Long-Read Genome Sequencing Provides Molecular Insights into Scavenging and Societal Complexity in Spotted Hyena *Crocuta crocuta*

Yong Shao,*[,†,1] Xiao-Bo Wang,[†,2,3] Mei-Ling Zhang,[†,4] Yan Liu,[†,5] Sheng Wang,[1] Bao-Lin Zhang,[1] Min-Min Yang,[1] Ming-Hai Yang,[5] Ting Jia,[5] Tian-Chun Pu,[5] Yan Lu,[5] He Liu,[5] Zhe Xu,[5] Bo Li,[5] Ning Liu,[5] Violet Magoma Onsongo,[1] Dong-Dong Wu,[1] Cheng-Lin Zhang,*[5] Jue Ruan,*[2] and Yan Li*[6]

[1]State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, China

[2]Agricultural Genomics Institute, Chinese Academy of Agricultural Sciences, Shenzhen, China

[3]State Key Laboratory for Conservation and Utilization of Subtropical Agro-Bioresources, Guangxi University, Nanning, China

[4]Department of Acute Infectious Diseases Control and Prevention, Yunnan Provincial Centre for Disease Control and Prevention, Kunming, China

[5]Beijing Zoo, Beijing, China

[6]State Key Laboratory for Conservation and Utilization of Bio-resource in Yunnan, School of Life Sciences, Yunnan University, Kunming, China

[†]These authors contributed equally to this work.

***Corresponding authors:** E-mails: shaoyong@mail.kiz.ac.cn; zhch6465@263.net; ruanjue@caas.cn; liyan0910@ynu.edu.cn.

Associate editor: Jian Lu

Abstract

The spotted hyena (*Crocuta crocuta*) is a large and unique terrestrial carnivore. It is a particularly fascinating species due to its distinct phenotypic traits, especially its complex social structure and scavenging lifestyle, with associated high dietary exposure to microbial pathogens. However, the underlying molecular mechanisms related to these phenotypes remain elusive. Here, we sequenced and assembled a high-quality long-read genome of the spotted hyena, with a contig N50 length of ~13.75 Mb. Based on comparative genomics, immunoglobulin family members (e.g., *IGKV4-1*) showed significant adaptive duplications in the spotted hyena and striped hyena. Furthermore, immune-related genes (e.g., *CD8A*, *LAG3*, and *TLR3*) experienced species-specific positive selection in the spotted hyena lineage. These results suggest that immune tolerance between the spotted hyena and closely related striped hyena has undergone adaptive divergence to cope with prolonged dietary exposure to microbial pathogens from scavenging. Furthermore, we provided the potential genetic insights underlying social complexity, hinting at social behavior and cognition. Specifically, the RECNE-associated genes (e.g., *UGP2* and *ACTR2*) in the spotted hyena genome are involved in regulation of social communication. Taken together, our genomic analyses provide molecular insights into the scavenging lifestyle and societal complexity of spotted hyenas.

Key words: **spotted hyena, long-read genome, social complexity, immune tolerance.**

Article

Background

Although Hyaenidae comprised only four extant species (aardwolf, spotted hyena, striped hyena, and brown hyena), representing the last remnants of this family, it possesses a remarkable ecological diversity (Werdelin and Solounias 1991; Koepfli et al. 2006; Watts and Holekamp 2007). For example, the modern hyenas occupy diverse habitats (i.e., deserts, montane forests, savannas) and distinct dietary niches, including specialized insectivores, predators, and, most intriguingly, scavengers (Gusset and Burgener 2005; Watts and Holekamp 2007). Scavengers serve as important carrion consumers in ecosystem food chains, impacting the spread of diseases, such as anthrax, and reducing fly populations and fetid odors by quickly devouring carcasses (Koenig 2006; Benbow and Tarone 2015). In contrast to the aardwolf (*Proteles cristata*), a termite-feeding specialist, striped hyenas (*Hyaena hyaena*) and brown hyenas (*Parahyaena brunnea*) are scavengers that forage primarily on the carcasses of large vertebrates (Watts and Holekamp 2007), wheras spotted hyenas (*Crocuta crocuta*) occupy both a scavenging niche and predatory position at the top of the food chain (Gade 2006; Hayward 2006; Koepfli et al. 2006).

Compared with other vertebrates, scavengers possess a strong immune defense system because of their feeding

lifestyles, suggesting that they have evolved specific adaptive mechanisms to prevent infection from microbes in their food (de la Lastra and de la Fuente 2007; Chung et al. 2015). Although a previous study has examined the immune system of scavenging vultures (Chung et al. 2015), little is known about the genetic variations and molecular mechanisms involved in the regulation of immune responses in hyaenids, which represent a rare branch of mammalian scavengers. To date, most studies on hyaenid species have focused on their evolutionary history, conservation biology, and social structure (Holekamp et al. 2012; Westbury et al. 2018; 2020).

In addition to immune system evolution in hyaenids, studies have indicated that spotted hyenas possess a more complex social structure than their close relatives (i.e., striped hyena, brown hyena, and aardwolf) and other carnivores. This complexity is comparable with that of primates (i.e., baboons and macaques) (Holekamp et al. 2012). The social complexity of spotted hyenas is embodied by communication signals (e.g., a rich repertoire of olfactory, acoustic, and visual signals), large group sizes (i.e., 9–70 individuals in each clan), matriarchal and hierarchical organization, and high social intelligence (Holekamp et al. 2007; Arsznov et al. 2010; Holekamp et al. 2012). In clans, female spotted hyenas are more socially dominant than males and more aggressive in aspects of behavior (Holekamp 2006). Morphologically, females are ∼10% larger than males in body size and have masculinized external genitalia (Holekamp 2006), the formation of which may involve hormone system balance of androgens and estrogens (Cunha et al. 2014). At present, however, the genetic and molecular factors involved in social structure complexity in spotted hyenas remain unclear.

Here, in contrast to a previous short-read assembly (Yang et al. 2020), we assembled a high-quality long-read genome for the spotted hyena and used comparative genomics to gain molecular insights into social complexity and immune tolerance in the spotted hyena. Furthermore, our improved genome assembly offers insight into the potential molecular mechanisms underlying the unique phenotypes of this species.

Results

Sequencing, Assembly, and Annotation of Spotted Hyena Genome

Utilizing ∼64.49× coverage data (∼155.57 Gb) from a single-molecule real-time genome sequencing platform (PacBio Sequel) (supplementary table 1, Supplementary Material online), we obtained a high-quality de novo genome of the spotted hyena with a contig N50 of ∼13.75 Mb using FALCON (Chin et al. 2016) (fig. 1A and supplementary table 2, Supplementary Material online). The genome assembly was ∼2.39 Gb in size, approaching the 2.41 Gb obtained from K-mer analysis (supplementary fig. 1 and supplementary table 3, Supplementary Material online).

We mapped ∼50× Illumina short reads representing ∼123 Gb to this assembly, with a relatively high mapping rate (∼99.3%) and homozygous single nucleotide polymorphism (SNP) ratio approximating 4.37×10^{-6}. Conserved synteny of genomes was observed between the spotted hyena and cat (http://asia.ensembl.org/) (fig. 1B). These analyses supported the high base accuracy of our genome assembly. In contrast to the previous short-read assembly of the spotted hyena (Yang et al. 2020), our assembly showed improved quality (fig. 1C and supplementary fig. 2, Supplementary Material online). Additionally, Core Eukaryotic Genes Mapping Approach (CEGMA) and Benchmarking Universal Single-Copy Orthologs (BUSCO) were applied to evaluate the completeness of our assembly. Both analyses (96.4% in CEGMA and 95.4% in BUSCO) suggested high-quality assembly and genomic completeness (supplementary tables 4 and 5, Supplementary Material online). Based on de novo (Augustus, Genscan, SNAP, and GlimmerHMM) and homology (cat, dog, tiger, panda, and human) annotations, we retrieved 21,468 protein-coding genes (supplementary fig. 3 and supplementary table 6, Supplementary Material online).

Evolution and Functional Roles of Transposable Elements

In animals, transposable elements (TEs) play important roles in shaping genomic and phenotypic evolution (Chuong et al. 2017). Here, we profiled evolutionary characteristics of TEs in the spotted hyena genome. Long interspersed nuclear element-1 (LINE-1) accounted for the highest proportion (∼22.12%) of the genome, closely approaching that of cat (∼20.42%) and dog (∼20.22%) (fig. 2A and supplementary fig. 4 and supplementary table 7, Supplementary Material online). Intriguingly, we found that the top-three TE subclass LTR/ERVL-MaLR (∼2.31%) in the spotted hyena genome was at least 1.5 times as long as those of closely related non-Hyaenidae species, that is, cat (∼1.43%) and dog (∼1.51%). The contents of another TE subclass (top five) (fig. 2A), DNA/hAT-Charlie, accounted for ∼1.41% of the spotted hyena genome, which was at least 1.8 times higher than that of the cat (∼0.77%) and dog (∼0.78%) (fig. 2A).

For extant Hyaenidae species, genomes of the aardwolf, spotted hyena (this study), and striped hyena have been released (Allio et al. 2021; Westbury et al. 2018). To compare potential expansion of the two types of TEs, we downloaded the genomes of the aardwolf and striped hyena and analyzed TE content across diverse species. In addition to integrating genomes of the cat and dog, we also analyzed LTR/ERVL-MaLR and DNA/hAT-Charlie in the cow as a contrasting background species. We identified a potential pattern, whereby Hyaenidae species exhibited significantly higher expanded TEs than the closely related non-Hyaenidae species (i.e., cat, dog, and cow), for example, LTR/ERVL-MaLR ($P = 0.0296$, unpaired one-tailed t-test) and DNA/hAT-Charlie ($P = 0.008$, unpaired one-tailed t-test) (fig. 2B). We concluded that the two types of TEs in the Hyaenidae species genomes experienced a burst insertion in ∼11–13 Ma (fig. 2C). The burst dating of LTR/ERVL-MaLR and DNA/hAT-Charlie approximated the divergence time (∼13 Ma) of the most recent common ancestor of Hyaenidae species (Westbury et al. 2021), suggesting that the two types of TEs underwent expansion in Hyaenidae.

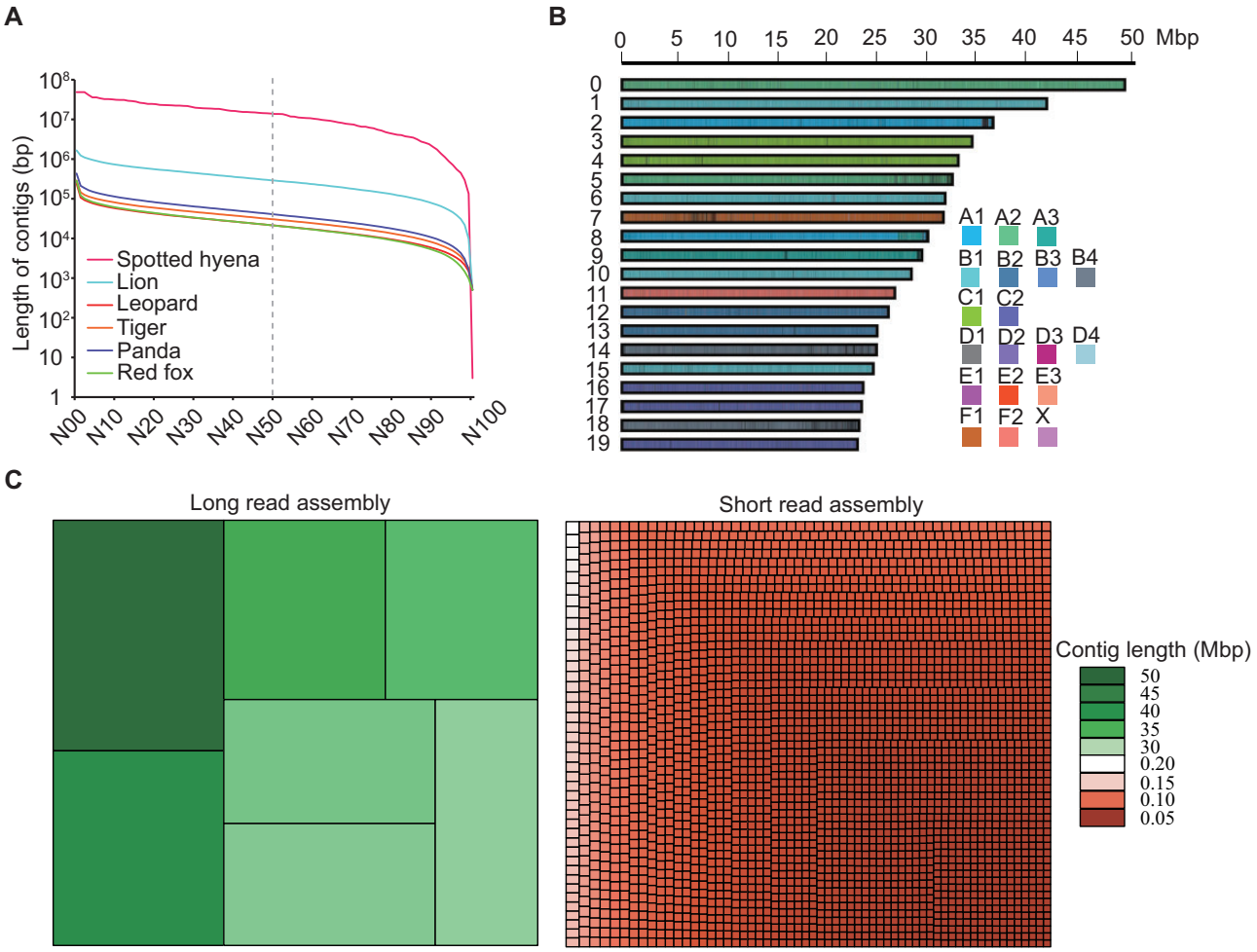

FIG. 1. Contiguity and completeness of spotted hyena genome assembly. (*A*) Assembly *N* (*x*) plot shows percentage of genome (*x*-axis) containing contigs of at least *x* kb (*y*-axis). The contig N50 was highlighted by the gray dotted line. (*B*) Syntenic relationship between spotted hyena and cat genomes. Bars represent top 20 longest contigs of spotted hyena genome from top to bottom. Bar colors represent different cat chromosomes (legend on right). (*C*) Treemaps of long-read and short-read spotted hyena assemblies, respectively. Rectangles represent longest contigs, accounting for ~255 Mb (~10%) of assembly.

Expansion of LTR/ERVL-MaLR and DNA/hAT-Charlie may hold important biological implications for shaping phenotypic evolution in Hyaenidae species. To test this, we downloaded the expression matrices of 30 normal human tissues from the Genotype-Tissue Expression project (GTEx, https://commonfund.nih.gov/GTEx/) to obtain tissue-specific expressed genes in each tissue according to *t*-statistics (supplementary fig. 5, Supplementary Material online) (Finucane et al. 2018). Based on in silico analysis (Fang et al. 2020), tissue-specific expressed genes are highly correlated across diverse species (e.g., humans and cattle). Therefore, we utilized tissue-specific expressed genes in humans to determine whether there was significant overlap between genes with LTR/ERVL-MaLR and DNA/hAT-Charlie insertions in each Hyaenidae species and tissue-specific expressed genes across diverse tissues. We used the OrthoMCL algorithm (Li et al. 2003) to identify pairwise single-copy orthologous genes between each Hyaenidae species and humans. Interestingly, genes with LTR/ERVL-MaLR insertions in the spotted hyena showed a significant overlap with tissue-specific expressed genes from six tissues, including the bladder ($P = 0.02$, Fisher's exact test), cervix uteri ($P = 0.02$, Fisher's exact test), nerve ($P = 0.01$, Fisher's exact test), pituitary ($P = 0.02$, Fisher's exact test), testis ($P = 4.00E\text{-}06$, Fisher's exact test), and thyroid ($P = 3.10e\text{-}03$, Fisher's exact test). Furthermore, genes with DNA/hAT-Charlie insertions in the spotted hyena showed a significant overlap with tissue-specific expressed genes from three tissues, including the bladder ($P = 0.01$, Fisher's exact test), blood vessel ($P = 0.03$, Fisher's exact test), and testis ($P = 2.90e\text{-}03$, Fisher's exact test) (fig. 2*D*). We also observed a similar enrichment pattern overlap in the two other Hyaenidae species (e.g., striped hyena and aardwolf) (fig. 2*D*), with tissues showing significant TE enrichment mainly involved in the urinary system, reproduction, and hormone regulation (fig. 2*D*). We ranked the top 5% of genes based on TE insertion length in each gene locus and overlapped the tissue-specific expressed genes for each Hyaenidae species, and still found a similar LTR/ERVL-MaLR enrichment pattern in the pituitary across Hyaenidae (fig. 2*D*).

Our analysis suggests that the expanded TEs (e.g., LTR/ERVL-MaLR and DNA/hAT-Charlie) in Hyaenidae species

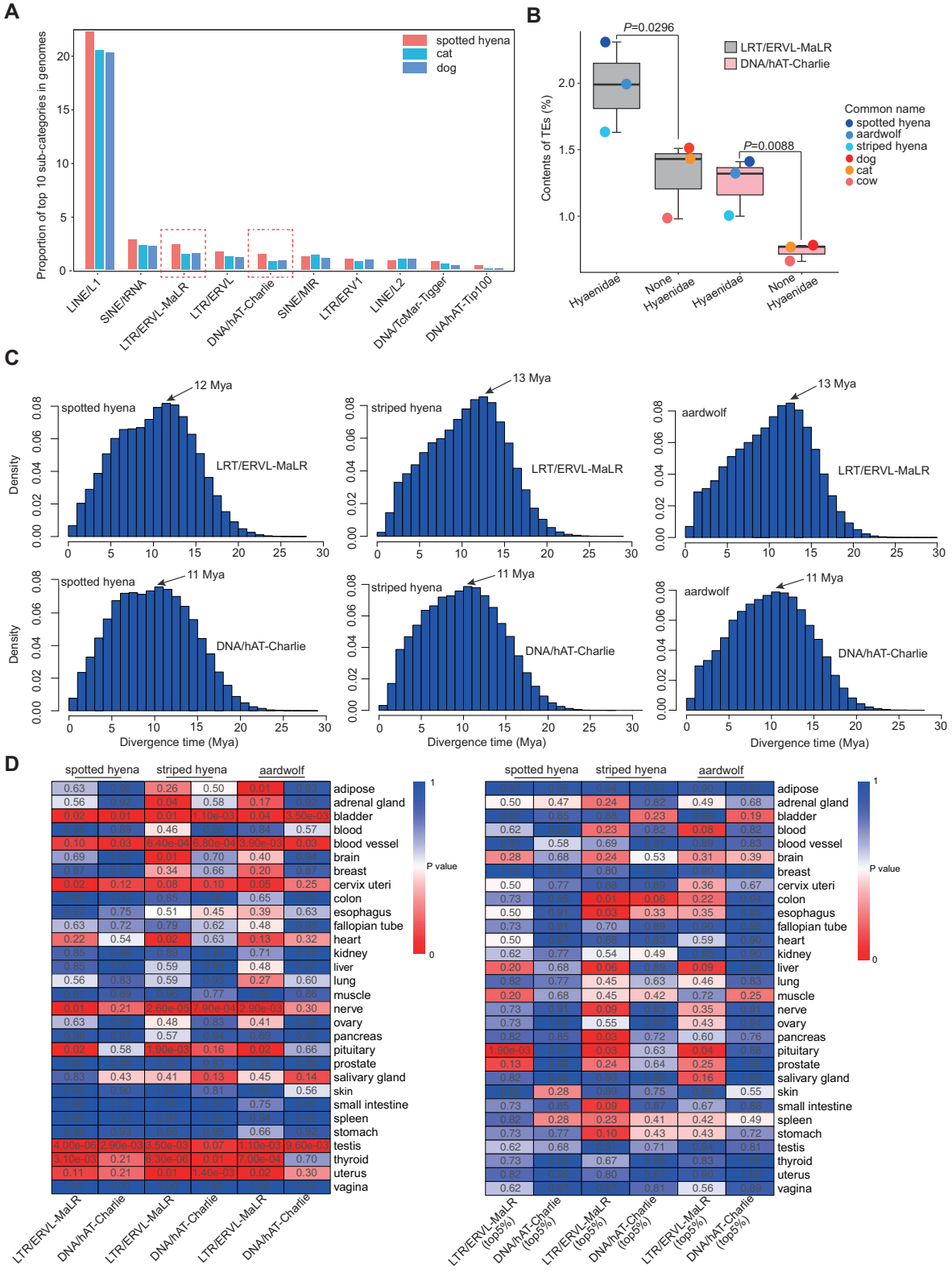

Fig. 2. Expansion of TEs in Hyaenidae species genomes. (*A*) Top ten TE subclasses in spotted hyena genome. Subclasses were compared with those of cat and dog. Red dashed box indicates candidate expanded TE subcategories in the spotted hyena genome. (*B*) Comparative analyses of genome contents of LTR/ERVL-MaLR and DNA/hAT-Charlie between Hyaenidae and non-Hyaenidae closely related species. Significance test ($P < 0.05$) was performed using unpaired one-tailed *t*-test. (*C*) Divergence time of LTR/ERVL-MaLR and DNA/hAT-Charlie in Hyaenidae species genomes. The TE burst divergence time is highlighted by the black arrow. Kimura nucleotide distance of masked regions against consensus sequences were automatically estimated by RepeatMasker, and TE ages were inferred under a mutation rate of 1.1×10^{-8} per generation (human) (Roach et al. 2010). (*D*) Significance of gene overlap between genes with expanded TE insertions in each Hyaenidae species and tissue-specific genes. Top 5% of genes were ranked by LTR/ERVLMaLR and DNA/hAT-Charlie insertion length in each gene locus. *P*-values of Fisher's exact test are shown in box.

may have played roles in the development and evolution of the urinary system, reproduction, and hormone regulation.

Expansion of Immune-Related Gene Families in Spotted Hyena Genome

Changes in gene family size potentially contribute to important genetic factors in phenotypic evolution (Qiu et al. 2012). Using the OrthoMCL algorithm (Li et al. 2003), we identified 11,753 homologous gene families that were shared by four species (cat, tiger, leopard, and spotted hyena) (fig. 3A). Furthermore, compared with other species (i.e., cat, tiger, leopard, dog, fox, panda, horse, and cow), 69 gene families were potentially specific to the spotted hyena. These specific gene families comprised 148 genes, 71 of which were identified by known InterPro domains (supplementary tables 8 and 9, Supplementary Material online). The spotted hyena-specific gene families were significantly over-represented in certain functional categories: that is, alternative splicing (UP_KEYWORDS, $P = 0.003$), synapse assembly (GO:0007416, $P = 0.007$), neuron projection development (GO:0031175, $P = 0.018$), and cell adhesion molecules (CAMs) (hsa04514, $P = 0.021$) (supplementary table 10, Supplementary Material online). Thus, these spotted hyena-specific gene families appear to be primarily involved in the nervous system.

We next identified 333 gene families that were substantially expanded in the spotted hyena compared with the most recent common ancestor of the spotted hyena and (cat [tiger, leopard]) (fig. 3B). Functional enrichment analyses revealed that expanded gene families in the spotted hyena were significantly related to immune categories: that is, immunoglobulin domain (UP_KEYWORDS, $P = 7.28E-17$), viral capsid (GO:0019028, $P = 2.46E-06$), viral envelope protein (UP_KEYWORDS, $P = 1.33E-04$), regulation of immune response (GO:0050776, $P = 3.22E-04$), virion (UP_KEYWORDS, $P = 0.0018$), negative regulation of lymphocyte activation (GO:0051250, $P = 0.0027$), viral process (GO:0016032, $P = 0.0042$), immunity (UP_KEYWORDS, $P = 0.0123$), and natural killer cell-mediated cytotoxicity (GO:0042267, $P = 0.0158$) (fig. 3C and supplementary table 11, Supplementary Material online).

Moreover, we identified 380 and 362 substantially expanded gene families in striped hyena and aardwolf, respectively (supplementary fig. 6, Supplementary Material online). For these expanded gene families in three Hyaenidae species, we found that 47 (in spotted hyena), 41 (in striped hyena), and 50 (in aardwolf) expanded gene families were involved in immune-related functional categories identified by the DAVID Functional Annotation Bioinformatics Microarray Analysis (Huang da et al. 2009a, 2009b) (fig. 3D). Further, we noticed that $\geq$62% of the immune-related expanded gene families in the Hyaenidae species were species-specific (i.e., 62% of 47 immune-related expanded gene families in spotted hyena, 71% of 41 immune-related expanded gene families in striped hyena, and 62% of 50 immune-related expanded gene families in aardwolf) (fig. 3D), suggesting that the evolution of immune-related gene families diverged during Hyaenidae species evolution. Additionally, we observed five expanded gene families, including *IGKV4-1* (immunoglobulin kappa variable 4-1), with 17 and 16 copies in the two scavengers compared with that in the termite-feeding aardwolf (fig. 3E), implying important functional roles in regulating immune system evolution in the spotted and striped hyenas. In addition, *IGKV4-1* and other genes (e.g., *IFITM1*, *IGLV1-40*, *IGHV3-48*, and *HLA-A*) in the segmental duplications of the spotted hyena genome (supplementary fig. 7 and supplementary table 12, Supplementary Material online) were also significantly enriched in immune-related functional terms, especially immunoglobulins (supplementary table 13, Supplementary Material online). The immune systems of animal scavengers are likely to have been molded by selective pressures associated with surviving microbial assaults from their food (Flies et al. 2012).

Adaptive Evolution of DNA Damage/Repair and Immune Genes in the Spotted Hyena Genome

Positive selection of genes plays important roles in phenotypic evolution of species (Bakewell et al. 2007; Shao et al. 2015). We identified 7,562 high-confidence single-copy one-to-one orthologous genes in 11 species, including the cat, tiger, leopard, spotted hyena, striped hyena, aardwolf, dog, fox, panda, horse, and cow (supplementary fig. 8, Supplementary Material online). We utilized the branch-site model under a likelihood ratio test to identify positively selected genes (PSGs) in the spotted hyena lineage. Compared with the striped hyena and aardwolf (56 and 143 PSGs, respectively) (supplementary tables 14 and 15, Supplementary Material online), we identified 70 PSGs in the spotted hyena (supplementary fig. 9 and supplementary table 16, Supplementary Material online). We compared the percentage distribution of tissue-specific expressed genes in the spotted hyena PSGs across 30 normal human tissues. The PSGs in the spotted hyena showed higher biased expression in eight tissues (fig. 4A), with digestion-related organs (e.g., pancreas and small intestine), immune-related organs (e.g., spleen), and metabolic organs (e.g., liver) preferentially expressing spotted hyena PSGs (fig. 4A). These results suggest potential molecular factors of adaptive evolution underlying digestion, metabolism, and immune systems in spotted hyenas (fig. 4B).

We also observed a significant PSG overlap (*SHPRH*, *TLE6*, *ZC3H18*, *CEP126*, *MILR1*, and *CORO7*) between the spotted and striped hyenas ($P = 1.1e-05$, Fisher's exact test) (fig. 4C), which participate in DNA repair [e.g., *SHPRH* (GO:0006281)], disease mutation (e.g., *TLE6*; UP_KEYWORDS), and immune processes (e.g., *MILR1*; immunoglobulin domain) (fig. 4D and supplementary table 17, Supplementary Material online). Indeed, due to their specific feeding behavior, spotted and striped hyenas are repeatedly exposed to pathogens from the carcasses they scavenge (Flies et al. 2014; Zhou et al. 2020). Chronic exposure to pathogens may have resulted in adaptive changes in various pathways, including DNA damage and repair, disease resistance, and immunity. Although congruent genes (e.g., *SHPRH* and *TLE6*) showed positive selection in the

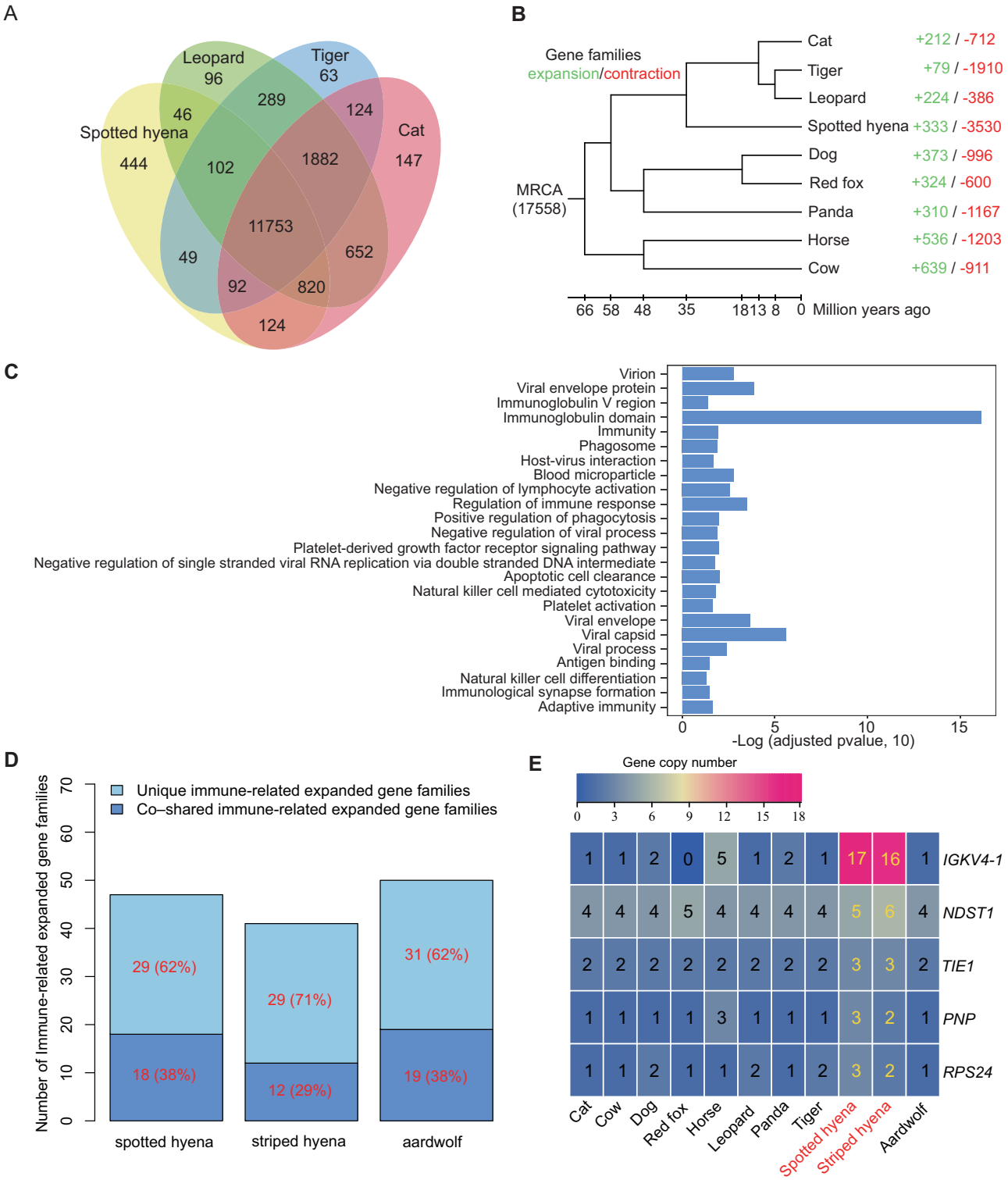

Fig. 3. Evolution of gene families in spotted hyena genome. (*A*) Venny plot of gene families across four species. (*B*) Expansion and contraction of gene families among nine species (cat, tiger, leopard, spotted hyena, dog, red fox, panda, horse, and cow). (*C*) Functional enrichment analysis of expanded gene families in the spotted hyena genome. The immune-related functional categories are shown in this barplot. (*D*) Comparative genomics analyses of immune-related expanded gene families in three Hyaenidae species. Light blue represents species-specific expanded gene families, and dark blue represents co-shared expanded gene families with at least one species. Number or proportion of unique immune-related expanded gene families or co-shared immune-related expanded gene families in all immune-related expanded gene families for each species is shown in red. (*E*) Five common immune-related expanded gene families in both scavenger genomes. Copy numbers in spotted hyena and striped hyena are highlighted in yellow.

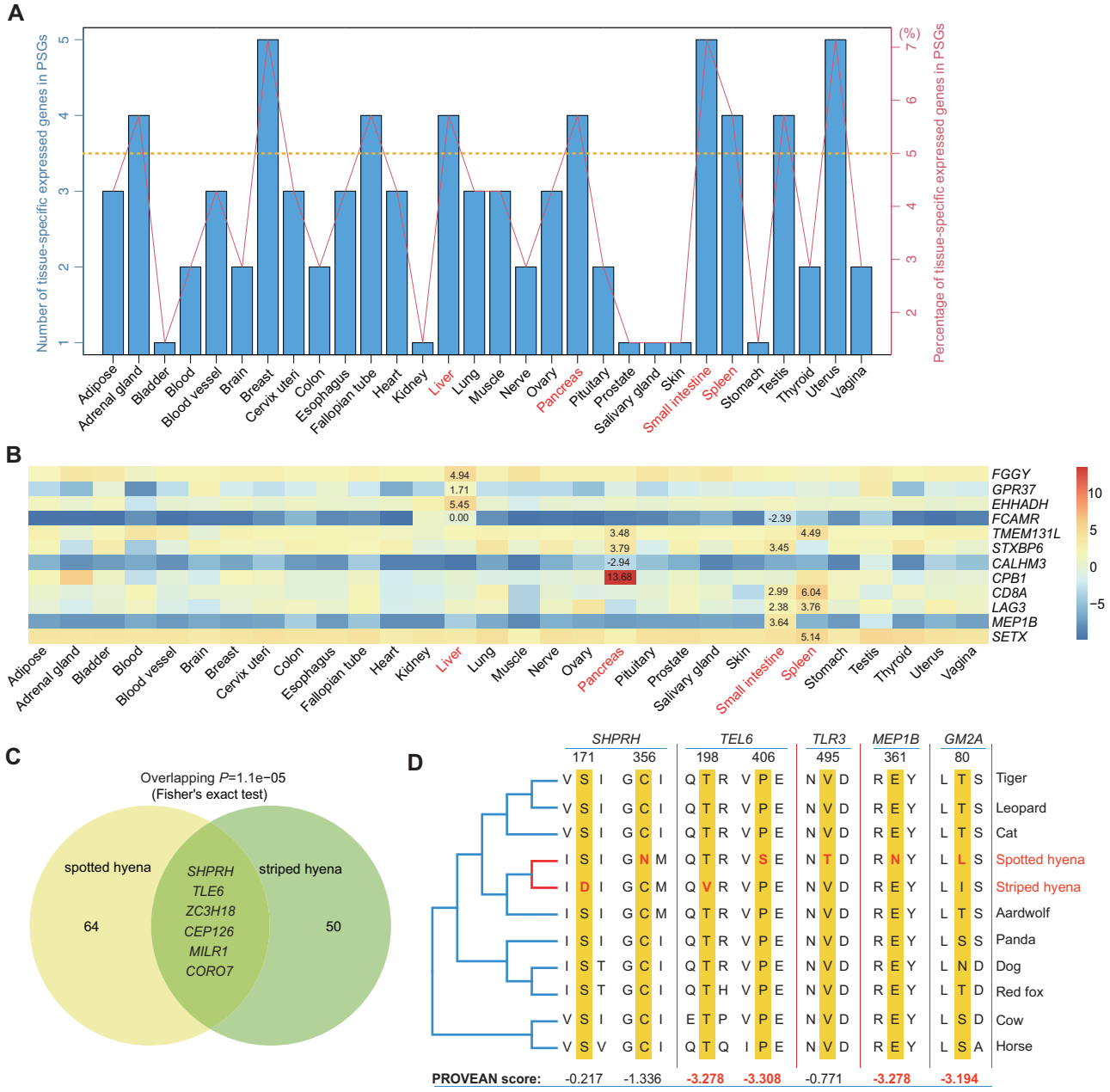

Fig. 4. Comparative analysis of PSGs in spotted hyena lineage. (*A*) Statistical analysis of tissue-specific expressed PSGs in spotted hyena genome. Yellow dotted line represents a percentage (5%) of tissue-specific expressed genes in PSGs in spotted hyena. In total, eight tissues showed a higher percentage of tissue-specific expressed genes in the spotted hyena PSGs. Digestion/immune/metabolism organs are highlighted in red. (*B*) Expression heatmap of tissue-specific expressed genes (in liver, spleen, small intestine, and pancreas) across 30 human tissues. Gene expression level was normalized by $\log_2$(expression). (*C*) Overlapping analysis of PSGs between spotted hyena and striped hyena. Overlapping *P*-value was calculated by Fisher' exact test ($P < 0.05$). (*D*) Sequence alignments of candidate PSGs in spotted hyena and striped hyena. Genes *SHPRH* and *TEL6* were co-shared as PSGs in spotted hyena and striped hyena with different positively selected sites. Examples of unique PSGs in spotted hyena include *TLR3*, *MEP1B*, and *GM2A*. PROVEAN score is also shown for each positively selected site. PROVEAN scores with <-2.5 are highlighted in red bold. Cat coding sequences (Ensembl v93) were regarded as the reference to decide the sequence alignment order.

two scavengers' lineages, their positively selected sites were located in different regions (fig. 4*D*), suggesting specific divergence of adaptive response mechanisms between spotted and striped hyenas to their long-term scavenging lifestyle. The genetic mechanisms of adaptive divergence were also reflected by the species-specific PSGs in the spotted hyena (64 species-specific PSGs) and striped hyena (50 species-specific PSGs) (fig. 4*C*). By searching functional annotations in the DAVID Bioinformatics Resources 6.8 database (Huang da et al. 2009a, 2009b) (e.g., UP_KEYWORDS, Gene Ontology (GO), and KEGG_PATHWAY), we found seven PSGs related to immunity (*CD8A*, *LAG3*, *TLR3*, *VSIG10*, *FCAMR*, *VSIG10*, and *PAK6*), three PSGs related to responses to toxins or toxin transport (*MEP1B*, *PON1*, and *MTMR12*), and three PSGs associated with lysosomes (*GM2A*, *GNS*, and *NAGLU*) specific to the spotted hyena lineage (fig. 4*D* and supplementary table

18, Supplementary Material online). Among these candidates, *CD8A*, *LAG3*, and *MEP1B* were specifically expressed in immune or digestion-related organs (spleen and small intestine) (fig. 4*B*). As ingested carrion contains large amounts of toxic metabolites and pathogens (Tanner et al. 2010), these species-specific PSGs in the spotted hyena may help them deal with toxic substance metabolism during digestion and improve immunity for pathogen resistance. Thus, we concluded that spotted and striped hyenas may have evolved different adaptive evolutionary strategies in response to their common scavenging lifestyles, which are continually exposed to pathogenic microorganisms.

Rapidly Evolving Conserved Noncoding Elements Potentially Contributed to Evolution Relating to Social Behavior in Spotted Hyena

The evolution of regulatory elements, such as conserved noncoding elements (CNEs), has played a crucial role in the phenotypic adaptations of animals to specific environments (Leal and Cohn 2016; Peng et al. 2020). Here, based on a previous comparative genomics methodology (Roscito et al. 2018) (fig. 5*A* and supplementary fig. 10, Supplementary Material online), we identified rapidly evolving CNEs (RECNEs) in the Hyaenidae species with a mean RECNE length of ∼26 bp (fig. 5*B* and supplementary fig. 11 and supplementary tables 19–21, Supplementary Material online). In total, 324, 329, and 583 lineage-specific RECNEs were obtained in the spotted hyena, striped hyena, and aardwolf, respectively (fig. 5*C*), and we retrieved 65 lineage-specific RECNE-associated genes with a distance of ≤5kb to the RECNEs in the spotted hyena (supplementary table 22, Supplementary Material online) compared with the striped hyena and aardwolf (supplementary tables 23 and 24, Supplementary Material online).

Functional enrichment analyses indicated that the lineage-specific RECNE-associated genes in Hyaenidae species evolution were mainly involved in gene transcription and regulation (supplementary tables 25–27, Supplementary Material online). Intriguingly, for the spotted hyena lineage-specific RECNE-associated genes, the most significantly enriched category "GO: 0007411∼axon guidance" ($P = 0.001$), which included five genes (*EFNA3*, *DCC*, *FEZ2*, *NRXN3*, and *ATOH1*) related to neural system functions involving migration of axon growth cones, did not emerge in the enrichment analyses of the striped hyena or aardwolf (supplementary tables 25–27, Supplementary Material online). Typically, the *EFNA3* gene is involved in autism (Marchese et al. 2016), which is characterized by impaired social communication (Fernandez and Scherer 2017; Galvez-Contreras et al. 2017; Szoko et al. 2017). Compared with the striped hyena and aardwolf, we observed a short distance (504 bp) between *EFNA3* and RECNE (scaffold486: 168,079–168,129; length = 50 bp) in the spotted hyena lineage and high sequence divergence in RECNE with eight mutations (supplementary fig. 12, Supplementary Material online). Based on the GO and Human Phenotype Database (supplementary table 28, Supplementary Material online), we identified a series of RECNE-associated genes in the spotted hyena involved in social behavior, learning, and cognition (e.g., *UGP2*, *ACTR2*, *ARL6*, *NSD1*). Of note, from the Human Phenotype Database, *UGP2* is involved in impaired social interactions (supplementary table 28, Supplementary Material online) and its loss in the human brain can lead to severe epileptic encephalopathy (Perenthaler et al. 2020). Another lineage-specific RECNE in the spotted hyena harbored a 10-bp deletion compared with other two Hyaenidae species and nine outgroup species without complex social behaviors (fig. 5*D*) and the RECNE-associated gene, *ACTR2*, involving in synaptic plasticity is related to improved learning (González-Martín et al. 2021). Furthermore, Illumina short-read mapping demonstrated the accuracy of this RECNE and excluded the possibility of genome assembly error (supplementary fig. 13, Supplementary Material online). Our results suggest that these lineage-specific RECNEs possibly harbor a revised functional role in the spotted hyena, differentiated from that of other species, and contribute to the regulatory roles of those genes in social behaviors, for example, *EFNA3*, *UGP2*, and *ACTR2*. The functional implications of these candidates suggest that at least some RECNEs likely regulate the complexity of social behavior in spotted hyena communities. However, further experiments are needed to validate the functional roles of these rapidly evolving CNEs in regulating societal complexity.

Discussion

In this study, we sequenced and assembled a high-quality spotted hyena genome. Compared with previous spotted hyena genome assembly, with a contig N50 of 21.301 kb (Yang et al. 2020), our genome assembly was improved by long-read sequencing (PacBio Sequel), resulting in a contig N50 length of ∼13.75 Mb. The improved spotted hyena genome permits further discovery of the evolutionary features of the genome and better understanding of the genetic underpinnings of specific phenotypic evolution.

Evolution of Immune Tolerance in Spotted Hyena

Scavenging animals play key roles in various ecosystems by the removal of carcasses, thus preventing the spread of disease (Chung et al. 2015). The scavenging lifestyle also suggests immune system tolerance, indicating that these animals have potentially undergone adaptive evolution in response to chronic exposure to pathogens. For example, studies on obligate scavenging birds (i.e., turkey and cinereous vultures) have explored the potential genetic mechanisms related to the prevention of infection from microbes that exist in their diet, and several immune genes have been reported to be under positive selection (i.e., *AHSG* and *BCL6*) (Chung et al. 2015; Zhou et al. 2019).

However, compared with scavenging birds, the underlying molecular mechanisms related to immune tolerance to pathogens in mammalian scavengers remain unclear. Hyaenidae species, especially brown, striped, and spotted hyenas, are excellent models for testing the potential evolutionary mechanisms underlying their immune systems. Although striped and spotted hyenas have undergone next-generation (short-read) genomic sequencing, details on the genetic

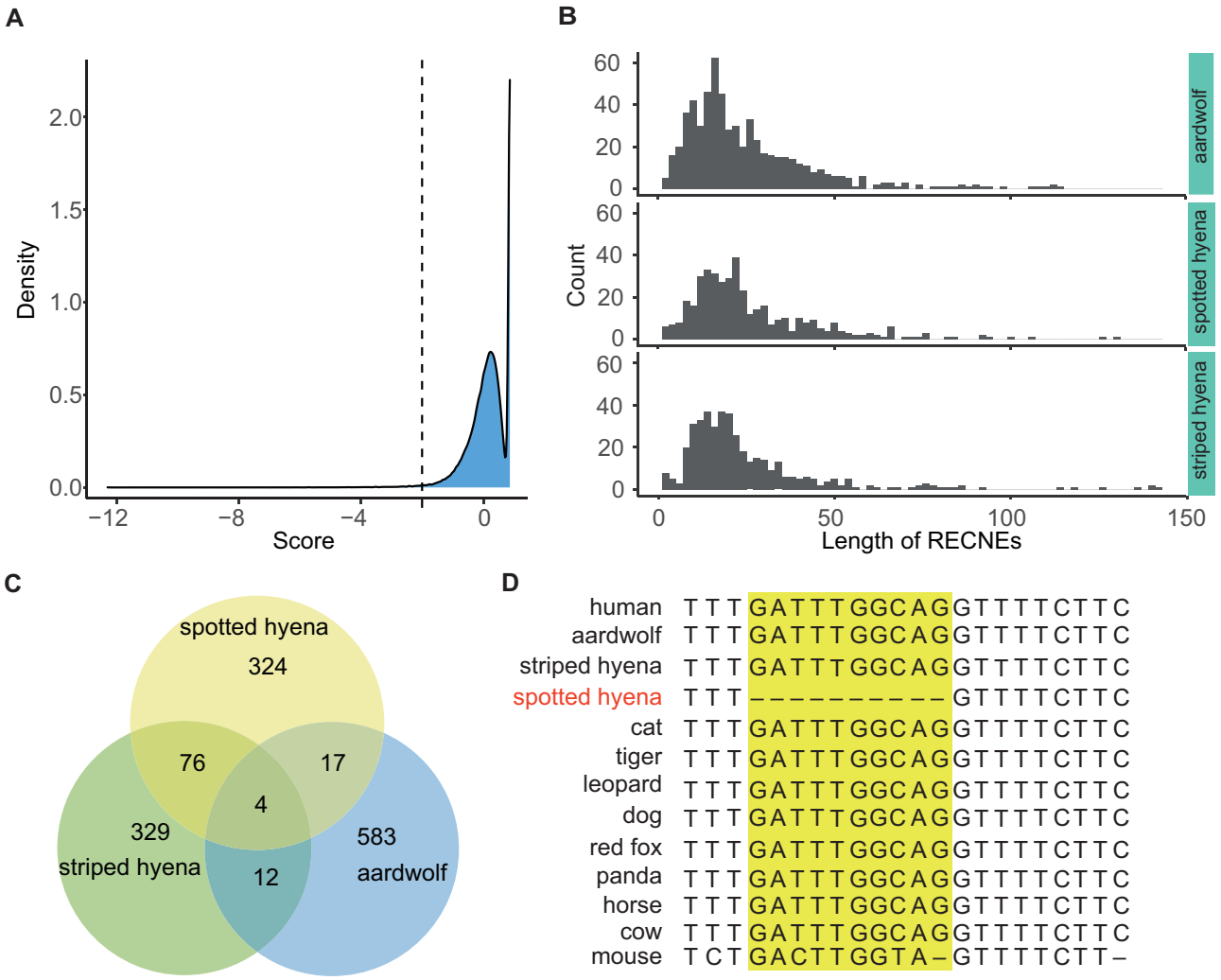

Fig. 5. Identification and functional analysis of RECNEs in spotted hyena lineage. (*A*) Identification of RECNEs in spotted hyena lineage using global *Z*-score method. Dotted line represents the top 1% of global *Z*-scores. (*B*) Length distribution of RECNEs for three Hyaenidae species. (*C*) Comparative genomics analysis of RECNEs in three Hyaenidae species. (*D*) Multiple species sequence alignment of RECNEs with the neighboring gene *ACTR2* in the spotted hyena genome. RECNE is located at the region (scaffold33: 8,759,259–8759271; length = 12 bp) in the spotted hyena genome. The 10-bp deletion sequence in the spotted hyena genome is highlighted in yellow.

mechanisms of adaptive evolution, especially for the immune system, are still incomplete (Westbury et al. 2018; Yang et al. 2020; Zhou et al. 2020). In this study, we constructed an improved long-read genome assembly for the spotted hyena to explore the potential genetic basis of immune system evolution. Our analyses hint that the evolution of immune tolerance to pathogenic microorganisms in spotted hyenas may be involved in multiple strategies functioning on immune-related genes, such as gene duplications and adaptive sequence changes.

Immunoglobulins play important roles in recognizing an immense variety of different pathogens and responding dynamically as pathogens evolve during infection (Maizels 2005). Here, gene family analyses indicated that immunoglobulin family members (e.g., *IGKV4-1*) have undergone significant adaptive duplications in the spotted hyena and striped hyena genomes. We also detected that another immunoglobulin gene, *FCAMR*, experienced positive selection in the spotted hyena lineage with high expression in the liver (fig. 4*B*). *FCAMR* encodes a receptor for IgA and IgM and can modulate antigen-driven immune responses in mice (Choi et al. 2013). These analyses suggest that immunoglobulin genes have undergone adaptive evolution in the spotted hyena immune system. Significantly, other immune-related genes have also experienced positive selection in the spotted hyena, such as *TLR3* (Toll-like receptor 3), which evolved under strong purifying selection throughout human history and plays a fundamental role in pathogen recognition and activation of innate immunity (Muzio et al. 2000; Zhang et al. 2013), suggesting a vital role in the evolution of immune tolerance in the spotted hyena. Additionally, we found that certain PSGs, especially *MEP1B* (Van Spaendonk et al. 2017) with high tissue-specific expression in the small intestine and with potential functional changes in the positively selected site (PROVEAN < -2.5), are associated with toxin transport, which is important for detoxification and metabolism of toxic substances. We also found a series of PSGs involved in the biological processes of DNA damage and repair. These results

imply metabolism/repair-associated genes also underwent adaptive evolution to cope with the scavenging lifestyle. Our analyses further suggest that evolutionary mechanisms of immune tolerance in spotted and striped hyenas have experienced adaptive divergences at the molecular level.

Societal Complexity of Spotted Hyena

Spotted hyenas are large terrestrial carnivores that occur throughout sub-Saharan Africa. The complexity of spotted hyena societies is comparable with that of cercopithecine primates, particularly in regard to group size, structure, and patterns of competition and cooperation, and far exceeds that of any other carnivore (Holekamp et al. 2007). Computed tomography has also indicated that the anatomy of the spotted hyena brain resembles that of primates with a large frontal cortex, which is involved in the regulation of social behavior (Dunbar and Bever 2010; Adolphs 2001; Amodio and Frith 2006; Holekamp et al. 2007). This suggests that spotted hyenas may possess a relatively high social cognitive ability. Based on RECNE analysis, we unexpectedly found that adjacent genes (e.g., *EFNA3*, *UGP2*, and *ACTR2*) of certain RECNEs in the spotted hyena lineage were strongly related to neural system function, especially social communication, learning, and cognition. For example, the learning-related gene, *ACTR2*, was also adjacent to the RECNE with a 10-bp deletion in the spotted hyena genome, hinting at its potential functional role in regulating social behaviors. However, further functional experiments are needed to validate this possibility. Our analyses hint at the potential genetic factors regulating social complexity in spotted hyenas.

Conclusions

In this study, we produced a high-quality long-read genome assembly for the spotted hyena. Furthermore, we unraveled the evolutionary features of its genome and potential molecular/genetic mechanisms of several unique phenotypes, namely complex societal structure and immune system tolerance, in the spotted hyena lineage. We also highlighted specific functions of TEs (e.g., LTR/ERVL-MaLR and DNA/hAT-Charlie) with high content in Hyaenidae species genomes. This study provides an excellent basis for future functional experiments as well as phenotypic and evolutionary biology studies on spotted hyenas.

Materials and Methods

Sample Collection and Single-Molecule Real-Time Sequencing

All animal experiments were approved by the Ethics and Experimental Animal Committee of the Kunming Institute of Zoology, Chinese Academy of Sciences, China (Approval ID: SMKX-20200120-27). Genomic DNA was isolated from blood samples of a male spotted hyena (*Crocuta crocuta*) in Beijing Zoo, China. The large-insert PacBio library preparation was conducted following the Pacific Biosciences recommended protocols. In brief, high-quality DNA was fragmented randomly, then purified and further repaired. The sequencing adapters were ligated, and failed fragments were removed. A 20-kb library was constructed and further sequenced by PacBio Sequel.

Long-Read Genome Assembly of Spotted Hyena

The ~64.49× PacBio Sequel subreads were applied to produce a de novo assembly using FALCON (Chin et al. 2016). The assembled genome was further polished by mapping the subreads to the contigs of the assembly using Quiver (Chin et al. 2013). Base accuracy of the assembly was evaluated by paired-end short reads from the Illumina platform. Assembly completeness was estimated using CEGMA (https://github.com/KorfLab/CEGMA_v2) and BUSCO (http://busco.ezlab.org/). Genome size was estimated by the k-mer algorithm.

Genome Annotation of Spotted Hyena

The de novo and homology strategies were utilized to produce repetitive sequences for the spotted hyena genome. We constructed a de novo repeat library using RepeatModeler (v1.0.8) (http://www.repeatmasker.org/RepeatModeler/) and ran RepeatMasker (v4.0.7) (http://www.repeatmasker.org/) for the spotted hyena assembly using the constructed de novo library. To obtain homologous repeat annotations, we ran RepeatMasker against RepBase (v20170127) (https://www.girinst.org/repbase/), and the repeat annotations of spotted hyena were combined from the two strategies. Similarly, we also used de novo and homology-based algorithms to annotate the protein-coding genes of the spotted hyena genome, and the last annotation version was obtained from their integrations. In brief, Augustus (v3.0.3) (Stanke and Waack 2003) and SNAP (Korf 2004) were applied to predict genes in the repeat-masked spotted hyena genome, and PASA (v2014-04-17) (Haas et al. 2003) was utilized to train gene model parameters. For homology-based annotations, protein sequences from five species (cat, tiger, dog, panda, and human) were downloaded from the Ensembl genome browser (http://asia.ensembl.org/index.html) and were aligned to the spotted hyena assembly using tblastn with an *e*-value $< 10^{-5}$. GenBlastA (v1.0.138) (She et al. 2009) was used to cluster adjacent HSPs from the same protein alignments, and GeneWise (v2.4.1) (Birney et al. 2004) was used to identify the precise gene structures. EvidenceModeler (v2012-06-25) (Haas et al. 2008) was applied to integrate all predicted genes of different strategies and form a high-confidence gene set. The functional annotations were assigned and integrated using diverse databases including Kyoto Encyclopedia of Genes and Genomes (KEGG, https://www.genome.jp/kegg/), SwissProt, TrEMBL (https://www.uniprot.org/), GO (http://geneontology.org/), and InterPro (http://www.ebi.ac.uk/interpro/).

Identification of Gene Families and Evaluation of Divergence Time

Coding sequences of eight species (cat, tiger, leopard, dog, fox, panda, horse, and cow) were downloaded from the Ensembl genome browser and, together with that of the spotted hyena, were used to construct gene families. Briefly, we removed CDS sequences with fewer than 30 encoded amino acids, premature stop codons, and nontriplet lengths. OrthoMCL (v2.0.9)

(Li et al. 2003) was used to produce gene family categories. Cafe (https://hahnlab.sitehost.iu.edu/software.html) was used to detect the expansion and contraction of gene families. Significantly expanded gene families for each lineage were identified by Viterbi $P < 0.05$. The sequences of each single-copy gene family from the nine species were aligned by MUSCLE (v3.8.31) (https://drive5.com/muscle5/) and then concatenated. We utilized RAxML (v8.2.9) (https://cme.h-its.org/exelixis/web/software/raxml/index.html) under 100 bootstrap replicates and a PROTGAMMAAUTO model to construct a phylogenetic tree. Based on the above constructed phylogenetic topology, MCMCtree in PAML (v4.4) (http://abacus.gene.ucl.ac.uk/software/paml.html) was used to analyze divergence time across species lineages, and fossil calibration points were determined using TimeTree (http://www.timetree.org/).

Segmental Duplication Analyses

Comparison of the whole-genome assembly was conducted to confirm segmental duplications. Self-alignment of the repeat-masked genome sequence was performed with LASTZ (v1.03.73) (–identity = 90 –notrivial) (https://www.geneious.com/plugins/lastz-plugin/). The resulting alignments with lengths larger than 1 kb and maximum simultaneous gaps ≤100 bp were regarded as segmental duplications.

Enrichment Analyses of Expanded TEs in Diverse Tissues

We downloaded the gene expression matrix of 30 tissues from 7,862 human samples from GTEx (https://commonfund.nih.gov/GTEx/). We further identified the single-copy orthologous genes between humans and spotted hyenas using OrthoMCL (v.2.0.9) (Li et al. 2003). The expression matrix of single-copy orthologous genes between humans and spotted hyenas was extracted using an in-house script. We then identified tissue-specific genes for each tissue based on *t*-statistics. Briefly, we performed *t*-statistics on single-copy orthologous genes between the target tissue samples and all other samples, and then ranked the *t* values from high to low. According to rank, the top 5% genes based on *t* values were defined as tissue-specific genes. We also calculated the distributions of expanded TEs in single-copy orthologous gene loci, including 5-kb flanking regions, exons, and introns. We sorted the length ratio (length of expanded TEs located in gene locus/gene locus length) from high to low and identified the top 5% genes enriched in expanded TEs. Gene overlap between genes involved in expanded TEs and tissue-specific genes was determined by Fisher's exact test, with $P < 0.05$ (https://www.bioconductor.org/packages/release/bioc/html/GeneOverlap.html).

Identification and Analysis of Positively Selected Genes

The single-copy orthologous genes among cat, tiger, leopard, dog, fox, panda, horse, spotted hyena, striped hyena, aardwolf, and cow were identified by OrthoMCL (v2.0.9) (Li et al. 2003). The single-copy orthologous gene sequences across the nine species were aligned by MUSCLE (v3.8.31) (https://drive5.com/muscle5/), and low-quality aligned regions were further trimmed by Gblocks (v0.91b) with default parameters. The aligned genes with CDS lengths <100 bp were removed for downstream evolutionary analyses. Based on a reliable constructed species-guided tree topology, the branch-site model in PAML (v4.4) (http://abacus.gene.ucl.ac.uk/software/paml.html) and the likelihood rate test were utilized to detect PSGs in the spotted hyena, striped hyena, and aardwolf lineages based on the single-copy orthologous genes with $P < 0.05$ (χ^2 test). The Bayes empirical Bayes algorithm was applied to calculate the posterior probabilities for inferred positively selected sites. Furthermore, we removed PSGs with positively selected sites located at the starting 1–5 amino acids of CDS and PSGs with the distance between adjacent positively selected sites ≤5 amino acids. To further elucidate functional implications on mutations, we used PROVEAN (http://provean.jcvi.org) to predict potential impacts on positively selected sites. PROVEAN provides a score to evaluate the potential effect of the substitution. A cutoff score of ≤−2.5 suggests a high probability that the mutation experienced a functional change (Choi et al. 2012).

Identification of Rapidly Evolving CNEs in Spotted Hyena Lineage

We used LAST (v992) (Kiełbasa et al. 2011) to conduct whole-genome alignments among 11 species, including human, mouse, cow, horse, spotted hyena, dog, cat, tiger, leopard, fox, and panda. Each genome was aligned to the human genome (as reference) using the lastal command in LAST (v992) (Kiełbasa et al. 2011) with the parameter "-m 100 -E 0.05." We used the last-split, maf-swap, and maf-sort commands on each pairwise mapping result to generate ordered MAF-format alignment. All pairwise genome alignments were processed by MULTIZ (v11.2) (Blanchette et al. 2004) to combine into multiple alignments. To detect CNEs, we used PhastCons from rphast (v1.6.9) (Hubisz et al. 2011) with parameters "estimate.rho = TRUE." PhastCons needs a phylogenetic tree with neutral branch lengths as input, so we used phyloFit from the same package (parameters "–EM –precision HIGH –subst-mod REV") to estimate branch lengths based on 4-fold degenerated third codon positions. We removed those results that overlapped with coding regions. To detect rapidly evolving CNEs in the spotted hyena, we used a previously described method (Roscito et al. 2018). Briefly, the forward genomics method (Prudent et al. 2016) was used to calculate sequence identities between the ancestral and CNE sequences of each species. We calculated local *Z*-scores by comparing the sequence identities of the spotted hyena with those of the cat, tiger, leopard, striped hyena, and aardwolf. The global *Z*-scores were calculated between the sequence identities of the spotted hyena and all other species. The sequence identities of CNEs were further used to compute global and local *Z*-scores. A CNE with top 1% global and local *Z*-scores was regarded as a rapidly evolving CNE.

GO/Pathway Enrichment Analysis

DAVID (v6.8) (https://david.ncifcrf.gov/summary.jsp) was used for functional enrichment analysis, including disease,

functional categories, GO, pathways, and protein domains for candidate gene lists under a current background-*Homo sapiens* with a two tailed corrected Fisher's exact test ($P < 0.05$).

Supplementary Material

Supplementary data are available at *Molecular Biology and Evolution* online.

Acknowledgments

This work was supported by the Second Tibetan Plateau Scientific Expedition and Research (STEP) Program (2019QZKK05010703), the Strategic Priority Research Program of the Chinese Academy of Sciences (XDA2004010302), and the Young Academic and Technical Leader Raising Foundation of Yunnan Province.

Author Contributions

Y.S., C.L.Z., J.R., and Y.L. designed and led the project. Y.S. analyzed the genome and drafted the article. X.B.W. performed the assembly, annotation, and comparative genomics analyses of genomic data, and wrote the methods. M.L.Z. and Y.L. conducted genomic analyses and plotted figures. S.W., B.L.Z., M.M.Y., M.H.Y., T.J., T.C.P., Y.L., H.L., Z.X., B.L., and N.L. analyzed the sample and finished DNA experiments. D.D.W. and V.M.O. discussed the results and edited the manuscript. All authors approved the final version of the manuscript.

Data Availability

The PacBio genome assembly and annotation were deposited at the Genome Warehouse in the National Genomics Data Center (http://bigd.big.ac.cn/gwh/) under BioProject accession code: PRJCA004316 (accession code: GWHAZPN00000000). The genome short-read sequencing data were deposited at the Sequence Read Archive (SRA) database of the NCBI under BioProject accession code: PRJNA693355.

References

Amodio DM, Frith CD. 2006. Meeting of minds: the medial frontal cortex and social cognition. *Nat Rev Neurosci*. 7(4):268–277.

Arsznov BM, Lundrigan BL, Holekamp KE, Sakai ST. 2010. Sex and the frontal cortex: a developmental CT study in the spotted hyena. *Brain Behav Evol*. 76(3–4):185–197.

Bakewell MA, Shi P, Zhang J. 2007. More genes underwent positive selection in chimpanzee evolution than in human evolution. *Proc Natl Acad Sci U S A*. 104(18):7489–7494.

Benbow MT, Tarone AM. 2015. Carrion ecology, evolution, and their applications. Boca Raton (FL): CRC Press.

Birney E, Clamp M, Durbin R. 2004. GeneWise and Genomewise. *Genome Res*. 14(5):988–995.

Blanchette M, Kent WJ, Riemer C, Elnitski L, Smit AF, Roskin KM, Baertsch R, Rosenbloom K, Clawson H, Green ED, et al. 2004. Aligning multiple genomic sequences with the threaded blockset aligner. *Genome Res*. 14(4):708–715.

Chin CS, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Clum A, Copeland A, Huddleston J, Eichler EE, et al. 2013. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nat Methods*. 10(6):563–569.

Chin CS, Peluso P, Sedlazeck FJ, Nattestad M, Concepcion GT, Clum A, Dunn C, O'Malley R, Figueroa-Balderas R, Morales-Cruz A, et al. 2016. Phased diploid genome assembly with single-molecule real-time sequencing. *Nat Methods*. 13(12):1050–1054.

Choi S-C, Wang H, Tian L, Murakami Y, Shin D-M, Borrego F, Morse HC, Coligan JE. 2013. Mouse IgM Fc receptor, *FCMR*, promotes B cell development and modulates antigen-driven immune responses. *J Immunol*. 190(3):987–996.

Choi Y, Sims GE, Murphy S, Miller JR, Chan AP. 2012. Predicting the functional effect of amino acid substitutions and indels. *PLoS One*. 7(10):e46688.

Chung O, Jin S, Cho YS, Lim J, Kim H, Jho S, Kim HM, Jun J, Lee H, Chon A, et al. 2015. The first whole genome and transcriptome of the cinereous vulture reveals adaptation in the gastric and immune defense systems and possible convergent evolution between the Old and New World vultures. *Genome Biol*. 16:215.

Chuong EB, Elde NC, Feschotte C. 2017. Regulatory activities of transposable elements: from conflicts to benefits. *Nat Rev Genet*. 18(2):71–86.

Cunha GR, Risbridger G, Wang H, Place NJ, Grumbach M, Cunha TJ, Weldele M, Conley AJ, Barcellos D, Agarwal S, et al. 2014. Development of the external genitalia: perspectives from the spotted hyena (*Crocuta crocuta*). *Differentiation* 87(1–2):4–22.

de la Lastra JM, de la Fuente J. 2007. Molecular cloning and characterisation of the griffon vulture (*Gyps fulvus*) toll-like receptor 1. *Dev Comp Immunol*. 31(5):511–519.

Dunbar RIM, Bever J. 2010. Neocortex size predicts group size in carnivores and some insectivores. *Ethology* 104(8):695–708.

Fang L, Cai W, Liu S, Canela-Xandri O, Gao Y, Jiang J, Rawlik K, Li B, Schroeder SG, Rosen BD, et al. 2020. Comprehensive analyses of 723 transcriptomes enhance genetic and biological interpretations for complex traits in cattle. *Genome Res*. 30(5):790–801.

Fernandez BA, Scherer SW. 2017. Syndromic autism spectrum disorders: moving from a clinically defined to a molecularly defined approach. *Dialogues Clin Neuro*. 19:353–371.

Finucane HK, Reshef YA, Anttila V, Slowikowski K, Gusev A, Byrnes A, Gazal S, Loh PR, Lareau C, Shoresh N, et al. 2018. Heritability enrichment of specifically expressed genes identifies disease-relevant tissues and cell types. *Nat Genet*. 50(4):621–629.

Flies AS, Grant CK, Mansfield LS, Smith EJ, Weldele ML, Holekamp KE. 2012. Development of a hyena immunology toolbox. *Vet Immunol Immunopathol*. 145(1–2):110–119.

Flies AS, Maksimoski MT, Mansfield LS, Weldele ML, Holekamp KE. 2014. Characterization of Toll-like receptors 1–10 in spotted hyenas. *Vet Res Commun*. 38(2):165–170.

Gade DW. 2006. Hyenas and humans in the horn of Africa. *Geogr Rev*. 96(4):609–632.

Galvez-Contreras AY, Campos-Ordonez T, Gonzalez-Castaneda RE, Gonzalez-Perez O. 2017. Alterations of growth factors in autism and attention-deficit/hyperactivity disorder. *Front Psychiatry*. 8:126.

González-Martín A, Moyano T, Gutiérrez DA, Carvajal FJ, Cerpa W, Hanley JG, Gutiérrez RA, Álvarez AR. 2021. c-Abl regulates a synaptic plasticity-related transcriptional program involved in memory and learning. *Prog Neurobiol*. 205:102122.

Gusset M, Burgener N. 2005. Estimating larger carnivore numbers from track counts and measurements. *Afr J Ecol*. 43(4):320–324.

Haas BJ, Delcher AL, Mount SM, Wortman JR, Smith RK, Hannick LI, Maiti R, Ronning CM, Rusch DB, Town CD, et al. 2003. Improving the Arabidopsis genome annotation using maximal transcript alignment assemblies. *Nucleic Acids Res*. 31(19):5654–5666.

Haas BJ, Salzberg SL, Zhu W, Pertea M, Allen JE, Orvis J, White O, Buell CR, Wortman JR. 2008. Automated eukaryotic gene structure annotation using EVidenceModeler and the program to assemble spliced alignments. *Genome Biol*. 9(1):R7.

Hayward MW. 2006. Prey preferences of the spotted hyaena (*Crocuta crocuta*) and degree of dietary overlap with the lion (*Panthera leo*). *J Zool*. 270(4):606–614.

Holekamp KE. 2006. Spotted hyenas. *Curr Biol*. 16(22):R944–R945.

Holekamp KE, Sakai ST, Lundrigan BL. 2007. Social intelligence in the spotted hyena (*Crocuta crocuta*). *Philos Trans R Soc Lond B Biol Sci*. 362(1480):523–538.

Holekamp KE, Smith JE, Strelioff CC, Van Horn RC, Watts HE. 2012. Society, demography and genetic structure in the spotted hyena. *Mol Ecol*. 21(3):613–632.

Huang da W, Sherman BT, Lempicki RA. 2009a. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res*. 37(1):1–13.

Huang da W, Sherman BT, Lempicki RA. 2009b. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc*. 4(1):44–57.

Hubisz MJ, Pollard KS, Siepel A. 2011. PHAST and RPHAST: phylogenetic analysis with space/time models. *Brief Bioinform*. 12(1):41–51.

Kiełbasa SM, Wan R, Sato K, Horton P, Frith MC. 2011. Adaptive seeds tame genomic sequence comparison. *Genome Res*. 21(3):487–493.

Koenig R. 2006. Vulture research soars as the scavengers' numbers decline. *Science* 312(5780):1591–1592.

Koepfli KP, Jenks SM, Eizirik E, Zahirpour T, Van Valkenburgh B, Wayne RK. 2006. Molecular systematics of the Hyaenidae: relationships of a relictual lineage resolved by a molecular supermatrix. *Mol Phylogenet Evol*. 38(3):603–620.

Korf I. 2004. Gene finding in novel genomes. *BMC Bioinformatics*. 5:59.

Leal F, Cohn MJ. 2016. Loss and re-emergence of legs in snakes by modular evolution of sonic hedgehog and *HOXD* enhancers. *Curr Biol*. 26(21):2966–2973.

Li L, Stoeckert CJ Jr, Roos DS. 2003. OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome Res*. 13(9):2178–2189.

Maizels N. 2005. Immunoglobulin gene diversification. *Annu Rev Genet*. 39:23–46.

Marchese M, Valvo G, Moro F, Sicca F, Santorelli FM. 2016. Targeted gene resequencing (astrochip) to explore the tripartite synapse in autism-epilepsy phenotype with macrocephaly. *Neuromolecular Med*. 18(1):69–80.

Muzio M, Bosisio D, Polentarutti N, D'amico G, Stoppacciaro A, Mancinelli R, van't Veer C, Penton-Rol G, Ruco LP, Allavena P, et al. 2000. Differential expression and regulation of toll-like receptors (TLR) in human leukocytes: selective expression of *TLR3* in dendritic cells. *J Immunol*. 164(11):5998–6004.

Peng C, Ren JL, Deng C, Jiang D, Wang J, Qu J, Chang J, Yan C, Jiang K, Murphy RW, et al. 2020. The genome of Shaw's sea snake (*Hydrophis curtus*) reveals secondary adaptation to its marine environment. *Mol Biol Evol*. 37(6):1744–1760.

Perenthaler E, Nikoncuk A, Yousefi S, Berdowski WM, Alsagob M, Capo I, van der Linde HC, van den Berg P, Jacobs EH, Putar D, et al. 2020. Loss of *UGP2* in brain leads to a severe epileptic encephalopathy, emphasizing that bi-allelic isoform-specific start-loss mutations of essential genes can cause genetic diseases. *Acta Neuropathol*. 139(3):415–442.

Prudent X, Parra G, Schwede P, Roscito JG, Hiller M. 2016. Controlling for phylogenetic relatedness and evolutionary rates improves the discovery of associations between species' phenotypic and genomic differences. *Mol Biol Evol*. 33(8):2135–2150.

Qiu Q, Zhang G, Ma T, Qian W, Wang J, Ye Z, Cao C, Hu Q, Kim J, Larkin DM, et al. 2012. The yak genome and adaptation to life at high altitude. *Nat Genet*. 44(8):946–949.

Roach JC, Glusman G, Smit AF, Huff CD, Hubley R, Shannon PT, Rowen L, Pant KP, Goodman N, Bamshad M, et al. 2010. Analysis of genetic inheritance in a family quartet by whole-genome sequencing. *Science* 328(5978):636–639.

Roscito JG, Sameith K, Parra G, Langer BE, Petzold A, Moebius C, Bickle M, Rodrigues MT, Hiller M. 2018. Phenotype loss is associated with widespread divergence of the gene regulatory landscape in evolution. *Nat Commun*. 9(1):4737.

Shao Y, Li JX, Ge RL, Zhong L, Irwin DM, Murphy RW, Zhang YP. 2015. Genetic adaptations of the plateau zokor in high-elevation burrows. *Sci Rep*. 5:17262.

She R, Chu JS, Wang K, Pei J, Chen N. 2009. GenBlastA: enabling BLAST to identify homologous gene sequences. *Genome Res*. 19(1):143–149.

Stanke M, Waack S. 2003. Gene prediction with a hidden Markov model and a new intron submodel. *Bioinformatics*. 19(Suppl 2):ii215–ii225.

Szoko N, McShane AJ, Natowicz MR. 2017. Proteomic explorations of autism spectrum disorder. *Autism Res*. 10(9):1460–1469.

Tanner JB, Zelditch ML, Lundrigan BL, Holekamp KE. 2010. Ontogenetic change in skull morphology and mechanical advantage in the spotted hyena (*Crocuta crocuta*). *J Morphol*. 271(3):353–365.

Van Spaendonk H, Ceuleers H, Witters L, Patteet E, Joossens J, Augustyns K, Lambeir AM, De Meester I, De Man JG, De Winter BY. 2017. Regulation of intestinal permeability: the role of proteases. *World J Gastroenterol*. 23(12):2106–2123.

Watts HE, Holekamp KE. 2007. Hyena societies. *Curr Biol*. 17(16):R657–R660.

Werdelin L, Solounias N. 1991. The Hyaenidae: taxonomy, systematics and evolution. *Fossils Strata*. 30:1–104.

Westbury MV, Hartmann S, Barlow A, Preick M, Ridush B, Nagel D, Rathgeber T, Ziegler R, Baryshnikov G, Sheng G, et al. 2020. Hyena paleogenomes reveal a complex evolutionary history of cross-continental gene flow between spotted and cave hyena. *Sci Adv*. 6(11):eaay0456.

Westbury MV, Hartmann S, Barlow A, Wiesel I, Leo V, Welch R, Parker DM, Sicks F, Ludwig A, Dalén L, et al. 2018. Extended and continuous decline in effective population size results in low genomic diversity in the world's rarest hyena species, the brown hyena. *Mol Biol Evol*. 35(5):1225–1237.

Westbury MV, Le Duc D, Duchêne DA, Krishnan A, Prost S, Rutschmann S, Grau JH, Dalén L, Weyrich A, Norén K, et al. 2021. Ecological specialization and evolutionary reticulation in extant hyaenidae. *Mol Biol Evol*. 38(9):3884–3897.

Yang C, Li F, Xiong Z, Koepfli K-P, Ryder O, Perelman P, Li Q, Zhang G. 2020. A draft genome assembly of spotted hyena, *Crocuta crocuta*. *Sci Data*. 7(1):126–126.

Zhang S-Y, Herman M, Ciancanelli MJ, Pérez de Diego R, Sancho-Shimizu V, Abel L, Casanova J-L. 2013. *TLR3* immunity to infection in mice and humans. *Curr Opin Immunol*. 25(1):19–33.

Zhou C, Liu Y, Qiao L, Lan Y, Price M, Meng Y, Yang N, Yue B. 2020. Genome-wide analyses provide insights into the scavenging lifestyle of the striped hyena (*Hyaena hyaena*). *DNA Cell Biol*. 39(10):1872–1885.

Zhou C, Wang G, Yu H, Geng Y, Wu W, Tu H, Price M, Fan Z, Meng Y, Yue B. 2019. Genome-wide analysis reveals the genomic features of the turkey vulture (*Cathartes aura*) as a scavenger. *Mol Genet Genomics*. 294(3):679–692.