OXFORD

Article

Genetic boundary and gene flow between 2 parapatric subspecies of brown rats

Lei Zhao[a,b], Jian-Xu Zhang[a,b,*], and Yao-Hua Zhang[a,*]

[a]State Key Laboratory of Integrated Management of Pest Insects and Rodents in Agriculture, Institute of Zoology, Chinese Academy of Sciences, Beichen West Road 1-5, Chaoyang District, Beijing 100101, China and [b]CAS Center for Excellence in Biotic Interactions, University of Chinese Academy of Sciences, Beijing, 100049, China

*Address correspondence to Yao-Hua Zhang and Jian-Xu Zhang. E-mail: zhangyh@ioz.ac.cn and zhangjx@ioz.ac.cn.

Handling editor: Zhi-Yun Jia

Abstract

Two parapatric *Rattus norvegicus* subspecies, *R. n. humiliatus* (RNH) and *R. n. caraco* (RNC), are classified according to morphological divergence and are mainly distributed in North and Northeast China. Here, we aimed to explore the population genetic structure, genetic boundary, and gene flow in these rats using 16 microsatellite loci. Structure analysis and principal component analysis revealed 3 ancestral clusters. We found that the intermediate cluster exhibited higher genetic diversity and a lower inbreeding coefficient than the other 2 clusters. The genetic differentiation between the 3 clusters was significant but weak, with a higher F_{ST} value being observed between the clusters on both sides. The subspecies boundary inferred from microsatellite markers may indicate the existence of an admixture or hybridization area covering Liaoning, Inner Mongolia, and Jilin Provinces, rather than corresponding to the administrative provincial boundaries between Liaoning and Jilin. The RNH and RNC subspecies presented moderate gene exchange and an asymmetric bidirectional gene flow pattern, with higher gene flow from the RNH subspecies to the RNC subspecies, constraining speciation. Such genetic characteristics might be explained by biological processes such as dispersal ability, mate choice, and dynamic lineage boundaries.

Key words: genetic boundary, gene flow, population genetic structure, *Rattus norvegicus*

The brown rat (*Rattus norvegicus*) is the most common global pest of mankind and the most successful commensal mammal on the planet and was the earliest model animal to be domesticated for biomedical research (Musser and Carleton 2005). In China, brown rats are distributed across all provinces except Tibet. Based on distinguishable morphological differences (e.g., in body size and hair color), the brown rat is mainly classified into 4 geographical subspecies: *R. n. norvegicus*, *R. n. soccer*, *R. n. humiliatus* (RNH), and *R. n. caraco* (RNC; Wu 1982; Wang 2003; Musser and Carleton 2005). In accordance with morphological taxonomy, *R. n. norvegicus* is distributed in southeastern coastal areas and adjacent islands, such as Guangdong Province, Fujian Province, and Hainan Island, and has the largest body size among the 4 rat subspecies. *R. n. soccer*, which exhibits a smaller body size, inhabits the southern region of the Huaihe River basin and the western region of Taihang Mountain; this subspecies exhibits the widest distribution of the 4 rat subspecies and extends throughout most of Northwest, Southwest, and Central China. RNH possesses the smallest body size and ranges from the North China Plains across the Great Wall and the Yan Mountains to the southern NE China Plains (Liaoning Province; Wu 1982; Wang 2003). As a parapatric subspecies of RNH, RNC is found on the northern NE China Plains (Jilin Province and Heilongjiang Province) and exhibits asmaller body size than *R. n. soccer* and dark brown fur (Wu 1982).

In the field of population ecology, population genetic structure studies have contributed to estimating the spatial connectivity and

temporal stability of populations or species (Abdelkrim et al. 2005; Brouat et al. 2007). Population genetic structures have been extensively investigated in rodents; for example, genetic structure analysis revealed that the house mouse subspecies include *Mus musculus domesticus* (Mmd), *M. m. musculus* (Mmm), and *M. m. castaneus*, representing 3 evolutionary units (Schulte-Hostedde et al. 2001; Duvaux et al. 2011; Varudkar and Ramakrishnan 2015). In brown rats, global population genetic structures have also been analyzed on large geographical scales (Ness et al. 2012; Deinum et al. 2015; Puckett 2016). Comparative genome analysis of global rat populations provided insights into the migration and geographical origin of wild brown rats (Zeng et al. 2017; Puckett and Munshi-South 2019). Teng et al. (2017) used whole-genome SNP data to estimate the genetic divergence among wild rat populations across almost all of China and compared the genomic relationship with *R. nitidus*, a sister species (Teng et al. 2017). Emerging studies have evaluated the complex genetic structures and gene flow patterns associated with heterogeneous urban landscapes at a fine spatial scale in brown rats (Kajdacsi et al. 2013; Ortiz et al. 2017; Combs et al. 2018). However, we still lack investigatory studies of fine-scale geographical genetic structure across populations of brown rats in nature.

Microsatellite loci are a type of nucleotide marker that is used as a valuable and efficient molecular tool for estimating the dispersal behavior, genetic structure, and gene flow of pest populations (Abdelkrim et al. 2005; Gilabert et al. 2007; Russell et al. 2009; Abdelkrim et al. 2010; Kajdacsi et al. 2013). During differentiation, the inference and estimation of gene flow can convey important information about isolation and speciation (Duvaux et al. 2011). Mmd and Mmm, 2 parapatric subspecies of mice in Europe, were shown to form a hybrid zone after secondary contact and to exhibit asymmetric gene flow (significant gene flow from Mmd into Mmm but no gene flow into Mmd) and intermediate levels of reproductive isolation by using autosomal markers (Christophe and Baudoin 1998; Smadja and Ganem 2002; Geraldes et al. 2008; Teeter et al. 2008; Duvaux et al. 2011). In the current work, 16 autosomal microsatellite markers that were screened and identified in brown rats were used to investigate the fine-scale genetic substructure between 2 parapatric rat subspecies, RNH and RNC, and to infer their genetic diversity, genetic differentiation, genetic boundary, gene flow pattern and possible hybridization.

Materials and Methods

Sampling
The distribution ranges of the 4 rat subspecies are presented in Figure 1. For population genetic analysis, 64 wild brown rats from 24 sites (Figure 2) were sampled by using snap traps in settlements and farms on both sides of highways. All rats were trapped alongside approximated expressway lines with parallel railway lines from Harbin City in Heilongjiang Province southwestward to the southern suburb of the Beijing Municipality, which resulted from their commensal features (migrated mainly by virtue of human transportation lines). The sampling range passed through large cities such as Changchun and Siping in Jilin Province, Tieling, Jinzhou, and Huludao in Liaoning Province, Qinhuangdao, and Tangshan in Hebei Province and the Tianjin Municipality. Detailed sampling information is listed in Table 1.

Genotyping
The phenol–chloroform method was employed to isolate genomic DNA (Liu et al. 2011). Lysis buffer 400 μL (20 mM Tris-HCl, 5 mM EDTA, 400 mM NaCl, 1% SDS [m/V], 200 μg/mL proteinase K, pH = 8.0) was mixed with each DNA sample (0.5 cm tail tissue) in a 1.5 mL centrifuge tube and kept in a 55°C water bath overnight.

Highly purified DNA was extracted via a series of steps including digestion, decontamination, washing, precipitation, drying, and dissolution and then stored at −20°C until use. OD_{260}/OD_{280} values were measured to determine the concentration and purity of the DNA in a NanoDrop spectrophotometer (Gene Company Limited, Hongkong, China). Based on high conservation, polymorphism and repeatability, 16 microsatellites were screened, and each forward primer of satisfactory loci was labeled with the HEX, FAM, or TAMRA dye.

All loci but one were amplified using a 20 μL Polymerase Chain Reaction (PCR) system including 2 mM $MgCl_2$ (25 mM), eachdNTP at 0.15 mM (2.5 mM each), 2 μL of 10× PCR Buffer, each of the forward and reverse primers at 0.4 μM (10 μM), 100 ng genomic DNA, 0.8 U Takara Taq (5 U/μL, Takara, Dalian, China), and sterile water to the final total volume. The D1Wox31 loci was amplified with a 50 μL reaction system (34.75 μL sterile water, 3 μL $MgCl_2$, 4 μL of each dNTP, 5 μL of 10 × PCR Buffer, 1 μL of each forward primer and reverse primer, 1 μL DNA [100 ng/μL] and 0.25 μL Takara Taq). The PCR conditions were as follows: 94°C for 5 min; 35 cycles of 94°C for 30 s, Tm (Appendix Table A1) for 40 s and 72°C for 1 min; and a final step at 72°C for 10 min. The PCR products for all microsatellites were subjected to capillary electrophoresis on anABI 3130 Genetic instrument (Applied Biosystems, Life Technologies, USA). Genotyping was carried out repeatedly at least 3 times (i.e., the error rate of genotyping was <2%) according to GeneMarker software (Holland and Parson 2011).

Population genetic structure
Prior to any analyses, all satisfactory loci were tested for Hardy–Weinberg equilibrium (HWE) and linkage disequilibrium (LD) in Arlequin 3.5 software (Excoffier and Lischer 2010). For the exact Hardy–Weinberg test, we used a Markov chain with 10,000 forecast chain lengths and 10,000 dememorization steps. In pairwise linkage disequilibrium analyses, we performed a permutations test using theExpectation Maximization (EM) algorithm with 10,000 permutations and set a significance level of 0.01. Using the Bayesian approach, we inferred the assignment of *R. norvegicus* and the pattern of population genetic structure in Structure 2.3 software (Pritchard et al. 2000; Evanno et al. 2005). For K (from 1 to 5) estimation and runs, we set 5 iterations for each K with a burn-in period of 100,000 iterations and 1,000,000 MCMC reps after burn-in. Principal component analysis (PCA) was also used to assess population division, which counted aggregate variables among all brown rats according to the allele frequency datasets.

Genetic diversity
The genetic statistics for all loci including allele numbers (Na), observed heterozygosity (Ho), expected heterozygosity (He), and the polymorphic information content (PIC) were calculated using Cervus 3.0 software (Kalinowski et al. 2007). Fstat 2.9.3 (Goudet 1995) was employed to calculate gene diversity (Hs), inbreeding coefficient (FIS), and allelic richness (An). Private allelic richness (Ap) was also tested in the Hp-Rare 1.0 program (Kalinowski 2004; Kalinowski 2005). To assess genetic differentiation among sampling sites, Arlequin 3.5 was used for pairwise F_{ST} calculation with 10,000 permutations and a significance level of 0.05. Furthermore, the genetic differentiation pattern was estimated by the isolation-by-distance (IBD) method with the Mantel test in the Arlequin 3.5 and Genepop 4.6 (Rousset 2008) programs. That is, we examined the correlation between Euclidean distance and pairwise F_{ST} matrices with 10,000 permutations number.
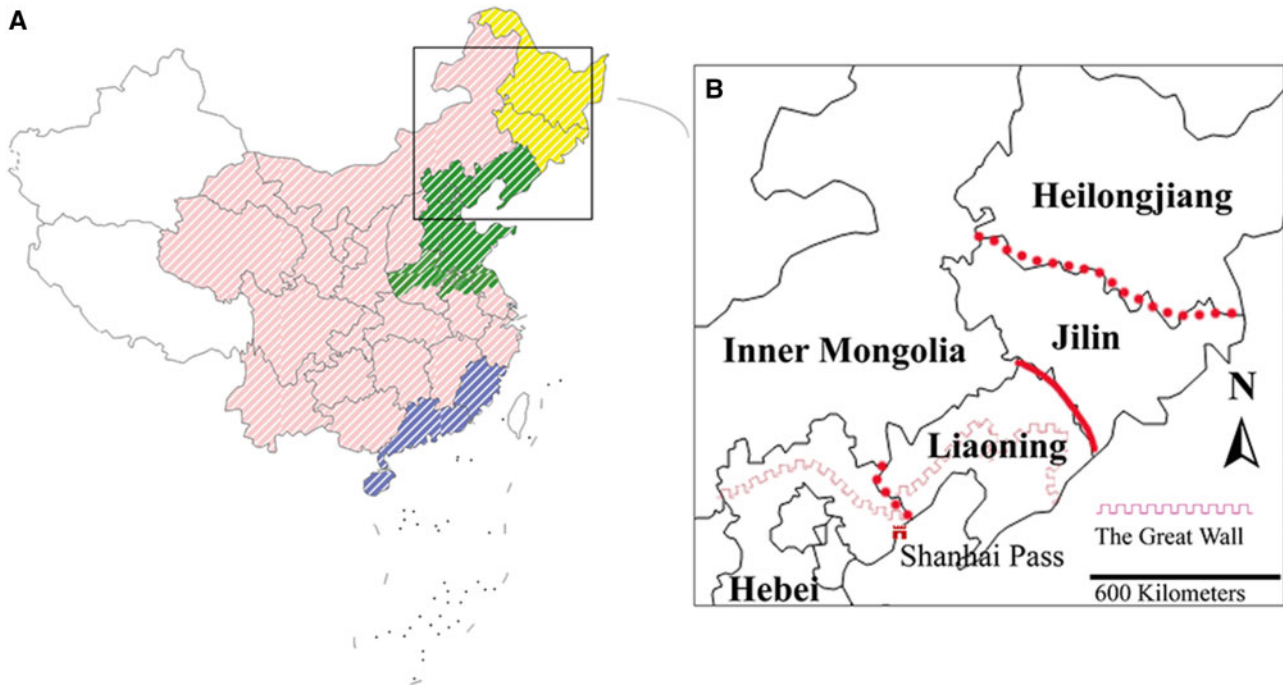
Figure 1. The distribution ranges of 4 *R. norvegicus* subspecies in China. Data from Wu (1982) and Wang (2003). (**A**) The ranges of *R. norvegicus* subspecies are indicated by hatching. Purple: *R. n. norvegicus*; pink: *R. n. soccer*; green: RNH, yellow: RNC. (**B**) The sampling ranges of the RNH and RNC subspecies in North China and Northeast China covering Beijing City, Tianjin City, Hebei, Liaoning, Inner Mongolia, Jilin, and Heilongjiang Provinces. The red solid line represents the morphological boundary, and the red dotted line represents the genetic boundary. The base maps are from Standard Map Service website (http://bzdt.ch.mnr.gov.cn/index.html).

Hybridization, migration, and gene flow

NewHybrids 1.1 (Anderson and Thompson 2002) was used to assess the categories of pure RNH, backcrosses with RNH, F1 hybrids, F2 hybrids, pure RNC, and backcrosses with RNC based on the assigned posterior probability of each *R. norvegicus*. For the MCMC chain run, the settings included 100,000 burn-in sweeps and 1,000,000 total sweeps. First-generation migrants were assessed using theUSEPOPINFO option in Structure 2.3. We calculated Q to quantify the proportion of ancestry for each *R. norvegicus* from RNH or RNC with a migration rate of 0.05. A brown rat was identified as a pure migrant if Qo (its capture site) < Qa (its assignment site). Lh and Lh/Lmax in the Geneclass 2.0 (Paetkau et al. 2004) program were also used here to detect first-generation migrants. We computed the likelihood of Lh and Lh/Lmax by using the options of simulated individuals ($N = 1{,}000$) and Type I error ($\alpha = 0.01$). For gene flow simulations, a basic model of isolation and migration implemented in theIsolation with Migration (IM) program (Hey and Nielsen 2004; Hey 2005) was applied to generate marginal posterior probability density estimates for each demographic parameter. We mainly assessed 6 parameters: θ_H (theta for the RNH subspecies), θ_C (theta for the RNC subspecies), θ_A (theta for the ancestral species), t (time of population splitting), m_{CH} (migration rate from RNC to RNH), and m_{HC} (migration rate from RNH to RNC), using a mutation model of SSM. For the IM analyses, we used 3 different sample sets to estimate the migration rate to guarantee reliable results. The sample sets of 10 individuals (5 from the southwestern distribution of putative RNH and 5 from the northeastern distribution of putative RNC), 41 individuals (20 from the southwestern distribution of putative RNH and 21 from the northeastern distribution of putative RNC), and 64 individuals were considered as 3 representatives for the calculation of gene flow between the putative RNH and RNC

subspecies. Initially, we carried out multiple runs with different random numbers of seeds and long enough runs to ensure mixing and a lowest ESS of not <50.

Isolation by resistance

A causal modeling method was used to estimate the influence of landscape features on gene flow (Cushman et al. 2006; Cushman et al. 2013). Mountain landscapes (i.e., the Yan Mountains and the Great Wall) were hypothesized to have potentially important effects on *R. norvegicus*, so we modeled landscape resistance as a function of the elevation variable. The landscape resistance value of this variable was assumed and assigned to grid cells with a range of 1 (least resistance to gene flow) to 100 (maximal resistance to gene flow). The generated resistance maps were transformed into resistance distance matrices using 4-neighbor and 8-neighbor connection schemes in the Circuitscape 4.0 program (Mcrae et al. 2013). Finally, simple Mantel tests and partial Mantel tests between genetic distance and resistance distance were performed in the zt program (Bonnet and Van De Peer 2002).

Results

Pattern of population genetic structure of the RNH and RNC subspecies

A total of 64 rats from 24 sampling sites were successfully genotyped using 16 microsatellite loci. In the HWE test, all these loci showed no significant deviation from HWE at the sampling sites. In addition, between all pairs of loci, there was no significant linkage disequilibrium. This confirmed the availability of our chosen microsatellite loci, for which the information and references (Steen 1999; Heiberg 2006; Bryda 2008) are listed in detail in Appendix Table A1. Bayesian clustering analyses determined the ancestry of each
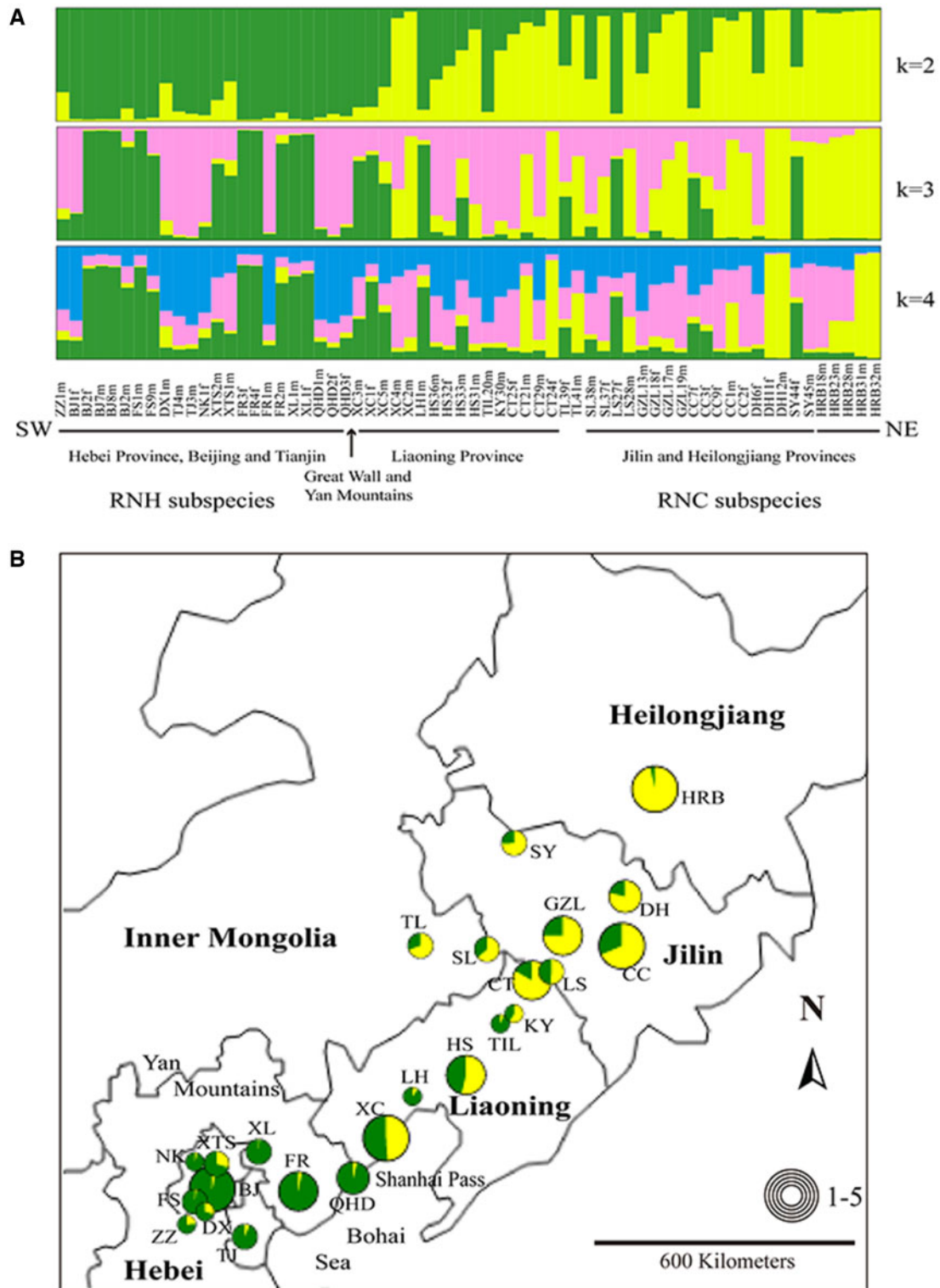
Figure 2. Sampling sites and population genetic structure of RNH and RNC across geographical gradients from the Beijing Municipality to Harbin City were determined. (**A**) Bayesian clustering analyses identified 2, 3, or 4 ancestral populations of *R. norvegicus*. Each vertical bar represents an *R. norvegicus* individual. (**B**) Sampling included 24 sites in Zhuozhou (ZZ), Beijing (BJ), Fangshan (FS), Daxing (DX), Tianjin (TJ), Nankou (NK), Xiaotangshan (XTS), Fengrun (FR), Xinglong (XL), Qinhuangdao (QHD), Xingcheng (XC), Linghai (LH), Heishan (HS), Tieling (TIL), Kaiyuan (KY), Changtu (CT), Tongliao (TL), Shuangliao (SL), Lishu (LS), Gongzhuling (GZL), Changchun (CC), Dehui (DH), Songyuan (SY), and Harbin (HRB) and 64 individuals. Green and yellow represent 2 different clusters. The sizes of the circles (from small to large) represent the sampling size (from 1 to 5) at each sampling site. The base map is from Standard Map Service website (http://bzdt.ch.mnr.gov.cn/index.html).

Table 1. The *R. norvegicus* individuals from each sampling site are listed

Site	N	East longitude	Northern latitude	Altitude (m)	Area
ZZ	1	115.98	39.43	40	Zhuozhou, Hebei
BJ	5	116.41	39.91	60	Beijing
FS	2	116.15	39.75	50	Fangshan, Beijing
DX	1	116.25	39.68	38	Daxing, Beijing
TJ	2	117.21	39.09	4	Tianjin
NK	1	116.15	40.23	93	Changping, Beijing
XTS	2	116.48	40.16	39	Changping, Beijing
FR	4	118.20	39.83	51	Tangshan, Hebei
XL	2	117.51	40.42	588	Chengde, Hebei
QHD	3	119.63	40.04	149	Qinhuangdao, Hebei
XC	5	120.42	40.37	31	Huludao, Liaoning
LH	1	121.26	41.13	118	Jinzhou, Liaoning
HS	4	122.20	41.56	12	Jinzhou, Liaoning
TIL	1	123.85	42.30	72	Tieling, Liaoning
KY	1	124.05	42.48	88	Tieling, Liaoning
CT	4	124.25	42.99	147	Tieling, Liaoning
TL	2	122.29	43.63	177	Tongliao, Inner Mongolia
SL	2	123.49	43.52	117	Siping, Jilin
LS	2	124.61	43.10	369	Siping, Jilin
GZL	4	124.89	43.55	226	Siping, Jilin
CC	5	125.55	43.87	269	Changchun, Jilin
DH	3	125.71	44.43	172	Changchun, Jilin
SY	2	124.39	45.26	126	Songyuan, Jilin
HRB	5	126.54	45.81	119	Harbin, Heilongjiang

The number of samples at each site is expressed as N. Specific geographical coordinates where the samples were collected, including the east longitude and Northern latitude and altitude (m) are shown. Sampling cities and provinces are provided in the last column.

brown rat from the ancestral populations defined by K. When $K = 2$, all sampled brown rats were divided into 2 ancestral populations. One cluster, colored green, included 27 individuals, which represented rat populations from the Beijing Municipality (DX1m and XTS1m were excluded owing to a membership proportion of 0.65~0.80), the Tianjin Municipality, Hebei Province, Liaoning Province (4 individuals), and Jilin Province (2 individuals). The other cluster, colored yellow, included rats from the Harbin population in Heilongjiang Province, Jilin Province (11 individuals), Inner Mongolia (1 individual), and Liaoning Province (7 individuals) (Figure 2A). The remaining 13 rats from Beijing (2 individuals), Liaoning Province (5 individuals), Inner Mongolia (1 individual), and Jilin Province (5 individuals) might have been derived from these 2 ancestral clusters. When $K = 3$, the green cluster grouped 12 individuals from the populations of Beijing (5 rats), Hebei (5 rats), and Liaoning (2 rats). The yellow cluster grouped 2 rats from Liaoning and 7 rats from the Jilin and Harbin populations (5 rats). The pink cluster (i.e., the intermediate cluster) grouped 18 rats from Beijing (3 individuals), Tianjin (2 individuals), Hebei (4 individuals), Liaoning (6 individuals), and Jilin (3 individuals). The remaining individuals were composed of 2 or 3 ancestral clusters. When $K = 4$, 11 rats belonged to the green cluster, and 5 rats belonged to the yellow cluster. No individuals exhibited a membership proportion ≥ 0.80 related to belonging to the pink cluster or the blue cluster. The Evanno method in Structure Harvester determined $K = 3$ to be the optimal K-value. In northern and northeastern China, *R. norvegicus* was considered to be divided into 3 ancestral populations.

In the genetic structure analyses, we also assessed the membership proportion of each sampling site of the RNH and RNC subspecies (Figure 2B). Along the geographical gradients, we found a steep transition southwest of Liaoning Province, where *R. norvegicus* individuals showed an approximately equal probability of belonging
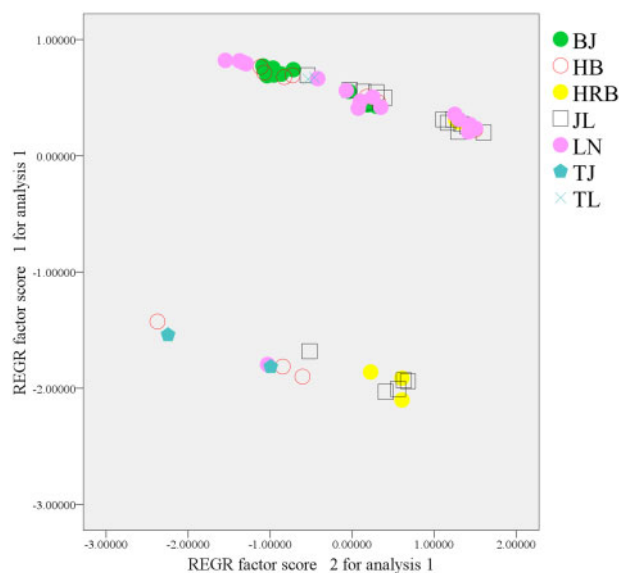

Figure 3. PCA of *R. norvegicus* from the Beijing Municipality to Harbin City was performed based on allele frequencies. Each *R. norvegicus* is represented by a dot, and 7 populations (BJ, Beijing; HB, Hebei; HRB, Harbin; JL, Jilin, LN, Liaoning; TJ, Tianjin; TL, Tongliao) are shown.

to the RNH subspecies or RNC subspecies. In dimension reduction analyses for the RNH and RNC subspecies, we extracted 2 principal components, which accounted for total variance of 45.85% (Figure 3). The Beijing population was grouped together, and individuals from Tianjin and the Harbin population were grouped together, respectively. In the Hebei population, FR1m, QHD1m, and QHD2f were clustered with the Tianjin rats, FR4f was clustered

Table 2. Genetic variations within each of the 3 clusters were examined

Parameter	Green cluster (mean $\pm$ SD)	Pink cluster (mean $\pm$ SD)	Yellow cluster (mean $\pm$ SD)
Sample size (n)	12	18	14
Ho	0.567 ± 0.251	0.677 ± 0.199	0.575 ± 0.207
He	0.745 ± 0.213	0.800 ± 0.177	0.738 ± 0.177
Mean PIC	0.683 ± 0.201	0.755 ± 0.147	0.676 ± 0.180
Hs	0.754 ± 0.217	0.804 ± 0.150	0.745 ± 0.180
Mean numbers of alleles (Na)	6.313 ± 2.182	8.750 ± 3.194	6.250 ± 2.049
An	6.203 ± 2.121	7.508 ± 2.394	5.898 ± 1.896
Ap	1.162 ± 0.830	1.596 ± 1.140	0.987 ± 0.997
FIS	0.237 ± 0.276	0.145 ± 0.211	0.227 ± 0.215

The statistical analyses of sample size (n) are presented, followed by the Ho, He, mean PIC, Hs, mean numbers of alleles (Na), An, Ap, and FIS in order.

Table 3. The migration and hybridization of *R. norvegicus* individuals were predicted via methods performed in the geneclass, structure, and NewHybrids programs

Sample	Sex	Age	Lh (P)	Lh/Lmax (P)	Qo	Qa	Pure green	Pure pink	Pure yellow	Hybrid or Backcross
ZZ1m	M	Adult		0.002			0.000		0.001	0.993
BJ1f	F	Adult	0.005							
BJ2m	M	Adult					0.444		0.000	0.533
FS9m	M	Adult					0.081	0.547		0.169
DX1m	M	Adult	**0.003**	**0.002**	0.456	0.461				
TJ4m	M	Adult	**0.000**		0.165	0.693				
TJ3m	M	Adult	**0.000**		0.129	0.731				
NK1f	F	Adult	**0.004**	**0.002**						
XTS2m	M	Adult					0.000		0.005	0.993
XTS1m	M	Adult					0.000		0.027	0.966
XC3m	M	Adult		0.008			0.070		0.000	0.928
XC1f	F	Adult					0.729		0.000	0.258
XC5m	M	Adult						0.712	0.002	0.257
XC4m	M	Adult						0.109	0.002	0.797
HS32f	F	Adult						0.690	0.000	0.261
HS33m	M	Subadult						0.570	0.000	0.399
HS31m	M	Adult						0.015	0.004	0.919
CT21m	M	Adult					0.000		0.236	0.731
TL39f	F	Adult	**0.000**	**0.000**			0.000		0.004	0.986
TL41m	M	Adult	**0.000**					0.074	0.249	0.628
SL37f	F	Adult		0.003				0.361	0.002	0.563
LS27f	F	Adult					0.003		0.000	0.988
LS28m	M	Adult						0.000	0.523	0.434
GZL18f	F	Adult						0.178	0.000	0.765
GZL17m	M	Adult						0.000	0.032	0.868
CC7f	F	Adult					0.021		0.000	0.964
CC3f	F	Adult						0.198	0.000	0.729
CC9f	F	Adult						0.005	0.000	0.983
CC2f	F	Adult						0.000	0.662	0.298
SY44f	F	Adult					0.006		0.000	0.979
HRB23m	M	Adult		0.002						

Sample names, sex, age, Lh (P), Lh/Lmax (P), Qo, Qa, pure green (rats only from the green cluster), pure pink (rats only from the pink cluster), pure yellow (rats only from the yellow cluster), and hybrids or backcrosses are listed in order. First-generation migrants detected by 2 or more methods are indicated in bold.

with the Harbin population, and the remaining 6 individuals (ZZ1m, FR3f, FR2m, XL1m, XL1f, and QHD3f) were clustered with the Beijing population. In the Jilin population, SL38m, LS27f, CC7f, DH11f, DH12m, and SY44f were clustered with Beijing population, and the remaining 12 individuals (SL37f, LS28m, GZL13m, GZL18m, GZL17m, GZL19m, CC3f, CC9f, CC1m, CC2f, DH6f, and SY45m) were clustered with the Harbin population. The individuals from Inner Mongolia and the Beijing population were grouped together. For the Liaoning population, HS32f was clustered with the Tianjin individuals; XC3m, HS36m, HS31m, CT29m, and CT24f were clustered with the Harbin population; and

XC1f, XC5m, XC4m, XC2m, LH1m, HS33m, TIL20m, KY30m, CT25f, and CT21m were clustered with the Beijing population.

Genetic variation and differentiation

Within each of the 3 clusters, we detected the levels and patterns of genetic variation, including Ho, He, PIC, Na, Hs, An, Ap, and FIS (Table 2). Based on our analyses, the pink cluster presented higher levels of Ho, He, PIC, Na, Hs, An, and Ap and lower FIS compared with both the green cluster and the yellow cluster. The Na in the pink cluster were significantly higher than those in the

Table 4. Gene flow between RNH and RNC was evaluated

	θ_H	θ_C	θ_A	m_{CH}	m_{HC}	$2N_H m_{CH}$	$2N_C m_{HC}$
Microsatellite dataset from 10 individuals							
HiPt	0.467	0.389	19.726	0.005	0.725	0.001	0.141
HPD90Lo	0.094	0.056	10.190	0.005	0.005	—	—
HPD90Hi	9.255	2.388	105.547	5.675	7.695	—	—
Microsatellite dataset from 41 individuals							
HiPt	8.594	7.552	21.535	0.005	0.155	0.021	0.585
HPD90Lo	3.741	2.617	12.031	0.005	0.005	—	—
HPD90Hi	13.851	11.740	74.716	1.445	1.455	—	—
Microsatellite dataset from total of 64 individuals							
HiPt	7.223	10.291	20.899	0.005	0.075	0.018	0.386
HPD90Lo	2.793	3.692	12.424	0.005	0.005	—	—
HPD90Hi	12.039	15.476	36.501	2.165	1.335	—	—

Three different sample sets represent different *R. norvegicus* numbers from RNH and RNC. θ_H, θ_C, and θ_A represent the theta values for RNH, RNC, and the common ancestor, respectively. m_{CH} expresses the rate per gene per generation from population RNC to RNH, and m_{HC} expresses the same rate in the opposite direction. $2N_H m_{CH}$ expresses gene flow from RNC to RNH, and $2N_C m_{HC}$ expresses gene flow from RNH to RNC. HiPt represents the greatest residence time, HPD90Lo represents the lowest value of the 90% highest posterior density interval, and HPD90Hi represents the highest value of the 90% highest posterior density interval.

green cluster ($P = 0.017$) and the yellow cluster ($P = 0.013$). For comparisons between sampling sites, the pairwise F_{ST} values ranged from 0.01 to 0.44 (Appendix Table A2). There were 29 pairwise F_{ST} values (BJ-FR, BJ-QHD, BJ-XC, BJ-HS, BJ-CT, BJ-GZL, BJ-DH, BJ-HRB, TJ-CC, XTS-CC, XTS-HRB, RF-XC, FR-HS, FR-CT, FR-GZL, FR-CC, FR-HRB, XL-CC, QHD-XC, QHD-HS, QHD-CT, QHD-HRB, XC-CT, XC-DH, XC-HRB, HS-CT, HS-HRB, CT-HRB, and TL-CC) showing significant differences ($P < 0.05$), 5 of which (BJ-CT, BJ-HRB, FR-CC, FR-HRB, and XC-HRB) showed extremely significant differences ($P < 0.01$). For the F_{ST} comparison among the 3 clusters, the pairwise F_{ST} values of the green cluster versus the pink cluster (0.046), the pink cluster versus the yellow cluster (0.045), and the green cluster versus the yellow cluster (0.095) reached a significant level ($P < 0.001$). Then, we calculated IBD among 24 rat populations from North and NE China, resulting in no statistical correlation between genetic distance and geographical distance along contiguous sampling sites from the Beijing Municipality to Harbin City according to either Arlequin ($P = 0.062$, $r = 0.108$) or Genepop ($P = 0.065$, $r = 0.081$).

Migration, hybridization, and gene flow estimates

More than two-thirds of the *R. norvegicus* individuals (44/64) presented a probability of being assigned to the 3 clusters ≥0.80. Migration results were calculated with the 2 programs (Table 3). In the Structure program, the membership coefficient for Qo and Qa identified DX1m, TJ4m, and TJ3m as migrants. In the Geneclass program, the L*h* method detected 7 migrants, while the L*h*/L*max* method again detected DX1m, NK1f, and TL39f plus 4 additional *R. norvegicus* migrants (Table 3). In the NewHybrids analyses of hybridization, each *R. norvegicus* individual was assigned a posterior probability. Twenty-five individuals exhibited a posterior probability of belonging to the hybrid or backcross category ≥ 0.169 (Table 3).

We obtained reliable gene flow estimates (Table 4) based on smooth single peak and convergent posterior density curves. Based on the autosomal dataset, gene flow was calculated based on 5 individuals of each RNH subspecies and RNC subspecies. The

migration rate estimated from RNC to RNH was close to negligible (0.001), while the estimated migration rate from RNH to RNC was 0.141. According to the calculations for 20 RNH individuals and 21 RNC individuals, the estimated gene flow from RNH to RNC was 0.585, whereas much lower gene flow (0.021) was observed in the opposite direction. In addition, based on all *R. norvegicus* individuals, we estimated apparent gene flow (0.386) from RNH to RNC and near-zero gene flow (0.018) from the RNC to RNH subspecies.

Isolation by resistance at the landscape scale

In the causal modeling results of the mountain landscapes, there was no significant association between the genetic and resistance distance matrices. When connecting raster cells to 4 neighbors, the simple Mantel test between genetic distance and resistance distance was not significant ($P = 0.251$, $r = -0.111$). The partial Mantel test between genetic distance and resistance distance when controlling for the effect of geographical distance was not significant ($P = 0.177$, $r = -0.150$). When connecting raster cells to 8 neighbors, the association between genetic distance and resistance distance in the simple Mantel test was not significant ($P = 0.243$, $r = -0.121$). The association between genetic distance and resistance distance according to the partial Mantel test when partialling out the effect of geographical distance was not significant ($P = 0.173$, $r = -0.160$).

Discussion

In this study, the 2 examined parapatric subspecies of brown rats exhibited a rough, complex population genetic structure. Clustering analyses identified 3 clusters of wild brown rats in North China and Northeast China based on 16 microsatellite markers. In detail, the green cluster included rats from Beijing, Hebei, and Liaoning; the yellow cluster included rats from Harbin, Jilin, and Liaoning; and the pink cluster included rats with a larger geographical range including Beijing, Tianjin, Hebei, Liaoning, and Jilin. Although these clusters with membership proportions for each rat of ≥0.80 were genetically separated from each other, the pink cluster did not present clear geographical separation from the green and yellow clusters. The PCA results showed that rats from Beijing, Tianjin, Hebei, and Inner Mongolia were grouped together, and the Harbin rats were grouped together. Interestingly, the rats from Liaoning Province and Jilin Province seemed to be a mixture of the Beijing population and the Harbin population. For example, CT25f and CT21m clustered with the Beijing population, whereas CT29m and CT24f clustered with the Harbin population. Similar cases occurred at other sampling sites (XC, HS, SL, LS, CC, DH, and SY). Wu (1982) classified wild brown rats in North China and Liaoning as putative RNH subspecies and those in Jilin and Heilongjiang as putative RNC subspecies according to their morphological differences (Wu 1982; Wang 2003). However, our current results regarding population genetic structure did not completely agree with the morphological classification of the RNH and RNC subspecies. Recent genome-wide SNP markers revealed distinct genetic divergence between the Harbin population (RNC core population) and the Beijing population (RNH core population) of brown rats, consistent with our current results in these 2 rat populations (Teng et al. 2017). We included more rat samples between the 2 core populations to show the incomplete genetic structure of the RNH and RNC subspecies, rather than the classification previously described by Wu (1982). Such disagreement might be partially explained by population genetic structure formed at fine spatial scales and a lack of isolation between rat populations (Ortiz et al.

2017; Combs et al. 2018). Thus, a rough, complex, incomplete population genetic structure of *R. norvegicus* has formed in the North and NE China.

Within the population genetic structure of the 2 parapatric rat subspecies, we found relatively high genetic diversity within each of the 3 clusters (Table 2). The genetic diversity levels observed here were compared with results reported for urban *R. norvegicus* individuals (He ranged from 0.67 to 0.78; Gardner-Santana et al. 2009). The levels of genetic diversity were also comparable with those obtained for *R. norvegicus* according to fine-scale genetic structures (Abdelkrim et al. 2010; Kajdacsi et al. 2013). Moreover, higher genetic diversity and a lower FIS were observed in the pink cluster than in the other 2 clusters. These results agreed with the extensive composition of rats in the pink cluster. We also detected significant and relatively weak genetic differentiation among either the sampling sites or the 3 clusters. The differentiation level between the green cluster and the yellow cluster was significantly higher than between the other 2 pairs (the green cluster versus the pink cluster and the pink cluster versus the yellow cluster). In the mountain landscape, commensal *R. rattus* showed lower genetic differentiation than non-commensal *R. satarae* (Varudkar and Ramakrishnan 2015). Together with the F_{ST} results obtained in this study, these results suggested a genetic differentiation pattern of commensal species at fine spatial scales. Isolation by distance inferred by 2 methods suggested no spatial autocorrelation between the geographical matrix and the genetic matrix. This pattern received support from significant genetic differentiation between rat populations located close to each other geographically (BJ-FR, RF-XC, QHD-XC, and HS-CT, etc.) and paralleled the results of the clustering analyses (Kajdacsi et al. 2013). Therefore, the results revealed high genetic diversity, low genetic differentiation, and no IBD pattern.

Along the geographical gradient in which the 2 parapatric subspecies were distributed, approximately one-third of the brown rats exhibited genetic admixture in addition to a complex population genetic structure. Among all admixed rats derived from 2 or 3 clusters, the majority (16/20) came from Liaoning and Jilin Provinces. Moreover, the admixture did not show a clinal pattern. In Liaoning Province and Jilin Province, the admixed rats that were geographically closer to the green cluster did not show a higher membership proportion in the green cluster. For the admixed rats of Liaoning and Jilin Provinces, the closer geographical distance to the yellow cluster and the higher membership proportion of this cluster were inconsistent with each other. This admixture could be related to behavioral traits (Brouat et al. 2007). First, the admixture was a result of hybridization among different clusters or their ancestors. Under laboratory conditions, captured RNH and RNC could interbreed and give birth to healthy hybrid offspring (Zhang J-X et al., unpublished data). We identified XTS2m, TL41m, and CC9f as hybrids because the 2 alleles of these rats came from the private alleles of 2 clusters. Additionally, these admixed rats were demonstrated to be hybrid individuals in our NewHybrids analyses, which revealed an absence of reproductive isolation between the 2 putative subspecies of *R. norvegicus*. Another possible reason might be that brown rats exhibit relatively strong dispersal abilities along transportation lines for human activities due to their commensal characteristics (Davis 1951; Taylor 1978; Taylor et al. 1982; Musser and Carleton 2005; Kajdacsi et al. 2013; Varudkar and Ramakrishnan 2015). The dispersal distances of the rats range from 260 to 2000 m in rural areas (Calhoun 1963; Macdonald and Fenn 1995). Here, DX1m, TJ3m, TJ4m, NK1f, and TL39f are presented as evidence of migrants.

Therefore, there seems to be extensive admixture in Liaoning Province and Jilin Province in North China and NE China.

As a result of admixture, a broader boundary has developed between the 2 parapatric rat subspecies. This geographical boundary seems to be an admixture/hybridization area, which is very different from the previous boundary of these rat subspecies (Wu 1982; Wang 2003). More than 30 years ago, Wu separated the putative RNH subspecies from the putative RNC subspecies at the boundary of Liaoning Province and Jilin Province based on their morphological differences (Wu 1982). The boundary of these 2 rat subspecies might be dynamic and shift over time due to the lack of reproductive isolation and asymmetric migration. On the contrary, morphological divergence does not always agree with the genetic divergence among populations or species (Gonçalves and Oliveira 2004; Hamidy et al. 2011). Therefore, the shifted boundary between the putative RNH subspecies and the putative RNC subspecies might be partially attributed to the discordance between genetic divergence and morphological divergence as a simple reflection of asymmetric gene flow and recent splits (Wood et al. 2014). Therefore, based on microsatellite data, we have discovered a broader genetic boundary that covers Liaoning Province, Inner Mongolia, and Jilin Province and might represent an admixture/hybridization zone between the putative RNH and RNC subspecies.

Between the putative RNH and RNC subspecies, a bidirectional, asymmetric gene flow pattern, and a moderate gene flow level were detected. Such bidirectional gene flow is in accord with the relatively low pairwise F_{ST} value and broad admixture observed in the population genetic structure (Crispo et al. 2011; Varudkar and Ramakrishnan 2015; Lagerholm et al. 2017). *Ceroxylon echinulatum* dispersal is believed to be more constrained in complex mountain ranges (Trenel et al. 2008). It has been demonstrated that the Atlas Mountains rather than the strait of Gibraltar are a significant biogeographic barrier to gene flow in *Mauremys leprosa* (Fritz et al. 2005).

However, human activities could directly or indirectly change contact patterns and have the potential to influence gene flow between populations (Crispo et al. 2011). Commensal *R. rattus* populations (abundant in human settlements) present higher migration rates because of their close association with humans (Varudkar and Ramakrishnan 2015). In this study, the bidirectional gene flow pattern observed in this commensal species might be explained by a scenario in which natural barriers (i.e., the Yan Mountains) and man-made construction (i.e., the Great Wall along the mountains) restrict the movement of the 2 rat subspecies; however, migration might be facilitated along man-made transportation lines (the roads, highways, and railways through the Shanhai Pass along the Great Wall). The West Liaoning Corridor between the Yan Mountains and the Bo Sea has allowed secondary contact and recent gene flow. However, isolation by resistance associated with the mountain landscapes (i.e., the Yan Mountains and the Great Wall) did not show any effect on gene flow here. It would be difficult to distinguish between these scenarios at this time. Mmd, Mmm, and *M. m. castaneus* began to diverge ~500,000 years ago. Accordingly, loci showed considerable variation in the obtained genealogical patterns. Some loci formed a monophyletic lineage within each subspecies, whereas other loci among these subspecies intermingled with each other (Geraldes et al. 2008). An alternative explanation might be incomplete lineage sorting caused by recent divergence between putative RNH and RNC subspecies, wherein a fraction of *R. norvegicus* genomes have remained undifferentiated and retained admixed regions or ancestral polymorphisms

(Gryseels et al. 2016). Moreover, gene flow exhibited a moderate asymmetric pattern with a higher value from putative RNH to putative RNC and a lower value from putative RNC to putative RNH. Similarly strong asymmetry of gene flow has also been found in a hybrid zone between 2 parapatric subspecies of mice (Teeter et al. 2008), with higher allele frequencies on the Mmm side (Geraldes et al. 2008; Duvaux et al. 2011). In addition, asymmetric gene flow based on autosomal microsatellite markers might be associated with genetic incompatibility, assortative mate preference or reproductive success and dispersal patterns between these 2 parapatric subspecies and might be influenced by hybrid zone dynamics (Christophe and Baudoin 1998; Smadja and Ganem 2002; Teeter et al. 2008). Overall, all this evidence confirmed the gene flow with characteristics of bidirectional and asymmetry.

Interaction processes and mechanisms among these 2 rat subspecies will deserve future studies.

Funding

Conflict of Interest Statement

The authors declare that they have no conflict of interest.

Appendix

Table A1. Microsatellite loci were chosen for population genetic analyses of RNH and RNC

Loci	Label	Primer sequence (F)	Primer sequence (R)	Tm (°C)	Multiplex partner	Source
D1Wox31	5'-HEX	5'-CATGCACACCCACTTACACAC-3'	5'-CCTATTAGAACTTCCCCCTTC-3'	57.8	D8Wox7, D12wox1	Heiberg (2006)
D8Wox7	5'-FAM	5'-GGTATACAAAGCCTCGTGCA-3'	5'-TGGGCTAAAGCTTATCCATTTA-3'	55.9	D1Wox31, D12wox1	Heiberg (2006)
D12wox1	5'-TAMRA	5'-GACATTAAGGGGTCTTCCTAAG-3'	5'-TATCTTTGCAACGCTGAGG-3'	55.9	D1Wox31, D8Wox7	Heiberg (2006)
D6wox2	5'-TAMRA	5'-CCAGTCCATACTTATCCATCTG-3'	5'-CATTTAGATAGGTGATAGATTCAG-3'	55.4	D2Wox27, D19Wox11	Heiberg (2006)
D2Wox27	5'-HEX	5'-GATAATTGACATGTCCAGTTCC-3'	5'-CTGGCTGATGGTAGGATGAG-3'	55.4	D6wox2, D19Wox11	Heiberg (2006)
D19Wox11	5'-FAM	5'-CTACCCACCCATCTATTCATCC-3'	5'-GTTTCCAGCACCCATGTCC-3'	55.4	D6wox2, D2Wox27	Heiberg (2006)
D4wox7	5'-TAMRA	5'-GATAGCATAAAATCCCTAGAGGTT-3'	5'-TGCATTTATCTGAAAACCATCAC-3'	60.8	D3Wox12, D10Wox11	Heiberg (2006)
D3Wox12	5'-HEX	5'-TATAGTAAGTTCGAGGCGCGG-3'	5'-AGGGGACCAGTGAGACTCAC-3'	60.8	D4wox7, D10Wox11	Heiberg (2006)
D10Wox11	5'-FAM	5'-TCATCTGGTGGGGACATAAC-3'	5'-GATGAACCAGCACATGGAAG-3'	60.8	D4wox7, D3Wox12	Heiberg (2006)
D9Wox23	5'-FAM	5'-CTGCAGGCCATCAGCATATGT-3'	5'-CAGTCATACATGCTGGCAGAC-3'	59.8	D1Wox23, D3Wox7	https://www.ncbi.nlm.nih.gov/probe/
D1Wox23	5'-HEX	5'-TCTGACCCATACTTGTACTTTGC-3'	5'-AATTTCTGCCTCTTTTTCTCAG-3'	59.8	D9Wox23, D3Wox7	Heiberg (2006)
D3Wox7	5'-TAMRA	5'-ATGTGTGGATTACTCACCAAGTG-3'	5'-CGAGCAATTGAACAGAAAGG-3'	59.8	D9Wox23, D1Wox23	Heiberg (2006)
D7Wox2	5'-TAMRA	5'-TGTGGCAAGTACATACATCCC-3'	5'-CTCTACCTGCCTGGTCTAGC-3'	59.8	D1Wox1	https://www.ncbi.nlm.nih.gov/probe/
D1Wox1	5'-FAM	5'-CCAATTTCTCAACATCACCCC-3'	5'-CAATATCCTATGATAGATGATGG-3'	59.8	D7Wox2	Heiberg (2006)
D8Rat54	5'-HEX	5'-AACCTGGGCCCTAAGGTAGA-3'	5'-CTGCAGCCATCAGCATATGT-3'	55.4	D2Rat3	Steen (1999)
D2Rat3	5'-TAMRA	5'-TCTCTGGTGGATACCTCACAAA-3'	5'-ACATGATGTGGAATTCCCCC-3'	55.4	D8Rat54	Bryda (2008)

The names of the loci, the fluorescent dye used for labeling the forward primer of each marker (label), the sequences of the forward primer (F), and reverse primer (R), the Tm for locus amplification, and the locus partners in the multiplex analyses are shown for each marker (multiplex partner). The cited references and database websites are provided in the last column (source).

Table A2. Pairwise F$_{ST}$ comparison between sampling sites was performed

Sites	ZZ	BJ	FS	DX	TJ	NK	XTS	FR	XL	QHD	XC	LH	HS	TIL	KY	CT	TL	SL	LS	GZL	CC	DH	SY	HRB
ZZ	0.00																							
BJ	0.08	0.00																						
FS	0.07	0.03	0.00																					
DX	0.13	0.18	0.19	0.00																				
TJ	0.14	0.13	0.12	0.28	0.00																			
NK	0.08	0.11	0.14	0.28	0.17	0.00																		
XTS	0.22	0.14	0.23	0.36	0.28	0.31	0.00																	
FR	0.10	**0.08***	0.10	0.22	0.18	0.06	0.23	0.00																
XL	0.08	0.10	0.10	0.19	0.18	0.14	0.19	0.11	0.00															
QHD	0.11	**0.08***	0.08	0.12	0.16	0.06	0.19	0.08	0.12	0.00														
XC	0.09	**0.09***	0.06	0.16	0.11	0.11	0.16	**0.10***	0.12	**0.12***	0.00													
LH	0.24	0.16	0.26	0.44	0.19	0.36	0.39	0.17	0.16	0.21	0.17	0.00												
HS	0.06	**0.09***	0.06	0.15	0.05	0.14	0.18	**0.11***	0.11	**0.11***	0.05	0.17	0.00											
TIL	0.08	0.10	0.06	0.18	0.09	0.12	0.23	0.17	0.13	0.05	0.11	0.30	0.08	0.00										
KY	0.02	0.16	0.15	0.35	0.15	0.14	0.31	0.16	0.17	0.14	0.10	0.36	0.13	0.14	0.00									
CT	0.08	**0.13****	0.14	0.14	0.16	0.15	0.23	**0.15***	0.18	**0.13***	**0.10***	0.24	**0.05***	0.14	0.11	0.00								
TL	0.07	0.16	0.16	0.25	0.19	0.17	0.28	0.14	0.14	0.15	0.10	0.21	0.10	0.17	0.17	0.09	0.00							
SL	0.09	0.08	0.07	0.21	0.10	0.03	0.21	0.11	0.14	0.09	0.07	0.22	0.04	0.07	0.09	0.04	0.09	0.00						
LS	0.06	0.10	0.06	0.17	0.11	0.12	0.16	0.14	0.14	0.11	0.07	0.16	0.04	0.09	0.03	0.04	0.11	0.05	0.00					
GZL	0.05	**0.10***	0.09	0.17	0.10	0.13	0.20	**0.11***	0.14	0.09	0.06	0.18	0.02	0.06	0.08	0.04	0.07	0.02	0.03	0.00				
CC	0.06	0.05	0.05	0.14	**0.09***	0.12	**0.14***	**0.10****	**0.12***	0.05	0.06	0.18	0.03	0.08	0.08	0.05	**0.14***	0.03	0.02	−0.01	0.00			
DH	0.09	**0.16***	0.17	0.18	0.19	0.17	0.26	0.16	0.14	0.11	**0.12***	0.14	0.08	0.18	0.12	0.09	0.08	0.11	0.08	0.07	0.06	0.00		
SY	0.02	0.09	0.06	0.13	0.12	0.12	0.20	0.07	0.06	0.07	0.02	0.18	0.00	0.09	0.09	0.01	0.09	0.03	0.04	−0.02	−0.02	0.07	0.00	
HRB	0.16	**0.15****	0.16	0.23	0.16	0.19	**0.26***	**0.14****	0.20	**0.14***	**0.12****	0.22	**0.09***	0.19	0.20	**0.11***	0.17	0.08	0.11	0.07	0.06	0.10	0.05	0.00

Significant differentiation ($P < 0.05$) is indicated with an asterisk (*) and bold type; extremely significant differences ($P < 0.01$) are indicated with 2 asterisks (**), and a bold type.

References

Abdelkrim J, Byrom AE, Gemmell NJ, 2010. Fine-scale genetic structure of mainland invasive *Rattus rattus* populations: implications for restoration of forested conservation areas in New Zealand. *Conserv Genet* **11**: 1953–1964.

Abdelkrim J, Pascal M, Calmet C, Samadi S, 2005. Importance of assessing population genetic structure before eradication of invasive species: examples from insular Norway rat populations. *Conserv Biol* **19**:1509–1518.

Anderson EC, Thompson EA, 2002. A model-based method for identifying species hybrids using multilocus genetic data. *Genetics* **160**:1217–1229.

Bonnet E, Van De Peer Y, 2002. Zt: a software tool for simple and partial Mantel tests. *J Stat Softw* **7**:1–12.

Brouat C, Loiseau A, Kane M, Ba K, Duplantier JM, 2007. Population genetic structure of two ecologically distinct multimammate rats: the commensal *Mastomys natalensis* and the wild *Mastomys erythroleucus* in southeastern Senegal. *Mol Ecol* **16**:2985–2997.

Bryda EC, 2008. Multiplex microsatellite marker panels for genetic monitoring of common rat strains. *J Am Assoc Lab Anim Sci* **47**:37–41.

Calhoun JB, 1963. *The Ecology and Sociology of the Norway Rat*. Bethesda (MD): US Department of Health.

Christophe N, Baudoin C, 1998. Olfactory preferences in two strains of wild mice, *Mus musculus musculus* and *Mus musculus domesticus*, and their hybrids. *Anim Behav* **56**:365–369.

Combs M, Puckett EE, Richardson J, Mims D, Munshi-South J, 2018. Spatial population genomics of the brown rat (*Rattus norvegicus*) in New York City. *Mol Ecol* **27**:83–98.

Crispo E, Moore JS, Lee-Yaw JA, Gray SM, Haller BC, 2011. Broken barriers: human-induced changes to gene flow and introgression in animals: an examination of the ways in which humans increase genetic exchange among populations and species and the consequences for biodiversity. *Bioessays* **33**: 508–518.

Cushman SA, Mckelvey KS, Hayden J, Schwartz MK, 2006. Gene flow in complex landscapes: testing multiple hypotheses with causal modeling. *Am Nat* **168**:486–499.

Cushman SA, Wasserman TN, Landguth EL, Shirk AJ, 2013. Re-evaluating causal modeling with Mantel tests in landscape genetics. *Diversity* **5**:51–72.

Davis DE, 1951. The relation between level of population and size and sex of Norway rats. *Ecology* **32**:462–464.

Deinum EE, Halligan DL, Ness RW, Zhang YH, Cong L et al., 2015. Recent evolution in *Rattus norvegicus* is shaped by declining effective population size. *Mol Biol Evol* **32**:2547–2558.

Duvaux L, Belkhir K, Boulesteix M, Boursot P, 2011. Isolation and gene flow: inferring the speciation history of European house mice. *Mol Ecol* **20**: 5248–5264.

Evanno G, Regnaut S, Goudet J, 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol* **14**:2611–2620.

Excoffier L, Lischer HE, 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol Ecol Resour* **10**:564–567.

Fritz U, Fritzsch G, Lehr E, Ducotterd JM, Muller A, 2005. The Atlas Mountains, not the Strait of Gibraltar, as a biogeographic barrier for *Mauremys leprosa* (Reptilia: Testudines). *Salamandra* **41**:97–106.

Gardner-Santana LC, Norris DE, Fornadel CM, Hinson ER, Klein SL et al., 2009. Commensal ecology, urban landscapes, and their influence on the genetic characteristics of city-dwelling Norway rats (*Rattus norvegicus*). *Mol Ecol* **18**:2766–2778.

Geraldes A, Basset P, Gibson B, Smith KL, Harr B et al., 2008. Inferring the history of speciation in house mice from autosomal, X-linked, Y-linked and mitochondrial genes. *Mol Ecol* **17**:5349–5363.

Gilabert A, Loiseau A, Duplantier J-M, Rahelinirina S, Rahalison L et al., 2007. Genetic structure of black rat populations in a rural plague focus in Madagascar. *Can J Zool* **85**:965–972.

Gonçalves P, Oliveira JD, 2004. Morphological and genetic variation between two sympatric forms of *Oxymycterus* (Rodentia: Sigmodontinae): an evaluation of hypotheses of differentiation within the genus. *J Mammal* **85**: 148–161.

Goudet J, 1995. A computer program to calculate F-statistics. *J Hered* **86**: 485–486.

Gryseels S, Gouy de Bellocq J, Makundi R, Vanmechelen K, Broeckhove J et al., 2016. Genetic distinction between contiguous urban and rural multimammate mice in Tanzania despite gene flow. *J Evol Biol* **29**:1952–1967.

Hamidy A, Matsui M, Shimada T, Nishikawa K, Yambun P et al., 2011. Morphological and genetic discordance in two species of Bornean *Leptobrachium* (Amphibia, Anura, Megophryidae). *Mol Phylogenet Evol* 61:904–913.

Heiberg AC, 2006. Reproductive success of bromadiolone–resistant rats in absence of anticoagulant pressure. *Pest Manag Sci* 62:862–871.

Hey J, 2005. On the number of new world founders: a population genetic portrait of the peopling of the Americas. *PLoS Biol* 3:e193.

Hey J, Nielsen R, 2004. Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of *Drosophila pseudoobscura and D. persimilis*. *Genetics* 167:747–760.

Holland MM, Parson W, 2011. GeneMarker(R) HID: a reliable software tool for the analysis of forensic STR data. *J Forensic Sci* 56:29–35.

Kajdacsi B, Costa F, Hyseni C, Porter F, Brown J et al., 2013. Urban population genetics of slum-dwelling rats (*Rattus norvegicus*) in Salvador, Brazil. *Mol Ecol* 22:5056–5070.

Kalinowski ST, 2004. Counting alleles with rarefaction: private alleles and hierarchical sampling designs. *Conserv Genet* 5:539–543.

Kalinowski ST, 2005. Hp-rare 1.0: a computer program for performing rarefaction on measures of allelic richness. *Mol Ecol Notes* 5:187–189.

Kalinowski ST, Taper ML, Marshall TC, 2007. Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Mol Ecol* 16:1099–1106.

Lagerholm VK, Noren K, Ehrich D, Ims RA, Killengreen ST et al., 2017. Run to the hills: gene flow among mountain areas leads to low genetic differentiation in the Norwegian lemming. *Biol J Linnean Soc* 121:1–14.

Liu Y, Jiang Y, Si Y, Kim JY, Chen ZF et al., 2011. Molecular regulation of sexual preference revealed by genetic studies of 5-HT in the brains of male mice. *Nature* 472:95–99.

Macdonald DW, Fenn MGP, 1995. Rat ranges in arable areas. *J Zool* 236: 349–353.

Mcrae BH, Shah VB, Mohapatra TK, 2013. *Circuitscape 4 User Guide*. The Nature Conservancy. Available from: http://www.circuitscape.org (accessed March 28, 2014).

Musser GG, Carleton MD, 2005. Superfamily Muroidea. In: Wilson DE, Reeder DM, editors. *Mammal Species of the World: A Taxonomic and Geographic Reference*. Baltimore (MD): Johns Hopkins University Press. 894–1531.

Ness RW, Zhang Y-H, Cong L, Wang Y, Zhang J-X et al., 2012. Nuclear gene variation in wild brown rats. *G3-Genes Genomes Genet* 2:1661–1664.

Ortiz N, Polop FJ, Andreo VC, Provensal MC, Polop JJ et al., 2017. Genetic population structure of the long-tailed pygmy rice rat (Rodentia, Cricetidae) at different geographic scales in the Argentinean Patagonia. *J Zool* 301: 215–226.

Paetkau D, Slade R, Burden M, Estoup A, 2004. Genetic assignment methods for the direct, real-time estimation of migration rate: a simulation-based exploration of accuracy and power. *Mol Ecol* 13:55–65.

Pritchard JK, Stephens M, Donnelly P, 2000. Inference of population structure using multilocus genotype data. *Genetics* 155:945–959.

Puckett EE, 2016. Global population divergence and admixture of the brown rat (*Rattus norvegicus*). *Proc Biol Sci* 283:20161762.

Puckett EE, Munshi-South J, 2019. Brown rat demography reveals pre-commensal structure in eastern Asia before expansion into southeast Asia. *Genome Res* 29:762–770.

Rousset F, 2008. GENEPOP'007: a complete re-implementation of the GENEPOP software for Windows and Linux. *Mol Ecol Resour* 8: 103–106.

Russell JC, Abdelkrim J, Fewster RM, 2009. Early colonisation population structure of a Norway rat island invasion. *Biol Invasions* 11: 1557–1567.

Schulte-Hostedde AI, Gibbs HL, Millar JS, 2001. Microgeographic genetic structure in the yellow-pine chipmunk (*Tamias amoenus*). *Mol Ecol* 10: 1625–1631.

Smadja C, Ganem G, 2002. Subspecies recognition in the house mouse: a study of two populations from the border of a hybrid zone. *Behav Ecol* 13: 312–320.

Steen RG, 1999. A high-density integrated genetic linkage and radiation hybrid map of the laboratory rat. *Genome Res* 9:AP1–AP8.

Taylor JM, Calaby JH, Deusen HMV, 1982. A revision of the Genus Rattus (Rodentia, Muridae) in the New Guinean region. *Bull Am Museum Nat Hist* 173:177–336.

Taylor K, 1978. Range of movement and activity of common rats (*Rattus norvegicus*) on agricultural land. *J Appl Ecol* 15:663–677.

Teeter KC, Payseur BA, Harris LW, Bakewell MA, Thibodeau LM et al., 2008. Genome-wide patterns of gene flow across a house mouse hybrid zone. *Genome Res* 18:67–76.

Teng HJ, Zhang YH, Shi CM, Mao FB, Cai WS et al, 2017. Population genomics reveals speciation and introgression between brown Norway rats and their sibling species. *Mol Biol Evol* 34:2214–2228.

Trenel P, Hansen MM, Normand S, Borchsenius F, 2008. Landscape genetics, historical isolation and cross-Andean gene flow in the wax palm *Ceroxylon echinulatum* (Arecaceae). *Mol Ecol* 17:3528–3540.

Varudkar A, Ramakrishnan U, 2015. Commensalism facilitates gene flow in mountains: a comparison between two *Rattus* species. *Heredity* 115: 253–261.

Wang Y, 2003. *A Complete Checklist of Mammal Species and Subspecies in China: A Taxonomic and Geographic Reference*. Beijing: China Forestry Publishing House.

Wood DA, Fisher RN, Vandergast AG, 2014. Fuzzy boundaries: color and gene flow patterns among parapatric lineages of the western shovel–nosed snake and taxonomic implication. *PLoS ONE* 9:e97494.

Wu D, 1982. On subspecific differentiation of brown rat (*Rattus norvegicus* Berkenhout) in China. *Acta Theriologica Sinica* 2: 107–112.

Zeng L, Ming C, Li Y, Su LY, Su YH et al., 2017. Out of southern east Asia of the brown rat revealed by large-scale genome sequencing. *Mol Biol Evol* 35: 149–158.