Genome evolution of blind subterranean mole rats: Adaptive peripatric versus sympatric speciation

Kexin Li[a,b,1,2], Shangzhe Zhang[a,1], Xiaoying Song[c,1], Alexandra Weyrich[d,1], Yinjia Wang[a], Xi Liu[a], Na Wan[a], Jianquan Liu[a], Matěj Lövy[e], Haihong Cui[f], Vladimir Frenkel[b], Avi Titievsky[g], Julia Panov[g,h], Leonid Brodsky[g], and Eviatar Nevo[b,2]

[a]State Key Laboratory of Grassland Agro-Ecosystem, Institute of Innovation Ecology, Lanzhou University, 730000 Lanzhou, China; [b]Institute of Evolution, University of Haifa, 3498838 Haifa, Israel; [c]School of Public Health, Lanzhou University, Lanzhou, 730000, China; [d]Leibniz Institute for Zoo and Wildlife Research, 10315 Berlin, Germany; [e]Faculty of Science, University of South Bohemia, 370 05 České Budejovice, Czech Republic; [f]Department of Gastroenterology, The 305 Hospital of PLA, 100017 Beijing, China; [g]Tauber Bioinformatics Research Center, Haifa 3498838, Israel; and [h]Sagol Department of Neurobiology, University of Haifa, Haifa 3498838, Israel

Contributed by Eviatar Nevo, October 8, 2020 (sent for review September 8, 2020; reviewed by Sergey Gavrilets and Eugene V. Koonin)

Speciation mechanisms remain controversial. Two speciation models occur in Israeli subterranean mole rats, genus *Spalax*: a regional speciation cline southward of four peripatric climatic chromosomal species and a local, geologic-edaphic, genic, and sympatric speciation. Here we highlight their genome evolution. The five species were separated into five genetic clusters by single nucleotide polymorphisms, copy number variations (CNVs), repeatome, and methylome in sympatry. The regional interspecific divergence correspond to Pleistocene climatic cycles. Climate warmings caused chromosomal speciation. Triple effective population size, N_e, declines match glacial cold cycles. Adaptive genes evolved under positive selection to underground stresses and to divergent climates, involving interspecies reproductive isolation. Genomic islands evolved mainly due to adaptive evolution involving ancient polymorphisms. Repeatome, including both CNV and LINE1 repetitive elements, separated the five species. Methylation in sympatry identified geologically chalk-basalt species that differentially affect thermoregulation, hypoxia, DNA repair, P53, and other pathways. Genome adaptive evolution highlights climatic and geologic-edaphic stress evolution and the two speciation models, peripatric and sympatric.

subterranean rodents | genomic sequencing | speciation models | methylation | repeatome

The remarkable ecological adaptation of mammals to the underground environment due to climatic change in Eocene times, approximately 50 million years ago (Mya), is one of nature's best-studied long-term evolutionary adaptive experiments. It involves mosaic evolution of regression, progression, and global convergent adaptations to their common, unique subterranean ecology (1). The genus *Spalax* (Spalacidae, Rodentia; Fig. 1*A*), originated in Asia Minor and displays an outstanding three-pronged adaptive climatic radiation into the Balkans, Ukraine, and Near East southward to North Africa, reflected in an increasing diploid set of chromosome numbers, from $2n = 36$ in Asia Minor to $2n = 62$ associated with high ecological stresses in all three prongs (1). The divergence of chromosomes resulted from Robertsonian chromosomal fission mutations (2, 3) and led to a southward ecological chromosomal speciation cline of the *Spalax ehrenbergi* superspecies in Israel. Ecogeographically, *Spalax* chromosome sets are increasing southward toward the Negev Desert. They are associated with extreme changes in ecological factors (1), especially the climate of Israel from the northern humid, cold Galilee and Golan region southward to the hot, dry Negev Desert, representing an ecology of increasing climatic aridity southward (*SI Appendix*, Table S1) (4).

Notably, *Spalax* has largely been known to speciate chromosomally (5), adaptively, allopatrically (i.e., separated distantly geographically without ongoing gene flow), or peripatrically (i.e., isolated relatively closely in peripheral populations surrounding the main range but without ongoing gene flow) (6), a kind of close allopatry. *Spalax galili* ($2n = 52$) is distributed in the humid, cool Upper Galilee and underwent further sympatric speciation (i.e., speciation in the same homeland [patria]), with ongoing gene flow between the abutting populations evolving into two different species due to geologic-edaphic (soil) divergence from *S. galili* ancestor species on Senonian (Upper Cretaceous) chalk rocks to *S. galili*, the derivative species evolving on Pleistocene volcanic basalt (7–9) (Fig. 1*B*). *S. golani* ($2n = 54$) inhabits the semihumid, cold northern Golan and Mount Hermon region (Fig. 1*C*). *Spalax carmeli* ($2n = 58$) ranges from the warm southern Golan to the humid and warm lower Galilee, Mount Carmel, and Coastal Plain and down to the Yarkon River (10) (Fig. 1*C*). *Spalax judaei* ($2n = 60$) occurs in the warm, dry southern Samaria and Judea Mountains, southern coastal plain, and northern Negev Desert (10) (Fig. 1*C*). Thus, the four chromosomal sibling species ($2n = 52, 54, 58, 60$) are distributed in four climatic regions based on a combination of temperature and humidity (10–12). In contrast, the sole sympatric species occurs locally in the same macroclimate but is divergent geologically

Significance

We substantiate genomically, repeatomically, and epigenomically the origin, demography, and timing of two divergent speciation models in the Israeli five blind subterranean species of *Spalax ehrenbergi* superspecies. Four species demonstrate a regional, chromosomal, peripatric, climatic aridity speciation model trending southward from the northern cold and humid Golan and Upper Galilee to the hot, dry Negev Desert. The fifth species shows a local, genic, geologic-edaphic, and sympatric speciation model demonstrating primary sympatric speciation in subterranean mammals. The five species are differentiated at multiple genomic levels and demonstrate different ecological mechanisms of speciation in the superspecies. Sympatric speciation may be common in nature. Numerous ecologically divergent microsites—geologic, edaphic, climatic, abiotic, biotic—abound globally, where selection overrules gene flow homogenization.

Author contributions: K.L. and E.N. designed research; K.L., X.S., N.W., M.L., and H.C. performed research; S.Z. contributed new reagents/analytic tools; K.L., A.W., Y.W., X.L., V.F., A.T., J.P., and L.B. analyzed data; and K.L., A.W., J.L., J.P., and E.N. wrote the paper.

Reviewers: S.G., University of Tennessee at Knoxville; and E.V.K., National Institutes of Health.

The authors declare no competing interest.

[1]K.L., S.Z., X.S., and A.W. contributed equally to this work.

[2]To whom correspondence may be addressed. Email: likexin@lzu.edu.cn or nevo@research.haifa.ac.il.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2018123117/-/DCSupplemental.

First published December 4, 2020.

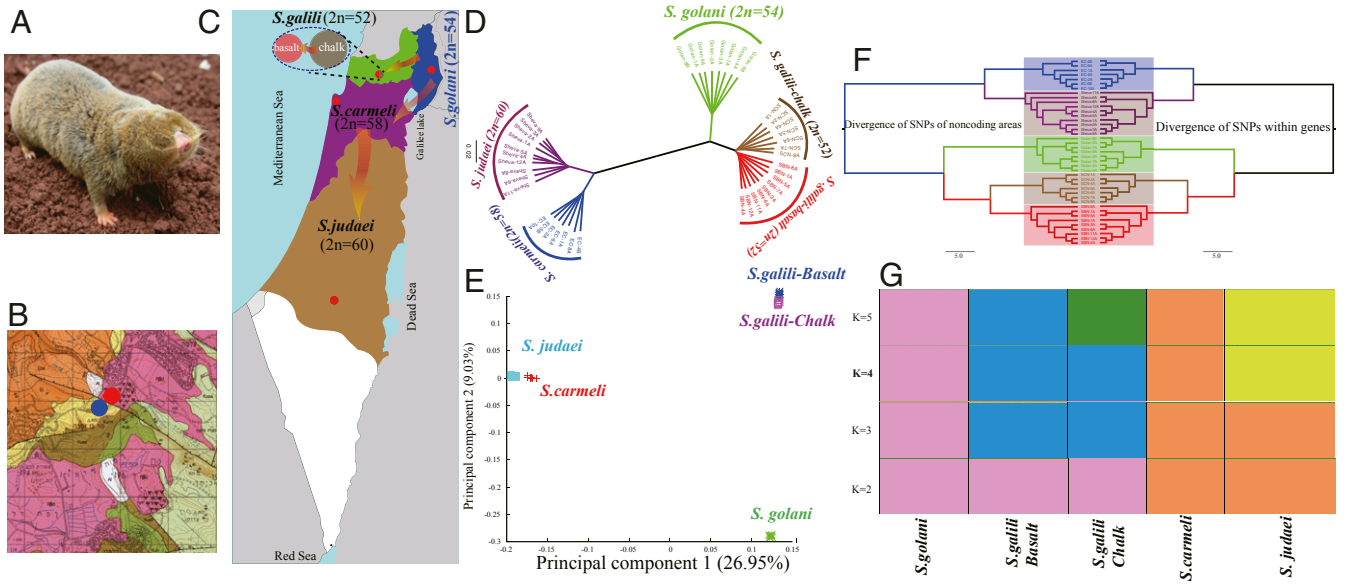

Fig. 1. The ecogeographic distribution and the genomic divergence among *Spalax* species in Israel. (*A*) Blind mole rat, *Spalax*. (*B*) Geological map of east Upper Galilee including the Evolution plateau. (*C*) Ecogeographic map of species distribution (from north to south marked in different colors) and sampling sites (red dots) of the four climatic and chromosomal peripatric species, and the fifth *S. galili* (*2n* = 52) marked in green, which diverged edaphically and genetically from sympatric species derivative *S. galili*_basalt) (*2n* = 52) marked with a red circle and its ancestor *S. galili*_chalk (*2n* = 52) marked with a blue circle, *S. golani* (*2n* = 54) marked in blue, *S. carmeli* (*2n* = 58) marked in violet, and *S. judaei* (*2n* = 60) marked in brown. (*D*) A neighbor-joining tree reconstructed with the allele shared matrix of SNPs of the five blind mole rat species populations. The scale bar represents the *p* distance. (*E*) Genetic clusters of the four species shown by PCA based on SNPs; only principal component 1 (26.95%) and principal component 2 (9.03%) are displayed. (*F*) Neighbor-joining tree based on the SNPs in coding (right side) and noncoding (left side) genomic regions. (*G*) Structure analysis of the five *Spalax* species. The number of putatively genetic populations (*K*) was defined from *K* = 2 to *K* = 5; each column (separated by white lines) denotes one individual.

and edaphically, living in rendzina soil on the chalk rock and in basalt soil on the volcanic basalt rock (Fig. 1 *B* and *C*).

Narrow hybrid zones separate the abutting species, increasing in breadth southward 320 m between *2n* = 52 and 58, 725 m between *2n* = 54 and 58, and a 2,825 m hybrid zone between *2n* = 58 and 60 (13), suggesting a southward speciation trend with increasing chromosomal numbers, correlated with increasing aridity (11) (Fig. 1*C*). This speciation trend provides evidence that the genus *Spalax* speciated chromosomally (5), adaptively (e.g., climatically), peripatrically (6), or sympatrically, which likely applies to the specific microsites in which a genetic/genomic divergence has been detected within a metapopulation with gene flow subdivided into two contrasting ecologies, chalk abutting with basalt (Fig. 1*B*).

Here we compare and contrast genomically two speciation models in *Spalax*. The regional, climatic, and peripatric speciation model (10) of the four chromosomal species, with the local, edaphic, sympatric speciation, within a population of *S. galili*, with limited gene flow between Pleistocene basalt abutting Senonian chalk at the "Evolution Plateau" in the eastern Upper Galilee (7–9, 14, 15) (Fig. 1*B*). Chromosomal rearrangements are widespread in animals and are thought to facilitate speciation through rapid reproductive isolation (5, 10–12). In Spalax, such rearrangements occurred due to postzygotic meiotic disturbances (16), followed by prezygotic reproductive isolation, olfaction (17, 18), vocal dialects (19), and seismic communication (20). However, the underlying adaptive genomic evolution in the correlated climatic changes, and differences between peripatric and sympatric speciation models, have remained largely unknown. We resequenced population genomes and repeatomes of the four chromosomal species, and the methylome of the *S. galili*_basalt and *S. galili*_chalk to address these evolutionary questions.

Results

Population Structure and Genetic Diversity. We conducted whole genome resequencing of five *Spalax* species (*SI Appendix*, Tables S2–S4) and removed closely related individuals according to relatedness (*SI Appendix*, Fig. S1). Notably, the individuals from each species were clustered together, but separated from other species (Fig. 1 *D*–*G*). Population divergence was also seen based on single nucleotide polymorphisms (SNPs) from noncoding genomic regions which are mirroring the coding regions (Fig. 1*F*). The three Northern (N) species (*S. golani*, *S. galili_chalk*, and *S. galili_basalt*) share more genetic variation than the two Southern (S) species (*S. carmeli, S. judaei*) (*SI Appendix*, Fig. S2). All of the individuals were separated into the S and N categories when the number of putatively genetic populations, *K*, was set to 2 in the structure analysis. A gradual speciation trend was observed for *S. galili* and *S. judaei* from *K* = 3 to *K* = 4 (Fig. 1*G*). No recombinants were detected by the structure analysis, indicating that gene flow between species was limited. *S. golani* showed the highest and *S. judaei* the lowest genetic diversity (*SI Appendix*, Table S4). The second highest genetic diversity was found in *S. carmeli*, followed by *S. galili*. These differential genetic diversities among species may have arisen through the duration of speciation and the exposure to ecological stresses (21).

Interspecies Divergence, Gene Flow, and Demographic History. The demography and divergence of the four species were further assessed by pairwise sequential Markovian coalescent (PSMC) analysis (22), which suggested that the common ancestor of the four species gave rise to two clades (*S. golani-S. galili* and *S. carmeli-S. judaei*) between 1.2 and 1.5 Mya (Fig. 2*A*), as was the case in all pregenomic analyses. In the first interglacial warming stage (0.8 to 0.3 Mya), the Northern ancestral clade diverged into two current species, *S. golani* and *S. galili*, in the North (Fig. 1 *D* and *G*), whereas the second clade diverged into the two Southern species, *S. carmeli* and *S. judaei*, in the second interglacial warming stage (0.2 to 0.1 Mya) (Figs. 2*A* and 3*A*). These findings suggest that climatic cycles triggered *Spalax* chromosomal speciation. We further detected three episodes of decline of the

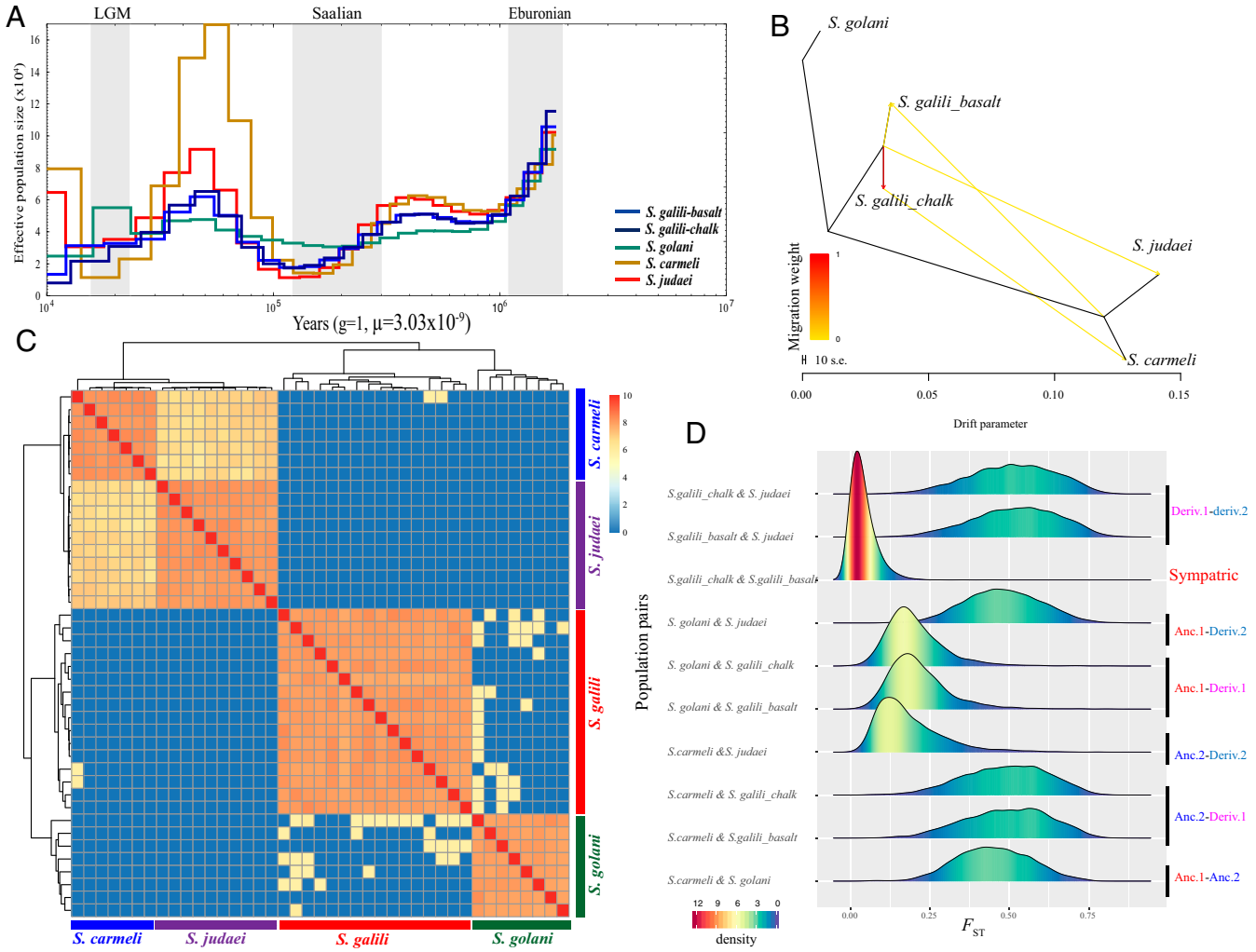

Fig. 2. Population demography of five *Spalax* species based on whole genome resequencing data. (*A*) Pairwise sequential Markovian coalescent results display the historical demography from 10 kya to ~ 2 Mya. The five color lines represent the estimated effective population size, with three population declines mirroring three glacial periods: Eburonian, Saalian, and LGM. (*B*) The maximum likelihood inferred tree of blind mole rat with mixture events. Arrows denotes migration which is colored according to its migration weight. The scale bar denotes the average SE of the entries in the sample covariance matrix. (*C*) Estimated shared haplotype between individuals. Heat map colors represent the total length of IBD blocks for each species after pairwise comparison. (*D*) Population pair genetic divergence (F_{ST}) between sympatric and peripatric speciation. Species pairs are listed on the left, and the relationship of each species pair is on the right. We define derivative 1 as the speciation of *S. galili* from *S. golani* and derivative 2 as the speciation of *S. judaei* from *S. carmeli*. The height of each curve is the density of F_{ST}. The *x*-axis represents F_{ST}.

effective population size, N_e, for each of the four chromosomal species or their ancestors. The first population decline occurred immediately after the Eburonian glacial period in Europe, approximately 1.3 Mya (Fig. 2*A*). The second population decline occurred around 300 kya, during the Saale glaciation period (0.3~0.20 Mya). The third population decline occurred during the last glacial maximum (LGM), approximately 20 kya. Recent distinct population expansions were found for the two Southern species approximately 50 kya (Fig. 2*A*), corresponding to climatic pluvial times, with enhanced rainfall and increased temperatures during this period from 55 to 52 kya, with a peak at 54 kya (23, 24). The vegetation increase in the warm and humid pluvial (interglacial) period (25) presumably provided sufficient food supplies for the expanding *Spalax* populations.

The genetic divergence, F_{ST}, between each species pair (*SI Appendix*, Table S5 and Fig. 2*D*) corresponds to phylogenetic and structure analyses, with high divergences between the three N species and the S species pair (Fig. 1 *D* and *E*). As expected, the smallest genomic distance was between the two sympatric species, the ancestor *S. galili*_chalk and the derivative *S. galili*_basalt (F_{ST} = 0.053), and the largest interspecies genomic distance was between *S. galili* and *S. judaei* (F_{ST} = 0.645) (Fig. 2*D* and *SI Appendix*, Table S5), which is negatively correlated with the width of hybrid zones between species. The hybrid zones separating the chromosomal species are under strong natural selection and decrease northward, from ~2,825 m between *S. judaei* and *S. carmeli* to ~320 m between *S. galili* and *S. carmeli* (11, 13, 26, 27), with strong selection against hybrids (11, 13). Similarly, genetic relationships among the four species were inferred by *D* statistics (28) (Fig. 3*B*), NetView P (29) (Fig. 3*D*), and TreeMix (30) (Fig. 2*B*). The BABA-ABBA and identity-by-descent (IBD) (Fig. 2*C*) tests demonstrated an extremely limited gene flow and only a few shared haplotypes. The three-population test (f_3) can provide a clear evidence of admixture, even if the gene flow occurred hundreds of years ago (31). We calculated the corresponding f_3 statistics (*SI Appendix*, Table S6) in pairwise comparisons for the 10 possible species combinations. A negative f_3 value indicates a complex history of the tested

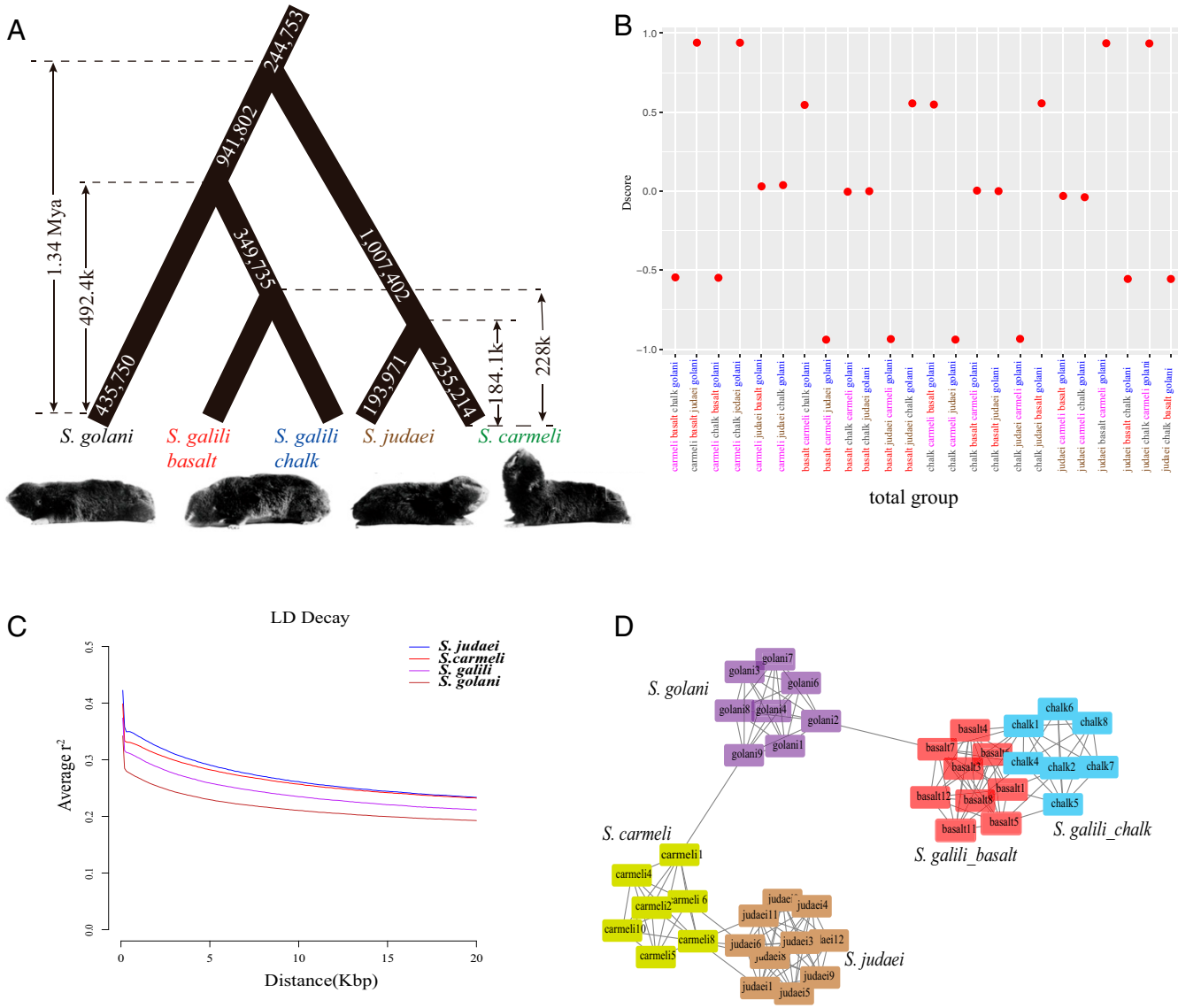

Fig. 3. Genetic relationships between and within species. (*A*) The schematic population demographic scenario that best fit our empirical data estimated by fastsimcoal2, the external numbers indicate spcies splitting time, and the internal numbers in white in the tree branches indicate effective population size. (*B*) D-statistics for different quadruples of blind mole rat species (P1–P3 and outgroup *S. golani*). Positive D values indicate that P1 shares more derived alleles with P_3 compared with P_2. (*C*) LD of the four chromosomal *Spalax* species. (*D*) Genetic network of the five *Spalax* species at $K = 10$ with minimum spanning tree based on 41,925,480 SNPs. Individuals are marked as rectangles while the edges that connect individuals denote the genetic relationships among individuals.

target population. All f_3 values and Z-scores were positive in all of the tested combinations, suggesting that no admixture occurred in the history of these species (*SI Appendix*, Table S6). Gene flow did not occur, probably because of the strong climatic adaptation of the four chromosomal species to four climatic regions (10). Both the f_3 test and the large F_{ST} suggest the current contact of *S. galili* and *S. carmeli* is secondary (*SI Appendix*, Table S6 and Fig. 2*D*).

Linkage disequilibrium (LD) of the four species drops rapidly, to below 0.3 within 5 kbp (Fig. 3*C*). *S. golani* shows the lowest LD, and *S. judaei* shows the highest LD, which is consistent with the strong selection exerted by the xeric environment on *S. judaei*.

To evaluate alternative divergence models (nine probable models in *SI Appendix*, Fig. S3) between the four chromosomal species, we used pairwise joint site frequency spectra to perform a composite likelihood comparison with fastsimcoal2 modeling software. The best demographic model was selected by the lowest delta likelihood and Akaike information criterion (32). The best-supported model (Fig. 3*A* and *SI Appendix*, Table S7) indicated that the common ancestor of the N bifurcatingly separated from the common ancestor of the S species approximately 1.34 Mya (95% highest posterior density [HPD] = 1.33 to 1.35 Mya) in the warm South Golan, south of the Afiq hybrid zone (13). Importantly, a fossil *Spalax*, presumably *S. carmeli*, was found in Ubadiyya, south of the lake of Galilee, from 1.4 Mya (33). The other bifurcation branched westward, producing *S. galili* from *S. golani* approximately 492.4 kya (95% HPD = 492.1 to 492.7 kya) in upper Galilee. After splitting from *S. golani*, *S. carmeli* speciated to *S. judaei* southward more recently, 184.1 kya (95% HPD = 184.2 to 183.9 kya). These datings are fully compatible with those obtained from PSMC results (Fig. 2*A*). No relatedness between *S. galili* and *S. carmeli* was detected by the network analysis (Fig. 3*D*), but both

of these species were found to be related to *S. golani,* suggesting an ancestral state.

Selective Sweep in the Five *Spalax* Species. To explore the adaptive evolution of the four chromosomal species, we conducted d_i tests to look for genes with high interspecific divergence and under positive selection driven by ecological stresses. Among 24,636 genes analyzed, a total of 1,256 genes were identified as evolving under selection (i.e., putatively selected genes [PSGs]) (*SI Appendix*, Fig. S4). This 5.09% of PSGs is plausible in view of the severe underground stresses, including darkness, hypoxia, hypercapnia, energetics, and pathogenicity (1, 10). We found 507 and 552 PSGs that were shared between the S and N species pairs, respectively. Furthermore, 365 genes were selected and shared in the four chromosomal species. Presumably, in each species, the same PSGs are involved in the adaptation to the same stresses that are characteristic of the underground lifestyle, such as hypoxia, hypercapnia, and darkness (*SI Appendix*, Fig. S4 and Dataset S1); however, there are also species-specific PSGs in each of the four chromosomal species that are likely involved in their unique adaptation to the divergent climates (*SI Appendix*, Fig. S4). Furthermore, the blind mole rat is a cancer-resistant animal (34, 35), conceivably linked with hypoxia resistance, that potentially could transform cancer resistance in medicine. *BCL7B*, a member of the *BCL7* gene family, is a tumor suppressor in humans (36) and is one PSG detected in all four *Spalax* species. Thus, the positive selection of *BCL7B* is consistent with the cancer resistance of this subterranean mammal, mediated by a concerted necrotic cell death mechanism (34). The tumor suppressor candidate 2 (*Tusc2*) was identified as a PSG in *S. galili*. It has been reported that ectopic expression of the *TUSC2* 3′-UTR inhibits cell proliferation, survival, migration, invasion, and colony formation and furthermore causes tumor cell death in humans (37). This gene was selected only in *S. galili*, probably because this species lives in a more hypoxic region, due to the much stronger winter rains in the northern Israel compared with southern Israel. Another gene, *F8*, coagulation factor VIII, which belongs to a group of proteins that are essential for the formation of blood clots, was positively selected in *S. galili* and *S. golani*, possibly because the aggression between individuals in populations of these species is stronger than that in the populations of *S. carmeli* and *S. judaei* (38). The *F8* gene product might be involved in keeping the animals from bleeding and promoting accelerated wound healing after fighting (39, 40). *Hspa14*, a heat shock protein family A (*Hsp70*) member, was also positively selected in *S. judaei* species, which extends to the hot, dry northern Negev Desert. *Hspa14* is down-regulated during heat stress, which could lower the rate of translation by slowing the release of properly folded proteins from the ribosome. Thus, it would contribute to the reduction of the protein synthesis burden during heat stress (41) in the hot and dry northern Negev Desert, where food resources for the blind mole rats are limited.

Reproductive isolation is necessary for speciation. In the current study, a number of genes related to male fertility and reproduction, including *Sox8*, *Spag5*, *Spata2l*, *Tex264*, *Tex28*, and *Tex38*, were found to be positively selected in the four chromosomal species. *Sox8* is a critical regulator of adult Sertoli cell function and male fertility (42). *SPATA2* is highly expressed in Sertoli cells of the adult mouse testis, and deletion of this gene attenuates fertility in male mice (43). These genes are potentially important for the blind mole rat adaptation and speciation underground (Dataset S1).

Genome Islands of Divergence between Species Pairs. The divergence of each population pair along the genome is highly heterogeneous (Fig. 4 *A–J*), and most of them are small, with size of 10 kb (*SI Appendix*, Fig. S5). The number of shared islands among the 10 population pairs ranged from 24 to 260 (Fig. 4*L* and *SI Appendix*, Table S9). Significantly elevated d_{xy} (Fig. 4 *K* and *M*) and LD (Fig. 4*O*) values were detected within F_{ST}_islands of all of the 10 population pairs (Fig. 4*M* and *SI Appendix*, Table S8), which is consistent with a model in which the island regions were derived from divergent sorting of adaptive evolution and ancient polymorphisms (44, 45). The genetic diversity (π) and population-scaled recombination rates were significantly lower in island regions (Fig. 4 *N* and *P* and *SI Appendix*, Table S8) compared with the backgrounds in all species pairs. Most of the Tajima's D values are strongly negative within genomic islands, indicating an excess of low-frequency variants (*SI Appendix*, Table S8). The shared islands between different populations pairs (Fig. 4*L* and *SI Appendix*, Table S9) suggest that they were apparently formed before the species split (44). Although the *S. carmeli* vs. *S. judaei* population pair displays the largest number of divergence islands, it shares the smallest number with all of the other pairs, suggesting the uniqueness of recent adaptive evolution to the drought and heat in southern Israel. The N and S clades diverged earliest, and the shared islands between them are probably from ancestry. This is also true between *S. carmeli* vs. *S. judaei* and between *S. galili*_basalt vs. *S. galili*_chalk.

Gene flow was mainly restricted between the peripatric chromosomal population pairs, and only slightly between the sympatric pair (Fig. 2*B*). However, there is no obvious difference in number of islands between them (*SI Appendix*, Table S10), suggesting that recent gene flow was not the major factor shaping the genomic islands (44). Gene Ontology (GO) enrichment of genes from islands (Fig. 5 *A–E*) shows that the comparison of *S. galili_chalk* and *S. galili_basalt* is related to angiogenesis, cancer, and autophagy (Fig. 5*D* and Dataset S2), and that between *S. carmeli* and *S. judaei* is related to water homeostasis, autophagy, and neurogenetics (Fig. 5*E* and Dataset S3).

Copy Number Variations and Repeatome.

Differences in copy number variations (CNV). CNV regions (CNVRs) varied among the species and showed four clusters by principal component analysis (PCA) (Fig. 6*A* and *SI Appendix*, Fig. S6), phylogenetic tree analysis (Fig. 6*B*), and heatmap analysis (Fig. 6*C*). *S. carmeli* harbored the largest number of CNVRs (2,873), followed by *S. golani* (2,209), *S. galili*_basalt (1,911), *S. galili*_chalk (1,761), and *S. judaei* (1,525). This was the same order of the total length of CNVRs, number of average CNVRs, species-unique CNVRs, and loss of CNVRs for these species (*SI Appendix*, Figs. S7, S8, and S10*B* and Tables S11–S14, S16). Most of the CNVRs were distributed in intergenic regions (*SI Appendix*, Table S15 and Fig. S9). We observed that the larger the effective population size, the greater the CNV (46).

The V_{ST} was used to estimate population differentiation, which is similar to $_{FST}$, ranges from 0 to 1. $V_{ST}$of all of the species pairs (*SI Appendix*, Figs. S10*A* and S11) showed the same trend as for F_{ST}, suggesting the same selection trend on different mutations. In some species population pairs, one-half of the V_{ST} values were extremely high, >0.5 (Dataset S4). Most of the genes with high V_{ST} values in pairwise species comparisons are known to occur in gene clusters and are related to digestion and metabolism, reproductive isolation, and local inflammatory reactions (47).

Significant differences were detected in the number of CNV calls between different species. The lowest CNV was found in *S. judaei* (*SI Appendix*, Fig. S10*B* and Table S14), which might optimize energetic balance given that this is the species with the lowest metabolic rate (48). However, a complementary explanation could be the reduced power to detect smaller CNVs and precise breakpoints in samples with a lower read depth (49). We found 332 CNVRs that overlapped with large segmental

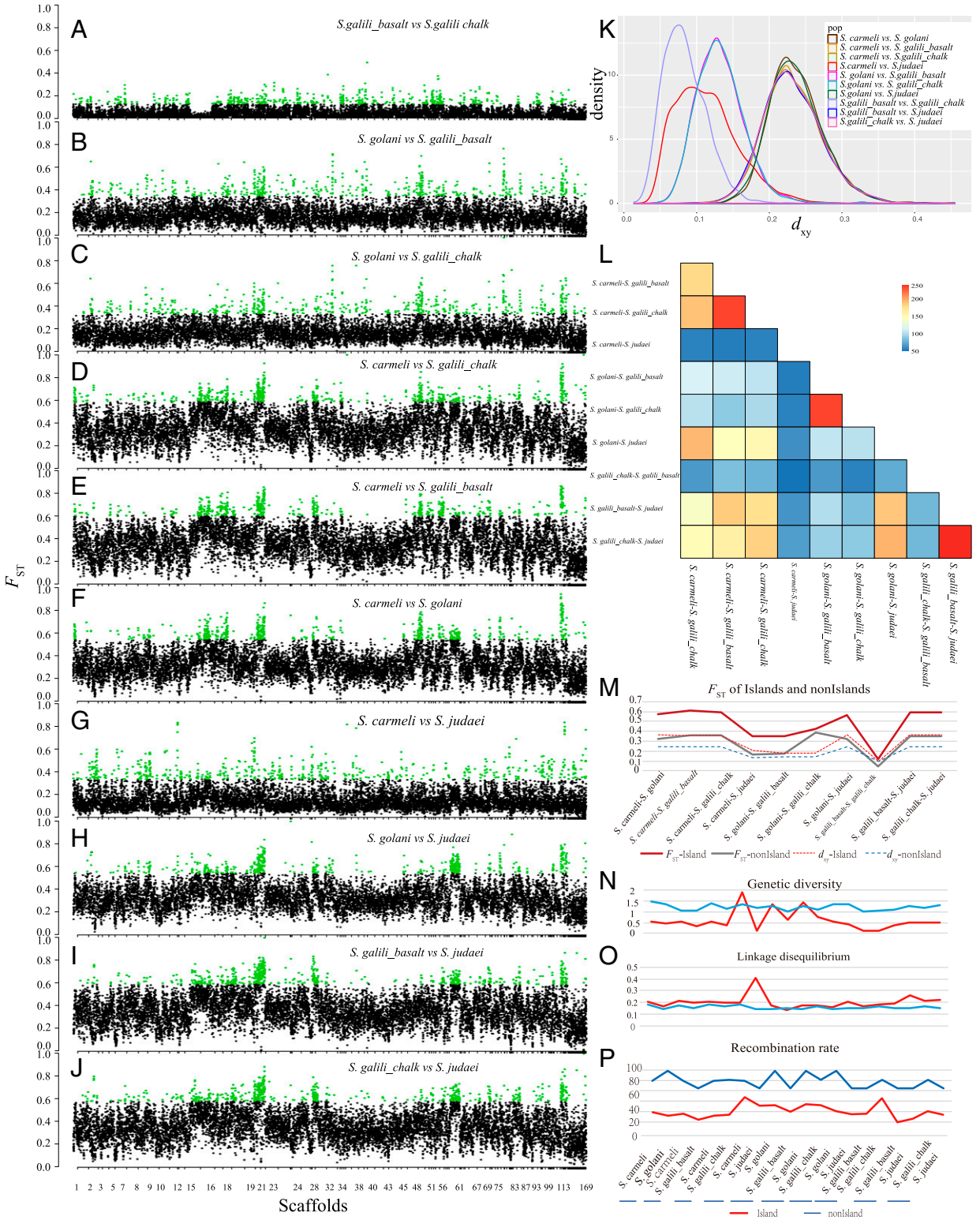

Fig. 4. Genome divergence between the sympatric and peripatric species. (*A*) F_{ST} of the sympatric species pair. (*B–J*) F_{ST} of the peripatric species pairs. (*K*) Absolute divergence (d_{xy}) of all linage pairs. (*L*) Number of shared islands between all of the linages. (*M*) F_{ST} and d_{xy}, (*N*) Genetic diversity, (*O*) LD, and (*P*) recombination rate within islands and its backgrounds.

duplications and 6,182 CNVRs that did not overlap with large segmental duplications (*SI Appendix*, Fig. S10*C*). The length distribution of CNV that overlapped with genes is different between the two kinds of CNVs (*SI Appendix*, Fig. S10*D*).

KEGG pathway enrichment analysis of CNV genes showed that CNV genes in xeric *S. judaei* were enriched in digestion, neurogenetics, and immunobiology (*SI Appendix*, Table S17). These are related to neurology (50), immunology (51), metabolism

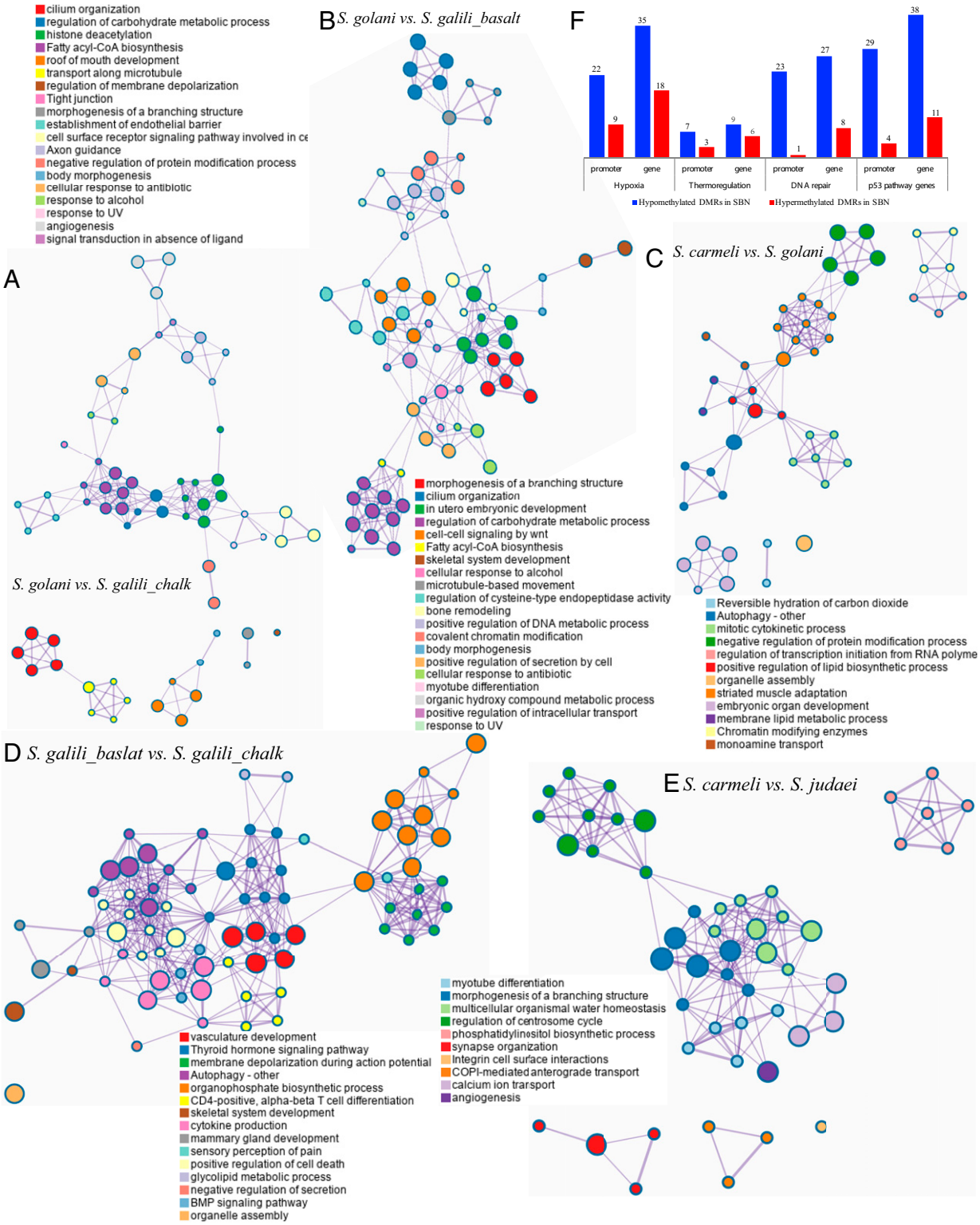

Fig. 5. GO of genes within islands and differential methylated regions. (*A*–*E*) GO of genes from islands of species pair (*A*) *S. golani* and *S. galili_chalk*, (*B*) *S. golani* and *S. galili_basalt*, (*C*) *S. carmeli* and *S. golani*, (*D*) *S. galili_basalt* and *S. galili_chalk*, and (*E*) *S. carmeli* and *S. judaei*. (*F*) Differential methylated genes between *S. galili_basalt* and *S. galili_chalk* within relevant pathways.

(52), and pathology (47), optimizing a network of adaptations to xeric hot stressful ecologies (4), where *S. judaei* is distributed.

Our analysis of genomic repetitive elements in *Spalax* species (see *SI Appendix, Materials and Methods*) revealed that the abundance of repetitive elements in genomes of the studied animals separates them into four clusters corresponding to the four chromosomal peripatric species (*SI Appendix*, Fig. S15 *A* and *B*). In addition, phylogenetic analysis based on the abundance of repetitive elements

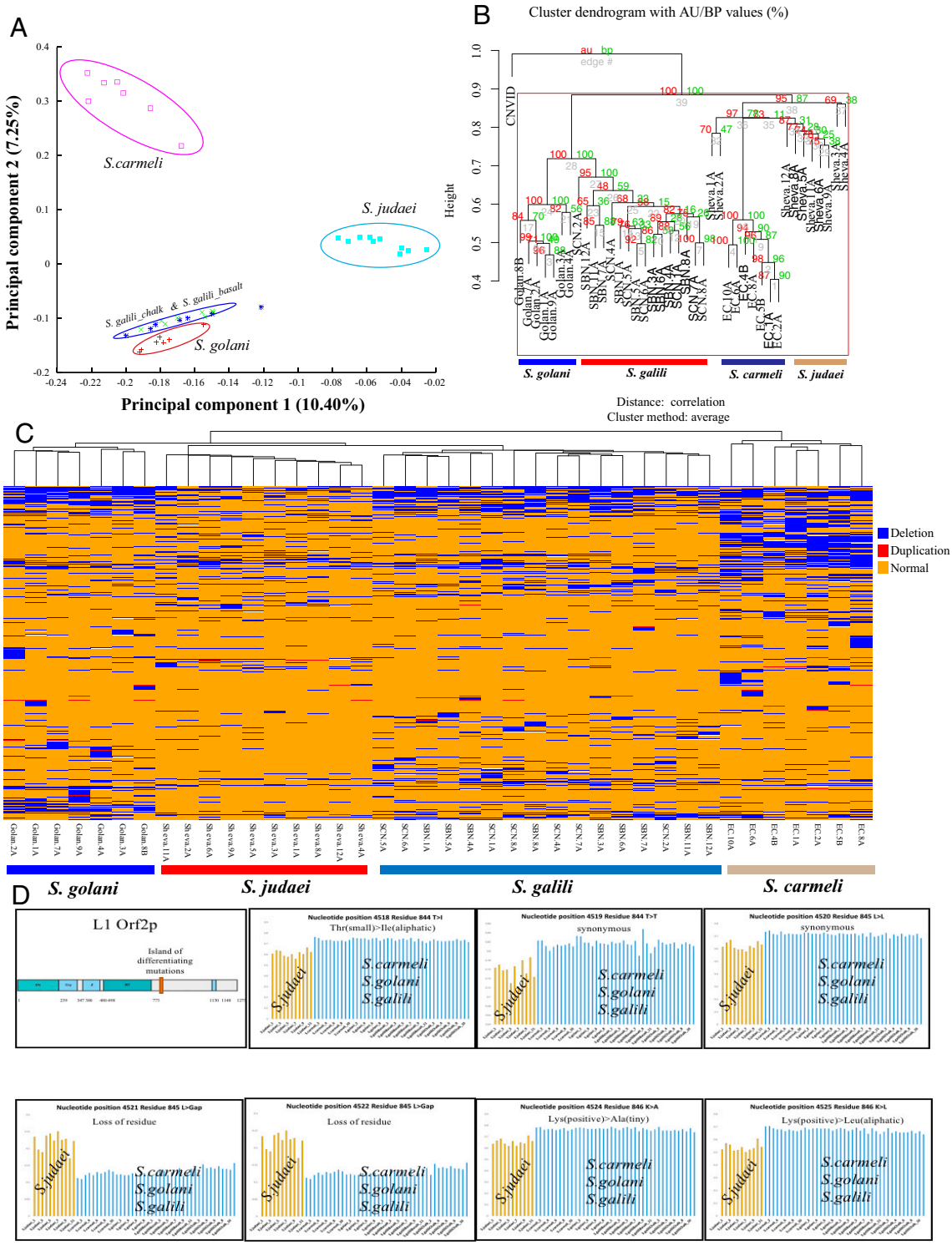

Fig. 6. CNV and retrotransposon LINE1 of the *S. ehrenbergi* superspecies. (*A*) PCA of CNV of the five *Spalax* species. (*B*) Cluster dendrogram of the four blind mole rat chromosomal species. (*C*) Heat map of CNV of the five species, highlighting deletion divergences. Orange, red, and blue denote autosomes without CNV, duplication, and deletion, respectively. (*D*) Mutations in retrotransposon LINE1. Orange bar in upper left panel denote the positions of mutated region on the L1Orf2p genomic sequence. The other panels are seven positions of L1 genomic sequence differentiating *S. judaei* species from other *Spalax* species. *Y* axis are frequencies of mutations in each sample. Note the difference in mutation frequencies between *S. judaei* compared with the other three chromosomal species.

taken as a continuous measurement yielded a similar separation of animals into four chromosomal *Spalax* species (*SI Appendix*, Fig. S15 *C* and *D*).

We studied differential mutations in repetitive elements across all pairs of studied *Spalax* species (*SI Appendix*, *Materials and Methods*) and found genomic positions in the LINE1 loci that have species-distinctive substitutions in several *Spalax* species (*SI Appendix*, Table S22). Polymorphisms of the repetitive elements, as well as their abundances in the genomes, are known to play

important roles in the adaptation of organisms to their environment (53). The most prominent species-distinctive substitutions are located in an island of mutations of the LINE-1 loci that is differentiating between the *S. judaei* species and all other *Spalax* species (Fig. 6*D*, *SI Appendix*, Table S22), possibly highlighting a genomic response to high coccid pathology (47) and other pathogens prevalent in xeric environments. The island of mutations that differentiates the *S. judaei* species from all other *Spalax* species matches human LINE1 genomic locus known to be important for transposition (54).

Epigenetic differences precede and accompany in sympatric speciation. Epigenetic differences in DNA methylation patterns expand the toolbox of adaptation. As expected for mammalian genomes, after mapping (*SI Appendix*, Table S18), methylated Cs mainly occurred in a CpG context (69.3 to 77.5%), whereas of all methylated Cs, only 0.3% were in a CHG context and 0.3 to 0.6% were in a CHH context) (C, cytosine; mC, methylated cytosine; in CHH and CHG, H stands for A, T, or C) (55, 56). Significantly differentially methylated CpGs (Fisher's exact test) were clustered, and differentially methylated regions (DMRs) were calculated by comparing the two sympatric species, *S. galili*_chalk and *S. galili*_basalt. Because promoter methylation has a main regulatory function, we focused on DMRs within promoters with strong methylation differences above 30% between chalk and basalt animals. We detected 129 DMRs, of which 43 were hypomethylated and 86 were hypermethylated in the *S. galili*_basalt (*SI Appendix*, Table S19). Of the 129 promoter DMRs, 114 were overlapping with genes, and 8 were still uncharacterized. The set of genes was incorporated into STRING for gene pathway analysis, enriched GO terms were identified, and relevant gene networks were analyzed (*SI Appendix*, Figs. S13 and S14 and Table S20). Compared with *S. galili*_basalt, *S. galili*_chalk is less methylated in liver in gene promoters, including genes important for acetylation, indicating greater gene activation, as both promoter hypomethylation and acetylation are reflective of gene activation status (*SI Appendix*, Table S20 and Figs. S13 and S14). In addition, DMRs were detected in genes involved in pathways relevant in hypoxia, hypercapnia, thermoregulation, DNA repair, cancer resistance, P53 pathways (Fig. 5*F*), heat shock proteins, and olfactory and taste receptors, and these genes also showed different expression levels in our previous study (9). This indicates a systemic function of these genes in response to differences in climate and edaphic factors due to their regulation in both liver and brain (9). In speciation processes, DNA methylation may act as a rapid adaptive mechanism of genomic regulation for initializing, compensating, supporting, or causing species divergence.

Discussion

Genome analysis has highlighted adaptive evolution and speciation of the four regional peripatric chromosomal species (1–3, 7–10) and the fifth local sympatric genic species of *S. galili* on basalt (7–9). All are good biological species adapted to climate or soil, respectively, despite narrow interspecies hybridization, where hybrids were strongly selected against (11). Regional Pleistocene chromosomal speciation events were associated with repeated climatic warmings (interglacial cycles, or pluvial high rainfall periods in the Near East), followed by three population size declines occurring during cold and dry glacial cycles (Fig. 2*A*). The divergent time estimated by whole genome sequencing (Figs. 2*A* and 3*A*) is similar to that of DNA-DNA hybridization, highlighting the initiation evolution of the *S. ehrenbergi* superspecies in Israel approximately 1.6 ± 0.3 Mya (57), apparently within the range of the genomic estimate, which is 1.34 Mya. No gene flow was found due to strong divergent climatic adaptation (10) irrespective of the interspecific hybrid zones (13) (Fig. 2*B*). General and species-unique cancer resistant genes complemented earlier-identified mechanisms mediating concerted necrotic cell death (34), associated with underground hypoxia-resistant genes. Remarkably, the V_{ST} of all species pairs (*SI Appendix*, Figs. S10A and S11) showed the same trend as F_{ST} on differential mutations (Dataset S2), i.e., adaptive evolution.

Most genes with high V_{ST} values occur in adaptive, speciation, and regulatory gene clusters related to metabolism, reproductive isolation (18–20), and inflammation (47). Generally, V_{ST} is much larger in *Spalax* than in mice (46), probably because of adaption to the high stresses of life underground. Repeatome selection of CNV and functional enrichment in xeric *S. judaei* occurred in genes rich in metabolism (52), neurogenesis (50), immunobiology (51), inflammation, and pathology, adaptive to xeric ecology.

The small number of islands shared among different population pairs is unlikely related solely to intrinsic genome characteristics, such as recombination rate, which is conserved between independent comparisons (44). This suggests the possible existence of divergence hitchhiking or background selection and/or adaptive recurrent selective sweeps. However, selection could lead to elevated F_{ST} but unchanged or decreased d_{xy}. Divergence hitchhiking may allow for the significantly lower recombination rate and higher d_{xy} of genomic islands, as sorting of adaptive evolution and ancient polymorphisms would reduce gene exchange in the surrounding divergent selected regions (58).

We began this paper by stating that the concepts of species and speciation modes are still contentious. Clearly, there are different kinds of species and different mechanisms of speciation. The species and speciation concepts must be broadened to accommodate many species types and mechanisms in nature. Nature's imaginative diversity creativity is not restricted only to within-species protein and DNA/RNA polymorphisms, but abounds in interspecies diversity in origin, structure, and evolution. The basic as-yet unresolved issue in the modes of speciation is between allopatric and sympatric speciation (59). Clearly, as *Spalax* exemplifies, species diversity largely matches ecological diversity in nature, both regionally and locally, climatically and edaphically, respectively. Moreover, we have demonstrated in our Evolution Canyon (60) and Evolution Plateau (7–9, 14, 15) models, a microclimatic interslope and geological-edaphic model, respectively, hot spots of sympatric speciation across life from bacteria to mammals. Since geological, edaphic, climatic, abiotic, and biotic contrasts in microsites abound globally, sympatric speciation might be a common speciation model in which selection overrules gene flow homogenization (61).

Materials and Methods

Pair-end sequencing was performed with an Illumina NovaSeq sequencing system. Clean reads were mapped against the reference genome using BWA, and SNPs were called by GATK. Phylogenetic tree, PCA, and structure were carried out using TreeBeST, GCTA, and frappe, respectively. F_{ST} and Tajima's D were calculated by VCFtools. Fluctuations in effective population size were calculated by PSMC, gene flow was assessed by Treemix, f_3 and D-statistics were calculated using AdmixTools. LD was calculated by PopLDdecay. IBD was estimated by Beagle. Gene enrichment was conducted by Metascape. CNV was called by CNVnator. Recombination rate was calculated by fastEPRR. Species divergence pattern and time were estimated by fastsimcoal2. Detailed information is available in *SI Appendix, Materials and Methods*.

Data Availability. Sequence Read Archive data have been deposited in China National Center for Bioinformation-National Genomics Data Center (accession numbers CRA003292 and CRA003332). All study data are included in the main text and *SI Appendix*.

ACKNOWLEDGMENTS. We thank Xu Zhang for the Perl scripts, Dorina Meneghini for bioinformatics support, and Jörns Fickel for scientific discussions. This project was supported by the National Natural Science Foundation of China (32071487), Lanzhou University's "Double First-Class" Guided Project Team Building-Funding-Research Startup Fee (K.L.), the Chang Jiang Scholars Program, the International Collaboration 111 Programme (BP0719040), and the Ancell–Teicher Research Foundation for Genetic and Molecular Evolution.

1. E. Nevo, *Mosaic Evolution of Subterranean Mammals: Regression, Progression, and Global Convergence* (Oxford University Press, 1999).
2. J. Wahrman, R. Goitein, E. Nevo, Mole rat *Spalax*: Evolutionary significance of chromosome variation. *Science* **164**, 82–84 (1969).
3. J. Wahrman, R. Goitein, E. Nevo, "Geographic variation of chromosome forms in *Spalax*, a subterranean mammal of restricted mobility" in *Comparative Mammalian Cytogenetics, K. Benirschke, Ed.* (Springer, New York, 1969), pp. 30–48.
4. A. Bitan, S. Rubin, *Climatic Atlas of Israel for Physical and Environmental Planning and Design* (Israel Ministry of Transport, 1991).
5. E. Nevo, *Speciation: Chromosomal Mechanisms* (In: eLS, John Wiley & Sons Ltd, Chichester, 2012).
6. E. Nevo, "Modes of speciation: The nature and role of peripheral isolates in the origin of species" in *Genetics, Speciation and the Founder Principle, L. V. Giddings, K. Y. Kaneshiro, W. W. Anderson, Eds.* (Oxford University Press, Oxford, 1989), pp. 205–236.
7. Y. Hadid *et al*., Possible incipient sympatric ecological speciation in blind mole rats (*Spalax*). *Proc. Natl. Acad. Sci. U.S.A.* **110**, 2587–2592 (2013).
8. K. Li *et al*., Sympatric speciation revealed by genome-wide divergence in the blind mole rat *Spalax*. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 11905–11910 (2015).
9. K. Li *et al*., Transcriptome, genetic editing, and microRNA divergence substantiate sympatric speciation of blind mole rat, *Spalax*. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 7584–7589 (2016).
10. E. Nevo, E. Ivanitskaya, A. Beiles, *Adaptive Radiation of Blind Subterranean Mole Rats: naming and Revisiting the Four Sibling Species of the Spalax Ehrenbergi Superspecies in Israel: Spalax galili (2n= 52), S. Golani (2n= 54), S. Carmeli (2n= 58), and S. judaei (2n= 60)* (Backhuys Publishers, 2001).
11. E. Nevo, Speciation in action and adaptation in subterranean mole rats: Patterns and theory. *Ital. J. Zool. (Modena)* **52**, 65–95 (1985).
12. E. Nevo, Evolutionary theory and processes of active speciation and adaptive radiation in subterranean mole rats, *Spalax ehrenbergi* superspecies, in Israel. *Evol. Biol.* **25**, 1–125 (1991).
13. E. Nevo, H. Bar-El, Hybridization and speciation in fossorial mole rats. *Evolution* **30**, 831–840 (1976).
14. M. Lövy *et al*., Habitat and burrow system characteristics of the blind mole rat *Spalax galili* in an area of supposed sympatric speciation. *PLoS One* **10**, e0133157 (2015).
15. M. Lövy, J. Šklíba, R. Šumbera, E. Nevo, Soil preference in blind mole rats in an area of supposed sympatric speciation: Do they choose the fertile or the familiar? *J. Zool. (Lond.)* **303**, 291–300 (2017).
16. J. Wahrman, C. Richler, R. Gamperl, E. Nevo, Revisiting *Spalax*: Mitotic and meiotic chromosome variability. *Isr. J. Zool.* **33**, 15–38 (1984).
17. E. Nevo, M. Bodmer, G. Heth, Olfactory discrimination as an isolating mechanism in speciating mole rats. *Experientia* **32**, 1511–1512 (1976).
18. G. Heth, E. Nevo, Origin and evolution of ethological isolation in subterranean mole rats. *Evolution* **35**, 259–274 (1981).
19. G. Heth, E. Frankenberg, E. Nevo, "Courtship" call of subterranean mole rats (*Spalax ehrenbergi*): Physical analysis. *J. Mammal.* **69**, 121–125 (1988).
20. E. Nevo, G. Heth, H. Pratt, Seismic communication in a blind subterranean mammal: a major somatosensory mechanism in adaptive evolution underground. *Proc. Natl. Acad. Sci. U.S.A.* **88**, 1256–1260 (1991).
21. E. Nevo, M. G. Filippucci, A. Beiles, Genetic polymorphisms in subterranean mammals (*Spalax ehrenbergi* superspecies) in the near east revisited: Patterns and theory. *Heredity* **72**, 465–487 (1994).
22. H. Li, R. Durbin, Inference of human population history from individual whole-genome sequences. *Nature* **475**, 493–496 (2011).
23. M. Bar-Matthews, A. Ayalon, A. Kaufman, Timing and hydrological conditions of Sapropel events in the Eastern Mediterranean, as evident from speleothems, Soreq Cave, Israel. *Chem. Geol.* **169**, 145–156 (2000).
24. A. Almogi-Labin *et al*., Climatic variability during the last ~90 ka of the southern and northern Levantine basin as evident from marine records and speleothems. *Quat. Sci. Rev.* **28**, 2882–2896 (2009).
25. W. G. Deuser, E. H. Ross, L. S. Waterman, Glacial and pluvial periods: Their relationship revealed by Pleistocene sediments of the Red Sea and Gulf of Aden. *Science* **191**, 1168–1170 (1976).
26. E. Ivanitskaya, L. Rashkovetsky, E. Nevo, Chromosomes in a hybrid zone of Israeli mole rats (*Spalax*, Rodentia). *Russ. J. Genet.* **46**, 1149–1151 (2010).
27. E. Nevo, K. Kishi, A. Beiles, Genetic polymorphism of urine deoxyribonuclease I isomerases of subterranean mole rats, *Spalax ehrenbergi* superspecies, in Israel: Ecogeographical patterns and correlates. *Biochem. Genet.* **28**, 561–570 (1990).
28. E. Y. Durand, N. Patterson, D. Reich, M. Slatkin, Testing for ancient admixture between closely related populations. *Mol. Biol. Evol.* **28**, 2239–2252 (2011).
29. E. J. Steinig, M. Neuditschko, M. S. Khatkar, H. W. Raadsma, K. R. Zenger, Netview p: A network visualization tool to unravel complex population structure using genome-wide SNPs. *Mol. Ecol. Resour.* **16**, 216–227 (2016).
30. J. K. Pickrell, J. K. Pritchard, Inference of population splits and mixtures from genome-wide allele frequency data. *PLoS Genet.* **8**, e1002967 (2012).
31. N. Patterson *et al*., Ancient admixture in human history. *Genetics* **192**, 1065–1093 (2012).
32. H. Bozdogan, Model selection and Akaike's information criterion (AIC): The general theory and its analytical extensions. *Psychometrika* **52**, 345–370 (1987).
33. E. Tchernov, *Succession of Rodent Faunas during the Upper Pleistocene of Israel*. (Paul Parey, 1968).
34. V. Gorbunova *et al*., Cancer resistance in the blind mole rat is mediated by concerted necrotic cell death mechanism. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 19392–19396 (2012).
35. I. Manov *et al*., Pronounced cancer resistance in a subterranean rodent, the blind mole-rat, *Spalax*: In vivo and in vitro evidence. *BMC Biol.* **11**, 91 (2013).
36. T. Uehara, E. Kage-Nakadai, S. Yoshina, R. Imae, S. Mitani, The tumor suppressor BCL7B functions in the Wnt signaling pathway. *PLoS Genet.* **11**, e1004921 (2015).
37. Z. J. Rutnam, W. W. Du, W. Yang, X. Yang, B. B. Yang, The pseudogene *TUSC2P* promotes TUSC2 function by binding multiple microRNAs. *Nat. Commun.* **5**, 2914 (2014).
38. E. Nevo, G. Heth, A. Beiles, Aggression patterns in adaptation and speciation of subterranean mole rats. *J. Genet.* **65**, 65–78 (1986).
39. N. Danial-Farran *et al*., Adaptive evolution of coagulation and blood properties in hypoxia tolerant *Spalax* in Israel. *J. Zool. (Lond.)* **303**, 226–235 (2017).
40. J. Šklíba *et al*., Activity of free-living subterranean blind mole rats *Spalax galili* (Rodentia: Spalacidae) in an area of supposed sympatric speciation. *Biol. J. Linn. Soc. Lond.* **118**, 280–291 (2016).
41. L. Sun *et al*., Transcriptome response to heat stress in a chicken hepatocellular carcinoma cell line. *Cell Stress Chaperones* **20**, 939–950 (2015).
42. M. K. O'Bryan *et al*., Sox8 is a critical regulator of adult Sertoli cell function and male fertility. *Dev. Biol.* **316**, 359–370 (2008).
43. J. Zhao *et al*., Deletion of *Spata2* by CRISPR/Cas9n causes increased inhibin alpha expression and attenuated fertility in male mice. *Biol. Reprod.* **97**, 497–513 (2017).
44. F. Han *et al*., Gene flow, ancient polymorphism, and ecological adaptation shape the genomic landscape of divergence among Darwin's finches. *Genome Res.* **27**, 1004–1015 (2017).
45. T. Ma *et al*., Ancient polymorphisms and divergence hitchhiking contribute to genomic islands of divergence within a poplar species complex. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E236–E243 (2018).
46. Ž. Pezer, B. Harr, M. Teschke, H. Babiker, D. Tautz, Divergence patterns of genic copy number variation in natural populations of the house mouse (*Mus musculus domesticus*) reveal three conserved genes with major population-specific expansions. *Genome Res.* **25**, 1114–1124 (2015).
47. L. Couch, D. W. Duszynski, E. Nevo, Coccidia (Apicomplexa), genetic diversity, and environmental unpredictability of four chromosomal species of the subterranean superspecies *Spalax ehrenbergi* (mole-rat) in Israel. *J. Parasitol.* **79**, 181–189 (1993).
48. E. Nevo, A. Shkolnik, Adaptive metabolic variation of chromosome forms in mole rats, *Spalax*. *Experientia* **30**, 724–726 (1974).
49. A. Abyzov, A. E. Urban, M. Snyder, M. Gerstein, CNVnator: An approach to discover, genotype, and characterize typical and atypical CNVs from family and population genome sequencing. *Genome Res.* **21**, 974–984 (2011).
50. E. Nevo, P. Pirlot, A. Beiles, Brain size diversity in adaptation and speciation of subterranean mole rats. *J. Zool. Syst. Evol. Res.* **26**, 467–479 (1988).
51. E. Nevo, J. Klein, Structure and evolution of *Mhc* in subterranean mammals of the superspecies *Spalax ehrenbergi* in Israel. *Prog. Clin. Biol. Res.* **335**, 397–411 (1990).
52. S. Yahav, S. Simson, E. Nevo, The effect of protein and salt loading on urinary concentrating ability in four chromosomal species of *Spalax ehrenbergi*. *J. Zool. (Lond.)* **222**, 341–347 (1990).
53. F. Carducci, M. Biscotti, M. Barucca, A. Canapa, Transposable elements in vertebrates: Species evolution and environmental adaptation. *Eur. Zool. J.* **86**, 497–503 (2019).
54. B. J. Wagstaff, M. Barnerssoi, A. M. Roy-Engel, Evolutionary conservation of the functional modularity of primate and murine LINE-1 elements. *PLoS One* **6**, e19672 (2011).
55. A. Weyrich *et al*., Whole genome sequencing and methylome analysis of the wild Guinea pig. *BMC Genomics* **15**, 1036 (2014).
56. V. Grandjean, R. Yaman, F. Cuzin, M. Rassoulzadegan, Inheritance of an epigenetic mark: The CpG DNA methyltransferase 1 is required for de novo establishment of a complex pattern of non-CpG methylation. *PLoS One* **2**, e1136 (2007).
57. F. M. Catzeflis, E. Nevo, J. E. Ahlquist, C. G. Sibley, Relationships of the chromosomal species in the Eurasian mole rats of the *Spalax ehrenbergi* group as determined by DNA-DNA hybridization, and an estimate of the spalacid-murid divergence time. *J. Mol. Evol.* **29**, 223–232 (1989).
58. S. Via, Divergence hitchhiking and the spread of genomic isolation during ecological speciation-with-gene-flow. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **367**, 451–460 (2012).
59. A. P. Hendry, Evolutionary biology: Speciation. *Nature* **458**, 162–164 (2009).
60. E. Nevo, Evolution in action: Adaptation and incipient sympatric speciation with gene flow across life at "Evolution Canyon", Israel. *Isr. J. Ecol. Evol.* **60**, 85–98 (2014).
61. E. Nevo, Selection overrules gene flow at "Evolution Canyons", Israel. *Adv. Genet. Res.* **5**, 67–89 (2011).