Journal of Zoological Systematics and Evolutionary Research

SUPPORTING INFORMATION

Evolutionary history of a tropical affinity species expanding across the Mediterranean Basin

Francisco Amorim, Orly Razgour, Vanessa Mata, Susana Lopes, Raquel Godinho, Carlos Ibáñez, Javier Juste, Stephen J. Rossiter, Pedro Beja, Hugo Rebelo

Table S1 – Mitochondrial primers designed for *COI*, *ATP6*, *CytB* and *D-Loop* regions. Forward and reverse primers are presented, as well as the amplicon position in the mitogenome and its length (excluding primers). Internal primers designed for highly degraded museum samples are also presented (*COI-mini* and *D-Loop-mini*).

Fragment	**Primer (5'-3')**		**Mitogenome position[1]**	**Length (bp)**
COI	Tten-COI-F	CCCAYGCAGGAGCTTCAGTT	5761-6373	613
	Tten-COI-R	TGTTAGGCCYCCTACTGTGA		
ATP6	Tten-ATP6-F	CAAMCAACTCCTTGAGAYAYC	7931-8620	690
	Tten-ATP6-R	CTATATGRTARGCATGTGTTTGG		
CytB	Tten-Cytb-F	CGCAGCCATAGYCATAGTC	14740-15303	564
	Tten-Cytb-R	CAAGACCGRRGTAATGTATATATAC		
D-Loop	Tten-Dloop-F	CGTAAACGGCATCTGGTTCT	16011-16349	339
	Tten-Dloop-R	GCGTATCTGATTRTGTTATATGTCCTG		
COI-mini	Tten-COI-mini-F	RGTCCTAGCAGCMGGAATCAC	621-718	98
	Tten-COI-mini-R	GGATTAGAATATAAACTTCGGGGTG		
D-Loop-mini	Tten-Dloop-mini-F	AATTAARGGTCCCAGGACATTCA	16278-16349	72
	Tten-Dloop-R	GCGTATCTGATTRTGTTATATGTCCTG		

(1) Mata *et al.* (2017) First complete mitochondrial genomes of molossid bats (Chiroptera: Molossidae). Mitochondrial DNA Part B, 2(1), 152–154. https://doi.org/10.1080/23802359.2017.1298419.

Table S2 – Microsatellite loci and primer sequences used in the study, including information on motif, concentration of forward (F) primer and tail, allele size range, observed (H_{obs}), expected (H_{exp}) heterozygosity, P values for deviation from Hardy-Weinberg equilibrium (P(HW)), and estimated frequency of null alleles. Concentration of reverse primer was 10x the forward primer.

	Primer name	motif	Sequence (5'-3'), F: 10 µM, R: 100 µM	Tail	Concentration (pM) Primer	Concentration (pM) Tail	Allele size range	Number of alleles	H_{obs}	H_{exp}	P (HW)	F (null allele)
Mix1	**TAD1**	ac	F: GGCTACCTGAGCTTCTTAGTCTTC R: TGTTGCTTTCTAGGCTGTTGC	FAM	0.11	4.00	145-183	15	0.86	0.87	0.073	0.01
	TAD5	tatc	F: ATTCCTGGTCAGTGCATCGT R: TCACTGACACAGATGATAATAGAGACA		0.19		322-354	9	0.49	0.80	0.00[*]	0.20
	TAD6	ac	F: TTCTTAACCACACTGGCCCT R: TGCAAGCCAAACATAAGTCG		0.11		221-247	13	0.67	0.88	0.00[ǂ]	0.13
	TAD23	ctat	F: GAGCACAGTTCTCCTGATTGG R: TTCAGTCCTAAACTAGGATCAAAGAC	VIC	0.08	0.80	132-156	7	0.76	0.76	0.73	0.00
	TAD2	ac	F: TCATCGGTATCCTCAAAGAAATG R: GCATTGATGGAGACAGAAATATAAA	NED	3.20	33.20	177-199	11	0.81	0.85	0.11	0.03
	TAD4	ag	F: ACCTTTGATTACGCAGCACC R: AATTTTGAGAAGTTTGAATGCCA		0.12		99-127	14	0.73	0.77	0.08	0.02
	TAD7	ac	F: AAAAGCAGATGGTCAGCAGG R: TGCCAGGCCTAAGATGCTAC	PET	0.12	1.20	96-126	13	0.85	0.84	0.52	0.00
Mix2	**TAD13**	atct	F: TTTCTCTTTTATGGGAAATCCA R: CCTCAGAAGTAGCCGCAAGT	FAM	0.08	1.60	183-219	10	0.78	0.81	0.46	0.02
	TAD19	ca	F: TGTTCAATTACACTGTTACATGCG R: CAACCTTGGTCAATTTTAAAGC		0.08		104-130	10	0.79	0.78	0.00[ǂ]	0.00
	TAD9	tg	F: CCCTTCTGCCAGTCCAAATA R: CATTGCCACCAAATTCCTCT	VIC	0.08	1.60	136-154	10	0.63	0.78	0.00[*]	0.09
	TAD18	gaag	F: CCTCAGCTATCTTGGGATTGTC R: AAGCACCACCGATCTGAATC		0.08		186-242	13	0.88	0.86	0.76	0.00
	TAD26	tg	F: TTGAAACACCACATTCAGTCG R: AGTTGAATAAGTTAATTGAAATAAGCC	NED	0.2	2.00	138-156	9	0.79	0.81	0.53	0.01
	TAD12	ag	F: CGCTAATCTCTGCTAATCTCTGG R: GCCTTACCTATACTTCGCTTGG	PET	0.12	3.60	172-200	14	0.83	0.83	0.18	0.00
	TAD24	ag	F: GATTGTTCCATTGTTATTCCTATTTG R: TGACTGTGGGCAGGTTGTTA		0.24		108-138	14	0.69	0.78	0.02[ǂ]	0.05

[*] Deviation from Hardy-Weinberg equilibrium (P < 0.05 no correction for multiple tests applied) in more than two out of the eight populations. These loci were removed from further analysis

[ǂ] Deviation from Hardy-Weinberg equilibrium (P < 0.05 no correction for multiple tests applied) in two or less populations. These loci were included in further analysis.

Table S3 – ABC Scenarios parameters for inferring 1) colonization analysis; and 2) demographic history analysis. Only the best scenarios for each analysis are presented. Parameters as follow: N – effective population size parameters, where *e* denotes equal population size for all populations, *Ib* and *Fr* correspond to Iberian and French populations respectively; t – time as number of generations where numbers represent the scenario and time step; r – admixture rate corresponding to time step *t*.

Colonization analysis (SC2)

Parameters	Priors		Results				
	Conditions	Distribution [min-max]	mean	median	mode	*Q5*	*Q95*
N_e		Uniform [$1E^3$ - $1E^6$]	$5.16E^4$	$4.99E^4$	$4.68E^4$	$3.18E^4$	7.70^4
$t_{2.1}$		Uniform [$1E^4$ - $2E^5$]	$6.06E^4$	$4.51E^4$	$2.50E^4$	$1.65E^4$	$1.60E^5$
$t_{2.2}$	$t_{2.2} < t_{2.1}$	Uniform [$1E^3$ - $2E^4$]	$1.27E^4$	$1.29E^4$	$1.32E^4$	$5.62E^3$	$1.92E^4$
$t_{2.3}$	$t_{2.3} \leq t_{2.2}$	Uniform [$1E^3$ - $2E^4$]	$3.93E^3$	$3.32E^3$	$2.53E^3$	$1.45E^3$	$8.56E^3$
$t_{2.4}$	$t_{2.4} \leq t_{2.1}$	Uniform [$1E^3$ - $2E^4$]	$5.34E^3$	$4.34E^3$	$3.04E^3$	$2.01E^3$	$1.26E^4$
$t_{2.5}$	$t_{2.5} \leq t_{2.1}$	Uniform [$1E^3$ - $2E^4$]	$6.16E^3$	$5.15E^3$	$3.62E^3$	$2.33E^3$	$1.39E^4$
$t_{2.6}$	$t_{2.6} < t_{2.4}$ $t_{2.6} \leq t_{2.5}$	Uniform [$1E^3$ - $2E^4$]	$1.96E^3$	$1.76E^3$	$1.53E^3$	$1.17E^3$	$3.37E^3$
$r_{2.1}$		Uniform [0.001 - 0.999]	4.72E-1	4.60E-1	3.63E-1	7.85E-2	9.03E-1
$r_{2.3}$		Uniform [0.001 - 0.999]	5.48E-1	5.57E-1	6.25E-1	1.66E-1	8.89E-1
$r_{2.6}$		Uniform [0.001 - 0.999]	5.07E-1	5.12E-1	5.47E-1	9.49E-2	9.01E-1

Demographic history analysis (SC2)

Parameters	Priors		Results				
	Conditions	Distribution [min-max]	mean	median	mode	*Q5*	*Q95*
N_{Fr1}		Uniform [$1E^2$ - $5E^5$]	$1.14E^3$	$8.90E^2$	$5.80E^2$	$2.92E^2$	$2.73E^3$
N_{Fr2}	$N_{Fr2} > N_{Fr1}$	Uniform [$1E^2$ - $1E^6$]	$1.76E^6$	$1.88E^6$	$2.00E^6$	$1.05E^6$	$1.99E^6$
N_{Ib}	$N_{Ib} > N_{Fr1}$	Uniform [$1E^2$ - $1E^6$]	$3.22E^5$	$2.41E^5$	$2.74E^4$	$2.74E^4$	$8.62E^5$
$t_{2.1}$		Uniform [10 - $5E^3$]	$1.45E^3$	$1.09E^3$	$7.93E^2$	$3.60E^2$	$3.84E^3$

Table S4 – Genetic differentiation at the mitochondrial DNA level between *Tadarida teniotis* geographical groups based on θ_{ST} values.

	Canary	**Iberia**	**Morocco**	**France**	**Italy**	**Greece**	**Anatolia**
Morocco	0.094						
Iberia	0.065	0.023					
France	**0.442****	**0.049***	**0.068***				
Italy	**0.343****	**0.093***	**0.105****	0.013			
Greece	**0.179***	0.016	**0.181***	0.031	**0.13***		
Anatolia	**0.831***	**0.72***	**0.843****	**0.819***	**0.533***	**0.841****	
Middle-East	**0.75****	**0.685****	**0.816****	**0.744****	**0.507***	**0.79****	0.142

*p<0.05; **p<0.001

Table S5 – Genetic differentiation at nuclear (microsatellite) level between *Tadarida teniotis* populations based on F_{ST} values. The bottom diagonal includes all individuals, while in the top diagonal values of F_{ST} were calculated after the removal of close relatives (TrioML > 0.5). Values above 0.1 (moderate to strong population differentiation) are marked in bold

	Canary	**Morocco**	**Iberia**	**France**	**Italy**	**Greece**	**Anatolia**	**Middle-east**
Canary		0.093	0.093	0.066	0.067	**0.144**	**0.130**	0.093
Morocco	0.067		-0.004	-0.002	-0.005	0.039	0.012	0.014
Iberia	0.067	-0.005		0.000	0.012	0.046	0.025	0.037
France	0.042	-0.002	0.000		0.005	0.038	0.005	0.004
Italy	0.047	-0.005	0.011	0.005		0.055	0.010	0.027
Greece	**0.106**	0.039	0.046	0.038	0.055		0.020	0.003
Anatolia	0.093	0.012	0.025	0.005	0.010	0.020		0.023
Middle East	0.074	0.014	0.037	0.004	0.027	0.003	0.023	

Table S6 – Mean LnP(K) and Delta K (Evanno method) results for the Bayesian clustering analysis with prior population information, as implemented in STRUCTURE HARVESTER. Bold highlights the largest value in the Delta K for K = 3.

K	**Reps**	**Mean LnP(K)**	**Stdev LnP(K)**	**Ln'(K)**	**|Ln''(K)|**	**Delta K**
1	10	-5842.78	0.38	NA	NA	NA
2	10	-5799.81	11.54	42.97	30.37	2.63
3	**10**	**-5787.21**	**12.01**	**12.60**	**108.49**	**9.03**
4	10	-5883.10	34.42	-95.89	103.69	3.01
5	10	-5875.30	36.12	7.80	38.18	1.06
6	10	-5905.68	43.03	-30.38	24.09	0.56
7	10	-5911.97	66.55	-6.29	15.39	0.23
8	10	-5902.87	40.91	9.10	32.29	0.79
9	10	-5926.06	81.41	-23.19	18.08	0.22
10	10	-5967.33	86.42	-41.27	NA	NA

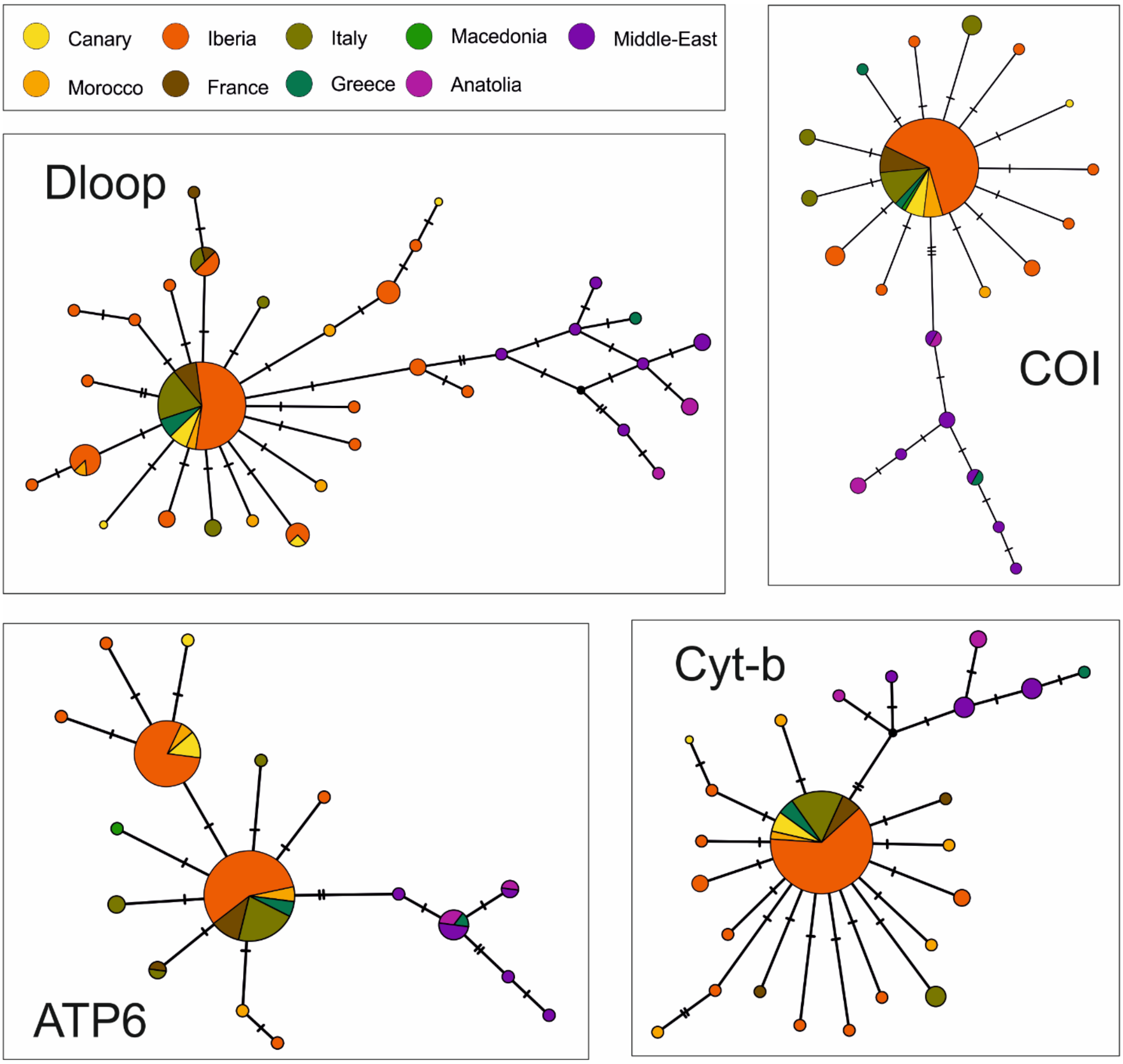

Figure S1 –Median-joining haplotype networks for *T. teniotis* mtDNA coding regions (cytochrome c oxidase subunit I, ATP6, and cytochrome b) and noncoding region (control region Dloop). All concatenated mtDNA sequences (n = 109) were used in the Median-joining networks in which branch lengths are not proportional to base-pair changes.

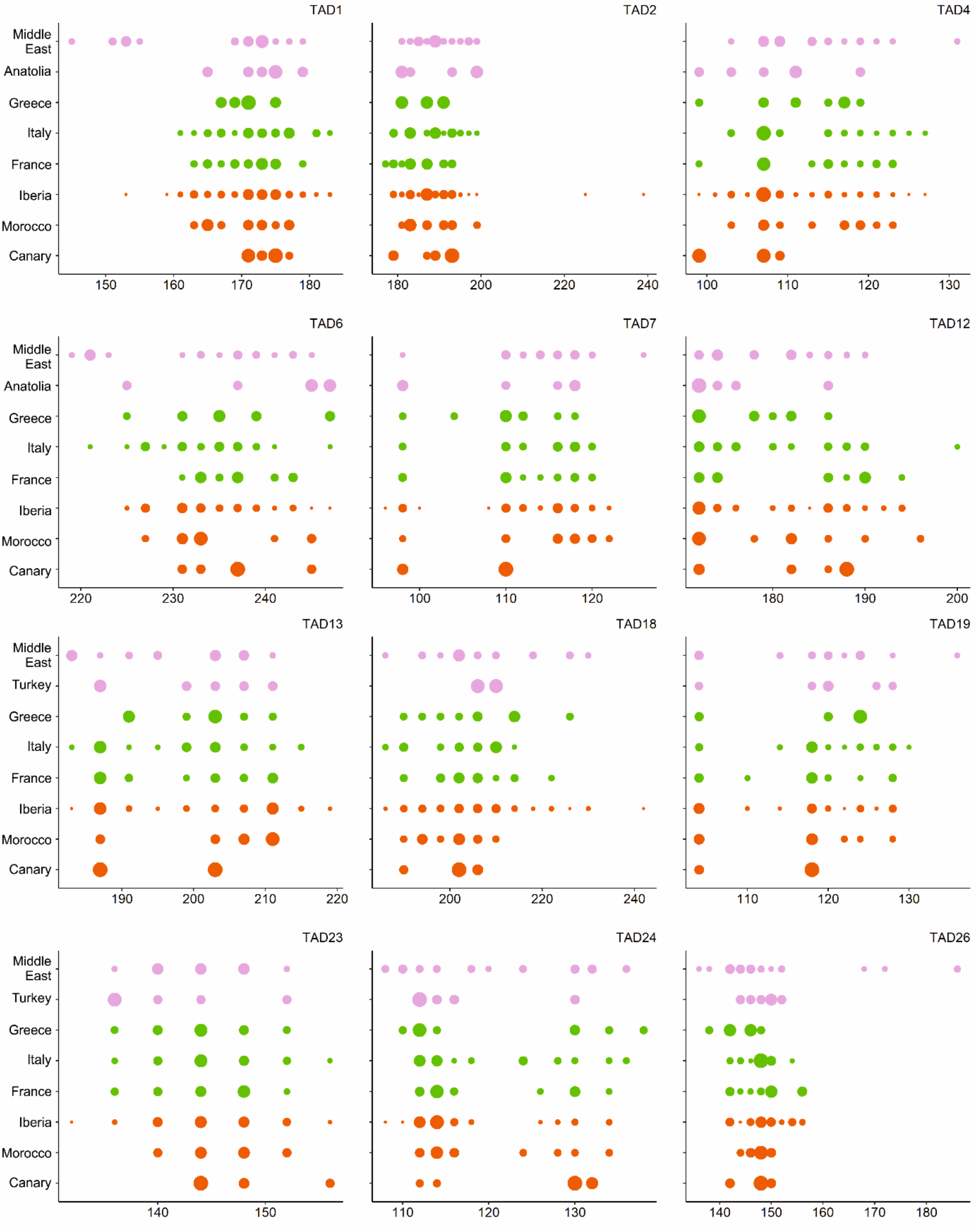

Figure S2 – Allele frequency per population and per locus corrected by sample size

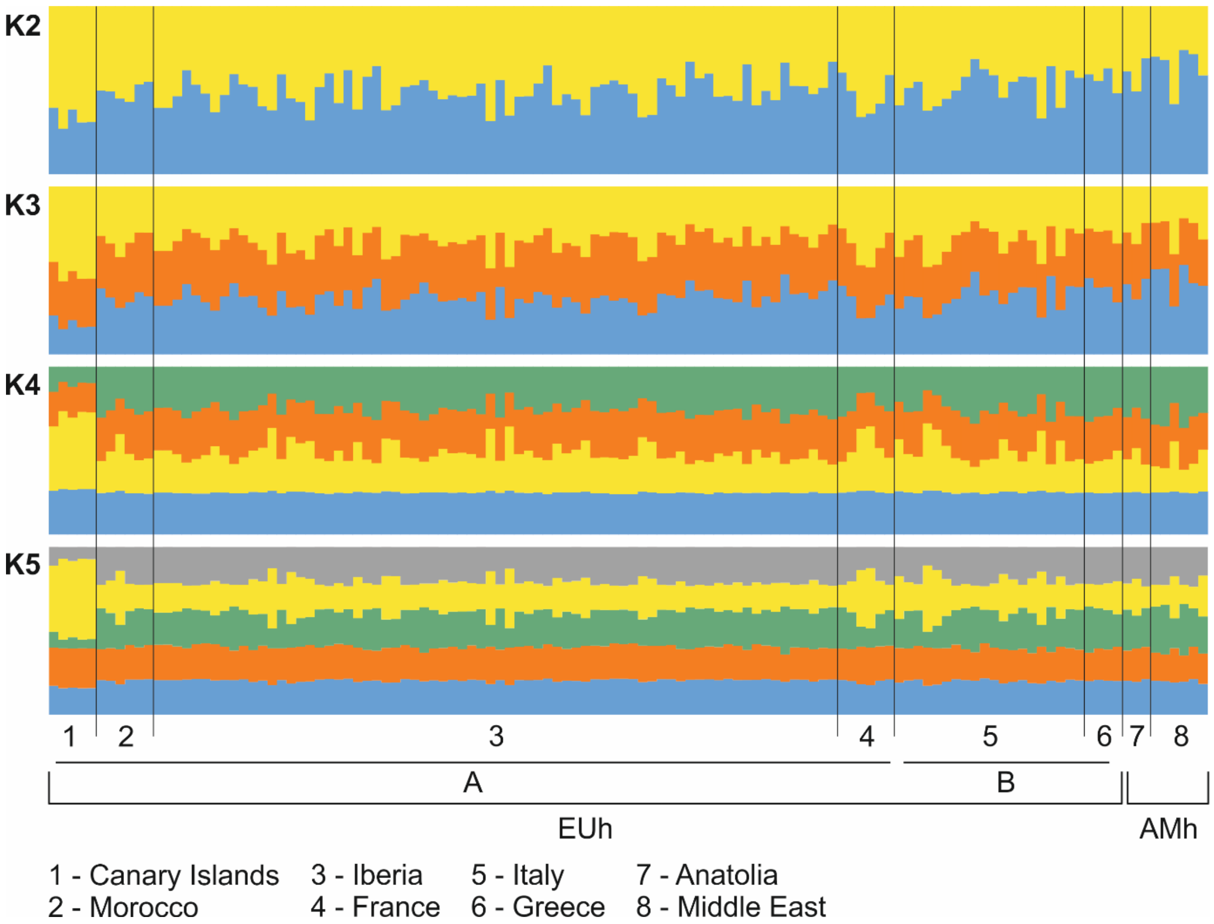

Figure S3 – *Tadarida teniotis* population structure based on the microsatellite data set. (A) Individual membership assignment plots resulting from STRUCTURE analysis without prior population information including all samples (K = 2 to K = 5).

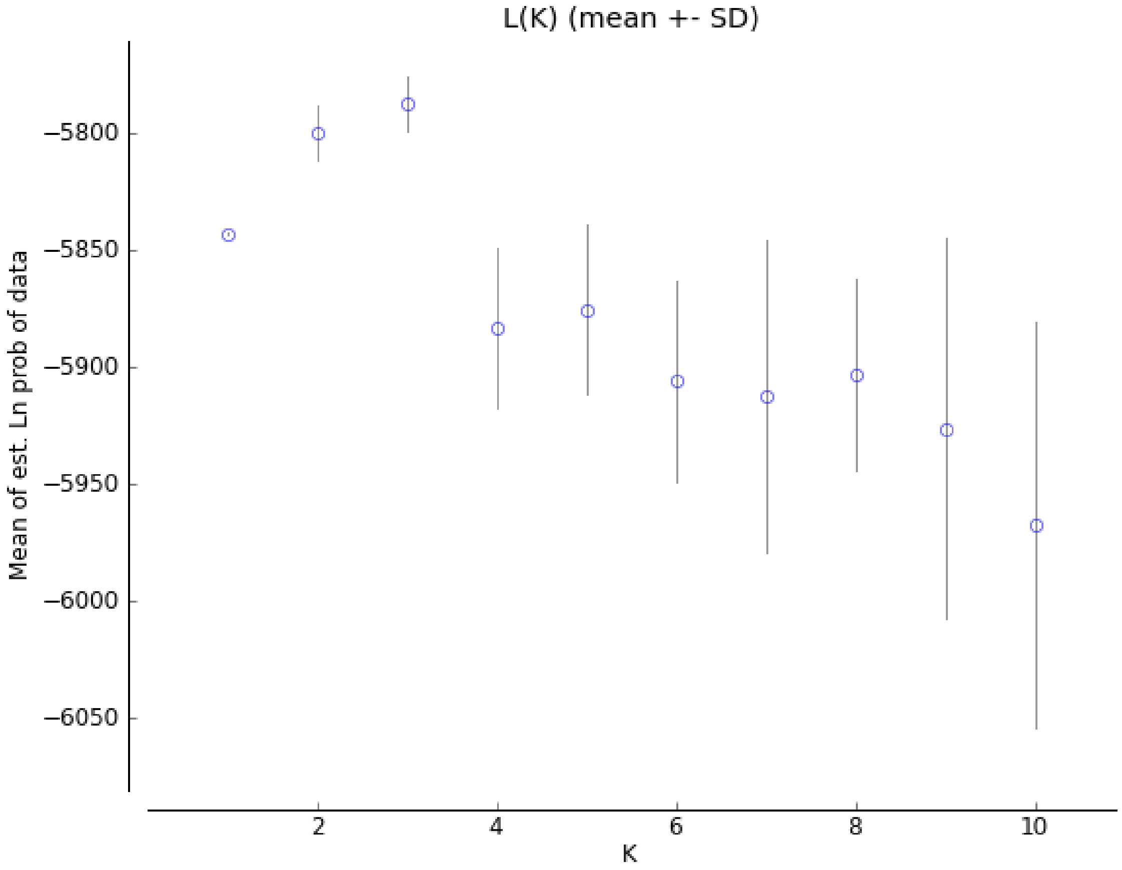

Figure S4 – Results of the STRUCTURE analysis for the *Tadarida teniotis* microsatellite dataset using a prior for population information. Log-likelihood probability for each K from 1–10 clusters, showing means and standard deviations of the ten replicate runs for each number of clusters.

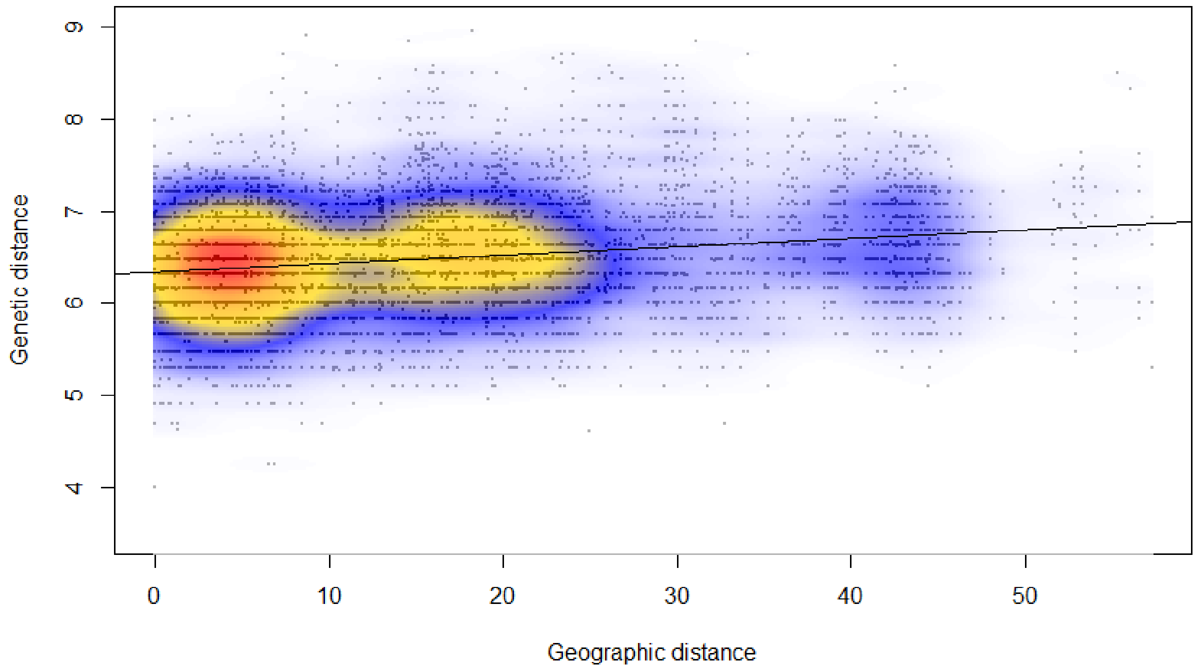

Figure S5 – Scatterplot of genetic and geographic distance. Colours represent points' density. The observed discontinuities in the cloud of points suggest a pattern of IBD, confirmed by the Monte-Carlo test with 1000 permutations (p = 0.001).

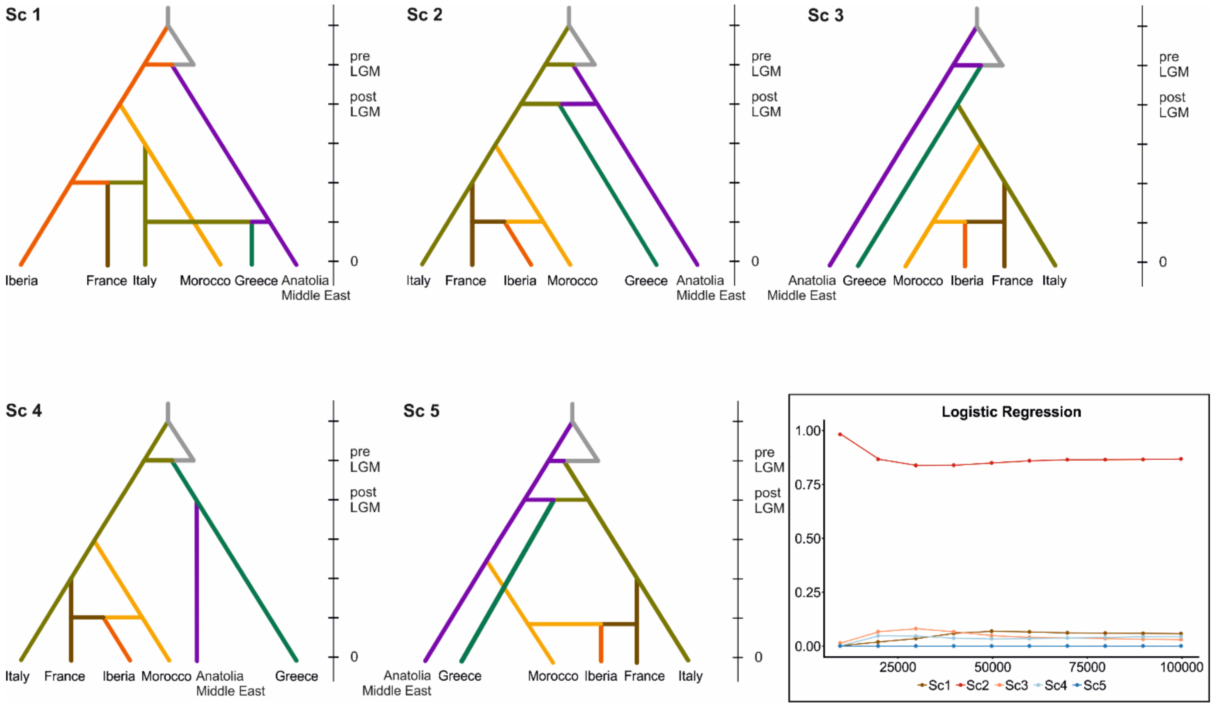

Figure S6 – Schematic representation of colonization scenarios compared using Approximate Bayesian Computation and the logistic regression of the posterior probabilities of the five scenarios as a function of number of simulated datasets. For further details, see Methods S1 and Table S4.

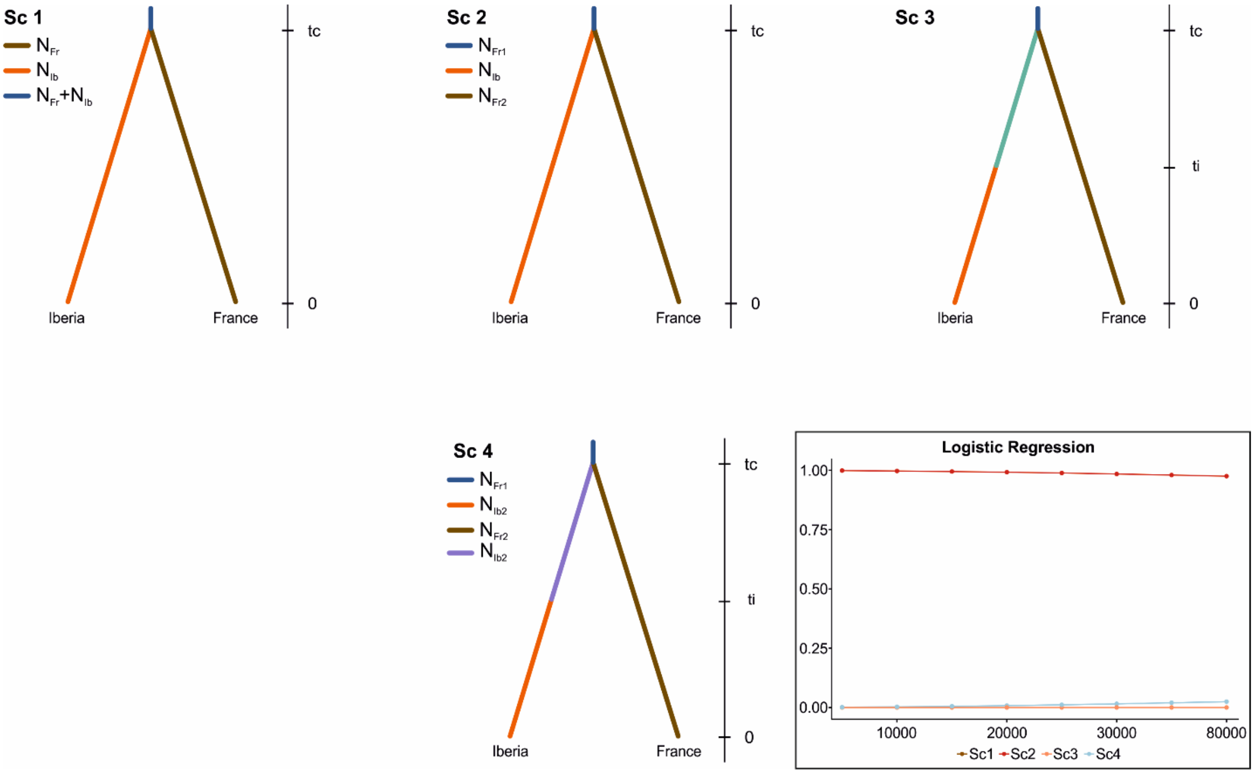

Figure S7 – Schematic representation of the demographic history scenarios of the western range compared using Approximate Bayesian Computation and the logistic regression of the posterior probabilities of the five scenarios as a function of number of simulated datasets. Subscripts *Ib* and *Fr* correspond to Iberian and French populations respectively. For further details, see Methods S1 and Table S4.