Title: Genomic approaches reveal an endemic sub-population of gray wolves in Southern China

Authors: Guo-Dong Wang[1,2&], Ming Zhang[3,4,5&], Xuan Wang[1,2,5], Melinda A. Yang[3,4], Peng Cao[3], Feng Liu[3], Heng Lu[6], Xiaotian Feng[3], Pontus Skoglund[7], Lu Wang[8], Qiaomei Fu[2,3,4#], and Ya-Ping Zhang[1,2#]

Affiliations:

[1]State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming 650223, China;
[2]Center for Excellence in Animal Evolution and Genetics, Chinese Academy of Sciences, Kunming 650223, China;
[3]Key Laboratory of Vertebrate Evolution and Human Origins of Chinese Academy of Sciences, IVPP, CAS, Beijing, 100044, China;
[4]Center for Excellence in Life and Paleoenvironment, Chinese Academy of Sciences, Beijing, 100044, China;
[5]University of Chinese Academy of Sciences, Beijing, 100049, China;
[6]Department of Molecular and Cell Biology, School of Life Sciences, University of Science and Technology of China, Hefei 230026, China;
[7]The Francis Crick Institute, London NW1 1AT, United Kingdom;
[8]Laboratory for Conservation and Utilization of Bio-Resources & Key Laboratory for Microbial Resources of the Ministry of Education, Yunnan University, Kunming 650091, China;

[&]These authors contribute equally to the work.

Correspondence:

Qiaomei Fu (fuqiaomei@ivpp.ac.cn) OR Ya-Ping Zhang (zhangyp@mail.kiz.ac.cn)

SUMMARY

Despite being one of the most widely distributed terrestrial mammals, the history of gray wolves (*Canis lupus*) in China is not well understood as their habitats have been destroyed with growing economic development. Using six specimens from wolf skins in Chinese Natural History museums, we sequenced their genome using a modified ancient DNA procedure. Using whole genome sequence analysis, we showed that gray wolves from Southern China (SC) derive from a single lineage, distinct from gray wolves from the Tibetan Plateau (*Canis lupus chanco*) and Northern China, suggesting that SC gray wolves may form a distinct sub-population. Of SC gray wolves, one wolf from Zhejiang carries a genetic component from a canid that must have diverged earlier from other wolves than jackals did, perhaps through gene flow from a population related to or further diverged from wolves than the dhole, a species distributed in Southern China and Southeast Asia. This may indicate that interspecific gene flow likely played an important role in shaping the speciation patterns and population structure in the genus *Canis*. Our study is the first to survey museum genomes of gray wolves from Southern China, revealing the presence of an endemic population with ancient interspecific gene flow from a population related to the dhole, and highlighting how sequencing the paleogenome from museum specimens can help us to study extinct species.

RESULTS AND DISCUSSION

The place of origin for domestic dogs (*Canis lupus familiaris*) remains a controversial question for the scientific community despite many efforts at studying dog domestication [1-7]. Geographic distribution, population structure, and genomic features of wild ancestors are essential factors to determine sources of domestication [8]. Gray wolves (*Canis lupus*) are the closest wild relative of dogs, and they are also one of the most widely distributed terrestrial mammals, originally inhabiting major parts of Eurasia, North America, and North Africa [9-11]. Previous studies suggested gray wolves have a complex history [12, 13], with subpopulation structure related to local niches [14-17] and long-term genetic admixture not only with dogs [5, 18], but also with coyotes [13, 16, 19-21]. In China, gray wolves were distributed across the mainland, including most southern regions [8, 22]. Genomic approaches using gray wolf specimens from Southern China may help to shed new light on the demographic history of gray wolves and domestic dogs.

From two Chinese Natural History museums, we obtained six historical wolf skin samples from gray wolves collected from Mainland China (Figure 1A, S1, and Table 1, detailed description in [8]). Four of them came from Southern China. Two gray wolves were from downstream of Yangtze River of Southern China, collected in Lin'an, Zhejiang province in 1974 and Jiangxi province in May 1974. Another two specimens were collected from Guizhou province near Southeast Asia in 1963. The last two specimens were collected from Northeast China, in Baoqing, Heilongjiang province on Jan 24th, 1957, and Baicheng, Jilin province on Feb 11th, 1956.

As skin samples were treated with chemicals reagents and underwent special processing for preservation during storage and exhibition in museums, we used a modified ancient DNA

(aDNA) protocol [23] to retrieve genetic material from the skin samples. In total, 35 genomic libraries were produced using a double stranded library preparation protocol [24, 25], and each was treated with uracil-DNA-glycosylase (UDG) and endonuclease (Endo VIII) to remove characteristic aDNA deamination [26] (Table S1). We sequenced the libraries using 2×150 bp reads on an Illumina HiSeq X platform.

All but one sample was sequenced to 0.15 -15-fold average coverage, too low a coverage to determine heterozygous sites, so we called haploid alleles from a randomly chosen sequence read at each position. We applied a filter where we ignored fragments with length less than 30, ignored the first and last two base pairs of each fragment, and required a base pair quality higher than 20 and mapping quality of at least 30. The Jiangxi wolf was sequenced to 37x (Table 1) – high enough to call heterozygotes. Thus, we applied the software GATK with the Unified Genotyper parameter to determine diploid calls [27]. We also included 79 canids from previous studies [28, 29], including an ancient gray wolf, Taimyr [30], 25 modern gray wolves chosen from regions overlapping the current ranges [3, 7, 18, 19, 31-34], 46 domestic dogs from all over the world [3, 31, 35, 36], two jackals [33], two coyotes [33], one red wolf [19], one dhole [37], and one Andean fox [28] (Figure 1, Table S2).

Phylogeny and population structure

To investigate the relationship of the newly sampled individuals to wolf and dog populations, we calculated pairwise allele-sharing distances among all pairs of wolf and dog populations. We applied a principal components analysis (PCA) to the resulting pairwise distance matrix using SMARTPCA [38]. The first principal component distinguishes between gray wolf and dog populations, while the second principal component distinguishes between East Asian and European dogs (Figure S2), consistent with previous studies [1, 3, 4]. To obtain increased

resolution, we redid the PCA excluding dog populations (Figure 1B). The resulting PCA shows that the two Guizhou wolves cluster with a Chinese wolf (labeled China) from the San Diego Zoo, whose origin is not recorded [18], and a gray wolf from Linfen, Shanxi, China (labeled Shanxi) sampled in 1988, near the border of Southern China [3]. We also find that the Jilin and Heilongjiang wolves cluster with gray wolves from Inner Mongolia and Liaoning. The Jiangxi wolf is closest to the Qinghai wolf, while the Zhejiang wolf is closest to the cluster containing the gray wolves from Inner Mongolia and Liaoning (Figure S1).

Using both a maximum likelihood (ML) tree (Figure 1C) and a neighbor-joining (NJ) tree with only wolves (Figure S3), we further find that the Zhejiang wolf forms a clade with the Guizhou, Shanxi, and China wolves (Figure 1C). We define these gray wolves as the gray wolves from Southern China (SC) and find that they are most closely related to the Qinghai, Tibet, and Jiangxi wolves, where the Tibetan gray wolf (*Canis lupus chanco*) is a gray wolf sub-species that occupies habitats on the Qinghai-Tibet Plateau [39] and adapts to high-altitude environments [32]. Other gray wolves from Northern Asia (NA) form a clade with SC and Tibetan gray wolves relative to Middle Eastern and European gray wolves, which form a distinct clade. The 35,000-year-old wolf from Taimyr Peninsula in northern Siberia joins at the base of the Eurasian wolf phylogeny, and American wolves separate the earliest from other wolves, consistent with [18].

We used TreeMix [40] to investigate the genetic relationship between historical and present-day wolves. TreeMix determines population structure using maximum likelihood trees and allows for both population splits and potential gene flow by using genome-wide allele frequency data and a Gaussian approximation of genetic drift. The maximum likelihood tree (Figure S4) without admixture (m=0) is consistent with previous patterns, where gray wolves

from East Asia form three groups: Guizhou and Zhejiang wolves form a clade with Shanxi and China wolves, the Jiangxi wolf forms a clade with Tibet and Qinghai wolves, and Jilin and Heilongjiang wolves, like other North Asian wolves, are outgroup to SC and Tibetan wolves. All three of these groups form a clade relative to non-Asian wolves.

We measured the shared genetic drift between each newly sequenced individual (X) and other dogs and wolves (Y) since their separation from an outgroup, *Dhole*, using $f_3(X, Y; Dhole)$ [37, 41], and found a similar pattern as above, where the Zhejiang wolf shares the most genetic drift with gray wolves from Guizhou and Jiangxi (Figure S5). We then used *D*-statistics [42] of the form *D(Fox, Test; X, Y)*, where X and Y are previously published wolves, to formally test the relationship these historical wolves have with different wolf populations. For Guizhou and Zhejiang wolves, we find that they share more alleles with the SC gray wolves (Shanxi and China) than with all other wolves, as *D(Fox, Guizhou/Zhejiang; X, SC)>0* ($10.8<Z<33.4$, Table S3). The Jiangxi wolf shares more allele with the Tibet and Qinghai gray wolves than other wolves, as *D(Fox, Jiangxi; X, Tibet/Qinghai)>0* ($11.7<Z<24.9$, Table S3), while Jilin and Heilongjiang wolves share the most alleles with NA gray wolves. These results support our above analyses, again grouping the Guizhou and Zhejiang wolves with the SC wolves, the Jiangxi wolf with the Tibetan wolves, and the Jilin and Heilongjiang wolves with the NA wolves.

In summary, our results revealed that the lowland Chinese wolves [20] consisted of two major populations: SC and NA wolves. Gray wolves from Zhejiang and Guizhou group most closely with and share the most genetic drift with SC wolves, which includes present-day populations in Shanxi and China. The Jilin and Heilongjiang gray wolves share the most

genetic similarity to the NA gray wolves, which is the other clade in Northern China and Eastern Russia.

Testing for admixture in gray wolves.

Using *D(Fox, X; Test, Y),* we find that all gray wolves (X) share more alleles with the Guizhou, Jilin, and Heilongjiang wolves (Test) than with the gray wolves from Tibet and Qinghai, i.e. *D(Fox, X; Test, Qinghai/Tibet)<0* (Table S5). The Jiangxi wolf shares more alleles with the Tibet and Qinghai wolves than other wolves (Table S3), and here, we find that the Tibet and Qinghai wolves share more alleles with the Jiangxi wolf, i.e. *D(Fox, Qinghai/Tibet; Jiangxi, X)<0* (-28.7<Z<-14.7, Table S5), emphasizing that the Jiangxi, Tibet, and Qinghai wolves form a clade. However, while *D(Fox, X; Jiangxi, Qinghai)~0* (-2.3<Z<1.8, Table S5), indicating a symmetric relationship as expected for the Jiangxi and Qinghai wolves forming a clade, we observe *D(Fox, X; Jiangxi, Tibet)<0* (Table S5), suggesting that the Jiangxi wolf has a connection to other wolves relative to the Tibetan wolf. Thus, we observe that the Tibet and Qinghai gray wolves act as an outgroup to most SC and NA gray wolf populations, and while the Jiangxi wolf forms a clade with the Tibet and Qinghai gray wolves, the Jiangxi wolf shows connections to non-Tibetan wolf populations. To test for admixture between the SC, NA, and Tibetan wolves, we used *f3(Test; X, Y)*, where a significantly negative value (|Z|>3) suggests that the Test population is a mixture of ancestry related to X and Y, two other wolf populations. Testing all combinations as both a source population and the admixed population, we found that *f3(Jiangxi/Qinghai; Tibetan, SC/NA)<0* (-19.101<Z<-9.141, Table S4), suggesting that both the Jiangxi and Qinghai wolves show evidence of ancestry from populations related to both Tibetan and SC gray wolves, explaining why the Jiangxi gray wolf shares a connection to non-Tibetan gray wolf populations.

In contrast to all other ancient gray wolves from China, the Zhejiang wolf shows a markedly different pattern, where all other wolf populations, including the Tibetan and Qinghai gray wolves and more distantly related wolves from further west, form a clade with each other relative to the Zhejiang wolf. That is, we observe that *D(Fox, X; Zhejiang, Y)>0* (9.1<Z<71.5, Table S5), where X and Y are all other gray wolves, including the Taimyr. Earlier, we found that *D(Fox, Zhejiang; X, Shanxi/China)>0* (Table S3), indicating that the Zhejiang wolf shows connections to the wolves from Shanxi and China. These results suggest that the Zhejiang wolf shows a close relationship to gray wolves from Shanxi and China, but that this wolf also possesses an ancestral component that is older than the common ancestral population of the Taimyr and all other gray wolves. The error rate for the Zhejiang wolf (0.4%) is higher than that estimated for other wolves sequenced in this study (0.1%-0.2%), likely because of its low coverage (Table 1). After simulating an error rate similar to that observed for the Zhejiang wolf in these other wolves, we find that our results remain consistent with our previous results. That is, the Zhejiang wolf shows a distinct pattern from that observed in other wolves for both lower and higher error rates.

We use the genomic data from canids typically outgroup to all wolves and dogs – the Dhole, Jackal, Coyote, and Red wolf – to understand how deeply the old component found in the Zhejiang wolf separated from other canid populations. Other canids separated from wolf populations very early, with the Dhole diverging earliest, followed by the Jackal and most recently the Coyote [36, 43, 44]. The Red wolf is genetically very similar to the Coyote and shows substantial gene flow from gray wolves [13, 19]. First, comparing the Jackal to wolves (X) and the Coyote, we find that for all wolves but the Zhejiang wolf, *D(Fox, Jackal; X, Coyote)* ranges from -0.04 to -0.02 (-20.2<Z<-12.6, Table S6), indicating a connection

between the Jackal and gray wolves. We find the reverse for the Zhejiang wolf, however, where the Jackal shares more alleles with the Coyote than with the Zhejiang wolf, i.e. *D(Fox, Jackal; Zhejiang, Coyote)=0.13* (Z=39.3, Table S6). We find similar results replacing the Coyote with the Red wolf. The large contrast between the results for the Zhejiang wolf compared to other gray wolves suggests that the old component came from a population that diverged deeply in the past, who separated prior to the common ancestor of jackals and coyotes.

We also observe that for all gray wolves, we find *D(Fox, Dhole; X, Jackal)>0* (Table S6), suggesting that gray wolves share a deep lineage older than the separation of the Jackal and Dhole or that there is a direct genetic connection between the Jackal and Dhole. However, while D ranges from 0.1 to 0.2 (4.5<Z<7.9, Table S6) for most gray wolves, using the Zhejiang wolf greatly increases the D value to 0.11 (Z=22.4, Table S6). We find that *D(Fox, Dhole; Zhejiang, Jackal)* remains significantly positive (Z=12.2) using only transversions, suggesting that the result for the Zhejiang wolf is not related to ancient DNA damage and likely reflects an unusual admixture history. If the Zhejiang wolf was no different from other gray wolves, especially the two Guizhou individuals to which they share the closest relationship (Table S5), we would expect to find that *D(Fox, Dhole; X, Zhejiang)~0*, which would indicate that the Zhejiang wolf and other gray wolves are similarly related to the Dhole. However, we observe that *D(Fox, Dhole; X, Zhejiang)<<0* (-23.4<Z<-12.9, Table S6), indicating that the Dhole shares more alleles with other gray wolves than with the Zhejiang gray wolf. These patterns suggest that the ancestral component found in the Zhejiang wolf came from a population that diverged earlier than the common ancestor of the Jackal and Dhole, which is older than the separation of the Jackal and Dhole from wolves, suggesting that the Zhejiang wolf possesses very deep ancestry not found in other gray wolves.

Using the tree model with no admixture from Treemix (Figure S6), we visualized the matrix of residuals (Figure S7) to determine how the estimated genetic relationship between each pair of canids fit the model. A high residual indicates that the pair does not fit the tree model and may be candidates for an admixture event. We find two candidate admixture events, the first between the Andean fox and the Zhejiang wolf (Figure 2); and the second between gray wolves from Portugal and Iberia. In a maximum-likelihood tree allowing two admixture events, admixture from the lineage leading to the Andean fox to the lineage leading to the Zhejiang wolf is included, while gray wolves from Portugal and Iberia are instead grouped into the same cluster (Figure S8). Admixture between the outgroup Andean fox and the Zhejiang wolf supports our conclusions from the D-statistic analysis (Table S5 and Table S6), in which the Zhejiang wolf possesses an ancestral component that came from a population that diverged earlier than the Jackal or Dhole did from wolves. The estimated value of the migration event in the Zhejiang individual is 12.3% ± 0.4% (P<2.2x10^{-308}). In the Treemix analysis, we used the Andean fox as an outgroup, whose distance from the included canids would result in weak phylogenetic constraints. In addition, we also used the F4-ratio test to estimate the proportion of this deep ancestry, and since it is older than the separation of the Jackal and Dhole, we used an unrooted phylogeny where the Fox is used as a proxy as the source of the deep ancestry. Thus, we estimate the proportion of ancestry related to the Fox, which is given by:

$$\frac{f_4(Dhole,\ Jackal;\ Coyote,\ X)}{f_4(Dhole,\ Jackal;\ Coyote,\ Andean\ Fox)}$$, where X is each gray wolf in turn.

Using this method, we found that the estimated admixture proportion of the deep ancestry for the Zhejiang wolf is 11.7% ± 0.5%, whereas all other grey wolves have an estimated admixture proportion close to zero. Thus, both the Treemix analysis and the F4-ratio test

support the presence of gene flow from an ancient canid population into the ancestors of the Zhejiang wolf.

Conclusion

The distribution of gray wolves in East Asia is controversial since some studies have claimed that gray wolves never existed [45, 46] or are now extinct from Southern China [47], while others sources, especially those based on Chinese literature, stated that they are present across all of mainland China [8]. In this study, we provide the first comparative genomic analysis of gray wolves from East Asia focusing particularly on wolves from Southern China, where some believed no gray wolves were distributed [46, 48]. Previously, Asian wolves could be divided into two populations: Tibetan gray wolves (*Canis lupus chanco*) and Chinese lowland wolves [20, 32]. Here, using ancient genome-wide data, we reconstruct the phylogeny and evaluate the population structure and shared genetic drift between East Asian gray wolves to show that they form three major groups, which we call Southern China, North Asia, and Tibetan, based on their geographic distribution. Interestingly, specimens from SC gray wolves were all collected from 1956 - 1988. Our results highlight that the population in Southern China is endemic, and with the fast growing economic development of China, it is paramount to protect and restore their ecological habitat. Through our study, we also emphasize the value of using paleogenomic approaches to study the numerous museum specimens available [49, 50], and we address the importance of using population genomics to determine current or future conservation efforts.

Finally, our analyses show that admixture played a large role between the different Asian wolves, and we highlight two instances here. First, we find that a wolf as far southeast as Jiangxi province shows evidence of being a mixture of Tibetan wolves and other wolves in

China. Second, we traced an unusual admixture event in the Zhejiang wolf. In many analyses, this wolf behaved similarly to other gray wolves in China, particularly those from Southern China. However, tests of admixture suggest that the Zhejiang wolf has gene flow from a canid that diverged earlier than the separation of wolves and jackals. D-statistic analysis suggests this may be from the Dhole, a species distributed in Southern China and Southeast Asia [51], or another canid that separated earlier than the divergence between wolves and the Dhole. Whatever the source of this ancestry, estimates of the admixture proportion from this deeply diverging population are estimated to be ~12%. Our results, taken together with previous research [52], reveal that canids are an ideal system in which to study how gene flow can shape speciation in a genus, and highlight the need for greater study of ancient gray wolf populations.

ACCESSION NUMBERS

Sequence data for six gray wolf genomes has been submitted to the Genome Sequence Archive (http://gsa.big.ac.cn/) under accession number PRJCA001135.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, nine figures, and six tables, which can be found with this article online at XXX.

AUTHOR CONTRIBUTIONS

Y.-P.Z. and Q.F. designed research. G.-D.W. and Q.F. managed the project. L.W. collected samples. P.C., F.L., M.Z., and X.F. performed DNA experiments. Q.F., X.F., M.Y., and P.S., did the data processing. Q.F., G.-D.W., M.Z., X.W., M.Y., and H.L. analyzed the data. G.-D.W., M.Y., Q.F., and Y.-P.Z. wrote the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

1. Frantz, L.A.F., Mullin, V.E., Pionnier-Capitan, M., Lebrasseur, O., Ollivier, M., Perri, A., Linderholm, A., Mattiangeli, V., Teasdale, M.D., Dimopoulos, E.A., et al. (2016). Genomic and archaeological evidence suggests a dual origin of domestic dogs. Science *352*, 1228-1231.

2. Shannon, L.M., Boyko, R.H., Castelhano, M., Corey, E., Hayward, J.J., McLean, C., White, M.E., Said, M.A., Anita, B.A., Bondjengo, N.I., et al. (2015). Genetic structure in village dogs reveals a Central Asian domestication origin. P Natl Acad Sci USA *112*, 13639-13644.

3. Wang, G.D., Zhai, W.W., Yang, H.C., Wang, L., Zhong, L., Liu, Y.H., Fan, R.X., Yin, T.T., Zhu, C.L., Poyarkov, A.D., et al. (2016). Out of southern East Asia: the natural history of domestic dogs across the world. Cell Res *26*, 21-33.

4. vonHoldt, B.M., Pollinger, J.P., Lohmueller, K.E., Han, E.J., Parker, H.G., Quignon, P., Degenhardt, J.D., Boyko, A.R., Earl, D.A., Auton, A., et al. (2010). Genome-wide SNP and haplotype analyses reveal a rich history underlying dog domestication. Nature *464*, 898-U109.

5. Liu, Y.H., Wang, L., Xu, T., Guo, X.M., Li, Y., Yin, T.T., Yang, H.C., Hu, Y., Adeola, A.C., Sanke, O.J., et al. (2018). Whole-Genome Sequencing of African Dogs Provides Insights into Adaptations against Tropical Parasites. Mol Biol Evol *35*, 287-298.

6. Thalmann, O., Shapiro, B., Cui, P., Schuenemann, V.J., Sawyer, S.K., Greenfield, D.L., Germonpre, M.B., Sablin, M.V., Lopez-Giraldez, F., Domingo-Roura, X., et al.

(2013). Complete Mitochondrial Genomes of Ancient Canids Suggest a European Origin of Domestic Dogs. Science *342*, 871-874.

7. Botigue, L.R., Song, S., Scheu, A., Gopalan, S., Pendleton, A.L., Oetjens, M., Taravella, A.M., Seregely, T., Zeeb-Lanz, A., Arbogast, R.M., et al. (2017). Ancient European dog genomes reveal continuity since the Early Neolithic. Nat Commun *8*.

8. Wang, L., Ma, Y.P., Zhou, Q.J., Zhang, Y.P., Savolaimen, P., and Wang, G.D. (2016). The geographical distribution of grey wolves (Canis lupus) in China: a systematic review. Zool Res *37*, 315-326.

9. Young, S.P., Goldman, E.A., and American Wildlife Foundation. (1944). The wolves of North America. Part-II. Classification of wolves, "First edition." Edition, (Washington, D.C.,: The American Wildlife Institute).

10. Wilson, D.E., and Reeder, D.M. (2005). Mammal species of the world : a taxonomic and geographic reference, 3rd Edition, (Baltimore: Johns Hopkins University Press).

11. Loog, L., Thalmann, O., Sinding, M.-H.S., Schuenemann, V.J., Perri, A., Germonpre, M., Bocherens, H., Witt, K.E., Samaniego Castruita, J.A., Velasco, M.S., et al. (2018). Modern wolves trace their origin to a late Pleistocene expansion from Beringia. bioRxiv.

12. Wayne, R.K., Lehman, N., Allard, M.W., and Honeycutt, R.L. (1992). Mitochondrial-DNA Variability of the Gray Wolf - Genetic Consequences of Population Decline and Habitat Fragmentation. Conserv Biol *6*, 559-569.

13. vonHoldt, B.M., Pollinger, J.P., Earl, D.A., Knowles, J.C., Boyko, A.R., Parker, H., Geffen, E., Pilot, M., Jedrzejewski, W., Jedrzejewska, B., et al. (2011). A genome-wide perspective on the evolutionary history of enigmatic wolf-like canids. Genome Res *21*, 1294-1305.

14.   Carmichael, L.E., Nagy, J.A., Larter, N.C., and Strobeck, C. (2001). Prey specialization may influence patterns of gene flow in wolves of the Canadian Northwest. Mol Ecol *10*, 2787-2798.

15.   Geffen, E., Anderson, M.J., and Wayne, R.K. (2004). Climate and habitat barriers to dispersal in the highly mobile grey wolf. Mol Ecol *13*, 2481-2490.

16.   Pilot, M., Greco, C., Vonholdt, B.M., Jedrzejewska, B., Randi, E., Jedrzejewski, W., Sidorovich, V.E., Ostrander, E.A., and Wayne, R.K. (2014). Genome-wide signatures of population bottlenecks and diversifying selection in European wolves. Heredity *112*, 428-442.

17.   Pilot, M., Jedrzejewski, W., Branicki, W., Sidorovich, V.E., Jedrzejewska, B., Stachura, K., and Funk, S.M. (2006). Ecological factors influence population genetic structure of European grey wolves. Mol Ecol *15*, 4533-4553.

18.   Freedman, A.H., Gronau, I., Schweizer, R.M., Ortega-Del Vecchyo, D., Han, E.J., Silva, P.M., Galaverni, M., Fan, Z.X., Marx, P., Lorente-Galdos, B., et al. (2014). Genome Sequencing Highlights the Dynamic Early History of Dogs. Plos Genet *10*.

19.   vonHoldt, B.M., Cahill, J.A., Fan, Z.X., Gronau, I., Robinson, J., Pollinger, J.P., Shapiro, B., Wall, J., and Wayne, R.K. (2016). Whole-genome sequence analysis shows that two endemic species of North American wolf are admixtures of the coyote and gray wolf. Sci Adv *2*.

20.   Fan, Z.X., Silva, P., Gronau, I., Wang, S.G., Armero, A.S., Schweizer, R.M., Ramirez, O., Pollinger, J., Galaverni, M., Del-Vecchyo, D.O., et al. (2016). Worldwide patterns of genomic variation and admixture in gray wolves. Genome Res *26*, 163-173.

21.   Monzon, J., Kays, R., and Dykhuizen, D.E. (2014). Assessment of coyote-wolf-dog admixture using ancestry-informative diagnostic SNPs. Mol Ecol *23*, 182-197.

22.   Smith, A.T., and Xie, Y. (2008). A guide to the mammals of China, (Princeton, N.J.: Princeton University Press).

23.   Dabney, J., Knapp, M., Glocke, I., Gansauge, M.T., Weihmann, A., Nickel, B., Valdiosera, C., Garcia, N., Paabo, S., Arsuaga, J.L., et al. (2013). Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. P Natl Acad Sci USA *110*, 15758-15763.

24.   Meyer, M., and Kircher, M. (2010). Illumina sequencing library preparation for highly multiplexed target capture and sequencing. Cold Spring Harb Protoc *2010*, pdb prot5448.

25.   Kircher, M., Sawyer, S., and Meyer, M. (2012). Double indexing overcomes inaccuracies in multiplex sequencing on the Illumina platform. Nucleic Acids Res *40*.

26.   Briggs, A.W., Stenzel, U., Johnson, P.L.F., Green, R.E., Kelso, J., Prufer, K., Meyer, M., Krause, J., Ronan, M.T., Lachmann, M., et al. (2007). Patterns of damage in genomic DNA sequences from a Neandertal. P Natl Acad Sci USA *104*, 14616-14621.

27.   DePristo, M.A., Banks, E., Poplin, R., Garimella, K.V., Maguire, J.R., Hartl, C., Philippakis, A.A., del Angel, G., Rivas, M.A., Hanna, M., et al. (2011). A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet *43*, 491-+.

28.   Auton, A., Li, Y.R., Kidd, J., Oliveira, K., Nadel, J., Holloway, J.K., Hayward, J.J., Cohen, P.E., Greally, J.M., Wang, J., et al. (2013). Genetic Recombination Is Targeted towards Gene Promoter Regions in Dogs. Plos Genet *9*.

29.   Tang, B., Zhou, Q., Dong, L., Li, W., Zhang, X., Lan, L., Zhai, S., Xiao, J., Zhang, Z., Bao, Y., et al. (2018). iDog: an integrated resource for domestic dogs and wild canids. Nucleic Acids Res.

30. Skoglund, P., Ersmark, E., Palkopoulou, E., and Dalen, L. (2015). Ancient Wolf Genome Reveals an Early Divergence of Domestic Dog Ancestors and Admixture into High-Latitude Breeds. Current Biology *25*, 1515-1519.

31. Marsden, C.D., Ortega-Del Vecchyo, D., O'Brien, D.P., Taylor, J.F., Ramirez, O., Vila, C., Marques-Bonet, T., Schnabel, R.D., Wayne, R.K., and Lohmueller, K.E. (2016). Bottlenecks and selective sweeps during domestication have increased deleterious genetic variation in dogs. P Natl Acad Sci USA *113*, 152-157.

32. Zhang, W.P., Fan, Z.X., Han, E.J., Hou, R., Zhang, L., Galaverni, M., Huang, J., Liu, H., Silva, P., Li, P., et al. (2014). Hypoxia Adaptations in the Grey Wolf (Canis lupus chanco) from Qinghai-Tibet Plateau. Plos Genet *10*.

33. Wang, X., Zhou, B.-W., Yin, T.-T., Chen, F.-L., Esmailizadeh, A., Turner, M.M., Poyarkov, A.D., Savolainen, P., Wang, G.-D., Fu, Q., et al. (2018). Canine transmissible venereal tumor genome reveals ancient introgression from coyotes to arctic sled dogs. bioRxiv, 350512.

34. Wang, G.D., Zhai, W.W., Yang, H.C., Fan, R.X., Cao, X., Zhong, L., Wang, L., Liu, F., Wu, H., Cheng, L.G., et al. (2013). The genomics of selection in dogs and the parallel evolution between dogs and humans. Nat Commun *4*.

35. Decker, B., Davis, B.W., Rimbault, M., Long, A.H., Karlins, E., Jagannathan, V., Reiman, R., Parker, H.G., Drogemuller, C., Corneveaux, J.J., et al. (2015). Comparison against 186 canid whole-genome sequences reveals survival strategies of an ancient clonally transmissible canine tumor. Genome Res *25*, 1646-1655.

36. Lindblad-Toh, K., Wade, C.M., Mikkelsen, T.S., Karlsson, E.K., Jaffe, D.B., Kamal, M., Clamp, M., Chang, J.L., Kulbokas, E.J., Zody, M.C., et al. (2005). Genome sequence, comparative analysis and haplotype structure of the domestic dog. Nature *438*, 803-819.

37.   Wang, G.-D., Shao, X.-J., Bai, B., Wang, J., Wang, X., Cao, X., Liu, Y.-H., Wang, X., Yin, T.-T., Zhang, S.-J., et al. (2018). Structural variation during dog domestication: insights from gray wolf and dhole genomes. National Science Review, nwy076-nwy076.

38.   Patterson, N., Price, A.L., and Reich, D. (2006). Population structure and eigenanalysis. Plos Genet *2*, 2074-2093.

39.   Pocock, R.I. (1941). The fauna of British India including Ceylon and Burma. Mammalia, Vol II. Carnivora., (Taylor and Francis, London.).

40.   Pickrell, J.K., and Pritchard, J.K. (2012). Inference of Population Splits and Mixtures from Genome-Wide Allele Frequency Data. Plos Genet *8*.

41.   Raghavan, M., Skoglund, P., Graf, K.E., Metspalu, M., Albrechtsen, A., Moltke, I., Rasmussen, S., Stafford, T.W., Orlando, L., Metspalu, E., et al. (2014). Upper Palaeolithic Siberian genome reveals dual ancestry of Native Americans. Nature *505*, 87-+.

42.   Patterson, N., Moorjani, P., Luo, Y.T., Mallick, S., Rohland, N., Zhan, Y.P., Genschoreck, T., Webster, T., and Reich, D. (2012). Ancient Admixture in Human History. Genetics *192*, 1065-+.

43.   Wayne, R.K., Geffen, E., Girman, D.J., Koepfli, K.P., Lau, L.M., and Marshall, C.R. (1997). Molecular systematics of the Canidae. Syst Biol *46*, 622-653.

44.   Macdonald, D.W., and Sillero-Zubiri, C. (2004). Biology and conservation of wild canids, (Oxford: Oxford University Press).

45.   Callaway, E. (2013). Dog genetics spur scientific spat. Nature *498*, 282-283.

46.   Larson, G., and Fuller, D.Q. (2014). The Evolution of Animal Domestication. Annu Rev Ecol Evol S *45*, 115-136.

47. Lau, M.W.N., Fellowes, J.R., and Chan, B.P.L. (2010). Carnivores (Mammalia: Carnivora) in South China: a status review with notes on the commercial trade. Mammal Rev *40*, 247-292.

48. Callan, R., Nibbelink, N.P., Rooney, T.P., Wiedenhoeft, J.E., and Wydeven, A.P. (2013). Recolonizing wolves trigger a trophic cascade in Wisconsin (USA). J Ecol *101*, 837-845.

49. Roy, M.S., Girman, D.J., Taylor, A.C., and Wayne, R.K. (1994). The Use of Museum Specimens to Reconstruct the Genetic-Variability and Relationships of Extinct Populations. Experientia *50*, 551-557.

50. Min-Shan Ko, A., Zhang, Y., Yang, M.A., Hu, Y., Cao, P., Feng, X., Zhang, L., Wei, F., and Fu, Q. (2018). Mitochondrial genome of a 22,000-year-old giant panda from southern China reveals a new panda lineage. Curr Biol *28*, R693-R694.

51. Iyengar, A., Babu, V.N., Hedges, S., Venkataraman, A.B., Maclean, N., and Morin, P.A. (2005). Phylogeography, genetic structure, and diversity in the dhole (Cuon alpinus). Mol Ecol *14*, 2281-2297.

52. Gopalakrishnan, S., Sinding, M.S., Ramos-Madrigal, J., Niemann, J., Samaniego Castruita, J.A., Vieira, F.G., Caroe, C., Montero, M.M., Kuderna, L., Serres, A., et al. (2018). Interspecific Gene Flow Shaped the Evolution of the Genus Canis. Curr Biol *28*, 3441-3449 e3445.

TABLES

Table 1. Information on samples in the present study.

Sample ID	Coverage	Snps	Sources	Location
Jilin	1.87	6,890,236	National Zoological Museum of China, Beijing, China	Jilin
Zhejiang	0.15	1,866,574		Zhejiang
Heilongjiang	1.61	5,438,606		Heilongjiang
Guizhou1	15.33	12,688,043	Natural History Museum of Zoology, Kunming, China	Guizhou
Guizhou2	12.39	12,968,150		Guizhou
Jiangxi	37.00	13,528,667		Jiangxi

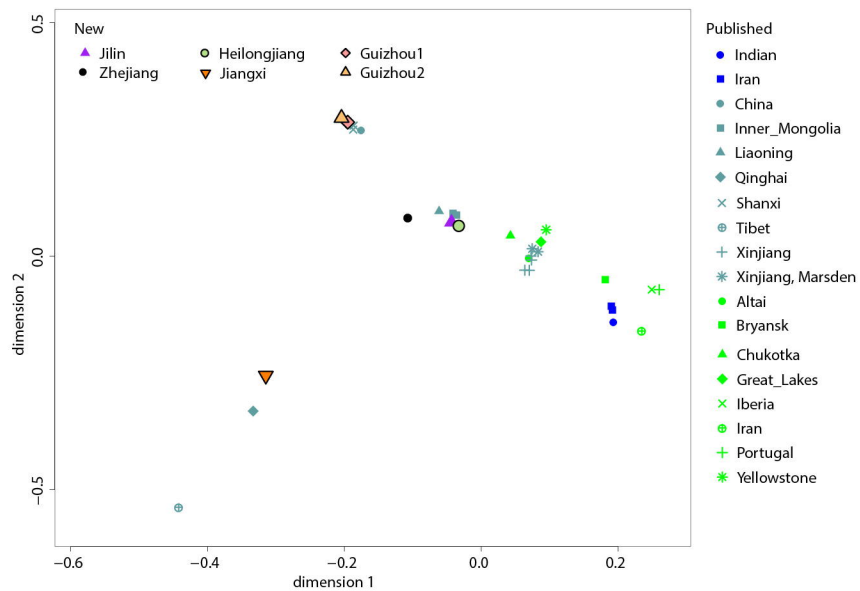## Figure legends

Figure 1. Population structure and phylogeny of 32 gray wolves. (A) Geographic locations, where the key is shared with the principal component analysis in (B). In (B), the label 'New' represents the six samples sequenced in the study, and the label 'Published' represents 25 samples from previous studies. (C) The maximum likelihood tree of 39 canids, where the Andean fox is used as an outgroup.

Figure 2. The maximum-likelihood tree based on TreeMix with m=1. The scale bar shows ten times the average standard error of the entries in the sample covariance matrix. We have used the prefix "Gray_Wolf_" for highlighting specimens of gray wolves.

A

B

C

Southern China

Tibetan

North Asian

Middle Eastern

European

New
Jilin
Zhejiang
Heilongjiang
Jiangxi
Guizhou1
Guizhou2

Published
Indian
Iran
China
Inner_Mongolia
Liaoning
Qinghai
Shanxi
Tibet
Xinjiang
Xinjiang, Marsden
Altai
Bryansk
Chukotka
Great_Lakes
Iberia
Iran
Portugal
Yellowstone

dimension 2
dimension 1

Guizhou1
Zhejiang
Guizhou2
Shanxi
Shanxi
China
Qinghai
Tibet
Jiangxi
Jilin
Heilongjiang
Liaoning
Inner mongolia
Inner mongolia
Chukotka, Russia
Xinjiang
Xinjiang
Xinjiang
Xinjiang
Xinjiang
Xinjiang
Alatai
Indian
Iran
Iran
Iran
Iberia
Portugal
Bryansk, Russia
Taimyr
Yellowstone
Great Lakes
Coyotes
Coyotes
Red wolf
Jackals
Jackals
Dhole
Andean fox

0.03