Vulture Genomes Reveal Molecular Adaptations Underlying Obligate Scavenging and Low Levels of Genetic Diversity

Dahu Zou,[1] Shilin Tian,[1] Tongzuo Zhang,[2] Nima Zhuoma,[3] Guosheng Wu,[4] Muyang Wang,[5] Lu Dong,[6] Stephen J. Rossiter,[7] and Huabin Zhao *[,1,3]

[1]Department of Ecology, Tibetan Centre for Ecology and Conservation at WHU-TU, Hubei Key Laboratory of Cell Homeostasis, College of Life Sciences, Wuhan University, Wuhan, China

[2]Key Laboratory of Adaptation and Evolution of Plateau Biota, Northwest Institute of Plateau Biology, Chinese Academy of Sciences, Xining, China

[3]Research Center for Ecology, College of Science, Tibet University, Lhasa, China

[4]Xining Wildlife Park of Qinghai Province, Xining, China

[5]State Key Laboratory of Desert and Oasis Ecology, Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences, Urumqi, China

[6]Ministry of Education Key Laboratory for Biodiversity and Ecological Engineering, College of Life Sciences, Beijing Normal University, Beijing, China

[7]School of Biological and Chemical Sciences, Queen Mary University of London, London, United Kingdom

***Corresponding author:** E-mail: huabinzhao@whu.edu.cn.

Associate editor: Bing Su

Abstract

Obligate scavenging on the dead and decaying animal matter is a rare dietary specialization that in extant vertebrates is restricted to vultures. These birds perform essential ecological services, yet many vulture species have undergone recent steep population declines and are now endangered. To test for molecular adaptations underlying obligate scavenging in vultures, and to assess whether genomic features might have contributed to their population declines, we generated high-quality genomes of the Himalayan and bearded vultures, representing both independent origins of scavenging within the Accipitridae, alongside a sister taxon, the upland buzzard. By comparing our data to published sequences from other birds, we show that the evolution of obligate scavenging in vultures has been accompanied by widespread positive selection acting on genes underlying gastric acid production, and immunity. Moreover, we find evidence of parallel molecular evolution, with amino acid replacements shared among divergent lineages of these scavengers. Our genome-wide screens also reveal that both the Himalayan and bearded vultures exhibit low levels of genetic diversity, equating to around a half of the mean genetic diversity of other bird genomes examined. However, demographic reconstructions indicate that population declines began at around the Last Glacial Maximum, predating the well-documented dramatic declines of the past three decades. Taken together, our genomic analyses imply that vultures harbor unique adaptations for processing carrion, but that modern populations are genetically depauperate and thus especially vulnerable to further genetic erosion through anthropogenic activities.

Key words: **vulture, conservation, scavenging, gastric acid, immune system.**

Introduction

The evolution of obligate scavenging in vultures has provided humans with essential and irreplaceable ecological services. By efficiently locating and consuming carrion, vultures play a pivotal role in decomposition and nutrient cycling (DeVault et al. 2003; Wilson and Wolkovich 2011). Moreover, they also suppress populations of other scavengers, including rats, insects, and microbes, including potentially dangerous pathogens such as rabies, anthrax, and bubonic plague (Buechley and Sekercioglu 2016b). Vultures also hold important cultural significance; for example, human populations inhabiting the Qinghai-Tibet Plateau have long practiced traditional sky burials, in which the deceased are offered for consumption by the Himalayan vulture (*Gyps himalayensis*) (Martin 1996; Maming et al. 2018). Despite their ecological and cultural importance, vultures are of considerable conservation concern. Of the 22 extant species of vulture worldwide, 16 (73%) are at risk of extinction, including those in the IUCN Red List category of Extinct, Critically Endangered, Endangered, Vulnerable, or Near Threatened, and 17 (77%) have decreasing population trends (IUCN 2020). As such, vultures are among the world's most endangered groups of birds (Buechley and Şekercioğlu 2016a).

Vultures also present a unique model system in which to study the evolution of dietary adaptations. These birds were traditionally classified as a single group on the basis of phenotypic similarity yet are now known to belong to two

separate lineages of the family Accipitridae (Old World vultures) as well as the family Cathartidae (New World vultures) (Buechley and Sekercioglu 2016b). Thus, obligate scavenging has evolved independently on at least three occasions in birds. To survive on carrion, vultures must overcome the risks of ingesting rotten flesh. Indeed, these birds frequently feed on the carcasses of animals that died of disease, and they are thus exposed to a range of pathogens that would be harmful to most species, including bacteria that lead to anthrax, tuberculosis, and brucellosis (Mendoza et al. 2018), as well as the toxic metabolites of microorganism such as botulinum toxins, cadaverine, and putrescine (Cope 2018).

Several ideas have been proposed to explain the ability of vultures to subsist on carrion without becoming unwell (Blumstein et al. 2017). First, obligate scavengers appear to have the highest stomach acidity (pH ~1.3) of all birds, and it has been suggested that this might serve to inactivate microorganisms (Beasley et al. 2015). In birds, gastrointestinal acid is produced by the proventriculus and is thought to have evolved as an ecological filter to kill microorganisms that would otherwise damage commensal microbiota. Supporting this hypothesis, there is some evidence that the *SST* gene in the gastric acid secretion pathway has been positively selected in Accipitrimorphae (Chung et al. 2015). Second, vultures might also derive some protection from their own microbiomes. In New World vultures, for example, it has been proposed that simple and highly conserved gut microbiota promote the breakdown of carrion (Roggenbuck et al. 2014), while in turkey vultures, the microbiomes of facial skin may offer protection from dermatitis and pneumonia, and gas gangrene and food poisoning (Mendoza et al. 2018). Finally, and perhaps least surprising, vultures are also thought to have strong immune systems compared to other vertebrates (de la Lastra and de la Fuente 2007). Although evidence for this idea remains limited, it has been shown that the toll-like receptor 1 (TLR1) of the griffon vulture differs from that of other birds, and it is posited that this confers immunological function (de la Lastra and de la Fuente 2007).

To test for molecular adaptations underlying obligate scavenging in vultures, and to assess whether genomic features might have contributed to their population declines, we sequenced and generated high-quality genome assemblies for the Himalayan vulture (*Gyps himalayensis*) and bearded vulture (*Gypaetus barbatus*)—representing both independent origins of scavenging within the Accipitridae—and a sister group of these taxa, the upland buzzard (*Buteo hemilasius*). Currently, low-coverage genome data are available for two other vultures, the cinereous vulture (Accipitridae) and the turkey vulture (Cathartidae) (Zhang et al. 2014; Chung et al. 2015). By comparing our data to published sequences from these and other birds, we examined whether the evolution of obligate scavenging has been accompanied by molecular adaptations for the detection and processing of carrion. As such, we hypothesized that signatures of positive selection would be seen in genes underlying gastric acid production, and immunity. We also used these new data to infer genome-wide patterns of heterozygosity and demographic histories, to assess whether recent declines in vulture numbers are related to a loss of genetic diversity.

Results and Discussion

Genomic Landscapes of the Three Birds

We generated three high-quality genomes. The Himalayan vulture (fig. 1A) was sequenced on a PacBio Sequel instrument, which produced 69.9 Gb qualified reads (supplementary table S1, Supplementary Material online). The genome size was estimated to be 1.26 Gb based on the k-mer spectrum (supplementary table S2, Supplementary Material online) and the final assembly was 1.19 Gb comprising 961 contigs with high contiguity and completeness (~58.7× mean coverage, contig N50 = 3.97 Mb; supplementary tables S1 and S3, Supplementary Material online). The genome size is similar to that of 48 bird genomes previously reported (Zhang et al. 2014). The bearded vulture (fig. 1B) was sequenced using the Illumina HiSeq platform with 10× Genomics. A total of 121-fold high-quality bases were generated across 1.29 Gb assembled sequence, with N50 values of 175.1 kb and 3.91 Mb, respectively, for contigs and scaffolds (supplementary tables S1 and S3, Supplementary Material online). For genome of the upland buzzard (fig. 1C), Illumina paired-end plus mate-pair sequencing yielded 205.15-Gb qualified reads (coverage of 168.71×), resulting in an assembly with size of 1.2 G, scaffold N50 of 5.83 Mb, and contig N50 of 32.78 kb (supplementary tables S1 and S3, Supplementary Material online).

Analyses of the three draft genomes showed that 95.5–96.3% of the 8,338 avian BUSCOs were complete (supplementary table S3, Supplementary Material online), and that GC content was 41.99–42.03% (supplementary table S4, Supplementary Material online). Whole-genome annotation for the Himalayan vulture was performed via three complementary methods (supplementary table S5, Supplementary Material online). Similar to the numbers in the best-characterized chicken genome, 17,849 protein-coding genes were predicted after filtering low-quality genes (<50aa or with internal stop codon) (supplementary table S6, Supplementary Material online). For the bearded vulture and the upland buzzard, the only homology-based prediction was conducted due to the lack of RNA samples. A total of 16,774 and 17,218 protein-coding genes were annotated for the two species respectively after filtering low-quality genes (supplementary table S5, Supplementary Material online). About 86% of genes were functionally annotated for the Himalayan vulture, while 95% for both the bearded vulture and the upland buzzard (supplementary table S6, Supplementary Material online).

Analysis of Selection and Convergence

To identify all putative positively selected genes (PSGs) and rapidly evolving genes (REGs) in vultures, we generated data sets with and without the inclusion of the low-quality turkey vulture genome; these data sets comprised 5,655 and 8,080 orthologous genes, respectively. The species tree topology used for selection analysis is shown in figure 2A. When

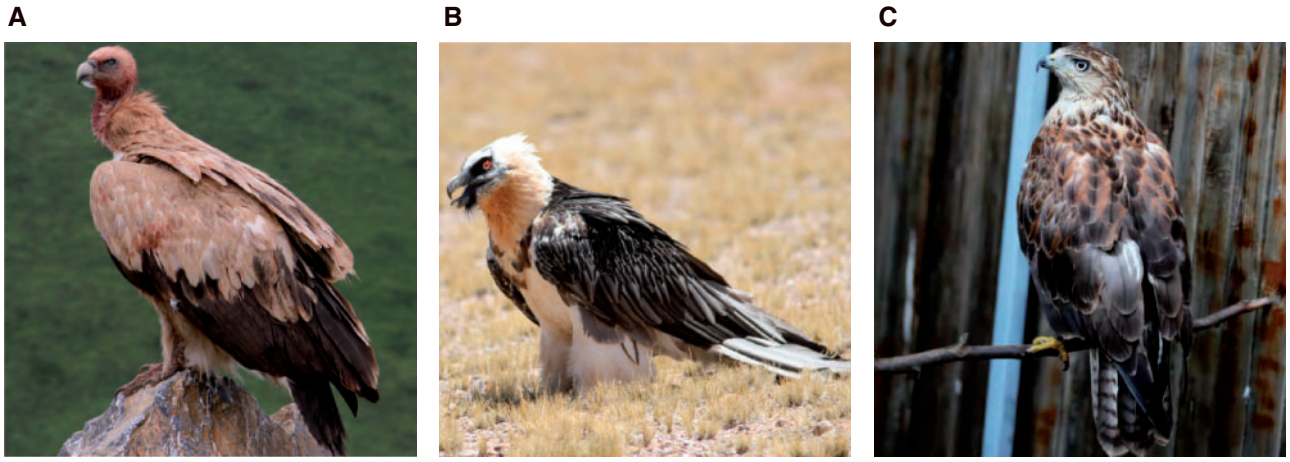

FIG. 1. Photographs of three birds. (*A*) Himalayan vulture (Image credit: Shiyi Tang); (*B*) Bearded vulture (Image credit: Hongfen Cao); (*C*) Upland buzzard (Image credit: Chang Dai).

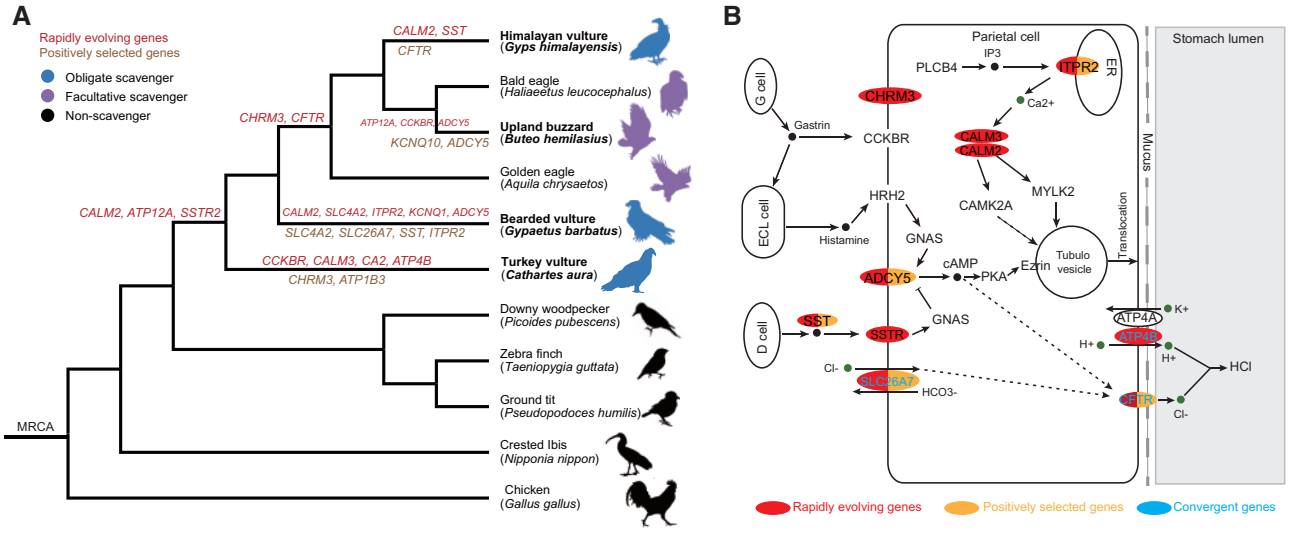

FIG. 2. Diagram of gastric acid secretion pathway evolution in the three vultures. (*A*) Phylogenetic relationship of birds used in tests for selection. Rapidly evolving genes are marked red, and positively selected genes are marked yellow. (*B*) Alterations in the gastric acid secretion pathway of vultures. Genes in red are rapidly evolving genes, genes in orange are under positive selection, and blue genes are under convergent evolution. *ATP4B* is a convergent gene between Himalayan vulture and Turkey vulture; *CFTR* is a convergent gene between Himalayan vulture and Bearded vulture; and *SLC26A7* is a convergent gene among all three vultures examined.

performing branch and branch-site model tests on the focal vulture, other vultures were removed from the background set of taxa. Our analysis yielded 70 PSGs (supplementary table S7, Supplementary Material online) and 321 REGs (supplementary table S8, Supplementary Material online) for Himalayan vulture, 93 PSGs (supplementary table S9, Supplementary Material online) and 352 REGs (supplementary table S10, Supplementary Material online) for the bearded vulture, and 109 PSGs (supplementary table S11, Supplementary Material online) and 281 REGs (supplementary table S12, Supplementary Material online) for the turkey vulture. Among these genes, three (*PIGR*, *NAA20*, and *SLC25A44*) are shared by all three vultures (supplementary fig. S1, Supplementary Material online). The results of the selection analysis for three control lineages (upland buzzard, ground tit, and downy woodpecker) are shown in supplementary table S13, Supplementary Material online. Between the two closely related vultures (Himalayan vs. Bearded vulture), positive selection or rapid evolution was detected in orthologous genes for 54 times and in paralogous genes for 60 times (supplementary table S14, Supplementary Material online). Between the pair of distantly related vultures (Himalayan vs. Turkey vulture), positive selection or rapid evolution was detected in orthologous genes for 24 times and in paralogous genes for 59 times; Between the other pair of distantly related vultures (Bearded vs. Turkey vulture), the corresponding numbers are 29 and 65 times, respectively (supplementary table S14, Supplementary Material online). Thus, distantly related scavengers tend to have more paralogous genes that are positively selected or rapidly evolved than orthologous genes, whereas closely related scavengers seem to have comparable numbers of such paralogous and

orthologous genes. For comparison, we also conducted selection analysis for the Upland buzzard (*Buteo hemilasius*) and found two PSGs (*KCNQ10*, *ADCY5*) and three REGs (*ATP12A*, *CCKBR*, and *ADCY5*) involved in the gastric acid secretion pathway (fig. 2A). However, only one of these genes (*ADCY5*) was also detected in one Old World vulture (Bearded vulture, *Gypaetus barbatus*) (fig. 2A). In addition, two PSG/REGs (*CALM2*, *SST*) are shared between the two Old World vultures but not in other birds (fig. 2A). Thus, the signatures seen in the two Old World vultures are likely convergent rather than earlier changes in the ancestral lineage of the two birds.

To test for convergence in vultures, we generated two data sets among 25 bird species in which the turkey vulture genome is of low quality and other genomes are of high quality: the first data set contained all 25 species that includes the turkey vulture (5,221 orthologous genes, supplementary fig. S2A, Supplementary Material online); and the second data set contained 24 species that excludes the turkey vulture (6,335 orthologous genes, supplementary fig. S2B, Supplementary Material online). In the first data set, we used the JTT-Fgene model (Zou and Zhang 2015) and compared the observed number of convergent amino acid substitutions among all three vultures examined (Himalayan vulture, bearded vulture, and turkey vulture) with the neutral expectations, and identified 10 genes that are under convergent evolution among the three vultures (supplementary table S15, Supplementary Material online, supplementary fig. S2A, Supplementary Material online). Based on the PCOC method (Rey et al. 2018), which considers shifts in amino acid preference instead of convergent substitutions, a total of 210 genes were found to have undergone convergent evolution (supplementary table S16, Supplementary Material online). Three genes (*RMDN1*, *OBSCN*, and *EPCAM*) were identified by both methods. In the second data set, convergence was examined between the two Old World vultures (Himalayan and bearded vulture). A total of 290 genes were identified under the JTT-Fgene model (FDR < 0.1, Poisson test, supplementary table S17, Supplementary Material online, supplementary fig. S2B, Supplementary Material online), whereas 295 genes were identified by the PCOC method (supplementary table S18, Supplementary Material online). Thirty genes were detected by both two methods (JTT-Fgene and PCOC, supplementary table S19, Supplementary Material online).

Adaptation in Gastric Acid Secretion Pathway

To determine whether the extreme stomach acidity that characterizes vultures has arisen via selection acting on enzymes controlling acid production, we estimated rates of molecular evolution in 22 genes that belong to the gastric acid secretion pathway (KEGG pathway: map04971) in 11 birds shown in figure 2A. We screened for two gene sets that show genetic signatures of adaptive evolution: 1) positively selected genes and 2) rapidly evolving genes. Of these, we found that the bearded vulture showed positive selection in four loci (*ADCY5*, *ITPR2*, *SLC26A7*, and *SST*) and rapid evolution in five (*ADCY5*, *CALM2*, *ITPR2*, *KCNQ1*, and *SLC4A2*), the Himalayan vulture showed positive selection in one locus (*CFTR*) and rapid evolution in two loci (*SST* and *SSTR2*), and Turkey vulture showed positive selection in two genes (*CHRM3* and *ATP1B3*), while four loci (*ATP4B*, *CALM3*, *CA2*, and *CCKBR*) showing rapid evolution (fig. 2A and B, supplementary table S20, Supplementary Material online). However, functional enrichment was not observed in the gastric acid secretion pathway for any set of genes (PSGs, REGs, and convergent genes) in our genome-wide analysis. We repeated these analyses for two nonscavenger control species and detected no gastric acid secretion genes under positive selection or rapid evolution in the ground tit, while three loci under positive selection (*ITPR3*, *KCNJ1*, and *MYLK3*) and none showing rapid evolution in the downy woodpecker. No gene was shared by three vultures and two control lineages. SLC26A7 is a basolateral Cl^-/HCO_3^- exchanger in gastric parietal cells and plays a major role in gastric acid secretion (Petrovic et al. 2003). *SST* encodes a regulatory peptide that inhibits gastric acid secretion by activation of the SSTR2 receptor (Lloyd et al. 1995; Patel 1999). We also found evidence that *SSTR2* underwent rapid evolution in the ancestral branch of Accipitridae. *CFTR* encodes the cystic fibrosis transmembrane conductance regulator, which regulates Cl^- secretion (Borowitz 2015). CHRM3 is a muscarinic receptor primarily expressed in gastrointestinal tract and regulate the secretion of gastric acid in parietal cells (Xie and Raufman 2016). *ATP4B*, encodes the β subunit (H/Kβ) of the gastric proton pump, the heterodimeric gastric H+/K+ ATPase, which transport H^+ into the GI tract (Nguyen et al. 2004). H/Kβ stabilizes H+/K+ ATPase alpha subunit (H/Kα) and is required for the enzymatic activity and in vivo function of the enzyme (Scarff et al. 1999).

In addition to uncovering selection, we also found evidence that three members of the gastric pathway have undergone convergent evolution in vultures: *ATP4B*, *CFTR*, and *SLC26A7* (fig. 2B). The first of these, *ATP4B*, encodes the β subunit (H/Kβ) of the gastric proton pump, is a rapidly evolving gene in Turkey vulture. The third convergent gene, *SLC26A7*, under selection in the bearded vulture, is a basolateral Cl^-/HCO_3^- exchanger (Petrovic et al. 2003). In an analysis toward conserved nonexonic elements (CNEs), four genes (*CAMK2G*, *KCNQ1*, *MYLK*, and *SLC26A7*) in the gastric acid secretion pathway were located by at least three CNEs that showed signals of convergence in the three vultures (Himalayan vulture, Bearded vulture, and Turkey vulture) (supplementary fig. S3, Supplementary Material online). *KCNQ1* is located by seven convergent CNEs is a channel gene for luminal K^+, which plays an essential role in gastric acid secretion (fig. 3A) (Grahammer et al. 2001). Five out of 60 CNEs near *SLC26A7* were detected to undergo convergent evolutionary rate shifts (fig. 3B). Among them, two CNEs located in introns of *SLC26A7* have faster rates in all the three vultures (fig. 3B). Introns have been reported to be able to enhance gene expression (Shaul 2017). This may indicate that *SLC26A7* has evolved to play a key role in gastric acid secretion not only in coding region but also in the regulatory region.

Although gastric acid secretion is a complex process involving numerous proteins (Schubert and Peura 2008),

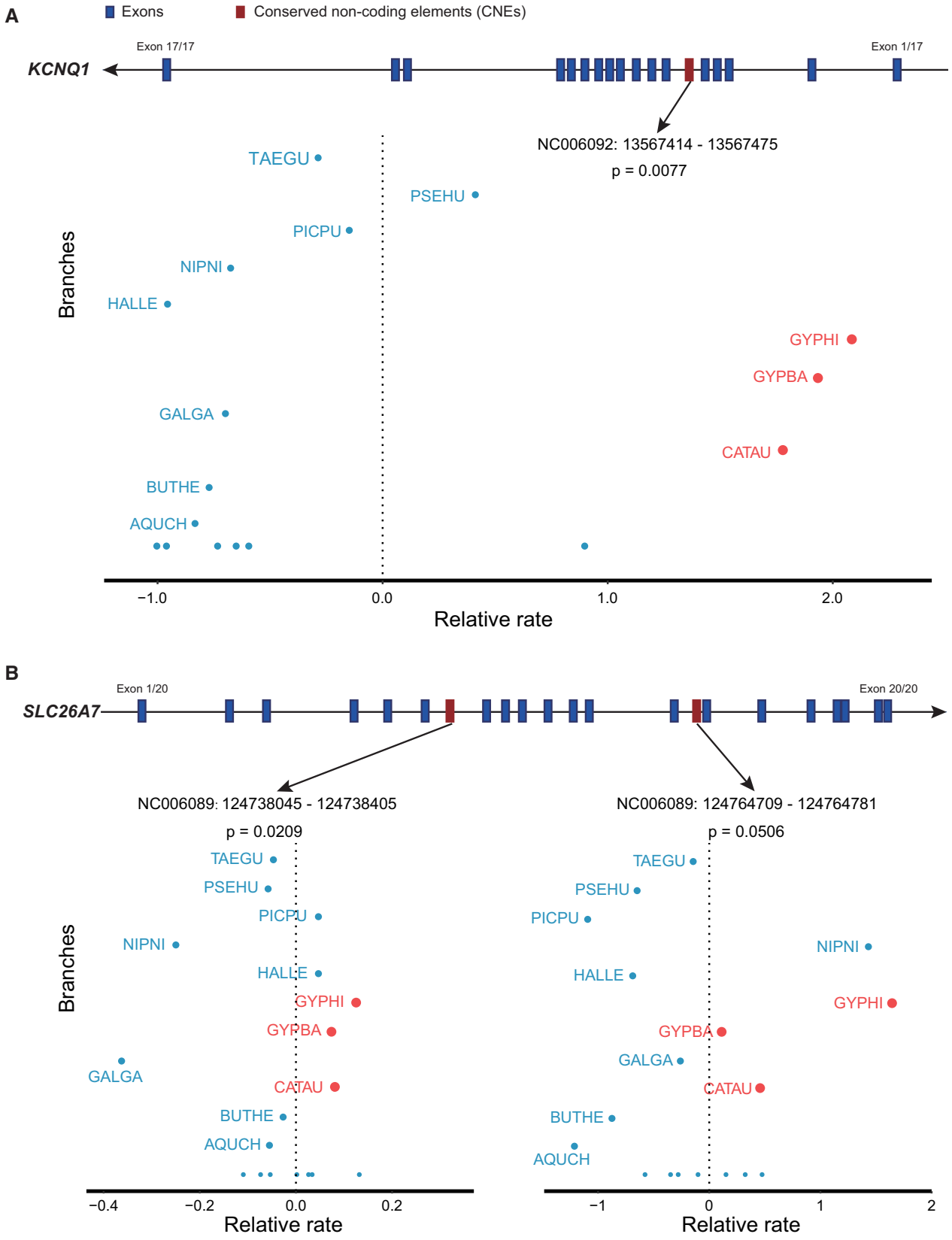

Fig. 3. Convergent acceleration of conserved nonexonic elements (CNEs) in the three vultures. (*A*) One convergent CNE in intron 5 of *KCNQ1* has a higher relative rate in all three vultures examined; (*B*) Two convergent CNEs located in introns 6 and 13 of *SLC26A7* have a higher relative rate in all three vultures examined. Information of CNEs (scaffold, start, and end) was included in its name. *P*-values refer to the Wilcoxon-rank sum test for rate acceleration on vulture branches when compared with other bird branches. AQUCH, *Aquila chrysaetos*; BUTHE, *Buteo hemilasius*; CATAU, *Cathartes aura*; GALGA, *Gallus gallus*; GYPHI, *Gyps himalayensis*; GYPBA, *Gypaetus barbatus*; HALLE, *Haliaeetus leucocephalus*; NIPNI, *Nipponia nippon*; PICPU, *Picoides pubescens*; PSEHU, *Pseudopodoces humilis*; TAEGU, *Taeniopygia guttata*.

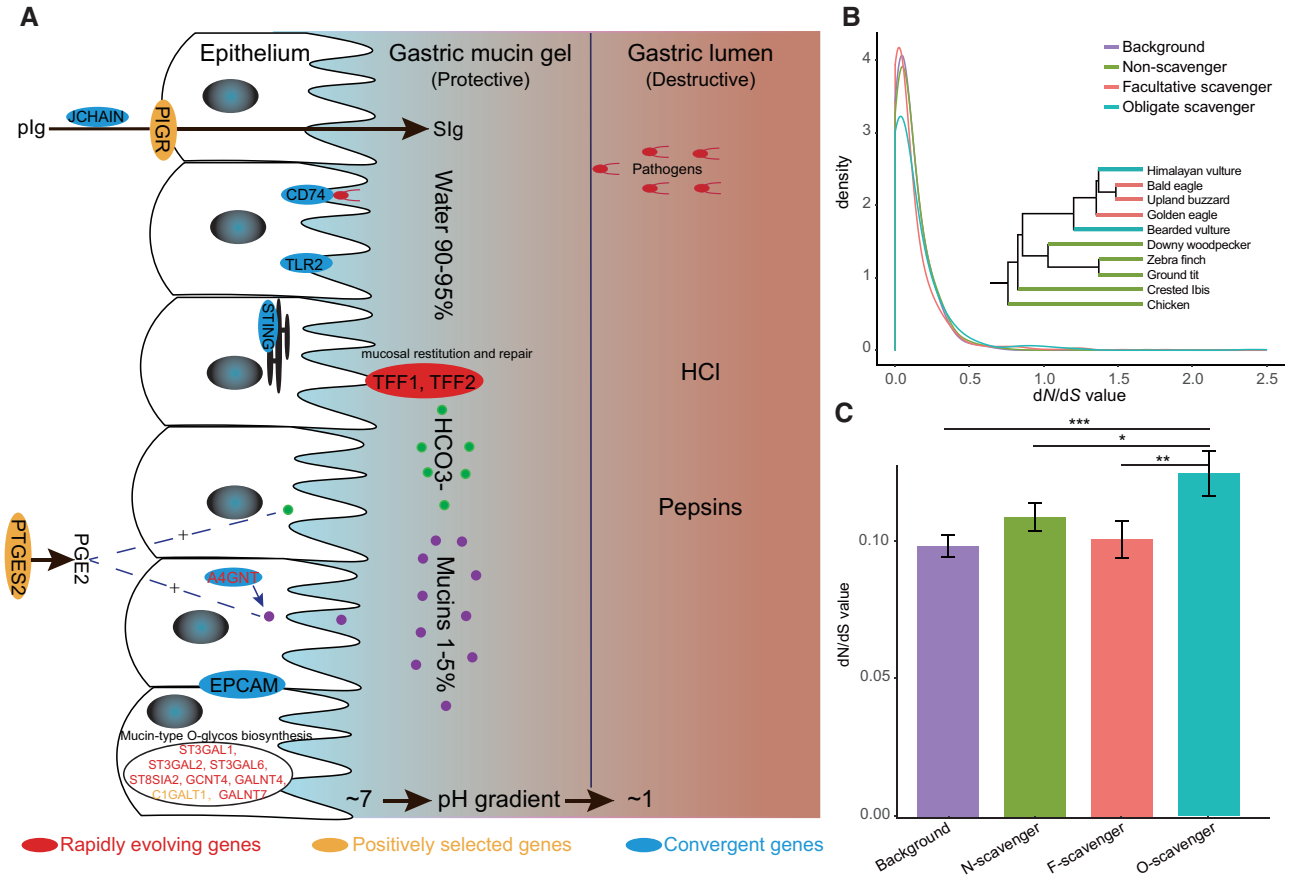

FIG. 4. Genomic features related to immune system of vultures. (*A*) Diagram of mucus defense system in vultures. Genes in red are rapidly evolving genes, genes in orange are under positive selection, and blue genes are under convergent evolution. pIg, polymeric immunoglobulin; Sig, secretory immunoglobulin. (*B*) Immune genes show higher dN/dS values in branches leading to vultures. A total of 754 genes in GO terms related to the immune response were retrieved, and we calculated the dN/dS ratios on the branches, leading to obligate scavengers, facultative scavengers, and non-scavengers, under the two-ratio branch model of PAML. Background dN/dS ratios are calculated under the one-ratio branch model of PAML. (*C*) Histogram of dN/dS (mean ± SE) in different dietary categories. The mean dN/dS values at the branches of obligate scavengers are significantly larger than background, non-scavenge, and facultative scavenge, respectively. ***$P < 0.01$, ** $P < 0.05$, *$P < 0.1$ (t test).

previous insights into the molecular basis of acidity in obligate scavengers have been limited to the detection of rapid evolution (two-ratio branch model of PAML) in a single gene (*SST*) in the cinereous vulture (Chung et al. 2015). Indeed, no genes associated with gastric acid secretion were detected in a similar study of the turkey vulture (Zhou et al. 2019). Gastric acid itself is mainly composed of hydrochloric acid (HCl), with protons (H^+) transported by gastric H^+, K^+-ATPase (ATP4A and ATP4B), and Cl^- transported by CFTR (Borowitz 2015). Additionally, different lines of evidence (from selection, convergence, and CNEs) support that *SLC26A7* has evolved at both sequence level and regulatory level. Thus, our findings of molecular adaptation in three key genes (*ATP4B*, *CFTR*, and *SLC26A7*), along with several other genes (*ADCY5*, *CA2*, *CALM3*, *CALM2*, *CCKBR*, *CAMK2G*, *ITPR2*, *KCNQ1*, *MYLK*, *SLC4A2*, *SST*, and *SSTR2*) of the gastric acid secretion pathway (fig. 2), provide compelling genetic basis for the evolution of high stomach acidity in vultures.

Adaptation of Defense System

Of our PSGs and REGs, several genes associated with mucus clearance were only recorded in the vulture group (fig. 4*A*). *PIGR* is a PSG shared by three vultures, two of which (turkey vulture and Himalayan vulture) show the same site under selection. This gene encodes a receptor that transports polymeric Ig (IgA or IgM)—one of the main immunological agents at mucosal barriers—from the basolateral surface of the epithelium to the apical side in gastrointestinal tract (Turula and Wobus 2018). *TFF1* and *TFF2* were found to be rapidly evolved at the ancestor branch of Accipitridae (figs. 2*A* and 4*A*). These two small secretory proteins encoding by *TFF1* and *TFF2* are expressed mainly in the gastrointestinal tract and play an important role in mucosal restitution and repair processes (Tolusic et al. 2018). *PTGES* is a PSG in the bearded vulture, which encodes a glutathione-dependent prostaglandin E (PGE) synthase. PGE2 has been reported to have a stimulatory effect on mucus secretion (Bersimbaev et al. 1985). Eight genes related to mucin-type O-glycos biosynthesis were identified in vultures: Himalayan vulture (1 PSG, 2 REGs), bearded vulture (4 REGs), and turkey vulture (1 REGs). We further compared these genes with those found in a group of mammalian scavenger, Hyaenidae (Westbury et al. 2021). Two genes (*NAA25* and *PTPN5*) were shared by Hyaenidae and the Himalayan vulture, six genes (*TP53BP2*,

HERC2, *STRADA*, *UBTF*, *PKP3*, and *SLC22A7*) for Hyaenidae and the bearded vulture, and two genes (*USF3* and *SLC26A3*) for Hyaenidae and the turkey vulture. Among them, the protein encoded by *TP53BP2* was found to be significantly correlated with gastric cancer susceptibility (Ju et al. 2005). This finding suggests that mammals and birds could share some molecular adaptations underlying scavenging. We next conducted GO enrichment analysis toward PSGs and REGs (*P*-value < 0.05). GO terms were retained for further analysis if they showed a *P*-value < 0.05 (Fisher's exact test), and if they did not occur in two control lineages. For the Himalayan vulture, we found that PSGs and REGs were significantly enriched in several immune-related terms and pathways (supplementary table S21, Supplementary Material online, supplementary fig. S4, Supplementary Material online): GO: 0050691∼regulation of defense response to virus by host; GO: 0039702∼viral budding via host ESCRT complex; GO: 0032715∼negative regulation of interleukin-6 production; GO: 0043029∼ T cell homeostasis; gga04115∼p53 signaling pathway. For the bearded vulture (supplementary table S22, Supplementary Material online, supplementary fig. S5, Supplementary Material online), significant enrichment was found in three terms (gga00512∼Mucin type O-Glycan biosynthesis, gga01130∼Biosynthesis of antibiotics, GO: 0032735∼positive regulation of interleukin-12 production). Of note, O-glycans contribute to maintaining the structure of mucins, a protein that covers and protects epithelial cells from pathogens and gastric acid (Guzman-Aranguez and Argueso 2010; Lang et al. 2007). For the turkey vulture (supplementary table S23, Supplementary Material online, supplementary fig. S6, Supplementary Material online), significant enrichment was detected in seven immune-related terms: GO: 0050727∼regulation of inflammatory response; GO : 0046718∼viral entry into host cell; GO: 0043922∼negative regulation by host of viral transcription; GO: 0043923∼positive regulation by host of viral transcription; GO: 00327432∼positive regulation of interleukin-2 production; GO: 0032757∼positive regulation of interleukin-8 production; GO: 0045088∼regulation of innate immune response. These terms are specific to vultures and thus may have implications for carrion feeding.

Of the convergent genes identified under the JTT-Fgene model among three vultures (supplementary table S15, Supplementary Material online, fig. 4*A*), three (*EPCAM*, *JCHAIN*, and *STING*) are strongly associated with immune defense, especially mucus defense system (Nochi et al. 2004; Li et al. 2020; Hu et al. 2021). The protein EPCAM encoded by the gene *EPCAM* is a transmembrane glycoprotein expressed on the epithelial cells and primarily known to mediate homotypic cell contacts in epithelia tissues (Nochi et al. 2004). Studies have shown that EPCAM may act as a physical homophilic interaction molecule between intestinal epithelial cells (IECs) and intraepithelial lymphocytes (IELs) at the mucosal epithelium for providing immunological barrier as a first line of defense against mucosal infection (Nochi et al. 2004). Furthermore, we also found *EPCAM* to be evolved convergently by the PCOC method (supplementary table S16, Supplementary Material online). JCHAIN is a small polypeptide, expressed by mucosal and glandular plasma cells, plays an important role in production of secretory antibodies (Johnsen et al. 2000). STING, also known as "stimulator of interferon genes," is a signaling protein that plays an essential role in controlling the transcription of several innate immune molecules, such as type I interferons (IFNs) and proinflammatory cytokines (Barber 2015). For convergent genes identified by the PCOC method, significant enrichment was observed in GO: 0032715∼negative regulation of interleukin-6 production (supplementary table S24, Supplementary Material online). *CD74* and *TLR2* are convergent genes between the two Old World vultures (fig. 4*A*), which may have critical roles in defensing pathogens in gastric tract (Smith et al. 2003; Barrera et al. 2005).

In addition, the decarboxylation of amino acids during the microbial breakdown of flesh produces biogenic amines (BAs) such as cadaverine and putrescine (Karovicova and Kohajdova 2005; Sherratt et al. 2006). Although these compounds play crucial roles in the physiology and development of eukaryotic cells, they are also toxic when consumed in excess, inducing nausea, headaches, rashes, and changes in blood pressure (Ladero et al. 2010). Since vultures appear to be unique in their ability to process food resources that are very rich in BAs (Sherratt et al. 2006), it is plausible that they may have experienced selection on amine oxidases, the key metabolic enzymes responsible for inactivating harmful exogenous BAs. To test this idea, ω values of the diamine oxidase gene (*DAO*) along each branch were estimated under the free-ratio model of PAML (Yang 2007). LRT test indicated that the free-ratio model is a better fit than the one-ratio model that assumes one same ω across all branches in interpreting our data set (chi-square test, *P*-value $= 0.0013$). We found that the *DAO* gene has rapidly evolved at least for two times. One is on the ancestral branch of Accipitridae ($\omega = 0.2327$) which coincided with the transition of nonscavenging to facultative scavenging; the other ($\omega = 0.2715$) is on the Himalayan vulture branch which represents the transition of facultative scavenging to obligate scavenging (supplementary fig. S7, Supplementary Material online). Rapid evolution has not occurred on the branch leading to the bearded vulture, which could be explained by its specialized feeding niche (bone-eating) (Buechley and Sekercioglu 2016b). Bones contain relatively less microorganisms and are difficult to degrade and thus may contain no or little biogenic amines.

To further assess whether the evolution of scavenging in Old World vultures is associated with selection in immune-related genes, we performed analyses of molecular evolution in 754 immune-related genes across 10 bird species (fig. 2*A*, without Turkey vulture). The distribution densities of the dN/dS values are shown in figure 4*B*. Comparing lineages with contrasting levels of scavenging, we found that obligate scavengers (Himalayan vulture and bearded vulture) were characterized by 11 genes (*CEBPB*, *HAND2*, *HSIL5R2*, *IFIT5*, *IL20RB*, *JCHAIN*, *NKX2-3*, *RNF26*, *SOX4*, and *TRAF3IP1*) showing positive selection with dN/dS ratio (ω) > 1. In comparison, facultative scavengers (bald eagle, golden eagle, and buzzard), and nonscavengers (chicken, crested ibis, downy woodpecker,

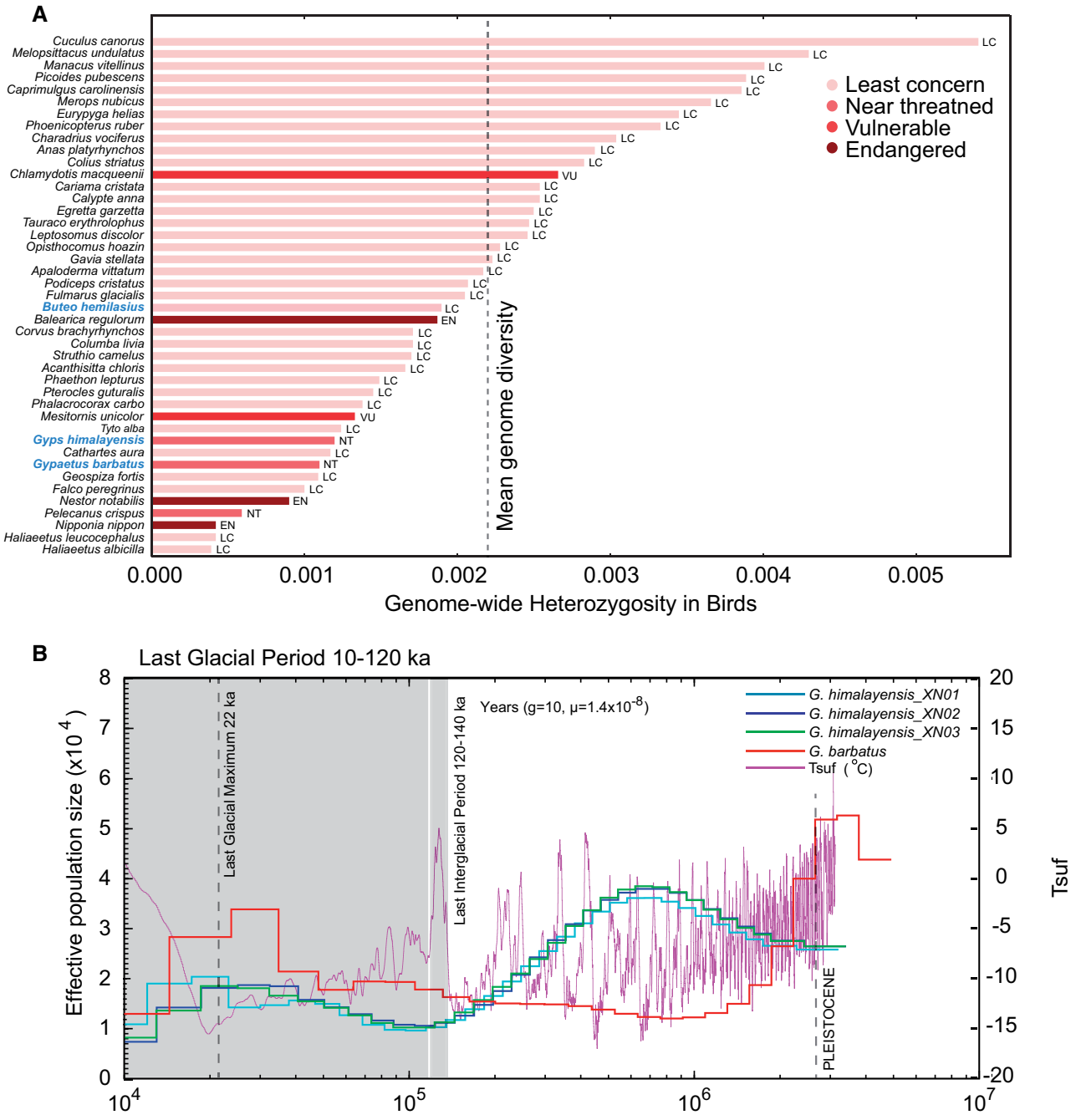

Fig. 5. Genetic diversity and demographic history of vultures. (*A*) Comparison of genome-wide heterozygosity in birds. (*B*) Demographic history of the two vultures (*G. himalayensis* and *G. barbatus*). Genomes of three individuals of *G. himalayensis* (Sample ID: XN01-XN03) and one individual of *G. barbatus* were analyzed. Tsuf, atmospheric surface air temperature.

and zebra finch) showed seven and two genes with $\omega > 1$, respectively. Compared with either the background or two other bird groups, obligate scavengers showed significantly higher ω values (Student's *t*-test) (fig. 4C), further suggesting that genetic changes in immune genes could help us explain how vultures adapted to obligate scavenging. As a control, the same analysis was performed on 451 reproductive genes (GO: 0000003). Similar to immune genes, reproductive genes also evolve more rapidly than almost all other gene categories (Dapper and Wade 2020). No significant difference in dN/dS values was observed between obligate scavengers and other bird groups (supplementary fig. S8, Supplementary Material online). This finding suggests that the evolution of scavenging in vultures is accompanied by the selection of immune genes.

Genetic Diversity and Demographic History of the Two Vultures

We estimated genome-wide heterozygosity for each of the three newly sequenced birds by mapping short reads back to their respective reference genomes, recording 0.0012 heterozygous sites/base pair for the Himalayan vulture (*Gyps himalayensis*), 0.0011 for the bearded vulture (*Gypaetus barbatus*),

and 0.0019 for the upland buzzard (*Buteo* hemilasius) (fig. 5A). By comparing these values to published data (Li et al. 2014) we found that genome-wide heterozygosity of the two vultures was similar to that reported for another vulture *Cathartes aura* (0.0012) (fig. 5A). Thus, these three vulture genomes all appear to show remarkably low diversity, around half of the mean genome diversity (0.0022) of the 42 birds examined (fig. 5A) (Li et al. 2014).

Effective population size (*Ne*) is one of the most important genetic parameters in population genetics (Wang et al. 2016), with a low *Ne* associated with a loss of genetic diversity and an increase in inbreeding (Bolton et al. 2018). We reconstructed the demographic histories of *G. himalayensis* and *G. barbatus* using the pairwise sequential Markovian coalescent (PSMC) method (Li and Durbin 2011). Our PSMC analysis used genomes of three individuals of *G. himalayensis* (XN01-XN03) and one individual of *G. barbatus* and revealed similar *Ne* fluctuations in both species (fig. 5B, supplementary fig. S9, Supplementary Material online). For *G. himalayensis*, the *Ne* showed two periods of population expansion, peaking at $\sim 7 \times 10^5$ years ago and 3×10^4 years ago, with subsequent declines to minima at $\sim 1 \times 10^5$ million years ago and present day (fig. 5B). The first population expansion coincided with the prolongation of glacial cycle (Muller and MacDonald 1997) and the second occurred after the beginning of the Last Glacial Period (LGP, 1×10^4–1.2×10^5 years ago) (Rohling et al. 1998). The demographic history of *G. barbatus* also showed two periods of population expansion and contraction, with *Ne* peaks $\sim 3 \times 10^6$ years ago and 3.5×10^4 years ago, again with evidence of a recent post-LGM decline at 2.2×10^4–1×10^4 years ago (fig. 5B).

We estimated genome-wide heterozygosity of the two focal vultures to be 0.0012 (Himalayan vulture) and 0.0011 (bearded vulture), both of which are lower than that of *Chlamydotis macqueenii* (macQueen's bustard, "Vulnerable") and *Balearica regulorum* (crowned crane, "Endangered"), two birds that listed by the IUCN within the recent past (fig. 1D) (Li et al. 2014). Vultures are among the most threatened avian groups in the world (Buechley and Sekercioglu 2016b) with 9 of the 22 species listed as "Critically Endangered" (IUCN Red List category), three as "Endangered" and, four as "Near Threatened" (Buechley and Şekercioğlu 2016a). The number of mature individuals of Himalayan and bearded vultures is estimated to be 66,000–334,000, and 1,300–6,700, respectively (IUCN 2020). Although both focal vulture species are classified as "Near Threatened," the low genetic diversity uncovered by our study indicates that they may be more vulnerable than currently appreciated. Our PSMC analysis also revealed that both vultures appear to have declined in population size since the Last Glacial Maximum (fig. 5B), predating the well-documented dramatic declines of the past three decades, especially in Asia and Africa (Buechley and Şekercioğlu 2016a). Thus, these two vultures may have experienced an ancient bottleneck at $\sim 1 \times 10^4$ years ago (fig. 5B), an observation that was also inferred from the genome of the Yangtze River dolphin (Zhou et al. 2013). It therefore seems probable that modern vulture populations are genetically depauperate due to a long-term loss of diversity, and for this reason, they might be especially vulnerable to ongoing human-induced declines, such as hunting, secondary poisoning by pesticides, infrastructure construction, and reduction of food availability due to changes in human lifestyle (DeVault et al. 2003; Wilson and Wolkovich 2011). Indeed, low genome-wide heterozygosity has been directly linked to inbreeding depression and a loss of fitness in other endangered birds (Li et al. 2014). Therefore, we suggest that more attention should be paid to vultures, and proactive management, such as restricting human activities near their nesting area, setting up artificial feeding sites, and standardizing monitoring and pharmacovigilance, should be considered to reduce the risk of extinction. Further genetic sequencing of the Himalayan and bearded vultures will allow assessments of recent population changes since the LGM, and thus a more detailed insight into the effects of human activities on populations of these keystone birds.

Conclusions

In this study, we generated high-quality genomes of the Himalayan and bearded vultures, representing both lineages of Old World vultures, alongside a sister taxon, the upland buzzard. Genome analyses showed that both vultures have remarkably low genetic diversity that is only about a half of the mean genetic diversity among all bird genomes examined, and predicted that both vultures have declined in population size since the Last Glacial Maximum. Moreover, we showed that the evolution of obligate scavenging in vultures may have been accompanied by widespread positive selection acting on genes underlying gastric acid production and immunity. In addition, we found evidence of parallel molecular evolution that involved shared amino acid replacements among divergent lineages of these scavengers. Our results provide important insights into vulture conservation and molecular adaptations underlying obligate scavenging.

Materials and Methods

Sampling and Genome Sequencing

We obtained blood from three captive Himalayan vultures (XN01, XN02, XN03) from Xining Wildlife Park, one bearded vulture (XJ01) from Taxkorgan Nature Reserve, Xinjiang, and one upland buzzard (BJ01) from the Beijing Raptor Rescue Center. Sampling was nonlethal and approved by the Wuhan University Institutional Animal Care and Use Committee. Genomic DNA was isolated from blood samples using Qiagen DNAeasy kits (Qiagen, Valencia, CA, USA).

For the Himalayan vulture, we used the individual XN01 to construct 500 bp and 20 kb libraries with an Illumina TruSeq Nano DNA Library Prep Kit and SMRTbell Template Prep Kit, respectively. We sequenced the 500 bp library on an Illumina HiSeq X platform (Genetron Health, Beijing), generating 41.7 Gb raw data. The 20 kb library was sequenced on a PacBio Sequel platform with Sequel SMRT cells 1 M v2 (Genome Center of Nextomics, Wuhan), generating 69.9 Gb subreads with an average length and N50 of 9.1 kb and 13.9 kb, respectively (supplementary table S1, Supplementary Material online; supplementary fig. S10,

Supplementary Material online, Supplementary Material online). For the bearded vulture, ~50kb insert library was processed on a 10× Genomics GemCode platform using ~1 ng input DNA. For the GEM reaction procedure during PCR, 16 bp barcodes were introduced into droplets. Barcode libraries were purified and sheared into 500 bp fragments for sequencing on the NovaSeq instrument. After sequencing, we used the supernova-2.0.0 software to take the FASTQ file containing the barcoded reads and build a graph-based assembly to produce a FASTA file suitable for downstream processing and analysis. In total, 175.55 Gb raw data was generated (supplementary table S1, Supplementary Material online). For the buzzard, we constructed short-insert-sized (250, 500, 800 bp) and mate-pair (2 kb and 5 kb) DNA libraries, generating a total of 249.25 Gb raw data on an Illumina HiSeq 2000 platform.

For the Illumina data, we used the following strategies to filter raw data: 1) reads were filtered to remove the adapters; 2) reads were trimmed to remove two low-quality bases at the 5′ end and three low-quality bases at the 3′ end; 3) filtered reads with N bases more than 10%; 4) filtered duplicated reads due to PCR amplification; 5) filtered reads with low-quality bases (≤5) more than 50%. For the PacBio data, subreads were filtered with the default parameters. Details of data are shown in supplementary table S1, Supplementary Material online.

De Novo Assembly for the Himalayan Vulture Genome

Genome size was predicted according to k-mer spectra. Using Jellyfish (Marcais and Kingsford 2011) (v2.1.3), 17-mers were counted as 28,921,310,653 from short clean reads with a k-mer depth of 23 (supplementary fig. S11, Supplementary Material online). Thus, the genome size was estimated to be ~1.26 G (supplementary table S2, Supplementary Material online). To assemble the Himalayan vulture genome assembly, we used the software Falcon (v0.2.2) (Chin et al. 2016) with the following parameters: length_cutoff =13,000, length_cutoff_pr = 13,000, max_diff = 60, max_cov= 60, min_cov= 2. The resulting assembly was 1.19 Gb and contig N50 is 3.96 Mb, with contig number of 961 (supplementary table S3, Supplementary Material online). The compositions of three bird genomes were shown in supplementary table S4, Supplementary Material online. Due to the high error rate of PacBio raw reads, we used two steps to improve data quality. First, long PacBio reads were mapped to the original assembly using the BLASR (Chaisson and Tesler 2012) with parameters: –bam –bestn 5 –minMatch 18 –nproc 4 –minSubreadLength 1000 –minAlnLength 500 –minPctSimilarity 70 –minPctAccuracy 70 –hitPolicy randombest –randomSeed 1. We also used Pilon v1.20 (Walker et al. 2014) to further correct the PacBio-corrected assembly with Illumina short reads for two rounds. To further evaluate the accuracy of the Himalayan vulture genome assembly, we aligned the short read data from whole-genome sequencing data against the genome assembly using the bowtie2 (Langmead and Salzberg 2012). Then, the resulting bam file was processed with picard-2.9.0 (http://broadinstitute.github.io/picard/, last accessed March 10, 2018) to remove PCR duplicates. After sorting with SAMtools v0.1.19 (Li et al. 2009), gatk-4.0.9.0 (McKenna et al. 2010) was applied to call SNP/INDEL. Using genomic homozygous mutations detected from the NGS data, we estimated that the genome accuracy at the base level reached 99.998% (1-homogeneous_site/genome_size).

Assembly for the Bearded Vulture Genome

To generate the de novo assembly, the supernova-2.1.1 (Weisenfeld et al. 2017) was used to assemble the 10× Genomics linked-read data with default parameters. We firstly constructed the basal graph using the de bruijn graph and DISCOVAR algorithm, by concatenating successive edges and deleting the shared K-1 bases at each junction. The supper graphs were connected when the overlap of two edges is longer than 200 bp. These steps were followed by connecting adjacent contigs into one scaffold based on one read pair and in one barcode type. The genome assembly was then phased after adjustment of orientation and order. Finally, a gap-filling step was performed by aligning 10X Genomics reads to the phased assembly using longranger align.

Assembly for the Upland Buzzard Genome

Primary genome assembly was generated by ALLPATHS-LG (Butler et al. 2008). Unipaths were built with the minimal overlap of $K = 96$. The short-insert reads (250 bp, 500 bp, and 800 bp) are merged into single "super-reads" after error correction. All these "super-reads" were then formed into an initial unipath graph which was expected to be highly accurate. However, it also contains gaps arising from bias in Illumina data. We filled gaps in the unipath graph using jumping reads (2 kb and 5 kb) by ignoring their pairing. This unipath graph was used to determine the fragment size distribution for jumping pairs reads, and estimate the distances between unipaths. Gapcloser (version 1.12) (Luo et al. 2012) was then used to conduct gap-filling. Scaffolds were finally improved by applying the fragScaff program (Adey et al. 2014).

Assessment of Genome Completeness and Genome Contents

Gene content of the three bird genomes was quantified with the Benchmarking Universal Single-Copy Orthologs (BUSCO) set (Simao et al. 2015) using the parameters of chicken in Augustus to validate the completeness (supplementary table S4, Supplementary Material online). We compared GC content of the three bird genomes with that of *Aquila chrysaetos*, *Gallus gallus*, *Haliaeetus leucocephalus*, and *Taeniopygia guttata* using quast-4.6.3 (Gurevich et al. 2013) (supplementary fig. S12, Supplementary Material online).

Evaluation of Genome Diversity

Cleaned reads were aligned to references with BWA (Li 2013) with default options, followed by removal of PCR duplicates using rmdup command in SAMtools (Li et al. 2009). Raw results of mapping were processed using the mpileup command of SAMtools with the -Q 30 option and SNPs were

called by "bcftools call" command. We counted the number of heterozygous sites in each genome and calculated genome diversity by the total number of heterozygous sites divided by its genome size.

Genome Annotation

Repeat Annotation

A combination method of similarity searching and de novo prediction was applied to identify repetitive elements in three bird genomes. Transposable elements were identified using RepeatMasker (Smit et al. 2015) and RepeatProteinMask against the Repbase transposable element library (Jurka 1998). We used RepeatScount, PILER-DF and RepeatModeler-1.0.5 (Smit and Hubley 2010) to construct a de novo transposable element library, which was then used by RepeatMasker to predict repeats. We predicted tandem repeats using TRF (Benson 1999). The Repbase-based annotations and de novo annotations were then merged. The two Old World vultures sequenced in this study seem to have more hAT-Charlie compared with other birds (supplementary fig. S13, Supplementary Material online).

Gene Prediction

For the bearded vulture and buzzard, only homology-based prediction was used while three different types of approaches were applied for the Himalayan vulture genome prediction, these results were shown in supplementary table S5, Supplementary Material online.

Homology-Based Prediction. We performed homology-based gene prediction with five steps. First, protein sequences of human, mouse, chicken, zebra finch, and bald eagle were aligned to the masked genome assemblies using genblastg, which uses tblastn hits to define high-quality gene models (She et al. 2011). The raw gene models were obtained in this step. Second, candidate gene regions were extracted from filtered and extended gene models. Third, the candidate gene regions were blasted against a query protein database to find the best match. Fourth, taking the best matches, final gene models were built using GeneWise (Birney et al. 2004). Finally, the gene models with the highest score for each candidate gene region were retained.

RNA-Seq Based Prediction. We also performed annotation based on transcripts. Raw RNA-seq data isolated from the Himalayan vulture blood were cleaned with trimmomatic-0.36 (Bolger et al. 2014) and assembled de novo using Trinity (Grabherr et al. 2011). The resulting transcripts were trimmed using seqclean to remove vectors, adaptors, and primers. Then, the tool Launch_PASA_pipeline.pl in PASA (Haas et al. 2003) was used to map clean transcripts to the Himalayan vulture genome. Gene models were extracted using the tool pasa_asmbls_to_training_set.dbi. Finally, using tophat (Trapnell et al. 2009), we mapped RNA-seq reads to the repeat-masked genome and retrieved information of intron hints using bam2hints in augustus-3.2.3.

Ab Initio Prediction. HMM parameters were trained under a standard procedure both for Augustus (Stanke et al. 2006), and SNAP (Korf 2004) with high-quality gene models derived from the step (b). Augustus was run with hints file that could highly improve the gene prediction.

Gene Models Merging. Gene models from three gene prediction methods were then integrated using EVM (Haas et al. 2008). We finally annotated 22,122 protein-coding genes in the Himalayan vulture genome. The gene number, gene length distribution, coding sequence (CDS) length distribution, exon length distribution, and intron length distribution were comparable with those in other three birds (supplementary table S5, Supplementary Material online).

Functional Annotation of Protein-Coding Genes

Gene functions were assigned according to the best match of the alignment to the SwissProt (Apweiler et al. 2010), using BLASTP (Altschul et al. 1990). We annotated motifs and domains by searching against publicly available databases, including Pfam, PRINTS, PROSITE, ProDom, and SMART using InterProScan (Quevillon et al. 2005). Gene Ontology (GO) terms for each gene were retrieved using go_enrichment package (https://github.com/enormandeau/go_enrichment, last accessed August 10, 2019). Furthermore, all genes were submitted to the online sever KAAS for KEGG pathway annotation (Moriya et al. 2007). Results are shown in supplementary table S6, Supplementary Material online.

Ortholog Definition and Alignment

To identify putative orthologs, we used the reciprocal best hit (rbh) (Wall et al. 2003). Here, we used the protein sequence of chicken as reference to conduct reciprocal blast with protein sequences of other bird species. The best-hit pairs were believed to be orthologs. The code used here is accessible at https://github.com/JinfengChen/Scripts (last accessed March 5, 2021). Orthologous genes were firstly aligned using prank (Loytynoja 2014) with parameters "-codon +F –termgap." Alignments were then processed by Fasparser (Sun 2018) and Gblocks (Castresana 2000) to remove ambiguously aligned blocks with default parameters and "-b5 = N -t = c -b0 = 10," respectively.

Identification of PSGs and REGs

Using the rbh methodology described above, a gene set with 8,080 orthologous genes were generated among the Himalayan vulture, bald eagle, upland buzzard, golden eagle, bearded vulture, sowny woodpecker, zebra finch, ground tit, crested Ibis, and chicken. The resulting alignment and tree topology shown in figure 2A (without Turkey vulture) were used as inputs for CODEML implemented in PAML v4.9e (Yang 2007). To identify PSGs, we used the branch-site model, in which an alternative model allows sites to be under positive selection on the foreground branch, whereas a null model limits sites to evolve neutrally or under purifying selection. To identify REGs, we used the branch model, in which the alternative model allows different rates for different branches and the null model assigns the same ratio to all branches. When performing the branch and branch-site model tests on the focal vulture, other vultures were removed from background

set of taxa. For example, to identify PSGs in the *Gyps himalayensis* lineage, we first removed *Gypaetus barbatus*, and then ran the branch-site model. Likelihood ratio tests (LRTs) were used to calculate *P*-values. We checked the distribution of *P*-values produced by LRT (supplementary fig. S14, Supplementary Material online) and observed an excess of *P*-values close to 1 (70–86% *P*-values greater than 0.98) for the branch-site model, in which false discovery rate (FDR) correction is largely inappropriate and too conservative (Potter et al. 2021). As a result, following a recent study (Potter et al. 2021), we newly added two post hoc filtering steps to reduce false positives in selection analysis with the branch-site model: 1) Uncorrected *P*-value generated by LRT was less than 0.05, and at least one site identified by Bayes empirical Bayes (BEB) analysis should belong to the site class 2a or 2b with a posterior probability >0.5; 2) Genes with a median interval between such BEB sites less than 10 amino acids were discarded, which helps reduce bias caused by alignment error. For the turkey vulture, a gene set containing 5,640 orthologous genes was established across 11 birds (fig. 2A). When conducting selection tests on the turkey vulture, the Himalayan vulture and bearded vulture were removed. Finally, we repeated these analyses for two nonscavenger control species (ground tit and downy woodpecker) to ensure our findings are specific to vultures.

Convergent Evolution among Vultures

To test for convergence among vultures, we expanded our data set to include 25 bird species for which high-quality genomes were available and the turkey vulture genome is of low quality. We applied a reciprocal best-hit strategy to generate two data sets: The first data set includes the turkey vulture (5,221 orthologous genes, supplementary fig. S2A, Supplementary Material online); and the second data set excludes the turkey vulture (6,335 orthologous genes, supplementary fig. S2B, Supplementary Material online). Protein sequences were aligned using prank as described previously. Signals of convergent evolution were detected using two methods: 1) method of Zhang and Kumar (Zhang and Kumar 1997; Zou and Zhang 2015). A site was assumed as a convergent site if amino acids of a focused node at that site are the same but different with their most recent ancestral amino acids. Amino acid sequences of internal nodes were reconstructed by CODEML in PAML. For each gene, the number of observed convergent site was compared with the neutral expectations derived from the JTT-fgene model, and the Poisson test was then used to evaluate the difference. 2) PCOC, which considers shifts in amino acid preference instead of convergent substitutions (Rey et al. 2018). In PCOC, a vector of amino acid frequencies that was called profile was modeled for each position and each branch. Two models were implemented in PCOC: convergent model and nonconvergent model. Under the convergent model, a site on convergent branches evolves under the profile different from that of nonconvergent branches. Under the nonconvergent model, a site is modeled to evolve under a same profile for both convergent and nonconvergent branches. PCOC then detects convergent sites by identifying the better fit between the two models.

Conserved Nonexonic Elements Analysis

To identify conserved nonexonic elements (CNEs), we started from the whole-genome alignment and used the phast package (Hubisz et al. 2011) to identify conserved elements. Analyses in this section were accomplished by a Python script which could be accessed at Github. Using LASTZ, 10 repeat-masked genomes (shown in fig. 2A) were aligned to the reference genome of chicken (GRCg6a). The Multiz program (Blanchette et al. 2004) was used to produce a multiple alignment. Then, we extracted 4-fold degenerate sites from our whole-genome alignment and used phyloFit to estimate a neutral phylogenetic model (nonconserved model). Next, we estimated rho (expected substitution rate of conserved elements relative to neutrality) using phastCons with the option of "–target-coverage 0.25 –expected-length 20 –estimate_rho," and we ran separately on nonoverlapping 10^{-7} bp windows of the input alignment. Conserved models for each window were combined by phyloBoot and then used for initial conserved element prediction. Exon regions were excluded from the highly conserved elements to generate CNEs, using the command "subtract" in BEDTools (https://bedtools.readthedocs. io/en/latest/, last accessed December 15, 2020). A final set of 847,300 CNEs with a minimum of 30-bp in length were identified. We mainly focused on CNEs ($n = 515{,}771$) which are located in the intergenic regions, in introns and within the 10-kb upstream or downstream flanking regions of genes. Trees with branch lengths for each CNE were generated by "baseml" in PAML. Convergent evolutionary rate shifts for CNEs were detected by RERconverge, a R package that can test for associations between genes' relative rates and traits of interest on a phylogeny (Kowalczyk et al. 2019).

Evolution of the High Gastric Acidity in Vultures

Vultures have stomachs with the highest acidity (pH = 1.0) where most microbes cannot survive. In order to identify candidate genes contributing to its extremely acidic stomachs, we select 37 genes in gastric acid secretion pathway (map04971) from the KEGG database. Then, we retrieved ortholog sequences of selected genes in the species of figure 2A, using the rbh methodology described above. A total of 22 homologous genes were retrieved. From the common ancestral branch of Accipitridae, PSGs and REGs were identified on all branches leading to the three vultures. When performing the branch and branch-site model tests on terminal branches of the three vultures, other vultures were excluded from background group of taxa. We also conducted selection analysis toward two control lineages (ground tit and downy woodpecker) to test whether selection detected in our analysis was specific to genes in vultures.

Evolution of Immune System in Vultures

We first retrieved 1,179 genes associated with immunity under GO: 0002376 ~ immune system process in chicken. Using immune genes of chicken as query, a total of 754 orthologous

genes were then defined among 10 birds (fig. 2A). Ten birds were then classified into three dietary groups: obligate scavenging group (Himalayan vulture and Turkey vulture), facultative scavenging group (Bald eagle, Golden eagle, and Buzzard), and nonscavenging group (Chicken, Crested ibis, Downy woodpecker, Zebra finch, and Ground tit) according to levels of scavenging (Buechley and Şekercioğlu 2016a; Sánchez-Zapata et al. 2010). Background dN/dS ratio (ω) was calculated under one-ratio branch model (model 1) of PAML. The ω values of the three dietary groups described above were calculated under the two-ratio branch model (model 2) implemented in PAML. The distribution densities and a box plot of the dN/dS values were shown (fig. 4). Mean ω values of different groups were compared using the Student's *t*-test.

Evolution of *DAO* Gene

We failed to find *DAO* gene in genomes of zebra finch, downy woodpecker and crested ibis. Therefore, we recovered the *DAO* gene in eight additional bird genomes currently accessible, they are *Accipiter nisus*, *Strigops habroptila*, *Pseudopodoces humilis*, *Calidris pugnax*, *Strix occidentalis*, *Nannopterum auritus*, *Anser indicus*, and *Chlamydotis undulata*. To estimate the lineage-specific evolutionary rate for each branch, codeml with the free-ratio model (model = 1) was run using the tree topology shown in supplementary figure S3, Supplementary Material online as the prior tree topology. The likelihood under this model was tested by likelihood ratio test (LRT) against the one-ratio model (model = 0). The significance of the LRT value was evaluated by the chi-square test with df = 14.

Demographic History Reconstruction

For assessing demographic histories of the Himalayan vulture, we supplemented our genome assembly (the individual XN01) with resequencing of two further individuals (XN02, XN03) sequenced at an average sequencing depth of 54×. Raw clean reads of these three individuals were mapping onto contigs (>50 kb) of the genome assembly. To remove contigs or scaffolds that belongs to sex chromosomes, we aligned newly sequenced genomes to the genome of chicken using lastz (Harris 2007). SAMtools (Li et al. 2009) and vcfutils.pl were then used to generate the consensus genomes for each individual. The PSMC program (Li and Durbin 2011) was used to reconstruct effective population size (Ne) fluctuations, with 100 bootstraps for each individual. For the last plotting step, 1.4×10^{-8} substitutions per site per generation (Zhang et al. 2014) and 10 years as generation time were specified (Gautschi 2001). For the bearded vulture, the same procedure and same parameters were used to reconstruct its demographic history. For both species, 100 rounds of bootstrapping analysis were performed, and the results were shown in supplementary figure S8, Supplementary Material online.

Data Availability

The whole-genome sequence data of *Gyps himalayensis*, *Gypaetus barbatus*, and *Buteo hemilasius* have been deposited in the Genome Warehouse in BIG Data Center (National Genomics Data Center Members and Partners 2020), Beijing Institute of Genomics (China National Center for Bioinformation), Chinese Academy of Sciences, under the accession number GWHBAOP00000000 for Himalayan vulture, GWHBAOQ00000000 for Bearded vulture, and GWHANWF00000000 for Upland buzzard. They are publicly accessible at http://bigd.big.ac.cn/gwh (last accessed February 29, 2021). Short-read data have also been deposited into the Genome Sequence Archive in BIG Data Center under accession number CRA003024. All these data have also been deposited into GenBank under BioProject number PRJNA648843. The scripts and pipelines used in this study have been deposited at github (https://github.com/BIGtigr/GenomicPipelines, last accessed April 25, 2021).

Supplementary Material

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

Acknowledgments

The authors thank Drs Hengwu Jiao and Kai Wang for helpful discussion on genome analysis. We thank Prof. Xin Lu (Wuhan University) for insightful comments on vulture conservation. This study was supported by the National Natural Science Foundation of China (31672272, 31722051), Natural Science Foundation of the Hubei Province (2019CFA075), and Plateau Ecology Youth Innovative Fund of Wuhan University (413100105).

Author Contributions

H.Z. conceived research, D.Z. performed research, T.Z., G.W., M.W., and L.D. provided blood samples of birds, D.Z., S.T., N.Z., and H.Z. analyzed the data, and D.Z., S.J.R., and H.Z. discussed results and wrote the manuscript.

References

Adey A, Kitzman JO, Burton JN, Daza R, Kumar A, Christiansen L, Ronaghi M, Amini S, Gunderson KL, Steemers FJ, et al. 2014. In vitro, long-range sequence information for de novo genome assembly via transposase contiguity. *Genome Res.* 24(12):2041–2049.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *J Mol Biol.* 215(3):403–410.

Apweiler R, Martin MJ, O'Donovan C, Magrane M, Alam-Faruque Y, Antunes R, Barrell D, Bely B, Bingley M, Binns D, et al. 2010. The Universal Protein Resource (UniProt) in 2010. *Nucleic Acids Res.* 38:D142–D148.

Barber GN. 2015. STING: infection, inflammation and cancer. *Nat Rev Immunol.* 15(12):760–770.

Barrera CA, Beswick EJ, Sierra JC, Bland D, Espejo R, Mifflin R, Adegboyega P, Crowe SE, Ernst PB, Reyes VE. 2005. Polarized expression of CD74 by gastric epithelial cells. *J Histochem Cytochem.* 53(12):1481–1489.

Beasley DE, Koltz AM, Lambert JE, Fierer N, Dunn RR. 2015. The evolution of stomach acidity and its relevance to the human microbiome. *PLoS One* 10(7):e0134116.

Benson G. 1999. Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res.* 27(2):573–580.

Bersimbaev RI, Tairov MM, Salganik RI. 1985. Biochemical mechanisms of regulation of mucus secretion by prostaglandin E2 in rat gastric mucosa. *Eur J Pharmacol.* 115(2-3):259–266.

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

Blanchette M, Kent WJ, Riemer C, Elnitski L, Smit AFA, 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.

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.

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30(15):2114–2120.

Bolton PE, Rollins LA, Brazill-Boast J, Maute KL, Legge S, Austin JJ, Griffith SC. 2018. Genetic diversity through time and space: diversity and demographic history from natural history specimens and serially sampled contemporary populations of the threatened Gouldian finch (Erythrura gouldiae). *Conserv Genet*. 19(3):737–754.

Borowitz D. 2015. CFTR, bicarbonate, and the pathophysiology of cystic fibrosis. *Pediatr Pulm*. 50:S24–S30.

Buechley ER, Şekercioğlu ÇH. 2016a. The avian scavenger crisis: looming extinctions, trophic cascades, and loss of critical ecosystem functions. *Biol Conserv*. 198:220–228.

Buechley ER, Sekercioglu CH. 2016b. Vultures. *Curr Biol*. 26(13):R560–R561.

Butler J, MacCallum I, Kleber M, Shlyakhter IA, Belmonte MK, Lander ES, Nusbaum C, Jaffe DB. 2008. ALLPATHS: de novo assembly of whole-genome shotgun microreads. *Genome Res*. 18(5):810–820.

Castresana J. 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol Biol Evol*. 17(4):540–552.

Chaisson MJ, Tesler G. 2012. Mapping single molecule sequencing reads using basic local alignment with successive refinement (BLASR): application and theory. *BMC Bioinformatics* 13:238.

Chin C-S, 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):1050–1056.

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.

Cope RB. 2018. Botulinum neurotoxins. *Veterin. Toxicol*. 743–757.

Dapper AL, Wade MJ. 2020. Relaxed selection and the rapid evolution of reproductive genes. *Trends Genet*. 36(9):640–649.

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.

DeVault TL, Rhodes OE, Shivik JA. 2003. Scavenging by vertebrates: behavioral, ecological, and evolutionary perspectives on an important energy transfer pathway in terrestrial ecosystems. *Oikos*. 102(2):225–234.

Gautschi BS. 2001. Conservation genetics of the bearded vulture (*Gypaetus barbatus*). University of Zurich.

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng QD, et al. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol*. 29(7):644–U130.

Grahammer F, Herling AW, Lang HJ, Schmitt-Graff A, Wittekindt OH, Nitschke R, Bleich M, Barhanin J, Warth R. 2001. The cardiac K+ channel KCNQ1 is essential for gastric acid secretion. *Gastroenterology*. 120(6):1363–1371.

Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUAST: quality assessment tool for genome assemblies. *Bioinformatics* 29(8):1072–1075.

Guzman-Aranguez A, Argueso P. 2010. Structure and biological roles of mucin-type O-glycans at the ocular surface. *Ocul Surf*. 8(1):8–17.

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.

Harris RS. 2007. Improved pairwise alignment of genomic DNA [Ph.D. Thesis]. The Pennsylvania State University.

Hu Q, Zhou Q, Xia X, Shao L, Wang M, Lu X, Liu S, Guan W. 2021. Cytosolic sensor STING in mucosal immunity: a master regulator of gut inflammation and carcinogenesis. *J Exp Clin Cancer Res*. 40(1):39.

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

IUCN. 2020. The IUCN Red List of Threatened Species. Version 2020-3. [cited 2021 Jan 25] Available from: https://www.iucnredlist.org.

Johnsen FE, Braathen R, Brandtzaeg P. 2000. Role of J chain in secretory immunoglobulin formation. *Scand J Immunol*. 52:644.

Ju H, Lee K-A, Yang M, Kim H-J, Kang CP, Sohn T-S, Rhee J-C, Kang C, Kim J-W. 2005. TP53BP2 locus is associated with gastric cancer susceptibility. *Int J Cancer*. 117(6):957–960.

Jurka J. 1998. Repeats in genomic DNA: mining and meaning. *Curr Opin Struc Biol*. 8(3):333–337.

Karovicova J, Kohajdova Z. 2005. Biogenic amines in food. *Chem Pap*. 59:70–79.

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

Kowalczyk A, Meyer WK, Partha R, Mao WG, Clark NL, Chikina M. 2019. RERconverge: an R package for associating evolutionary rates with convergent traits. *Bioinformatics* 35(22):4815–4817.

Ladero V, Calles-Enríquez M, Fernández M, Alvarez MA. 2010. Toxicological effects of dietary biogenic amines. *Curr Nutr Food Sci*. 6(2):145–156.

Lang TA, Hansson GC, Samuelsson T. 2007. Gel-forming mucins appeared early in metazoan evolution. *Proc Natl Acad Sci U S A*. 104(41):16209–16214.

Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods*. 9(4):357–U354.

Li H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv* 1303.3997v2.

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, Proc GPD; 1000 Genome Project Data Processing Subgroup. 2009. The sequence alignment/map format and SAMtools. *Bioinformatics* 25(16):2078–2079.

Li SB, Li B, Cheng C, Xiong ZJ, Liu QB, Lai JH, Carey HV, Zhang Q, Zheng HB, Wei SG, et al. 2014. Genomic signatures of near-extinction and rebirth of the crested ibis and other endangered bird species. *Genome Biol*. 15(12):557.

Li Y, Jin L, Chen T. 2020. The effects of secretory IgA in the mucosal immune system. *Biomed Res Int*. 2020:2032057.

Lloyd KCK, Wang J, Aurang K, Gronhed P, Coy DH, Walsh JH. 1995. Activation of somatostatin receptor subtype-2 inhibits acid-secretion in rats. *Am J Physiol-Gastr Liver Physiol*. 268(1):G102–G106.

Loytynoja A. 2014. Phylogeny-aware alignment with PRANK. *Methods Mol Biol*. 1079:155–170.

Luo R, Liu B, Xie Y, Li Z, Huang W, Yuan J, He G, Chen Y, Pan Q, Liu Y, et al. 2012. SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. *Gigascience* 1(1):18.

Maming R, Lee L, Yang X, Buzzard P. 2018. Vultures and sky burials on the Qinghai-Tibet Plateau. *Vul News*. 71(1):22–35.

Marcais G, Kingsford C. 2011. A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. *Bioinformatics* 27(6):764–770.

Martin D. 1996. On the cultural ecology of sky burial on the Himalayan Plateau. *East and West* 46:353–370.

McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, et al. 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res*. 20(9):1297–1303.

Mendoza MLZ, Roggenbuck M, Vargas KM, Hansen LH, Brunak S, Gilbert MTP, Sicheritz-Ponten T. 2018. Protective role of the vulture facial skin and gut microbiomes aid adaptation to scavenging. *Acta Veterin Scand*. 60(1):61.

Moriya Y, Itoh M, Okuda S, Yoshizawa AC, Kanehisa M. 2007. KAAS: an automatic genome annotation and pathway reconstruction server. *Nucleic Acids Res*. 35(Web Server issue):W182–W185.

Muller RA, MacDonald GJ. 1997. Glacial cycles and astronomical forcing. *Science* 277(5323):215–218.

National Genomics Data Center Members and Partners. 2020. Database Resources of the National Genomics Data Center in 2020. *Nucleic Acids Res*. 48:D24–D33.

Nguyen NV, Gleeson PA, Courtois-Coutry N, Caplan MJ, Van Driel IR. 2004. Gastric parietal cell acid secretion in mice can be regulated independently of H^+/K^+ ATPase endocytosis. *Gastroenterology* 127(1):145–154.

Nochi T, Yuki Y, Terahara K, Hino A, Kunisawa J, Kweon MN, Yamaguchi T, Kiyono H. 2004. Biological role of Ep-CAM in the physical interaction between epithelial cells and lymphocytes in intestinal epithelium. *Clin Immunol*. 113(3):326–339.

Patel YC. 1999. Somatostatin and its receptor family. *Front Neuroendocrinol*. 20(3):157–198.

Petrovic S, Ju X, Barone S, Seidler U, Alper SL, Lohi H, Kere J, Soleimani M. 2003. Identification of a basolateral $Cl^-/HCO3^-$ exchanger specific to gastric parietal cells. *Am J Physiol Gastrointest Liver Physiol*. 284(6):G1093–G1103.

Potter JHT, Davies KTJ, Yohe LR, Sanchez MKR, Rengifo EM, Struebig M, Warren K, Tsagkogeorga G, Lim BK, Dos Reis M, et al. 2021. Dietary diversification and specialisation in new world bats facilitated by early molecular evolution. *Mol Biol Evol*. doi: 10.1093/molbev/msab028.

Quevillon E, Silventoinen V, Pillai S, Harte N, Mulder N, Apweiler R, Lopez R. 2005. InterProScan: protein domains identifier. *Nucleic Acids Res*. 33(Web Server issue):W116–120.

Rey C, Gueguen L, Semon M, Boussau B. 2018. Accurate detection of convergent amino-acid evolution with PCOC. *Mol Biol Evol*. 35(9):2296–2306.

Roggenbuck M, Bærholm Schnell I, Blom N, Bælum J, Bertelsen MF, Sicheritz-Pontén T, Pontén TS, Sørensen SJ, Gilbert MTP, Graves GR, et al. 2014. The microbiome of New World vultures. *Nat Commun*. 5:5498.

Rohling EJ, Fenton M, Jorissen FJ, Bertrand P, Ganssen G, Caulet JP. 1998. Magnitudes of sea-level lowstands of the past 500,000 years. *Nature*. 394(6689):162–165.

Sánchez-Zapata JA, Eguía S, Blázquez M, Moleón M, Botella F. 2010. Unexpected role of ungulate carcasses in the diet of golden eagles *Aquila chrysaetos* in Mediterranean mountains. *Bird Study*. 57(3):352–360.

Scarff KL, Judd LM, Toh BH, Gleeson PA, van Driel IR. 1999. Gastric H+,K+-adenosine triphosphatase beta subunit is required for normal function, development, and membrane structure of mouse parietal cells. *Gastroenterology* 117(3):605–618.

Schubert ML, Peura DA. 2008. Control of gastric acid secretion in health and disease. *Gastroenterology* 134(7):1842–1860.

Shaul O. 2017. How introns enhance gene expression. *Int J Biochem Cell B*. 91:145–155.

She R, Chu JSC, Uyar B, Wang J, Wang K, Chen NS. 2011. genBlastG: using BLAST searches to build homologous gene models. *Bioinformatics* 27(15):2141–2143.

Sherratt TN, Wilkinson DM, Bain RS. 2006. Why fruits rot, seeds mold and meat spoils: a reappraisal. *Ecol Model*. 192(3-4):618–626.

Simao FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31(19):3210–3212.

Smit AFA, Hubley R. 2010. RepeatModeler Open-1.0. [cited 2018 Mar 2] Available from: www.repeatmasker.org/.

Smit AFA, Hubley R, Green P. 2015. RepeatMasker Open-4.0.2013-2015. [cited 2018 Mar 2] Available from: www.repeatmasker.org.

Smith MF Jr, Mitchell A, Li G, Ding S, Fitzmaurice AM, Ryan K, Crowe S, Goldberg JB. 2003. Toll-like receptor (TLR) 2 and TLR5, but not TLR4, are required for *Helicobacter pylori*-induced NF-kappa B activation and chemokine expression by epithelial cells. *J Biol Chem*. 278(35):32552–32560.

Stanke M, Keller O, Gunduz I, Hayes A, Waack S, Morgenstern B. 2006. AUGUSTUS: ab initio prediction of alternative transcripts. *Nucleic Acids Res*. 34(Web Server issue):W435–W439.

Sun YB. 2018. FasParser2: a graphical platform for batch manipulation of tremendous amount of sequence data. *Bioinformatics* 34(14):2493–2495.

Tolusic LM, Mihalj M, Koprivcic I, Lovric I, Novak S, Bijelic N, Baus-Loncar M, Belovari T, Kralik K, Pauzar B. 2018. Differential expression of *TFF* genes and proteins in breast tumors. *Acta Clin Croat*. 57:264–277.

Trapnell C, Pachter L, Salzberg SL. 2009. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* 25(9):1105–1111.

Turula H, Wobus CE. 2018. The role of the polymeric immunoglobulin receptor and secretory immunoglobulins during mucosal infection and immunity. *Viruses*. 10(5):237.

Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, Cuomo CA, Zeng QD, Wortman J, Young SK, et al. 2014. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One* 9(11):e112963.

Wall DP, Fraser HB, Hirsh AE. 2003. Detecting putative orthologs. *Bioinformatics* 19(13):1710–1711.

Wang J, Santiago E, Caballero A. 2016. Prediction and estimation of effective population size. *Heredity (Edinb)*. 117(4):193–206.

Weisenfeld NI, Kumar V, Shah P, Church DM, Jaffe DB. 2017. Direct determination of diploid genome sequences. *Genome Res*. 27(5):757–767.

Westbury MV, Le DD, Duchene DA, Krishnan A, Prost S, Rutschmann S, Grau JH, Dalen L, Weyrich A, Noren K, et al. 2021. Ecological specialisation and evolutionary reticulation in extant hyaenidae. *Mol Biol Evol*. doi: 10.1093/molbev/msab1055.

Wilson EE, Wolkovich EM. 2011. Scavenging: how carnivores and carrion structure communities. *Trends Ecol Evol*. 26(3):129–135.

Xie G, Raufman JP. 2016. Muscarinic receptor signaling and colon cancer progression. *J Cancer Metastasis Treat*. 2:195–200.

Yang Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol*. 24(8):1586–1591.

Zhang GJ, Li C, Li QY, Li B, Larkin DM, Lee C, Storz JF, Antunes A, Greenwold MJ, Meredith RW, et al.; Avian Genome Consortium. 2014. Comparative genomics reveals insights into avian genome evolution and adaptation. *Science* 346(6215):1311–1320.

Zhang JZ, Kumar S. 1997. Detection of convergent and parallel evolution at the amino acid sequence level. *Mol Biol Evol*. 14(5):527–536.

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

Zhou XM, Sun FM, Xu SX, Fan GY, Zhu KL, Liu X, Chen Y, Shi CC, Yang YX, Huang ZY, et al. 2013. Baiji genomes reveal low genetic variability and new insights into secondary aquatic adaptations. *Nat Commun*. 4:2708.

Zou Z, Zhang J. 2015. Are convergent and parallel amino acid substitutions in protein evolution more prevalent than neutral expectations? *Mol Biol Evol*. 32(8):2085–2096.