SUPPLEMENTARY MATERIALS

Supplementary Materials and Methods

Samples: Specimens were collected according to the guidelines of the American Society of Mammalogists and approved by the Institutional Animal Care and Use Committee (Sikes, 2016). The voucher specimens of *Typhlomys* consist of four adult males and six adult females collected from the Huangshan and Qingliangfeng mountains in Anhui Province, China. We followed a combination of morphological and molecular analyses to determine the specific status of the specimens. These were preliminarily identified based on morphology and distribution following Wang et al. (1996), and subsequently by molecular phylogenetic analysis.

DNA sequencing: DNA was extracted from muscle or liver samples using phenol-chloroform extraction (Sambrook et al., 1989). We amplified three genes in this study, including one mitochondrial gene (cytochrome *b* (cyt *b*) and two nuclear genes (growth hormone receptor, GHR (774 bp); interphotoreceptor retinoid-binding protein IRBP (1 052 bp)), which are widely used in phylogenetic studies of Asian rodent species (Cheng et al., 2017; Michaux et al., 2002). The PCR primers and program conditions are provided in Supplementary Table S3. The PCR products were sequenced at General Biosystems (Anhui) Co., Ltd., China. Sequences were edited and assembled using SeqMan (DNASTAR, Lasergene v7.1 reference for every software) and then aligned using Clustal X 1.8 (Thompson et al., 1997) in MEGA6 (Tamura et al., 2013) under default parameters. Sequences with insertions or deletions (indels) were sequenced at least twice with forward and reverse primers to confirm

variations. In addition, we obtained sequence information of other species from GenBank, and employed *Jaculus jaculus*, *Rattus rattus*, and *Myospalax aspalax* as outgroups (Supplementary Table S1).

Phylogenetic reconstructions: We used MrBayes v3.2.2 to conduct Bayesian tree estimations. Analysis was performed in three parts: i.e., 1) each gene was analyzed separately, 2) two nuclear segments were concatenated, and 3) *IRBP*, *GHR*, and cyt *b* sequences were concatenated. The aligned sequences were formatted using Clustal X 1.8 for construction of Bayesian inference. We defined data blocks based on genes and codon positions and estimated evolutionary models or partition schemes using MrModeltest 1.0b (Nylander, 2004). For Bayesian inference analyses of each gene dataset, we specified the best-fit model of nucleotide substitution in MrBayes and conducted Metropolis-coupled Markov chain Monte Carlo (MCMCMC) for 1 000 000 generations using one cold and three incrementally heated chains with trees sampled every 100 generations. For the combined-gene dataset, we sampled 3 000 000 generations of MCMCMC as above. The substitution parameters were allowed to vary independently between the two genes, allowing branch lengths to be estimated proportionally to each gene. We discarded the first 25% of all trees prior to stationarity, and the remaining samples were used to generate majority rule consensus trees. An average standard deviation of split frequencies of 0.01 was used for checking model stability (SD<0.01). We computed uncorrected pairwise genetic distances on the cyt *b* gene with the APE R package and dist.dna command (Paradis et al., 2004) (Supplementary Table S4).

Species delimitation and divergence time: Multilocus coalescent delimitation was run only with the two nuclear genes to avoid the species division results being over-dominated by the mitochondrial gene (Zhang et al., 2014). First, we used the Poisson tree processes (PTP) model to infer putative species boundaries on the concatenated nuclear gene Bayesian phylogenetic input tree, implemented in the bPTP server (Zhang et al., 2013). Second, we conducted coalescent-based species delimitation analyses of the phased nuclear genes using BPP v3.4 (Yang, 2015). We assigned the 39 individuals to six species, including five species (*T. cinereus, T. chapensis, T. daloushanensis, T. nanus* and *Typhlomys* sp. 1) and one previously potential candidate species (*T.* sp. 2), based on the splits results and Bayesian tree topology (Supplementary Table S5). With each algorithm (0 and 1), we used the inverse gamma prior for theta (ancestral population size)=(2, 1 000) and tau (root age)=(13, 1 000). We performed species delimitation analyses using a fixed guide tree (A10) derived from the concatenated sequences, and joint species delimitation and species-tree estimation (A11) (Yang, 2015). We ran 24 independent analyses using combinations of models A10 and A11, algorithms 1 and 2, and different priors to avoid mixing problems (Supplementary Table S5).

We estimated the *BEAST coalescent species tree using phased nuclear genes (*GHR* and *IRBP*) plus mitochondrial cyt *b*, implemented in BEAST v1.10 (Suchard et al., 2018). The best-fit evolutionary model was determined using bModeltest (Bouckaert & Drummond, 2017). Each BEAST analysis used four partition/substitution (HKY, HKY+G, GTR, GTR+I+G) models. Divergence time

analysis was composed of a random starting tree, a birth-death tree prior, and an uncorrelated lognormal relaxed molecular clock model (Drummond et al., 2006). Considering the fossil-calibrated age constraints, we used exponential distributions to account for uncertainty in fossil calibrations. Two calibration points derived from paleontological data were used: (1) The earliest known member of the crown clade of Myomorpha (Muroidea+Dipodidae; Rodrigues et al., 2010) at approximately 54 Ma was used as the split of the most recent common ancestor (MRCA) of all taxa to calibrate tree height and set up the prior using an exponential distribution, such that the mean age was at 54 Ma (Meredith et al., 2011). (2) We set the minimum age to 45 Ma based on the age of the Middle Eocene formation in China (Beard et al., 1994), and set up the prior using exponential distribution (offset=45, mean=4.64) following Cheng et al. (2017). Each analysis was run for 100 million generations and sampled every 10 000 generations. Convergence between runs was monitored by effective sample size (ESS) values indicative of adequate sampling (i.e., >200) in Tracer v1.6 (Rambaut et al., 2014).

Morphological analysis: External measurements were taken in the field and included: body mass (BM), head and body length (HB), tail length (TL), hind foot length (HL), and ear length (EL). Eleven cranial measurements were taken with Mitutoyo Digimatic calipers to the nearest 0.01 mm: greatest length of skull (GLS), condylobasal length (CBL), basal length (BL), interorbital breadth (IOB), braincase height (BCH), zygomatic width (ZMW), length between upper incisor and molar (LUIM), length of upper molar row (UML), crown breadth of 1st upper molars

(M^1–M^1), height of coronoid valley (HCV), and length between backmost notch point of mandibular and front of lower molars (LNM–FLM), and dental eruption and wear patterns were checked to age each specimen (Cheng et al., 2017). All samples were adult individuals and were included in morphometric analysis. All measurements were log-transformed and analyzed using principal component analysis (PCA) and canonical discriminant function analysis (DFA) in SPSS v20.0 (George & Mallery, 2011). The structures of the upper and lower molars were analyzed and described using the dental nomenclature of Qiu (1989), which was used successfully in the comparison of living and fossil species of *Typhlomys*.

REFERENCES

Beard KC, Qi T, Dawson MR, Wang BY, Li CK. 1994. A diverse new primate fauna from middle Eocene fissure-fillings in southeastern China. *Nature*, **368**(6472): 604–609.

Bouckaert RR, Drummond AJ. 2017. bModelTest: Bayesian phylogenetic site model averaging and model comparison. *BMC Evolutionary Biology*, **17**: 42.

Cheng F, He K, Chen ZZ, Zhang B, Wan T, Li JT, et al. 2017. Phylogeny and systematic revision of the genus *Typhlomys* (Rodentia, Platacanthomyidae), with description of a new species. *Journal of Mammalogy*, **98**(3): 731–743.

Drummond AJ, Ho SYW, Phillips MJ, Rambaut A. 2006. Relaxed phylogenetics and dating with confidence. *PLoS Biology*, **4**(5): e88.

George D, Mallery P. 2011. SPSS for Windows step by Step: a Simple Study Guide

and Reference, 17.0 Update. Boston: Allyn & Bacon.

Meredith RW, Janečka JE, Gatesy J, Ryder OA, Fisher CA, Teeling EC, Goodbla A, et al. 2011. Impacts of the cretaceous terrestrial revolution and KPg extinction on mammal diversification. *Science*, **334**(6055): 521–524.

Michaux JR, Chevret P, Filippucci MG, Macholan M. 2002. Phylogeny of the genus *Apodemus* with a special emphasis on the subgenus *Sylvaemus* using the nuclear IRBP gene and two mitochondrial markers: cytochrome *b* and 12S rRNA. *Molecular Phylogenetics and Evolution*, **23**(2): 123–136.

Nylander JAA. 2004. MrModeltest v2. Program Distributed by the Author. Evolutionary Biology Centre, Uppsala University.

Paradis E, Claude J, Strimmer K. 2004. APE: analyses of phylogenetics and evolution in R language. *Bioinformatics*, **20**(2): 289–290.

Qiu ZD. 1989. Fossil Platacanthomyids from the hominoid locality of Lufeng, Yunnan. *Vertebrat PalAsiatic*, **27**(4): 268–283. (in Chinese)

Rambaut A, Suchard MA, Xie D, Drummond AJ. 2014. Version 1.6, MCMC Trace Analysis Package. http://tree.bio.ed.ac.uk/software/tracer/.

Rodrigues HG, Marivaux L, Vianey–Liaud M. 2010. Phylogeny and systematic revision of Eocene Cricetidae (Rodentia, Mammalia) from Central and East Asia: on the origin of cricetid rodents. *Journal of Zoological Systematics and Evolutionary Research*, **48**(3): 259–268.

Sambrook J, Fritsch EF, Maniatis T. 1989. Molecular Cloning: A Laboratory Manual. 2nd ed. Cold Springs Harbor: Cold Spring Harbor Laboratory Press, 1659.

Sikes RS. 2016. 2016 Guidelines of the American Society of Mammalogists for the use of wild mammals in research and education. *Journal of Mammalogy*, **97**(3): 663–688.

Suchard MA, Lemey P, Baele G, Ayres DL, Drummond AJ, Rambaut A. 2018. Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evolution*, **4**(1): vey016.

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution*, **30**(12): 2725–2729.

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 Research*, **25**(24): 4876–4882.

Wang YX, Li CY, Chen ZP. 1996. Taxonomy, distribution and differentiation on *Typhlomys cinereus* (Platacanthomyidae, Mammalia). *Acta Theriologica Sinica*, **16**(1): 54–66. (in Chinese)

Yang ZH. 2015. The BPP program for species tree estimation and species delimitation. *Current Zoology*, **61**(5): 854–865.

Zhang C, Rannala B, Yang ZH. 2014. Bayesian species delimitation can be robust to guide–tree inference errors. *Systematic Biology*, **63**(6): 993–1004.

Zhang JJ, Kapli P, Pavlidis P, Stamatakis A. 2013. A general species delimitation method with applications to phylogenetic placements. *Bioinformatics*, **29**(22): 2869–2876.

Supplementary Table S1 Samples used for molecular phylogenetic analysis in this study

Species	Specimen voucher	Identification code	GenBank Accession Number		
			Cyt *b*	*GHR*	*IRBP*
T. cinereus	USNM238223	Tcin7-1	KX778363	KX778417	KX778444
	G12230	Tcin-guangdong	KX397283		
T. daloushanensis	KIZ033590	Tdal8-1	KX778382	KX778409	KX778436
	KIZ033594	Tdal8-2	KX778376	KX778403	KX778430
	KIZ033551	Tdal8-3	KX778377	KX778404	KX778431
	KIZ033595	Tdal8-4	KX778379	KX778406	KX778433
	KIZ033596	Tdal8-5	KX778381	KX778408	KX778435
	KIZ033589	Tdal8-6	KX778378	KX778405	KX778432
	KIZ033599	Tdal8-7	KX778385	KX778412	KX778439
	KIZ033600	Tdal8-8	KX778384	KX778411	KX778438
	KIZ033552	Tdal8-9	KX778380	KX778407	KX778434
	KIZ033553	Tdal8-10	KX778386	KX778413	KX778440
	KIZ033555	Tdal8-11	KX778387	KX778414	KX778441
	KIZ033556	Tdal8-12	KX778383	KX778410	KX778437
T. nanus	KIZ033584	Tnan9-1	KX778373	KX778400	KX778427
	KIZ033585	Tnan9-2	KX778374	KX778401	KX778428
	1112274	Tnan3-1	KX778375	KX778402	KX778429
T. chapensis	KIZ033587	Tcha1-1	KX778364	KX778391	KX778391

	KIZ033588	Tcha2-1	KX778365	KX778392	KX778392
	ZIN99914	Ty-145	KC209551	KJ949612	KJ949612
	ZIN99916	Ty-148	KC209552	KJ949613	KJ949613
	KIZ019150	Tcha11-1	KX778366	KX778393	KX778393
	KIZ019152	Tcha11-2	KX778367	KX778394	KX778394
	KIZ033589	Tcha12-1	KX778368	KX778395	KX778395
	KIZ031851	Tcha13-1	KX778369	KX778396	KX778396
	KIZ029295	Tcha13-2	KX778370	KX778397	KX778397
	KIZ033591	Tcha14-1	KX778371	KX778398	KX778398
Typhlomys huangshanensis Hu, **sp.**	AE1901HS01	Ths19-1	MT219901	MT232968	MT232972
nov.	AE1902HS02	Ths19-2	MT219902	MT232969	MT232973
	AE1902HS03	Ths19-3	MT219903	MT232970	MT232974
	AE1902HS04	Ths19-4	MT219904	MT232971	MT232975
Jaculus jaculus			KM397186	KM397231	KM397140
Myospalax aspalax			AF326272	GQ272599	AY326097
Rattus rattus			HM217733	AM910976	HM217746

Supplementary Table S2 Results and percentage of variance explained by principal component and discriminant function analyses

Variables	Component		Canonical axis		
	1	2	1	2	3
GLS	0.947	–0.003	–0.028	0.203	–0.240
CBL	0.928	–0.110	0.364	–0.163	0.016
BL	0.945	–0.107	–0.287	0.051	0.588
BCH	0.424	0.734	–0.293	–0.147	0.052
ZMW	0.954	–0.112	0.649	–0.612	0.283
IOB	0.173	0.890	–0.312	0.739	0.502
M^1–M^1	0.897	0.194	0.251	0.732	–0.660
LUIM	0.949	–0.031	–0.035	0.768	–0.064
UML	0.612	–0.123	0.346	–0.602	0.682
HCV	0.771	–0.416	0.280	–0.440	–0.327
LNM–FLM	0.736	0.136	0.439	–0.047	–0.005
Eigenvalues	6.987	1.613			
Total variance explained	63.52%	14.66%	59.7%	34.9%	4.4%

Supplementary Table S3 PCR program conditions for cyt *b*, *GHR*, and *IRBP* genes

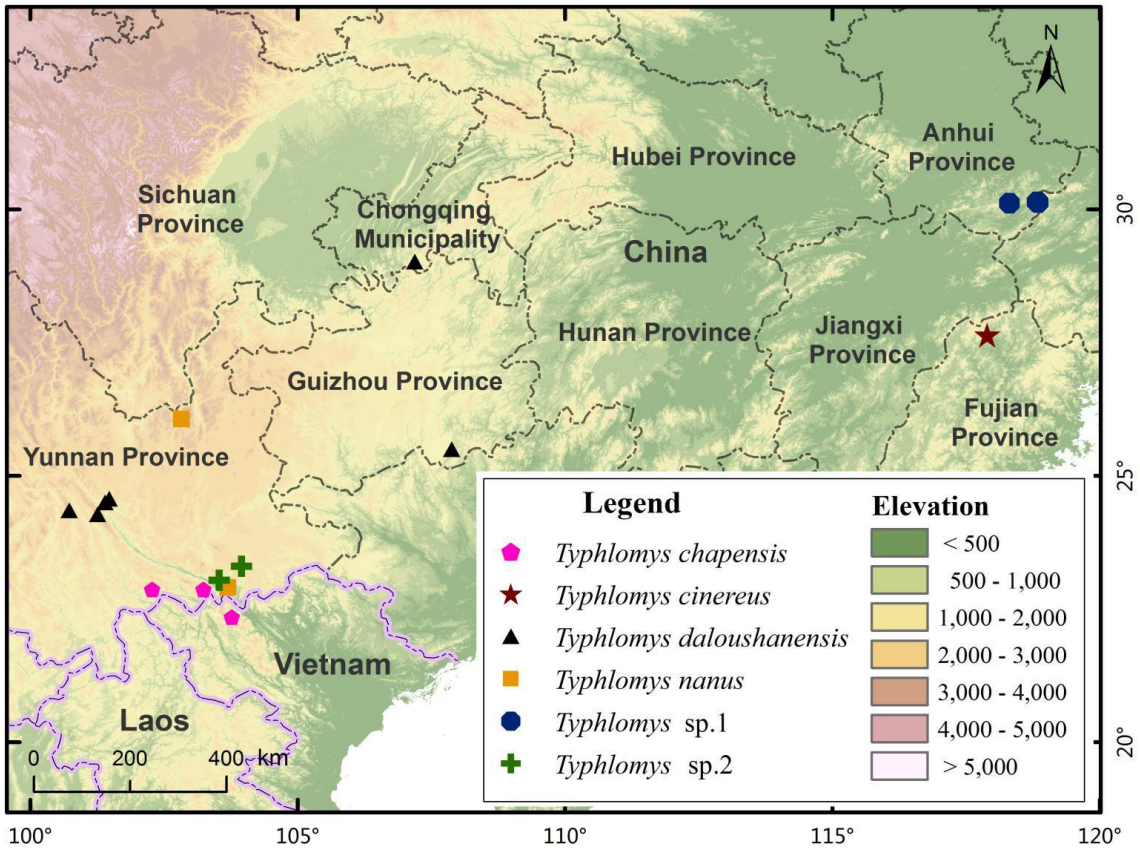
Locus	Primer	Initial denaturation	Denaturation	Annealing	Elongation	Terminal elongation
Cyt *b*	L14724-hk3, H15915-hk3 (Irwin *et al*., 1991)	2 min at 94°C	40 s at 94°C	30 s at 50°C	45 s at 72°C	5 min at 72°C
GHR	GHRF1, GHRendAlt (Jansa*et al*., 2009)	5 min at 95°C	30 s at 95°C	30 s at 52°C	1 min20s at 72°C	10 min at 72°C
IRBP	IRBPA, IRBPB (Cheng*et al*., 2017)	5 min at 95°C	30 s at 95°C	30 s at 61°C	1 min at 72°C	10 min at 72°C

Supplementary Table S4 Uncorrected pairwise genetic distances of cyt *b* sequences used in this study

Species/individual	1	2	3	4	5	6	7	8	9	10	11	12	13
1. QLF1911234													
2. QLF1911245	0.009												
3. QLF1911341	0.000	0.009											
4. QLF1911342	0.000	0.009	0.000										
5. QLF1911379	0.100	0.009	0.000	0.000									
6. AE1901HS01	0.004	0.013	0.004	0.004	0.004								
7. AE1902HS02	0.004	0.013	0.004	0.004	0.004	0.000							
8. AE1902HS03	0.004	0.014	0.004	0.004	0.004	0.000	0.000						
9. AE1902HS04	0.004	0.013	0.004	0.004	0.004	0.000	0.000	0.000					
10. T. cinereus	0.142	0.147	0.142	0.142	0.142	0.142	0.142	0.142	0.142				
11. *T. chapensis*	0.173	0.186	0.173	0.173	0.173	0.175	0.175	0.175	0.175	0.189			
12. *T.* sp. 2	0.160	0.170	0.160	0.160	0.160	0.160	0.160	0.160	0.160	0.146	0.149		
13. *T. daloushanensis*	0.172	0.178	0.173	0.173	0.173	0.173	0.173	0.173	0.173	0.173	0.186	0.129	
14. *T. nanus*	0.156	0.160	0.156	0.156	0.156	0.158	0.158	0.158	0.158	0.192	0.113	0.158	0.181

Supplementary Table S5 Summarized results of BPP analyses using different datasets, algorithms, models, and parameters

Supplementary Table S5 is listed as a separate file due to its large size.



Supplementary Figure S1 Sample localities of *Typhlomys* in this study