SUPPLEMENTARY INFORMATION for:

The population genomic structure of green turtles (*Chelonia mydas*) suggests a warm-water corridor for tropical marine fauna between the Atlantic and Indian oceans during the last interglacial

Jurjan P. van der Zee[1], Marjolijn J.A. Christianen[1,2], Martine Bérubé[1,3], Mabel Nava[4], Kaj Schut[4], Frances Humber[5], Alonzo Alfaro-Núñez[6], Leontine E. Becking[7,8], Per J. Palsbøll[1,3]

[1]Marine Evolution and Conservation, Groningen Institute for Evolutionary Life Sciences, University of Groningen, Nijenborg 7 9747 AG Groningen, The Netherlands
[2]Aquatic Ecology and Water Quality Management Group, Wageningen University & Research, P.O. Box 47, 6700 AA Wageningen, The Netherlands
[3]Center for Coastal Studies, 5 Holway Avenue, Provincetown, MA 02657, United States
[4]Sea Turtle Conservation Bonaire, P.O. Box 492, Kaya Korona 53 Kralendijk, Bonaire, The Caribbean Netherlands
[5]BlueVentures, Omnibus Business Centre, 41 North Road, London N7 9DP, United Kingdom
[6]Section for Evolutionary Genomics, Department of Biology, University of Copenhagen, Øster Farimagsgade 5, 1353 Copenhagen K, Denmark
[7]Wageningen Marine Research, Ankerpark 27 1781 AG Den Helder, The Netherlands
[8]Marine Animal Ecology Group, Wageningen University & Research, P.O. Box 338, 6700 AH Wageningen, The Netherlands

Corresponding author: Jurjan P. van der Zee (jurjanvanderzee@gmail.com)

Supplementary Table S1. Barcodes and adapters, total number of reads, reads without RAD-tags, low quality reads, the number of retained reads, and alignment rate per sample. 'BO' samples were collected in the Caribbean (CA), 'MG' samples in the Southwest Indian Ocean (SWO) and 'PI' samples in the East Atlantic (EA). Sample PI070009 aligned poorly (62.04%) and was excluded from subsequent analyses.

Samples	*Library preparation*		*Summary of read statistics*				*Alignment*
ID	Adapter-P1	Adapter-P2	Total reads	No RAD-tag	Low quality	Retained	Rate
BO150013	P1_01_GCATG_HindIII	P2_05_ACAGTG_MspI	10,968,858	118,292	4,895	10,845,671	95.03%
BO150077	P1_04_AAGTGA_HindIII	P2_05_ACAGTG_MspI	10,884,640	86,276	7,775	10,790,589	95.56%
BO150087	P1_05_ATTACA_HindIII	P2_05_ACAGTG_MspI	9,831,072	69,014	5,053	9,757,005	95.51%
BO150098	P1_07_AGAATGA_HindIII	P2_05_ACAGTG_MspI	14,811,236	113,624	9,893	14,687,719	95.22%
BO150106	P1_08_AGTTAAT_HindIII	P2_05_ACAGTG_MspI	15,539,356	81,416	7,636	15,450,304	94.94%
BO150125	P1_11_AGTGTTAA_HindIII	P2_05_ACAGTG_MspI	16,674,886	126,409	10,704	16,537,773	95.29%
BO150141	P1_07_AGAATGA_HindIII	P2_19_GTGAAA_MspI	12,587,234	89,705	72,255	12,425,274	95.65%
BO150161	P1_08_AGTTAAT_HindIII	P2_19_GTGAAA_MspI	22,739,154	118,525	56,528	22,564,101	96.05%
BO150164	P1_12_CACGACCA_HindIII	P2_05_ACAGTG_MspI	19,222,898	131,370	17,448	19,074,080	95.31%
BO150181	P1_11_AGTGTTAA_HindIII	P2_19_GTGAAA_MspI	15,461,152	114,085	68,323	15,278,744	95.72%
BO150189	P1_12_CACGACCA_HindIII	P2_19_GTGAAA_MspI	12,952,436	155,270	148,139	12,649,027	95.76%
BO150191	P1_01_GCATG_HindIII	P2_06_GCCAAT_MspI	21,271,650	164,182	7,114	21,100,354	95.92%
BO150212	P1_04_AAGTGA_HindIII	P2_06_GCCAAT_MspI	7,429,198	62,016	3,099	7,364,083	95.40%
MG060001	P1_05_ATTACA_HindIII	P2_12_CTTGTA_MspI	14,424,202	95,991	6,450	14,321,761	95.68%
MG060028	P1_07_AGAATGA_HindIII	P2_12_CTTGTA_MspI	9,489,532	69,357	4,316	9,415,859	94.14%
MG060030	P1_08_AGTTAAT_HindIII	P2_12_CTTGTA_MspI	14,554,756	80,933	6,138	14,467,685	95.63%
MG070026	P1_11_AGTGTTAA_HindIII	P2_12_CTTGTA_MspI	10,021,962	82,896	4,931	9,934,135	95.07%
MG070047	P1_12_CACGACCA_HindIII	P2_12_CTTGTA_MspI	7,809,722	57,184	6,281	7,746,257	94.68%
MG070052	P1_01_GCATG_HindIII	P2_19_GTGAAA_MspI	9,918,092	112,122	42,653	9,763,317	95.53%
MG070061	P1_04_AAGTGA_HindIII	P2_19_GTGAAA_MspI	8,281,428	65,868	33,994	8,181,566	94.01%
MG070069	P1_05_ATTACA_HindIII	P2_19_GTGAAA_MspI	6,671,990	47,819	46,094	6,578,077	94.88%
PI070001	P1_05_ATTACA_HindIII	P2_06_GCCAAT_MspI	23,376,394	170,206	9,279	23,196,909	95.32%
PI070002	P1_07_AGAATGA_HindIII	P2_06_GCCAAT_MspI	13,925,760	125,090	6,056	13,794,614	95.08%
PI070005	P1_08_AGTTAAT_HindIII	P2_06_GCCAAT_MspI	15,443,374	84,313	5,963	15,353,098	95.94%
PI070007	P1_11_AGTGTTAA_HindIII	P2_06_GCCAAT_MspI	9,314,390	79,482	4,022	9,230,886	95.65%
PI070008	P1_12_CACGACCA_HindIII	P2_06_GCCAAT_MspI	16,802,588	129,319	13,332	16,659,937	95.48%
PI070009	P1_01_GCATG_HindIII	P2_12_CTTGTA_MspI	12,684,248	119,312	5,389	12,559,547	62.04%
PI070011	P1_04_AAGTGA_HindIII	P2_12_CTTGTA_MspI	16,150,026	116,951	6,911	16,026,164	95.64%

Supplementary Table S2. Model parameter prior distributions and values (mean, minimum, maximum and delta; Δ) used in the MIGRATE-N analysis.

Parameter	Distribution	Mean	Minimum	Maximum	Δ
Θ	Uniform	0.005	0.00	0.01	0.001
M	Uniform	2500	0.00	5000	500
τ	Uniform	0.05	0.00	0.1	0.01
τ_σ	Uniform	0.05	0.00	0.1	0.01

Supplementary Table S3. Total number of SNPs across locations, SNPs per location, private alleles and nucleotide diversity (π) per region (A) prior to additional filtering and (B) after excluding SNPs with a mean depth across samples below 30x or above 300x and (C) additionally thinning SNPs within 100,000bp windows.

	Location	SNPs (total)	SNPs (per location)	Private alleles	π
A	CA	156,439	99,984	46,359	0.133
	EA	156,439	54,696	9,681	0.110
	SWO	156,439	83,634	42,579	0.130
B	CA	95,213	61,959	25,559	0.147
	EA	95,213	35,929	5,138	0.123
	SWO	95,213	54,404	25,588	0.145
C	CA	12,035	7,733	3,385	0.140
	EA	12,035	4,391	699	0.118
	SWO	12,035	6,643	3,310	0.138

Supplementary Table S4. The amount of missing data per sample for the data prior to additional filtering (*'Standard'*; i.e. standard STACKS output; Supplementary Figure S1), after excluding SNPs with mean depth across samples below 30x or above 300x (*'Depth-filtered'*) and additionally thinning SNPs within 100,000 bp windows (*'Unlinked'*).

Samples	Missing data		
	Standard	*Depth-filtered*	*Unlinked*
BO150013	2.6%	1.4%	1.5%
BO150077	2.5%	1.7%	1.8%
BO150087	3.4%	1.6%	1.6%
BO150098	1.5%	1.2%	1.4%
BO150106	2.1%	2.1%	1.6%
BO150125	1.7%	1.8%	1.7%
BO150141	3.3%	1.5%	1.6%
BO150164	1.0%	0.9%	0.9%
BO150161	1.8%	1.8%	1.7%
BO150181	2.2%	1.6%	1.6%
BO150189	2.4%	1.4%	1.5%
BO150191	1.5%	1.2%	1.1%
BO150212	6.5%	2.4%	2.0%
PI070001	1.0%	0.8%	0.9%
PI070002	1.7%	1.1%	1.1%
PI070005	1.7%	1.3%	1.0%
PI070007	3.6%	1.4%	1.4%
PI070008	0.9%	0.7%	0.9%
PI070011	1.8%	1.5%	1.6%
MG060001	0.8%	0.8%	1.0%
MG060028	2.1%	1.8%	2.2%
MG060030	1.4%	1.5%	1.3%
MG070026	2.4%	2.2%	2.5%
MG070047	1.6%	0.8%	1.0%
MG070052	3.7%	1.9%	2.1%
MG070061	5.5%	2.2%	2.1%
MG070069	5.4%	2.0%	2.4%

Supplementary Table S5. Pairwise genetic differentiation among the Caribbean (CA), East Atlantic (EA) and Southwest Indian Ocean (SWO) estimated using Hudson's F_{ST}, and Weir and Cockerham's θ, and the number of 'privately shared alleles' (i.e., alleles only found in the two pairwise compared locations).

Location 1	Location 2	Hudson's F_{ST}	Weir & Cockerham's θ	Shared alleles
CA	EA	0.110	0.096	1305
CA	SWO	0.159	0.153	946
EA	SWO	0.180	0.170	293

Supplementary Table S6. The mean number of SNPs and the mean, standard deviation (SD), lower (2.5%) and upper boundaries (97.5%) of the 95% confidence interval (CI) of estimates of Hudson's F_{ST} as a function of sample size (n = 1,000 simulated datasets per sample size).

Sample size	SNPs	Mean	SD	CI (2.5%)	CI (97.5%)
2	2049.0	0.091	0.009	0.073	0.109
3	2359.8	0.091	0.007	0.078	0.105
4	2576.6	0.091	0.006	0.080	0.102
5	2733.8	0.091	0.005	0.081	0.101
6	2866.6	0.091	0.005	0.082	0.100
7	2976.7	0.091	0.004	0.083	0.099
8	3067.9	0.091	0.004	0.083	0.099
9	3148.7	0.091	0.004	0.083	0.099
10	3221.6	0.091	0.004	0.084	0.098
15	3498.9	0.091	0.003	0.085	0.098
20	3689.4	0.091	0.003	0.085	0.097
25	3836.8	0.091	0.003	0.085	0.097
50	4279.1	0.091	0.003	0.085	0.097
75	4536.2	0.091	0.003	0.086	0.096
100	4711.9	0.091	0.003	0.086	0.096

Supplementary Table S7. The mean number of SNPs and the mean, standard deviation (SD), lower (2.5%) and upper boundaries (97.5%) of the 95% confidence interval (CI) of estimates of Weir and Cockerham's θ as a function of sample size (n = 1,000 simulated datasets per sample size).

Sample size	SNPs	Mean	SD	CI (2.5%)	CI (97.5%)
2	2049.0	0.091	0.010	0.071	0.110
3	2359.8	0.091	0.007	0.078	0.106
4	2576.6	0.091	0.006	0.080	0.103
5	2733.8	0.091	0.005	0.082	0.101
6	2866.6	0.091	0.005	0.082	0.100
7	2976.7	0.091	0.004	0.083	0.099
8	3067.9	0.091	0.004	0.083	0.099
9	3148.7	0.091	0.004	0.083	0.099
10	3221.6	0.091	0.004	0.083	0.098
15	3498.9	0.091	0.003	0.085	0.098
20	3689.4	0.091	0.003	0.085	0.097
25	3836.8	0.091	0.003	0.085	0.097
50	4279.1	0.091	0.003	0.085	0.097
75	4536.2	0.091	0.003	0.086	0.096
100	4711.9	0.091	0.003	0.086	0.096

Supplementary Table S8. Effective sample sizes (ESS) per parameter for each model per replicate run for a random sub-sample of 5,000 RAD loci.

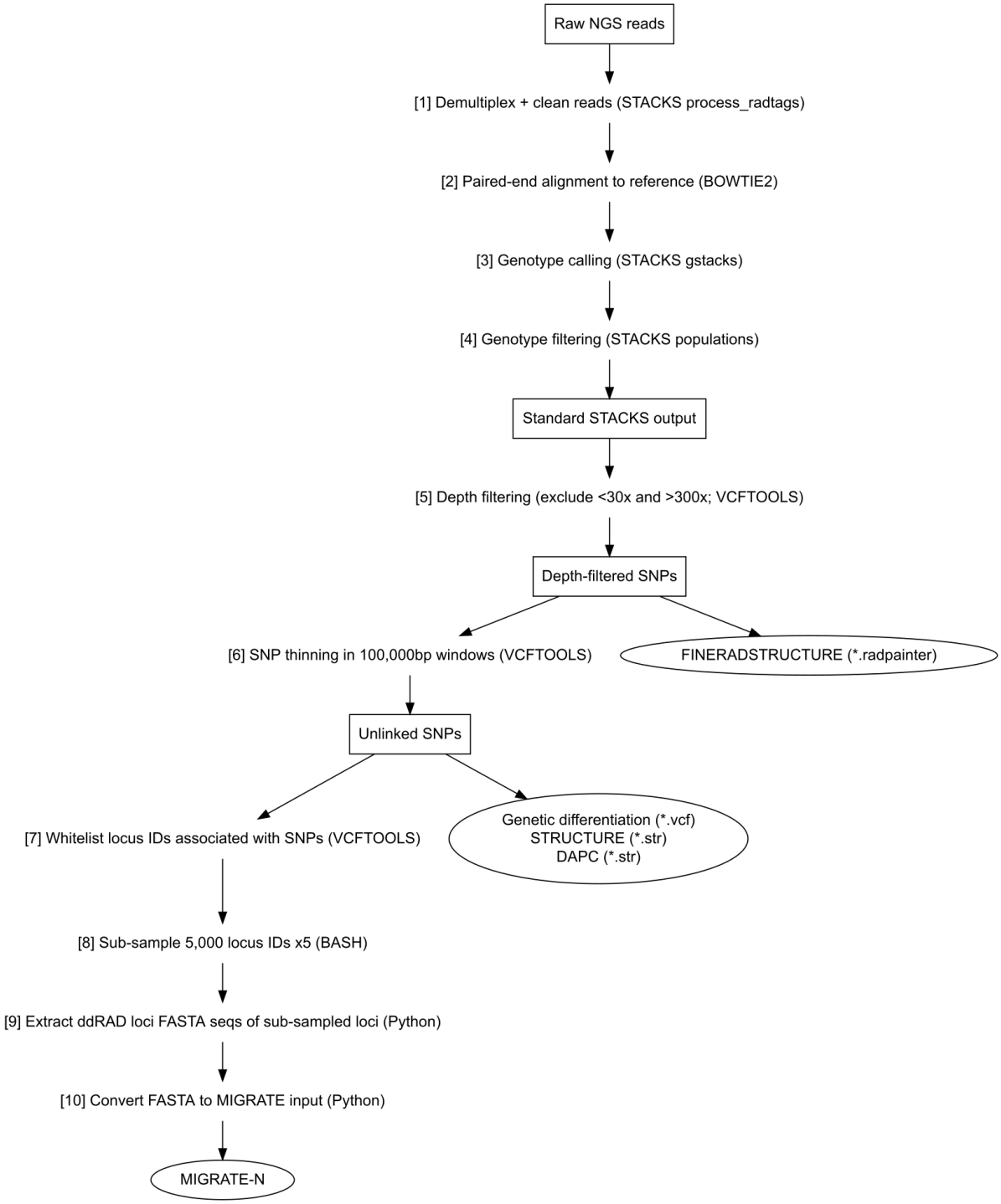
Model	Parameter	Sub-sample 1	Sub-sample 2	Sub-sample 3	Sub-sample 4	Sub-sample 5
1	θ_{SWO}	4.45E+09	4.47E+09	4.45E+09	4.47E+09	4.48E+09
1	θ_{CA}	4.34E+09	4.34E+09	4.36E+09	4.33E+09	4.34E+09
1	θ_{EA}	4.16E+09	4.16E+09	4.17E+09	4.17E+09	4.16E+09
1	$M_{CA->SWO}$	2.19E+09	2.19E+09	2.19E+09	2.20E+09	2.19E+09
1	$M_{EA->SWO}$	2.18E+09	2.18E+09	2.18E+09	2.19E+09	2.19E+09
1	$M_{SWO->CA}$	2.07E+09	2.06E+09	2.07E+09	2.07E+09	2.07E+09
1	$M_{EA->CA}$	2.00E+09	1.99E+09	2.00E+09	2.00E+09	2.00E+09
1	$M_{SWO->EA}$	2.19E+09	2.18E+09	2.19E+09	2.19E+09	2.19E+09
1	$M_{CA->EA}$	2.15E+09	2.15E+09	2.15E+09	2.16E+09	2.16E+09
1	Genealogies	3.47E+09	3.45E+09	3.45E+09	3.45E+09	3.45E+09
2	θ_{SWO}	4.66E+09	4.70E+09	4.67E+09	4.69E+09	4.73E+09
2	θ_{CA}	4.60E+09	4.60E+09	4.62E+09	4.59E+09	4.62E+09
2	θ_{EA}	3.54E+09	3.54E+09	3.55E+09	3.54E+09	3.57E+09
2	$M_{CA->SWO}$	2.46E+09	2.47E+09	2.46E+09	2.47E+09	2.49E+09
2	$M_{EA->SWO}$	2.35E+09	2.36E+09	2.35E+09	2.36E+09	2.38E+09
2	$M_{SWO->CA}$	2.40E+09	2.39E+09	2.40E+09	2.40E+09	2.42E+09
2	$M_{EA->CA}$	2.23E+09	2.23E+09	2.23E+09	2.23E+09	2.26E+09
2	$M_{SWO->EA}$	2.19E+09	2.19E+09	2.19E+09	2.20E+09	2.21E+09
2	$M_{CA->EA}$	2.22E+09	2.20E+09	2.21E+09	2.21E+09	2.23E+09
2	$\Delta_{EA->SWO}$	6.86E+09	6.86E+09	6.86E+09	6.86E+09	6.90E+09
2	$\Delta_{EA->CA}$	6.67E+09	6.67E+09	6.67E+09	6.66E+09	6.71E+09
2	Genealogies	3.49E+09	3.49E+09	3.49E+09	3.49E+09	3.51E+09
3	θ_{SWO}	3.68E+09	3.70E+09	3.78E+09	3.69E+09	3.70E+09
3	θ_{CA}	4.72E+09	4.72E+09	4.89E+09	4.71E+09	4.71E+09
3	θ_{EA}	4.46E+09	4.45E+09	4.61E+09	4.46E+09	4.46E+09
3	$M_{CA->SWO}$	2.19E+09	2.18E+09	2.26E+09	2.19E+09	2.19E+09
3	$M_{EA->SWO}$	2.10E+09	2.10E+09	2.17E+09	2.11E+09	2.11E+09
3	$M_{SWO->CA}$	2.43E+09	2.43E+09	2.52E+09	2.44E+09	2.44E+09
3	$M_{EA->CA}$	2.42E+09	2.42E+09	2.50E+09	2.42E+09	2.43E+09
3	$M_{SWO->EA}$	2.37E+09	2.36E+09	2.45E+09	2.37E+09	2.37E+09
3	$M_{CA->EA}$	2.47E+09	2.46E+09	2.55E+09	2.48E+09	2.48E+09
3	$\Delta_{EA->CA}$	6.84E+09	6.84E+09	7.05E+09	6.83E+09	6.83E+09
3	$\Delta_{SWO->EA}$	6.94E+09	6.96E+09	7.15E+09	6.94E+09	6.94E+09
3	Genealogies	3.53E+09	3.53E+09	3.64E+09	3.53E+09	3.53E+09
4	θ_{SWO}	4.74E+09	4.76E+09	4.73E+09	4.77E+09	4.77E+09
4	θ_{CA}	3.77E+09	3.77E+09	3.79E+09	3.77E+09	3.77E+09
4	θ_{EA}	4.49E+09	4.51E+09	4.51E+09	4.51E+09	4.51E+09
4	$M_{CA->SWO}$	2.42E+09	2.41E+09	2.41E+09	2.43E+09	2.42E+09
4	$M_{EA->SWO}$	2.46E+09	2.47E+09	2.46E+09	2.47E+09	2.47E+09
4	$M_{SWO->CA}$	2.21E+09	2.21E+09	2.20E+09	2.21E+09	2.22E+09
4	$M_{EA->CA}$	2.05E+09	2.05E+09	2.05E+09	2.06E+09	2.06E+09
4	$M_{SWO->EA}$	2.50E+09	2.50E+09	2.51E+09	2.51E+09	2.51E+09
4	$M_{CA->EA}$	2.38E+09	2.39E+09	2.39E+09	2.40E+09	2.39E+09
4	$\Delta_{CA->SWO}$	6.95E+09	6.95E+09	6.95E+09	6.95E+09	6.94E+09
4	$\Delta_{CA->EA}$	6.97E+09	6.96E+09	6.96E+09	6.97E+09	6.97E+09
4	Genealogies	3.52E+09	3.51E+09	3.51E+09	3.51E+09	3.50E+09

Supplementary Table S9. Marginal likelihoods for each model (1: island model without divergence; 2: CA and SWO diverged from EA; 3: CA and EA diverged from SWO; 4: CA and EA diverged from SWO) per replicate run for a random sub-sample of 5,000 RAD loci. The mean marginal likelihoods across replicate MIGRATE-N runs using different sub-samples of 5,000 RAD loci are also shown for each model. The best-supported model is highlighted in bold.

Model	Sub-sample 1	Sub-sample 2	Sub-sample 3	Sub-sample 4	Sub-sample 5	Mean
1: Island	-1639138	-1640123	-1641663	-1645143	-1648576	-1642928
2: CA ←EA → SWO	**-1629995**	**-1630927**	**-1632467**	**-1635943**	**-1639393**	**-1633745**
3: CA ←SWO → EA	-1630573	-1631587	-1632980	-1636614	-1640085	-1634368
4: EA ←CA → SWO	-1631734	-1632719	-1634234	-1637678	-1641204	-1635514

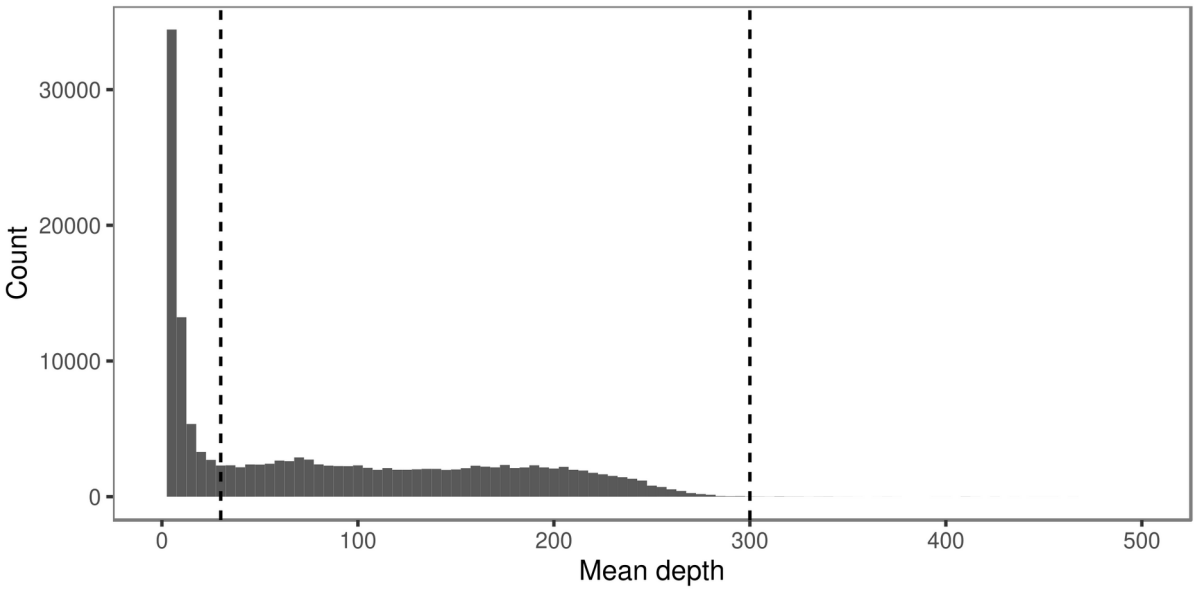
Supplementary Table S10. Mean, median and 95% credible interval (CI) of parameter estimates of the best-supported model. Population divergence times (τ) are in units of N_e * generations.

Parameter	Mean	Median	2.5%	97.5%
Θ_{SWO}	0.0041	0.0041	0.0038	0.0043
Θ_{CA}	0.0048	0.0048	0.0045	0.0050
Θ_{EA}	0.0031	0.0031	0.0029	0.0033
$M_{SWO \rightarrow CA}$	504.7	508.3	396.7	610.0
$M_{SWO \rightarrow EA}$	251.4	255.0	144.0	354.0
$M_{CA \rightarrow SWO}$	440.5	443.7	332.0	545.3
$M_{CA \rightarrow EA}$	234.1	237.7	128.0	338.0
$M_{EA \rightarrow SWO}$	710.0	713.0	598.7	817.3
$M_{EA \rightarrow CA}$	800.7	804.3	688.7	908.7
$\tau_{EA \rightarrow SWO}$	0.0345	0.0345	0.0316	0.0374
$\tau_{EA \rightarrow CA}$	0.0290	0.0290	0.0261	0.0320

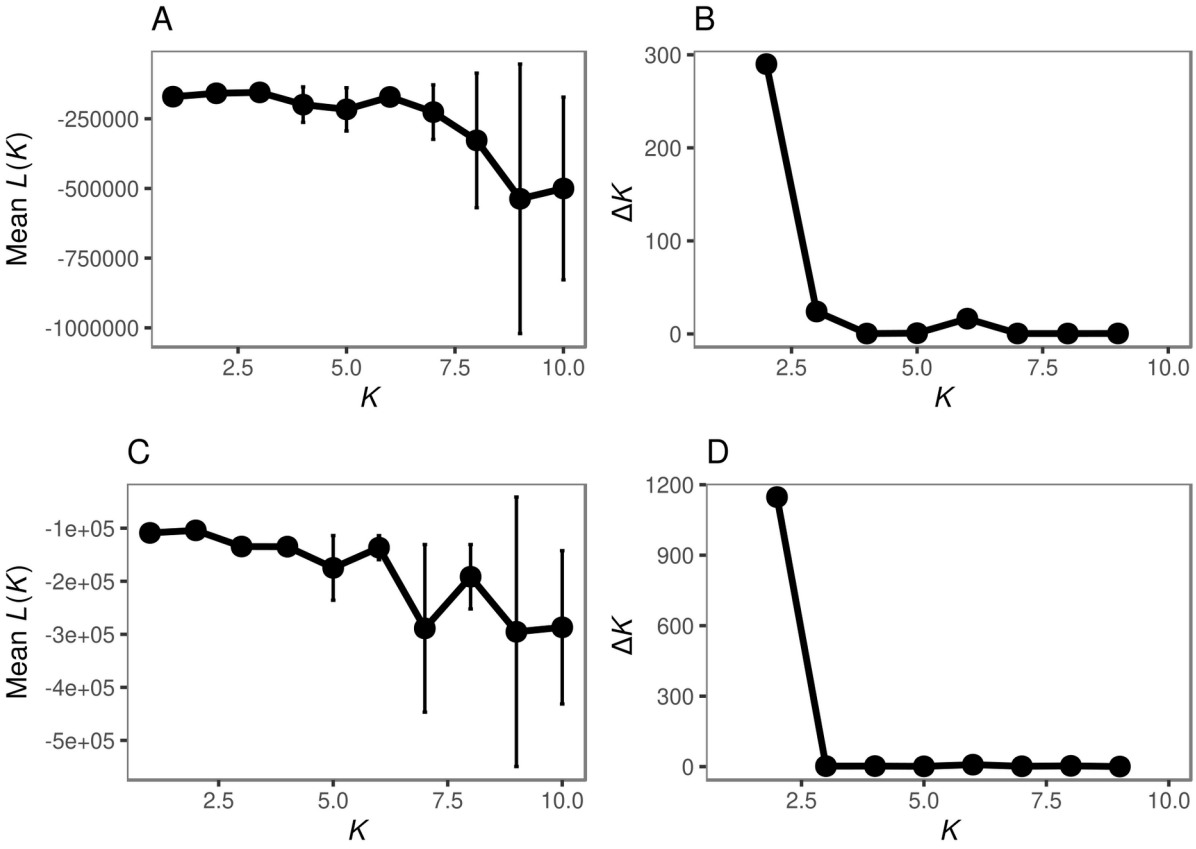


Supplementary Figure S1. (Continues on next page) Flowchart describing the steps of the bioinformatic pipeline, the data at key points in the pipeline and the analyses performed in the present study. Raw Illumina sequencing reads derived from 28 green turtle samples represented the starting point of the pipeline, which were [1] demultiplexed and cleaned and then [2] aligned to a reference genome in paired-end mode (reads without mates were aligned as unpaired reads). One East Atlantic sample (PI070009) aligned poorly (62%; Supplementary Table S1) and was excluded, resulting in the full data containing 27 samples (i.e. $N = 27$). Next, we performed [3] genotype calling using STACKS' *gstacks* module and [4] genotype filtering using STACKS' *populations* module (note: all SNPs per locus were retained in this step), producing the '*standard STACKS output*'. To filter SNPs with very low and high coverage (i.e. mean

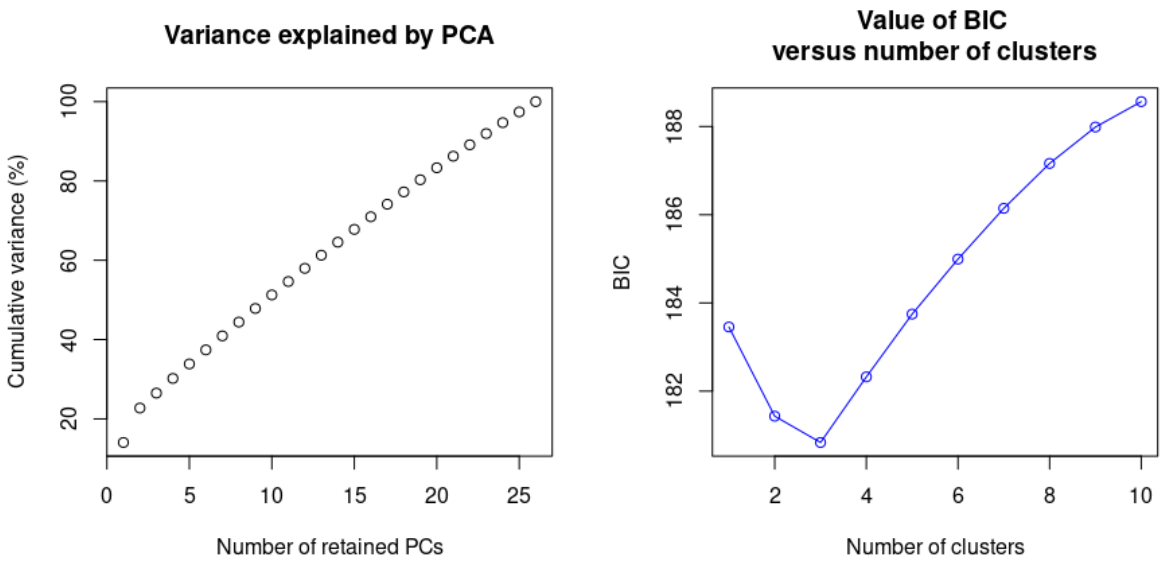
depth across samples; estimated using VCFTOOLS), we additionally [5] excluded SNPs with <30x and >300x mean depth across samples (resulting in the *depth-filtered data*). The depth-filtered data was used in the FINERADSTRUCTURE analysis. FINERADSTRUCTURE uses haplotype- and linkage information, hence loci were not removed based upon physical proximity here. The remaining analyses (genetic differentiation, STRUCTURE and DAPC) required unlinked SNPs, which is why we [6] performed 'thinning' within 100,000 bp windows using VCFTOOLS prior to these analyses. Thinning resulted in a minimum distance of 100,000 bp between SNPs, yielding putatively unlinked SNPs. Due to the computational costs of Bayesian MCMC approaches and the size of our dataset, the MIGRATE-N analysis was performed using five random subsets of 5,000 loci sampled from the ~12,000 unlinked SNPs. First, we [7] created a whitelist containing the locus IDs associated with each unlinked SNP. Next, we [8] randomly sampled 5,000 locus IDs, replicated five times, using a custom BASH script (resulting in five random sub-samples of the whitelist created in [7]). Next, we [9] used a custom Python script that did the following: for each of the five sub-sampled whitelists, go through the locus IDs and extract the FASTA sequence of that locus from the *populations.samples.fa* FASTA file produced by STACKS after step [4]. This resulting in five datasets each contained 5,000 FASTA sequences, which were [10] converted to MIGRATE-N input files using a slightly modified version (i.e., print statements were changed to make the script work with Python3) of the *stacks2mig.py* script provided with MIGRATE-N.



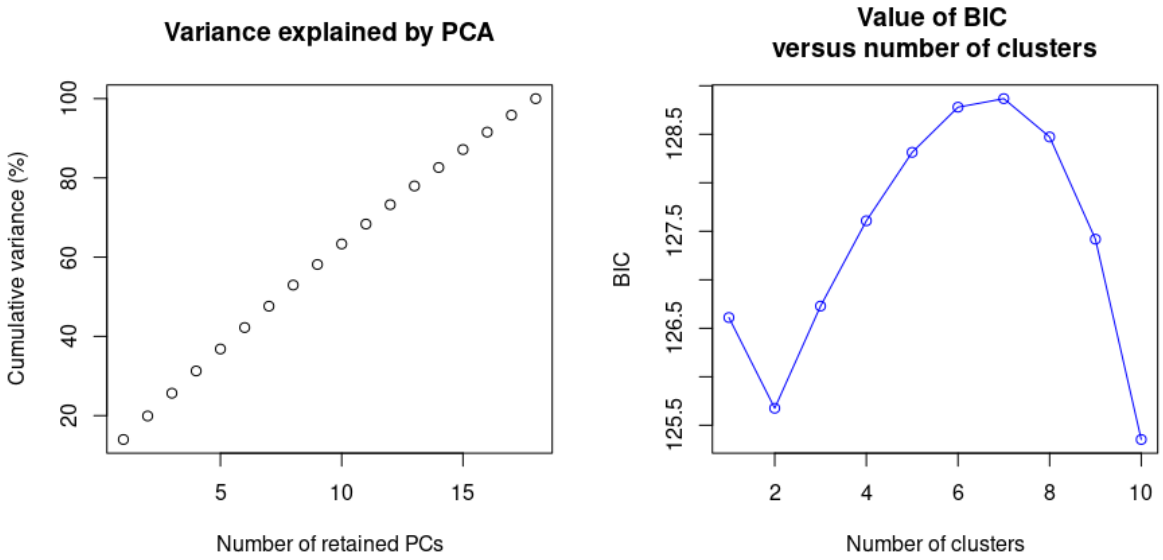
Supplementary Figure S2. Histograms of the distribution of mean sequencing depth per SNP averaged across all samples. Histogram bin-width was set to 5.



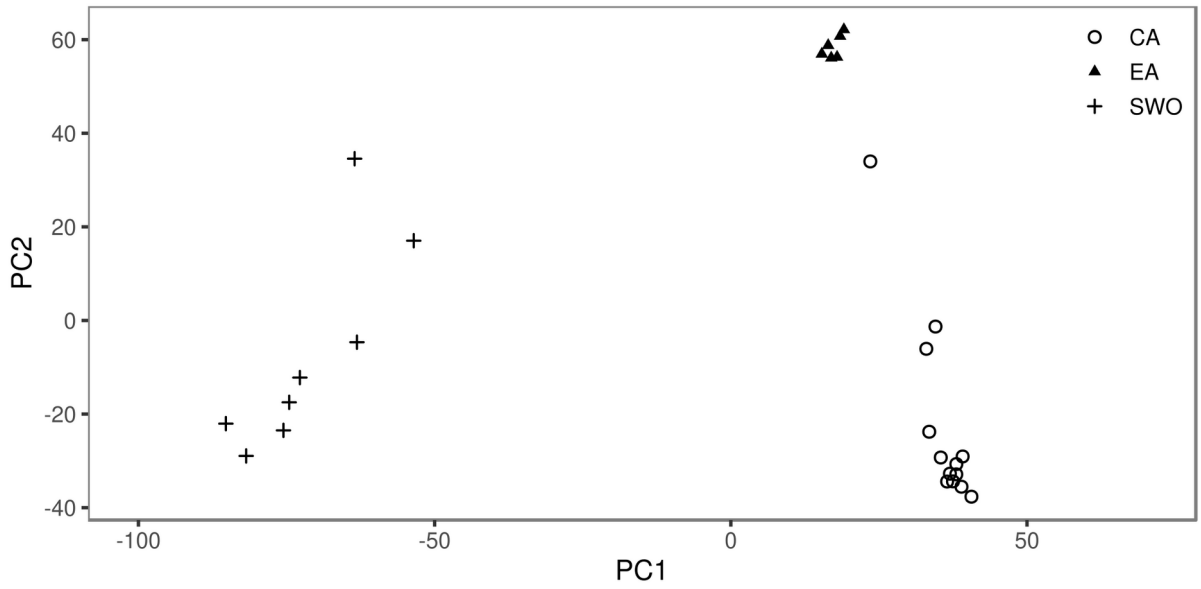
Supplementary Figure S3. Mean likelihood of *K* (left panels; A and C) and Δ*K* (right panels; B and D) for up to *K* = 10 clusters with 15 replicates per *K* estimated using STRUCTURE. Error bars depict standard deviations. Top panels (A and B): full data (*N* = 27); bottom panels (C and D): full data, Atlantic samples only (*N* = 19).



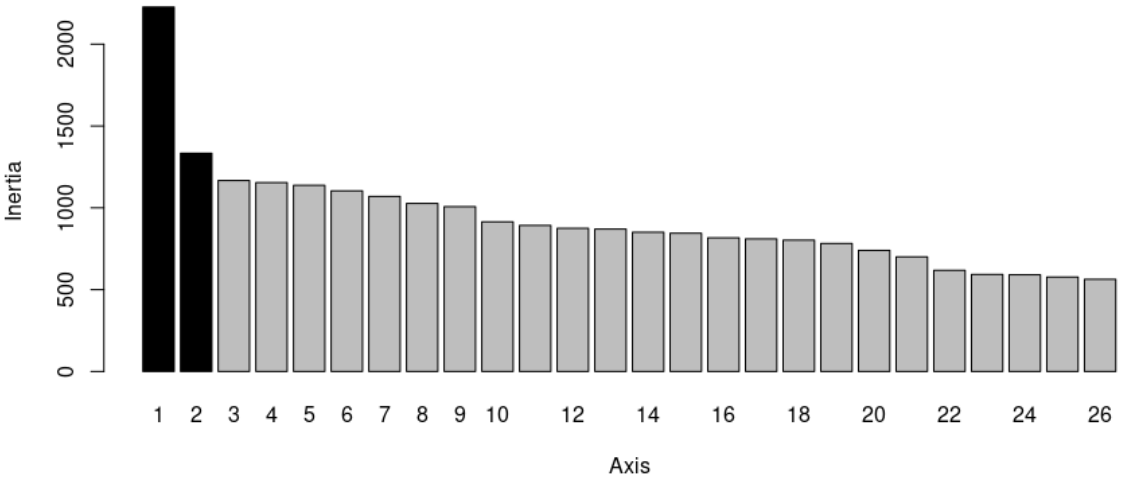
Supplementary Figure S4. Multivariate clustering results for the full data (CA, EA and SWO; *N* = 27) showing the cumulative variance (%) explained versus the number of retained principal components (PCs; left panel) and the BIC score versus the number of clusters (right panel), retaining 18 PCs and three clusters.



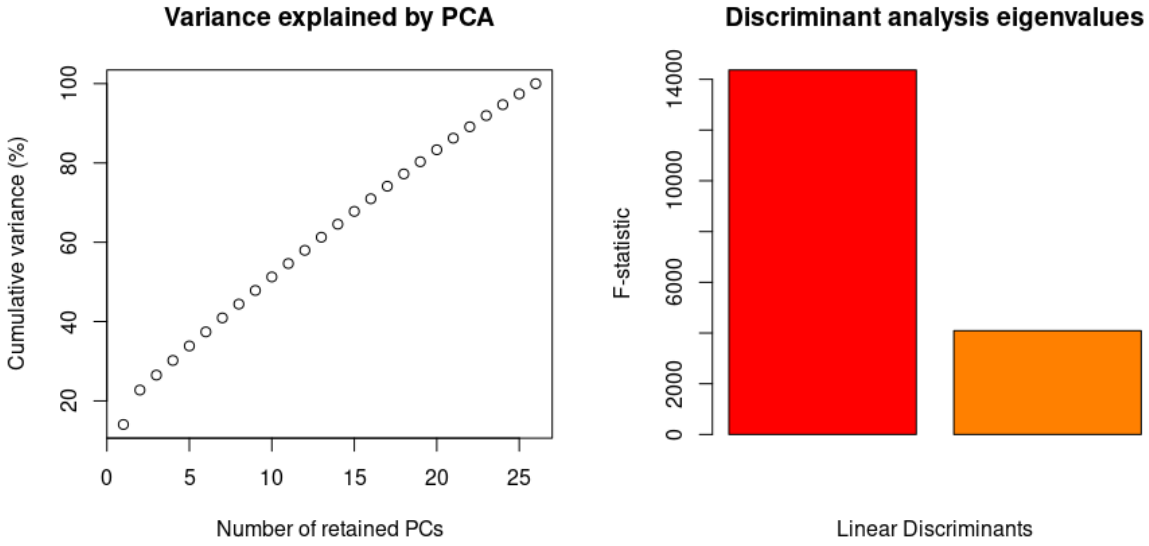
Supplementary Figure S5. Multivariate clustering results for the Atlantic data (CA and EA; *N* = 19) showing the cumulative variance (%) explained versus the number of retained principal components (PCs; left panel) and the BIC score versus the number of clusters (right panel), retaining 12 PCs and two clusters.



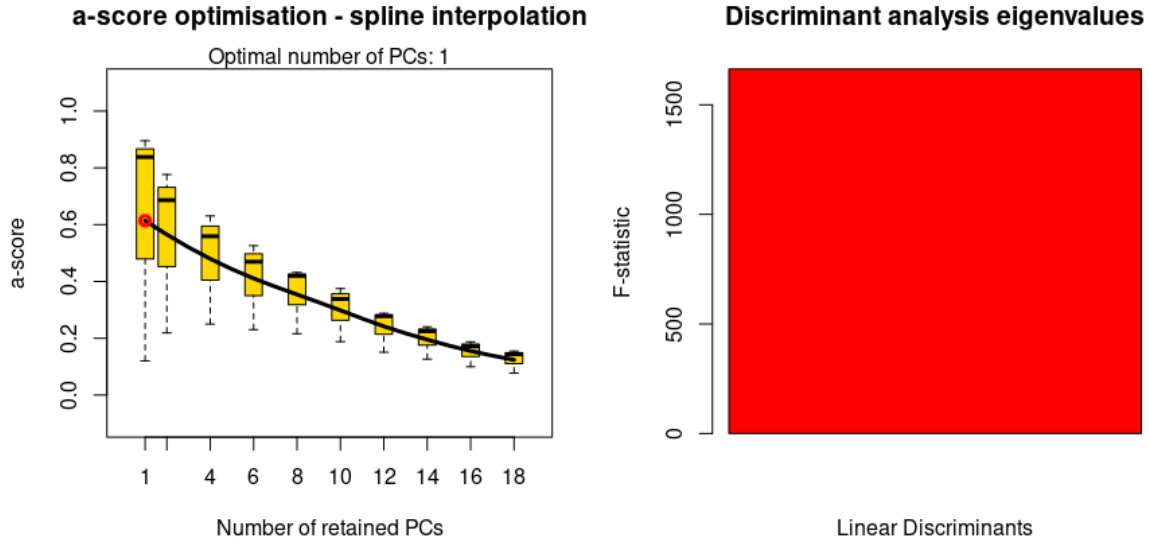
Supplementary Figure S6. Principal component analysis results, showing the first two principal components with individuals labelled by sampling location.



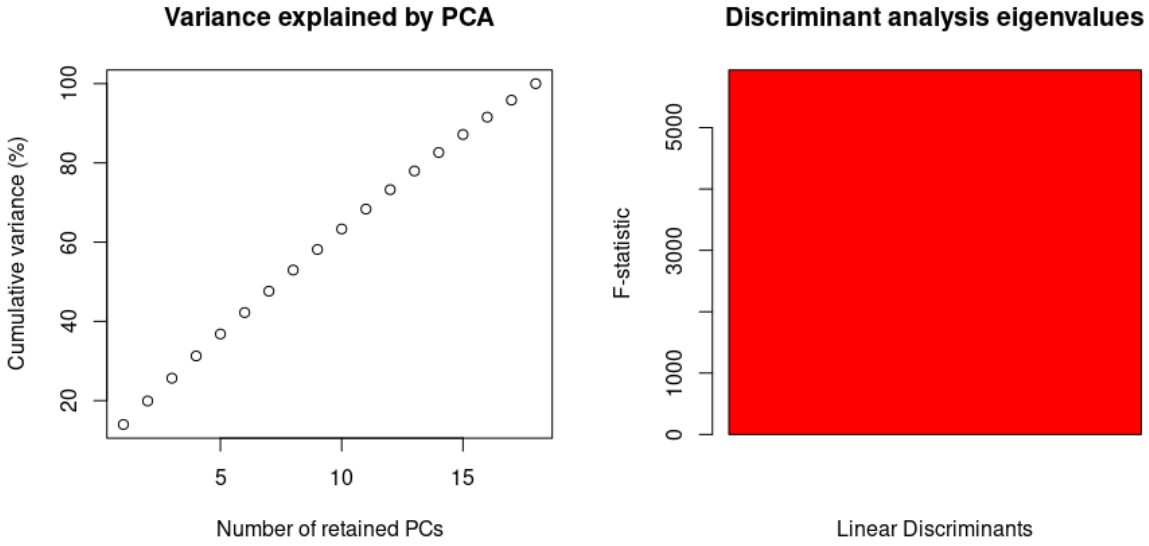
Supplementary Figure S7. Screeplot of total inertia (eigenvalues) across principal components. The break in inertia observed between the second and third principal component suggested most relevant structure was captured in the first two principal components (Jombart, 2009), which are highlighted in black.



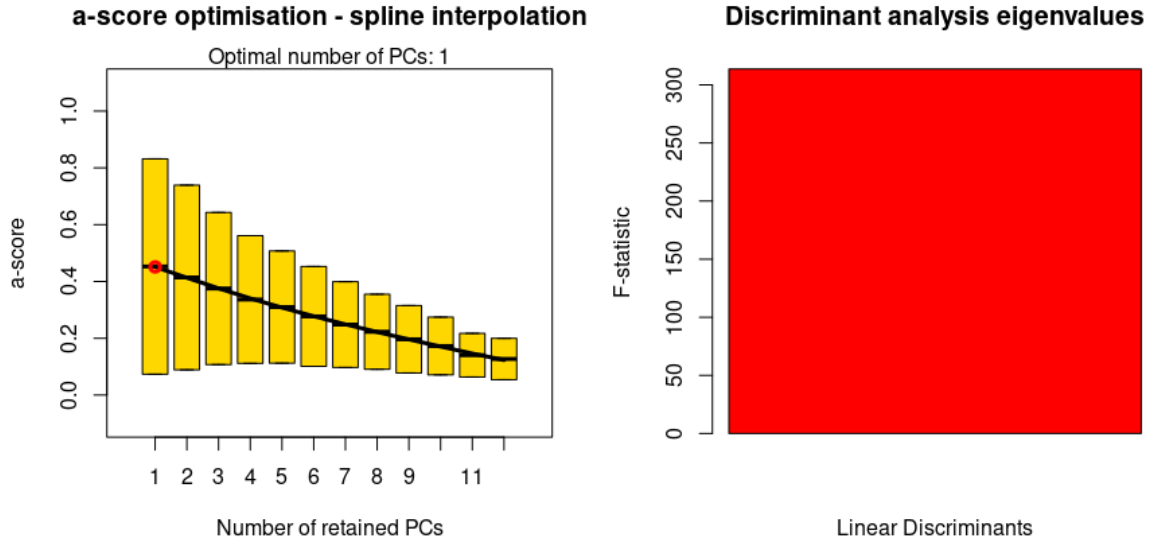
Supplementary Figure S8. DAPC results for the full data (CA, EA and SWO; *N* = 27); cumulative variance (%) explained versus the number of retained principal components (PCs; left panel) and the eigenvalues of the linear discriminants (right panel). We retained 18 PCs and two linear discriminants.



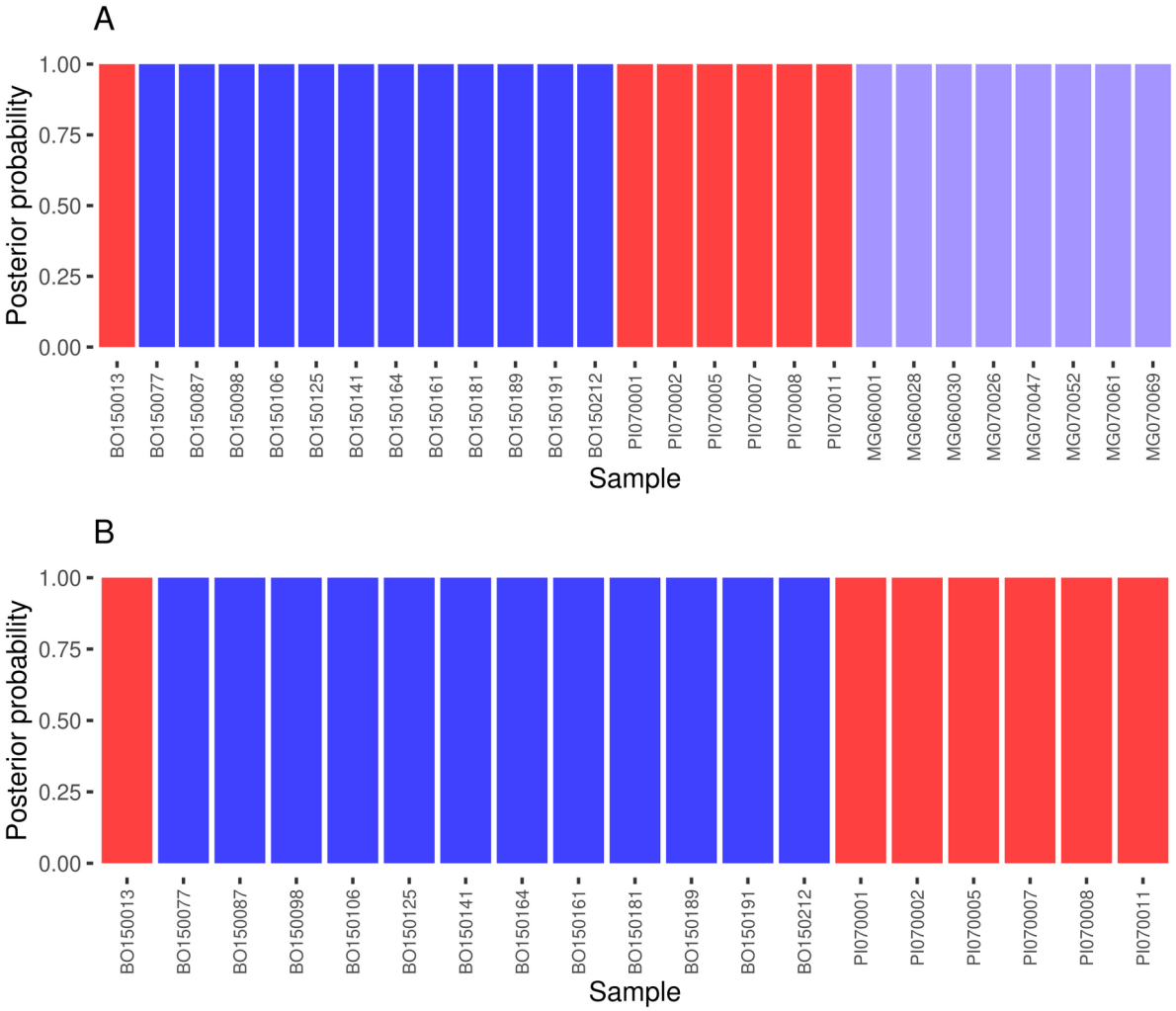
Supplementary Figure S9. DAPC results for the full data (CA, EA and SWO; *N* = 27); optimal α-scores versus the number of retained principal components (PCs; left table) and the eigenvalues of the linear discriminants (right panel). We retained one PC and one linear discriminant.



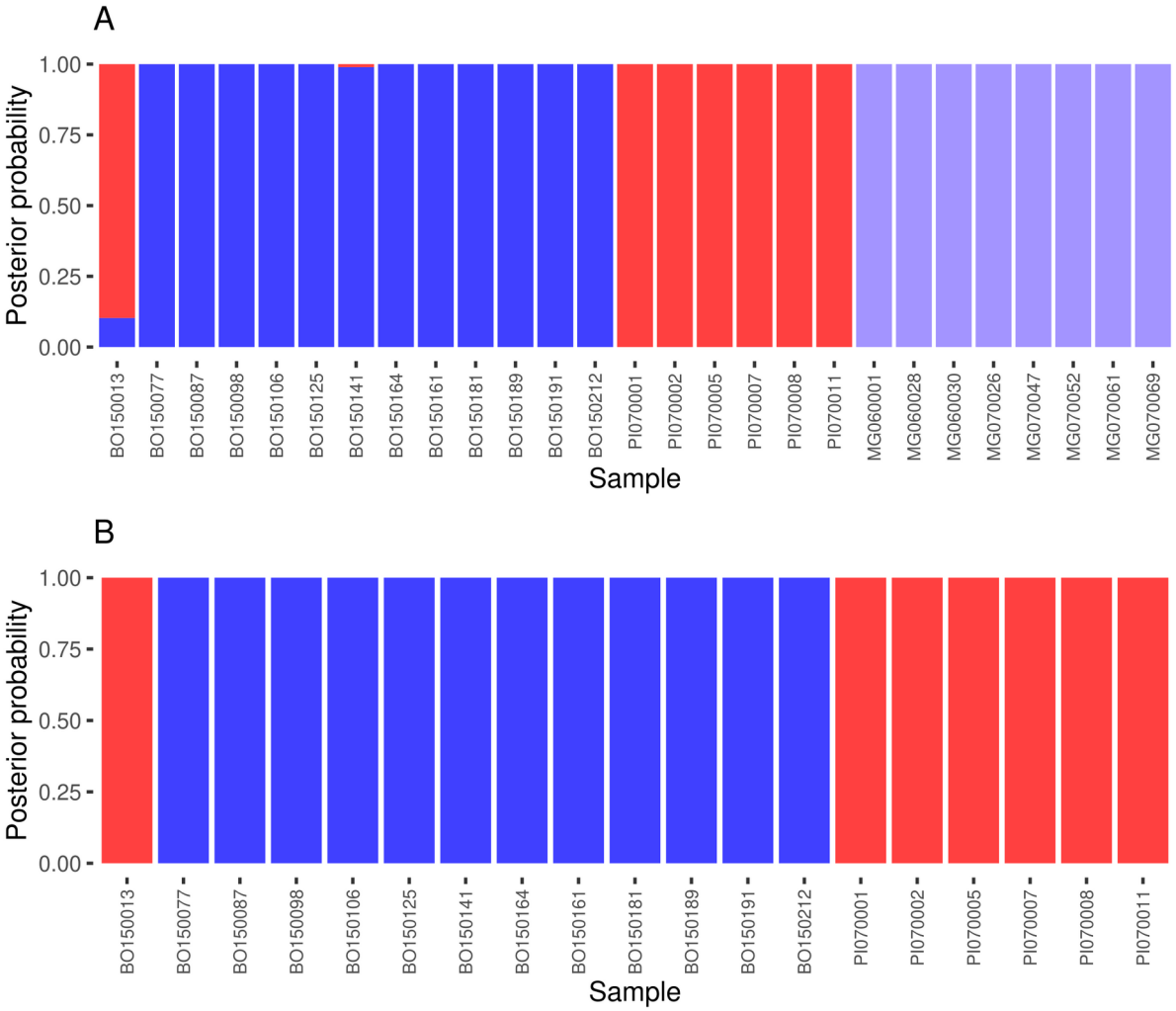
Supplementary Figure S10. DAPC results for the Atlantic data (CA and EA; *N* = 19); cumulative variance (%) explained versus the number of retained principal components (PCs; left panel) and the eigenvalues of the linear discriminants (right panel). We retained 12 PCs and one linear discriminant.



Supplementary Figure S11. DAPC results for the Atlantic data (CA and EA; *N* = 19); *o*ptimal α-scores versus the number of retained principal components (PCs; left table) and the eigenvalues of the linear discriminants (right panel). We retained one PC and one linear discriminant.



Supplementary Figure S12. Posterior group membership probabilities estimated via DAPC (retaining the number of PCs explaining ~80% of cumulative variance) for the A) full data (CA, EA and SWO; N = 27) and B) the Atlantic data (CA and EA; N = 19).



Supplementary Figure S13. Posterior group membership probabilities estimated via DAPC (retaining a single PC, which represented the optimal number of PCs according to α-score optimization) for the A) full data (CA, EA and SWO; *N* = 27) and B) Atlantic data (CA and EA; *N* = 19).