

Conservation challenges imposed by evolutionary history and habitat suitability shifts of endangered freshwater mussels under a global climate change scenario.

Reis J.¹, Perea S.^{1,2}, Gama, M. ³, Mendes S.L.^{1,4}, Sousa V.C.⁴, Lima C.S.⁵, Banha, F. ³ Gil, M.¹, Alvarez, M.G.¹, Anastácio, P. ³, Sousa-Santos C.⁵, Araujo R.^{6†},

Authors' addresses

¹*MARE - Marine and Environmental Sciences Centre/ARNET-Aquatic Research Network, Faculdade de Ciências da Universidade de Lisboa, Campo Grande, 1749-016 Lisbon, Portugal.*

²*Instituto de Biología, Departamento de Zoología, Universidad Nacional Autónoma de México, Tercer Circuito Exterior S/N. C.P. 04510. Ciudad de México, Mexico.*

³*MARE-Marine and Environmental Sciences Centre/ARNET - Aquatic Research Network.Universidade de Évora, Portugal.*

⁴*cE3c – Centre for Ecology, Evolution and Environmental Changes & CHANGE – Global Change and Sustainability Institute. Department of Animal Biology, Faculdade de Ciências da Universidade de Lisboa, Campo Grande, 1749-016 Lisboa, Portugal.*

⁵ *MARE – Marine and Environmental Sciences Centre/ARNET-Aquatic Research Network. ISPA – Instituto Superior de Ciências Psicológicas, Sociais e da Vida. Rua Jardim do Tabaco, 34 - 1149-041 Lisboa, Portugal.*

⁶ *Museo Nacional de Ciencias Naturales - CSIC, C/José Gutierrez Abascal, 2, 28006 Madrid, Spain.*

[†] Rafael Araujo is deceased.

Corresponding author:

Joaquim Reis

E-mail: joaquim.reis@fc.ul.pt

Address: *MARE-Marine and Environmental Sciences Centre/ARNET-Aquatic Research Network, Faculdade de Ciências da Universidade de Lisboa, Campo Grande, 1749-016 Lisbon, Portugal.*

Telephone: (+351) 964723114

Table S1 – Location, vouchers, and quality of genomic data (% missing data and mean depth coverage) for each sample.

Basin	Sub-basin	River/Sampling location (number of sampled/analyzed individuals after SNP-filtering)	Locality	Sample name	GenBank Accession Number (BioSample) (b)	% missing data	Mean depth coverage
Sado	Marateca	Marateca 1 (5/5)	Landeira	TM1_2 ^(a)	SRR28340993/ SAMN40447825	38.57	14.17
				TM1_3 ^(a)	SRR28340992/ SAMN40447826	41.63	12.49
				TM1_4	SRR28340981/ SAMN40447827	36.40	14.34
				TM1_8	SRR28340970/ SAMN40447828	37.37	14.31
				TM1_12	SRR28340959/ SAMN40447829	34.94	15.37
		Marateca 2 (5/5)	Landeira	TM2_1 ^(a)	SRR28340948/ SAMN40447830	45.77	10.05
				TM2_8 ^(a)	SRR28340937/ SAMN40447831	31.25	15.44
				TM2_10	SRR28340936/ SAMN40447832	39.34	14.21
				TM2_11	SRR28340935/ SAMN40447833	42.06	14.48
				TM2_14	SRR28340934/ SAMN40447834	45.77	12.09
	Sado	Sado (3/3)	Torre Vã	TS8_1 ^(a)	SRR28340991/ SAMN40447835	36.52	13.56
				TS8_2 ^(a)	SRR28340990/ SAMN40447836	34.33	14.77
				TS8_3	SRR28340989/ SAMN40447837	37.63	15.90
Mira	Torgal	Torgal (5/3)	Odemira	T_Torg_1 ^(a)	SRR28340988/ SAMN40447838	25.42	12.72
				T_Torg_2 ^(a)	SRR28340987/ SAMN40447839	82.48	5.74
				T_Torg_3	SRR28340986/ SAMN40447840	55.29	9.88
				T_Torg_4	SRR28340985/ SAMN40447841	26.12	13.17
				T_Torg_5	SRR28340984/ SAMN40447842	27.76	11.95
Lower Guadiana	Vascão	Vascão 1 (5/2)	Tavilhão	2.1-T1	SRR28340983/ SAMN40447843	31.94	14.09
				2.1-T2	SRR28340982/ SAMN40447844	28.83	12.89
				2.1-T7 ^(a)	SRR28340980/ SAMN40447845	65.82	8.73
				2.1-T8 ^(a)	SRR28340979/ SAMN40447846	53.49	18.65
				2.1-T9	SRR28340978/ SAMN40447847	-	2.97
		Vascão 2 (5/3)	Castelhanos	1.1-T1 ^(a)	SRR28340977/ SAMN40447848	39.16	9.98
				1.1-T3 ^(a)	SRR28340976/ SAMN40447849	44.67	8.63
				1.1-T7	SRR28340975/ SAMN40447850	69.80	7.55

Upper Guadiana	Odeleite	Odeleite (5/5)	Grainho	1.1-T8	SRR28340974/ SAMN40447851	-	3.41
				1.1-T9	SRR28340973/ SAMN40447852	38.93	9.91
				3.1-T1 ^(a)	SRR28340972/ SAMN40447853	32.26	13.38
				3.1-T2	SRR28340971/ SAMN40447854	30.01	14.76
				3.1-T3 ^(a)	SRR28340969/ SAMN40447855	32.08	14.97
				3.1-T4	SRR28340968/ SAMN40447856	27.56	15.60
				3.1-T9	SRR28340967/ SAMN40447857	28.48	14.02
				4.2-T2 ^(a)	SRR28340966/ SAMN40447858	34.94	15.93
				4.2-T4 ^(a)	SRR28340965/ SAMN40447859	28.20	14.33
				4.2-T5	SRR28340964/ SAMN40447860	61.93	17.65
	Chança	Barranco do Vidigão (6/4)	Vila Verde de Ficalho	4.2-T7	SRR28340963/ SAMN40447861	29.00	15.56
				4.2-T9	SRR28340962/ SAMN40447862	54.40	9.71
				4.2-D6	SRR28340961/ SAMN40447863	26.58	14.22
				2.2-T1 ^(a)	SRR28340960/ SAMN40447864	87.23	4.65
				TSPed_1 ^(a)	SRR28340958/ SAMN40447865	34.94	11.39
	Estena	Estena(5/4)	Ciudad Real	TECD_1 ^(a)	SRR28340957/ SAMN40447866	48.32	8.98
				TECD_2 ^(a)	SRR28340956/ SAMN40447867	27.30	19.51
				TECD_3	SRR28340955/ SAMN40447868	83.87	5.64
				TENE_1 ^(a)	SRR28340954/ SAMN40447869	39.39	10.84
			Navas de Estena	TENE_2 ^(a)	SRR28340953/ SAMN40447870	46.43	9.66
				TBul_1 ^(a)	SRR28340952/ SAMN40447871	45.78	14.07
			El torno	TBul_2 ^(a)	SRR28340951/ SAMN40447872	35.24	9.95
				TBul_3	SRR28340950/ SAMN40447873	29.96	13.76
				TBul_4	SRR28340949/ SAMN40447874	27.81	12.66
				TBul_5	SRR28340947/ SAMN40447875	27.87	16.18
	Bullaque	Bullaque (5/5)	Castillo del Milagro	TMil_1 ^(a)	SRR28340946/ SAMN40447876	27.45	16.77
				TMil_2 ^(a)	SRR28340945/ SAMN40447877	35.93	9.60
				TMil_3	SRR28340944/ SAMN40447878	23.50	17.25
				TMil_4	SRR28340943/ SAMN40447879	23.76	17.67
				TRui_1	SRR28340942/ SAMN40447880	32.24	12.25
	Guadiana	Guadiana (5/4)	Lagunas de Ruidera				
	Terges e Cobres	Terges e Cobres (1/0)	Vale de Russins				
	Ardila	São Pedro (1/1)	Safara				

TRui_2	SRR28340941/ SAMN40447881	30.68	13.95
TRui_3	SRR28340940/ SAMN40447882	22.93	15.69
TRui_4	SRR28340939/ SAMN40447883	54.09	9.59
TRui_5	SRR28340938/ SAMN40447884	34.31	18.85

-
- (a) samples used to build the catalog in stacks (M4m4n5, 2 individual/sub-basin).
(b) GenBank Accession numbers (Biosample Accession) will be made available upon publication.

Table S2. Genetic diversity parameters at different hierarchical levels (basins, sub-basins, and population/location).

Basin	Sub-basins	population	H _{exp} (a)	H _{obs} (b)	private sites	% polymorphic sites	FIS (d)
Sado	Sado (all)		0.108	0.100	7296	34.70	-
	Marateca	Marateca (all)	0.077	0.086	1628	21.90	-0.014
		Marateca 1	0.077	0.086	97	17.74	-0.0097
		Marateca 2	0.075	0.086	39	15.67	-0.0107
	Sado	Sado	0.128	0.143	230	24.52	-0.0045
Mira	Torgal	Torgal	0.065	0.073	1427	13.48	-0.0099
Guadiana	Guadiana (all)		0.253	0.143	33494	87.53	-
	<i>Lower Guadiana (all)</i>		0.229	0.121	11101	63.56	-
	Odeleite	Odeleite	0.111	0.120	3377	25.78	-0.0073
	Vascão	Vascão (all)	0.109	0.116	3365	24.34	-0.0068
		Vascão (Tavilhão)	0.097	0.114	278	16.00	-0.0056
		Vascão (Castelhanos)	0.104	0.118	379	18.94	-0.0050
	Chança	Barranco do Vidigao	0.119	0.128	1988	25.15	-0.0088
	<i>Upper Guadiana</i>		0.225	0.158	9602	66.95	-
	Ardila	San Pedro	0.109	0.167	9	16.71	-
	Estena	Estena	0.138	0.143	427	27.64	0.0028
	Bullaque	Bullaque (all)	0.192	0.172	1872	53.99	0.0619
		Milagro	0.181	0.155	245	38.60	0.0536

Bullaque	0.184	0.185	259	41.53	0.0036
Guadiana Ruidera	0.135	0.150	1067	27.80	-0.018

- (a) Expected heterozygosity.
- (b) Observed heterozygosity.
- (c) Nucleotide diversity.
- (d) Inbreeding coefficient.

Table S3. Genetic differentiation (FST pairwise comparisons) for 13 populations based on Weir and Cockerham's estimator.

	Torgal	Sado	Marateca 2	Marateca 1	Ruidera	Milagro	Bullaque	Estena	São Pedro	Barranco do Vidigão	Vascão Castelhanos	Vascão Tavilhão	Odeleite
Torgal	-	0.630	0.745	0.747	0.563	0.440	0.430	0.574	0.565	0.597	0.676	0.668	0.628
Sado		-	0.183	0.218	0.561	0.357	0.351	0.450	0.325	0.548	0.562	0.541	0.561
Marateca 2			-	-0.09	0.613	0.520	0.508	0.610	0.606	0.667	0.697	0.684	0.676
Marateca 1				-	0.598	0.494	0.485	0.602	0.606	0.655	0.697	0.681	0.661
Ruidera					-	0.257	0.228	0.383	0.300	0.479	0.498	0.480	0.522
Milagro						-	0.033	0.146	0.065	0.374	0.367	0.342	0.432
Bullaque							-	0.217	0.097	0.365	0.365	0.343	0.417
Estena								-	0.297	0.441	0.492	0.479	0.476
São Pedro									-	0.348	0.442	0.437	0.402
Barranco do Vidigão										-	0.553	0.526	0.558
Vascao Tavilhão											-	-0.09	0.009

Vascão	-	0.554
Castelhanos		
Odeleite	0.	-

* value is significant after Bonferroni's correction (p-value = 0.0006).

Table S4 – Analysis of Molecular Variance (AMOVA) to evaluate the partitioning of genetic variation (structure) among different sampling schemes, based on a pairwise distance matrix between samples.

Source of Variation	d. f.	Sum of Squares	Variance Components	Percentage of Genetic Variation
Among sub-basins	9	115380.930	944.36407	43.98*
Among locations within sub-basins	3	7107.994	215.26808	10.02*
Among individuals within locations	44	19355.233	-547.80855	-25.51
within individuals	57	87524.000	1535.50877	71.51*
Total	113	229368.158	2147.33237	

* significant at 0.05 p-level.

Table S5 –Average Model evaluation using True skill statistic (TSS) for each of the algorithms and modelled taxa.

	<i>U. tumidiformis</i>		<i>Squalius</i>	
Models*	Average	SE	Average	SE
SRE	0.792	0.030	0.77	0.01
CTA	0.962	0.017	0.98	0.00
RF	0.999	0.000	0.99	0.00
FDA	0.992	0.001	0.96	0.00
GLM	0.989	0.005	0.97	0.01
GBM	0.993	0.003	0.98	0.00
GAM	0.990	0.007	0.99	0.00
ANN	0.997	0.001	0.99	0.00
Average	0.96	0.01	0.95	0.00

* Models computed: GLM (generalized linear model), GAM (generalized additive model), CTA (classification tree analysis), FDA (flexible discriminant analysis), ANN (artificial neural networks), RF (random forest for classification and regression), GBM (generalized boosted regression model) and SRE (surface range envelope).

Table S6 – Percentage of suitable and unsuitable areas in the Iberian Peninsula for the selected taxa and the modeled timeframes and shared socioeconomic pathways (SSPs).

Species	Timeframe	SSP	Suitable (%)	Unsuitable (%)	% reduction*
<i>Unio tumidiformis</i>	Present		7.17	92.83	
	2021-2040	245	0.11	99.89	99
	2021-2040	585	0.06	99.94	99
<i>Squalius</i>	Present		87.07	12.93	
	2021-2040	245	81.37	18.63	19
	2021-2040	585	79.40	20.60	21
	2041-2060	245	67.30	32.70	33
	2061-2080	245	58.54	41.46	42
	2081-2100	245	79.16	20.84	21

* Percentage reduction of suitable areas for each timeframe modelled when compared with current predictions.

Figure S1. Cross-entropy for each number of K ancestral populations inferred with sNMF.

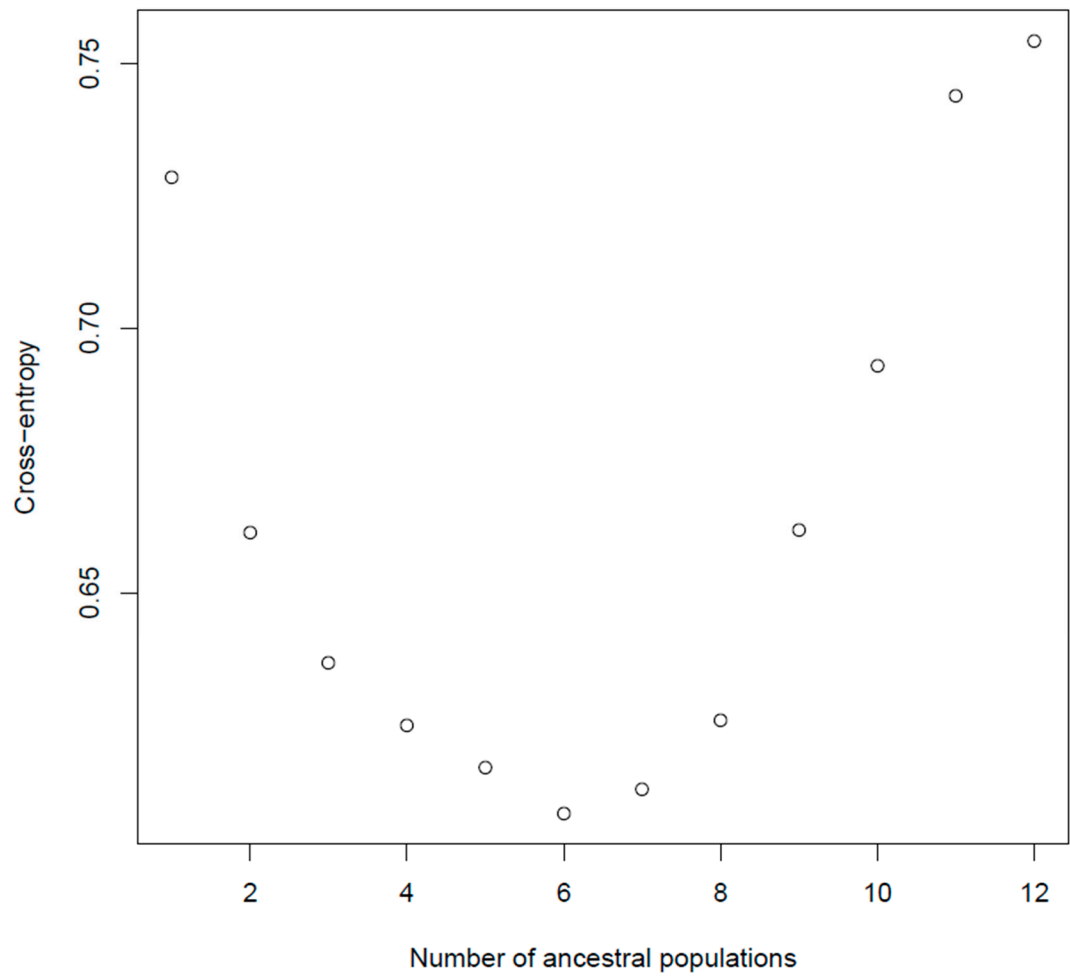
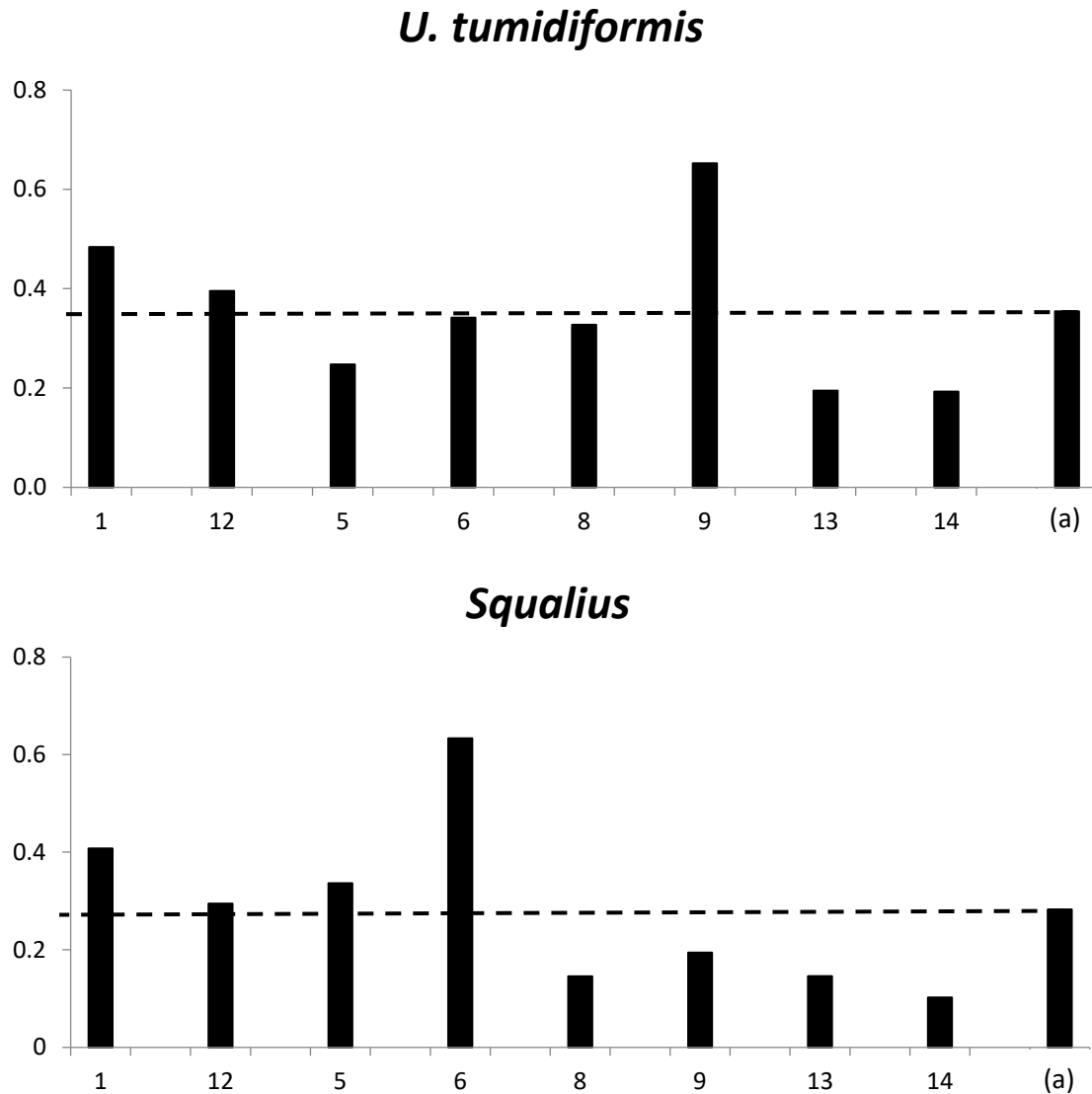


Figure S2 - Relative importance of the eight environmental variables (BIO1 = Annual Mean Temperature; BIO12 = Annual Precipitation; BIO5 = Max Temperature of Warmest Month; BIO6 = Min Temperature of Coldest Month; BIO8 = Mean Temperature of Wettest Quarter; BIO9 = Mean Temperature of Driest Quarter; BIO13 = Precipitation of Wettest Month; BIO14 = Precipitation of Driest Month) used to predict the distribution of *U. tumidiformis* and its hosts.



(a) The last column represents the mean value of relative importance obtained from eight different modelling algorithms. Individual variables with relative importance above this value (dashed line) were assumed as important in determining suitability according to the models used.