Short Note

Phylogeny and distribution of Palaearctic chipmunks *Eutamias* (Rodentia: Sciuridae)

Andrey A. Lissovsky[1,*], Ekaterina V. Obolenskaya[1], Deyan Ge[2], Qisen Yang[2]

[1]*Zoological Museum of Moscow State University, Bolshaya Nikitskaya 6, Moscow 125009, Russia*
[2]*Key Laboratory of Zoological Systematics and Evolution, Institute of Zoology, Chinese Academy of Sciences, Beichen West Road, Chaoyang District, Beijing 100101, China*

Abstract

We examined a cranial morphometric data set consisting of 940 specimens, as well as 156 complete sequences of the cytochrome b gene of *Eutamias sibiricus*. The geographic coordinates of 378 (northern chipmunk: 318; Chinese: 50; Korean: 10) localities were used for distribution modelling. Both morphology and genetics support existence of three subspecies of *E. sibiricus*: northern *E. s. sibiricus*, Chinese *E. s. senescens*, and Korean *E. s. barberi*. Chinese chipmunks form sister clade to northern chipmunks. Only 5.8% of the specimens have no strong similarities with representatives of their own geographical form in terms of cranial shape; an overwhelming majority of these specimens comes from the region of three taxa distributions junction. We expect two potential contact zones between Chinese and northern subspecies: in the south of the central part of the Great Khingan Range and in the central and south-eastern part of Liaoning, China.

Palaearctic chipmunks are represented by one species — *Eutamias sibiricus* (Laxmann, 1769). Siberian chipmunk has a wide distribution across major parts of Asia, crossing several climatic and vegetation zones. The species is associated with different types of forests across the range. This species is absent in tree-less landscapes far from forests; however, in arid zones, chipmunk distribution is associated with rich flood-plain vegetation (Obolenskaya, 2008). A recent review of taxonomic diversity within this species discovered the existence of three allopatric taxa of provisional subspecific rank (Obolenskaya et al., 2009). One of these sub-species, *E. s. sibiricus*, inhabits Russia, Mongolia, Japan (Hokkaido), and Northeast China. The second, *E. s. barberi* Johnson and Jones, 1955, is found in the Korean Peninsula. Lastly, *E. s. senescens* (Miller, 1898) is distributed throughout central China to the south of Liaoning province. These three subspecies have strong morphological (craniometrical and fur colouration) differences (Obolenskaya et al., 2009), which allow for identification by appearance. Only two of them were studied genetically: *E. s. sibiricus* and *E. s. barberi* (Obolenskaya et al., 2009; Koh et al., 2010; Pisanu et al., 2013). Difference in mitochondrial and nuclear genes between Korean and northern chipmunks are high, comparable with and exceeding interspecies distances in other chipmunk species (Obolenskaya et al., 2009; Patterson and Norris, 2016). Hence, it is not surprising that the three Eurasian chipmunks are cited as three potentially separate species (Patterson and Norris, 2016). In this paper, we focus on analysing the phylogenetic position of the Chinese chipmunk *E. s. senescens* and search for the potential overlap of distribution ranges of the three forms under discussion.

We analysed 156 complete sequences of *E. sibiricus* cytochrome b gene (1140 bp) (Supplement S1). New specimen of *E. s. senescens* was collected in Baishuijiang, Gansu, China; GenBank accession KX989534. Primers (L14723 and H15915) and PCR conditions were used after Irwin et al. (1991). We used 19 sequences from all available chipmunk species as an outgroup. A Bayesian analysis was performed in MrBayes 3.2.5 (Ronquist et al., 2012) with 15000000 generations (the standard deviations of split frequencies were 0.003; potential scale reduction factors were equal to 1.0; stationarity was examined in Tracer v1.6; Rambaut et al., 2014), 25% burn in, two runs with five independent chains, a sampling frequency of 1000, and the GTR + I + Γ model was separate for each codon position. Maximum likelihood distances were calculated on the basis of 500 data replicates in Treefinder (Jobb, 2011) using separate models for three codon positions.

We studied the sample comprised 940 intact skulls with a known geographical position (Supplement S2). This sample contained new 135 chipmunks from China compared to the previous studies. Seventeen measurements were taken with an accuracy of 0.1 mm (Obolenskaya et al., 2009). All calculations were carried out on log transformed measurements to linearize age variation. We performed a cluster analysis based on Mahalonobis distances using the unweighted pair group method with arithmetic mean; age and sample size biases were corrected following Obolenskaya et al. (2009). Only 98 geographic samples with n > 2 (688 specimens in total) were used in the cluster analysis. The same 688 specimens were used as learning sample in canonical discriminant analysis in order to calculate the posterior probability of the assignment of all specimens to one of the clusters. Cluster identifier after the cluster analysis (northern, Chinese, and Korean) was used as the grouping variable. Data were processed using standard algorithms implemented in STATISTICA 13.0 (Dell

*Corresponding author

Email address: andlis@zmmu.msu.ru (Andrey A. Lissovsky)

Inc. 2015) and several algorithms written by AAL under Statistica Visual Basics.

The geographic coordinates of 378 museum localities (northern chipmunk: 318; Chinese: 50; Korean: 10) were used for distribution modelling. This sample contained 318 localities of our craniometrical data set (Supplement S2) and additionally 60 localities of museum specimens with broken skulls (northern chipmunk: 54; Chinese: 6). These three subspecies datasets were tested separately and combined. We used Maxent software (Phillips and Dudik, 2008) for modelling. Eighty-six environmental variables were used as predictors: WorldClim 19 "bioclimatic" variables (http://www.worldclim.org) (Hijmans et al., 2005), altitude, slopes curvature and steepness, and 63 MODIS generalised average monthly data layers (nine months of 2004 per seven spectral bands; http://glcf.umiacs.umd.edu/data). All raster layers were converted to the equal-area Mollweide projection with a cell size of 2000 m. We didn't remove correlations among environmental variables, since Maxent has internal algorithms of feature selection (Elith et al., 2011). Maxent was executed with the following parameters: auto features or linear+quadratic features, maximum of 50000 iterations, and 20000 background points, with the remaining parameters being set to their default values. Each model was subsampled (20 times with 20% and 20 times with 40% as a test sample) to evaluate the stationarity of spatial extent. Since museum data is spatially biased, the correcting layer was created on the basis of all localities of representatives of *Ochotona*, *Microtus* s.l. and *Eutamias* from the same museums. We used three equal-range categories of relative likelihoods to prepare illustrations with the lower limit selected as the "maximum training sensitivity and specificity threshold" (Liu et al., 2013).

Previous studies reported the existence of three major morphological groups within Siberian chipmunks; two of these groups (Korean and northern) have strong genetic differences. After our analysis of enhanced data, cluster analysis also displays three well-separated clusters that completely correspond to northern, Chinese, and Korean chipmunks (Fig. 1).

Our specimen from Central China is identical to other Chinese chipmunks in fur coloration. It occupies a sister position to a pet shop specimen EU050995 (Fig. 2), thus supporting our previous proposition of the central Chinese origin of the latter (Obolenskaya et al., 2009). Both Chinese chipmunks form sister clade to northern chipmunks; this node is well supported. Net between-group ML distance between northern and Chinese chipmunks is 9.3±1.4%; northern and Korean: 14.5±4.8%; Korean and Chinese: 11.4±2.0%. Thus, all three provisional subspecies of chipmunks are supported both morphologically and genetically. Our information on genetic variation of the Chinese clade is minimal; therefore, new sampling efforts are needed for understanding its variation. However, it is less possible that new samples will fill the gap between northern and Chinese clades as a result of polymorphism. Such high variation would several times exceed the known variation in well studied northern and Korean clades. There is also no ground to suspect high internal geographic variation within Chinese chipmunks, since morphological analysis found it to be quite homogenous taxon (Fig. 1).
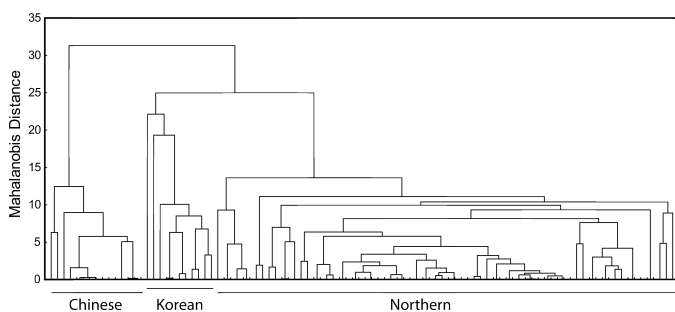
Figure 1 – The results of a hierarchical cluster analysis of craniometric features of geographical samples of *Eutamias sibiricus*.
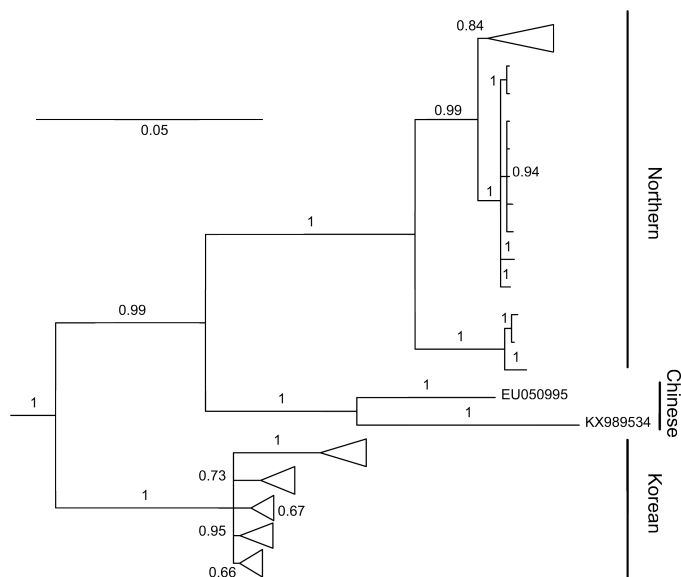

Figure 2 – Consensus tree of Bayesian analysis based on Cyt b sequences. Numbers on branches indicate Bayesian probabilities. The outgroup is not shown.

Posterior probabilities of chipmunk identification fit the distribution of three chipmunks' taxa. Only six specimens from the learning sample were identified incorrectly on the basis of posterior probabilities; twenty-eight have posterior probabilities below 0.75. Posterior probabilities of specimens, which were not included in cluster analysis, are as follows: identification of five of these specimens did not correspond to the geographical position and 11 were below 0.75. Thus, only 1.3% of the total sample was "misidentified", while 4.5% had "doubtful" identification. An overwhelming majority of specimens with problematic identification comes from the region of three taxa distributions junction (Fig. 3). Such abundance of morphologically intermediate specimens in the region of possible contact could be the result of hybridisation or the influence of similar ecological conditions on similar morphotypes. The last hypothesis seems to be less sound, as we did not find any correlation between morphological difference and geographical distance between specimens (Fig. 1). However, both hypotheses should be tested prior to any taxonomical output.

Although our distribution modelling was based upon unequal sample size for the three subspecies, spatial patterns of three distribution models were very stable in analyses of all subsamples and selected feature types. Therefore, we assume our result is not a consequence of model over complexity, which can lead to outlining any noisy set of points. In general, models reflect published data on chipmunk distribution (Obolenskaya, 2008) but have been slightly larger than simple outlines of known occurrences (Fig. 3). There are some isolated portions of the potential distribution that are not currently inhabited by chipmunks (for example, north of the Tian Shan Range). The analysis of Chinese chipmunks converged at 1860 iterations with an AUC value of 0.9898 (auto features), Korean: at 320 iterations and with AUC=0.9986, northern: at 27380 iterations, and with AUC=0.9381. The species distribution model, calculated for the whole dataset, is very similar in spatial extent to the sum of three separate models, AUC=0.9256. There is minimal spatial overlap of suitable habitats for the three taxa of chipmunks under the selected threshold. Although using of thresholds can affect the spatial pattern, it is not an obstacle to our main task — search for the potential overlap of subspecies distribution ranges. Since we cut some habitats of lower suitability, the spatial contact zones revealed in this study are the most perspective for future field studies.

Korean and northern chipmunks border at the north of the Korean Peninsula. The border stretches along the northern shore of the East Korean Bay. A potential distribution of Korean chipmunks penetrates the eastern part of China's Liaoning Province. This border region also contains some scattered isolated habitats of northern chipmunks. There is no overlap or spatial contact between the potential distribution of
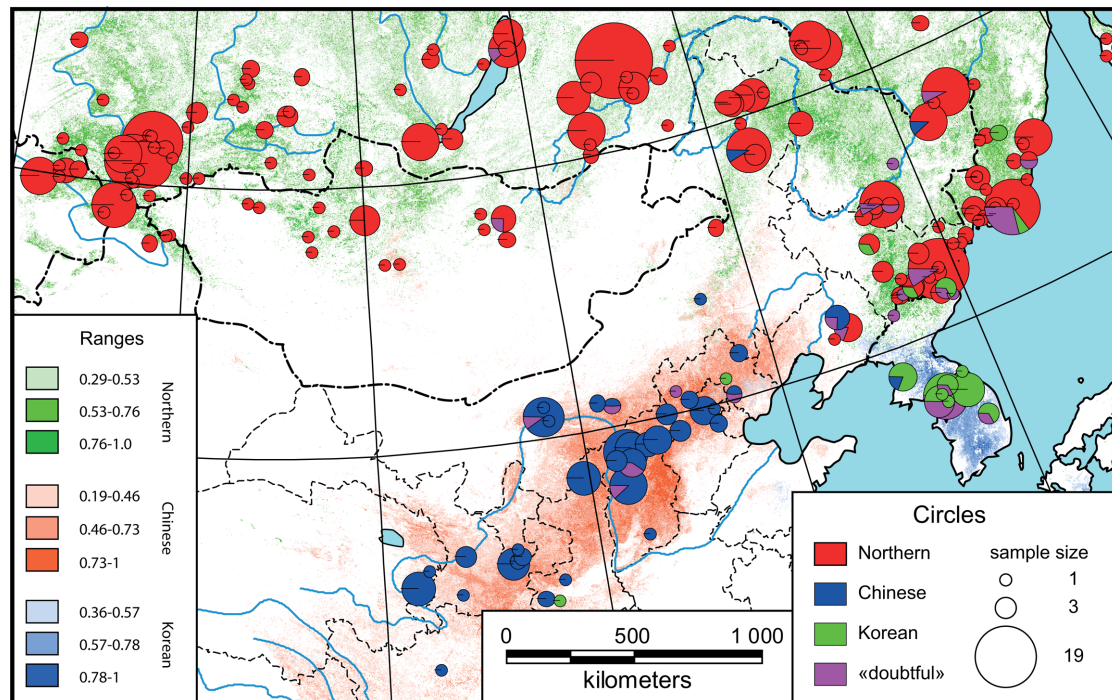
Figure 3 – Craniometrical identification of specimens of *Eutamias sibiricus* on the basis of posterior probabilities of belonging to one of three groups: northern, Chinese, or Korean. Specimens with maximal posterior probability below 0.75 are shown with magenta. The size of circles is proportional to the sample size. Samples are shown on the background of species distribution models that were modelled separately for the three taxa of chipmunks. Three colour gradations in each taxon correspond to three equal ranges of relative likelihoods of species occurrence, which roughly outline the spatial structure of potential habitat suitability.

Chinese and Korean chipmunks. Spatial contact between Chinese and northern chipmunks is predicted in two regions. The first is situated in the central part of Liaoning Province, with a scattered distribution of suitable habitats for both chipmunks in this region. This contact region is supported by morphological data; two localities of northern chipmunks and one of Chinese chipmunks are known here. The two chipmunks are separated by a distance of approximately 70–80 km. According to our models, Chinese chipmunks inhabit Liao River valley, while northern chipmunks live in higher elevations of the adjoining remote spurs of Changbai Mts. The second contact region lies along the Great Khingan Range. In general, the western slope of the Great Khingan Range possesses a more suitable environment for northern chipmunks, while the eastern slope is more suitable for Chinese chipmunks. However, a large zone of potential contact occupies the southern third of the Great Khingan Range.

We can speculate that modern distribution pattern has a relation to partial Pleistocene glaciation of the Manchurian region. Potential contact zones between northern and Chinese chipmunks in Great Khingan and between northern and Korean chipmunks on the southern limit of the Changbai Mountains are situated in the regions of Pleistocene glaciations (Sun, 1982). Glaciers or periglacial tree-less landscapes should have isolated Chinese chipmunks, as well as removed northern chipmunks from larger parts of the Changbai Mountains. The contact zone adjoining the Liao River valley have been affected by cold periods as a result of sea regression linked to climate aridization. Thus, the three taxa of chipmunks under discussion were likely more spatially separated during at least some of the glacial events of the Pleistocene period.

In conclusion, we continue to persist on the prematurity of separating Siberian chipmunks into three species. The major reason for this is the absence of studies on their hybridisation or concordance between mitochondrial and nuclear sets of genes, or any studies on possible contact zones. The present study gives the opportunity of starting such an investigation. Although contact zones between the three taxa under discussion are still unknown in nature, we found two regions that should be primarily checked: the south of the central part of the Great Khingan Range and the central and south-eastern part of the Liaoning province of China.

References

Elith J., Phillips S.J., Hastie T., Dudik M., Chee Y.E., Yates C.J., 2011. A statistical explanation of MaxEnt for ecologists. Diversity and Distributions 17(1): 43–57.

Hijmans R.J., Cameron S.E., Parra J.L., Jones P.G., Jarvis A., 2005. Very high resolution interpolated climate surfaces for global land areas. International Journal of Climatology 25(15): 1965–1978.

Irwin D.M., Kocher T.D., Wilson A.C., 1991. Evolution of the cytochrome b gene of mammals. Journal of molecular evolution 32(2): 128–44.

Jobb G., 2011. TREEFINDER version of March 2011. Munich, Germany. Retrieved from www.treefinder.de

Koh H.S., Zhang M., Bayarlkhagva D., Ham E.J., Kim J.S., Jang K.H., Park N.J., 2010. Concordant Genetic Distinctness of the Phylogroup of the Siberian Chipmunk from the Korean Peninsula (*Tamias sibiricus barberi*), Reexamined with Nuclear DNA c-myc Gene Exon 2 and mtDNA Control Region Sequences. Biochemical Genetics 48(7-8): 696–705.

Liu C., White M., Newell G., 2013. Selecting thresholds for the prediction of species occurrence with presence-only data. Journal of Biogeography 40(4): 778–89.

Obolenskaya E.V., 2008. Distribution patterns of the Siberian chipmunk (*Tamias sibiricus* Laxmann, 1769). Archives of the Zoological museum of Moscow state university 49: 265–79.

Obolenskaya E.V., Lee M-Y., Dokuchaev N.E., Oshida T., Lee M-S., Lee H., Lissovsky A.A., 2009. Diversity of Palaearctic chipmunks (*Tamias*, Sciuridae). Mammalia 73(4): 281–98.

Patterson B.D., Norris R.W., 2016. Towards a uniform nomenclature for ground squirrels: the status of the Holarctic chipmunks. Mammalia 80(3): 241–51.

Phillips S.J., Dudik M., 2008. Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. Ecography 31(2): 161–75.

Pisanu B., Obolenskaya E.V., Baudry E., Lissovsky A.A., Chapuis J-L., 2013. Narrow phylogeographic origin of five introduced populations of the Siberian chipmunk *Tamias* (*Eutamias*) *sibiricus* (Laxmann, 1769) (Rodentia: Sciuridae) established in France. Biological Invasions 15: 1201–1207.

Rambaut A., Suchard M.A., Xie D., Drummond A.J., 2014. Tracer v1.6, Available from http://beast.bio.ed.ac.uk/Tracer.

Ronquist F., Teslenko M., van der Mark P., Ayres D.L., Darling A., Höhna S., Larget B., Liu L., Suchard M.A., Huelsenbeck J.P., 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Systematic biology 61(3): 539–42.

Sun J., 1982. Classification of the quaternary glacial stages in Jilin province. Acta Geologica Sinica 2: 174–88.

Associate Editor: P. Colangelo

Supplemental information

Additional Supplemental Information may be found in the online version of this article:

Supplement S1 List of specimens analysed in genetic study.

Supplement S2 List of specimens analysed in morphometric study.