Summary results of a genetic-based investigation of cobia (*Rachycentron canadum*)

J. McDowell, H. Brightman, H. Small, and S. Musick

SEDAR58-SID-03

Submitted: 5 April 2018
Revised: 6 April 2018
Revised: 23 April 2018

Please cite this document as:

McDowell, J., H. Brightman, H. Small, S. Musick. 2018. Summary results of a genetic-based investigation of cobia (*Rachycentron canadum*). SEDAR58-SID-03. SEDAR, North Charleston, SC. 21 pp.

Summary of the results of a genetic-based investigation of cobia (*Rachycentrum canadum*)

SEDAR Working Paper SEDAR58-SID-03

Jan McDowell[1]
Heidi Brightman[1]
Hamish Small[2]
Susanna Musick[3]

Department of Fisheries Science[1]
Department of Aquatic Health Sciences[2]
Marine Advisory Services[3]
Virginia Institute of Marine Science

April 2018

MATERIAL AND METHODS

Microsatellite Marker Development and Optimization

High molecular weight DNA from a cobia captured off the coast of Virginia was used to create a 400 base pair (bp) insert genomic library. The resulting fragments were sequenced using a PGM™ Hi-Q™ Sequencing Kit on an Ion Torrent PGM sequencer using an Ion 318™ chip (Ion Torrent Systems, Inc., Guilford, CT). The FastQC software (Andrews 2010) integrated into the Galaxy Project platform (Giardine et al. 2005; Blankenberg et al. 2010; Goecks et al. 2010) was used to assess the quality of the resulting sequences, remove sequencing artifacts, and filter out sequences below 50 bp in length. The remaining sequences were trimmed to exclude positions 1-9 and all bases over 400 bp and then filtered by quality score to exclude those in which 50% of the sequence length had a quality < 20 (base call accuracy <99%). Exported sequence files were further filtered for the presence of perfect tetranucleotide repeats, resulting in the identification over 5,000 potential microsatellite loci using the Msatcommander 1.0.8 software (Faircloth 2008). Primers were designed for ~1000 of the identified loci using the Primer3 software program (Koressaar and Remm 2007; Untergasser et al. 2012) and 41 primer pairs were ordered and tested for amplification of a product of the predicted length from DNA samples isolated from cobia captured off Virginia and Mississippi.).

All primer pairs were initially assessed and optimized using gradient PCR on a Bio-Rad C1000 thermal cycler (Bio-Rad, Hercules, CA) using standard protocols. Each 5 µl PCR reaction contained 1x PCR Buffer (Qiagen), 1.5 mM $MgCl_2$, 200 µM of each dNTP, 0.125 µM forward and reverse primers, 0.5 unit of Taq polymerase (Qiagen), and 0.5 µl genomic DNA. Four cobia DNA samples, two from Virginia and two from Mississippi were used for initial testing. Samples were amplified with an initial denaturation temperature of 94 ºC for 3 min, followed by 35 cycles at 94 ºC for 1 min, 48-65 ºC for 1 min, 72 ºC for 1 min, with a final elongation step at 72 ºC for 7 min. Amplified products were visualized to confirm the presence of a single amplification product of correct size by agarose gel electrophoresis (1.5 % w/v), stained with ethidium bromide and viewed under a UV light source. Markers found to reliably amplify test DNA samples (see above) were further evaluated using an expanded panel of 8 cobia DNA samples each from Virginia and Mississippi to assess amplification consistency, levels of polymorphism, and conformance to the expectations of Hardy-Weinberg Equilibrium (HWE). PCR reactions were carried out as above except for the addition of a labeled fluorescent primer (either FAM, VIC, NED, PET). The resulting fluorescently labeled PCR products were separated on an ABI 3130xl Prism Genetic Analyzer (Applied Biosystems, Foster City, CA) with a GeneScan 500-Liz size standard (Applied Biosystems, Foster City, CA). The chromatic peaks for each microsatellite locus were sized using the GeneMarker AFLP/Genotyping Software, ver. 1.75 (SoftGenetics, State

College, PA). From the original 41 primer pairs bench tested, we identified 23 loci that that amplify a single product and display variability among samples. These primers have been multiplexed with an additional 5 primer pairs from a previous publication (Pruett et al. 2005).

Sample Collection and DNA Isolation

Geographic sampling ranged from the Rappahannock River, Virginia, to Louisiana (Figure 1). Fin clips were taken by members of the Virginia Game Fish Tagging Program, cooperating recreational fishermen or, in the case of Louisiana, by dock samplers and commercial spearfishers. All fin clips were immediately submerged in 95% ethanol for storage and shipment until DNA could be extracted. All available collection information (date, location, length, weight, fish sex, etc.) was recorded on the accompanying data sheets. DNA was extracted from fin clip tissue samples using either the DNeasy Tissue Kit (Qiagen, Valencia, CA) or the Quick-DNA™ Universal Kit (Zymo Research, Irvine, CA). Briefly, 2-3 mm fin clip sub-samples were incubated in lysis buffer (Longmire et al. 1997) for 2 hours at room temperature to facilitate removal of residual ethanol prior to extraction following the manufacturers protocol. All DNA samples were quantified using a NanoDrop 2000 (Thermo Scientific, West Palm Beach, Florida), and stored at -20 °C.

Microsatellite Markers

Following optimization, primer pairs were multiplexed into panels using the Type-it® Microsatellite PCR Kit (Qiagen) and one of seven unique fluorescently labeled primer. Once optimized, samples from each location were amplified using the multiplexed primer sets and alleles were sized as described above. Each multiplex reaction contained 1x Type-it Multiplex PCR Master Mix, 1x Q-Solution, 0.05 μM of the forward primer, 0.2 μM of the reverse primer, 0.2 μM of the fluorescent primer, 0.5 μl genomic DNA and water to a final volume of 6 μl. Amplifications were performed with an initial denaturation temperature of 95 ºC for 5 min, followed by 28 cycles at 95 ºC for 30 sec, annealing for 90 sec at 49 to 64 ºC (multiplex dependent), extension at 72 ºC for 30 sec, with a final elongation step at 60 ºC for 30 min. The resulting fluorescently labeled PCR products were separated on an ABI 3130xl Prism Genetic Analyzer (Applied Biosystems, Foster City, CA) with a GeneScan 500-Liz size standard (Applied Biosystems, Foster City, CA). The chromatic peaks for each microsatellite locus were sized using the GeneMarker AFLP/Genotyping Software, ver. 1.75 (SoftGenetics, State College, PA). All alleles were independently sized twice and the results were compared to check for errors.

After alleles had been sized for each locus, the Micro-Checker 2.2.3 software (Van

Oosterhout et al. 2004) was used to check for the presence of null alleles and evidence of scoring errors. The Genepop'007 software package (Rousset 2008) was used to test for deviations of genotypic distributions from HWE expectations (F_{IS}, exact tests, Guo *et al.* 1992). Summary statistics including number of alleles, allele frequencies and observed ad expected heterozygosity were generated using GenAlEx 6.5 (Peakall and Smouse 2012). To evaluate evidence for the presence of population structure, the Arlequin software package v3.5.2.2 (Excoffier and Lischer 2010) was used to calculate Weir & Cockerham's (1984) unbiased estimator of Wrights F-statistics and to conduct an analysis of molecular variance (AMOVA, Excoffier *et al.* 1992) based on alternate geographic groupings. Significance was assessed via 10,000 permutations of the data. A principal components analysis (PCA) was performed using the Adegenet software (Jombart 2008; Jombart et al. 2010). To test data formed distinct genetic groups, clustering of the genotype data was evaluated using the Bayesian modeling-based algorithm implemented in Structure v2.3.4 (Pritchard et al. 2000, Falush, Stephens, & Pritchard, 2007; Hubisz, Falush, Stephens, & Pritchard, 2009, Hubisz et al. 2009). Data was grouped into Gulf of Mexico (GOM), Virgina and North Carolina (Mid-A) and Florida (EFL). Since the F_{ST} values among the three samples were low, sampling location was used as a prior. Structure simulations were performed using an admixture model of ancestry, correlated allele frequencies, K=2 and a burn-in of 100,000 followed by 500,000 Markov chain Monte Carlo iterations. Additional iterations with a higher K values were not informative and are not shown.

Mitochondrial DNA (mtDNA)

Primers spanning the partial cytochrome *b* (cyt*b*), complete tRNA-Thr and tRNA-Pro genes, and partial control region (Khongchatee, A. and Phinchongsakuldit, J., unpublished data, available in GENBANK) were used to amplify and sequence a 648 bp mitochondrial region of mitochondrial DNA (mtDNA). The FaBox software (Villesen 2007) was used to collapse sequences into haplotypes and create input files for the Arlequin software package (Excoffier and Lischer 2010). Arlequin was used to generate descriptive statistics including mean number of pairwise differences (k), haplotype diversity (H), and nucleotide sequence diversity (π) and to perform analysis of population pairwise Φ_{ST}, AMOVA (Excoffier et al. 1992). Statistical significance was assessed based on 10,000 permutations of the data. The PopART software (Leigh and Bryant 2015) was used to reconstruct and visualize the genealogical relationships among sequences using the Minimum Spanning Network algorithm (Bandelt et al. 1999).

RESULTS AND DISCUSSION

Microsatellite Analysis

In total, 427 cobia DNA samples were analyzed across 28 polymorphic microsatellite loci. Of this total 310 samples were taken from cobia captured in Virginia and North Carolina waters. Samples included fish captured from the mouth of the Rappahannock River (RRVA; n=29), the Eastern Shore of Virginia off Cape Charles (ESVA; n=65), the mouth of the York River (YRVA; n=104), lower Chesapeake Bay (LBVA; n=112) and off Oregon Inlet and Avon, North Carolina (OINC; n=8). Additional Atlantic samples were taken off Stuart, Florida (EFL; n=14). In the Gulf of Mexico, samples were taken at two time points representing a Gulf temporal replicate. Cobia were sampled off Biloxi, Mississippi in 2003 (GMMS; n=48) and off Cocodrie and Venice, Louisiana in 2017 (GMLA; n=47) (Figures 1a, 1b). All loci were polymorphic, with the number of alleles across all samples ranging from 3 at locus *Rca* 1B-A10 to 25 at locus *Rca* 1B-D10. Markers were in conformance to the expectations of Hardy-Weinberg equilibrium (HWE) with the exception of *Rca* 1B-F06, which had significant global heterozygote deficit across samples, most likely due to the presence of null alleles. All subsequent analyses were conducted excluding this locus. Summary statistics are in Table 1. Samples were initially analyzed holding all geographic sample collections separate to look for evidence of genetic structure within geographic sampling regions and to decide whether samples from different locations within a region could be combined (Table 2). There were no significant differences between sample collections from the mid-Atlantic region (Virginia and North Carolina); F_{ST} values ranged from 0-0.006. Within the Gulf of Mexico region, the F_{ST} calculated between Mississippi and Louisiana was 0 consistent with being sampled from a single genetic population. The GMLA sample was significantly different from 4 of 5 mid-Atlantic collections and GMMS was significantly different from 3 of 5 mid-Atlantic collections ($P< 0.001$). The sample from Florida was not significantly different from either Gulf of Mexico sample, however it was significantly different from RRVA and OINC, the two smallest mid-Atlantic collections at $P< 0.05$. A PCA revealed a slight separation between mid-Atlantic and Gulf of Mexico samples (Figure 2).

Based on the results of the pairwise comparisons, samples were combined into mid-Atlantic and Gulf of Mexico collections. The small eastern Florida sample was held separately due to the uncertainty of placement. The FST value for the Gulf of Mexico vs. mid-Atlantic comparison was 0.0055, $P < 0.0001$. The comparison of the EFL collection with both Gulf and mid-Atlantic samples was non-significant; each resulted in an F_{ST} of 0.0023 ($P > 0.05$). Consistent with the F_{ST} values, the Structure analysis clustered the GOM and Mid-A into separate groups. The EFL samples taken off Stuart, FL (n=14) had a higher level of shared ancestry with the GOM sample (Figure 3). These results are consistent with the presence of genetically distinct populations of cobia in the Gulf and

mid-Atlantic regions.

mtDNA analysis

In total, 161 cobia mitochondrial DNA sequences were examined across a subset of samples from all geographic locations. All sequences were edited to a final length of 648 bp, resulting in 31 unique haplotypes with 23 variable sites including 16 parsimony informative sites. The most common haplotype, haplotype 1, was recovered 71 times (44% of sequences), and was recovered in all locations (Table 3). The second most common haplotype, haplotype 9 was recovered 19 times (11.8% of all samples) and was recovered in all locations except North Carolina and Florida, which had the smallest number of samples sequenced (7 and 14 respectively). Haplotype diversity (H) was moderately high in all geographic samples ranging from 0.64 in samples from Gulf of Mexico to 0.88 in Eastern Shore, Virginia. The mean number of pairwise differences between sequences (k) ranged from 0.86 in Oregon Inlet, North Carolina, to 9.12 in Eastern Shore, Virginia. Nucleotide diversity (π) ranged from 0.014 in Eastern Shore, Virginia to 0.001 in Oregon Inlet, North Carolina (Table 4). A minimum spanning network indicated the presence of haplotypes unique to the mid-Atlantic and Gulf (Figure 4). A global test of differentiation among samples based on the distribution of haplotypes and 10, 000 permutations of the data was significant ($P < 0.0001$) whether holding geographic sampling locations separately or when sampling locations were combined into Gulf of Mexico, east Florida, and mid-Atlantic regions. The population pairwise FST values between Gulf and mid-Atlantic region was 0.061 ($P= 0.012$). These results are consistent with the presence of genetically distinct populations of cobia in the Gulf and mid-Atlantic regions.

REFERENCES

Bandelt, H.-J., and A. R. Peter Forster, and ¨hl. 1999. Median-Joining Networks for Inferring Intraspeci c Phylogenies. Molecular Biology 16(1):37–48.

Blankenberg, D., G. Von Kuster, N. Coraor, G. Ananda, R. Lazarus, M. Mangan, A. Nekrutenko, and J. Taylor. 2010. Galaxy: A web-based genome analysis tool for experimentalists.

Excoffier, L., and H. E. L. Lischer. 2010. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. Molecular Ecology Resources 10(3):564–567.

Excoffier, L., P. E. Smouse, and J. M. Quattro. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. Genetics 131(2):479–491.

Faircloth, B. C. 2008. MSATCOMMANDER: Detection of microsatellite repeat arrays and automated, locus-specific primer design. Molecular Ecology Resources 8(1):92–94.

Falush, D., M. Stephens, and J. K. Pritchard. 2007. Inference of population structure using multilocus genotype data: Dominant markers and null alleles. Molecular Ecology Notes 7(4):574–578.

Giardine, B., C. Riemer, R. C. Hardison, R. Burhans, L. Elnitski, P. Shah, Y. Zhang, D. Blankenberg, I. Albert, J. Taylor, W. Miller, W. J. Kent, and A. Nekrutenko. 2005. Galaxy: A platform for interactive large-scale genome analysis. Genome Research 15(10):1451–1455.

Goecks, J., A. Nekrutenko, and J. Taylor. 2010. Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. Genome biology 11(8):R86.

Guo, S. W., S. W. Guo, E. a Thompson, and E. a Thompson. 1992. Performing the exact test of Hardy-Weinberg proportion for multiple alleles. Biometrics 48(2):361–372.

Hubisz, M. J., D. Falush, M. Stephens, and J. K. Pritchard. 2009. Inferring weak population structure with the assistance of sample group information. Molecular Ecology Resources 9(5):1322–1332.

Koressaar, T., and M. Remm. 2007. Enhancements and modifications of primer design program Primer3. Bioinformatics 23(10):1289–1291.

Leigh, J. W., and D. Bryant. 2015. Popart: Full-Feature Software for Haplotype Network Construction. Methods in Ecology and Evolution 6(9):1110–1116.

Van Oosterhout, C., W. F. Hutchinson, D. P. M. Wills, and P. Shipley. 2004. MICRO-CHECKER: Software for identifying and correcting genotyping errors in microsatellite data. Molecular Ecology Notes 4(3):535–538.

Peakall, R., and P. E. Smouse. 2012. GenALEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research-an update. Bioinformatics 28(19):2537–2539.

Pritchard, J. K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. Genetics 155(2):945–959.

Pruett, C. L., E. Saillant, M. A. Renshaw, J. C. Patton, C. E. Rexroad, and J. R. Gold. 2005. Microsatellite DNA markers for population genetic studies and parentage assignment in cobia, Rachycentron canadum. Molecular Ecology Notes 5(1):84–86.

Rousset, F. 2008. GENEPOP'007: A complete re-implementation of the GENEPOP software for Windows and Linux. Molecular Ecology Resources 8(1):103–106.
Untergasser, A., I. Cutcutache, T. Koressaar, J. Ye, B. C. Faircloth, M. Remm, and S. G. Rozen. 2012. Primer3-new capabilities and interfaces. Nucleic Acids Research 40(15).
Villesen, P. 2007. FaBox: An online toolbox for FASTA sequences. Molecular Ecology Notes 7(6):965–968.
Weir, B. S., and C. C. Cockerham. 1984. Estimating F-statistics for the analysis of population structure. Evolution 38(6):1358–1370.

Table 1. Sample Size (N), No. Alleles (N_a), No. Effective Alleles (N_e), Information Index (I), Observed Heterozygosity (H_o), Expected Heterozygosity (H_e), Probability of Conformance to HWE (P_{HWE}), and Unbiased Expected Heterozygosity (uH_e), and Fixation Index (F) for each locus. Rappahannock River (RRVA), the Eastern Shore of Virginia off Cape Charles (ESVA), the mouth of the York River (YRVA), lower Chesapeake Bay (LBVA), Oregon Inlet, North Carolina (OINC), Mississippi (GMMS), Louisiana (GMLA), eastern Florida (EFL).

Pop		A10	C06	D10	H09	RC3888	RC2868	RC1129	RC1685	RC3970
LBVA	**N**	111	111	112	111	111	108	112	110	110
	Na	4	19	22	15	12	10	13	5	8
	Ne	1.541	10.540	14.732	8.813	6.528	3.633	6.927	1.240	5.531
	I	0.664	2.537	2.828	2.377	2.064	1.602	2.130	0.436	1.824
	Ho	0.333	0.919	0.902	0.919	0.829	0.750	0.830	0.191	0.809
	He	0.351	0.905	0.932	0.887	0.847	0.725	0.856	0.193	0.819
	uHe	0.353	0.909	0.936	0.891	0.851	0.728	0.859	0.194	0.823
	F	0.050	-0.015	0.033	-0.037	0.021	-0.035	0.030	0.012	0.012
ESVA	**N**	61	65	65	65	64	64	65	64	65
	Na	4	18	21	15	12	11	12	5	9
	Ne	1.826	9.165	14.594	8.102	5.535	4.450	7.484	1.194	5.338
	I	0.800	2.416	2.827	2.300	1.944	1.804	2.200	0.402	1.823
	Ho	0.508	0.908	0.969	0.877	0.859	0.781	0.815	0.172	0.815
	He	0.452	0.891	0.931	0.877	0.819	0.775	0.866	0.162	0.813
	uHe	0.456	0.898	0.939	0.883	0.826	0.781	0.873	0.164	0.819
	F	-0.123	-0.019	-0.041	0.000	-0.049	-0.008	0.059	-0.059	-0.003
RRVA	**N**	29	29	28	29	29	29	29	29	29
	Na	3	13	17	13	8	9	11	2	8
	Ne	1.275	8.806	11.615	8.626	6.050	4.014	5.741	1.109	5.374
	I	0.418	2.343	2.630	2.284	1.896	1.663	2.003	0.204	1.810
	Ho	0.241	0.862	0.893	0.931	0.862	0.724	0.931	0.103	0.897
	He	0.216	0.886	0.914	0.884	0.835	0.751	0.826	0.098	0.814
	uHe	0.220	0.902	0.931	0.900	0.849	0.764	0.840	0.100	0.828
	F	-0.118	0.027	0.023	-0.053	-0.033	0.036	-0.127	-0.055	-0.102
YRVA	**N**	100	104	104	103	103	103	104	104	104
	Na	4	17	22	15	14	9	11	5	10
	Ne	1.448	10.052	13.436	8.457	6.283	3.508	6.662	1.340	5.424
	I	0.612	2.465	2.817	2.307	2.043	1.518	2.096	0.579	1.897
	Ho	0.300	0.952	0.923	0.883	0.893	0.767	0.827	0.269	0.856
	He	0.309	0.901	0.926	0.882	0.841	0.715	0.850	0.254	0.816
	uHe	0.311	0.905	0.930	0.886	0.845	0.718	0.854	0.255	0.820
	F	0.031	-0.057	0.003	-0.002	-0.062	-0.073	0.027	-0.062	-0.049
OINC	**N**	8	8	8	8	8	8	8	8	8
	Na	2	7	10	5	6	7	6	1	4
	Ne	1.438	5.818	8.533	4.267	4.923	2.415	4.000	1.000	2.844
	I	0.483	1.841	2.220	1.511	1.684	1.333	1.548	0.000	1.180

	Ho	0.375	0.875	1.000	0.875	0.875	0.625	0.750	0.000	1.000
	He	0.305	0.828	0.883	0.766	0.797	0.586	0.750	0.000	0.648
	uHe	0.325	0.883	0.942	0.817	0.850	0.625	0.800	0.000	0.692
	F	-0.231	-0.057	-0.133	-0.143	-0.098	-0.067	0.000	#N/A	-0.542
EFL	**N**	14	14	14	14	14	14	14	14	14
	Na	4	12	15	12	8	7	8	3	7
	Ne	1.347	8.711	11.529	9.333	5.521	5.521	6.222	1.244	3.733
	I	0.559	2.317	2.571	2.345	1.883	1.822	1.927	0.409	1.516
	Ho	0.286	0.929	1.000	1.000	0.929	0.857	0.929	0.071	0.714
	He	0.258	0.885	0.913	0.893	0.819	0.819	0.839	0.196	0.732
	uHe	0.267	0.918	0.947	0.926	0.849	0.849	0.870	0.204	0.759
	F	-0.109	-0.049	-0.095	-0.120	-0.134	-0.047	-0.106	0.636	0.024
GMLA	**N**	47	47	47	47	47	47	47	47	47
	Na	3	18	23	14	11	10	12	6	10
	Ne	1.599	11.657	16.609	10.620	6.375	5.222	7.889	1.465	6.536
	I	0.600	2.629	2.936	2.452	2.028	1.873	2.189	0.701	2.029
	Ho	0.404	0.915	1.000	0.894	0.787	0.872	0.894	0.298	0.894
	He	0.375	0.914	0.940	0.906	0.843	0.809	0.873	0.317	0.847
	uHe	0.379	0.924	0.950	0.916	0.852	0.817	0.883	0.321	0.856
	F	-0.079	-0.001	-0.064	0.013	0.066	-0.079	-0.023	0.061	-0.055
GMMS	**N**	48	48	47	48	47	40	48	47	47
	Na	3	16	17	15	11	11	9	2	12
	Ne	1.488	10.332	14.252	9.952	6.384	3.587	6.857	1.235	5.297
	I	0.584	2.469	2.727	2.439	2.030	1.668	2.044	0.339	1.945
	Ho	0.250	0.917	0.936	1.000	0.894	0.725	0.875	0.213	0.936
	He	0.328	0.903	0.930	0.900	0.843	0.721	0.854	0.190	0.811
	uHe	0.331	0.913	0.940	0.909	0.852	0.730	0.863	0.192	0.820
	F	0.238	-0.015	-0.007	-0.112	-0.060	-0.005	-0.024	-0.119	-0.154

Pop		RC4656	RC4878	RC5046	RC1559	RC3318	RC1830	RC3767	RC4502	RC4713
LBVA	**N**	112	111	112	111	111	111	112	112	112
	Na	12	13	6	6	6	9	6	14	16
	Ne	3.750	7.617	1.941	3.913	1.301	5.306	2.036	8.837	10.480
	I	1.620	2.196	0.923	1.412	0.537	1.818	1.009	2.316	2.469
	Ho	0.768	0.865	0.491	0.730	0.243	0.811	0.527	0.821	0.902
	He	0.733	0.869	0.485	0.744	0.231	0.812	0.509	0.887	0.905
	uHe	0.737	0.873	0.487	0.748	0.232	0.815	0.511	0.891	0.909
	F	-0.047	0.004	-0.013	0.020	-0.053	0.001	-0.035	0.074	0.003
ESVA	**N**	65	65	65	65	65	65	65	65	65
	Na	10	12	6	4	6	8	6	12	18
	Ne	4.057	7.216	1.726	3.110	1.404	4.528	2.110	8.009	8.720
	I	1.622	2.114	0.871	1.252	0.644	1.689	1.028	2.198	2.443
	Ho	0.800	0.877	0.415	0.662	0.277	0.831	0.477	0.831	0.908
	He	0.753	0.861	0.421	0.678	0.288	0.779	0.526	0.875	0.885
	uHe	0.759	0.868	0.424	0.684	0.290	0.785	0.530	0.882	0.892
	F	-0.062	-0.018	0.012	0.025	0.037	-0.066	0.094	0.051	-0.025
RRVA	**N**	29	29	29	29	29	29	29	29	29
	Na	7	9	3	4	4	7	4	11	13
	Ne	4.698	5.644	1.482	3.426	1.494	4.369	1.664	9.141	8.715
	I	1.717	1.908	0.616	1.307	0.688	1.626	0.769	2.281	2.333
	Ho	0.862	0.759	0.345	0.552	0.345	0.724	0.414	0.897	0.862
	He	0.787	0.823	0.325	0.708	0.331	0.771	0.399	0.891	0.885
	uHe	0.801	0.837	0.331	0.721	0.336	0.785	0.406	0.906	0.901
	F	-0.095	0.078	-0.060	0.221	-0.043	0.061	-0.037	-0.007	0.026
YRVA	**N**	104	104	104	104	104	104	104	104	104
	Na	9	15	6	6	6	9	7	12	16
	Ne	4.375	6.644	1.761	3.705	1.282	4.894	1.890	8.632	9.336
	I	1.679	2.100	0.879	1.411	0.510	1.770	0.942	2.235	2.386
	Ho	0.769	0.856	0.394	0.731	0.221	0.740	0.500	0.846	0.923
	He	0.771	0.849	0.432	0.730	0.220	0.796	0.471	0.884	0.893
	uHe	0.775	0.854	0.434	0.734	0.221	0.800	0.473	0.888	0.897
	F	0.003	-0.007	0.088	-0.001	-0.005	0.069	-0.062	0.043	-0.034
OINC	**N**	8	8	8	8	8	8	8	8	8
	Na	5	7	3	4	3	5	3	10	8
	Ne	4.000	4.741	1.662	3.459	1.471	4.000	2.133	7.529	6.737
	I	1.494	1.754	0.703	1.305	0.602	1.474	0.900	2.166	1.981
	Ho	0.625	0.750	0.500	0.875	0.375	0.750	0.500	1.000	1.000
	He	0.750	0.789	0.398	0.711	0.320	0.750	0.531	0.867	0.852
	uHe	0.800	0.842	0.425	0.758	0.342	0.800	0.567	0.925	0.908
	F	0.167	0.050	-0.255	-0.231	-0.171	0.000	0.059	-0.153	-0.174
EFL	**N**	14	14	14	14	14	14	14	14	14
	Na	4	10	4	6	3	7	4	10	12
	Ne	2.780	6.125	1.574	4.356	1.338	5.297	1.574	7.127	10.316
	I	1.198	2.038	0.736	1.567	0.490	1.782	0.736	2.115	2.397

	Ho	0.714	0.929	0.429	0.786	0.143	0.929	0.429	0.857	1.000
	He	0.640	0.837	0.365	0.770	0.253	0.811	0.365	0.860	0.903
	uHe	0.664	0.868	0.378	0.799	0.262	0.841	0.378	0.892	0.937
	F	-0.116	-0.110	-0.175	-0.020	0.434	-0.145	-0.175	0.003	-0.107
GMLA	**N**	47	47	47	47	47	47	47	47	47
	Na	10	13	6	5	7	9	6	12	16
	Ne	4.705	5.883	1.995	2.765	1.544	4.005	2.421	7.875	11.475
	I	1.774	2.041	1.005	1.192	0.806	1.648	1.163	2.216	2.589
	Ho	0.830	0.830	0.553	0.638	0.362	0.830	0.574	0.809	0.957
	He	0.787	0.830	0.499	0.638	0.352	0.750	0.587	0.873	0.913
	uHe	0.796	0.839	0.504	0.645	0.356	0.758	0.593	0.882	0.923
	F	-0.054	0.000	-0.109	0.000	-0.027	-0.106	0.021	0.074	-0.049
GMMS	**N**	46	47	47	46	47	47	48	47	48
	Na	10	11	4	7	7	8	5	12	17
	Ne	4.666	6.565	1.644	3.574	1.584	4.327	2.496	8.399	9.309
	I	1.788	2.044	0.777	1.445	0.843	1.660	1.126	2.248	2.463
	Ho	0.674	0.872	0.362	0.696	0.426	0.851	0.625	0.787	0.833
	He	0.786	0.848	0.392	0.720	0.368	0.769	0.599	0.881	0.893
	uHe	0.794	0.857	0.396	0.728	0.372	0.777	0.606	0.890	0.902
	F	0.142	-0.029	0.077	0.034	-0.155	-0.107	-0.043	0.106	0.066

Pop		RC4507	RC5036	RC2112	RC3212	RC849	RC3450	RC4840	RC4426	RC781
LBVA	**N**	111	112	111	110	111	111	111	111	111
	Na	15	7	15	6	5	3	12	7	4
	Ne	7.648	3.737	7.594	1.565	1.725	1.135	2.214	3.576	1.372
	I	2.317	1.507	2.262	0.709	0.753	0.261	1.323	1.478	0.548
	Ho	0.838	0.679	0.901	0.336	0.495	0.126	0.550	0.712	0.234
	He	0.869	0.732	0.868	0.361	0.420	0.119	0.548	0.720	0.271
	uHe	0.873	0.736	0.872	0.363	0.422	0.120	0.551	0.724	0.273
	F	0.036	0.074	-0.038	0.068	-0.179	-0.059	-0.002	0.012	0.137
ESVA	**N**	65	65	65	65	65	65	65	65	65
	Na	14	7	13	5	4	3	11	7	3
	Ne	7.204	3.844	6.771	1.712	1.462	1.150	2.401	3.759	1.263
	I	2.247	1.532	2.149	0.793	0.554	0.288	1.444	1.514	0.415
	Ho	0.769	0.708	0.846	0.415	0.323	0.138	0.600	0.769	0.215
	He	0.861	0.740	0.852	0.416	0.316	0.131	0.583	0.734	0.208
	uHe	0.868	0.746	0.859	0.419	0.319	0.132	0.588	0.740	0.210
	F	0.107	0.044	0.007	0.001	-0.022	-0.061	-0.028	-0.048	-0.034
RRVA	**N**	29	29	29	29	29	29	29	29	29
	Na	11	6	12	4	5	2	7	6	4
	Ne	5.408	3.867	6.865	1.857	1.660	1.035	2.431	3.617	1.236
	I	1.944	1.508	2.162	0.767	0.801	0.087	1.245	1.484	0.422
	Ho	0.690	0.759	0.828	0.552	0.448	0.034	0.655	0.621	0.207
	He	0.815	0.741	0.854	0.461	0.398	0.034	0.589	0.724	0.191
	uHe	0.829	0.754	0.869	0.469	0.405	0.034	0.599	0.736	0.194
	F	0.154	-0.023	0.031	-0.196	-0.127	-0.018	-0.113	0.142	-0.084
YRVA	**N**	104	104	104	104	104	104	104	104	104
	Na	17	7	13	5	6	3	11	7	5
	Ne	6.869	3.968	8.611	1.660	1.540	1.081	1.769	3.433	1.459
	I	2.272	1.547	2.321	0.763	0.686	0.188	1.091	1.420	0.637
	Ho	0.798	0.798	0.837	0.442	0.317	0.077	0.433	0.692	0.298
	He	0.854	0.748	0.884	0.398	0.351	0.075	0.435	0.709	0.315
	uHe	0.859	0.752	0.888	0.400	0.352	0.075	0.437	0.712	0.316
	F	0.066	-0.067	0.054	-0.112	0.095	-0.030	0.005	0.023	0.053
OINC	**N**	8	8	7	8	8	8	8	8	8
	Na	6	4	8	3	2	1	7	5	2
	Ne	4.267	2.844	5.765	1.293	1.969	1.000	3.879	2.327	1.133
	I	1.576	1.180	1.909	0.463	0.685	0.000	1.629	1.160	0.234
	Ho	0.875	0.500	1.000	0.250	0.625	0.000	0.875	0.625	0.125
	He	0.766	0.648	0.827	0.227	0.492	0.000	0.742	0.570	0.117
	uHe	0.817	0.692	0.890	0.242	0.525	0.000	0.792	0.608	0.125
	F	-0.143	0.229	-0.210	-0.103	-0.270	#N/A	-0.179	-0.096	-0.067
EFL	**N**	14	14	14	14	14	14	14	14	14
	Na	11	4	10	4	2	2	6	6	2
	Ne	6.323	2.562	7.127	2.190	1.324	1.237	2.596	3.806	1.324
	I	2.107	1.089	2.112	0.968	0.410	0.340	1.265	1.530	0.410

	Ho	0.786	0.714	0.929	0.714	0.286	0.214	0.500	0.643	0.143
	He	0.842	0.610	0.860	0.543	0.245	0.191	0.615	0.737	0.245
	uHe	0.873	0.632	0.892	0.563	0.254	0.198	0.638	0.765	0.254
	F	0.067	-0.172	-0.080	-0.315	-0.167	-0.120	0.187	0.128	0.417
GMLA	**N**	47	47	47	47	47	47	47	47	47
	Na	14	7	14	6	2	3	10	7	3
	Ne	5.962	3.215	8.447	2.205	1.559	1.137	2.008	3.848	1.506
	I	2.164	1.401	2.329	0.992	0.544	0.266	1.182	1.527	0.607
	Ho	0.851	0.745	0.894	0.489	0.298	0.128	0.596	0.766	0.340
	He	0.832	0.689	0.882	0.546	0.359	0.121	0.502	0.740	0.336
	uHe	0.841	0.696	0.891	0.552	0.362	0.122	0.507	0.748	0.340
	F	-0.023	-0.081	-0.014	0.104	0.169	-0.058	-0.187	-0.035	-0.013
GMMS	**N**	48	48	48	48	47	48	48	47	47
	Na	14	8	13	6	4	3	7	8	5
	Ne	5.346	3.746	8.948	2.134	1.501	1.284	2.647	4.083	1.457
	I	2.064	1.504	2.352	0.997	0.612	0.413	1.323	1.630	0.655
	Ho	0.813	0.646	0.896	0.542	0.383	0.208	0.625	0.766	0.362
	He	0.813	0.733	0.888	0.531	0.334	0.221	0.622	0.755	0.314
	uHe	0.821	0.741	0.898	0.537	0.337	0.223	0.629	0.763	0.317
	F	0.001	0.119	-0.009	-0.019	-0.147	0.058	-0.005	-0.014	-0.153

Table 2. Population pairwise FST values based 27 microsatellite loci for cobia (*Rachycentron canadum*) for all geographic samples within a region held separately. Bolded values were significant at P<0.001, underlined values were significant at P<0.05 based on 10,000 permutations of the data. Rappahannock River (RRVA), the Eastern Shore of Virginia off Cape Charles (ESVA), the mouth of the York River (YRVA), lower Chesapeake Bay (LBVA), off Oregon Inlet, North Carolina (OINC), Mississippi (GMMS), Louisiana (GMLA), eastern Florida (EFL).

	LBVA	ESVA	RRVA	YRVA	OINC	EFL	GMLA	GMMS
LBVA	0.0000							
ESVA	0.0008	0.0000						
RRVA	0.0008	0.0002	0.0000					
YRVA	-0.0010	0.0007	-0.0008	0.0000				
OINC	0.0006	0.0040	0.0060	0.0037	0.0000			
EFL	0.0011	0.0051	-0.0024	0.0036	0.0113	0.0000		
GMLA	**0.0055**	**0.0047**	**0.0047**	*0.0058*	**0.0131**	0.0031	0.0000	
GMMS	**0.0052**	**0.0051**	0.0022	**0.0051**	0.0053	0.0008	-0.0013	0.0000

Table 3. Distribution of mtDNA haplotypes. Rappahannock River (RRVA), the Eastern Shore of Virginia off Cape Charles (ESVA), the mouth of the York River (YRVA), lower Chesapeake Bay (LBVA), Oregon Inlet, North Carolina (OINC), Gulf of Mexico (GOM), eastern Florida (EFL).

	RRVA	YRVA	LBVA	ESVA	OINC	EFL	GOM
Hap_1	7	18	10	4	4	5	23
Hap_2	3	6	1	5	0	0	0
Hap_3	1	0	0	0	0	0	0
Hap_4	1	4	1	1	0	3	1
Hap_5	1	2	2	0	0	0	0
Hap_6	1	0	1	0	0	0	0
Hap_7	1	1	0	0	0	0	0
Hap_8	1	0	0	1	0	0	0
Hap_9	3	5	4	2	0	0	5
Hap_10	1	1	0	1	0	0	0
Hap_11	0	0	1	0	0	0	0
Hap_12	0	0	1	0	0	1	0
Hap_13	0	3	1	1	1	0	0
Hap_14	0	0	1	0	0	0	0
Hap_15	0	1	0	0	0	0	0
Hap_16	0	0	0	1	0	0	0
Hap_17	0	0	0	1	0	0	0
Hap_18	0	0	0	0	1	1	1
Hap_19	0	0	0	0	1	0	0
Hap_20	0	0	0	0	0	1	1
Hap_21	0	0	0	0	0	1	0
Hap_22	0	0	0	0	0	1	0
Hap_23	0	0	0	0	0	1	0
Hap_24	0	0	0	0	0	0	1
Hap_25	0	0	0	0	0	0	1
Hap_26	0	0	0	0	0	0	1
Hap_27	0	0	0	0	0	0	1
Hap_28	0	0	0	0	0	0	1
Hap_29	0	0	0	0	0	0	1
Hap_30	0	0	0	0	0	0	1
Hap_31	0	0	0	0	0	0	1

Table 4. Haplotype diversity (H), Mean number of pairwise differences (K) and nucleotide diversity (π) for mtDNA sequences. Rappahannock River (RRVA), the Eastern Shore of Virginia off Cape Charles (ESVA), the mouth of the York River (YRVA), lower Chesapeake Bay (LBVA), Oregon Inlet, North Carolina (OINC), Gulf of Mexico, (GOM), eastern Florida (EFL).

	H	K	π
RRVA	0.8579 +/- 0.0623	5.847368 +/- 2.916797	0.009024 +/- 0.005027
YRVA	0.7707 +/- 0.0551	5.320732 +/- 2.621638	0.008211 +/- 0.004494
ESVA	0.8750 +/- 0.0576	9.125000 +/- 4.417649	0.014082 +/- 0.007630
LBVA	0.7945 +/- 0.0766	2.924901 +/- 1.591142	0.004514 +/- 0.002738
OINC	0.7143 +/- 0.1809	0.857143 +/- 0.681989	0.001329 +/- 0.001211
EFL	0.8571 +/- 0.0774	1.373626 +/- 0.899284	0.002130 +/- 0.001565
GOM	0.6451 +/- 0.0851	1.379217 +/- 0.867887	0.002138 +/- 0.001495

Table 3. Population pairwise FST values based on sequencing of mitochondrial DNA (mtDNA) in cobia (*Rachycentron canadum*) with sample collections within region held separately. Bolded values were significant at P<0.001, underlined values were significant at P<0.05 based on 10,000 permutations of the data. Rappahannock River (RRVA), the Eastern Shore of Virginia off Cape Charles (ESVA), the mouth of the York River (YRVA), lower Chesapeake Bay (LBVA), Oregon Inlet, North Carolina (OINC), Gulf of Mexico (GOM), eastern Florida (EFL).

	RRVA	LBVA	YRVA	ESVA	OINC	EFL	GOM
RRVA	0.0000						
LBVA	0.0070	0.0000					
YRVA	-0.0355	0.0123	0.0000				
ESVA	0.0309	<u>0.1832</u>	0.0566	0.0000			
OINC	0.0195	-0.0168	0.0180	0.1745	0.0000		
EFL	0.0602	<u>0.0617</u>	0.0610	<u>0.2287</u>	0.0066	0.0000	
GOM	<u>0.0902</u>	-0.0062	<u>0.0776</u>	**0.3261**	-0.0051	<u>0.0922</u>	0.0000

Figure 1a. Sampling locations and numbers of cobia (*Rachycentron canadum*) samples used in this study. Samples were collected off Virginia (VA, n=310) and North Carolina (OINC; n=8). Additional Atlantic samples were taken off Stuart, Florida (EFL; n=14). In the Gulf of Mexico (GOM, n=95 total), samples were taken off Biloxi MS (n=48) and Cocodrie and Venice, Louisiana (n=47).

Figure 1b. Sampling locations and numbers of cobia (*Rachycentron canadum*) samples collected Virginia waters include: Rappahannock River (RRVA; n=29), the Eastern Shore of Virginia off Cape Charles (ESVA; n=65), the mouth of the York River (YRVA; n=104), lower Chesapeake Bay (LBVA; n=112).

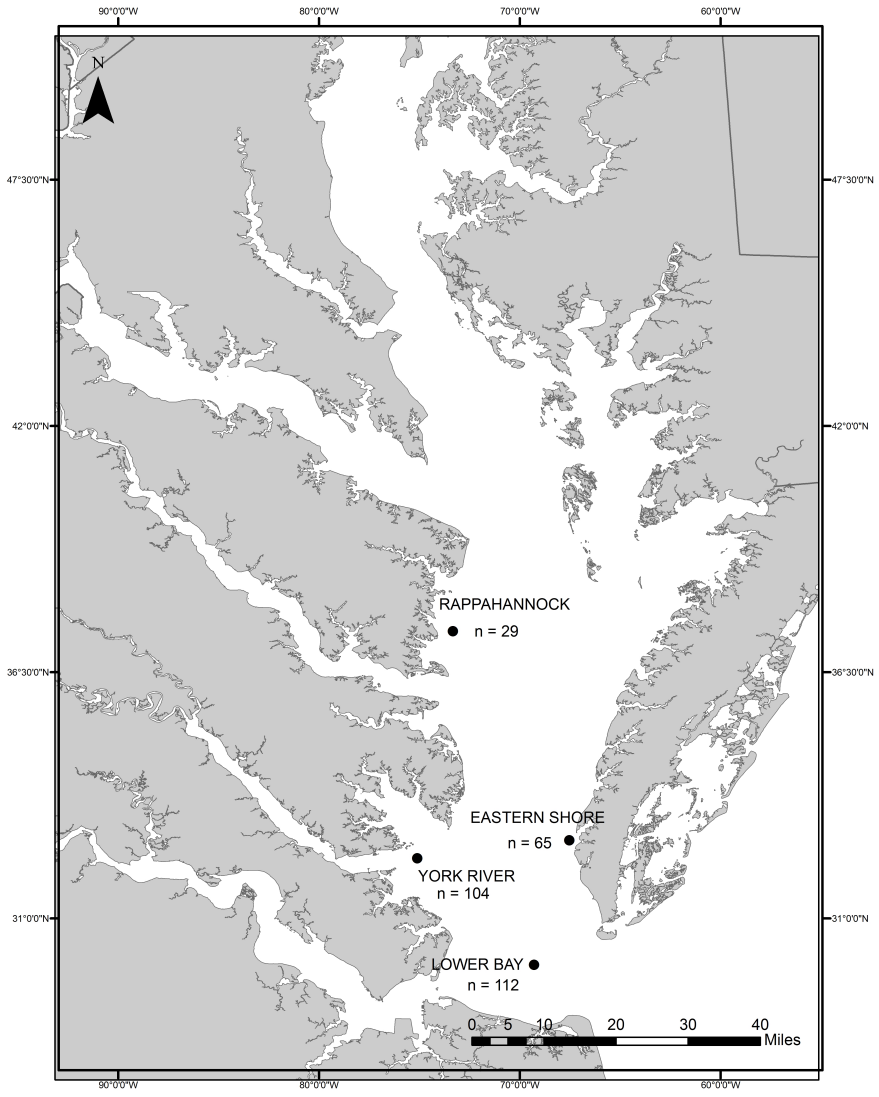

Figure 2. Principal components analysis based on microsatellite data. Mid-Atlantic (mid-Atl, red dots) DNA samples comprised: Rappahannock River, Virginia (RRVA; n=29), the Eastern Shore of Virginia off Cape Charles (ESVA; n=65), the mouth of the York River, Virginia (YRVA; n=104), lower Chesapeake Bay, Virginia (LBVA; n=112), and offOregon Inlet, North Carolina (OINC; n=8). Gulf of Mexico (GOM, blue squares) comprised samples taken off Biloxi, Mississippi (GMMS; n=48) and off Cocodrie and Venice, Louisiana (GMLA; n=47). Eastern Florida (EFL, green triangles) represents samples taken off Stuart, Florida (n=14).

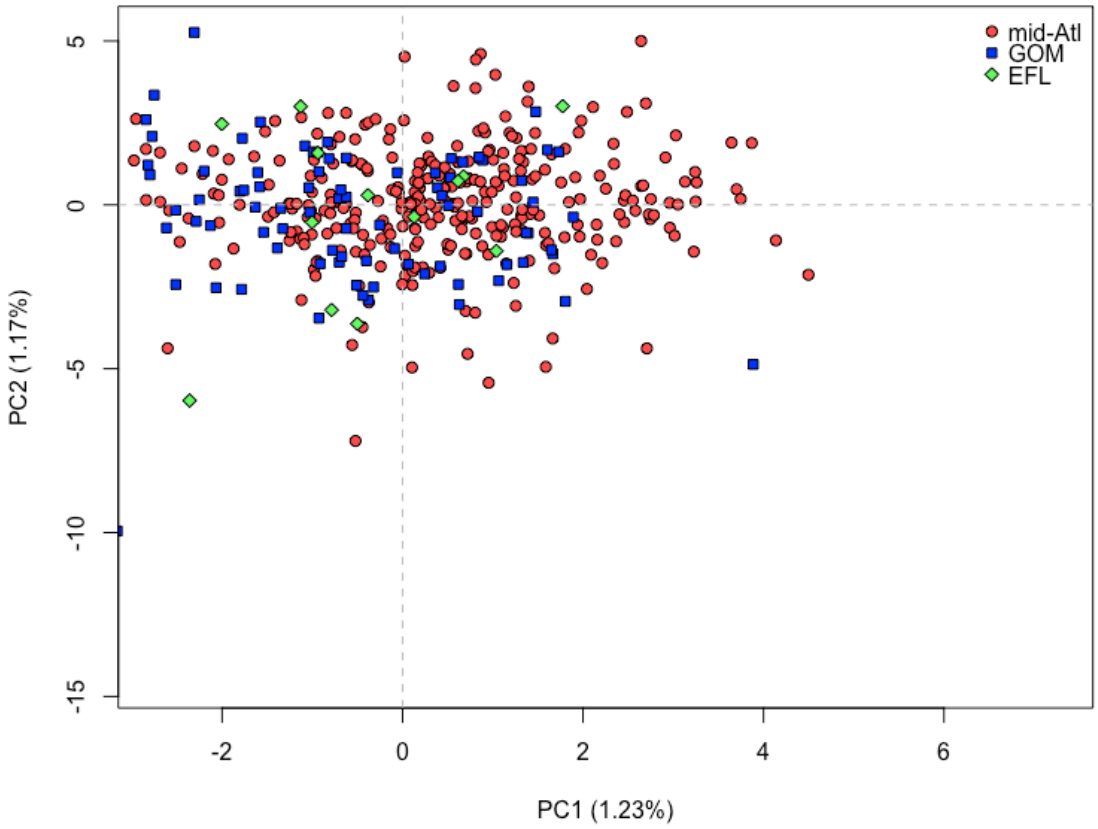

Figure 3. Barplots generated from Structure results based on microsatellite genotype data. Gulf of Mexico (GOM), Virgina and North Carolina (Mid-A) and Florida (EFL) samples.

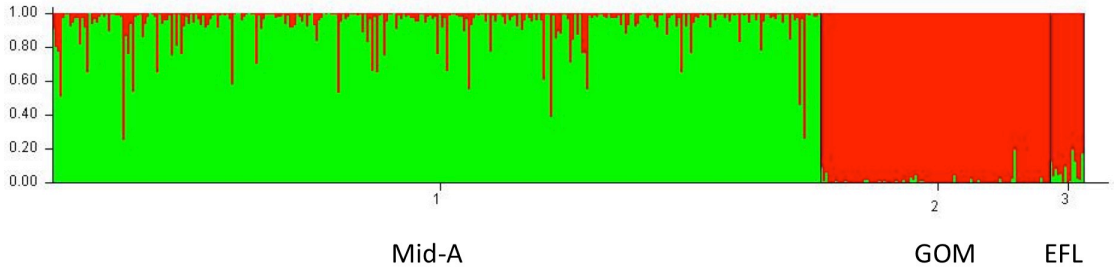

Figure 4. Minimum spanning network of the relationship among mtDNA haplotypes.