

Article

Open Access

Comparative genomics of widespread and narrow-range white-bellied rats in the *Niviventer niviventer* species complex sheds light on invasive rodent success

Xin-Lai Wu^{1,2,6}, Dan-Ping Mu^{2,3}, Qi-Sen Yang², Yu Zhang², Yu-Chun Li⁴, Anderson Feijó², Ji-Long Cheng², Zhi-Xin Wen², Liang Lu⁵, Lin Xia², Zhi-Jun Zhou^{1,6,*}, Yan-Hua Qu^{2,*}, De-Yan Ge^{2,*}

¹ Key Laboratory of Zoological Systematics and Application of Hebei Province, College of Life Sciences, Hebei University, Baoding, Hebei 071002, China

² Key Laboratory of Zoological Systematics and Evolution, Institute of Zoology, Chinese Academy of Sciences, Beichen West Road, Chaoyang District, Beijing 100101, China

³ Xinjiang Key Laboratory of Biological Resources and Genetic Engineering, College of Life Science and Technology, Xinjiang University, Urumqi, Xinjiang 830046, China

⁴ Marine College, Shandong University, Weihai, Shandong 264209, China

⁵ State Key Laboratory for Infectious Diseases Prevention and Control, National Institute for Communicable Disease Control and Prevention, Chinese Center for Disease Control and Prevention, Beijing 102206, China

⁶ Hebei Basic Science Center for Biotic Interaction, Institute of Life Science and Green Development, Hebei University, Baoding, Hebei 071002, China

ABSTRACT

Widespread species that inhabit diverse environments possess large population sizes and exhibit a high capacity for environmental adaptation, thus enabling range expansion. In contrast, narrow-range species are confined to restricted geographical areas and are ecologically adapted to narrow environmental conditions, thus limiting their ability to expand into novel environments. However, the genomic mechanisms underlying the differentiation between closely related species with varying distribution ranges remain poorly understood. The *Niviventer niviventer* species complex (NNSC), consisting of highly abundant wild rats in Southeast Asia and China, offers an excellent opportunity to investigate these questions due to the presence of both widespread and narrow-range species that are phylogenetically closely related. In the present study, we combined ecological niche modeling with phylogenetic analysis, which suggested that sister species cannot be both widespread and dominant within the same geographical region. Moreover, by assessing heterozygosity, linkage disequilibrium decay, and Tajima's *D* analysis, we found that widespread species exhibited higher genetic diversity than narrow-range species. In addition, by exploring the “genomic islands of speciation”,

we identified 13 genes in highly divergent regions that were shared by the two widespread species, distinguishing them from their narrow-range counterparts. Functional annotation analysis indicated that these genes are involved in nervous system development and regulation. The adaptive evolution of these genes likely played an important role in the speciation of these widespread species.

Keywords: *Niviventer niviventer* species complex; Widespread species; Narrow-range species; Speciation; Genomic islands of differentiation; Phylogenomics

INTRODUCTION

Environmental change and human activity during the Anthropocene epoch have had a profound impact on the original geographical distribution of various species. This has resulted in significant shifts in their ranges and colonization of new areas, leading to considerable disruption of ecological dynamics within local communities. Consequently, the conservation of biodiversity and formulation of policy planning and management strategies face substantial challenges (Daru et al., 2021). Furthermore, species with different distributions

This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Copyright ©2023 Editorial Office of Zoological Research, Kunming Institute of Zoology, Chinese Academy of Sciences

Received: 24 March 2023; Accepted: 12 September 2023; Online: 13 September 2023

Foundation items: This work was supported by the Guangdong Provincial Key R&D Program (2022B1111040001), the Second Tibetan Plateau Scientific Expedition and Research Program (2019QZKK0402/2019QZKK0501), and National Natural Science Foundation of China (32170426)

*Corresponding authors, E-mail: zhijunzhou@hbu.edu.cn; quyh@ioz.ac.cn; gedy@ioz.ac.cn

may shift their ranges to different degrees. For example, species with broad distribution patterns tend to demonstrate robust dispersal capabilities and ability to adapt to diverse environments. As a result, these species possess heightened adaptability and a reduced risk of extinction. Conversely, species with limited geographical ranges tend to be confined to isolated ecological niches and display preferences for specific habitats (Slatyer et al., 2013). Thus, the ranges of widespread species tend to increase while those of narrow-range species tend to decrease under environmental change. This process involves certain species undergoing range contraction, while others undergo range expansion into previously unoccupied areas, often establishing themselves as invasive species and displacing native species that may lose their original habitats due to environmental disturbance (McKinney & Lockwood, 1999). Their impact on local ecosystems and biodiversity has received considerable attention (Hickling et al., 2006; Hill et al., 1999; Thomas et al., 2001), with the identification of various factors that affect species range size, including climate fluctuations, topographic heterogeneity, evolutionary history, land area, dispersal ability, annual fecundity, and body size (Gaston, 2009; Hawkins & Diniz-Filho, 2006; Laube et al., 2013; Lester et al., 2007; Pither, 2003; Sandel et al., 2011). However, the genomic mechanisms underlying the differentiation between widespread species and their closely related narrow-range congeneric species remain largely unexplored.

In recent decades, due to their rapid range expansion, brown rats (*Rattus norvegicus*), Asian house rats (*R. tanezumii*), and black rats (*R. rattus*) have demonstrated the highest invasion success among mammals (Puckett et al., 2016; Suzuki et al., 2019; Teng et al., 2017). They not only pose a substantial threat to agricultural and forestry production but, as commensal rodents, they also increase the risks to human health and ecosystem sustainability by serving as carriers of viruses and diseases, such as plague, hantavirus, Anjozorobe virus, and other zoonotic infections (Liu et al., 2017; Rabiee et al., 2018; Raharinosy et al., 2018). In addition to the aforementioned *Rattus* species, white-bellied rats (*Niviventer* Marshall, 1976) represent one of the most widely distributed and abundant genera of wild rats in Asia, with 23 recognized species, including 18 species in China (Ge et al., 2018, 2021a; Zhang et al., 2016). Notably, these rats are known vectors for many diseases and viruses, including hantavirus (Dai et al., 2019; Ge et al., 2016; Hu et al., 2014; Lin et al., 2012; Wang et al., 2000), *Borrelia*-causing Lyme disease (Masuzawa et al., 2001), tsutsugamushi disease (Latif et al., 2017), and babesiosis (Saito-Ito et al., 2008). Inference from demographic history has shown a marked reduction in range for species living in mid- to high-altitude regions, whereas Southeast Asian species have experienced considerable population and range expansion (Ge et al., 2021b). Of note, *Niviventer confucianus*, a habitat generalist within this genus, has shown a remarkable increase in population size and range since the late Quaternary, particularly in central and northern China (Ge et al., 2019). However, this population increase and range expansion is accompanied by a higher risk of rodent-borne diseases (Young et al., 2014).

Phylogenetic studies have revealed the existence of four distinct species complexes within *Niviventer*, including the *Niviventer andersoni* species complex (NASC), *Niviventer niviventer* species complex (NNSC), *Niviventer fulvescens*

species complex (NFSC), and *Niviventer eha* species complex (NESC) (Ge et al., 2018; He & Jiang, 2015; Lu et al., 2015; Zhang et al., 2016). Among these complexes, three expanded outwards from the southeastern Qinghai-Xizang Plateau, while one remains restricted to the southeastern edge of the Qinghai-Xizang Plateau (Ge et al., 2021b). The NNSC (*N. confucianus*, *N. bukit*, *N. tenaster*, *N. gladiusmaculus*, *N. pianmaensis*, *N. niviventer*, *N. lotipes*, *N. sacer*, *N. culturatus*, and *N. coninga*) represents the most abundant group of small mammals in Southeast Asia and China, providing an excellent model for studying the genomic differentiation between widespread and narrow-range species. Within the NNSC, *N. confucianus* and *N. lotipes* are widely distributed, while all other species are confined to limited geographic ranges. Notably, *N. confucianus* is widely distributed from the Indo-China Peninsula and adjacent mountains near the southeastern Qinghai-Xizang Plateau to the Loess Plateau and northern China, while *N. lotipes* is the most common faunal component in southeastern China. These widespread species may have benefited from the well-documented extinction of mega-fauna and the range contraction and population decline of large mammals since the late Pleistocene (Bartlett et al., 2016; Gill, 2014). In contrast, many species within the NNSC are confined to small regions. For example, *N. sacer* is narrowly distributed within Shandong, China, *N. gladiusmaculus*, *N. pianmaensis*, and *N. niviventer* are known only from the southeast region of the Qinghai-Xizang Plateau, *N. bukit* and *N. tenaster* are confined to Southeast Asia, and *N. culturatus* and *N. coninga* are endemic to Taiwan, China. Thus, we aimed to explore the potential genomic divergence between these widespread and narrow-range species, shedding light on their evolutionary trajectories and adaptations to diverse or specific environmental conditions.

In the present study, we conducted whole-genome sequencing of 50 individuals belonging to the *Niviventer* genus. Using single nucleotide polymorphism (SNP) data extracted from these genomes, we analyzed the phylogenomic structure and species differentiation within the NNSC. Through the integration of ecological niche models, our analysis revealed that the two widespread species were not closely related but exhibited partial sympatry in the NNSC. Subsequently, we performed genome scanning to assess the genomic differentiation between the widespread and narrow-range species. The primary objective of this study was to reveal the genomic mechanisms underlying the differentiation of widespread and narrow-range species in the NNSC using large-scale genomic data.

MATERIALS AND METHODS

Sampling

Tissue samples were collected from 50 white-bellied rats within the *Niviventer* genus, including eight NNSC species ($n=44$) and one NASC species ($n=6$). NNSC species, including *N. niviventer* ($n=3$), *N. tenaster* ($n=4$), *N. gladiusmaculus* ($n=4$), *N. bukit* ($n=6$), *N. pianmaensis* ($n=5$), *N. lotipes* ($n=6$), *N. confucianus* ($n=10$), and *N. sacer* ($n=6$), and NASC species *N. andersoni* ($n=6$) were used as the outgroup. All samples were collected at altitudes of 425 to 2830 m and were preserved in 95% ethanol at -80°C before DNA extraction. The molecular voucher specimens are preserved in the National Zoological Museum, Institute of

DNA extraction and sequencing

The cetyltrimethylammonium bromide (CTAB) method was used to extract genomic DNA from muscle tissue. The extracted genomic DNA was sequenced by the Beijing Berry (China) sequencing company. DNA concentration, purity, and integrity were tested using a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific, USA) and agarose electrophoresis. High-quality DNA samples were selected for library construction and sequencing. Whole-genome sequencing was performed using the Illumina HiSeq X Ten platform (Illumina, USA). From each sample, 1.5 µg of genomic DNA was sheared into 350 bp fragments using a Covaris instrument for Illumina library preparation. The fragmented DNA was end-polished, A-tailed, and ligated using a TruSeq Nano DNA HT Sample Preparation Kit (Illumina, USA). We used the AMPure XP system to purify polymerase chain reaction (PCR) products, with the libraries then analyzed using the Agilent 2100 Bioanalyzer (Agilent, USA). The Illumina HiSeq X Ten platform (Illumina, USA) was employed to sequence the constructed libraries and generate paired-end 150 bp reads.

Genomic data quality control

The raw read data of all samples were tested and filtered using fastp v0.20.1 (Chen et al., 2018). Low-quality bases with a Phred score < 20 were clipped from the 5' and 3' ends of the reads. Adapters and low-quality and duplicated reads were filtered out. FastQC v0.11.9 (Schmieder & Edwards, 2011) was used to measure the quality of clean data.

Read mapping and SNP identification

The 150 bp paired-end reads were mapped to the *Rattus norvegicus* genome (GCF_015227675.2_mRatBN7.2 from NCBI) by BWA v0.7.17 (Li & Durbin, 2009). Mapping results were then converted into BAM format and sorted with SAMtools v1.12, which also removed duplicate reads. If multiple read pairs had identical external coordinates, only the pair with the highest mapping quality was retained. After mapping, we performed SNP calling using BCFtools v1.12 (Li, 2011) and VCFtools v0.1.16. High-quality SNPs retained for further analysis had to meet the following criteria: (1) mean coverage depth ≥ 5, (2) missing rates < 30%, (3) genotype quality (GQ) > 20 and minor allele count (MAC) > 3, and (4) double allele retained.

Phylogenetics, genetic structure, and demographic parameters

After filtering, we generated a set of SNPs for subsequent analyses. The phylogenetic tree of the nine species was constructed using the maximum-likelihood (ML) approach in IQ-TREE v1.6.12 (Nguyen et al., 2015) with multiple nucleotide substitution models. The tree was inferred based on high-quality SNPs for 44 NNSC individuals, with six *N. andersoni* individuals set as the outgroup and with 1 000 bootstraps. Next, an individual-based neighbor-joining (NJ) tree was constructed for all samples based on the nucleotide *P*-distance matrix using VCF2Dis (<https://github.com/BGI-shenzhen/VCF2Dis>), with the phylogenetic relationships between individuals then inferred using FastMe v2.0 (Lefort et al., 2015). The NJ tree was rooted using the six *N. andersoni* individuals as the outgroup and visualized using FigTree v1.4.3 (<http://tree.bio.ed.ac.uk/software/figtree/>).

Principal component analysis (PCA) of the whole-genome SNPs for all 44 individuals was performed using Plink v1.90b6.21 (Purcell et al., 2007). Genetic structure and admixture proportion were assessed using default settings in ADMIXTURE v1.3.0 (Alexander & Lange, 2011). The number of assumed genetic clusters (*K*) ranged from 2 to 9. We implemented linkage disequilibrium (LD) pruning using the parameter --indep-pairwise (50 5 0.2) in PLINK. The r^2 for LD was calculated for SNPs using PopLDdecay (Zhang et al., 2019) with the parameter (-MaxDist 50). Heterozygosity of each individual was calculated using Plink v1.90b6.21 and Tajima's *D* value for each species was calculated using VCFtools in a 100 kb non-overlapping sliding window.

Ecological niche modeling to predict niche suitability under current climate

To predict the potential range size of eight NNSC species, we built ecological niche models using the MaxEnt v3.4.1 (Muscarella et al., 2014) algorithm implemented in the ENMeval R package. Nineteen temperature- and precipitation-related variables were downloaded from WorldClim (<http://www.worldclim.org>) at 10 min resolution for all periods (Supplementary Table S1). To minimize the bias introduced by high correlations between environmental factors and their impact on the potential distribution range size, we eliminated environmental factors with Pearson coefficients exceeding 0.8, resulting in the retention of eight uncorrelated environmental factors for subsequent analyses (Supplementary Figure S1). Species occurrence was limited to confirmed museum specimens or locations of sequenced individuals (Supplementary Table S3). We did not include *N. bukit* and *N. tenaster* as narrow-range species due to their limited distribution records and rare distribution in China. After excluding duplicated localities, a total of 10 000 pseudoabsence points were randomly selected for model training from a self-defined area, which included the entire distribution of the NNSC.

For each species, we used ENMeval wrapped with five feature classes ("L", "LQ", "H", "LQH", and "LQHP", L=linear, Q=quadratic, H=hinge, and P=product) and regularization multipliers (RM, from 0.5 to 5 with an increment of 1). We estimated the best-fitting models based on the smallest Akaike information criterion (AIC). To calculate the distribution range size, we converted the raw MaxEnt output into a logistic output with a range from 0 to 1. Finally, the logistic output data were used to calculate suitable area (km²) for each species.

Detection of selective sweeps

To identify genome-wide selective sweeps associated with widespread species, we used two different methods (F_{ST} and π ratio) to identify highly differentiated regions between widespread and narrow-range NNSC species, with a window size of 100 kb and step size of 10 kb. After Z-transformation of F_{ST} values, the regions in both the top 1% of $Z(F_{ST})$ values and top 1% of π ratios were chosen as selective sweep regions. To clarify the location of highly differentiated SNPs in the genome and corresponding functional genes, we used SnpEff v5.1 (Cingolani et al., 2012) to annotate SNP loci within the highly differentiated region (top 1%) between the two compared species. These highly differentiated genes were used for functional enrichment analysis by ShinyGO v0.75 (Ge et al., 2020), with $P < 0.05$ used as the threshold for significantly enriched pathways and functions.

RESULTS

Sequencing and SNP calling

Genome sequencing of the 50 NNSC and *N. andersoni* samples generated a total of 2 375.60 Gb of 150 bp paired-end reads, with an average depth of 17.73× per individual. Average genome depth was 15.10× (11.52–23.66) for *N. bukit*, 15.68× (12.04–32.92) for *N. confucianus*, 20.65× (12.01–43.81) for *N. gladiusmaculus*, 16.97× (10.79–39.12) for *N. lotipes*, 20.91× (13.21–35.09) for *N. niviventer*, 19.20× (13.80–39.25) for *N. pianmaensis*, 17.55× (11.83–42.75) for *N. sacer*, and 20.20× (11.53–40.44) for *N. tenaster*. An average of 96.14% of the NNSC sequencing reads were successfully mapped to the *R. norvegicus* reference genome. A total of 355 746 980 SNPs were obtained after SNP calling using SAMtools v1.12 (Li et al., 2009). A final set of 138 826 884 SNPs was retained after filtering with VCFtools v0.1.16 (Danecek et al., 2011). The average number of SNPs per species was 10 806 233 (Figure 1B; Supplementary Tables S3, S4).

Ecological niche modeling of NNSC

Based on the current distribution records of eight species within the NNSC, we predicted their potential niches using ecological niche modeling. Results showed that *N. confucianus* and *N. lotipes* exhibited the largest distribution ranges (>200 000 km²), while the remaining six species showed ranges of <100 000 km² (Figure 1C). We thus defined *N. confucianus* and *N. lotipes* as widespread species and *N. gladiusmaculus*, *N. pianmaensis*, *N. niviventer*, and *N. sacer* as narrow-range species.

Phylogenetic reconstruction of genomic structure of NNSC

The whole-genome ML tree, using *N. andersoni* as the outgroup, showed that *N. sacer* formed a monophyletic lineage and was sister to *N. confucianus* (Figure 1D). The two widespread species were not closely related, but their distributions overlapped in central and southeastern China. The reconstructed NJ tree indicated that the eight NNSC species were separated from each other and clustered in well-supported monophyletic clades, with the structure of the NJ and ML trees based on whole-genome SNPs showing similarity (Figure 2A). This result was confirmed by PCA, although *N. pianmaensis* and *N. tenaster* were clustered together and not clearly separated (Figure 2B). Population structure analysis using ADMIXTURE, which estimates individual ancestry and admixture proportions assuming K ancestral populations, indicated that the best-supported cluster was K=7. ADMIXTURE analysis also revealed a small amount of gene flow between *N. lotipes* and *N. confucianus* (Figure 2C). Regarding heterozygosity, *N. lotipes* (4.77E–2) demonstrated the largest heterozygosity, followed by *N. confucianus* (4.38E–2), while *N. gladiusmaculus* (1.49E–2) showed the lowest heterozygosity, followed by *N. tenaster* (1.59E–2), indicating high genetic diversity between *N. lotipes* and *N. confucianus* (Figure 2E). For Tajima's D, all species showed negative values, suggesting that the range of these species may expand in the future (Figure 2F). LD comparison among the NNSC members showed that *N. confucianus* and *N. lotipes* had more rapid LD decay than other species (Figure 2D). Overall, these results suggest that *N. confucianus* and *N. lotipes* (widespread species) have higher genetic

diversity (larger effective population sizes) than narrow-range species in the NNSC.

Detection signature of selection between widespread and narrow-range species

To clarify the genomic differences between widespread and narrow-range species, F_{ST} and π ratio sweep analyses were used to detect common selective sweep regions between these species within the NNSC (Figure 3A). Notably, the widespread species were highly differentiated from the narrow-range species on chromosome 13 (Figure 3B, C). Using global F_{ST} , we scanned for genomic regions with extreme allele frequency differentiation. The top 1% of F_{ST} values revealed numerous candidate genes (486–740) in the different groups after annotation and removal of repeats. In addition, for the different groups, functional analysis of the candidate genes based on Gene Ontology (GO) categories identified two common functions (G protein-coupled receptor signaling pathway and nervous system process) (Figure 3D, E; Supplementary Tables S5–S9). Comparing the widespread species *N. confucianus* with narrow-range species, we identified the most candidate genes in the CG (*N. confucianus* vs. *N. gladiusmaculus*) group (740) and the fewest candidate genes in the CP (*N. confucianus* vs. *N. pianmaensis*) group (486). Among these genes, 34 were shared by the four species pairs (Supplementary Figure S2A). Comparing the widespread species *N. lotipes* with narrow-range species, we identified the most candidate genes in the LG (*N. lotipes* vs. *N. gladiusmaculus*) group (803) and the fewest in the LS (*N. lotipes* vs. *N. sacer*) group (463). Among these genes, 45 were shared by the four species pairs (Supplementary Figure S2B). In addition, we identified 22 shared candidate genes in these two sets of comparisons (*N. confucianus* vs. four narrow-range species and *N. lotipes* vs. four narrow-range species) (Supplementary Figure S2C and Table S10).

Regarding the π ratio results, we identified 38 significantly differentiated genes between the widespread and narrow-range species (Supplementary Figure S3 and Table S11), which were mainly enriched in neuronal development, nervous system development, and sympathetic nervous system development (Figure 4A; Supplementary Tables S12–S19). Most of these genes are highly expressed in the brain and are essential for brain development (Allen et al., 2005; Anney et al., 2010; Howell et al., 1997; Ito et al., 2018; Laughlin et al., 2011; Stefansson et al., 2009), suggesting that cognitive improvement aids in exploratory behavior and may play a key role in species expansion.

Finally, based on F_{ST} and π ratio, we identified 13 candidate genes (*LOC100360923*, *LOC100362382*, *LOC100911899*, *LOC102553684*, *LOC103691598*, *LOC120096228*, *LOC120096412*, *LOC120096444*, *LOC682479*, *LOC304725*, *HSPD-PS8*, *LOC679734*, and *RGD1561440*) between the widespread and narrow-range species (Supplementary Figure S4 and Table S20). Interestingly, these candidate genes were all located on chromosome 13 within an approximately 5 Mb region (Figure 4B), with the haplotype map of two genes (*LOC103691598* and *LOC304725*) showing significant differences between the widespread and narrow-range species (Figure 4C). Subsequently, based on the NJ tree constructed using SNPs, these three genes were differentiated between the widespread and narrow-range species (Figure 4D).

We next combined all widespread individuals as one group

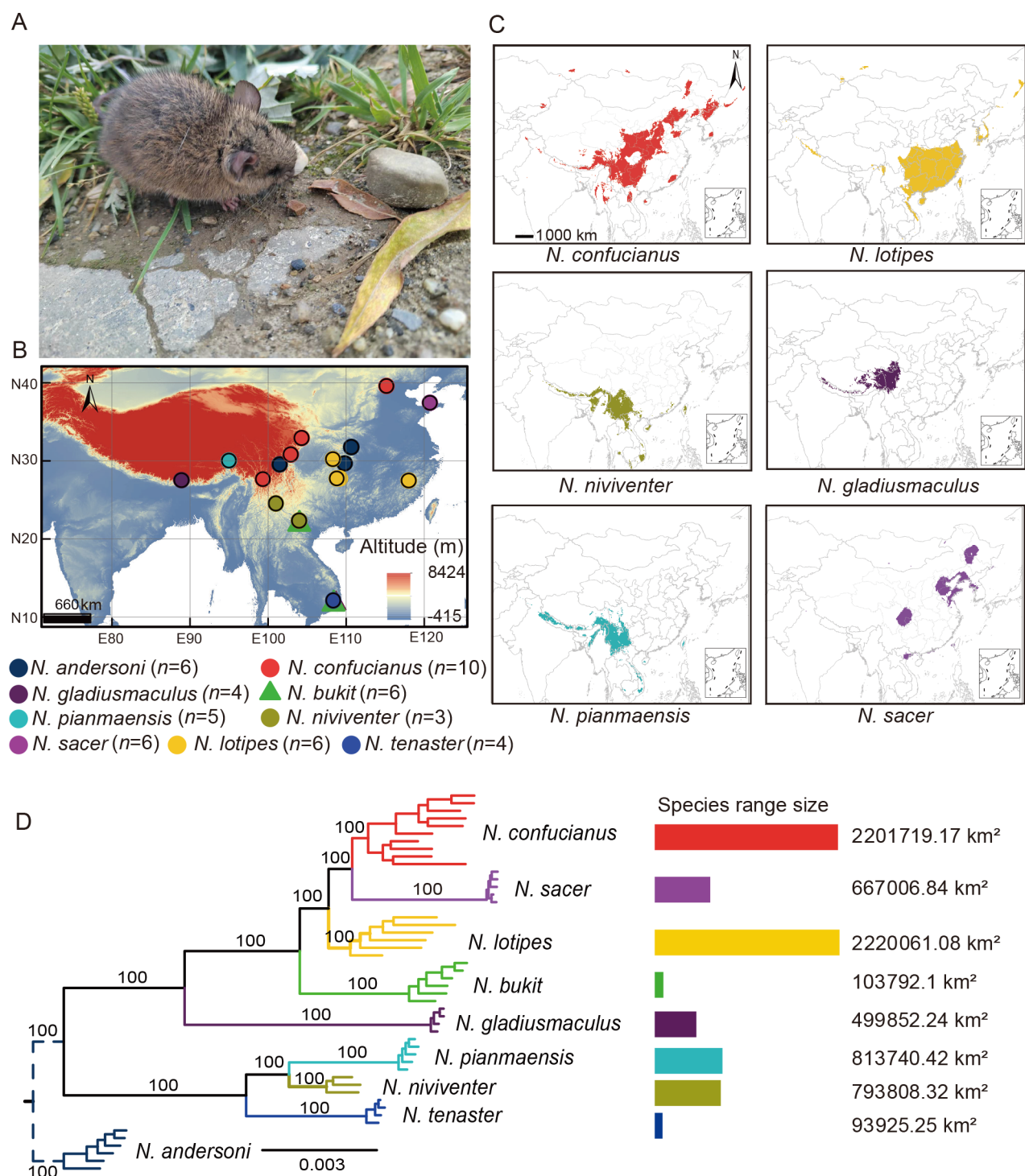


Figure 1 Geographic distribution and phylogenetic relationships in *Niviventer nivienter* species complex (NNSC)

A: *Niviventer confucianus*. Photo by De-Yan Ge. **B:** Geographical distribution of NNSC samples used in our study. Different species are represented by different colors. **C:** Potential distribution range of widespread and narrow-range species predicted by MaxEnt. **D:** ML tree of NNSC reconstructed based on whole-genome SNPs and overall distribution area of different species in NNSC. Length of column chart indicates size of potential distribution range for different species calculated by MaxEnt.

and all narrow-range individuals as another group. Results showed that the highly differentiated genes between the widespread and narrow-range species were mainly located on chromosome 13 when calculating the F_{ST} values of the two groups (Supplementary Figure S5A), with these genes mainly enriched in neuronal development, system development, and neuron projection development (Supplementary Figure S5B). These results are consistent with the above findings.

DISCUSSION

Phylogenetic structure of NNSC

The taxonomic classification of *N. sacer* within the NNSC remains a subject of debate. Historically, *N. sacer* was considered a subspecies of *N. confucianus* (Allen, 1926; Thomas, 1908; Wang, 2003). After conducting multilocus phylogenetic analysis using nuclear genes, Zhang et al.

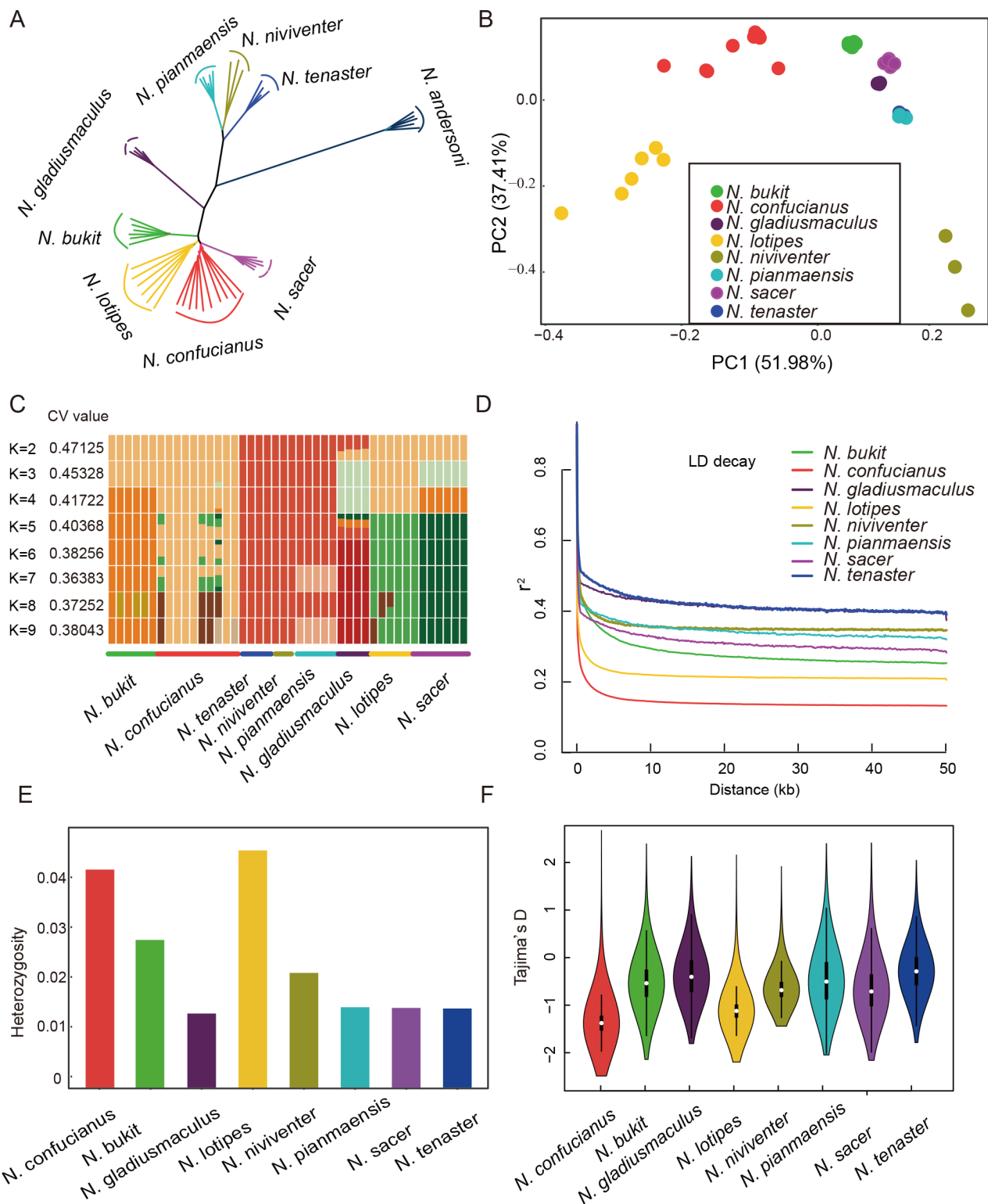


Figure 2 Species structure and genetic relationships of NNSC

A: NJ tree of NNSC with *Niviventer andersoni* as an outgroup. B: Plots of principal component 1 versus principal component 2 (PC1 vs. PC2) for NNSC. C: Clustering solution inferred using ADMIXTURE. Color in each column represents ancestral proportions, with presumed group sizes from K=2 to K=9. D: LD decay for each species. Same number of individuals was chosen randomly for each species to calculate r^2 . E: Comparisons of heterozygosity among eight species in NNSC. Heterozygosity of widespread species (*N. confucianus* and *N. lotipes*) was higher than that of small-range species. F: Genome-wide Tajima's D of each species based on a 100 kb sliding window.

(2016) reported *N. sacer* as an independent phylogenetic clade, although they did not consider the genetic difference to reach the species level. Furthermore, based on analysis of phylogenetic relationships, genetic distances, chromosomal characteristics, and morphology, Li et al. (2020) suggested

that *N. sacer* should be reclassified as a distinct species rather than a subspecies of *N. confucianus*. However, these previous studies relied on limited and short DNA fragments for their analyses. Therefore, in our study, we utilized whole-genome data to reconstruct the phylogenetic tree of the NNSC, which

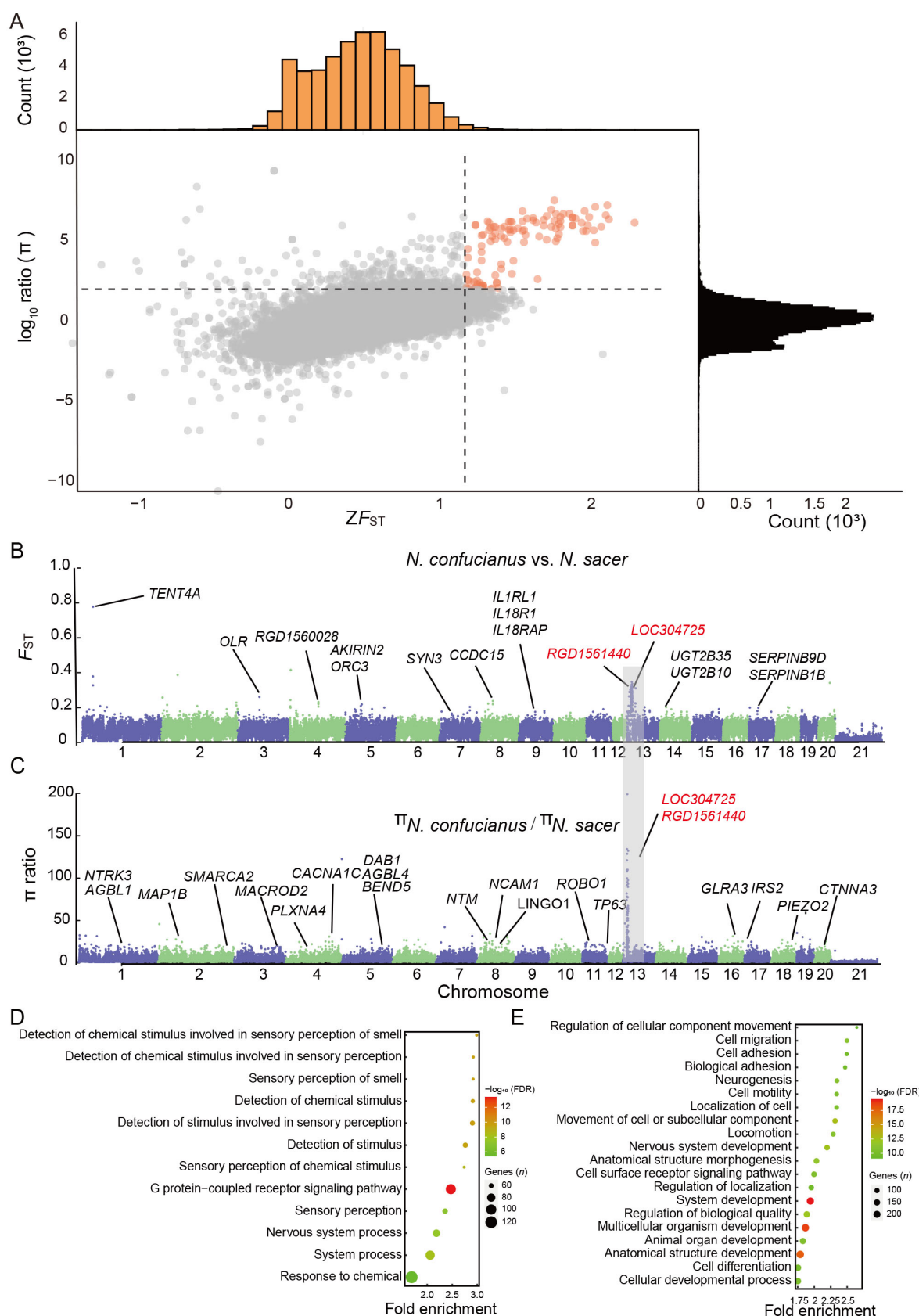


Figure 3 Selection signals related to differentiation between widespread and narrow-range species genomes

A: Distribution of ZF_{ST} and $\log_{10}$ ratio (π , *N. confucianus*/ π , *N. sacer*) values (indicated by orange and black, respectively) calculated in 100 kb sliding windows with 10 kb steps between *N. confucianus* and *N. sacer*. Data points in red (corresponding to top 1% of empirical ZF_{ST} distribution with values >1.18 and top 1% of empirical $\log_{10}$ ratio (π) distribution with values >2.098) are genomic regions under positive selection between *N. confucianus* and *N. sacer*. B: Selection signals between *N. confucianus* and *N. sacer* detected by F_{ST} . C: Selection signals between *N. confucianus* and *N. sacer* detected by π ratio. D, E: GO enrichment analyses of highly differentiated genes between *N. confucianus* and *N. sacer* detected by F_{ST} (left) and π ratio (right).

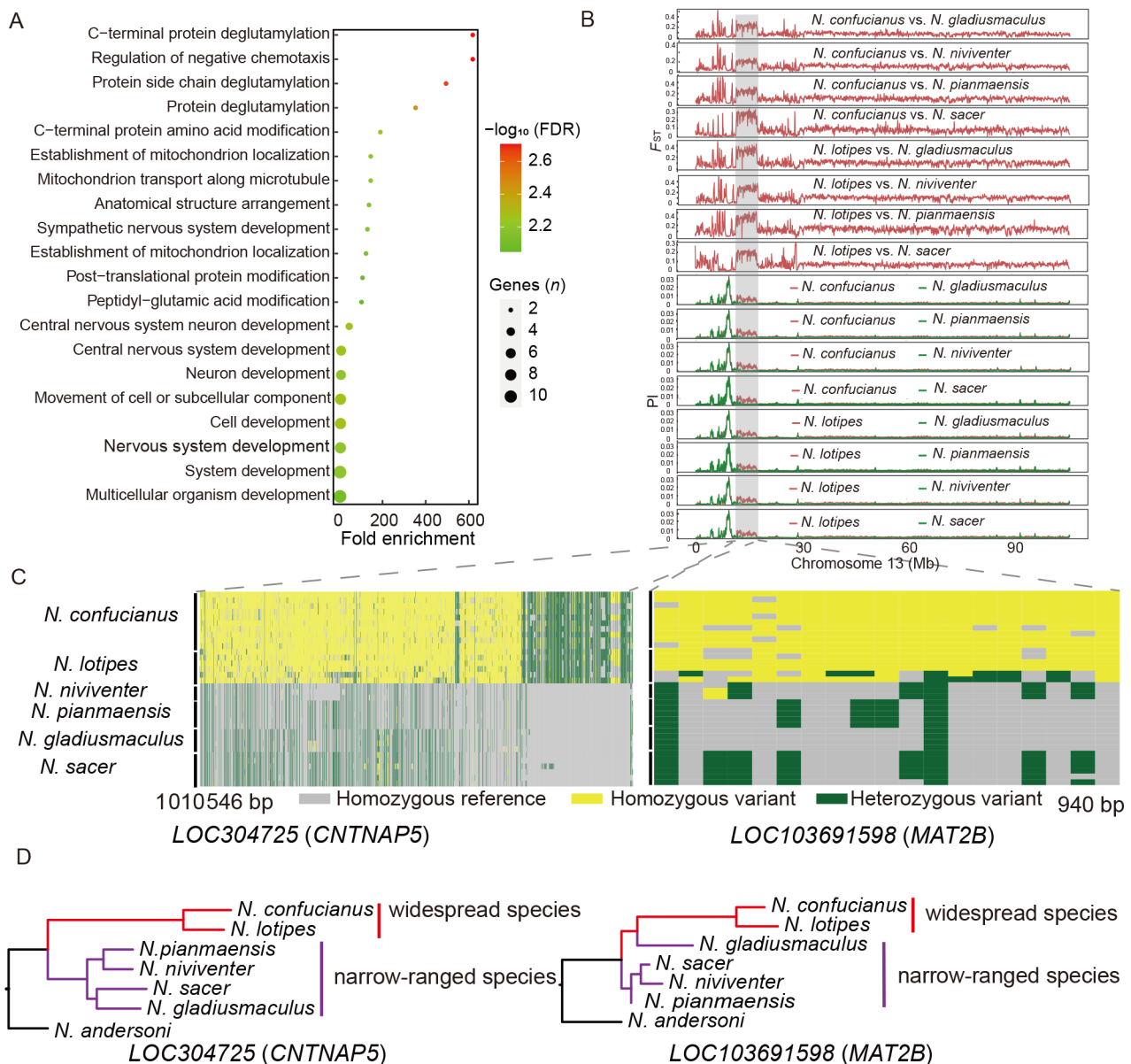


Figure 4 Genes associated with divergence between widespread and narrow-range species

A: GO enrichment analyses of 38 shared genes. B: Different approaches identifying highly differentiated regions on chromosome 13 between widespread and narrow-range species. C: Genetic map of LOC304725 and LOC103691598. D: Gene tree of LOC304725 and LOC103691598.

strongly supported *N. sacer* as an independent species. Subsequent NJ tree analysis based on genome-wide SNP, PCA, and ADMIXTURE analyses showed a significant separation between *N. sacer* and *N. confucianus*. Interestingly, the phylogenetic trees revealed a substructure within *N. confucianus* across geographic locations, with six individuals from southern China and four individuals from northern China clustered into two different lineages. Consistently, Ge et al. (2019) also found that many *N. confucianus* samples from three different geographical regions were clustered into three distinct lineages. Our phylogenetic findings further support the substantial genomic differentiation of *N. confucianus* across various geographic locations, thus corroborating previous studies.

Sister species cannot be widespread and dominant in the same geographical region

Interspecific competition refers to the shared use of limited resources, such as food, breeding sites, and habitats, by

multiple sympatric species (Arthur, 1982). Gause (1934) demonstrated that if the resource requirements of two species completely overlap, they cannot coexist within the same landscape, leading to the exclusion of one species by the other. Building upon this understanding, Hardin (1960) proposed the well-known competitive exclusion principle in ecology, which suggests that two closely related species face considerable challenges in occupying the same or similar ecological niches, with most unable to coexist long-term due to intense competition for limited food, space, or other environmental resources.

Our phylogenetic analysis identified *N. confucianus* and *N. sacer* as sister species and *N. pianmaensis* and *N. niviventer* as sister species. Interestingly, we did not find sister species that were both widespread in the same geographical region. Ecological niche modeling suggested a slight range overlap between the two widespread species (*N. confucianus* and *N. lotipes*). We propose that the ability of both species to become

widespread is facilitated by their non-overlapping distribution ranges, which reduces competition for food, breeding sites, and other resources and contributes to their dominance within their respective distribution ranges. Thus, sister species cannot both become widespread in the same region due to competition for limited resources, as supported by multiple examples in nature. For instance, the wild boar (*Sus scrofa*), a globally distributed ecological generalist, has experienced a substantial population increase in many regions (Erkinaro et al., 1982). In contrast, its sister species, the bearded pig (*S. barbatus*), remains restricted to specific areas in Sumatra, Borneo, Malay Peninsula, and certain smaller islands in Southeast Asia, and is currently classified as vulnerable due to population decline. Sister species *Lepus tolai* and *L. yarkandensis* show a similar pattern, with the former widely distributed in China, including northeastern, northwestern, and middle China, while the latter is only found in the Tarim Basin of western China (Gao, 1983). In addition, the striped field mouse (*Apodemus agrarius*) boasts an extensive range spanning the Palearctic and Indomalaya regions (Karaseva et al., 1992), while its sister species, *A. chevrieri*, is only native to southwest China. Therefore, based on phylogenetic analysis, ecological niche modeling, and the principle of competitive exclusion, we hypothesize that sister species cannot be both widespread and dominant within the same geographical region.

Genomic differentiation between widespread and narrow-range species

The alteration of range sizes plays a significant role in the evolutionary history of species. Members within the NNSC are recognized as highly abundant small mammals in Southeast Asia and China. Recent investigations have revealed the dispersal of the NNSC across southeastern, central, and northern regions of China (Ge et al., 2021b). Our findings suggested the occurrence of gene flow between *N. lotipes* and *N. confucianus*, implying potential hybridization between these two widespread species during the process of population expansion in overlapping distribution areas. Notably, compared with narrow-range species, widespread species in the NNSC exhibited higher levels of genetic diversity, which may contribute to their adaptability to changing environments and rapid expansion (Krehenwinkel et al., 2015).

Based on analysis of highly divergent outlier genomic regions in both widespread and narrow-range species, we found that genes in the top 1% of F_{ST} outlier regions were related to nervous system processes, G protein-coupled receptor signaling pathway, sensory perception of smell, and detection of chemical stimuli involved in sensory perception of smell. Certainly, increased cognitive abilities would aid rodents in adapting to changes in the external environment, while improved olfactory abilities would enable rats to survive in different habitats. Moreover, G protein-coupled receptors have been linked to population differentiation in rodents and ecological adaptation in house mice (*Mus musculus*) (Janoušek et al., 2015; Staubach et al., 2012). Furthermore, using the π ratio, we observed the presence of positively selected immune-related genes in widespread species, indicating that these species experienced strong selection pressure from external pathogens during range expansion.

Elucidating the selection signatures in widespread species is important for understanding adaptive evolution in wild rodents in response to different environments. In the current

study, we identified several crucial genes that exhibited high expression levels in the brain, including *SYN3*, *DAB1*, *LINGO1*, *MACROD2*, and *MAP1B*. Several of these key nuclear genes, such as synapsin III (*SYN3*), are expressed in the hippocampus, striatum, and neocortex (Laughlin et al., 2011), with *SYN3* expression found to be positively correlated with behavioral flexibility (Moore et al., 2021). Moreover, Disabled-1 (*DAB1*), a cytoplasmic adaptor protein, regulates neuronal migration during mammalian brain development (Howell et al., 1997; Sheldon et al., 1997). Leucine-rich repeat and Ig domain containing 1 (*LINGO1*), another gene overexpressed in the brain, is associated with an increased risk of essential tremor (ET) (Stefansson et al., 2009). MACRO domain containing 2 (*MACROD2*) is a neurodevelopmental disorder-related mono-ADP-ribosylhydrolase (Ito et al., 2018) involved in the etiology of neurodevelopmental, neurological, and neuropsychiatric diseases (Anney et al., 2010; Lionel et al., 2011; Xu et al., 2009). Microtubule-associated protein 1B (*MAP1B*) participates in several neuronal developmental processes, such as neuronal migration, neurite growth, and growth cone function, and is associated with several neurological disorders, such as fragile X syndrome (Lu et al., 2004) and giant axonal neuropathy (Allen et al., 2005). Interestingly, we observed significant genomic differences between the widespread and narrow-range species on chromosome 13. Within this chromosome, the *LOC304725* gene is similar to contactin-associated protein-like 5 (*CNTNAP5*). *CNTNAP5*, which belongs to the neurexin gene family, is associated with epilepsy, schizophrenia, autoimmune encephalitis, and multiple sclerosis (Pagnamenta et al., 2010; Zou et al., 2017). Therefore, our findings suggest that adaptive evolution of these genes may facilitate range expansion in widespread species. The superior environmental adaptability and expansion capacity of *N. confucianus* and *N. lotipes* are key factors driving their dominance as rodent species in China.

In recent years, increasing studies have explored the factors that may contribute to the successful invasion and wide distribution of certain rodent species. Many of these studies have suggested that genes involved in the nervous system, G protein-coupled receptor signaling pathway, and immunity may facilitate local adaptation and range expansion (Chen et al., 2021; Harpak et al., 2021; Zeng et al., 2018), consistent with our findings. Moreover, we also observed that genes related to sensory perception of smell and detection of chemical stimulus involved in sensory perception of smell underwent positive selection in the widespread species. It is important to note that the expansion of these species may have significant consequences. They may compete with native species for food and ecological niches, thereby placing native species at risk of population decline due to resource scarcity. Additionally, these invasive species can serve as hosts for a variety of zoonotic viruses, posing potential risks to human safety.

In conclusion, based on phylogenetic tree construction from genomic data and prediction of potential distribution ranges, our study revealed several important findings. Firstly, we determined that sister species within the same geographical region cannot be both widespread and dominant. Second, we revealed that genetic diversity was higher in the widespread species, such as *N. confucianus* and *N. lotipes*, than in narrow-range species, indicating robust adaptability to environmental changes. Finally, by comparing widespread and

narrow-range species, we identified several candidate genes (*SYN3*, *DAB1*, *LINGO1*, *MACROD2*, *MAP1B*, *LOC304725*, and *LOC103691598*) that were highly expressed in the brain and located in “genomic islands of speciation”. These genes are involved in hindbrain structural organization, cerebellum structural organization, and neuronal development and differentiation. Collectively, our findings provide valuable insights into the genomic mechanisms underlying the differentiation of widespread and narrow-range species.

DATA AVAILABILITY

The original genomic sequencing reads can be downloaded from the NCBI (PRJNA1026848), China National Center for Bioinformation (PRJCA020058), and Science Data Bank databases (DOI: 10.57760/sciencedb.j00139.00066).

SCIENTIFIC FIELD SURVEY PERMISSION INFORMATION

Permission for field surveys was granted by the Science and Technology Department of Sichuan Province in the Second Tibetan Plateau Scientific Expedition and Research Program from 2019–2023.

SUPPLEMENTARY DATA

Supplementary data to this article can be found online.

COMPETING INTERESTS

The authors declare that they have no competing interests.

AUTHORS' CONTRIBUTIONS

D.Y.G. and X.L.W. designed the research. D.Y.G., X.L.W., Z.X.W., A.F., J.L.C., L.X., Q.S.Y., Y.C.L., L.L., and D.P.M. participated in sample collection. X.L.W. and D.Y.G. performed data collection and analyses. X.L.W., D.Y.G., Z.J.Z., Y.Z., and Y.H.Q. wrote the paper. All authors read and approved the final version of the manuscript.

ACKNOWLEDGMENTS

We thank Xuan-Zhao Li from Hebei University for help with data analysis.

REFERENCES

- Alexander DH, Lange K. 2011. Enhancements to the ADMIXTURE algorithm for individual ancestry estimation. *BMC Bioinformatics*, **12**: 246.
- Allen E, Ding JQ, Wang W, et al. 2005. Gigaxonin-controlled degradation of *MAP1B* light chain is critical to neuronal survival. *Nature*, **438**(7065): 224–228.
- Allen GM. 1926. Rats (Genus *Rattus*) from the Asiatic Expeditions. New York: The American Museum of Natural History.
- Anney R, Klei L, Pinto D, et al. 2010. A genome-wide scan for common alleles affecting risk for autism. *Human Molecular Genetics*, **19**(20): 4072–4082.
- Arthur W. 1982. The evolutionary consequences of interspecific competition. *Advances in Ecological Research*, **12**: 127–187.
- Bartlett LJ, Williams DR, Prescott GW, et al. 2016. Robustness despite uncertainty: regional climate data reveal the dominant role of humans in explaining global extinctions of Late Quaternary megafauna. *Ecography*, **39**(2): 152–161.
- Chen SF, Zhou YQ, Chen YR, et al. 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*, **34**(17): i884–i890.
- Chen Y, Zhao L, Teng HJ, et al. 2021. Population genomics reveal rapid genetic differentiation in a recently invasive population of *Rattus norvegicus*. *Frontiers in Zoology*, **18**(1): 6.
- Cingolani P, Platts A, Wang LL, et al. 2012. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w¹¹¹⁸; iso-2; iso-3. *Fly*, **6**(2): 80–92.
- Dai XX, Jian C, Li N, et al. 2019. Characterization of the L genome segment of an orthohantavirus isolated from *Niviventer confucianus*. *Archives of Virology*, **164**(2): 613–616.
- Danecek P, Auton A, Abecasis G, et al. 2011. The variant call format and VCFtools. *Bioinformatics*, **27**(15): 2156–2158.
- Daru BH, Davies TJ, Willis CG, et al. 2021. Widespread homogenization of plant communities in the Anthropocene. *Nature Communications*, **12**(1): 6983.
- Erkinaro E, Heikura K, Lindgren E, et al. 1982. Occurrence and spread of the wild boar (*Sus scrofa*) in eastern Fennoscandia. *Memoranda Societatis pro Fauna et Flora Fennica*, **58**(2): 39–47.
- Gao YT. 1983. Current studies on the Chinese Yarkand hare. *Acta Zoologica Fennica*, **174**: 23–25.
- Gaston KJ. 2009. Geographic range limits of species. *Proceedings of the Royal Society B: Biological Sciences*, **276**(1661): 1391–1393.
- Gause GF. 1934. Experimental analysis of Vito Volterra's mathematical theory of the struggle for existence. *Science*, **79**(2036): 16–17.
- Ge DY, Feijó A, Abramov AV, et al. 2021a. Molecular phylogeny and morphological diversity of the *Niviventer fulvescens* species complex with emphasis on species from China. *Zoological Journal of the Linnean Society*, **191**(2): 528–547.
- Ge DY, Feijó A, Wen ZX, et al. 2021b. Demographic history and genomic response to environmental changes in a rapid radiation of wild rats. *Molecular Biology and Evolution*, **38**(5): 1905–1923.
- Ge DY, Lu L, Abramov AV, et al. 2019. Coalescence models reveal the rise of the white-bellied rat (*Niviventer confucianus*) following the loss of Asian Megafauna. *Journal of Mammalian Evolution*, **26**(3): 423–434.
- Ge DY, Lu L, Xia L, et al. 2018. Molecular phylogeny, morphological diversity, and systematic revision of a species complex of common wild rat species in China (Rodentia, Murinae). *Journal of Mammalogy*, **99**(6): 1350–1374.
- Ge SX, Jung D, Yao RN. 2020. ShinyGO: a graphical gene-set enrichment tool for animals and plants. *Bioinformatics*, **36**(8): 2628–2629.
- Ge XY, Yang WH, Pan H, et al. 2016. Erratum to: Fugong virus, a novel hantavirus harbored by the small oriental vole (*Eothenomys eleusis*) in China. *Virology Journal*, **13**: 75.
- Gill JL. 2014. Ecological impacts of the late Quaternary megaherbivore extinctions. *New Phytologist*, **201**(4): 1163–1169.
- Hardin G. 1960. The competitive exclusion principle. *Science*, **131**(3409): 1292–1297.
- Harpak A, Garud N, Rosenberg NA, et al. 2021. Genetic adaptation in New York City rats. *Genome Biology and Evolution*, **13**(1): evaa247.
- Hawkins BA, Diniz-Filho JAF. 2006. Beyond Rapoport's rule: evaluating range size patterns of New World birds in a two-dimensional framework. *Global Ecology and Biogeography*, **15**(5): 461–469.
- He K, Jiang XL. 2015. Mitochondrial phylogeny reveals cryptic genetic diversity in the genus *Niviventer* (Rodentia, Muroidea). *Mitochondrial DNA*, **26**(1): 48–55.
- Hickling R, Roy DB, Hill JK, et al. 2006. The distributions of a wide range of taxonomic groups are expanding polewards. *Global Change Biology*, **12**(3): 450–455.
- Hill JK, Thomas CD, Huntley B. 1999. Climate and habitat availability determine 20th century changes in a butterfly's range margin. *Proceedings of the Royal Society B: Biological Sciences*, **266**(1425): 1197–1206.
- Howell BW, Hawkes R, Soriano P, et al. 1997. Neuronal position in the developing brain is regulated by *mouse disabled-1*. *Nature*, **389**(6652): 733–737.
- Hu XQ, Li SG, Liu H, et al. 2014. Diversity and distribution of host animal species of hantavirus and risk to human health in Jiuhua Mountain Area, China. *Biomedical and Environmental Sciences*, **27**(11): 849–857.
- Ito H, Morishita R, Mizuno M, et al. 2018. Biochemical and morphological

- characterization of a neurodevelopmental disorder-related mono-ADP-Ribosylhydrolase, MACRO domain containing 2. *Developmental Neuroscience*, **40**(3): 278–287.
- Janoušek V, Munclinger P, Wang LY, et al. 2015. Functional organization of the genome may shape the species boundary in the house mouse. *Molecular Biology and Evolution*, **32**(5): 1208–1220.
- Karaseva EV, Tikhonova GN, Bogomolov PL. 1992. Distribution of the striped field-mouse (*apodemus-agrarius*) and peculiarities of its ecology in different parts of its range. *Zoologicheskyy Zhurnal*, **71**(6): 106–115.
- Krehenwinkel H, Rödder D, Tautz D. 2015. Eco-genomic analysis of the poleward range expansion of the wasp spider *Argiope bruennichi* shows rapid adaptation and genomic admixture. *Global Change Biology*, **21**(12): 4320–4332.
- Latif A, Liu BY, Chen Z, et al. 2017. *Orientia tsutsugamushi* infection in rodents in Anhui Province of China. *Infection, Genetics and Evolution*, **56**: 14–18.
- Laube I, Korntheuer H, Schwager M, et al. 2013. Towards a more mechanistic understanding of traits and range sizes. *Global Ecology and Biogeography*, **22**(2): 233–241.
- Laughlin RE, Grant TL, Williams RW, et al. 2011. Genetic dissection of behavioral flexibility: reversal learning in mice. *Biological Psychiatry*, **69**(11): 1109–1116.
- Lefort V, Desper R, Gascuel O. 2015. FastME 2.0: a comprehensive, accurate, and fast distance-based phylogeny inference program. *Molecular Biology and Evolution*, **32**(10): 2798–2800.
- Lester SE, Ruttenberg BI, Gaines SD, et al. 2007. The relationship between dispersal ability and geographic range size. *Ecology Letters*, **10**(8): 745–758.
- Li H. 2011. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics*, **27**(21): 2987–2993.
- Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, **25**(14): 1754–1760.
- Li H, Handsaker B, Wysoker A, et al. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, **25**(16): 2078–2079.
- Li YY, Li YQ, Li HT, et al. 2020. *Niviventer confucianus sacer* (Rodentia, Muridae) is a distinct species based on molecular, karyotyping, and morphological evidence. *ZooKeys*, **959**: 137–159.
- Lin XD, Wang W, Guo WP, et al. 2012. Cross-species transmission in the speciation of the currently known *Murinae*-associated hantaviruses. *Journal of Virology*, **86**(20): 11171–11182.
- Lionel AC, Crosbie J, Barbosa N, et al. 2011. Rare copy number variation discovery and cross-disorder comparisons identify risk genes for ADHD. *Science Translational Medicine*, **3**(95): 95ra75.
- Liu DY, Liu J, Liu BY, et al. 2017. Phylogenetic analysis based on mitochondrial DNA sequences of wild rats, and the relationship with Seoul virus infection in Hubei, China. *Virologica Sinica*, **32**(3): 235–244.
- Lu L, Ge DY, Chesters D, et al. 2015. Molecular phylogeny and the underestimated species diversity of the endemic white-bellied rat (Rodentia: Muridae: *Niviventer*) in Southeast Asia and China. *Zoologica Scripta*, **44**(5): 475–494.
- Lu R, Wang HP, Liang Z, et al. 2004. The fragile X protein controls microtubule-associated protein 1B translation and microtubule stability in brain neuron development. *Proceedings of the National Academy of Sciences of the United States of America*, **101**(42): 15201–15206.
- Masuzawa T, Takada N, Kudeken M, et al. 2001. *Borrelia sinica* sp. nov., a Lyme disease-related *Borrelia* species isolated in China. *International Journal of Systematic and Evolutionary Microbiology*, **51**(Pt 5): 1817–1824.
- McKinney ML, Lockwood JL. 1999. Biotic homogenization: a few winners replacing many losers in the next mass extinction. *Trends in Ecology & Evolution*, **14**(11): 450–453.
- Moore A, Linden J, Jentsch JD. 2021. *Syn3* gene knockout negatively impacts aspects of reversal learning performance. *eNeuro*, **8**(5): ENEURO.0251–21.2021.
- Muscarella R, Galante PJ, Soley-Guardia M, et al. 2014. ENMeval: an R package for conducting spatially independent evaluations and estimating optimal model complexity for MAXENT ecological niche models. *Methods in Ecology and Evolution*, **5**(11): 1198–1205.
- Nguyen LT, Schmidt HA, Von Haeseler A, et al. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating Maximum-likelihood phylogenies. *Molecular Biology and Evolution*, **32**(1): 268–274.
- Pagnamenta AT, Bacchelli E, De Jonge MV, et al. 2010. Characterization of a family with rare deletions in CNTNAP5 and DOCK4 Suggests novel risk loci for autism and dyslexia. *Biological Psychiatry*, **68**(4): 320–328.
- Pither J. 2003. Climate tolerance and interspecific variation in geographic range size. *Proceedings of the Royal Society B: Biological Sciences*, **270**(1514): 475–481.
- Puckett EE, Park J, Combs M, et al. 2016. Global population divergence and admixture of the brown rat (*Rattus norvegicus*). *Proceedings of the Royal Society B: Biological Sciences*, **283**(1841): 20161762.
- Purcell S, Neale B, Todd-Brown K, et al. 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. *The American Journal of Human Genetics*, **81**(3): 559–575.
- Rabiee MH, Mahmoudi A, Siahparvar R, et al. 2018. Rodent-borne diseases and their public health importance in Iran. *PLoS Neglected Tropical Diseases*, **12**(4): e0006256.
- Raharinosy V, Olive MM, Andriamiarimanana FM, et al. 2018. Geographical distribution and relative risk of Anjozorobe virus (Thailand orthohantavirus) infection in black rats (*Rattus rattus*) in Madagascar. *Virology Journal*, **15**(1): 83.
- Saito-Ito A, Takada N, Ishiguro F, et al. 2008. Detection of Kobe-type *Babesia microti* associated with Japanese human babesiosis in field rodents in central Taiwan and southeastern mainland China. *Parasitology*, **135**(6): 691–699.
- Sandel B, Arge L, Dalsgaard B, et al. 2011. The influence of late quaternary climate-change velocity on species endemism. *Science*, **334**(6056): 660–664.
- Schmieder R, Edwards R. 2011. Quality control and preprocessing of metagenomic datasets. *Bioinformatics*, **27**(6): 863–864.
- Sheldon M, Rice DS, D'Arcangelo G, et al. 1997. *Scrambler* and *yotari* disrupt the *disabled* gene and produce a *reeler*-like phenotype in mice. *Nature*, **389**(6652): 730–733.
- Slatyer RA, Hirst M, Sexton JP. 2013. Niche breadth predicts geographical range size: a general ecological pattern. *Ecology Letters*, **16**(8): 1104–1114.
- Staubach F, Lorenc A, Messer PW, et al. 2012. Genome patterns of selection and introgression of haplotypes in natural populations of the house mouse (*Mus musculus*). *PLoS Genetics*, **8**(8): e1002891.
- Stefansson H, Steinberg S, Petursson H, et al. 2009. Variant in the sequence of the *LINGO1* gene confers risk of essential tremor. *Nature Genetics*, **41**(3): 277–279.
- Suzuki TA, Phifer-Rixey M, Mack KL, et al. 2019. Host genetic determinants of the gut microbiota of wild mice. *Molecular Ecology*, **28**(13): 3197–3207.
- Teng HJ, Zhang YH, Shi CM, et al. 2017. Population genomics reveals speciation and introgression between brown Norway rats and their sibling species. *Molecular Biology and Evolution*, **34**(9): 2214–2228.
- Thomas CD, Bodsworth EJ, Wilson RJ, et al. 2001. Ecological and evolutionary processes at expanding range margins. *Nature*, **411**(6837): 577–581.
- Thomas O. 1908. The duke of Bedford's zoological exploration in Eastern Asia.- X. List of mammals from the provinces of Chih-li and Shan-si, N. China. *Proceedings of the Zoological Society of London*, **78**(3): 635–646.

- Wang H, Yoshimatsu K, Ebihara H, et al. 2000. Genetic diversity of hantaviruses isolated in China and characterization of novel hantaviruses isolated from *Niviventer confucianus* and *Rattus rattus*. *Virology*, **278**(2): 332–345.
- Wang YX. 2003. A Complete Checklist of Mammal Species and Subspecies in China: A Taxonomic and Geographic Reference. Beijing: China Forestry Publishing House, 41–46. (in Chinese)
- Xu B, Woodroffe A, Rodriguez-Murillo L, et al. 2009. Elucidating the genetic architecture of familial schizophrenia using rare copy number variant and linkage scans. *Proceedings of the National Academy of Sciences of the United States of America*, **106**(39): 16746–16751.
- Young HS, Dirzo R, Helgen KM, et al. 2014. Declines in large wildlife increase landscape-level prevalence of rodent-borne disease in Africa. *Proceedings of the National Academy of Sciences of the United States of America*, **111**(19): 7036–7041.
- Zeng L, Ming C, Li Y, et al. 2018. Out of Southern East Asia of the brown rat revealed by large-scale genome sequencing. *Molecular Biology and Evolution*, **35**(1): 149–158.
- Zhang B, He K, Wan T, et al. 2016. Multi-locus phylogeny using topotype specimens sheds light on the systematics of *Niviventer* (Rodentia, Muridae) in China. *BMC Evolutionary Biology*, **16**(1): 261.
- Zhang C, Dong SS, Xu JY, et al. 2019. PopLDdecay: a fast and effective tool for linkage disequilibrium decay analysis based on variant call format files. *Bioinformatics*, **35**(10): 1786–1788.
- Zou Y, Zhang WF, Liu HY, et al. 2017. Structure and function of the contactin-associated protein family in myelinated axons and their relationship with nerve diseases. *Neural Regeneration Research*, **12**(9): 1551–1558.