Support Information of the United States of the United States of the United States of the United States of the U

Hadid et al. 10.1073/pnas.1322301111

SI Text

The microsite dubbed Evolution Canyon (EC) in lower Nahal Oren, Mount Carmel, Israel (32°43′N/34°58′E), is a Plio-Pleistocene canyon. It is a natural laboratory of evolution in action, a model for studying biodiversity evolution, adaptation, and speciation across life (a full reference list is available at [http://](http://evolution.haifa.ac.il) evolution.haifa.ac.il, Nevo list of "Evolution Canyon" publications). The EC rocks are Upper Cenomanian limestones; the regional climate is Mediterranean (1) (mean annual rainfall ca. 600 mm, potential evapotranspiration 1,700 mm, and mean August and January temperatures 28 °C and 13 °C, respectively) (2). The soil is terra rossa on both slopes with more organic material on the "European" north-facing slope (3) of the canyon. The EC runs in an east–west direction. The opposite slopes are separated by only 100 m at the bottom and 400 m at the top. They receive different insulation owing to their orientation and dips (4). The higher insulation is on the south-facing slope, or "African" slope (AS), which is a tropical, hot, xeric, open park-forest of evergreen Ceratonia siliqua–Pistacia lentiscus, with dominant savanna-like grasslands. The lower insulation is on the north-facing slope, or "European" slope (ES), which is a temperate, cool, mesic, shady, dense, Mediterranean maquis forest of evergreen Quercus calliprinos and deciduous Pistacia palaestina (5). The AS is generally richer in terrestrial species and genetic diversity than the ES, which is richer in species associated reproductively with water, and are, generally, less genetically diverse (1, 6).

SI Materials and Methods

Molecular Analysis. Isolation of genomic DNA and mtDNA sequencing. We extracted DNA from tails or stools of the captured animals (permission no. 31933/2008 by the Israel Nature and Park Authority) by using DNeasy Tissue Kit (Qiagen) or DNAzol kit, following the manufacturer's instructions (Table S1). Before sequencing, we identified the longest possible mtDNA fragments in the Acomys cahirinus genome by means of the long-arm PCR (LA-PCR). The used templates were reported in GenBank (12S rRNA, HQ652130; NADH3+ tRNA, Arg+NADH4L+ NADH4- U83803; Cyt b EF187818; and D-loop, FJ415595). The LA-PCR was conducted in a 20-μL reaction volume that contains 10 μL ready-mix, 20 pM of each primer, and about 60 ng genome DNA. The PCR conditions were set as 94 °C for the first 5 min, followed by 35 cycles of 94 °C denaturation for 30s, 60 °C annealing for 60s, and 68 °C extension for 3 min. Following LA-PCR, the sequencing of the complete mtDNA molecule was done stepwise with each step represented by about 500- to 600 bp-long segments. The procedure resulted in 67 partly overlapping fragments, which were used to infer the complete sequence (Tables S2 and S9). Sequencing was performed on a 3100 DNA Analyzer (Applied Biosystems).

Amplified fragment length polymorphism. To unfold the amplified fragment length polymorphism (AFLP), the isolated genomic DNA was digested using the enzyme combination EcoR1 and Mse1, then ligated to double-stranded EcoRI (E-) and MseI (M-) adaptors (7). The resulting fragments were preamplified with nonselective primers, where the ligated adaptors served as target sites for primer annealing. Eight selected primer combinations were used for AFLP amplification. The selected EcoRI (E-) primers were labeled with a florescent dye (6-Fam, Vic, Ned, and Pet). PCR reactions were carried out according to the original protocol (7). Amplification products were visualized by Sequencher 3130X (Applied Biosystems). Allele identification and genotyping were determined directly from the chromatographs using Genemapper software (Applied Biosystems). Amplification products were scored as discrete character states (presence/ absence) and recorded as band frequencies (allele frequency).

Bioinformatics Software and Statistical Analyses. mtDNA sequencing. The complete mtDNA sequences obtained and Cytb sequences downloaded from GenBank (Table S6) were aligned by means of ClustalW (8), which is a part of the Mega4 (9) package. Statistically, the most fitting substitution models to data were identified in Hyphy (10) by using the Akaike Information Criterion (11). Maximum likelihood phylogram and chronogram were constructed by means of PhyML_3 (12). SplitsTree software (13) was used for network analysis that is present in the form of spinning networks. Haplotype and nucleotide diversities, Hd and Pi, respectively, were estimated by means of the program DNAsp (14). The differences between haplotype frequencies of the pooled sample and pairwise comparisons of haplotype frequencies were tested by using binomial probabilities (15). RDP3 (16) was used to test for the presence of homologous recombinations in the sequences. The presence of negative or positive selection on the constructed neighbor-joining phylogenetic tree was done by means of the tests available in the Datamonkey softwares (17). The PARRIS likelihood test (18) was used for identifying sites with $dN/ds > 1$, a synonymous rate variation. The integrative test for positive or negative selection was performed by the same software. Cohen's unweighted kappa (19) was used to measure concordance in a categorical setting. Median joining network was constructed by means of median joining network software (www.fluxus-engineering.com/).

AFLP and mtDNA analyses as binary data. The AFLP genetic data were tested for consistency by means of the AFLPsurv (20). Hickory (21) was used to estimate interslope distance (Θ^I and Θ^{II}) and genetic diversity (Hs). Θ^I corresponds directly to Wright's FST, ΘII that measuring divergence among contemporaneous populations is similar to Nei's GST (22), and Hs is defined as average panmictic heterozygosity. The credible intervals of these parameters were estimated by the setting of 5,000 cycles for burning and 100,000 cycles of Markov chain Monte Carlo. The interslope differences in polymorphic loci (Pb) and allelic richness (P) were estimated by the program Aflpdiv based on rarefaction-based correction for sample sizes (23). We also constructed a dendogram based on AFLP data by means of SplitsTree version 4 (24). To indicate the interslope differential selection we used MatSAM (25).

5. Nevo E, Fragman O, Dafni A, Beiles A (1999) Biodiversity and interslope divergence of vascular plants caused by microclimatic differences at "Evolution Canyon", Lower Nahal Oren, Mount Carmel, Israel. Isr J Plant Sci 47(1):49–59.

- 6. Pavlícek T, Nevo E (2008) Local species richness distribution at "Evolution Canyon" microsite, Mt. Carmel, Israel. Ecologia Mediterranea 34:25–34.
- 7. Vos P, et al. (1995) AFLP: A new technique for DNA fingerprinting. Nucleic Acids Res 23(21):4407–4414.
- 8. Thompson JD, Gibson TJ, Higgins DG (2002) Multiple sequence alignment using ClustalW and ClustalX. Current Protocols in Bioinformatics (Wiley Online Library), 00:2.3.1–2.3.22.

^{1.} Nevo E (1995) Asian, African and European biota meet at "Evolution Canyon", Israel: Local tests of global biodiversity and genetic diversity patterns. Proc Biol Sci 262:149–155.

^{2.} Goldreich Y (2003) The Climate of Israel: Observation, Research, and Application (Springer, New York).

^{3.} Nevo E, et al. (1998) Edaphic interslope and valley bottom differences at "Evolution Canyon", Lower Nahal Oren, Mount Carmel, Israel. Catena 33:241–254.

^{4.} Pavlícek T, Sharon D, Kravchenko V, Saaroni H, Nevo E (2003) Microclimatic interslope differences underlying biodiversity contrasts in "Evolution Canyon", Mt. Carmel, Israel. Isr J Earth Sci 52:1–9.

- 9. Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol Biol Evol 24(8):1596–1599.
- 10. Pond SL, Frost SD, Muse SV (2005) HyPhy: Hypothesis testing using phylogenies. Bioinformatics 21(5):676–679.
- 11. Akaike H (1974) A new look at the statistical model identification. IEEE Trans Automat Contr 119:716–723.
- 12. Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol 52(5):696–704.
- 13. Huson DH (1998) SplitsTree: Analyzing and visualizing evolutionary data. Bioinformatics 14(1):68–73.
- 14. Librado P, Rozas J (2009) DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25(11):1451–1452.
- 15. Lowry R (2011) Binomial probabilities. VassarStats: Website for Statistical Computations. Available at<http://vassarstats.net/>.
- 16. Martin DP, et al. (2010) RDP3: A flexible and fast computer program for analyzing recombination. Bioinformatics 26(19):2462–2463.
- 17. Delport W, Poon AFY, Frost SDW, Kosakovsky Pond SL (2010) Datamonkey 2010: A suite of phylogenetic analysis tools for evolutionary biology. Bioinformatics 26(19):2455–2457.
- 18. Scheffler K, Martin DP, Seoighe C (2006) Robust inference of positive selection from recombining coding sequences. Bioinformatics 22(20):2493–2499.
- 19. Lowry R (2011) Kappa as measure concordance in categorical sorting. Available at [http://faculty.vassar.edu/lowry/VassarStats.html.](http://faculty.vassar.edu/lowry/VassarStats.html)
- 20. Vekemans X (2002) Aflp-surv 1.0. A Program for Genetic Analysis with AFLP (and RAPD) Population Data (Laboratoire de Génétique et d'Ecologie Végétales, Université Libre de Bruxelles, Brussels).
- 21. Holsinger KE, Lewis PO (2007) Hickory: A Package for Analysis of Population Genetic Data (Department of Ecology and Evolutionary Biology, University of Connecticut, Storrs, CT).
- 22. Holsinger KE (1999) Analysis of genetic diversity in geographically structured populations: A Bayesian perspective. Hereditas 130:245–255.
- 23. Coart E, Van Glabeke S, Petit RJ, Van Bockstaele E, Roldán-Ruiz I (2005) Range wide versus local patterns of genetic diversity in hornbeam (Carpinus betulus L.). Conserv Genet 6:259–273.
- 24. Huson DH, Bryant D (2008) User Manual for SplitsTree4 V4. 10 (Image, Rochester, NY), pp 1–57.
- 25. Joost S, Kalbermatten M, Bonin A (2008) Spatial analysis method (sam): A software tool combining molecular and environmental data to identify candidate loci for selection. Mol Ecol Resour 8(5):957–960.
- 26. Schluter PM, Harris S (2006) Analysis of multilocus fingerprinting data sets containing missing data. Mol Ecol Notes 6:569–572.

Fig. S1. (A) G-banded karyotype of a male Acomys from the south-facing slope of Nahal Oren: 2n = 38, Number of fragments (NF) = 70. All 34 karyotyped animals in both slopes had the same structure (G, C, and AgNOR banding) of karyotype based on 34 karyotypes at EC. (B) An example of AFLP profile of two animals representing each slope. Two upper profiles represent AS animals (A41 and A52) and the two lower profiles represent the ES animals (E23 and E2). (C) AFLP-based neighbor joining dendogram of A. cahirinus samples from EC made by FAMD (26). A, "African" slope (AS); E, "European" slope (ES).

Serial	Original			Date	DNA
no.	serial	Station	Gender	(catch)	purification
1	A01	2	М	10/29/07	DNAzole-tail
2	A02	2		5/26/09	DNAzole-tail
3	A03	$\overline{2}$	м	10/19/07	Stool kit
4	A04	2	м	7/2/07	DNAzole-tail
5	A05	1	F	5/28/04	Stool kit
6	A06	3	м	9/4/04	DNAzole-tail
7	A07	2	F	4/9/08	Stool kit
8	A08	2	м	10/31/07	Stool kit
9	A09	3	м	10/31/07	Stool kit
10	A10	2	м	2.07.07	Stool kit
		2			DNAzole-tail
11	A100			5/26/09 5/26/09	
12	A101	2			DNAzole-tail
13	A102	2		5/26/09	DNAzole-tail
14	A11	2	М	4/9/08	DNAzole-tail
15	A12	$\overline{2}$	м	4/9/08	Stool kit
16	A13	$\overline{2}$	м	4/9/08	Stool kit
17	A14	3	м	4/10/08	Stool kit
18	A14	2		5/26/09	DNAzole-tail
19	A15	3		4/10/08	Stool kit
20	A16	3	м	4/17/08	Stool kit
21	A20	2		5/26/09	DNAzole-liver
22	A21	2		5/26/09	DNAzole-tail
23	A32	2		5/26/09	DNAzole-tail
24	A33	$\overline{2}$		5/26/09	DNAzole-tail
25	A40	2		5/26/09	DNAzole-tail
26	A41	2		5/26/09	DNAzole-tail
27	A50	$\overline{2}$		5/26/09	DNAzole-tail
28	A51	2		11/26/08	DNAzole-tail
29	A52	2		11/26/08	DNAzole-tail
30	A53	2		12/1/08	DNAzole-tail
31	A54	2		12/1/08	DNAzole-tail
32	A55	2		12/1/08	DNAzole-tail
33	A56	2		12/1/08	DNAzole-tail
34	A57	2		12/1/08	DNAzole-tail
35	A58	2		12/1/08	DNAzole-tail
36	A59	2		12/1/08	DNAzole-tail
37	A60	2		9/16/09	DNAzole-tail
38	A61	$\overline{2}$		9/16/09	DNAzole-tail
39	E ₁	6	F	5/16/07	Stool kit
40	E10	5	м	5/16/07	DNAzole-tail
41	E11	7	м	5/29/04	Stool kit
42	E12	7	м	6/30/07	DNAzole-Tail
43	E13	7	м	7/26/04	Stool kit
44	E14	5	м	11/29/05	DNAzole-tail
45	E15	7	F	5/15/04	Stool kit
46	E16	6	F	8/2/05	Stool kit
47	E20	6		5/26/09	DNAzole-tail
48	E21	5		5/26/09	DNAzole-tail
49	E22	5		5/26/09	DNAzole-tail
50	E23	6		5/26/09	DNAzole-tail
51	E2	6		5/26/09	DNAzole-tail
52	E3		М	6/30/07	Stool kit
53	E8		М	6/14/07	
54	E9	6	м	5/16/07	Stool kit

Table S1. Acomys animals examined, laboratory serial numbers, station locality, sex, collection date, and DNA extraction method

mtDNA			No. of variable
sequence	Position	Base pairs	sites
tRNA-Phe	$1 - 67$	67	
12SrRNA	68-1013	946	2
tRNA-Val	1014-1080	67	
16SrRNA	1081-2640	1,560	2
tRNA-Leu(UUR)	2641-2715	75	
NADH1	2716-3672	957	4
tRNA-Ile	3671-3739	69	1
tRNA-GIn	3737-3807	71	
tRNA-Met	3812-3880	69	
NADH ₂	3881-4922	1,042	2
tRNA-Trp	4918-4982	65	
tRNA-Ala	4986-5054	69	
tRNA-Asn	5057-5127	71	1
tRNA-Ile	5128-5158	31	
tRNA-Cys	5159-5219	61	
tRNA-Tyr	5227-5292	66	
COI	5294-6838	1,545	3
tRNA-Ser(UCN)	6836-6804	69	
tRNA-Asp	6908-6975	68	
COII	6977-7660	684	1
tRNA-Lys	7664–7726	63	
ATP8	7728-7931	204	1
ATP6	7889-8569	681	3
COIII	8569-9352	784	1
tRNA-Gly	9353-9420	68	
NADH3	9421-9768	348	
tRNA-Arg	9771-9837	67	
NADH4L	9838-10134	297	1
NADH4	10128-11499	1,372	6
tRNA-His	11500-11569	70	
tRNA-Ser(AGY)	11570-11626	57	1
tRNA-Leu(CUN)	11627-11697	71	
NADH5	11698-13530	1,833	3
NADH ₆	13509-14029	521	1
tRNA-Glu	14030-14102	73	1
cytb	14103-15243	1,141	5
tRNA-Thr	15244-15311	68	
tRNA-Pro	15312-15379	68	
D-loop	15380-16218	839	5

Table S2. Identification of mtDNA sequences and their basic parameters

Table S3. Identified haplotypes from complete mtDNA sequences of 17 A. cahirinus individuals in EC

We defined our reference (REF) as ES dominant haplotype Y1 and A32.

Table S4. Identified genetic variability at 17 complete mtDNA sequences of A. cahirinus at EC

Blanks correspond to the same nucleotide as the reference. Ns correspond to Nonsynonymous mutation. Letters correspond to DNA base pairs (T, thymine; G, guanine; C, cytosine; and A, adenine). *E08, E14, A32, E02, E10, E21, and E03.

† .1=insertion.

PNAS PNAS

h, number of haplotypes; Hd, haplotype diversity; M, number of samples; S, number of variable sites; π, nucleotide diversity.

Species Location		Accession no.	
A. cahirinus	Cairo, Egypt	Z96051	
A. cahirinus	Cairo, Egypt	AJ233953	
A. dimidiatus	South Sinai, Egypt	AJ2339595	
A. dimidiatus	Mount Sinai, Egypt	Z96060	
A. dimidiatus	Dead Sea	AJ233953	
A. dimidiatus	Saudi Arabia	Z96061	
A. dimidiatus	Saudi Arabia	Z96062	
A. ignitus	Tanzania	JN247674	
A. airensis	Eqypt	AJ012021	
A. percivali	Tanzania	EF187818	
A. wilsoni	Tanzania	EF187801	
A. minos	Greece	AJ233961	
A. cilicicus	Turkey	AJ233957	
A. nesiotes	Cyprus	AJ233952	

Table S6. Downloaded Cytb sequences of nine different Acomys species from GenBank

Table S7. European female mate choice

PNAS PNAS

Estral European female no. 7 choosing between 11 alternative "African" and "European" males in six experiments.

Table S8. Capture-recapture data of Acomys in Evolution Canyon for the period from March 1996 to April 1999

The number of traps in a month on average was 178. Studied by Elena Ivanitskaya and David Uekin.

	mtDNA		
Primer name	region name	Position	Primer sequence
LA4F	12S rRNA	315–335	ACCGCGGTCATACGATTAAC
LA1F	12S rRNA	683-702	TTCAGCAAACCCTCAAAAGG
LA3R	12S rRNA	744–763	CACCTCATGGGCTACACCTT
LA33F	12S rRNA	860-885	GCTTAATTGAATAGAGCAATGAAGTG
LA4AF	12S rRNA	898-917	CCGTCACCCTCCTCAAACTA
LA38F	tRNA-Val	1039-1058	TGGCCTACACCCAGAAGACT
LA25R	tRNA-Val	1075-1094	GGGCTGTTTTTGGTTCAAAG
LA48	16S rRNA	1402-1421	GAAATGCCTAACGAGCTTGG
LA39	16S rRNA	1559-1578	GTCAAAACCCCAAGGACAGA
LA40	16S rRNA	1950-1969	ACTGCCTGCCCAGTGACTAA
LA9F	16rRNA	2320-2340	TGATCAACGGACCAAGTTACC
LA41	16rRNA	2344-2363	GGGATAACAGCGCAATCCTA
LA42	16rRNA	2562-2581	CCTTTCCATAAGCGCTTTCA
LA33R	16rRNA	2590-2613	CATTGTGAAATTGAGATTTTTCAT
LA49	NADH 1	2807-2826	TGGGACCTTTTCGTAGTTGC
LA14F	NADH 1	2991-3010	CACCCCCTAGCCAACATAAA
LA43	NADH 1	3367-3386	GCCCCTTTGCCTTATTCTTC
LA16F	tRNA-GIn	3739-3759	GGACAACAGGAATTGAACCTG
LA19F	NADH ₂	4200-4219	CCTAGCACCCTTCCACACAT
LA18F	NADH 2	4430-4449	TCGCCCACATAGGATGAATA
LA20F	NADH 2	4884-4905	AATTCTCCCCTTAACACCCCTA
LA21F	COX ₁	5416-5435	GGACAACCAGGGGCTCTATT
LA12F	COX I	5912-5932	TGCGGTCGGTTAATAATATGG
LA8F	COX I	5919-5938	TTAGGTTGCGGTCGGTTAAT
LA29F	COX I	6026-6045	TCCTGATTCTCCCTGGGTTT
LA34F	COX I	6351-6371	AACCGGAATTGTACTGTCAAA
LA30	COX I	6440-6459	TGGGGGCTGTATTTGCTATC
LA4AR	COX I	6627-6649	CCATGTTTCAAGTTGTGTAAGCA
LA35F	COX I	6802-6821	ACCTTCGAAGAACCCACCTA
LA31	COX II	7273-7292	ATGGGCCACCAATGATACTG
LA6R	COX II	7558-7578	GCCACAAATTTCTGAGCATTG
LA46	ATPase 6	8163-8182	CACGCCTACCACCCAACTAT
LA32	ATPase 6	8337-8355	TATCCAACCCATGGCCCTA
LA36F	ATPase 6	8915-8934	ACCACCTGCAGGAATCTCAC
LA44	COX III	9006-9025	GCCCACCATAGCCTCATAGA
LA36R	COX III	9133-9152	TGAGCCATAGATGCCATCAG
LA45	COX III	9254-9273	AAAACACCACTTTGGCTTCG
LA2F	NADH 3	9423-9442	AACATACTCCTGGCCGTCAC
LA12R	NADH3	9451-9747	TAATTAGGAGAAGTGACAGGGTGA
LA22F	NADH 4	9519-9538	GCAAACCCCTATGAATGTGG
LA26F	NADH 4	10148-10169	TCAACCATACTTCTCCCCCTAA
LA ₁ R	NADH 4	10354-10373	TATGGTTTTGGCTGGCTAGG
LA2AF	NADH 4	10531-10550	GGGCAATCAAACAGAACGAT
LA5R	NADH 4	10684-10706	GGATGTTGTTTGATCATGTTGTG
LA22R	NADH 4	10784-10803	CCTGCGATTGGAGCTTCTAC
LA7R	NADH 4	10784-10803	CCTGCGATTGGAGCTTCTAC
LA8R	NADH 4	11280-11299	TGATGGGTTGGATCAAGAGA
LA10F	tRNA-Ser	11567-11589	CCAAGAAAGAATACAAGGACTGC
LA23F	NADH 5	12282-12301	TGGGAAATGCAACAAATCCT
LA13F	NADH 5	12358-12377	GAAAATCCGCTCAATTTGGA
LA15F	NADH 5	12795-12814	TTCACCACCTCCTGCCTAAC
LA47	NADH 5	13217-13236	CTCAAACCTAACCCCCAACA
LA17F	NADH 5	13418-13438	GCACATCTCCAATCAAAAAGG
LA27F	NADH 6	13841-13867	ATACCTCCCAAATAAATTAGAAACACC
LA13R	tRNA-Glu	14031-14055	GCCATAGTTAAGTGCTATGTAGGAA
LA24F	cyt b	14114-14133	CGAAAAACACACCCACTCCT
LA2AR	cyt b	14114-14133	AGGAGTGGGTGTGTTTTTCG
LA3F	cyt b	14114-14133	CGAAAAACACACCCACTCCT
LA10R	cyt b	14389-14408	AGATGCCTCGTCCTACGTGT
LA37F	cyt b	14588-14608	TGAATTTGAGGTGGGTTTTCA
LA28F	cyt b	14697-14716	TCCTTCACGAAACTGGCTCT
LA3AF	cyt b	15126-15145	CCGTAGAACACCCCTTCATC

Table S9. List of primers used for complete mtDNA sequencing in A. cahirinus

Table S9. Cont.

