	0.12	Atto-550	AGGGTATGGCATGATAAG	CTCTACATTTCATCAGACAG
Svi6	G36962	(AC)6	115–165	19	0.61	0.50	HEX	CATATTATGTAGAGTGCAGACCC	TGAGCTTCACCTCATATTCC
Svi18	G36964	(AC)18	132–182	10	0.67	0.59	6-FAM	GATCTGTAAACTCCAGCGTG	CTTAAGCTGCTCAGCATCCAGG
SviL9	AF144742	(CA)18AA(CA)3A(AC)4	161–223	10	0.17	0.17	Atto-550	TACTGTTCACTTATCTATCC	TGTATGTGTGTGTTCATGT
PflaL8	AF211833	(TG)39	167–203	16	0.72	0.51	Atto-565	GCCTTATTGTGTGACTTATCG	GGATCTTTCACTTTTTCTTTCAG
PflaL9	AF211834	(TG)24	182–214	4	0.65	0.52	HEX	GTTAGTGTGAAAGAAGCATCTGC	TGGGAAATGTGGTCAGCGGC
SviL7	AF144740	(TG)22	201–249	17	0.64	0.56	6-FAM	GATGTGCATACATTTACTCC	GCTTTAATCTGCTGAGAAC
PflaL2	AF211827	(CA)23	209–229	7	0.45	0.30	Atto-565	GTAAAGGAGAAAGCCTTAAC	TAGCATGACTGGCAAATG
Za121	HM622316	(CT)9	227-235	4	0.63	0.60	HEX	CAAAGTCATGAACGAGCTGC	AGCCAGGACCACTCTGTGAG
Za038	HM622298	(AC)11	107-130	6	0.80	0.77	6-FAM	TGAATCGCTGCTCTTTCTCA	TATGCAATTACATCGGAGCG
Svi4	G36961	(AC)16	120–166	15	0.70	0.65	Atto-550	ACAAATGCGGGCTGCTGTTC	GATCGCGGCACAGATGTATTG
Za024	HM622294	(AC)7	127–139	4	0.47	0.43	HEX	TGAACCTCCCTATCCCCTCT	TCTTTTCCACAGCAGGAAGC
Za138	HM622317	(AC)8	135-148	5	0.27	0.43	Atto-565	TTCTTTATACAAGAGGAATAGTTGCAG	TTTTTGTGATTGTGCTATTTTAAAGG
Za113	HM622314	(CA)11	169-229	9	0.70	0.85	HEX	ACCACGCACAATCACTCGTA	CCTGGCTTTACCCAGAAACA
Za237	HM622342	(CA)10	171-178	5	0.57	0.54	6-FAM	ATCTCAAGTCATGGGGCATC	GGTCCTCTGGTGCAGCTATAA
Za179	HM622329	(TCT)9	171-196	4	0.37	0.35	Atto-565	ATTTCCCATTGCGGGATTA	GGATTCTTGCATGCTTTGGT
Za144	HM622319	(AC)8	199-228	8	0.70	0.80	Atto-550	GCCCACAATAGCACCGTAAT	TTTGTGAATGTGAGTGAGAGTCAG
Za199	HM622334	(TCT)13	201-234	7	0.67	0.74	6-FAM	CCTTCCCCTCAAAAGCATGT	AGGAAATGGAAAGGGAATGC
Za207	HM622337	(GT)13	222-237	5	0.67	0.64	Atto-565	GGATTCCAGAAGCAAAAGAGG	TGGGACAAGGCTTTAACCAC

Polymerase chain reactions (PCR) were performed in 12.5 µl total volume, with the following cycling conditions. **1st Multiplex**: initial denaturation at 95°C for 5 min, 35cycles of 30 sec at 95 °C, 90 sec at 59°C, 90 sec at 72°C and a final extension for 30 min at 68°C. **2nd Multiplex:** initial denaturation at 95°C for 3 min, 30cycles of 15 sec at 95 °C, 30 sec at 58°C, 90 sec at 72°C and a final extension for 10 min at 72°C.

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).

Biological material

DNA extractions have been completed for all domesticated samples/populations following standard protocols (salt precipitation, Miller et al., 1988). In total, DNA was extracted and analyzed from 439 fish

samples (Table 4.1.2); we also used a wild population of 53 fish from Hungary as a reference for all population genetics parameters.

Table 4.1.2 List of domesticated pikeperch samples and number of fish per sample that were genotyped and analyzed; the first population marked in red is of wild origin.

A/a	Population	Sample size
1	*Gyori Elore, HTSZ, Hungary*	*53*
2	Szabolsi, Halaszati Kft, Hungary	50
3	VanMecklen, Holland, Aquapri A/S, Danemark	54
4	Czech Rep., Aquapri A/S, Danemark	38
5	Excellence fish, Hollande, Aquapri A/S, Danemark	14
6	Hungary, Aquapri A/S, Danemark	74
7	Mosso, Aquapri A/S, Danemark	19
8	IfB, Potsdam, Germany	48
9	FGFRI Kainuu fisheries research station, Finland	31
10	FGFRI Laukaa Fish Farm, Finland	20
11	ASIALOR, France	31
12	INAGRO, Belgium (German origin)	30
13	INAGRO, Belgium (Dutch origin)	30

Results and Discussion

The Qiagen multiplex PCR kit was used to optimize PCR conditions and to determine the feasibility of working with the two multiplexes (12-plex and 10-plex) and finally to determine that feasibility of developing one single multiplex as a powerful molecular tool for genotyping. Furthermore, the Qiagen multiplex PCR kit gives the advantage of maximal transferability of molecular protocols between labs. Unfortunately, the Qiagen multiplex PCR kit did not provide a single optimized multiplex that could be used for genotyping and the following two multiplexes were developed and optimized:

- 1st Multiplex: with loci Svi18, PflaL3, Za138 and Za199.

- 2nd Multiplex: with loci Za038, Svi4, Za024, Za237, Za144, Za207 and PflaL9

In domesticated stocks, the loss of genetic variability within the first generations of breeding practices limits the potential for future genetic improvement from selection practices. Considering a long term breeding program, ensuring sufficient genetic variation within populations is fundamental, because it determines the potential of adaptation to hostile changes in environmental /rearing conditions. Basic population genetics parameters (allelic richness, heterozygosity indices, inbreeding coefficients) were calculated for both wild and domesticated stocks

The number of alleles per locus ranged from 6-7 (PflaL3 and PflaL9, respectively) to 20 (Za138) (Tables 4.1.3 to 4.1.5). Therefore, microsatellite loci showed relatively high levels of polymorphism even though some samples were monomorphic (exhibited only one allele) for some loci like for Za199 and PflaL9 in the Excellence fish of Aquapri A/S (population 5), locus Za237 in Kainuu fisheries research station (population 9), and Za144 in Laukaa Fish Farm (population 10).

For the thirteen populations analyzed, the least number of alleles was encountered in Aquapri's VanMecklen (2.64), Aquapri's Excellence fish and Laukaa Fish Farm (2.73) and the greatest in Hungarian Aquapri's (7.91) and Halaszati Kft (7.55) stocks, which were greater than that in wild Hungarian stock (6.00). Likewise, expected heterozygosity (H_E) ranged from 0.3198 (in Aquapri's Excellence fish) to 0.7163 (in Aquapri's Hungarian fish) (Table 4.1.6)

Table 4.1.3 Number of alleles per locus; populations numbers follow those in Table 4.1.2.

	Population													
Locus	1	2	3	4	5	6	7	8	9	10	11	12	13	Total
PflaL3	4	5	3	3	2	5	2	3	3	2	5	4	4	6
Svi18	6	12	3	4	4	13	4	4	4	3	6	7	5	18
Za199	5	5	2	3	1	6	3	4	4	3	5	9	7	14
Za138	9	14	3	3	2	13	2	7	4	2	6	9	6	20
PflaL9	3	4	2	3	1	3	4	5	3	5	3	5	4	7
Svi4	7	6	4	5	5	8	5	6	5	4	4	5	5	15
Za024	7	6	2	3	2	7	2	6	4	3	4	4	2	12
Za038	6	8	3	4	4	9	3	4	4	3	4	5	3	11
Za144	8	8	3	3	4	8	3	9	4	1	8	8	4	17
Za207	4	6	2	3	3	6	3	7	4	2	6	5	3	11
Za237	7	9	2	2	2	9	2	5	1	2	3	5	3	13

Table 4.1.4 Unbiased gene diversity per locus and population calculated using FSTAT 2.9.3 (Goudet 1995); populations numbers follow those in Table 4.1.2.

	Population												
Locus	1	2	3	4	5	6	7	8	9	10	11	12	13
PflaL3	0,647	0,661	0,632	0,645	0,071	0,677	0,491	0,509	0,498	0,505	0,661	0,744	0,648
Svi18	0,683	0,703	0,62	0,713	0,766	0,784	0,525	0,481	0,651	0,528	0,732	0,76	0,692
Za199	0,504	0,711	0,404	0,22	0	0,655	0,481	0,305	0,649	0,646	0,672	0,762	0,736
Za138	0,855	0,881	0,619	0,555	0,091	0,811	0,053	0,73	0,578	0,439	0,679	0,813	0,648
PflaL9	0,623	0,617	0,358	0,218	0	0,627	0,642	0,541	0,494	0,695	0,284	0,603	0,661
Svi4	0,748	0,727	0,567	0,687	0,613	0,78	0,599	0,783	0,695	0,654	0,584	0,71	0,701
Za024	0,631	0,583	0,491	0,196	0,2	0,541	0,053	0,538	0,621	0,496	0,629	0,486	0,334
Za038	0,638	0,693	0,624	0,53	0,544	0,733	0,528	0,528	0,625	0,653	0,713	0,771	0,593
Za144	0,758	0,812	0,61	0,519	0,604	0,82	0,66	0,611	0,57	0	0,709	0,853	0,59
Za207	0,67	0,715	0,5	0,566	0,589	0,661	0,648	0,586	0,525	0,295	0,69	0,652	0,668
Za237	0,782	0,82	0,019	0,477	0,159	0,845	0,102	0,525	0	0,429	0,604	0,664	0,615

Table 4.1.5 Allelic richness per locus and population calculated using FSTAT 2.9.3 (Goudet 1995); populations' numbers follow those in Table 4.1.2.

	Population													
Locus	1	2	3	4	5	6	7	8	9	10	11	12	13	Total
PflaL3	3.327	4.074	2.994	2.994	1.714	3.93	2	2.213	2.868	2	3.862	3.989	3.541	4.664
Svi18	4.724	6.677	2.996	3.95	4	7.724	3.484	3.218	3.688	2.5	5.02	5.465	4.194	8.064
Za199	3.893	4.543	1.999	2.457	1	4.542	2.903	2.858	3.314	2.999	4.259	5.92	5.347	6.145
Za138	7.514	9.134	2.99	2.774	1.909	7.418	1.526	5.214	3.303	2	4.556	6.736	4.761	8.855
PflaL9	2.988	3.184	1.997	2.354	1	2.987	3.896	3.435	2.942	4.255	2.51	4.101	3.931	4.297
Svi4	5.536	4.737	3.161	3.883	4.354	5.554	3.835	5.707	4.33	3.554	3.171	4.044	4.534	7.622
Za024	4.429	3.368	2	2.311	2	3.515	1.526	4.233	3.314	2.756	3.529	3.534	1.997	5.028
Za038	4.783	4.472	2.994	2.526	3.429	5.242	2.526	2.742	3.833	2.999	3.921	4.917	2.992	6.051
Za144	6.147	5.905	2.978	2.263	3.64	5.998	2.999	4.665	3.089	1	5.626	6.757	3.827	7.979
Za207	3.221	5.072	2	2.804	3	4.924	2.997	4.02	2.645	1.996	4.939	4.194	2.999	6.202
Za237	5.272	6.606	1.185	2	1.978	6.731	1.782	3.351	1	2	2.991	4.089	2.981	5.663

Table 4.1.6 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**.**

A/a	Population	No Fish	Mean No Alleles	H_E	H_O	Fis
1	Gyori Elore, HTSZ, Hungary	53	6.00	0.6787	0.7325	-0.06881*
2	Szabolsi, Halaszati Kft, Hungary	50	7.55	0.7121	0.6712	0.06807*
3	VanMecklen, Holland, Aquapri A/S, Denmark	54	2.64	0.4921	0.7084	-0.43179*
4	Czech Rep., Aquapri A/S, Denmark	38	3.27	0.4781	0.5157	-0.0650*4
5	Excellence fish, Holland, Aquapri A/S, Denmark	14	2.73	0.3198	0.3922	-0.18608*
6	Hungary, Aquapri A/S, Denmark	74	7.91	0.7163	0.7193	0.00275
7	Mosso, Aquapri A/S, Denmark	19	3.00	0.4224	0.4149	0.04555
8	IfB, Potsdam, Germany	48	5.45	0.5516	0.5447	0.02399
9	FGFRI Kainuu fisheries research station, Finland	31	3.64	0.5291	0.6003	-0.11825*
10	FGFRI Laukaa Fish Farm, Finland	20	2.73	0.4754	0.5801	-0.19545*
11	ASIALOR, France	31	4.91	0.6215	0.5909	0.06584*
12	INAGRO, Belgium (German origin)	30	6.00	0.7003	0.8057	-0.13374*
13	INAGRO, Belgium (Dutch Origin)	30	4.18	0.6152	0.6196	0.01019

A wide range of F_{IS} values were observed in the 13 populations analyzed (Tables 4.1.6 and 4.1.7). In principle, **positive** F_{IS} values indicate that individuals in a population are **more related** than you would expect under a model of random mating, whereas **negative** F_{IS} values indicate that individuals in a population are **less related** than you would expect under a model of random mating. The F_{IS} values are high and significant for Halaszati Kft (0.068), Aquapri's Mosso (0.0455) and ASIALOR (0.0658) 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 4.1.7 F_{IS} values per locus and population; populations numbers follow those in Table 4.1.2. Asterisks indicate significance at p=0.05. The last row shows F_{IS} values per populations when locus PflaL3 is excluded due to null alleles.

Locus	1	2	3	4	5	6	7	8	9	10	11	12	13
PflaL3	0,096	0,023	-0,26	-0,183	0	-0,108	0,036	-0,003	0,028	0,307	0,512	-0,3	0,537
Svi18	-0,409	0,199	-0,105	0,078	-0,305	-0,048	-0,003	-0,31	-0,091	-0,137	-0,366	0,254	-0,156
Za199	0,064	-0,202	-0,377	0,641	NA	-0,011	-0,094	0,002	-0,192	-0,238	0,136	-0,225	0,016
Za138	-0,032	0,33	-0,615	-0,091	0	-0,015	0	0,256	-0,211	-0,44	0,287	-0,025	-0,184
PflaL9	-0,242	0,027	-0,293	0,395	NA	0,03	0,344	0,219	0,15	-0,08	0,091	-0,27	0,043
Svi4	0,003	-0,047	-0,431	-0,142	-0,516	0,133	-0,668	0,138	-0,439	-0,53	-0,298	-0,267	-0,207
Za024	0,022	0,02	0,057	0,193	1	0,037	0	0,112	0,117	-0,109	-0,166	-0,029	-0,033
Za038	0,024	-0,06	-0,603	-0,341	-0,051	-0,05	0,402	0,012	-0,291	-0,302	0,252	0,265	-0,013
Za144	-0,116	0,069	-0,64	0,037	-0,3	0,007	0,465	-0,104	0,094	NA	0,013	-0,051	0,096
Za207	-0,161	0,039	-1	-0,289	-0,189	0,088	-0,138	-0,046	-0,23	-0,188	0,112	-0,483	0,052
Za237	0,01	0,244	0	-0,158	-0,048	-0,021	-0,029	-0,175	NA	-0,166	0,145	-0,456	-0,03
All	-0,069*	0,068*	-0,432*	-0,065*	-0,186*	0,003	0,046	0,024	-0,118*	-0,195*	0,066*	-0,134*	0,01
All-10L	-0.084*	0.072*	-0.454*	-0.049	-0.190*	0.013*	0.047	0.026	-0.132*	-0.248*	0.019	-0.116*	-0.045

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.

Moreover, one of the above microsatellite loci (PflaL3) showed significant probability ($P > 0.05$) of "large allele dropout" or "stuttering". However, when this locus is excluded from the analysis the F_{IS} values are slightly changed but remain significant in any case (see bottom of Table 4.1.7).

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: higher values for biallelic single-nucleotide polymorphisms (SNPs) than for multi-allelic microsatellites, low values among high-diversity populations viewed as substantially distinct, and low values for populations that differ primarily in rare alleles. Due to these reasons, several authors argued that F_{ST} measures may be poor measures of genetic differentiation when the level of diversity is high. When we estimate population differentiation across samples using the F_{ST} estimate by Weir & Cockerham's (1984) we see that the smallest values are between Hungarian samples (wild-1 and domesticated-2) and we also show (as expected) close relationship (F_{ST} <0.11) of the above two populations with the Aquapri's Hungarian one (pop 6) (Table 4.1.8). We also see a close relationship of Aquapri's VanMecklen (pop 6) with the Czech population from the same company and that from IfB Potsdam (F_{ST} <0.15). Lastly, we see a close relationship between ASIALOR and INAGRO's Belgian samples (F_{ST} =0.14) and Aquapri's Mosso sample with that from INAGRO's Dutch samples (F_{ST} =0.16).

Table 4.1.8 Pairwise Population Theta (θ) calculations [Weir & Cockerham's (1984) F_{ST}'s] between the thirteen populations; populations numbers follow those in Tables 4.1.2 and 4.1.6.

		2	3	4	5	6	7	8	9	10	11	12	13
1	Gyori Elore, HTSZ, Hungary	0,09	0,27	0,25	0,31	0,11	0,36	0,24	0,35	0,37	0,20	0,19	0,28
2	Szabolsi, Halaszati Kft, Hungary	-------	0,26	0,25	0,29	0,01	0,33	0,23	0,32	0,34	0,19	0,15	0,24
3	VanMecklen, Holland, Aquapri A/S, Danemark	-------	---------	0,12	0,27	0,26	0,29	0,15	0,36	0,44	0,15	0,24	0,30
4	Czech Rep., Aquapri A/S, Denmark	-------	---------	---------	0,20	0,27	0,34	0,11	0,39	0,47	0,17	0,21	0,32
5	Excellence fish, Hollande, Aquapri A/S, Denmark	-------	---------	---------	---------	0,30	0,42	0,21	0,44	0,51	0,25	0,27	0,33
6	Hungary, Aquapri A/S, Denmark	-------	---------	---------	---------	---------	0,33	0,23	0,31	0,34	0,19	0,16	0,24
7	Mosso, Aquapri A/S, Denmark	-------	---------	---------	---------	---------	---------	0,25	0,24	0,37	0,22	0,24	0,16
8	IfB, Potsdam, Germany	-------	---------	---------	---------	---------	---------	---------	0,32	0,39	0,14	0,20	0,26
9	FGFRI Kainuu fisheries research station, Finland	-------	---------	---------	---------	---------	---------	---------	---------	0,21	0,28	0,22	0,22
10	FGFRI Laukaa Fish Farm, Finland	-------	---------	---------	---------	---------	---------	---------	---------	---------	0,34	0,28	0,25
11	ASIALOR, France	-------	---------	---------	---------	---------	---------	---------	---------	---------	---------	0,14	0,22
12	INAGRO, Belgium (German origin)	-------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	0,19
13	INAGRO, Belgium (Dutch origin)	-------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------

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. 4.1.1).

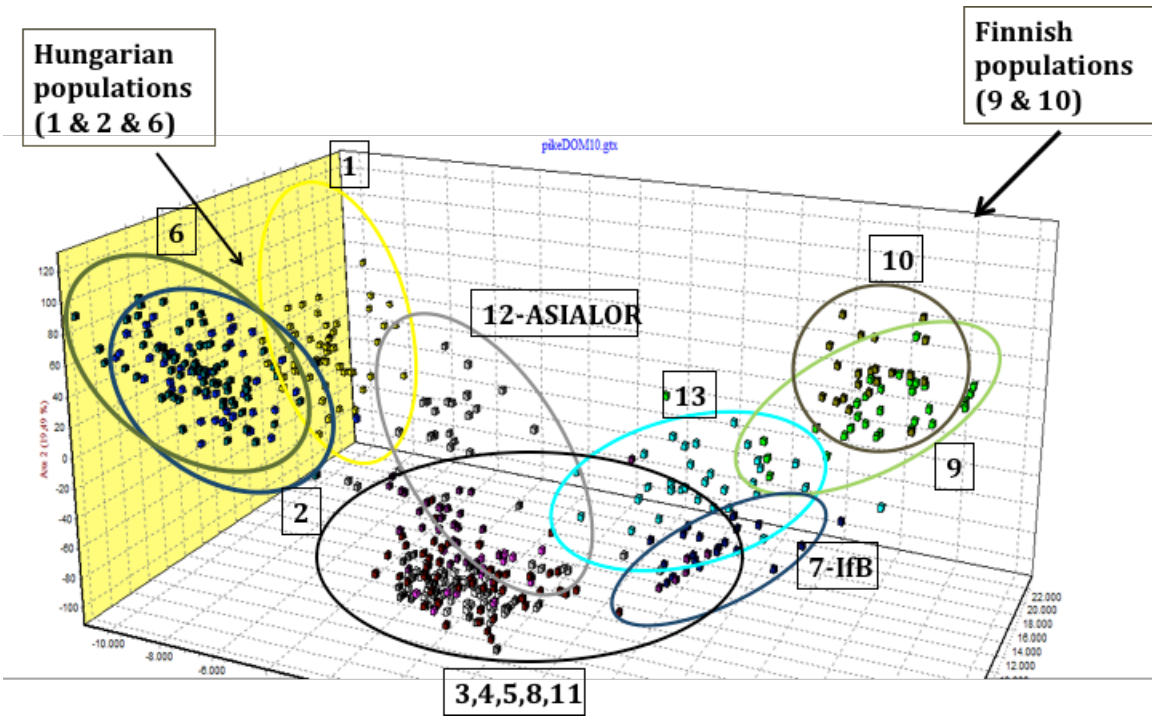

Figure 4.1.1 Factorial Correspondence Analysis (FCA) for all thirteen population and ten loci using the GENETIX v. 4.05 (Belkhir et al., 2004) software; populations numbers follow those in Tables 4.1.2 and 4.1.6.

Deviations: There were no major deviations from the general outline in the DOW. One minor deviation refers to the development of two instead of one single microsatellite multiplex to genotype at first the domesticated stocks available and the wild stocks in the future (D4.2). Unfortunately the size range of the loci used and the optimum PCR conditions for each locus did not permit that the eleven loci were combined and used in a single multiplex. In the long term this may lead to slightly higher financial (consumables) and labor costs (twice the number of PCRs to prepare), but it will have absolutely **no impact either on the scientific deliverables or on the project overall**.

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.

Borer, S.O., Miller, L.M., and Kapuscinski, A.R. (1999). Microsatellites in walleye *Stizostedion vitreum* . Molecular Ecology *8*, 336-338.

Dubut, V., Grenier, R., Meglécz, E., Chappaz, R., Costedoat, C., Danancher, D., Descloux, S., Malausa, T., Martin, J.-F., Pech, N., *et al.* (2010). Development of 55 novel polymorphic microsatellite loci for the critically endangered *Zingel asper* L. (Actinopterygii: Perciformes: Percidae) and cross-species amplification in five other percids. European Journal of Wildlife Research *56*, 931-938.

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

Kohlmann, K., and Kersten, P. (2008). Isolation and characterization of nine microsatellite loci from the pike-perch, *Sander lucioperca* (Linnaeus, 1758). Molecular Ecology Resources *8*, 1085-1087.

Leclerc, E., Mailhot, Y., Mingelbier, M., and Bernatchez, L. (2008). The landscape genetics of yellow perch (*Perca flavescens*) in a large fluvial ecosystem. Molecular Ecology *17*, 1702-1717.

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.

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.

Van Oosterhout, C., Hutchinson, W.F., Wills, D.P.M., and Shipley, P. (2004). micro-checker: software for identifying and correcting genotyping errors in microsatellite data. Molecular Ecology Notes *4*, 535-538.

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

Wirth, T., Saint-Laurent, R., and Bernatchez, L. (1999). Isolation and characterization of microsatellite loci in the walleye (*Stizostedion vitreum*), and cross-species amplification within the family Percidae. Molecular Ecology *8*, 1960-1962.


Co-funded by the Seventh Framework Programme of the European Union