NPAFC
Doc. 913
Rev. _____

A CASE STUDY OF TWO GENETIC MARKERS FOR INTER-LABORATORY COLLABORATION: SNPS PROVIDE TRANSPORTABILITY WITHOUT STANDARDIZATION

by
Daria Zelenina[1],
Anastasia Khrustaleva[1],
Alexander Volkov[1],
Chris Habicht[2],
Christian Smith[2],
James Seeb[2].

[1]VNIRO, Russian Federal Research Institute of Fisheries & Oceanography, Federal Fisheries Agency of Russia, 17 V. Krasnoselskaya, Moscow, Russia
[2]Alaska Department of Fish and Game, 333 Raspberry Road, Anchorage, USA

Submitted to the
North Pacific Anadromous Fish Commission

By

RUSSIA

October 2005

THIS PAPER MAY BE CITED IN THE FOLLOWING MANNER:
Daria Zelenina, Anastasia Khrustaleva, Alexander Volkov**,** Chris Habicht, Christian Smith, James Seeb.. 2005. A Case Study of Two Genetic Markers for Inter-Laboratory Collaboration: SNPs Provide Transportability without Standardization. (NPAFC Doc. 913) 14 p. Russian Federal Research Institute of Fisheries & Oceanography, Federal Fisheries Agency of Russia, VNIRO, 17 V. Krasnoselskaya, Moscow, Russia.

INTRODUCTION

Genetic data have been used to partition stock components of sockeye salmon (*Oncorhynchus nerka*) in mixed stock samples on geographically regional scales for many years. Different genetic methods have been used for this purpose including allozymes (Seeb et al, 2000), microsatellites (Beacham et al., 2005), mtDNA variation (Brykov et al., 2003; 2005) and RAPDs (Zelenina et al, 2006). Although the genetic structure detected by these different methods has been similar, DNA methods have produced higher resolution in identifying stock compositions than allozymes. However, reproducibility of data using these DNA methods is only obtained on identical equipment using the same reagents. As the mixed stock analysis needs expand to broad geographical areas such as the Bering Sea, larger baselines are required. These larger baselines are most efficiently collected by multiple laboratories to create standardized Pacific Rim baselines. Creating such baselines using these DNA methods requires laboratories to share standard samples to calibrate each lab. Shipping samples internationally can be complicated or impossible. In Russia, for example, recent changes in the regulations for exporting biological materials have made in practically impossible to share samples with outside laboratories. An alterative approach to solving this problem is to find a method that will be completely reproducible in any laboratory. In this respect SNP polymorphism analysis seems to be very promising.

In this study, independent analysis of the same samples was carried out in two laboratories using SNP as well as microsatellites. We examined variations in two SNP and four microsatellite loci. It should be particularly noticed that the laboratory equipment as well as reagents including Taq-polymerase and fluorescent labels were different. Only the primer sequences were identical.

Given the consistency in allele scores between the American and Russian laboratory for the two SNP loci, we include SNP data for the same loci for 5 other collections from Russia.

MATERIALS AND METHODS

Tissue (liver) samples were collected from adult fish in three sockeye salmon populations in West Kamchatka, Russia. Each sample was divided into two parts: one of which was analyzed in Moscow, Russia (A) in Division of Molecular Genetics of Aquatic Animals of VNIRO and the other in Anchorage, United States (B) in Alaska Department of Fish and Game. In addition, five collections of sockeye salmon samples collected by Nataly Varnavskaya (KamchatNIRO) were screened for the two SNP loci (see below) by the United States.

DNA isolation.

(A) DNA was isolated by phenol-chloroform extraction and ethanol precipitation according to Maniatis et al (1982). DNA concentration was measured on "SPECTRAmax PLUS^{384}" (Molecular Devices) and normalized up to approximately 20 ng/μl.

(B) Genomic DNA was extracted using a DNeasy® 96 Tissue Kit by QIAGEN®, (Valencia, CA). No DNA quantification or normalization was performed.

Microsatellite analysis.

Four loci were used for comparative investigation: One108, One109, One111 and One 114. All of them were tetranucleotide repeats. The primers for PCR amplification were synthesized according to Olsen et al. (2000).

(A) The analysis was based on touchdown multiplex PCR with fluorescent labeled primers. Three loci - One108, One109 and One114 - were amplified in one reaction (multiplex-1) and the locus One111 was combined with two other loci – OMM1082 (Rexroad III, 2002) and Ots-253 (Williamson et al, 2002) (multiplex 2).

Forward primers in each pair were labeled on their 5' end with fluorescent dyes as follows: One108 and OMM1082 – 6-**JOE** (yellow), One109 and One111 – **FAM** (green), One114 and Ots253 – **TAMRA** (orange).

Multiplex 1 was performed in a final volume of 15 μl containing about 60 ng of DNA; 1x PCR-buffer (Silex, Russia); 3 mM $MgCl_2$; 60 μM each dATP, dGTP, dCTP and dTTP; 1,5 pmol of each labeled primer; 3 pmol of each unlabeled primer and 0,3 μl of Hot-Taq DNA polymerase (Silex, Russia).

Multiplex 2 was performed in a final volume of 15 μl containing about 60 ng of DNA; 1x PCR-buffer (Silex, Russia); 2 mM $MgCl_2$; 60 μM each dATP, dGTP, dCTP and dTTP; 0,75 pmol of Ots-253-F-TAMRA; 1,5 pmol Ots-253-R, OMM1082-F-JOE and One111-F-FAM; 3 pmol of OMM1082-R and One111-R; 0,3 μl of Hot-Taq DNA polymerase (Silex, Russia).

PCR conditions were the same for both multiplexes: initial denaturation and polymerase activation at 95°C for 10 minutes, 12 cycles at 95°C for 5 seconds, 68°C (decreasing for 0,5 °C per cycle) for 15 seconds, 70°C for 20 seconds followed by 25 cycles at 95°C for 5 seconds, 62°C for 15 seconds, 70°C for 20 seconds and final elongation at 70°C for 20 minutes.

8% polyacrylamide nondenaturing gel (Maniatis et al, 1982) was used for optimizing of reaction conditions and for analysis of One111 locus. The allele sizes were determined using Phoretix 1D Professional (Nonlinear Dynamics, United Kingdom) according to 20 bp Ladder (Bio-Rad).

The analysis of all the loci was performed on SpectruMedix SCE 9610 Genetic Analyzer capillary electrophoresis system. Each sample consisted of 1μl of PCR product, 15 μl of HiDi formamide and 0,5 μl of internal size standard. The latter was synthesized according to Symonds and Lloyd, 2004. We applied Genospectrum Software (SpectruMedix LLC) for allele size determination.

(B) Microsatellite assays generally followed the automated methods of Olsen et al. (1998) with the following modifications. One108, One109, One111, and One 114 (Olsen et al. 2000) were amplified in three multiplex PCR reactions using a MJ Research PTC-225 Thermocycler. One108 and One109 were multiplexed with One102 (Olsen et al. 2000) and uSat60 (Estoup et al. 1993). One111 was multiplexed with Omy77 (Morris et al. 1996), Ots107 (Nelsen and Beacham 1999) and Ots103 (Small et al. 1998). Finally, One114 was multiplexed with One103, and One112 (Olsen et al. 2000). Microsatellites were size fractionated on an Applied Biosystems 3730 capillary DNA sequencer. Alleles for each locus was visualized and separated into bin sets using ABI GeneMapper software v3.0. Data was tabulated for importing into statistical software according to the methods described in Olsen et al. (2003).

SNP analysis

Two SNP loci One_MHC2_190 and One_MHC2_251 were used in our study. The primers for this analysis were developed in Alaska Department of Fish and Game (unpublished data).

(A) Our analysis was carried out according to Taq-Man principles. As there is no real-time PCR machine at our disposal so we had to think out some alternative way of allelic discrimination detection. Thus a whole procedure differed a lot from the original one.

The first stage of optimization recommended by the Taq-Man protocol consists in determining the most suitable concentrations of forward and reverse primers. We

checked the quantity of PCR-products on 2% agarose gels and thus chose what primer concentration is better to use.

The oligonucleotide probes held fluorescent dyes **R6G** (yellow, 6-carboxy-rodamin) or **Cy5** (red, Cy5) linked to 5' end and non-fluorescent quencher on its 3'end.

The reaction for One_MHC2_190 was performed in a final volume of 15 μl containing about 60 ng of DNA; 1x PCR-buffer (Silex, Russia); 1,2 mM $MgCl_2$; 60 μM each dATP, dGTP, dCTP and dTTP; 300 nM of the PCR primers; 100 nM of R6G probe; 150 nM of Cy5 probe; 0,3 μl of Hot-Taq DNA polymerase (Silex, Russia).

The reaction for One_MHC2_251 was performed in a final volume of 15 μl containing about 60 ng of DNA; 1x PCR-buffer (Silex, Russia); 1,2 mM $MgCl_2$; 60 μM each dATP, dGTP, dCTP and dTTP; 400 nM of the PCR primers; 250 nM of R6G probe; 150 nM of Cy5 probe;0,3 μl of Hot-Taq DNA polymerase (Silex, Russia).

An initial denaturation and polymerase activation at 95°C for 10 minutes was followed by 40 cycles of: 92°C for 15 seconds and 60°C for 1 min. The fluorescence was detected using variable mode imager Typhoon 8600 (Molecular dynamics). For R6G detection we used 532 nm laser with Cy3 555 emission filter and for Cy5 – 635 nm laser with Cy5 emission filter. After scanning we performed spectral deconvolution between channels.

(B) Genotyping assays for each SNP were run on an ABI7900HT real-time sequence detection system in 384-well reaction plates, with two wells in each plate as negative controls (no-template), and two wells as positive controls (one for each allele). Each of the 384 wells contained 2.5μL 2X TaqMan PCR cocktail (ABI), 900nM of the PCR primers and 200nM of the probes, and 0.15μL template DNA in a total reaction volume of 5μL. An initial denaturation of 10 min at 95°C will be followed by 24 to 45 cycles of: 92°C for 15 sec and annealing / extension temperature for 1 min. All cycling was conducted using a ramp speed of 1°C per second. Allelic discrimination analysis was performed using Sequence Detection Software 2.0 (ABI).

RESULTS

Microsatellite analysis.

All of the loci from the two collections (130 samples) were analyzed in Anchorage as well in Moscow using capillary electrophoresis but the equipment and reagents were different. In addition, in Moscow the One111 locus was analyzed by use of non-denaturing polyacrylamide gel electrophoresis.

The allele sizes for One111 differed depending on the analysis platform, but the ABI and the SpectruMedix methods provided data that were consistently two base-pairs different at all aleles (Figure 1). The Page method, by contrast, provided some outliers and consistently more variation throughout the allele range. For the other three loci (One108, One109 and One114), the ABI and the SpectruMedix methods provided consistent deviations in the numbers of base-pair differences for each allele (Table 1).

We observed 18 cases (3%) where the alleles detected using the ABI and the SpectruMedix method did not follow the pattern in Table 1. Six of these cases appear to be the result of switching two fish – when these two fish were reversed, all the alleles followed the pattern described in Table 1. In five cases, the fish were scored heterozygotes with the ABI and homozygotes for one of these alleles with the SpectruMedix and in two cases the fish were scored heterozygotes with the SpectruMedix and homozygotes for one of these alleles with the ABI. These types of results are consistent with allelic dropout. One allele detected by the SpectruMedix for One109 was intermediate between the 188 and 196 allele (192), but was observed as a homozygote 196 using the ABI (allele adjusted to SpectruMedix). In the remaining four cases, the individuals were scored for different heterozygotes and may have been due to incorrect reading of capillary electrophoresis data occurring because of mirror bands.

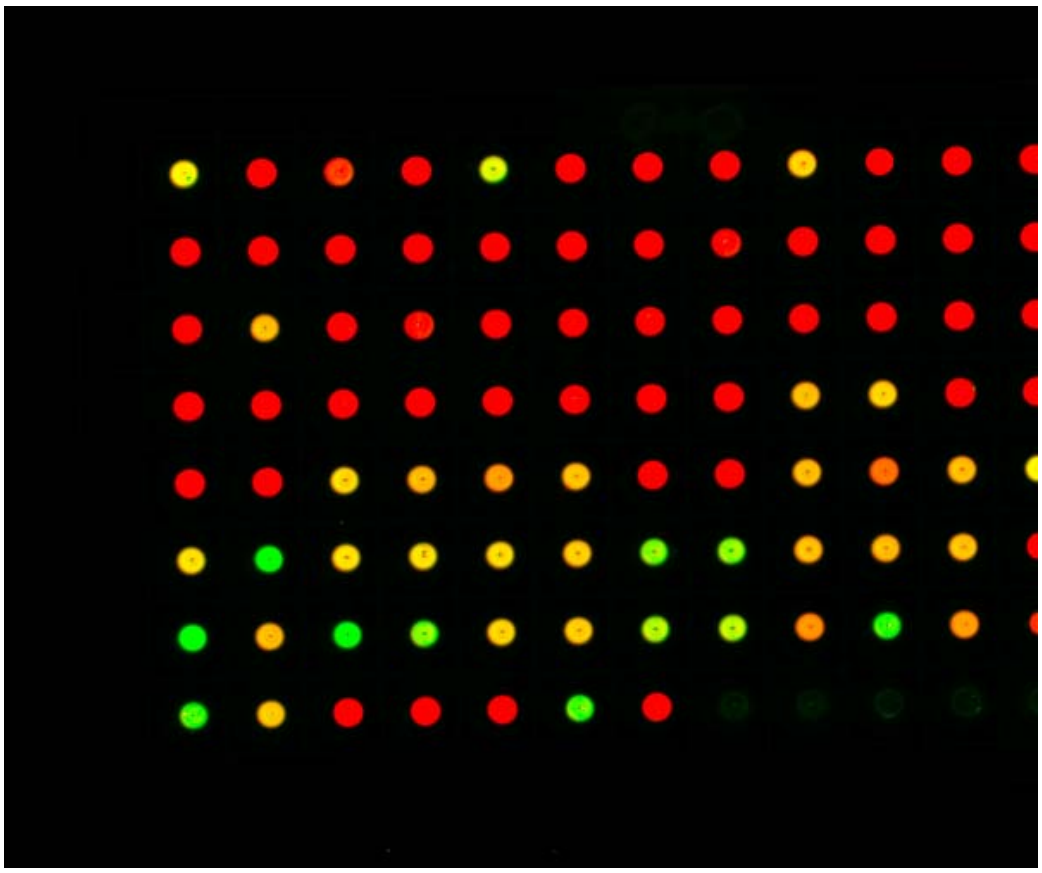
SNP analysis

The SNP detection methods used in Moscow differed from previously published methods in the salmonid literature. We therefore present a colored picture of a scanned plate to illustrate the results of this method.

DNA samples for One_MHC2_251 SNP-locus analysis were put into the plate for amplification in following order.

	1	2	3	4	5	6	7	8	9	10	11	12
a	KP 365	KP 366	KP 367	KP 368	KO 238	KP 370	KP 371	KP 371	KP 372	KP 373	KP 374	KP 375
b	KP 377	KP 378	KP 379	KP 380	KP 381	KP 382	KP 383	KP 383	KP 384	KP 385	KP 386	KP 387
c	KP 389	KP 390	KP 391	KP 392	KP 393	KP 394	KP 395	KP 395	KP 396	KP 397	KP 398	KP 399
d	KP 401	KP 402	KP 403	KP 404	KP 405	KP 406	KP 407	KP 407	KP 408	KP 409	KP 410	KP 411
e	KP 413	KP 414	KO 196	KO 197	KO 198	KO 199	KO 200	KO 200	KO 201	KO 202	KO 203	KO 204
f	KO 206	KO 207	KO 208	KO 209	KO 210	KO 211	KO 212	KO 212	KO 213	KO 214	KO 215	KO 216
g	KO 218	KO 219	KO 220	KO 221	KO 223	KO 224	KO 225	KO 225	KO 226	KO 227	KO 228	KO 229
h	KO 231	KO 232	KO 233	KO 235	KO 236	KO 237	KP 369					

After scanning the plate under the conditions mentioned above we obtained the following picture.



RED spot corresponds to homozygous T/T, GREEN spot - to homozygous C/C and the YELLOW one – to heterozygous T/C.

More accurate data were obtained by means of FluorSep 2.2 and Fragment Analysis 1.1 software. The fluorescence intensity was measured in RFU (relative fluorescence unit) and according to the ratio of intensities for both channels the genotype was determined.

The results for One_MHC2_190 SNP-locus were obtained in the same way.

American and Russian data are identical except for one individual where the American and Russian data showed opposite homozygotes at the One_MHC2_251-locus. This corresponds to 99.5% concordance between the two laboratories. An additional five collections from Russia were screened for the two SNP loci and allele frequencies are presented in Table 2.

DISCUSSION

The goal of this paper was not to make any investigation of Kamchatka sockeye salmon population structure but just to compare the results obtained in different laboratories on the same samples. We used in our study two methods being from our point of view the most suitable for solving the stock identification tasks – microsatellites and SNP.

All of the microsatellite data collected on the ABI and the SpectruMedix methods provided the basis for standardization between laboratories but only if sample exchange is done. The Page does not appear to be standardizable to either the ABI or the SpectruMedix method and appears to be unreliable for One111 (the only microsatellite tested).

In the case of One111, the consistent nature across all alleles in the difference in base-pair numbers between the ABI and the SpectruMedix methods, allows for exchange of a few individuals to standardize these data. However, data from the other three loci (One108, One109 and One114), from the ABI and the SpectruMedix methods demonstrates the need to exchange samples representing all alleles in order to standardize these loci between these two methods (Table 1). The need to have representatives of all alleles in the standardization is due to the variation in base-pair differences depending on which allele is being adjusted. In order to keep standardization up-to-date on loci that behave in this fashion, sample exchanges would have to occur every time a laboratory observes a new allele.

The SNP data, on the other hand is automatically standardized. The nucleotide determination depends neither on equipment nor on reagents used. The difficulty of sending samples internationally and the logistical problems of keeping up with standardization as new loci and alleles are identified with microsatellites makes the SNP approach very convenient for

performing independent analysis of pacific salmon samples collected in different countries as well as at sea.

LITERATURE

Banks, M. A., M. S. Blouin, B. A. Baldwin, V. K. Rashbrook, H. A. Fitzgerald, S. M. Blankenship and D. Hedgecock. 1999. Isolation and inheritance of novel microsatellites in Chinook salmon (*Oncorhynchus tschawytscha*). The Journal of Heridity 90: 281-288.

Beacham, T. D., B. McIntosh and C. MacConnachie. 2005. Population structure and stock identification of sockeye salmon (*Oncorhynchus nerka*) in coastal lakes in British Columbia, Canada. Canadian Journal of Zoology 83: 834-844.

Brykov, Vl. A., N. E. Polyakova and A. V. Podlesnykh. 2003. Divergence of mitochondrial DNA in populations of sockeye salmon *Oncorhynchus nerka* Walbaum from Azabach'e Lake (Kamchatka). Russian J. Genetics 39: 1432-1437.

Brykov, Vl. A., N. E. Polyakova, A.V. Podlesnykh, E. V. Golub', A.P.Golub' and O. L. Zhdanova. 2005. The effect of reproduction biotopes on the genetic differentiation of populations of sockeye salmon *Oncorhynchus nerka.* Russian J. Genetics 41: 509-517.

Estoup, A. P. Presa, F. Kreig, C. Vaiman and R. Guyomard. 1993. (CT)n and (GT)n microsatellites: a new class of genetic markers for *Salmo trutta* (brown trout). Heredity 71: 488-496.

Maniatis, T., E. E. Fritsch and J. Sambrook. 1982. Molecular Cloning, A laboratory manual. Cold Spring Harbor Laboratory.

Morris, D. B., K. R. Richard and J. M. Wright. 1996. Microsatellites from rainbow trout (*Oncorhynchus mykiss*) and their use for genetic study of salmonids. Canadian Journal of Fisheries and Aquatic Ecology 53: 120-126.

Nelson, R. J., and T. D. Beacham. 1999. Isolation and cross species amplification of microsatellite loci useful for study of Pacific salmon. J. Animal Genetics 30: 228-229.

Olsen, J. B., L. W. Seeb, P. Bentzen, and J. E. Seeb. 1998. Genetic interpretation of broad-scale microsatellite polymorphism in odd-year pink salmon. Trans. Am. Fish. Soc. 127(4): 535-550.

Olsen, J. B., S. L. Wilson, E. J. Kretschmer, K. C. Jones and J. E. Seeb. 2000. Characterization of 14 tetranucleotide microsatellite loci derived from sockeye salmon. Molecular Ecology 9: 2185-2187.

Olsen, J. B., C. Habicht, and J. E. Seeb. 2003. Moderately and highly polymorphic microsatellites provide discordant estimates of population divergence in sockeye salmon, *Oncorhynchus nerka*. Environmental Biology of Fishes (accepted and in press).

Rexroad III, C. E., R. L. Coleman, W. K. Hershberger and J. Killefer. 2002. Rapid communication: Thirty-eight polymorphic microsatellite markers for mapping in rainbow trout. J. Animal Sciences 80: 541-542.

Seeb, L. W., C. Habicht, W. D. Templin, K. E. Tarbox, R. Z. Davis, L. K. Brannian, and J. E. Seeb. 2000. Genetic diversity of sockeye salmon of Cook Inlet, Alaska, and its application to management of populations affected by the *Exxon Valdez* oil spill. Transactions of the American Fisheries Society 129: 1223-1249.

Small, M. P., T. D. Beacham, R. E. Withler and R. J. Nelson. 1998. Discriminating coho salmon (*Oncorhynchus kisutch*) populations within the Fraser River, British Columbia, using microsatellite DNA markers. Molecular Ecology 7:141-155.

Symonds, V. V. and A. M. Lloyd. 2004. A simple and inexpensive method for producing fluorescently labeled size standard. Molecular Ecology Notes: 4: 768-771.

Williamson, K. S., J. F. Cordes and B. May. 2002. Characterization of microsatellite loci in chinook salmon (*Oncorhyncus tshawytscha*) and cross-species amplification in other salmonids. Molecular Ecology Notes 2: 17-19.

Zelenina, D.A., A. M. Khrustaleva and A. A. Volkov. 2006. RAPD-PCR and microsattelite loci polymorphism comparative investigation of West Kamchatka sockeye salmon (*Oncorhynchus nerka*) population structure. Russian J. Genetics 42: in press.

Table 1. Alleles observed for three microsatellite loci at the Alaska Department of Fish and Game Gene Conservation Lab using an ABI 3730 (ABI) and the Division of Molecular Genetics of Aquatic Animals of VNIRO using a SpectruMedix (SpectruMedix). The differences in allele size between the two methods is noted in the "Δ" columns. Note the variation in Δ values both between and within loci. In the fourth locus screened by both laboratories, One111, ABI alleles were always two bases shorter than SpectruMedix alleles for all 200 fish assayed (see Figure 1).

Locus	Laboratory - Method		
	ABI	SpectruMedix	
One108			
	175	180	5
	179	184	5
	184	188	4
	192	196	4
	195	202	7
	199	206	7
	203	210	7
	208	214	6
	212	218	6
	215	222	7
	219	226	7
	223	230	7
	227	234	7
	231	238	7
One109			
	121	124	3
	125	128	3
	129	132	3
	133	136	3
	137	140	3
	141	144	3
	146	148	2
	150	152	2
	154	156	2
	159	160	1
	163	164	1
	167	168	1
	171	172	1
	175	176	1
	179	180	1
One114			

210	212	2
214	216	2
218	220	2
221	224	3
225	228	3
229	232	3
233	236	3
237	240	3
241	244	3
245	248	3
249	252	3
253	256	3
257	260	3
261	264	3
265	268	3
269	272	3
273	276	3
277	280	3
281	284	3
289	292	3

Table 2.

Allele frequencies for SNPs One_MHC2_190 and One_MHC2_251 for the three collections for sockeye salmon spawning aggregates from Russia screened in this project and from five collections provided by Nataly Varnavskaya (KamchatNIRO). Automatic standardization of SNP data across laboratories and methods allows for easy integration of data from multiple laboratories.

Population	Year	One_MHC2_190			One_MHC2_251		
		N	G	T	N	C	T
Palana River	2002	49	0.46	0.54	49	0.05	0.95
Bolshaya River	1998	23	0.76	0.24	26	0.62	0.38
Ozernaya River	2002	50	0.31	0.69	50	0.56	0.44
Kamchatka River - Late Run	1998	87	0.04	0.96	87	0.95	0.05
Kamchatka River - Early Run	1998	78	0.08	0.92	77	0.93	0.07
Hapiza River -Early	1998	86	0.98	0.02	78	0.91	0.09
Hapiza River - Late	1998	16	0.78	0.22	48	0.98	0.02

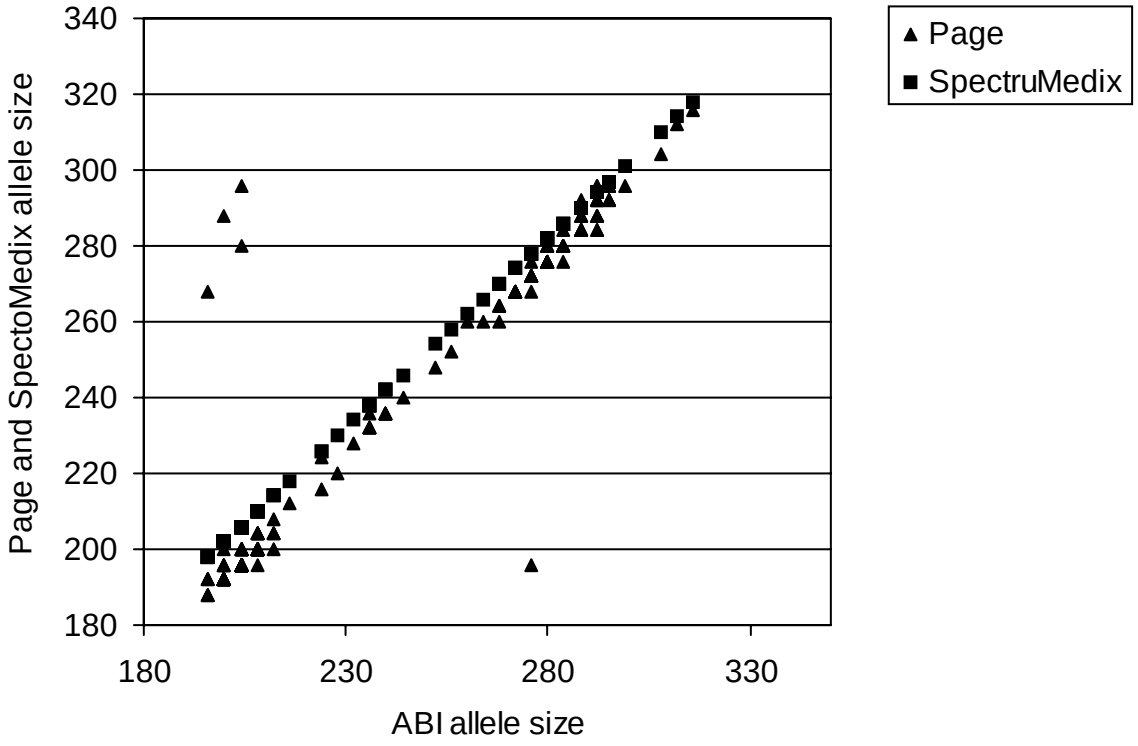


Figure 1. Allele sizes for One111 as obtained by the Page and the SpectruMedix methods (Y-axis) in relation to allele sizes detected using the ABI method (X-axis) on 200 fish from Palana and Ozernaya Rivers. Notice the correlation between the ABI and the SpectruMedix data, where the alleles are consistently two base-pairs larger using the SpectruMedix method. Page method provided some outliers and consistently more variation throughout the allele range making this method not standardizable among laboratories and unreliable for this microsatellite.