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

Kexin Li1,b,1,2, Shangzhe Zhang1,b,1, Xiaoying Song1,a, Alexandra Weyrich1,d,1, Yinjia Wang1, Xi Liu1,1, Na Wan1, Jianquan Liu2, Matěj Lövy3, Haihong Cui2, Vladimir Frenkel3, Avi Titievsky9, Julia Panov5,9,10, Leonid Brodsky5,9, and Eviatar Nevo6,2

1State Key Laboratory of Grassland Agro-Ecosystem, Institute of Innovation Ecology, Lanzhou University, 730000 Lanzhou, China; 2Institute of Evolution, University of Haifa, 3498838 Haifa, Israel; 3School of Public Health, Lanzhou University, Lanzhou, 730000, China; 4Leibniz Institute for Zoo and Wildlife Research, 10315 Berlin, Germany; 5Faculty of Science, University of South Bohemia, 370 05 Ceské Budejovice, Czech Republic; 6Department of Gastroenterology, The 305 Hospital of PLA, 100017 Beijing, China; 7Tauber Bioinformatics Research Center, Haifa 3498838, Israel; and 8Sagol 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 dine 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 corresponds 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 sympathy identified geologically chal-kauskalnt 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.

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.

---

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. 1A), originated in Asia Minor and displays an outstanding three-pronged adaptive radiation into the Balkans, Ukraine, and Near East southward to North Africa, reflected in an increasing diplodid 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 clade 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*
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. 1C). 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. 1B).

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. 1B). 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. 1F). The three Northern (N) species (S. golani, S. gali-li_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. golani and S. judaei from K = 3 to K = 4 (Fig. 1G). 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. golani. 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. 2A), 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. 2A and 3A). 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 denote 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 (FST) 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 FST. The x-axis represents FST.

The genetic divergence, $F_{ST}$, between each species pair (SI Appendix, Table S5 and Fig. 2D) 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. 2D 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. 3B), NetView P (29) (Fig. 3D), and TreeMix (30) (Fig. 2B). The BABA-ABBA and identity-by-descent (IBD) (Fig. 2C) 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 species, as expected for species with significant interbreeding.
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. 2D).

Linkage disequilibrium (LD) of the four species drops rapidly, to below 0.3 within 5 kbp (Fig. 3C). 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. 3A 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. 24). No relatedness between S. galili and S. carmeli was detected by the network analysis (Fig. 3D), but both

---

**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 species 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 P2 compared with P3. (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.
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*<sub>ℓ</sub> tests to look for genes with high interspecific divergence and under positive selection driven by ecological stressors. 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, hypoxia, 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. galli*. It has been reported that ectopic expression of the *TUSC2* 3′-UTR inhibits cell proliferation, survival, migration, invasiveness and colony formation and furthermore causes tumor cell death in humans (37). This gene was selected only in *S. galli*, 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. galli* 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*, *Spata2*, *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. 4A-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. 4L](#) and [SI Appendix, Table S9](#)). Significantly elevated *d*<sub>ST</sub> ([Fig. 4K and M](#)) and *LD* ([Fig. 4O](#)) values were detected within *F*<sub>ST</sub>-islands of all of the 10 population pairs ([Fig. 4M](#) 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 (θ) and population-scaled recombination rates were significantly lower in island regions ([Fig. 4N 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. 4L](#) 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. galli* basalt vs. *S. galli* chalk.

Gene flow was mainly restricted between the peripatric chromosomal population pairs, and only slightly between the sympatric pair ([Fig. 2B](#)). 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. 5A–E](#)) shows that the comparison of *S. galli* chalk and *S. galli* basalt is related to angiogenesis, cancer, and autophagy ([Fig. 5D and Dataset S2](#)), and that between *S. carmeli* and *S. judaei* is related to water homeostasis, autophagy, and neurogenetics ([Fig. 5E and Dataset S3](#)).

**Copy Number Variations and Repeatedome.**

**Differences in copy number variations (CNV).** CNV regions (CNVRs) varied among the species and showed four clusters by principal component analysis (PCA) ([Fig. 6A and SI Appendix, Fig. S6](#)), phylogenetic tree analysis ([Fig. 6B](#)), and heatmap analysis ([Fig. 6C](#)). *S. carmeli* harbored the largest number of CNVRs (2,873), followed by *S. golani* (2,209), *S. galli* basalt (1,911), *S. galli* 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 S10B 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*<sub>ST</sub> was used to estimate population differentiation, which is similar to *F*<sub>ST</sub>, ranges from 0 to 1. *V*<sub>ST</sub> of all of the species pairs ([SI Appendix, Figs. S104 and S11](#)) showed the same trend as for *F*<sub>ST</sub>, suggesting the same selection trend on different mutations. In some species population pairs, one-half of the *V*<sub>ST</sub> values were extremely high, >0.5 ([Dataset S4](#)). Most of the genes with high *V*<sub>ST</sub> 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. S10B 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...
duplications and 6,182 CNVRs that did not overlap with large segmental duplications (SI Appendix, Fig. S10C). The length distribution of CNV that overlapped with genes is different between the two kinds of CNVs (SI Appendix, Fig. S10D).

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...
(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
taken as a continuous measurement yielded a similar separation of animals into four chromosomal \textit{Spalax} species (SI Appendix, Fig. S15).

We studied differential mutations in repetitive elements across all pairs of studied \textit{Spalax} species (SI Appendix, Materials and Methods) and found genomic positions in the LINE1 loci that have species-distinctive substitutions in several \textit{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. 6D, 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. 5F), 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 special cases, 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. 24). The divergent time estimated by whole genome sequencing (Figs. 24 and 34) is similar to that of DNA-DNA hybridization, highlighting the initiation evolution of the *S. ehrenbergi* superspecies in Israel approximately 1.6 ± 0.3 Ma (57), apparently within the range of the genomic estimate, which is 1.34 Ma. No gene flow was found due to strong divergent climatic adaptation (10) irrespective of the interspecific hybrid zones (13) (Fig. 2B). 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 *VST* of all species pairs (*SI Appendix, Figs. S10A and S11*) showed the same trend as *FST* on differential mutations (Dataset S2), i.e., adaptive evolution.

Most genes with high *VST* values occur in adaptive, speciation, and regulatory gene clusters related to metabolism, reproductive isolation (18–20), and inflammation (47). Generally, *FST* 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 *FST* but unchanged or decreased *dxy*. Divergence hitchhiking may allow for the significantly lower recombination rate and higher *dxy* 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).

**Data Availability.** Sequence Read Archive data have been deposited in China National Center for Bioinformation-National Genomics Data Center (accession numbers CRA003292 and CRA003322). All study data are included in the main text and *SI Appendix*.

**ACKNOWLEDGMENTS.** We thank Xu Zhang for the Perl scripts, Dorina Mivesghini for bioinformatics support, and Jorns 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.

**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 conducted using TreeBeeST, OCTA, and frappe, respectively. *FST* and Tajima’s *D* were calculated by VCFtools. Fluctuations in effective population size were calculated by PSMC, gene flow was assessed by T-reemix, *f*<sub>D</sub> 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 fastmoca2. Detailed information is available in *SI Appendix, Materials and Methods*.

2. J. Wahrman, R. Gafni, E. Nevo, Molecular clock: Speciation with admixture. Evolutionary significance of chro-


