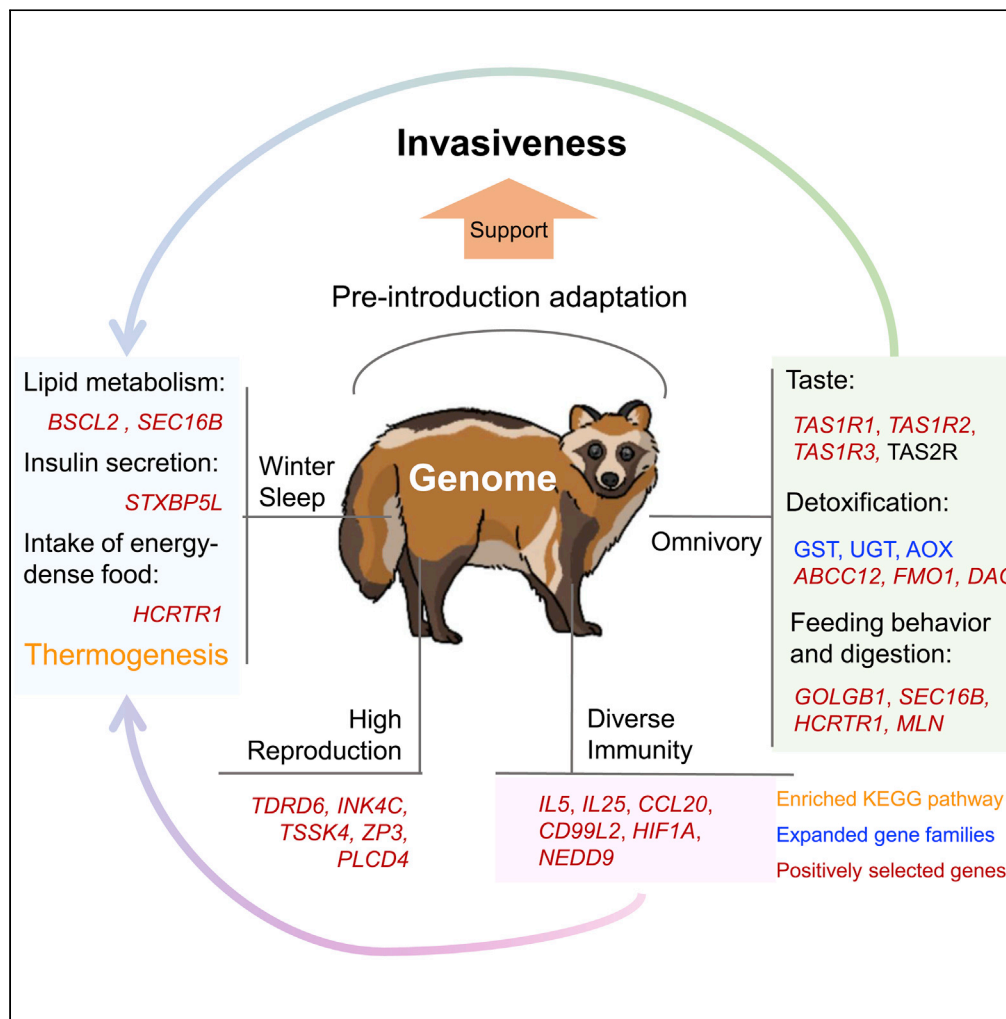## Article

The chromosome-scale genome of the raccoon dog: Insights into its evolutionary characteristics

Tianming Lan, Haimeng Li, Shangchen Yang, ..., Chuyu Lin, Huan Liu, Zhijun Hou

xu_daniel@163.com (Y.X.)
linchuyu@znhaidao.com (C.L.)
liuhuan@genomics.cn (H.L.)
houzhijundb@163.com (Z.H.)

Highlights

Generated a high-quality chromosome-scale reference genome of the raccoon dog

Identified and validated sex-linked genomic regions by resequencing 38 individuals

Genomic basis for omnivory, high reproduction, immune system, and winter sleep

Pre-introduction adaptation facilitated its successful invasion from Asia to Europe

iScience

Article

The chromosome-scale genome of the raccoon dog: Insights into its evolutionary characteristics

Tianming Lan,[1,2,9] Haimeng Li,[2,4,9] Shangchen Yang,[5,9] Minhui Shi,[2,4] Lei Han,[3] Sunil Kumar Sahu,[2] Yaxian Lu,[3] Jiangang Wang,[2] Mengchao Zhou,[3] Hui Liu,[6] Junxuan Huang,[2] Qing Wang,[2,4] Yixin Zhu,[2,4] Li Wang,[5] Yanchun Xu,[1,3,*] Chuyu Lin,[7,*] Huan Liu,[1,2,8,*] and Zhijun Hou[1,3,10,*]

SUMMARY

The raccoon dog (*Nyctereutes procyonoides*) is an invasive canid species native to East Asia with several distinct characteristics. Here, we report a chromosome-scale genome of the raccoon dog with high contiguity, completeness, and accuracy. The intact taste receptor genes, expanded gene families, and positively selected genes related to digestion, absorption, foraging, and detoxification likely support the omnivory of raccoon dogs. Several positively selected genes and raccoon dog-specific mutations in *TDRD6* and *ZP3* genes may explain their high reproductivity. Enriched GO terms in energy metabolism and positively selected immune genes were speculated to be closely related to the diverse immune system of raccoon dogs. In addition, we found that several expanded gene families and positively selected genes related to lipid metabolism and insulin resistance may contribute to winter sleep of the raccoon dog. This high-quality genome provides a valuable resource for understanding the evolutionary characteristics of this species.

INTRODUCTION

The raccoon dog (*Nyctereutes procyonoides*) is the only species in the genus Nyctereutes. It is a medium-sized canid that prefers various habitats such as gardens, marshlands, river valleys, and damp forests with rich undergrowth (Kauhala and Kowalczyk, 2011). The native distribution area of raccoon dogs includes China, Korea, North Vietnam, Japan, Mongolia, and East Siberia (Pitra et al., 2009; Ward and Wurster-Hill, 1990) (Figure 1A). The range of this species has continuously expanded over the last 100 years, and currently covers most of Eurasia, making the raccoon dog a prominent invasive species in Europe and posing a great risk to local ecosystems and public health.

Globally, there are six subspecies (Ellerman and Morrison-Scott, 1951; Hong et al., 2018; Ward and Wurster-Hill, 1990), including *N. procyonoides ussuriensis* (originally found in eastern China and southeastern Siberia), *N. procyonoides procyonoides* (northern Indochina and China), *N. procyonoides orestes* (southwestern China), *N. procyonoides koreensis* (Korean Peninsula), *N. procyonoides viverrinus* (Japan, except Hokkaido), and *N. procyonoides albus* (Hokkaido, Japan). The *N. procyonoides viverrinus* and *N. procyonoides albus* subspecies were suggested to be recognized as distinct species considering the large difference in the number of chromosomes between Japanese raccoon dogs (2n = 38 + 3–4 B) and other raccoon dogs (2n = 54 + 2–3 B) (Nie et al., 2003).

There are several distinct evolutionary characteristics in raccoon dogs. This species is reported to have a high reproductive rate with a relatively young age of sexual maturity and large average litter size (8–10 individuals) (Helle and Kauhala, 1995; Kauhala, 1996; Kauhala and Kowalczyk, 2011; Kowalczyk et al., 2009). The raccoon dog is a typical omnivore with a diverse diet that includes food ranging from fruits to insects, frogs, birds, and small mammals (Drygala et al., 2013; Kauhala and Kowalczyk, 2011). This helps raccoon dogs shift their diets according to the different habitats and seasons, making the raccoon dog a species with high plasticity. Another unique habit of raccoon dogs among canids is winter sleep under harsh conditions, which can protect raccoon dogs from food deprivation and cold weather (Asikainen et al., 2004). Before winter sleep, raccoon dogs fatten themselves and their body weight will be correspondingly increased (Nieminen et al., 2002). Unlike typical hibernation, the body temperature of the raccoon dog
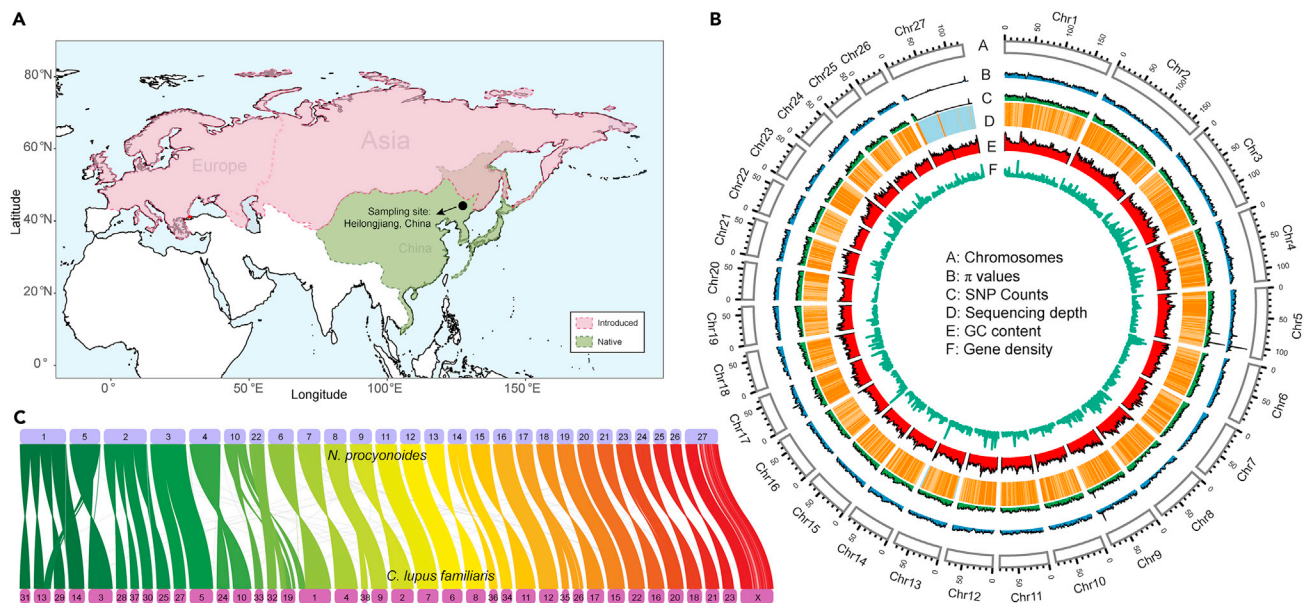
Figure 1. The distribution of native (green) and introduced/invasive (pink) areas of raccoon dogs, the landscape of the raccoon dog genome, and the chromosome-scale synteny analysis of the raccoon dog and domestic dog

(A) The map describes the current distribution of raccoon dogs; here, we show both the native and introduced areas (https://www.cabi.org/isc/datasheet/72656).

(B) The genomic landscape of the raccoon dog genome. A: The 27 chromosomes of the raccoon dog genome; B: population level genetic diversity (π) calculated by a 500 kb window; C: SNP density across the genome (500 kb window); D: sequencing depth (X) calculated by 500 kb window; E: GC content (%); F: gene density calculated by 500 kb window.

(C) The chromosome-scale synteny analysis between the raccoon dog genome and the domestic dog genome, which was visualized using RectChr v1.27 (https://github.com/BGI-shenzhen/RectChr).

is close to normal (Asikainen et al., 2004). However, little is known about the possible genetic basis of these biological characteristics.

In addition, raccoon dogs are regarded as reservoir hosts for many pathogens, which poses a great threat to public health, because related diseases can also be transmitted to humans and other animals (Kauhala and Kowalczyk, 2011; Kjaer et al., 2021; Laurimaa et al., 2016; Sutor et al., 2014). The raccoon dog is a very important vector of rabies, as a large percentage of rabies cases are found in raccoon dogs, even higher than that in red foxex (Kauhala and Kowalczyk, 2011). In addition, the raccoon dogs have also been found to be hosts of SARS-CoV (Guan et al., 2003), H5N1 virus (Qi et al., 2009), canine distemper virus (Aoyagi et al., 2000), porcine circoviruses (Song et al., 2019), and Amdoparvovirus (Shao et al., 2014). The raccoon dog is also a reservoir species of many parasites, including *Alaria alata*, *Echinococcu multilocularis*, *Sarcoptes scabiei*, and *Trichinella* spp. (Kauhala and Kowalczyk, 2011). Other tick-borne pathogens, such as *Anaplasma phagocytophilum*, are also detected in raccoon dogs (Kjaer et al., 2021). We anticipated that the immune system of raccoon dogs is diverse.

A high-quality genome is a valuable genetic resource to explore the possible genetic basis for the biological features of a species. Chueca et al. recently reported a genome assembly of the raccoon dog, but with a female individual (Chueca et al., 2021). In this study, we assembled a chromosome-scale genome of raccoon dogs by using a male individual collected from Heilongjiang, China. We explored the possible genetic basis of the biological characteristics of raccoon dogs, including immunity, reproduction, omnivory, winter sleep, and invasiveness. The findings from this study will provide a useful and valuable genomic resource for future research on the evolution, ecology, and management of wild populations of this species.

RESULTS

Improved genome assembly

A total of 385.94 Gb long reads generated by the Pacific Biosciences (PacBio) platform, 175.52 Gb whole-genome sequencing (WGS) short reads and 203.52 Gb of Hi-C reads were used for the chromosome-scale

Table 1. Statistics for the sequencing data, genome assembly, and annotation of the raccoon dog genome

Item	Category	Number
Sequencing data	PacBio (Gb)	385.94
	WGS (Gb)	175.52
	Hi-C (Gb)	203.52
	RNA-seq (5 organs) (Gb)	101.70
Assembly	Estimated genome size (Gb)	3.21
	Assembled genome size (Gb)	2.38
	Contig N50 (Mb)	41.87
	Scaffold N50 (Mb)	83.70
	Longest scaffold (Mb)	177.96
Annotation	GC content (%)	41.33
	Repeat sequences (%)	35.11
	Number of protein-coding genes	20,000
	Number of functionally annotated genes	19,973

genome assembly (Table 1). We first assembled one primary genome with a contig N50 of 23.99 Mb by using error-corrected PacBio subreads. Then, this primary genome was polished by PacBio subreads and WGS short reads. Finally, Hi-C reads were used for concatenating primary contigs into chromosome-scale assembly. The raccoon dog genome contains 27 pairs of chromosomes (2n = 54), including 26 pairs of autosomes and one pair of sex chromosomes (Nie et al., 2003). The size of the final assembled nuclear genome was 2.38 Gb, and the genome consisted of 218 scaffolds after redundancy removal, with 2.32 Gb assigned to 27 chromosomes (Table S1, Figure S1, and Figure 1B). The scaffold N50 of this chromosome-scale assembly was 41.87 Mb and the GC content of this genome was 41.33%, which is very similar to that of the domestic dog (CanFam3.1, GCA_000002285.2) and red fox (VulVul2.2, GCA_003160815.1) genomes (Table 1).

To evaluate the completeness of the genome, we first performed the Benchmarking Universal Single-Copy Orthologs (BUSCO) (Simão et al., 2015) analysis by using the mammalia_odb10 database. We found that 96.4% of the 9,226 BUSCO genes were complete in the genome, and only 2.8% and 0.8% were identified as missing and fragmented, respectively (Table S2). To compare with the previously published genome (GCA_905146905.1, Rac 1.0 hereafter), we also used the laurasiatheria_odb10 database to perform the BUSCO analysis (Table S2). The BUSCO score of our genome was slightly higher than that of the Rac 1.0 genome. In addition, 96.28% of the transcript data (kidney) and 99.26% of the WGS short reads were mapped onto our final assembly (Table S3). We also downloaded the RNA-seq data of the Rac 1.0 genome and mapped it against its own genome and the genome we assembled. We found that the mapping rate on our genome (96.24%) was higher than that on the Rac 1.0 genome (85.75%). The above assessments showed that our assembly represents a more complete genome with high quality and contiguity.

To further evaluate the accuracy of our assembly at the chromosome level and detect the fusion and fission events, we performed an interspecies synteny analysis between our assembled genome and the domestic dog genome (CanFam3.1, GCA_000002285.2). Overall, we found high collinearity between our assembled genome and the domestic dog genome with clear one-to-one syntenic blocks (Figure 1C). We also found nine fusion and three fission events in the synteny result. For example, Chr1 is the largest chromosome of the raccoon dog, which was found to be a fusion of Chr13, Chr14, Chr29, and Chr31 of the domestic dog genome. Chr1, Chr13, and Chr19 of the domestic dog genome were split into two chromosomes in the raccoon dog genome. All these fission and fusion events were surprisingly consistent with the findings of a previous karyotypic study (Becker et al., 2011), indicating the accuracy of our assembled genome at the chromosome level.

Genome annotation

We first identified 835.98 Mb of repetitive elements in our assembled genome, representing 35.11% of the total genome size. These repeat elements included LINEs (22.71%), LTRs (12.00%), DNA elements (2.40%), SINEs (1.23%), and other repeats (0.34%) (Tables S4–S6). Our genome's total repeat length was found to be

somewhat larger than that of the Rac 1.0 genome (34.04%), demonstrating the superiority of our genome's longer contigs in assembling highly repeated regions.

We sequenced 101.70 Gb of RNA-seq data and generated 175,521 full-length transcripts to support the gene annotation of our genome. We finally identified 20,000 protein-coding genes with high confidence by combining high-quality transcript mapping, homology-based protein alignment, and *ab initio* predictions. The average exon length, intron length, and gene length were 176.40 bp, 4481.96 bp, and 38.22 kb, respectively (Figure S6). Among our predicted genes, 14,913 (74.57%) were supported by transcript data. The BUSCO score of our predicted gene set was evaluated as 92.7% with the mammalia_odb10 database, showing high completeness (Table S2). Finally, 19,973 (99.87%) genes were functionally annotated in at least one of the five databases we used (Table S7 and Figure S7). In addition, 781 rRNAs, 995 miRNAs, 2,280 snRNAs, and 146,963 tRNAs were predicted in our genome (Table S8).

Identification of sex-linked regions

Considering that the Rac 1.0 genome was assembled using a female individual, we determined both X- and Y-linked scaffolds with multiple lines of evidence. We first ranked the average sequencing coverage of all 218 scaffolds (Table S9) and found that Chr27 and Scaffold30 presented sequencing coverages with nearly half of the whole genome level (Figure S8). We then regarded these two genome regions as derived from sex-linked chromosomes. A corresponding low level of genetic diversity was also found in these scaffolds, which further supported our identification of sex chromosomes (Table S10). To further separate the Y- and X-linked regions, we performed synteny analysis between our primarily identified sex-linked regions and the X/Y chromosomes of the domestic dog genome (Figures 2A and 2B). We found 797 and 25 protein-coding genes in the domestic dog X and Y chromosomes, respectively, that were successfully mapped onto Chr27 and Scaffold30 (Table S11) in the raccoon dog genome. The assembled X and Y chromosomes of the raccoon dog presented high collinearity with the X and Y chromosomes of the domestic dog. We therefore concluded that Chr27 (127.44 Mb) and Scaffold30 (3.24 Mb) were the X chromosome and Y-linked scaffold, respectively.

To further validate our identification, we mapped the whole genome resequencing data from 38 individuals of known sex (5 females and 33 males) to our assembled genome (Table S12). As we expected, the sequencing depth of Chr27 and Scaffold30 in the male individuals was nearly half that of the autosomes (Figures 2C–2E). For the female individuals, however, the depths of Chr27 were nearly the same as those of the autosomes (Figures 2C–2E). These results further supported that our identification of the sex-linked regions was accurate. This is the first time we identified X and Y sex-linked genomic regions in a raccoon dog reference genome, which will be a valuable resource for future studies.

Expansion of gene families in the raccoon dog genome

To explore the genomic adaptations of raccoon dogs, we performed a comprehensive comparative genomic analysis with 17 other Carnivora and mammalian species (Table S13). We focused on biological characteristics such as reproduction, immunity, omnivory, and winter sleep, which are prerequisites for raccoon dogs to spread and adapt to habitats. To better understand the possible genetic basis of the omnivorous diet, we included herbivores, omnivores, and carnivores.

We found 430 expanded gene families in the raccoon dog genome compared with the common ancestor of the raccoon dog, red fox, and arctic fox (Figure 3A). GO enrichment analysis of these expanded gene families showed that a large proportion of significantly enriched GO terms were closely related to energy metabolism, including carbohydrate metabolic process (GO:0,005,975, p = 0.0025), ATP metabolic process (GO:0,046,034, p = 2.92×10^{-26}), ATP synthesis coupled proton transport (GO:0,015,986, p = 5.55×10^{-46}), energy coupled proton transport, down electrochemical gradient (GO:0,015,985, p = 5.55×10^{-46}), and malate metabolic process (GO:0,006,108, p = 5.47×10^{-8}) (Figures 3B–3D, Table S14). We also found an enriched GO term that was related to detoxification (drug metabolic process, GO:0017144, p = 3.77×10^{-23}) and lipid metabolism (GO:0055102, p = 0.0034). In the 37 significantly enriched KEGG pathways (Table S15, Figure 3E), we found that eight pathways were closely related to energy metabolism, including the citrate cycle (map00020, p = 1.76×10^{-4}), glycolysis (map00010, p = 1.33×10^{-44}), oxidative phosphorylation (map00190, p = 7.21×10^{-18}), pyruvate metabolism (map00620, p = 1.95×10^{-12}), carbon metabolism (map01200, p = 8.95×10^{-34}), thermogenesis (map04714, p = 8.36E-10), HIF-1 signaling (map04066, p = 7.96×10^{-38}), and propanoate metabolism (map00640, p = 0.0014) pathways. Four significantly enriched pathways were found to be immune-related, including base
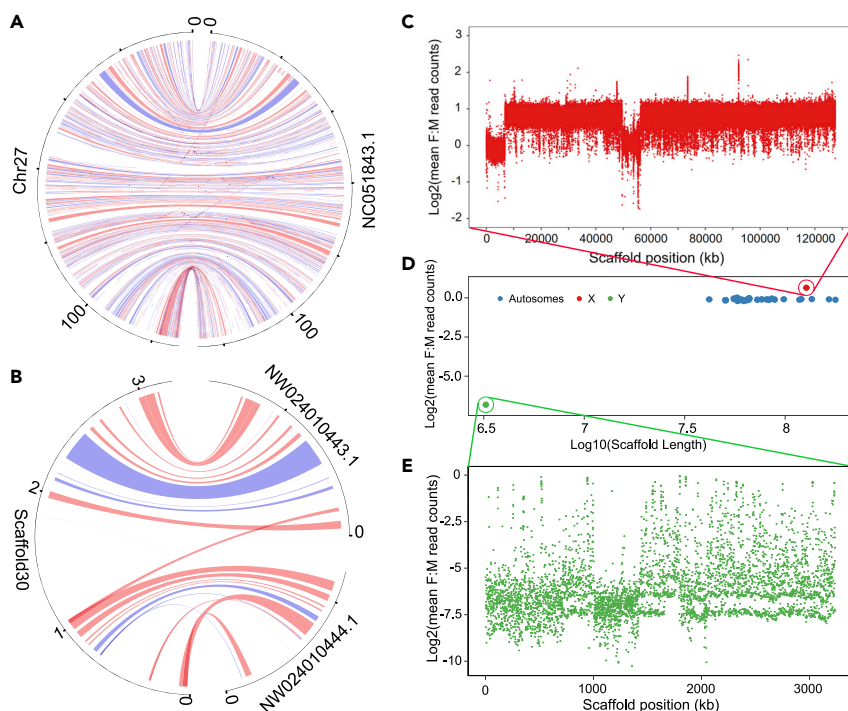
Figure 2. Identification of the X chromosome and Y-linked scaffold

(A) Synteny analysis of genes on the X chromosome between raccoon dog and domestic dog. Red lines indicate the genes in the dog genome mapped to the positive strand of the raccoon dog genome, and blue lines indicate the genes in the dog genome mapped to the negative strand of the raccoon dog genome.

(B) Synteny analysis of genes on the Y chromosome between raccoon dog and domestic dog. The red line and blue line are the same as those in (A).

(C) The ratio of sequencing depth between female and male individuals in a 500 bp window across the X chromosome.

(D) The ratio of the sequencing depth of each chromosome-scale scaffold between female and male individuals. The red dot represents the X chromosome, the green dot represents the Y-linked scaffold, and the blue dots are scaffolds from autosomes. The expected ratio is 1:1, the X chromosome is expected to have a higher ratio and the Y chromosome is expected to have a lower ratio.

(E) The ratio of sequencing depth between female and male individuals in a 500 bp window across the Y-linked scaffold.

excision repair (map03410, p = 1.22×10^{-72}), necroptosis (map04217, p = 2.47×10^{-56}), spliceosome (map03040, p = 3.65E-20), and proteasome (map03050, p = 1.13×10^{-6}). We also found some enriched pathways that were related to detoxification (drug metabolism-cytochrome P450, map00982, p = 0.0035 and glutathione metabolism, map00480, p = 0.0076) and the biosynthesis of amino acids (map01230, p = 1.54×10^{-32}). In addition, olfactory transduction (map04740, p = 5.79×10^{-5}) and salivary secretion (map04970, p = 3.52×10^{-39}) were also significantly enriched pathways in the raccoon dog genome (Figure 3E and Table S15).

Evolution of genes associated with omnivory, reproduction, immunity, and winter sleep

Raccoon dogs are opportunistic omnivores, as they mainly consume easily found food. Their diet comprises small mammals, insects, fruit, plant seeds and leaves, and carrion comprise (Hirasawa et al., 2006; Mulder, 2012). Toxic secondary metabolites usually exist in plants (Xia et al., 2021) and toxic byproducts are also widely distributed in rotting food (Blumstein et al., 2017). In addition, poisonous substances such as plant alkaloids are typically bitter in taste (Li and Zhang, 2014). Therefore, we focused on detoxification and taste-receptor-related gene families. In the expanded gene families, we found some typical detoxification gene families, including glutathione S-transferase (Hayes and Pulford, 1995; Sayyab et al.), protein tyrosine phosphatase (UGT) (Bock, 2016), and aldehyde oxidase (Chang et al., 2010; Guo et al.). Interestingly, we found that the GST gene family was expanded in the raccoon dog genome in all groups, including omnivores, herbivores, and carnivores (Figures 4A and 4B). Sweet taste, umami taste, and bitter taste are mediated by the taste receptors TAS1R2/TAS1R3, TAS1R1/TAS1R3, and TAS2Rs, respectively. In this study, we found that *TAS1R1, TAS1R2, TAS1R3,* and several TAS2R genes were intact and functional in
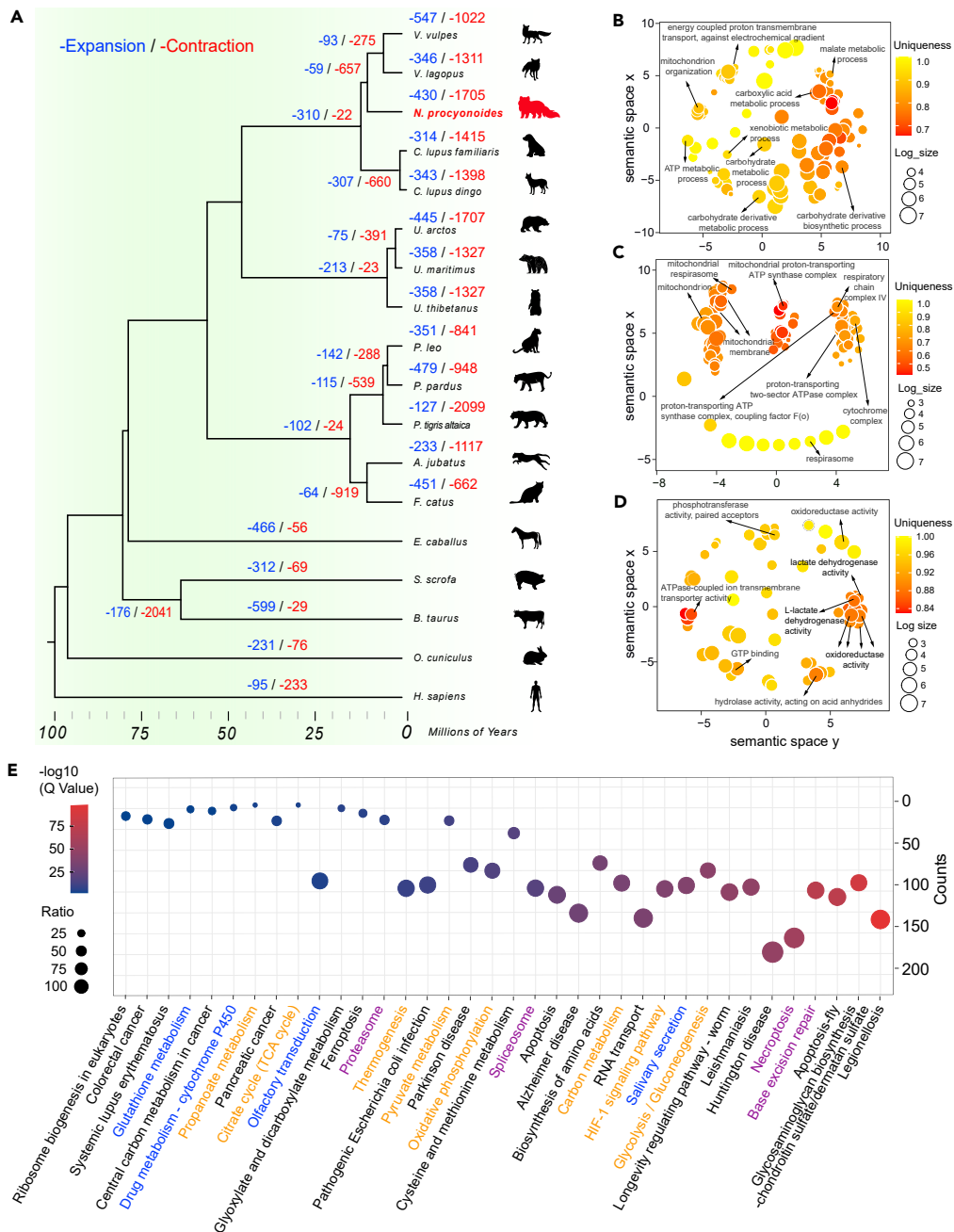
Figure 3. Comparative genomics analysis and enrichment analysis of expanded gene families in the raccoon dog genome

(A) The phylogenetic relationship of 18 species and the estimated divergence time. Numbers on the branch of the phylogenetic tree represent the significantly expanded (blue) and contracted (red) gene families.

(B–D) Clusters of significantly overrepresented GO items for biological process (B), cellular component (C), and molecular function (D) by REVIGO for expanded gene families in the raccoon dog genome. Semantic similar GO terms clustered together.

(E) Significantly enriched KEGG pathways in the raccoon dog genome compared with the other 17 species. Blue: pathways related to the omnivorous diet. Orange: pathways related to energy metabolism. Purple: pathways related to immunity.
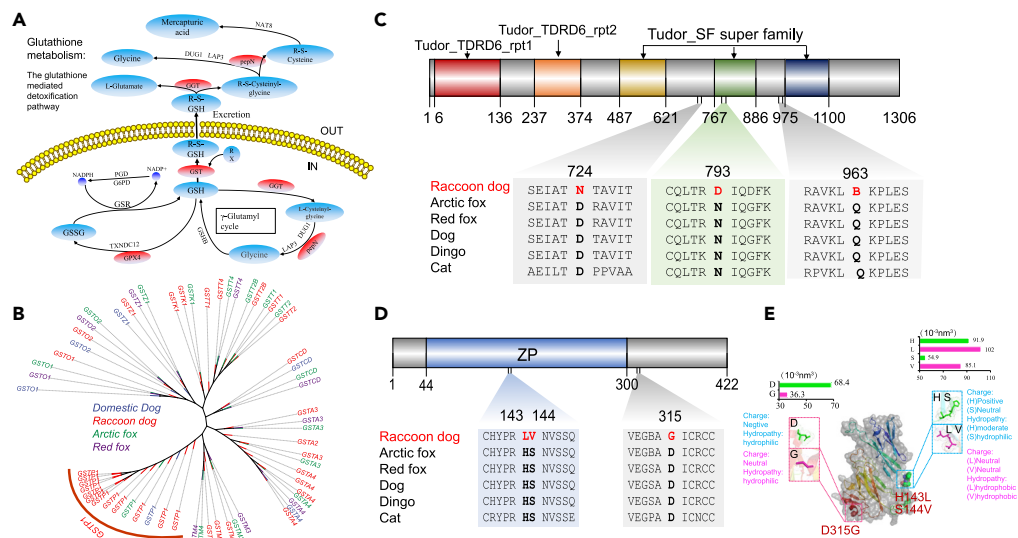
Figure 4. The possible genetic basis for detoxification and high reproduction in the raccoon dog

(A) The glutathione-mediated detoxification pathway. The ellipse in red represents enzymes encoded by expanded gene families.

(B) The phylogenetic tree of the GST gene family was constructed by the maximum likelihood method. Blue: Domestic dog; Red: Raccoon dog; Green: Arctic fox; Purple: Red fox. The GSTP1 genes in the raccoon dog genome are obviously expanded.

(C) Raccoon dog-specific amino acid changes in the *TDRD6* gene. N793D was found in the Tudor_SF superfamily.

(D) Raccoon dog-specific amino acid changes in the *ZP3* gene. Two substitutions were found to be in the ZP domain region.

(E) Three-dimensional view of the ZP3 protein, highlighting raccoon dog-specific amino acid changes. The zoomed in pink amino acids are the raccoon dog-specific amino acids and the green amino acids are the those predicted from the dog. The bar plot shows the residual volume of amino acids.

the raccoon dog genome (Table S16). Selection analysis showed that three detoxification-related genes, *ABCC12, FM O 1,* and *DAO,* were under positive selection. Feeding behavior and digestion-related genes (*GOLGB1, SEC16B,* and *MLN*) were also found to be under positive selection (Table S17).

Although we did not find significantly enriched GO or KEGG pathways that were directly related to reproduction in expanded gene families, five reproductive genes, *TDRD6, CDKN2C, ZP3, PLCD4,* and *TSSK4,* were found to be positively selected in the raccoon dog genome. These genes are closely related to spermatogenesis and acrosomal reactions. In the *TDRD6* and *ZP3* genes, we found that each gene contains three raccoon dog-specific amino acid changes (Figures 4C and 4D). The mutation N793D is located in the Tudor superfamily domain of the *TDRD6* gene, and the S144V mutation was found in the ZP domain region of the *ZP3* gene. We then predicted the effects of these mutations on protein function by using the Protein Variation Effect Analyzer (Provean) web server (Choi and Chan, 2015). Five of the six mutations in the *TDRD6* and *ZP3* genes were predicted to be neutral by Provean, and we did not find damaging effects in the amino acid changes in the domain regions (Table S18). We further predicted the three-dimensional structure of the ZP3 protein. Although these amino acid substitutions did not change the main structure of this protein, the physicochemical properties of the amino acids changed, which may contribute to the high reproduction of the raccoon dog (Figure 4E).

By comparison with the above mentioned 17 species, we screened 206 genes that were under positive selection. Thirty genes, accounting for 14.56% of the total positively selected genes (PSGs), were immune-related, including *IL5, IL25, CCL20, CD99L2, HIF1A,* and *NEDD9* (Table S19). We also performed a comparative genomic analysis on all 30 genes with other canines. Thirty-five (35) amino acid changes harbored in 18 genes were found to be raccoon dog-specific changes, and 21 mutations were found to be located in functional domain regions (Table S20, Figure S9). To test the reliability of these mutations, we checked the read mapping of these mutations and found that all mutations were real and well supported by read mapping. Further inspection found that 32 of these mutations were fixed in the population
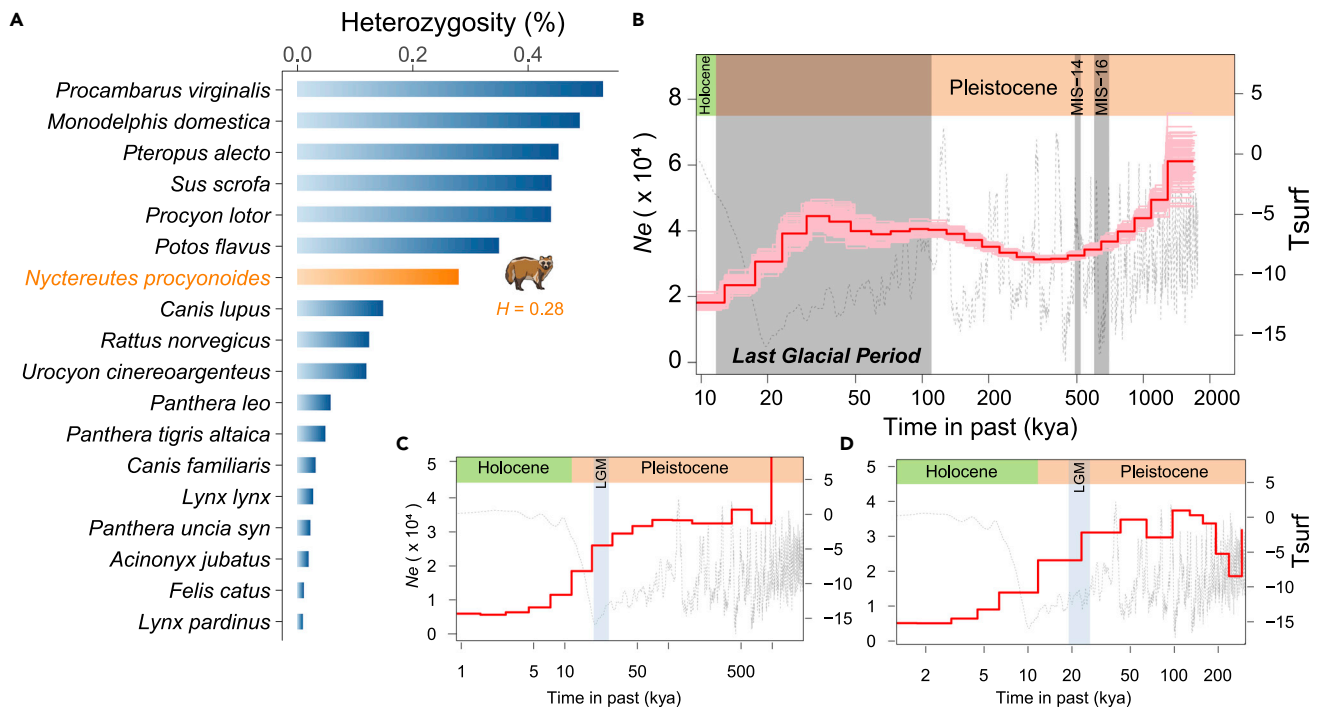
Figure 5. Population history and genome-wide heterozygosity

(A) The genomic heterozygosity of the raccoon dog and 17 other species. The genomic heterozygosity of 17 species was collected from published data.

(B) The population history of the raccoon dog inferred by PSMC with 100 bootstraps. The red line represents the estimated effective population size (Ne), and the 100 thin red lines represent the PSMC estimates of 100 randomly resampled from the original sequence. Tsurf: atmospheric surface air temperature relative to the present. The mutation rate (μ) and generation interval (g) used here were 1.0×10^{-8} and 3 years, respectively.

(C) The recent population history of the raccoon dog inferred by MSMC2 with four individuals. LGM: Last Glacial Maximum. We used the same μ and g as used in the PSMC analysis.

(D) The recent population history of the raccoon dog inferred by SMC++ with 38 individuals. We used the same μ and g as used in the PSMC analysis.

(Table S21). We also found that 10 tumor suppressor genes (*RERG*, *BRCA1*, *FETUB*, *RNF20*, *MY O 18B*, *RBM5*, *BMP3*, *LACTB*, *UNC5A*, and *PCK1*) (Table S22) and 6 energy metabolism-related genes (*CDKAL1*, *TRIM63*, *GALNT13*, *PRKAG3*, *RPUSD4*, and *MRPL19*) were under positive selection (Table S23).

To further explore the possible genetic basis of winter sleep in raccoon dogs, we performed selection analysis by comparing raccoon dogs with other 14 other non-hibernating animals (Table S24). In total, we found that 194 genes were under positive selection (Table S25). As expected, we found 21 immune-related genes (Table S26) and five energy metabolism-related genes (Table S27). Genes with feeding behaviors were also found to be under positive selection (*HCRTR1*, *GOLGB1*, and *MLN*) (Table S28). In particular, we found three genes related to insulin secretion (*STXBP5L*) and lipid metabolism (*SEC16B* and *BSCL2*) (Table S28).

Population history and genetic diversity

Although the genetic diversity of raccoon dogs has been explored by using mtDNA and microsatellite loci (Hong et al., 2018; Slaska and Grzybowska-Szatkowska, 2011; Slaska et al., 2010), this is the first study reporting genetic diversity evaluation at the whole-genome scale. To facilitate comparison with other species, we collected the heterozygosity (*H*) data of 17 other species, including the Felidae, Canidae, and some other invasive species (Table S29). The genome-wide *H* of the raccoon dog was identified as 0.28 (Figure 5A). This *H* is lower than some other invasive species; however, it is obviously higher than many Felidae and Canidae species listed in this study (Figure 5A and Table S29).

We reconstructed the history of effective population size (*Ne*) of the raccoon dog from ~1.6 million years ago (Mya) to nearly 1,000 years ago (kya), by combining the PSMC (Li and Durbin, 2011), MSMC2 (Schiffels

and Durbin, 2014), and SMC++ (Terhorst et al., 2017) methods. The PSMC result showed that the *Ne* gradually declined from ~1.6 Mya to ~320 kya and then experienced an expansion from ~320 kya to ~100 kya. The *Ne* was relatively stable between ~100 kya and 50 kya, and the population declined again to 10 kya (Figure 5B). The PSMC usually shows poor performance for inferring recent population history (Li and Durbin, 2011). We therefore used MSMC2 to infer a more recent population history, and the results showed that *Ne* presented a continuous decline from 50 kya to 1.0 kya without any rebounding (Figure 5C). Since the switch error rate induced by phasing influences the accuracy of MSMC2, we further validated the results from MSMC2 by using SMC++, which is independent of the phasing process. As we expected, the trend in *Ne* inferred by SMC++ was extremely similar to that inferred by MSMC2 (Figure 5D).

DISCUSSION

Improved genome assembly and annotation

A chromosome-scale genome assembly with high completeness, accuracy, and contiguity can provide a solid foundation for genetic research (Rhie et al., 2021). Although a draft genome of the raccoon dog has been recently reported (Chueca et al., 2021), the authors used a female individual for genome assembly, which lacked data for the Y chromosome. The raccoon dog genome assembly needs to be further improved. Here, we report a comprehensively improved chromosome-scale genome by combining PacBio and Hi-C technologies for this species. This genome showed 1.19-fold and 1.55-fold improvements in N50 values of the contig and scaffold over those of the Rac 1.0 genome (Chueca et al., 2021), respectively, indicating substantial improvement in this chromosome-scale genome assembly in terms of contiguity. The karyotypic study of the raccoon dog showed 26 pairs of autosomes, one pair of allosomes, and 2 to 3 B chromosomes (Nie et al., 2003). Here, we identified 27 chromosomes by Hi-C scaffolding, which was consistent with the karyotyping result. Surprisingly, all fission and fusion events between the domestic dog and raccoon dog genomes were highly consistent with those found in a previously published karyotypic study (Nie et al., 2003), indicating the high accuracy of this assembly at the chromosome level. In addition, we identified a 127.44 Mb X chromosome and 3.24 Mb Y-linked region with multiple lines of evidence. This is the first time we identified a genome-wide Y-linked region in the raccoon dog. The genome size estimated in this study was ~3.21 Gb (Figure S10), which is similar to that estimated by flow cytometry (Chueca et al., 2021), indicating the existence of B chromosomes. However, we failed to assemble the B chromosomes probably due to clusters of telomeric sequences along the whole B chromosome (Bugno-Poniewierska et al., 2014). For the gene annotation, although the gene number we identified is much less than that in the Rac 1.0 genome, the number we annotated is more comparable with those of other Canidae species (Kukekova et al., 2018; Peng et al., 2021). Altogether, we presented a much-improved raccoon dog genome, which will be a valuable resource for future studies.

Genetic basis for the omnivorous diet of the raccoon dog

The raccoon dog is a typical omnivorous canid species. We found that genes encoding receptors that mediate sweet, bitter, and umami tastes all exist in the raccoon dog genome and are structurally intact. As in humans, the most typical representatives of omnivores, these functional taste receptor genes provide the basic genetic basis for the omnivorous diet (Gravina et al., 2013). The sweet and umami taste receptors help the raccoon dog to forage fruits and meat, and the bitter taste receptor plays an important role in recognizing poisonous substances such as plant alkaloids that are typically bitter in taste (Li and Zhang, 2014). However, many toxic secondary metabolites in plants (Xia et al., 2021) and toxic byproducts in rotting food (Blumstein et al., 2017) cannot be tasted. GST and CYP450 are reported to be the main enzymes that catalyze the phase I and phase II reactions of many toxic substances in organisms (Zhao et al., 2020). The expanded gene families in the raccoon dog were enriched in one detoxification-related GO term (drug metabolic process) and two KEGG pathways (drug metabolism-cytochrome P450 and glutathione metabolism). The cytochromes P450 (CYP) superfamily is a large enzymatic protein family in both animals and plants that participates in multiple and critical physiological processes (Danielson, 2002; Xu et al., 2015). One of the most important functions of the CYP450 superfamily is the degradation of toxic substances, which has been reported in many species (Gui et al., 2020; Li et al., 2018). Glutathione (GSH) plays an important role in detoxification by forming GSH-toxin conjugates (either with xenobiotic or endogenous electrophilic compounds) (DeLeve and Kaplowitz, 1991; Ketterer et al., 1983) (Figure 4A). Although some conjugations can occur spontaneously, conjugations of GSH to a wide variety of toxic substances are catalyzed by GST, a powerful detoxification enzyme (DeLeve and Kaplowitz, 1991). Interestingly, we found that the GST gene family was significantly expanded in the raccoon dog genome compared with omnivorous, herbivorous, and carnivorous animals. GST can affect the activity of many antioxidant enzymes to help

remove contaminants and reactive oxygen species (Yao et al., 2020). In addition, we also found that the *ABCC12* gene was under positive selection compared with other canids. The *ABCC12* gene is a member of the ABC gene family, and the proteins encoded by this gene family play a role in detoxification (Dermauw and Van Leeuwen, 2014). Although the exact function of *ABCC12* is unknown, this gene is a member of the ABCC subfamily, which belongs to the group of multidrug resistance protein (MRPs) (Ono et al., 2007; Whitlock and Leslie, 2020), which play important roles in the function of GST pumps (Muller et al., 1996). Taken together, from these results, we inferred that GST, which is involved in glutathione metabolism, plays an important role in physiological processes associated with detoxification, perhaps even being more important than the CYP 450 superfamily in raccoon dogs. We also found two other detoxification-related genes that were under positive selection, *FM O 1* and *DAO*. FMO genes can be found across the whole animal kingdom, in which they play roles in the detoxification of xenobiotics by catalyzing the conversion of heteroatom-containing chemicals to excretable metabolites (Sehlmeyer et al., 2010). The *DAO* gene encodes diamine oxidase, which also plays an important role in dietary histamine detoxification to reduce the uptake of histamine by enterocytes (Yoshikawa et al., 2019). All these detoxification-related genes, gene families, and pathways are possible candidates for the genetic basis of the raccoon dog's broad diet.

High reproduction rate from a genetic perspective

The raccoon dog has a high reproductive rate with an average litter size of 8–10 individuals (Helle and Kauhala, 1995; Kauhala, 1996; Kauhala and Kowalczyk, 2011; Kowalczyk et al., 2009), which is larger than that of the red fox, its close relative (Kauhala, 1996). Although we did not find significantly enriched GO and KEGG pathways that are directly related to reproduction in the raccoon dog genome, we still found several reproduction-associated genes that were under positive selection. The *TDRD6* (Vasileva et al., 2009), *INK4C* (Zindy et al., 2001), and *TSSK4* (Wang et al., 2015) genes were reported to be involved in spermatogenesis. The development of spermatids from round to elongated morphology will be abrogated in mice with *TDRD6* gene knockout (Vasileva et al., 2009). *INK4C* and *INK4D* are two essential genes for male fertility and collaborate in regulating spermatogenesis. The deletion of *INK4C* and *INK4D* in mice results in male infertility (Zindy et al., 2001). Mutation in the *TSSK4* gene was also found to be harmful to spermatogenesis in infertile Chinese men (Su et al., 2008). The *ZP3* (Wassarman, 1999) and *PLCD4* (Fukami et al., 2001) genes play key roles in the acrosome reaction. The acrosome reaction is the fusion process of sperm and egg cells, and ZP3 is considered the sperm receptor of the oocyte and triggers the acrosome reaction (Jungnickel et al., 2007); *PLCD4* is responsible for mediating this zona pellucida-induced acrosome reaction (Fukami et al., 2001). We speculated that these genes may help to maintain high reproduction in raccoon dogs, especially for males. The raccoon dog-specific amino acid changes were also found in the *ZP3* and *TDRD6* genes (Figures 4C and 4D), which may also play positive roles in the function of spermatogenesis and acrosome reaction. However, experimental analysis is necessary for further validation. In addition, many of these genes were pleiotropic (Lutful Kabir et al., 2015; Sayyab et al., 2016), and these genes and mutations were only candidates for future validation and studies on the genetic basis of high reproduction of raccoon dogs.

Genomic basis for the diverse immune system

The raccoon dog is a well-known reservoir host for many pathogens (Aoyagi et al., 2000; Guan et al., 2003; Kauhala and Kowalczyk, 2011; Kjaer et al., 2021; Qi et al., 2009; Song et al., 2019), but raccoon dogs with many pathogens are often asymptomatic, without signs of illness (Han et al., 2010; Yang et al., 2021b). We expected that the raccoon dog has a diverse immune system. We found several significantly enriched KEGG pathways related to the immune system. The base excision repair pathway is essential for the diversification of antigen receptors that shape the adaptive immune response (Stratigopoulou et al., 2020). Necroptosis can release danger signals to provoke the immune system for the clearance of pathogens, and can affect innate immunity by actively inducing the death of infected cells (Cho, 2020). The spliceosome is reported to have links with immune signaling, and some core spliceosome components present immune functions in cells (Yang et al., 2021a). The proteasome is considered necessary for the important immune functions of activated CD4$^+$ T cells and proteasome inhibition suppresses unwanted and deregulated immune responses mediated by T cells (Berges et al., 2008). Although the primary functions of these enriched KEGG pathways are not immune functions, it is still possible to improve the immune system of raccoon dogs. Interestingly, a large proportion (14.56%) of PSGs were found to be immune-related, which was much large than the proportions for other functions. The functions of these genes are widely related to many innate and adaptive immune system processes (Table S19). Interestingly, raccoon dog-specific amino

acid changes were found in 18 of the 30 immune genes and seemed to be retained by positive selection (Tables S20 and S21 and Figure S9). Further population genomic analysis showed that most of these amino acid changes were fixed in the population, even though the genetic diversity of the raccoon dog was high (Figure 5A), indicating the potential advantages of these mutations to raccoon dog survival. We did not predict the effects of these mutations on the protein function, because it is hard to conclude that a single mutation is harmful or helpful to the function of a certain protein. However, multiple mutations in a single gene may collaborate to aid their immune adaptation to new habitats, but this prediction should still be experimentally validated in future works. Although certain types of cancer have been reported in raccoon dogs, we did not find a comprehensive study on the investigation of the cancer incidence statistics in this species. A series of tumor suppressor genes found in the PSGs indicated the possible strong cancer survival ability of raccoon dogs.

It is energetically very demanding for the immune system to defend against pathogens. The synthesis of a cytokine requires the breakdown of 1,150 glucose molecules to generate 2,300 ATP molecules (Straub et al., 2010). The energy demand of the immune system usually represents 25%–30% of basal metabolism (Hortova-Kohoutkova et al., 2021; Straub et al., 2010). Balancing the energetic trade-off between immunity and other physiological energy-consuming processes is critically important for animal survival (Ganeshan et al., 2019). The cost of the immune response for innate and adaptive immunity is variable. The systemic innate immune response is the most resource-demanding, followed by cell-mediated immunity, and humoral responses cost the least when compared with innate and cell-mediated defenses due to inflammation being rarely induced (Lee and Klasing, 2004). A strategy to favor a less costly adaptive immune system and reduce the high resource-demanding innate immune response should be beneficial for allocating more energy for reproduction and growth (Lee and Klasing, 2004). However, we found that most PSGs (immune-related) were involved in innate immunity and induced an inflammatory response with fever, a loss of appetite and fatigue, which was energetically expensive. The large number of expanded gene families and PSGs related to energy metabolism in the raccoon dog genome may play important roles in supplying energy for their immune response.

Several genetic factors may be jointly involved in the winter sleep of raccoon dogs

The most unique characteristic of raccoon dogs among canids is winter sleep (Asikainen et al., 2004). Raccoon dogs will fatten themselves before winter sleep (Nieminen et al., 2002). First, the omnivorous diet is very conducive to raccoon dogs to foraging and fattening. Except for taste receptor and detoxification-related genes, the two enriched olfactory transduction and salivary secretion pathways were also closely related to foraging behaviors and digestion (Jiahuan et al., 2018; Pedersen et al., 2018). The *GOLGB1* gene plays a crucial role in the development of the mammalian palate (Lan et al., 2016). The *HCRTR1* gene is the receptor gene of orexin, which is involved in foraging behavior and the intake of energy-dense food (Barson, 2020). The *MLN* gene encodes a small peptide hormone, *motilin*, which is secreted by the small intestine and regulates gastrointestinal motility and contractions, and stimulates hunger signaling (Tack et al., 2016). A genome-wide association study also detected a significant variant that was closest to the *MLN* gene in a study of the seasonal onset of hibernation in the 13-lined ground squirrel (Grabek et al., 2019). In addition, we also found that several lipid metabolism-related genes were under positive selection. The *SEC16B* gene in mice is required for lipid absorption, and is reported to be closely related to obesity in the human population (Sahibdeen et al., 2018; Shi et al., 2021). The *BSCL2* gene is also involved in lipid metabolism, and the deletion of *BSCL2* in mature white and brown adipose tissue triggers cAMP/PKA-mediated lipolysis and fatty acid oxidation resulting in adipose tissue loss (Zhou et al., 2020, 2022). These genes may facilitate intensive foraging and fat accumulation before winter sleep.

Unlike typical hibernation, the body temperature of raccoon dogs is close to normal during winter sleep (Asikainen et al., 2004). We found that the expanded gene families were enriched in thermogenesis (map04714, p = 8.36E-10) which may help raccoon dogs to maintain a near-normal body temperature. During winter sleep, fuel use shifts from glucose to lipids. Bears exhibit insulin resistance during hibernation and regain insulin sensitivity in the spring (Rigano et al., 2017). The positively selected gene *STXBP5L* is reported to be a negative regulator of insulin secretion, which may be involved in the usage shift of fuel in raccoon dogs during winter sleep. In addition, a large number of PSGs related to immunity may play an important role in preventing raccoon dog from becoming infected with the pathogen during winter sleep.

Historical declining population and currently high genetic diversity

Although the raccoon dog is considered a successful invasive species with a strong ability to survive, its effective population size was detected to be continuously declining from ~50 kya. The population decline accelerated from ~28 kya to 11 kya, which may be caused by the harsh climate during the Last Glacial Maximum (Denton et al., 2010). However, despite the climate becoming warmer at the beginning of the Holocene, the effective size of the raccoon dog population was still declining. Considering the frequent human activity within the most recent 10 kya, we speculate that the population decline of the raccoon dog in its evolutionary history might in part be due to human activities and climate change. In addition, changes in predator and prey abundance, habitat connectivity, and habitat productivity may also influence the fluctuation of the raccoon dog population.

Generally, population size is positively correlated with the level of genetic diversity (Ouborg et al., 2010). Interestingly, the genetic diversity of the raccoon dog is high, which was consistent with previous reports (Lavrov, 1971; Pitra et al., 2009), indicating a current healthy population. The historical long-term population decline of this species seems to indicate that this species may not be as tolerant and plastic as we expected for such a successful invader. The success of invasion of a certain species is not directly related to its population size but is closely related to the "preadapted" characteristics and post-invasive adaptation (North et al., 2021). We cannot investigate the post-invasive adaptation of raccoon dogs after invasion due to the lack of global samples. The "preadapted" biological characteristics, such as an omnivorous diet, diverse immune system, high reproduction rate, and winter sleep, could contribute to their successful invasiveness, even though raccoon dogs are experiencing a long-term population decline.

Limitations of the study

Although we assembled a high-quality chromosome-scale genome of the raccoon dog, we did not assemble the B chromosomes, which may contribute to important biological functions. In addition, functional analysis should be further performed to validate the candidate genes and raccoon dog-specific mutations we found in this study to be associated with omnivory, reproduction, and immunity.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
  - Nucleic acid extraction, DNA library preparation and sequencing
  - Genome assembly, annotation, and assessment
  - Phylogeny reconstruction with genome-wide single-copy genes
  - Gene family expansion and contraction
  - Positively selected genes (PSGs)
  - Genome-wide variants calling and quality control
  - Genomic diversity and population demography
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2022.105117.

AUTHOR CONTRIBUTIONS

DECLARATION OF INTERESTS

SUPPORTING CITATIONS

The following reference appears in the supplemental information: Abascal et al., 2016; Anthoni et al., 2007; Arroyo et al., 2016; Aslanukov et al., 2005; Bhatnagar et al., 2011; CABI, 2019; Cho et al., 2013; Coppo et al., 2016; Corbett-Detig et al., 2015; Cotton et al., 2022; Demeyer et al., 2019; Ding et al., 2020; Dobrynin et al., 2015; Dolgachev et al., 2018; Funakoshi et al., 2015; Guo et al., 2021; Gutekunst et al., 2018; Hitomi et al., 2019; Huber et al., 2008; Ikutani et al., 2012; Jacquet et al., 2020; Jamsai et al., 2017; Kallianpur et al., 2014; Keckesova et al., 2017; Kisiel et al., 2013; Korman et al., 2008; Lee et al., 2019; Leffler et al., 2012; Li et al., 2021; Lindblad-Toh et al., 2005; Liu et al., 2015; Maciejewska-Skrendo et al., 2019; Mehawej et al., 2020; Mikkelsen et al., 2007; Moore et al., 2008; Murphy, 2020; Nishimura et al., 2021; Nishioka et al., 2002; Palazon et al., 2014; Palmer et al., 2017; Peris-Moreno et al., 2020; Pickering and Booty, 2021; Qin et al., 2020; Ranasinghe and Eri, 2018; Readinger et al., 2009; Robinson et al., 2016; Ryan et al., 2012; Salzer et al., 2016; Santoni et al., 2013; Schenkel et al., 2007; Seo et al., 2005; Sheikh et al., 2015; Shema et al., 2008; Silver and Livingston, 2012; Smith, 2010; Strobl et al., 2011; Tsuchiya et al., 2021; Tuo et al., 2019; Valizadeh et al., 2015; Wang et al., 2016; Wang et al., 2018; Warren et al., 2015; Yang et al., 2016; Yoshinaka et al., 2019; Zhan et al., 2020; Zhang et al., 2012; Zhang et al., 2013; Zhao et al., 2017; Zhu et al., 2018.

REFERENCES

Aoyagi, T., Sato, Y., Matsuura, S., and Wada, H. (2000). Listeriosis in a raccoon dog (Nyctereutes procyonoides) associated with canine distemper. J. Vet. Med. Sci. 62, 639–641. https://doi.org/10.1292/jvms.62.639.

Abascal, F., Corvelo, A., Cruz, F., Villanueva-Cañas, J.L., Vlasova, A., Marcet-Houben, M., Martínez-Cruz, B., Cheng, J.Y., Prieto, P., Quesada, V., et al. (2016). Extreme genomic erosion after recurrent demographic bottlenecks in the highly endangered Iberian lynx. Genome Biol. 17, 251–319.

Anthoni, H., Zucchelli, M., Matsson, H., Müller-Myhsok, B., Fransson, I., Schumacher, J., Massinen, S., Onkamo, P., Warnke, A., Griesemann, H., et al. (2007). A locus on 2p12 containing the co-regulated MRPL19 and C2ORF3 genes is associated to dyslexia. Hum. Mol. Genet. 16, 667–677. https://doi.org/10.1093/hmg/ddm009.

Arroyo, J.D., Jourdain, A.A., Calvo, S.E., Ballarano, C.A., Doench, J.G., Root, D.E., and Mootha, V.K. (2016). A genome-wide CRISPR death screen identifies genes essential for oxidative phosphorylation. Cell Metab. 24, 875–885. https://doi.org/10.1016/j.cmet.2016.08.017.

Aslanukov, A., Bhowmick, R., Guruju, M., Oswald, J., Raz, D., Bush, R.A., Sieving, P., Lu, X., Bock, C.B., and Ferreira, P.A. (2005). RanBP2 modulates Cox11 and hexokinase I activities and haploinsufficiency of RanBP2 causes deficits in hexokinase I and glucose metabolism. PLoS genetics. PLoS Genet. preprint, e177. https://doi.org/10.1371/journal.pgen.0020177.eor.

Asikainen, J., Mustonen, A.M., Hyvärinen, H., and Nieminen, P. (2004). Seasonal physiology of the wild raccoon dog (Nyctereutes procyonoides). Zoolog. Sci. 21, 385–391. https://doi.org/10.2108/zsj.21.385.

Barson, J.R. (2020). Orexin/hypocretin and dysregulated eating: promotion of foraging behavior. Brain Res. 1731, 145915. https://doi.org/10.1016/j.brainres.2018.08.018.

Becker, S.E.D., Thomas, R., Trifonov, V.A., Wayne, R.K., Graphodatsky, A.S., and Breen, M. (2011). Anchoring the dog to its relatives reveals new evolutionary breakpoints across 11 species of the Canidae and provides new clues for the role of B chromosomes. Chromosome Res. 19, 685–708. https://doi.org/10.1007/s10577-011-9233-4.

Berges, C., Haberstock, H., Fuchs, D., Miltz, M., Sadeghi, M., Opelz, G., Daniel, V., and Naujokat, C. (2008). Proteasome inhibition suppresses essential immune functions of human CD4+ T cells. Immunology 124, 234–246. https://doi.org/10.1111/j.1365-2567.2007.02761.x.

Benjamini, Y., and Yekutieli, D. (2005). False discovery rate–adjusted multiple confidence intervals for selected parameters. J. Am. Stat. Assoc. 100, 71–81. https://doi.org/10.1198/016214504000001907.

Benson, G. (1999). Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Res. 27, 573–580. https://doi.org/10.1093/nar/27.2.573.

Bhatnagar, S., Oler, A.T., Rabaglia, M.E., Stapleton, D.S., Schueler, K.L., Truchan, N.A., Worzella, S.L., Stoehr, J.P., Clee, S.M., Yandell, B.S., et al. (2011). Positional cloning of a type 2 diabetes quantitative trait locus; tomosyn-2, a negative regulator of insulin secretion. PLoS Genet. 7, e1002323. https://doi.org/10.1371/journal.pgen.1002323.

Birney, E., Clamp, M., and Durbin, R. (2004). GeneWise and genomewise. Genome Res. 14, 988–995.

Blumstein, D.T., Rangchi, T.N., Briggs, T., De Andrade, F.S., and Natterson-Horowitz, B. (2017). A systematic review of carrion eaters' adaptations to avoid sickness. J. Wildl. Dis. 53, 577–581. https://doi.org/10.7589/2016-07-162.

Bock, K.W. (2016). The UDP-glycosyltransferase (UGT) superfamily expressed in humans, insects

and plants: animal-plant arms-race and co-evolution. Biochem. Pharmacol. *99*, 11–17. https://doi.org/10.1016/j.bcp.2015.10.001.

Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics *30*, 2114–2120.

Bugno-Poniewierska, M., Solek, P., Wronski, M., Potocki, L., Jezewska-Witkowska, G., and Wnuk, M. (2014). Genome organization and DNA methylation patterns of B chromosomes in the red fox and Chinese raccoon dogs. Hereditas *151*, 169–176. https://doi.org/10.1111/hrd2.00049.

CABI (2019). Invasive species compendium: nyctereutes procyonoides (raccoon dog). https://www.cabi.org/isc/datasheet/72656#6F91F1E1-592B-4800-91A5-AE0EB88BC750.

Campbell, M.S., Holt, C., Moore, B., and Yandell, M. (2014). Genome annotation and curation using MAKER and MAKER-P. Curr. Protoc. Bioinformatics *48*, 4.11.1–4.11.39. https://doi.org/10.1002/0471250953.bi0411s48.

Capella-Gutierrez, S., Silla-Martínez, J.M., and Gabaldón, T. (2009). trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics *25*, 1972–1973. https://doi.org/10.1093/bioinformatics/btp348.

Chang, J.S., Yoon, I.H., Lee, J.H., Kim, K.R., An, J., and Kim, K.W. (2010). Arsenic detoxification potential of aox genes in arsenite-oxidizing bacteria isolated from natural and constructed wetlands in the Republic of Korea. Environ. Geochem. Health *32*, 95–105. https://doi.org/10.1007/s10653-009-9268-z.

Cho, Y.S., Hu, L., Hou, H., Lee, H., Xu, J., Kwon, S., Oh, S., Kim, H.-M., Jho, S., Kim, S., et al. (2013). The tiger genome and comparative analysis with lion and snow leopard genomes. Nat. Commun. *4*, 2433–2437.

Chen, F.Z., You, L.J., Yang, F., Wang, L.N., Guo, X.Q., Gao, F., Hua, C., Tan, C., Fang, L., Shan, R.Q., et al. (2020). CNGBdb: China National GeneBank DataBase. Hereditas *42*, 799–809. https://doi.org/10.16288/j.yczz.20-080.

Cho, Y.S. (2020). Roles of necroptosis in immune responses. Biomed. Res. Rev. *4*, 5.

Choi, Y., and Chan, A.P. (2015). PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels. Bioinformatics *31*, 2745–2747. https://doi.org/10.1093/bioinformatics/btv195.

Coppo, M., Chinenov, Y., Sacta, M.A., and Rogatsky, I. (2016). The transcriptional coregulator GRIP1 controls macrophage polarization and metabolic homeostasis. Nat. Commun. *7*, 12254. https://doi.org/10.1038/ncomms12254.

Corbett-Detig, R.B., Hartl, D.L., and Sackton, T.B. (2015). Natural selection constrains neutral diversity across a wide range of species. PLoS Biol. *13*, e1002112.

Cotton, T.R., Cobbold, S.A., Bernardini, J.P., Richardson, L.W., Wang, X.S., and Lechtenberg, B.C. (2022). Structural basis of K63-ubiquitin chain formation by the Gordon-Holmes syndrome RBR E3 ubiquitin ligase RNF216. Mol. Cell *82*, 598–

615.e8. https://doi.org/10.1016/j.molcel.2021.12.005.

Chueca, L.J., Kochmann, J., Schell, T., Greve, C., Janke, A., Pfenninger, M., and Klimpel, S. (2021). De novo genome assembly of the raccoon dog (Nyctereutes procyonoides). Front. Genet. *12*, 658256. https://doi.org/10.3389/fgene.2021.658256.

Danielson, P.B. (2002). The cytochrome P450 superfamily: biochemistry, evolution and drug metabolism in humans. Curr. Drug Metab. *3*, 561–597. https://doi.org/10.2174/1389200023337054.

DeLeve, L.D., and Kaplowitz, N. (1991). Glutathione metabolism and its role in hepatotoxicity. Pharmacol. Ther. *52*, 287–305. https://doi.org/10.1016/0163-7258(91)90029-l.

Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., Handsaker, R.E., Lunter, G., Marth, G.T., Sherry, S.T., et al. (2011). The variant call format and VCFtools. Bioinformatics *27*, 2156–2158.

De Bie, T., Cristianini, N., Demuth, J.P., and Hahn, M.W. (2006). CAFE: a computational tool for the study of gene family evolution. Bioinformatics *22*, 1269–1271. https://doi.org/10.1093/bioinformatics/btl097.

Demeyer, A., Skordos, I., Driege, Y., Kreike, M., Hochepied, T., Baens, M., Staal, J., and Beyaert, R. (2019). MALT1 proteolytic activity suppresses autoimmunity in a T cell intrinsic manner. Front. Immunol. *10*, 1898. https://doi.org/10.3389/fimmu.2019.01898.

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*, 491–498.

Denton, G.H., Anderson, R.F., Toggweiler, J.R., Edwards, R.L., Schaefer, J.M., and Putnam, A.E. (2010). The last glacial termination. Science *328*, 1652–1656. https://doi.org/10.1126/science.1184119.

Ding, S., Zhang, H., Zhao, X., Dang, J., and Li, G. (2020). UNC5A, an epigenetically silenced gene, functions as a tumor suppressor in non-small cell lung cancer. Saudi J. Biol. Sci. *27*, 3009–3017. https://doi.org/10.1016/j.sjbs.2020.09.023.

Dobrynin, P., Liu, S., Tamazian, G., Xiong, Z., Yurchenko, A.A., Krasheninnikova, K., Kliver, S., Schmidt-Küntzel, A., Koepfli, K.-P., Johnson, W., et al. (2015). Genomic legacy of the African cheetah, Acinonyx jubatus. Genome Biol. *16*, 277–320.

Dolgachev, V., Panicker, S., Balijepalli, S., McCandless, L.K., Yin, Y., Swamy, S., Suresh, M.V., Delano, M.J., Hemmila, M.R., Raghavendran, K., and Machado-Aranda, D. (2018). Electroporation-mediated delivery of FER gene enhances innate immune response and improves survival in a murine model of pneumonia. Gene Ther. *25*, 359–375. https://doi.org/10.1038/s41434-018-0022-y.

Dudchenko, O., Batra, S.S., Omer, A.D., Nyquist, S.K., Hoeger, M., Durand, N.C., Shamim, M.S., Machol, I., Lander, E.S., Aiden, A.P., and Aiden,

E.L. (2017). De novo assembly of the Aedes aegypti genome using Hi-C yields chromosome-length scaffolds. Science *356*, 92–95.

Durand, N.C., Shamim, M.S., Machol, I., Rao, S.S.P., Huntley, M.H., Lander, E.S., and Aiden, E.L. (2016). Juicer provides a one-click system for analyzing loop-resolution Hi-C experiments. Cell Syst. *3*, 95–98.

Dermauw, W., and Van Leeuwen, T. (2014). The ABC gene family in arthropods: comparative genomics and role in insecticide transport and resistance. Insect Biochem. Mol. Biol. *45*, 89–110. https://doi.org/10.1016/j.ibmb.2013.11.001.

Drygala, F., Werner, U., and Zoller, H. (2013). Diet composition of the invasive raccoon dog (Nyctereutes procyonoides) and the native red fox (Vulpes vulpes) in north-east Germany. Hystrix *24*, 190.

Ellerman, J.R., and Morrison-Scott, T.C.S. (1951). Checklist of Palearctic and Indian Mammals, 1758 to 1946. British Museum (Natural History).

Flynn, J.M., Hubley, R., Goubert, C., Rosen, J., Clark, A.G., Feschotte, C., and Smit, A.F. (2020). RepeatModeler2 for automated genomic discovery of transposable element families. Proc. Natl. Acad. Sci. USA *117*, 9451–9457. https://doi.org/10.1073/pnas.1921046117.

Freed, D., Aldana, R., Weber, J.A., and Edwards, J.S. (2017). The sentieon genomics tools-a fast and accurate solution to variant calling from next-generation sequence data. Preprint at bioRxiv. https://doi.org/10.1101/115717.

Freedman, A.H., Gronau, I., Schweizer, R.M., Ortega-Del Vecchyo, D., Han, E., Silva, P.M., Galaverni, M., Fan, Z., Marx, P., Lorente-Galdos, B., et al. (2014). Genome sequencing highlights the dynamic early history of dogs. PLoS Genet. *10*, e1004016.

Funakoshi, S., Shimizu, T., Numata, O., Ato, M., Melchers, F., and Ohnishi, K. (2015). BILL-cadherin/cadherin-17 contributes to the survival of memory B cells. PLoS One *10*, e0117566. https://doi.org/10.1371/journal.pone.0117566.

Fukami, K., Nakao, K., Inoue, T., Kataoka, Y., Kurokawa, M., Fissore, R.A., Nakamura, K., Katsuki, M., Mikoshiba, K., Yoshida, N., and Takenawa, T. (2001). Requirement of phospholipase Cdelta4 for the zona pellucida-induced acrosome reaction. Science *292*, 920–923. https://doi.org/10.1126/science.1059042.

Ganeshan, K., Nikkanen, J., Man, K., Leong, Y.A., Sogawa, Y., Maschek, J.A., Van Ry, T., Chagwedera, D.N., Cox, J.E., and Chawla, A. (2019). Energetic trade-offs and hypometabolic states promote disease tolerance. Cell *177*, 399–413.e12. https://doi.org/10.1016/j.cell.2019.01.050.

Grabek, K.R., Cooke, T.F., Epperson, L.E., Spees, K.K., Cabral, G.F., Sutton, S.C., Merriman, D.K., Martin, S.L., and Bustamante, C.D. (2019). Genetic variation drives seasonal onset of hibernation in the 13-lined ground squirrel. Commun. Biol. *2*, 478. https://doi.org/10.1038/s42003-019-0719-5.

Gravina, S.A., Yep, G.L., and Khan, M. (2013). Human biology of taste. Ann. Saudi Med. *33*,

217–222. https://doi.org/10.5144/0256-4947.2013.217.

Guan, D., McCarthy, S.A., Wood, J., Howe, K., Wang, Y., and Durbin, R. (2020). Identifying and removing haplotypic duplication in primary genome assemblies. Bioinformatics 36, 2896–2898.

Guo, Y., Jiang, F., Kong, L., Wu, H., Zhang, H., Chen, X., Zhao, J., Cai, B., Li, Y., Ma, C., et al. (2021). OTUD5 promotes innate antiviral and antitumor immunity through deubiquitinating and stabilizing STING. Cell. Mol. Immunol. 18, 1945–1955. https://doi.org/10.1038/s41423-020-00531-5.

Gutekunst, J., Andriantsoa, R., Falckenhayn, C., Hanna, K., Stein, W., Rasamy, J., and Lyko, F. (2018). Clonal genome evolution and rapid invasive spread of the marbled crayfish. Nat. Ecol. Evol. 2, 567–573. https://doi.org/10.1038/s41559-018-0467-9.

Guan, Y., Zheng, B.J., He, Y.Q., Liu, X.L., Zhuang, Z.X., Cheung, C.L., Luo, S.W., Li, P.H., Zhang, L.J., Guan, Y.J., et al. (2003). Isolation and characterization of viruses related to the SARS coronavirus from animals in southern China. Science 302, 276–278. https://doi.org/10.1126/science.1087139.

Gui, F., Lan, T., Zhao, Y., Guo, W., Dong, Y., Fang, D., Liu, H., Li, H., Wang, H., Hao, R., et al. (2020). Genomic and transcriptomic analysis unveils population evolution and development of pesticide resistance in fall armyworm Spodoptera frugiperda. Protein Cell. https://doi.org/10.1007/s13238-020-00795-7.

Guo, X., Chen, F., Gao, F., Li, L., Liu, K., You, L., Hua, C., Yang, F., Liu, W., Peng, C., et al. (2020). CNSA: A Data Repository for Archiving Omics Data (Database-The Journal of Biological Databases and Curation). Database 2020. https://doi.org/10.1093/database/baaa055.

Han, J.I., Lee, S.J., Jang, H.J., and Na, K.J. (2010). Asymptomatic Babesia microti-like parasite infection in wild raccoon dogs (Nyctereutes procyonoides) in South Korea. J. Wildl. Dis. 46, 632–635. https://doi.org/10.7589/0090-3558-46.2.632.

Haas, B.J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P.D., Bowden, J., Couger, M.B., Eccles, D., Li, B., Lieber, M., et al. (2013). De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat. Protoc. 8, 1494–1512. https://doi.org/10.1038/nprot.2013.084.

Haas, B.J., Salzberg, S.L., Zhu, W., Pertea, M., Allen, J.E., Orvis, J., White, O., Buell, C.R., and Wortman, J.R. (2008). Automated eukaryotic gene structure annotation using EVidenceModeler and the Program to Assemble Spliced Alignments. Genome Biol. 9, R7–R22.

Han, Y., and Wessler, S.R. (2010). MITE-Hunter: a program for discovering miniature inverted-repeat transposable elements from genomic sequences. Nucleic Acids Res. 38, e199. https://doi.org/10.1093/nar/gkq862.

Hayes, J.D., and Pulford, D.J. (1995). The glutathione S-transferase supergene family:

regulation of GST and the contribution of the isoenzymes to cancer chemoprotection and drug resistance. Crit. Rev. Biochem. Mol. Biol. 30, 445–600. https://doi.org/10.3109/10409239509083491.

Helle, E., and Kauhala, K. (1995). Reproduction in the raccoon dog in Finland. J. Mammal. 76, 1036–1046.

Hitomi, Y., Nakatani, K., Kojima, K., Nishida, N., Kawai, Y., Kawashima, M., Aiba, Y., Nagasaki, M., Nakamura, M., and Tokunaga, K. (2019). NFKB1 and MANBA confer disease susceptibility to primary biliary cholangitis via independent putative primary functional variants. Cell. Mol. Gastroenterol. Hepatol. 7, 515–532. https://doi.org/10.1016/j.jcmgh.2018.11.006.

Hirasawa, M., Kanda, E., and Takatsuki, S. (2006). Seasonal food habits of the raccoon dog at a western suburb of Tokyo. Mamm. Stud. 31, 9–14.

Hong, Y., Kim, K.S., Kimura, J., Kauhala, K., Voloshina, I., Goncharuk, M.S., Yu, L., Zhang, Y.P., Sashika, M., Lee, H., and Min, M.S. (2018). Genetic diversity and population structure of East Asian raccoon dog (Nyctereutes procyonoides): genetic features in central and marginal populations. Zoolog. Sci. 35, 249–259. https://doi.org/10.2108/zs170140.

Hortova-Kohoutkova, M., Lázničková, P., and Frič, J. (2021). How immune-cell fate and function are determined by metabolic pathway choice: the bioenergetics underlying the immune response. Bioessays 43, e2000067. https://doi.org/10.1002/bies.202000067.

Hu, J., Fan, J., Sun, Z., and Liu, S. (2020). NextPolish: a fast and efficient genome polishing tool for long-read assembly. Bioinformatics 36, 2253–2255.

Huber, C., Mårtensson, A., Bokoch, G.M., Nemazee, D., and Gavin, A.L. (2008). FGD2, a CDC42-specific exchange factor expressed by antigen-presenting cells, localizes to early endosomes and active membrane ruffles. J. Biol. Chem. 283, 34002–34012. https://doi.org/10.1074/jbc.M803957200.

Ikutani, M., Yanagibashi, T., Ogasawara, M., Tsuneyama, K., Yamamoto, S., Hattori, Y., Kouro, T., Itakura, A., Nagai, Y., Takaki, S., and Takatsu, K. (2012). Identification of innate IL-5-producing cells and their role in lung eosinophil regulation and antitumor immunity. J. Immunol. 188, 703–713. https://doi.org/10.4049/jimmunol.1101270.

Jacquet, S., Pontier, D., and Etienne, L. (2020). Rapid evolution of HERC6 and duplication of a chimeric HERC5/6 gene in rodents and bats suggest an overlooked role of HERCs in mammalian immunity. Front. Immunol. 11, 605270. https://doi.org/10.3389/fimmu.2020.605270.

Jamsai, D., Watkins, D.N., O'Connor, A.E., Merriner, D.J., Gursoy, S., Bird, A.D., Kumar, B., Miller, A., Cole, T.J., Jenkins, B.J., and O'Bryan, M.K. (2017). In vivo evidence that RBM5 is a tumour suppressor in the lung. Sci. Rep. 7, 16323. https://doi.org/10.1038/s41598-017-15874-9.

Jiahuan, R., Wenhao, S., Xiaofan, G., Wei, S., Shanjie, Z., Maolong, H., Haifeng, W., and Guangxu, L. (2018). Ocean acidification impairs foraging behavior by interfering with olfactory neural signal transduction in black sea bream, Acanthopagrus schlegelii. Front. Physiol. 9, 1592. https://doi.org/10.3389/fphys.2018.01592.

Jones, P., Binns, D., Chang, H.-Y., Fraser, M., Li, W., McAnulla, C., McWilliam, H., Maslen, J., Mitchell, A., Nuka, G., et al. (2014). InterProScan 5: genome-scale protein function classification. Bioinformatics 30, 1236–1240.

Jungnickel, M.K., Sutton, K.A., Wang, Y., and Florman, H.M. (2007). Phosphoinositide-dependent pathways in mouse sperm are regulated by egg ZP3 and drive the acrosome reaction. Dev. Biol. 304, 116–126. https://doi.org/10.1016/j.ydbio.2006.12.023.

Kauhala, K. (1996). Reproductive strategies of the racoon dog and the red fox in Finland. Acta Theriol. 41, 51–58.

Kauhala, K., and Kowalczyk, R. (2011). Invasion of the raccoon dog Nyctereutes procyonoides in Europe: history of colonization, features behind its success, and threats to native fauna. Curr. Zool. 57, 584–598. https://doi.org/10.1093/czoolo/57.5.584.

Kallianpur, A.R., Jia, P., Ellis, R.J., Zhao, Z., Bloss, C., Wen, W., Marra, C.M., Hulgan, T., Simpson, D.M., Morgello, S., et al. (2014). Genetic variation in iron metabolism is associated with neuropathic pain and pain severity in HIV-infected patients on antiretroviral therapy. PLoS One 9, e103123. https://doi.org/10.1371/journal.pone.0103123.

Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K.F., von Haeseler, A., and Jermiin, L.S. (2017). ModelFinder: fast model selection for accurate phylogenetic estimates. Nat. Methods 14, 587–589. https://doi.org/10.1038/nmeth.4285.

Katoh, K., and Standley, D.M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. 30, 772–780. https://doi.org/10.1093/molbev/mst010.

Keckesova, Z., Donaher, J.L., De Cock, J., Freinkman, E., Lingrell, S., Bachovchin, D.A., Bierie, B., Tischler, V., Noske, A., Okondo, M.C., et al. (2017). LACTB is a tumour suppressor that modulates lipid metabolism and cell state. Nature 543, 681–686. https://doi.org/10.1038/nature21408.

Ketterer, B., Coles, B., and Meyer, D.J. (1983). The role of glutathione in detoxication. Environ. Health Perspect. 49, 59–69.

Kjaer, L.J., Jensen, L.M., Chriél, M., Bødker, R., and Petersen, H.H. (2021). The raccoon dog (Nyctereutes procyonoides) as a reservoir of zoonotic diseases in Denmark. Int. J. Parasitol. Parasites Wildl. 16, 175–182. https://doi.org/10.1016/j.ijppaw.2021.09.008.

Kisiel, J.B., Li, J., Zou, H., Oseini, A.M., Strauss, B.B., Gulaid, K.H., Moser, C.D., Aderca, I., Ahlquist, D.A., Roberts, L.R., and Shire, A.M. (2013). Methylated bone morphogenetic protein 3 (BMP3) gene: evaluation of tumor suppressor function and biomarker potential in biliary cancer.

J. Mol. Biomark. Diagn. 4, 1000145. https://doi.org/10.4172/2155-9929.1000145.

Koren, S., Walenz, B.P., Berlin, K., Miller, J.R., Bergman, N.H., and Phillippy, A.M. (2017). Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. Genome Res. 27, 722–736.

Korf, I. (2004). Gene finding in novel genomes. BMC Bioinf. 5, 59. https://doi.org/10.1186/1471-2105-5-59.

Korman, B.D., Kastner, D.L., Gregersen, P.K., and Remmers, E.F. (2008). STAT4: genetics, mechanisms, and implications for autoimmunity. Curr. Allergy Asthma Rep. 8, 398–403. https://doi.org/10.1007/s11882-008-0077-8.

Kowalczyk, R., Zalewski, A., Jędrzejewska, B., Ansorge, H., and Bunevich, A.N. (2009). Reproduction and mortality of invasive raccoon dogs (Nyctereutes procyonoides) in the białowieża primeval forest (eastern Poland). Ann. Zool. Fenn. 46, 291–301.

Kukekova, A.V., Johnson, J.L., Xiang, X., Feng, S., Liu, S., Rando, H.M., Kharlamova, A.V., Herbeck, Y., Serdyukova, N.A., Xiong, Z., et al. (2018). Red fox genome assembly identifies genomic regions associated with tame and aggressive behaviours. Nat. Ecol. Evol. 2, 1479–1491. https://doi.org/10.1038/s41559-018-0611-6.

Lan, Y., Zhang, N., Liu, H., Xu, J., and Jiang, R. (2016). Golgb1 regulates protein glycosylation and is crucial for mammalian palate development. Development 143, 2344–2355. https://doi.org/10.1242/dev.134577.

Laurimaa, L., Süld, K., Davison, J., Moks, E., Valdmann, H., and Saarma, U. (2016). Alien species and their zoonotic parasites in native and introduced ranges: the raccoon dog example. Vet. Parasitol. 219, 24–33.

Lavrov, N.P. (1971). Results of raccoon dog introductions in different parts of the Soviet Union (in Russian). Trudy Kafedry Biologii MGZPI 29, 101–160.

Lee, N., Kim, D., and Kim, W.U. (2019). Role of NFAT5 in the immune system and pathogenesis of autoimmune diseases. Front. Immunol. 10, 270. https://doi.org/10.3389/fimmu.2019.00270.

Leffler, E.M., Bullaughey, K., Matute, D.R., Meyer, W.K., Segurel, L., Venkat, A., Andolfatto, P., and Przeworski, M. (2012). Revisiting an old riddle: what determines genetic diversity levels within species? PLoS Biol. 10. 10.1371/journal.pbio.1001388.

Lee, K.A., and Klasing, K.C. (2004). A role for immunology in invasion biology. Trends Ecol. Evol. 19, 523–529. https://doi.org/10.1016/j.tree.2004.07.012.

Li, D., and Zhang, J. (2014). Diet shapes the evolution of the vertebrate bitter taste receptor gene repertoire. Mol. Biol. Evol. 31, 303–309. https://doi.org/10.1093/molbev/mst219.

Li, H., and Durbin, R. (2011). Inference of human population history from individual whole-genome sequences. Nature 475, 493–496.

Li, S., Zhu, S., Jia, Q., Yuan, D., Ren, C., Li, K., Liu, S., Cui, Y., Zhao, H., Cao, Y., et al. (2018). The

genomic and functional landscapes of developmental plasticity in the American cockroach. Nat. Commun. 9, 1008. https://doi.org/10.1038/s41467-018-03281-1.

Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. Preprint at arXiv. https://doi.org/10.48550/arXiv.1303.3997.

Li, H., Coghlan, A., Ruan, J., Coin, L.J., Hériché, J.K., Osmotherly, L., Li, R., Liu, T., Zhang, Z., Bolund, L., et al. (2006). TreeFam: a curated database of phylogenetic trees of animal gene families. Nucleic Acids Res. 34, D572–D580. https://doi.org/10.1093/nar/gkj118.

Li, H., and Durbin, R. (2010). Fast and accurate long-read alignment with Burrows–Wheeler transform. Bioinformatics 26, 589–595.

Li, W., Guan, X., Sun, B., and Sun, L. (2021). A novel microRNA of Japanese flounder regulates antimicrobial immunity involving a bacteria-binding CSF3. Front. Immunol. 12, 723401. https://doi.org/10.3389/fimmu.2021.723401.

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.

Liu, Y., Zhang, Q., Ding, Y., Li, X., Zhao, D., Zhao, K., Guo, Z., and Cao, X. (2015). Histone lysine methyltransferase Ezh1 promotes TLR-triggered inflammatory cytokine production by suppressing Tollip. J. Immunol. 194, 2838–2846. https://doi.org/10.4049/jimmunol.1402087.

Lowe, T.M., and Eddy, S.R. (1997). tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. 25, 955–964.

Lutful Kabir, F.M., Alvarez, C.E., and Bird, R.C. (2015). Canine mammary carcinomas: a comparative analysis of altered gene expression. Vet. Sci. 3, E1. https://doi.org/10.3390/vetsci3010001.

Maciejewska-Skrendo, A., Cięszczyk, P., Chycki, J., Sawczuk, M., and Smółka, W. (2019). Genetic markers associated with power athlete status. J. Hum. Kinet. 68, 17–36. https://doi.org/10.2478/hukin-2019-0053.

Majoros, W.H., Pertea, M., and Salzberg, S.L. (2004). TigrScan and GlimmerHMM: two open source ab initio eukaryotic gene-finders. Bioinformatics 20, 2878–2879. https://doi.org/10.1093/bioinformatics/bth315.

Manni, M., Berkeley, M.R., Seppey, M., Simao, F.A., and Zdobnov, E.M. (2021). BUSCO update: novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes. Mol. Biol. Evol. 38, 4647–4654. https://doi.org/10.1016/b978-0-12-816434-1.00002-4.

Mehawej, C., Khalife, H., Hanna-Wakim, R., Dbaibo, G., and Farra, C. (2020). DNMT3B deficiency presenting as severe combined immune deficiency: a case report. Clin. Immunol. 215, 108453. https://doi.org/10.1016/j.clim.2020.108453.

Mikkelsen, T.S., Wakefield, M.J., Aken, B., Amemiya, C.T., Chang, J.L., Duke, S., Garber, M., Gentles, A.J., Goodstadt, L., Heger, A., et al. (2007). Genome of the marsupial Monodelphis domestica reveals innovation in non-coding sequences. Nature 447, 167–177.

Moore, S.W., Sidler, D., and Zaahl, M.G. (2008). The ITGB2 immunomodulatory gene (CD18), enterocolitis, and Hirschsprung's disease. J. Pediatr. Surg. 43, 1439–1444. https://doi.org/10.1016/j.jpedsurg.2007.12.057.

Mount, D.W. (2007). Using the basic local alignment search tool (BLAST). Cold Spring Harb. Protoc. 2007, pdb.top17.

Murphy, S. (2020). PEX1 and PEX7 genes are necessary for completion of immune pathways and survival post-infection. STEM Fellowship J. 6, 5–11.

Mulder, J. (2012). A review of the ecology of the raccoon dog (Nyctereutes procyonoides) in Europe. Lutra 55, 27.

Muller, M., de Vries, E.G., and Jansen, P.L. (1996). Role of multidrug resistance protein (MRP) in glutathione S-conjugate transport in mammalian cells. J. Hepatol. 24 (Suppl 1), 100–108.

Nawrocki, E.P., Kolbe, D.L., and Eddy, S.R. (2009). Infernal 1.0: inference of RNA alignments. Bioinformatics 25, 1335–1337.

Nie, W., Wang, J., Perelman, P., Graphodatsky, A.S., and Yang, F. (2003). Comparative chromosome painting defines the karyotypic relationships among the domestic dog, Chinese raccoon dog and Japanese raccoon dog. Chromosome Res. 11, 735–740. https://doi.org/10.1023/b:chro.0000005760.03266.29.

Nguyen, L.T., Schmidt, H.A., von Haeseler, A., and Minh, B.Q. (2015). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol. Biol. Evol. 32, 268–274. https://doi.org/10.1093/molbev/msu300.

Nishimura, K., Tanaka, T., Takemura, S., Tatsumi, K., and Wanaka, A. (2021). SNX25 regulates proinflammatory cytokine expression via the NF-kappaB signal in macrophages. PLoS One 16, e0247840. https://doi.org/10.1371/journal.pone.0247840.

Nishioka, M., Kohno, T., Tani, M., Yanaihara, N., Tomizawa, Y., Otsuka, A., Sasaki, S., Kobayashi, K., Niki, T., Maeshima, A., et al. (2002). MYO18B, a candidate tumor suppressor gene at chromosome 22q12.1, deleted, mutated, and methylated in human lung cancer. Proc. Natl. Acad. Sci. USA 99, 12269–12274. https://doi.org/10.1073/pnas.192445899.

Nieminen, P., Mustonen, A.M., Asikainen, J., and Hyvärinen, H. (2002). Seasonal weight regulation of the raccoon dog (Nyctereutes procyonoides): interactions between melatonin, leptin, ghrelin, and growth hormone. J. Biol. Rhythms 17, 155–163. https://doi.org/10.1177/074873040201700206.

North, H.L., McGaughran, A., and Jiggins, C.D. (2021). Insights into invasive species from whole-genome resequencing. Mol. Ecol. 30, 6289–6308. https://doi.org/10.1111/mec.15999.

Ono, N., Van der Heijden, I., Scheffer, G.L., Van de Wetering, K., Van Deemter, E., De Haas, M., Boerke, A., Gadella, B.M., De Rooij, D.G., Neefjes, J.J., et al. (2007). Multidrug resistance-associated protein 9 (ABCC12) is present in mouse and boar sperm. Biochem. J. 406, 31–40. https://doi.org/10.1042/BJ20070292.

Ouborg, N.J., Pertoldi, C., Loeschcke, V., Bijlsma, R.K., and Hedrick, P.W. (2010). Conservation genetics in transition to conservation genomics. Trends Genet. 26, 177–187. https://doi.org/10.1016/j.tig.2010.01.001.

Palazon, A., Goldrath, A.W., Nizet, V., and Johnson, R.S. (2014). HIF transcription factors, inflammation, and immunity. Immunity 41, 518–528. https://doi.org/10.1016/j.immuni.2014.09.008.

Palmer, C.J., Bruckner, R.J., Paulo, J.A., Kazak, L., Long, J.Z., Mina, A.I., Deng, Z., LeClair, K.B., Hall, J.A., Hong, S., et al. (2017). Cdkal1, a type 2 diabetes susceptibility gene, regulates mitochondrial function in adipose tissue. Mol. Metab. 6, 1212–1225. https://doi.org/10.1016/j.molmet.2017.07.013.

Peris-Moreno, D., Taillandier, D., and Polge, C. (2020). MuRF1/TRIM63, master regulator of muscle mass. Int. J. Mol. Sci. 21, E6663. https://doi.org/10.3390/ijms21186663.

Pedersen, A., Sørensen, C.E., Proctor, G.B., and Carpenter, G.H. (2018). Salivary functions in mastication, taste and textural perception, swallowing and initial digestion. Oral Dis. 24, 1399–1416. https://doi.org/10.1111/odi.12867.

Peng, Y., Li, H., Liu, Z., Zhang, C., Li, K., Gong, Y., Geng, L., Su, J., Guan, X., Liu, L., et al. (2021). Chromosome-level genome assembly of the Arctic fox (Vulpes lagopus) using PacBio sequencing and Hi-C technology. Mol. Ecol. Resour. 21, 2093–2108. https://doi.org/10.1111/1755-0998.13397.

Pitra, C., Schwarz, S., and Fickel, J. (2009). Going west—invasion genetics of the alien raccoon dog Nyctereutes procynoides in Europe. Eur. J. Wildl. Res. 56, 117–129. https://doi.org/10.1007/s10344-009-0283-2.

Pflug, J.M., Holmes, V.R., Burrus, C., Johnston, J.S., and Maddison, D.R. (2020). Measuring genome sizes using read-depth, k-mers, and flow cytometry: methodological comparisons in beetles (Coleoptera). G3 10, 3047–3060. https://doi.org/10.1534/g3.120.401028.

Pickering, R.J., and Booty, L.M. (2021). NLR in eXile: emerging roles of NLRX1 in immunity and human disease. Immunology 162, 268–280. https://doi.org/10.1111/imm.13291.

Qi, X., Li, X., Rider, P., Fan, W., Gu, H., Xu, L., Yang, Y., Lu, S., Wang, H., and Liu, F. (2009). Molecular characterization of highly pathogenic H5N1 avian influenza A viruses isolated from raccoon dogs in China. PLoS One 4, e4682. https://doi.org/10.1371/journal.pone.0004682.

Qin, M., Wu, S., Li, A., Zhao, F., Feng, H., Ding, L., and Ruan, J. (2019). LRScaf: improving draft genomes using long noisy reads. BMC Genom. 20, 955–1012.

Qin, T., Jia, Y., Liu, Y., Dai, R., Zhou, L., Okada, S., Tsumura, M., Ohnishi, H., Kato, Z., Kanegane, H.,

et al. (2020). A novel homozygous mutation destabilizes IKKbeta and leads to human combined immunodeficiency. Front. Immunol. 11, 517544. https://doi.org/10.3389/fimmu.2020.517544.

Ranasinghe, R., and Eri, R. (2018). CCR6–CCL20-Mediated immunologic pathways in inflammatory bowel disease. Gastrointest. Disord. 1, 15–29. https://doi.org/10.3390/gidisord1010003.

Readinger, J.A., Mueller, K.L., Venegas, A.M., Horai, R., and Schwartzberg, P.L. (2009). Tec kinases regulate T-lymphocyte development and function: new insights into the roles of Itk and Rlk/Txk. Immunol. Rev. 228, 93–114. https://doi.org/10.1111/j.1600-065X.2008.00757.x.

Rhie, A., McCarthy, S.A., Fedrigo, O., Damas, J., Formenti, G., Koren, S., Uliano-Silva, M., Chow, W., Fungtammasan, A., Kim, J., et al. (2021). Towards complete and error-free genome assemblies of all vertebrate species. Nature 592, 737–746. https://doi.org/10.1038/s41586-021-03451-0.

Rigano, K.S., Gehring, J.L., Evans Hutzenbiler, B.D., Chen, A.V., Nelson, O.L., Vella, C.A., Robbins, C.T., and Jansen, H.T. (2017). Life in the fat lane: seasonal regulation of insulin sensitivity, food intake, and adipose biology in brown bears. J. Comp. Physiol. B 187, 649–676. https://doi.org/10.1007/s00360-016-1050-9.

Robinson, J.A., Ortega-Del Vecchyo, D., Fan, Z., Kim, B.Y., vonHoldt, B.M., Marsden, C.D., Lohmueller, K.E., and Wayne, R.K. (2016). Genomic flatlining in the endangered island fox. Curr. Biol. 26, 1183–1189.

Ryan, M.T., Hamill, R.M., O'Halloran, A.M., Davey, G.C., McBryan, J., Mullen, A.M., McGee, C., Gispert, M., Southwood, O.I., and Sweeney, T. (2012). SNP variation in the promoter of the PRKAG3 gene and association with meat quality traits in pig. BMC Genet. 13, 66. https://doi.org/10.1186/1471-2156-13-66.

Salzer, E., Santos-Valente, E., Keller, B., Warnatz, K., and Boztug, K. (2016). Protein kinase C delta: a gatekeeper of immune homeostasis. J. Clin. Immunol. 36, 631–640. https://doi.org/10.1007/s10875-016-0323-0.

Santoni, G., Farfariello, V., Liberati, S., Morelli, M.B., Nabissi, M., Santoni, M., and Amantini, C. (2013). The role of transient receptor potential vanilloid type-2 ion channels in innate and adaptive immune responses. Front. Immunol. 4, 34. https://doi.org/10.3389/fimmu.2013.00034.

Sahibdeen, V., Crowther, N.J., Soodyall, H., Hendry, L.M., Munthali, R.J., Hazelhurst, S., Choudhury, A., Norris, S.A., Ramsay, M., and Lombard, Z. (2018). Genetic variants in SEC16B are associated with body composition in black South Africans. Nutr. Diabetes 8, 43. https://doi.org/10.1038/s41387-018-0050-0.

Sayyab, S., Viluma, A., Bergvall, K., Brunberg, E., Jagannathan, V., Leeb, T., Andersson, G., and Bergström, T.F. (2016). Whole-genome sequencing of a canine family trio reveals a FAM83G variant associated with hereditary footpad hyperkeratosis. G3 6, 521–527. https://doi.org/10.1534/g3.115.025643.

Schiffels, S., and Durbin, R. (2014). Inferring human population size and separation history from multiple genome sequences. Nat. Genet. 46, 919–925. https://doi.org/10.1038/ng.3015.

Schenkel, A.R., Dufour, E.M., Chew, T.W., Sorg, E., and Muller, W.A. (2007). The murine CD99-related molecule CD99-like 2 (CD99L2) is an adhesion molecule involved in the inflammatory response. Cell Commun. Adhes. 14, 227–237. https://doi.org/10.1080/15419060701755966.

Sehlmeyer, S., Wang, L., Langel, D., Heckel, D.G., Mohagheghi, H., Petschenka, G., and Ober, D. (2010). Flavin-dependent monooxygenases as a detoxification mechanism in insects: new insights from the arctiids (Lepidoptera). PLoS One 5, e10435.

Seo, S., Asai, T., Saito, T., Suzuki, T., Morishita, Y., Nakamoto, T., Ichikawa, M., Yamamoto, G., Kawazu, M., Yamagata, T., et al. (2005). Crk-associated substrate lymphocyte type is required for lymphocyte trafficking and marginal zone B cell maintenance. J. Immunol. 175, 3492–3501. https://doi.org/10.4049/jimmunol.175.6.3492.

Shao, X.Q., Wen, Y.J., Ba, H.X., Zhang, X.T., Yue, Z.G., Wang, K.J., Li, C.Y., Qiu, J., and Yang, F.H. (2014). Novel amdoparvovirus infecting farmed raccoon dogs and arctic foxes. Emerg. Infect. Dis. 20, 2085–2088. https://doi.org/10.3201/eid2012.140289.

Shi, R., Lu, W., Tian, Y., and Wang, B. (2021). Intestinal SEC16B modulates obesity by controlling dietary lipid absorption. Preprint at bioRxiv. https://doi.org/10.1101/2021.12.07.471468.

Sheikh, F., Dickensheets, H., Pedras-Vasconcelos, J., Ramalingam, T., Helming, L., Gordon, S., and Donnelly, R.P. (2015). The interleukin-13 receptor-alpha1 chain is essential for induction of the alternative macrophage activation pathway by IL-13 but not IL-4. J. Innate Immun. 7, 494–505. https://doi.org/10.1159/000376579.

Shema, E., Tirosh, I., Aylon, Y., Huang, J., Ye, C., Moskovits, N., Raver-Shapira, N., Minsky, N., Pirngruber, J., Tarcic, G., et al. (2008). The histone H2B-specific ubiquitin ligase RNF20/hBRE1 acts as a putative tumor suppressor through selective regulation of gene expression. Genes Dev. 22, 2664–2676. https://doi.org/10.1101/gad.1703008.

Silver, D.P., and Livingston, D.M. (2012). Mechanisms of BRCA1 tumor suppression. Cancer Discov. 2, 679–684. https://doi.org/10.1158/2159-8290.CD-12-0221.

Simão, F.A., Waterhouse, R.M., Ioannidis, P., Kriventseva, E.V., and Zdobnov, E.M. (2015). BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics 31, 3210–3212.

Slaska, B., and Grzybowska-Szatkowska, L. (2011). Analysis of the mitochondrial haplogroups of farm and wild-living raccoon dogs in Poland. Mitochondrial DNA 22, 105–110. https://doi.org/10.3109/19401736.2011.624603.

Slaska, B., Zieba, G., Rozempolska-Rucinska, I., Jezewska-Witkowska, G., and Jakubczak, A.

(2010). Evaluation of genetic biodiversity in farm-bred and wild raccoon dogs in Poland. Folia Biol. *58*, 195–199. https://doi.org/10.3409/fb58_3-4.195-199.

Smith, T.J. (2010). Insulin-like growth factor-I regulation of immune function: a potential therapeutic target in autoimmune diseases? Pharmacol. Rev. *62*, 199–236. https://doi.org/10.1124/pr.109.002469.

Song, T., Hao, J., Zhang, R., Tang, M., Li, W., Hui, W., Fu, Q., Wang, C., Xin, S., Zhang, S., et al. (2019). First detection and phylogenetic analysis of porcine circovirus type 2 in raccoon dogs. BMC Vet. Res. *15*, 107. https://doi.org/10.1186/s12917-019-1856-2.

Stratigopoulou, M., van Dam, T.P., and Guikema, J.E.J. (2020). Base excision repair in the immune system: small DNA lesions with big consequences. Front. Immunol. *11*, 1084. https://doi.org/10.3389/fimmu.2020.01084.

Straub, R.H., Cutolo, M., Buttgereit, F., and Pongratz, G. (2010). Energy regulation and neuroendocrine-immune control in chronic inflammatory diseases. J. Intern. Med. *267*, 543–560. https://doi.org/10.1111/j.1365-2796.2010.02218.x.

Stanke, M., Steinkamp, R., Waack, S., and Morgenstern, B. (2004). AUGUSTUS: a web server for gene finding in eukaryotes. Nucleic Acids Res. *32*, W309–W312. https://doi.org/10.1093/nar/gkh379.

Strobl, B., Stoiber, D., Sexl, V., and Mueller, M. (2011). Tyrosine kinase 2 (TYK2) in cytokine signalling and host immunity. Front. Biosci. *16*, 3214–3232. https://doi.org/10.2741/3908.

Suyama, M., Torrents, D., and Bork, P. (2006). PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. Nucleic Acids Res. *34*, W609–W612. https://doi.org/10.1093/nar/gkl315.

Su, D., Zhang, W., Yang, Y., Deng, Y., Ma, Y., Song, H., and Zhang, S. (2008). Mutation screening and association study of the TSSK4 Gene in Chinese infertile men with impaired spermatogenesis. J. Androl. *29*, 374–378. https://doi.org/10.2164/jandrol.107.004598.

Sutor, A., Schwarz, S., and Conraths, F.J. (2014). The biological potential of the raccoon dog (Nyctereutes procyonoides, Gray 1834) as an invasive species in Europe-new risks for disease spread? Acta Theriol. *59*, 49–59. https://doi.org/10.1007/s13364-013-0138-9.

Tack, J., Deloose, E., Ang, D., Scarpellini, E., Vanuytsel, T., Van Oudenhove, L., and Depoortere, I. (2016). Motilin-induced gastric contractions signal hunger in man. Gut *65*, 214–224. https://doi.org/10.1136/gutjnl-2014-308472.

Tarailo-Graovac, M., and Chen, N. (2009). Using RepeatMasker to identify repetitive elements in genomic sequences. Curr. Protoc. Bioinformatics. Chapter 4, Unit 4.10.

Terhorst, J., Kamm, J.A., and Song, Y.S. (2017). Robust and scalable inference of population history from hundreds of unphased whole genomes. Nat. Genet. *49*, 303–309. https://doi.org/10.1038/ng.3748.

Tsuchiya, M.T.N., Dikow, R.B., Koepfli, K.P., Frandsen, P.B., Rockwood, L.L., and Maldonado, J.E. (2021). Whole-genome sequencing of procyonids reveals distinct demographic histories in kinkajou (Potos flavus) and northern raccoon (Procyon lotor). Genome Biol. Evol. *13*, evaa255. https://doi.org/10.1093/gbe/evaa255.

Tuo, L., Xiang, J., Pan, X., Hu, J., Tang, H., Liang, L., Xia, J., Hu, Y., Zhang, W., Huang, A., et al. (2019). PCK1 negatively regulates cell cycle progression and hepatoma cell proliferation via the AMPK/p27(Kip1) axis. J. Exp. Clin. Cancer Res. *38*, 50. https://doi.org/10.1186/s13046-019-1029-y.

Valizadeh, A., Khosravi, A., Zadeh, L.J., and Parizad, E.G. (2015). Role of IL-25 in immunity. J. Clin. Diagn. Res. *9*, OE01–OE04. https://doi.org/10.7860/JCDR/2015/12235.5814.

Vasileva, A., Tiedau, D., Firooznia, A., Müller-Reichert, T., and Jessberger, R. (2009). Tdrd6 is required for spermiogenesis, chromatoid body architecture, and regulation of miRNA expression. Curr. Biol. *19*, 630–639. https://doi.org/10.1016/j.cub.2009.02.047.

Wang, X., Wei, Y., Fu, G., Li, H., Saiyin, H., Lin, G., Wang, Z., Chen, S., and Yu, L. (2015). Tssk4 is essential for maintaining the structural integrity of sperm flagellum. Mol. Hum. Reprod. *21*, 136–145. https://doi.org/10.1093/molehr/gau097.

Ward, O.G., and Wurster-Hill, D.H. (1990). Nytereutes procyonoides. Mamm. Species *358*, 1–5.

Wassarman, P.M. (1999). Mammalian fertilization: molecular aspects of gamete adhesion, exocytosis, and fusion. Cell *96*, 175–183. https://doi.org/10.1016/s0092-8674(00)80558-9.

Wang, K., Xu, R., Snider, A.J., Schrandt, J., Li, Y., Bialkowska, A.B., Li, M., Zhou, J., Hannun, Y.A., Obeid, L.M., et al. (2016). Alkaline ceramidase 3 deficiency aggravates colitis and colitis-associated tumorigenesis in mice by hyperactivating the innate immune system. Cell Death Dis. *7*, e2124. https://doi.org/10.1038/cddis.2016.36.

Wang, Q., Wang, C., Li, N., Liu, X., Ren, W., Wang, Q., and Cao, X. (2018). Condensin Smc4 promotes inflammatory innate immune response by epigenetically enhancing NEMO transcription. J. Autoimmun. *92*, 67–76. https://doi.org/10.1016/j.jaut.2018.05.004.

Warren, W.C., Jasinska, A.J., García-Pérez, R., Svardal, H., Tomlinson, C., Rocchi, M., Archidiacono, N., Capozzi, O., Minx, P., Montague, M.J., et al. (2015). The genome of the vervet (Chlorocebus aethiops sabaeus). Genome Res. *25*, 1921–1933.

Whitlock, B.D., and Leslie, E.M. (2020). Efflux Transporters in Anti-Cancer Drug Resistance: Molecular and Functional Identification and Characterization Of Multidrug Resistance Proteins (MRPs/ABCCs). Drug Efflux Pumps in

Cancer Resistance Pathways: From Molecular Recognition and Characterization to Possible Inhibition Strategies in Chemotherapy (Elsevier), pp. 31–65. https://doi.org/10.1016/b978-0-12-816434-1.00002-4.

Xia, J., Guo, Z., Yang, Z., Han, H., Wang, S., Xu, H., Yang, X., Yang, F., Wu, Q., Xie, W., et al. (2021). Whitefly hijacks a plant detoxification gene that neutralizes plant toxins. Cell *184*, 3588. https://doi.org/10.1016/j.cell.2021.06.010.

Xu, Z., and Wang, H. (2007). LTR_FINDER: an efficient tool for the prediction of full-length LTR retrotransposons. Nucleic Acids Res. *35*, W265–W268. https://doi.org/10.1093/nar/gkm286.

Xu, J., Wang, X.-y., and Guo, W.-z. (2015). The cytochrome P450 superfamily: key players in plant development and defense. J. Integr. Agric. *14*, 1673–1686. https://doi.org/10.1016/s2095-3119(14)60980-1.

Yang, H., Beutler, B., and Zhang, D. (2021a). Emerging roles of spliceosome in cancer and immunity. Protein Cell *13*, 559–579. https://doi.org/10.1007/s13238-021-00856-5.

Yang, S., He, Y., Chen, X., Kalim, U., Wang, Y., Yang, S., Qi, H., Cheng, H., Lu, X., Wang, X., et al. (2021b). Viral metagenomics reveals diverse viruses in the feces samples of raccoon dogs. Front. Vet. Sci. *8*, 693564. https://doi.org/10.3389/fvets.2021.693564.

Yang, L., Liu, C.C., Zheng, H., Kanekiyo, T., Atagi, Y., Jia, L., Wang, D., N'Songo, A., Can, D., Xu, H., et al. (2016). LRP1 modulates the microglial immune response via regulation of JNK and NF-kappaB signaling pathways. J. Neuroinflammation *13*, 304. https://doi.org/10.1186/s12974-016-0772-7.

Yang, Z. (2007). Paml 4: phylogenetic analysis by maximum likelihood. Mol. Biol. Evol. *24*, 1586–1591. https://doi.org/10.1093/molbev/msm088.

Yoshinaka, T., Kosako, H., Yoshizumi, T., Furukawa, R., Hirano, Y., Kuge, O., Tamada, T., and Koshiba, T. (2019). Structural basis of mitochondrial scaffolds by prohibitin complexes: insight into a role of the coiled-coil region. iScience *19*, 1065–1078. https://doi.org/10.1016/j.isci.2019.08.056.

Yao, X., Zhang, F., Qiao, Z., Yu, H., Sun, S., Li, X., Zhang, J., and Jiang, X. (2020). Toxicity of thifluzamide in earthworm (Eisenia fetida). Ecotoxicol. Environ. Saf. *188*, 109880. https://doi.org/10.1016/j.ecoenv.2019.109880.

Yoshikawa, T., Nakamura, T., and Yanai, K. (2019). Histamine N-methyltransferase in the brain. Int. J. Mol. Sci. *20*, E737. https://doi.org/10.3390/ijms20030737.

Zhao, S., Wang, B., Zhong, Z., Liu, T., Liang, T., and Zhan, J. (2020). Contributions of enzymes and gut microbes to biotransformation of perfluorooctane sulfonamide in earthworms (Eisenia fetida). Chemosphere *238*, 124619. https://doi.org/10.1016/j.chemosphere.2019.124619.

Zhan, K., Liu, R., Tong, H., Gao, S., Yang, G., Hossain, A., Li, T., and He, W. (2020). Fetuin B overexpression suppresses proliferation,

migration, and invasion in prostate cancer by inhibiting the PI3K/AKT signaling pathway. Biomed. Pharmacother. *131*, 110689. https://doi.org/10.1016/j.biopha.2020.110689.

Zhang, G., Cowled, C., Shi, Z., Huang, Z., Bishop-Lilly, K.A., Fang, X., Wynne, J.W., Xiong, Z., Baker, M.L., Zhao, W., et al. (2013). Comparative analysis of bat genomes provides insight into the evolution of flight and immunity. Science *339*, 456–460.

Zhang, Y.G., Wu, S., Xia, Y., Chen, D., Petrof, E.O., Claud, E.C., Hsu, W., and Sun, J. (2012). Axin1 prevents Salmonella invasiveness and inflammatory response in intestinal epithelial cells. PLoS One *7*, e34942. https://doi.org/10.1371/journal.pone.0034942.

Zhao, W., Ma, N., Wang, S., Mo, Y., Zhang, Z., Huang, G., Midorikawa, K., Hiraku, Y., Oikawa, S., Murata, M., and Takeuchi, K. (2017). RERG suppresses cell proliferation, migration and angiogenesis through ERK/NF-kappaB signaling pathway in nasopharyngeal carcinoma. J. Exp. Clin. Cancer Res. *36*, 88. https://doi.org/10.1186/s13046-017-0554-9.

Zhu, M., Foreman, D.P., O'Brien, S.A., Jin, Y., and Zhang, W. (2018). Phospholipase D in TCR-mediated signaling and T cell activation. J. Immunol. *200*, 2165–2173. https://doi.org/10.4049/jimmunol.1701291.

Zhou, H., Li, J., Su, H., Li, J., Lydic, T.A., Young, M.E., and Chen, W. (2022). BSCL2/Seipin deficiency in hearts causes cardiac energy deficit and dysfunction via inducing excessive lipid

catabolism. Clin. Transl. Med. *12*, e736. https://doi.org/10.1002/ctm2.736.

Zhou, H., Xu, C., Lee, H., Yoon, Y., and Chen, W. (2020). Berardinelli–Seip congenital lipodystrophy 2/SEIPIN determines brown adipose tissue maintenance and thermogenic programing. Mol. Metab. *36*, 100971. https://doi.org/10.1016/j.molmet.2020.02.014.

Zindy, F., den Besten, W., Chen, B., Rehg, J.E., Latres, E., Barbacid, M., Pollard, J.W., Sherr, C.J., Cohen, P.E., and Roussel, M.F. (2001). Control of spermatogenesis in mice by the cyclin D-dependent kinase inhibitors p18(Ink4c) and p19(Ink4d). Mol. Cell Biol. *21*, 3244–3255. https://doi.org/10.1128/MCB.21.9.3244-3255.2001.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Chemicals, peptides, and recombinant proteins		
TRIzol reagent	Invitrogen, USA	Cat#15596 - 026
Critical commercial assays		
Monarch HMW DNA Extraction Kit	NEB, Ipswich, England	Cat#T3060L
SMRTbell Template Prep Kit 1.0	Pacific Biosciences, CA, USA	Cat#100-259-100
Qiagen Blood & Cell Culture DNA Mini Kit	Qiagen, United States	Cat#13323
Deposited data		
Nyctereutes procyonoides reference genome	This paper	CNSA: CNP0002053
Bos Taurus reference genome	Ensembl	ARS-UCD1.2
Ursus thibetanus thibetanus reference genome	Ensembl	ASM966005v1
Acinonyx jubatus reference genome	NCBI	GCF_003709585.1_Aci_jub_2
Ursus arctos reference genome	NCBI	GCF_003584765.1_ASM358476v1
Sus scrofa reference genome	Ensembl	Sscrofa11.1
Panthera leo reference genome	Ensembl	PanLeo1.0
Panthera pardus reference genome	Ensembl	PanPar1.0
Vulpes lagopus reference genome	NCBI	GCF_018345385.1_ASM1834538v1
Ursus maritimus reference genome	Ensembl	UrsMar_1.0
Canis lupus dingo reference genome	Ensembl	ASM325472v1
Panthera tigris altaica reference genome	Ensembl	PanTig1.0
Vulpes Vulpes reference genome	Ensembl	VulVul2.2
Homo sapiens reference genome	Ensembl	GRCh38
Felis catus reference genome	Ensembl	Felis_catus_9.0
Equus caballus reference genome	Ensembl	EquCab3.0
Canis lupus familiaris reference genome	Ensembl	CanFam3.1
Oryctolagus cuniculus reference genome	Ensembl	OryCun2.0
Software and algorithms		
Canu (v2.0)	(Koren et al., 2017)	https://github.com/marbl/canu
NextPolish (v1.3.1)	(Hu et al., 2020)	https://github.com/Nextomics/NextPolish
Purge_dups (v1.2.5)	(Guan et al., 2020)	https://github.com/dfguan/purge_dups
LRScaf (v1.1.8)	(Qin et al., 2019)	https://github.com/shingocat/lrscaf
BWA (v0.7.17)	(Li and Durbin, 2010)	http://bio-bwa.sourceforge.net/
Juicer (v1.5)	(Durand et al., 2016)	https://github.com/aidenlab/juicer
3d-DNA (v190716)	(Dudchenko et al., 2017)	https://github.com/aidenlab/3d-dna
BUSCO (v5.2.2)	(Manni et al., 2021)	https://busco.ezlab.org/
LTR finder (v1.0.6)	(Xu and Wang, 2007)	https://github.com/xzhub/LTR_Finder
MITE-hunter (v4.07)	(Han and Wessler, 2010)	https://github.com/jburnette/MITE-Hunter
RepeatModeler2 (v2.0.1)	(Flynn et al., 2020)	http://www.repeatmasker.org/RepeatModeler/
RepeatMasker (v4.0.5)	(Tarailo-Graovac and Chen, 2009)	https://www.repeatmasker.org/
Tandem Repeats Finder (v4.07)	(Benson, 1999)	https://tandem.bu.edu/trf/trf.html
GlimmerHMM (v3.0.1)	(Majoros et al., 2004)	https://ccb.jhu.edu/software/glimmerhmm/
Augustus (v3.0.3)	(Stanke et al., 2004)	https://bioinf.uni-greifswald.de/augustus/
SNAP (v11/29/2013)	(Korf, 2004)	https://github.com/KorfLab/SNAP

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Trimmomatic (v0.30)	(Bolger et al., 2014)	http://www.usadellab.org/cms/?page=trimmomatic
Trinity (v2.13.2)	(Haas et al., 2013)	https://github.com/trinityrnaseq/trinityrnaseq
PASA (v2.0.2)	(Haas et al., 2008)	https://github.com/PASApipeline/PASApipeline
Blastall (v2.2.26)	(Mount, 2007)	http://gensoft.pasteur.fr/docs/blast/2.2.26/
GeneWise (v2.4.1)	(Birney et al., 2004)	https://www.ebi.ac.uk/Tools/psa/genewise/
MAKER pipeline (v3.01.03)	(Campbell et al., 2014)	https://www.yandell-lab.org/software/maker.html
tRNAscan-SE (v2.0.9)	(Lowe and Eddy, 1997)	https://github.com/UCSC-LoweLab/tRNAscan-SE
INFERNAL (v1.1.1)	(Nawrocki et al., 2009)	http://eddylab.org/infernal/
InterProScan (v5.52-86.0)	(Jones et al., 2014)	https://github.com/ebi-pf-team/interproscan
MAFFT (v.7.310)	(Katoh and Standley, 2013)	https://mafft.cbrc.jp/alignment/software/
PAL2NAL (v14)	(Suyama et al., 2006)	http://www.bork.embl.de/pal2nal/
Trimal (v1.4.1)	(Capella-Gutierrez et al., 2009)	http://trimal.cgenomics.org/
IQTREE (v1.6.12)	(Nguyen et al., 2015)	http://www.iqtree.org/
Treefam (v1.4)	(Li et al., 2006)	http://www.treefam.org/
CAFE (v4.2.1)	(De Bie et al., 2006)	https://hahnlab.github.io/CAFE/src_docs/html/index.html
PAML (v4.8)	(Yang, 2007)	http://evomics.org/resources/software/molecular-evolution-software/paml/
Picard (v2.1.1)	N/A	https://github.com/broadinstitute/picard
Sentieon (v202010.01)	(Freed et al., 2017)	https://support.sentieon.com
VCFtools (v4.1)	(Danecek et al., 2011)	http://vcftools.sourceforge.net/
PSMC (v0.6.5)	(Li and Durbin, 2011)	https://github.com/lh3/psmc
MSMC2 (v2.1.1)	(Schiffels and Durbin, 2014)	https://github.com/stschiff/msmc2
SMC++ (v1.15.4)	(Terhorst et al., 2017)	https://github.com/popgenmethods/smcpp
Other		
Raw and analyzed data	This paper	CNSA: CNP0002053

RESOURCE AVAILABILITY

Lead contact
Further information and requests for reagents should be directed to and will be fulfilled by the Lead Contact, Prof. Zhijun Hou (houzhijundb@163.com).

Materials availability
This study did not generate new unique reagents.

Data and code availability
The data that support the findings in this study have been deposited into CNGB Sequence Archive (CNSA) (Guo et al., 2020) of China National GeneBank DataBase (CNGBdb) (Chen et al., 2020) with accession number CNP0002053.

All original codes are included in the supplemental information.

Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

An adult male raccoon dog (one-year-old) who died of natural causes was collected from Harbin, Heilongjiang, China for the genome assembly. This individual is captive born and from the subspecies of

N. procyonoides ussuriensis. We immediately stored tissue samples in liquid nitrogen after quick dissection to perform high-quality DNA and RNA isolation. Samples from 5 organs, including the heart, lung, spleen, liver, and kidney were used for RNA sequencing. The muscle sample was used for Pacific Bioscience (PacBio) sequencing and genome survey. The liver sample was used for Hi-C sequencing. We also collected blood samples of another 38 individuals from Harbin Hualong Fur Farm for whole genome resequencing. Sample collection, experiments and research design in this study were all approved by the Institutional Review Board of BGI (BGI-IRB E22001). We strictly followed the guidelines from BGI-IRB for all procedures conducted in this study.

METHOD DETAILS

Nucleic acid extraction, DNA library preparation and sequencing

High-molecular-weight genomic DNA was isolated from the muscle sample using the Monarch HMW DNA Extraction Kit (NEB, Ipswich, England). The SMRTbell Template Prep Kit 1.0 (Pacific Biosciences, CA, USA) was used for SMRTbell library preparation based on DNA fragments with an average size of 20 kb. To obtain the 20 kb DNA fragments, we first extracted the total genomic DNA and detected the completeness of the DNA by agarose gel electrophoresis, and then the genomic DNA was randomly broken into fragments of about 20 kp length by Covaris ultrasonic crusher. The DNA library was then subjected to the PacBio Sequel II platform for long-read sequencing. For Hi-C library construction, we first performed the crosslinking with formaldehyde for the liver sample, and then one Hi-C library was constructed by using the dpnII restriction endonuclease. Genomic DNAs from the 38 blood samples for resequencing and the muscle sample for genome survey were isolated with Qiagen Blood & Cell Culture DNA Mini Kit (Qiagen, United States). The 5 organs were used for RNA isolated with TRlzol reagent (Invitrogen, USA) following the manufacturer's guidelines. The extracted RNA was then fragmented into 200–400 bp by heat digestion with a divalent metal cation by adjusting the time of incubation, and then reverse-transcribed to cDNA for library preparation. 45 short insert size libraries (5 for cDNA, 40 for genomic DNA) were finally constructed according to the manufacturer's instruction of the MGI platform (MGI, Shenzhen, China). All 45 libraries were finally sequenced on the DNBSEQ-T1 platform with 100-bp paired-end sequencing strategy.

Genome assembly, annotation, and assessment

To estimate the genome size of the raccoon dog, we calculated the distribution of the frequency of single nucleotide depth across the genome with 82.21 Gb short reads. The genome size was estimated by dividing the peak depth by the total sequencing base number (Pflug et al., 2020). The primary genome was assembled by Canu (Koren et al., 2017) (v2.0) with PacBio long reads. This primary assembly was polished with both DNBSEQ short reads and PacBio long reads by three rounds using NextPolish (Hu et al., 2020) (v1.3.1) to correct sequence errors. We used purge_dups (Guan et al., 2020) (v1.2.5) to remove haplotypic duplication through sequence similarity and read depth. LRScaf (Qin et al., 2019) (v1.1.8) was then used for the linking of contigs into longer scaffolds. To obtain the chromosome-scale genome, we mapped Hi-C reads to the primary genome by Burrows-Wheeler Aligner (BWA, v0.7.17) (Li and Durbin, 2010), and then used Juicer (Durand et al., 2016) (v1.5) for Hi-C data quality control, finally used 3d-DNA pipeline (Dudchenko et al., 2017) (v190716) to concatenate the scaffolds to the chromosome-scale genome. BUSCO (Manni et al., 2021) (v5.2.2) analysis was performed to evaluate the completeness of the final chromosome-scale genome with both laurasiatheria_odb10 and mammalia_odb10 data set.

A combination of *de novo* and homolog-based methods was used for repeat annotation. We firstly used the LTR finder (Xu and Wang, 2007) (v1.0.6), MITE-hunter (Han and Wessler, 2010) (v4.07), and RepeatModeler2 (Flynn et al., 2020) (v2.0.1) to annotate *de novo* repeats. These identified repeats were then merged into the RepBase as known elements. Finally, we performed RepeatMasker (Tarailo-Graovac and Chen, 2009) (v4.0.5) to classify and identify transposable elements with a conserved BLASTN search against the prepared RepBase library. Tandem repeats were also annotated by Tandem Repeats Finder (Benson, 1999) (v4.07) software. We then masked all repeat elements for gene annotation.

Protein-coding genes were predicted using *de novo*, homology-based and transcript mapping approaches. GlimmerHMM (Majoros et al., 2004) (v3.0.1), Augustus (Stanke et al., 2004) (v3.0.3), and SNAP (Korf, 2004) (v11/29/2013) were used for *de novo* gene prediction. For RNA-seq based prediction, we first filtered RNA-seq data by Trimmomatic (Bolger et al., 2014) (v0.30). Transcripts were assembled using Trinity (Haas et al., 2013) (v2.13.2) based on clean RNA-seq data. Program to Assemble Spliced Alignments (PASA) (Haas et al., 2008) (v2.0.2) was finally used to align transcript against the raccoon dog genome to obtain

gene structures. Homology-based prediction was carried out by mapping protein sequences of UniProt database (release-2020_05), *Homo sapiens*, *Mus musculus*, *Canis lupus familiaris*, *Nyctereutes procyonoides* (Chueca et al., 2021), and *Vulpes vulpes* to the raccoon dog genome using the Blastall (Mount, 2007) (v2.2.26) with an E-value cut-off of 1e-5. Then the GeneWise (Birney et al., 2004) (v2.4.1) was used to predict gene models by analyzing alignment results. The final gene set representing RNA-seq, homology, and *de novo* predicted genes was generated by performing MAKER pipeline (Campbell et al., 2014) (v3.01.03). Blastall (Mount, 2007) (v2.2.26) was used to identify ribosomal RNA (rRNA) sequences by aligning human rRNA sequences to the raccoon dog genome. Transfer RNA (tRNA) was predicted by tRNAscan-SE (Lowe and Eddy, 1997) (v2.0.9). Both microRNA (miRNA) and small nuclear RNA (snRAN) were predicted using INFERNAL v1.1.1 (Nawrocki et al., 2009) by aligning the raccoon dog genome sequences to the Rfam database. Functional annotation was performed by performing BLAST search against the SwissProt, TrEMBL and Kyoto Encyclopedia of Genes and Genomes (KEGG) database with an E-value cut-off of 1e-5. InterProScan (Jones et al., 2014) (v5.52-86.0) was used to predict motifs and domains, as well as Gene ontology (GO) terms.

Phylogeny reconstruction with genome-wide single-copy genes

Homologous genes of 18 species were determined by performing all-to-all BLASTP analysis for proteins from each species with the parameter of "-evalue 1e-5". We identified 6,936 single-copy genes shared by these 18 species. A phylogenetic tree was then constructed based on these single copy genes with the following procedures: 1) We used MAFFT (Katoh and Standley, 2013) (v.7.310) to do multiple amino acid sequence alignments for each single-copy gene orthogroup; 2) The aligned sequences of the amino acid were converted to aligned DNA sequences using PAL2NAL (Suyama et al., 2006) (v14); 3) Gaps were removed using the trimal (Capella-Gutierrez et al., 2009) (v1.4.1) software; 4) We calculated the best-fit substitution model using ModelFinder (Kalyaanamoorthy et al., 2017); 5) We constructed a maximum-likelihood (Warren et al.) phylogenetic tree by IQTREE (Nguyen et al., 2015) (v1.6.12) with concatenated super-genes. For the GST gene family, we also used the same methods to generate the ML phylogenetic tree with only Canidae species.

Gene family expansion and contraction

We clustered all annotated genes to identify gene families with hierarchically clustering on a sparse graph by the software Treefam (Li et al., 2006) (v1.4). We identified 20,699 gene families in the above-mentioned 18 genomes, and 16,403 were in the raccoon dog genome. We then used CAFE (De Bie et al., 2006) (v4.2.1) to detect contracted and expanded gene families. The random birth and death model was used to estimate the size of gene families at ancestral nodes. A family-wise $p < 0.05$ was set as the cut-off value. We carried out the KEGG and GO enrichment analysis on expanded gene families with all annotated genes as the background. To avoid an inaccurate chi-square test, we performed Fisher's exact test with an expected gene count of below five. We finally adjusted p values using the Benjamini-Hochberg method (Benjamini and Yekutieli, 2005) with multiple tests by specifying a false discovery rate with q-value less than 0.05.

Positively selected genes (PSGs)

The phylogenetic tree and 6,936 single-copy genes identified in the section of "phylogeny reconstruction with genome-wide single-copy genes" were subjected to the selection analysis. PSGs were identified in the raccoon dog genome by comparing these single copy orthologs with the above-mentioned 17 species. We calculated the non-synonymous to synonymous substitutions ratio (dN/dS) under the branch-site model in the CodeML of PAML (Yang, 2007) (v4.8) software, with raccoon dog as the foreground. We performed the likelihood ratio test (LRT) for Lnl values of each model pair and obtained p values with chi-square test. The positively selected sites were identified with a false discovery rate corrected p value less than 0.05.

Genome-wide variants calling and quality control

We mapped the whole genome resequencing data of 38 individuals to the raccoon dog genome by using the BWA (Li, 2013) *mem* method with default parameters. We then used the Picard package (v2.1.1) for sorting, reordering and deduplication of alignment files for variants calling. Sentieon (Freed et al., 2017) (v202010.01) DNAseq Haplotyper was used to call variants for each sample independently to generate the genomic Variant Call Format (gVCF) files. The following joint variant calling was carried out by Sentieon DNAseq GVCFtyper with the 38 gVCF files to finally create a common VCF file. To facilitate downstream analysis, we removed InDels and multi-allelic variants and then performed hard filtering with "QD < 2.0 ||

FS > 60.0 || MQ < 40.0 || MQRankSum < −12.5 || ReadPosRankSum < −8.0 –filter-name snp_filter"(DePristo et al., 2011).

Genomic diversity and population demography

Autosomal variants were selected to calculate the genome-wide *H* and genetic diversity (π) using VCFtools (Danecek et al., 2011) (v4.1). The population level π was calculated with 500-kb sliding window by setting the parameter "*-window-pi 500000*". For inference of population history, we first used the Pairwise Sequentially Markovian Coalescent (PSMC) (Li and Durbin, 2011) (v0.6.5) to infer the fluctuation of effective population size over its evolutionary history under the pattern of "4 + 25×2 + 4+6". To show the robustness of this estimation, we performed 100 bootstrap replicates for the same individual. We visualized the PSMC result by scaling the time to the real years by using the mutation rate (μ) of 1.0×10^{-8} substitution per site per generation and a generation time (g) of three years (Freedman et al., 2014). PSMC has been shown to be inappropriate for inferring more recent population history (e.g. more recent than 20 kya for human populations) (Li and Durbin, 2011), Therefore, we performed MSMC2 (Schiffels and Durbin, 2014) (v2.1.1) to infer the more recent population history (especially the time period from 1 kya to 10 kya) of the raccoon dog with four randomly selected individuals. Considering switch errors induced by the phasing process which will further bias the MSMC2 inference, we also performed the SMC++ (Terhorst et al., 2017) (v1.15.4) analysis based on the population variations from all 38 individuals to validate the result from MSMC2. For both MSMC2 and SMC++, we used the same generation interval and mutation rate used in the PSMC to show the results.

QUANTIFICATION AND STATISTICAL ANALYSIS

Quantification and statistical analysis used in the genome assembly and comparative genome analysis can be found in the method details.