

Positively selected genes in the hoary bat (*Lasiurus cinereus*) lineage: prominence of thymus expression, immune and metabolic function, and regions of ancient synteny

Robert S. Cornman and Paul M. Cryan

Fort Collins Science Center, U.S. Geological Survey, Fort Collins, CO, United States of America

ABSTRACT

Background: Bats of the genus *Lasiurus* occur throughout the Americas and have diversified into at least 20 species among three subgenera. The hoary bat (*Lasiurus cinereus*) is highly migratory and ranges farther across North America than any other wild mammal. Despite the ecological importance of this species as a major insect predator, and the particular susceptibility of lasiurine bats to wind turbine strikes, our understanding of hoary bat ecology, physiology, and behavior remains poor.

Methods: To better understand adaptive evolution in this lineage, we used whole-genome sequencing to identify protein-coding sequence and explore signatures of positive selection. Gene models were predicted with Maker and compared to seven well-annotated and phylogenetically representative species. Evolutionary rate analysis was performed with PAML.

Results: Of 9,447 single-copy orthologous groups that met evaluation criteria, 150 genes had a significant excess of nonsynonymous substitutions along the *L. cinereus* branch ($P < 0.001$ after manual review of alignments). Selected genes as a group had biased expression, most strongly in thymus tissue. We identified 23 selected genes with reported immune functions as well as a divergent paralog of *Steep1* within suborder Yangochiroptera. Seventeen genes had roles in lipid and glucose metabolic pathways, partially overlapping with 15 mitochondrion-associated genes; these adaptations may reflect the metabolic challenges of hibernation, long-distance migration, and seasonal variation in prey abundance. The genomic distribution of positively selected genes differed significantly from background expectation by discrete Kolmogorov–Smirnov test ($P < 0.001$). Remarkably, the top three physical clusters all coincided with islands of conserved synteny predating Mammalia, the largest of which shares synteny with the human cat-eye critical region (CECR) on 22q11. This observation coupled with the expansion of a novel *Tbx1*-like gene family may indicate evolutionary innovation during pharyngeal arch development: both the CECR and *Tbx1* cause dosage-dependent congenital abnormalities in thymus, heart, and head, and craniodysmorphism is associated with human orthologs of other positively selected genes as well.

Submitted 23 September 2021

Accepted 25 February 2022

Published 17 March 2022

Corresponding author

Robert S. Cornman,

rcornman@usgs.gov

Academic editor

Madhava Meegaskumbura

Additional Information and
Declarations can be found on
page 24

DOI 10.7717/peerj.13130

© Copyright

2022 Cornman and Cryan

Distributed under

Creative Commons CC-BY 4.0

OPEN ACCESS

Subjects Genetics, Genomics, Molecular Biology, Zoology

Keywords Hoary bat, Adaptation, Positive selection, Immunogenetics, Thymus, Conserved synteny, Cat-eye critical region, Tbx gene family, Chrna9

INTRODUCTION

Bats are a highly diverse clade of mammals exhibiting well-known innovations such as flight, echolocation, and extremes of metabolism and longevity. They have important ecological roles including pollination, seed dispersal, and insect predation, and as a group are understudied and threatened (Adams & Pedersen, 2013). Recent intensification of stresses in North America related to white-nose syndrome (Hoyt, Kilpatrick & Langwig, 2021) and wind-turbine fatalities (Frick et al., 2017) have heightened the need for research to support bat management, yet many species are challenging to census and study.

A species of growing concern for U.S. management agencies (Hein, Hale & Straw, 2021) is the hoary bat (*Lasiurus cinereus*), a tree bat comprised of two disjunct populations in North America and South America. Individuals are highly migratory, traveling thousands of kilometers per year in search of prey or mates (Cryan, Stricker & Wunder, 2014). Little is known about their hibernacula or the physiological character of hibernation in the species, and while the effective population size has historically been large (Korstian, Hale & Williams, 2015; Sovic, Carstens & Gibbs, 2016; Pylant et al., 2016; Cornman et al., 2021), evidence of range contraction (Rodhouse et al., 2019) and high turbine-associated mortality (Frick et al., 2017) suggest population pressures are currently high. Importantly, wind-turbine mortality has a strong behavioral component as individuals actively approach the moving blades (Cryan et al., 2014) for reasons that remain unclear. It has been hypothesized that the turbines mimic a natural stimulus (e.g., Corcoran & Weller, 2018) and technologies to counter this attraction are being explored (e.g., Weaver et al., 2020). Interestingly, a novel microcall has recently been identified from acoustic analysis of *L. cinereus* echolocation (Corcoran & Weller, 2018), although the phylogenetic distribution and consistency in form of this behavior remain to be fully determined. Elucidating the evolutionary timing and developmental or physiological bases of these and other distinctive traits can help researchers better understand how ecological niches evolve (e.g., Oliver et al., 2010; Zhang et al., 2019b; Davies et al., 2020) and potentially their sensitivity to perturbation. Genes underlying recent physiological adaptations could be particularly relevant as biomarkers of organism condition, and standing variation at such loci may influence future adaptation to anthropogenic pressures (Bitter et al., 2019).

Whereas genomic resources have begun to appear for the hoary bat in support of population analyses (Pinzari et al., 2020; Cornman et al., 2021), evolutionary comparisons with related bat genomes have not yet been explored. While we do not propose specific hypotheses to be addressed, descriptive assessments of protein-evolutionary rates provide the groundwork that could help define future hypotheses on the molecular basis of adaptation in the hoary bat lineage. Indeed, comparative genomics has revealed molecular mechanisms of adaptation, novelty, and constraint across diverse evolutionary clades (Kosiol et al., 2008; Corradi et al., 2010; Manceau et al., 2010; Lindblad-Toh et al., 2011; Roux et al., 2014; Neafsey et al., 2015; Hughes et al., 2018; Beichman et al., 2019). The Bat1K project (Teeling et al., 2018), for example, has begun to produce insights into bat evolutionary genomics (Jebb et al., 2020).

In this study, we annotated a preliminary gene set based on protein homology and performed branch-specific tests of evolutionary rate of coding sequences (the estimated ratio of nonsynonymous substitutions to synonymous substitutions). We found a large number of significant tests after manual review of initially significant alignments. Searches of gene annotation databases and the literature for human and mouse orthologs identified several major themes including immune function, metabolic homeostasis, and mitochondrial localization. Available gene expression data for mouse orthologs indicated biased expression of positively selected genes in the thymus, a key site of adaptive immunity (*Thapa & Farber, 2019*). Three physical clusters of positively selected genes were noted, all of which coincide with genomic islands of conserved gene content and order (ancient synteny) predating Mammalia; two of these date to the origins of Tetrapoda. The largest cluster of positively selected genes was syntenic with the human cat-eye critical region (CECR) on 22q11, which is associated with a spectrum of congenital developmental disorders (*Footz et al., 2001*). The genes we identified are candidates for further vetting by genome sequencing of congeners and other close relatives, comparative transcriptomics, and intraspecific polymorphism scans.

MATERIALS AND METHODS

DNA extraction and sequencing

Genome sequencing and assembly are detailed in *Cornman et al. (2021)*. Briefly, blood was drawn from an adult female specimen within 12 h *post mortem* and shipped cold to MedGenome Inc. for DNA extraction and 10X linked-read library preparation. Sequencing was performed on the Illumina NovaSeq platform in a paired 150-bp layout. A total of 594,097,777 read pairs were generated together with the associated linking index reads (available from the National Center for Biotechnology Information (NCBI) under BioProject accession [PRJNA601154](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA601154)). Two iterations of linked-read assembly were performed with the supernova assembly program on the Yeti supercomputing resource of the U.S. Geological Survey (<https://www.usgs.gov/core-science-systems/sas/arc/usgs-yeti-supercomputer>). The final diploid assembly had coverage metrics in line with manufacturer recommendations. Two ‘pseudohaploid’ sequences, between which unphased haplotype blocks are arbitrarily split (see *Weisenfeld et al., 2017* for details), were deposited in NCBI’s Genome database (accessions [GCA_011751065.1](https://www.ncbi.nlm.nih.gov/nuccore/GCA_011751065.1) and [GCA_011751095.1](https://www.ncbi.nlm.nih.gov/nuccore/GCA_011751095.1)). Of the two reference pseudohaplotypes, we arbitrarily chose to annotate the pseudohaplotype associated with accession [GCA_011751065.1](https://www.ncbi.nlm.nih.gov/nuccore/GCA_011751065.1). However, our analysis was performed on the original scaffolds output by the supernova assembly program (available from *Cornman et al., 2020*), not the slightly more fragmented assembly generated during submission to NCBI. Fragmentation occurs because NCBI policy requires scaffolds be split on strings of 10 or more ambiguous bases when the gap size is not explicitly estimated, but this threshold may not be a strong indication of assembly ambiguity in linked-read data.

Gene annotation

Annotation was performed with Maker v. 2.31.10 (Cantarel *et al.*, 2008). Repeat masking was performed within Maker using a licensed RepBase database (v. 24.12; Jurka *et al.*, 2005) and RepeatMasker (Tarailo-Graovac & Chen, 2009) database Dfam3.1, in addition to *de novo* repeat discovery. As no transcriptomic data were available for *L. cinereus*, predictions were based on protein homology only, using all Chiroptera identical protein groups (downloaded from NCBI on January 29, 2020) as the reference database. Diamond (Buchfink, Xie & Huson, 2015) was selected as the alignment package and exonerate v. 2.2.0 (Slater & Birney, 2005) was selected to create homology-based gene models. A total of 20,663 protein models were generated (Cornman & Cryan, 2022), which is comparable to the 19,728 coding models and 1,713 pseudogenes listed in the Ensembl genome database (Yates *et al.*, 2020) for *M. lucifugus*. The predicted proteins included 84.9% of ‘universal’ mammalian orthogroups, as assessed with Busco 5.1.3 (Seppey, Manni & Zdobnov, 2019) using OrthoDB version 10 (Kriventseva *et al.*, 2019), of which 98.7% were complete and single copy. In comparison, 96% of the same benchmark set were detected for the related vespertilionid bat *Eptesicus fuscus* (see below), of which 99.6% were complete and single-copy. We did not annotate non-coding features such as non-coding RNAs, transfer RNAs, or ribosomal RNAs. Gene predictions in GFF format along with the corresponding coding sequences are available from Cornman & Cryan (2022).

Evolutionary rate analysis

RefSeq mRNA models and their protein translations were downloaded from NCBI for *M. lucifugus* (assembly accession GCA_000147115.1; Suborder Yangochiroptera, Family Vespertilionidae) (Lindblad-Toh *et al.*, 2011) and six other bat species (Jebb *et al.*, 2020): *E. fuscus* (assembly accession GCA_000308155.1, EptFus1.0; Yangochiroptera, Vespertilionidae), *Hipposideros armiger* (assembly accession GCA_001890085.1, ASM189008v1; Dong *et al.*, 2016; Yinpterochiroptera, Hipposideridae), *Phyllostomus discolor* (assembly accession GCA_004126475.3 mPhyDis1.pri.v3; Yangochiroptera, Phyllostomidae), *Pteropus vampyrus* (assembly accession GCA_000151845.2 Pvam_2.0; Yinpterochiroptera, Pteropodidae), *Rhinolophus ferrumequinum* (assembly accession GCA_004115265.2, mRhiFer1_v1.p; Yinpterochiroptera, Rhinolophidae), and *Rousettus aegyptiacus* (assembly accession GCA_014176215.1 mRouAeg1.p; Yinpterochiroptera, Pteropodidae). These accessions represent most full-genome derived proteomes available for bats at the time of analysis, after excluding closely related species such as *Myotis* congeners. Sequences were filtered on locus name to include only the first listed isoform when multiple isoforms of a locus were present.

To identify orthologous groups of sequences and generate nucleotide alignments for analysis of substitution rates, we adapted the workflow of Santagata (2021). Orthologous groups were first identified using Orthofinder (Emms & Kelly, 2019) and filtered to include only orthogroups with a single representative from every species and at most a single paralog in one species other than *L. cinereus*. The number of “perfect” single-copy, complete orthogroups identified was 9,064, with an additional 466 “near-perfect” orthogroups with a single paralog in a species other than *L. cinereus*. Only perfect or

near-perfect orthogroups were used because closely related paralogs can evolve non-independently by gene conversion, neofunctionalization, or subfunctionalization (Casola, Ganote & Hahn, 2010; Sandve, Rohlfs & Hvidsten, 2018), distorting the estimate of background evolutionary rate. Orthogroups lacking orthologs in any species were also excluded due to the reduced power to estimate evolutionary rate.

Near-perfect orthologous groups were trimmed of the inferred paralog using phylopypruner (Thalén, 2018). Orthologous protein sequences were aligned with mafft (Katoh *et al.*, 2002). As protein-coding models occasionally contain stop codons (*e.g.*, due to annotation error or evidence of stop codon read-through), we converted in-frame stop codons to ambiguity characters. This conversion was performed because mafft treats stop characters in protein coding sequence as gaps, which can result in a loss of correspondence between protein and nucleotide sequences. Nucleotide alignments were reconstituted from protein alignments and formatted for PAML using pal2nal (Suyama, Torrents & Bork, 2006). Orthogroups with sequences less than 100 amino acids in length were discarded as too short for analysis.

Evolutionary rate analysis assumes that the gene tree has the same topology as the species tree and paralogs and orthologs are not conflated. It further assumes orthology of each aligned codon, which may be uncertain for some parts of an alignment. Error can arise from poorly conserved or low-complexity regions, incorrect 5' or 3' gene termini, missed exons, and frameshift errors in the reference genome. Gross errors are particularly likely at alignment edges because the corresponding exons may have low coding content and thus more likely to be missed by protein-alignment methods. We therefore trimmed alignments from both the 5' and 3' ends until the first completely conserved amino-acid triplet was observed in each direction. This approach is likely conservative, entailing a tradeoff between alignments that are shorter and enriched in relatively well-conserved protein regions versus alignments containing more errors and requiring greater manual evaluation. While there is a clear need to mask ambiguous sequence alignments that bias evolutionary rate estimation, methods that balance this need against the loss of informative sites remain an area of active investigation (Di Franco *et al.*, 2019). For example, Beichman *et al.* (2019) manually reviewed over 600 candidates for positive selection after applying four such methods, resulting in only 18 accepted tests, and found that different masking approaches gave divergent results. We therefore followed Jebb *et al.* (2020) in using the T-Coffee alignment consistency score as a guideline for validating all manually revised alignments as well as all alignments with a significant PAML result (see below), requiring a minimum score of 950 (the maximum possible score is 1,000) (Notredame, Higgins & Heringa, 2000; Chang, Di Tommaso & Notredame, 2014). Similar considerations led us to filter orthogroup alignments based on the tree length (substitutions per codon) calculated by PAML: we excluded 134 orthogroups with tree lengths greater than 10, as by inspection these outliers were highly enriched in alignment error (Supplemental File S1).

We used the codeml program of PAML 4.9i (Yang, 2007) to test for positive selection acting on the *L. cinereus* branch of the multilocus species tree estimated by Orthofinder, after unrooting the Orthofinder-inferred tree with the R package *ape*

(Paradis, Claude & Strimmer, 2004). PAML models 1 and 2 were first fit to the tree topology assuming all rate classes applied to all branches. Model 1 assumes a fraction of sites p_0 have evolutionary rate $\omega_0 < 1$ (purifying selection), with the remaining sites evolving neutrally ($\omega_1 = 1$). Model 2 estimates two evolutionary rates and two proportions, dividing sites into those evolving under purifying selection ($\omega_0 < 1$) or positive selection ($\omega_2 > 1$), with the balance evolving under neutrality ($\omega_1 = 1$). Model 2 was then re-fit after designating the *L. cinereus* branch as the foreground branch with a potentially different value of ω_2 than the background branches (for which ω_2 is fixed at 1). For each pair of models fit to a given alignment, twice the difference in the natural log of the reported likelihoods constitutes a test metric that can be compared to the χ^2 probability distribution (Yang, 2007) with one degree of freedom. Although Model 1 is not the recommended null model for detecting branch-specific selection, we estimated its likelihood to confirm it was not appreciably greater than Model 2 estimates, which might suggest poor estimate convergence.

Initially significant alignments were manually checked for misaligned or low-complexity sequence. Gapped low-complexity regions were deleted between the nearest 5' and 3' conserved codons. Runs of amino-acid substitutions unique to a particular branch were considered errors in the predicted primary sequence and were also deleted between the nearest 5' and 3' conserved codons. Gross misalignments were not reanalyzed if little nonsynonymous variation was evident on the *L. cinereus* branch and were discarded as false positives. Because of the need for manual masking of poorly aligned regions, resulting in many initial test metrics being judged invalid on inspection or requiring a revised test (see "Results"), the standard approach to false discovery correction is unsatisfactory. This is because the orthogroups with highest test statistics are enriched in alignment error: ranking test results by test statistic and revising the significance threshold based on test number would therefore have resulted in their preferential acceptance at the expense of orthogroups with more biologically plausible test statistics. We addressed the lack of a straightforward correction procedure by setting a relatively strict $\alpha = 0.001$ for the likelihood ratio test. Summary metrics and test results for the 9,447 groups of orthologous genes ("orthogroups") identified by the workflow are provided in [Supplemental File S1](#). Gene annotations, initial orthogroup alignments, and manually revised orthogroup alignments are provided in [Supplemental Files S2–S4](#).

We compared PAML results with those of the 'adaptive' branch-specific model implemented by the program aBSREL of the HyPhy package (Smith et al., 2015). This program tests whether a foreground branch is evolving under positive diversifying selection while allowing separate background rates to potentially be calculated for other branches in the tree. This is potentially a more parameter-rich model depending on the outcome of the adaptive algorithm, with each additional parameter necessarily estimated from fewer data. The aBSREL method was developed for larger alignments such as viral data sets (Smith et al., 2015) and is only used here to check for divergent behavior of the two methods. For our objective of identifying loci positively selected in the *L. cinereus* lineage and given the relatively small number of sequences available per locus, we believe estimating a single rate of purifying selection from all background branches with

PAML is the preferred abstraction because more data are used to estimate that parameter and contrast it with the *L. cinereus* rate.

P-values estimated by PAML were generally lower (more significant) than those calculated by aBSREL (Supplemental File S5). The two methods were often concordant but exhibited a qualitative pattern in that aBSREL *P*-values were either low and log-linear with PAML *P*-values or much higher and unrelated to the PAML *P*-values. This duality was most striking for longer alignments (Supplemental File S5), whereas there was little concordance between *P*-values estimated by the two methods for the shortest alignments. The dual distribution presumably reflects the extent to which the adaptive aBSREL algorithm attributes evolutionary change to background rather than foreground branches of the phylogeny (a single background rate is assumed by PAML). If some background branches are estimated by aBSREL to have relatively high evolutionary rates, the hypothesis of selection on the foreground branch (*L. cinereus*) should be correspondingly less supported, particularly as alignment lengths decrease.

For each tested orthogroup, the *M. lucifugus* RefSeq protein ID was used to extract the corresponding NCBI Gene database record (Brown *et al.*, 2015), which was parsed to obtain the mouse and human ortholog identifiers, when available. We used the DAVID functional annotation tool (Huang *et al.*, 2007), accessed 12/21/2020, to test for significant enrichment, using the default categories implemented for mouse. From the XML-formatted entries of the NCBI Gene database, we also extracted gene names, gene symbols, summary comments, and expression data for 30 mouse tissues from the MODENCODE project (Yue *et al.*, 2014). To avoid repeated tests of tissues with highly similar expression profiles, we calculated all pairwise Pearson correlations between tissues and randomly removed one tissue of each pair having a correlation coefficient above 0.90, resulting in the exclusion of five tissues. Including highly correlated tissues would have falsely inflated the apparent number of independent tests, thereby making false discovery corrections too severe.

The cumulative distribution function of genes in consecutive windows was computed with the *ecdf* function in R (R Core Team, 2019) and the observed distribution of significant genes compared to this background expectation using the discrete Kolmogorov–Smirnov test of Dimitrova, Kaishev & Tan (2020). *L. cinereus* genes were assigned genomic coordinates based on the midpoint of the gene model, and tabulated in consecutive windows of 1 MB, with the final window on each scaffold truncated in length based on the residual sequence (genes annotated on short scaffolds less than 1 MB in length were excluded from this analysis). Scaffolds were ordered arbitrarily by name as this has no bearing on the test outcome. Rejection of the null hypothesis indicates that the observed distribution of significant genes differs from expected, but does not specify which genes contribute to the divergence.

Phylogenetic analyses were performed with MegaX (Kumar *et al.*, 2018). Synteny plots were based on protein similarity and computed with Mummer v.4.0.0 (Marçais *et al.*, 2018), with a seed length of 35 and nonunique matches filtered. Protein secondary structure was predicted with TMHMM v. 2.0 (Sonnhammer, von Heijne & Krogh, 1998).

RESULTS AND DISCUSSION

Genes under positive selection

Evolutionary rates of predicted *L. cinereus* coding sequences and their inferred orthologs in seven other bat species (orthogroups) were estimated with PAML. Of 9,447 orthogroups tested, 487 were initially significant and further reviewed manually, of which 90 were accepted as free of obvious errors affecting the *L. cinereus* branch. An additional 78 alignments were discarded because they exceeded a nonsynonymous tree-length threshold of 10 (see “Materials and Methods”), whereas 105 alignments were discarded on the basis of evident errors, with insufficient nonsynonymous variation in credible regions of the alignment to justify re-analysis. Gross alignment errors typically manifested as a combination of long gaps and runs of divergent residues unique to a single taxon that would be more parsimoniously explained by errors in the primary sequence prediction. In a few instances, the *L. cinereus* gene appeared to be paralogous based on *post hoc* BLASTP searches against all Chiroptera proteins in the *nr* database. Those searches revealed the *L. cinereus* sequence to be more similar to other bat proteins not included in the study than to sequences in the assigned orthogroup cluster. Errors of paralogy presumably arise because the true ortholog was lost or unannotated in *L. cinereus*. The remaining 214 initially significant alignments were manually trimmed and re-analyzed; of these, the *L. cinereus*-specific positive selection model was accepted for 60. One of these 60 significant orthogroup alignments (OG0012876, see [Supplemental File S1](#)) had a revised alignment length of 99 codons, below our 100-codon threshold for initial tests, but we chose not to exclude it from the second round of manual testing because the full protein was only slightly longer (102 amino acids) and invariant in the other seven species, whereas the *L. cinereus* sequence had four nonconservative substitutions not present in any other bat BLASTP protein match in *nr*.

In total, 150 tested alignments (1.6%) were considered significant, which is relatively high compared to some studies in other mammals. For example, [Beichman et al. \(2019\)](#) identified 18 positively selected genes in the ancestor of sea otters and [Kosiol et al. \(2008\)](#) identified a total of 144 significant genes across eight different branch-specific tests of mammalian genomes (*i.e.*, an average of 18 unique genes per mammalian lineage tested). On the other hand, [Shen et al. \(2010\)](#) identified 100 positively selected genes in the lineage leading to bats with other mammals as the background, and [Hawkins et al. \(2019\)](#) identified 181 positively selected genes from a combined analysis of diverse bat genomic and transcriptomic sequence. Thus, our results further support high rates of adaptive protein evolution in bats relative to other mammalian clades (but see [Jebb et al., 2020](#)).

For 25 of the 150 genes identified as under positive selection, the evolutionary rate ω_2 of positively selected genes was estimated to be one rather than greater than one. The average length of these 25 alignments was substantially shorter than the other 125 (455.6 vs 638.5 codons, respectively). While genes with ω_2 estimates of one may be false positives that are actually evolving neutrally, ω_2 estimates strongly trade off with the estimated proportion of sites (p_2) to which the rate applies ([Supplemental File S6](#)), suggesting that the joint estimation of these two parameters is challenging. Furthermore, a comparable

number of genes (29) had effectively infinite ω_2 estimates (*i.e.*, the programmatic maximum of 999), which seems no more realistic than an ω_2 estimate of one. We therefore considered only the overall model likelihoods in detecting positive selection and did not further require that ω_2 estimates exceed some threshold. Of course, we did require that the ω_2 estimate not be less than one, which is unlikely but formally possible even when the positive selection model is accepted (Yang, 2007).

We found a significant association between single-exon genes and the frequency of positive selection ($P = 4.28E-5$, Fisher's 2×2 exact test with $df = 1$; Supplemental File S7). A cautious view of this finding is that genes with high evolutionary rates are enriched in weakly constrained retrogenes, which are initially single-exon and may be unexpressed. However, transcript and protein expression data are ultimately required to evaluate gene activity and the evolutionary dynamics of retained retrogenes has been shown to be shaped by diversifying and purifying selection mechanisms (Vinckenbosch, Dupanloup & Kaessmann, 2006; Gayral *et al.*, 2007). Thus, neutral evolution is not necessarily expected for retrogenes, and in any event the statistical test we performed rejects a hypothesis of neutral evolution in favor of adaptive evolution. We also note that the number of single-exon genes in *L. cinereus* may be methodologically inflated, because noncoding exons are not readily predictable based on protein homology alone. Single-exon genes do generally appear to have higher evolutionary rates than multiexon genes regardless of these ascertainment issues (Shabalina *et al.*, 2010), so the statistical association we observed is consistent with multiple causes.

Relative expression data were available for 137 of the 143 mouse orthologs associated with positively selected genes and for 8,410 of the remaining orthogroups (Supplemental File S8). Figure 1A shows the proportion of genes in each of these two groups that have peak expression in the corresponding tissue, sorted by the intergroup difference. Figure 1B shows the average expression of genes in each group (normalized to their maximum) by tissue type, sorted by Student's *t* statistic for the corresponding difference of means. By both standards, the largest difference in expression between significant and nonsignificant genes was in thymus tissue. After pooling all other tissues into a single bin for comparison, Fisher's exact test confirmed a significantly higher incidence of peak expression in thymus for positively selected genes ($P = 0.023$). Mean expression of positively selected genes was significantly higher than nonselected genes in both thymus and heart (two-sided *t*-test with equal variance assumed, $P < 0.05$) after Benjamini–Hochberg correction for multiple tests. Biased expression in thymus of positively selected genes suggests positive selection on immunodevelopment, as the thymus is the site of T cell maturation (reviewed by Thapa & Farber, 2019), defects of thymus development are associated with immune dysfunction (*e.g.*, Bockman, 1997), and immunosenescence is associated with regression of the thymus with age (Palmer, 2013; Rezzani *et al.*, 2014).

Genomic clusters of significant genes

The cumulative distribution of significant genes along scaffolds (Fig. 2A) differed from that expected based on the distribution of all genes ($P = 1.99E-13$ by discrete Kolmogorov–Smirnov test). To identify clusters of significant genes that likely contributed

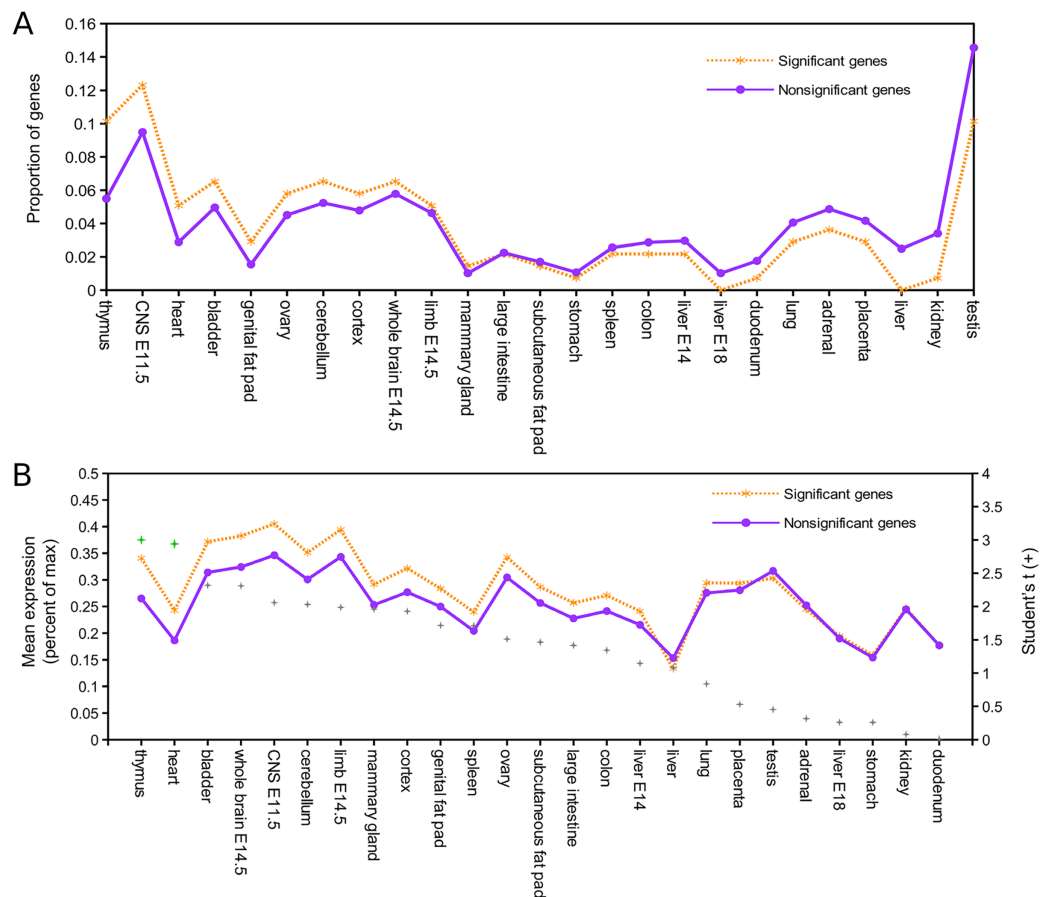


Figure 1 Genes positively selected in the *Lasiurus cinereus* lineage have biased expression in the mouse thymus. (A) Proportion of mouse orthologs of selected genes that have peak relative expression in each tissue. The proportion is significantly higher in thymus for positively selected genes than for other genes when all other tissues are binned together (see text for details). (B) The mean relative expression of positively selected genes is highest in thymus. The difference in means is significant at $P < 0.05$ (Student's t -test) for both thymus and heart after Benjamini-Hochberg correction for multiple tests. [Full-size !\[\]\(1663bb69f307a960345edb0e712f8c02_img.jpg\) DOI: 10.7717/peerj.13130/fig-1](https://doi.org/10.7717/peerj.13130/fig-1)

to this result, we tabulated the number of significant genes in 5-Mb windows and identified three windows containing three or more selected genes (Fig. 2B). Interestingly, all three clusters are near the end of their respective genomic scaffolds (Fig. 2C), although it remains to be confirmed whether they are actually subtelomeric in *L. cinereus*. One cluster contains five significant genes with midpoints in a 2-Mb window, situated near the terminus of scaffold 133, which at 109 Mb in length is the longest scaffold of the assembly. A second cluster contains three genes with midpoints within a 2-Mb window near the end of another long scaffold, scaffold 111. A third cluster contains six genes with midpoints within a 3-Mb window on scaffold 533. Two of these three clusters of positively selected genes lie in genomic regions of conserved synteny with other tetrapods, whereas the third is syntenic with other mammals (Supplemental File S9). While regions of conserved synteny are expected due to the variable persistence of ancestral vertebrate linkage groups (Simakov *et al.*, 2020), the overlap between positive selection and conserved synteny in

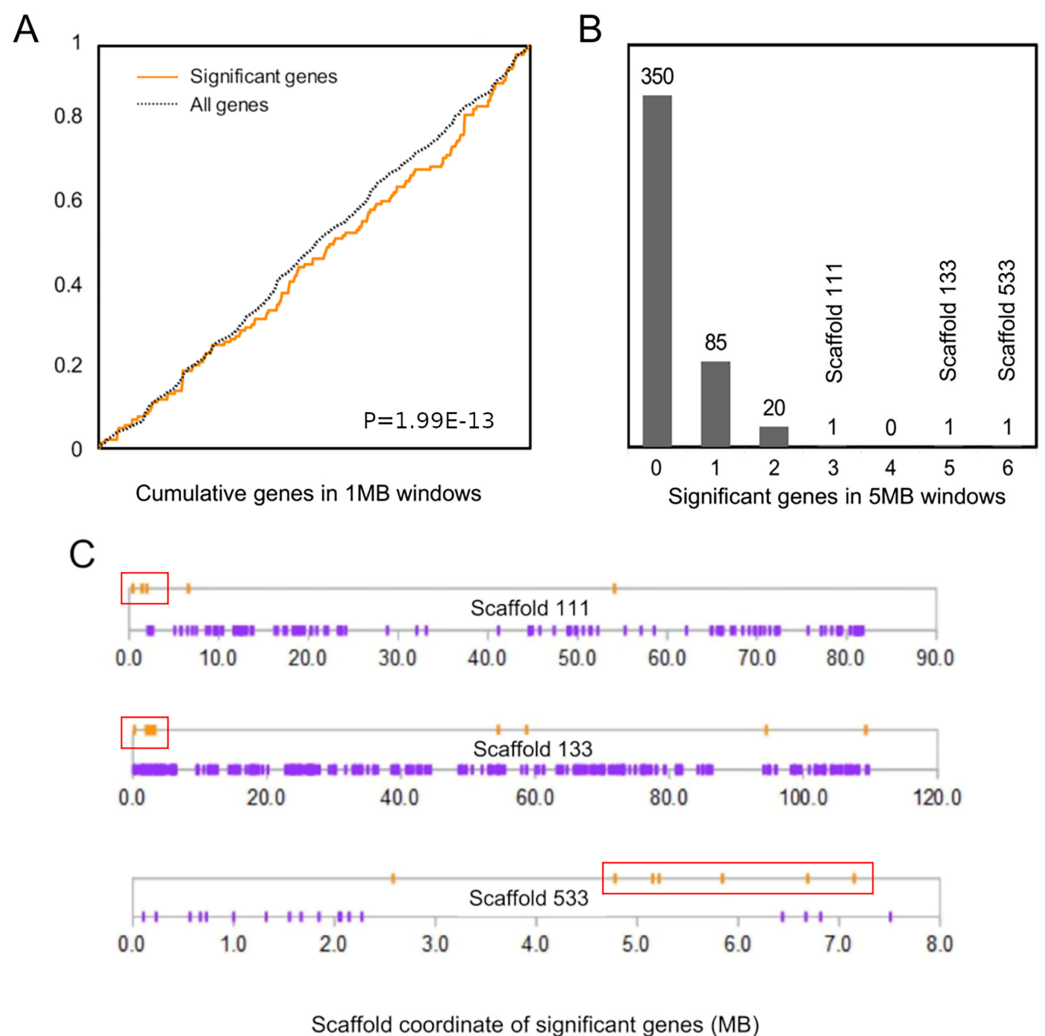


Figure 2 Positively selected genes are clustered in the *Lasiurus cinereus* genome. (A) The cumulative distribution of significant genes along scaffolds differs significantly from that of all tested orthogroups by discrete Kolmogorov–Smirnov test. (B) Distribution of the counts of significant genes in consecutive 5-Mb windows. The scaffold locations are shown for the three windows with three or more significant genes. (C) Schematic of the distribution of significant genes (orange marks) and nonsignificant genes (purple marks) by coordinate position on the three scaffolds marked in (B). Red boxes indicate the selected genes within 5-Mb windows marked in panel (B). [Full-size !\[\]\(ba1b80118482ccef74a5d718ca4d7242_img.jpg\) DOI: 10.7717/peerj.13130/fig-2](https://doi.org/10.7717/peerj.13130/fig-2)

L. cinereus is striking and suggests that selection may have targeted co-evolving genes maintained in syntenic blocks. Factors hypothesized to favor the maintenance of synteny include co-regulated expression *via* local chromatin structure or shared promoter elements (Reimegård *et al.*, 2017; Rokas *et al.*, 2018) and epistatic effects of multigene haplotypes on fitness (“co-adapted gene complexes”, *e.g.*, Dobzhansky & Pavlovsky, 1958; Michalak, 2008; Hittinger *et al.*, 2010).

The cluster of positively selected genes on scaffold 133 includes orthologs of the mouse genes *Fam173a*, *Lmf1*, *Tsr3*, *Cramp1l*, and *Ndufb10* (Supplemental File S10). Three of the five genes have roles at the nexus of redox homeostasis and metabolism, as *Fam173a* and *Ndufb10* are both mitochondrial membrane proteins that affect rates of

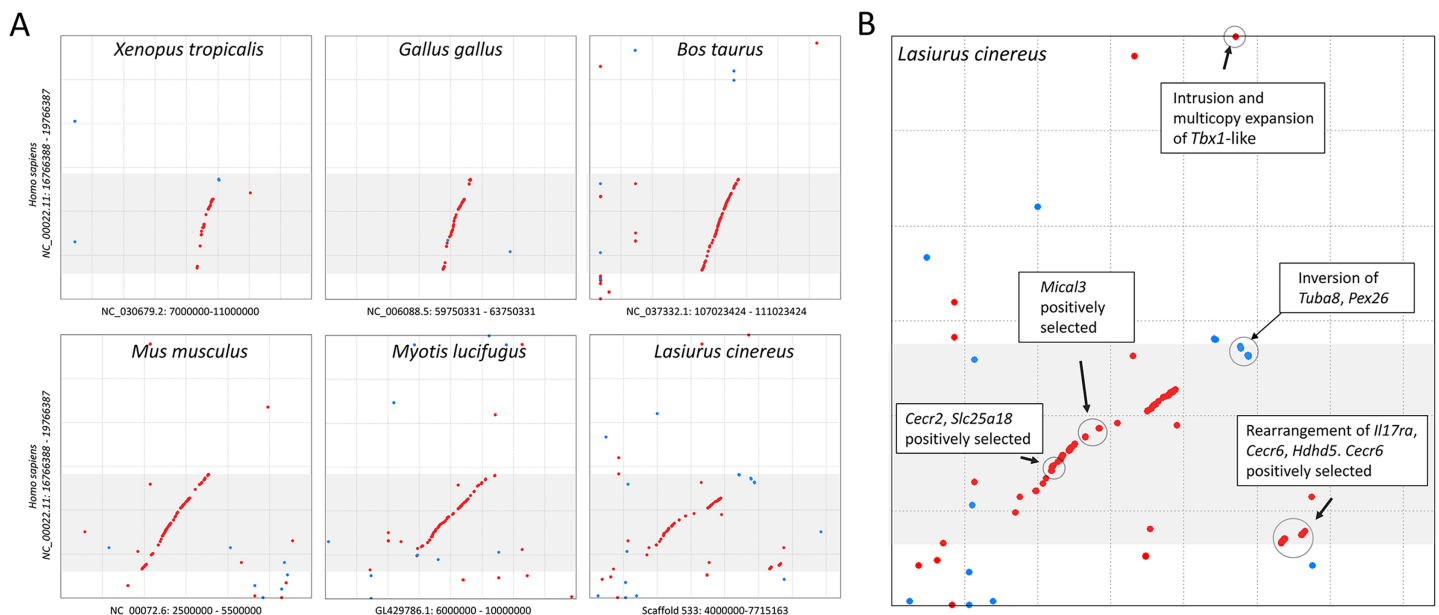


Figure 3 Positively selected genes clustered on scaffold 533 are syntenic with the human cat-eye critical region (CECR). (A) Protein-based alignment between human chromosome 22 in the vicinity of the CECR (gray-shaded region) and six other species. Red and blue dots indicate homology in the same and reverse orientations, respectively. (B) Positively selected genes and structural variation in the syntenic region in *Lasiurus cinereus*. Full-size DOI: 10.7717/peerj.13130/fig-3

mitochondrial respiration and *Lmf1* is generally responsive to redox stress as it relates to membrane-protein folding. Specifically, *Fam173a* encodes a mitochondrial lysine-specific methyltransferase that targets adenine nucleotide translocase (Malecki et al., 2019), and NADH dehydrogenase (ubiquinone) 1 beta subcomplex 10 (*Ndufb10*) is an accessory subunit of respiratory complex I important for complex assembly (Friederich et al., 2017). *Lmf1* (lipase maturation factor 1) is important for regulating plasma lipid levels and responds to redox stress on the endoplasmic reticulum (ER) (Mao et al., 2014). It functions as a molecular chaperone enabling proper folding and activity of ER membrane proteins. The other two genes in the cluster on scaffold 133 do not share an apparent functional association with the previous three genes. *Cramp1l* encodes a chromatin remodeling domain well conserved in vertebrates. *Tsr3* (20S rRNA accumulation factor) contributes to the maturation of 18S rRNA via the post-transcriptional modification of a conserved pseudouridine-like nucleotide (Meyer et al., 2016).

The cluster of positively selected genes on scaffold 533 includes *Rps13*, *Cecr2*, *Slc25A18*, *Mical3*, and *Cecr6*. All but the first listed of these genes are landmarks of the cat-eye critical region (CECR) on human 22q11 that is syntenic with other mammals (Puech et al., 1997; Footz et al., 2001). Alignments of the human region to scaffold 533 and scaffold GL42786.1 of *M. lucifugus* confirm the region is syntenic with bats as well (Fig. 3). Synteny is also shared with other tetrapods, such as chicken and *Xenopus*, but is not apparent with zebrafish. The overall conservation of synteny is belied by multiple types of structural variation unique to the *L. cinereus* lineage (Fig. 3, Supplemental File S11). For example, *L. cinereus* orthologs of the human CECR genes *Pex26*, *Il17ra*, and *Hdhd5* are present in a

rearranged order not found in any other species examined. Several genes annotated in *L. cinereus* were not present in at least some of the other bat species included in the study. Three genes appear to have undergone a copy-number expansion in the *L. cinereus* lineage: a *Tbx1* homolog, *Btf3*, and *Tuba8*, although only one gene copy of each is present on scaffold 533 (these expansions are discussed further below). A mixture of conserved synteny for a subgroup of genes and dynamic structural evolution also characterizes the human-mouse comparative alignment (Puech et al., 1997), which, coupled with the frequent clinical observation of genomic rearrangements of 22q11, suggests that the core set of CECR genes are maintained in physical proximity by purifying selection rather than a low rate of structural mutation *per se*. However, some other apparent novelties in *L. cinereus* are actually annotation artifacts on closer inspection (Supplemental File S11). For example, *Atp6v1e1* and *Bid* were not annotated in *L. cinereus*, but homologous sequences were detected by BLASTX searches of the genomic scaffold, whereas other genes putatively unique to *L. cinereus* based on orthogroup clustering are actually fragmented annotations of single genes, leading to orthogroup artifacts.

A third cluster of three positively selected genes on scaffold 111 exhibits partially conserved synteny with diverse tetrapods (Supplemental File S12). The genes are periostin (*Postn*), FRAS1-related extracellular matrix protein 2 (*Frem2*), and proline and serine-rich protein 1 (*Proser1*). A fourth positively selected gene encoding saccin (*Sacs*) is an additional 3-Mb away on scaffold 111 and exhibits conserved synteny in some but not all lineages (Supplemental File S12). However, in all species examined, *Postn*, *Frem2*, and *Proser1* were retained in the same order and relative orientation, even as other elements of chromosomal synteny were lost. No functional relationship among these four genes was apparent to us.

We used both automated and manual approaches to investigate functions of positively selected genes. The DAVID annotation webtool (Huang et al., 2007) did not identify significant clustering of ontology terms among the 143 mouse orthologs submitted, after correction for multiple testing. This contrasts with some other studies that have detected enrichment of annotation categories in positive selection analyses of various mammalian groups (Kosiol et al., 2008; Hawkins et al., 2019; Beichman et al., 2019). We also manually categorized significant genes based on diverse annotation data within NCBI Gene database entries for mouse and human orthologs (Supplemental File S13), supplemented with literature searches. Of 143 positively selected genes with mouse orthologs, we identified 23 genes (16.2%) with a role in immunity or disease, 17 genes (11.9%) related to metabolic homeostasis and energetics, 15 genes (10.6%) involved in mitochondrial respiration or biogenesis, and six genes (4.2%) with behavioral phenotypes. These categorizations were based on *post hoc* assessments of commonalities that in our judgement show particular relevance to hoary bat ecology. However, as a uniform process was not applied *a priori* to all tested orthogroups, no enrichment test can be performed for these manual classifications.

Selected genes with immune function and identification of a *Steep1* paralog

Immune function was the most prominent theme among selected genes (Supplemental File S13), consistent with patterns of selection on mammal genes generally (Kosiol et al., 2008) and bats specifically (Zhang et al., 2013; Hawkins et al., 2019; Jebb et al., 2020). While most listed genes are directly related to pathogen detection and immune response, some were included that are only indirectly related, in that their protein products are known to be co-opted by pathogens during infection (e.g., *Hnrnpul1*, *Nxf1*, *Eif5b*). We also note that genes classified as relating to immune processes often have diverse additional functions. For example, immune functions have been ascribed to selected genes involved in phosphatidylinositol signaling (*Pik3r5*, *Pik3ap1*, and *Myo10*) and inositol triphosphate signaling (*Dapp1* and *Plcb2*), yet both signaling pathways regulate cellular homeostasis generally (Koyasu, 2003; Hogan, Lewis & Rao, 2009). Nonetheless, several commonalities exist among these immune-classified genes that support their biological relevance as a group. At least four genes regulate or are regulated by interferons, including *Josd1* (Wang et al., 2017), *Psmb9* (Kataoka et al., 2021), *Usp21* (Chen et al., 2017), and *Dapp1* (Onyilagha et al., 2015). Six genes are involved in surface antigen processing and signaling, including the aforementioned *Psmb9* as a component of the immunoproteasome that processes MHC Class I antigens (Singal, Ye & Quadri, 1995). *B4gat1* is involved in synthesizing poly-N-acetyllactosamine, which contributes to diverse glycoprotein structures including antigens (Ujita et al., 1999; Lee, Kohler & Pfeffer, 2009), *Lrmp* is involved in antigen processing within immune cells (Snyder et al., 1997), *Chst2* is involved in L-selectin mediated leukocyte adhesion (Li et al., 2001), and *Itgb2* encodes a major leukocyte adhesion protein (Arnaout, 1990). *Cul7* negatively regulates two somatic mechanisms of immunoglobulin diversification in B lymphocytes, somatic hypermutation and recombinant class-switching (Luo et al., 2019). The previously mentioned *Dapp1* mediates antigen-specific conjugation of B and T cells in mice (Al-Alwan et al., 2010). We speculate that the energetic demands of migration, hibernation, and solitary roosting behavior on hoary bats favor altered immunoresponse dynamics in lasiurine bats compared with ancestral Vespertilionids.

We initially identified *Steep1* as another immunity gene under positive selection. *Steep1* encodes a transport chaperone of STING, a critical regulator of interferon expression that responds to signatures of viral infection (Ishikawa, Ma & Barber, 2009; Holm et al., 2012; Härtlova et al., 2015; Liu et al., 2019b). The putative *Steep1* ortholog of *L. cinereus* had seven unique amino-acid substitutions among 208 aligned positions, whereas the other bat species were invariant except at gapped (masked) positions. However, additional BlastP searches showed that the *L. cinereus* sequence is actually paralogous to *Steep1* and thus the initial selection test was invalid. These homology searches identified a novel paralog within suborder Yangochiroptera only that is orthologous to *Ensembl*-annotated gene ENSMLUG00000026380. Although the coding sequence of this *M. lucifugus* gene occurs on a single exon, suggesting a retrogene, an initial noncoding exon is annotated and

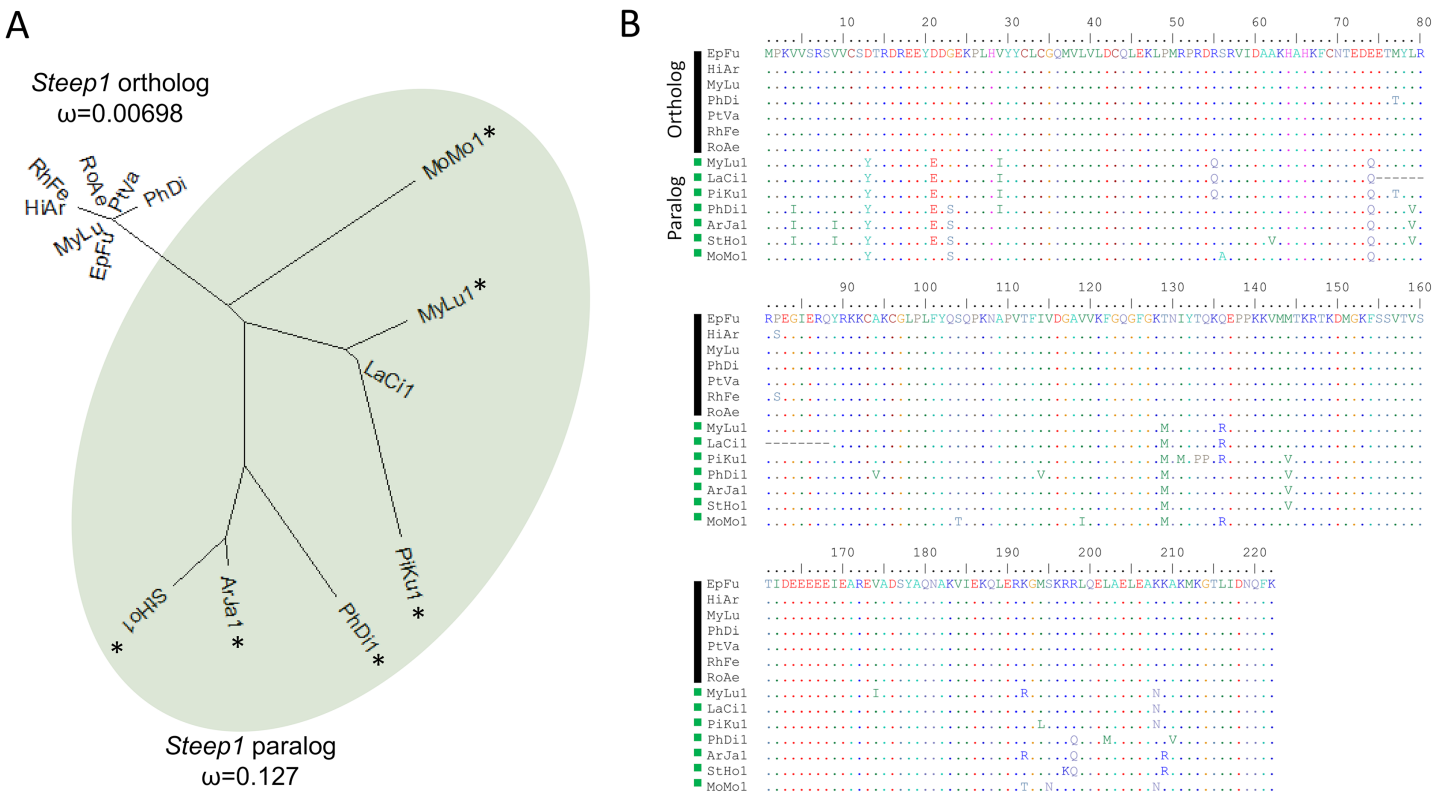


Figure 4 A derived *Steep1* paralog is positively selected within Yangochiroptera. (A) Neighbor-joining dendrogram of the *Steep1* ortholog together with the *Steep1* paralog (shaded in green), with distances based on nonsynonymous amino-acid substitutions only. The evolutionary rate parameter ω computed under PAML Model 0 is shown for each group of orthologous sequences. The likelihoods of PAML Model 2 were computed with and without branch labels denoting the *Steep1* paralog as the foreground branch. The labeled model was significantly more likely than the unlabeled model or Model 1. (B) A protein alignment of bat *Steep1* orthologs and the novel paralogous sequences in Yangochiroptera. Four-letter species codes for bats are: EpFu = *Eptesicus fuscus*, MyLu = *Myotis lucifugus*, HiAr = *Hipposideros armiger*, RhFe = *Rhinolophus ferrumequinum*, RoAe = *Rousettus aegyptiacus*, PtVA = *Pteropus vampyrus*, PhDi = *Phyllostomus discolor* [*Steep 1* paralog clade] MoMo = *Molossus molossus*, MyLu = *M. lucifugus*, LaCi = *Lasiurus cinereus*, PiKu = *Pipistrellus kuhlii*, PhDi = *P. discolor*, ArJa = *Artibeus jamaicensis*, StHo = *Sturnira hondurensis*. Full-size DOI: 10.7717/peerj.13130/fig-4

supported by RNAseq evidence in the Gene database entries of all putative orthologs we identified (Fig. 4, Supplemental File S14).

While gene duplication is a common evolutionary occurrence, the potential for alternative regulation of STING *via* a *Steep1* paralog is of particular interest given that STING is a powerful mediator of innate immunity, particularly to viruses (Clayton & Munir, 2020), and one that has evolved dampened activity in bats (Xie et al., 2018). Indeed, human symptoms associated with infection by the coronavirus SARS-CoV-2 have been characterized as a detrimental STING response (Berthelot & Lioté, 2020). We therefore performed a revised test of positive selection in which we labeled the duplicated orthologous group as the foreground branch with the ancestral *Steep1* as the background (Fig. 4). The resulting test of diversifying selection acting on the duplicated *Steep1* was highly significant ($P = 6.95E-9$).

Interestingly, a processed retrogene of *Steep1* is present in both human (ENSG00000233170) and mouse (ENSMUSG00000048188), of independent origins based

on their genomic locations and absence in sister taxa. The human ‘pseudogene’ lies within an intron of *Erlin2* in both Old World and New World monkeys (e.g., ENSANAG00000019514) but is absent in the tree shrew genome (Supplemental Figure S15A). While it is beyond the scope of this study to review annotation evidence in all relevant genomes, a survey of representative New World and Old World primates (Supplemental Figure S15B) reveals that the derived ‘pseudogene’ on human chromosome 8 is well conserved despite a presumed origin circa 50 Mya, based on molecular clock dating of New and Old World primate divergence (Perez et al., 2013). As with the paralog in Yangochiroptera, Gene database entries suggest that the primate *Steep1* retrogene is transcribed above the background level of surrounding intron sequence (Supplemental Figure S15A, <https://www.ncbi.nlm.nih.gov/gene/666594>). The functions and evolutionary dynamics of retained *Steep1* retrogenes may collectively merit further exploration.

Energetics

Bats have very high metabolic rates, primarily due to the energetic demands of flight (O’Shea et al., 2014) and positive selection on cellular respiration genes has previously been shown for bats as a whole relative to other mammals (Shen et al., 2010). *L. cinereus* is further challenged by its extreme migratory behavior and seasonality of habitat and of prey. It is reasonable to suspect that innovations related to heterothermy, torpor, hibernation, and migratory capacity (e.g., Davis, 1970; Cryan & Wolf, 2003; Geiser, 2013; McGuire, Fenton & Guglielmo, 2013; Weller et al., 2016) would entail selective pressure on genes related to energetics and metabolic homeostasis.

Of the 17 positively selected genes we identified relating to lipid metabolism and energetics (Supplemental File S13), at least five have roles in regulating adipogenesis. Mesenteric estrogen dependent adipogenesis (*Medag*) is adipogenic (Zhang, Chen & Sairam, 2012) whereas zinc-finger transcription factor *Zfp521* is a negative regulator of adipocyte differentiation from stem cells (Chiarella et al., 2018). Subunit *Med23* of the Mediator Complex is directly required to transduce insulin signaling so that adipocyte differentiation can be initiated by *Krox20* (Wang et al., 2009). *Mtch2* promotes both adipogenesis (Jiang et al., 2019) and lipid accumulation (Bar-Lev et al., 2016) and is strongly expressed in white adipocytes (Kulyté et al., 2011). Genetic variation in *Mtch2* has been linked to human obesity (Kulyté et al., 2011). *Mtch2* is particularly of interest with regard to hibernation physiology, as differences in *Mtch2* transcript or protein abundance are associated with hibernation states of brown bear (*Ursos arctos*) (Chazarin et al., 2019) and ground squirrel (*Ictidomys tridecemlineatus*) (Ballinger et al., 2016). A combined transcriptomic – proteomic approach associated *Zfp512* abundance with increased intramuscular fat deposition and adipogenesis in pigs (Ma et al., 2020).

At least eight positively selected genes have roles in modulating the sensitivity of differentiated cells to metabolic hormones (Supplemental File S13). Regulatory factor X6 (*Rfx6*) has strongly biased expression in K cells and regulates their secretion of gastric inhibitory protein, thereby increasing insulin sensitivity and lipid accumulation (Suzuki et al., 2013). SH2B adaptor protein 1 (*Sh2b1*) is a potent determinant of fat accumulation that is responsive to leptin signaling; *Sh2b1* deletion in mice reduces brown fat function

and alters body temperature (Jiang et al., 2020). Rap guanine exchange factor *Rapgef4* mediates cyclic AMP stimulation of insulin secretion by pancreatic β -cells (Song et al., 2013; Henquin & Nenquin, 2014) whereas deletion of replication initiator 1 (*Repin1*) promotes insulin sensitivity via glucose stimulation (Kunath et al., 2016). Transcription factor *E2f8* promotes gluconeogenesis and when overexpressed impairs insulin sensitivity in hepatocytes (Chen et al., 2019). *Perlecan* inhibits adipocyte growth, reduces insulin sensitivity, and promotes oxidative muscle fiber development (Yamashita et al., 2018). In cultured islet cells, *Gjb4* inhibits proliferation, insulin secretion, and apoptosis, and the gene is highly expressed in a mouse model of obesity (NZO), in contrast to the leptin-deficient obesity strain B6-ob/ob (Gässler et al., 2020). Deletion or inhibition of phosphodiesterase 10A (*Pde10a*) limits the onset of diet-induced obesity and insulin resistance in mice (Nawrocki et al., 2014).

Three positively selected genes encode lipid metabolizing enzymes or proteins that directly modulate their activity. Phospholipid phosphatase 1 (*Plpp1*) encodes a membrane-bound protein that supports uptake of lipids from the extracellular space (Brindley & Waggoner, 1998), lipase maturation factor 1 (*Lmf1*) is required for the correct folding of certain lipases including hepatic lipase and lipoprotein lipase, but also has a more general role in lipid homeostasis in relation to redox stress as mentioned previously (Ehrhardt, Bedoya & Péterfy, 2014). *Etfbkmt* is a methyltransferase that alters the activity of the 'hub' protein ETF- β (Rhein et al., 2014), which as a heterodimer with ETF- α delivers electrons from diverse donors to the membrane-bound electron transport chain (Henriques et al., 2021). The ETF heterodimer controls the flux of fatty-acid oxidation and may modulate energy production during fasting or otherwise when fatty-acid oxidation is upregulated.

Two positively selected genes were involved in the behavioral regulation of metabolism by the hypothalamus. The ligands of neuropeptides B/W receptor 1 (*Npbwr1*) strongly affect feeding behavior of mice in a photoperiod-dependent manner (Tanaka et al., 2003). Male mice that were *Npbwr1*-null exhibited adult-onset obesity that increased with age and when fed a high-fat diet, were hyperphagic, and had decreased locomotor activity (Ishii, Fei & Friedman, 2003). Plasma levels of glucose, leptin, and insulin were all elevated in these males whereas female nulls did not show metabolic defects. Single-minded homolog (*Sim1*) is essential for the developing paraventricular nucleus of the hypothalamus (Michaud et al., 1998) and experimental *Sim1* deficiency, conditional knockout, and ablation of *Sim1* neurons all cause altered feeding behavior (Holder et al., 2004; Xi et al., 2012; Tolson et al., 2014). Genetic variation at the *Sim1* locus is associated with human obesity risk as well (Ramachandrapa et al., 2013; Liu et al., 2019a).

Fifteen genes associated with mitochondria were positively selected, some of which overlap with the lipid homeostasis genes already described, which is to be expected given the central role of mitochondria in energy production. Selected genes encoding mitochondrial membrane transporter components include *Slc25a1*, *Slc25a18*, *Mtch2*, and *Fam173a*, whereas selected genes encoding electron transport proteins include *Cyc1*, *Uqcrrf1*, *Ndufa12*, *Ndufb10*, and *Etfbkmt*. Genes important for mitochondrial maintenance include *Peo1* (Sarzi et al., 2007), *Prorp* (Howard et al., 2012), *Pum2* (D'Amico

et al., 2019), and *Capn3*. *Capn3* mutations cause heterogeneous forms of autosomal recessive limb-girdle muscular dystrophy in humans, due in part to abnormal mitochondrial biogenesis in skeletal muscle (Kramerova *et al.*, 2009; Yalvac *et al.*, 2017).

Sensory and behavioral gene functions

Perhaps surprising given previous results (Kosiol *et al.*, 2008; Jebb *et al.*, 2020), we did not identify positively selected genes with unambiguous roles in sensory perception, although a strong candidate is *Chrna9*. This gene encodes the $\alpha 9$ nicotinic cholinergic receptor subunit, which is strongly expressed in sensory hair cells of the ear (Scheffer *et al.*, 2015; Liu *et al.*, 2018). The gene has been studied in the contexts of auditory stimulus modulation by cochlear hairs (e.g., Boero *et al.*, 2020) and maintenance of orientation and balance by vestibular hairs (e.g., Poppi *et al.*, 2018). For example, analyses of mouse mutants with altered or lost function suggest that *Chrna9* modulates the “gain” (stimulus strength) of various auditory signals by altering cochlear sensitivity, including sensitivity to sound frequency, to persistent background noise, and to extreme noise (Taranda *et al.*, 2009; Wedemeyer *et al.*, 2018; Boero *et al.*, 2020). We find selection on *Chrna9* intriguing given the challenges to sensory discrimination associated with echolocation (Corcoran & Moss, 2017) and the recent identification of a novel microcall form of echolocation by *L. cinereus* (Corcoran & Weller, 2018). However, *Chrna9* is expressed in diverse mammalian tissues and has also been implicated in the perception of pain (Vincler *et al.*, 2006) and psychological stress response (Mohammadi, Burton & Christie, 2017). Whether positive selection on *Chrna9* is related to sensory hair function in *L. cinereus* therefore remains to be determined. Additionally, acetylcholine receptor subunits form diverse heteromeric combinations that have distinct properties (Millar, 2003), such that the phenotypic significance of the *L. cinereus Chrna9* sequence may depend on the particular channel complex in which it is constituted.

Several radical amino-acid changes unique to the *L. cinereus* CHRNA9 protein (Fig. 5) occur in the extracellular N-terminal region that contains the acetylcholine binding loops (Taly *et al.*, 2009) and an otherwise conserved glycosylation site is uniquely abolished in this region (the substitution was confirmed by inspecting raw reads aligned to the reference). Glycosylation sites are important both for cellular trafficking during protein maturation and for assembly of multimeric complexes, but it is difficult to predict the effects of their abolishment (Wanamaker & Green, 2005). Substitutions in the extracellular region of CHRNA9 have been shown to alter calcium permeability of $\alpha 9\alpha 10$ cholinergic receptors in mammalian cochlear hair cells (Lipovsek *et al.*, 2014), although the two mutations studied in that work are not evident in *L. cinereus* or in bats generally (Fig. 5). Importantly, the long intracellular spacer region linking transmembrane domains is highly variable among bat species, likely due to reduced functional constraint, potentially biasing the selection test. We therefore repeated the PAML selection test on the N-terminal extracellular region only and confirmed a significant result ($P = 0.00089$, $\omega_2 = 1.000$). Functional analyses (cf. Lipovsek *et al.*, 2014; Marcovich *et al.*, 2020) of bat *Chrna9* haplotypes could be employed to potentially identify divergent properties of *L. cinereus* sequences.



Figure 5 Protein substitutions unique to the *L. cinereus* Chrna9 gene relative to protein structure and functional sites. Black squares denote bat sequences included in the selection test, with additional bat sequences and the human sequence also included for comparison. Alignment coordinates are shaded orange, pink, or blue according to the secondary structure predicted by TMHMM (Krogh *et al.*, 2001). Gaps occur at each end of the alignment because the tested orthogroup alignment was shorter than the full human protein sequence, for which glycosylation sites (red stars) are annotated in the corresponding NCBI accession record. The second glycosylation site has been abolished by an N→K amino acid change in the *L. cinereus* sequence. See text for additional details.

Full-size DOI: 10.7717/peerj.13130/fig-5

Some of the positively selected genes we identified have documented behavioral phenotypes in human or mouse, including impacts on diurnality, feeding behavior, and social interactions. For example, polymorphisms of the largely uncharacterized gene

D430041D05Rik were associated with diurnal preference in a human twin study (Melroy-Greif et al., 2017) and had previously been associated with suicidal tendencies in bipolar patients (Willour et al., 2012). The *Timeless* gene encodes a circadian clock component; a nonsense mutation alters photic entrainment and causes familial advanced sleep phase – a shifting forward of diurnal activity but with normal photoperiod – in human and mouse (Kurien et al., 2019). Several positively selected genes affecting metabolic homeostasis are also linked to behavioral phenotypes when manipulated, including the genes *Sim1* and *Npbwr1* that were discussed above in the context of behavioral control of energy balance. *Sim1* expressing neurons enable normal melanocortin-4-associated sexual behaviors in male mice (Semple & Hill, 2018) and linked genetic variants are associated with erectile dysfunction in human (Jorgenson et al., 2018). *Npbwr1* knockdowns altered stress-response behaviors and circadian rhythm in mice, including prolonged and aggressive interactions with unknown “intruder” mice (Nagata-Kuroiwa et al., 2011). A human study associated a variant *Npbwr1* allele with altered perceptions of, and reactions to, angry facial expressions (Watanabe et al., 2012). In addition to effects on insulin sensitivity, *Rapgef4*-deficient mice exhibit reduced social interactions and ultrasonic vocalizations but no impairment of memory, learning, or nonsocial behaviors (Srivastava et al., 2012). *Pde10a* is highly expressed in neurons of the striatum (Seeger et al., 2003), which are important for limb coordination and motivated learning, and *Pde10a* knockdown impairs motivated learning in mice *via* both positive and negative stimuli (Piccart et al., 2011). Collectively, these behavioral phenotypes are intriguing with respect to the ecology of hoary bats, which are conspicuous in roosting predominantly in isolation (Shump & Shump, 1982), exhibiting sex biased migration that may relate to divergent thermoregulatory strategies (Cryan & Wolf, 2003; McGuire, Fenton & Guglielmo, 2013), and frequent use of facultative torpor during migration (McGuire, Jonasson & Guglielmo, 2014). More speculatively, circadian rhythm and patterns of exploratory or anxious movement are believed to contribute to migratory propensity in birds (Ramos, Delmore & Liedvogel, 2017 and references therein) and could have similar effects in bats.

Genes associated with craniodysmorphism

A remarkable number of positively selected genes in *L. cinereus* have orthologs in human associated with craniodysmorphism when deleteriously mutated. They include five genes associated with the CECR and DGCR regions mapped on human 22q11. Cat-eye syndrome is a dosage-sensitive syndrome arising from segmental or supernumerary duplication (Footz et al., 2001). While defects are congenital, individuals may be developmentally mosaic for the chromosomal duplication and thus phenotypic penetrance is highly variable. Mild craniofacial dysmorphism, particularly affecting the eyes and ears, is characteristic of the syndrome but other conditions include heart and renal defects and anal fistula. Of the five positively selected genes syntenic with the CECR, *Slc25A18* and *Cecr2* are most strongly implicated in disease etiology. *Slc25A18* is a mitochondrial membrane transport channel and *Cecr2* contains a bromodomain associated with chromatin-based transcriptional regulation. Additional positively selected genes associated

with craniodysmorphism include *Frem2*, which underlies the cryptophthalmia and syndactyly of Fraser syndrome (Van Haelst et al., 2008; Zhang et al., 2019a), *Perlecan*, which is implicated in Schwarz–Jampel syndrome (Stum et al., 2006), and *Cul7*, a cause of 3M syndrome (Huber et al., 2005). These human syndromes may ultimately reflect developmental pleiotropy (Lindsay et al., 2001), i.e., impacts on tissues that arise from a common primordium. Our findings most likely signal developmental novelty affecting the embryonic pharyngeal arch that gives rise to thymus, heart, and head, which would be consistent with the expression biases shown in Fig. 1.

Expansion of a *Tbx1*-like gene in *L. cinereus*

As mentioned previously, orthogroup copy numbers in the CECR region (Supplemental File S11) initially suggested expansions of the genes *Tbx1*, *Btf3*, and *Tuba8* within the *L. cinereus* lineage. Most copies of these three orthogroups are dispersed in the *L. cinereus* genome and have multiple exons, i.e., they are not processed retrogenes or tandem duplicates. However, some annotations are fragmented and we cannot assess the activity or function of individual gene copies from the present data. While a general investigation of copy number evolution in *L. cinereus* is outside the scope of this work, the causal link between *Tbx1* dosage and developmental phenotypes in human and mouse that overlap with CECR and DGCR phenotypes (discussed below) suggests a connection between positive selection and gene amplification in the region. Gene amplification is an important but often ephemeral mechanism of adaptation in response to selection on gene expression levels in bacteria (Tomanek et al., 2020) and invertebrates (Bass et al., 2013) but its role in vertebrates is less clear.

Two of the potentially expanded genes, *Tbx1* and *Tuba8*, are also present on human 22q11, and *Tbx1* dysregulation is considered causal for DiGeorge syndrome (Merscher et al., 2001), whereas *Tuba8* is adjacent to and dysregulated by DGCR haploinsufficiency but not implicated in the syndrome itself (Dantas et al., 2019). In contrast, the functional human ortholog of *Btf3* is located on chromosome 5 and is a basal transcription factor necessary for transcription initiation (Zheng et al., 1990). *Btf3* pseudogenes appear to be common in mammalian evolution as numerous presumed pseudogenes are listed in the Ensembl database for both human and mouse. *Tbx1* is a key transcription factor in many developmental pathways, including cardiac morphogenesis (Alfano et al., 2019) and otocyst (inner ear) development (Raft et al., 2004), and is an antagonist of retinoic acid pathways (Koop et al., 2014). In human, *Tbx1* haploinsufficiency is associated with congenital defects of the heart, hypoplasia of the thymus, hypoparathyroidism, and facial dysmorphism (Du, de la Morena & van Oers, 2020).

Alignment of the *Tbx1* orthogroup sequences revealed that the genes annotated in *L. cinereus* are structurally distinct from *Tbx1* annotated in the other bat species, although a similar homolog is also found in *E. fuscus*. Only the 5' portion of the *Tbx* domain is present in these “*Tbx1*-like” sequences, and the sequences are shorter than other *Tbx* homologs. BLASTX searches demonstrated that the true *Tbx1* ortholog is in fact present in *L. cinereus* but unannotated on scaffold 536, which is colinear with the *Tbx1*-containing region of *E. fuscus* (Supplemental File S16). A phylogeny of the *Tbx1*-like sequences

together with BlastP matches from the other vespertilionid bats studied demonstrates that the novel gene family is not a mis-annotation of expected *Tbx* homologs, and that it is expanded only in the *L. cinereus* lineage (Supplemental File S16). Importantly, both deletion and overexpression of *Tbx1* in mouse generate DiGeorge-like phenotypes in a dose-dependent manner (Liao et al., 2004), suggesting that amplification of the novel *Tbx1*-like gene could also drive a dosage-sensitive phenotype related to the positively selected genes in the region.

Limitations and caveats of this study

While our annotations are based on a draft genome uninformed by RNA-seq data, the numbers of genes and orthogroups identified were on par with expectation and the number of positively selected genes identified was comparable to some recent studies in bats (Shen et al., 2010; Hawkins et al., 2019). Nonetheless, there are general limitations and biases to identifying positive selection based on excess nonsynonymous mutation alone. Coding-sequence prediction is vulnerable to sequencing error in the underlying genome, which disrupts features such as splicing signals and reading frames, thereby reducing the number of codons available for analysis or introducing alignment errors. On the other hand, sequencing error should not directly bias evolutionary rate estimates because it impacts synonymous and nonsynonymous sites proportionally, and also because the other analyzed genomes are similarly affected. Alignment ambiguity is another major source of uncertainty, particularly in more variable protein regions, resulting in incorrectly inferred codon orthology and substitution rates. Even when the primary sequence is correct and complete, substitution rates must be estimated for abstracted models with simple parameterizations. Excess nonsynonymous substitution as a criterion of positive selection may be inadequate to detect many actual instances of positive selection, in part because the functional significance of accrued amino-acid changes need not be related to their number *per se* (e.g., a single nonsynonymous substitution strongly alters STING activity in bats (Xie et al., 2018)), and also because the excess is evolutionarily transient, such that statistical power depends on the particular set of taxa investigated.

Beyond these ascertainment issues, inferring the biological relevance of positively selected genes presents a further challenge. Pairwise divergence between bats and rodents is high (Jebb et al., 2020), such that the transferability of functional annotations from model organisms may be more tenuous than for many other mammals. Studies of mammalian gene function also depend heavily on gross manipulations in mouse and genetic analyses of clinical disorders in human, yet deleterious phenotypes induced by severe alterations of gene expression or sequence may be dissimilar to positively selected phenotypes in nature. Despite these caveats, a number of independent factors support the biological relevance of the genes we identified. While we manually categorized genes after the fact and ascertainment biases likely exist among functional categories (e.g., phenotypes relevant to human health are more likely to be identified), the strong representation of genes with immune function or mitochondrial function is consistent with the literature (Kosiol et al., 2008; Shen et al., 2010; Zhang et al., 2013; Hawkins et al., 2019), particularly as it relates to interferon regulation (Zhang et al., 2013; Xie et al., 2018).

Furthermore, both biased expression (higher expression in thymus) and biased genomic location (physical clustering) were observed among positively selected genes, with no plausible connection between these variables and false positive rates. Moreover, all three clusters of positively selected genes discussed in the text occur within regions of ancient synteny, further supporting the view that the genomic clustering we detected is biologically relevant and not due to chance.

CONCLUSIONS

Hoary bats migrate vast distances, are challenging to census, and have poorly characterized hibernacula. As a result, their ecology remains poorly understood despite increasing conservation concerns. Here we have used a relatively rapid, low-cost approach to identify candidate targets of positive selection since divergence of the hoary bat lineage from other Vespertilionid groups, suggesting that the present-day ecology of this group rests in part on substantial evolutionary innovation. Functional annotations and expression data for selected genes strengthen the view that immunity and metabolism are important targets in hoary bat evolution. However, the observation that positively selected genes cluster in regions of ancient synteny is to our knowledge unprecedented in vertebrates and suggests that architecturally constrained regions are not precluded from diversifying selection and may even be hotspots of adaptation.

Future directions

Assemblies and their annotation are inherently progressive and we expect additional data to further refine the *L. cinereus* gene set. We consider the coding sequences used here preliminary, to be superseded by RNA-supported annotations of a chromosome-level assembly (recently made available by DNAZoo (<https://www.dnazoo.org>) but not presently available from NCBI). In turn, orthogroup delineation can be enhanced by incorporating local genomic synteny in addition to protein similarity, such as was applied by *Jebb et al. (2020)* to six bat genomes. For individual candidate genes, the underlying primary sequences should be confirmed through manual annotation if necessary and selection tests buttressed with additional phylogenetic representation. The expression of genes that remain strong candidates after this additional vetting should be characterized specifically in *L. cinereus* and background species. While many bat species are not candidates for captive rearing or experimental manipulation, behavioral and physiological tests may sometimes be tractable, as might immune challenges. *In vitro* or *in vivo* tests of protein function may also be feasible in some cases, such as has been used to investigate acetylcholine receptor function (*Marcovich et al., 2020*).

Future work may move beyond single-copy orthogroups to explore genes with more complicated evolutionary histories as well as the expansion and contraction of multi-gene families. These approaches will benefit greatly from ongoing genome sequencing initiatives for bats (*Teeling et al., 2018*). In addition to refining the analysis of protein evolutionary rate, other statistical signatures of natural selection can be explored that reveal more recent selection pressures, including population genomic scans for selective sweeps (*Pinzari et al., 2020*). Resequencing of opportunistically sampled bats obtained

from wind-turbine facilities and targeted sampling of congeners and disjunct *L. cinereus* populations would provide a powerful framework for identifying more recent episodes of selection, for example signatures of turbine-imposed selection if such exists. Such strategies might also reveal shifts in the neuroendocrine control of metabolism in island populations with altered seasonality and reduced migration.

ACKNOWLEDGEMENTS

Selection test methods applied in this study were facilitated using best practices and scripts provided as part of a Next Generation Sequencing-based workshop sponsored by the National Science Foundation (Award # 1744877 to S. Santagata, Long Island University). We are grateful to reviewers who provided their time and expertise to improve the manuscript. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

The authors received no external funding for this work.

Competing Interests

The authors declare that they have no competing interests.

Author Contributions

- Robert S. Cornman conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft.
- Paul M. Cryan conceived and designed the experiments, authored or reviewed drafts of the paper, and approved the final draft.

Data Availability

The following information was supplied regarding data availability:

The hoary bat genome sequence that was annotated is available from the ScienceBase repository: Cornman, R.S., Fike, J.M., Oyler-McCance, S.J., and Cryan, P.M., 2021, Genetic variation in hoary bats (*Lasiurus cinereus*) assessed from archived samples: U.S. Geological Survey data release, <https://doi.org/10.5066/P9VSG54Z>.

The gene annotations in gff format and orthogroup sequence alignments in FASTA format are available in the [Supplemental Files](#).

Supplemental Information

Supplemental information for this article can be found online at <http://dx.doi.org/10.7717/peerj.13130#supplemental-information>.

REFERENCES

Adams RA, Pedersen SC. 2013. *Bat evolution, ecology, and conservation*. Berlin: Springer.

- Al-Alwan M, Hou S, Zhang T, Makondo K, Marshall AJ. 2010. Bam32/DAPP1 promotes B cell adhesion and formation of polarized conjugates with T cells. *The Journal of Immunology* **184**(12):6961–6969 DOI [10.4049/jimmunol.0904176](https://doi.org/10.4049/jimmunol.0904176).
- Alfano D, Altomonte A, Cortes C, Bilio M, Kelly RG, Baldini A. 2019. Tbx1 regulates extracellular matrix-cell interactions in the second heart field. *Human Molecular Genetics* **28**(14):2295–2308 DOI [10.1093/hmg/ddz058](https://doi.org/10.1093/hmg/ddz058).
- Arnaout MA. 1990. Structure and function of the leukocyte adhesion molecules CD11/CD18. *Blood* **75**:1037–1050.
- Ballinger MA, Hess C, Napolitano MW, Bjork JA, Andrews MT. 2016. Seasonal changes in brown adipose tissue mitochondria in a mammalian hibernator: from gene expression to function. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology* **311**(2):R325–R336 DOI [10.1152/ajpregu.00463.2015](https://doi.org/10.1152/ajpregu.00463.2015).
- Bar-Lev Y, Moshitch-Moshkovitz S, Tsarfaty G, Kaufman D, Horev J, Resau JH, Tsarfaty I. 2016. Mimp/Mtch2, an obesity susceptibility gene, induces alteration of fatty acid metabolism in transgenic mice. *PLOS ONE* **11**(6):e0157850 DOI [10.1371/journal.pone.0157850](https://doi.org/10.1371/journal.pone.0157850).
- Bass C, Zimmer CT, Riveron JM, Wilding CS, Wondji CS, Kausmann M, Field LM, Williamson MS, Nauen R. 2013. Gene amplification and microsatellite polymorphism underlie a recent insect host shift. *Proceedings of the National Academy of Sciences USA* **110**(48):19460–19465 DOI [10.1073/pnas.1314122110](https://doi.org/10.1073/pnas.1314122110).
- Beichman AC, Koepfli K-P, Li G, Murphy W, Dobrynin P, Kliver S, Tinker MT, Murray MJ, Johnson J, Lindblad-Toh K, Karlsson EK, Lohmueller KE, Wayne RK. 2019. Aquatic adaptation and depleted diversity: a deep dive into the genomes of the sea otter and giant otter. *Molecular Biology and Evolution* **36**(12):2631–2655 DOI [10.1093/molbev/msz101](https://doi.org/10.1093/molbev/msz101).
- Berthelot J-M, Lioté F. 2020. COVID-19 as a STING disorder with delayed over-secretion of interferon-beta. *EBioMedicine* **56**:102801 DOI [10.1016/j.ebiom.2020.102801](https://doi.org/10.1016/j.ebiom.2020.102801).
- Bitter MC, Kapsenberg L, Gattuso J-P, Pfister CA. 2019. Standing genetic variation fuels rapid adaptation to ocean acidification. *Nature Communications* **10**(1):1–10 DOI [10.1038/s41467-019-13767-1](https://doi.org/10.1038/s41467-019-13767-1).
- Bockman DE. 1997. Development of the thymus. *Microscopy Research and Technique* **38**:209–215 DOI [10.1002/\(ISSN\)1097-0029](https://doi.org/10.1002/(ISSN)1097-0029).
- Boero LE, Castagna VC, Terreros G, Moglie MJ, Silva S, Maass JC, Fuchs PA, Delano PH, Elgoyhen AB, Gómez-Casati ME. 2020. Preventing presbycusis in mice with enhanced medial olivocochlear feedback. *Proceedings of the National Academy of Sciences USA* **117**(21):11811–11819 DOI [10.1073/pnas.2000760117](https://doi.org/10.1073/pnas.2000760117).
- Brindley DN, Waggoner DW. 1998. Mammalian lipid phosphate phosphohydrolases. *Journal of Biological Chemistry* **273**(38):24281–24284 DOI [10.1074/jbc.273.38.24281](https://doi.org/10.1074/jbc.273.38.24281).
- Brown GR, Hem V, Katz KS, Ovetsky M, Wallin C, Ermolaeva O, Tolstoy I, Tatusova T, Pruitt KD, Maglott DR, Murphy TD. 2015. Gene: a gene-centered information resource at NCBI. *Nucleic Acids Research* **43**(D1):D36–D42 DOI [10.1093/nar/gku1055](https://doi.org/10.1093/nar/gku1055).
- Buchfink B, Xie C, Huson DH. 2015. Fast and sensitive protein alignment using DIAMOND. *Nature Methods* **12**(1):59–60 DOI [10.1038/nmeth.3176](https://doi.org/10.1038/nmeth.3176).
- Cantarel BL, Korf I, Robb SM, Parra G, Ross E, Moore B, Holt C, Alvarado AS, Yandell M. 2008. MAKER: an easy-to-use annotation pipeline designed for emerging model organism genomes. *Genome Research* **18**:188–196 DOI [10.1101/gr.6743907](https://doi.org/10.1101/gr.6743907).
- Casola C, Ganote CL, Hahn MW. 2010. Nonallelic gene conversion in the genus *Drosophila*. *Genetics* **185**(1):95–103 DOI [10.1534/genetics.110.115444](https://doi.org/10.1534/genetics.110.115444).

- Chazarin B, Storey KB, Ziemianin A, Chanon S, Plumel M, Chery I, Durand C, Evans AL, Arnemo JM, Zedrosser A, Swenson JE, Gauquelin-Koch G, Simon C, Blanc S, Lefai E, Bertile F. 2019. Metabolic reprogramming involving glycolysis in the hibernating brown bear skeletal muscle. *Frontiers in Zoology* 16:12 DOI 10.1186/s12983-019-0312-2.
- Chen Y, Wang L, Jin J, Luan Y, Chen C, Li Y, Chu H, Wang X, Liao G, Yu Y, Teng H, Wang Y, Pan W, Fang L, Liao L, Jiang Z, Ge X, Li B, Wang P. 2017. p38 inhibition provides anti-DNA virus immunity by regulation of USP21 phosphorylation and STING activation. *Journal of Experimental Medicine* 214:991–1010 DOI 10.1084/jem.20161387.
- Chen Y, Yu D, Wang L, Du S. 2019. Identification of E2F8 as a transcriptional regulator of gluconeogenesis in primary mouse hepatocytes. *Biochemistry* 84(12–13):1529–1536 DOI 10.1134/S0006297919120125.
- Chang JM, Di Tommaso P, Notredame C. 2014. TCS: a new multiple sequence alignment reliability measure to estimate alignment accuracy and improve phylogenetic tree reconstruction. *Molecular Biology and Evolution* 31(6):1625–1637 DOI 10.1093/molbev/msu117.
- Chiarella E, Aloisio A, Codispoti B, Nappo G, Scicchitano S, Lucchino V, Montalcini Y, Camarotti A, Galasso O, Greco M, Gasparini G, Mesuraca M, Bond HM, Morrone G. 2018. ZNF521 has an inhibitory effect on the adipogenic differentiation of human adipose-derived mesenchymal stem cells. *Stem Cell Reviews and Reports* 14:901–914 DOI 10.1007/s12015-018-9830-0.
- Clayton E, Munir M. 2020. Fundamental characteristics of bat interferon systems. *Frontiers in Cellular and Infection Microbiology* 10:762 DOI 10.3389/fcimb.2020.527921.
- Corcoran AJ, Moss CF. 2017. Sensing in a noisy world: lessons from auditory specialists, echolocating bats. *Journal of Experimental Biology* 220(24):4554–4566 DOI 10.1242/jeb.163063.
- Corcoran AJ, Weller TJ. 2018. Inconspicuous echolocation in hoary bats (*Lasiurus cinereus*). *Proceedings of the Royal Society B: Biological Sciences* 285(1878):20180441 DOI 10.1098/rspb.2018.0441.
- Cornman RS, Cryan PM. 2022. *Gene annotations for the hoary bat (Lasiurus cinereus) and alignments with other bat gene sets for evolutionary analysis*. Reston: U.S. Geological Survey.
- Cornman RS, Fike JA, Oyler-McCance SJ, Cryan PM. 2021. Historical effective population size of North American hoary bat (*Lasiurus cinereus*) and challenges to estimating trends in contemporary effective breeding population size from archived samples. *PeerJ* 9(2):e11285 DOI 10.7717/peerj.11285.
- Cornman RS, Jennifer FM, Oyler-McCance SJ, Cryan PM. 2020. *Genetic variation in hoary bats (Lasiurus cinereus) assessed from archived samples*. Reston: U.S. Geological Survey.
- Corradi N, Pombert J-F, Farinelli L, Didier ES, Keeling PJ. 2010. The complete sequence of the smallest known nuclear genome from the microsporidian *Encephalitozoon intestinalis*. *Nature Communications* 1(1):1–7 DOI 10.1038/ncomms1082.
- Cryan PM, Stricker CA, Wunder MB. 2014. Continental-scale, seasonal movements of a heterothermic migratory tree bat. *Ecological Applications* 24(4):602–616 DOI 10.1890/13-0752.1.
- Cryan PM, Gorresen PM, Hein CD, Schirmacher M, Diehl R, Huso M, Hayman DTS, Fricker P, Bonaccorso F, Johnston DH, Heist K, Dalton D. 2014. Behaviors of bats at wind turbines. *Proceedings of the National Academy of Sciences USA* 111:15126–15131.
- Cryan PM, Wolf BO. 2003. Sex differences in the thermoregulation and evaporative water loss of a heterothermic bat, *Lasiurus cinereus*, during its spring migration. *Journal of Experimental Biology* 206(19):3381–3390 DOI 10.1242/jeb.00574.

- D'Amico D, Mottis A, Potenza F, Sorrentino V, Li H, Romani M, Lemos V, Schoonjans K, Zamboni N, Knott G, Schneider BL, Auwerx J. 2019. The RNA-binding protein PUM2 impairs mitochondrial dynamics and mitophagy during aging. *Molecular Cell* 73(4):775–787 DOI 10.1016/j.molcel.2018.11.034.
- Dantas AG, Santoro ML, Nunes N, de Mello CB, Pimenta LSE, Meloni VA, Soares DCQ, Belangero SN, Carvalheira G, Kim CA, Melaragno MI. 2019. Downregulation of genes outside the deleted region in individuals with 22q11.2 deletion syndrome. *Human Genetics* 138(1):93–103 DOI 10.1007/s00439-018-01967-6.
- Davies KT, Yohe LR, Almonte J, Sánchez MK, Rengifo EM, Dumont ER, Sears KE, Dávalos LM, Rossiter SJ. 2020. Foraging shifts and visual preadaptation in ecologically diverse bats. *Molecular Ecology* 29(10):1839–1859 DOI 10.1111/mec.15445.
- Davis WH. 1970. Hibernation: ecology and physiological ecology. *Biology of Bats* 1:265–300 DOI 10.1016/B978-0-12-758001-2.50013-7.
- Di Franco A, Poujol R, Baurain D, Philippe H. 2019. Evaluating the usefulness of alignment filtering methods to reduce the impact of errors on evolutionary inferences. *BMC Evolutionary Biology* 19(1):21 DOI 10.1186/s12862-019-1350-2.
- Dimitrova DS, Kaishev VK, Tan S. 2020. Computing the Kolmogorov–Smirnov distribution when the underlying CDF is purely discrete, mixed, or continuous. *Journal of Statistical Software* 95(10):1–42 DOI 10.18637/jss.v095.i10.
- Dobzhansky T, Pavlovsky O. 1958. Interracial hybridization and breakdown of coadapted gene complexes in *Drosophila paulistorum* and *Drosophila willistoni*. *Proceedings of the National Academy of Sciences of the United States of America* 44(6):622–629 DOI 10.1073/pnas.44.6.622.
- Dong D, Lei M, Hua P, Pan Y-H, Mu S, Zheng G, Pang E, Lin K, Zhang S. 2016. The genomes of two bat species with long constant frequency echolocation calls. *Molecular Biology and Evolution* msw231(1):20–34 DOI 10.1093/molbev/msw231.
- Du Q, de la Morena MT, van Oers NS. 2020. The genetics and epigenetics of 22q11.2 deletion syndrome. *Frontiers in Genetics* 10:1365 DOI 10.3389/fgene.2019.01365.
- Ehrhardt N, Bedoya C, Péterfy M. 2014. Embryonic viability, lipase deficiency, hypertriglyceridemia and neonatal lethality in a novel LMF1-deficient mouse model. *Nutrition & Metabolism* 11(1):37 DOI 10.1186/1743-7075-11-37.
- Emms DM, Kelly S. 2019. OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biology* 20(1):1–14 DOI 10.1186/s13059-019-1832-y.
- Footz TK, Brinkman-Mills P, Banting GS, Maier SA, Riazi MA, Bridgland L, Hu S, Birren B, Minoshima S, Shimizu N, Pan HQ, Nguyen T, Fang F, Fu Y, Ray L, Wu H, Shaull S, Phan S, Yao Z, Chen F, Huan A, Hu P, Wang Q, Loh P, Qi S, Roe BA, McDermid HE. 2001. Analysis of the cat eye syndrome critical region in humans and the region of conserved synteny in mice: a search for candidate genes at or near the human chromosome 22 pericentromere. *Genome Research* 11:1053–1070 DOI 10.1101/gr.154901.
- Frick W, Baerwald EF, Pollock J, Barclay R, Szymanski J, Weller T, Russell A, Loeb S, Medellin R, McGuire L. 2017. Fatalities at wind turbines may threaten population viability of a migratory bat. *Biological Conservation* 209:172–177 DOI 10.1016/j.biocon.2017.02.023.
- Friederich MW, Erdogan AJ, Coughlin CR, Elos MT, Jiang H, O'Rourke CP, Lovell MA, Wartchow E, Gowan K, Chatfield KC, Chick WS, Spector EB, Van Hove JLK, Riemer J. 2017. Mutations in the accessory subunit NDUFB10 result in isolated complex I deficiency and illustrate the critical role of intermembrane space import for complex I holoenzyme assembly. *Human Molecular Genetics* 26:702–716 DOI 10.1093/hmg/ddw431.

- Gässler A, Quiclet C, Kluth O, Gottmann P, Schwerbel K, Helms A, Stadion M, Wilhelmi I, Jonas W, Ouni M, Mayer F, Spranger J, Schürmann A, Vogel H. 2020. Overexpression of Gjb4 impairs cell proliferation and insulin secretion in primary islet cells. *Molecular Metabolism* 41:101042 DOI 10.1016/j.molmet.2020.101042.
- Gayral P, Caminade P, Boursot P, Galtier N. 2007. The evolutionary fate of recently duplicated retrogenes in mice. *Journal of Evolutionary Biology* 20(2):617–626 DOI 10.1111/j.1420-9101.2006.01245.x.
- Geiser F. 2013. Hibernation. *Current Biology* 23(5):R188–R193 DOI 10.1016/j.cub.2013.01.062.
- Härtlova A, Erttmann SF, Raffi FA, Schmalz AM, Resch U, Anugula S, Lienenklaus S, Nilsson LM, Kröger A, Nilsson JA, Ek T, Weiss S, Gekara NO. 2015. DNA damage primes the type I interferon system via the cytosolic DNA sensor STING to promote anti-microbial innate immunity. *Immunity* 42:332–343 DOI 10.1016/j.immuni.2015.01.012.
- Hawkins JA, Kaczmarek ME, Müller MA, Drost C, Press WH, Sawyer SL. 2019. A metaanalysis of bat phylogenetics and positive selection based on genomes and transcriptomes from 18 species. *Proceedings of the National Academy of Sciences USA* 116:11351–11360 DOI 10.1073/pnas.1814995116.
- Hein C, Hale A, Straw B. 2021. *Technical report: acoustic and genetic approaches for informing population status and trends of migratory tree bats*. Oak Ridge: Department of Energy, Office of Scientific and Technical Information.
- Henquin J-C, Nenquin M. 2014. Activators of PKA and Epac distinctly influence insulin secretion and cytosolic Ca²⁺ in female mouse islets stimulated by glucose and tolbutamide. *Endocrinology* 155(9):3274–3287 DOI 10.1210/en.2014-1247.
- Henriques BJ, Olsen RKJ, Gomes CM, Bross P. 2021. Electron transfer flavoprotein and its role in mitochondrial energy metabolism in health and disease. *Gene* 776(Pt. 3):145407 DOI 10.1016/j.gene.2021.145407.
- Hittinger CT, Gonçalves P, Sampaio JP, Dover J, Johnston M, Rokas A. 2010. Remarkably ancient balanced polymorphisms in a multi-locus gene network. *Nature* 464(7285):54–58 DOI 10.1038/nature08791.
- Hogan PG, Lewis RS, Rao A. 2009. Molecular basis of calcium signaling in lymphocytes: STIM and ORAI. *Annual Review of Immunology* 28(1):491–533 DOI 10.1146/annurev.immunol.021908.132550.
- Holder JL Jr, Zhang L, Kublaoui BM, DiLeone RJ, Oz OK, Bair CH, Lee Y-H, Zinn AR. 2004. Sim1 gene dosage modulates the homeostatic feeding response to increased dietary fat in mice. *American Journal of Physiology-Endocrinology and Metabolism* 287:E105–E113 DOI 10.1152/ajpendo.00446.2003.
- Holm CK, Jensen SB, Jakobsen MR, Cheshenko N, Horan KA, Moeller HB, Gonzalez-Dosal R, Rasmussen SB, Christensen MH, Yarovinsky TO, Rixon FJ, Herold BC, Fitzgerald KA, Paludan Søren R. 2012. Virus-cell fusion as a trigger of innate immunity dependent on the adaptor STING. *Nature Immunology* 13:737–743 DOI 10.1038/ni.2350.
- Howard MJ, Lim WH, Fierke CA, Koutmos M. 2012. Mitochondrial ribonuclease P structure provides insight into the evolution of catalytic strategies for precursor-tRNA 5' processing. *Proceedings of the National Academy of Sciences USA* 109(40):16149–16154 DOI 10.1073/pnas.1209062109.
- Hoyt JR, Kilpatrick AM, Langwig KE. 2021. Ecology and impacts of white-nose syndrome on bats. *Nature Reviews Microbiology* 19(3):196–210 DOI 10.1038/s41579-020-00493-5.
- Huang DW, Sherman BT, Tan Q, Kir J, Liu D, Bryant D, Guo Y, Stephens R, Baseler MW, Lane HC, Lempicki RA. 2007. DAVID bioinformatics resources: expanded annotation database

and novel algorithms to better extract biology from large gene lists. *Nucleic Acids Research* 35(Suppl. 2):W169–W175 DOI 10.1093/nar/gkm415.

- Huber C, Dias-Santagata D, Glaser A, O'Sullivan J, Brauner R, Wu K, Xu X, Pearce K, Wang R, Uzielli MLG, Dagonneau N, Chemaitilly W, Superti-Furga A, Santos HD, Mégarbané A, Morin G, Gillessen-Kaesbach G, Hennekam R, der Burgt IV, Black GCM, Clayton PE, Read A, Merrer ML, Scambler PJ, Munnich A, Pan ZQ, Winter R, Cormier-Daire V. 2005. Identification of mutations in CUL7 in 3-M syndrome. *Nature Genetics* 37:1119–1124 DOI 10.1038/ng1628.
- Hughes GM, Boston ES, Finarelli JA, Murphy WJ, Higgins DG, Teeling EC. 2018. The birth and death of olfactory receptor gene families in mammalian niche adaptation. *Molecular Biology and Evolution* 35(6):1390–1406 DOI 10.1093/molbev/msy028.
- Ishii M, Fei H, Friedman JM. 2003. Targeted disruption of GPR7, the endogenous receptor for neuropeptides B and W, leads to metabolic defects and adult-onset obesity. *Proceedings of the National Academy of Sciences USA* 100(18):10540–10545 DOI 10.1073/pnas.1334189100.
- Ishikawa H, Ma Z, Barber GN. 2009. STING regulates intracellular DNA-mediated, type I interferon-dependent innate immunity. *Nature* 461(7265):788–792 DOI 10.1038/nature08476.
- Jebb D, Huang Z, Pippel M, Hughes GM, Lavrichenk 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 Sébastien J, 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 DOI 10.1038/s41586-020-2486-3.
- Jiang L, Su H, Wu X, Shen H, Kim M-H, Li Y, Myers MG, Owyang C, Rui L. 2020. Leptin receptor-expressing neuron Sh2b1 supports sympathetic nervous system and protects against obesity and metabolic disease. *Nature Communications* 11(1):1–13 DOI 10.1038/s41467-020-15328-3.
- Jiang Q, Sun B, Liu Q, Cai M, Wu R, Wang F, Yao Y, Wang Y, Wang X. 2019. MTCH2 promotes adipogenesis in intramuscular preadipocytes via an m6A-YTHDF1-dependent mechanism. *The FASEB Journal* 33(2):2971–2981 DOI 10.1096/fj.201801393RRR.
- Jorgenson E, Matharu N, Palmer MR, Yin J, Shan J, Hoffmann TJ, Thai KK, Zhou X, Hotaling JM, Jarvik GP, Ahituv N, Wessells H, Van Den Eeden SK. 2018. Genetic variation in the SIM1 locus is associated with erectile dysfunction. *Proceedings of the National Academy of Sciences USA* 115:11018–11023 DOI 10.1073/pnas.1809872115.
- Jurka J, Kapitonov VV, Pavlicek A, Klonowski P, Kohany O, Walichiewicz J. 2005. Repbase update, a database of eukaryotic repetitive elements. *Cytogenetic and Genome Research* 110(1–4):462–467 DOI 10.1159/000084979.
- Kataoka S, Kawashima N, Okuno Y, Muramatsu H, Miwata S, Narita K, Hamada M, Murakami N, Taniguchi R, Ichikawa D, Kitazawa H, Suzuki K, Nishikawa E, Narita A, Nishio N, Yamamoto H, Fukasawa Y, Kato T, Yamamoto H, Natsume J, Kojima S, Nishino I, Taketani T, Ohnishi H, Takahashi Y. 2021. Successful treatment of a novel type I interferonopathy due to a de novo PSMB9 gene mutation with a Janus kinase inhibitor. *Journal of Allergy and Clinical Immunology* 148:639–644 DOI 10.1016/j.jaci.2021.03.010.
- Katoh K, Misawa K, Kuma K, Miyata T. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research* 30(14):3059–3066 DOI 10.1093/nar/gkf436.

- Koop D, Chen J, Theodosiou M, Carvalho JE, Alvarez S, De Lera AR, Holland LZ, Schubert M. 2014. Roles of retinoic acid and Tbx1/10 in pharyngeal segmentation: amphioxus and the ancestral chordate condition. *EvoDevo* 5(1):1–16 DOI 10.1186/2041-9139-5-36.
- Korstian JM, Hale AM, Williams DA. 2015. Genetic diversity, historic population size, and population structure in 2 North American tree bats. *Journal of Mammalogy* 96(5):972–980 DOI 10.1093/jmammal/gyv101.
- Kosiol C, Vinař T, da Fonseca RR, Hubisz MJ, Bustamante CD, Nielsen R, Siepel A. 2008. Patterns of positive selection in six mammalian genomes. *PLoS Genetics* 4(8):e1000144 DOI 10.1371/journal.pgen.1000144.
- Koyasu S. 2003. The role of PI3K in immune cells. *Nature Immunology* 4(4):313–319 DOI 10.1038/ni0403-313.
- Kramerova I, Kudryashova E, Wu B, Germain S, Vandeborne K, Romain N, Haller RG, Verity MA, Spencer MJ. 2009. Mitochondrial abnormalities, energy deficit and oxidative stress are features of calpain 3 deficiency in skeletal muscle. *Human Molecular Genetics* 18(17):3194–3205 DOI 10.1093/hmg/ddp257.
- Kriventseva EV, Kuznetsov D, Tegenfeldt F, Manni M, Dias R, Simão FA, Zdobnov EM. 2019. OrthoDB v10: sampling the diversity of animal, plant, fungal, protist, bacterial and viral genomes for evolutionary and functional annotations of orthologs. *Nucleic Acids Research* 47(D1):D807–D811 DOI 10.1093/nar/gky1053.
- Krogh A, Larsson B, Von Heijne G, Sonnhammer EL. 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *Journal of Molecular Biology* 305(3):567–580 DOI 10.1006/jmbi.2000.4315.
- Kulyté A, Rydén M, Mejhert N, Dungner E, Sjölin E, Arner P, Dahlman I. 2011. MTCH2 in human white adipose tissue and obesity. *The Journal of Clinical Endocrinology & Metabolism* 96(10):E1661–E1665 DOI 10.1210/jc.2010-3050.
- Kumar S, Stecher G, Li M, Niyaz C, Tamura K. 2018. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution* 35(6):1547–1549 DOI 10.1093/molbev/msy096.
- Kunath A, Hesselbarth N, Gericke M, Kern M, Dommel S, Kovacs P, Stumvoll M, Blüher M, Klötting N. 2016. Repin1 deficiency improves insulin sensitivity and glucose metabolism in db/db mice by reducing adipose tissue mass and inflammation. *Biochemical and Biophysical Research Communications* 478(1):398–402 DOI 10.1016/j.bbrc.2016.07.038.
- Kurien P, Hsu P-K, Leon J, Wu D, McMahon T, Shi G, Xu Y, Lipzen A, Pennacchio LA, Jones CR, Fu Y-H, Ptáček LJ. 2019. TIMELESS mutation alters phase responsiveness and causes advanced sleep phase. *Proceedings of The National Academy of Sciences of The United States of America* 116:12045–12053 DOI 10.1073/pnas.1819110116.
- Lee PL, Kohler JJ, Pfeffer SR. 2009. Association of β -1, 3-N-acetylglucosaminyltransferase 1 and β -1, 4-galactosyltransferase 1, trans-Golgi enzymes involved in coupled poly-N-acetyllactosamine synthesis. *Glycobiology* 19(6):655–664 DOI 10.1093/glycob/cwp035.
- Li X, Tu L, Murphy PG, Kadono T, Steeber DA, Tedder TF. 2001. CHST1 and CHST2 sulfotransferase expression by vascular endothelial cells regulates shear-resistant leukocyte rolling via l-selectin. *Journal of Leukocyte Biology* 69:565–574 DOI 10.1189/jlb.69.4.565.
- Liao J, Kochilas L, Nowotschin S, Arnold JS, Aggarwal VS, Epstein JA, Brown MC, Adams J, Morrow BE. 2004. Full spectrum of malformations in velo-cardio-facial syndrome/DiGeorge syndrome mouse models by altering Tbx1 dosage. *Human Molecular Genetics* 13(15):1577–1585 DOI 10.1093/hmg/ddh176.

- Lindblad-Toh K, Garber M, Zuk O, Lin MF, Parker BJ, Washietl S, Kheradpour P, Ernst J, Jordan G, Mauceli E, Ward LD, Lowe CB, Holloway AK, Clamp M, Gnerre S, Alföldi J, Beal K, Chang J, Clawson H, Cuff J, Di Palma F, Fitzgerald S, Flicek P, Guttman M, Hubisz MJ, Jaffe DB, Jungreis I, Kent WJ, Kostka D, Lara M, Martins AL, Massingham T, Moltke I, Raney BJ, Rasmussen MD, Robinson J, Stark A, Vilella AJ, Wen J, Xie X, Zody MC, Worley KC, Kovar CL, Muzny DM, Gibbs RA, Warren WC, Mardis ER, Weinstock GM, Wilson RK, Birney E, Margulies EH, Herrero J, Green ED, Haussler D, Siepel A, Goldman N, Pollard KS, Pedersen JS, Lander ES, Kellis M. 2011. A high-resolution map of human evolutionary constraint using 29 mammals. *Nature* 478:476–482 DOI 10.1038/nature10530.
- Lindsay EA, Vitelli F, Su H, Morishima M, Huynh T, Pramparo T, Jurecic V, Ogunrinu G, Sutherland HF, Scambler PJ, Bradley A, Baldini A. 2001. Tbx1 haploinsufficiency in the DiGeorge syndrome region causes aortic arch defects in mice. *Nature* 410(6824):97–101 DOI 10.1038/35065105.
- Lipovsek M, Fierro A, Pérez EG, Boffi JC, Millar NS, Fuchs PA, Katz E, Elgoyhen AB. 2014. Tracking the molecular evolution of calcium permeability in a nicotinic acetylcholine receptor. *Molecular Biology and Evolution* 31(12):3250–3265 DOI 10.1093/molbev/msu258.
- Liu H, Chen L, Giffen KP, Stringham ST, Li Y, Judge PD, Beisel KW, He DZ. 2018. Cell-specific transcriptome analysis shows that adult pillar and Deiters' cells express genes encoding machinery for specializations of cochlear hair cells. *Frontiers in Molecular Neuroscience* 11:356 DOI 10.3389/fnmol.2018.00356.
- Liu G, Hu Y, Han Z, Jin S, Jiang Q. 2019a. Genetic variant rs17185536 regulates SIM1 gene expression in human brain hypothalamus. *Proceedings of the National Academy of Sciences USA* 116(9):3347–3348 DOI 10.1073/pnas.1821550116.
- Liu D, Wu H, Wang C, Li Y, Tian H, Siraj S, Sehgal SA, Wang X, Wang J, Shang Y, Jiang Z, Liu L, Chen Q. 2019b. STING directly activates autophagy to tune the innate immune response. *Cell Death & Differentiation* 26(9):1735–1749 DOI 10.1038/s41418-018-0251-z.
- Luo Y, Liu Y, Wu L, Ma X, Liu Q, Huang F, Zhang X, Zhang Y, Zhang J, Luo H, others. 2019. CUL7 E3 ubiquitin ligase mediates the degradation of activation-induced cytidine deaminase and regulates the Ig class switch recombination in B lymphocytes. *The Journal of Immunology* 203:269–281 DOI 10.4049/jimmunol.1900125.
- Ma C, Wang W, Wang Y, Sun Y, Kang L, Zhang Q, Jiang Y. 2020. TMT-labeled quantitative proteomic analyses on the longissimus dorsi to identify the proteins underlying intramuscular fat content in pigs. *Journal of Proteomics* 213(2):103630 DOI 10.1016/j.jprot.2019.103630.
- Małeckı JM, Willemen HL, Pinto R, Ho AY, Moen A, Eijkelkamp N, Falnes PØ. 2019. Human FAM173A is a mitochondrial lysine-specific methyltransferase that targets adenine nucleotide translocase and affects mitochondrial respiration. *Journal of Biological Chemistry* 294(31):11654–11664 DOI 10.1074/jbc.RA119.009045.
- Manceau M, Domingues VS, Linnen CR, Rosenblum EB, Hoekstra HE. 2010. Convergence in pigmentation at multiple levels: mutations, genes and function. *Philosophical Transactions of the Royal Society B: Biological Sciences* 365(1552):2439–2450 DOI 10.1098/rstb.2010.0104.
- Mao HZ, Ehrhardt N, Bedoya C, Gomez JA, DeZwaan-McCabe D, Mungrue IN, Kaufman RJ, Rutkowski DT, Péterfy M. 2014. Lipase maturation factor 1 (lmf1) is induced by endoplasmic reticulum stress through activating transcription factor 6 α (Atf6 α) signaling. *Journal of Biological Chemistry* 289(35):24417–24427 DOI 10.1074/jbc.M114.588764.
- Marçais G, Delcher AL, Phillippy AM, Coston R, Salzberg SL, Zimin A. 2018. MUMmer4: a fast and versatile genome alignment system. *PLOS Computational Biology* 14(1):e1005944 DOI 10.1371/journal.pcbi.1005944.

- Marcovich I, Moglie MJ, Carpaneto Freixas AE, Trigila AP, Franchini LF, Plazas PV, Lipovsek M, Elgoyhen AB. 2020. Distinct evolutionary trajectories of neuronal and hair cell nicotinic acetylcholine receptors. *Molecular Biology and Evolution* 37(4):1070–1089 DOI 10.1093/molbev/msz290.
- McGuire LP, Fenton MB, Guglielmo CG. 2013. Seasonal upregulation of catabolic enzymes and fatty acid transporters in the flight muscle of migrating hoary bats, *Lasiurus cinereus*. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology* 165(2):138–143 DOI 10.1016/j.cbpb.2013.03.013.
- McGuire LP, Jonasson KA, Guglielmo CG. 2014. Bats on a budget: torpor-assisted migration saves time and energy. *PLOS ONE* 9(12):e115724 DOI 10.1371/journal.pone.0115724.
- Melroy-Greif WE, Gizer IR, Wilhelmsen KC, Ehlers CL. 2017. Genetic influences on evening preference overlap with those for bipolar disorder in a sample of Mexican Americans and American Indians. *Twin Research and Human Genetics* 20(6):499–510 DOI 10.1017/thg.2017.62.
- Merscher S, Funke B, Epstein JA, Heyer J, Puech A, Lu MM, Xavier RJ, Demay MB, Russell RG, Factor S, Tokooya K, Jore BS, Lopez M, Pandita RK, Lia M, Carrion D, Xu H, Schorle H, Kobler JB, Scambler P, Wynshaw-Boris A, Skoultschi AI, Morrow BE, Kucherlapati R. 2001. TBX1 is responsible for cardiovascular defects in velo-cardio-facial/DiGeorge syndrome. *Cell* 104:619–629 DOI 10.1016/S0092-8674(01)00247-1.
- Meyer B, Wurm JP, Sharma S, Immer C, Pogoryelov D, Kötter P, Lafontaine DL, Wöhnert J, Entian K-D. 2016. Ribosome biogenesis factor Tsr3 is the aminocarboxypropyl transferase responsible for 18S rRNA hypermodification in yeast and humans. *Nucleic Acids Research* 44(9):4304–4316 DOI 10.1093/nar/gkw244.
- Michalak P. 2008. Coexpression, coregulation, and cofunctionality of neighboring genes in eukaryotic genomes. *Genomics* 91(3):243–248 DOI 10.1016/j.ygeno.2007.11.002.
- Michaud JL, Rosenquist T, May NR, Fan C-M. 1998. Development of neuroendocrine lineages requires the bHLH-PAS transcription factor SIM1. *Genes & Development* 12(20):3264–3275 DOI 10.1101/gad.12.20.3264.
- Millar N. 2003. Assembly and subunit diversity of nicotinic acetylcholine receptors. *Biochemical Society Transactions* 31(4):869–874 DOI 10.1042/bst0310869.
- Mohammadi SA, Burton TJ, Christie MJ. 2017. α 9-nAChR knockout mice exhibit dysregulation of stress responses, affect and reward-related behaviour. *Behavioural Brain Research* 328(10):105–114 DOI 10.1016/j.bbr.2017.04.005.
- Nagata-Kuroiwa R, Furutani N, Hara J, Hondo M, Ishii M, Abe T, Mieda M, Tsujino N, Motoike T, Yanagawa Y, Kuwaki T, Yamamoto M, Yanagisawa M, Sakurai T, Zhuang X. 2011. Critical role of neuropeptides B/W receptor 1 signaling in social behavior and fear memory. *PLOS ONE* 6:e16972 DOI 10.1371/journal.pone.0016972.
- Nawrocki AR, Rodriguez CG, Toolan DM, Price O, Henry M, Forrest G, Szeto D, Keohane CA, Pan Y, Smith KM, Raheem IT, Cox CD, Hwa J, Renger JJ, Smith SM. 2014. Genetic deletion and pharmacological inhibition of phosphodiesterase 10A protects mice from diet-induced obesity and insulin resistance. *Diabetes* 63:300–311 DOI 10.2337/db13-0247.
- Neafsey DE, Waterhouse RM, Abai MR, Aganezov SS, Alekseyev MA, Allen JE, Amon J, Arcà B, Arensbarger P, Artemov G, Assour LA, Basseri H, Berlin A, Birren BW, Blandin SA, Brockman AI, Burkot TR, Burt A, Chan CS, Chauve C, Chiu JC, Christensen M, Costantini C, Davidson VLM, Deligianni E, Dottorini T, Dritsou V, Gabriel SB, Guelbeogo WM, Hall AB, Han MV, Hlaing T, Hughes DST, Jenkins AM, Jiang X, Jungreis I, Kakani EG, Kamali M, Kempainen P, Kennedy RC, Kirmitzoglou IK, Koekemoer LL,

- Laban N, Langridge N, Lawniczak MKN, Lirakis M, Lobo NF, Lowy E, MacCallum RM, Mao C, Maslen G, Mbogo C, McCarthy J, Michel K, Mitchell SN, Moore W, Murphy KA, Naumenko AN, Nolan T, Novoa EM, O'Loughlin S, Oringanje C, Oshaghi MA, Pakpour N, Papathanos PA, Peery AN, Povelones M, Prakash A, Price DP, Rajaraman A, Reimer LJ, Rinker DC, Rokas A, Russell TL, Sagnon N'Fale, Sharakhova MV, Shea T, Simão FA, Simard F, Slotman MA, Somboon P, Stegny V, Struchiner CJ, Thomas GWC, Tojo M, Topalis P, Tubio Jé MC, Unger MF, Vontas J, Walton C, Wilding CS, Willis JH, Wu Y-C, Yan G, Zdobnov EM, Zhou X, Catteruccia F, Christophides GK, Collins FH, Cornman RS, Crisanti A, Donnelly MJ, Emrich SJ, Fontaine MC, Gelbart W, Hahn MW, Hansen IA, Howell PI, Kafatos FC, Kellis M, Lawson D, Louis C, Luckhart S, Muskavitch MAT, Ribeiro JM, Riehle MA, Sharakhov IV, Tu Z, Zwiebel LJ, Besansky NJ. 2015. Highly evolvable malaria vectors: the genomes of 16 Anopheles mosquitoes. *Science* 347:1258524 DOI 10.1126/science.1258522.
- Notredame C, Higgins DG, Heringa J. 2000. T-Coffee: A novel method for fast and accurate multiple sequence alignment. *Journal of Molecular Biology* 302(1):205–217 DOI 10.1006/jmbi.2000.4042.
- Oliver TA, Garfield DA, Manier MK, Haygood R, Wray GA, Palumbi SR. 2010. Whole-genome positive selection and habitat-driven evolution in a shallow and a deep-sea urchin. *Genome Biology and Evolution* 2:800–814 DOI 10.1093/gbe/evq063.
- Onyilagha C, Jia P, Jayachandran N, Hou S, Okwor I, Kuriakose S, Marshall A, Uzonna JE. 2015. The B cell adaptor molecule Bam32 is critically important for optimal antibody response and resistance to Trypanosoma congolense infection in mice. *PLOS Neglected Tropical Diseases* 9(4):e0003716 DOI 10.1371/journal.pntd.0003716.
- O'Shea TJ, Cryan PM, Cunningham AA, Fooks AR, Hayman DT, Luis AD, Peel AJ, Plowright RK, Wood JL. 2014. Bat flight and zoonotic viruses. *Emerging Infectious Diseases* 20(5):741–745 DOI 10.3201/eid2005.130539.
- Palmer DB. 2013. The effect of age on thymic function. *Frontiers in Immunology* 4:316 DOI 10.3389/fimmu.2013.00316.
- Paradis E, Claude J, Strimmer K. 2004. APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* 20(2):289–290 DOI 10.1093/bioinformatics/btg412.
- Perez SI, Tejedor MF, Novo NM, Aristide L. 2013. Divergence times and the evolutionary radiation of New World monkeys (Platyrrhini, Primates): an analysis of fossil and molecular data. *PLOS ONE* 8(6):e68029 DOI 10.1371/journal.pone.0068029.
- Piccart E, Gantois I, Laeremans A, de Hoogt R, Meert T, Vanhoof G, Arckens L, D'Hooge R. 2011. Impaired appetitively as well as aversively motivated behaviors and learning in PDE10A-deficient mice suggest a role for striatal signaling in evaluative salience attribution. *Neurobiology of Learning and Memory* 95(3):260–269 DOI 10.1016/j.nlm.2010.11.018.
- Pinzari CA, Kang L, Michalak P, Jermiin LS, Price DK, Bonaccorso FJ. 2020. Analysis of genomic sequence data reveals the origin and evolutionary separation of Hawaiian hoary bat populations. *Genome Biology and Evolution* 12(9):1504–1514 DOI 10.1093/gbe/evaa137.
- Poppi LA, Tabatabaee H, Drury HR, Jobling P, Callister RJ, Migliaccio AA, Jordan PM, Holt JC, Rabbitt RD, Lim R, Brichta AM. 2018. ACh-induced hyperpolarization and decreased resistance in mammalian type II vestibular hair cells. *Journal of Neurophysiology* 119(1):312–325 DOI 10.1152/jn.00030.2017.
- Puech A, Saint-Jore B, Funke B, Gilbert DJ, Sirotkin H, Copeland NG, Jenkins NA, Kucherlapati R, Morrow B, Skoultschi AI. 1997. Comparative mapping of the human 22q11 chromosomal region and the orthologous region in mice reveals complex changes in gene

- organization. *Proceedings of the National Academy of Sciences USA* **94**(26):14608–14613 DOI [10.1073/pnas.94.26.14608](https://doi.org/10.1073/pnas.94.26.14608).
- Pylant CL, Nelson DM, Fitzpatrick MC, Gates JE, Keller SR. 2016.** Geographic origins and population genetics of bats killed at wind-energy facilities. *Ecological Applications* **26**(5):1381–1395 DOI [10.1890/15-0541](https://doi.org/10.1890/15-0541).
- R Core Team. 2019.** *R: a language and environment for statistical computing*. Vienna: The R Foundation for Statistical Computing. Available at <https://www.R-project.org/>.
- Raft S, Nowotschin S, Liao J, Morrow BE. 2004.** Suppression of neural fate and control of inner ear morphogenesis by Tbx1. *Development* **131**(8):1801–1812 DOI [10.1242/dev.01067](https://doi.org/10.1242/dev.01067).
- Ramachandrapa S, Raimondo A, Cali AM, Keogh JM, Henning E, Saeed S, Thompson A, Garg S, Bochukova EG, Brage S, Trowse V, Wheeler E, Sullivan AE, Dattani M, Clayton PE, Datta V, Bruning JB, Wareham NJ, O’Rahilly S, Peet DJ, Barroso I, Whitelaw ML, Farooqi IS. 2013.** Rare variants in single-minded 1 (SIM1) are associated with severe obesity. *The Journal of Clinical Investigation* **123**:3042–3050 DOI [10.1172/JCI68016](https://doi.org/10.1172/JCI68016).
- Ramos JSL, Delmore KE, Liedvogel M. 2017.** Candidate genes for migration do not distinguish migratory and non-migratory birds. *Journal of Comparative Physiology A* **203**(6–7):383–397 DOI [10.1007/s00359-017-1184-6](https://doi.org/10.1007/s00359-017-1184-6).
- Reimegård J, Kundu S, Pendle A, Irish VF, Shaw P, Nakayama N, Sundström JF, Emanuelsson O. 2017.** Genome-wide identification of physically clustered genes suggests chromatin-level co-regulation in male reproductive development in *Arabidopsis thaliana*. *Nucleic Acids Research* **45**(6):3253–3265 DOI [10.1093/nar/gkx087](https://doi.org/10.1093/nar/gkx087).
- Rezzani R, Nardo L, Favero G, Peroni M, Rodella LF. 2014.** Thymus and aging: morphological, radiological, and functional overview. *Age* **36**(1):313–351 DOI [10.1007/s11357-013-9564-5](https://doi.org/10.1007/s11357-013-9564-5).
- Rhein VF, Carroll J, He J, Ding S, Fearnley IM, Walker JE. 2014.** Human METTL20 methylates lysine residues adjacent to the recognition loop of the electron transfer flavoprotein in mitochondria. *Journal of Biological Chemistry* **289**(35):24640–24651 DOI [10.1074/jbc.M114.580464](https://doi.org/10.1074/jbc.M114.580464).
- Rodhouse TJ, Rodriguez RM, Banner KM, Ormsbee PC, Barnett J, Irvine KM. 2019.** Evidence of region-wide bat population decline from long-term monitoring and Bayesian occupancy models with empirically informed priors. *Ecology and Evolution* **9**(19):11078–11088 DOI [10.1002/ece3.5612](https://doi.org/10.1002/ece3.5612).
- Rokas A, Wisecaver JH, Lind AL, Rokas A, Wisecaver JH, Lind AL. 2018.** The birth, evolution and death of metabolic gene clusters in fungi, The birth, evolution and death of metabolic gene clusters in fungi. *Nature Reviews Microbiology* **16**(12):731–744 DOI [10.1038/s41579-018-0075-3](https://doi.org/10.1038/s41579-018-0075-3).
- Roux J, Privman E, Moretti S, Daub JT, Robinson-Rechavi M, Keller L. 2014.** Patterns of positive selection in seven ant genomes. *Molecular Biology and Evolution* **31**(7):1661–1685 DOI [10.1093/molbev/msu141](https://doi.org/10.1093/molbev/msu141).
- Sandve SR, Rohlfs RV, Hvidsten TR. 2018.** Subfunctionalization versus neofunctionalization after whole-genome duplication. *Nature Genetics* **50**(7):908–909 DOI [10.1038/s41588-018-0162-4](https://doi.org/10.1038/s41588-018-0162-4).
- Santagata S. 2021.** Select_Test. Available at https://github.com/Santagata/Select_Test (accessed 20 September 2021).
- Sarzi E, Goffart S, Serre V, Chrétien D, Slama A, Munnich A, Spelbrink JN, Rötig A. 2007.** Twinkle helicase (PEO1) gene mutation causes mitochondrial DNA depletion. *Annals of Neurology* **62**(6):579–587 DOI [10.1002/ana.21207](https://doi.org/10.1002/ana.21207).

- Scheffer DI, Shen J, Corey DP, Chen Z-Y. 2015. Gene expression by mouse inner ear hair cells during development. *Journal of Neuroscience* 35(16):6366–6380 DOI 10.1523/JNEUROSCI.5126-14.2015.
- Seeger TF, Bartlett B, Coskran TM, Culp JS, James LC, Krull DL, Lanfear J, Ryan AM, Schmidt CJ, Strick CA, Varghese AH, Williams RD, Wylie PG, Menniti FS. 2003. Immunohistochemical localization of PDE10A in the rat brain. *Brain Research* 985:113–126 DOI 10.1016/S0006-8993(03)02754-9.
- Semple E, Hill JW. 2018. Sim1 neurons are sufficient for MC4R-mediated sexual function in male mice. *Endocrinology* 159(1):439–449 DOI 10.1210/en.2017-00488.
- Seppy M, Manni M, Zdobnov EM. 2019. BUSCO: assessing genome assembly and annotation completeness. In: *Gene Prediction*, Berlin: Springer, 227–245.
- Shabalina SA, Ogurtsov AY, Spiridonov AN, Novichkov PS, Spiridonov NA, Koonin EV. 2010. Distinct patterns of expression and evolution of intronless and intron-containing mammalian genes. *Molecular Biology and Evolution* 27(8):1745–1749 DOI 10.1093/molbev/msq086.
- Shen Y-Y, Liang L, Zhu Z-H, Zhou W-P, Irwin DM, Zhang Y-P. 2010. Adaptive evolution of energy metabolism genes and the origin of flight in bats. *Proceedings of the National Academy of Sciences USA* 107(19):8666–8671 DOI 10.1073/pnas.0912613107.
- Shump KA, Shump AU. 1982. *Lasiurus borealis*. *Mammalian species* 183:1–6 DOI 10.2307/3503878.
- Simakov O, Marlétaz F, Yue J-X, O’Connell B, Jenkins J, Brandt A, Calef R, Tung C-H, Huang T-K, Schmutz J, Satoh N, Yu J-K, Putnam NH, Green RE, Rokhsar DS. 2020. Deeply conserved synteny resolves early events in vertebrate evolution. *Nature Ecology & Evolution* 4:820–830 DOI 10.1038/s41559-020-1156-z.
- Singal DP, Ye M, Quadri SA. 1995. Major Histocompatibility-encoded human proteasome LMP2: genomic organization and a new form of mRNA. *Journal of Biological Chemistry* 270(4):1966–1970 DOI 10.1074/jbc.270.4.1966.
- Slater GSC, Birney E. 2005. Automated generation of heuristics for biological sequence comparison. *BMC Bioinformatics* 6(1):1–11 DOI 10.1186/1471-2105-6-1.
- Smith MD, Wertheim JO, Weaver S, Murrell B, Scheffler K, Kosakovsky Pond SL. 2015. Less is more: an adaptive branch-site random effects model for efficient detection of episodic diversifying selection. *Molecular Biology and Evolution* 32(5):1342–1353 DOI 10.1093/molbev/msv022.
- Snyder HL, Bačík I, Bennink JR, Kearns G, Behrens TW, Bächli T, Orlowski M, Yewdell JW. 1997. Two novel routes of transporter associated with antigen processing (TAP)-independent major histocompatibility complex class I antigen processing. *The Journal of Experimental Medicine* 186(7):1087–1098 DOI 10.1084/jem.186.7.1087.
- Song W-J, Mondal P, Li Y, Lee SE, Hussain MA. 2013. Pancreatic β -cell response to increased metabolic demand and to pharmacologic secretagogues requires EPAC2A. *Diabetes* 62(8):2796–2807 DOI 10.2337/db12-1394.
- Sonnhammer EL, von Heijne G, Krogh A. 1998. A hidden Markov model for predicting transmembrane helices in protein sequences. *Proceedings, International Conference on Intelligent Systems for Molecular Biology* 6:175–182.
- Sovic MG, Carstens BC, Gibbs HL. 2016. Genetic diversity in migratory bats: results from RADseq data for three tree bat species at an Ohio windfarm. *PeerJ* 4(6):e1647 DOI 10.7717/peerj.1647.
- Srivastava DP, Jones KA, Woolfrey KM, Burgdorf J, Russell TA, Kalmbach A, Lee H, Yang C, Bradberry MM, Wokosin D, Moskal JR, Casanova MF, Waters J, Penzes P. 2012. Social,

- communication, and cortical structural impairments in Epac2-deficient mice. *Journal of Neuroscience* 32(34):11864–11878 DOI 10.1523/JNEUROSCI.1349-12.2012.
- Stum M, Davoine C-S, Vicart S, Guillot-Noël L, Topaloglu H, Carod-Artal FJ, Kayserili H, Hentati F, Merlini L, Urtizberea JA, Hammouda EL-H, Quan PC, Fontaine B, Nicole S. 2006.** Spectrum of HSPG2 (Perlecan) mutations in patients with Schwartz-Jampel syndrome. *Human Mutation* 27(11):1082–1091 DOI 10.1002/humu.20388.
- Suyama M, Torrents D, Bork P. 2006.** PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Research* 34(Web Server):W609–W612 DOI 10.1093/nar/gkl315.
- Suzuki K, Harada N, Yamane S, Nakamura Y, Sasaki K, Nasteska D, Joo E, Shibue K, Harada T, Hamasaki A, Toyoda K, Nagashima K, Inagaki N. 2013.** Transcriptional regulatory factor X6 (Rfx6) increases gastric inhibitory polypeptide (GIP) expression in enteroendocrine K-cells and is involved in GIP hypersecretion in high fat diet-induced obesity. *Journal of Biological Chemistry* 288(3):1929–1938 DOI 10.1074/jbc.M112.423137.
- Taly A, Corringer P-J, Guedin D, Lestage P, Changeux J-P. 2009.** Nicotinic receptors: allosteric transitions and therapeutic targets in the nervous system. *Nature Reviews Drug Discovery* 8(9):733–750 DOI 10.1038/nrd2927.
- Tanaka H, Yoshida T, Miyamoto N, Motoike T, Kurosu H, Shibata K, Yamanaka A, Williams SC, Richardson JA, Tsujino N, Garry MG, Lerner MR, King DS, O'Dowd BF, Sakurai T, Yanagisawa M. 2003.** Characterization of a family of endogenous neuropeptide ligands for the G protein-coupled receptors GPR7 and GPR8. *Proceedings of the National Academy of Sciences USA* 100:6251–6256 DOI 10.1073/pnas.0837789100.
- Tarailo-Graovac M, Chen N. 2009.** Using RepeatMasker to identify repetitive elements in genomic sequences. *Current Protocols in Bioinformatics* 25(1):4–10 DOI 10.1002/0471250953.bi0410s25.
- Taranda J, Maison SF, Ballesterio JA, Katz E, Savino J, Vetter DE, Boulter J, Liberman MC, Fuchs PA, Elgoyhen AB. 2009.** A point mutation in the hair cell nicotinic cholinergic receptor prolongs cochlear inhibition and enhances noise protection. *PLOS Biology* 7(1):e1000018 DOI 10.1371/journal.pbio.1000018.
- Teeling EC, Vernes SC, Dávalos LM, Ray DA, Gilbert MTP, Myers E, Bat1K Consortium. 2018.** Bat biology, genomes, and the Bat1K project: to generate chromosome-level genomes for all living bat species. *Annual Review of Animal Biosciences* 6:23–46 DOI 10.1146/annurev-animal-022516-022811.
- Thalén F. 2018.** PhyloPyPruner: tree-based orthology inference for phylogenomics with new methods for identifying and excluding contamination. Available at <https://lup.lub.lu.se/student-papers/search/publication/8963554> (accessed 20 September 2021).
- Thapa P, Farber DL. 2019.** The role of the thymus in the immune response. *Thoracic Surgery Clinics* 29(2):123–131 DOI 10.1016/j.thorsurg.2018.12.001.
- Tolson KP, Gemelli T, Meyer D, Yazdani U, Kozlitina J, Zinn AR. 2014.** Inducible neuronal inactivation of Sim1 in adult mice causes hyperphagic obesity. *Endocrinology* 155(7):2436–2444 DOI 10.1210/en.2013-2125.
- Tomanek I, Grah R, Lagator M, Andersson A, Bollback JP, Tkačik G, Guet CC. 2020.** Gene amplification as a form of population-level gene expression regulation. *Nature Ecology & Evolution* 4(4):612–625 DOI 10.1038/s41559-020-1132-7.
- Ujita M, McAuliffe J, Hindsgaul O, Sasaki K, Fukuda MN, Fukuda M. 1999.** Poly-N-acetyllactosamine synthesis in branched N-glycans is controlled by complemental branch specificity of i-extension enzyme and β 1, 4-galactosyltransferase I. *Journal of Biological Chemistry* 274(24):16717–16726 DOI 10.1074/jbc.274.24.16717.

- Van Haelst M, Maiburg M, Baujat G, Jadeja S, Monti E, Bland E, Pearce K, Fraser Syndrome Collaboration Group, Hennekam R, Scambler P. 2008. Molecular study of 33 families with Fraser syndrome new data and mutation review. *American Journal of Medical Genetics Part A* 146:2252–2257 DOI 10.1002/ajmg.a.32440.
- Vinckenbosch N, Dupanloup I, Kaessmann H. 2006. Evolutionary fate of retroposed gene copies in the human genome. *Proceedings of the National Academy of Sciences USA* 103(9):3220–3225 DOI 10.1073/pnas.0511307103.
- Vincler M, Wittenauer S, Parker R, Ellison M, Olivera BM, McIntosh JM. 2006. Molecular mechanism for analgesia involving specific antagonism of $\alpha 9\alpha 10$ nicotinic acetylcholine receptors. *Proceedings of the National Academy of Sciences USA* 103(47):17880–17884 DOI 10.1073/pnas.0608715103.
- Wanamaker CP, Green WN. 2005. N-linked glycosylation is required for nicotinic receptor assembly but not for subunit associations with calnexin. *Journal of Biological Chemistry* 280(40):33800–33810 DOI 10.1074/jbc.M501813200.
- Wang W, Huang L, Huang Y, Yin J, Berk AJ, Friedman JM, Wang G. 2009. Mediator MED23 links insulin signaling to the adipogenesis transcription cascade. *Developmental Cell* 16(5):764–771 DOI 10.1016/j.devcel.2009.04.006.
- Wang X, Zhang L, Zhang Y, Zhao P, Qian L, Yuan Y, Liu J, Cheng Q, Xu W, Zuo Y, Guo T, Yu Z, Zheng H. 2017. JOSD1 negatively regulates type-I interferon antiviral activity by deubiquitinating and stabilizing SOCS1. *Viral Immunology* 30:342–349 DOI 10.1089/vim.2017.0015.
- Watanabe N, Wada M, Irukayama-Tomobe Y, Ogata Y, Tsujino N, Suzuki M, Furutani N, Sakurai T, Yamamoto M. 2012. A single nucleotide polymorphism of the neuropeptide B/W receptor-1 gene influences the evaluation of facial expressions. *PLOS ONE* 7:e35390 DOI 10.1371/journal.pone.0035390.
- Weaver SP, Hein CD, Simpson TR, Evans JW, Castro-Arellano I. 2020. Ultrasonic acoustic deterrents significantly reduce bat fatalities at wind turbines. *Global Ecology and Conservation* 24:e01099 DOI 10.1016/j.gecco.2020.e01099.
- Wedemeyer C, Vattino LG, Moglie MJ, Ballesterio J, Maison SF, Di Guilmi MN, Taranda J, Liberman MC, Fuchs PA, Katz E, Elgoyhen AB. 2018. A gain-of-function mutation in the $\alpha 9$ nicotinic acetylcholine receptor alters medial olivocochlear efferent short-term synaptic plasticity. *Journal of Neuroscience* 38(16):3939–3954 DOI 10.1523/JNEUROSCI.2528-17.2018.
- Weisenfeld NI, Kumar V, Shah P, Church DM, Jaffe DB. 2017. Direct determination of diploid genome sequences. *Genome research* 27:757–767 DOI 10.1101/gr.214874.116.
- Weller TJ, Castle KT, Liechti F, Hein CD, Schirmacher MR, Cryan PM. 2016. First direct evidence of long-distance seasonal movements and hibernation in a migratory bat. *Scientific Reports* 6(1):1–7 DOI 10.1038/srep34585.
- Willour VL, Seifuddin F, Mahon PB, Jancic D, Pirooznia M, Steele J, Schweizer B, Goes FS, Mondimore FM, MacKinnon DF, Perlis RH, Lee PH, Huang J, Kelsoe JR, Shilling PD, Rietschel M, Nöthen M, Cichon S, Gurling H, Purcell S, Smoller JW, Craddock N, DePaulo JR Jr, Schulze TG, McMahon FJ, Zandi PP, Potash JB. 2012. A genome-wide association study of attempted suicide. *Molecular Psychiatry* 17:433–444 DOI 10.1038/mp.2011.4.
- Xi D, Gandhi N, Lai M, Kublaoui BM. 2012. Ablation of Sim1 neurons causes obesity through hyperphagia and reduced energy expenditure. *PLOS ONE* 7:e36453 DOI 10.1371/journal.pone.0036453.

- Xie J, Li Y, Shen X, Goh G, Zhu Y, Cui J, Wang L-F, Shi Z-L, Zhou P. 2018. Dampened STING-dependent interferon activation in bats. *Cell Host & Microbe* 23(3):297–301 DOI 10.1016/j.chom.2018.01.006.
- Yalvac ME, Amornvit J, Braganza C, Chen L, Hussain S-RA, Shontz KM, Montgomery CL, Flanigan KM, Lewis S, Sahenk Z. 2017. Impaired regeneration in calpain-3 null muscle is associated with perturbations in mTORC1 signaling and defective mitochondrial biogenesis. *Skeletal Muscle* 7(1):1–18 DOI 10.1186/s13395-017-0146-6.
- Yamashita Y, Nakada S, Yoshihara T, Nara T, Furuya N, Miida T, Hattori N, Arikawa-Hirasawa E. 2018. Perlecan, a heparan sulfate proteoglycan, regulates systemic metabolism with dynamic changes in adipose tissue and skeletal muscle. *Scientific Reports* 8(1):1–14 DOI 10.1038/s41598-018-25635-x.
- Yang Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. *Molecular Biology and Evolution* 24(8):1586–1591 DOI 10.1093/molbev/msm088.
- Yates AD, Achuthan P, Akanni W, Allen J, Allen J, Alvarez-Jarreta J, Amode MR, Armean IM, Azov AG, Bennett R, Bhai J, Billis K, Boddu S, Marugán JC, Cummins C, Davidson C, Dodiya K, Fatima R, Gall A, Giron CG, Gil L, Grego T, Haggerty L, Haskell E, Hourlier T, Izuogu OG, Janacek SH, Juettemann T, Kay M, Lavidas I, Le T, Lemos D, Martinez JG, Maurel T, McDowall M, McMahan A, Mohanan S, Moore B, Nuhn M, Oheh DN, Parker A, Parton A, Patricio M, Sakthivel MP, Abdul Salam AI, Schmitt BM, Schuilenburg H, Sheppard D, Sycheva M, Szuba M, Taylor K, Thormann A, Threadgold G, Vullo A, Walts B, Winterbottom A, Zadissa A, Chakiachvili M, Flint B, Frankish A, Hunt SE, Iisley G, Kostadima M, Langridge N, Loveland JE, Martin FJ, Morales J, Mudge JM, Muffato M, Perry E, Ruffier M, Trevanion SJ, Cunningham F, Howe KL, Zerbino DR, Flicek P. 2020. Ensembl 2020. *Nucleic Acids Research* 48:D682–D688 DOI 10.1093/nar/gkz966.
- Yue F, Cheng Y, Breschi A, Vierstra J, Wu W, Ryba T, Sandstrom R, Ma Z, Davis C, Pope BD, Shen Y, Pervouchine DD, Djebali S, Thurman RE, Kaul R, Rynes E, Kirilusha A, Marinov GK, Williams BA, Trout D, Amrhein H, Fisher-Aylor K, Antoshechkin I, DeSalvo G, See LH, Fastuca M, Drenkow J, Zaleski C, Dobin A, Prieto P, Lagarde J, Bussotti G, Tanzer A, Denas O, Li K, Bender MA, Zhang M, Byron R, Groudine MT, McCleary D, Pham L, Ye Z, Kuan S, Edsall L, Wu Y-C, Rasmussen MD, Bansal MS, Kellis M, Keller CA, Morrissey CS, Mishra T, Jain D, Dogan N, Harris RS, Cayting P, Kawli T, Boyle AP, Euskirchen G, Kundaje A, Lin S, Lin Y, Jansen C, Malladi VS, Cline MS, Erickson DT, Kirkup VM, Learned K, Sloan CA, Rosenbloom KR, Lacerda de Sousa B, Beal K, Pignatelli M, Flicek P, Lian J, Kahveci T, Lee D, James Kent W, Ramalho-Santos M, Herrero J, Notredame C, Johnson A, Vong S, Lee K, Bates D, Neri F, Diegel M, Canfield T, Sabo PJ, Wilken MS, Reh TA, Giste E, Shafer A, Kutayavin T, Haugen E, Dunn D, Reynolds AP, Neph S, Humbert R, Scott Hansen R, De-Brujin M, Selleri L, Rudensky A, Josefowicz S, Samstein R, Eichler EE, Orkin SH, Levasseur D, Papayannopoulou T, Chang KH, Skoultschi A, Gosh S, Distech C, Treuting P, Wang Y, Weiss MJ, Blobel GA, Cao X, Zhong S, Wang T, Good PJ, Lowdon RF, Adams LB, Zhou XQ, Pazin MJ, Feingold EA, Wold B, Taylor J, Mortazavi A, Weissman SM, Stamatoyannopoulos JA, Snyder MP, Guigo R, Gingeras TR, Gilbert DM, Hardison RC, Beer MA, Ren B, The Mouse ENCODE Consortium. 2014. A comparative encyclopedia of DNA elements in the mouse genome. *Nature* 515:355–364 DOI 10.1038/nature13992.
- Zhang G, Cowled C, Shi Z, Huang Z, Bishop-Lilly KA, Fang X, Wynne JW, Xiong Z, Baker ML, Zhao W, Tachedjian M, Zhu Y, Zhou P, Jiang X, Ng J, Yang L, Wu L, Xiao J, Feng Y, Chen Y, Sun X, Zhang Y, Marsh GA, Cramer G, Broder CC, Frey KG, Wang L-F, Wang J. 2013.

Comparative analysis of bat genomes provides insight into the evolution of flight and immunity. *Science* **339**:456–460 DOI [10.1126/science.1230835](https://doi.org/10.1126/science.1230835).

Zhang H, Chen X, Sairam M. 2012. Novel genes of visceral adiposity: identification of mouse and human mesenteric estrogen-dependent adipose (MEDA)-4 gene and its adipogenic function. *Endocrinology* **153**(6):2665–2676 DOI [10.1210/en.2011-2008](https://doi.org/10.1210/en.2011-2008).

Zhang X, Wang D, Dongye M, Zhu Y, Chen C, Wang R, Long E, Liu Z, Wu X, Lin D, Chen J, Lin Z, Wang J, Li W, Li Y, Li D, Lin H. 2019a. Loss-of-function mutations in *FREM2* disrupt eye morphogenesis. *Experimental Eye Research* **181**:302–312 DOI [10.1016/j.exer.2019.02.013](https://doi.org/10.1016/j.exer.2019.02.013).

Zhang Z, Mu Y, Shan L, Sun D, Guo W, Yu Z, Tian R, Xu S, Yang G. 2019b. Divergent evolution of TRC genes in mammalian niche adaptation. *Frontiers in Immunology* **10**:871 DOI [10.3389/fimmu.2019.00871](https://doi.org/10.3389/fimmu.2019.00871).

Zheng X, Black D, Chambon P, Egly J. 1990. Sequencing and expression of complementary DNA for the general transcription factor BTF3. *Nature* **344**(6266):556–559 DOI [10.1038/344556a0](https://doi.org/10.1038/344556a0).