

Lineage Divergence and Historical Gene Flow in the Chinese Horseshoe Bat (*Rhinolophus sinicus*)

Xiuguang Mao^{1,2}, Guimei He¹, Junpeng Zhang¹, Stephen J. Rossiter^{2*}, Shuyi Zhang^{1*}

1 Institute of Molecular Ecology and Evolution, Institute for Advanced Studies in Multidisciplinary Science and Technology, East China Normal University, Shanghai, China, **2** School of Biological and Chemical Sciences, Queen Mary University of London, London, United Kingdom

Abstract

Closely related taxa living in sympatry provide good opportunities to investigate the origin of barriers to gene flow as well as the extent of reproductive isolation. The only two recognized subspecies of the Chinese rufous horseshoe bat *Rhinolophus sinicus* are characterized by unusual relative distributions in which *R. s. septentrionalis* is restricted to a small area within the much wider range of its sister taxon *R. s. sinicus*. To determine the history of lineage divergence and gene flow between these taxa, we applied phylogenetic, demographic and coalescent analyses to multi-locus datasets. MtDNA gene genealogies and microsatellite-based clustering together revealed three divergent lineages of *sinicus*, corresponding to Central China, East China and the offshore Hainan Island. However, the central lineage of *sinicus* showed a closer relationship with *septentrionalis* than with other lineages of *R. s. sinicus*, in contrary to morphological data. Paraphyly of *sinicus* could result from either past asymmetric mtDNA introgression between these two taxa, or could suggest *septentrionalis* evolved *in situ* from its more widespread sister subspecies. To test between these hypotheses, we applied coalescent-based phylogenetic reconstruction and Approximate Bayesian Computation (ABC). We found that *septentrionalis* is likely to be the ancestral taxon and therefore a recent origin of this subspecies can be ruled out. On the other hand, we found a clear signature of asymmetric mtDNA gene flow from *septentrionalis* into central populations of *sinicus* yet no nuclear gene flow, thus strongly pointing to historical mtDNA introgression. We suggest that the observed deeply divergent lineages within *R. sinicus* probably evolved in isolation in separate Pleistocene refugia, although their close phylogeographic correspondence with distinct eco-environmental zones suggests that divergent selection might also have promoted broad patterns of population genetic structure.

Citation: Mao X, He G, Zhang J, Rossiter SJ, Zhang S (2013) Lineage Divergence and Historical Gene Flow in the Chinese Horseshoe Bat (*Rhinolophus sinicus*). PLoS ONE 8(2): e56786. doi:10.1371/journal.pone.0056786

Editor: Dmitry A. Filatov, University of Oxford, United Kingdom

Received: August 23, 2012; **Accepted:** January 15, 2013; **Published:** February 25, 2013

Copyright: © 2013 Mao et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was funded by a grant awarded to S. Zhang under the Key Construction Program of the National “985” Project and “211” Project. X. Mao was supported by China Postdoctoral Science Foundation funded project and Marie Curie Actions-International Incoming Fellowships (IIF). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. S. Rossiter was supported by a Royal Society Fellowship (UK).

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: s.j.rossiter@qmul.ac.uk (SJR); syzhang@bio.ecnu.edu.cn (SZ)

Introduction

Closely related taxa that have sympatric or parapatric ranges provide good opportunities to investigate the origins of barriers to gene flow and the extent of reproductive isolation. Although gene flow has a powerful homogenizing effect that can hamper the divergence of populations into new taxa [1] by preventing the evolution of strong reproductive isolation [2], several recent studies suggest divergence with gene flow is not uncommon, not only in the process of speciation but also after speciation is complete [3,4,5,6]. Indeed there is good evidence that introgressive hybridization may occur between closely related species on secondary contact after a period of allopatric isolation [7,8]. By allowing for genetic exchange during divergence, the isolation-with-migration (IM) model has been developed to disentangle the relative importance of complete isolation and divergence with gene flow in the process of taxon divergence [9].

As with many temperate taxa, tropical and subtropical species underwent radical range shifts during Pleistocene glacial-interglacial cycles [10,11,12]. Recent studies have shown that Pleistocene climatic fluctuations and associated sea level changes have contributed to intraspecific diversification of many taxa from East

Asia [12,13]. Several refugia have been reported in the east and southwestern plateau of China [13,14,15], which appear to have promoted population divergence and speciation [10]. In addition to global climate cycles, the biota of East Asia was further impacted by severe climatic and environmental changes due to the uplift of the Qinghai-Tibetan Plateau (QTP). This process not only dramatically modified the topography of East Asia, leading to the current higher elevations in the west [16], but also gave rise to the East-Asia monsoon, which further enhanced the climatic difference between the east and south-west [16].

The horseshoe bats (family Rhinolophidae) number over 70 species and show highest diversity in the Old World tropics and subtropics [17]. The Chinese rufous horseshoe bat *Rhinolophus sinicus* is a common species in East Asia with a wide longitudinal range that encompasses the Qinghai-Tibetan Plateau. Its sister species, *R. thomasi* and *R. rouxii*, occur further south and west, respectively, and together with a recently described cryptic species of *R. rouxii* from South India (*R. indorouxii* sp. nov) [18], these taxa constitute the only members of the so-called *rouxii*-group [17] (also see Figure 1). Previously, *R. sinicus* was regarded as a subspecies of *R. rouxii* based on craniodental features [19], but was later elevated

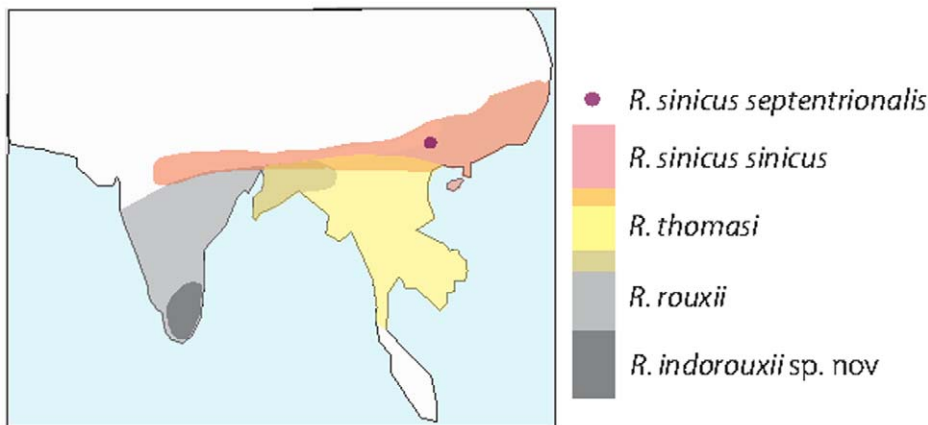


Figure 1. Map showing the distributions of the member species of the *R. rouxii*-group, and the ranges of the focal subspecies of *Rhinolophus sinicus*.

doi:10.1371/journal.pone.0056786.g001

to a distinct species following detailed phenetic, molecular [20] and cytogenetic analyses [21,22].

In spite of its wide distribution, *R. sinicus* is currently recognized as consisting of just two subspecies (see Figure 1): *R. s. sinicus* and *R. s. septentrionalis*, which can be distinguished on the basis of morphological characters; in particular, *septentrionalis* has a larger body size than *sinicus* (see [17] and Figure 2a in this study). These two recognized subspecies have unusual relative distributions that raise questions about the origin of *septentrionalis*. Currently the subspecies *sinicus* is thought to range from the Himalayas, through northern Vietnam and regions of Sichuan to southeastern China. In contrast, *septentrionalis* is thought to be highly restricted to Yunnan Province in China [17] and is completely surrounded by its more widespread sister taxon (see Figure 1). Three alternative hypotheses might account for the unusual asymmetric distributions of these two taxa. The first and most simple explanation is that *septentrionalis* and *sinicus* are not true sister taxa and thus the current taxonomy is incorrect. Second, they are sister taxa and *septentrionalis* has evolved *in situ* via divergence from *sinicus*, in which case the former will appear to be derived from the latter in phylogenetic analyses, and will harbor a subset of the genetic variation. If this is true we might also expect historical gene flow during divergence. The third hypothesis is that *septentrionalis* evolved allopatrically, possibly related to Pleistocene glaciation and/or uplift of the Qinghai-Tibetan Plateau. In this scenario, the current sympatric distributions will have arisen from secondary contact.

To distinguish between these three broad hypotheses, we sampled individuals of *R. sinicus* from across China, including both putative subspecies, and generated multilocus datasets consisting of two mitochondrial genes, five nuclear genes and nine microsatellite loci. We then conducted a range of phylogenetic, demographic and coalescent analyses that together provide information on the history of lineage divergence and gene flow among these taxa.

Materials and Methods

Ethics Statement

Our procedure of sampling consisted of taking wing membrane biopsies from animals, and was approved by the National Animal Research Authority, East China Normal University (approval ID 20080209). Bats were captured using mist nets at cave exits, and the bats were measured, sampled and subsequently released *in situ*.

Bat Capture and Measurement, Sample Collection and DNA Extraction

We collected data from individuals of *R. sinicus* from 41 localities across the Chinese mainland, including the known range of *septentrionalis* in Yunnan Province, and the offshore Hainan Island (Figure 2c and Table 1). Bats were captured in mist nets set at cave exits and, for each individual, a wing membrane biopsy was taken and stored in 95% ethanol at -20°C until genomic DNA was extracted using Qiagen kits. Additionally, the forearm of each individual was measured with dial calipers, and the body mass obtained using Pesola spring scales. For a subset of bats, we also recorded the echolocation call resting frequency while in the hand using the Avisoft UltraSoundGate 116Hnb kit (Avisoft, Berlin, Germany) connected to a laptop. Spectrograms were analysed using Avisoft-SASLab Pro software (Avisoft) and the constant frequency of the second harmonic was extracted from a series of three consecutive pulses.

The taxonomic status of each individual was initially established on the basis of forearm length following reference [17], and validated retrospectively with molecular data and echolocation call frequencies (see Results). This led to us classifying the bats into four geographical groups, each containing several localities as follows: East (AJ, JF, GG and GD), Central (SC and HG), Hainan Island (HND) and Yunnan (YN), of which the first three were *sinicus* and the latter was *septentrionalis* (also Figure 2c).

DNA Sequencing and Phylogenetic Analyses

We amplified and sequenced two mitochondrial genes (*Cytb* and *ND1*), three autosomal (*Chd1*, *SWS1* and *TH1*), one X-chromosomal (*USP9x*) and one Y-chromosomal (*Dbx*) genes. Information on the samples and primers for these markers are summarized in Table 1 and Table 2, respectively. PCR was performed in 50 μl reaction mixtures containing 10–50 ng DNA, 0.25 mM of each primer and 25 μl Premix Taq polymerase (TaKaRa). The thermal profile for each marker has been described previously (see references in Table 2). PCRS were carried out on a PTC-220 thermal cycler (Bio-Rad).

For sequencing, we used the forward primer of each mitochondrial gene and both primers of each nuclear gene. The products were analysed on an ABI PRISM 3700 automated sequencer (Applied Biosystems). Heterozygous sites in the nuclear gene sequences were identified by the presence of clear double overlapping peaks in both the forward and reverse sequence

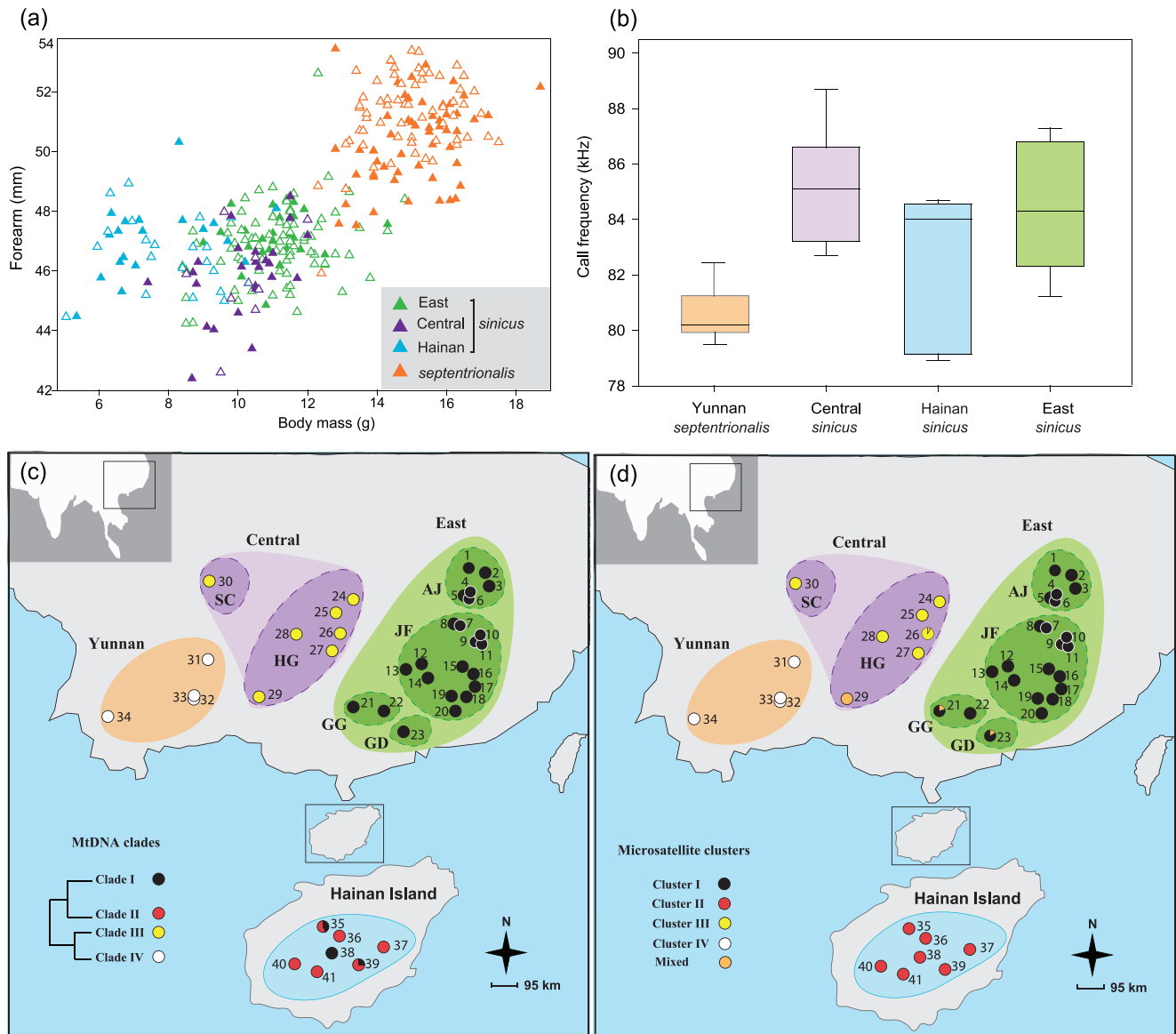


Figure 2. Morphological and echolocation data for bats with corresponding membership in mtDNA clades and microsatellite clusters. (a) Forearm (mm) and body mass (g) data measured for 396 individuals. Open and filled triangles correspond to female and male bats, respectively. (b) Boxplots of the echolocation call frequency (kHz) recorded from 130 individuals. (c) Map showing the sample sites of *R. sinicus* in this study. Forty-one sampling localities were classified into eight regions, depicted by dashed lines. These regions were further divided into four groups coded by colours (East: green; Central: purple; Hainan Island (HND): blue; Yunnan (YN): orange). Populations are presented as pie charts in which individuals are coloured based on the membership of mtDNA clades (I, II, III and IV). The relationships between clades are shown on the left side of the map (see details in Figure 2). (d) Populations are presented as pie charts in which individuals are coloured based on the membership of microsatellite clusters (I, II, III, IV and mixed, see Figure 4).
doi:10.1371/journal.pone.0056786.g002

chromatograms. Haplotypes were resolved unambiguously if there was only one heterozygous site, and, for sequences with more than one heterozygous site, we resolved haplotypes probabilistically using PHASE version 2.1 [23]. The sequences were aligned using CLUSTAL_X 1.83 [24] and edited by eye with BIOEDIT version 7.0.0 [25]. All sequences obtained in this study were deposited in GenBank (accession numbers: JN650618–JN651091 and JN651094–JN651146). For each region the number of haplotypes, haplotype diversity and nucleotide diversity was calculated using DnaSP 4.0 [26].

Phylogenetic analyses were performed for both *Cytb* and concatenated mtDNA sequences (*Cytb* and *ND1*). To investigate

the relationship between the two recognized subspecies of *R. sinicus* we also incorporated *Cytb* data from two outgroups: *R. thomasi* from the *rouxii*-group, and *R. affinis* from a different clade (accession numbers FJ85215 and DQ297582 respectively). To establish phylogenetic consensus we implemented maximum parsimony (MP) and neighbor joining (NJ) in PAUP* 4.0b10 [27], Bayesian Inference (BI) in MrBayes 3.1.2 [28] and maximum-likelihood (ML) in PhyML [29]. For MP, we used heuristic searching with tree-bisection-reconnection and branch swapping. Characters were treated as equally weighted and unordered. Node support was estimated based on 1000 bootstraps. We used MODELTEST 3.0 [30] to determine the best-fit substitution models to be

Table 1. Summary of the number and sampling locality of individuals used in the molecular analyses.

No	Locality	Coordinates	Code	<i>Cytb</i>	<i>ND1</i>	<i>Chd1</i>	<i>SWS1</i>	<i>THY</i>	<i>USP9x</i>	Microsatellites
1	Qingyang, Anhui	N30:20:511 E117:50:128	AJ	1	1	1	1	1	1	1
2	Jingxian, Anhui	N30:26:785 E118:24:783	AJ	6	2	1	2	1	1	6
3	Huashanmiku, Anhui	N29:45:199 E118:23:331	AJ	7	2	4	2	1	1	11
4	Sanling mountain, Jiangxi	N29:22:112 E117:34:324	AJ	8	1	6	2	1	1	12
5	Qingfeng cave, Jiangxi	N29:22:262 E117:39:357	AJ	5	1	1	1	1	1	5
6	Qinhui cave, Jiangxi	N29:22:662 E117:32:335	AJ	5	1	2	–	1	1	6
7	Longhu mountain, Jiangxi	N28:04:107 E116:58:227	JF	5	1	2	2	1	1	6
8	Lijia country, Jiangxi	N28:06:726 E116:59:282	JF	8	1	5	3	1	1	13
9	Wuyishan baohuqu, Fujian	N27:44:000 E117:40:000	JF	10	2	9	5	2	2	26
10	Wuyishan tiliqiao, Fujian	N27:44:543 E117:29:953	JF	1	–	1	1	1	1	2
11	Wuyishan Yanzijiao, Fujian	N27:48:511 E117:42:505	JF	1	1	1	1	–	–	2
12	Taihe, Jiangxi province	N26:36:151 E114:12:734	JF	8	2	3	2	–	1	16
13	Jinggang mountain, Jiangxi	N26:31:215 E115:06:610	JF	2	1	2	–	1	1	2
14	Xingguo, Jiangxi	N26:19:314 E115:35:229	JF	3	2	3	1	1	1	5
15	Taining, Fujian	N26:42:236 E117:29:867	JF	4	1	5	4	1	1	11
16	Jiangle, Fujian	N26:39:537 E117:34:387	JF	4	2	2	1	1	2	4
17	Mingxi, Fujian	N26:21:440 E117:11:398	JF	2	1	–	–	1	1	2
18	Yongan, Fujian	N25:51:500 E117:17:000	JF	1	1	1	–	1	1	2
19	Liancheng, Fujian	N25:12:404 E117:15:066	JF	4	2	3	2	2	2	4
20	Shanghang, Fujian	N25:15:020 E116:49:009	JF	2	1	2	1	1	1	4
21	Guilin, Guangxi	N25:16:278 E110:17:009	GG	8	2	2	3	4	6	10
22	Ruyuan, Guangdong	N24:59:086 E113:08:523	GG	1	1	–	1	1	1	1
23	Luofushan, Guangdong	N23:15:589 E114:03:656	GD	5	2	6	3	3	2	6
24	Zhangjijie, Hunan	N29:21:410 E110:34:783	HG	7	2	1	2	1	2	10
25	Yongshun, Hunan	N29:03:720 E109:38:358	HG	1	1	–	–	2	2	4
26	Jishou, Hunan	N28:18:208 E109:39:175	HG	10	3	2	1	3	3	12
27	Fenghuang, Hunan	N27:59:580 E109:33:786	HG	1	1	–	–	–	1	2
28	Wuchuan, Guizhou	N28:34:237 E107:54:058	HG	2	1	1	1	1	–	1
29	Anlong, Guizhou	N25:16:577 E105:31:931	HG	2	1	1	1	1	1	1
30	Emeishan, Sichuan	N29:34:803 E103:24:708	SC	3	2	3	2	2	2	3
31	Huize, Yunnan	N26:42:000 E103:30:000	YN	7	1	1	1	2	1	11
32	Jiuxiang, Yunnan	N25:07:000 E103:22:000	YN	1	1	1	–	1	1	1
33	Fumin, Yunnan	N25:11:796 E102:27:863	YN	3	–	–	–	2	–	3
34	Yongde, Yunnan	N24:21:427 E099:02:161	YN	10	6	5	3	9	4	11
35	Yinggeling, Hainan	N19:04:982 E109:33:107	HND	7	2	8	9	3	2	19
36	Wuzhishan, Hainan	N18:46:309 E109:31:012	HND	2	1	3	3	1	1	3
37	Qiongzong, Hainan	N18:49:589 E110:00:435	HND	–	–	3	3	2	2	13
38	Baoqing, Hainan	N18:42:237 E109:41:528	HND	1	1	1	1	1	1	3
39	Lingshui, Hainan	N18:38:586 E109:57:636	HND	3	3	1	1	1	1	5
40	Jianfengling, Hainan	N18:47:275 E108:57:364	HND	1	1	1	–	1	1	1
41	Maogan, Hainan	N18:36:306 E109:26:776	HND	3	1	–	–	–	–	3
	Total			165	59	94	66	61	56	263

doi:10.1371/journal.pone.0056786.t001

HKY+G [G = 0.092] for *Cytb* and HKY+G [G = 0.102] for concatenated mtDNA. For BI, we performed two simultaneous runs of Metropolis-coupled Markov chain Monte Carlo (MCMC) analysis using the chosen molecular evolution model parameters, each comprising four chains and 10 million generations. Trees and parameters were sampled every 100 generations, and the first 25% of the sampled trees were discarded as a burn-in. For NJ and ML,

node supports were estimated based on 1000 bootstraps. The best topology was compared with several alternative candidate hypotheses using PAUP and CONSEL [31]. Log-likelihoods of site-pattern trees estimated with PAUP were used by CONSEL to calculate the P-values for several statistical tests for which only the Approximately Unbiased (AU) test and Shimodaira-Hasagawa (SH) test [32] are presented here.

Table 2. Details of primers used in this study.

Molecular markers	ID	Length(bp)	Primers (5'→3')	Reference
Cytochrome <i>b</i>	<i>Cytb</i>	491	F: TAGAATATCAGCTTTGGGTG R: AAATCACCGTTGACTTCAAC	[68]
NADH dehydrogenase 1	<i>ND1</i>	957	F: CCTCGATGTTGGATCAGG R: GTATGGGCCCGATAGCTT	[69]
Nucleosome remodeling factor	<i>Chd1</i>	424	F:GATAAARTCAGARACAGACCTTAGACG R: TTTGGCATTACCTGYACTCC	[70]
Short-wavelength-sensitive opsin	<i>SWS1</i>	510	F: CACAGGCTATGGTGCTGACTT R: GCCCGTGGGGATGGCTATTGA	[62]
Thyrotropin	<i>THY</i>	381	F: GGGTATGTAGTTCATCTTACTTC R: GGCATCCTGGTATTCTACAGTCTTG	[71]
Ubiquitin specific protease 9 X	<i>USP9x</i>	591	F: GGCAGACAGGTTGATGAC TTGA R: AGGTCTGCAACTTGC CAAAGG AA	[70]
DEAD box RNA helicase Y	<i>Dby</i>	234	F: GGTCCAGGAGARGCTTTGAA R: CAGCCAATTCTCTGTTGGG	[70]

doi:10.1371/journal.pone.0056786.t002

At the intraspecific level, bifurcating trees may not be the most appropriate way to represent genealogical relationships among haplotypes [33], and therefore, we also constructed networks based on *Cytb* and concatenated mtDNA haplotypes using the maximum parsimony method implemented in TCS version 1.21 [34]. Ambiguous connections were resolved based on the criteria of topology, frequency and geography [35].

To estimate the time to the most recent common ancestor (TMRCA) for each mtDNA (*Cytb*) clade (see Figure S2), we used a Bayesian MCMC approach implemented in BEAST version 1.4.6 [36]. A HKY+G [G = 0.01] substitution model for *Cytb* dataset without outgroup estimated by MODELTEST, and a relaxed-clock model with an uncorrelated lognormal distribution for the substitution rate, were applied. We used a substitution rate of 0.02/Myr based on a divergence rate of 4%/Myr for *Cytb* obtained for *Pipistrellus* bats [37]. We performed two independent runs of 10^7 generations, each with a burn-in of 10^6 generations, and sampled every 1000 steps. These two runs were then combined in TRACER version 1.4 [38], which were also used to examine the effective sample size (ESS) for each parameter.

For each nuclear gene, phylogenetic relationships among reconstructed haplotypes were displayed using haplotype networks. To infer the species phylogeny from these multiple nuclear sequences, as well as infer each gene tree, we used the Bayesian approach in BEST 2.0 that takes account the coalescent process [39,40]. For this analysis, run at the Computational Biology Service Unit at Cornell University, model parameters for each marker were estimated using MODELTEST. The priors for theta were an inverse gamma distribution (3, 0.003) and the priors for gene mutation were a uniform distribution (0.5, 1.5). We performed two runs of Metropolis-coupled MCMC, each comprising four chains and 20 million generations. Trees and parameters were sampled every 100 generations, and the first 25% of the sampled trees were discarded as a burn-in. A consensus tree was constructed from the estimated posterior distribution and used as the estimated species tree [39].

Hypothesis Testing of Lineage Divergence

To test probabilistically among alternative hypotheses for the history of lineage divergence, we used the approximate Bayesian

computation procedure (ABC, [41]) in DIY ABC v1.0.4.37 [42]. ABC analyses were applied to sequences of the concatenated mtDNA data as well as three informative nuclear genes (*THY* was excluded because it contains indels and has less informative sites). We tested the following three competing scenarios: (1) *septentrionalis* arose recently via divergence from Central *sinicus* following a more ancient divergence between the common ancestor of these taxa, and lineages of *sinicus* from the East and Hainan. We also incorporated a bottleneck early in the history of *septentrionalis*, because of its small population size. This scenario corresponds approximately to our hypothesis 2 in the Introduction. Scenario (2) is the same as (1) but without the bottleneck. In scenario (3) *septentrionalis* was considered the ancestor of all *R. sinicus* with a subsequent divergence between Central *sinicus* and East/Hainan *sinicus* (see details in Figure S1). In the first scenario, a population size with the prior of $10-10^4$ individuals and with a bottleneck of 1 to 1000 generations was used as the founding population for *septentrionalis*. A total of 3 million datasets were simulated, of which the 1% closest to the observed data were used to estimate the relative posterior probabilities of each scenario via a logistic regression approach [43]. Equal prior probability was set for each scenario.

Microsatellite Genotyping

Nine primer pairs developed for *R. sinicus* [44] were used to screen a total of 263 bats. PCRs were undertaken in 15 µl reaction volumes (containing 50–100 ng genomic DNA, 0.25 mM forward primer (labeled with FAM, HEX or TAMRA), 0.25 mM unlabeled reverse primer, 0.2 mM of each dNTP, 0.2 U Hotstar Taq DNA polymerase (Qiagen) and 1.5 mM MgCl₂) with the thermal profile: initial denaturation step at 95°C for 15 min, followed by 35 cycles of 30 s at 94°C, 30 s at (55–66°C) and 30 s at 72°C and final extension step at 72°C for 20 min. Sample information for each locality is summarized in Table 1. The products were visualized on an ABI 3730 sequencer and analyzed with GeneMapper v3.7 (ABI).

Only individuals genotyped at over seven loci were used in subsequent analyses. Data were checked for possible null alleles and genotyping errors with Micro-checker v2.2.3 [45]. GENEPOP v3.4 [46] was used to test for deviation from Hardy–

Weinberg equilibrium (HWE) and linkage equilibrium (LE) among all loci across all populations. To quantify genetic diversity per region, we calculated allelic richness (standardized for $n = 3$) using FSTAT [47], and we calculated the mean number of alleles per locus, the number of private alleles and expected and observed heterozygosity using GDA [48].

Microsatellite-based Population Genetic Analyses

To assess the relationship among populations, we used STRUCTURE 2.2 [49] to assign individuals to clusters, for cases of 2 to 8 clusters (K) in the data, the upper limit corresponding to the number of regions in the study data [12,14]. We performed ten replicate runs (10^6 iterations and with an initial burn-in of 10^5 iterations per run) and used the admixture model. The most likely number of clusters was inferred based on the method described in [50]. Similar replicate runs were grouped based on the symmetric similarity coefficient of >0.9 using the Greedy algorithm in CLUMPP [51] and visualized using DISTRUCT 1.1 [52]. To complement the results of STRUCTURE, we used GENELAND [53] and TESS [54], which can incorporate the spatial coordinates of individuals when uncovering population structure. For both analyses, we performed ten independent runs from $K = 2$ to 8, with 10^6 iterations and a burn-in of 10%.

Finally, we repeated ABC analyses for microsatellite data, again using DIY ABC v1.0.4.37 [42]. Details of the scenarios tested are described above and shown in Figure S1. In this analysis, a total of 1.5 million datasets were simulated, of which the 1% closest to the observed data were used to estimate the relative posterior probabilities of each scenario via a logistic regression approach.

Estimates of Gene Flow

To test for gene flow between *septentrionalis* and adjacent populations of *sinicus*, we used the isolation with migration (IM) model [9] implemented in the program IM [55]. The IM model assumes that the populations under study are panmictic, and that the markers are selectively neutral and free from recombination. We tested for recombination by calculating the minimum number of recombination events (R_m) based on four-gamete test [56] and we tested for neutrality using the McDonald–Kreitman test [57] and the Hudson–Kreitman–Aguade test [58]. These analyses were conducted in DnaSP. We did not detect deviation from neutral expectations for all markers; however, there was some evidence of recombination for sections of sequence of *Cytb*, *Chd1* and *SWS1*, and only nonrecombined regions were used in the subsequent IM analyses.

IM analyses were performed for four datasets: (i) mtDNA (*Cytb*), (ii) combined ncDNA (*USP9x*, *Chd1*, *SWS1* and *THY*), (iii) combined mtDNA and ncDNA, and (iv) microsatellites. To incorporate differences in the effective population size of the markers, we set inheritance scalars as follows: 0.25 for *Cytb*, 0.75 for the X-linked locus *USP9x* and 1 for the three other nuclear genes (*Chd1*, *SWS1* and *THY*) and also for the microsatellites. Scaled directional migration rates ($m_1 = m_1/u$ and $m_2 = m_2/u$; u is mutation rate per locus per year) were estimated between *septentrionalis* and adjacent populations of *sinicus*. Preliminary runs with large parameter intervals were used to determine the starting values of prior distributions for the subsequent runs. Three final runs were conducted with different random seeds and a single chain with a burn-in of 10^7 steps for 2×10^8 iterations. For the sequence datasets, each run included five MCMCs with ten multiple chain-swapping attempts, and a geometric heating scheme with the heating terms h_1 set to 0.05. For the microsatellites dataset, each run included eight MCMCs with 25

multiple chain-swapping attempts, and a geometric heating scheme with the heating terms h_1 and h_2 set to 0.05 and 2.

Results

In total, we captured and measured 396 bats, of which we recorded the echolocation calls of 130 individuals and undertook genetic analyses on 263 individuals.

Morphological Data and Echolocation Call Frequency

Individual bats initially assigned to *sinicus* showed similar body size and echolocation call frequencies across the study region, with considerable overlap among bats captured from the regions Central, East and Hainan Island. In contrast, bats captured in Yunnan Province around the type locality of *septentrionalis* were characterized as having larger forearms and lower echolocation call frequencies than bats from neighbouring areas, consistent with the presence of a distinct taxon in this area. For a summary of these findings see Figure 2a and b.

Mitochondrial DNA Sequences

The alignment of *Cytb* (491 bp) from 165 individuals contained 56 haplotypes with 66 polymorphic sites, and the alignment of *ND1* (957 bp) from 59 individuals contained 28 haplotypes with 88 polymorphic sites. No premature stop codons were observed in either of these genes, suggesting they are functional mitochondrial genes rather than nuclear copies. When concatenated, the mtDNA alignment spanned 1448 bp, and data from 59 individuals contained 40 haplotypes with 124 polymorphic sites.

Phylogenetic tree and network analyses indicated that *R. sinicus* is a monophyletic grouping to the exclusion of the closely related congener *R. thomasi* (see Figure S2), supporting the sister relationship of the two subspecies *sinicus* and *septentrionalis*. Within *R. sinicus*, all tree-building methods recovered similar topologies in which four well-supported clades were resolved, corresponding closely to East, Hainan Island, Central and Yunnan, hereafter referred to as clades I, II, III and IV, respectively (Figures 2 and 3). Two haplotypes from Hainan (clade II) were exceptions to this trend, and were classified with haplotypes from the East (clade I). Surprisingly, *sinicus* bats from the Central area showed a closer relationship with *septentrionalis* than with other *sinicus* lineages. The AU and SH tests both supported the current topology and were able to reject the alternative phylogenetic hypothesis ($[(IV,(III,II,I))]$; AU test, $P < 0.001$; SH test, $P = 0.03$). Consistent with the tree results, network analysis for concatenated mtDNA recovered three networks with a 95% connection limit, corresponding to clade I, II and III+IV, respectively (Figure 3).

For the *Cytb* dataset, the samples from Yunnan contained the lowest level of nucleotide diversity followed by those from the Central area (Table 3). In contrast, the samples from Hainan and the East showed very high levels of nucleotide diversity, indicating a long and stable evolutionary history or just a mixture of different evolutionary lineages. The latter scenario is supported by the observation of divergent lineages within the East samples and of several haplotypes seen in Hainan samples also nested within the East clade.

Bayesian estimates of the TMRCA performed in BEAST provided reliable estimates of all parameters with ESS >500 . The inferred TMRCA for all *Cytb* sequences was around 1.5 MYA, corresponding to the early Quaternary. TMRCA estimates obtained for all *septentrionalis* (clade VI) and for all Central *sinicus* (clade III) were similar to each other, both estimated to be around 0.3 MYA. This date was more recent than that of the East *sinicus* (clade I) but older than Hainan (clade II) (see details in Table 4).

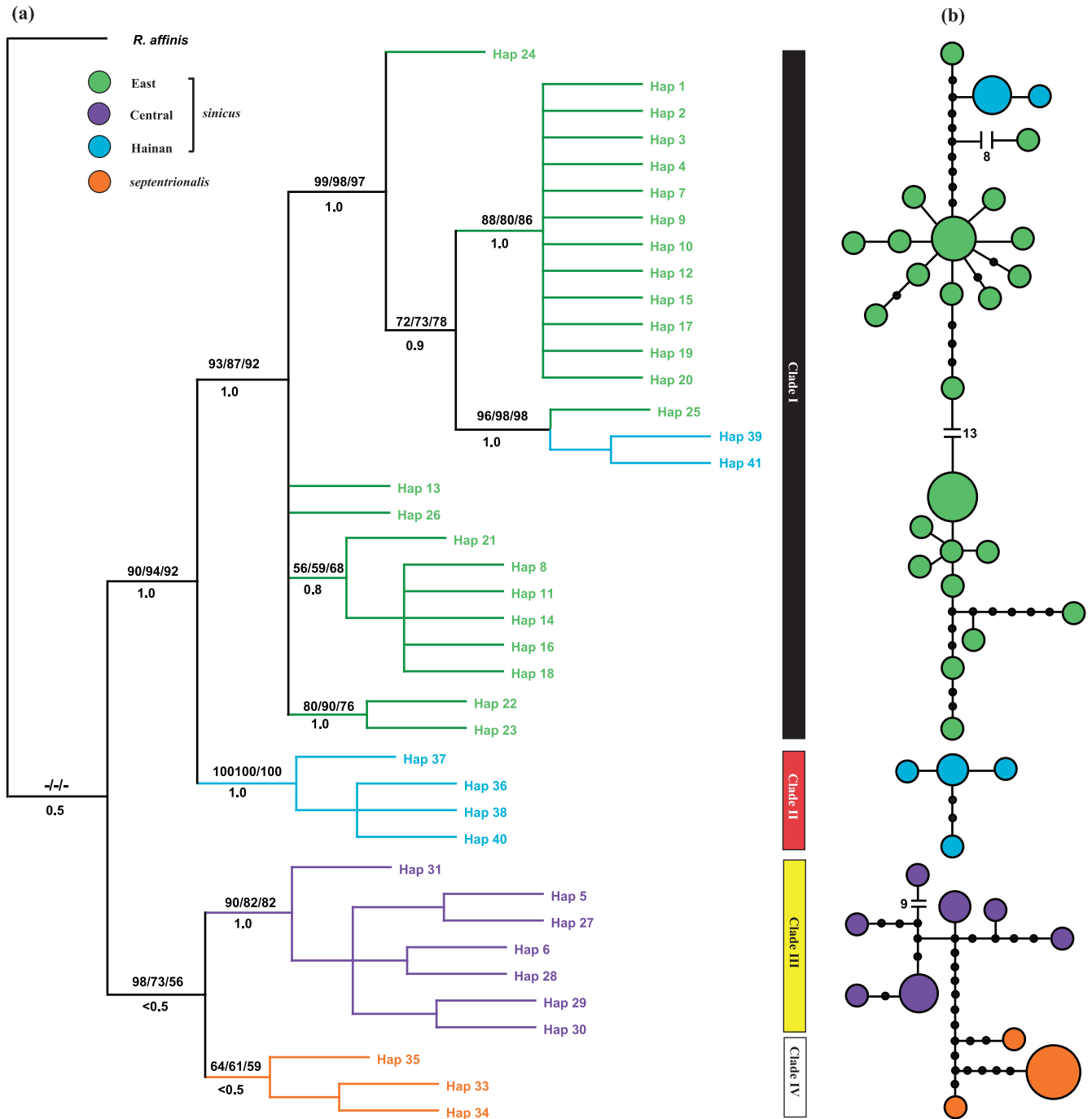


Figure 3. Maximum parsimony tree and statistical parsimony network based on concatenated mtDNA sequences. (a) The phylogenetic tree. Node support is indicated with MP, NJ and ML bootstrap values and Bayesian posterior probabilities (given above and below the branch, respectively). (b) The network with a 95% connection limit. Each circle represents a single haplotype and the circle size is scaled by the haplotype frequency. Filled black circles represent missing or unsampled haplotypes. The numbers in the network represent mutational steps between haplotypes. doi:10.1371/journal.pone.0056786.g003

Nuclear DNA Sequence Analyses

Within *R. sinicus* all nuclear markers were polymorphic except for *Dby* which was invariant across the bats screened here and, therefore, discarded prior to genetic analyses. The alignment of *Chd1* (424 bp) from 94 individuals included 5 haplotypes based on 4 polymorphic sites and 11 indels. The alignment of *SWS1* sequences (510 bp) from 66 individuals included 33 haplotypes

based on 32 polymorphic sites and 20 indels. The alignment of *THY* (381 bp) from 61 individuals included 12 haplotypes based on 6 polymorphic sites and 10 indels. The alignment of *USP9x* sequences (591 bp) from 56 individuals included 12 haplotypes based on 12 polymorphic sites. Levels of nucleotide diversity were not consistent across four ncDNA markers (see Table 3). Overall, Yunnan and the Central area showed the highest mean levels of

Table 3. Summary of genetic diversity for five gene sequences and eight microsatellite loci in eight regions of *R. sinicus*.

Genes	Parameters\regions	AJ	JF	GG	GD	East	HG	SC	Central	HND	YN
<i>Cytb</i>	N1	32	55	9	5	101	23	3	26	17	21
	n	8	19	4	4	32	8	3	11	7	5
	h	0.724	0.893	0.694	0.900	0.902	0.581	1.000	0.674	0.772	0.681
	π (%)	0.214	0.674	0.328	0.855	0.652	0.328	0.407	0.449	1.480	0.376
<i>Chd1</i>	N2	15	38	2	6	61	4	3	7	17	7
	n	2	2	1	2	2	1	1	1	2	2
	h	0.495	0.483	–	0.571	0.480	–	–	–	0.382	0.286
	π (%)	0.117	0.114	–	0.135	0.113	–	–	–	0.992	0.067
<i>SWS1</i>	N3	9	23	4	3	39	5	2	7	17	4
	n	8	13	3	2	18	5	2	7	7	4
	h	0.956	0.913	0.833	0.667	0.911	1.000	1.000	1.000	0.772	1.000
	π (%)	0.643	0.621	0.295	0.131	0.584	2.630	0.984	2.110	0.402	2.940
<i>THY</i>	N4	6	14	5	3	28	8	2	10	8	14
	n	2	2	2	1	3	5	2	6	2	5
	h	0.286	0.366	0.524	–	0.340	0.833	0.667	0.803	0.530	0.450
	π (%)	0.075	0.096	0.150	–	0.092	0.656	0.175	0.537	0.140	0.128
<i>USP9x</i>	N5	6	16	7	2	31	9	2	11	8	6
	n	1	2	6	1	6	6	2	6	1	1
	h	–	0.125	0.952	–	0.488	0.889	1.000	0.873	–	–
	π (%)	–	0.021	0.354	–	0.123	0.291	0.338	0.283	–	–
	Mean π (%)	–	–	–	–	0.228	–	–	0.733	0.384	0.784
Microsatellites	N6	40	95	11	6	152	30	3	33	43	23
	A	6.4	9.5	5.6	4.8	11.9	9.8	3.5	12.3	8.9	10.0
	H _E	0.655	0.645	0.677	0.627	0.673	0.759	0.700	0.759	0.698	0.807
	H _O	0.653	0.625	0.693	0.583	0.636	0.713	0.625	0.705	0.680	0.750
	R _S	3.15	3.18	3.36	3.24	3.36	3.99	3.63	3.98	3.47	4.06
	Private alleles	3(0)	5(0)	2(1)	0	–	12(1)	3(3)	–	5(0)	8(3)

N1–N6, sample sizes for genes and microsatellite loci; n, number of haplotypes observed; h, haplotype diversity; π (%), nucleotide diversity; mean π (%), mean nucleotide diversity of four nuclear genes. A, mean number of alleles per locus; H_E, expected heterozygosity; H_O, observed heterozygosity; R_S, allele richness standardized by three individuals. The numbers of private alleles with >10% frequency are shown in parentheses.

doi:10.1371/journal.pone.0056786.t003

nucleotide diversity (π (%), 0.784 and 0.733, respectively). The levels recorded in the East were significantly lower (0.228; $t = -4.667$, $df = 3$, $P = 0.018$), and the Hainan population was

Table 4. Mean estimates of TMRCA with 95% credible intervals for each clade or group of clades based on *Cytb* data.

Clade	Mean (Myr)	95% C.I. (Myr)
clade I	0.48	0.24–0.78
clade II	0.18	0.05–0.37
clade I +II	0.99	0.53–1.57
clade III	0.33	0.16–0.52
clade IV	0.32	0.12–0.53
clade III+IV	0.48	0.23–0.76
clade I+III+IV	1.30	0.72–2.07
clade II+III+IV	1.20	0.64–1.94
clade I+II+III+IV	1.51	0.84–2.42

doi:10.1371/journal.pone.0056786.t004

also lower than the population from Yunnan, although this was not significant (0.384; $t = -1.826$, $df = 3$, $P = 0.165$).

Haplotype networks reconstructed for *Chd1* and *SWS1* both resolved two sub-networks, broadly corresponding to *sinicus* and *septentrionalis* (Figure 4a and b). The exception is that one *SWS1* haplotype from Central *sinicus* showed a closer relationship to *septentrionalis* haplotypes than to other haplotypes of *sinicus*. Between the sub-networks there were two transversions in the *Chd1* network and over ten mutational steps in the *SWS1* network. Within *sinicus*, haplotypes of Hainan were seen to be very divergent from those from the East and Central; specifically, in the *Chd1* network, most individuals from Hainan were separated from those of other *sinicus* with ten indels, while in the *SWS1* network all Hainan haplotypes clustered together. In contrast to *Chd1* and *SWS1*, the network analysis for *THY* and *USP9x* did not recover clear subspecies, and instead, individuals of *septentrionalis* were mixed with Central and Hainan *sinicus* (Figure 4c and d).

The species tree estimated from the four nuclear markers in BEST supported the reciprocal monophyly of *septentrionalis* and *sinicus* with high node support (Figure 4e). Moreover, East *sinicus* showed a closer relationship with Hainan *sinicus* than with Central *sinicus*, consistent with the mtDNA tree.

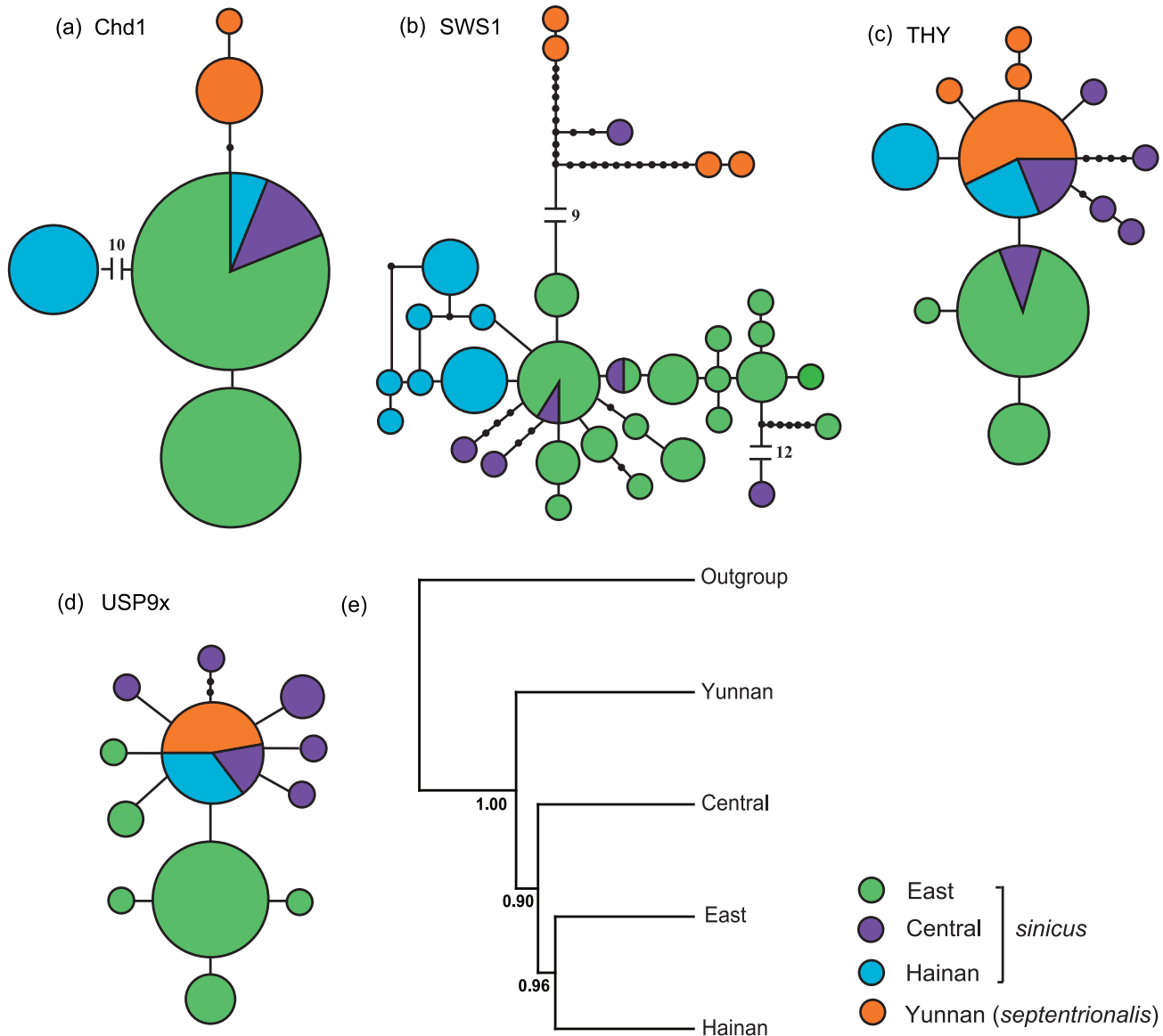


Figure 4. Statistical parsimony networks for four nuclear markers and species tree. Networks for *Chd1* (a), *SWS1* (b), *THY* (c) and *USP9x* (d). Each circle in the network represents a single haplotype and the circle size is scaled by the haplotype frequency. Filled black circles represent missing or unsampled haplotypes. The numbers in the network represent mutational steps between haplotypes. (e) Species tree of *R. sinicus* lineages and outgroups estimated on the basis of four nuclear genes. Node support is indicated with posterior probabilities. doi:10.1371/journal.pone.0056786.g004

Microsatellite-based Population Genetic Analyses

No evidence for allele dropout and scoring error due to stuttering was detected using Micro-checker except for one locus (SM24) which was deleted in the subsequent analysis. We detected no significant departure from Hardy-Weinberg equilibrium after Bonferroni correction, although significant departure from linkage equilibrium was detected between SF77 and SN30 in several locations of AJ and JF.

Genetic diversity decreased broadly from the west (i.e. Yunnan and the Central area) to the east (i.e. East and Hainan) (see Table 3). Private alleles with the frequency over 10% were observed in GG, HG, SC and YN. Among them, one allele with a frequency of 28% in Yunnan was private to *septentrionalis*, and thus private to this subspecies.

Clustering at a range of values of *K* using STRUCTURE not only revealed substantial phylogeographic structure but also

recovered hierarchical relationships among populations. Individuals of East *sinicus* were firstly separated from others at *K* = 2 (Figure 5), so conflicted with mtDNA and ncDNA sequence analyses in which Hainan showed a closer relationship with East *sinicus* than with Central *sinicus*+*septentrionalis* (Figures 3 and 4). This subdivision showed the highest value of ΔK (Figure S3). However, further substructuring was observed when increasing the value of *K*. Specifically, at *K* = 3, Hainan was separated from Central *sinicus*+*septentrionalis*. Further clustering was observed within East at *K* = 4. At *K* = 5, *septentrionalis* and Central *sinicus* finally became separate clusters. When increasing the value of *K*, further clusters were only detected in East *sinicus* but the clustering patterns were generally uninformative. Clustering conducted in GENELAND and TESS revealed similar patterns as STRUCTURE at *K* = 2 and 3 (data not shown), although *septentrionalis* was separated from the Central *sinicus* at *K* = 4

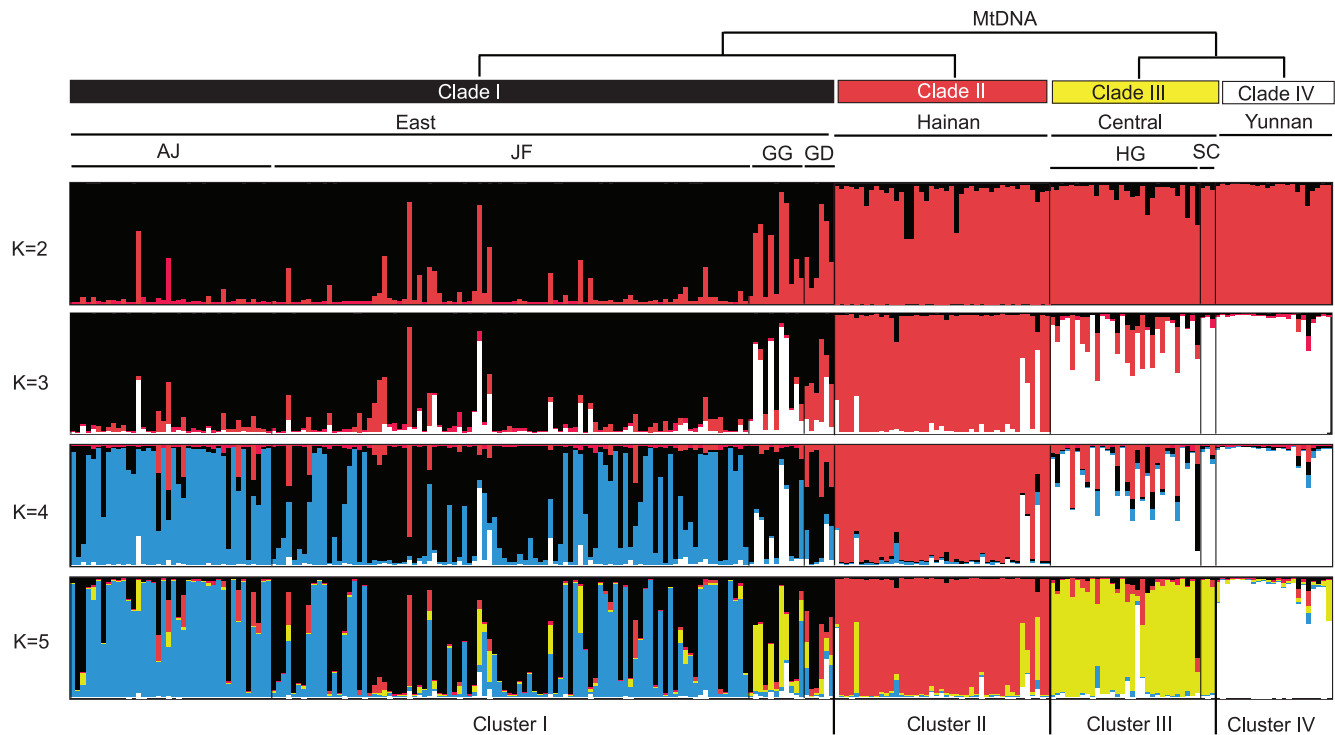


Figure 5. Clustering of *R. sinicus* individuals in STRUCTURE based on microsatellite genotypes. Clusters are shown for values of $K=2$ to 5 inclusive. The mtDNA clade relationships are shown above the four cluster plots.
doi:10.1371/journal.pone.0056786.g005

(Figures S4 and S5). The four clusters detected were generally concordant with the four mtDNA clades and four geographic groups (East, Central, Yunnan and Hainan), respectively (Figures 2c and d, S4 and S5).

ABC Analysis

ABC analysis provided good support for the scenario of an ancient divergence between *septentrionalis* and *sinicus* with later divergence between Central *sinicus* and East/Hainan *sinicus* based on both DNA sequences and microsatellites datasets (scenario 3 in Figure S1, DNA sequences data, posterior probability = 1, the type I error rate is 0.25; microsatellite data, posterior probability = 1, the type error rate is 0.64).

Estimates of Gene Flow

Three independent IM analyses produced similar posterior distributions with effective sample sizes of >500 for all parameters, suggesting convergence on the true stationary distribution. Posterior probability distributions for migration rate based on four classes of datasets all had clear peaks and bounds within the prior distributions (Figure 6). The posterior mode and 90% credible intervals for migration rates are shown in Table 5. Estimates of gene flow using IM were markedly different among the four classes of markers used. Based on mtDNA, a moderate level of unidirectional gene flow was observed from *septentrionalis* (Yunnan) to the Central population of *sinicus*. In contrast, very restricted or no gene flow was detected between Yunnan and Central in either direction when IM analyses incorporated microsatellite data or ncDNA sequences, either alone or with mtDNA (see Table 5 and Figure 6).

Discussion

The Origin of *septentrionalis* and its Relationship with *sinicus*

Morphological, echolocation and nuclear DNA sequence data collected from *Rhinolophus sinicus* from across China supported the occurrence of the two subspecies *septentrionalis* and *sinicus*. Furthermore, phylogenetic results showed that these together form a monophyletic group and are thus true sister taxa although samples from its related taxa (e.g. *R.thomasi* and *R.rouxii*) will be needed to confirm this suggestion in the future. Therefore, of our three original hypotheses proposed to account for the origin and restricted range of *septentrionalis* nested within *sinicus*, we are able to reject the first scenario that the current taxonomy is incorrect, and so must consider the alternative explanations.

MtDNA gene trees and microsatellite-based clustering together revealed three divergent lineages within the subspecies *sinicus*, corresponding to Central China, East China and the offshore Hainan Island. However, contrary to morphological data, the Central lineage showed an unexpected closer relationship with *septentrionalis* than with other lineages of *sinicus*. The paraphyletic arrangement of *sinicus* with respect to *septentrionalis* could arise if, as stated in Hypothesis 2, the latter subspecies had recently evolved *in situ* from a Central lineage of its more widespread sister subspecies. However, phylogenetic discordance can also result from other processes such as incomplete lineage sorting of ancestral polymorphism or past mtDNA introgression [59,60], which first need to be ruled out before accepting a recent origin of *septentrionalis*.

Several pieces of evidence in our study suggest that incomplete lineage sorting is an unlikely cause of the observed *sinicus* mtDNA paraphyly. First, the mtDNA haplotypes causing discordance were not distributed randomly across the range of *sinicus* as would be

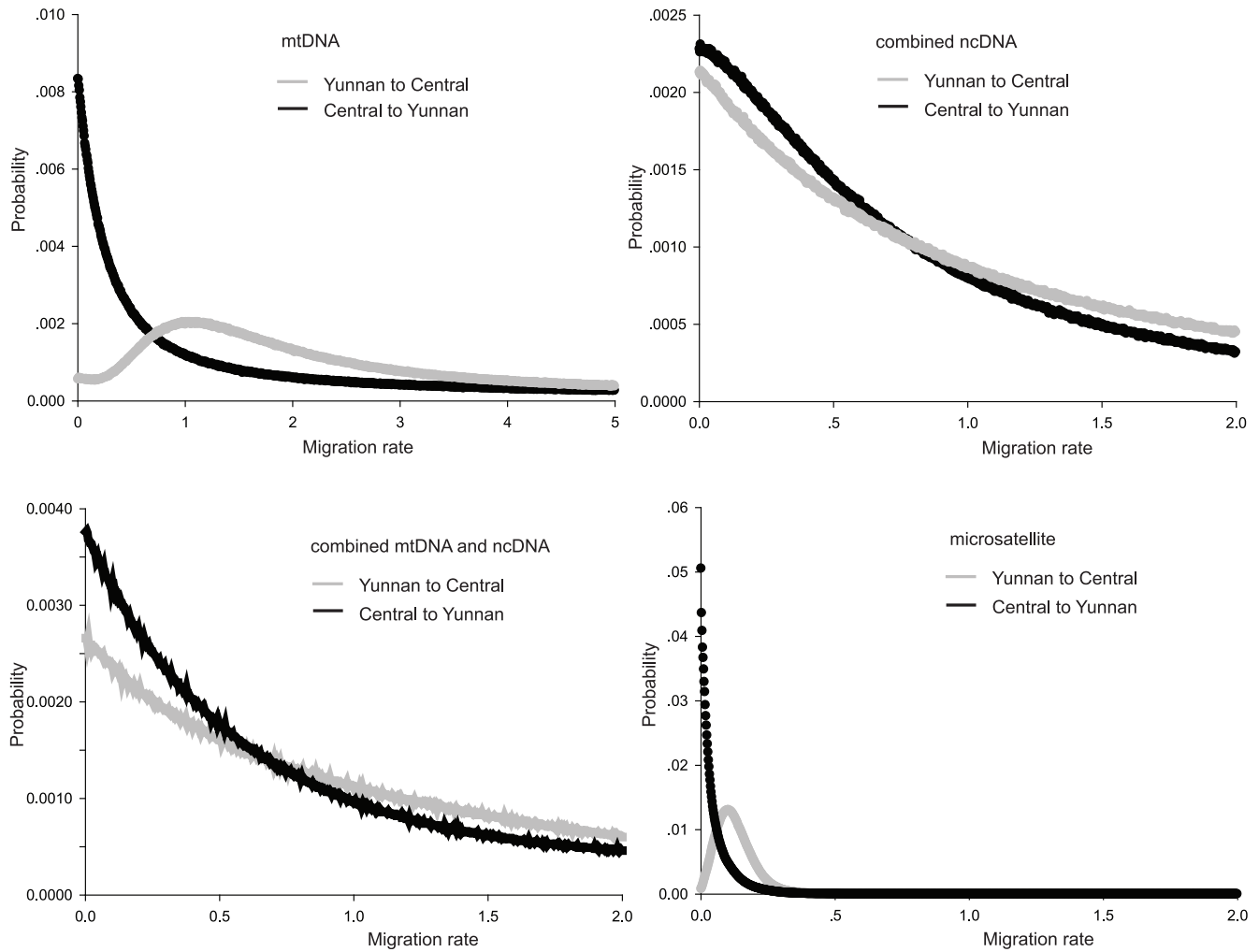


Figure 6. The marginal posterior probability distributions of migration rates between Central *sinicus* and *septentrionalis*. The IM analysis was performed for mtDNA, combined nuclear gene sequences, combined mtDNA and nuclear gene sequences and microsatellites, respectively.

doi:10.1371/journal.pone.0056786.g006

expected if this had been due to stochastic incomplete lineage sorting [2], but instead appeared in the parapatric regions of these two subspecies. Additionally, if incomplete lineage sorting had led to mtDNA paraphyly in *sinicus*, then we might expect to see similar patterns in the genealogies of nuclear genes given their larger effective population sizes [61], whereas we observed almost exclusive monophyly of each subspecies based on the markers *Chd1* and *SWS1* (Figure 4a and b). At the same time, however, such conflicts between mtDNA and ncDNA networks, plus the

geographically localized nature of the discordant haplotypes, add considerable weight to the possibility of asymmetrical mtDNA introgression between female *septentrionalis* in Yunnan and male *sinicus* in the adjoining Central areas. This scenario also received support from the results of the IM analysis that indicated moderate mtDNA gene flow from *septentrionalis* to Central *sinicus* (Table 5). Similar cases of historical and contemporary introgression among related horseshoe bat taxa have been reported previously, and also appear to occur asymmetrically [12,62].

Table 5. IM estimates of posterior mode and 90% credible intervals for directional migration rates (m_1 and m_2).

Migration rate/markers	MtDNA	Combined ncDNA	Combined mtDNA and ncDNA	Microsatellite
Central vs YN				
m_1	0.792	0.005	0.003	0.068
90% CI	0.298–1.478	0.001–1.497	0.001–1.661	0.008–0.162
m_2	0.002	0.005	0.002	0.002
90% CI	0.001–0.414	0.001–1.603	0.001–0.818	0.001–0.131

doi:10.1371/journal.pone.0056786.t005

For more formal tests of the hypotheses of a derived origin of *septentrionalis* versus introgression with *sinicus*, we applied coalescent-based phylogenetic reconstruction and Approximate Bayesian computation (ABC). From the former analysis, we found that contrary to expectations, the estimated species tree showed *septentrionalis* to be ancestral to *sinicus* (Figure 4e) and, therefore, a recent origin of this former subspecies (hypothesis 2) can be ruled out. It is noteworthy that *septentrionalis* was also found to show higher nuclear genetic diversity than regional populations of *sinicus* (Table 3), and was also characterized by a private yet frequent (28%) microsatellite allele. Such findings contradict predictions of a recent origin of *septentrionalis* but are consistent with a long evolutionary history of *septentrionalis*. ABC analysis also strongly supported the scenario of *septentrionalis* as the ancestor of all *R. sinicus* bats, although the type I error rate was high based on the microsatellite data. On the other hand, IM results offered further evidence of asymmetric mtDNA gene flow from *septentrionalis* into central populations of *sinicus* yet no nuclear gene flow, so again strongly pointing to historical mtDNA introgression.

The History and Possible Causes of Lineage Divergence in *R. sinicus*

Accepting an ancient origin of *septentrionalis*, with subsequent mtDNA introgressive hybridisation with *sinicus*, raises questions about the likely forces that have led to the relative distributions of these taxa. Our results from mtDNA and microsatellites both revealed high genetic differentiation and strong phylogeographic structuring within *R. sinicus*, with four independent lineages resolved, i.e. *septentrionalis* and three *sinicus* lineages from Central, East and Hainan (see Figure 2c and d). Similar to the congeneric horseshoe bat *R. affinis* [12] in which the divergence dates between each mtDNA clade corresponded well to the period of glaciations in Pleistocene, strong phylogeographic differentiation within *R. sinicus* might result from the effects of Pleistocene climatic fluctuations. Specifically, TMRCA obtained for all *R. sinicus* *Cytb* sequences was estimated to be around 1.5 million years BP, while TMCRA estimates for *septentrionalis* and Central *sinicus* were similar (95% CI 140,000–500,000 years BP). This latter period coincides with the penultimate glacial period of Pleistocene [63] when numerous species will have undergone range contractions into multiple refugia where isolation could have promoted genetic divergence via drift [10]. Evidence for refugia in the east and southwestern plateaus of China has been reported by previous phylogeographic studies, e.g. in frogs [13] and in bats [14,15]. In our study, a refugium in the southwest is indicated by high nuclear diversity observed in *septentrionalis*, and one in the East by both high mtDNA diversity and an ancient TMRCA estimate for bats in this region. Similarly, the Central area harbours high genetic diversity (mtDNA and ncDNA), shows morphological distinctiveness [64], and represents an independent lineage in phylogenetic analyses, so might have originated from yet another refugial area.

Given the distribution of *R. sinicus*, the observed strong genetic differentiation among lineages might have also been caused, at least in part, by geological and ecological differences between regions associated with the dramatic uplift of Qinghai-Tibetan Plateau (QTP) in the Quaternary [16]. This phenomenon had major consequences for latitudinal climatic clines between the east

and the west of East Asia [16], and led to the formation of three marked altitudinal steps that also correspond to different ecological zones (see Figure 1 in ref. [15]). Relating this topography to our focal populations, *septentrionalis* corresponds to the edge of step 1, the Central *sinicus* population to step 2, and the East *sinicus* population to step 3. Because each period of uplift of the QTP had different effects on different regions of East Asia [65], divergent selection might have been a main driver of lineage divergence as well as a barrier to gene flow (see also [66]). Further detailed analysis of the environmental variables in different regions will be needed to test the possible occurrence of divergent selection [67], and work is needed to establish whether past distributions and population sizes also fit with these explanations.

Supporting Information

Figure S1 Graphic representation of the three scenarios for the history of lineage divergence used in the DIY ABC analysis. Graphs show the logistic regression, showing the posterior probabilities of three tested scenarios. N_5 corresponds to a bottleneck lasting a time of db generations. (EPS)

Figure S2 Maximum parsimony tree and network based on *Cytb* haplotypes including one sequence of *R. thomasi*. For the tree, node support is indicated with MP, NJ and ML bootstrap values and Bayesian posterior probabilities (given above and below the branch, respectively). For the network, each circle represents a single haplotype and the circle size is scaled by the haplotype frequency. Filled black circles represent missing or unsampled haplotypes. One unfilled circle represents haplotype of *R. thomasi*. The numbers in the network represent mutational steps between haplotypes. (EPS)

Figure S3 The plots of the log-likelihood value $L(K)$ (a) and delta K value (b) based on ten runs for each value of K from 2 to 8. (EPS)

Figure S4 Geographic distributions of genetic clusters identified by GENELAND at $K = 4$. (EPS)

Figure S5 Geographic distributions of genetic clusters identified by TESS at $K = 4$. (EPS)

Acknowledgments

We thank Zhu Guangjian, Weili, Xudong, Meng Fanxing and Yuan Xinpu for assistance with field collection. We also thank Dmitry Filatov and two anonymous reviewers for helpful comments that improved an earlier version of the manuscript.

Author Contributions

Bat species sampling: JZ. Conceived and designed the experiments: XM SR SZ. Performed the experiments: XM GH. Analyzed the data: XM SR. Contributed reagents/materials/analysis tools: SZ. Wrote the paper: XM SR.

References

- Slatkin M (1987) Gene flow and the geographic structure of natural populations. *Science* 236: 787–792.
- Coyne JA, Orr HA (2004) *Speciation*. Sinauer Associates, Sunderland, Massachusetts.
- Niemiller ML, Fitzpatrick BM, Miller BT (2008) Recent divergence with gene flow in Tennessee cave salamanders (Plethodontidae, Gyrinophilus) inferred from gene genealogies. *Mol Ecol* 17: 2258–2275.

4. Stadler T, Arunyawat U, Stephan W (2008) Population genetics of speciation in two closely related wild tomatoes (*Solanum* Section *Lycopersicon*). *Genetics* 178: 339–350.
5. Kondo B, Peters JL, Rosensteel BB, Omland KE (2008) Coalescent analyses of multiple loci support a new route to speciation in birds. *Evolution* 62: 1182–1191.
6. Yeung CKL, Tsai PW, Chesser TR, Lin RC, Yao CT, et al. (2011) Testing founder effect speciation: divergence population genetics of the Spoonbills *Platalea regia* and *Pl. minor* (Threskiornithidae, Aves). *Mol Biol Evol* 28: 473–482.
7. Barton NH, Hewitt GM (1985) Analysis of hybrid zones. *Annu Rev Ecol S* 16: 113–148.
8. Mallet J (2005) Hybridization as an invasion of the genome. *Trends Ecol Evol* 20: 229–237.
9. Nielsen R, Wakeley J (2001) Distinguishing migration from isolation: a Markov chain Monte Carlo approach. *Genetics* 158: 885–896.
10. Hewitt GM (2004) Genetic consequences of climatic oscillations in the Quaternary. *Proc R Soc B* 359: 183–195.
11. Moussalli A, Moritz C, Williams SE, Carnaval AC (2009) Variable responses of skinks to a common history of rainforest fluctuation: concordance between phylogeography and palaeo-distribution models. *Mol Ecol* 18: 483–499.
12. Mao XG, Zhu GJ, Zhang SY, Rossiter SJ (2010a) Pleistocene climatic cycling drives intra-specific diversification in the intermediate horseshoe bat (*Rhinolophus affinis*) in Southern China. *Mol Ecol* 19: 2754–2769.
13. Zhang H, Yan J, Zhang G, Zhou K (2008) Phylogeography and demographic history of Chinese black-spotted frog populations (*Pelophylax nigromaculata*): evidence for independent refugia expansion and secondary contact. *BMC Evol Biol* 8: 21.
14. Flanders J, Jones G, Benda P, Dietz C, Zhang SY, et al. (2009) Phylogeography of the greater horseshoe bat, *Rhinolophus ferrumequinum*: contrasting results from mitochondrial and microsatellite data. *Mol Ecol* 18: 306–318.
15. You Y, Sun K, Xu L, Wang L, Jiang TL, et al. (2010) Pleistocene glacial cycle effects on the phylogeography of the Chinese endemic bat species, *Myotis davidii*. *BMC Evol Biol* 10: 208.
16. Zhang D, Fengquan L, Jianmin B (2000) Eco-environmental effects of the Qinghai-Tibet Plateau uplift during the Quaternary in China. *Environ Geol* 39: 1352–1358.
17. Csorba G, Ujhelyi P, Thomas N (2003) Horseshoe Bats of the World (Chiroptera: Rhinolophidae). Alana Books, Shropshire, UK.
18. Chattopadhyay B, Garg KM, AK VK, Doss D PS, Ramakrishnan U, et al. (2012) Sibling species in South Indian populations of the rufous horse-shoe bat *Rhinolophus rouxi*. *Conserv Genet* 13: 1435–1445.
19. Andersen K (1905) On the bats of the *Rhinolophus macrotis* group, with descriptions of two new forms. *Ann Mag Nat Hist* 16: 289–292.
20. Thomas NM (1997) A systematic review of selected Afro-Asiatic Rhinolophidae (Mammalia: Chiroptera): an evaluation of taxonomic methodologies. Unpublished PhD. Thesis, Harrison Zoological Museum, sevenoaks.
21. Zhang W (1985) A study on the karyotypes in four species of bat (*Rhinolophus*). *Acta Theriol Sinica* 5: 95–101.
22. Mao XG, Nie WH, Wang JH, Su WT, Ao L, et al. (2007) Karyotype evolution in *Rhinolophus* bats (Rhinolophidae, Chiroptera) illuminated by cross-species chromosome painting and G-banding comparison. *Chromosome Res* 15: 835–848.
23. Stephens M, Smith N, Donnelly P (2001) A new statistical method for haplotype reconstruction from population data. *Am J Hum Genet* 68: 978–989.
24. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* 25: 4876–4882.
25. Hall TA (1999) BIOEDIT: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser* 41: 95–98.
26. Rozas J, Sanchez-De I, Barrio JC, Messeguer X, Rozas R (2003) DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* 19: 2496–2492.
27. Swofford DL (2002) PAUP* Ver 4.0.B10. Phylogenetic analysis using parsimony and other methods. Sinauer and Associates, Sunderland, Massachusetts.
28. Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19: 1572–1574.
29. Guindon S, Gascuel O (2003) Simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol* 52: 696–704.
30. Posada D, Crandall KA (1998) MODELTEST: testing the model of DNA substitution. *Bioinformatics* 14: 817–818.
31. Shimodaira H, Hasegawa M (1999) Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Mol Biol Evol* 16: 1114–1116.
32. Shimodaira H, Hasegawa M (2001) CONSEL: for assessing the confidence of phylogenetic tree selection. *Bioinformatics* 17: 1246–1247.
33. Posada D, Crandall KA (2001) Intraspecific gene genealogies: trees grafting into networks. *Trends Ecol Evol* 16: 37–45.
34. Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. *Mol Ecol* 9: 1657–1660.
35. Pfenninger M, Posada D (2002) Phylogeographic history of the land snail *Candidula unifasciata* (Helicellinae, Stylomatophora): fragmentation, corridor migration, and secondary contact. *Evolution* 56: 1776–1788.
36. Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol* 7: 214.
37. Hulva P, Horáček I, Strelkov PP, Benda P (2004) Molecular architecture of *Pipistrellus pipistrellus*/*Pipistrellus pygmaeus* complex (Chiroptera: Vespertilionidae): further cryptic species and Mediterranean origin of the divergence. *Mol Phylogenet Evol* 32: 1023–1035.
38. Rambaut A, Drummond AJ (2007) Tracer Version 1.4. Available: <http://tree.bio.ed.ac.uk/software/tracer>.
39. Liu L (2008) BEST: Bayesian estimation of species trees under the coalescent model. *Bioinformatics* 24: 2542–2543.
40. Liu L, Edwards SV (2009) Phylogenetic analysis in the anomaly zone. *Syst Biol* 58: 452–460.
41. Beaumont MA, Zhang W, Balding DJ (2002) Approximate Bayesian computation in population genetics. *Genetics* 162: 2025–2035.
42. Cornuet JM, Ravigné A, Estoup A (2010) Inference on population history and model checking using DNA sequence and microsatellite data with the software DIYABC (v. 1.0). *BMC Bioinformatics* 11: 401.
43. Cornuet JM, Santos F, Beaumont MA, Robert CP, Marin JM, et al. (2008) Inferring population history with DIY ABC: a user-friendly approach to approximate Bayesian computation. *Bioinformatics* 24: 2713–2719.
44. Liu WC, Zhang JS, Hua PY, Zhang SY, Rossiter SJ (2009) Development and characterisation of novel microsatellite markers from the Chinese rufous horseshoe bat (*Rhinolophus sinicus*) with cross-species amplification in closely related taxa. *Mol Ecol Resour* 9: 183–185.
45. Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) Micro-checker: software for identifying and correcting genotyping errors in microsatellite data. *Mol Ecol Notes* 4: 535–538.
46. Raymond M, Rousset F (1995) GENEPOP (version 1.2): population genetics software for exact test and ecumenicism. *J Hered* 86: 248–249.
47. Goudet J (1995) FASTAT version 1.2: a computer program to calculate F-statistics. *J Hered* 86: 485–486.
48. Lewis PO, Zaykin D (2001) GENETIC DATA ANALYSIS: Computer program for the analysis of allelic data Version 1.0 (d16c). University of Connecticut, Storrs, Connecticut.
49. Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155: 945–959.
50. 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.
51. Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* 23: 1801–1806.
52. Rosenberg NA (2004) Distruct: a program for the graphical display of population structure. *Mol Ecol Notes* 4: 137–138.
53. Guillot G, Mortier F, Estoup A (2005) GENELAND: a computer package for landscape genetics. *Mol Ecol Notes* 5: 712–715.
54. Chen C, Durand E, Forbes F, François O (2007) Bayesian clustering algorithms ascertaining spatial population structure: a new computer program and a comparison study. *Mol Ecol Notes* 7: 747–756.
55. 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.
56. Hudson RR, Kaplan N (1985) Statistical properties of the number of recombination events in the history of a sample of DNA sequences. *Genetics* 111: 147–164.
57. McDonald JH, Kreitman M (1991) Adaptive protein evolution at the Adh locus in *Drosophila*. *Nature* 351: 652–654.
58. Hudson RR, Kreitman M, Aguade M (1987) A test of neutral molecular evolution based on nucleotide data. *Genetics* 116: 153–159.
59. Ballard JWO, Whitlock MC (2004) The incomplete natural history of mitochondria. *Mol Ecol* 13: 729–744.
60. Funk DJ, Omland KE (2003) Species-level paraphyly and polyphyly: frequency, causes, and consequences, with insights from animal mitochondrial DNA. *Annu Rev Ecol Evol S* 34: 397–423.
61. Palumbi SR, Cipriano F, Hare MP (2001) Predicting nuclear gene coalescence from mitochondrial data: the three-times rule. *Evolution* 55: 859–868.
62. Mao XG, Zhang JP, Zhang SY, Rossiter SJ (2010b) Historical male-mediated introgression in horseshoe bats revealed by multi-locus DNA sequence data. *Mol Ecol* 19: 1352–1366.
63. Shi YF, Cui ZJ, Su Z (2006) The Quaternary Glaciations and Environmental Variations in China. Hebei Science and Technology Press, Shijiazhuang.
64. Xu WX, Zhou ZM, Zhang JQ, Wu Y, Li YC, et al. (2005) Study on geographical variations of skull morphology of *Rhinolophus sinicus*. *Sichuan J Zool* 24: 467–472.
65. Cheng J, Liu XQ, Gao ZJ, Tang DX, Yue JW (2001) Effect of the Tibetan Plateau uplifting on geological environment of the Yunnan Plateau. *Geosci* 15: 290–296.
66. Nosil P, Crespi BJ (2004) Does gene flow constrain adaptive divergence or vice versa? A test using ecomorphology and sexual isolation in *Timema cristinae* Walking-Sticks. *Evolution* 58: 102–112.
67. Nosil P, Funk DJ, Ortiz-Barrionto D (2009) Divergent selection and heterogeneous divergence. *Mol Ecol* 18: 375–402.
68. Li G, Jones G, Rossiter SJ, Chen SF, Parsons S, et al. (2006) Phylogenetics of small horseshoe bats from East Asia based on mitochondrial DNA sequence variation. *J Mammal* 87: 1234–1240.

69. Cao Y, Janke A, Waddell PJ, Westerman M, Takenaka O, et al. (1998) Conflict among individual mitochondrial proteins in resolving the phylogeny of eutherian orders. *J Mol Evol* 47: 307–322.
70. Lim BK, Engstrom MD, Bickham JW, Patton JC (2008) Molecular phylogeny of New World sheath-tailed bats (Emballonuridae: Diclidurini) based on loci from the four genetic transmission systems in mammals. *Biol J Linn Soc* 93: 189–209.
71. Matthee CA, van Vuuren BJ, Bell D, Robinson TJ (2004) A molecular supermatrix of the rabbits and hares (Leporidae) allows for the identification of five intercontinental exchanges during the Miocene. *Syst Biol* 53: 433–447.