

# Online Research @ Cardiff

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository: <http://orca.cf.ac.uk/105674/>

This is the author's version of a work that was submitted to / accepted for publication.

Citation for final published version:

Zhou, Xuming, Meng, Xuehong, Liu, Zhijin, Chang, Jiang, Wang, Boshi, Li, Mingzhou, Orozco Ter Wengel, Pablo, Tian, Shilin, Wen, Changlong, Wang, Ziming, Garber, Paul A, Pan, Huijuan, Ye, Xinping, Xiang, Zuofu, Bruford, Michael William, Edwards, Scott V., Cao, Yinchuan, Yu, Shuancang, Gao, Lianju, Cao, Zhisheng, Liu, Guangjian, Ren, Baoping, Shi, Fanglei, Peterfi, Zalan, Li, Dayong, Li, Baoguo, Jiang, Zhi, Li, Junsheng, Gladyshev, Vadim N., Li, Ruiqiang and Li, Ming 2016. Population genomics reveals low genetic diversity and adaptation to hypoxia in snub-nosed monkeys. *Molecular Biology and Evolution* 33 (10) , pp. 2670-2681. 10.1093/molbev/msw150 file

Publishers page: <http://dx.doi.org/10.1093/molbev/msw150>  
<<http://dx.doi.org/10.1093/molbev/msw150>>

Please note:

Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher's version if you wish to cite this paper.

This version is being made available in accordance with publisher policies. See <http://orca.cf.ac.uk/policies.html> for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.



*Title:*

## **Population genomics reveals low genetic diversity and adaptation to hypoxia in snub-nosed monkeys**

---

*Authors:*

Xuming Zhou<sup>1,2†\*</sup>, Xuehong Meng<sup>3†</sup>, Zhijin Liu<sup>1†</sup>, Jiang Chang<sup>4†</sup>, Boshi Wang<sup>1†</sup>, Mingzhou Li<sup>5†</sup>, Pablo Orozco-terWengel<sup>6</sup>, Shilin Tian<sup>3,5</sup>, Changlong Wen<sup>7</sup>, Ziming Wang<sup>1</sup>, Paul A Garber<sup>8,9</sup>, Huijuan Pan<sup>10</sup>, Xinping Ye<sup>11</sup>, Zuofu Xiang<sup>12</sup>, Michael W. Bruford<sup>6</sup>, Scott V. Edwards<sup>13</sup>, Yinchuan Cao<sup>3</sup>, Shuancang Yu<sup>7</sup>, Lianju Gao<sup>3</sup>, Zhisheng Cao<sup>3</sup>, Guangjian Liu<sup>1</sup>, Baoping Ren<sup>1</sup>, Fanglei Shi<sup>1</sup>, Zalan Peterfi<sup>2</sup>, Dayong Li<sup>14</sup>, Baoguo Li<sup>15</sup>, Zhi Jiang<sup>3</sup>, Junsheng Li<sup>4</sup>, Vadim N. Gladyshev<sup>2</sup>, Ruiqiang Li<sup>3\*</sup>, Ming Li<sup>1\*</sup>

*Author affiliation:*

<sup>1</sup> Key laboratory of Animal Ecology and Conservation Biology, Institute of Zoology, Chinese Academy of Sciences, Beijing 100101, China.

<sup>2</sup> Division of Genetics, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115, USA

<sup>3</sup> Novogene Bioinformatics Institute, Beijing 100083, China.

<sup>4</sup> State key laboratory of Environmental Criteria and Risk Assessment, Chinese Research Academy of Environmental Sciences, Beijing 100012, China.

<sup>5</sup> College of Animal Science and Technology, Sichuan Agricultural University, Ya'an 625014, China.

<sup>6</sup> Biodiversity and Sustainable Places Research Institute, Cardiff School of Biosciences, Cardiff University, Cardiff CF10 3T, UK.

<sup>7</sup> Beijing Vegetable Research Center (BVRC), Beijing Academy of Agricultural and Forestry Sciences, Beijing 100097, China

<sup>8</sup> Department of Anthropology, University of Illinois at Urbana-Champaign, Urbana-Champaign, Illinois 61801, USA.

<sup>9</sup> Program in Ecology and Evolutionary Biology, University of Illinois at Urbana-Champaign, Urbana-Champaign, Illinois 61801, USA.

<sup>10</sup> College of Nature Conservation, Beijing Forestry University, Beijing 100083, China.

<sup>11</sup> School of Life Sciences, Shaanxi Normal University, Xi'an 710062, China.

<sup>12</sup> College of Life Science and Technology, Central South University of Forestry and Technology, Changsha 410004, China.

<sup>13</sup> Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138, USA

<sup>14</sup> College of life Sciences, China West Normal University, Nanchong, Sichuan 637002, China

<sup>15</sup> College of Life Sciences, Northwest University, Xi'an 710069, China.

†These authors contributed equally to this work.

---

\* *Corresponding author:*

Dr. **Ming Li** ([lim@ioz.ac.cn](mailto:lim@ioz.ac.cn))

Dr. **Ruiqiang Li** ([lirq@novogene.cn](mailto:lirq@novogene.cn))

Dr. **Xuming Zhou** ([xmzhou0103@163.com](mailto:xmzhou0103@163.com))

## Abstract

2 Snub-nosed monkeys (genus *Rhinopithecus*) are a group of endangered colobines  
endemic to South Asia. Here, we re-sequenced the whole genomes of 38 snub-nosed  
4 monkeys representing four species within this genus. By conducting population  
genomic analyses, we observed an similar load of deleterious variation in snub-nosed  
6 monkeys living in both smaller and larger populations and found that genomic  
diversity was lower than that reported in other primates. Reconstruction of  
8 *Rhinopithecus* evolutionary history suggested that episodes of climatic variation over  
the past 2 million years, associated with glacial advances and retreats and population  
10 isolation, have shaped snub-nosed monkey demography and evolution. We further  
identified several hypoxia-related genes under selection in *R. bieti* (black snub-nosed  
12 monkey), a species that exploits habitats higher than any other nonhuman primate.  
These results provide the first detailed and comprehensive genomic insights into  
14 genetic diversity, demography, genetic burden and adaptation in this radiation of  
endangered primates.

16

---

18 *Keywords:*

20 Snub-nosed monkeys | Population genomics | Demographics | Genetic diversity |  
Genetic adaptation

22

## Introduction

2        Snub-nosed monkeys (*Rhinopithecus* spp.) represent a diverse radiation of  
endangered leaf-eating monkeys (colobines or foregut fermenters) that exploit a diet  
4        composed principally of leaves, seeds, bark and lichen (IUCN 2011). These primates  
were once widely distributed across Asia but now are limited to mountain forests in  
6        south China, Myanmar, and Vietnam (Kirkpatrick 1998; Quan and Xie 2002). Five  
extant species with different census population sizes are recognized: *R. roxellana*, the  
8        Sichuan or golden snub-nosed monkey, ~ 25,000 individuals; *R. brelichi*, the gray  
snub-nosed monkey, ~800 individuals; *R. bieti*, the black or Yunnan snub-nosed  
10        monkey, ~2,000 individuals; *R. avunculus*, the Tonkin snub-nosed monkey, ~200  
individuals; and *R. strykeri*, the Myanmar snub-nosed monkey, ~300 individuals (**fig.**  
12        **1a** and **supplementary fig. S1**) (Quan and Xie 2002). Three species, the golden, gray  
and black, are endemic to China, while a fourth species, the Tonkin snub-nosed  
14        monkey is restricted to northern Vietnam. The Myanmar snub-nosed monkey, which  
was discovered from Myanmar (Burma) in 2011, has now also been reported to range  
16        into southwestern China (Nujiang prefecture) (Geissmann et al. 2011). Genetic  
evidence suggests that *Rhinopithecus* represents the most basal lineage of Asian  
18        colobines, however, evolutionary relationships within the genus *Rhinopithecus*,  
remain unclear (Liedigk et al. 2012). During the Pleistocene changes in rates of gene  
20        flow, population isolation, and periods of population expansion and population  
reduction, associated with major climatic (glacial) and geological events (uplifting of  
22        the Tibetan plateau), make this clade an instructive model for investigating questions

concerning evolutionary genetics and adaptation.

2        Moreover, the black snub-nosed monkey which is confined to high altitude  
forests from 3,400 and 4,600 meters above sea level within a narrow region of the  
4 Tibetan plateau between the Yangtze and Mekong rivers (98°37' to 99°41'E, 26°14' to  
29°20'N) has the highest elevational range of any primate species (Long et al. 1994,  
6 1996; Li et al. 2002). Across their range, *R. bieti* is exposed to significant challenges  
associated with hypothermia, hypoxia and locating calorie-rich food during winter  
8 (2-3 months) when temperatures commonly fall below -10° C. In response to these  
challenges, *R. bieti* exhibits a set of derived anatomical and behavioral traits such as  
10 increased body mass, predominantly black pelage and a behavioral pattern  
characterized by travel to locations above the tree line and preferred use of southern  
12 mountain slopes to increase exposure to solar radiation (Quan and Xie 2002)  
(**supplementary table S1**). It has been proposed that the tectonic migration of the  
14 Hengduan Mountains, located in southwest China and northern Myanmar, and  
concomitant environmental perturbations associated with periods of glacial expansion  
16 and glacial retreat during the last 1.9-2.8 mya (million years ago) are likely to have  
led to population isolation events in snub-nosed monkeys (Pan and Jablonski 1987;  
18 Liedigk et al. 2012). However, the genetic implications of population isolation and the  
genomic basis of adaptations of black snub-nosed monkeys to high altitude  
20 environments have not been determined (Yu et al. 2011).

Here, we sequenced 38 wild snub-nosed monkey genomes (four *Rhinopithecus*  
22 species). By conducting extensive species and population comparisons, we have

reconstructed the phylogenetic relationships, demographic history and species ancient  
2 distributions, as well as have determined the effects of changes in population size and  
the relative frequency of deleterious mutations across the genus. We also examined  
4 selected regions of the genome that indicate genetic changes that facilitated the  
exploitation of high altitude forests and recent population divergence among  
6 snub-nosed monkey species.

## 8 **Results and Discussion**

### **Phylogeny and genetic diversity**

10 Genome mapping of 42 individuals (27 golden, four gray, two Myanmar and nine  
black snub-nosed monkeys) across the snub-nosed monkey's 3.05 Gb genome  
12 resulted in an average of 98.6 % sequencing coverage with ~10.39-fold and  
~25.38-fold coverage depth for 35 and seven individuals, respectively (Zhou et al.  
14 2014) (**supplementary tables S2-S7** and **figs. S2-S4**). The neighbor-joining (NJ) tree  
of snub-nosed monkeys features two clusters, with golden and gray snub-nosed  
16 monkeys, both northern species, grouping together and black and Myanmar  
snub-nosed monkeys ('Himalayan' species) forming the other cluster (**fig. 1b** and  
18 **supplementary fig. S5**). The principal components analysis (PCA) (Patterson et al.  
2006) also recovered these groupings, with the first and second eigenvectors separate  
20 *R. roxellana*, *R. brelichi* and *R. strykeri/R. bieti* (**fig. 1c** and **supplementary table S8**).  
Within golden snub-nosed monkeys, only individuals from the Shennongjia (SNJ)  
22 mountains clustered together and formed a monophyletic group distinct from the  
remaining populations (Sichuan/Gansu populations, including Minshan [MS] and

Qionglai [QLA] mountains, and Qinling mountains [QL], termed as non-SNJ cluster)  
2 **(fig. 1b and supplementary fig. S5)**. Both PCA and *frappe* analyses (Tang et al. 2005)  
recapitulated these findings. In particular, the first eigenvector separated the two  
4 genetic lineages (SNJ and non-SNJ clusters) (variance explained = 5.98%,  
Tracy-Widom  $P = 1.09 \times 10^{-8}$ ), but the second eigenvector failed to separate the  
6 Qinling population from the other populations (variance explained = 3.65%,  
Tracy-Widom  $P = 1.09 \times 10^{-8}$ ) **(supplementary fig. S6 and table S9)**. Cross validation  
8 analysis using ADMIXTURE (Alexander et al. 2009) offers support for two clusters  
**(supplementary figs. S7)**. Using *frappe*, we found a division between the SNJ and  
10 the non-SNJ (QL-MS-QLA) when  $K = 2$ , with slight shared ancestry sequences  
between these two groups **(fig. 1d and supplementary fig. S8)**. Although three  
12 distinct clusters were found at  $K = 3$ , this result was not corroborated by  
cross-validation and was inconsistent with the population's biogeographic distribution,  
14 i.e. Qinling (QL), Minshan-Qionglai (MS-QLA), and Shennongjia (SNJ) mountains,  
indicating strong gene flow between the Qinling (QL) and Minshan-Qionglai  
16 (MS-QLA) populations.

Compared to other primate taxa, the heterozygosity of snub-nosed monkeys,  
18 defined as the number of heterozygous SNPs divided by total callable sequence in  
each individual, at 0.015%-0.068%, was found to be generally lower than in humans  
20 (0.08%-0.12%) and great apes (0.065%-0.178%) (Prado-Martinez et al. 2013), and  
was similar to that reported in endangered carnivores, such as the Amur Tiger, African  
22 Lion and Snow Leopard (0.048%-0.0073%) (Cho et al. 2013) **(supplementary figs.**



**S9-S10, supplementary table S10**). In contrast, the heterozygous SNP rate of the  
2 snub-nosed monkeys was higher than the Yangtze River dolphin (0.0121%) (Zhou et  
al. 2013) and Chinese alligator (0.015%) (Wan et al. 2013), both of which are  
4 considered critically endangered. The low genetic diversity of snub-nosed monkeys is  
consistent with their small extant population sizes, which may be related to an  
6 evolutionary history of environmental change associated with severe forest  
fragmentation across isolated mountain terrain (Pan and Jablonski 1987; Quan and  
8 Xie 2002). Within snub-nosed monkeys, the heterozygosity of the two northern  
species (golden:  $0.041\% \pm 7.16 \times 10^{-5}$  [mead  $\pm$  standard deviations] and gray:  
10  $0.068\% \pm 5.40 \times 10^{-5}$ ) was higher than that of the Himalayan species (black:  
 $0.032\% \pm 1.02 \times 10^{-5}$  and Myanmar:  $0.015\% \pm 2.05 \times 10^{-5}$ ) ( $P < 0.05$ ) (**supplementary**  
12 **fig. S10**), suggesting differences in population histories, with smaller population sizes  
in the Himalayan species.

14

### **Demographic history and correlation with climate change**

16 To explore the demographic histories of snub-nosed monkeys, we inferred historical  
population size changes using the PSMC model (Li and Durbin 2011) based on the  
18 distribution of heterozygous sites across the genome. We found that individuals from  
the same species showed a highly similar pattern of historical fluctuation in effective  
20 population sizes ( $N_e$ ) (**fig. 2a** and **supplementary fig. S11**). However, the historical  
trends in  $N_e$  for black, gray and Myanmar snub-nosed monkeys showed a very  
22 different pattern from those of golden snub-nosed monkeys, with continuing declines

of  $N_e$  in the first three species and two bottlenecks and two expansions in golden  
2 snub-nosed monkeys (Zhou et al. 2014). Within golden snub-nosed monkeys,  
individuals from the same habitats or mountain region showed a similar pattern of  
4 historical fluctuation of  $N_e$  with very few exceptions (two individuals), possibly due to  
migration or hybridization that occurred in these populations (**fig. 2b**). The  
6 post-bottleneck population expansion after  $10^5$  years was not found in individuals  
from the Shennongjia (SNJ) and southern Qinling mountain populations, consistent  
8 with their reduction in heterozygosity when compared to other populations.

For golden snub-nosed monkeys, we also inferred the time of the divergence event  
10 between the Shennongjia (SNJ) and Qinling-Minshan-Qionglai (non-SNJ)  
populations, as well as recent demographic fluctuations in population size. Using the  
12 joint site frequency spectrum between SNJ and non-SNJ and the diffusion  
approximations method (Gutenkunst et al. 2009), the IM-pre model, identified the  
14 split between SNJ and non-SNJ populations around 4.60 thousand years ago (Upper  
Pleistocene and Holocene) (**fig. 3a, supplementary table S11 and fig. S12**), similar to  
16 the timing of separation revealed by the SM model (split with migration) and PSMC  
analyses. Relative cross coalescence rates (CCR) estimated by a MSMC (multiple  
18 sequentially Markovian coalescent) model (Schiffels and Durbin 2014) revealed that  
the complete split between the SNJ and non-SNJ populations occurred 10 kya  
20 (thousand years ago), near the end of the last glacial maximum (LGM, 25-15 kya) (**fig.**  
**3b**). In the early stages of the LGM (~20 kya), substantial alpine glaciations (for  
22 example, Gongga glacial II in Qinghai-Tibetan Plateau, 15-20 kya) (Zheng et al. 2002)

2 resulted in extensive loss and fragmentation of forested habitats across the snub-nosed  
3 monkey distribution, likely leading to population isolation. Population studies using  
4 allozyme loci polymorphisms (Li et al. 2003) also proposed a recent bottleneck for  
5 golden snub-nosed monkeys, in that case within the last 15,000 years.

6 During the Pleistocene, climatic fluctuations and periodic uplifting of the  
7 Qinghai-Tibetan Plateau had important effects on biodiversity and changes in the  
8 geographic distribution of mammals in southwest China (He and Jiang 2014). For  
9 example, it has been suggested that the LIG (Last Interglacial Maximum) was an  
10 unfavorable period for mole shrews (*Anourosorex squamipes*) because of increased  
11 seasonality in temperature extremes. During this period, mole shrews survived in  
12 isolated interglacial refugia mainly located in the southern subregion and then rapidly  
13 expanded during the LGM (He et al. 2016). The amount and distribution of suitable  
14 habitat for snub-nosed monkeys also is likely to have been affected by these climatic  
15 events. To assess the role of climate using species distribution models and  
16 palaeoclimatic data on temperature and precipitation, we estimated suitable habitat of  
17 golden snub-nosed monkeys during the time periods of 120-140 kya (LIG), 21 kya  
18 (LGM), 6 kya (Mid-Holocene), and the present (**fig. 3c** and **supplementary table**  
19 **S12**). The model revealed that suitable habitat during the LIG was more widely  
20 available than during both the colder LGM and warmer mid Holocene. Specifically,  
21 climate change during the LIG period resulted in an expansion of suitable habitat in  
22 northward areas (**fig. 3c**). This result parallels patterns of population size change in  
23 golden snub-nosed monkeys estimated from the PSMC model, which suggested an

expansion in population size beginning at 0.1 mya and a bottleneck occurring at the  
2 LGM. Strikingly, the suitable habitat of golden snub-nosed monkeys expanded toward  
warmer southern areas during the LGM; by contrast, the fossil record implied that  
4 population have moved into colder northern areas during the mid-Holocene (**fig. 3c**).  
Given that seasonal extremes in temperature have a significant effect on forest  
6 productivity (Ni et al. 2001), climate change likely dramatically contributed to the  
shifts in the geographic distribution of golden snub-nosed monkey populations from  
8 the LIG to the LGM. In addition, combined with genetic evidence of population  
history, these results are consistent with severe population isolation of golden  
10 snub-nosed monkeys during the LGM, followed by subsequent *in situ* survival in their  
extant ranges.

12

### **Signatures of deleterious variation and positive selection**

14 As demonstrated above and reported in a previous study (Zhou et al. 2014),  
levels of heterozygosity and demographic trajectories varied substantially among  
16 snub-nosed monkey species and populations. Genome-wide linkage disequilibrium  
(LD) also varied markedly within *Rhinopithecus*; specifically, the average distance  
18 over which LD decayed to half of its maximum value in gray (>100 kb) and black  
snub-nosed monkeys (~66.7 kb) was substantially longer than that of golden  
20 snub-nosed monkeys (~17.5 kb) (**fig. 4a**). Because deleterious mutations should be  
removed more efficiently in larger populations (Kimura 1983; Charlesworth 2009;  
22 Xue et al. 2015), we examined the genetic load of deleterious variation in different

snub-nosed monkeys. We found a similar number of derived missense and  
2 loss-of-function (LoF) alleles among golden, gray and black snub-nosed monkeys (**fig.**  
**4b**). We did not observe a larger number of derived missense alleles in gray  
4 snub-nosed monkeys and black snub-nosed monkeys compared to golden snub-nosed  
monkeys (**fig. 4b**). Thus, despite their larger census population size, selection was not  
6 more effective at removing moderately deleterious mutations in golden snub-nosed  
monkeys, a pattern also seen in humans (Do et al. 2015). This may either be because  
8 golden snub-nosed populations were still too small for selection to effectively remove  
deleterious mutations or that ancient bottlenecks had an impact on purifying and  
10 background selection (Lynch 2006). We found that black snub-nosed monkeys had  
more LoF variants than golden snub-nosed monkeys. However, the ratio of  
12 homozygous LoF genotypes in black snub-nosed monkeys is higher than in golden  
snub-nosed monkeys (**fig. 4c**), which is consistent with the assumption that LoF  
14 variants are more likely to be strongly deleterious and are more frequently exposed in  
a homozygous state (Glémin 2003). We also found LoF variants differing significantly  
16 in homozygous and heterozygous tracts (**fig. 4d**), which is consistent with that  
recessive LoF variants are likely more deleterious in the homozygous rather than the  
18 heterozygous state, where they are masked.

Finally, we looked at variation in genes or genomic regions that were possibly  
20 under selection and adaptation in snub-nosed monkeys. We first focused on black  
snub-nosed monkeys. This species inhabits the highest known mountain elevations  
22 among *Rhinopithecus* and among non-human primates generally (Long et al. 1996; Li

et al. 2002). Given that the divergence between black snub-nosed monkeys and other  
2 snub-nosed monkeys was relatively recent, we used a population genetic methodology  
to identify selective regions and adaptations. Using fixation ( $F_{ST}$ ) and genetic  
4 diversity ( $\theta_{\pi}$ ), we identified 19 genes, containing a total of 2.16 Mb of genomic  
sequence, with maximum  $F_{ST}$  values (top 5%) and minimum  $\theta_{\pi 1}/\theta_{\pi 2}$  (top 5%) in black  
6 snub-nosed monkeys (**fig. 5a** and **supplementary table S13**). Of these potentially  
selected genes, *ADAM9*, which encodes a cell-surface membrane glycoprotein, is a  
8 member of the ADAM (a disintegrin and metalloprotease domain) family (**fig. 5b**).  
Members of this family have been implicated in a variety of biological processes  
10 involving cell-cell and cell-matrix interactions, muscle development and neurogenesis  
(Primakoff and Myles 2000). Several ADAM family members have been identified as  
12 positively selected genes in Tibetan yak and ground tit, compared to their lowland  
relatives (Qiu et al. 2012; Cai et al. 2013). With regard to *ADAM9*, previous study has  
14 showed that the mRNA and protein expression levels of this gene are elevated by the  
regulation of reactive oxygen species under hypoxic stress conditions in prostate  
16 cancer cells (Shen et al. 2013). In addition, hypoxia (1% oxygen) can also induce the  
expression of *ADAM9* in gastric cancer cells (Kim et al. 2014). Interestingly, *ADAM9*  
18 has also been identified as one of the positively selected genes in Tibetan chicken  
(Zhang et al. 2016), suggesting a common strategy for hypoxia adaption has been  
20 utilized in both two species. Also, it is well documented that that an unfolded protein  
response (UPR) is involved in  $O_2$ -sensitive signaling pathways caused by hypoxia in  
22 cancer (Wouters et al. 2008) and moderate to severe hypoxia will affect the

pulmonary vascular remodeling and cardiac muscle function (Julian 2007).

2 Accordingly, we found another two genes positively selected in the black snub-nosed  
monkey (**supplementary tables S14 and S15**) that are associated with such pathways,  
4 i.e., Ubiquitin mediated proteolysis (*HERC1*) and Cardiac muscle contraction  
(*SLC9A6*).

6 To examine the local adaptations of golden snub-nosed monkey populations, we  
measured selected regions across genome between the SNJ and non-SNJ golden  
8 snub-nosed monkey populations, which are geographically and genetically distinct.  
We found twice as many genomic regions were under positive selection in the SNJ  
10 population (161 genes containing 17.86 Mb of sequence, or 0.625% of the genome) as  
there were in the non-SNJ populations (64 genes containing 7.77 Mb, or 0.272% of  
12 the genome) (**fig. 5c and supplementary table S14**). Genes embedded in selected  
regions of the SNJ and non-SNJ populations were significantly associated with lipid  
14 metabolism process, such as negative regulation of very-low-density lipoprotein  
particle remodeling (2 genes,  $P = 2.0 \times 10^{-4}$ , Fisher exact test), steroid metabolic  
16 process (8 genes,  $P = 2.5 \times 10^{-3}$ , Fisher exact test) and cellular lipid catabolic process  
(4 genes,  $P = 4.4 \times 10^{-3}$ , Fisher exact test) (**supplementary tables S16-S17**).  
18 Moreover, several genes (including *GTF2I*, *TRIPL1*, *DTNB*, *CATSPER3*, *PPIL6*,  
*RBBP8*) that under selection in SNJ populations associated with adult height and  
20 organ development in humans. This may be related to the fact that golden snub-nosed  
monkeys living in Shennongjia mountain differ from other golden snub-nosed  
22 monkey populations in possessing extremely reduced or absent nasal bones and fusion

of the premaxilla superior to the piriform aperture, a longer tail and longer forelimbs  
2 (supplementary tables S18-S19) (Kirkpatrick 1998; Wang et al. 1998; Mittermeier et  
al. 2013). Interestingly, one of those genes (i.e., *GTF2I*) is located in the causal region  
4 (7q11.23) of the Williams-Beuren syndrome, a microdeletion or  
contiguous-gene-deletion disorder, resulting in a distinctive, ‘elfin’ facial appearance,  
6 along with a flat nasal bridge (Francke 1999; Pober 2010). Moreover, given that some  
traits characteristic of the SNJ population, for example their long tail, are also found  
8 in other species of snub-nosed monkeys, we speculate that the natural selection on  
genes corresponding to these common characters has relaxed in the non-SNJ  
10 population of golden snub-nosed monkeys, leading to an increase in the frequency of  
this polymorphism.

12

## Conclusions

14 In the present study, we first provide a comprehensive analysis of genomic  
polymorphism within the genus *Rhinopithecus* and then reconstruct phylogenetic  
16 relationships across snub-nosed monkey genomes. Our results indicate that  
biogeography and climate change critically shaped the demographic history of  
18 snub-nosed monkeys. We found low genetic diversity within each species of this  
primate clade, likely the result of historical bottlenecks caused by glacial and tectonic  
20 events the Pleistocene. This is consistent with biogeographical and fossil information  
indicating that snub-nosed monkeys were once widely distributed throughout China  
22 and parts of South Asia but that currently all five endangered species are confined to



isolated mountain forests. We also found that a larger census population size of golden  
2 snub-nosed monkeys did not result in a reduction of burden of deleterious mutations  
compared to other snub-nosed monkeys. Given the limited remaining population size  
4 of each snub-nosed monkey species, it is likely that continued habitat loss and  
fragmentation resulting from climate change and human interference will lead to  
6 further population isolation and decline.

We found evidence of positive selection in the genome of the black snub-nosed  
8 monkeys for several genes that potentially facilitate an adaptive response to hypoxia.  
We do caution, however, that genomic variation can be affected by many evolutionary  
10 processes and that without functional experiments, it is difficult to distinguish natural  
selection from genetic drift on specific loci, including candidate genes listed in the  
12 present study. This study highlights the value of genomic data in assessing genetic  
diversity, measuring the power of natural selection, and identifying local adaptations.

14

## **Materials and Methods**

### **16 Sampling and Sequencing**

This study was approved by the Animal Ethics Committee of the Institute of Zoology,  
18 Chinese Academy of Sciences. The muscle and dried skin samples were collected  
from individuals who died of natural causes in Nature Reserves. The blood samples  
20 were collected from injured individuals who were treated by veterinarians from the  
rescue station of the nature reserves or zoos. The procedure of blood collection was in  
22 strict accordance with the Animal Ethics Procedures and Guidelines of the People's

Republic of China. Our sample included 26 golden, three gray, one Myanmar and  
2 eight black snub-nosed monkeys. For golden snub-nosed monkeys, our samples  
included each of the five extant mountain populations, i.e., Sichuan/Gansu (GS or SG),  
4 Minshan (MS), Qinling mountains (QL), Qionglai mountain (QLA) and Shennongjia  
(SNJ). Individuals from the same location were avoided in order to reduce the  
6 potential effect of kinship on genetic diversity. Genomic DNA of 38 snub-nosed  
monkeys was extracted from blood or muscle using a Gentra Puregene Blood Kit  
8 (Qiagen) according to the manufacturer's instructions. Sequencing libraries were  
constructed with insert size 500 bp or 300 bp, and then sequenced on an Illumina  
10 HiSeq 2000 platform using paired-end (PE) 100 bp reads. The average sequencing  
depth was  $\sim 30\times$  per sample for the 3 golden snub-nosed monkey individuals and  
12  $\sim 10\times$  per sample for the other 35 individuals (**supplementary table S3**).  
Resequencing reads of snub-nosed monkeys in this study have been deposited in  
14 NCBI Sequence Read Archive (SRA) under accession SRP051679. Four previously  
published snub-nosed monkey genome data were retrieved from NCBI under  
16 accession SRP033389.

### 18 **Read mapping and variant calling**

Adapter and low quality reads (e.g. Reads with  $\geq 10\%$  unidentified nucleotides (N);  
20 reads with  $> 10$  nt aligned to the adapter, allowing  $\leq 10\%$  mismatches;  
bases having phred quality  $< 30$ ) were removed from the raw data using inhouse Perl  
22 script. The remaining high quality reads were aligned to the golden snub-nosed

1 monkey (GSM) genome (Zhou et al. 2014) using Burrows-Wheeler algorithm  
2 implemented in BWA software (Li and Durbin 2009) with parameters ‘BWA aln -o 1  
-e 10’. In order to reduce mismatch generated by PCR amplification before  
4 sequencing, duplicated reads were removed using Samtools (version 1.2) (Li et al.  
2009).

6 We performed variant calling on population-scale for all samples using Samtools  
(version 1.2) mpileup with parameters ‘-q 1 -C 50 -S -D -m 2 -F 0.002’ and using  
8 bcftools (with parameters ‘view -cvg’) to view variants (Li et al. 2011). The following  
filtering steps were applied in order to obtain high quality variants: 1) minimum read  
10 depth was 8; 2) maximum read depth was 30; 3) minimum root-mean-square (RMS)  
mapping quality was 20; and 4) SNP within 3 bp around a mapping gap to be filtered.  
12 Other parameters in ‘bcftools/vcfutils.pl varFilter’ are default. Genotype quality less  
than 20 and variants with allele frequency less than 10% of the 42 samples also were  
14 filtered. To assess sequence accuracy, we genotyped 100 randomly selected variants in  
38 snub-nosed monkeys using a PCR-based KASP (Kbioscience Allele-specific  
16 Polymorphism) genotyping approach. Of the 2,330 loci successfully genotyped, 95.45%  
SNPs (2,224 loci) were consistent with the SNPs identified from resequencing,  
18 demonstrating the high quality and reliability of our variant calls.

## 20 **Heterozygosity analysis**

Genetic diversity was assessed using heterozygosity in each individual.  
22 Heterozygosity ( $h$ ) for each individual was calculated as the number of heterozygous

sites divided by the total number of callable sites across the whole genome.

2

### **Phylogenetics, population structure and principal components analysis**

4 Phylogenetic trees were inferred using neighbor joining (NJ) method and  
implemented in TreeBest (<http://treesoft.sourceforge.net/treebest.shtml>) with 500  
6 bootstraps. Population structure within golden snub-nosed monkeys was inferred by  
ADMIXTURE (Alexander et al. 2009) with 10-fold cross-validation (Alexander and  
8 Lange 2011). The ancestral sequences of snub-nosed monkeys were inferred using the  
program *frappe* (Tang et al. 2005), which implements an expectation-maximization  
10 algorithm. We did not assume any prior information about their structures in these  
analyses and predefined the number of genetic clusters  $K$  from 2 to 5. The maximum  
12 iteration of the expectation-maximization algorithm was set to 10,000.

The principal component analysis (PCA) of 42 samples was conducted using  
14 EIGENSOFT3.0 (Patterson et al. 2006) and the significance of eigenvectors were  
determined with the Tracey-Widom test in the EIGENSOFT 3.0 (Patterson et al.  
16 2006). The numbers of individuals differ greatly between species and this can  
potentially influence the PCA. Therefore, we repeated the PCA analyses by randomly  
18 selecting 9 and 4 golden snub-nosed monkey individuals, respectively (five replicates  
for each). The repeated analyses supported the original results.

20

### **Demographic history reconstruction**

22 Changes in effective population size were inferred from a haploid pairwise

sequentially Markovian coalescence (PSMC) model (Li and Durbin 2011) with  
2 parameters '-g 10 -R -u 1.36e-8 -m 100 -X 2000000 -p "4+25\*2+4+6"'. Several  
gradient levels for the parameter '-p' were set (from 26 free interval parameters with  
4 60 atomic time intervals to 30 free interval parameters with 68 atomic time intervals)  
to test the performance of PSMC model in golden snub-nosed monkeys. The  
6 estimated TMRCA is in units of  $2N_0$  time and the mean generation time was set at 10  
years and  $\mu$  (unit: nt/generation) was estimated as  $1.36 \times 10^{-8}$  according to our  
8 previous estimation (Zhou et al. 2014). Given that the reduced genetic variation in X  
chromosomes would affect the inference of population genetic information, we  
10 excluded SNPs located in X chromosomal regions and kept autosomal SNPs for  
PSMC analyses and later MSMC analyses. We aligned the scaffolds of the golden  
12 snub-nosed assembly to the human (genome hg19) using LASTZ (Harris 2007) and  
excluded the longest scaffold that showed a syntenic relationship with the human X  
14 chromosome.

The recent demographic history of golden snub-nosed monkeys was inferred using  
16  $\partial a \partial i$  (Gutenkunst et al. 2009). Many divergence models between two genetic  
populations, including the Standard neutral model (SNM), the Instantaneous size  
18 change followed by exponential growth with and without population split model, Split  
into two populations with migration (SM) model, Isolation-with-migration model with  
20 exponential pop growth (IM) model, and Isolation-with-migration model with  
exponential pop growth and a size change prior to split (IM-pre) were considered. The  
22 IM\_pre and SM models were chosen as the optimal models by comparing the

optimized model likelihood. The ancestral population size ( $N_A$ ) was estimated on the  
2 basis of the optimized parameters and the estimated mutation rate. Nonparametric  
bootstrapping was performed 100 times to determine the variance of each parameter  
4 using equal numbers of loci sampled without replacement from the original data set.  
The relative cross coalescence rates (CCR) for the Shennongjia (SNJ) and  
6 Qinling-Minshan-Qionglai (non-SNJ) populations divergence within Golden  
snub-nosed monkeys was estimated under MSMC model (Schiffels and Durbin 2014).

8

### **Species distribution modeling**

10 Geographical information on species presence ( $n=127$ ) was obtained from the  
State and local records in China since 1616 and published in previous studies (Quan  
12 and Xie 2001; Li et al. 2002, 2003b). A total of 25 fossil localities were drawn from  
the literature (Pan 1995; Hijmans et al. 2005) to collate the modelled past species  
14 distribution areas. Generally, the fossil localities represent a subset, rather than an  
exhaustive search of the literature available, and modelled ranges consequently  
16 represent a subset of the entire past distribution of the species.

We developed the present-day Species Distribution Models (SDMs) based on  
18 current climate data, and then projected the SDMs to three periods: mid-Holocene  
(~6000 years BP), last glacial maximum (LGM, ~21000 years BP), and last  
20 interglacial (LIG, ~120000 - 140000 years BP). Nineteen bio-climatic variables used  
for model construction and projections were obtained from the WorldClim 1.4

database at a spatial resolution of 2.5 arc-minutes (Hijmans et al. 2005). The  
2 bio-climatic layers for the present-day were based on spatially interpolated values of  
temperature and precipitation gathered from weather stations around the world from  
4 1950-2000 (Pan 1995), whereas the variables for both the period mid-Holocene and  
LGM were generated from the climate reconstructions based on CCSM4  
6 (<http://www.cesm.ucar.edu/>) (Collins et al. 2006) and MIROC-ESM (Otto-Bliesner et  
al. 2006), and the variables for LIG were downscaled from simulations with Global  
8 Climate Models (GCMs) (Otto-Bliesner et al. 2006).

SDMs were run at each grid size in R-based BIOMOD for the golden snub-nosed  
10 monkey, and each probability distribution was predicted across the study area. We  
divided the occurrence data into two subsets: the first, a random sample from 70% of  
12 the total database was used to calibrate (train) the models, whereas the second,  
comprising the remaining data, was used to evaluate (test) model predictions (Fielding  
14 and Bell 1997). The eight methods included in the testing framework were: artificial  
neural networks (ANN) (Venables and Ripley 2013); classification tree analysis (CTA)  
16 (Breiman et al. 1984); random forest (RF) (Breiman 2011); generalized linear models  
(GLM) (McCullagh and Nelder 1989); generalized boosting models (GBM)  
18 (Ridgeway 1999); generalized additive models (GAM) (Hastie and Tibshirani 1990),  
mixture discriminate analysis (MDA) (Hastie 1994), and Maximum Entropy (MaxEnt)  
20 (Phillips et al. 2006). To assess the quality of predictions, we randomly split the  
dataset using 70% of the data to calibrate the models and the remaining 30% for

evaluation (Barbet-Massin et al. 2012). Since only occurrence records were available  
2 for the species, we used 1000 randomly generated pseudo-absence points in  
constructing SDMs (Elith et al. 2006). Model performance at each grid size was  
4 assessed using Cohen's kappa statistic (Cohen 1990), the area under the curve (AUC)  
of receiver operating characteristics (ROC) (Fielding and Bell 1997), and true skills  
6 statistics (TSS) (Allouche et al. 2006). From the output of all eight species  
distribution models, we generated ensemble projections based on the weighted mean  
8 of predicted probabilities of presence (Thuiller 2003).

## 10 **Linkage-disequilibrium (LD) analyses**

To assess linkage disequilibrium (LD) pattern in snub-nosed monkey species, the  
12 coefficient of determination between any two loci in each population was calculated  
using Haploview (Barrett et al. 2005) with parameters '-maxdistance 500 -dprime  
14 -minGeno 0.75 -minMAF 0.05 -hwcutoff 0.001'. For plotting LD decay curves,  
variants with a highly similar physical distance (500 kb) were merged into groups  
16 and calculation of average  $r^2$  for each group. Considering the samples were not equal  
in three snub-nosed monkeys, we randomly chose 9 and 4 golden snub-nosed  
18 monkeys, respectively and repeated the LD analyses 10 times.

## 20 **Genetic load and purging of deleterious mutations**

The genetic load and the power of removing deleterious mutations were  
22 measured using the methods proposed for gorillas (Xue et al. 2015). We used



ANNOVAR (Wang et al. 2010) to annotate and classify coding sequence variants based on the following sequence ontology terms: LoF variants: stop\_gain, frameshift\_variant, inframe\_insertion, inframe\_deletion; Missense variants: missense\_variant; Synonymous variants: synonymous\_variant. We determined the ancestral allele at each position on the basis of comparison to the macaque genome (MMUL 1.0). In order to estimate the relative excess of missense and deficit of LoF variants in *R. bieti* (black snub-nosed monkey) and *R. brelichi* (gray snub-nosed monkey) compared to *R. roxellana* (golden snub-nosed monkey): for each pairwise comparison, we first calculated the number of alleles ('derived alleles') found in one population (for example, population A), but not the other (for example, population B). Then, the relative number of derived alleles in each comparison were determined by  $R_{A,B}(C) = L_{A,B}(C) / L_{B,A}(C)$ , where  $L_{A,B}(C)$  was estimated from:

$$L_{A,B}(C) = \frac{\sum_{i \in C} f_i^A (1 - f_i^B)}{\sum_{j \in I} f_j^A (1 - f_j^B)}$$

When we estimated  $L_{A,B}(C)$ , at each site  $i$ , the observed derived allele frequency in population A was defined as  $f^A_i = d^A_i / n^A_i$ , where  $n^A_i$  is the total number of alleles and  $d^A_i$  is the number of derived alleles identified in population A. The  $f^B_i$  was similarly defined in population B. In the above formula, C and I indicate category of protein-coding sites and a set of intergenic sites respectively. Totally, 3.66 Mb, 3.53Mb and 4.88Mb intergenic sites were selected in black, gray and golden snub-nosed monkeys, respectively.

To investigate the ratio of homozygous LoF genotypes in each subspecies of

sub-nosed monkeys, we also counted the number of LoF segregating sites ( $n_{\text{hom}}$ )  
2 where at least one individual is homozygous for the derived allele in each subspecies.  
In total, we found 296 such sites in *R. bieti*, 181 in eastern *R. brelichi* and 271 in *R.*  
4 *roxellana*. In order to normalize the numbers of samples in each species, we also  
counted  $n_{\text{hom}}$  in each of 10,000 random samples of segregating synonymous sites. The  
6 LoF mutations occurring as homozygous tracts within individuals were normalized by  
the rate of synonymous variant sites in the same regions.

8

### **Selective signals in black and golden snub-nosed monkeys**

10 The polymorphism levels in each group were quantified by pairwise nucleotide  
diversity ( $\theta_{\pi}$ ) (Tajima 1983, equation A3) and the genetic differentiation in two  
12 populations was quantified by pairwise  $F_{\text{ST}}$  ( $=[H_{\text{T}} - H_{\text{S}}] / H_{\text{T}}$ , Weir and Cockerham  
1984). Both  $\theta_{\pi}$  and  $F_{\text{ST}}$  were calculated by a sliding window method (100 kb windows  
14 sliding in 10 kb steps). Variants were filtered when the minor allele frequency was  
less than 5% and the missing genotypes frequency was more than 50%. For  
16 comparing groups (groups 1 and 2), the regions with maximum  $F_{\text{ST}}$  values (top 5%)  
and minimum  $\theta_{\pi 1} / \theta_{\pi 2}$  (top 5%) were identified as selected regions for group1, and vice  
18 versa. We used DAVID (Huang et al. 2009) to perform GO term enrichment analysis  
on the candidate selective gene located in selected regions (DAVID analysis was run  
20 June 25, 2016).

### **22 Supplementary Material**

Supplementary figures S1–S12, tables S1–S19, and references are available at

2 Molecular Biology and Evolution online (<http://www.mbe.oxfordjournals.org/>).

#### 4 **Acknowledgements**

We greatly thank the Shennongjia National Nature Reserve, Baimaxueshan National  
6 Nature Reserve, Gaoligongshan National Nature Reserve (Nujiang), Beijing Zoo,  
Beijing Wildlife Park and Shanghai Wildlife Park for permission to sample. Thanks  
8 Dr. Wen Xiao in Dali University for providing the samples of Myanmar snub-nosed  
monkey, and Dr. Nguyen Van Truong in FFI Vietnam for providing the photo of  
10 Tonkin snub-nosed monkey. We wish to acknowledge the contributions of Chrissie  
McKenney, Sara Garber and Jennifer Garber.

12 This project was supported by National Key Program (2016YFC0503200 and  
2013BAD03B02) from Ministry of Science and Technology of China, the Natural  
14 Science Foundation of China (No. 31270420, 41371071 and 31130061), Hubei  
Province Key Laboratory of Conservation Biology of Shennongjia Golden Monkey  
16 and State Forestry Administration of China to Ming Li.

## References

- 2 Alexander DH, Lange K. 2011. Enhancements to the ADMIXTURE algorithm for  
individual ancestry estimation. *BMC bioinformatics*. 2:246.
- 4 Alexander DH, Novembre J, Lange K. 2009. Fast model-based estimation of ancestry  
in unrelated individuals. *Genome Res*. 19:1655-1664.
- 6 Allouche O, Tsoar A, Kadmon R. 2006. Assessing the accuracy of species distribution  
models: prevalence, kappa and the true skill statistic (TSS). *J Appl Ecol*.  
8 43:1223-1232.
- Alvarez K, de Andrés MC, Takahashi A, Oreffo RO. 2014. Effects of hypoxia on  
10 anabolic and catabolic gene expression and DNA methylation in OA  
chondrocytes. *BMC Musculoskelet Disord*. 15:431.
- 12 Anderson LL, Mao X, Scott BA, Crowder CM. 2009. Survival from hypoxia in *C.*  
*elegans* by inactivation of aminoacyl-tRNA synthetases. *Science*. 323:630-633.
- 14 Barbet-Massin M, Thuiller W, Jiguet F. 2012. The fate of European breeding birds  
under climate, land-use and dispersal scenarios. *Global Change Biol*.  
16 18:881-890.
- Barrett JC, Fry B, Maller J, Daly MJ. 2005. Haploview: analysis and visualization of  
18 LD and haplotype maps. *Bioinformatics*. 21:263-265.
- Breiman L, Friedman J, Stone CJ, Olshen RA. 1984. Classification and regression  
20 trees. CRC press.
- Breiman L. 2011. Random forests. *Machine learning*. 45:5-32.
- 22 Cai Q, Qian X, Lang Y, Luo Y, Xu J, Pan S, Hui Y, Gou C, Cai Y, Hao M, et al. 2013.

- Genome sequence of ground tit *Pseudopodoces humilis* and its adaptation to high  
2 altitude. *Genome Biol.* 14:R29.
- Charlesworth B. 2009. Fundamental concepts in genetics: effective population size  
4 and patterns of molecular evolution and variation. *Nat. Rev. Genet.* 10:195–205.
- Cho YS, Hu L, Hou H, Lee H, Xu J, Kwon S, Oh S, Kim HM, Jho S, Kim S, et al.  
6 2013. The tiger genome and comparative analysis with lion and snow leopard  
genomes. *Nat Commun.* 4:2433.
- 8 Cohen J. 1990. A coefficient of agreement for nominal scales. *Edu Psychol Measure.*  
20:37-46.
- 10 Collins WD, Rasch PJ, Boville BA, Hack JJ, McCaa JR, Williamson DL, Briegleb BP,  
Bitz CM, Lin SJ, Zhang M. 2006. The formulation and atmospheric simulation  
12 of the Community Atmosphere Model version 3 (CAM3). *J Climate.*  
19:2144-2161.
- 14 Do R, Balick D, Li H, Adzhubei I, Sunyaev S, Reich D. 2015. No evidence that  
selection has been less effective at removing deleterious mutations in Europeans  
16 than in Africans. *Nat Genet.* 47:126-131.
- Elith J, Graham CH, Anderson RP, Dudik M, Ferrier S, Guisan A, Hijmans RJ,  
18 Huettmann F, Leathwick JR, Lehmann A, et al. 2006. Novel methods improve  
prediction of species' distributions from occurrence data. *Ecography.*  
20 29:129-151.
- Fielding AH, Bell JF. 1997. A review of methods for the assessment of prediction  
22 errors in conservation presence/absence models. *Environ Conserv.* 24:38-49.

- 2 Francke U. 1999. Williams-Beuren syndrome: genes and mechanisms. *Hum Mol Genet.* 8:1947-1954.
- 4 Geissmann T, Lwin N, Aung SS, Aung TN, Aung ZM, Hla TH, Grindley M, Momberg F. 2011. A new species of snub-nosed monkey, genus *Rhinopithecus* Milne-Edwards, 1872 (Primates, Colobinae), from northern Kachin state, northeastern Myanmar. *Am J Primatol.* 73:96-107.
- 8 Glémin S. 2003. How are deleterious mutations purged? Drift versus nonrandom mating. *Evolution.* 57:2678-2687.
- 10 Gutenkunst RN, Hernandez RD, Williamson SH, Bustamante CD. 2009. Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data. *PLoS Genet.* 5:e1000695.
- 12 Harris R. 2007. Improved pairwise alignment of genomic DNA. Ph.D. thesis, The Pennsylvania State University.
- 14 Hastie T, Tibshirani R, Buja A. 1994. Flexible discriminant analysis by optimal scoring. *J Am Stat Assoc.* 89:1255-1270.
- 16 Hastie TJ, Tibshirani RJ. 1990. Generalized additive models. Chapman & Hall: CRC Monographs on Statistics & Applied Probability.
- 18 He K, Hu NQ, Chen , Li JT, Jiang XL. 2016. Interglacial refugia preserved high genetic diversity of the Chinese mole shrew in the mountains of southwest China. *Heredity.* 116: 23-32.
- 20 He K, Jiang XL. 2014. Sky islands of southwest China. I: an overview of phylogeographic patterns. *Chin Sci Bull.* 59:585-597.
- 22

- Hijmans RJ, Cameron S, Parra J, Jones P, Jarvis A. 2005. Very high resolution  
2 interpolated climate surfaces for global land areas. *Int J Climatol*. 25:1965-1978.
- Huang DW, Sherman BT, Lempicki RA. 2009. Bioinformatics enrichment tools: paths  
4 toward the comprehensive functional analysis of large gene lists. *Nucleic Acids  
Res*. 37:1-13.
- 6 Julian RJ. 2007. The response of the heart and pulmonary arteries to hypoxia, pressure,  
and volume. A short review. *Poult Sci*. 86:1006-1011.
- 8 Kim JM, Jeung HC, Rha SY, Yu EJ, Kim TS, Shin YK, Zhang X, Park KH, Park SW,  
Chung HC, et al. 2014. The effect of disintegrin-metalloproteinase ADAM9 in  
10 gastric cancer progression. *Mol Cancer Ther*. 13:3074-3085.
- Kimura M. 1983. *The Neutral Theory of Molecular Evolution*. Cambridge Univ. Press,  
12 Cambridge.
- Kirkpatrick RC. 1998. Ecology and behavior in snub-nosed and douc langurs. In:  
14 Jablonski, N.G. (Ed.), *The Natural History of the Doucs and Snub-nosed  
Monkeys*. World Scientific Press, Singapore. pp. 155-190.
- 16 Li B, Jia Z, Pan R, Ren B. 2003a. Changes in distribution of the snub-nosed monkey  
in China. In *Primates in Fragments*. Springer US. pp. 29-51.
- 18 Li BG, Pan RL, Oxnard CE. 2002. Extinction of snub-nosed monkeys in China during  
the past 400 years. *Int J Primatol*. 23(6):1227-1244.
- 20 Li H. 2011. A statistical framework for SNP calling, mutation discovery, association  
mapping and population genetical parameter estimation from sequencing data.  
22 *Bioinformatics* 27:2987-2993.

- Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler  
2 transform. *Bioinformatics*. 25:1754-1760.
- Li H, Durbin R. 2011. Inference of human population history from individual whole  
4 genome sequences. *Nature*. 475:493-496.
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G,  
6 Durbin R, 1000 Genome Project Data Processing Subgroup. 2009. The sequence  
alignment/map format and SAMtools. *Bioinformatics*. 25:2078-2079.
- 8 Li H, Meng SJ, Men ZM, Fu YX, Zhang YP. 2003b. Genetic diversity and population  
history of golden monkeys (*Rhinopithecus roxellana*). *Genetics*. 164:269-275.
- 10 Li M, Liang B, Feng Z, Tamate HB. 2001. Molecular phylogeny of Sichuan  
snub-nosed monkey (*Rhinopithecus roxellae*) with cytochrome *b* sequences.  
12 *Primates*. 42:153-160.
- Li M, Liu Z, Gou J, Ren B, Pan R, Su Y, Funk SM, Wei F. 2007. Phylogeography and  
14 population structure of the golden monkeys (*Rhinopithecus roxellana*): inferred  
from mitochondrial DNA sequences. *Am J Primatol*. 69:1195-1209.
- 16 Li M, Tian S, Jin L, Zhou G, Li Y, Zhang Y, Wang T, Yeung CK, Chen L, Ma J, et al.  
2013. Genomic analyses identify distinct patterns of selection in domesticated  
18 pigs and Tibetan wild boars. *Nat Genet*. 45:1431-1438.
- Li M, Wei F, Huang C, Pan R, de Ruiter J. 2004. Phylogeny of snub-nosed monkeys  
20 inferred from mitochondrial DNA, cytochrome B, and 12S rRNA sequences. *Int  
J Primatol*. 25:861-873.
- 22 Liedigk R, Yang M, Jablonski NG, Momberg F, Geissmann T, Lwin N, Hla TH, Liu Z,



- Wong B, Ming L, et al. Evolutionary history of the odd-nosed monkeys and the  
2 phylogenetic position of the newly described Myanmar snub-nosed monkey  
*Rhinopithecus strykeri*. *PLoS One*. 7:e37418.
- 4 Long YC, Kirkpatrick RC, Zhong T, Xiao L. 1994. Report on the distribution,  
population, and ecology of the Yunnan snub-nosed monkey (*Rhinopithecus bieti*)  
6 *Primates* 35:241-250.
- Long YC, Kirkpatrick C, Zhong T, Xiao L. 1996. Status and conservation strategy of  
8 the Yunnan snub-nosed monkey. *Chinese Biodiversity*. 4:145-152.
- Luo M, Liu Z, Pan H, Zhao L, Li M. 2012. Historical geographic dispersal of the  
10 golden snub-nosed monkey (*Rhinopithecus roxellana*) and the influence of  
climatic oscillations. *Am J Primatol*. 74(2):91-101.
- 12 Lynch M. 2006. The origins of eukaryotic gene structure. *Mol Biol Evol*.  
23:450-468.
- 14 McCullagh P, Nelder JA. 1989. Generalized linear models (Vol. 37). CRC press.
- Mittermeier RA, Wilson DE, Rylands AB. 2013. Handbook of the Mammals of the  
16 World: Primates. Lynx Edicions.
- Ni J, Zhang XS, Scurlock MO. 2001. Synthesis and analysis of biomass and net  
18 primary productivity in Chinese forests. *Ann. For. Sci.* 58: 351- 384.
- Otto-Bliesner BL, Marshall SJ, Overpeck JT, Miller GH, Hu A. 2006. Simulating  
20 Arctic climate warmth and icefield retreat in the last interglaciation. *Science*.  
311:1751-1753.
- 22 Pan YR, Jablonski NG. 1987. The age and geographical distribution of fossil

- cercopithecids in China. *Hum Evol.* 2:59-69.
- 2 Pan YR. 1995. Fossil primates discovered in China. Primate Research and  
Conservation, China Forestry Publishing House, Beijing, pp. 99-105.
- 4 Patterson N, Price AL, Reich D. 2006. Population structure and eigenanalysis. *PLoS*  
*Genet.* 2:e190.
- 6 Peters DG, Ning W, Chu TJ, Li CJ, Choi AM. 2006. Comparative SAGE analysis of  
the response to hypoxia in human pulmonary and aortic endothelial cells. *Physiol*  
8 *Genomics.* 26:99-108.
- Phillips SJ, Anderson RP, Schapire RE. 2006. Maximum entropy modeling of species  
10 geographic distributions. *Ecol Model.* 190:231-259.
- Pober BR. 2010. Williams-Beuren syndrome. *N Engl J Med.* 362:239-252.
- 12 Prado-Martinez J, Sudmant PH, Kidd JM, Li H, Kelley JL, Lorente-Galdos B,  
Veeramah KR, Woerner AE, O'Connor TD, Santpere G, et al. 2013. Great ape  
14 genetic diversity and population history. *Nature.* 499:471-475.
- Primakoff P, Myles DG. 2000. The ADAM gene family: surface proteins with  
16 adhesion and protease activity. *Trends Genet.* 16:83-87.
- Qiu Q, Zhang G, Ma T, Qian W, Wang J, Ye Z, Cao C, Hu Q, Kim J, et al. 2012. The  
18 yak genome and adaptation to life at high altitude. *Nat Genet.* 44:946-949.
- Quan GQ, Xie JY. 2002. Research on the golden monkey. Science and Education  
20 Publishing House, Beijing. pp. 79-92.
- Ridgeway G. 1999. The state of boosting. *Comput Sci Stat.* 31:172-181.
- 22 Schiffels S, Durbin R. 2014. Inferring human population size and separation history

- from multiple genome sequences. *Nat Genet.* 46:919-925.
- 2 Shah AN, Cadinu D, Henke RM, Xin X, Dastidar RG, Zhang L. 2011. Deletion of a  
subgroup of ribosome-related genes minimizes hypoxia-induced changes and  
4 confers hypoxia tolerance. *Physiol Genomics.* 43:855-872.
- Shen Z, Kauttu T, Seppänen H, Vainionpää S, Ye Y, Wang S, Mustonen H,  
6 Puolakkainen P. 2013. Both macrophages and hypoxia play critical role in  
regulating invasion of gastric cancer in vitro. *Acta Oncol.* 52:852-860.
- 8 Tajima F. 1983. Evolutionary relationship of DNA sequences in finite populations.  
*Genetics.* 105:437-460.
- 10 Tang H, Peng J, Wang P, Risch NJ. 2005. Estimation of individual admixture:  
analytical and study design considerations. *Genet Epidemiol.* 28:289-301.
- 12 Thuiller W. 2003. BIOMOD-optimizing predictions of species distributions and  
projecting potential future shifts under global change. *Global Change Biol.*  
14 9:1353-1362.
- Venables WN, Ripley BD. 2013. Modern applied statistics with S-PLUS. Springer  
16 Science & Business Media.
- Wan QH, Pan SK, Hu L, Zhu Y, Xu PW, Xia JQ, Chen H, He GY, He J, Ni XW, et al.  
18 2013. Genome analysis and signature discovery for diving and sensory properties  
of the endangered Chinese alligator. *Cell Res.* 23:1091-1105.
- 20 Wang YX, Jiang XL, Li DW. 1998. Classification and distribution of the extant  
subspecies of golden snub-nosed monkey (*Rhinopithecus roxellana*), in The  
22 Natural History of the Doucs and Snub-nosed Monkeys, edited by N.G. Jablon

ski World Scientific Publishing, Singapore. pp. 53-64.

- 2 Wang K, Li M, Hakonarson H. 2010. ANNOVAR: functional annotation of genetic  
variants from high-throughput sequencing data. *Nucleic Acids Res.* 38:e164.
- 4 Weir BS, Cockerham CC. 1984. Estimating F-statistics for the analysis of population  
structure. *Evolution.* 38:1358-1370.
- 6 Winzeler EA, Shoemaker DD, Astromoff A, Liang H, Anderson K, Andre B,  
Bangham R, Benito R, Boeke JD, Bussey H, et al. 1999. Functional  
8 characterization of the *S. cerevisiae* genome by gene deletion and parallel  
analysis. *Science.* 285:901-906.
- 10 Wouters BG, Koritzinsky M. 2008. Hypoxia signalling through mTOR and the  
unfolded protein response in cancer. *Nat Rev Cancer.* 8:851-864.
- 12 Xue Y, Prado-Martinez J, Sudmant PH, Narasimhan V, Ayub Q, Szpak M, Frandsen P,  
Chen Y, Yngvadottir B, Cooper DN, et al. 2015. Mountain gorilla genomes  
14 reveal the impact of long-term population decline and inbreeding. *Science.*  
348:242-245.
- 16 Yu L, Wang X, Ting N, Zhang Y. 2011. Mitogenomic analysis of Chinese snub-nosed  
monkeys: Evidence of positive selection in NADH dehydrogenase genes in  
18 high-altitude adaptation. *Mitochondrion.* 11:497-503.
- Zhang Q, Gou W, Wang X, Zhang Y, Ma J, Zhang H, Zhang Y, Zhang H. 2016.  
20 Genome Resequencing Identifies Unique Adaptations of Tibetan Chickens to  
Hypoxia and High-Dose Ultraviolet Radiation in High-Altitude Environments.  
22 *Genome Biol Evol.* 8:765-776.

Zheng B, Xu Q, Shen Y. 2002. The relationship between climate change and

2 Quaternary glacial cycles on the Qinghai-Tibetan Plateau: review and  
speculation. *Quat Int.* 2002; 97-98:93-101.

4 Zhou X, Sun F, Xu S, Fan G, Zhu K, Liu X, Chen Y, Shi C, Yang Y, Huang Z, et al.

2013. Baiji genomes reveal low genetic variability and new insights into  
6 secondary aquatic adaptations. *Nat Commun.* 4:2708.

Zhou X, Wang B, Pan Q, Zhang J, Kumar S, Sun X, Liu Z, Pan H, Lin Y, Liu G, et al.

8 Whole-genome sequencing of the snub-nosed monkey provides insights into  
folivory and evolutionary history. *Nat Genet.* 46:1303-1310.

10

12

14

16

18

20

**Figure Legend:**

2 **Fig. 1 | Geographic distribution, phylogenetics, and population structure of**  
3 **snub-nosed monkeys.** (a) Current geographic distribution of snub nosed  
4 monkeys. (b) The neighbor-joining tree constructed from the allele-shared matrix  
5 of variants of 42 snub-nosed monkeys (crab-eating macaque, *Macaca*  
6 *fascicularis*, is the outgroup). (c) The Principal components analysis (PCA) of 42  
7 snub-nosed monkeys. (d) Genetic structure of the golden snub-nosed monkeys  
8 inferred by *frappe* analysis. The number of populations (*K*) from 2 to 5 is shown.  
9 The sample location for each individual is also indicated.

10

11 **Fig. 2 | Ancestral population size of snub-nosed monkeys.** (a) Ancestral population  
12 size was inferred using a PSMC model for each individual from gray, black, and  
13 Myanmar snub-nosed monkeys. Generation time (*g*) = 10 years, and neutral  
14 mutation rate per generation ( $\mu$ ) =  $1.36 \times 10^{-8}$ . (b) Ancestral population size  
15 was inferred for each golden snub-nosed monkey individual from six current  
16 habitats or mountains. PSMC of black and Myanmar snub-nosed monkeys  
17 shown here are under 2% and 4% FDR, respectively.

18

19 **Fig. 3 | Recent demographic history and species distribution models of golden**  
20 **snub-nosed monkeys.** (a)  $\partial a \partial i$  simulations showing the recent demographic  
21 history of golden snub-nosed monkeys using IM-pre models. Y in brackets

denotes years and  $I$  in brackets denotes individuals. **(b)** Relative cross  
2 coalescence rates (CCR) for the Shennongjia (SNJ) and  
Qinling-Minshan-Qionglai (non-SNJ) populations divergence within Golden  
4 snub-nosed monkeys. **(c)** Species distribution models of golden snub-nosed  
monkeys. Present occurrences (Upper left), fossil records (green triangles) and  
6 the Mid-Holocene (~6 kya), LGM (~21 kya) and LIG (~120-140 kya) species  
distribution predictions of golden snub-nosed monkeys made using occurrence  
8 records (black dots) and the climatic and elevation variables. Areas in red  
indicate the most suitable environmental conditions and those in blue are  
10 background areas (values ranging from 0 to 1). The red line is the value of  
seasonal changes in temperature (standard deviation  $\times$  100), which has the  
12 highest contribution among the 19 climatic variables used.

14 **Fig. 4 | The linkage disequilibrium and genetic load.** **(a)** LD patterns. **(b)** Relative  
number ( $R_{x/y}$ ) of derived alleles at LoF (shade) and missense (blank) sites in  
16 pair-wise comparing groups among gray, black, and golden snub-nosed monkeys.  
Error bars represent  $\pm 2$  SD. **(c)** The scaled number of LoF variant sites (hexagon)  
18 in each population where at least one sample is homozygous for the derived  
allele. The distributions of the synonymous sites using the same statistic are  
20 shown in box. Whiskers show 5<sup>th</sup> and 95<sup>th</sup> percentiles. **(d)** The rate of LoF  
variants relative to synonymous variants in homozygous tracts for each sample

(red circles) and in nonhomozygous regions (green diamonds). Horizontal bars  
2 denote population means;  $P$  values received from Kolmogorov-Smirnov test.

4 **Fig. 5 | Genomic regions with selective sweep signals.** (a) The  $\theta_\pi$  ratios (black  
snub-nosed monkeys /other snub-nosed monkeys) and  $F_{ST}$  values of the  
6 scaffolds where 19 positively selected genes in black snub-nosed monkey  
located. (b) Example of genes with strong selective sweep signals in black  
8 snub-nosed monkeys.  $\theta_\pi$  ratios and  $F_{ST}$  values are plotted using a 10 kb sliding  
window. (c) Distribution of  $\theta_\pi$  ratios (SNJ /Non-SNJ) and  $F_{ST}$  values, which are  
10 calculated in 100 kb windows sliding in 10 kb steps. Data points located to the  
left and right of the left and right vertical dashed lines, respectively  
12 (corresponding to the 5% left and right tails of the empirical  $\theta_\pi$  ratio distribution),  
and above the horizontal dashed line (the 5% right tail of the empirical  $F_{ST}$   
14 distribution) were identified as selected regions for the SNJ (Red points) and  
non-SNJ (blue points) populations. (d) Box plot of  $F_{ST}$  values (left) and  $\theta_\pi$  ratios  
16 (SNJ /non-SNJ) (right) for genomic regions of the SNJ and nonSNJ populations  
of golden snub-nosed monkeys that have undergone positive selection versus the  
18 whole genome. Boxes denote the interquartile range (IQR) between the first and  
third quartiles (25th and 75th percentiles, respectively) and the line inside the  
20 box denotes the median. Whiskers denote the lowest and highest values within  
1.5 times IQR from the first and third quartiles, respectively. Outliers beyond the  
22 whiskers are shown as black dots.