The following supplements accompany the article

Population connectivity and phylogeography of the Mediterranean endemic skate *Raja polystigma* and evidence of its hybridization with the parapatric sibling *R. montagui*

Nadja Frodella, Rita Cannas, Alessandro Velonà, Pierluigi Carbonara, Edward D. Farrell, Fabio Fiorentino, Maria Cristina Follesa, Germana Garofalo, Farid Hemida, Cecilia Mancusi, Marco Stagioni, Nicola Ungaro, Fabrizio Serena, Fausto Tinti*, Alessia Cariani

*Corresponding author: fausto.tinti@unibo.it

Marine Ecology Progress Series 554: 99–113 (2016)

Supplement 1. Additional tables and figures

Table S1 - Data on the analysed specimens. The table shows the sampling area and individual codes as well as the identification of individuals based on morphological characters, nuclear and mitochondrial results. RP = *Raja polystigma*, RM = *Raja montagui*. See Excel file at www.int-res.com/articles/suppl/m554p099_supp.xlsx

Table S2 - PCR conditions for the 7 microsatellite loci by El Nagar et al. (2010) optimized for the locus amplification in *Raja polystigma* and *R. montagui*

Locus	Fluorescent Label	Accession Number	Primers (5'-3')	Core sequence	Ta (°C)
LERI24	TET	CV221951	F: GCACGTACGCAGAATTTGAA R: CCGGCACGTGTAATTTAAGG	(TC)8	52
LERI26	TET	CV068031	F: GGAGCAGCAGTGAGGACAAT R: CTCCTACCGTCATGCCTCAT	(GA)12	48
LERI27	TET	CV068389	F: AACTGGGCAACTGACCACA R: AACGTTCTGGGTGCTGCTAC	(CT)15	54
LERI34	HEX	CO050073	F: CTTGCAATCTTTTGCCGAGT R: GTTCATCGGCCTCTTGATGT	(GT)11	52
LERI44	FAM	EE991287	F: CAGCGAGTAAACACCGACCT R: TGCGATGATCTTGAAAGACG	(GT)11	56
LERI50	FAM	DR713467	F: AATAATTGTGCCTCTTTGAGACAT R: CACAGGGAACGCAATACCTT	(TA)11	50
LERI63	FAM	CV221951	F: TTTTGATCGGCTGCAAAAAT R: CGGACTGTATAATGTGTACCAACC	(TC)8	53

Table S3 - Summary statistics of the microsatellite dataset. A= number of alleles, Ar= allelic richness, Ho= observed heterozygosity, He= expected heterozygosity, NA= estimate of null allele frequency. HWE = test of deviation from HW equilibrium *=P<0.05, **=P<0.01, ***=P<0.001. Codes of population samples are given as in Table 1.

Locus	*Raja montagui*		*Raja polystigma*						
	WI	**AL**	**AL**	**NT**	**ST**	**SI**	**ES**	**WS**	**AD**
	N = 25	N = 4	N = 12	N = 35	N = 9	N = 7	N = 12	N = 12	N = 6
LERI24									
N	19	4	11	35	9	7	12	12	6
251				0.2571	0.2222		0.0417		
253			0.4545		0.0556		0.2083	0.2500	0.1667
255	0.1316								
259			0.1364	0.0857	0.0556	0.1429	0.1250	0.3750	0.3333
261			0.0455						
263	0.1842								
265	0.1579	0.2500			0.1111	0.1429			
267	0.5263	0.7500	0.3636	0.6571	0.5556	0.7143	0.6250	0.3750	0.5000
A	4	2	4	3	5	3	4	3	3
Ar	3.339	2.000	3.116	2.451	3.526	2.670	2.954	2.923	2.907
He	0.6643	0.4286	0.6710	0.5019	0.6601	0.4835	0.5725	0.6848	0.6667
Ho	0.5263	0.5000	0.7273	0.5429	0.2222	0.2857	0.6667	0.6667	0.5000
NA	0.055	0.000	0.000	0.000	0.217	0.146	0.000	0.000	0.046
HWE					*				
LERI26									
N	24	4	11	35	9	7	12	12	5
134				0.0143					
136							0.0417		
140	0.5000	0.3750	0.8636	0.4429	0.5000	0.4286	0.4167	0.5833	0.4000
142									0.2000
144	0.2083	0.2500							
146	0.2917	0.3750	0.0909	0.5143	0.5000	0.5714	0.5000	0.3750	0.4000
A	0.0000	0.0000	0.0455	0.0286	0.0000	0.0000	0.0417	0.0417	0.0000
Ar	2.820	3.000	1.970	2.323	2.000	2.000	2.662	2.325	2.978
He	0.6348	0.7500	0.2554	0.5462	0.5294	0.5275	0.5978	0.5399	0.7111
Ho	0.8333	0.7500	0.0909	0.7714	0.7778	0.8571	0.9167	0.6667	1.000
NA	0.000	0.000	0.158	0.000	0.000	0.000	0.000	0.000	0.000
HWE			*	**			*		
LERI27									
N	18	4	9	34	7	7	12	12	5
206							0.0417		0.4000
208			0.1667	0.1618	0.2143	0.2143	0.2500	0.0833	
214				0.0735					
216			0.0556	0.3088	0.0714	0.2857	0.3333	0.4167	

Locus	*Raja montagui*		*Raja polystigma*						
	WI	**AL**	**AL**	**NT**	**ST**	**SI**	**ES**	**WS**	**AD**
	N = 25	N = 4	N = 12	N = 35	N = 9	N = 7	N = 12	N = 12	N = 6
218	0.0278		0.0556	0.0147					
220	0.3889	0.1250	0.5556	0.2794	0.5000	0.3571	0.2500	0.3333	
222	0.2222		0.1111	0.1176	0.2143	0.0714	0.1250	0.1667	0.1000
224	0.1389	0.2500	0.0556	0.0441		0.0714			0.1000
226	0.2222	0.3750							0.4000
228		0.2500							
A	5	4	6	7	6	5	5	4	4
Ar	3.746	4.000	3.892	4.238	3.462	4.070	3.920	3.372	3.600
He	0.7508	0.8214	0.6797	0.7906	0.7033	0.7912	0.7790	0.7101	0.7333
Ho	0.3333	0.7500	0.6667	0.5000	0.5714	0.5714	0.6667	0.4167	0.6000
NA	0.232	0.000	0.000	0.151	0.059	0.091	0.028	0.148	0.001
HWE	***			***					
LERI34									
N	23	4	11	31	9	7	12	11	4
270			0.1364	0.1129		0.1429	0.1667	0.0909	
274	0.0435								
278		0.1250	0.8636	0.8710	0.8333	0.7857	0.8333	0.8636	1.000
280	0.2826	0.5000		0.0161	0.1667	0.0714		0.0455	
282	0.3478	0.1250							
284	0.0435								
286	0.2609	0.2500							
288	0.0217								
A	6	4	2	3	2	3	2	3	1
Ar	3.670	4.000	1.764	1.769	1.853	2.407	1.829	1.970	1.000
He	0.7430	0.7500	0.2468	0.2322	0.2941	0.3846	0.2899	0.2554	0.0000
Ho	0.6522	0.7500	0.2727	0.2581	0.1111	0.4286	0.1667	0.2727	0.0000
NA	0.030	0.078	0.000	0.000	0.111	0.000	0.112	0.000	0.001
HWE									
LERI44									
(N)	21	4	12	29	8	5	9	11	5
287	0.0238		0.4167	0.2931	0.5625	0.1000	0.4444	0.1364	0.5000
289	0.0476	0.2500				0.1000			
291	0.1190		0.2083	0.1034			0.1111	0.1364	0.1000
293	0.0238	0.1250							
295	0.0476	0.1250		0.0690					
297	0.4762	0.2500	0.2500	0.3966	0.3750	0.7000	0.3333	0.5909	0.4000
299	0.2619	0.2500	0.1250	0.1034		0.1000	0.1111	0.1364	
315				0.0172					
317				0.0172	0.0625				
A	7	5	4	7	3	4	4	4	3
Ar	3.681	5.000	3.557	3.886	2.497	3.400	3.399	3.291	2.800

Locus	*Raja montagui*		*Raja polystigma*						
	WI	**AL**	**AL**	**NT**	**ST**	**SI**	**ES**	**WS**	**AD**
	N = 25	N = 4	N = 12	N = 35	N = 9	N = 7	N = 12	N = 12	N = 6
He	0.7015	0.8929	0.7355	0.7429	0.5750	0.5333	0.7059	0.6234	0.6444
Ho	0.4762	0.7500	0.5000	0.6552	0.2500	0.6000	0.5556	0.6364	0.6000
NA	0.128	0.157	0.100	0.054	0.182	0.000	0.023	0.018	0.000
HWE	*			*	*				
LERI63									
(N)	22	4	12	35	8	7	12	12	5
283			0.3750	0.2286	0.3750		0.2500	0.2500	0.3000
289			0.0833						0.1000
291	0.1818		0.0833	0.0714	0.0625	0.0714	0.1250	0.3333	0.3000
293	0.0455								
295	0.0682								
297	0.5909	1.000	0.4583	0.6857	0.5625	0.9286	0.6250	0.4167	0.3000
299	0.0455			0.0143					
301	0.0682								
A	6	1	4	4	3	2	3	3	4
Ar	3.419	1.000	3.120	2.469	2.497	1.571	2.664	2.919	3.800
He	0.6184	0.0000	0.6630	0.4791	0.5750	0.1429	0.5543	0.6812	0.8000
Ho	0.3182	0.0000	0.7500	0.4857	0.3750	0.1429	0.6667	0.7500	0.4000
NA	0.196	0.000	0.000	0.000	0.054	0.000	0.000	0.000	0.180
HWE	***								
A_{mean}	5.1667	3.1667	3.8333	4.6667	3.1667	3.1667	3.6667	3.3333	3.0000
Ar_{mean}	3.446	3.167	2.903	2.856	2.639	2.686	2.905	2.800	2.848
He_{mean}	0.6855	0.6071	0.5419	0.5488	0.5562	0.4772	0.5832	0.5825	0.5926
Ho_{mean}	0.5233	0.5833	0.5013	0.5355	0.3846	0.4810	0.6065	0.5682	0.5167
HWE	***			***	**				

Table S4 - Detailed results of the analyses performed to assess evidence of hybridization/introgression between *Raja polystigma* and *R. montagui*. Table S4A defines the assignment criteria, while Table S4B details the results obtained for each individual. See Excel file at www.int-res.com/articles/suppl/m554p099_supp.xlsx

Table S5 - Mean genetic distances (expressed as F-statistics indexes and estimated by AMOVA) between *R. polystigma* and *R. montagui* (A) and within *R. polystigma* (B) at the mitochondrial gene fragments and microsatellites.

Comparison/Markers	% variation	F statistics	P value
A: Two groups (*R. montagui* vs *R .polystigma*)			
COI			
Among groups	89.79	Φct = 0.90	0.035 ± 0.006
Among populations within groups	2.44	Φsc = 0.24	0.000 ± 0.000
Within populations	7.77	Φst = 0.92	0.000 ± 0.000
CR			
Among groups	89.82	Φct = 0.90	0.012 ± 0.003
Among populations within groups	0.29	Φsc = 0.03	0.130 ± 0.009
Within populations	9.89	Φst = 0.90	0.000 ± 0.000
16S			
Among groups	85.90	Φct = 0.86	0.016 ± 0.001
Among populations within groups	1.90	Φsc = 0.14	0.000 ± 0.000
Within populations	12.20	Φst = 0.88	0.000 ± 0.000
Microsatellites			
Among groups	13.74	Fct = 0.137	0.035± 0.006
Among populations within groups	3.9	Fsc = 0.045	0.000 ± 0.000
Within populations	82.35	Fst = 0.176	0.000 ± 0.000
B: *R. polystigma*			
COI			
Among populations	20.4	Φst = 0.20	0.000 ± 0.000
Within populations	79.6		
CR			
Among populations	2.21	Φst = 0.02	NS
Within populations	97.79		
16S			
Among populations	11.80	Φst = 0.12	0.000 ± 0.000
Within populations	88.20		
Microsatellites			
Among populations	4.65	Fst = 0.053	0.000 ± 0.000
Within populations	95.35		

Table S6 - Pairwise genetic distances (expressed as Fst) based on the microsatellite data among population samples of *Raja polystigma* and *R. montagui*. Significant values after FDR correction are in bold; α=0.03. Negative values are set to zero. Codes of population samples are given as in Table 1.

		R. polystigma							*R. montagui*
		AD	SI	ST	NT	ES	WS	AL	AL
R. polystigma	SI	**0.140**							
	ST	0.054	0.049						
	NT	**0.087**	0.006	0.002					
	ES	0.037	0.013	0	0				
	WS	0.048	0.060	0.055	**0.060**	0.021			
	AL	**0.089**	**0.146**	0.057	**0.098**	**0.043**	**0.049**		
R. montagui	AL	**0.187**	**0.117**	**0.192**	**0.171**	**0.175**	**0.220**	**0.261**	
	WI	**0.161**	**0.158**	**0.181**	**0.168**	**0.176**	**0.161**	**0.205**	0.036

Table S7 - Frequency of COI sequence variants in the population samples. Sequence variants belonging to the Clades P and M are indicated in black and in red, respectively.

COI haplotype	AD N = 7	SI N = 10	ST N = 10	NT N = 22	ES N = 18	WS N = 19	AL N = 20	WI N = 30	Total N = 132
COI/01			5	1	1				7
COI/02		4	1	5	6	3	4		23
COI/03		3	4	1	2	3	7		20
COI/04		2		5		6	1		14
COI/05						2			2
COI/06					4				4
COI/07						1			1
COI/08				1					1
COI/09	7								7
COI/10						2			2
COI/11		1		9	5	2	1		18
COI/12							3	26	29
COI/13								1	1
COI/14								1	1
COI/15								1	1
COI/16								1	1

Table S8 - Frequency of CR sequence variants in the population samples. Sequence variants belonging to the Clades P and M are indicated in black and in red, respectively.

haplotype	AD N = 5	SI N = 5	ST N = 10	NT N = 65	ES N = 16	NS N = 3	WS N = 24	SS N =8	AL N = 13	WI N= 26	Total N = 176
CR/01									3	24	27
CR/02										1	1
CR/03										1	1
CR/04				2					1		3
CR/05		1		4	1		2		1		9
CR/06		2	1	13	2		3	2	5		28
CR/07	5	2	4	28	6	2	7	3	4		61
CR/08				1							1
CR/09				1	1						2
CR/10				4			1				5
CR/11				4							4
CR/12				1							1
CR/13				1							1
CR/14			5	2	1						8
CR/15				1							1
CR/16				1							1
CR/17				2							2
CR/18					2	1	6	1			10
CR/19					2		1	1			4
CR/20					1						1
CR/21							2	1			3
CR/22							2				2

Table S9 - Frequency of 16S sequence variants in the population samples. Sequence variants belonging to the Clades P and M are indicated in black and in red, respectively.

haplotype	AD	SI	ST	NT	ES	NS	WS	SS	AL	WI	Total
	N = 4	N = 7	N = 10	N = 64	N = 18	N = 6	N = 26	N =13	N = 14	N= 27	N = 189
16S/01							5	1			6
16S/02						1	4				5
16S/03			1								1
16S/04		1									1
16S/05				1							1
16S/06				1							1
16S/07	4	6	9	34	12	1	14	6	11		97
16S/08				1		1	2	1			5
16S/09				26	6	3	1	5			41
16S/10				1							1
16S/11									3	27	30

Table S10 - Frequency of haplotypes for concatenated sequences (COI+CR+16S) in the population samples. Haplotypes belonging to the Clades P and M are indicated in black and in red, respectively.

haplotype	AD	SI	ST	NT	ES	WS	AL	WI	Total
	N = 4	N = 5	N = 10	N = 18	N = 14	N = 9	N = 10	N = 21	N = 91
CM/01	4								4
CM/02		1	3	1		1	4		10
CM/03			1	2	2		2		7
CM/04							1		1
CM/05				1			1		2
CM/06							2	18	20
CM/07				7	4				11
CM/08				1					1
CM/09				1					1
CM/10				1					1
CM/11				1					1
CM/12				1					1
CM/13			5	1					6
CM/14				1					1
CM/15		1							1
CM/16		1							1
CM/17		1							1
CM/18		1							1
CM/19			1						1
CM/20					1				1
CM/21					1				1
CM/22					1				1
CM/23					1				1
CM/24					1				1
CM/25					1				1
CM/26					1				1
CM/27					1				1
CM/28						1			1
CM/29						1			1
CM/30						1			1
CM/31						1			1

CM/32		1		1
CM/33		1		1
CM/34		1		1
CM/35		1		1
CM/36			1	1
CM/37			1	1
CM/38			1	1

Table S11 - Pairwise COI genetic distances (expressed as Φst) among population samples. Significant values after FDR correction are in bold; α=0.04. Negative values are set to zero. Codes of population samples are given as in Table 1.

		Clade P							Clade M
		AD	SI	ST	NT	ES	WS	AL	AL
Clade P	SI	**0.620**							
	ST	**0.629**	0.107						
	NT	**0.561**	0.093	**0.293**					
	ES	**0.496**	0.004	**0.168**	0.021				
	WS	**0.477**	0	**0.201**	**0.100**	**0.062**			
	AL	**0.629**	0	0.151	**0.211**	**0.116**	0.028		
Clade M	AL	**1**	**0.922**	**0.899**	**0.917**	**0.892**	**0.894**	**0.919**	
	WI	**0.988**	**0.964**	**0.958**	**0.952**	**0.944**	**0.944**	**0.960**	0

Table S12 - Pairwise genetic distances (expressed as Φst) based on the CR (above the diagonal) and 16S (below the diagonal) fragments among population samples. Significant values after FDR correction are in bold; α=0.02. Negative values are set to zero. Codes of population samples are given as in Table 1.

		Clade P									Clade M	
		AD	SI	ST	NT	ES	NS	SS	WS	AL	AL	WI
Clade P	AD		0.160	**0.417**	0	0.042	0.190	0.092	0.135	**0.201**	**1**	**0.992**
	SI	0		0.152	0	0	0	0	0	0	**0.911**	**0.976**
	ST	0	0.025		**0.153**	**0.132**	0.203	0.123	**0.183**	0.101	**0.904**	**0.969**
	NT	0.145	0.076	**0.217**		0	0	0	**0.033**	0	**0.818**	**0.933**
	ES	0.111	0.014	**0.196**	0		0	0	0	0	**0.930**	**0.976**
	NS	0.148	0.067	**0.289**	0	0.015		0	0	0	0.975	**0.989**
	SS	0.073	0.012	0.178	0	0	0		0	0	**0.844**	**0.958**
	WS	0	0.036	0.047	**0.222**	**0.153**	**0.157**	**0.118**		0.027	**0.875**	**0.945**
	AL	0	0.069	0,010	**0.220**	**0.232**	**0.370**	**0.213**	0.048		**0899**	**0.966**
Clade M	AL	1	**0.896**	**0.961**	**0.844**	**0.890**	**0.774**	**0.823**	**0.837**	**1**		0
	WI	**1**	**0.972**	**0.987**	**0.884**	**0.949**	**0.942**	**0.929**	**0.909**	**1**	0	

Fig. S1 – Phylogenetic relationships among haplotypes obtained with Bayesian approach (MrBayes). Relationships are represented at the three sequence markers as separate (A: COI; B: CR; C: 16S) and concatenated (D). The values above the nodes refer to the posterior probabilities, while below the nodes are the bootstrap values of the PhyML analyses (see text for details). Variants/Haplotypes belonging to the Clades P and M are indicated in black and in red, respectively.

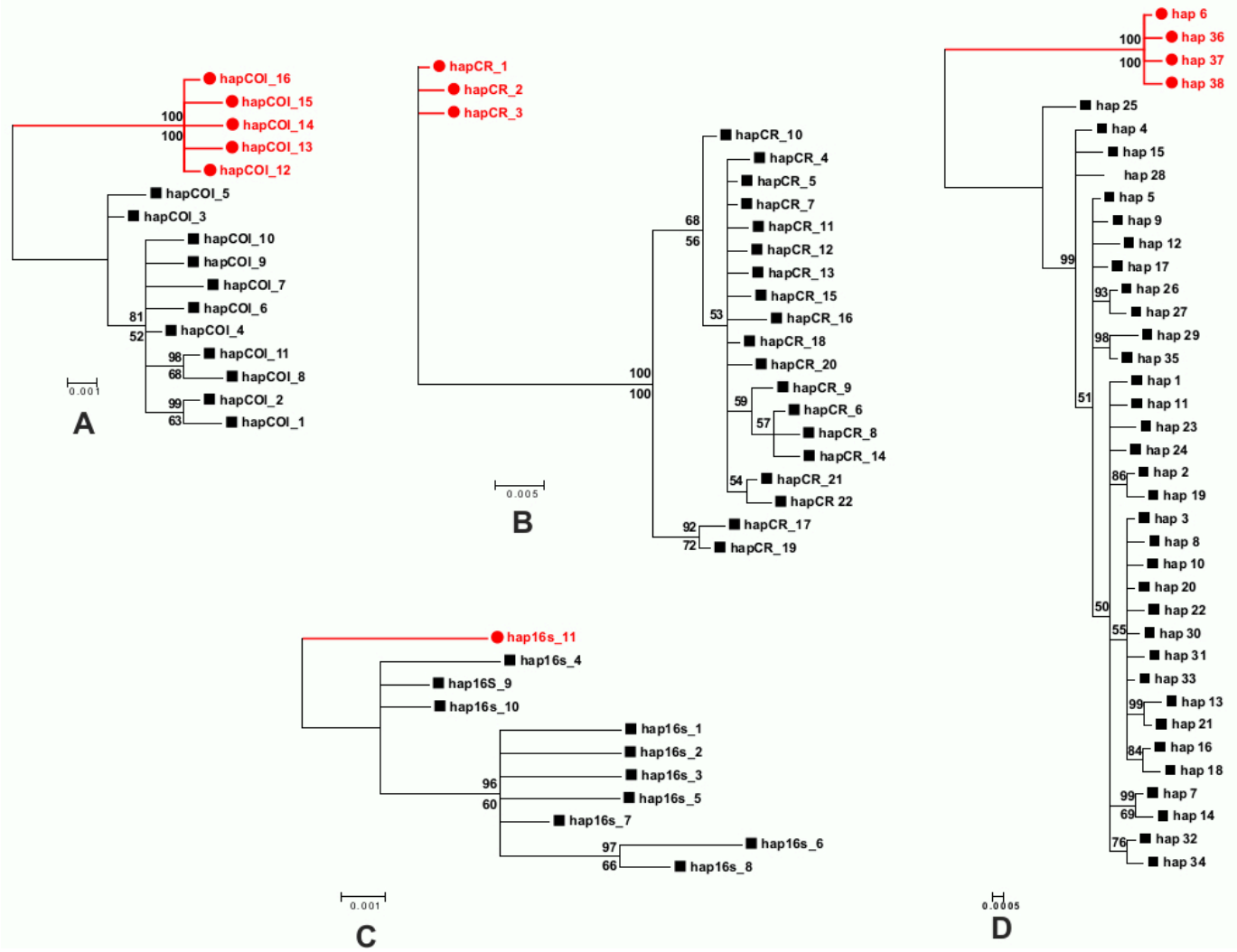


Fig. S2 - Principal Coordinates Analysis (PCoA) of the *Raja polystigma* and *R. montagui* population samples. Scatter plots built on the first two principal coordinates (coordinate 1, x axis; coordinate 2, y axis) based on the nucleotide variation at markers COI (Fig. S2A), CR (Fig. S2B) and 16S (Fig. S2C). Codes of population samples are given as in Table 1. The barplot of Eigenvalues of the first and second coordinates and the stress coefficient (d) are reported in each PCoA graph.

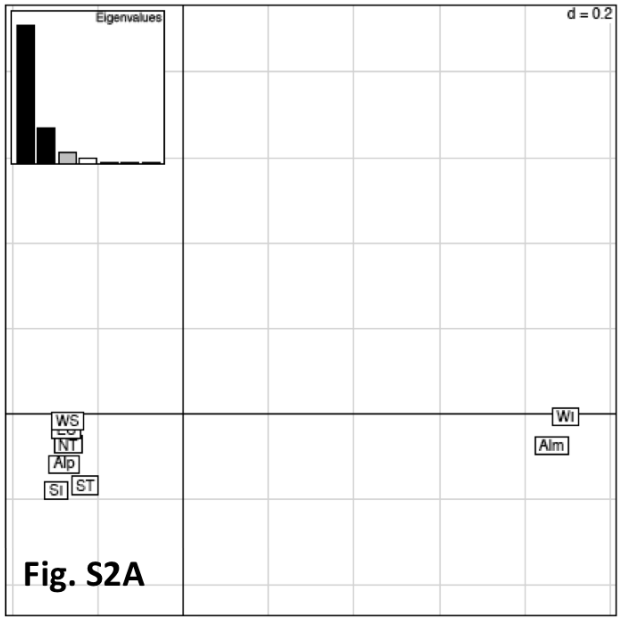


Fig. S2A

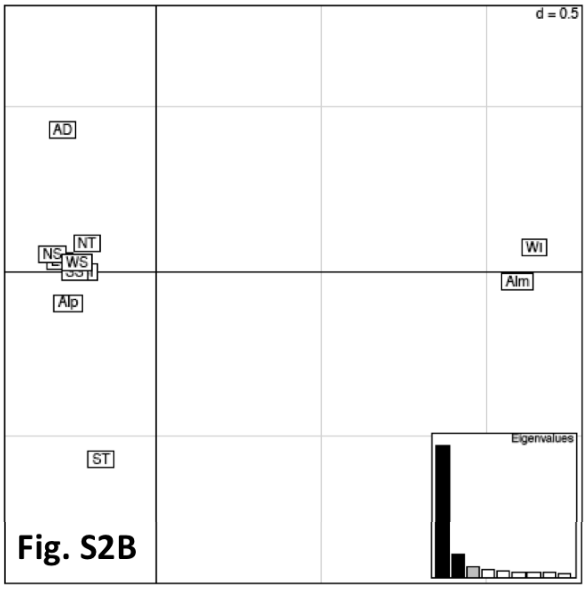


Fig. S2B

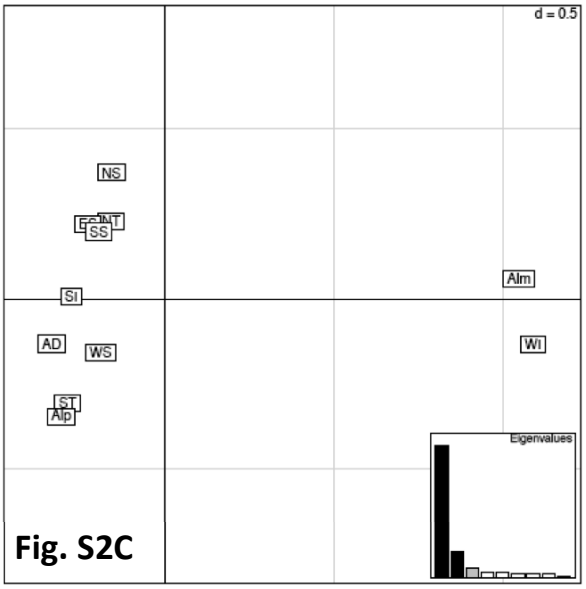
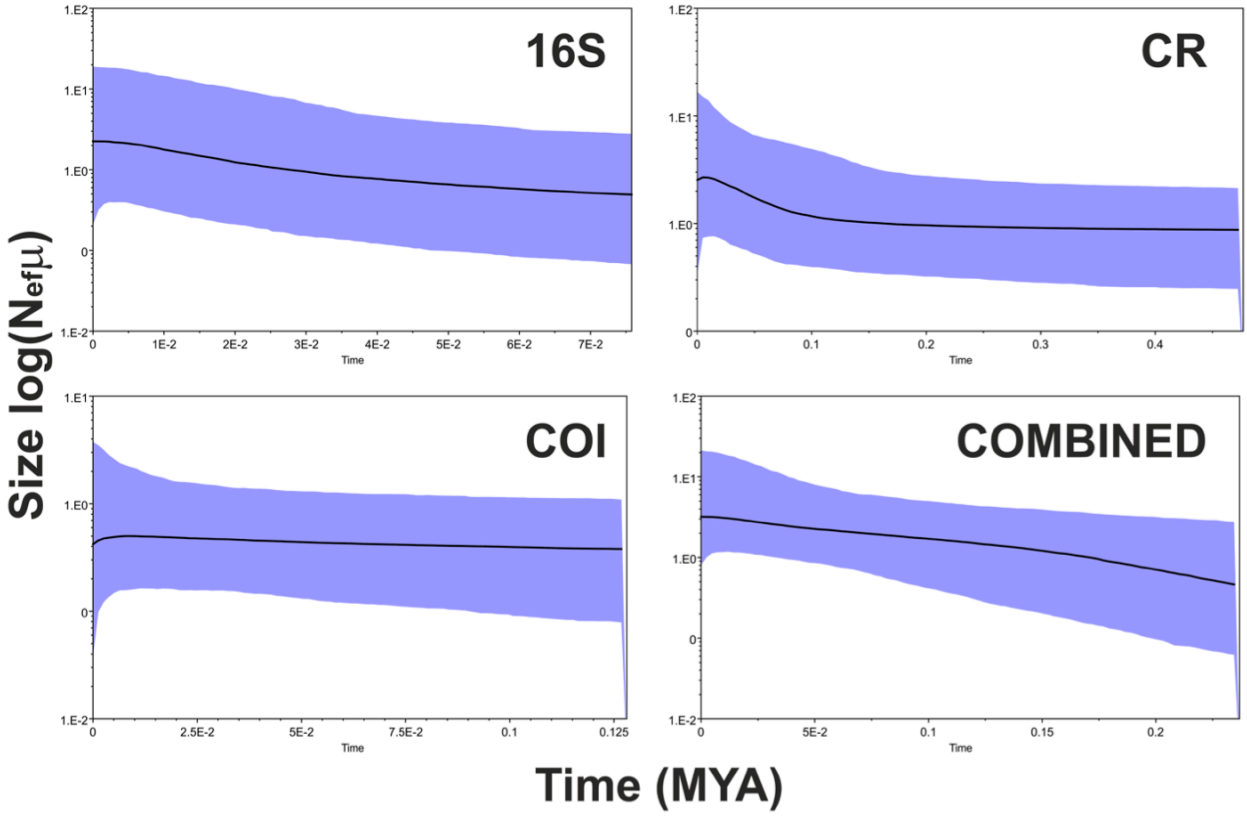


Fig. S2C

Fig. S3 - Bayesian Skyline Plots showing changes in the female effective population sizes ($N_{ef}\mu$) during time (MYA=million years ago). Black lines represent the median estimates of $N_{ef}\mu$, while grey lines the upper and the lower 95% highest posterior density (HPD) limits.



Supplement 2. Molecular protocols and data analysis methods

PCR conditions optimized for the amplification of loci in Raja polystigma *and* R. montagui

All PCR reactions were performed in either a Personal or a T-Gradient thermocycler (Biometra). The PCR reactions of COI and 16S gene fragments were carried out in a total volume of 50 μL containing 5 μL of template DNA (~40ng), 5 μL of 10X reaction buffer (Invitrogen), 4 μL of 10 mM dNTP mixture, 2.5 μL of each 10mM primer, 3 μL of 50mM $MgCl_2$, and 1.25 U recombinant *Taq* polymerase (Invitrogen). The temperature profile included an initial denaturation at 94°C for 3 min, followed by 34 cycles of denaturation at 94°C for 30 s, annealing at 54°C for 30s, elongation at 72°C for 1 min, and a final elongation step at 72°C for 7 min. The thermal profile for the CR amplification was as in Valsecchi et al. (2005).

The PCR reactions of the microsatellite loci were carried out in a total volume of 10 μL containing 1-3 μL of DNA template (~10ng), 1 μL of 10x reaction buffer, 0.8 μL of 10 mM dNTP mixture, 0.5 μL of each 10mM primer (the forward primer was fluorescent labeled), 0.5 μL of 50 mM $MgCl_2$ and 0.25 U of recombinant *Taq* polymerase (Invitrogen). The temperature profile was an initial denaturation step at 94°C for 3 min, followed by 30 cycles of denaturation at 94°C for 30 s, annealing at 48-56°C for 30 s, elongation at 72°C for 30 s. The final elongation was at 72°C for 10 min.

Clustering of individuals and hybridization analyses

The Bayesian clustering analyses was performed using STRUCTURE 2.3.4 (Pritchard et al. 2000, Hubisz et al. 2009) and a stepwise approach with three tests. The Test1 was carried out assuming an admixture ancestry model with the geographical origin of samples as prior information (LOCPRIOR models), associated with a correlated allele frequencies model. For each simulation of K (1-10), five independent replicates were run, setting a burn-in period of 250,000 iterations and 1,000,000 iterations for the MCMC. The true K was inferred using Evanno's Δk and Pritchard's average log probability methods (Pritchard et al. 2000, Evanno et al. 2005), both implemented in the STRUCTURE HARVESTER v.0.6.93 web application (Earl & Von Holdt 2012). Once the most likely number of clusters was selected, a supplemental run was performed with the same settings for additional ten independent replicates at the selected K.

The Test2 was carried out with the same settings and priors of the first test but also adding the population of origin of selected individuals as prior to assist ancestry estimation for the other individuals, applying the PFROMPOPFLAGONLY option. In this way, it was possible to update the allele frequencies, P, using only a pre-specified subset of the individuals to be regarded as the "reference" set (pre- assigned POPFLAG = 1), chosen on the basis of both the results of the first test and the species-specific mtDNA haplotype. Applying a good balance between power and accuracy, as defined in simulations trials using HYBRIDLAB v1.0 (data not shown, Nielsen et al. 2006), a $Q \geq 0.90$ was set as threshold. Each individual with a proportion of membership $Q \geq 0.90$ and 95% of Credible Interval (CI) falling within the range 0.8–1.0 was considered as a purebred (Negri et al. 2013), while individuals with $Q < 0.90$ but $Q \geq 0.80$ were considered as putative purebred. Hybrid/admixed individuals were classified as those with assignment probabilities between 50% and 80%. Finally, it was considered as unclassifiable any individual whose Q value may have indicated hybridization/admixture but the credibility intervals included values 1.00, and/or the credible intervals for two categories overlapped irrespectively from its Q (see details Table S4). The Q and CI values were obtained as the mean values from ten independent runs with K=2, using the same settings as in the first test.

The Test3 used priors to test for migrant or hybrid/introgressed individuals. By defining the population of origin as prior for all individuals, this pre-defined groups info was used for evaluating whether any individuals in the sample were immigrants to their supposed populations, or have recent immigrant ancestors. The input file was built by setting POPFLAG = 1 for all individuals, asking the program to test whether each individual had an immigrant ancestor in the last two generations (GENSBACK = 2).These latter settings inferred the posterior probability of individuals being correctly assigned to the a priori defined population, and the probability of having ancestry in the other population (i.e. in the other species). As recommended by the authors, the MIGRPRIOR value, indicated as ν in (Pritchard et al. 2000), was set to 0.05 to allow for some misclassification. The Q values were the mean values derived from ten independent runs with K=2 with settings as in the first test.

An additional Test4 was realized with NEWHYBRIDS (Anderson & Thompson 2002) to infer estimates of admixture proportions and hybrids ancestry. This software calculates Bayesian posterior probabilities (qn) that individuals fall within particular user-defined hybrid categories (purebred, F1, F2, backcross, etc.) based

on the genotypic information. The analyses were performed specifying prior information on reference, purebred *R. montagui* and *R. polystigma* individuals (as defined in the STRUCTURE Test2) by means of the "z" option in the NEWHYBRIDS input file. Model priors were set to "Jeffreys-like" for both the mixing proportions and the allele frequencies. For all individuals, probabilities of belonging to four (purebred 1, purebred 2, F1 and F2) genotype frequency classes were estimated. All results were based on 50,000 MCMC sweeps following a burn-in period of 50,000, with ten independent replicates for each series.

Mitochondrial polymorphism analysis

Sequences were checked and edited with the software MEGA6 (Tamura et al. 2013).

The number of haplotypes (N_h) and the haplotype (h) and nucleotide diversity (π) (Nei 1987) were estimated using DNAsp v.5 (Librado & Rozas 2009). Mean interspecific sequence divergence was calculated with MEGA 6 (Tamura et al. 2013), after the complete deletion of all ambiguous positions for each pair of sequences, as the between-species mean p-distance.

Haplotype trees

The relationships among haplotypes were investigated with Bayesian and Maximum Likelihood approaches using MrBayes v 3.1 (Huelsenbeck and Ronquist, 2001; Ronquist and Huelsenbeck, 2003) and PhyML v3.0 (Guindon et al., 2010), respectively. In MrBayes the analyses were performed using two parallel runs of 2 million generations each, using four chains, sampling every 100 generations, burnin 0.25, and saving branch lengths. The performance of the analyses was evaluated using the software Tracer 1.6 (Rambaut et al., 2014). Maximum-likelihood (ML) analyses of the mtDNA were performed in PhyML under 100 replications, using the best fit model the data identified by JModelTest v2. Trees were visualized with MEGA (Figure S1).

Species and pairwise population differentiation

The genetic differentiation among the putative species and population samples (pairwise F_{st}s) were computed using the substitution model of Tamura (Tamura 1992). This is the closest model implemented in Arlequin to the optimal evolutionary model HKY (Hasegawa et al. 1985) selected by the software JModelTest v2 (Darriba et al. 2012) based on the Bayesian Information Criterion (BIC).

Statistical significance (α=0.05) of overall and pairwise F-statistics values was obtained after 10,000 permutations and adjusted using false discovery rate (FDR) correction for multiple comparisons using the method of Benjamini & Hochberg as implemented in SGOF+ (Carvajal-Rodriguez & de Uña-Alvarez 2011).

Principal coordinate analysis

Pairwise Φsts between population samples were transformed into Euclidean matrices through the addition of smallest positive constant (Cailliez 1983), and used to reconstruct scatter plots of Principal Coordinates Analysis (PCoA) using the packages ade4 (Dray & Dufour 2007) and ape (Paradis et al. 2004) in R environment 3.0.2 (R-Core-Team 2013).

Demographic analyses

Bayesian skyline plot (Drummond et al. 2005) was obtained in BEAST 1.75 (Drummond & Rambaut 2007) using a strict molecular clock and a mutation rate of 0.005/million years (Chevolot et al. 2006b), and the optimal model of nucleotide evolution (HKY) selected with JModelTest. We performed a Markov Chain Monte Carlo (MCMC) run of 50,000,000 generations sampled every 5,000 generations with the first 10% of the sampled points removed as burn-in. The quality of the run was assessed by effective sample size (ESS) > 200 for each parameter using Tracer 1.6 (Rambaut et al., 2014). The same software was used to produce the skyline plots.

References

Anderson EC, Thompson EA (2002) A model-based method for identifying species hybrids using multilocus genetic data. Genetics 160:1217-1229

Cailliez F (1983) The analytical solution of the additive constant problem. Psychometrika 48:305-308

Carvajal-Rodriguez A, de Uña-Alvarez J (2011) Assessing significance in high-throughput experiments by sequential goodness of fit and q-value estimation. PloS one 6

Chevolot M, Hoarau G, Rijnsdorp AD, Stam WT, Olsen JL (2006b) Phylogeography and population structure of thornback rays (Raja clavata L., Rajidae). Molecular Ecology 15:3693-3705

Darriba D, Taboada GL, Doallo R, Posada D (2012) jModelTest 2: more models, new heuristics and parallel computing. Nat Meth 9:772-772

Dray S, Dufour A (2007) The ade4 package: implementing the duality diagram for ecologists. Journal of statistical software 22:1-20

Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. BMC evolutionary biology 7:214

Drummond AJ, Rambaut A, Shapiro B, Pybus OG (2005) Bayesian coalescent inference of past population dynamics from molecular sequences. Molecular biology and evolution 22:1185-1192

Earl DA, Von Holdt B (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation genetics resources 4:359-361

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

Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O (2010) New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. Systematic Biology 59:307-321

Hasegawa M, Kishino H, Yano T (1985) Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. Journal of molecular evolution 22:160-174

Hubisz MJ, Falush D, Stephens M, Pritchard JK (2009) Inferring weak population structure with the assistance of sample group information. Molecular ecology resources 9:1322-1332

Huelsenbeck JP, Ronquist F (2001) MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17:754-755

Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25:1451-1452

Negri A, Pellegrino I, Mucci N, Randi E, Tizzani P, Meneguz PG, Malacarne G (2013) Mitochondrial DNA and microsatellite markers evidence a different pattern of hybridization in red-legged partridge (Alectoris rufa) populations from NW Italy. European Journal of Wildlife Research 59:407-419

Nei M (1987) Molecular evolutionary genetics. Columbia University Press

Nielsen EE, Bach LA, Kotlicki P (2006) HYBRIDLAB (version 1.0): a program for generating simulated hybrids from population samples. Molecular Ecology Notes 6:971-973

Paradis E, Claude J, Strimmer K (2004) APE: analyses of phylogenetics and evolution in R language. Bioinformatics 20:289-290

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945-959

R-Core-Team (2013) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/)

Rambaut A, Suchard MA, Xie D & Drummond AJ (2014) Tracer v1.6, Available from http://beast.bio.ed.ac.uk/Tracer

Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19:1572-1574

Tamura K (1992) Estimation of the number of nucleotide substitutions when there are strong transition-transversion and G+C-content biases. Molecular biology and evolution 9:678-687

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. Molecular biology and evolution 30:2725-2729

Valsecchi E, Pasolini P, Bertozzi M, Garoia F, Ungaro N, Vacchi M, Sabelli B, Tinti F (2005) Rapid Miocene-Pliocene dispersal and evolution of Mediterranean rajid fauna as inferred by mitochondrial gene variation. Journal of Evolutionary Biology 18:436-446