

Positive selection on secretory and structural components of salivary glands within the ecologically diverse bat family Phyllostomidae

Michael Vandewege (✉ mike.vandewege@gmail.com)

Temple University <https://orcid.org/0000-0003-1115-9087>

Cibele G. Sotero-Caio

Charles University

Caleb D. Phillips

Texas Tech University

Research article

Keywords: Submandibular glands, selection, molecular evolution, immunity

Posted Date: October 21st, 2019

DOI: <https://doi.org/10.21203/rs.2.16268/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

The leaf-nosed bats (Phyllostomidae) are outliers among chiropterans with respect to the unusually high diversity of dietary strategies within the family. Salivary glands, owing to their functions and high ultrastructural variability among lineages, are proposed to have played an important role during the phyllostomid radiation. Salivary gland secretory products directly interact with food materials and pathogens which can provide selective interactions. To identify genes underlying salivary gland functional diversification, we sequenced submandibular gland transcriptomes from phyllostomid species representative of divergent dietary strategies. From the assembled transcriptomes, we identified and tested 3,266 single gene orthologs and identified 57 evolving under positive selection. All enriched gene ontology terms were related to defense against other organisms and the cell membrane/extracellular environment. Based on GO annotation, many of the defense-related loci under selection form secretory products. Salivary glands perform a complex array of tasks, and we identified positive selection on additional proteins bound to membranes and in the extracellular environment with multiple functions. Although illuminating exact function is difficult, results suggest that regulatory release of secretory products, not the products themselves, is disproportionately shaped by positive selection on coding sequence. Identified instances of selection on such cellular components may help explain the ultrastructural variability of salivary glands previously documented.

Background

Bats (Order Chiroptera) make up 20% of mammalian diversity and several groups have been alluded to as examples of extreme phenotypic and genomic changes leading to species diversity [1–7]. In particular, the leaf-nosed bats (Phyllostomidae) include more than 200 extant species, and certainly represent the most ecologically diverse mammal family as a wide range of dietary strategies (i.e. insectivory, sanguivory, nectarivory, frugivory, carnivory, and omnivory) are practiced [8, 9]. Morphometric and evolutionary models suggest the ecological and phenotypic diversity among lineages and species are associated with strong selection on dietary specialization features [3, 10–12].

Bats exhibit several characteristics rendering them interesting to examine, most uniquely their ability to fly and echolocate. Furthermore, bats act as vectors for zoonotic diseases including SARS coronavirus, Ebola, Nipah virus, and rabies [13–15]. A lesser examined characteristic thought to be directly linked to their dietary multiplicity and adaptation is the anatomical diversity of their salivary glands [16, 17]. In mammals, there are typically three major pairs of salivary glands: the submandibular, parotid, and sublingual glands, and all use intracellular processes that involve the synthesis, modification, packaging and secretion of proteins in membrane-bound granules. These glands are considered part of the digestive system as they secrete digestive enzymes, however their products perform additional functions including antimicrobial resistance and biochemical communication [18–22].

Tandler and Phillips [23] have shown that among 55 species of bats, secretory products were correlated with diet, especially in insectivorous species. Further, Phillips et al. [24] found evidence that salivary

glands play a role in lipid metabolism in *Myotis lucifugus*. Because of this wide variation and their direct link to immunity, diet and reproduction, SMGs and their products are hypothesized to play an important role in the adaptive radiation of mammals [25].

Linking genetic variation and gene products to selection and adaptation is still a challenge in evolutionary biology. More recently, the implementation of next-generation sequencing has facilitated the search for signatures of adaptation through DNA and/or RNA comparative analyses. Indeed, sequencing transcriptomes offers an in-depth, data rich method to identify selective pressures that have influenced the evolution of tissue-specific genes. Of interest are genes that have undergone positive selection to adapt physiological, immunological and ecological processes to new environments [26, 27].

Selection analyses among bat genes have generally been limited to few species for specific purposes [28, 29]. Hawkins et al. [26] analyzed orthologs from 18 different bat genomes and transcriptomes and found most genes under selection were related to immunity and collagen production. Here, Phyllostomidae was examined specifically because their extensive and relatively rapid radiation to new feeding strategies. We sequenced the SMG transcriptomes of nine phyllostomid bats representing different subfamilies and different diets, and through analysis of orthologs characterized how selection on coding sequence has shaped SMG evolution.

Results

To investigate putative proteins underlying the evolution of SMGs we performed RNASeq on 11 Chiroptera species: nine phyllostomids with varying dietary strategies, and two outgroups representing the bat ancestral state of insectivory (figure 1). Among the 11 samples, between 53,757 and 105,666 transcripts were assembled per species. Over 87% of the paired-reads mapped back to each assembly indicating most reads were used. Among the transcripts, we predicted between 17,820 and 54,578 ORFs, and among the ORFs, we found between 5,392 and 9,580 unique hits to the SwissProt database (Additional file 2: Table S1). The most similar isoform to the reference sequence was chosen to represent the annotation. SwissProt annotations were not used *a priori* for ortholog clustering. However, if after clustering there were multiple SwissProt annotations in a single gene ortholog group, the most common annotation was used to describe the ortholog group.

In all, OrthoMCL grouped 79,817 annotations into 10,267 orthologous groups that included between 2 and 45 members (Additional file 1: Figure S1A). Among the orthologs groups, 4,800 groups represented single gene families present in seven or more species and almost half of these families were found in all 11 species (Additional file 1: Figure S1B). Among these 4,800 groups, 4,667 (97%) had the same SwissProt annotation among orthologs, suggesting overall consistency between OrthoMCL clustering and Trinotate annotations. In the remaining 133 groups, there was one dominant gene annotation and the outliers were likely a result of improper classification by Trinotate, given BLAST hits are often closely related paralogs and not true orthologs, which seemed to be the case in most instances. Therefore, the annotations were manually corrected.

To additionally test if the ortholog predictions were generally accurate and DNA sequences largely reflected species relationships, we generated a phylogenetic tree from the concatenated alignments of 500 randomly sampled orthologs. Although individual gene trees may not reflect a species trees, we expected that a consensus would emerge from large volume of DNA sequences if orthologs were accurately predicted and aligned. We found that the resulting concatenated tree accurately reflected previous species trees generated by [3, 8, 49], with high bootstrap support at each node (figure 1).

Adaptation among SMG proteins

Among the 4,800 single gene orthologs, 3,266 had alignments with enough shared sites to produce output from CODEML. After correcting for FDR, we found 57 genes where models of evolution allowing positive selection were significantly better fit to the data than neutrality in both the M2a and M8 tests (Table S2).

To contrast background salivary gene functions to those under selection, we summarized the GO terms of the 3,266 tested genes using the REVIGO server (figure 2A). Overall most of these gene products are thought to be localized to the cytosol, but many were also positioned in the plasma membrane and extracellular region. The majority of terms in the “molecular function” category was protein binding, zinc ion binding and DNA binding. Background cellular processes were also apparent from the molecular function GO term analysis, i.e. transcription, transcription regulation and protein regulation. Similarly, common biological processes identified were related to transcription regulation, cell death, and protein synthesis and transport. By contrast, the terms of the genes under selection assigned to the three major categories (molecular function, cellular component and biological process) were largely different from those obtained in our global assessment of protein function (figure 2B). These proteins were largely receptors (molecular function), in the extracellular exosome or plasma membrane (cellular component) that are a part of the immune response (biological process). Interestingly, terms related to diet, like carbohydrate or lipid metabolism, were not associated with genes under positive selection.

We used PANTHER’s over-representation test to determine whether any terms were significantly over-represented among the genes under selection. There were no molecular function terms that were significantly more common among the genes under selection. However, there were significantly over-represented cellular component and biological process terms. The most enriched biological process term was antimicrobial humoral response, but all enriched biological process terms were related to the defense against other organisms (figure 3A). Consistent with receptor-like proteins under selection and salivary glands having a primary function in secretion, terms associated with membrane surfaces and extracellular space were enriched in the cellular component category (figure 3B). Upon survey of the multiple GO terms assigned to each gene we found that 11 of the 57 genes were involved in the immune response (Additional file 3: Table S2). Moreover, eight of the 11 (73%) immunity-related loci had secretory annotations, and a binomial test was used to quantify enrichment for secretory products among immune genes under selection ($p = 0.012$). Of the remaining 46, 17 genes were membrane bound, three were secreted, and two had terms for both membranes and secretion.

Discussion

Salivary glands secrete products that have important biological role in diet/digestion, oral health, and communication. They exhibit a high degree of morphological variation even among individuals of the same species and their role in the adaptation of mammals to novel niches has been largely under investigated. Additionally, salivary proteins are among the first to directly encounter food and introduced pathogens, which lends biological importance in the context of adaptive evolution. Here, we isolated and sequenced the genes expressed from the SMGs of phyllostomid bats, the most ecologically diverse family of mammals, to shed light into how positive natural selection has driven adaptation within the group. Among the sequenced transcriptomes we were able to annotate between 5,392 and 9,580 expressed proteins, which were successfully clustered as 4,800 single gene orthologs. Out of the 3,266 single gene orthologs tested, only a small proportion, 1.7% of genes, were under selection. Consistently, Hawkins et al. [26] yielded the same frequency of genes under selection from the transcriptomes and genomes of 18 bat species using multiple tissue sources (e.g. spleen, trigeminal ganglia, inner ear, embryos, etc. transcriptomes). This suggests that overall the genes expressed in SMGs are not necessarily subjected to stronger selective pressures in coding regions than genes expressed in other tissues. Given the diverse function and morphology of salivary glands previously shown [23], current results could indicate a role for regulatory evolution in SMG diversification, which can be assessed with expanded sample sizes per lineage.

Salivary glands perform a complex array of tasks and can present the most varied cellular ultrastructure in distinct taxa [23]. The evolution of cellular structure and the resulting array of phenotypes could be attributed to gene regulation, alternative splicing, and gene sequence evolution. The great morphological variation observed in the ultrastructure of phyllostomid bat salivary gland granules led Phillips et al. [25] to speculate that adaptation to novel niches drives salivary gland evolution. SMG acinar cells are responsive to a variety of extracellular signals which can affect gene expression, protein synthesis, and protein modifications, and this responsiveness can be driven by the density and distribution of cell surface receptors [16]. Consistently, we found that out of the 46 non-immune genes, 25 were either a membrane bound, secreted, or both. Unfortunately, membrane receptors could be linked to many functions. For example, basigin (BASI) is a globally expressed receptor linked to spermatogenesis, tumor progression, neural network formation and odontogenesis, among many other processes [50]. Therefore, it was difficult to parse and identify the function of each receptor expressed in salivary glands. However, molecular changes among receptors would likely alter the regulation and release of secretory products with a variety of functions.

A notable aspect of the ontological enrichment for genes with known roles in immunological or secretory processes provides indication about the specific cellular settings in which they function. For example, salivary glands are known to function in immunity through the localized function of immune cells [51]. Eleven of the 57 proteins inferred to be affected by positive selection were linked to immunity and defense, and eight of these had a secretory component (Additional file 3: Table S2). These 11 genes had common functions associated with the innate and humoral immune response and stimulating the

activation of immune cells to viral infected sites. For example, among these genes, were BPIA1 and BPIA2, both of which are secreted and known to localize to upper airways and prevent biofilm formation by gram-negative bacteria [52, 53]. A high exposure to pathogens unique to bats could explain this result, but positive selection among immune genes is not uncommon in animals [54–56]. Defense and immunity genes typically have a high evolutionary rate attributed to a continuous arms race between pathogens and host immune response [57–59], and our data provide insight about how positive selection has shaped bat oral tract defense against introduced pathogens.

Interestingly, Schneeberger et al [60] found that total white blood cell (WBC) counts in neotropical bats varied among bat species according to their dietary niches, where carnivorous bats had the highest WBC counts, insectivores had the lowest, and frugivorous had intermediate amounts. They hypothesized carnivorous species are at the highest risk of obtaining infectious diseases, as parasite and pathogen transmission is expected to be facilitated from prey, especially when those are relatively related species. Nevertheless, frugivorous species may still face a notable risk of acquiring pathogens by ingesting contaminated plant matter [61–63], and plantivores will also consume insects as a non-major dietary component. We examined the percentage of immune genes expressed in each species to determine if the same pattern was also present (Additional file 2: Table S1), but found that carnivorous and sanguivorous species expressed the fewest immune genes at 3.5 and 3.1%, respectively, and the highest was only 4.45% in *M. lucifugus*. This suggests that generally the same immune genes are expressed across lineages, but selection is shaping gene sequences in species-specific context.

Conclusions

Salivary glands are informative target organs to study the link between genetic variation and signatures of potential adaptations. Analysis of orthologs expressed in salivary transcriptomes enables detection of positive selection that helps explain the wide variation in cellular structures related to the regulation of salivary products released into the secretory ducts. Moreover, the identification of immunity-related genes under selection in the SMG is consistent with the necessity of hosts to continuously fend pathogens to which they are exposed in natural environments.

Methods

Nine species were chosen to maximize representation of the phylogenetic and dietary diversity of Phyllostomidae (figure 1). We also included two insectivorous bats, *Myotis lucifugus* from Vespertilionidae and *Pteronotus parnellii* from Mormoopidae, as outgroups. Tissues were extracted and frozen in liquid nitrogen within five minutes after euthanasia. Additional details from tissue loans provided by the Natural Science Research Laboratory (NSRL) of the Museum of Texas Tech University can be found in Additional file 1: Table S1.

RNA isolation, sequencing, and assembly

RNA was extracted from SMGs of each bat using Trizol (Invitrogen, Carlsbad, CA, USA) following manufacturer protocols. Oligo-dT magnetic beads were used to enrich for mRNA strands with poly-A tails and a strand-specific paired-end cDNA library was prepared using a ScriptSeq kit (Epicentre, Madison WI USA). Libraries were sequenced on Illumina platform (see Additional file 1: Table S1). For each species, paired-end reads were filtered for quality using Trimmomatic 0.36 [30] with the parameters LEADING:20 TRAILING:20 SLIDINGWINDOW:1:30 MINLEN:75 (MINLEN:50 was used for *D. rotundus*), and *de novo* transcript assemblies were constructed from clean read pairs with Trinity v2.7 [31] using default settings.

Gene discovery and annotation

Putative open reading frames and translated peptide sequences were identified using TransDecoder [32] and the resulting peptide sequences were processed through the Trinotate pipeline to identify functional properties and Gene Ontology (GO) annotations associated with biological processes, molecular functions, and cellular components. To summarize, peptide sequences were queried against the SwissProt database [33] using BLASTP [34] and the Pfam [35] database using hmmer 3.2.1 [36]. Peptides were also scanned for a signal peptide using signalP [37], and transmembrane domains using tmhmm 2.0 [38].

Orthology clustering

Orthology assignment is still a major challenge in bioinformatics and evolutionary biology. Here, we developed a process to filter out similar transcripts produced by Trinity. The first step was to assume similar sequences would have similar Trinotate annotations. We parsed the Trinotate output to identify the best sequence representative for each unique SwissProt annotation. To choose the best gene representative among multiple coding sequences (CDSs) with the same SwissProt annotation, we multiplied the length of the CDS to the percent identity of the SwissProt hit. This metric correlated with E value, but effectively acted as a bit score when E values were identical. The CDS with the highest metric was chosen to represent the annotation. If a CDS did not have a SwissProt annotation, it was removed. We then ran combined best sequences from all species, and ran this dataset through the OrthoMCL [39] pipeline to identify orthologous groups. Only single gene ortholog groups were used in downstream selection tests.

Poor ortholog assignment can influence sequence relationships in a phylogeny, and since the relationships among phyllostomids are robust, a reasonable test of ortholog assignment would be to reconstruct a phylogenetic tree and determine if the resulting tree reflects previously described relationships. Therefore, we reconstructed a phylogenetic tree from 500 randomly sampled single gene orthologs shared among all 11 individuals. Each orthologous group was translated, aligned using LINSi parameters in MAFFT [40] and reverse translated to construct a codon alignment. Resulting alignments were concatenated and we used RAxML [41] to find the best tree from the un-partitioned dataset using the ML and rapid bootstrapping algorithm, a GTRGAMMA model of nucleotide substitution, and with 100 starting trees.

Identifying genes under selection

Single gene orthologous groups that were found in seven or more individuals were tested for evidence of selection using the maximum likelihood approach described by Goldman and Yang [42]. Amino acid sequences were aligned using LINSi parameters in MAFFT [40], and the resulting alignment was reverse translated. Each gene tree was individually estimated using RAxML. We used CODEML in PAML [43] to estimate the role of selection on gene evolution by comparing the rate of nonsynonymous substitution per nonsynonymous site (d_N) to the rate of synonymous substitution per synonymous site (d_S). The d_N/d_S ratio can be used as a sensitive measure of selective pressure; however, in most cases the overall d_N/d_S is less than one and only a few amino acid sites are evolving quickly. Therefore, to determine whether a gene was evolving adaptively, we calculated the likelihood of models that allow d_N/d_S to vary among codon sites (M1a, M2a, M7, and M8). For all genes we used likelihood ratio tests (LRTs) to compare nested models that allow and disallow codon site d_N/d_S to be greater than one (M2a v M1a, M8 v M7), and to test for significant differences between nested models [44]. If both models that allow d_N/d_S rates to be greater than one (M2a and M8) were significantly better fits to the data (LRT, $p < 0.05$), we inferred these genes were evolving under positive selection. We performed a false discovery rate correction (FDR) on the p-values resulting from the M2a v M1a LRT because the M2a tests were more conservative than M8 and resulted in fewer LTR tests with $p < 0.05$. FDR was estimated using the qvalue function in the “qvalue” R module [45, 46].

Functional analysis of genes under selection

For all single gene orthologs tested for selection we counted the number of reoccurring GO terms. We then used the GO term proportion to summarize the general function of genes tested in the SMGs of phyllostomids using REVIGO [47], which nests redundant and similar terms from long GO term lists by semantic clustering. We repeated the same analysis using GO terms described from orthologs under positive selection. To test for GO term enrichment for the genes under selection, PANTHER [48] was applied to identify statistically overrepresented GO terms from the list of genes under positive selection using the genes tested as the reference list. We used the Fisher’s exact test and applied the FDR correction and statistical significance was determined below an adjusted p-value < 0.05 . We estimated whether a protein was membrane bound, a receptor, immune related, or secreted by searching for specific GO terms. If “secretion”, “extracellular space” or “extracellular region” was present among GO terms, we annotated the gene as secreted. If “immune”, “defense”, or “antimicrobial” was present, the gene was annotated as an immune. If “membrane” was present we called the gene membrane bound and if “receptor” was present, we called the gene a receptor (Additional file 3: Table S2).

Declarations

Acknowledgments

We would like to thank Robert Baker and Carlton Phillips for previous work that inspired this research. This work was not possible without the previous efforts of TTU students, faculty and staff towards sample collection and curation. Therefore, we are thankful for Nicté Ordoñez, Maria Sagot, Heath Garner, Robert Bradley, and the Texas Tech University Natural Science Research Laboratory (TTU NSRL) for tissue loans. We also acknowledge the High Performance Computing Center (HPCC) at Texas Tech University at Lubbock for providing HPC resources. URL: <http://cmsdev.ttu.edu/hpcc>

Funding

CGSC was supported with a fellowship (PDE) by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), and by a postdoctoral fellowship (PNPD) from Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Brazil during the development of this study. These funding agencies had no role in any experimental aspect of this study.

Availability of data and materials

Raw RNASeq files are available from the SRA under the BioProject Accession PRJNA451361.

Author's contributions

MWV conducted the bioinformatics, analyzed the results, and drafted the manuscript. CGSC interpreted the data and revised the manuscript. CDP conceived the study, conducted the assembly and annotation of transcriptomes, and revised the manuscript. All authors have read and approved the final manuscript.

Ethics approval and consent to participate

RNA was extracted from subsamples of tissue loans provided by the TTU NSRL. No animals were collected for this study.

Consent for publication

Not applicable

Competing interests

The authors declare that they have no competing interests

Additional Files

Additional File 1. Figure 1. A) Results of OrthoMCL clustering from unique sequences among all 11 species. B) Number of single gene orthologs that included between 7 and 12 species.

Additional File 2. Table S1. Attributes associated with the samples used in this study.

Additional File 3. Table S2. The list of genes found under selection in CODEML site tests with functional annotations.

References

1. Pritham EJ, Feschotte C. Massive amplification of rolling-circle transposons in the lineage of the bat *Myotis lucifugus*. *Proc Natl Acad Sci*. 2007;104:1895–900.
2. Ray DA, Pagan HJT, Thompson ML, Stevens RD. Bats with hATs: evidence for recent DNA transposon activity in genus *Myotis*. *Mol Biol Evol*. 2007;24:632–9.
3. Dumont ER, Dávalos LM, Goldberg A, Santana SE, Rex K, Voigt CC. Morphological innovation, diversification and invasion of a new adaptive zone. *Proc R Soc B Biol Sci*. 2011;279:1797–805.
4. Hayden S, Bekaert M, Goodbla A, Murphy WJ, Dávalos LM, Teeling EC. A cluster of olfactory receptor genes linked to frugivory in bats. *Mol Biol Evol*. 2014;31:917–27.
5. Platt RN, Vandewege MW, Kern C, Schmidt CJ, Hoffmann FG, Ray DA. Large numbers of novel miRNAs originate from DNA transposons and are coincident with a large species radiation in bats. *Mol Biol Evol*. 2014;31:1536–45.
6. Sotero-Caio CG, Volleth M, Hoffmann FG, Scott L, Wichman HA, Yang F, et al. Integration of molecular cytogenetics, dated molecular phylogeny, and model-based predictions to understand the extreme chromosome reorganization in the Neotropical genus *Tonatia* (Chiroptera: Phyllostomidae). *BMC Evol Biol*. 2015;15:220.
7. Phillips CD, Baker RJ. Secretory gene recruitments in vampire bat salivary adaptation and potential convergences with sanguivorous leeches. *Front Ecol Evol*. 2015;3:122.
8. Baker RJ, Solari S, Cirranello A, Simmons NB. Higher Level Classification of Phyllostomid Bats with a Summary of DNA Synapomorphies. *Acta Chiropterologica*. 2016;18:1–38.
9. Cirranello A, Simmons NB, Solari S, Baker RJ. Morphological diagnoses of higher-level phyllostomid taxa (Chiroptera: Phyllostomidae). *Acta Chiropterologica*. 2016;18:39–71.
10. Monteiro LR, Nogueira MR. Evolutionary patterns and processes in the radiation of phyllostomid bats. *BMC Evol Biol*. 2011;11:137.
11. Rossoni DM, Assis APA, Giannini NP, Marroig G. Intense natural selection preceded the invasion of new adaptive zones during the radiation of New World leaf-nosed bats. *Sci Rep*. 2017;7:11076.
12. Rojas D, Vale Á, Ferrero V, Navarro L. When did plants become important to leaf-nosed bats? Diversification of feeding habits in the family Phyllostomidae. *Mol Ecol*. 2011;20:2217–28.
13. Calisher CH, Childs JE, Field HE, Holmes K V, Schountz T. Bats: important reservoir hosts of emerging viruses. *Clin Microbiol Rev*. 2006;19:531–45.
14. Olival KJ, Hosseini PR, Zambrana-Torrel C, Ross N, Bogich TL, Daszak P. Host and viral traits predict zoonotic spillover from mammals. *Nature*. 2017;546:646–50.
15. Smith I, Wang L-F. Bats and their virome: an important source of emerging viruses capable of infecting humans. *Curr Opin Virol*. 2013;3:84–91.

16. Phillips CJ, Weiss A, Tandler B. Plasticity and patterns of evolution in mammalian salivary glands: comparative immunohistochemistry of lysozyme in bats. *Eur J Morphol.* 1998;36 Suppl:19–26.
17. Tandler B, Gresik EW, Nagato T, Phillips CJ. Secretion by striated ducts of mammalian major salivary glands: Review from an ultrastructural, functional, and evolutionary perspective. *Anat Rec.* 2001;264:121–45.
18. Dobrosielski-Vergona K. *Biology of the salivary glands.* Boca Raton: CRC Press; 1993.
19. Talley HM, Laukaitis CM, Karn RC. Female preference for male saliva: implications for sexual isolation of *Mus musculus* subspecies. *Evolution.* 2001;55:631–4.
20. Vandewege MW, Phillips CJ, Wickliffe JK, Hoffmann FG. Evolution of the ABPA subunit of androgen-binding protein expressed in the submaxillary glands in New and Old World rodent taxa. *J Mol Evol.* 2013;76:324–31.
21. Bloss J, Acree TE, Bloss JM, Hood WR, Kunz TH. Potential use of chemical cues for colony-mate recognition in the big brown bat, *Eptesicus fuscus*. *J Chem Ecol.* 2002;28:819–34.
22. Safi K, Kerth G. Secretions of the interaural gland contain information about individuality and colony membership in the Bechstein's bat. *Anim Behav.* 2003;65:363–9.
23. Tandler B, Phillips CJ. Microstructure of mammalian salivary glands and its relationship to diet. *Gland Mech Salivary Secret.* 1998;10:21–35.
24. Phillips CJ, Phillips CD, Goecks J, Lessa EP, Sotero-Caio CG, Tandler B, et al. Dietary and flight energetic adaptations in a salivary gland transcriptome of an insectivorous bat. *PLoS One.* 2014;9:e83512.
25. Phillips CJ, Tandler B. Salivary glands, cellular evolution, and adaptive radiation in mammals. *Eur J Morphol.* 1996;34:155–61.
26. Hawkins JA, Kaczmarek ME, Müller MA, Drosten C, Press WH, Sawyer SL. A metaanalysis of bat phylogenetics and positive selection based on genomes and transcriptomes from 18 species. *Proc Natl Acad Sci U S A.* 2019;116:11351–60.
27. Daugherty MD, Malik HS. Rules of Engagement: Molecular Insights from Host-Virus Arms Races. *Annu Rev Genet.* 2012;46:677–700.
28. Zhang G, Cowled C, Shi Z, Huang Z, Bishop-Lilly KA, Fang X, et al. Comparative analysis of bat genomes provides insight into the evolution of flight and immunity. *Science.* 2013;339:456–60.
29. Shen Y-Y, Liang L, Zhu Z-H, Zhou W-P, Irwin DM, Zhang Y-P. Adaptive evolution of energy metabolism genes and the origin of flight in bats. *Proc Natl Acad Sci U S A.* 2010;107:8666–71.
30. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics.* 2014;30:2114–20.
31. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol.* 2011;29:644–52.
32. Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, et al. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and

- analysis. *Nat Protoc.* 2013;8:1494–512.
33. Dimmer EC, Huntley RP, Alam-Faruque Y, Sawford T, O'Donovan C, Martin MJ, et al. The UniProt-GO Annotation database in 2011. *Nucleic Acids Res.* 2012;40 Database issue:D565–70.
 34. Altschul SF, Madden TL, Schäffer A a., Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res.* 1997;25:3389–402.
 35. Finn RD, Mistry J, Tate J, Coggill P, Heger A, Pollington JE, et al. The Pfam protein families database. *Nucleic Acids Res.* 2010;38 suppl_1:D211–22.
 36. Eddy SR. Accelerated Profile HMM Searches. *PLoS Comput Biol.* 2011;7:e1002195.
 37. Petersen TN, Brunak S, von Heijne G, Nielsen H. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat Methods.* 2011;8:785–6.
 38. Krogh A, Larsson B, von Heijne G, Sonnhammer EL. Predicting transmembrane protein topology with a hidden markov model: application to complete genomes. *J Mol Biol.* 2001;305:567–80.
 39. Li L, Stoeckert CJ, Roos DS. OrthoMCL: Identification of Ortholog Groups for Eukaryotic Genomes. *Genome Res.* 2003;13:2178–89.
 40. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol.* 2013;30:772–80.
 41. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics.* 2014;30:1312–3.
 42. Goldman N, Yang Z. A codon-based model of nucleotide substitution for protein-coding DNA sequences. *Mol Biol Evol.* 1994;11:725–36.
 43. Yang Z. PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol.* 2007;24:1586–91.
 44. Yang Z. Likelihood ratio tests for detecting positive selection and application to primate lysozyme evolution. *Mol Biol Evol.* 1998;15:568–73.
 45. Storey JD, Tibshirani R. Statistical significance for genomewide studies. *Proc Natl Acad Sci.* 2003;100:9440–5.
 46. Storey JD. qvalue: Q-value estimating for false discovery rate control. R package version 2.14.0. <http://github.com/jdstorey/qvalue>. 2018.
 47. Supek F, Bošnjak M, Škunca N, Šmuc T. REVIGO Summarizes and Visualizes Long Lists of Gene Ontology Terms. *PLoS One.* 2011;6:e21800.
 48. Mi H, Poudel S, Muruganujan A, Casagrande JT, Thomas PD. PANTHER version 10: expanded protein families and functions, and analysis tools. *Nucleic Acids Res.* 2016;44:D336–42.
 49. Rojas D, Warsi OM, Dávalos LM. Bats (Chiroptera: Noctilionoidea) Challenge a Recent Origin of Extant Neotropical Diversity. *Syst Biol.* 2016;65:432–48.
 50. Muramatsu T, Miyauchi T. Basigin (CD147): a multifunctional transmembrane protein involved in reproduction, neural function, inflammation and tumor invasion. *Histol Histopathol.* 2003;18:981–7.

51. Fábíán TK, Hermann P, Beck A, Fejérdy P, Fábíán G. Salivary Defense Proteins: Their Network and Role in Innate and Acquired Oral Immunity. *Int J Mol Sci.* 2012;13:4295–320.
52. Prokopovic V, Popovic M, Andjelkovic U, Marsavelski A, Raskovic B, Gavrovic-Jankulovic M, et al. Isolation, biochemical characterization and anti-bacterial activity of BPIFA2 protein. *Arch Oral Biol.* 2014;59:302–9.
53. Liu Y, Bartlett JA, Di ME, Bomberger JM, Chan YR, Gakhar L, et al. SPLUNC1/BPIFA1 Contributes to Pulmonary Host Defense against *Klebsiella pneumoniae* Respiratory Infection. *Am J Pathol.* 2013;182:1519–31.
54. Xiao J, Zhong H, Liu Z, Yu F, Luo Y, Gan X, et al. Transcriptome analysis revealed positive selection of immune-related genes in tilapia. *Fish Shellfish Immunol.* 2015;44:60–5.
55. Roux J, Privman E, Moretti S, Daub JT, Robinson-Rechavi M, Keller L. Patterns of Positive Selection in Seven Ant Genomes. *Mol Biol Evol.* 2014;31:1661–85.
56. van der Lee R, Wiel L, van Dam TJP, Huynen MA. Genome-scale detection of positive selection in nine primates predicts human-virus evolutionary conflicts. *Nucleic Acids Res.* 2017;45:10634–48.
57. Jiggins FM, Kim KW. A screen for immunity genes evolving under positive selection in *Drosophila*. *J Evol Biol.* 2007;20:965–70.
58. Viljakainen L, Evans JD, Hasselmann M, Rueppell O, Tingek S, Pamilo P. Rapid evolution of immune proteins in social insects. *Mol Biol Evol.* 2009;26:1791–801.
59. Kosiol C, Vinař T, da Fonseca RR, Hubisz MJ, Bustamante CD, Nielsen R, et al. Patterns of positive selection in six mammalian genomes. *PLoS Genet.* 2008;4:e1000144.
60. Schneeberger K, Czirják GÁ, Voigt CC. Measures of the Constitutive Immune System Are Linked to Diet and Roosting Habits of Neotropical Bats. *PLoS One.* 2013;8:e54023.
61. Fogarty R, Halpin K, Hyatt AD, Daszak P, Mungall BA. Henipavirus susceptibility to environmental variables. *Virus Res.* 2008;132:140–4.
62. Chua KB, Lek Koh C, Hooi PS, Wee KF, Khong JH, Chua BH, et al. Isolation of Nipah virus from Malaysian Island flying-foxes. *Microbes Infect.* 2002;4:145–51.
63. Vitone ND, Altizer S, Nunn CL. Body size, diet and sociality influence the species richness of parasitic worms in anthropoid primates. *Evol Ecol Res.* 2004;6:183–99.

Figures

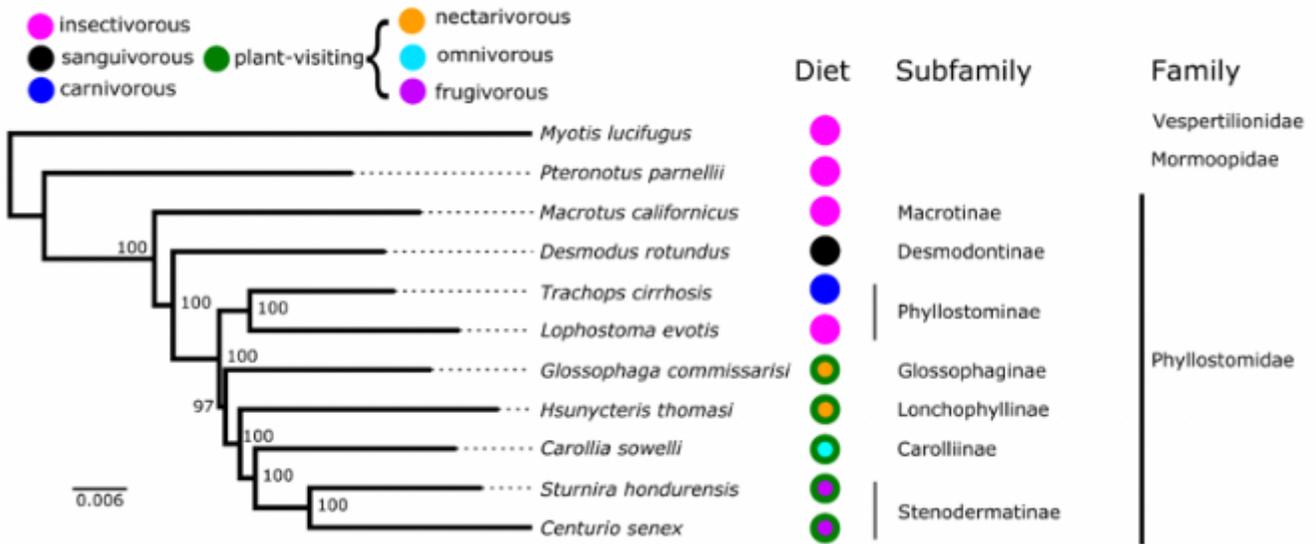


Figure 1

Schematic of species sampled, their relationships, and general dietary strategies.

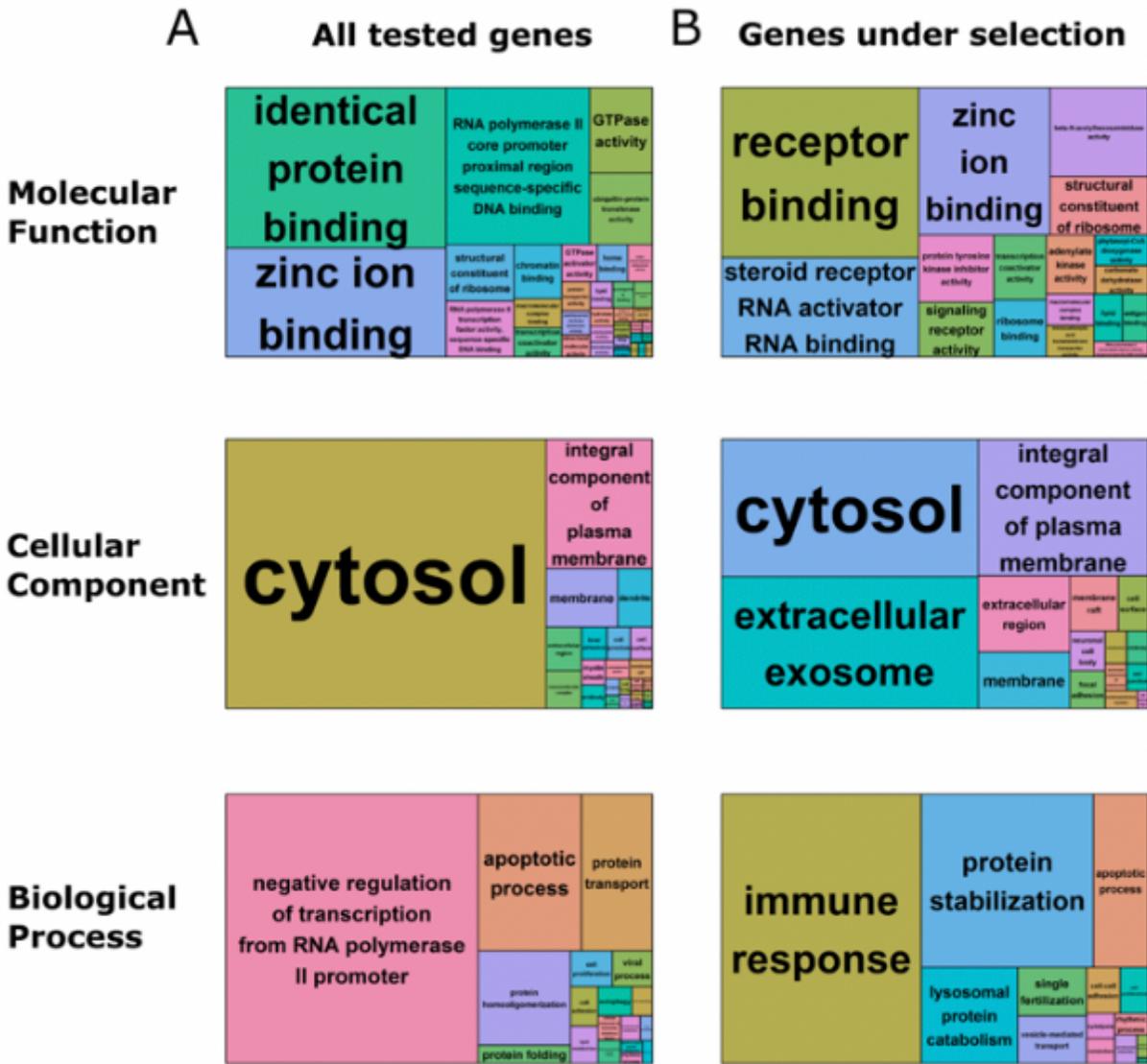


Figure 2

Tree map produced from the analysis of redundant GO terms in REVIGO for A) all GO terms linked to genes tested and B) GO terms linked to genes under selection.

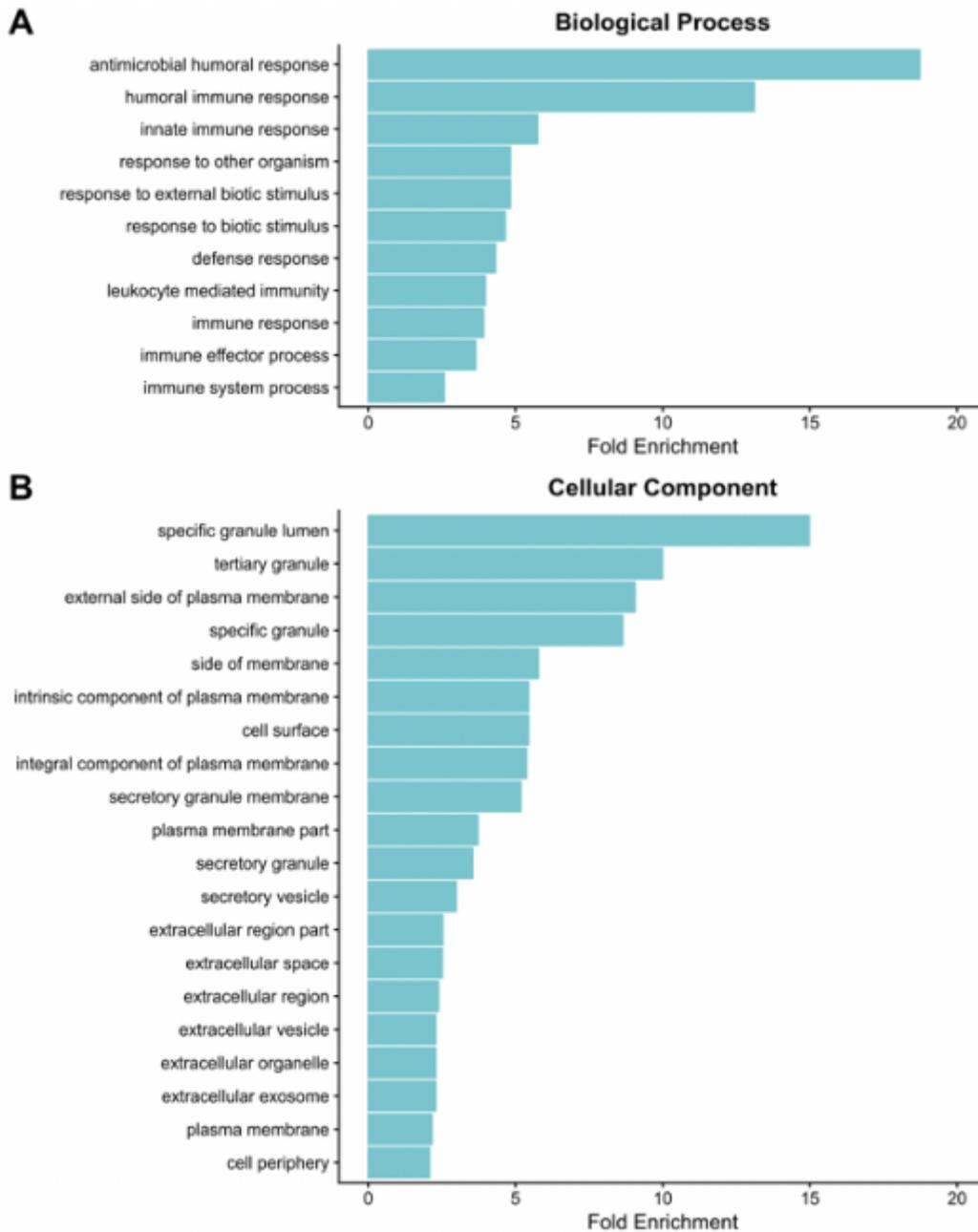


Figure 3

Fold enrichment for A) Biological Processes and B) Cellular Components GO terms that were statistically over-represented among genes under positive natural selection.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [supplementalFigures.docx](#)
- [SuppTableS2.xlsx](#)

- [SuppTableS1.xlsx](#)