JSE Journal of Systematics and Evolution

doi: 10.1111/jse.13059

Research Article

Genomic and functional insights into dietary diversification in New World leaf-nosed bats (Phyllostomidae)

Yiran Xu[1†], Yingcan Li[1†], Huiqiao Hu[1], Hengwu Jiao[2], and Huabin Zhao[1*]

[1]College of Life Sciences, Key Laboratory of Biodiversity and Environment on the Qinghai-Tibetan Plateau of the Ministry of Education, Wuhan University, Wuhan 430072, China

[2]School of Life Sciences, Central China Normal University, Wuhan 430079, China

[†]These authors contributed equally to this work.

*Author for correspondence. E-mail: huabinzhao@whu.edu.cn

Received 8 October 2023; Accepted 15 January 2024

Abstract The most significant driver of adaptive radiation in the New World leaf-nosed bats (Phyllostomidae) is their remarkably diverse feeding habits, yet there remains a notable scarcity of studies addressing the genetic underpinnings of dietary diversification in this family. In this study, we have assembled a new genome for a representative species of phyllostomid bat, the fringe-lipped bat (*Trachops cirrhosis*), and integrated it with eight published phyllostomid genomes, along with an additional 10 genomes of other bat species. Comparative genomic analysis across 10 200 orthologus genes has unveiled that those genes subject to divergent selection within the Phyllostomidae clade are notably enriched in metabolism-related pathways. Furthermore, we identified molecular signatures of divergent selection in the bitter receptor gene *Tas2r1*, as well as 14 genes involved in digesting key nutrients such as carbohydrates, proteins, and fats. In addition, our cell-based functional assays conducted on *Tas2r1* showed a broader spectrum of perception for bitter compounds in phyllostomids compared to nonphyllostomid bats, suggesting functional diversification of bitter taste in Phyllostomidae. Together, our genomic and functional analyses lead us to propose that divergent selection of genes associated with taste, digestion and absorption, and metabolism assumes a pivotal role in steering the extreme dietary diversification in Phyllostomidae. This study not only illuminates the genetic mechanisms underlying dietary adaptations in Phyllostomidae bats but also enhances our understanding of their extraordinary adaptive radiation.

Key words: bats, dietary diversification, evolutionary genomics.

1 Introduction

Adaptive radiation is the swift diversification of numerous species, each endowed with distinct adaptive traits, occurring within a relatively brief span of time; this phenomenon underscores the profound impact of natural selection and opportunism in shaping the origin and evolution of life on Earth, rendering it a prominent subject of research in the field of evolutionary biology (Schluter, 2000). Adaptive radiation signifies a substantial augmentation in diversity and the corresponding fitness of ecological roles within a lineage, often accompanied by the accumulation of unusually high levels of morphological, physiological, behavioral differences, and ecological differentiation (Givnish, 2015). Notable examples of adaptive radiation encompass Darwin's finches, East African cichlids, Hawaiian honeycreepers, and Australian marsupials. These species have rapidly adapted to a wide array of ecological niches by altering traits such as feeding habits, behavior, and morphology (Schluter, 2000). Adaptive radiation can instigate divergent selection, an evolutionary strategy in which species stemming from a common ancestor accumulate variations that enhance their adaptation when exposed to diverse selection pressures and evolutionary forces (Schluter, 2001; Rundle & Nosil, 2005; Gautam, 2020).

Feeding habits play a pivotal role in steering the adaptive radiation of animals, exerting influence on various aspects of their ecology, behavior, morphology, and physiology (Karasov & Douglas, 2013). A classic example of adaptive radiation can be observed in the case of Darwin's finches inhabiting the Galapagos Islands. These 15 finch species, which diverged from a common ancestor ~2–3 Ma, display substantial variations in beak shape, a feature intricately tied to their respective feeding habits (Grant, 1999; Grant & Grant, 2011). Recent comparative genomic studies have unveiled the critical role of the *ALX1* gene in regulating the diversity of beak shapes among Darwin's finches, thus driving the evolution of distinct feeding habits (Lamichhaney

et al., 2015). Similarly, East African cichlids, found in the Great Lakes of Eastern Africa, represent another vertebrate group that has undergone remarkable adaptive radiation. They exhibit stark dissimilarities in coloration, body shape, teeth, jaws, and more, with corresponding shifts in their feeding habits encompassing insects, debris, algae, phytoplankton, zooplankton, mollusks, and fish (Seehausen, 2006; Koblmüller et al., 2008). Comparative genomics studies have revealed that a combination of factors, including gene duplications, substantial differences in noncoding elements, accelerated evolution of coding genes, and gene expression variations associated with transposons, collectively propel the adaptive radiation of East African cichlids. Notably, genetic disparities related to the development of teeth and mandibles are likely significant contributors to the differentiations of East African cichlids (Brawand et al., 2014).

Much like Darwin's finches, the phyllostomid bats (Family Phyllostomidae), commonly known as New World leaf-nosed bats, also represent a classic example of adaptive radiation (Fleming et al., 2020). Contrasted with their two closest relatives, Mormoopidae (comprising 18 species in two genera) and Noctilionidae (with two species in one genus), the Phyllostomidae family has undergone rapid and extensive evolution of ecological and behavioral traits spanning ~30 million years (Dumont et al., 2012; Fleming et al., 2020). This extraordinary diversification has yielded 11 subfamilies, 60 genera, and 216 species within the Phyllostomidae family (Fleming et al., 2020). One of the most pivotal factors fueling adaptive radiation in phyllostomids is their dietary diversification. These bats boast an exceptionally wide array of food sources, encompassing insects, blood, fruits, nectar, and even various vertebrates such as fish, birds, and other bats (Fleming et al., 2020). Depending on the proportions of different foods in their diets, they can be classified into various groups, including insectivores, sanguivores, carnivores, omnivores, frugivores, and nectarivores. In comparison, phylogenetic analysis suggests that the present-day New World leaf-nosed bats evolved from a common insectivorous ancestor (Datzmann et al., 2010; Arbour et al., 2019).

In the earlier stages of research concerning the dietary diversification of New World leaf-nosed bats, the primary focus lay in physiological and morphological aspects. For example, scientists examined the activity of digestive enzymes and the morphology of kidneys in 16 phyllostomid species, observing significant variations in these traits corresponding to differences in feeding habits (Schondube et al., 2001). In recent years, delving into the molecular mechanisms that underlie dietary diversification in phyllostomids has emerged as a novel line of research. Five research teams have successively reported findings concerning the relationship between molecular evolution of the short-wavelength-sensitive opsin 1 (*SWS1*) and diversification of feeding habits in phyllostomid bats, their collective research has unveiled notable differences in the selection pressure on *SWS1* among phyllostomids with varying feeding habits, even leading to *SWS1* becoming a pseudogene due to relaxed selection in all three species of vampire bats (Gutierrez et al., 2018; Kries et al., 2018; Li et al., 2018; Wu et al., 2018; Simoes et al., 2019). More recently, researchers have utilized transcriptome data for comparative analyses and detected robust selection signals in the ancestral branch of the Phyllostomidae family; however, they found fewer selection signals in certain branches exhibiting shifts in feeding habits. These findings suggest that early evolutionary events may have played a role in driving the observed dietary diversification within the Phyllostomidae family (Potter et al., 2021).

Omnivorous leaf-nosed bats have broad and diverse diets and represent a classic taxon reflecting the diversity of diets among leaf-nosed bats. Thus, they could represent distinct types of omnivory. *Tonatia saurophila*, for instance, exhibits a heightened preference for insects in its diet, while *Trachops cirrhosus* displays a larger preference for vertebrates; in contrast, *Phyllostomus discolor* leans more towards a plant-based diet (Wilman et al., 2014). As a result, *T. cirrhosus* is acknowledged as a classic example of omnivores among leaf-nosed bats, owing to their broad and diverse diets combined with highly specialized hunting strategies (Fleming et al., 2020).

To investigate the genetic basis of dietary diversification in phyllostomid bats, we postulated a hypothesis that these bats experienced functional modifications in taste perception, digestive system, and metabolic pathways. To test this hypothesis, we carried out clade model analyses, taking the entire Phyllostomidae clade as the foreground, with the aim of identifying genetic signatures of divergent selection in genes linked to the aforementioned functional categories.

2 Material and Methods

2.1 Genome sequencing

The genetic material of the fringe-lipped bat *Trachops cirrhosis* (Catalog number: M-277702) was loaned from the American Museum of Natural History. Genomic DNA (1.5 μg) was extracted from the liver using Qiagen DNAeasy kits (Qiagen, Valencia, CA, USA). For Illumina pair-end sequencing, we generated short-insert libraries (250 and 500 bp) using the Truseq Nano DNA HT Sample Prep Kit (Illumina, San Diego, CA, USA), following the manufacturer's instructions, and subsequently sequenced them on the Illumina HiSeq X platform. For 10× Genomics sequencing, we performed the GEM reaction procedure on a GemCode platform using ~1 ng of input DNA (~50 kb insert length). During the polymerase chain reaction amplification of long inserts, we introduced 16 bp barcodes into droplets. Afterward, the long-insert DNA libraries were purified and sheared into 500 bp fragments to construct barcoded libraries for sequencing on Novaseq instruments. Following sequencing, we utilized supernova-2.0.0 software (Weisenfeld et al., 2017) to obtain FASTQ files containing barcoded reads and build a graph-based assembly, resulting in FASTA files suitable for subsequent processing and analysis.

To ensure the generation of high-quality data, we employed the following parameters for filtering low-quality reads, which included base-calling duplicates and adapter contamination: reads with ≥10% unidentified nucleotides "N"; more than 10 nucleotides aligned to the adaptor, allowing for ≤10% mismatches; and over 50% of bases with a Phred quality of <5. Ultimately, we generated 362.2 Gb of high-quality data (Table S1).

2.2 De novo assembly

We first estimated the genome size utilizing a K-mer method. We employed Jellyfish v2.2.9 (Marçais & Kingsford, 2011) to extract 17-mers information from the high-quality short

inserts (250 and 500 bp). The result of this estimation indicated that the genome size of the *T. cirrhosis* bat is ~2.13 Gb (Fig. S1; Table S2).

To assemble this new genome, 10× Genomics linked reads were assembled using supernova-2.0.0 software (Weisenfeld et al., 2017) with default parameters. Specifically, a de Bruijn graph was firstly constructed with $K = 48$ (K-mer size), by connecting successive edges that overlap each other and removing the shared K-1 bases at each junction. Next, the super graphs were formed using read pairs that share barcodes and have overlap length of more than 200-bp. Subsequently, reads-sharing barcodes with those that have assembled at contig ends, as well as reads located at nodes not yet connected, were recruited to extend edges in the super graphs. The genome assembly was generated after scaffold extension and assembly phasing. To optimize the assembly, we further employed Gapcloser v1.12 (Luo et al., 2012) for gap-filling, utilizing Illumina's short-insert paired-end reads. Finally, we applied the fragScaff program (Adey et al., 2014) to further interconnect the scaffolds generated in the preceding step, resulting in the formation of superscaffolds (Table S3).

2.3 Genome completeness assessment

To estimate the base-level accuracy and completeness, Merqury v1.3 was utilized for evaluating the quality value (QV) of genome assembly based on a K-mer based approach (Rhie et al., 2020). We also assessed the genome's completeness (Table S4) by employing the Benchmarking Universal Single-Copy Orthologs (BUSCO v4.1.4) set (Simão et al., 2015), which was used to search for 9226 mammalian BUSCO genes (mammalia_odb10).

2.4 Genome annotation

2.4.1 Repeat annotation

We employed a combination of methods, including similarity searching and ab initio prediction, to predict transposable elements (TEs) within the *T. cirrhosis* genome (Table S5). For homology annotation, we utilized RepeatMasker v4.0.5 and RepeatProteinMask v4.0.5, both with default parameters, to explore the Repbase TE library (Bergman & Quesneville, 2007). To construct a de novo repeat library for ab initio TE prediction, we employed LTR FINDER v1.07 (Xu & Wang, 2007), PILER v1.0 (Edgar & Myers, 2005), RepeatScout v1.0.5 (Price et al., 2005), and RepeatModeler v1.73 (Smit et al., 2008). This newly created library was subsequently utilized by RepeatMasker to facilitate the ab initio TE prediction process. In addition, we identified tandem repeats in the genome through the use of TRF v4.09 (Benson, 1999).

2.4.2 Gene annotation

We employed a comprehensive approach, combining homology prediction, transcript prediction, and ab initio prediction methods, to accurately predict the gene structure (Table S6). For our homology-based prediction, we utilized the bat reference protein set, as described in Tian et al. (2023). To identify potential genes, we employed tblastn (Camacho et al., 2009), searching the target genome with the protein set as queries. Subsequently, we employed Solar v0.9.6 (Sorting Out Local Alignment Results) to consolidate blast hits (Yu et al., 2006). To predict the gene structures within the corresponding genomic regions that were hit by proteins, we employed the GeneWise v2.4.1 pipeline (Birney et al., 2004). Regarding transcript prediction, we aligned the same RNA-seq data used by Tian et al. (2023) to the genome using TopHat v2.1.1 (Kim et al., 2013), which allowed us to identify putative exonic regions and splicing junctions. We then employed Cufflinks v2.2.1 (Trapnell et al., 2010) to assemble gene models from the mapped reads. For our ab initio prediction, we utilized Augustus v3.3.2 (Stanke & Waack, 2003), GlimmerHMM v3.0.4 (Majoros et al., 2004), and SNAP (Korf, 2004) to predict coding regions within repeat-masked genomes. All of these tools were trained using the PASA-T-set gene model. Additionally, we used GeneID v1.4 (Guigó, 1998) and GeneScan (Burge & Karlin, 1997) to directly predict gene models within the genome. To integrate the results from both ab initio prediction and homology-based prediction, we employed EvidenceModeler v1.1.1 (Haas et al., 2008). In this integration, we assigned weights to each piece of evidence as follows: Homology-set > Cufflinks-set > Augustus > GeneID = SNAP = GlimmerHMM = GeneScan. This comprehensive approach ensured a robust prediction of the gene structure, combining multiple methods to increase accuracy and reliability.

2.5 Phylogenetic analysis

2.5.1 Ortholog identification and alignment

In this study, we employed OrthoFinder v2.5.2 to identify putative one-to-one orthologous genes in a set of 21 genomes, which included nine leaf-nosed bats, 10 non-leaf-nosed bats, as well as human and mouse (as detailed in Table S7) (Emms & Kelly, 2019; Gutiérrez-Guerrero et al., 2020; Jebb et al., 2020; Wang et al., 2020; Blumer et al., 2022; Tian et al., 2023). A total of 1896 one-to-one orthologs were identified and extracted from the set of 21 examined species. Subsequently, we aligned the protein sequences of each ortholog using prank v.170427 with default parameters (Löytynoja, 2014). To complement these protein alignments, we generated back-translated coding DNA sequence (CDS) alignments, guided by the protein alignments, utilizing a Perl script called Epal2nal.pl v13 (Zhang et al., 2012). Following this alignment process, we isolated conserved CDS alignment blocks using the alignment trim module of FasParser v2.13.0. This extraction was performed with specific parameters, namely "Window size for trimming blocks=21" and "Window size for Gap regions=21," which were chosen to facilitate further phylogenomic analyses (Sun, 2017).

2.5.2 Phylogenetic tree inference

We concatenated the CDS alignments of the orthologs to create a supergene sequence that encompassed 3 096 477 codon positions, ensuring unambiguous alignment. For the construction of a maximum likelihood (ML) tree (Fig. S2), we employed IQ-TREE v2.0.3, using the supergene sequence as input and specified the following parameters: "-m MFP -bb 1000 –bnni." The best-fit model selected for this analysis was GTR + F + I + R7 (Hoang et al., 2017; Minh et al., 2020). To consider the impact of incomplete lineage sorting and other issues associated with the concatenation method in phylogenetic inference, we also utilized the species tree summary method in ASTRAL v.5.7.1 (Zhang et al., 2018) to infer a species tree. In detail, we first inferred the ML trees

for each gene using IQ-TREE with default parameters. Next, we summarized the distribution of ML trees and estimated the species tree using ASTRAL. At the same time, we also generated a density tree (Fig. S3A) using ggtree to show consensus and variations among all gene trees (Yu et al., 2017). The topologies of both the ML tree and the species tree (Fig. S3B) align coherently with the well-established phylogeny (Hao et al., 2023b).

2.5.3 Divergence time estimation

We employed the MCMCTree program within PAML v4.8 to conduct Bayesian estimation of species divergence times. Our input consisted of fourfold degenerate sites and a phylogenetic tree that featured six fossil calibration points (Springer et al., 2003; Teeling et al., 2005; Wang et al., 2020). In the program, we utilized the independent rates model (clock = 2) and the GTR substitution model (model = 7). Our MCMC analyses were executed for a total of 10 000 000 iterations, with samples taken every 100 iterations. We discarded the initial 2 000 000 iterations as burn-in (Yang, 2007). Finally, we utilized FigTree v1.4.4 (http://tree.bio.ed.ac.uk/software/figtree) to visualize the phylogenetic trees along with the estimated divergence times.

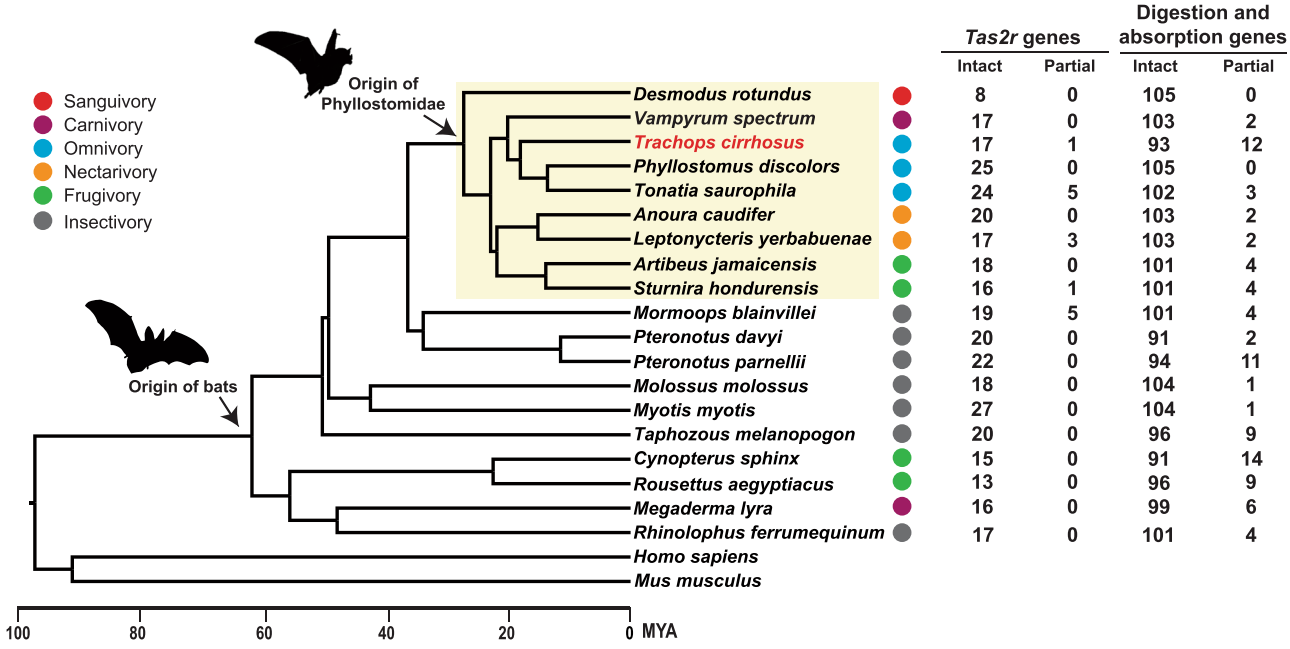
2.6 Selection tests

2.6.1 Genome-wide screening

Based on previous studies (Wilman et al., 2014; Fleming et al., 2020), we categorized Phyllostomidae bats into five diet groups: sanguivory (Common vampire bat, *Desmodus rotundus*), carnivory (Spectral bat, *Vampyrum spectrum*), omnivory (Fringe-lipped bat, *T. cirrhosis*; Stripe-headed round-eared bat, *Tonatia saurophila*; Pale spear-nosed bat, *Phyllostomus discolor*), frugivory (Jamaican fruit bat, *Artibeus jamaicensis*; Honduran yellow-shouldered bat, *Sturnira hondurensis*), and nectarivory (Lesser long-nosed bat, *Leptonycteris yerbabuenae*; Tailed tailless bat, *Anoura caudifer*). This categorization was done for ease of subsequent analysis.

Using a reciprocal Blast method with an E-value threshold of 1e−05 across 21 species (Fig. 1), we generated a gene set comprising 10 200 orthologous genes, ensuring that each ortholog included at least one species from each diet group and the outgroup. Subsequently, we created CDS alignments and identified conserved CDS alignment blocks using the method described earlier.

To identify potential genes that might be under divergent selection contributing to the dietary diversification of Phyllostomidae, we employed the clade model for the entire Phyllostomidae clade. This was achieved using CODEML in PAML 4.7 to examine whether ω (d_N/d_S ratio) differed between Phyllostomidae (foreground) and the non-Phyllostomidae (background). The clade model C (CmC), capable of detecting selection pressure on multiple branches simultaneously by allowing a set of sites to vary between two or more branches along a phylogeny, was used to test whether selection pressures varied along specific branches within the Phyllostomidae. This analysis was compared with the null-model M2a_rel, which restricts divergence within the foreground clade but allows unconstrained variation of ω (Bielawski & Yang, 2004; Yang, 2007; Weadick & Chang, 2012). Following a likelihood ratio test with a χ^2 distribution for the alternative models against the null models, we applied false discovery rate (FDR) correction to the results. Genes meeting



	Tas2r genes		Digestion and absorption genes	
	Intact	Partial	Intact	Partial
Desmodus rotundus	8	0	105	0
Vampyrum spectrum	17	0	103	2
Trachops cirrhosus	17	1	93	12
Phyllostomus discolors	25	0	105	0
Tonatia saurophila	24	5	102	3
Anoura caudifer	20	0	103	2
Leptonycteris yerbabuenae	17	3	103	2
Artibeus jamaicensis	18	0	101	4
Sturnira hondurensis	16	1	101	4
Mormoops blainvillei	19	5	101	4
Pteronotus davyi	20	0	91	2
Pteronotus parnellii	22	0	94	11
Molossus molossus	18	0	104	1
Myotis myotis	27	0	104	1
Taphozous melanopogon	20	0	96	9
Cynopterus sphinx	15	0	91	14
Rousettus aegyptiacus	13	0	96	9
Megaderma lyra	16	0	99	6
Rhinolophus ferrumequinum	17	0	101	4
Homo sapiens				
Mus musculus				

Fig. 1. Phylogenetic tree of the Phyllostomidae family showing genes associated with bitter taste (*Tas2rs*) alongside those involved in digestion and absorption pathways. Branch lengths of the tree represent divergence times estimated by MCMCTree. Circles with different colors represent feeding habits of bats. Specie marked in red is the newly sequenced species. Bootstrap values and branch lengths generated by IQtree are shown in the Fig. S2, where all nodes within the tree received 100% bootstrap support, with the exception of a single node that received 95% bootstrap support.

the criteria of FDR < 0.05, with ω in the foreground being greater than in the background, and a site class proportion > 0, were considered as the divergently selected genes. Subsequently, we uploaded the genes under divergent selection to Metascape (Zhou et al., 2019) for gene ontology (GO) enrichment analysis using all human genes in the GO database as background genes. We further used the R package ggplot2 to visualize the GO enrichment results (Wickham, 2016).

2.6.2 Bitter taste perception

We initiated this analysis by annotating and naming *Tas2r* genes in the 19 bat species (Fig. 1). To achieve this, we followed the analytical approach outlined by Jiao et al. (2018). First, we employed tblastn to search for *Tas2r* genes across the genomes of these 19 bats. As query sequences, we used complete *Tas2r* protein sequences from humans, mice, chickens, and zebrafish. We retained BLAST hits that had an e-value of 1e−10 or lower and a sequence length exceeding 300-bp for subsequent analysis. Next, we conducted open reading frame (ORF) identification on the identified hit sequences using a custom in-house script (https://github.com/xuyirano609/PipelineForGenomeAnalysisOfPhyllostomidae/tree/main/IdentificationOfTas2rGenes). A complete *Tas2r* gene was defined as one that possessed both a proper initiation codon and a stop codon, with an encoded amino acid length exceeding 270 amino acids. Moreover, we assessed all full-length candidate genes for the presence of seven transmembrane domains using the TMHMM method (Hallgren et al., 2022). Hits with disrupted ORFs were classified as pseudogenes, while those with truncated ORFs due to factors such as incomplete genome sequencing or poor genome assembly were categorized as partial *Tas2r* genes, in accordance with the criteria established by Jiao et al. (2018). To assign names and classifications to the candidate *Tas2r* genes, we employed the ML method to construct a gene tree encompassing all bat *Tas2r* genes. As reference points, we selected representative genes from 28 *Tas2r* clades in Boreoeutheria, following the approach outlined elsewhere (Hayakawa et al., 2014).

We conducted selection tests on four *Tas2r* genes (*Tas2r1*, *Tas2r2*, *Tas2r3*, and *Tas2r7*), as these genes were present in each diet group. Initially, we designated the ancestral branch of Phyllostomidae and performed tests using both the branch model and the improved branch-site model. In the branch model, the alternative hypothesis posits that the ω value of the foreground branches differs from that of the background branches, while the null hypothesis assumes equal ω values across all branches (Yang, 1998; Yang & Nielsen, 1998). For the branch-site model, the alternative hypothesis suggests a difference in ω between the foreground and background branches, along with the presence of positively selected sites. The null hypothesis aligns with the alternative model but fixes the ω of foreground branches at 1 (Yang & Nielsen, 2002; Zhang et al., 2005). In the next step, we extended our analysis to the clade model encompassing the entire Phylostomidae clade, following the same methodology as described in our genome-wide analysis. It is important to note that for *Tas2r1* and *Tas2r2*, we utilized the well-established species tree as input, while for *Tas2r3* and *Tas2r7*, we employed the gene tree generated through the ML method as input. This approach was chosen because we observed that *Tas2r3* and *Tas2r7* exist in multiple copies within individual species.

2.6.3 Digestive system

We initiated this analysis by conducting tblastn searches across the genomes of the 19 bats (Fig. 1), using human genes involved in protein digestion and absorption (map04974), carbohydrate digestion and absorption (map04973), and fat digestion and absorption (map04975) pathways as query sequences. We retained only the top blast hits, following the methodology outlined previously (Kanehisa & Goto, 2000). Furthermore, to predict the gene structure, we employed Exonerate v2.58.3 (Slater & Birney, 2005) to align the protein sequences with the best hit sequences. In cases where genes were only partially present in individual species, we further examined their fragments, seeking fragments with an identity rate higher than 70% within the genome to annotate the gene as completely as possible. Additionally, for genes that cannot be adequately annotated in individual species using human genes as references (typically due to the inability to identify the start codon or the presence of low-matching rate fragments), we employed annotated genes as queries for subsequent annotation efforts. Finally, we retained a total of 105 annotated genes that were found in at least one species in each diet group. Subsequently, we conducted the clade model analyses on these 105 genes to identify divergent selection signatures in the entire Phylostomidae clade.

2.7 Functional assays of *Tas2r1* and *Tas2r3*

2.7.1 Bitter compounds

We selected a total of 23 bitter compounds for functional assays, which included four artificial compounds and 19 natural compounds. All of these compounds were dissolved in DPBS buffer (Thermo Fisher) with a pH of 7.4. The highest concentrations, along with the source and catalog information for each bitter compound, are detailed in Table S8.

2.7.2 Construction of bat Tas2r1 and Tas2r3 expression vectors

The complete protein-coding sequences of bat *Tas2r1* and *Tas2r3* were chemically synthesized and subsequently inserted into the pcDNA3.1(+) vector, utilizing 5′-*Eco*RI and 3′-*Not*I restriction sites. To enhance protein expression in the in vitro experiments, we carried out codon optimization and incorporated a Kozak sequence before the start codon of the bat *Tas2r1* and *Tas2r3* coding sequences. The integrity of all DNA constructs was confirmed through Sanger sequencing.

2.7.3 Functional assays

The functional assays were performed following the method previously described (Hao et al., 2023a). HEK293-derived Peak Rapid Cells (PK) were cultured at 37 °C in Opti-MEM, supplemented with 6% fetal bovine serum. For calcium imaging, cells were seeded at a density of 50 000 cells per well in 96-well microplates. After 14 h, the cells were co-transfected with plasmids encoding bat Tas1r1 and Gα16-gust44 using Lipofectamine 2000. Cells used as a negative control were solely transfected with Gα16-gust44. Following a 6-h incubation, the medium was replaced once, and after an additional 18 h, the cells were washed with

DPBS. Subsequently, they were stained with Fluo-4 AM and Pluronic F-127 in the dark for 1 h. After being washed three times with DPBS, the cells were further incubated for 30 min. The microplates were then placed into a FlexStation III system to measure fluorescence changes. The FlexStation III system settings included recording fluorescence every 2 s for a total of 200 s, with the appropriate concentration of bitter compounds added at the initial 30 s. Calcium mobilization ($\Delta F/F$) was calculated as the percentage change in the fluorescence signal. This was determined by taking the difference (ΔF) between the peak fluorescence and the baseline fluorescence and dividing it by the baseline fluorescence (F). The response to bitter compounds was quantified by independently replicating all results at least three times and calculating the average of the independent replicates.

3 Results

3.1 De novo assembly of *Trachops cirrhosis* genome

The genome of the phyllostomid bat, *T. cirrhosus*, was sequenced using the Illumina HiSeq platform in conjunction with 10× Genomics technology, resulting in a total of 221.41 gigabases (Gb) and 140.83 Gb of qualified reads, respectively (Table S1). Subsequently, a genome assembly with a size of 2.3 Gb was generated, achieving a mean coverage of ~157.49 times (refer to Section 2.2 for more details). This assembly displayed a scaffold N50 of 581.2 kb and a contig N50 of 40.6 kb (Table S3). The QV of the genome was estimated to be 48.57 (error rate: 1.39×10^{-5}) using a K-mer-based approach. The assessment of the assembly using BUSCO (Benchmarking Universal Single-Copy Orthologs) indicated that 83.2% of the 9226 mammalian BUSCOs were complete (Table S4). To predict gene structures, a combination of homology prediction, transcript prediction, and ab initio prediction methods was employed. After filtering out genes of low quality (those with fewer than 50 amino acids or containing internal stop codons), a total of 21 758 protein-coding genes were successfully generated (Table S6).

3.2 Analysis of divergent selection at the genome-wide level

We compiled coding gene sequences from a total of 19 bat species, including nine phyllostomids and 10 other bats, as well as *Homo sapiens* and *Mus musculus* (Fig. 1). This comprehensive dataset enabled us to generate multiple sequence alignments for 10 200 homologous genes (refer to the Section 2.6.1 for more details). The species tree used for selection analysis was constructed using a data set of 1896 one-to-one orthologs (Figs. 1, S2, S3). We applied the clade model to identify genes that might be undergoing divergent selection, thereby contributing to the dietary diversification observed in phyllostomids. Following FDR correction and site class proportion assessment, our analyses revealed the identification of 824 genes undergoing divergent selection within the Phyllostomidae clade (Table S9).

Functional enrichment analyses of the genes under divergent selection unveiled their significant enrichment in 20 GO terms (Fig. S4; Table S10). Notably, eight of these terms were closely related to metabolism, encompassing processes such as amide metabolic process (GO:0043603), fatty acid metabolic process (GO:0006631), steroid metabolic process (GO:0008202), lipid catabolic process (GO:0016042), protein catabolic process (GO:0030163), carbohydrate transport (GO:0008643), nucleoside bisphosphate metabolic process (GO:0033865), and mRNA metabolic process (GO:0016071) (Fig. 2; Table S11). Of particular significance, 184 genes, constituting 22.3% of the genes under divergent selection, were associated with metabolism-related pathways (Table S12), highlighting the critical role of divergent selection in shaping genes related to metabolism during the dietary diversification of phyllostomids. In addition to dietary diversification, we cannot rule out other physiological adaptations that are associated with divergent selection of metabolism related genes.

3.3 Molecular adaptations of bat *Tas2r* genes

Taste is closely related to dietary selection (Jiao et al., 2018). Vertebrate *Tas2r*s encode receptors for bitter taste, granting the ability to detect bitter compounds. Given that toxic substances, like plant alkaloids and insect defensive secretions, commonly exhibit a bitter taste, the ability to sense bitterness is regarded as a crucial natural defense mechanism in animals (Hong & Zhao, 2014). We identified a total of 558 *Tas2r* genes across 19 bat species (refer to Section 2.6.2 for more details). These genes contained 349 intact genes, averaging 18 per species, with a median of 18 and ranging from eight to 27 (Figs. 1, S5; Table S13). Additionally, we found 15 partial genes, with an average of one per species, a median of 0, and a range from 0 to 5 (Fig. 1). We conducted an analysis to examine the signatures of accelerated evolution, positive selection, and divergent selection for four *Tar2r* genes (*Tas2r1*, *Tas2r2*, *Tas2r3*, and *Tas2r7*) that were present across various feeding habits in phyllostomid bats. The results revealed that all four genes did not show a significant positive selection signal when assessed using the branch-site model. However, *Tas2r1* exhibited signatures of accelerated evolution and divergent selection when evaluated with the branch model and the clade model, respectively. In contrast, *Tas2r3* displayed the signature of divergent selection when analyzed using the clade model (Table S14).

We further conducted cell-based functional assays to demonstrate functional changes in *Tas2r1* and *Tas2r3* genes among phyllostomid bats. These assays involved assessing the responsiveness of these two receptors to 23 bitter compounds, comprising 19 natural and four synthetic bitter compounds (Table S8). The quantification of calcium mobilization, based on relative fluorescence changes, revealed that the Tas2r1 receptor in phyllostomid bats exhibited a broader spectrum of perception for bitter substances when compared to nonphyllostomid bats (Fig. 3). Conversely, the Tas2r3 receptors did not demonstrate the ability to perceive any of the 23 bitter substances across all bat species. When combined with the previous study (Lu et al., 2021), we found that phyllostomid bats primarily possess the capability to detect six distinct bitter substances. These substances include two synthetic compounds (difenidol hydrochloride and denatonium benzoate) and four natural compounds (camphor, papaverine, esculine, and chloramphenicol) (Figs. 3, S6; Table S8). Among these, carnivorous bats can perceive four bitter substances, while

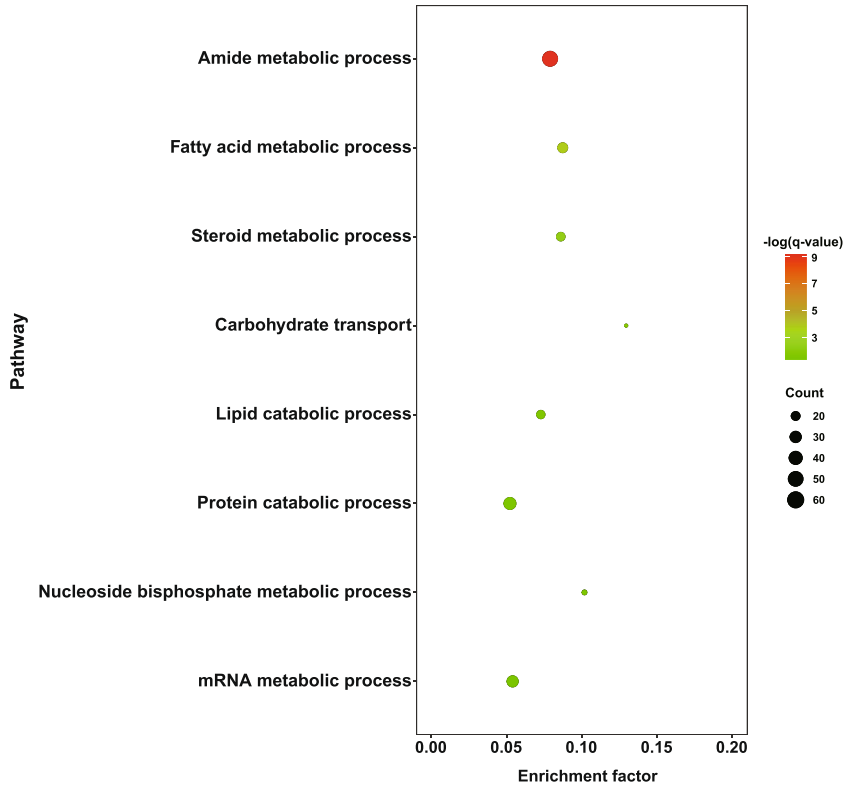


Fig. 2. Gene Ontology enrichment for metabolism-related terms among genes exhibiting divergent selection. The vertical axis denotes the pathway names, while the horizontal axis represents the enrichment factor. The size of each circle corresponds to the number of genes enriched within the pathway. The color of the circle signifies the significance of enrichment in the pathway, as indicated by the *q*-value, representing an adjusted *P*-value. The enrichment factor is calculated as the ratio of the gene count in the input to the gene count in the reference.

omnivorous bats, nectar-eating bats, and fruit-eating bats can perceive three bitter substances. Vampire bats, on the other hand, can perceive two bitter substances (Figs. 3, S6; Table S8). Our findings highlight differences both in the quantity and types of bitter substances perceived by the Tas2r1 receptor, as well as in the capability to sense bitter compounds, among Phyllostomidae bats with varying feeding habits. These observations are consistent with the divergent selection signatures of the *Tas2r1* gene.

3.4 Divergent selection of digestive system

The diverse food sources within the Phyllostomidae family result in significant differences in nutritional intake, such as high-protein diet (sanguivory, carnivory), high-carbohydrate diet (frugivory, nectarivory) and high-fat diet (carnivory). These dietary differences are considerable challenges to the digestive system of phyllostomids. To investigate the contribution of divergent selection of genes associated with the digestive system to the dietary diversification of Phyllostomidae, we manually annotated a total of 105 genes (Fig. 1; Table S15) involved in protein, carbohydrate and fat digestion and absorption among all bat species (refer to the Section 2.6.3 for more details). Subsequently, we performed the clade model analyses on the dataset of 105 genes, uncovering 14 genes that exhibited signatures of divergent selection. Notably, three of these genes are involved in protein digestion and absorption (*PGA3*, *CPB1*, and *MEP1B*), five are associated with carbohydrate digestion and absorption (*LCT*, *SI*, *SLC5A1*, *SLC2A2*, and *SLC2A5*), and six are linked to fat digestion and absorption (*PLA2G5*, *CD36*, *SLC27A4*, *ABCG5*, *MOGAT2*, and *MOGAT3*) (Fig. 4; Table S16).

4 Discussion

In this study, we reported a novel genome assembly of *Trachops cirrhosis*, a representative omnivorous leaf-nosed bat. Through comparative genomic analysis, we found that phyllostomid bats had undergone extensive divergent selection in taste, digestive system, and metabolism, which are associated with their dietary diversification.

In our analysis, we included three omnivorous phyllostomid bat species that effectively represent the diverse distribution of food types within an omnivorous diet. To enhance our study, we sequenced the genome of *T. cirrhosus*, which shows a larger preference for vertebrates and thereby strengthens our dataset. Considering the scarcity of *T. cirrhosus* samples and the inability of their DNA quality to meet the requirements for long-read sequencing, we used a genome assembly strategy combining 10× Genomics and Illumina pair-end sequencing.

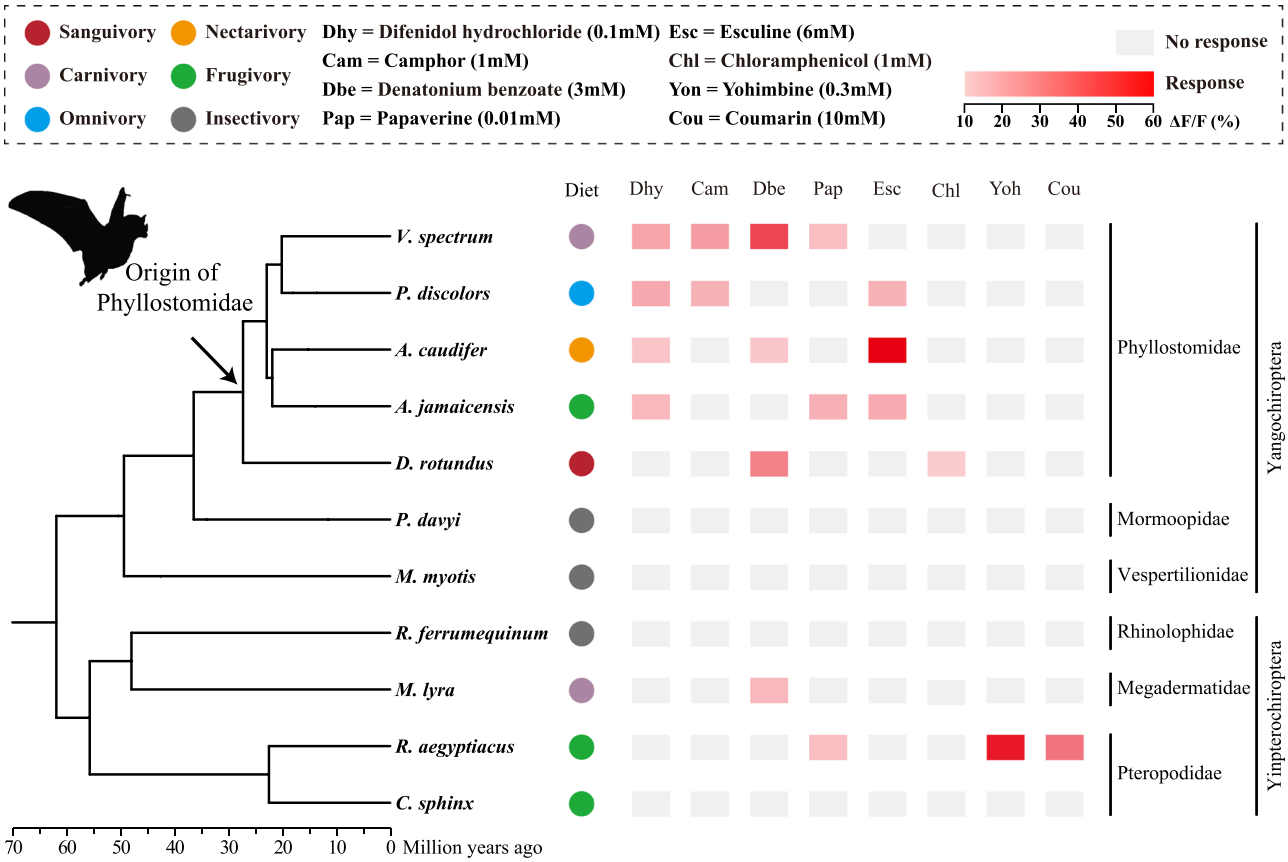


Fig. 3. Functional diversification of Tas2r1 receptors in Phyllostomidae. HEK293 cells transiently expressing bat *Tas2r1* genes were subjected to functional assays to assess their responses to eight bitter compounds. The responses of bat Tas2r1 receptors to these bitter substances were quantified using calcium mobilization (Δ*F*/*F*) values, visually depicted by rectangles of varying colors. Specifically, gray rectangles indicate no response to bitter substances, while red rectangles represent a response. The Δ*F*/*F* values, differentiated by distinct colors, are illustrated on the Δ*F*/*F* bar. Additionally, circles of varying colors correspond to the feeding habits of bats.

The evolution of taste is closely related to dietary changes, and it is generally believed that mammals can avoid potentially toxic foods by sensing bitter taste through bitter taste receptors, because poisonous substances usually taste bitter (Glendinning, 1994; Feng et al., 2014; Jiao et al., 2021). We examined the functional changes of the bitter taste receptors Tas2r1 and Tas2r3 during the dietary diversification of Phyllostomidae by molecular evolutionary analysis of bitter taste receptor genes and cell-based functional assays. Our investigation revealed that despite *Tas2r3* showing the signature of divergent selection, its encoding receptors of all bats were unable to perceive the bitter substances examined in this study. This could potentially be attributed to the limited perception range of these receptors, as the Tas2r3 receptor in humans only sensed a single synthetic compound chloroquine, an antimalarial drug (Meyerhof et al., 2009). Tas2r1 receptors in phyllostomid bats demonstrated a broader range of bitter substance perception. However, this heightened bitter taste perception ability was not observed in their close relatives, the Mormoopidae bats. We speculated that the bitter taste perception ability of Tas2r1 receptor in Phyllostomidae bats was a functional innovation that occurred in the common ancestor of Phyllostomidae bats, and its function had also changed during dietary diversification of Phyllostomidae bats, although this hypothesis is awaiting additional evidence for validation. Jiao et al. (2018) found no clear correlation between bat diet and the number of *Tas2r* genes, therefore, we did not examine the relationship between the diet of Phyllostomidae bats and the number of bitter taste receptors in this study.

Digestive enzymes adapt to diet as nutrient intake alters (Corring, 1980). The nutrient intake of Phyllostomidae bats shifted across various dietary patterns, transitioning from an insect-eating diet to meat-eating, fruit-eating, and nectar-eating diets (Potter et al., 2021). Consequently, the activities of a series of enzymes were altered, among which the activities of intestinal sucrase and maltase were increased in frugivorous and nectarivous bats, whereas vampire bats virtually absent intestinal maltase and sucrase activity, and the enzyme activity of trehalase in insectivorous bats was significantly higher than that of bats with other feeding habits (Schondube et al., 2001; Jiao et al., 2019). We demonstrated the contribution of the digestive system to the dietary diversification of Phyllostomidae by analyzing the divergent selection signatures of genes involved in

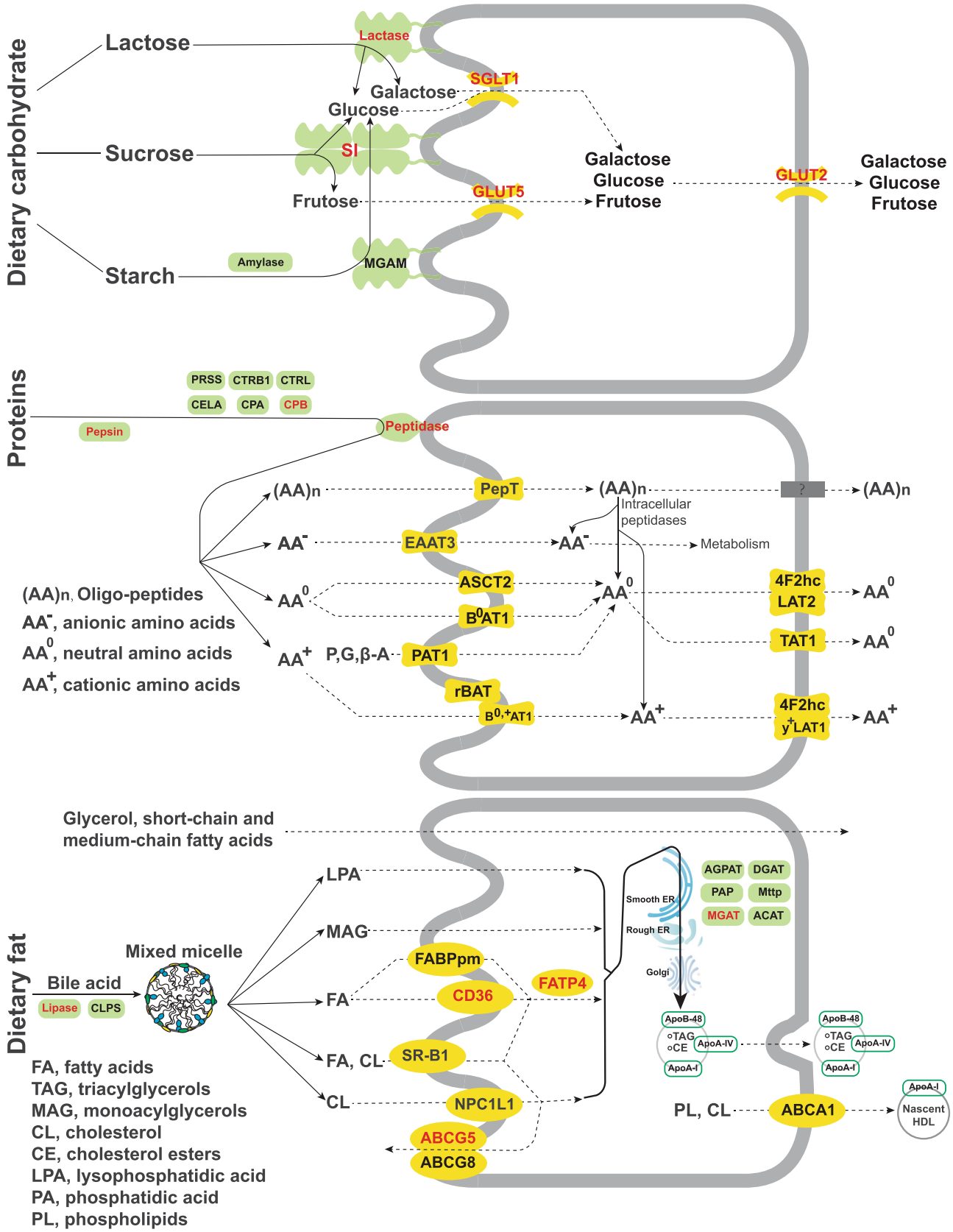


Fig. 4. Continued

carbohydrate, protein and fat digestion and absorption pathways. Specifically, three genes associated with protein digestion and absorption include *PGA3*, *CPB1*, and *MEP1B*. In each bat species, *PGA3* is present as a single copy, while in humans, there exists a cluster of related genes (*PGA3*, *PGA4*, *PGA5*); *PGA3* encodes for Pepsin (as depicted in Fig. 4), which functions in the breakdown of dietary proteins (Athauda et al., 1989). *CPB1*, which encodes a carboxypeptidase (CPB), plays a vital role in breaking down peptide chains and releasing free amino acids (Aloy et al., 1998). *MEP1B* encodes for a multidomain zinc metalloprotease (peptidase in Fig. 4), which is involved in the hydrolysis of various peptide and protein substrates (Marín, 2015). Only three genes under divergent selection were identified in protein digestion and absorption pathways, implying that the pathway may have not undergone many functional changes, and the study of aminopeptidase N activity seemed to support this notion, as its activity showed no difference between animalivorous and plant-eating bats (Fleming et al., 2020). Although few genes involved in protein digestion and absorption underwent divergent selection, we found more genes related to the protein catabolic process were under divergent selection in our genome-wide analysis. Five genes involved in the carbohydrate digestion and absorption pathway were identified. These genes encompass *LCT*, which encodes lactase (lactase in Fig. 4). Lactase is an integral component of the plasma membrane and is responsible for the hydrolysis of lactose into galactose and glucose (Kuokkanen et al., 2006). Furthermore, *SI* encodes sucrase-isomaltase (SI), which is indispensable for the digestion of various dietary carbohydrates, including starch, sucrose and isomaltose (Kuokkanen et al., 2006). The divergent selection feature of *SI* can be supported by the above-mentioned activity changes of maltase and sucrase in bats with different feeding habits. Additionally, three genes *SLC5A1, SLC2A2*, and *SLC2A5* encode sodium-dependent glucose transporter (SGLT1), glucose transporter (GLUT2), and fructose transporter (GLUT5), respectively. These transporters play pivotal roles in facilitating the transcellular transport of glucose and fructose (Barone et al., 2009; Thorens, 2015; Han et al., 2022). Although paracellular transport were considered to be dominant in glucose transport in bats, and the proportion of glucose uptake can reach 70% (Caviedes-Vidal et al., 2008; Price et al., 2013, 2015; Brun et al., 2014; Rodriguez-Peña et al., 2016). We speculate that it is plausible that the functions of the hexose transporter in Phyllostomatidae bats were altered under extremely different carbohydrate diets. Six genes within the fat digestion and absorption pathway were examined, featuring *PLA2G5*, encoding a secretory phospholipase (lipase in Fig. 4). This enzyme catalyzes the hydrolysis of membrane phospholipids, generating lysophospholipids and free fatty acids (Chen et al., 1994). Moreover, two genes, *CD36* and *SLC27A4*, encode fatty acid transporters (*CD36* and *FATP4*, respectively), which are crucial for the transport of long-chain fatty acids across the plasma membrane (Schwenk et al., 2010). The ATP-binding cassette transporter-coding gene, *ABCG5*, plays a pivotal role in mediating the selective transport of dietary plant sterols and cholesterol across the enterocytes (Yu et al., 2014). Lastly, *MOGAT2* and *MOGAT3* are implicated in the resynthesis of triacylglycerol in enterocytes, a process crucial for dietary fat absorption in the small intestine (as depicted by "MGAT" in Fig. 4) (Cheng et al., 2003; Zhang et al., 2014). In this study, three divergently selected genes associated with the fat digestion and absorption pathway were involved in the transport of long-chain fatty acids and cholesterol, which implied that phyllostomids may deal with fat intake differences through functional changes of genes involved in lipid transport.

Through genome-wide identification, we discovered the functional changes of metabolism-related genes in the process of dietary diversification in Phyllostomidae. Three genes we identified, *GSTT4*, *SOD1*, and *CAT*, which respectively encode glutathione-S-transferase theta 4, superoxide dismutase 1, and catalase, have been confirmed to show differential enzyme activities in sanguivorous and frugivorous bats in the Phyllostomidae family, among which the activity of the glutathione-S-transferase was higher in kidneys of vampire bats, and the activities of superoxide dismutase and catalase were higher in fruit-eating bats (Linhares et al., 2021).

5 Conclusion

In summary, we have presented a novel genome assembly of *Trachops cirrhosis*, an exemplary omnivorous bat within the Phyllostomidae family. Through the integration of genomic data from eight other phyllostomid bats and 10 non-Phyllostomidae bats, we conducted comparative genomic analyses aimed at elucidating the genetic basis of dietary diversification in Phyllostomidae. Our study has revealed extensive divergent selection signatures within genes associated with taste, digestion, absorption, and metabolism-related pathways. Subsequent cell-based functional assays have not only confirmed but also validated the functional diversification of the bitter taste receptor Tas2r1, highlighting its crucial role in bitter taste perception specific to Phyllostomidae. Our genomic and functional analyses has led us to propose that the divergent selection of genes related to taste, digestion, absorption, and metabolism plays a pivotal role in steering the remarkable dietary diversification observed in Phyllostomidae. This work not only sheds light on the genetic mechanisms underlying dietary adaptations in Phyllostomidae bats but also significantly

Fig. 4. Diagram depicting the evolution of digestive system-related pathways in Phyllostomidae. From top to bottom, the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways are carbohydrate digestion and absorption, protein digestion and absorption, fat digestion and absorption. In each pathway, proteins that are encoded by genes under divergent selection are marked red, the shapes of the cyan background represent the enzymes, the yellow background shapes represent the transport proteins, the solid arrows represent the metabolic process of the substance, and the dotted arrows represent the transfer direction of substance.

advances our comprehension of their extraordinary adaptive radiation linked to dietary diversification.

Acknowledgements

This study was supported by National Key Research and Development Program of China (2021YFF0702004), National Natural Science Foundation of China (32270436), and Natural Science Foundation of the Hubei Province (2023AFA015).

Data Availability Statement

All genome sequencing raw reads have been deposited into the NCBI Sequence Read Archive and GenBank under the BioProject accession number PRJNA1025356.

References

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

Aloy P, Catasús L, Villegas V, Reverter D, Vendrell J, Avilés FX. 1998. Comparative analysis of the sequences and three-dimensional models of human procarboxypeptidases A1, A2 and B. *Biological Chemistry* 379: 149–155.

Arbour JH, Curtis AA, Santana SE. 2019. Signatures of echolocation and dietary ecology in the adaptive evolution of skull shape in bats. *Nature Communications* 10: 2036.

Athauda SB, Tanji M, Kageyama T, Takahashi K. 1989. A comparative study on the NH2-terminal amino acid sequences and some other properties of six isozymic forms of human pepsinogens and pepsins. *Journal of Biochemistry* 106: 920–927.

Barone S, Fussell SL, Singh AK, Lucas F, Xu J, Kim C, Wu X, Yu Y, Amlal H, Seidler U, Zuo J, Soleimani M. 2009. Slc2a5 (Glut5) is essential for the absorption of fructose in the intestine and generation of fructose-induced hypertension. *Journal of Biological Chemistry* 284: 5056–5066.

Benson G. 1999. Tandem repeats finder: A program to analyze DNA sequences. *Nucleic Acids Research* 27: 573–580.

Bergman CM, Quesneville H. 2007. Discovering and detecting transposable elements in genome sequences. *Briefings in Bioinformatics* 8: 382–392.

Bielawski JP, Yang Z. 2004. A maximum likelihood method for detecting functional divergence at individual codon sites, with application to gene family evolution. *Journal of Molecular Evolution* 59: 121–132.

Birney E, Clamp M, Durbin R. 2004. Genewise and genomewise. *Genome Research* 14: 988–995.

Blumer M, Brown T, Freitas MB, Destro AL, Oliveira JA, Morales AE, Schell T, Greve C, Pippel M, Jebb D, Hecker N, Ahmed AW, Kirilenko BM, Foote M, Janke A, Lim BK, Hiller M. 2022. Gene losses in the common vampire bat illuminate molecular adaptations to blood feeding. *Science Advances*. 8: eabm6494.

Brawand D, Wagner CE, Li YI, Malinsky M, Keller I, Fan S, Simakov O, Ng AY, Lim ZW, Bezault E. 2014. The genomic substrate for adaptive radiation in African cichlid fish. *Nature* 513: 375–381.

Brun A, Price ER, Gontero-Fourcade MN, Fernandez-Marinone G, Cruz-Neto AP, Karasov WH, Caviedes-Vidal E. 2014. High paracellular nutrient absorption in intact bats is associated with high paracellular permeability in perfused intestinal segments. *The Journal of Experimental Biology* 217: 3311–3317.

Burge C, Karlin S. 1997. Prediction of complete gene structures in human genomic DNA. *Journal of Molecular Biology* 268: 78–94.

Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. 2009. BLAST+: Architecture and applications. *BMC Bioinformatics* 10: 421.

Caviedes-Vidal E, Karasov WH, Chediack JG, Fasulo V, Cruz-Neto AP, Otani L. 2008. Paracellular absorption: A bat breaks the mammal paradigm. *PLoS One* 3: e1425.

Chen J, Engle SJ, Seilhamer JJ, Tischfield JA. 1994. Cloning and recombinant expression of a novel human low molecular weight Ca(2+)-dependent phospholipase A2. *The Journal of Biological Chemistry* 269: 2365–2368.

Cheng D, Nelson TC, Chen J, Walker SG, Wardwell-Swanson J, Meegalla R, Taub R, Billheimer JT, Ramaker M, Feder JN. 2003. Identification of acyl coenzyme A: monoacylglycerol acyltransferase 3, an intestinal specific enzyme implicated in dietary fat absorption. *The Journal of Biological Chemistry* 278: 13611–13614.

Corring T. 1980. The adaptation of digestive enzymes to the diet: Its physiological significance. *Reproduction, Nutrition, Development* 20: 1217–1235.

Datzmann T, von Helversen O, Mayer F. 2010. Evolution of nectarivory in phyllostomid bats (Phyllostomidae Gray, 1825, Chiroptera: Mammalia). *BMC Evolutionary Biology* 10: 1–14.

Dumont ER, Dávalos LM, Goldberg A, Santana SE, Rex K, Voigt CC. 2012. Morphological innovation, diversification and invasion of a new adaptive zone. *Proceedings of the Royal Society B: Biological Sciences* 279: 1797–1805.

Edgar RC, Myers EW. 2005. PILER: Identification and classification of genomic repeats. *Bioinformatics* 21: i152–158.

Emms DM, Kelly S. 2019. OrthoFinder: Phylogenetic orthology inference for comparative genomics, *Genome Biology* 20: 238.

Feng P, Zheng J, Rossiter SJ, Wang D, Zhao H. 2014. Massive losses of taste receptor genes in toothed and baleen whales. *Genome Biology and Evolution* 6: 1254–1265.

Fleming TH, Dávalos LM, Mello MAR. 2020. *Phyllostomid bats: A unique mammalian radiation*. Chicago: University of Chicago Press.

Gautam P. 2020. Divergent evolution. In: Vonk J, Shackelford T eds. *Encyclopedia of animal cognition and behavior*. Cham: Springer International Publishing. 1–8.

Givnish TJ. 2015. Adaptive radiation versus "radiation" and "explosive diversification": Why conceptual distinctions are fundamental to understanding evolution. *New Phytologist* 207: 297–303.

Glendinning JI. 1994. Is the bitter rejection response always adaptive? *Physiology & Behavior* 56: 1217–1227.

Grant PR. 1999. *Ecology and evolution of Darwin's finches*. Princeton: Princeton University Press.

Grant PR, Grant BR. 2011. *How and why species multiply: The radiation of Darwin's finches*. Princeton: Princeton University Press.

Guigó R. 1998. Assembling genes from predicted exons in linear time with dynamic programming. *Journal of Computational Biology* 5: 681–702.

Gutiérrez-Guerrero YT, Ibarra-Laclette E, Martínez Del Río C, Barrera-Redondo J, Rebollar EA, Ortega J, León-Paniagua L, Urrutia A, Aguirre-Planter E, Eguiarte LE. 2020. Genomic consequences of dietary diversification and parallel evolution due to nectarivory in leaf-nosed bats. *Gigascience* 9: giaa059.

Gutierrez E, Schott RK, Preston MW, Loureiro LO, Lim BK, Chang BS. 2018. The role of ecological factors in shaping bat cone opsin evolution. *Proceedings of the Royal Society B: Biological Sciences* 285: 20172835.

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 Biology* 9: R7.

Hallgren J, Tsirigos KD, Pedersen MD, Almagro Armenteros JJ, Marcatili P, Nielsen H, Krogh A, Winther O. 2022. DeepTMHMM predicts alpha and beta transmembrane proteins using deep neural networks. *BioRxiv*: 2022.04.08.487609.

Han L, Qu Q, Aydin D, Panova O, Robertson MJ, Xu Y, Dror RO, Skiniotis G, Feng L. 2022. Structure and mechanism of the SGLT family of glucose transporters. *Nature* 601: 274–279.

Hao X, Jiao H, Zou D, Li Q, Yuan X, Liao W, Jiang P, Zhao H. 2023a. Evolution of bitter receptor genes and ontogenetic dietary shift in a frog. *Proceedings of the National Academy of Sciences of the United States of America* 120: e2218183120.

Hao X, Lu Q, Zhao H. 2023b. A molecular phylogeny for all 21 families within Chiroptera (bats). *Integrative Zoology*. https://doi.org/10.1111/1749-4877.12772. Online ahead of print.

Hayakawa T, Suzuki-Hashido N, Matsui A, Go Y. 2014. Frequent expansions of the bitter taste receptor gene repertoire during evolution of mammals in the Euarchontoglires clade. *Molecular Biology and Evolution* 31: 2018–2031.

Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS. 2017. UFBoot2: Improving the Ultrafast Bootstrap approximation. *Molecular Biology and Evolution* 35: 518–522.

Hong W, Zhao H. 2014. Vampire bats exhibit evolutionary reduction of bitter taste receptor genes common to other bats. *Proceedings of the Royal Society B: Biological Sciences* 281: 20141079.

Jebb D, Huang Z, Pippel M, Hughes GM, Lavrichenko K, Devanna P, Winkler S, Jermiin LS, Skirmuntt EC, Katzourakis A, Burkitt-Gray L, Ray DA, Sullivan KAM, Roscito JG, Kirilenko BM, Dávalos LM, Corthals AP, Power ML, Jones G, Ransome RD, Dechmann DKN, Locatelli AG, Puechmaille SJ, Fedrigo O, Jarvis ED, Hiller M, Vernes SC, Myers EW, Teeling EC. 2020. Six reference-quality genomes reveal evolution of bat adaptations. *Nature* 583: 578–584.

Jiao H, Wang Q, Wang B-J, Li K, Lövy M, Nevo E, Li Q, Su W, Jiang P, Zhao H. 2021. Local adaptation of bitter taste and ecological speciation in a wild mammal. *Molecular Biology and Evolution* 38: 4562–4572.

Jiao H, Wang Y, Zhang L, Jiang P, Zhao H. 2018. Lineage-specific duplication and adaptive evolution of bitter taste receptor genes in bats. *Molecular Ecology* 27: 4475–4488.

Jiao H, Zhang L, Xie H-W, Simmons NB, Liu H, Zhao H. 2019. Trehalase gene as a molecular signature of dietary diversification in mammals. *Molecular Biology and Evolution* 36: 2171–2183.

Kanehisa M, Goto S. 2000. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Research* 28: 27–30.

Karasov WH, Douglas AE. 2013. Comparative digestive physiology. *Comprehensive Physiology* 3: 741.

Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. 2013. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biology* 14: R36.

Koblmüller S, Schliewen UK, Duftner N, Sefc KM, Katongo C, Sturmbauer C. 2008. Age and spread of the haplochromine cichlid fishes in Africa. *Molecular Phylogenetics and Evolution* 49: 153–169.

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

Kries K, Barros MA, Duytschaever G, Orkin JD, Janiak MC, Pessoa DM, Melin AD. 2018. Colour vision variation in leaf-nosed bats (Phyllostomidae): Links to cave roosting and dietary specialization. *Molecular Ecology* 27: 3627–3640.

Kuokkanen M, Kokkonen J, Enattah NS, Ylisaukko-Oja T, Komu H, Varilo T, Peltonen L, Savilahti E, Jarvela I. 2006. Mutations in the translated region of the lactase gene (LCT) underlie congenital lactase deficiency. *American Journal of Human Genetics* 78: 339–344.

Lamichhaney S, Berglund J, Almén MS, Maqbool K, Grabherr M, Martinez-Barrio A, Promerová M, Rubin C-J, Wang C, Zamani N. 2015. Evolution of Darwin's finches and their beaks revealed by genome sequencing. *Nature* 518: 371–375.

Li L, Chi H, Liu H, Xia Y, Irwin DM, Zhang S, Liu Y. 2018. Retention and losses of ultraviolet-sensitive visual pigments in bats. *Scientific Reports* 8: 11933.

Linhares BS, Ribeiro SP, de Freitas RMP, Puga L, Sartori SSR, Freitas MB. 2021. Aspects regarding renal morphophysiology of fruit-eating and vampire bats. *Zoology* 144: 125861.

Löytynoja A. 2014. Phylogeny-aware alignment with PRANK. *Methods in Molecular Biology* 1079: 155–170.

Lu Q, Jiao H, Wang Y, Norbu N, Zhao H. 2021. Molecular evolution and deorphanization of bitter taste receptors in a vampire bat. *Integrative Zoology* 16: 659–669.

Luo R, Liu B, Xie Y, Li Z, Huang W, Yuan J, He G, Chen Y, Pan Q, Liu Y, Tang J, Wu G, Zhang H, Shi Y, Liu Y, Yu C, Wang B, Lu Y, Han C, Cheung DW, Yiu SM, Peng S, Xiaoqian Z, Liu G, Liao X, Li Y, Yang H, Wang J, Lam TW, Wang J. 2012. SOAPdenovo2: An empirically improved memory-efficient short-read de novo assembler. *Gigascience* 1: 18.

Majoros WH, Pertea M, Salzberg SL. 2004. TigrScan and GlimmerHMM: Two open source ab initio eukaryotic gene-finders. *Bioinformatics* 20: 2878–2879.

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

Marín I. 2015. Origin and diversification of Meprin proteases. *PLoS One* 10: e0135924.

Meyerhof W, Batram C, Kuhn C, Brockhoff A, Chudoba E, Bufe B, Appendino G, Behrens M. 2009. The molecular receptive ranges of human *TAS2R* bitter taste receptors. *Chemical Senses* 35: 157–170.

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. *Molecular Biology and Evolution* 37: 1530–1534.

Potter JHT, Davies KTJ, Yohe LR, Sanchez MKR, Rengifo EM, Struebig M, Warren K, Tsagkogeorga G, Lim BK, Dos Reis M, Dávalos LM, Rossiter SJ. 2021. Dietary diversification and specialization in Neotropical bats facilitated by early molecular evolution. *Molecular Biology and Evolution* 38: 3864–3883.

Price AL, Jones NC, Pevzner PA. 2005. De novo identification of repeat families in large genomes. *Bioinformatics* 21: i351–i358.

Price ER, Brun A, Fasulo V, Karasov WH, Caviedes-Vidal E. 2013. Intestinal perfusion indicates high reliance on paracellular nutrient absorption in an insectivorous bat *Tadarida brasiliensis*. *Comparative Biochemistry and Physiology. Part A, Molecular & Integrative Physiology* 164: 351–355.

Price ER, Brun A, Gontero-Fourcade M, Fernández-Marinone G, Cruz-Neto AP, Karasov WH, Caviedes-Vidal E. 2015. Intestinal water

absorption varies with expected dietary water load among bats but does not drive paracellular nutrient absorption. *Physiological and Biochemical Zoology* 88: 680–684.

Rhie A, Walenz BP, Koren S, Phillippy AM. 2020. Merqury: Reference-free quality, completeness, and phasing assessment for genome assemblies. *Genome Biology* 21: 245.

Rodriguez-Peña N, Price ER, Caviedes-Vidal E, Flores-Ortiz CM, Karasov WH. 2016. Intestinal paracellular absorption is necessary to support the sugar oxidation cascade in nectarivorous bats. *The Journal of Experimental Biology* 219: 779–782.

Rundle HD, Nosil P. 2005. Ecological speciation. *Ecology Letters* 8: 336–352.

Schluter D. 2000. *The ecology of adaptive radiation*. Oxford: Oxford University Press.

Schluter D. 2001. Ecology and the origin of species. *Trends in Ecology & Evolution* 16: 372–380.

Schondube JE, Herrera-M LG, Martínez del Rio C. 2001. Diet and the evolution of digestion and renal function in phyllostomid bats. *Zoology* 104: 59–73.

Schwenk RW, Holloway GP, Luiken JJ, Bonen A, Glatz JF. 2010. Fatty acid transport across the cell membrane: regulation by fatty acid transporters. *Prostaglandins Leukot Essent Fatty Acids* 82: 149–154.

Seehausen O. 2006. African cichlid fish: A model system in adaptive radiation research. *Proceedings of the Royal Society B: Biological Sciences* 273: 1987–1998.

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

Simoes BF, Foley NM, Hughes GM, Zhao H, Zhang S, Rossiter SJ, Teeling EC. 2019. As blind as a bat? Opsin phylogenetics illuminates the evolution of color vision in bats. *Molecular Biology and Evolution* 36: 54–68.

Slater GSC, Birney E. 2005. Automated generation of heuristics for biological sequence comparison. *BMC Bioinformatics* 6: 31.

Smit AF, Hubley R, Green P. 2008. RepeatModeler Open-1.0[online]. Available from www.repeatmasker.org. [accessed 13 October 2020].

Springer MS, Murphy WJ, Eizirik E, O'Brien SJ. 2003. Placental mammal diversification and the Cretaceous-Tertiary boundary. *Proceedings of the National Academy of Sciences of the United States of America* 100: 1056–1061.

Stanke M, Waack S. 2003. Gene prediction with a hidden Markov model and a new intron submodel. *Bioinformatics* 19(Suppl 2): ii215–i225.

Sun YB. 2017. FasParser: A package for manipulating sequence data. *Zoological Research* 38: 110–112.

Teeling EC, Springer MS, Madsen O, Bates P, O'Brien SJ, Murphy WJ. 2005. A molecular phylogeny for bats illuminates biogeography and the fossil record. *Science* 307: 580–584.

Thorens B. 2015. GLUT2, glucose sensing and glucose homeostasis. *Diabetologia* 58: 221–232.

Tian S, Zeng J, Jiao H, Zhang D, Zhang L, Lei CQ, Rossiter SJ, Zhao H. 2023. Comparative analyses of bat genomes identify distinct evolution of immunity in Old World fruit bats. *Science Advances* 9: eadd0141.

Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, Salzberg SL, Wold BJ, Pachter L. 2010. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nature Biotechnology* 28: 511–515.

Wang K, Tian S, Galindo-González J, Dávalos LM, Zhang Y, Zhao H. 2020. Molecular adaptation and convergent evolution of frugivory in Old World and neotropical fruit bats. *Molecular Ecology* 29: 4366–4381.

Weadick CJ, Chang BS. 2012. An improved likelihood ratio test for detecting site-specific functional divergence among clades of protein-coding genes. *Molecular Biology and Evolution* 29: 1297–1300.

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

Wickham H. 2016. *ggplot2: Elegant eraphics for data analysis*. Cham: Springer International Publishing.

Wilman H, Belmaker J, Simpson J, de la Rosa C, Rivadeneira MM, Jetz W. 2014. EltonTraits 1.0: Species-level foraging attributes of the world's birds and mammals: Ecological Archives E095-178. *Ecology* 95: 2027.

Wu J, Jiao H, Simmons NB, Lu Q, Zhao H. 2018. Testing the sensory trade-off hypothesis in New World bats. *Proceedings of the Royal Society B: Biological Sciences* 285: 20181523.

Xu Z, Wang H. 2007. LTR_FINDER: An efficient tool for the prediction of full-length LTR retrotransposons. *Nucleic Acids Research* 35: W265–W268.

Yang Z. 1998. Likelihood ratio tests for detecting positive selection and application to primate lysozyme evolution. *Molecular Biology and Evolution* 15: 568–573.

Yang Z. 2007. PAML 4: Phylogenetic analysis by maximum likelihood. *Molecular Biology and Evolution* 24: 1586–1591.

Yang Z, Nielsen R. 1998. Synonymous and nonsynonymous rate variation in nuclear genes of mammals. *Journal of Molecular Evolution* 46: 409–418.

Yang Z, Nielsen R. 2002. Codon-substitution models for detecting molecular adaptation at individual sites along specific lineages. *Molecular Biology and Evolution* 19: 908–917.

Yu G, Smith DK, Zhu H, Guan Y, Lam TT-Y. 2017. ggtree: An r package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods in Ecology and Evolution* 8: 28–36.

Yu XH, Qian K, Jiang N, Zheng XL, Cayabyab FS, Tang CK. 2014. ABCG5/ABCG8 in cholesterol excretion and atherosclerosis. *Clinica Chimica Acta* 428: 82–88.

Yu XJ, Zheng HK, Wang J, Wang W, Su B. 2006. Detecting lineage-specific adaptive evolution of brain-expressed genes in human using rhesus macaque as outgroup. *Genomics* 88: 745–751.

Zhang C, Rabiee M, Sayyari E, Mirarab S. 2018. ASTRAL-III: Polynomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinformatics* 19: 153.

Zhang J, Nielsen R, Yang Z. 2005. Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. *Molecular Biology and Evolution* 22: 2472–2479.

Zhang J, Xu D, Nie J, Cao J, Zhai Y, Tong D, Shi Y. 2014. Monoacylglycerol acyltransferase-2 is a tetrameric enzyme that selectively heterodimerizes with diacylglycerol acyltransferase-1. *The Journal of Biological Chemistry* 289: 10909–10918.

Zhang Z, Xiao J, Wu J, Zhang H, Liu G, Wang X, Dai L. 2012. ParaAT: A parallel tool for constructing multiple protein-coding DNA alignments. *Biochemical and Biophysical Research Communications* 419: 779–781.

Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, Benner C, Chanda SK. 2019. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nature Communications* 10: 1523.

Supplementary Material

The following supplementary material is available online for this article at http://onlinelibrary.wiley.com/doi/10.1111/jse.13059/suppinfo:

Fig. S1. A preliminary genome survey of *Trachops cirrhosus* based on k-mer analysis ($k = 17$) with Illumina paired-end data.

Fig. S2. ML phylogeny generated by IQ tree (log-likelihood = −15 369 407.9053). Numbers at the nodes represent bootstrap values.

Fig. S3. Topology of the species tree. **A**, A density tree generated using ggtree. **B**, A species tree topology (final normalized quartet score: 0.904) inferred by ASTRAL based on the species tree summary method. The branch lengths of the tree represent coalescent units, and the numbers at the nodes indicate support values.

Fig. S4. Gene ontology (GO) enrichment for genes under divergent selection.

Fig. S5. Gene tree constructed with all *Tas2r* genes. The red background represents the *Tas2r* genes in humans. The abbreviation of a bat's Latin name consists of the first letter of its genus name and the first three letters of its species name.

Fig. S6. Quantitative analysis of responses of bat *Tas2r1* receptors to seven bitter compounds (mean ± SE, ***$P < 0.001$, **$P < 0.01$, *$P < 0.05$, one-way ANOVA).

Table S1. Summary of sequencing data.

Table S2. Estimation of the *Trachops cirrhosus* genome size using K-mer analysis.

Table S3. Summary of the *Trachops cirrhosus* genome assembly.

Table S4. Completeness assessment for the *Trachops cirrhosus* genome.

Table S5. Summary of the transposable elements (TEs).

Table S6. Statistics of gene structures.

Table S7. Genome data sources of all analyzed species.

Table S8. Responses of 23 different bitter compounds with *Tas2r1* receptors and their corresponding maximum concentrations in functional assays.

Table S9. Genes under divergent selection ($n = 824$).

Table S10. Gene ontology (GO) enrichment of the genes under divergent selection.

Table S11. Metabolism-related gene ontology (GO) terms.

Table S12. Genes involved in eight metabolism-related gene ontology (GO) terms ($n = 184$).

Table S13. Distribution of *Tas2r* genes in bats.

Table S14. Adaptive evolution of *Tas2r* genes in Phyllostomidae.

Table S15. Genes involved in digestion and absorption ($n = 105$).

Table S16. Genes undergoing divergent selection involved in digestive system.