




Ecological Specialization and Evolutionary Reticulation in Extant Hyaenidae

Michael V. Westbury ^{1,2,*} Diana Le Duc,^{3,4} David A. Duchêne,^{2,5} Arunkumar Krishnan,^{6,7} Stefan Prost ^{8,9} Sereina Rutschmann,¹ Jose H. Grau,^{1,10} Love Dalén,^{11,12} Alexandra Weyrich ¹³ Karin Norén,¹⁴ Lars Werdelin,¹⁵ Fredrik Dalerum,^{14,16,17} Torsten Schöneberg,¹⁸ and Michael Hofreiter¹

¹Department of Mathematics and Natural Sciences, Institute of Biochemistry and Biology, University of Potsdam, Potsdam, Germany

²Section for Evolutionary Genomics, The GLOBE Institute, University of Copenhagen, Copenhagen, Denmark

³Institute of Human Genetics, University Medical Center Leipzig, Leipzig, Germany

⁴Department of Evolutionary Genetics, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany

⁵Research School of Biology, Australian National University, Canberra, ACT, Australia

⁶National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, MD, USA

⁷Department of Biological Sciences, Indian Institute of Science Education and Research (IISER) Berhampur, Odisha, India

⁸LOEWE-Center for Translational Biodiversity Genomics, Senckenberg, Frankfurt, Germany

⁹South African National Biodiversity Institute, National Zoological Garden, Pretoria, South Africa

¹⁰Amedes Genetics, amedes Medizinische Dienstleistungen, Berlin, Germany

¹¹Centre for Palaeogenetics, Stockholm, Sweden

¹²Department of Bioinformatics and Genetics, Swedish Museum of Natural History, Stockholm, Sweden

¹³Department of Evolutionary Genetics, Leibniz Institute for Zoo and Wildlife Research (IZW), Berlin, Germany

¹⁴Department of Zoology, Stockholm University, Stockholm, Sweden

¹⁵Department of Palaeobiology, Swedish Museum of Natural History, Stockholm, Sweden

¹⁶Research Unit of Biodiversity (UO-CSIC-PA), Mieres Campus, University of Oviedo, Mieres, Asturias, Spain

¹⁷Mammal Research Institute, Department of Zoology and Entomology, University of Pretoria, South Africa

¹⁸Rudolf Schönheimer Institute of Biochemistry, Molecular Biochemistry, Medical Faculty, Leipzig, Germany

*Corresponding author: E-mail: mvwestbury@gmail.com

Associate editor: Michael Rosenberg

Abstract

During the Miocene, Hyaenidae was a highly diverse family of Carnivora that has since been severely reduced to four species: the bone-cracking spotted, striped, and brown hyenas, and the specialized insectivorous aardwolf. Previous studies investigated the evolutionary histories of the spotted and brown hyenas, but little is known about the remaining two species. Moreover, the genomic underpinnings of scavenging and insectivory, defining traits of the extant species, remain elusive. Here, we generated an aardwolf genome and analyzed it together with the remaining three species to reveal their evolutionary relationships, genomic underpinnings of their scavenging and insectivorous lifestyles, and their respective genetic diversities and demographic histories. High levels of phylogenetic discordance suggest gene flow between the aardwolf lineage and the ancestral brown/striped hyena lineage. Genes related to immunity and digestion in the bone-cracking hyenas and craniofacial development in the aardwolf showed the strongest signals of selection, suggesting putative key adaptations to carrion and termite feeding, respectively. A family-wide expansion in olfactory receptor genes suggests that an acute sense of smell was a key early adaptation. Finally, we report very low levels of genetic diversity within the brown and striped hyenas despite no signs of inbreeding, putatively linked to their similarly slow decline in effective population size over the last ~2 million years. High levels of genetic diversity and more stable population sizes through time are seen in the spotted hyena and aardwolf. Taken together, our findings highlight how ecological specialization can impact the evolutionary history, demographics, and adaptive genetic changes of an evolutionary lineage.

Key words: hyena, genome, phylogenomics, genetic diversity, adaptation, comparative genomics.

© The Author(s) 2021. Published by Oxford University Press on behalf of the Society for Molecular Biology and Evolution.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

Open Access

Introduction

Originating ~25 million years ago (Ma) during the Late Oligocene (Werdelin and Solounias 1991), Hyaenidae was once a highly diverse family within the order Carnivora. During the Late Miocene, there were dozens of species occurring across Europe, Asia, Africa, and North America (Werdelin and Solounias 1991). However, this diversity has since been severely reduced to four extant genera, each of which contains only a single species, of which only one also ranges outside Africa (fig. 1A). Extant hyenas consist of three species of bone-cracking hyenas: the brown hyena (*Parahyaena brunnea*), the striped hyena (*Hyaena hyaena*), and the spotted hyena (*Crocuta crocuta*), and the insectivorous aardwolf (*Proteles cristata*). The aardwolf is unique amongst extant hyenas, not only in its feeding behavior but also in its morphology and is believed to originate from “dog-like” hyena ancestors that were less derived than the extant bone-cracking species (Werdelin and Solounias 1991). The oldest fossils showing insectivorous traits that include the vestigial, peg-like cheek dentition of the aardwolf date back approximately 2 Ma (Hendey 1974). However, aardwolf-sized postcranial material dated to approximately 4 Ma (Werdelin and Dehghani 2011) suggests that evolution toward insectivory may have begun somewhat earlier. This, and the absence of aardwolf-like fossils in the Miocene indicates that the current aardwolf morphology only evolved within the last 4 Ma and leads to the question of what underlying genetic mechanisms may have allowed the aardwolf to evolve into a successful insectivore.

The bone-cracking hyenas can further be separated into the more obligate scavenging species (brown and striped hyenas) and the spotted hyena, which largely relies on hunting but is also a facultative scavenger. These scavenging species play a vital role in maintaining the health of the ecosystem by facilitating nutrient cycling and influencing disease dynamics (Beasley et al. 2015; Benbow et al. 2015). However, despite their importance in their ecosystems, the genetic underpinnings of how these species adapted to such a potentially pathogen-rich feeding strategy remain unknown. Adaptations to this feeding behavior may include highly sensitive olfactory senses (Benbow et al. 2015) and/or specific physiological mechanisms, like specialized digestive and enhanced immune systems (Blumstein et al. 2017). In addition to a lack of knowledge on the genomic underpinnings surrounding the adaptation of these species to their current lifestyles, their joint and respective evolutionary and demographic histories are also unknown. Insights into these may help us understand how the extant species survived to the present day and also assess their resilience to future perturbations.

To this end, genomic analyses can provide profound insights into the evolutionary histories of extant species. To date, only two studies have investigated the evolutionary histories of species within the Hyaenidae based on genome data (Westbury et al. 2018, 2020). One study on the brown hyena (Westbury et al. 2018) revealed exceptionally low levels of genetic diversity without any obvious detrimental effects on

the survivability of the species, putatively caused by a long, slow, and continual decline in effective population size (N_e) over the last million years. Another study, on the spotted hyena, revealed gene flow between modern African spotted hyena and Eurasian cave hyenas, an extinct sister lineage of spotted hyenas within the genus *Crocuta* (Westbury et al. 2020). These studies set the stage for future research on the evolutionary histories of extant Hyaenidae, including whether interspecies gene flow was a widespread phenomenon within Hyaenidae, and what role demographic history played in producing the diversity seen in these species today.

Here, we set out to explore whole nuclear genomes to investigate the evolutionary histories of the remnant Hyaenidae species. We aimed to uncover whether interspecific gene flow was a widespread phenomenon within Hyaenidae, and what genomic underpinnings may have allowed the aardwolf to become a specialized insectivore and the bone-cracking species to feed on carrion. Moreover, we set out to find the current levels of genetic diversity and offer explanations to current levels of diversity through investigating the respective demographic histories of each species over the last ~2 million years.

Results

Evolutionary Relationships

We inferred a dated phylogeny to gain insight into when the extant lineages within Hyaenidae diverged from one another (fig. 1B). The aardwolf diverged first at ~13.2 Ma [95% credibility interval (CI) 8.9–18.6 Ma], followed by the spotted hyena, ~5.8 Ma (95% CI 4.5–8.0 Ma). The brown and striped hyenas diverged most recently ~4.5 Ma (95% CI 4.1–5.4 Ma). However, such interspecific divergences are often complicated by postdivergence gene flow (Campbell et al. 2018). Therefore, to test for interspecific admixture within extant Hyaenidae, we investigated the potential incongruence among species tree estimates under different models of genome evolution. The phylogenetic estimate under the multi-species coalescent model placed the aardwolf as sister to the other three species of hyenas (fig. 2A, i). However, a locus concatenation method placed the spotted hyena as sister to the three other species (fig. 2A, ii). We further examined these results using the gene and site concordance factors of the data and inferences (Minh, Hahn, et al. 2020). Analyses of concordance factors showed that a majority of the whole gene alignments support the former resolution (48% versus ≤25% for other resolutions) (fig. 2B). However, the majority of single nucleotide variants supported the latter resolution (45% versus ≤35% for other resolutions) (fig. 2C). These results and investigations into the proportion of variable sites within loci (Supplementary fig. S1, Supplementary Material online) showed that a minority, yet highly informative set of loci, support the spotted hyena as sister to the three other hyenas. Further investigations into the site-concordance factors in genes showing topology (in fig. 2A) showed relatively equal site-concordance factors for both alternative topologies (16.28% vs. 14.49%).

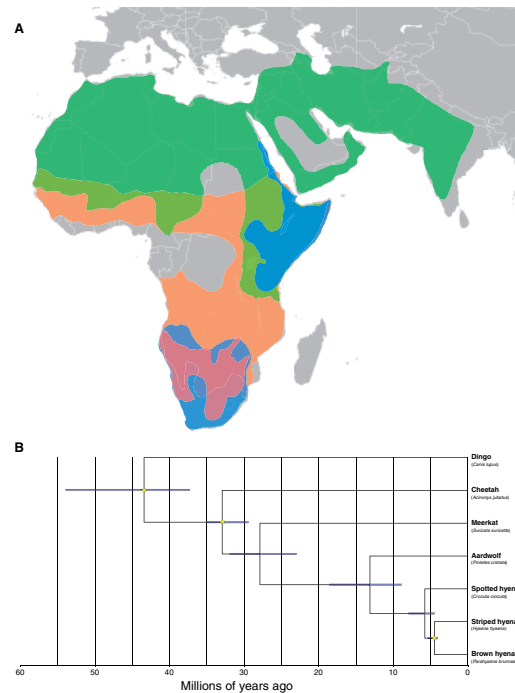


Fig. 1. Distribution ranges and dated phylogeny of extant hyenas. (A) Map showing the IUCN distribution ranges of the extant hyena species. Striped hyena—green, spotted hyena—orange, brown hyena—red, aardwolf—blue. (B) Molecularly dated phylogenomic tree built using a Bayesian relaxed-clock analysis on the ASTRAL species tree using three fossil calibrations with soft maximum bounds (indicated as yellow stars). Blue horizontal bars represent 95% credibility intervals of node times.

Additional details on events of reticulation were explored via species tree reconstruction under the multispecies network coalescent. We used a Bayesian implementation of this model that samples gene trees with their coalescent times from sequence alignments (Wen and Nakhleh 2018; Chi, Zhang et al. 2018). The analyses showed that the aardwolf originally diverged from other lineages of hyenas but underwent admixture with the ancestral lineages around the time of the split of the spotted hyena from the stem lineage of the bone-cracking hyenas (fig. 2D).

Genomic Adaptations

To better understand which genetic adaptations occurred in Hyaenidae, we performed comparative genomic analyses to uncover genes under selection as well as a gene family expansions. Out of the 9,400 1:1 orthologous genes in our data set, we uncovered 38 genes showing significant signs of positive selection in the bone-cracking lineage. In total, 35 of these had known annotations (Supplementary material table S1, Supplementary Material online). Furthermore, seven of these genes showed highly significant signs of positive selection ($P > 0.005$). We also found signs of significant positive selection in 214 genes, 184 of which had known annotations (Supplementary material table S2, Supplementary Material online), in the aardwolf lineage. Of these, 60 were highly significant. Analyses of olfactory receptor (OR) copy numbers revealed a substantial increase in both alpha and gamma family ORs in Hyaenidae (Supplementary material figs S2 and S3, Supplementary Material online). All hyenas showed at least 25–30% increase in alpha OR, and an approximately

45–50% increase in gamma OR compared with what can be seen in dogs. Furthermore, the alpha OR repertoire in hyenas was larger than that observed in the other carnivoran species analyzed here (cat, dog, and tiger). This result of a Hyaenidae-wide expansion in OR was confirmed using a subsequent de novo method implemented to avoid assembly or annotation biases (Supplementary material fig. S4, Supplementary Material online). The same analysis also found expansions in immunity-related genes in the bone-cracking lineage, and lipocalins and the UDP glucuronosyltransferase family (UGT) in the aardwolf lineage (Supplementary material fig. S4, Supplementary Material online).

Genetic Diversity

We used autosomal heterozygosity as a proxy to estimate the genetic diversity of the four hyena species. We calculated autosomal heterozygosity using two different methods. The first method was implemented using ANGSD (Korneliussen et al. 2014) to allow for comparability to previous studies (Westbury et al. 2018; Westbury, Petersen, et al. 2019). The second method employed ROHan (Renaud et al. 2019). Furthermore, ROHan also calculates runs of homozygosity (ROH) that can be used to estimate levels of inbreeding within each individual, independent of genetic diversity levels. Based on the chosen method, absolute values of average heterozygosity were different, but the relative numbers were comparable (Supplementary material tables S3, S4, S5, and S6, Supplementary Material online). However, we note that depending on the mapping reference, ROHan global heterozygosity estimates changed. When mapping the striped and

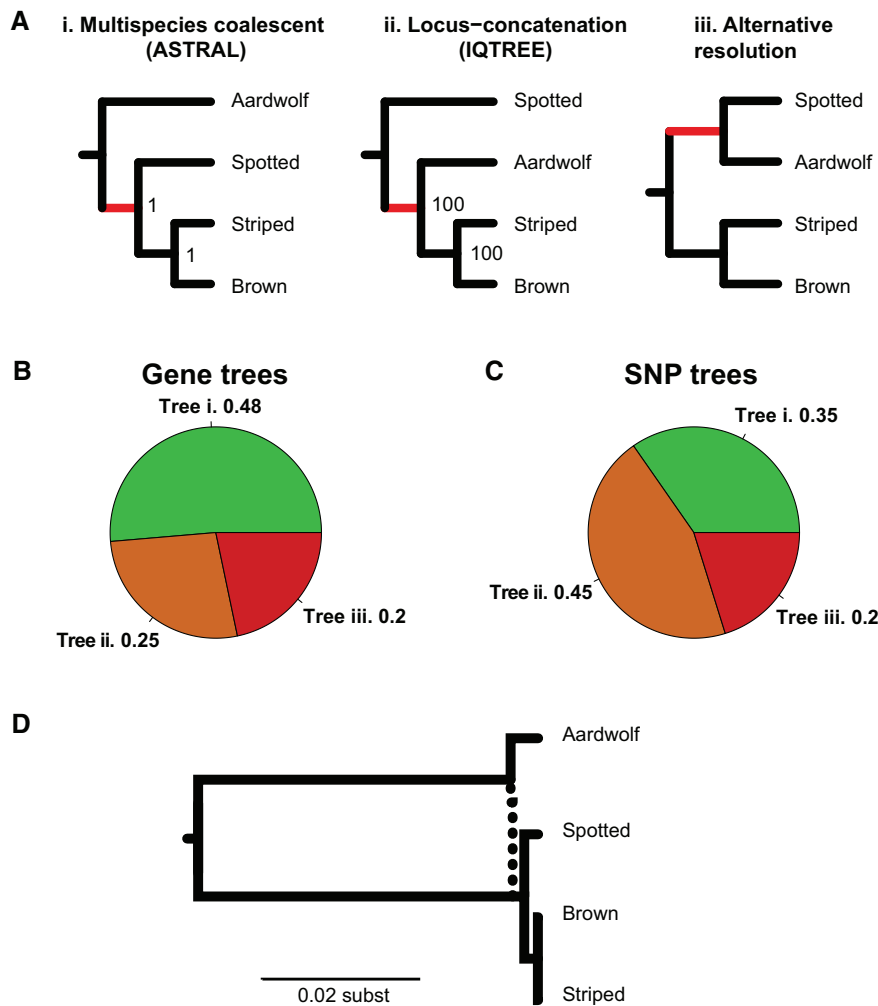


Fig. 2. Support for the resolutions of the earliest divergence among extant species of hyenas. (A) Fully bifurcating species tree estimates. Analyses assuming a fully bifurcating process support either the (i) aardwolf or (ii) spotted hyenas as the first to split from other extant lineages. Nodes in trees (i) and (ii) are labeled by their corresponding statistical support. (B) Gene concordance factors for the three alternative topologies (i, ii, and iii). Numbers show the proportion of decisive gene trees containing that branch. (C) Site concordance factors for the three alternative topologies (i, ii, and iii). Numbers show the proportion of decisive alignment sites supporting a branch in the reference tree. (D) Multispecies network coalescent estimate calculated using PhyloNet allowing for a single introgression event.

brown hyenas to the spotted hyena assembly, heterozygosity levels were higher than when mapping to the striped hyena assembly suggesting that reference genome selection may influence heterozygosity estimates. This pattern was also seen when mapping the spotted hyena to the striped hyena assembly, as heterozygosity levels were higher than when mapping to the spotted hyena assembly. However, despite this trend, all comparisons had overlapping confidence intervals.

The brown hyena had the lowest levels of heterozygosity, closely followed by the striped hyena, then the spotted hyena, and finally the aardwolf (Supplementary material table S3, Supplementary Material online). Relative to the other species included in this study, both the brown and striped hyena displayed very low levels, whereas the spotted hyena and aardwolf had medium to high levels of heterozygosity (fig. 3A). We further analyzed the distribution of heterozygosity across the genomes of the four hyena species in non-

overlapping windows of 500 kb (fig. 3B). Both the aardwolf and spotted hyena had similar broad distributions across their genomes, although the spotted hyena had a much higher percentage of 500 kb windows with less than 0.001% heterozygosity compared with the aardwolf. In contrast, both the striped hyena and brown hyena had a much narrower distributions in heterozygosity with a skew in their distributions toward 0. The skew toward 0 led to further investigation whether inbreeding may have caused the low levels of heterozygosity within the brown and striped hyenas (Supplementary material tables S4, S5, and S6, Supplementary Material online). To explore this, we calculated the percentage of the genome in ROH putatively due to inbreeding, specifying a 1 Mb window with a heterozygosity proportion of less than either 1×10^{-5} or 5×10^{-5} . When specifying a value of 1×10^{-5} , we did not find evidence for inbreeding in either the brown or striped hyena genomes. When specifying a value of 5×10^{-5} , we did find some

ROH in both the brown hyena (0.36% of the genome) and striped hyena (2.40%). However, both of these values were less than that found in the spotted hyena (5.12%). The aardwolf showed little to no signs of ROH regardless of the ROH heterozygosity value or mapping reference used.

Demographic Histories

To better contextualize the genetic diversity results, we used a pairwise sequential Markovian coalescence (PSMC) model (Li and Durbin 2011) and modeled the demographic history of each extant hyena species over the last ~ 2 million years (fig. 4). All species showed a unique pattern of demography over the last 2 million years but at the same time shared some similarities. The aardwolf showed the highest effective population size (N_e) through time relative to the other species with a relatively stable population size starting from ~ 2 Ma until ~ 500 thousand years ago (kya), where it began to increase until a sharp decrease at ~ 100 kya. The spotted hyena had the next highest N_e through time with a fluctuating N_e and a sharp decrease at ~ 100 kya. Both the brown and spotted hyenas show the lowest N_e through time. The striped hyena shows an increase in N_e at ~ 500 kya, whereas the brown hyena continued to decline. Both showed a general, slow decline from ~ 2 Ma, again followed by a sharp decrease ~ 100 kya.

Discussion

Phylogenetic analyses revealed the same topology as that found in previous studies based on mitochondrial DNA and small numbers of nuclear genes (Rohland et al. 2005; Koepfli et al. 2006; Westbury, De Cahsan, et al. 2019; Westbury et al. 2020). Our results confirm a comparatively deep divergence of the aardwolf from the other species (~ 13 Ma). This divergence date approximately coincides with the first appearance of Hyaenidae in Africa (Werdelin and Sanders 2010), tentatively suggesting that all living Hyaenidae today can be traced to an African origin. However, unlike previous studies, we found a more recent divergence among the bone-cracking hyenas (~ 6 Ma) (fig. 1), albeit with overlapping credibility intervals. Our divergence date aligns much closer to the first appearances of canids in Africa ~ 8 Ma (foxes) (de Bonis et al. 2007) and *Eucyon* (the precursor to *Canis*), which appeared possibly as early as ~ 6 Ma ago (Werdelin et al. 2015). The appearance of these competitors may have been the catalyst driving the evolution of the bone-cracking hyenas into the bone-cracking niche and allowed the persistence of these species while all dog-like hyenas (excluding the aardwolf) went extinct. Similar to the timing of the evolution of the bone-cracking hyena morph, we suggest that the adaptation to insectivory of the aardwolf could also have been driven by competition from the expanding dog-like clades of canids. The resulting ecological specialization may have allowed the aardwolf lineage to persist while all other dog-like hyenas were outcompeted.

Subsequent phylogenomic analyses supported an introgression event from the aardwolf lineage into lineages close to the divergence of the bone-cracking hyena lineages, and most likely disproportionately with populations of the stem

ancestor of striped and brown hyenas (fig. 2). While this appears likely to have resulted from an introgression event from the aardwolf lineage into the stem ancestor of the striped and brown hyenas, other phenomena leading to this pattern could be (1) introgression between the spotted hyena lineage and an extinct unsampled sister lineage to the four hyenas studied here or (2) gene flow between the proto-aardwolf lineage to the stem of the bone-crushing lineage, with the spotted hyena losing more of the shared alleles through incomplete lineage sorting (ILS). As we see such high discordance factors in our phylogenomic analyses, a combination of the above-mentioned phenomena may have been at play during the evolutionary history of extant Hyaenidae. This result is also consistent with the relatively broad credibility interval of the timing of the split of the aardwolf from other lineages of hyena (fig. 1) and the short branches in coalescent time units separating lineages of hyenas.

The deep divergence of the aardwolf and its substantially different ecology suggest introgression to be unlikely. However, considering the divergence confidence intervals, i.e. 8.9 Ma for the aardwolf (lower 5%) and 8 Ma for the spotted hyena (upper 5%), it becomes more realistic. Moreover, there are no traces of any of the current aardwolf adaptations to insectivory prior to 4 Ma despite the extensive fossil record of Miocene hyenas. These observations suggest that introgression occurred prior to the split of the brown and striped hyenas ~ 4.5 Ma. This would also suggest that, at the time of gene flow, these species may not have been as evolutionarily or even ecologically distinct as they are today and the specific insectivorous adaptations of the aardwolf did not originate until the aardwolf lineage had ceased exchanging genetic material with the bone-cracking lineage.

The evolution of the aardwolf morphology within the last 2–4 Ma leads to the question of what underlying genetic mechanisms may have allowed the aardwolf to evolve into a successful insectivore subsisting on a diet almost exclusively consisting of termites (Matsebula et al. 2009). Investigations into the genes showing highly significant signs of positive selection in the aardwolf lineage revealed a number of genes putatively linked to aardwolf specific adaptations. We found five genes with known functions related to craniofacial development, putatively key in the formation of the unique aardwolf skull. These genes code for glycyl-tRNA synthetase (*GARS*), guanosine monophosphate reductase (*GMPR*), and stress-induced phosphoprotein 1 (*STIP1*), smoothed (*SMO*), and 3'-phosphoadenosine 5'-phosphosulfate synthetase 2 (*PAPSS2*). *GARS* loss-of-function gene variants have an impact on the developmental phenotype in humans including growth retardation, a large calvaria, and a high-arched palate (Oprescu et al. 2017). Less is known about *GMPR*, since fewer studies have focused on this gene. However, a 1 Mb de novo interstitial deletion in 6p22.3 chromosomal region, which contains the *GMPR* gene, was shown to cause some craniofacial malformations, suggesting that there may be a link between this gene and craniofacial development (Di Benedetto et al. 2013). A large genomic screen identified *STIP1* to be in gene sets responsible for the branchial arch

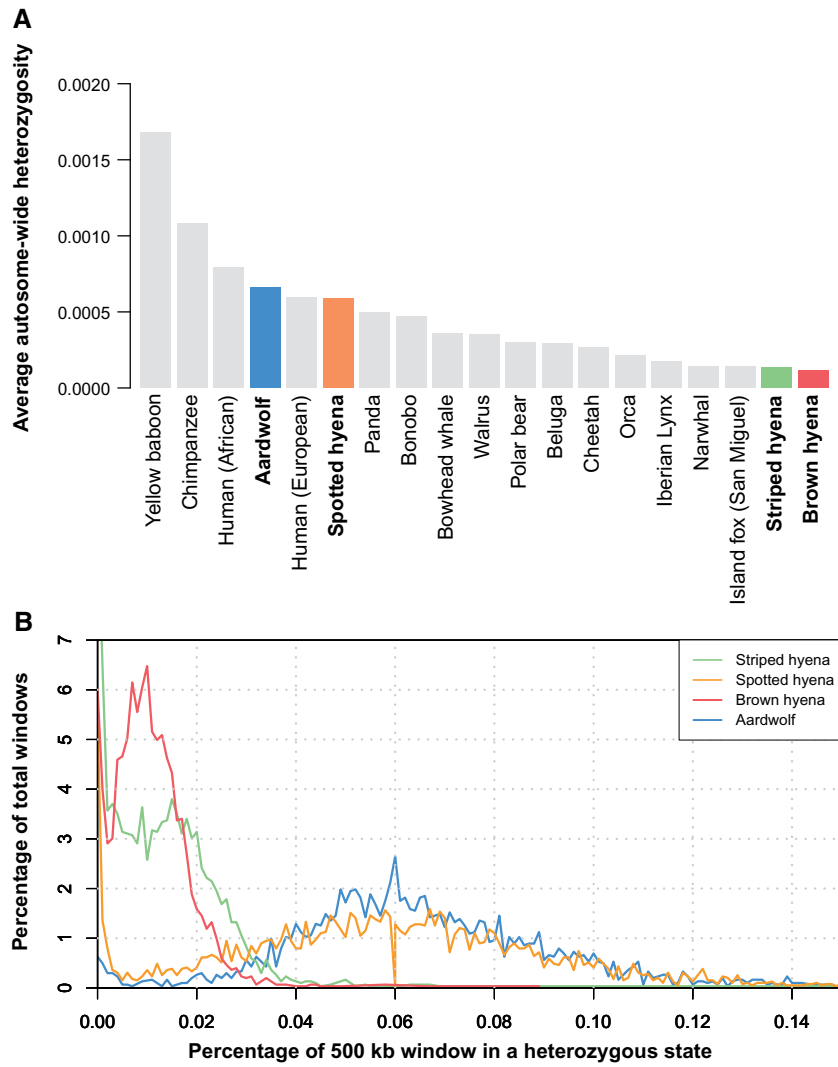


FIG. 3. Comparative levels of heterozygosity within Hyaenidae. (A) Average autosome-wide heterozygosity estimates for a number of mammalian species. (B) Heterozygosity across the autosomes of the four Hyaenidae species calculated using 500-kb sliding windows.

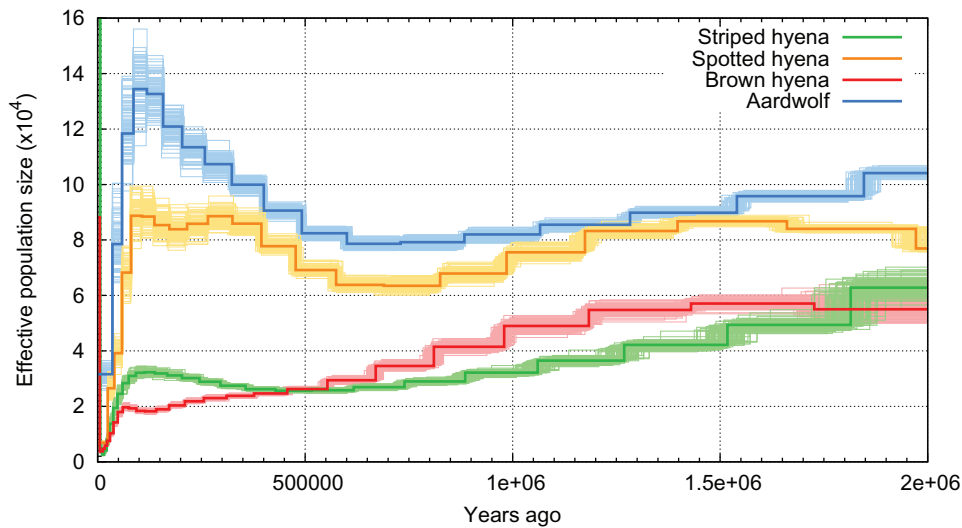


FIG. 4. Demographic history of the four Hyaenidae species over the last 2 Myr, estimated using the pairwise sequentially Markovian coalescent (PSMC) model.

formation, also suggesting a role of this gene in craniofacial development (Fowles et al. 2003). Conditional null *SMO* mice had severe malformations in craniofacial structures developing from neural crest-derived elements (Jeong et al. 2004). Mice having two copies of an autosomal recessive mutant allele of *Papss2* have been characterized as having foreshortened limbs, a short stout tail, and a complex craniofacial phenotype (Ford-Hutchinson et al. 2005). Furthermore, we also note one gene involved in skin barrier function (*DSC1*) (Chidgey et al. 2001). Skin barrier function may have been a key adaptation allowing aardwolves to transition to the insectivorous niche as the defense secretions of soldier *Trinervitermes* termites are known to be highly toxic, and despite this, the majority of the aardwolf's diet consists of *Trinervitermes* (Richardson and Levitan 1994). Finally, evidence for an expansion in the lipocalin and the UDP glucuronosyltransferase (*UGT*) gene families suggests important roles in the evolution of the aardwolf. Lipocalins produce major allergens that can induce severe anaphylactic reactions in humans when bitten by hematophagous insects (Paddock et al. 2001) while pseudogenization of *UGTs* in the domestic cat are linked to an inability to process drug toxins (Court and Greenblatt 2000). Although currently speculative, the aardwolf may have recruited these gene families as a defence mechanism against the termite toxins, enabling it to feed despite being bitten.

Our results suggest that the ability of the three scavenging hyena species to feed on potentially infectious carrion as well as break down large quantities of bone was facilitated by genetic adaptations to the immune system and digestion. A heightened immunity would allow these species to feed on rotten flesh containing many potential pathogens ensuring a relatively reliable food source (Blumstein et al. 2017). When fresh food was scarce, individuals that could survive on what was available may have had a competitive advantage over others. As immune-relevant genes are commonly found to be under positive selection due to their generally rapid evolutionary rate (Kosiol et al. 2008), it is difficult to discern whether these genes were flagged because of their faster background evolutionary rates or whether they are truly linked to adaptation to scavenging. However, our findings of a number of immune related gene families being expanded in the bone-cracking hyenas does add weight to the hypothesis that selection on immunity was key in the adaptation to scavenging.

Despite these results, investigations into genes showing the most significant signs of positive selection, suggest that genes involved in the gastrointestinal system were under stronger selection than immune genes in general; in particular *ASH1L*, *PTPN5*, *PKP3*, and *AQP10*. In mice, *ASH1L* has been shown to be significantly downregulated in the colon during T cell-mediated colitis (Fang et al. 2012), and also has a potential connection to T cell-mediated autoimmune diseases. The *PTPN5* gene has been linked to viral gastritis in humans, an infectious disease of the gastrointestinal system that involves inflammation of the stomach lining caused by viruses (Genecards.org). Studies have shown *PKP3* to undergo molecular processes associated with inflammatory responses in the gut (Sklyarova et al. 2015). Finally, *AQP10* appears to

contribute to liquid transport in the gastrointestinal tract (Li et al. 2005). All four of these genes suggest that specific gastrointestinal adaptations were vital in allowing hyenas to adapt to feeding on carrion and digesting bone. Moreover, in addition to their functions in the gut, variants of *PTPN5* have been shown to also cause increased bone mineralization in mice (<https://www.mousephenotype.org/>) and therefore may have also played a role in strong tooth/jaw development required for bone cracking. Finally, *PKP3* has been shown to also play a role in inflammation of the skin (Sklyarova et al. 2008), an adaptation that may have been vital to avoid disease while feeding on carrion.

The most parsimonious explanation for the observed expanded OR repertoire in all extant species would be that it was also expanded in their common ancestor, and consequently that an acute sense of smell may have been a driving force allowing for the survival of the family since the Miocene. This result was expected in the bone-cracking species due to their carrion feeding lifestyle as it could also be assumed that to find carrion before competitors, these hyenas would have required an acute sense of smell. However, the result was somewhat unexpected in the aardwolf as, on the surface, it could be argued that the aardwolf does not need as acute a sense of smell as the other species. However, the retention of this expanded OR repertoire may have been put to use in other ways (e.g. territory marking and mating, Richardson 1991; Marneweck et al. 2015), or it could have been maintained if there were no strong selective pressures against it.

While both the aardwolf and spotted hyena showed relatively high levels of diversity compared with a number of other mammalian species, both the striped and brown hyenas seem to have exceptionally low levels of genetic diversity. Similar to the brown hyena (Westbury et al. 2018), this low level of diversity in the striped hyena was likely not caused by recent inbreeding, inferred from the low occurrence of ROH (Supplementary material tables S4, S5, S6, Supplementary Material online). These results raise the question which demographic changes through time led to the different present day genetic diversity levels in these species. Aardwolf and spotted hyena not only show relatively high levels of N_e through time, surprisingly, in light of their different ecologies, they also display similar fluctuations in population size, characterized by a slow decrease from ~ 1.5 Ma followed by a recovery/increase ~ 500 kya. The striped hyena on the other hand shows a gradual decrease in N_e from about 2 Ma until a slight increase/plateau at a low level ~ 500 kya. This gradual decrease is similar to what is seen in the brown hyena and may explain the ability of the species to retain such low levels of diversity without a detectable signal of inbreeding. We suggest that the parallel trajectories of genetic diversity in the striped and brown hyenas may be related to their similar social and ecological characteristics. Partly relying on carrion, being mostly solitary (Watts and Holekamp 2007), and being found in arid parts of the Afrotropical realm to the exclusion of the wetter equatorial zone (fig. 1) may have resulted in low carrying capacities and consequently limited population sizes.

The striped hyena has the largest geographic distribution of all hyena species, ranging from eastern Africa to India

(AbiSaid and Dloniak 2015). Very little is known about the range wide population structure and population densities of this species, but our results suggest that they may persist at very low densities across their range. Unfortunately, the genome used in this study was obtained from a sample of a captive individual (Tierpark Berlin), wherefore it is unknown whether the genetic diversity of this animal reflects that of a wild individual. It was a minimum fourth generation zoo individual, with ancestors from Yerewan zoo, Bratislava zoo, and an unknown location. The potential mixed ancestry of this individual may provide a reason as to why there are no signs of inbreeding but if that was the case, the low heterozygosity found in this individual would be even more astonishing, with heterozygosity levels potentially even lower in wild bred individuals.

While the demographic trajectories between the high and low diversity species pairs were largely different through time, all four species showed a dramatic decline in $N_e \sim 100$ kya. A similar decline was also reported in a large dataset of African ruminant genomes and was proposed to have coincided with the expansion of humans and the resultant hunting of these large herbivores (Chen et al. 2019). Moreover, this was previously reported for the spotted hyena and was hypothesized to be linked to increased competition with humans and exacerbated by a decline in prey availability (Westbury et al. 2020), which could also hold true for brown and striped hyenas. However, we also see a dramatic decline at approximately the same time in the aardwolf, a specialized insectivore expected to be in no direct competition with early humans ~ 100 kya. Even though humans may have hunted aardwolves, a perhaps more probable scenario would be that there were other environmental mechanisms besides humans (e.g. climate induced environmental change) leading to large decreases in N_e of African species ~ 100 kya.

Materials and Methods

Data Acquisition

Aardwolf Data Generation

To generate a mapped aardwolf genome, we extracted DNA from a single aardwolf. The individual was captured for an ecological research project on Benfontein game reserve outside of Kimberley, central South Africa (28°52' S, 24°50' E) (Marneweck et al. 2015). The reserve covers approximately 11,400 ha and consists mainly of dry Karoo shrubland, arid grasslands, and Kalahari thornveld (Dalerum et al. 2017). The animal was initially captured in 2008 under a permit from the animal care and use committee of the University of Pretoria (EC031-07) and permits from the provincial government in the Northern Cape (FAUNA 846/2009, FAUNA 847/2009). Genetic samples consisted of a small piece of skin collected at the ear tip which was stored in 95% ethanol at -20°C.

We extracted DNA from the tissue sample using a Zymo genomic DNA clean and concentrator extraction kit (Zymo Research, USA), following the manufacturer's protocol. DNA extracts were built into Nextera Illumina sequencing libraries (Inqaba Biotec, Pretoria) and sequenced on an Illumina HiSeq

X at the National Genomics Institute (NGI) Stockholm, Sweden using 2x150bp paired-end reads.

Previously Published Data

For use in our analyses, we downloaded the previously published nuclear genome assemblies of the striped hyena (Genbank accession: GCA_003009895.1) (Westbury et al. 2018) and the spotted hyena (Genbank accession: GCA_008692635.1) (Yang et al. 2020). We further downloaded the respective Illumina raw reads from each assembly as well as the raw reads from the resequenced brown hyena (Genbank accession: SRS2398897), which was previously studied by mapping to the striped hyena (Westbury et al. 2018).

Mapping of Raw Illumina Reads

Raw reads were all treated comparably before being mapped to a specific reference genome, which varied depending on the analysis. We used Cutadapt v1.8.1 (Martin 2011) to trim Illumina adapter sequences from the ends of reads and remove reads shorter than 30 bp. We merged overlapping read pairs using FLASH v1.2.1 (Magoč and Salzberg 2011) and default parameters. We mapped the trimmed merged and unmerged reads to the respective reference sequences using BWA v0.7.15, the mem algorithm, and default parameters (Li and Durbin 2009). We further processed the mapped reads using SAMtools v1.3.1 (Li et al. 2009) to remove duplicates and reads of low mapping quality (< 30).

Consensus Sequence Construction

For the comparative genomic analyses, we used the two already assembled genomes (striped hyena and spotted hyena) and the two resequenced genomes mapped to the striped hyena (brown hyena and aardwolf). To decide which resultant mapping file to use, we mapped the brown hyena and aardwolf to both the striped hyena and spotted hyena assemblies and compared mapping statistics. As both species mapped more efficiently to the striped hyena genome (aardwolf - 457,078,015 reads/72,698,723,765 bp mapped to striped hyena vs. 454,576,330 reads/72,003,685,729 bp mapped to spotted hyena, brown hyena - 635,096,073 reads/98,210,957,289 bp mapping to striped hyena vs. 613,078,338 reads/94,209,031,605 bp mapped to spotted hyena), we built consensus sequences of these for further comparative and phylogenomic analyses. To build the consensus sequences we used a consensus base call approach (-doFasta 2) in ANGSD v0.913 (Korneliusson et al. 2014) and specified the following parameters: minimum mapping and base quality of 25 (-minQ 25 -minMapQ 25), only consider reads mapping to a single location (-uniqueOnly 1), and remove low quality secondary alignments (-remove_bads 1).

Spotted Hyena Transcriptome

For the purpose of improving downstream whole genome annotations, we also included the assembled transcriptome of a single spotted hyena individual. Three tissues (brain, olfactory tissue and testis) of a male juvenile spotted hyena were sampled postmortem, cooled, and frozen. From total RNA samples poly(A)-RNA was isolated, converted into

cDNA by random-priming and cDNA libraries prepared using Illumina TrueSeq adapters by vertis Biotechnologie AG (Freising, Germany). Libraries were paired-end sequenced (PE 2 × 100 bp and PE 2 × 150 bp) at the Friedrich Löffler Institute (Jena, Germany) on the Illumina Genome Analyzer Iix.

We trimmed adapter sequences from the raw Illumina reads using Cutadapt and quality trimmed the 3' end with a quality threshold of 23 and a minimum length of 35 bp using sickle (Joshi and Fass 2011). We applied the FRAMA pipeline (Bens et al. 2016) to assemble the trimmed RNA-seq data to an annotated mRNA assembly, incorporating Trinity (Grabherr et al. 2011) as the assembler. The domestic cat transcriptome served as reference for transcript gene symbol assignment, based on best bidirectional BLASTn hits against CDSs. The resultant orthologous reference transcript with the best aligned hit was used for scaffolding, CDS inference, and splitting of fusions. As output, we retrieved the resultant transcript sequences in the 5' to 3' orientation.

Repeat Masking and Gene Annotation

Next, we carried out repeat and gene annotations on the four genomes. We first masked repeats in each genome using a combination of ab initio repeat finding and homology-based repeat annotation using RepeatModeler (<http://www.repeat-masker.org>) and RepeatMasker (<http://www.repeatmasker.org>), respectively. For homology-based repeat annotation we used the mammal repeat consensus sequences from Repbase (Bao et al. 2015). During this step we did not mask simple repeats beforehand to improve mapping during the homology-based annotation. We annotated genes using ab initio gene prediction, as well as protein and transcriptome homology using the pipeline Maker2 (Holt and Yandell 2011). Simple repeats were soft-masked using Maker2 (Holt and Yandell 2011), to allow for more efficient mapping during gene annotation. Ab initio gene prediction was carried out using SNAP (Korf 2004) and Augustus (Stanke and Waack 2003). For the protein homology-based annotation step, we combined proteins from the previously annotated domestic cat (*Felis catus*; GCF_000181335.2) and domestic dog (*Canis lupus familiaris*; GCF_000002285.3). We further used the spotted hyena transcriptome for the homology-based annotation step (provided to Maker using the parameter “est=” for the annotation of the spotted hyena, and “altest=” for all other hyena species).

Species Tree and Molecular Dating

For producing a dated phylogenetic tree, we downloaded the assembled transcripts of two additional species from within Feliformia, cheetah (*Acinonyx jubatus*, GCF_003709585.1), and meerkat (*Suricata suricatta*, GCF_006229205.1), and the dingo (*Canis lupus dingo*, GCF_003254725.1) as outgroup. We searched for 1:1 orthologous gene sequences found in all four hyena genomes, as well as the three outgroup species using ProteinOrthov5.11 (Lechner et al. 2011). Preliminary coding sequence alignments were built for each ortholog using MACSE v2.03 (Ranwez et al. 2018). Fast trees were built using maximum likelihood and the GTR + Γ substitution model in

IQTREE v2.0 (Minh, Schmidt, et al. 2020). These trees were fed to TreeShrink (Mai and Mirarab 2018) for identifying any sequences that could mislead branch length estimates due to difficulties in assembly, annotation, or alignment. Orthologs were aligned again by excluding the flagged taxa. Codons with more than 50% missing taxa were excluded, and all orthologs were checked by eye for major alignment issues. Possible violations to the assumptions of the GTR + Γ model were tested using PhyloMAad (Duchêne et al. 2018) and IQTREE (Naser-Khdour et al. 2019), and any loci flagged as model-inadequate were excluded from further analyses. This processing led to 1,219 loci with >1.7 M aligned sites.

Using the filtered sequence alignments, we implemented two approaches to estimate the species tree of hyenas. We first estimated phylogenetic trees for each locus using a maximum likelihood and the best-fitting substitution model from the GTR+F + Γ +I family (Kalyanamoorthy et al. 2017) using IQTREE (Nguyen et al. 2015). Branch supports were calculated using an approximate likelihood ratio test (Anisimova and Gascuel 2006). These trees were used for subsequent species tree estimation under the multispecies coalescent in ASTRAL v5.6 (Chao, Zhang et al. 2018), with analyses replicated using either the raw locus trees or by collapsing branches with statistical support <50. We also concatenated all loci to perform a supermatrix approach assuming that gene tree discordance is entirely due to the stochastic inference error associated with having a finite sample size. This analysis was partitioned by locus with independently selected substitution models and assuming proportional variation in branch lengths across loci (Duchêne et al. 2020).

Molecular dating analyses were performed on the ASTRAL species tree using three fossil calibrations with soft maximum bounds. We constrained the age of the common ancestor of the brown and striped hyena to occur between 4.05 Ma based on the earliest *Parahyaena* fossil (*P. howelli*) (Werdelin 2003; Werdelin and Dehghani 2011; Werdelin and Manthi 2012) and 5.2 Ma based on the most recent putative *Hyaena/Parahyaena* ancestor, *Ikelohyaena abronia* (Werdelin et al. 1994). A second constraint was placed on the Hyaenidae-Felidae split to occur between 29 and 35 Ma, based on the reasoning described by Barnett et al. (Barnett et al. 2005). Lastly, we constrained the basal divergence of the carnivore clade, assuming it occurred prior to 37.1 Ma based on the age of the fossil *Hesperocyon gregarius* (Meredith et al. 2011).

Bayesian dating analyses were performed using MCMCtree in PAML v4.8 (Yang 2007), with improved efficiency by implementing approximate likelihood computation (Thorne et al. 1998). Loci that were discordant with the species tree topology were excluded from dating analyses to avoid violation of the tree prior (Angelis and Dos Reis 2015) and misleading branch length estimates (Mendes and Hahn 2016). The remaining loci were partitioned into each of the three codon positions. The molecular evolution of each codon position was described by a GTR + Γ substitution model and an uncorrelated gamma prior on rates across lineages, while a birth-death process was used as the prior for the branching times. We estimated the posterior distribution using Markov chain Monte Carlo (MCMC) sampling. We drew MCMC

samples every 1×10^3 steps over 1×10^7 steps, excluding a burn-in phase of 1×10^6 steps. Convergence to the stationary distribution was verified by comparing the parameter estimates of two independent runs. All parameters were found to have effective sample sizes above 1000, as estimated using the R package coda (Plummer et al. 2006).

Additional analyses were performed to disentangle any differences between species tree estimation methods (multi-species coalescent versus supermatrix ML method). We estimated the frequency of gene-trees that included each of the tree-quartets observed in the estimated species-trees, also known as gene concordance factors (Baum 2007). Similarly, site concordance factors measure the proportion of sites that support quartets estimated in the species tree (Minh, Hahn, et al. 2020). These two are measures of the decisiveness of the data for a particular phylogenetic resolution, and provide a comprehensive description of disagreement across the data. Concordance factors can also be compared with the support for the two alternative resolutions of each quartet, known as the two discordance factors. We obtained gene- and site concordance factors using IQTREE (Minh, Hahn, et al. 2020) from the tree estimated using the multispecies coalescent, after confirming that the concordant and discordant quartets of this estimate also included all the resolutions present in the species tree estimate obtained from concatenation of loci.

We further explored whether results may be driven by reference bias due to mapping the aardwolf data to the striped hyena assembly. We calculated the proportion of variable sites in each gene tree corresponding to two different Hyaenidae topologies (i.e. (((Striped, Brown), Spotted), Aardwolf), or (((Striped, Brown), Aardwolf), Spotted). We also repeated the gene- and site concordance factors analysis only using the gene tree that showed the (((Striped, Brown), Spotted), Aardwolf) topology.

In order to more comprehensively explore the history of the hyena lineages, we performed a Bayesian species tree estimation allowing for an event of reticulation under the multispecies network coalescent model (Wen and Nakhleh 2018; Chi, Zhang et al. 2018). We sampled gene trees and their coalescent times from the full set of locus alignments as implemented using the MCMC_SEQ function of PhyloNet v3.8 (Wen et al. 2018), including exclusively sequences from the four species of hyenas. The posterior distribution was estimated using Markov chain Monte Carlo (MCMC) sampling with samples drawn every 5×10^4 steps across 10^8 steps. A burn-in phase of 2×10^7 steps was excluded, and convergence to the stationary distribution was verified by confirming that all parameters had effective sample sizes above 200, using the R package coda (Plummer et al. 2006). We verified that results were identical across two independent runs, and reported the maximum a posteriori (MAP) network from the output.

Sex Chromosome Alignments

For downstream genetic diversity and demographic history analyses, we first needed to determine which scaffolds were most likely autosomal in origin. To do this we found putative sex chromosome scaffolds for the two de novo assembled

genomes and removed them from future analyses. We found putative sex chromosome scaffolds by aligning the assemblies to the domestic cat X (Genbank accession: NC_018741.3) and human Y (Genbank accession: NC_000024.10) chromosomes. Alignments were performed using satsuma synteny (Grabherr et al. 2010) and utilizing default parameters.

Ka/Ks Calculation

To determine branch specific selection in hyenas we estimated ω -values (Ka/Ks substitution ratios) for 9,400 1:1 orthologs in hyenas (aardwolf, spotted hyena, striped hyena, and brown hyena) and the domestic cat [Ensembl v90 annotation (Zerbino et al. 2018)]. We employed a previously described strategy (Le Duc et al. 2015) and performed selection analysis on 1,319 complete genes that did not contain any frameshift indels in the alignment and the longest stretch of at least 200 uninterrupted aligned bases from the other 8,081 genes. Briefly, we used the CODEML program under a branch model (Yang 2007) twice independently, once setting the aardwolf as the foreground branch, and the other with all bone-cracking species (spotted, striped, and brown hyenas) as foreground branches. This model was compared via a likelihood ratio test (1 degree of freedom) to the one-ratio model (model = 0, NSsites = 0) used to estimate the same ω ratio for all branches in the phylogeny. To access functionality of the genes under positive selection, we further looked into the associated MGI phenotypes (Genecards.org) of genes showing highly significant signs of positive selection ($X^2 > 7.9$, $p = 0.005$) in these lineages.

Olfactory Receptor Expansions

To estimate the “olfactory ability” of extant Hyaenidae, we curated the OR repertoire in the hyena genomes using established methodologies, as described earlier (Niimura 2013a, 2013b) and as performed in a previous study (Le Duc et al. 2015). To this end, we employed a multi-step process as follows: (1) Functional ORs of various mammalian genomes (dog, mouse, opossum, horse, and human) were downloaded from previous studies (Niimura and Nei 2007; Niimura 2009; Niimura et al. 2014, 2018). These functional ORs were utilized to generate manually curated alignments (of at least 40% identity) that served as seeds to build discrete HMM profiles using the HMMER package (Finn et al. 2011). These HMM profiles were utilized for standalone HMM searches against the curated proteome datasets of all Hyaenidae. All recovered hits within the default inclusion threshold (0.01) were considered as putative ORs. (2) Conversely, the proteome datasets of all Hyaenidae were utilized for standalone Pfam search (El-Gebali et al. 2019) and RPS-BLAST search as implemented in the CDD database (Lu et al. 2020). All sequences that recovered the 7tm_4 domain (HMM profile for the olfactory domain in the Pfam database, Pfam-ID: PF13853) and cd13954 (PSSM for the olfactory domain in CDD database) as their best scoring hits in the respective searches, were considered as putative ORs. (3) Putative ORs obtained from steps 1 and 2, and functional ORs downloaded from previous studies were clustered using CD-HIT (Huang et al. 2010) into groups at 40% sequence identity threshold and selected

candidate sequences from each group were utilized as seeds (diverse starting points) for several standalone TBLASTN searches (cut-off E -value of 1×10^{-9}) against individual genomes of all Hyaenidae. These searches recovered additional hits and ensured the recovery of hidden or missed ORs from any unannotated nucleotide sequences in the gene prediction/annotation pipelines. (4) A nonredundant dataset was obtained from steps 1–3 using CD-HIT (97% identify), and these putative ORs were first reviewed using PSI-BLAST searches (Altschul et al. 1997) against the NCBI-NR database and were further verified through phylogenetic analysis using nonolfactory Class-A GPCRs as outgroup (Niimura 2013a, 2013b). Protein sequences of all hyenas that recovered 7tm_1 (Pfam-ID: PF00001) as their best hit in the standalone Pfam search were obtained and categorized as Class-A non-olfactory GPCRs and utilized for Maximum-Likelihood (ML) trees computed using the IQ-TREE software (Nguyen et al. 2015). In all ML trees, the putative ORs formed an unambiguous monophyletic clade with high bootstrap support (>90%), distinct from the non-OR GPCRs.

To assign sub-type (Type-I and Type-II ORs) and subgroup level classification (α , β , γ , etc.) (Niimura 2009) of all identified ORs from each hyena genome, BLASTP searches were performed against a curated standalone BLAST database comprising previously classified ORs from vertebrate genomes (Niimura and Nei 2007; Niimura 2009; Niimura et al. 2014, 2018). The database included all identified and classified vertebrate ORs from previous studies, and sequences were tagged based on their category (such as α , β , and γ). Based on the BLASTP searches, all ORs were putatively categorized into each group, and the classification was further substantiated using phylogenetic comparisons.

For comparative analysis of ORs among carnivoran (dog, cat, and tiger) and mammalian genomes, we utilized the previously classified and functional ORs as aforementioned. The cat and the tiger genomes (for which the previous OR annotation is unavailable) were downloaded from Ensembl (*Felis catus*_8.0: GCA_000181335.3, release-92) and NCBI databases (*Panthera tigris*: GCA_000464555.1), respectively, and ORs were identified using the above procedures. To provide a comparative perspective on the OR expansion in the extant Hyaenidae relative to other vertebrates, we constructed an ML tree using full-length intact type-1 ORs (α , β , and γ group ORs) from all analyzed hyenas, carnivoran (dog, cat, and tiger) and other mammalian genomes (human and mouse). The best-fit substitution model for the curated alignment of ORs was estimated using ModelFinder (Kalyaanamoorthy et al. 2017) incorporated in the IQ-TREE software. ML tree topologies were derived using the edge-linked partition model as implemented in the IQ-TREE software, and branch supports were obtained using the ultrafast bootstrap method (1000 replicates) (Minh et al. 2013).

De Novo Gene Repertoire Expansion Analysis

To cross validate genes under positive selection and the OR repertoire expansions in hyenas, we conducted a secondary de novo approach to determine gene families with putative repertoire expansions. This was done to avoid the influence of

differing assembly or annotation quality. For this we blasted ORFs found in the raw sequencing reads against the dog proteome (Uniprot ID: UP000002254). Raw forward reads of six carnivores [*Canis lupus* (SRA accession code: SRR8926752), *Panthera tigris* (SRA accession code: SRR5591010), *Ursus arctos* (SRA accession code: SRR830337), *Suricata suricatta* (SRA accession code: SRR11434616), *Paradoxurus hermaphroditus* (SRA accession code: SRR11431891), *Cryptoprocta ferox* (SRA accession code: SRR11097184)] and our hyenas were quality checked, converted to fasta, and all ORFs with a minimal length of 90 were extracted and blasted to the dog proteome (Supplementary material fig. S5, Supplementary Material online). We additionally included hyena genomic data for training the downstream machine learning model. These included *Parahyena brunnea* (concatenation of multiple low coverage genomes: SRA accessions: SRR5886631, SRR5886634, SRR5886635, SRR5886638, SRR5886630, and SRR5886636), *Hyaena hyaena* (SRA accession: SRR11430567), and *Crocota crocuta* (SRA accessions: Namibia: SRR9914662, Ghana: SRR9914663). ORF extraction was performed using EMBOSS getorf (Rice et al. 2000) and protein to protein blasting was done using BLASTp (Nguyen and Lavenier 2009) reporting hits with an e -score over 1×10^5 . The resulting alignment tables were processed in R to remove secondary hits of each read, and normalized by the size of each library. The average size and coverage of the sequencing of the libraries was calculated, using the “estimate size factors” implemented in the DESeq2 R package (Love et al. 2014), on a set of ~ 170 highly conserved BUSCO (Waterhouse et al. 2018) genes in the metazoa ortholog set (Supplementary material figs S6 and S7, Supplementary Material online). The normalized gene set was analyzed with the machine learning randomforest algorithm (Liaw and Wiener 2002) (R package randomforest with option ntree = 100,000 and otherwise default options) to find expanded gene repertoires unique to the Hyaenidae, as well as to the aardwolf, and to the bone-cracking lineage. The 200 genes showing the highest levels of differentiation were further analyzed for GO terms.

Genetic Diversity

We used two different methods to estimate the heterozygosity of the four hyena genomes included in our study. To be able to directly compare to previously published heterozygosity results from a wide range of mammalian species (Westbury et al. 2018; Westbury, Petersen, et al. 2019), we followed the parameters first set out by Westbury et al. 2018. In brief, this was performed using ANGSD on the mapped bam files subsampled down to 20x (-downsample parameter). We applied the following filters, calculate genotype likelihoods using the SAMtools algorithm (-GL 1), -setMinDepthInd 5, -minmapq 25, -minq 25, -uniqueonly 1, adjust quality scores around indels (-baq 1), produce a folded SFS as output (-fold 1), and remove all scaffolds of less than 1 Mb and that aligned to the sex chromosomes (-rf command). We performed this using all species mapped to the striped hyena assembly apart from the spotted hyena which was mapped to the spotted hyena assembly. We further

calculated genome-wide heterozygosity and runs of homozygosity (ROH) using ROHan (Renaud et al. 2019). We did this for all species twice independently, once with all genomes mapped to the striped hyena assembly and once mapped to the spotted hyena assembly. We ran ROHan using default parameters that specify a window size of 1 Mb and ROH if said window has an average heterozygosity of less than 1×10^{-5} . We additionally reran the ROHan analysis specifying an ROH cutoff of 5×10^{-5} with the aardwolf mapped to both the spotted and striped hyena assemblies, spotted hyena mapped to the spotted hyena assembly, and both the brown and striped hyena mapped to the striped hyena assembly.

Demographic History

To investigate the respective demographic histories of each species, we performed a PSMC model (Li and Durbin 2011) on the diploid genomes of the four hyena species. For this, we used all species mapped to the striped hyena assembly apart from the spotted hyena that was mapped to the spotted hyena assembly. We called diploid genome sequences using SAMtools and BCftools (Narasimhan et al. 2016), specifying a minimum quality score of 20, a minimum coverage of 10, and a maximum coverage of 100. We removed scaffolds found to align to sex chromosomes in the previous step and scaffolds shorter than 1 Mb during the diploid sequence construction step. We ran PSMC specifying atomic intervals $4 + 25 \times 2 + 4 + 6$ and performed 100 bootstrap replicates to investigate support for the resultant demography. To calibrate the plot, we calculated a Hyaenidae average mutation rate. We did this by calculating the average pairwise distance for each species pair and dividing that by $2 \times$ our estimated mean divergence times calculated above (Supplementary material tables S7 and S8, Supplementary Material online). We calculated the pairwise distances twice independently, once with all species mapped to the spotted hyena and once with all the species mapped to the striped hyena and took the average of these numbers. We used ANGSD to calculate the pairwise distances using a consensus base call approach (-doIBS 2) and applied the following filters: -makeMatrix 1, -minMapQ 25, -minQ 25, -uniqueOnly 1, -remove_bads 1, -setmindepthind 5, -minind 4, and only include scaffolds not aligning to the sex chromosomes and larger than 1 Mb in length. The average Hyaenidae mutation rate was taken from the average of these numbers. This gave us a mutation rate of 7.7×10^{-10} per year, only slightly slower than the commonly implemented human mutation rate of 1×10^{-9} per year (Li and Durbin 2011). To calculate the generational mutation rate, we assumed a generation length of six years for all species (Westbury et al. 2018, 2020) giving a generational mutation rate of 4.6×10^{-9} .

Acknowledgments

We would like to thank Andrew Heidel, Pablo Santos, Marion East, Heribert Hofer, Oliver Hoener, Bettina Wachter, Simone Sommer, Martin Bens, and Matthias Platzer for providing the assembled spotted hyena transcriptome to aid in annotating the genomes. We also thank Dorina Menegheni and Bernd Senf for the transcriptome data transfer. We thank Armanda

Bastos for allowing the use of her laboratory for the aardwolf DNA extractions. We also acknowledge sequencing support from the Swedish National Genomics Infrastructure (NGI) at the Science for Life Laboratory, which is funded by the Swedish Research Council and the Knut and Alice Wallenberg Foundation, and UPPMAX for access to computational infrastructure. Main funding for the project was provided through ERC consolidator grant # 310763 (GeneFlow) to M.H. D.A.D. was funded by an Australian Research Council grant (DE190100544). D.L.D. is funded through “Clinician Scientist Programm, Medizinische Fakultät der Universität Leipzig.” A.W. is funded through “Leibniz Competition Fund (SAW-2018-IZW-3-EpiRank)”

Author Contributions

Conceptualization: M.V.W., M.H.; Formal analysis: M.V.W., D.L.D., D.A.D., A.K., S.P., J.H.G.; Investigation: M.V.W.; Writing—Original Draft: M.V.W.; Writing—Review & Editing: M.V.W., L.W., F.D., M.H.; Resources: K.N., F.D., L.D., M.H.; Data curation: M.V.W., S.R., A.W.; Supervision: T.S., M.H.; Funding Acquisition: M.H.

Data Availability

The raw reads of the aardwolf can be found under Genbank Bioproject ID: PRJNA694835. The gene annotation files have been uploaded to the University of Copenhagen’s Electronic Research Data Archive (ERDA) and can be accessed at the following link: https://sid.erd.dk/cgi-sid/lis.py?share_id=BX2AOyFyaN.

References

- AbiSaid M, Dloniak SMD. 2015. *Hyaena hyaena*. The IUCN Red List of Threatened Species 2015: e.T10274A45195080. Available from: <https://dx.doi.org/10.2305/IUCN.UK.2015-2.RLTS.T10274A45195080.en>. Accessed January 2020.
- Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25(17):3389–3402.
- Angelis K, Dos Reis M. 2015. The impact of ancestral population size and incomplete lineage sorting on Bayesian estimation of species divergence times. *Curr Zool* 61(5):874–885.
- Anisimova M, Gascuel O. 2006. Approximate likelihood-ratio test for branches: a fast, accurate, and powerful alternative. *Syst Biol* 55(4):539–552.
- Bao W, Kojima KK, Kohany O. 2015. Repbase Update, a database of repetitive elements in eukaryotic genomes. *Mob DNA* 6:11.
- Barnett R, Barnes I, Phillips MJ, Martin LD, Harington CR, Leonard JA, Cooper A. 2005. Evolution of the extinct sabretooths and the American cheetah-like cat. *Curr Biol* 15(15):R589–R590.
- Baum DA. 2007. Concordance trees, concordance factors, and the exploration of reticulate genealogy. *Taxon* 56(2):417–426.
- Beasley JC, Olson ZH, DeVault TL. 2015. Ecological role of vertebrate scavengers. In: Benbow ME, Tomberlin JK, Tarone AM, editors. Carrion ecology, evolution and their applications. Boca Raton (FL): CRC Press. p. 107–127.
- Benbow EM, Tomberlin JK, Tarone AM. 2015. Carrion ecology, evolution, and their applications. Boca Raton (FL): CRC Press.
- Bens M, Sahn A, Groth M, Jahn N, Morhart M, Holtze S, Hildebrandt TB, Platzer M, Szafranski K. 2016. FRAMA: from RNA-seq data to annotated mRNA assemblies. *BMC Genom* 17(1):54.

- Blumstein DT, Rangchi TN, Briggs T, De Andrade FS, Natterson-Horowitz B. 2017. A systematic review of carrion eaters' adaptations to avoid sickness. *J Wildl Dis* 53(3):577–581.
- de Bonis L, Peigné S, Likius A, Mackaye HT, Vignaud P, Brunet M. 2007. The oldest African fox (*Vulpes ruffautae* n. sp., Canidae, Carnivora) recovered in late Miocene deposits of the Djurab desert, Chad. *Naturwissenschaften* 94(7):575–580.
- Campbell CR, Poelstra JW, Yoder AD. 2018. What is speciation genomics? The roles of ecology, gene flow, and genomic architecture in the formation of species. *Biol J Linn Soc Lond* 124(4):561–583.
- Chen L, Qiu Q, Jiang Y, Wang K, Lin Z, Li Z, Bibi F, Yang Y, Wang J, Nie W, et al. 2019. Large-scale ruminant genome sequencing provides insights into their evolution and distinct traits. *Science* 364(6446):eaav6202.
- Chidgey M, Brakebusch C, Gustafsson E, Cruchley A, Hail C, Kirk S, Merritt A, North A, Tselepis C, Hewitt J, et al. 2001. Mice lacking desmocollin 1 show epidermal fragility accompanied by barrier defects and abnormal differentiation. *J Cell Biol* 155(5):821–832.
- Court MH, Greenblatt DJ. 2000. Molecular genetic basis for deficient acetaminophen glucuronidation by cats: UGT1A6 is a pseudogene, and evidence for reduced diversity of expressed hepatic UGT1A isoforms. *Pharmacogenet Genom* 10:355–369.
- Dalerum F, de Vries JL, Pirk CWW, Cameron EZ. 2017. Spatial and temporal dimensions to the taxonomic diversity of arthropods in an arid grassland savannah. *J Arid Environ* 144:21–30.
- Di Benedetto D, Di Vita G, Romano C, Giudice ML, Vitello GA, Zingale M, Grillo L, Castiglia L, Musumeci SA, Fichera M. 2013. 6p22.3 deletion: report of a patient with autism, severe intellectual disability and electroencephalographic anomalies. *Mol Cytogenet* 6(1):4.
- Duchêne DA, Duchêne S, Ho SYW. 2018. PhyloMA: efficient assessment of phylogenomic model adequacy. *Bioinformatics* 34(13):2300–2301.
- Duchêne DA, Tong KJ, Foster CSP, Duchêne S, Lanfear R, Ho SYW. 2020. Linking Branch Lengths across sets of loci provides the highest statistical support for phylogenetic inference. *Mol Biol Evol* 37(4):1202–1210.
- El-Gebali S, Mistry J, Bateman A, Eddy SR, Luciani A, Potter SC, Qureshi M, Richardson LJ, Salazar GA, Smart A, et al. 2019. The Pfam protein families database in 2019. *Nucleic Acids Res* 47(D1):D427–D432.
- Fang K, Zhang S, Glawe J, Grisham MB, Keivil CG. 2012. Temporal genome expression profile analysis during t-cell-mediated colitis: identification of novel targets and pathways. *Inflamm Bowel Dis* 18(8):1411–1423.
- Finn RD, Clements J, Eddy SR. 2011. HMMER web server: interactive sequence similarity searching. *Nucleic Acids Res* 39(suppl):W29–W37.
- Ford-Hutchinson AF, Ali Z, Seerattan RA, Cooper DML, Hallgrímsson B, Salo PT, Jirik FR. 2005. Degenerative knee joint disease in mice lacking 3'-phosphoadenosine 5'-phosphosulfate synthetase 2 (*Paps2*) activity: a putative model of human *PAPSS2* deficiency-associated arthritis. *Osteoarthritis Cartil* 13(5):418–425.
- Fowles LF, Bennetts JS, Berkman JL, Williams E, Koopman P, Teasdale RD, Wicking C. 2003. Genomic screen for genes involved in mammalian craniofacial development. *Genesis* 35(2):73–87.
- Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, et al. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol* 29(7):644–652.
- Grabherr MG, Russell P, Meyer M, Mauceli E, Alfoldi J, Di Palma F, Lindblad-Toh K. 2010. Genome-wide synteny through highly sensitive sequence alignment: satsuma. *Bioinformatics* 26(9):1145–1151.
- Hendey QB. 1974. The late Cenozoic Carnivora of the south-western Cape Province. *Annls S Afr Mus*. 63:1–369.
- Holt C, Yandell M. 2011. MAKER2: an annotation pipeline and genome-database management tool for second-generation genome projects. *BMC Bioinformatics* 12(1):491.
- Huang Y, Niu B, Gao Y, Fu L, Li W. 2010. CD-HIT Suite: a web server for clustering and comparing biological sequences. *Bioinformatics* 26(5):680–682.
- Jeong J, Mao J, Tenzen T, Kottmann AH, McMahon AP. 2004. Hedgehog signaling in the neural crest cells regulates the patterning and growth of facial primordia. *Genes Dev* 18(8):937–951.
- Joshi NA, Fass JN. 2011. Sickle: a sliding-window, adaptive, quality-based trimming tool for FastQ files (Version 1.33)[Software].
- Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS. 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat Methods* 14(6):587–589.
- Koepfli K-P, 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(1):59.
- Korneliusson TS, Albrechtsen A, Nielsen R. 2014. ANGSD: analysis of next generation sequencing data. *BMC Bioinformatics* 15(1):356.
- Kosiol C, Vinař T, da Fonseca RR, Hubisz MJ, Bustamante CD, Nielsen R, Siepel A. 2008. Patterns of positive selection in six mammalian genomes. *PLoS Genet* 4(8):e1000144.
- Lechner M, Findeiß S, Steiner L, Marz M, Stadler PF, Prohaska SJ. 2011. Proteinortho: detection of (co-)orthologs in large-scale analysis. *BMC Bioinformatics* 12(1):124.
- Le Duc D, Renaud G, Krishnan A, Almén MS, Huynen L, Prohaska SJ, Ongyerth M, Bitarello BD, Schiöth HB, Hofreiter M, et al. 2015. Kiwi genome provides insights into evolution of a nocturnal lifestyle. *Genome Biol* 16(1):147.
- Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* 25(14):1754–1760.
- Li H, Durbin R. 2011. Inference of human population history from individual whole-genome sequences. *Nature* 475(7357):493–496.
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup. 2009. The sequence alignment/map format and SAMtools. *Bioinformatics* 25(16):2078–2079.
- Li H, Kamiie J, Morishita Y, Yoshida Y, Yaoita E, Ishibashi K, Yamamoto T. 2005. Expression and localization of two isoforms of AQP10 in human small intestine. *Biol Cell* 97(11):823–829.
- Liaw A, Wiener M. 2002. Classification and regression by randomForest. *R News* [Internet] 2:18–22. Available from: <https://CRAN.R-project.org/doc/Rnews/>. Accessed September 2020.
- Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15(12):550.
- Lu S, Wang J, Chitsaz F, Derbyshire MK, Geer RC, Gonzales NR, Gwadz M, Hurwitz DI, Marchler GH, Song JS, et al. 2020. CDD/SPARCLE: the conserved domain database in 2020. *Nucleic Acids Res* 48(D1):D265–D268.
- Magoč T, Salzberg SL. 2011. FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* 27(21):2957–2963.
- Mai U, Mirab S. 2018. TreeShrink: fast and accurate detection of outlier long branches in collections of phylogenetic trees. *BMC Genom* 19(S5):272.
- Marneweck D, Cameron EZ, Ganswindt A, Dalerum F. 2015. Behavioural and endocrine correlates to the aardwolf mating system. *Mammal Biol—Zeitschrift Säugetierkunde* 80(1):31–38.
- Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *Embnet J* 17(1):10–12.
- Matsebula SN, Monadjem A, Roques KG, Garcelon DK. 2009. The diet of the aardwolf, *Proteles cristatus* at Malolotja Nature Reserve, western Swaziland. *Afr J Ecol* 47(3):448–451.
- Mendes FK, Hahn MW. 2016. Gene tree discordance causes apparent substitution rate variation. *Syst Biol* 65(4):711–721.
- Meredith RW, Janečka JE, Gatesy J, Ryder OA, Fisher CA, Teeling EC, Goodbla A, Eizirik E, Simão TLL, Stadler T, et al. 2011. Impacts of the cretaceous terrestrial revolution and KPg extinction on mammal diversification. *Science* 334(6055):521–524.
- Minh BQ, Hahn MW, Lanfear R. 2020. New methods to calculate concordance factors for phylogenomic datasets. *Mol Biol Evol* 37(9):2727–2733.

- Minh BQ, Nguyen MAT, von Haeseler A. 2013. Ultrafast approximation for phylogenetic bootstrap. *Mol Biol Evol* 30(5):1188–1195.
- Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, Lanfear R. 2020. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol Biol Evol* 37(5):1530–1534.
- Narasimhan V, Danecek P, Scally A, Xue Y, Tyler-Smith C, Durbin R. 2016. BCFtools/ROH: a hidden Markov model approach for detecting autozygosity from next-generation sequencing data. *Bioinformatics* 32(11):1749–1751.
- Naser-Khdour S, Minh BQ, Zhang W, Stone EA, Lanfear R. 2019. The prevalence and impact of model violations in phylogenetic analysis. *Genome Biol Evol* 11(12):3341–3352.
- Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol* 32(1):268–274.
- Nguyen VH, Lavenier D. 2009. PLAST: parallel local alignment search tool for database comparison. *BMC Bioinformatics* 10(1):329.
- Niimura Y. 2009. On the origin and evolution of vertebrate olfactory receptor genes: comparative genome analysis among 23 chordate species. *Genome Biol Evol* 1:34–44.
- Niimura Y. 2013a. Identification of chemosensory receptor genes from vertebrate genomes. *Methods Mol Biol* 1068:95–105.
- Niimura Y. 2013b. Identification of olfactory receptor genes from mammalian genome sequences. *Methods Mol Biol* 1003:39–49.
- Niimura Y, Nei M. 2007. Extensive gains and losses of olfactory receptor genes in mammalian evolution. *PLoS One* 2(8):e708.
- Niimura Y, Matsui A, Touhara K. 2014. Extreme expansion of the olfactory receptor gene repertoire in African elephants and evolutionary dynamics of orthologous gene groups in 13 placental mammals. *Genome Res* 24(9):1485–1496.
- Niimura Y, Matsui A, Touhara K. 2018. Acceleration of olfactory receptor gene loss in primate evolution: possible link to anatomical change in sensory systems and dietary transition. *Mol Biol Evol* 35(6):1437–1450.
- Opreescu SN, Chepa-Lotrea X, Takase R, Golas G, Markello TC, Adams DR, Toro C, Gropman AL, Hou Y-M, Malicdan MCV, et al. 2017. Compound heterozygosity for loss-of-function GARS variants results in a multisystem developmental syndrome that includes severe growth retardation. *Hum Mutat* 38(10):1412–1420.
- Paddock CD, McKerrrow JH, Hansell E, Foreman KW, Hsieh I, Marshall N. 2001. Identification, cloning, and recombinant expression of proclinalin, a major triatomine allergen. *J Immunol* 167(5):2694–2699.
- Plummer M, Best N, Cowles K, Vines K. 2006. CODA: convergence diagnosis and output analysis for. *MCMC R News* 6:7–11.
- Ranwez V, Douzery EJP, Cambon C, Chantret N, Delsuc F. 2018. MACSE v2: toolkit for the alignment of coding sequences accounting for Frameshifts and stop codons. *Mol Biol Evol* 35(10):2582–2584.
- R Core Team. 2013. R: a language and environment for statistical computing. Available from: <http://ftp.uvigo.es/CRAN/web/packages/dplr/vignettes/intro-dplr.pdf>
- Renaud G, Hanghøj K, Korneliusen TS, Willerslev E, Orlando L. 2019. Joint estimates of heterozygosity and runs of homozygosity for modern and ancient samples. *Genetics* 212(3):587–614.
- Rice P, Longden I, Bleasby A. 2000. EMBOSS: the European Molecular Biology Open Software Suite. *Trends Genet* 16(6):276–277.
- Richardson PRK. 1991. Territorial significance of scent marking during the non-mating season in the aardwolf *Proteles cristatus* (Carnivora: proteleidae). *Ethology* 87(1–2):9–27.
- Richardson PRK, Levitan CD. 1994. Tolerance of aardwolves to defense secretions of *Trinervitermes trinervoides*. *J Mammal* 75(1):84–91.
- Rohland N, Pollack JL, Nagel D, Beauval C, Airvaux J, Pääbo S, Hofreiter M. 2005. The population history of extant and extinct hyenas. *Mol Biol Evol* 22(12):2435–2443.
- Sklyarova T, Bonné S, D'Hooge P, Denecker G, Goossens S, De Rycke R, Borgonie G, Bösl M, van Roy F, van Hengel J. 2008. Plakophilin-3-deficient mice develop hair coat abnormalities and are prone to cutaneous inflammation. *J Invest Dermatol* 128(6):1375–1385.
- Sklyarova T, van Hengel J, Van Wouterghem E, Libert C, Van Roy F, Vandenbroucke RE. 2015. Hematopoietic plakophilin-3 regulates acute tissue-specific and systemic inflammation in mice. *Eur J Immunol* 45(10):2898–2910.
- Stanke M, Waack S. 2003. Gene prediction with a hidden Markov model and a new intron submodel. *Bioinformatics* 19(Suppl 2):ii215–ii225.
- Thorne JL, Kishino H, Painter IS. 1998. Estimating the rate of evolution of the rate of molecular evolution. *Mol Biol Evol* 15(12):1647–1657.
- Waterhouse RM, Seppey M, Simão FA, Manni M, Ioannidis P, Klioutchnikov G, Kriventseva EV, Zdobnov EM. 2018. BUSCO applications from quality assessments to gene prediction and phylogenomics. *Mol Biol Evol* 35(3):543–548.
- Watts HE, Holekamp KE. 2007. Hyena societies. *Curr Biol* 17(16):R657–R660.
- Wen D, Nakhleh L. 2018. Coestimating reticulate phylogenies and gene trees from multilocus sequence data. *Syst Biol* 67(3):439–457.
- Wen D, Yu Y, Zhu J, Nakhleh L. 2018. Inferring phylogenetic networks using PhyloNet. *Syst Biol* 67(4):735–740.
- Werdelin L. 2003. Carnivora from the Kanapoi hominid site, Turkana Basin, northern Kenya. *Contrib Sci* 498:115–132.
- Werdelin L, Dehghani R. 2011. Carnivora. In: Harrison T, editor. Paleontology and geology of laetoli: human evolution in context. Vol. 2: Fossil hominins and the associated fauna. Berlin, Germany: Springer.
- Werdelin L, Lewis ME, Haile-Selassie Y. 2015. A critical review of African species of *Eucyon* (Mammalia; Carnivora; Canidae), with a new species from the Pliocene of the Woranso-Mille Area, Afar Region, Ethiopia. *Palaeontology* 1(1):33–40.
- Werdelin L, Manthi FK. 2012. Carnivora from the Kanapoi hominid site, northern Kenya. *J Afr Earth Sci* 64:1–8.
- Werdelin L, Sanders WJ. 2010. Cenozoic mammals of Africa. Berkeley (CA): University of California Press.
- Werdelin L, Solounias N. 1991. The Hyaenidae: taxonomy, systematics and evolution. *Fossil Strata* 30:1–104.
- Werdelin L, Turner A, Solounias N. 1994. Studies of fossil hyaenids: the genera *Hyaenictis* Gaudry and *Chasmaporthetes* Hay, with a reconsideration of the Hyaenidae of Langebaanweg, South Africa. *Zool J Linn Soc* 111(3):197–217.
- Westbury MV, De Cahsan B, Dalerum F, Norén K, Hofreiter M. 2019. Aardwolf population diversity and phylogenetic positioning inferred using complete mitochondrial genomes. *S Afr J Wildl Res* 49:27–33.
- 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, Petersen B, Garde E, Heide-Jørgensen MP, Lorenzen ED. 2019. Narwhal genome reveals long-term low genetic diversity despite current large abundance size. *iScience* 15:592–599.
- 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.
- Yang Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol* 24(8):1586–1591.
- Zerbino DR, Achuthan P, Akanni W, Amode MR, Barrell D, Bhai J, Billis K, Cummins C, Gall A, Girón CG, et al. 2018. Ensembl 2018. *Nucleic Acids Res* 46(D1):D754–D761.
- Zhang C, Ogilvie HA, Drummond AJ, Stadler T. 2018. Bayesian inference of species networks from multilocus sequence data. *Mol Biol Evol* 35(2):504–517.
- Zhang C, Rabiee M, Sayyari E, Mirarab S. 2018. ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinformatics* 19(S6):153.