Letter (Discoveries)

Whole-genome sequencing of African dogs provides insights into adaptations against tropical parasites

Yan-Hu Liu[1,2*], Lu Wang[1*], Tao Xu[3*]，Xiaomin Guo[3], Yang Li[3], Ting-Ting Yin[2,4], He-Chuan Yang[5],Yang Hu[6], Adeniyi C. Adeola[2,4], Oscar J. Sanke[7], Newton O. Otecko[2,4], Meng Wang[1], Yaping Ma[1], Olaogun S. Charles[8], Mikkel-Holger S. Sinding[9,10], Shyam Gopalakrishnan[9], José Alfredo Samaniego[9], Anders J. Hansen[9], Carlos Fernandes[11], Philippe Gaubert[12], Jane Budd[13], Philip M. Dawuda[14], Eli Knispel Rueness[15], Lubin Jiang[6], Weiwei Zhai[5], M.T.P Gilbert[9,16,17], Min-Sheng Peng[2,4], Xiaopeng Qi[3¶], Guo-Dong Wang[2,4¶], Ya-Ping Zhang[1,2,4,¶]

[1]State Key Laboratory for Conservation and Utilization of Bio-Resources in Yunnan, and School of Life Sciences, Yunnan University, Kunming, 650091,Yunnan, China; [2]State Key Laboratory of Genetic Resources and Evolution, and Yunnan Laboratory of Molecular Biology of Domestic Animals, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, 650223, China; [3]Key Laboratory of Animal Models and Human Disease Mechanisms of Chinese Academy of Sciences/Key Laboratory of Bioactive Peptides of Yunnan Province, Kunming Institute of Zoology, Kunming, 650223, Yunnan, China; [4]Sino-Africa Joint Research Center, Chinese Academy of Sciences, Kunming, 650223, Yunnan, China; [5]Human Genetics, Genome Institute of Singapore, A*STAR, 60 Biopolis Street, Genome #02-01, 138672, Singapore; [6]Institut Pasteur of Shanghai, CAS Key Laboratory of Molecular Virology and Immunology, Chinese Academy of Sciences, Shanghai 200031, China; [7]Taraba State Ministry of Agriculture and Natural Resources, Jalingo, Taraba State, Nigeria; [8]Department of Veterinary Medicine, University of Ibadan, Nigeria; [9]Natural History Museum of Denmark, Øster Voldgade 5-7, 1350 Copenhagen, Denmark; [10]Natural History Museum, University of Oslo, PO Box 1172 Blindern, N-0318 Oslo, Norway; [11]CE3C – Centre for Ecology, Evolution and Environmental Changes, Departamento de Biologia Animal, Faculdade de Ciências, Universidade de Lisboa, 1749-016 Lisboa, Portugal; [12]Institut des Sciences de l'Evolution de Montpellier (ISEM) – UM-CNRS-IRD-EPHE, Université de Montpellier, Place Eugène Bataillon – CC 64, 34095 Montpellier cedex 05,France; [13]Breeding

Centre for Endangered Arabian Wildlife, PO Box 29922, Sharjah, United Arab Emirates; [14]Department of Veterinary Surgery and Theriogenology, College of Veterinary Medicine, University of Agriculture Makurdi, Benue State, Nigeria; [15]Centre for Ecological and Evolutionary Synthesis (CEES), Dept. of Biosciences, University of Oslo, PO Box 1166 Blindern, N-0316 Oslo, Norway; [16]Trace and Environmental DNA Laboratory, Department of Environment and Agriculture, Curtin University, Perth, Australia; [17]NTNU University Museum, N-7491 Trondheim, Norway;

* These authors contributed equally to this work.

¶ **Correspondence:** qixiaopeng@mail.kiz.ac.cn**,** wanggd@mail.kiz.ac.cn and zhangyp@mail.kiz.ac.cn

Abstract

Natural selection in domestic dogs is of great interest in evolutionary biology since dogs have migrated to every inhabited continent of the world alongside humans, and adapted to diverse environments. Here, we explored their demographic history and genetic basis of adaptation to the tropical African environment using whole genome analyses of 19 African indigenous dogs from Nigeria. Demographic analysis suggests that the ancestors of these dogs migrated into Africa from Eurasia 14,000 years ago and underwent a severe founder effect before population expansion. Admixture analysis further reveals that African dog genomes contain about 1.88%-3.50% introgression from African golden wolves (*Canis anthus*). Population genetic analysis identifies 50 positively selected genes linked with immunity, angiogenesis, ultraviolet protection, as well as insulin secretion and sensitivity that may contribute to adaptation to tropical conditions. One of the positively selected genes, *ADGRE1* (adhesion G protein-coupled receptor E1), has also been found to be association with severe malaria resistance in African human populations. Functional assessments showed that *ADGRE1* provides protective host defense against *Plasmodium* infections. This result, together with the fact that the inflammatory response to canine babesiosis is similar to complicated falciparum malaria in humans, support the dogs as a model for the study of malaria control and treatment.

Keywords: African dogs; adaptive evolution; anti-parasite; demographic history.

Introduction

The domestic dog (*Canis lupus familiaris*) dispersed across the planet alongside human migrations, and in doing so must have had to undergo adaptation to diverse environments. Natural selection played key roles in shaping the fitness of the domestic dog to these environments, with the underlying mechanisms being of great interest in evolutionary biology (Freedman et al. 2016). For example, previous studies on the adaptation of Tibetan dogs to the hypoxic Tibet Plateau consistently revealed *EPAS1* and *HBB* as candidate genes for hypoxia adaptation (Gou et al. 2014; Wang et al. 2014). Moreover, Wang et al. (2014) found that the hemoglobin levels were very similar between Tibetan dogs and human, suggesting a convergent evolution between dogs and humans living on the Tibetan Plateau.

The African continent is prone to a number of selective forces that have led to unique human genetic adaptations. One of the most pervasive is malaria – in most African regions malaria is endemic, and World Health Organization (WHO) estimates indicate that in 2015, 88% of global malaria cases and 90% of global malaria deaths occurred in Africa (WHO 2015). Previous studies have identified several genes associated with immune responses that protect some African populations against malaria (Kusi et al. 2008; Machado et al. 2010; Band et al. 2013; Kariuki et al. 2013; Apinjoh et al. 2014; Carstensen et al. 2014). In Africa, there also exists canine babesiosis, a malaria-like disease caused by the tick-borne parasite *Babesia*, which, like *Plasmodium*, is an apicomplexan hemoprotozoan parasite (Sasaki et al. 2007). It has been noted that infection rates of *B. canis rossi* and *B. canis vogeli* are lower in dogs from Nigeria than in those in South Africa or Sudan (Sasaki et al. 2007), suggesting that there might be differential local adaptation against the parasite across the African continent. A further well-documented series of genomic encoded human adaptations relate to protection against strong ultraviolet radiation (UVR) in the equatorial region (Norton et al. 2006; Norton et al. 2007; Jablonski and Chaplin 2013), including maintaining a dark skin and kinky hair texture. The latter potentially confer protection to the brain against thermal stress (Jablonski and Chaplin 2014). The African cattle also revealed signatures of selection for environmental adaptive traits including coat color, heat tolerance, and tick resistance (Kim et al. 2017).

Domestic dogs are exotic to Africa, having first been domesticated from Eurasian gray wolves between 15,000 and 40,000 years ago (Vilà et al. 1997; Germonpré et al. 2009). Little is currently known about the origin of African domestic dogs. Besides that, wolves have been considered to be absent in Africa (Mech 1981; Wang et al. 2016b). Although genome-wide analyses have estimated that a subset of ancestral dogs migrated towards the Middle East, Europe, and into Africa around 15 000 years ago (Wang et al. 2016), the oldest archaeological evidence for African dogs was found in Egypt, dating ca. 6300-5600 BC (Mitchell 2015). Therefore, considerable gaps exist in our understanding of the history of domestic African dogs. For example, a recent study based on 382 African dog mitochondrial D-loop sequences found that West African dogs share a sub-haplogroup with southwest European dogs and inferred possible gene flow between West African dogs and southwest European dogs around 12,000 years ago (Adeola et al. 2016). However, mitochondrial DNA (mtDNA) studies are limited as they only account for a small proportion of the genetic material of an individual. While analysis of whole-genome data could further contribute to our understanding, currently whole-genome information is only available for four African village dogs. Hence, additional nuclear genomes of African village dogs are needed to further investigate their origin and diversity. Moreover, dogs in Africa may possess adaptations related to protective mechanisms against selective forces like babesiosis and strong UVR. Such adaptations are largely unknown, but by identifying genomic regions under selection, genes possibly associated with adaptive responses could be revealed.

Whole genome sequencing provides a powerful, holistic approach to understand the demographic history and natural selection of populations and species. Here, we sequenced the genomes of 15 indigenous dogs from Nigeria and four African golden wolves (*Canis anthus*) (Gaubert et al. 2012; Rueness et al. 2012; Koepfli et al. 2015) to study their demographic history and adaptive evolution, as well as possible gene flow between the two species.

Results

Sample collection and whole genome sequencing

We sampled 15 indigenous dogs from six states in Nigeria (supplementary Table S1) and four African golden wolves from Algeria, Egyptian Sinai, Morocco, and Senegal. We performed whole-genome sequencing to an average depth of 13× for each sample after removing PCR redundancy (supplementary Table S1). Since tissue samples of three African golden wolves were collected from dead subjects, DNA damage was possible. We therefore analyzed C to T and G to A mutation enrichment near the end of the sequence reads (Binladen et al. 2006) , which confirmed no DNA degradation of these samples (supplementary Fig. S1).

We also incorporated 36 published whole-genome sequencing data of dogs and gray wolves (Fig. 1A: four Nigerian village dogs, one African ancient breed -Basenji, seven gray wolves, 14 European breeds and Middle Eastern village dogs, and 10 indigenous dogs from Yingjiang in southern China) (Auton et al. 2013a; Wang et al. 2013; Freedman et al. 2014; Gou et al. 2014; Wang et al. 2016a). These samples cover all major Eurasian dog and wolf groups. After strict filtering, we identified ~24.4 autosomal million SNPs for further analysis (supplementary Table S2).

Population structure and phylogenetic analysis

Principal component analysis (PCA) was performed to explore the relationships among the 55 individuals. In a two-dimensional plot of the genotypes, there was a clear separation between dogs, Eurasian gray wolves (EGW), and African golden wolves (AGW) (Fig. 1B). The dogs split into three groups along the second dimension: (i) Chinese indigenous dogs from Yingjiang (CIDY), (ii) Middle Eastern village dogs (MEVD) and European breeds (EB), and (iii) the Basenji together with the Nigerian indigenous dogs (NID). We combined our dataset with previous SNP array data for an expanded dataset (Vonholdt et al. 2011) and observed a clustering of dog populations according to their geographic affinities (supplementary Fig. S2).

A maximum-likelihood phylogenetic tree agreed with the PCA results. At the base of the tree, dogs and Eurasian wolves were separated from African golden wolves (Fig. 1C,

supplementary Fig. S3). Dogs then diverged from Eurasian wolves and split into three clades. Chinese indigenous dogs from Yingjiang firstly split from other dogs, suggesting that the Yingjiang dogs may represent a very ancient divergence. The Middle Eastern village dogs and European breeds formed a distinct sister clade to African dogs, suggesting that the latter had an evolutionary history independent of the main process of dog domestication and diversification.

To explore the genetic relationships among individuals, we performed a structure analysis using a block relaxation algorithm to cluster individuals into different numbers of groups (supplementary Fig. S4). Five different clusters were inferred: (i) African golden wolves, (ii) Eurasian wolves, (iii) Yingjiang indigenous dogs, (iv) Middle Eastern village dogs and European breeds, and (v) African dogs (Fig. 1D). This is consistent with the result of PCA and phylogenetic analysis.

Genetic diversity

In accordance with the above results, we grouped individuals into five populations (African golden wolves, Eurasian wolves, European breeds and Middle Eastern village dogs, Yingjiang indigenous dogs, and African dogs) and estimated nuclear diversity using the parameter $\theta\pi$. As shown in Fig. 1E, there were significant differences ($P<0.001$) in the levels of genetic diversity between populations. Both African golden wolves and Eurasian gray wolves exhibited high nuclear diversity (mean: 2.2×10^{-3} and 1.8×10^{-3}, respectively). Yingjiang indigenous dogs had the highest diversity among the three dog populations, at ~77.8% of the diversity in Eurasian gray wolves. While the genetic diversity of Middle Eastern village dogs and European breeds was lower than that of Yingjiang indigenous dogs, it was higher than in Nigerian indigenous dogs. These results suggest that the ancestors of the Nigerian indigenous dogs might have undergone a founder effect after their divergence from Eurasian dogs.

Admixture and demographic history

To test for gene flow between populations, we performed D-statistics analysis (Patterson

et al. 2012). The results indicate gene flow between African golden wolves and Nigerian indigenous dogs (D=-0.013 to -0.019, Z>=8.66), as well as between Yingjiang indigenous dogs and European breeds together with Middle Eastern village dogs (D=-0.057, Z=20.68) (supplementary Table S3).

Because the D-statistics could not identify the direction of gene flow between African golden wolves and dogs, we used the joint site frequency spectrum (SFS) approach implemented in fastsimcoal2 to simulate demographic history (Excoffier et al. 2013). A total of 12 models of historical divergence and various gene flows scenarios were fitted to the allele-frequency spectrum of five populations (supplementary Fig. S5). A demographic model (Fig. 2) in which mutual gene flow existed between African golden wolves and Nigerian dogs produced a better fit than the alternative models (supplementary Table S4). In the best fitting model, domestication of dogs occurred in the last 31,100 years, and the divergent time of Chinese dogs and other dogs is approximated at 16,600 years ago (supplementary Table S5). This analysis further indicated that about 14,800 years ago [95% CI 13,368-14,832], dogs migrated to Africa and introgressed with African golden wolves, followed by a 1.3 fold population expansion.

F4-ratio test were performed to calculate the proportion of gene flow between African golden wolves and African dogs (Brisbin and Bustamante 2012). As result, F4-ratio estimated about 5% ancestry from AGWs in NIDs (supplementary Table S6). PCAmix was used to identify segments of the African dogs' genome belonging to AGWs or EBs (Brisbin and Bustamante 2012). The ratio of the ancestral AGWs in African dogs range from 1.86% to 2.94% (average 2.45%), which is similar to the estimates from of fastsimcoal2 (1.43%, 95% CI [1.88%-3.50%], supplementary Table S5).The introgressed regions were about ~ 12.62 kb in size, and dispersed randomly across the genome. The lengths of regions range from 11.77kb to 12.26kb on average (supplementary Fig. S6).

Adaptive selection

We used Sweep Detector (SweeD) and the population branch statistic (PBS) (Yi et al. 2010; Pavlidis et al. 2013) to identify positive selection in African indigenous dogs.

Considering the top 1% outliers in each method for each gene annotation, SweeD identified 399 genes and the PBS approach found 580 genes. Fifty genes are common in the two methods and as considered as potential candidate genes (supplementary Table S7).

Gene ontology (GO) identified significant over-representation of genes involved in local environmental adaptation (Table 1). A number of relevant observations from other studies on the activity of these genes hint at their possible functions. For instance, enzyme activator activity (GO: 0008047) could be affected by infection. The expression profile of pig lung tissues post-inoculation with *Actinobacillus pleuropneumoniae* showed a significant representation of genes belonging to this GO term (Zuo et al. 2012). Endoplasmic reticulum (GO: 0005783) plays key roles in important processes like protein transport and energy metabolism. The mRNA expression of GO: 0005783 in mice is altered after heat treatment (Yu et al. 2011). To detect the convergent evolution between African human and dogs, we compared our list of GO terms in dogs with that from African humans compiled by Barreiro et al 2008. After performing GO enrichment analysis on the positively selected genes (PSGs) in African humans (Barreiro et al. 2008), we identified 34 enriched GO terms that showed statistical significance (supplementary Table 7). Interestingly, two terms: GTPase regulator activity (GO:0030695) and nucleoside-triphosphatase regulator activity (GO:0060589), both exist in the overlapping set between the two species (150.00 fold enrichment at a significance of P=6.42E-5) (Barreiro et al. 2008).

Among the fifty positively selected genes identified in African dogs, three of them may play roles in innate or adaptive immune responses: adhesion G protein-coupled receptor E1 (*ADGRE1*), caspase recruitment domain family member 9 (*CARD9*), and vav guanine nucleotide exchange factor 1 (*VAV1*) (Mesecke et al. 2011; Hafalla et al. 2012; Kariuki et al. 2013). The human homologues of three genes are highly expressed in spleen and bone marrow (supplementary Fig. S7).In the *ADGRE1* homolog in mice (*F4/80*) is required for the differentiation of antigen-specific $CD8^+$ regulatory T cells (Lin et al. 2005). A large multicenter case-control study revealed that *ADGRE1* is associated with severe malaria (Network 2014). *ADGRE1* has also been linked with complex malaria-associated seizures in African children (Kariuki et al. 2013). Moreover, *ADGRE1* has been associated with

hyperpyrexia, severe malaria anaemia, and uncomplicated malaria in a Cameroonian population (Apinjoh et al. 2014), and its expression was induced in women with placental malaria infection (Muehlenbachs et al. 2007). The expression of *ADGRE1* in mice is significantly increased by approximately 3.2-fold after eight days following infection with *Plasmodium chabaudi* (Al-Quraishy et al. 2013).

Because positive selection can be difficult to distinguish from genetic drift produced by neutral processes related to a specific demographic history (Freedman et al. 2016; Wu et al. 2016). We performed simulations to generate 40kb DNA sequences for 1,000,000 repetitions to calculate PBS, and to 1mb sequences for 10,000 times to cacluate Sweed, based on demographic history as the posterior distributions (Ewing and Hermisson 2010). We then used these data to test whether the observed values of PBS and Sweed for *ADGRE1* gene could be due to genetic drift instead of selection. As a result, the probability of the observed values of PBS and Sweed attributable to genetic drift was only 4E-4 and 3.17E-3, respectively (supplementary Fig. S8), implying that the genetic pattern observed on *ADGRE1* was not due to genetic drift but a target of positive selection.

ADGRE1 gene defense against *Plasmodium* infections

The dog *ADGRE1* gene is 3,489 bps long and contains 12 calcium-binding EGF domains (EGFCA), which exceeds that of mice (6 EGFCA) and humans (7 EGFCA) (supplementary Fig. S9). For an in-depth examination of the role of dog ADGRE1 protein in the host defense against parasitic infection, we performed ectopic expression of dog *ADGRE1* gene in a murine macrophage cell line (RAW 264.7 macrophages) (Fig. 3A). Phagocytosis of pathogens by macrophages induces innate immune response, which in turn activates adaptive immunity. Although the murine homolog of *ADGRE1* (F4/80) is predominantly expressed on eosinophils in mice, it is widely used as a cell surface marker of macrophage populations in mice. The precise role of F4/80 in innate and adaptive immunity remains elusive (Murray and Wynn 2011). To determine whether the dog membrane protein ADGRE1 is involved in the phagocytosis of *Plasmodium* parasites, RAW 264.7 macrophages transfected with *GFP* and *ADGRE1* were infected with *Plasmodium berghei ANKA*, and the expression of *18S rRNA*

and *Hsp70* specific to *P. berghei ANKA* was analyzed by real-time quantitative PCR. The expression levels of *Plasmodium 18S rRNA* and *Hsp70* were significantly higher in *ADGRE1* transfected cells compared to those transfected with *GFP* (Fig. 3B). This result indicates that the phagocytosis of *Plasmodium* parasites is increased in the presence of dog ADGRE1.

Type I interferon plays important roles in various infectious diseases, including malaria. Intracellular DNA and RNA sensors such as cGAS, MDA5, and TLR7 are crucial activators for type I interferon production in response to *Plasmodium* infection (Gun et al. 2013; Yu et al. 2016). To determine whether increased intracellular parasites in *ADGRE1* transfected cells results in a higher amount of type I interferon production, we analyzed *Ifnb* and *Cxcl10* expression in *GFP* and *ADGRE1* transfected cells and found that *ADGRE1* indeed promoted *Ifnb* and *Cxcl10* expression during *Plasmodium* infection (Fig. 3C). Guanylate binding proteins (GBPs) are type I interferon inducible proteins associated with pathogen-containing vacuoles and are crucial for host defense against pathogens (Degrandi et al. 2008; Haldar et al. 2013). In line with the increased phagocytosis and type I interferon production in *ADGRE1* transfected cells, the expression of *Gbp1* and *Gbp3* was markedly increased (Fig. 3C). Type I interferon signaling plays a central role in neutrophil activation and malaria pathogenesis by mediating the production of proinflammatory cytokines such as IL-1 members and IL-6, and neutrophil migration (Rocha et al. 2015). Accordingly, the expression of *Il1a*, *Il1b* and *Il6* was increased in *ADGRE1* transfected cells in response to *Plasmodium* infection (Fig. 3C). In contrast, the expression of chemokine *Cxcl1* was comparable between *GFP* and *ADGRE1* transfected cells (Fig. 3C). Moreover, the other critical innate immune responses during pathogenic infection including autophagy, lysosomal biogenesis, and apoptotic cell death pathways were not significantly affected by ectopic expression of *ADGRE1* in response to *Plasmodium* infection (Fig. 3D).

To identify the SNP which possibly enhances the function of *ADGRE1*, we used Fst values to identify highly differentiated sites between African dogs and European breeds, and took the top 1% of sites for annotation. We found a nonsynonymous G655A mutation within the *ADGRE1* gene. This mutation had a very high allele frequency in African dogs (89.5%) compared to 25.0% in European breeds (Fig. 4A). We subsequently explored whether this

mutation contributes to elevated resistance to malaria in African dogs through the ectopic expression of mutated *ADGRE1-G655A* in RAW 264.7 macrophages and analysis of its effect on innate immune response during parasite infection. Interestingly, the expression of *Plasmodium Hsp70*, host *Ifnb* and its inducible genes *Gbp1*, *Gbp2*, *Gbp3*, and *Gbp5*, and proinflammatory cytokine gene *Il6* were higher for *ADGRE1-G655A* than for *ADGRE1* (Fig. 4B). Collectively, these results suggest that dog ADGRE1 protein mediates *Plasmodium* internalization and type I interferon production, the latter being a master regulator of the immune response against malaria in dogs.

Besides the three genes involved in immune responses, we also detected 13 PSGs from 50 candidate genes that are functionally involved in angiogenesis, ultraviolet protection and the secretion sensitivity to insulin. Collagen and calcium binding EGF domains 1 (*CCBE1*), plexin domain containing 2 (*PLXDC2*), and Ras and Rab interactor 2 (*RIN2*) play crucial roles in angiogenesis (Hogan et al. 2009; Sandri et al. 2012; Cheng et al. 2014). Angiogenesis is extremely important for thermoregulation of warm-blooded animals (Arens and Zhang 2006). We also identified seven genes associated with DNA repair and melanin accumulation that we hypothesize may provide protection from the negative effects of exposure to ultraviolet light. Those genes were the ubiquitin carboxyl-terminal esterase L3 (*UCHL3*), LON peptidase N-terminal domain and ring finger 1 (*LONRF1*), SNF2-related CREBBP activator protein (*SRCAP*), ankyrin repeat domain 32 (*ANKRD32*), dystonin (*DST*), membrane-bound transcription factor peptidase site 1 (*MBTPS1*), and phospholipid phosphatase-related protein type 5 (*LPPR5*) (Lalonde et al. 2005; Sano et al. 2006; Matsuoka et al. 2007; Brandl et al. 2009; Weger et al. 2011; Park et al. 2013; Räschle et al. 2015). Several PSGs such as potassium voltage-gated channel interacting protein 1 (*KCNIP1*), nucleobindin-1 (*NUCB1*), and an enhancer of mRNA decapping 3 (*EDC3*) were associated with insulin secretion and sensitivity (Heun-Sik et al. 2014; Ramesh et al. 2015; Kim et al. 2016). In humans, it has been argued that these genes relate to specific traditional African diets and are linked to the current health challenges among African American populations as explained by the thrifty gene hypothesis (Marshall 2005).

Discussion

African dogs are a distinct population

In this study, we applied whole genome sequencing and performed population genomic analysis of 15 African indigenous dogs and four African golden wolves. We observed a split into three major dog groups: (i) southern Chinese indigenous dogs, (ii) a cluster containing Middle Eastern village dogs and European breeds, and (iii) African dogs (Fig. 1B, C, and D). The ancestors of the African dogs studied here were estimated to have entered the African continent about 14,000 years ago (Fig. 2). This is consistent with the findings of earlier studies on dog population structure and history based on mtDNA (Adeola et al. 2016), and full nuclear genomes (Wang et al. 2016). Interestingly, human Y chromosome haplotypes provide evidence of human migration to the Mediterranean coast of North Africa at the end of the Pleistocene (Underhill et al. 2001), and genome-wide SNPs analysis suggested a ‘‘back-to-Africa’’ migration more than 12,000 years ago (Henn et al. 2012). Therefore, it is plausible that the entry of dogs into Africa was in the company of the migrating humans. We detected gene flow from both African golden wolves and European breeds/Middle Eastern village dogs into the African dogs. A previous mtDNA study showed that West African dogs share a sub-haplogroup with southwestern European dogs, possibly due to gene flow events (Adeola et al. 2016).

African dogs have the lowest genetic diversity

In our study, the genetic diversity and effective population size of African indigenous dogs are the lowest among the major dog groups (Fig. 1E and Fig. 2). Genetic diversity was higher in the Middle Eastern village dogs/European breeds, and highest in the Chinese dogs. The lowest genetic diversity of African dogs may be due to in the availability of samples only from Nigeria out of the vast African continent. But the dogs from Yingjiang, only a small county in China, have the highest diversity. This genetic diversity gradient is derived from studies of mtDNA (Savolainen et al. 2002; Pang et al. 2009), Y chromosome (Ding et al. 2012), and whole-genome (Wang et al. 2016a). Moreover, the effective population sizes of dogs also show a similar gradual upward trend from African dogs, to European breeds/Middle

Eastern village dogs, and then southern Chinese dogs (Fig. 2). One study reported similar mtDNA haplotype diversity in African and East Asian village dogs (Boyko et al. 2009). But another study of dog mtDNA diversity found that 318 African village dogs represented 41 haplotypes, while 281 dogs sampled from southern China showed 71 haplotypes (Pang et al. 2009).

Convergent evolution between African dogs and humans

The tropical zones of Africa present considerably different environmental conditions than the Eurasian range. These conditions include a hotter climate, increased exposure to UVR, and a wide breadth of novel parasites and pathogens. With dogs sharing a common environment, food, and immunologic profiles with human (Storb and Thomas 1985), natural selection may work on a similar evolutionary direction among the two species. In this study, genes associated with insulin secretion and sensitivity, immunity, angiogenesis, and ultraviolet protection showed adaptive selection (supplementary Table S8). These genes were significantly clustered in seven GO terms (Table 1). Two of these GO terms also existed in the GO terms of African human selection genes. Both terms modulate the rate of guanosine triphosphate (GTP) hydrolysis. GTP is involved in processes like signal transduction, genetic translation, energy transfer within the cell. Moreover, the two terms also associated with climate-mediated selection in sheep (Lv et al. 2014). This suggests possible convergent evolution between African dogs and humans under the same environmental exposures.

Similar evolutionary adaptations have been reported in some African and African-derived human populations. For example, African American children have higher level of insulin and insulin resistance than those of various non-Africans populations (Svec et al. 1992; Arslanian and Suprasongsin 1997; Arslanian et al. 1997; Gower et al. 1999). Also, many African populations are generally characterized by darker integumentary phenotypes as a protective adaptation against strong UVR (Norton et al. 2006; Norton et al. 2007; Jablonski and Chaplin 2013). Other studies have demonstrated evidence of selection in genes associated with immune responses that protect African human populations from malaria and other diseases (Kusi et al. 2008; Machado et al. 2010; Band et al. 2013; Carstensen et al. 2014). In particular,

the *ADGRE1* gene, one of the strongly positively selected genes in African dogs, has been proved to be associated with immune response to malaria in African humans (Kariuki et al. 2013). In a previous case-control study, a relationship between *ADGRE1* and malaria-associated seizures in African children was revealed (Kariuki et al. 2013). Our study suggests that *ADGRE1* contributes to defense against *Plasmodium* infection in African dogs which could be a suitable model for biomedical research for malaria.

In summary, based on whole-genome analyses, this study expands our understanding of the genetic diversity, evolutionary history, and tropical adaptation of African dogs. The inflammatory-response against canine babesiosis is similar to that of complicated falciparum malaria in humans (Reyers et al. 1998), and dogs generally share approximately 360 diseases with humans (Shearin and Ostrander 2010). There are several advantages of *Canis familiaris* as a model for genetic susceptibility to disease, like greater homology to human, susceptibility to many diseases with humans variants (Shearin and Ostrander 2010; Boyko 2011). The Lupa project aimed to enhance the use of the dog as an effective model to study common complex diseases in human and has got some achievements so far (Shearin and Ostrander 2010; Boyko 2011; Lequarre et al. 2011). Dogs are thus not only a useful species for mapping disease loci, but also an excellent model for biomedical research of human diseases.

Materials and Methods

Sample collection and sequencing

We sampled 15 indigenous dogs from six different states in Nigeria (Oyo, Ondo, Akwa Ibom, Cross River, Taraba, and Ekiti) and four wolves from different countries in Africa (Algeria, Egypt, Morocco and Senegal). Blood samples were collected from each dog and African golden wolf from Egypt. Tissue samples were collected from other African golden wolves accidentally knocked down on roads. Total genomic DNA was extracted from blood samples using the phenol-chloroform method, and 1-3 μg of DNA from each individual was sheared into fragments using 200–800 bps with the Covaris system. Tissue samples were extracted using the DNeasy Blood and Tissue kit (Qiagen, Hilden, Germany), and DNA was

fragmented into 300-400 bp fragments using a Bioruptor NGS sonicator (Diagenode, Denville, NJ, USA). The DNA fragments were then sequenced using the Illumina Hiseq 2000 or 2500 platforms. Sequencing data from this study have been submitted to the Genome Sequence Archive (GSA, http://gsa.big.ac.cn/) under project number PRJCA000335.

Sequence data pre-processing and variant calling

Raw sequence reads were mapped to the dog reference genome (Canfam3) using the BWA-MEM version 0.7.10-r789 (Li 2013). Reads with identical start/end points were filtered using PICARD (version 1.87). Sequences were then locally realigned and base-recalibrated using the Genome Analysis Tool Kit (GATK, version 2.5-2-gf57256b) (Depristo et al. 2011). Specifically, after genome alignment and removing PCR duplicates, the distribution of misincorporation near the ends of the reads were carried out by mapDamage2.0 (Jónsson et al. 2013). Then, variant calling of sequence data were handled using the UnifiedGenotypeCaller in GATK. During base and variant recalibration, a list of known SNPs/indels was downloaded from the Ensembl database to serve as a training set.

Genetic diversity and population structure

Genetic diversity was calculated from a non-overlapping 40 kb windows across the genome using VCFtools v0.1.12b (Danecek et al. 2011). Principal component analysis was carried out using the smartPCA program from the EIGENSOFT package v5.0.1 (Patterson et al. 2006). Maximum-likelihood phylogenetic tree was built by SNPhylo (Lee et al. 2014) and dhole (*Cuon alpinus*) was used as the outgroup. After thinning to a single SNP per 50 kb window, population structure analysis was performed using the block relaxation algorithm implemented in the ADMIXTURE software (Alexander et al. 2009).

Evolutionary history

SNPs located 10kbs away from genes, were used to convert SFS by easySFS (https://github.com/isaacovercast/easySFS#easysfs). To mitigate the effect of linkage disequilibrium, we took one SNP every 10kb. Demographic history was simulation by

fastsimcoal2 (Excoffier et al. 2013). Mutation ratio was set to 6.6E-9 per sites per generation and generation time as 3 years (Kumar and Subramanian 2002; Wang et al. 2013; Wang et al. 2016a). We used a recombination of 9.7E-9 (Wong et al. 2010). Alternative models of historical events were fitted to the joint SFS of dogs, grey wolves, and African golden wolves. For each model, we ran the programme 50 times with varying starting points to ensure convergence, and retained the fitting model with the highest likelihood. Demographic estimates were obtained from 100,000 simulations per likelihood estimation (-n100, 000, -N100, 000), 40 Expectation/Conditional Maximization (ECM) cycle (-L40) and 50 runs per data set. The best model was selected through the maximum likelihood likelihoods value and Akaike information criterion (Excoffier et al. 2013).

PCAmix

Given the genotype information across the genome for each individual, we used SHAPEIT (v2.r790) to phase the genotypes into associated haplotypes with parameters: windows 0.5, effective-size 83600 and genetic maps from (http://autonlab.einstein.yu.edu/dog_recomb/) (Auton et al. 2013b). We then performed PCAmix to estimate the segments of the African dogs' genome belonging to AGW and EB&MEVD used haplotype data. Because the sample size of AGW (4 individuals) is greatly smaller than the samples size of EB&MEVD (14 individuals), we split EB&MEVD into four individuals per groups to before running PCAmix to avoid the effect of various simple size variation. We subsequently averaged the ratios and genomic segment length from AGW in African dogs. Segments with confidence >=0.9 were keep and others were labeled as 'undecided'.

Positive selection

SweeD (v3.2.12) was used to identify regions of the genome that showed the strongest signals of selective sweeps (Pavlidis et al. 2013). Dhole was used to distinguish the unfolded SNPs among the African dogs' SNPs. The grid size was set to the number of SNPs on each chromosome. PBS was calculated for African dogs, gray wolves, and European breeds, using 40 kb window size and 20kb stepwise increments (Yi et al. 2010). We extracted the outliers

that ranked in the top 1% of PBS or of the SweeD likelihood for subsequent separate gene annotations. For gene annotations, we extracted the genes which formed part of or completely overlapped with the top1% PBS windows or contained the top 1% Sweed sites. The genes present in both gene sets were considered significant candidate genes under positive selection.

Gene Ontology analysis was carried out using DAVID v6.7 (Huang et al. 2009). The terms with P-values < 0.05 were considered significantly enriched. The terms with P-values < 0.05 were considered significantly enriched. To detect convergent evolution between African human and dogs, selection regions of African humans were obtained from a previous publication (Barreiro et al. 2008). After gene annotation, GO enrichment analysis was done as dogs.

Simulation

We performed simulations for PBS to generate 40kb DNA sequences for 1,000,000 times based on the demographic history as the posterior distributions by the software msms (Ewing and Hermisson 2010). For each run, we calculated PBS values by 40kb windows from the simulated data sets. For Sweed, we simulated 1MB sequences for 10,000 times. African golden wolves were used as outgroup to distinguish folded or unfolded of African dogs' SNPs, and then SweeD was performed using unfolded site in African dogs for each 1MB sequence. P-values were got by calculating the proportion sets in simulation data larger than observed value.

Plasmid construction and retroviral infection

The full-length cDNA of the dog *ADGRE1* gene was amplified from cDNA of dog liver tissue using the primers *ADGRE1* forward (5'-AAATAGATCTATGTGGAGCTTCAACTTGCTCCTC-3'), and *ADGRE1* reverse (5'-AAATGCGGCCGCTTAATCCGTCTTAGAAGTGGAGGGG-3').

The PCR product was digested with BglII and NotI (restriction sites are underlined in the primer sequences) and cloned into the retroviral expression vector MSCV2.2 to obtain MSCV-ADGRE1. The cloned construct was verified by DNA sequencing. For generation of

the *ADGRE1-G655A*, the 1072 bp of the 5' portion of the *ADGRE1* gene was synthesized and cloned into MSCV-ADGRE1 by replacing the analogous fragment. Retrovirus production in 293T cells and infection of RAW 264.7 macrophages was performed as previously described (Qi et al. 2013). GFP positive cells were isolated by fluorescence activated cell sorting (FACS) and expanded for further analysis.

Plasmodium culture and infection of cells

P. berghei ANKA was cultured as described previously (Jiang et al. 2013). All animal experiments were conducted in accordance with the guidelines of animal care from the Institut Pasteur of Shanghai, Chinese Academy of Sciences and were approved by the Animal Care and Use Committee of the Kunming Institute of Zoology, Chinese Academy of Sciences. For parasite infection, infected red blood cells from mice were suspended in phosphate-buffered saline and subsequently administered to the RAW 264.7 macrophages for stimulation as described previously (Yao et al. 2016).

Real time quantitative PCR

The total RNA was isolated from RAW 264.7 macrophages using TRIzol reagent (Invitrogen), and cDNA was reverse transcribed using Superscript III reverse transcriptase (Invitrogen). Real time quantitative PCR was performed using the CFX96 Real-Time System (BIO-RAD). Primer sequences are listed in supplementary Table S9.

Immunoblot analysis and antibodies

Samples were separated by 12% SDS-PAGE and then electrophoretically transferred onto polyvinylidene fluoride membranes. Membranes were blocked with 5% nonfat milk and then incubated overnight in primary antibody at 4°C. The following primary antibodies were used: anti-LC3B (NB600-1384; Novus Biologicals), anti-caspase 3 (9661S, 9491S, and 8592S; Cell Signaling Technology), anti-GFP antibody (SC-9996, Santa Cruz Biotechnology), anti-TFEB (A303-673A; Bethyl Laboratories, Inc.), and anti-GAPDH (5174S; Cell Signaling Technology). The secondary antibodies used were HRP-labeled anti-rabbit antibodies

(Jackson ImmunoResearch Laboratories, Inc.).

Statistical analysis

Data was given as means ± SD. Statistical analyses were performed using two-tailed Student t tests. P-values <0.05 were considered significant.

Acknowledgements

This work was supported by grants from the NSFC (91531303), and the Breakthrough Project of Strategic Priority Program of the Chinese Academy of Sciences (XDB13000000), the Chinese Academy of Sciences (CXJJ-17-M141) and (Y602381081), Sino-Africa Joint Research Center, Chinese Academy of Sciences (SAJC201611). ERC Consolidator Grant (681396) 'Extinction Genomics', Marie Skłodowska-Curie Actions Fellowship (655732) 'Wherewolf' and Carlsbergfondet Fellowship (CF14-0995). G-D W and M-S P are supported by the Youth Innovation Promotion Association, Chinese Academy of Sciences. N.O.O. is supported by the CAS-TWAS President's Fellowship Program for Doctoral Candidates. CF acknowledges support from Fundação para a Ciência e Tecnologia (FCT, MCTES, Portugal) and Faculdade de Ciências da Universidade de Lisboa through, respectively, the Ciência 2007 contract C2007-UL-342-CBA1 and the contract as Invited Assistant Professor.

References

Adeola AC, Ommeh SC, Song J-J, Charles OS, Sanke OJ, Yin T-T, Wang G-D, Wu S-F, Zhou Z-Y, Lichoti JK. 2016. A cryptic mitochondrial DNA link between North European and West African dogs. J Genet Genomics.

Al-Quraishy S, Dkhil MA, Abdel-Baki AAS, Delic D, Santourlidis S, Wunderlich F. 2013. Genome-wide screening identifies Plasmodium chabaudi-induced modifications of DNA methylation status of Tlr1 and Tlr6 gene promoters in liver, but not spleen, of female C57BL/6 mice. Parasitology research. 112: 3757-3770.

Alexander DH, Novembre J, Lange K. 2009. Fast model-based estimation of ancestry in unrelated individuals. Genome Res. 19: 1655-1664.

Apinjoh TO, Anchang-Kimbi JK, Njua-Yafi C, Ngwai AN, Mugri RN, Clark TG, Rockett KA, Kwiatkowski DP, Achidi EA, Consortium M. 2014. Association of candidate gene polymorphisms and TGF-beta/IL-10 levels with malaria in three regions of Cameroon: a case–control study. Malar J. 13: 236.

Arens E, Zhang H. 2006. The skin's role in human thermoregulation and comfort. In: Pan N, Gibson P, editors. Thermal and Moisture Transport in Fibrous Materials. Woodhead. p. 560-602.

Arslanian S, Suprasongsin C. 1997. Differences in the in vivo insulin secretion and sensitivity of healthy black versus white adolescents. J Pediatr. 130: 440-443.

Arslanian S, Suprasongsin C, Janosky JE. 1997. Insulin secretion and sensitivity in black versus white prepubertal healthy children. J Clin Endocrinol Metab. 82: 1923-1927.

Auton A, Li YR, Kidd J, Oliveira K, Nadel J, Holloway JK, Hayward JJ, Cohen PE, Greally JM, Wang J. 2013a. Genetic recombination is targeted towards gene promoter regions in dogs. PLoS Genet. 9: e1003984.

Auton A, Li YR, Kidd J, Oliveira K, Nadel J, Holloway JK, Howard JJ, Cohen PE, Greally JM, Wang J. 2013b. Genetic recombination is targeted towards gene promoter regions in dogs. arXiv preprint arXiv:1305.6485.

Band G, Le QS, Jostins L, Pirinen M, Kivinen K, Jallow M, Sisayjoof F, Bojang K, Pinder M, Sirugo G. 2013. Imputation-Based Meta-Analysis of Severe Malaria in Three African Populations. PLoS Genet. 9: e1003509.

Barreiro LB, Laval G, Quach H, Patin E, Quintanamurci L. 2008. Natural selection has driven population differentiation in modern humans. Nat genet. 40: 340-345.

Binladen J, Wiuf C, Gilbert MT, Bunce M, Barnett R, Larson G, Greenwood AD, Haile J, Ho SY, Hansen AJ. 2006. Assessing the fidelity of ancient DNA sequences amplified from nuclear genes. Genetics. 172: 733-741.

Boyko AR. 2011. The domestic dog: man's best friend in the genomic era. Genome Biol. 12: 216-216.

Boyko AR, Boyko RH, Boyko CM, Parker HG, Castelhano M, Corey L, Degenhardt JD, Auton A, Hedimbi M, Kityo R. 2009. Complex population structure in African village dogs and its implications for inferring dog domestication history. Proc Natl Acad Sci U S A. 106: 13903-13908.

Brandl K, Rutschmann S, Li X, Du X, Xiao N, Schnabl B, Brenner DA, Beutler B. 2009. Enhanced sensitivity to DSS colitis caused by a hypomorphic Mbtps1 mutation disrupting the ATF6-driven unfolded protein response. Proc Natl Acad Sci U S A. 106: 3300-3305.

Brisbin A, Bustamante CD. 2012. PCAdmix: Principal Components-Based Assignment of Ancestry along Each Chromosome in Individuals with Admixed Ancestry from Two or More Populations. Human Biol. 84: 343.

Carstensen T, Tekola-Ayele F, Pagani L, Tachmazidou I, Hatzikotoulas K, Karthikeyan S, Iles L, Pollard MO, Choudhury A, Ritchie GRS. 2014. The African Genome Variation Project Shapes Medical Genetics in Africa. Nature. 517: 327-332.

Cheng G, Zhong M, Kawaguchi R, Kassai M, Al-Ubaidi M, Deng J, Ter-Stepanian M, Sun H. 2014. Identification of PLXDC1 and PLXDC2 as the transmembrane receptors for the multifunctional factor PEDF. Elife. 3: e05401.

Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST. 2011. The variant call format and VCFtools. Bioinformatics. 27: 2156-2158.

Degrandi D, Konermann C, Beuter-Gunia C, Kresse A, Würthner J, Kurig S, Beer S, Pfeffer K. 2008. Extensive characterization of IFN-induced GTPases mGBP1 to mGBP10 involved in host defense. J Immunol. 179: 7729-7740.

Depristo MA, Eric B, Ryan P, Garimella KV, Maguire JR, Christopher H, Philippakis AA, Guillermo DA, Rivas MA, Matt H. 2011. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat genet. 43: 491-498.

Ding ZL, Oskarsson M, Ardalan A, Angleby H, Dahlgren LG, Tepeli C, Kirkness E, Savolainen P, Zhang YP. 2012. Origins of domestic dog in southern East Asia is supported by analysis of Y-chromosome DNA. Heredity. 108: 507-514.

Ewing GB, Hermisson J. 2010. MSMS: A Coalescent Simulation Program Including Recombination, Demographic Structure, and Selection at a Single Locus. Bioinformatics. 26: 2064-2065.

Excoffier L, Dupanloup I, Huertasánchez E, Sousa VC, Foll M. 2013. Robust Demographic Inference from

Genomic and SNP Data. PLoS Genet. 9: e1003905.
Freedman AH, Gronau I, Schweizer RM, Ortega-Del Vecchyo D, Han E, Silva PM, Galaverni M, Fan Z, Marx P, Lorente-Galdos B. 2014. Genome sequencing highlights the dynamic early history of dogs. PLoS Genet. 10: e1004016.
Freedman AH, Schweizer RM, Vecchyo DO, Han E, Davis BW, Gronau I, Silva PM, Galaverni M, Fan Z, Marx P. 2016. Demographically-Based Evaluation of Genomic Regions under Selection in Domestic Dogs. PLoS Genet. 12.
Gaubert P, Bloch C, Benyacoub S, Abdelhamid A, Pagani P, Couloux A, Dufour S. 2012. Reviving the African Wolf Canis lupus lupaster in North and West Africa: A Mitochondrial Lineage Ranging More than 6,000 km Wide. Plos One. 7: e42740.
Germonpré M, Sablin MV, Stevens RE, Hedges RE, Hofreiter M, Stiller M, Després VR. 2009. Fossil dogs and wolves from Palaeolithic sites in Belgium, the Ukraine and Russia: osteometry, ancient DNA and stable isotopes. J Archaeol Sci. 36: 473-490.
Gou X, Wang Z, Li N, Qiu F, Xu Z, Yan D, Yang S, Jia J, Kong X, Wei Z. 2014. Whole genome sequencing of six dog breeds from continuous altitudes reveals adaption to high-altitude hypoxia. Genome Res. 24: 1308-1315.
Gower BA, Nagy TR, Goran MI. 1999. Visceral fat, insulin sensitivity, and lipids in prepubertal children. Diabetes. 48: 1515-1521.
Gun SY, Claser C, Tan KS, Rénia L. 2013. Interferons and interferon regulatory factors in malaria. Mediators Inflamm. 2014: 243713.
Hafalla JCR, Burgold J, Dorhoi A, Gross O, Ruland J, Kaufmann SH, Matuschewski K. 2012. Experimental cerebral malaria develops independently of caspase recruitment domain-containing protein 9 signaling. Infect Immun. 80: 1274-1279.
Haldar AK, Saka HA, Piro AS, Dunn JD, Henry SC, Taylor GA, Frickel EM, Valdivia RH, Coers J. 2013. IRG and GBP host resistance factors target aberrant, "non-self" vacuoles characterized by the missing of "self" IRGM proteins. PLoS Pathog. 9: e1003414.
Henn BM, Botigué LR, Gravel S, Wang W, Brisbin A, Byrnes JK, Fadhlaoui-Zid K, Zalloua PA, Moreno-Estrada A, Bertranpetit J. 2012. Genomic Ancestry of North Africans Supports Back-to-Africa Migrations. PLoS Genet. 8: e1002397.
Heun-Sik L, Sanghoon M, Jun Ho Y, Meehee L, Yeong HM, Young-Jin K, Bok-Ghee H, Jeong-Min K, Bong-Jo K. 2014. Genome-wide copy number variation study reveals KCNIP1 as a modulator of insulin secretion. Genomics. 104: 113–120.
Hogan BM, Bos FL, Bussmann J, Witte M, Chi NC, Duckers HJ, Schulte-Merker S. 2009. Ccbe1 is required for embryonic lymphangiogenesis and venous sprouting. Nat genet. 41: 396-398.
Huang DW, Sherman BT, Lempicki RA. 2009. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc. 4: 44-57.
Jónsson H, Ginolhac A, Schubert M, Johnson PLF, Orlando L. 2013. mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. Bioinformatics. 29: 1682-1684.
Jablonski NG, Chaplin G. 2013. Epidermal pigmentation in the human lineage is an adaptation to ultraviolet radiation. J Hum Evol. 65: 671-675.
Jablonski NG, Chaplin G. 2014. The Evolution of Skin Pigmentation and Hair Texture in People of African Ancestry. Dermatol Clin. 32: 113-121.
Jiang L, Mu J, Zhang Q, Ni T, Srinivasan P, Rayavara K, Yang W, Turner L, Lavstsen T, Theander TG, et al. 2013. PfSETvs methylation of histone H3K36 represses virulence genes in Plasmodium falciparum. Nature. 499: 223-227.

Kariuki SM, Rockett K, Clark TG, Reyburn H, Agbenyega T, Taylor TE, Birbeck GL, Williams TN, Newton CR. 2013. The genetic risk of acute seizures in African children with falciparum malaria. Epilepsia. 54: 990-1001.
Kim HS, Han T-Y, Yoo Y-M. 2016. Melatonin-Mediated Intracellular Insulin during 2-Deoxy-d-glucose Treatment Is Reduced through Autophagy and EDC3 Protein in Insulinoma INS-1E Cells. Oxid Med Cell Longev. 2016.
Kim J, Hanotte O, Mwai OA, Dessie T, Bashir S, Diallo B, Agaba M, Kim K, Kwak W, Sung S. 2017. The genome landscape of indigenous African cattle. Genome Biol. 18: 34.
Koepfli KP, Pollinger J, Godinho R, Robinson J, Lea A, Hendricks S, Schweizer RM, Thalmann O, Silva P, Fan Z. 2015. Genome-wide Evidence Reveals that African and Eurasian Golden Jackals Are Distinct Species. Curr Biol. 25: 2158-2165.
Kumar S, Subramanian S. 2002. Mutation rates in mammalian genomes. Proc Natl Acad Sci U S A. 99: 803-808.
Kusi KA, Gyan BA, Goka BQ, Dodoo D, Obeng-Adjei G, Troye-Blomberg M, Akanmori BD, Adjimani JP. 2008. Levels of soluble CD163 and severity of malaria in children in Ghana. Clin Vaccine Immunol. 15: 1456-1460.
Lalonde R, Marchetti N, Strazielle C. 2005. Primary neurologic screening and motor coordination of Dst dt-J mutant mice (dystonia musculorum) with spinocerebellar atrophy. Physiol Behav. 86: 46-51.
Lee TH, Hui G, Wang X, Kim C, Paterson AH. 2014. SNPhylo: a pipeline to construct a phylogenetic tree from huge SNP data. BMC genomics. 15: 120-121.
Lequarre A, Andersson L, Andre C, Fredholm M, Hitte C, Leeb T, Lohi H, Lindbladtoh K, Georges M. 2011. LUPA: A European initiative taking advantage of the canine genome architecture for unravelling complex disorders in both human and dogs. Veterinary Journal. 189: 155-159.
Li H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv preprint arXiv:1303.3997.
Lin H-H, Faunce DE, Stacey M, Terajewicz A, Nakamura T, Zhang-Hoover J, Kerley M, Mucenski ML, Gordon S, Stein-Streilein J. 2005. The macrophage F4/80 receptor is required for the induction of antigen-specific efferent regulatory T cells in peripheral tolerance. J Exp Med. 201: 1615-1625.
Lv FH, Agha S, Kantanen J, Colli L, Stucki S, Kijas JW, Joost S, Li MH, Marsan PA. 2014. Adaptations to Climate-Mediated Selective Pressures in Sheep. Molecular Biology & Evolution. 31: 3324.
Machado P, Rui P, Rocha AM, Manco L, Fernandes N, Miranda J, Ribeiro L, Rosário VED, Amorim A, Gusmão L. 2010. Malaria: looking for selection signatures in the human PKLR gene region. Br J Haematol. 149: 775–784.
Marshall MC. 2005. Diabetes in African Americans. Postgrad Med J. 81: 734-740.
Matsuoka S, Ballif BA, Smogorzewska A, McDonald ER, Hurov KE, Luo J, Bakalarski CE, Zhao Z, Solimini N, Lerenthal Y. 2007. ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage. Science. 316: 1160-1166.
Mech LD. 1981. The wolf : the ecology and behavior of an endangered species: University of Minnesota Press.
Mesecke S, Urlaub D, Busch H, Eils R, Watzl C. 2011. Integration of Activating and Inhibitory Receptor Signaling by Regulated Phosphorylation of Vav1 in Immune Cells. Sci Signal. 4: ra36.
Mitchell P. 2015. Did disease constrain the spread of domestic dogs (Canis familiaris) into Sub-Saharan Africa? Azania: Archaeological Research in Africa. 50: 1-44.
Muehlenbachs A, Fried M, Lachowitzer J, Mutabingwa TK, Duffy PE. 2007. Genome-wide expression analysis of placental malaria reveals features of lymphoid neogenesis during chronic infection. J Immunol. 179: 557-565.
Murray PJ, Wynn TA. 2011. Protective and pathogenic functions of macrophage subsets. Nat Rev Immunol. 11: 723-737.
Network MGE. 2014. Reappraisal of known malaria resistance loci in a large multicenter study. Nat genet. 46: 1197-1204.
Norton HL, Kittles RA, Parra E, Mckeigue P, Mao X, Cheng K, Canfield VA, Bradley DG, Mcevoy B, Shriver MD.

2007. Genetic evidence for the convergent evolution of light skin in Europeans and East Asians. Mol Biol Evol. 24: 710-722.

Norton HL, Mao X, Shriver MD, Cheng KC. 2006. SLC24A5, a putative cation exchanger, affects pigmentation in zebrafish and humans. Science. 310: 1782-1786.

Pang JF, Kluetsch C, Zou XJ, Zhang AB, Luo LY, Angleby H, Ardalan A, Ekstrom C, Skollermo A, Lundeberg J, et al. 2009. mtDNA data indicate a single origin for dogs south of Yangtze River, less than 16,300 years ago, from numerous wolves. Mol Biol Evol. 26: 2849-2864.

Park MN, Choi JA, Lee K-T, Lee H-J, Choi B-H, Kim H, Kim T-H, Cho S, Lee T. 2013. Genome-wide Association Study of Chicken Plumage Pigmentation. Asian-Australas J Anim Sci. 26: 1523-1528.

Patterson N, Moorjani P, Luo Y, Mallick S, Rohland N, Zhan Y, Genschoreck T, Webster T, Reich D. 2012. Ancient admixture in human history. Genetics. 192: 1065-1093.

Patterson N, Price AL, David R. 2006. Population Structure and Eigenanalysis. PLoS Genet. 2: 2074--2093.

Pavlidis P, Daniel I, Alexandros S, Nikolaos A. 2013. SweeD: Likelihood-Based Detection of Selective Sweeps in Thousands of Genomes. Mol Biol Evol. 30: 2224-2234.

Qi X, Hong J, Chaves L, Zhuang Y, Chen Y, Wang D, Chabon J, Graham B, Ohmori K, Li Y, et al. 2013. Antagonistic regulation by the transcription factors C/EBPα and MITF specifies basophil and mast cell fates. Immunity. 39: 97-110.

Räschle M, Smeenk G, Hansen RK, Temu T, Oka Y, Hein MY, Nagaraj N, Long DT, Walter JC, Hofmann K. 2015. Proteomics reveals dynamic assembly of repair complexes during bypass of DNA cross-links. Science. 348: 1253671.

Ramesh N, Mohan H, Unniappan S. 2015. Nucleobindin-1 encodes a nesfatin-1-like peptide that stimulates insulin secretion. Gen Comp Endocrinol. 216: 182-189.

Reyers F, Leisewitz A, Lobetti R, Milner R, Jacobson L. 1998. Canine babesiosis in South Africa: more than one disease. Does this serve as a model for falciparum malaria. Ann Trop Med Parasitol. 92: 503-511.

Rocha BC, Marques PE, Leoratti FM, Junqueira C, Pereira DB, Antonelli LR, Menezes GB, Golenbock DT, Gazzinelli RT. 2015. Type I Interferon Transcriptional Signature in Neutrophils and Low-Density Granulocytes Are Associated with Tissue Damage in Malaria. Cell Rep. 13: 2829-2841.

Rueness EK, Asmyhr MG, Sillerozubiri C, Macdonald DW, Bekele A, Atickem A, Stenseth NC. 2012. The Cryptic African Wolf: Canis aureus lupaster Is Not a Golden Jackal and Is Not Endemic to Egypt. Plos One. 6: e16385.

Sandri C, Caccavari F, Valdembri D, Camillo C, Veltel S, Santambrogio M, Lanzetti L, Bussolino F, Ivaska J, Serini G. 2012. The R-Ras/RIN2/Rab5 complex controls endothelial cell adhesion and morphogenesis via active integrin endocytosis and Rac signaling. Cell Res. 22: 1479-1501.

Sano Y, Furuta A, Setsuie R, Kikuchi H, Wang Y-L, Sakurai M, Kwon J, Noda M, Wada K. 2006. Photoreceptor cell apoptosis in the retinal degeneration of Uchl3-deficient mice. Am J Pathol. 169: 132-141.

Sasaki M, Omobowale O, Tozuka M, Ohta K, Matsuu A, Nottidge HO, Hirata H, Ikadai H, Oyamada T. 2007. Molecular survey of Babesia canis in dogs in Nigeria. J Vet Med Sci. 69: 1191-1193.

Savolainen P, Zhang Y-p, Luo J, Lundeberg J, Leitner T. 2002. Genetic evidence for an East Asian origin of domestic dogs. Science. 298: 1610-1613.

Shearin AL, Ostrander EA. 2010. Leading the way: canine models of genomics and disease. Dis Model Mech. 3: 27-34.

Storb R, Thomas ED. 1985. Graft‐versus‐Host Disease in Dog and Man: The Seattle Experience. Immunological Reviews. 88: 215-238.

Svec F, Nastasi K, Hilton C, Bao W, Srinivasan SR, Berenson GS. 1992. Black-white contrasts in insulin levels during pubertal development. The Bogalusa Heart Study. Diabetes. 41: 313-317.

Underhill PA, Passarino G, Lin AA, Shen P, Lahr MM, Foley RA, Oefner PJ, Cavalli-Sforza LL. 2001. The phylogeography of Y chromosome binary haplotypes and the origins of modern human populations. Ann Hum Genet. 65: 43-62.

Vilà C, Savolainen P, Maldonado JE, Amorim IR, Rice JE, Honeycutt RL, Crandall KA, Lundeberg J, Wayne RK. 1997. Multiple and ancient origins of the domestic dog. Science. 276: 1687-1689.

Wang G-D, Fan R-X, Zhai W, Liu F, Wang L, Zhong L, Wu H, Yang H-C, Wu S-F, Zhu C-L. 2014. Genetic convergence in the adaptation of dogs and humans to the high-altitude environment of the Tibetan plateau. Genome Biol Evol. 6: 2122-2128.

Wang G-d, Zhai W, Yang H-c, Fan R-x, Cao X, Zhong L, Wang L, Liu F, Wu H, Cheng L-g. 2013. The genomics of selection in dogs and the parallel evolution between dogs and humans. Nat Commun. 4: 1860.

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

Wang L, Ma YP, Zhou QJ, Zhang YP, Savolaimen P, Wang GD. 2016b. The geographical distribution of grey wolves (Canis lupus) in China: a systematic review. Zool Res. 37: 315.

Weger BD, Sahinbas M, Otto GW, Mracek P, Armant O, Dolle D, Lahiri K, Vallone D, Ettwiller L, Geisler R. 2011. The light responsive transcriptome of the zebrafish: function and regulation. PloS one. 6: e17080.

WHO. 2015. World Malaria Report 2015. World Malaria Report. 30: 189–206.

Wong AK, Ruhe AL, Dumont BL, Robertson KR, Guerrero G, Shull SM, Ziegle JS, Millon LV, Broman KW, Payseur BA. 2010. A comprehensive linkage map of the dog genome. Genetics. 184: 595-605.

Wu H, Liu YH, Wang GD, Yang CT, Otecko NO, Liu F, Wu SF, Wang L, Yu L, Zhang YP. 2016. Identifying molecular signatures of hypoxia adaptation from sex chromosomes: A case for Tibetan Mastiff based on analyses of X chromosome. Scientific Reports. 6: 35004.

Yao X, Wu J, Lin M, Sun W, He X, Channe G, Silvia B, Long CA, Wang R, Su XZ. 2016. Increased CD40 Expression Enhances Early STING-Mediated Type I Interferon Response and Host Survival in a Rodent Malaria Model. PLoS Pathog. 10: e1005930.

Yi X, Liang Y, Huerta-Sanchez E, Jin X, Cuo ZXP, Pool JE, Xu X, Jiang H, Vinckenbosch N, Korneliussen TS. 2010. Sequencing of 50 human exomes reveals adaptation to high altitude. Science. 329: 75-78.

Yu J, Liu F, Yin P, Zhu X, Cheng G, Wang N, Lu A, Luan W, Zhang N, Li J. 2011. Integrating miRNA and mRNA expression profiles in response to heat stress-induced injury in rat small intestine. Funct Integr Genomics. 11: 203-213.

Yu X, Cai B, Wang M, Tan P, Ding X, Wu J, Li J, Li Q, Liu P, Xing C, et al. 2016. Cross-Regulation of Two Type I Interferon Signaling Pathways in Plasmacytoid Dendritic Cells Controls Anti-malaria Immunity and Host Mortality. Immunity. 45: 1093-1107.

Zuo Z, Cui H, Li M, Peng X, Zhu L, Zhang M, Ma J, Xu Z, Gan M, Deng J. 2012. Transcriptional Profiling of Swine Lung Tissue after Experimental Infection with Actinobacillus pleuropneumoniae. Int J Mol Sci. 14: 10626-10660.

Tables

Table 1. GO terms for genes identified by both SweeD and PBS

Category	GO Term	No of genes	*P* value	Citation on environmental adaptation and immunity
CC	0030054~cell junction	5	0.035	
CC	0005783~endoplasmic reticulum	6	0.026	(Yu et al. 2011)
MF	0008047~enzyme activator activity	5	0.010	(Zuo et al. 2012; Lv et al. 2014)
MF	0030695~GTPase regulator activity	5	0.019	(Barreiro et al. 2008; Lv et al. 2014)
MF	0005096~GTPase activator activity	4	0.018	(Lv et al. 2014)
MF	0060589~nucleoside-triphosphatase regulator activity	5	0.020	(Barreiro et al. 2008; Lv et al. 2014)
MF	0070011~peptidase activity, acting on L-amino acid peptides	5	0.049	

Figure legends

Fig. 1. Population structure and genetic diversity of the canids analyzed in this study

(A) Geographic locations of the 55 canids studied. (B) Principal component analysis. EGW: Eurasian gray wolves; AGW: African golden wolves; CIDY: Chinese indigenous dogs from Yingjiang; MEVD: Middle Eastern village dogs; EB: European breeds; NID: Nigerian indigenous dogs. (C) Phylogenetic tree using bootstrapping analysis. (D) Structure analysis of the 55 canids. (E) Genetic diversity for the five inferred canid groups.

Fig. 2. Demographic history of wolves and dogs

Demographic history was inferred of African golden wolves (AGW), gray wolves, Yingjiang indigenous dogs, European breeds and Middle Eastern village dogs, and Nigerian indigenous dogs using fastsimcoal2. Mutation rate were set to 2.2×10^{-9} per year and a generation time of 3 years was used (Wang et al. 2016a). Divergent times are shown in the right side of the diagram.

Fig. 3. Dog *ADGRE1* promotes innate immune response during *Plasmodium* infection.

(A) Immunoblot analysis of empty vector GFP protein (left) and quantitative RT-PCR analysis of *ADGRE1* gene (right) in (GFP) and *ADGRE1* transfected RAW 264.7 macrophages. (B) Quantitative RT-PCR analysis of *Plasmodium berghei ANKA* specific *18S rRNA* and *Hsp70* in *GFP* and *ADGRE1* transfected RAW 264.7 cells after parasite infection. (C) Expression analysis of *Ifnb*, *Cxcl10*, *Gbp1*, *Gbp3*, *Il1a*, *Il1b*, *Il6* and *Cxcl1* in *GFP* and *ADGRE1* transfected RAW 264.7 cells during parasite infection by quantitative RT-PCR. (D) Immunoblot analysis of LC3, TFEB, caspase 3 and GAPDH (loading control) in *GFP* and *ADGRE1* transfected RAW 264.7 cells during parasite infection. Data are representative of three independent experiments for A and B, and two independent experiments for C and D. Data are means ± SD. ***, $P < 0.001$; ****, $P < 0.0001$; ns, not significant.

Fig. 4. The positive effect of the *ADGRE1-G655A* SNP in African dogs during *Plasmodium* infection.

(A) The predominant SNPs of the *ADGRE1* gene in European and African dogs, respectively. (B) Expression analysis of *Plasmodium berghei ANKA* specific *Hsp70,* host genes *Ifnb*, *Gbp1*, Gbp2, *Gbp3*, Gbp5, *Il6* and *Il1b* in *ADGRE1* and *ADGRE1-G655A* transfected RAW 264.7 cells during parasite infection by quantitative RT-PCR. Data are representative of two independent experiments. Data are means ± SD. *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$; ns, not significant.

Fig. 1.

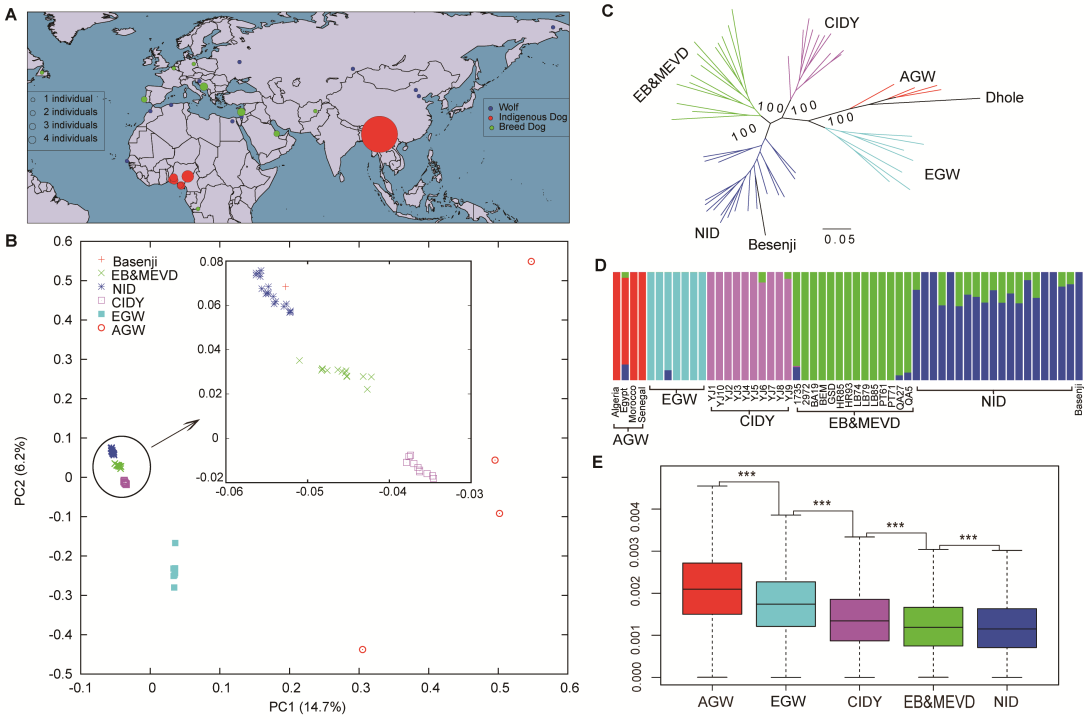

Fig. 2.

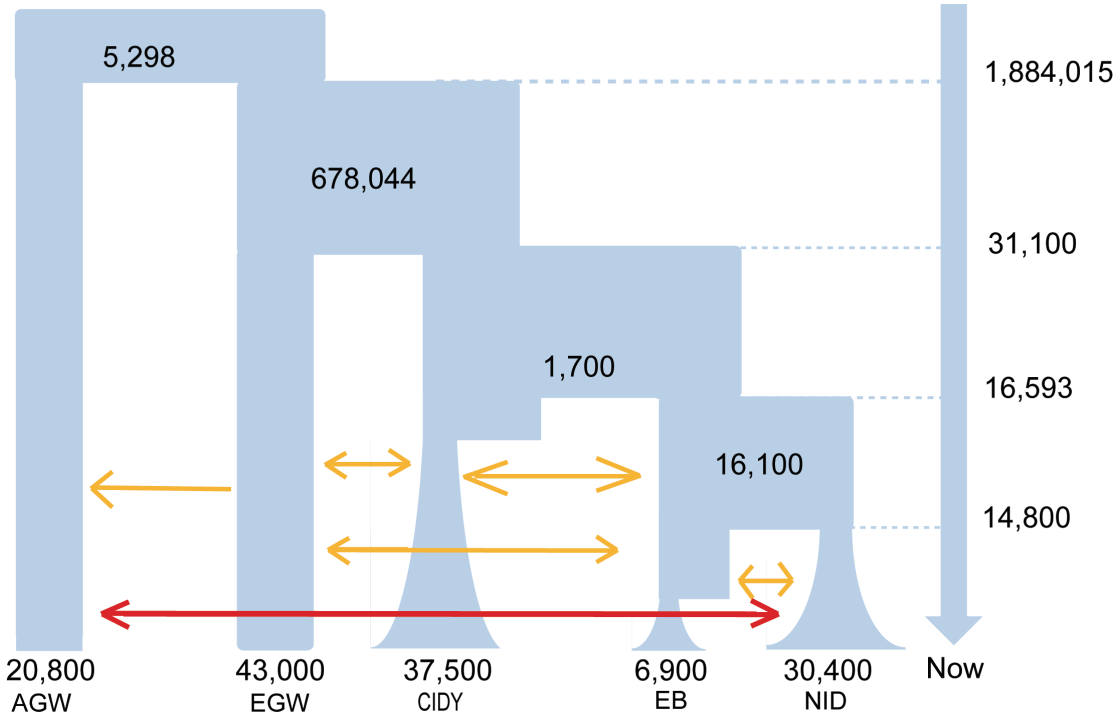

Fig. 3.

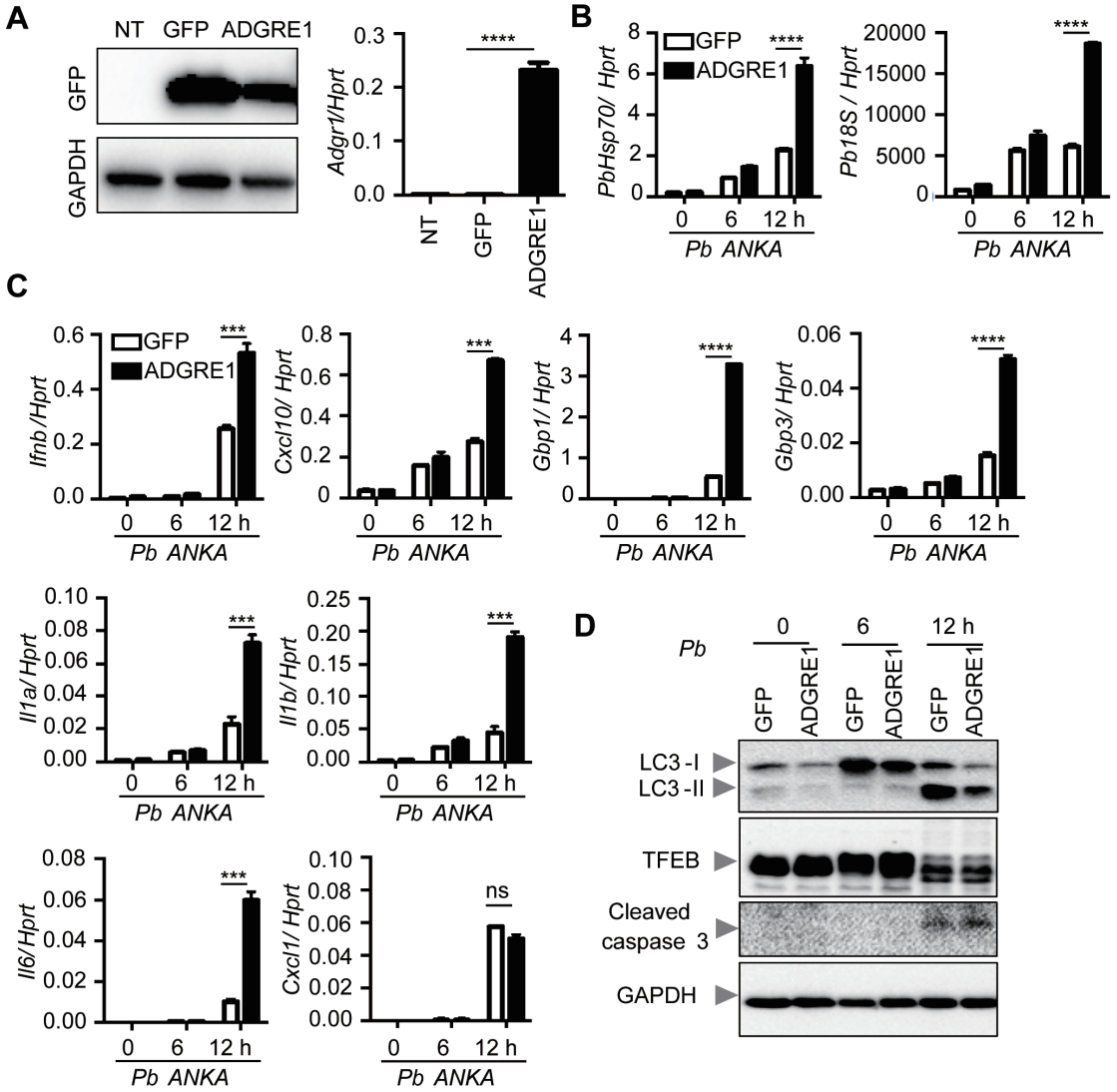

Fig. 4.