BUSTED-PH: Gene-wide detection of trait-associated positive selection

Authors

Avery Selberg[1], Nathan Clark[2], Maria Chikina[3], Sergei L Kosakovsky Pond[4]

Affiliations

[1] Temple University, Philadelphia, PA
[2] Department of Biological Sciences, University of Pittsburgh, Pittsburgh, PA 15260
[3] Department of Computational and Systems Biology, University of Pittsburgh, Pittsburgh, PA 15260

Abstract

Introduction

When multiple species or clades experience similar natural selection pressures they often independently develop similar adaptations. This process, known as phenotypic convergence, occurs in traits such as social behavior in butterflies (Cicconardi et al. 2025), endothermy in fish (Melendez-Vazquez et al. 2025), and the ability to ferment wine in yeast (Onetto et al. 2025). Convergent evolution of traits creates a unique setting where similar natural selection pressures have operated on multiple independent “replicates”, boosting statistical power to detect its action.

Convergent evolutionary signal can be detected with tools such as `RERConverge` (Kowalczyk et al. 2019), `TraitRate` (Halabi et al. 2021), `PhyloAcc` (Hu et al. 2019), and `Forward Genomics` (Prudent et al. 2016). The main source of signal comes from detecting elevated or reduced evolutionary rates in convergent species, compared to the rest of the species, either in genes, or noncoding regions of DNA. While these methods are very useful in detecting specific genes under evolutionary constraint, rate-based methods struggle to distinguish whether a gene is changing under positive selection or from reduced evolutionary constraint (Kowalczyk et al. 2020). Another prominent application is detection of relaxed selection or loss of constraint, such as loss of vision in subterranean mammals (Partha et al. 2017) or loss of flight in birds (Sackton et al. 2019).

To identify genes associated with new traits, codon models have been widely used (Kosakovsky Pond and Frost 2005; Murrell et al. 2015; Wertheim et al. 2015). These models detect positive selection by estimating the ratio of the rate of nonsynonymous substitutions (d_N) to the rate of synonymous substitutions (d_S), where synonymous changes are assumed to be neutral. If d_N/d_S (or ω) is greater than one, adaptation is presumed to be occurring (Nei and Gojobori 1986). Branch-site codon models in particular are popular because of their increased statistical power and their ability to separate foreground branches (with the trait) and background branches (without the trait). Researchers typically test for evidence of positive selection on the foreground branches to identify adaptive changes linked to the phenotype. However, it is equally critical to test for positive selection on the background branches to ensure the evolutionary signal is specific to the convergently evolving phenotype and is not instead found across the entire tree due to other processes (Kowalczyk et al. 2021).

Although protocols have been developed to test for positive selection specific to foreground branches (Kowalczyk et al. 2021), there is currently no comprehensive method to test for convergent evolution that considers positive selection in both the foreground and background set of branches. Further, d_N/d_S methods are rather sensitive to alignment or other errors, and this must be accounted for in large-scale analyses. Here, we expanded the BUSTED-E method (Selberg et al. 2025), designed to allow detection of selection in the presence of residual alignment error, to three separate tests: (i) a test for positive selection on the foreground branches, (ii) a test for positive selection on the background branches, and a (iii) test for the difference between the foreground and background branches. When tests (i) and (iii) are significant, and test (ii) is not, we infer that convergent evolution is associated with the trait of interest. We validate BUSTED-PH on positive control genes with well-established roles in adaptation and apply it to a genome-wide dataset of echolocating mammals to identify novel candidates linked to auditory function. Using simulation studies, we further evaluate model performance under varying levels of selection strength and phenotypic replication. Together, these analyses demonstrate that BUSTED-PH offers a robust and interpretable approach for detecting trait-associated adaptation across phylogenies. Of note, since we released BUSTED-PH as a `HyPhy` package module (Kosakovsky Pond et al. 2020) in 2023, it has already been cited in multiple studies of adaptive evolution, across different taxonomic scales and biological hypotheses (Supplemental Table 1).

Methods

Statistical methodology

We model codon evolution using a finite state continuous time Markov model based on the Muse-Gaut family of codon models (Muse and Gaut 1994). BUSTED-PH builds on the previous BUSTED models (Murrell et al. 2015; Wisotsky et al. 2020; Lucaci et al. 2023; Selberg et al. 2025), originally developed to test for episodic gene-level positive selection. BUSTED-PH has already been made available and integrated into the HyPhy framework, leading to its adoption in several recent phylogenetic analyses (Supplemental Table 1). BUSTED-PH uses the same

instantaneous rate (Q) matrix as in BUSTED-S (Wisotsky et al. 2020), where each the instantaneous substitution rate between codon i and codon j is given by:

$$q_{ij} = \begin{cases} \alpha^s \theta_{ij} \pi_j^p & \text{1-step synonymous change} \\ \alpha^s \omega^{bs} \theta_{ij} \pi_j^p & \text{1-step nonsynonymous change} \\ 0 & \text{Multi-step change} \\ -\sum_{l \neq i} q_{il}, & \text{No substitution} \end{cases}$$

where ω is the substitution rate ratio associated with category *k*, θ is the nucleotide substitution rate parameter for codons *i* and *j*, and π denotes position specific (*p*) equilibrium frequency of the target nucleotide (θ) from the CF3x4 corrected empirical estimator (Pond et al. 2010). At each branch and each site, ω is an independent draw from a *k* (default *k* = 3) generate discrete distribution; this feature allows selective pressures to vary from site to site and branch to branch. All model parameters are estimated using maximum likelihood and direct numerical optimization in HyPhy (Kosakovsky Pond et al. 2020). This model includes synonymous rate variation by default (Wisotsky et al. 2020); however, a non-synonymous rate variation *Q* matrix as given by (Murrell et al. 2015) can also be used by including `--srv No` when calling the model. The analysis also supports multi-nucleotide substitutions via the `--multiple-hits` flag, although we do not use this version of the method here; an interested reader should see (2023).

The BUSTED-PH model incorporates phenotypes by partitioning species, *a priori*, into phenotype-specific foreground and background (FG and BG) subsets of branches to estimate two separate ω distributions for each branch subset. Optionally, a third subset of “nuisance” branches can be designated to accommodate branches without a phenotype designation, or otherwise not directly used for inference.

Consider a phylogenetic tree that is separated into distinct *FG* and *BG* sets of species, based on a binary phenotype. The BUSTED-PH model allows for independent selective regimes, S_{FG} or S_{BG}. Each selective regime is described by a distribution of ω substitution rate ratios with *k* discrete bins ($\omega_1 \leq \omega_{k-1} \ldots \leq 1 \leq \omega_k$) For a given site and a branch in *F* (or *B*), ω is modeled as a random independent (of other branches) draw from S_{FG} (or S_{BG}), and the site phylogenetic likelihood is computed as the expectation over all possible draws. *k* is a parameter that is selected a priori based on our previous or typical model complexity needed to model gene evolution (Murrell et al. 2015), but can be adjusted, or selected by sequential search and a model-ranking criterion (e.g., AIC). Lastly, if the error sink is enabled, we augment the ω distribution with an error rate class ($100 \leq \omega_e$, with total weight not exceeding 1%), inferred separately for FG and BG branches. For details, please see Selberg et al. (2025).

Hypothesis testing and implementation

To test for FG-specific positive selection, three likelihood ratio tests (LRT) are used (Table 1). First, a test for evidence of episodic diversifying selection (EDS) in *FG* branches, with a likelihood ratio test used to compare model 1 and model 2. Second, a test for evidence of positive selection in *BG* branches, with a LRT comparing model 1 and model 3. Critical values of the LRT for these two comparisons are derived from a 50:50 mixture model of χ_0^2 and χ_2^2 (Self and Liang 1987; Wisotsky et al. 2020); this mixture is generally conservative. Third, a test for the difference between *FG* and *BG* branches with a LRT comparing model 1 and model 4 (Table 1, Figure 1). Critical values for this final LRT are derived from the χ^2_{2k-1} distribution.

Model	Explanation	Purpose
1	Independent ω distributions in the *FG* and *BG* branches	Universal alternative model: compare to models 2, 3, and 4 using likelihood ratio test
2	Enforce $\omega_k \leq 1$ in *FG* branches, *BG* branches are unconstrained	Null model in likelihood ratio test for positive selection in the *FG* branches
3	Enforce $\omega_k \leq 1$ in *BG* branches, *FG* branches are unconstrained	Null model in likelihood ratio test for positive selection in the *BG* branches
4	Entire tree has a single ω distribution; no separate *FG* and *BG* subsets of branches	Null model in likelihood ratio test for for whether ω distributions are different between *FG* and *BG* branches

Table 1. Models run in BUSTED-PH to detect FG-specific positive selection. Models 2, 3, and 4 serve as null models, and model 1 serves as the universal alternative model for likelihood ratio test comparisons.

We define a gene as exhibiting trait-associated EDS if it meet three criteria:

1. Statistically significant evidence of positive selection on trait-associated (foreground) branches ($p < C$),
2. No evidence of positive selection on trait-absent (background) branches ($p > C$), and
3. A statistically significant difference between the foreground and background ω (dN/dS) distributions, consistent with a change in selective pressure linked to the phenotype ($p < C$).

C is a significance level selected a priori, and is applied to p-values following the 3-way Holm-Bonferroni correction (C = 0.05 by default).

Optionally, criterion (2) may be relaxed to instead require that evidence for positive selection is stronger in the foreground than in the background, allowing for the identification of trait-associated EDS even when some background signal is present; however, we did not adopt that approach in this study. To prioritize candidate genes for follow-up, we applied a heuristic score based on the statistical support for each component of the test:

$$\text{Ranking score} = -\log_{10}(\max(10^{-8}, p_{FG})) - \log_{10}(\max(10^{-8},, p_{DIFF})) + \log_{10}(\max(10^{-8}, p_{BG}))$$

where p_{FG}, p_{DIFF}, and p_{BG} are the p-values for selection in the foreground, the difference between foreground and background, and selection in the background, respectively.

Empirical data analysis

Application to established positive controls

To validate the behavior of BUSTED-PH on well-characterized examples of molecular adaptation, we analyzed three canonical positive control genes previously shown, in multiple studies, to exhibit convergent or lineage-specific adaptive evolution linked to phenotypic traits. For each gene, we used published multiple sequence alignments and associated binary (presence/absence) trait annotations from prior studies (summarized in Table 2). We applied BUSTED-PH to test for episodic diversifying selection (EDS) associated with the specified trait and used BUSTED-E to account for residual alignment error.

For Prestin (SLC26A5) and the ATP alpha-1 subunit (ATP1), we used alignments and foreground/background annotations from (Fukushima and Pollock 2023). For SEMG2, we used the alignment and phenotype mapping from the TraitRELAX study (Halabi et al. 2021). Species annotations can be seen in Supplemental Figure 1. All analyses were run using HyPhy version ≥2.5.73, using the `--error-sink Yes --rates 2` options.

Genome wide analysis of echolocation in mammals

To identify genes potentially associated with the evolution of echolocation in mammals, we applied BUSTED-PH to a genome-wide dataset of orthologous protein-coding genes from 120 mammalian species, including both echolocating and non-echolocating species (Hecker and Hiller 2020; Redlich et al. 2024). We annotated species as echolocating (foreground) or non-echolocating (background), as shown in Supplemental Figure 2 (Gould 1965; Holland et al. 2004; Liu et al. 2014). Protein-coding sequence alignments were made by extracting protein-coding portions of the 120-species whole-genome alignment using the coordinates of UCSC ‘canonical’ human gene models for genome version hg38 (Hecker and Hiller 2020). Those exons were then pulled from the MAF alignment file with the `sub.msa` function of the RPHAST package (Hubisz et al. 2011), and the human reading frame was enforced through the codon.clean.msa function of the same package.

Software availability

BUSTED-PH is implemented in `HyPhy` versions 2.5.73 or later (Kosakovsky Pond et al. 2020). BUSTED-PH is executed using the `hyphy busted-ph` command. We created interactive notebooks for visualization of BUSTED-PH results (https://observablehq.com/collection/@hyphy/busted-ph). These notebooks include interactive results from this paper as well as notebooks to upload your own BUSTED-PH `JSON` output file(s) to visualize.

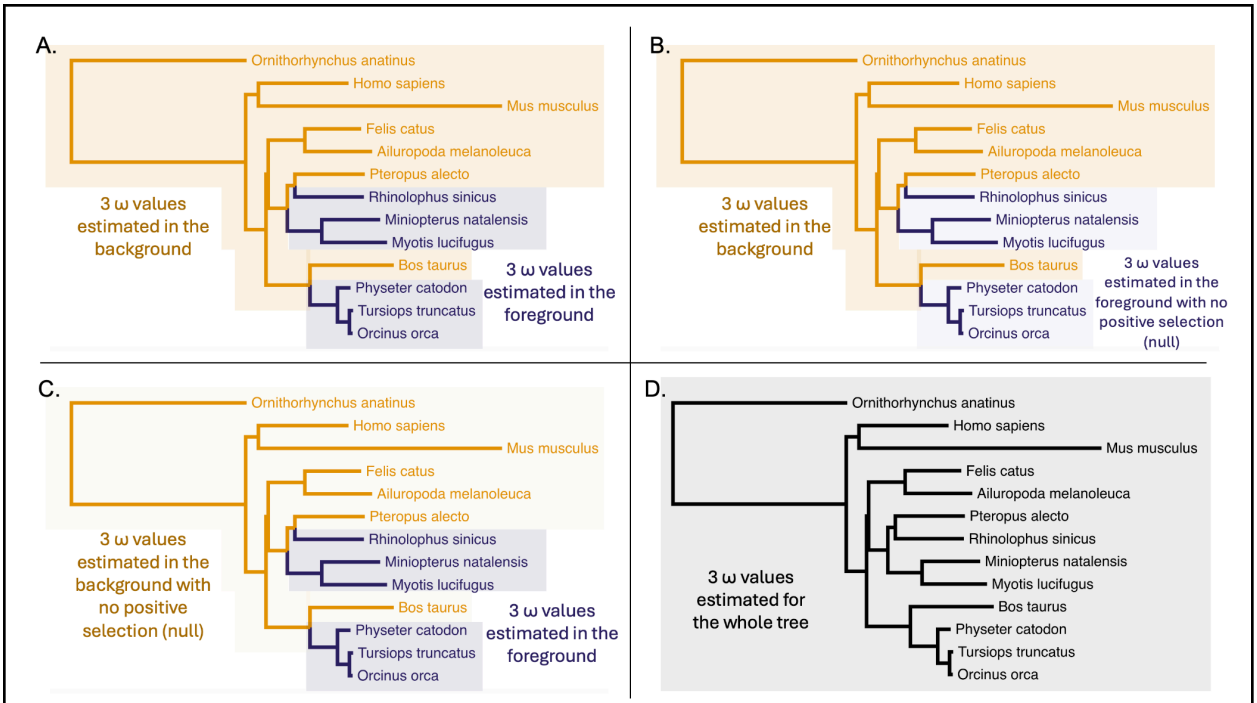

Figure 1. Schematic of the four models BUSTED-PH uses. Model A serves as the universal alternative model and is compared to models B,C, and D with a likelihood ratio test. Individual tests are denoted: the test for positive selection in FG (LRT with models A and B), the test for positive selection in the BG (LRT with models A and C), and the test for the difference between the FG and BG (LRT with models A and D).

Results

Application to established positive controls

To illustrate applications of BUSTED-PH, we selected three genes with broadly recognized and well-characterized evidence of genetic association with a phenotype. For all three genes, BUSTED-PH finds evidence of episodic diversifying selection (EDS) associated with a binary (present/absent) trait, *i.e.*, returns expected results on several canonical positive controls. We classified genes as showing evidence of EDS if they met the following criteria: (i) evidence of positive selection in the foreground, (ii) no evidence of positive selection in the background, and (iii) a statistically significant difference between the foreground and background selection profiles. We used BUSTED-PH with K=2 non-error ω rate classes because K=3 resulted in overfitting (zero-weight rate, or collapsed rate classes, confirmed by AIC-c comparison between the K=2 and K=3 models).

Prestin (SLC26A5). This gene was shown to be important for high frequency sensitivity in the inner ear of echolocating bats (Li et al. 2008), and later, also in echolocating marine mammals (Li et al. 2010), and has since become a prototypical example of convergent evolution. Applying BUSTED-PH, the alignment, and phenotype annotation from the Pollock (2023) paper (Table 2), we find a clear signal of EDS in the echolocating species, with a ~1% fraction of the alignment

subject to EDS, no evidence of EDS in the background branches, and significant difference between ω distributions on foreground and background branches. This alignment contains residual error (Selberg et al. 2025) on background lineages. Examples of such alignment errors are shown in Supplemental Figure 3.

ATP alpha-1 subunit. Convergent substitutions in the alpha-1 subunit of the Na+/K+ ATPase have been convincingly tied to resistance to ouabain – a plant derived cardiac glycoside, used for example as a poisoning agent in arrow tips, and other cardenolides (Dobler et al. 2012; Mohammadi et al. 2022). Applying BUSTED-PH to the alignment of ATP1 sequences from insects and phenotype annotation of glycoside resistance from the the Pollock (2023) paper (Table 2), we see a clear signal of EDS in the resistant species, with a ~1% fraction of the alignment subject to EDS, pervasive conservation on background lineages, and a marked difference between ω distributions on foreground and background branches. This alignment also contains residual error (Selberg et al. 2025) on background lineages (Supplemental Figure 3).

SEMG2. In species subject to high levels of sperm competition, several studies have found evidence of accelerated evolution and diversifying selection on this ejaculate protein, which is the main structural component of semen coagulum. Applying BUSTED-PH to the alignment of SEMG2 sequences and phenotype annotation of polyandry from the Halabi et al. (2021) paper (Table 2), there is a clear signal of EDS in the polyandryous species, with a large (~40%) fraction of the alignment subject to EDS, and a marked difference between ω distributions on foreground and background branches. This alignment also contains residual error (Selberg et al. 2025), this time on foreground lineages (Supplemental Figure 3).

Gene	LRT p-value	ω_1 (weight)	ω_2 (weight)	ω_E (weight)	Branches (L)
Prestin / echolocation, 128 sequences, 758 codons					
Foreground*	**0.015**	0.09 (98.8%)	7.15 (1.2%)	N.D	19 (0.19)
Background	0.48	0.04 (96.9%)	1.06 (3.1%)	201 (0.043%)	228 (3.13)
Shared	**<0.001**	0.05 (98.8%)	2.6 (1.2%)	287 (0.028%)	247 (3.32)
ATP1 alpha / Cardiac glycoside resistance, 29 sequences, 1045 codons					
Foreground*	**<0.001**	0.01 (99.3%)	5.2 (0.7%)	N.D	20 (0.71)
Background	0.50	0.01 (98.3%)	1.00 (1.7%)	203 (0.115%)	35 (7.11)
Shared	**<0.001**	0.01 (98.3%)	1.00 (1.7%)	100 (0.052%)	55 (7.82)
SEMG2 / polygynandrous mating patterns, 23 sequences, 586 codons					
Foreground*	**<0.001**	1.00 (59.0%)	4.58 (40.8%)	100 (0.24%)	8 (0.72)
Background	0.20	1.00 (29.9%)	1.24 (70.1%)	N.D	33 (0.34)
Shared	**0.02**	0.57 (42.9%)	2.25 (57.1%)	N.D	41 (1.06)

Table 2. Exemplar genes previously for evidence of convergent evolution or selection intensification. Inferred ω distributions for foreground, background, and shared branches are shown for each model, together with the number of branches (cumulative lengths in subs/site in parentheses) in each set. P-values are computed using asymptotic LRT distributions; the null models are (1) $\omega_3 \leq 1$ for Foreground and Background models, 2 degrees of freedom, and (2) ω distributions are the same on foreground and background branches for the Shared model, 5 degrees of freedom. The error component fitted is shown in the ω_E column. (*) All species with the corresponding phenotypes were added to the foreground set, as were all internal nodes where every descendant had the phenotype present (`all-descendant` labeling strategy); labeled trees are shown in Supplementary Figure 1.

Genome-wide analysis of echolocation in mammals

To evaluate the performance of BUSTED-PH in a genome-wide context, tested 18,940 orthologous gene alignments from a 120-species mammalian whole-genome alignment (Hecker and Hiller 2020; Redlich et al. 2024), focusing on echolocation as the binary phenotype of interest. This analysis parallels the study by Allard et al. (2025), which tested for convergently evolving amino acids under directional selection. Across the genome, BUSTED-PH identified 240 genes with evidence of phenotype-associated EDS. As expected, genes with phenotype-associated EDS exhibited elevated values of ω_+ (Figure 2E) and the foreground branch lengths were on average longer for genes experiencing phenotype associated EDS (Figure 2C), relative to the genome-wide background.

Top-ranked genes (Supplemental Table 2), prioritized based on the strength of phenotype-associated selection, include CDH23, a gene with a known role in hearing (Bolz et al. 2001; Bork et al. 2001), and ACTBL2, a beta-actin gene not previously characterized in this context. Although other beta-actin genes are implicated in hearing loss (Belyantseva et al. 2003; Zhu et al. 2003), the role of ACTBL2 in sensory function remains unknown. Prestin (SLC26A5), previously analyzed as a positive control, ranks fifth in this genome-wide screen, verifying its established association with echolocation. Of the 100 top-ranking genes from Allard et al. (2025), only four—CDH23, LOXHD1, SLC26A5, and TMC1—were shared by BUSTED-PH as genes under phenotype-associated EDS, highlighting the methodological differences between approaches that target directional versus diversifying selection.

STRING functional enrichment analysis (Szklarczyk et al. 2023) of the 240 phenotype-associated EDS genes revealed significant overrepresentation of terms related to sensory perception and signal transduction. Gene Ontology (GO) terms significantly overrepresented among these genes include sensory perception (GO:0007600; p = 0.0210), detection of chemical stimulus involved in sensory perception (GO:0050907; p = 0.0210), and nervous system process (GO:0050877; p = 0.0496). Notably, genes annotated with olfactory receptor activity (GO:0004984; p = 0.0170) and G protein–coupled receptor activity (GO:0004930; p = 0.0307) were also enriched. Although many of these genes are classified as olfactory receptors, this may reflect shared molecular mechanisms between olfactory and auditory systems—particularly in the use of large, rapidly evolving receptor families. Supporting this interpretation, we observed enrichment for genes expressed in outer hair cells of the cochlea (BTO:0003666; p = 0.0147), which are essential for active sound amplification in mammals and a core component of echolocation. Together, these results suggest that echolocating species have experienced recurrent positive selection on genes involved in

peripheral sensory detection and neural signal transduction, consistent with evolutionary pressures on the auditory system.

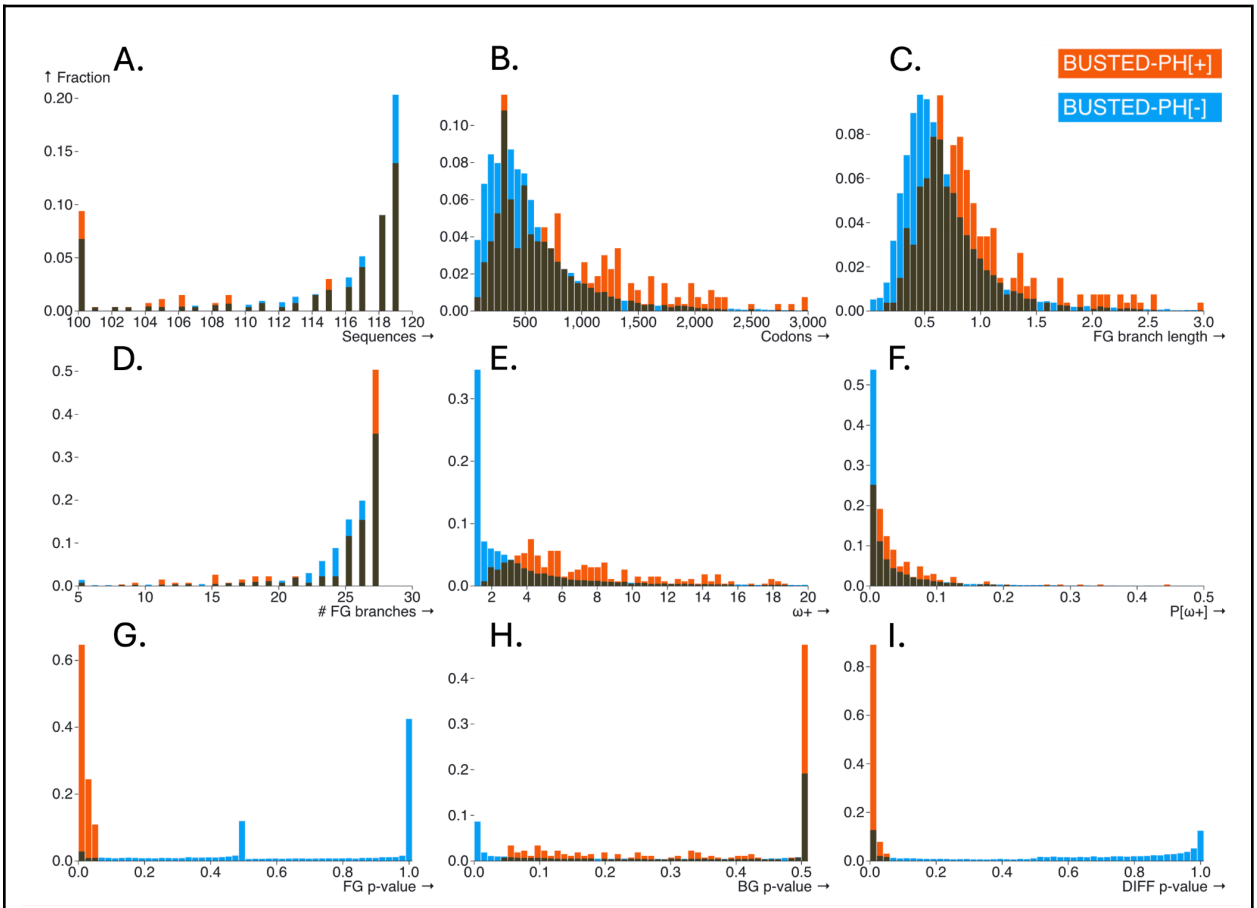

Figure 2. Gene feature distributions based on BUSTED-PH selection analyses. Genes with significant evidence of positive selection associated with the echolocation phenotype (BUSTED-PH[+]) are shown in orange; genes without significant association (BUSTED-PH[−]) are shown in blue. The y-axis in each panel shows the fraction of genes within each category falling into the corresponding x-axis bin: A. Number of sequences per gene. B. Gene length (number of codons). C. Total length of foreground branches in the gene tree. D. Number of foreground branches (i.e., echolocating species per gene tree). E. Estimated value of ω_+ (the dN/dS ratio for the positively selected site class). F. Estimated proportion of sites ($P[\omega_+]$) under positive selection. G. p-values from the BUSTED-PH test for positive selection in the foreground. H. p-values from the BUSTED-PH test for positive selection in the background. I. p-values from the BUSTED-PH test comparing selection between foreground and background branches.

Simulations

We evaluated BUSTED-PH using XX simulation scenarios to assess power and Type I error, and robustness to some model misspecification.

Parametric simulations under the Branch-Site Random Effects substitution model
These simulations evaluate BUSTED-PH performance when the evolutionary model used by the method is correctly specified, i.e., used to generate the sequence data.

Example datasets.

Nectar-consuming birds.

We took the tree of 45 bird species from Osipova et al. (2024), and used four nectar-consuming bird clades to designate foreground species (Supplemental Figure 4), and generated three *null* (N) and ten *power* (P) simulation scenarios (Table 3, `Birds` scenarios). The following observations were made. Simulation labels also encode the number of foreground clades (first integer) and the simulation scenario (integer following '-').

1. BUSTED-PH has well-controlled false positive rates (`N/4` simulations), both in cases when there is no EDS on FG branches (`N/4-2`, and `N/4-3` scenarios), or when there is EDS both on BG and FG branches (`N/4-1` scenario).
2. The power to detect EDS associated with FG lineages varies with:
    a. Effect size, (i.e., strength of selection on FG branches). Power is high (PH rate = 0.91–1.00) when ω > 5, and lower (PH rate = 0.20–0.52) when ω ≤ 2.
    b. Number of foreground clades. Detection improves when more independent origins of the phenotype are included. For example, scenarios with four FG clades (`P/4`) had higher PH rates (0.52–1.00) than scenarios with only one FG clade (`P/1`), which had PH rates ≤ 0.35. This is also influenced by the total branch lengths of FG.
3. Detection of a significant difference in ω distributions between FG and BG is a key driver of overall power. In scenarios where ω differed clearly between groups, this test was nearly always significant (Diff = 0.97–1.00), whereas it remained low in null simulations (Diff ≤ 0.01).
4. BUSTED-PH is robust to mild selection on background branches, as seen in scenario N/4-1, where ω > 1 on a subset of BG branches did not cause spurious phenotype-associated signal (PH rate = 0.00).

Overall, BUSTED-PH shows high specificity and good power when the number of independent phenotype transitions and selection strength are sufficient, supporting its use for genome-wide screens in comparative genomic datasets.

Exemplar genes.

We used the topologies, branch lengths, substitution biases, codon frequencies, and site counts from the three exemplar genes in Table 2. We simulated 100 replicates using the exact inferred distributions for ω (`P/1` scenarios) on FG and BG, and 100 replicates where a single distribution of ω was used for the entire tree (Shared model, `N/1` scenarios). For the latter case, the rate of detection for the *Difference* in selection between branches was low, as expected, since there was no difference between the clades. For the former case, the following observations are of interest.

1. **Prestin**. Because the simulated proportion of ω>1 on FG was low (~1%) **and** the total length of FG branches is short (0.2 substitutions), the power to detect selection in this case is low (18/100). However, there was nearly perfect 98/100 power to detect differences between FG and BG. The overall BUSTED-PH detection rate was limited by the FG power (18%).
2. **ATP**, detection of selection on the foreground branches was high (81/100). No EDS on BG was simulated, hence detection was low (3/100). The power to detect differences between FG and BG was high (79/100), and the overall BUSTED-PH detection rate was also respectable at 68%.
3. **SEMG2**, detection of selection on the foreground branches was very high (100/100); because 40% of the alignment was assigned to the positive selection regime. However, because there was also weak positive selection simulated on BG (ω=1.24), 34% of the replicates reported selection on BG, reducing the overall power to 52%. This is an interesting case because a positive result on empirical data still contained weak evidence of selection on BG, which is unearthed by simulations.

Scenario	Branches		Length		dN/dS distribution		Detection (p≤0.05)			PH rate
	BG	FG	BG	FG	BG	FG	FG	BG	Diff	
Birds/P/4-1	63	24	2.34	0.57	0.1:0.5/0.5:0.4/1:0.1	0.1:0.5/0.5:0.4/**5:0.1**	*0.95*	0.00	*0.99*	*0.94*
Birds/P/4-2	63	24	2.34	0.57	0.1:0.5/0.5:0.4/1:0.1	0.1:0.525/0.5:0.425/**5:0.05**	*0.55*	0.01	*0.80*	*0.52*
Birds/P/4-3	63	24	2.34	0.57	0.1:0.5/0.5:0.4/1:0.1	0.1:0.545/0.5:0.445/**5:0.01**	*0.00*	0.00	*0.01*	*0.00*
Birds/P/4-4	63	24	2.34	0.57	0.1:0.5/0.5:0.4/1:0.1	0.1:0.5/0.5:0.4/**10:0.1**	*1.00*	0.00	*1.00*	*1.00*
Birds/P/2-1	77	10	2.61	0.30	0.1:0.5/0.5:0.4/1:0.1	0.1:0.5/0.5:0.4/**5:0.1**	*0.75*	0.01	*0.97*	*0.73*
Birds/P/2-2	77	10	2.61	0.30	0.1:0.5/0.5:0.4/1:0.1	0.1:0.5/0.5:0.4/**10:0.1**	*0.99*	0.00	*0.99*	*0.98*
Birds/P/2-3	77	10	2.61	0.30	0.1:0.5/0.5:0.4/1:0.1	0.1:0.45/0.5:0.35/**5:0.2**	*0.91*	0.00	*1.00*	*0.91*
Birds/P/1-1	82	5	2.80	0.11	0.1:0.5/0.5:0.4/1:0.1	0.1:0.5/0.5:0.4/**5:0.1**	*0.20*	0.00	*0.66*	*0.20*
Birds/P/1-2	82	5	2.80	0.11	0.1:0.5/0.5:0.4/1:0.1	0.1:0.5/0.5:0.4/**10:0.1**	*0.35*	0.00	*0.97*	*0.35*
Birds/P/1-3	82	5	2.80	0.11	0.1:0.5/0.5:0.4/1:0.1	0.1:0.45/0.5:0.35/**5:0.2**	*0.46*	0.00	*0.97*	*0.46*

ATP/P/1	35	20	2.11	0.45	0.01:0.983/1:0.017/ 203:0.00115 (E)	0.01:0.993/**5.2:0.007**	*0.81*	0.03	*0.79*	*0.68*
SEMG/P/1	33	8	0.35	1.12	1.00:0.30/**1.24:0.70**	1.00:0.59/**4.58:0.4**/100:0. 0024 (E)	*1.00*	0.34	*0.75*	*0.52*
Prestin/P/1	228	19	3.13	0.20	0.04:0.969/1.06:0.031/ 201:0.00043 (E)	0.09:0.988/**7.15:0.012**	*0.18*	0.00	*0.98*	*0.18*
ATP/N/1	35	20	2.11	0.45	0.01:0.983/1:0.017/ 100:0.0052 (E)	0.01:0.983/1:0.017/ 100:0.0052 (E)	0.00	0.00	0.03	0.00
SEMG/N/1	33	8	0.35	1.12	0.57:0.43/**2.25:0.57**	0.57:0.43/**2.25:0.57**	*0.80*	*0.83*	0.00	0.00
Prestin/N/1	228	19	3.13	0.20	0.05:0.988/**2.6:0.012** / 287(0.00028,E)	0.05:0.988/**2.6:0.012**/ 287 (0.00028,E)	0.00	0.00	0.00	0.00
Birds/N/4-1	63	24	2.34	0.57	0.1:0.5/0.5:0.4/**5:0.1**	0.1:0.5/0.5:0.4/**5:0.1**	*0.98*	*1.00*	0.00	0.00
Birds/N/4-2	63	24	2.34	0.57	0.1:0.5/0.5:0.4/1:0.1	0.1:0.5/0.5:0.4/1:0.1	0.00	0.00	0.00	0.00
Birds/N/4-3	63	24	2.34	0.57	0.25:0.4/0.7:0.4/1.0:0.2	0.1:0.5/0.5:0.4/1:0.1	0.00	0.01	*0.99*	0.00

Table 3. BUSTED-PH performance on data simulated under the correct model (BS-REL), using either the reference tree from Osipova et al. (2024) or Shultz and Sackton (2019) with all (**/4**), two (**/2**) or one (**/1**) nectar consuming clades specified as foreground; or using one of the three datasets from Table 2 as templates for replicate generation. **/P** denote power simulations, **/N** denote null simulations. BG: background, FG: foreground. **Branches**: the number of tree branches in the corresponding set. **Length**: the cumulative length (subs/site) of branches in the corresponding set. **dN/dS distribution**: the distribution of dN/dS applied to the corresponding branch set; values with dN/dS > 1 are shown in bold. **Detection**: the rate of detection (/100 replicates) for the following tests, each at p≤0.05 (following within-alignment Holm-Bonferroni correction); **FG**: EDS on foreground branches, **BG**: EDS on background branches, **Diff**: dN/dS distributions are different between FG and BG. Values in italics are expected to be high, values that are underlined are expected to be low. **PH rate**: the fraction of alignments where the BUSTED-PH composite test holds (EDS on FG, no EDS on BG, significant difference between FG and BG).

Model misspecification.

We analyzed data simulated with `evolver` (Yang 2007) using parameters from (Kowalczyk et al. 2021). As shown in Table 4, BUSTED-PH does not find evidence of association with a trait under the pervasive selection scenario (`PAML/N/1`) and does under the selective regime where only the FG is under EDS (`PAML/P/1`). Thus, it automatically accounts for the non-specificity of selection in the scenario which called for drop out testing in the original manuscript. Furthermore, because the underlying models used for simulations are different from the random effects model of BUSTED-PH, this set of simulations illustrates some degree of robustness.

Scenario	Branches		Length		dN/dS distribution		Detection (p≤0.05)			PH rate
	BG	FG	BG	FG	BG	FG	FG	BG	Diff	
PAML/P/1	10	3	3.85	1.15	0.1:0.1/1.0:0.5	0.1:0.05/1.0:05/**5.0:0.9**	*1.00*	0.00	*1.00*	*1.00*
PAML/N/1	10	3	3.85	1.15	0.1:0.05/1:0.05/**5:0.9**	0.1:0.05/1:0.05/**5:0.9**	*1.00*	*1.00*	0.00	0.00

Table 4. BUSTED-PH performance on data simulated under a model that's different from the BS-REL model, drawn from previous publications (Kowalczyk et al. 2021). **/P** denote power simulations, **/N** denote null simulations. BG: background, FG: foreground. **Branches**: the number of tree branches in the corresponding set. **Length**: the cumulative length (subs/site) of branches in the corresponding set. **dN/dS distribution**: the distribution of dN/dS applied to the corresponding branch set; values with dN/dS > 1

are shown in bold. **Detection**: the rate of detection (/100 replicates) for the following tests, each at p≤0.05 (following within-alignment Holm-Bonferroni correction); **FG**: EDS on foreground branches, **BG**: EDS on background branches, **Diff**: dN/dS distributions are different between FG and BG. Values in italics are expected to be high, values that are underlined are expected to be low. **PH rate**: the fraction of alignments where the BUSTED-PH composite test holds (EDS on FG, no EDS on BG, significant difference between FG and BG).

Data availability

All data and code used for this paper are publicly available, specific alignments used for this analysis are found https://github.com/nclark-lab/ComparativeData/wiki. BUSTED-PH is available as part of the HyPhy package (≥2.5.73) which can be downloaded at https://github.com/veg/hyphy along with additional scripts for processing. Notebooks for data visualization are available at https://observablehq.com/@spond/busted. Data and code will remain accessible to ensure reproducibility.

Discussion

We introduce BUSTED-PH, a branch-site codon model designed to identify genes where the action of episodic diversifying selection (EDS) is associated with a binary trait. BUSTED-PH advances previous similar efforts in convergent evolution detection by explicitly modeling and contrasting positive selection across phenotype-defined foreground (FG) and background (BG) branches. Our results from positive controls, genome-wide scans, and simulation studies collectively demonstrate that BUSTED-PH is a statistically powerful and biologically interpretable tool for detecting trait-associated adaptation.

Application of BUSTED-PH to three canonical genes—Prestin, ATP1A1, and SEMG2—demonstrated the model's ability to recover well-established instances of adaptive evolution linked to convergent phenotypes. All three genes exhibited significant evidence of EDS restricted to trait-associated foreground lineages, where BUSTED-PH detected positive selection on the foreground species, no positive selection detected on background species and statistically significant difference between FG and BG selection profiles. These results validate both the sensitivity and specificity of the three-part BUSTED-PH test and highlight the model's robustness even when alignments contain residual error, as flagged by BUSTED-E (Selberg et al. 2025).

Our genome-wide scan for echolocation-linked adaptation in mammals identified 240 genes with evidence of phenotype-associated episodic diversifying selection (EDS). Several high-ranking candidates—including CDH23, SLC26A5 (Prestin), and TMC1—have well-established roles in

mammalian hearing, lending strong biological support to our findings. Prestin itself was ranked fifth in the entire genome, which demonstrates that BUSTED-PH is a promising method for discovery in other biological systems. Moreover, functional enrichment analysis revealed significant overrepresentation of terms related to sensory perception, olfactory receptor activity, and signal transduction, suggesting shared molecular mechanisms across sensory modalities. These results complement previous work based on directional selection (Allard et al. 2025) by identifying an overlapping but distinct set of genes evolving under episodic positive selection.

Simulation studies confirmed the empirical findings and provided important insights into the performance of BUSTED-PH under controlled conditions. The model maintained low false positive rates across all null scenarios, including challenging cases with mild background selection. Power analyses showed that detection of FG-specific EDS increases with both the number of independent phenotype origins and the strength of selection. Importantly, tests for a difference between FG and BG ω distributions emerged as the most reliable and informative component of the model, often serving as the decisive factor in trait association.

BUSTED-PH fills a methodological gap in the detection of convergent molecular adaptation by offering an integrated framework that (1) explicitly tests for selection in both FG and (2) BG branches, and (3) statistically contrasts their selective regimes. This three-test approach increases interpretability and reduces the risk of spurious trait associations arising from background adaptation or mis-specified phylogenies. The approach also fulfills the need for a branch-site method that assures a gene's selective regime is different on the FG branches and therefore unique to the trait or pressure of interest (Kowalczyk et al. 2021). Previous approaches based on branch-site models alone could not reliably filter genes under positive selection in most species across the phylogeny without respect to the phenotype of interest (Fu et al. 2025). Alternative approaches used combinations of methods to identify positive selection unique to the FG (Pollard et al. 2024; Guo et al. 2025). BUSTED-PH provides the ability to assure FG specificity in one unified and powerful set of models.

Together, our results establish BUSTED-PH as a powerful tool for identifying molecular signatures of phenotypic convergence in a codon-aware framework. Suggested uses of BUSTED-PH are in genome-wide scans to identify genes responsive to convergent selective pressures, imposed by an environment, trait, or other evolutionary challenge specific to a set of branches and species. It is also an appropriate method to test specific hypotheses about functional change in a protein or pathway thought to have influenced a convergent or lineage-specific change. By accounting for selection in both foreground and background lineages and emphasizing statistical contrast, BUSTED-PH improves our ability to distinguish true trait-associated adaptation from confounding evolutionary signals. This model will be valuable for comparative genomic studies seeking to uncover the genetic basis of complex traits in natural systems.

References

Allard JB, Sharma S, Patel R, Sanderford M, Tamura K, Vucetic S, Gerhard GS, Kumar S. 2025. Evolutionary sparse learning reveals the shared genetic basis of convergent traits. *Nat. Commun.* 16:3217.

Armitage DW, Alonso-Sánchez AG, Coy SR, Cheng Z, Hagenbeek A, López-Martínez KP, Phua YH, Sears AR. 2025. Adaptive pangenomic remodeling in the Azolla cyanobiont amid a transient microbiome. *ISME J.* 19:wraf154.

Barkdull M, Moreau CS. 2023. Worker Reproduction and Caste Polymorphism Impact Genome Evolution and Social Genes Across the Ants. *Genome Biol. Evol.* 15:evad095.

Belyantseva IA, Boger ET, Friedman TB. 2003. Myosin XVa localizes to the tips of inner ear sensory cell stereocilia and is essential for staircase formation of the hair bundle. *Proc. Natl. Acad. Sci.* 100:13958–13963.

Berger CA, Steinberg DK, Copeman LA, Tarrant AM. 2025. Comparative analysis of the molecular starvation response of Southern Ocean copepods. *Mol. Ecol.* 34:e17371.

Bolz H, von Brederlow B, Ramírez A, Bryda EC, Kutsche K, Nothwang HG, Seeliger M, del C-Salcedó Cabrera M, Vila MC, Molina OP, et al. 2001. Mutation of CDH23, encoding a new member of the cadherin gene family, causes Usher syndrome type 1D. *Nat. Genet.* 27:108–112.

Bork JM, Peters LM, Riazuddin S, Bernstein SL, Ahmed ZM, Ness SL, Polomeno R, Ramesh A, Schloss M, Srisailpathy CR, et al. 2001. Usher syndrome 1D and nonsyndromic autosomal recessive deafness DFNB12 are caused by allelic mutations of the novel cadherin-like gene CDH23. *Am. J. Hum. Genet.* 68:26–37.

Cicconardi F, McLellan CF, Seguret A, McMillan WO, Montgomery SH. 2025. Convergent Molecular Evolution Associated With Repeated Transitions to Gregarious Larval Behavior in Heliconiini. *Mol. Biol. Evol.* 42:msaf179.

Dhakal RR, Harkess A, Wolf PG. 2025. Chromosome Numbers and Reproductive Life Cycles in Green Plants: A Phylotranscriptomic Perspective. *Plant Direct* 9:e70044.

Dobler S, Dalla S, Wagschal V, Agrawal AA. 2012. Community-wide convergent evolution in insect adaptation to toxic cardenolides by substitutions in the Na,K-ATPase. *Proc. Natl. Acad. Sci.* 109:13040–13045.

Donaldson ML, Barkdull M, Moreau CS. 2025. Comparative genomics analyses reveal selection on neuronal and cuticular hydrocarbon genes is associated with aggression in ants (Hymenoptera: Formicidae). *Ann. Entomol. Soc. Am.* 118:37–58.

Dong S, Li X, Liu Q, Zhu T, Tian A, Chen N, Tu X, Ban L. 2025. Comparative genomics uncovers evolutionary drivers of locust migratory adaptation. *BMC Genomics* 26:203.

Fu R, Yan L, Wang S, Wei D, Wu Q, Liu X, Xiang M. 2025. Genomic Insights Into Convergent Evolution: Adaptation to Rocky Habitats in Rock-Inhabiting Fungi. *Mol. Biol. Evol.* 42:msaf249.

Fukushima K, Pollock DD. 2023. Detecting macroevolutionary genotype–phenotype associations using error-corrected rates of protein convergence. *Nat. Ecol. Evol.* 7:155–170.

Gould E. 1965. Evidence for Echolocation in the Tenrecidae of Madagascar. *Proc. Am. Philos. Soc.* 109:352–360.

Guo X, Jirimutu, Ming L, Wang Z. 2025. Deciphering genetic adaptations of Old World camels through comparative genomic analyses across all camelid species. *iScience* [Internet] 28. Available from: https://www.cell.com/iscience/abstract/S2589-0042(25)00738-2

Halabi K, Karin EL, Guéguen L, Mayrose I. 2021. A Codon Model for Associating Phenotypic Traits with Altered Selective Patterns of Sequence Evolution. *Syst. Biol.* 70:608–622.

Hecker N, Hiller M. 2020. A genome alignment of 120 mammals highlights ultraconserved element variability and placenta-associated enhancers. *GigaScience* 9:giz159.

Holland RA, Waters DA, Rayner JMV. 2004. Echolocation signal structure in the Megachiropteran bat Rousettus aegyptiacus Geoffroy 1810. *J. Exp. Biol.* 207:4361–4369.

Hu Z, Sackton TB, Edwards SV, Liu JS. 2019. Bayesian Detection of Convergent Rate Changes of Conserved Noncoding Elements on Phylogenetic Trees. *Mol. Biol. Evol.* 36:1086–1100.

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

Kopania EEK, Thomas GWC, Hutter CR, Mortimer SME, Callahan CM, Roycroft E, Achmadi AS, Breed WG, Clark NL, Esselstyn JA, et al. 2025. Sperm competition intensity shapes divergence in both sperm morphology and reproductive genes across murine rodents. *Evolution* 79:11–27.

Kosakovsky Pond SL, Frost SDW. 2005. Not So Different After All: A Comparison of Methods for Detecting Amino Acid Sites Under Selection. *Mol. Biol. Evol.* 22:1208–1222.

Kosakovsky Pond SL, Poon AFY, Velazquez R, Weaver S, Hepler NL, Murrell B, Shank SD, Magalis BR, Bouvier D, Nekrutenko A, et al. 2020. HyPhy 2.5—A Customizable Platform for Evolutionary Hypothesis Testing Using Phylogenies. *Mol. Biol. Evol.* 37:295–299.

Kowalczyk A, Chikina M, Clark NL. 2021. A cautionary tale on proper use of branch-site models to detect convergent positive selection. Evolutionary Biology Available from: http://biorxiv.org/lookup/doi/10.1101/2021.10.26.465984

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

Kowalczyk A, Partha R, Clark NL, Chikina M. 2020. Pan-mammalian analysis of molecular constraints underlying extended lifespan.Kaeberlein M, Tautz D, editors. *eLife* 9:e51089.

Li G, Wang J, Rossiter SJ, Jones G, Cotton JA, Zhang S. 2008. The hearing gene Prestin reunites echolocating bats. *Proc. Natl. Acad. Sci.* 105:13959–13964.

Li Y, Liu Z, Shi P, Zhang J. 2010. The hearing gene Prestin unites echolocating bats and whales. *Curr. Biol.* 20:R55–R56.

Liu Z, Qi F-Y, Zhou X, Ren H-Q, Shi P. 2014. Parallel Sites Implicate Functional Convergence of the Hearing Gene Prestin among Echolocating Mammals. *Mol. Biol. Evol.* 31:2415–2424.

Lucaci AG, Zehr JD, Enard D, Thornton JW, Kosakovsky Pond SL. 2023. Evolutionary Shortcuts via Multinucleotide Substitutions and Their Impact on Natural Selection Analyses. *Mol. Biol. Evol.* 40:msad150.

Ludington AJ, Hammond JM, Breen J, Deveson IW, Sanders KL. 2023. New chromosome-scale genomes provide insights into marine adaptations of sea snakes (Hydrophis: Elapidae). *BMC Biol.* 21:284.

Melendez-Vazquez F, Lucaci AG, Selberg A, Clavel J, Rincon-Sandoval M, Santaquiteria A, White WT, Drabeck D, Carnevale G, Duarte-Ribeiro E, et al. 2025. Ecological interactions and genomic innovation fueled the evolution of ray-finned fish endothermy. *Sci. Adv.* 11:eads8488.

Mohammadi S, Herrera-Álvarez S, Yang L, Rodríguez-Ordoñez M del P, Zhang K, Storz JF, Dobler S, Crawford AJ, Andolfatto P. 2022. Constraints on the evolution of toxin-resistant Na,K-ATPases have limited dependence on sequence divergence. *PLOS Genet.* 18:e1010323.

Mulhair PO, Crowley L, Boyes DH, Lewis OT, Holland PWH. 2023. Opsin Gene Duplication in Lepidoptera: Retrotransposition, Sex Linkage, and Gene Expression. *Mol. Biol. Evol.* 40:msad241.

Murrell B, Weaver S, Smith MD, Wertheim JO, Murrell S, Aylward A, Eren K, Pollner T, Martin DP, Smith DM, et al. 2015. Gene-Wide Identification of Episodic Selection. *Mol. Biol. Evol.* 32:1365–1371.

Muse SV, Gaut BS. 1994. A likelihood approach for comparing synonymous and nonsynonymous nucleotide substitution rates, with application to the chloroplast genome. *Mol. Biol. Evol.* 11:715–724.

Nei M, Gojobori T. 1986. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol. Biol. Evol.* 3:418–426.

Onetto CA, Ward CM, Varela C, Hale L, Schmidt SA, Borneman AR. 2025. Genetic and phenotypic diversity of wine-associated Hanseniaspora species. *FEMS Yeast Res.* 25:foaf031.

Osipova E, Ko M-C, Petricek KM, Sin SYW, Brown T, Winkler S, Pippel M, Jarrells J, Weiche S, Mosbech M-B, et al. 2024. Convergent and lineage-specific genomic changes contribute to adaptations in sugar-consuming birds. :2024.08.30.610474. Available from: https://www.biorxiv.org/content/10.1101/2024.08.30.610474v1

Partha R, Chauhan BK, Ferreira Z, Robinson JD, Lathrop K, Nischal KK, Chikina M, Clark NL. 2017. Subterranean mammals show convergent regression in ocular genes and enhancers, along with adaptation to tunneling.Odom DT, editor. *eLife* 6:e25884.

Pollard MD, Meyer WK, Puckett EE. 2024. Convergent relaxation of molecular constraint in herbivores reveals the changing role of liver and kidney functions across mammalian diets. *Genome Res.* 34:2176–2189.

Pond SK, Delport W, Muse SV, Scheffler K. 2010. Correcting the Bias of Empirical Frequency Parameter Estimators in Codon Models. *PLOS ONE* 5:e11230.

Prudent X, Parra G, Schwede P, Roscito JG, Hiller M. 2016. Controlling for Phylogenetic Relatedness and Evolutionary Rates Improves the Discovery of Associations Between Species' Phenotypic and Genomic Differences. *Mol. Biol. Evol.* 33:2135–2150.

Redlich R, Kowalczyk A, Tene M, Sestili HH, Foley K, Saputra E, Clark N, Chikina M, Meyer WK, Pfenning AR. 2024. RERconverge Expansion: Using Relative Evolutionary Rates to Study Complex Categorical Trait Evolution. *Mol. Biol. Evol.* 41:msae210.

Ruesink-Bueno IL, Drews A, O'Connor EA, Westerdahl H. 2024. Expansion of MHC-IIB Has Constrained the Evolution of MHC-IIA in Passerines. *Genome Biol. Evol.* 16:evae236.

Sackton TB, Grayson P, Cloutier A, Hu Z, Liu JS, Wheeler NE, Gardner PP, Clarke JA, Baker AJ, Clamp M, et al. 2019. Convergent regulatory evolution and loss of flight in paleognathous birds. *Science* 364:74–78.

Selberg A, Clark NL, Sackton TB, Muse SV, Lucaci AG, Weaver S, Nekrutenko A, Chikina M, Pond SLK. 2025. Minus the Error: Testing for Positive Selection in the Presence of Residual Alignment Errors. *eLife* [Internet] 14. Available from: https://elifesciences.org/reviewed-preprints/106921

Self SG, Liang K-Y. 1987. Asymptotic Properties of Maximum Likelihood Estimators and Likelihood Ratio Tests Under Nonstandard Conditions. *J. Am. Stat. Assoc.* 82:605–610.

Shultz AJ, Sackton TB. 2019. Immune genes are hotspots of shared positive selection across birds and mammals.Landry CR, Tautz D, editors. *eLife* 8:e41815.

Singh A, Pope NS, López-Uribe MM. 2025. Shifts in bee diet breadths are associated with gene gains and losses and positive selection across olfactory receptors. *G3 GenesGenomesGenetics* 15:jkaf105.

Szklarczyk D, Kirsch R, Koutrouli M, Nastou K, Mehryary F, Hachilif R, Gable AL, Fang T, Doncheva NT, Pyysalo S, et al. 2023. The STRING database in 2023: protein–protein association networks and functional enrichment analyses for any sequenced genome of interest. *Nucleic Acids Res.* 51:D638–D646.

Wertheim JO, Murrell B, Smith MD, Kosakovsky Pond SL, Scheffler K. 2015. RELAX: Detecting Relaxed Selection in a Phylogenetic Framework. *Mol. Biol. Evol.* 32:820–832.

Wisotsky SR, Kosakovsky Pond SL, Shank SD, Muse SV. 2020. Synonymous Site-to-Site Substitution Rate Variation Dramatically Inflates False Positive Rates of Selection

Analyses: Ignore at Your Own Peril. *Mol. Biol. Evol.* 37:2430–2439.

Yang Z. 2007. PAML 4: Phylogenetic Analysis by Maximum Likelihood. *Mol. Biol. Evol.* 24:1586–1591.

Yusuf LH, Saldívar Lemus Y, Thorpe P, Macías Garcia C, Ritchie MG. 2023. Genomic Signatures Associated with Transitions to Viviparity in Cyprinodontiformes. *Mol. Biol. Evol.* 40:msad208.

Zhu M, Yang T, Wei S, DeWan AT, Morell RJ, Elfenbein JL, Fisher RA, Leal SM, Smith RJH, Friderici KH. 2003. Mutations in the gamma-actin gene (ACTG1) are associated with dominant progressive deafness (DFNA20/26). *Am. J. Hum. Genet.* 73:1082–1091.

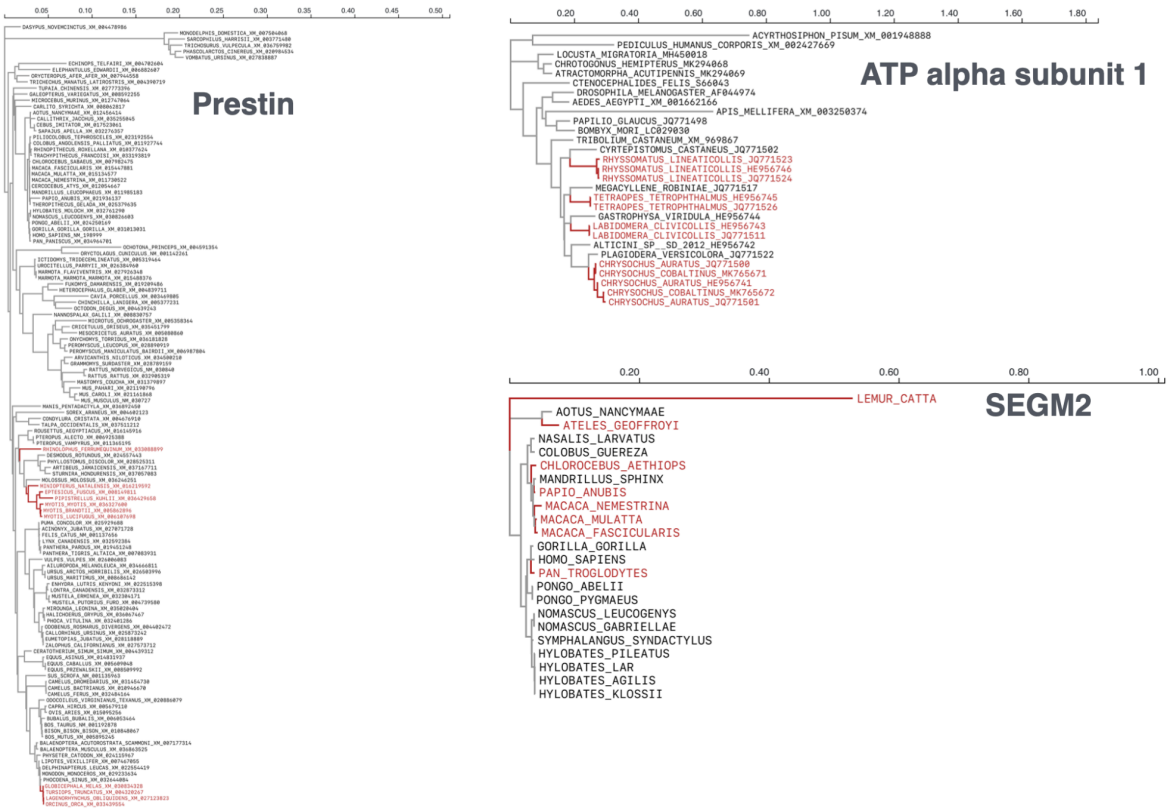

Figure S1. Phylogenetic trees with foreground branches shown in red for the three empirical datasets in Table 2.

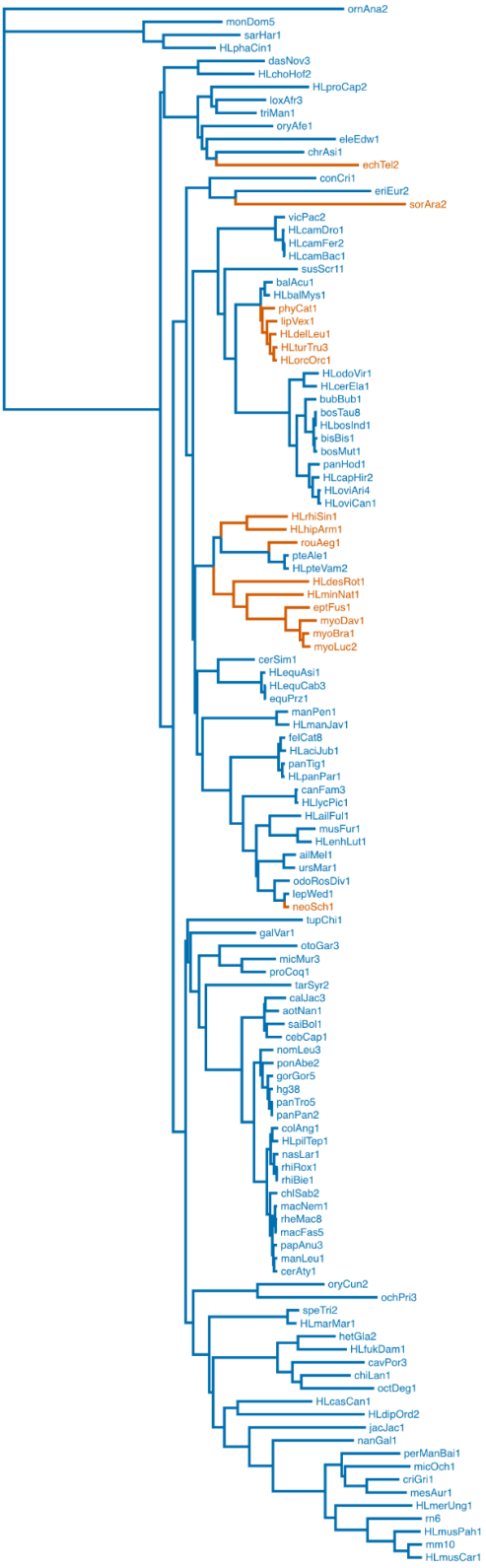

Figure S2. Phylogenetic trees for echolocating mammal study, with foreground species shown in orange.

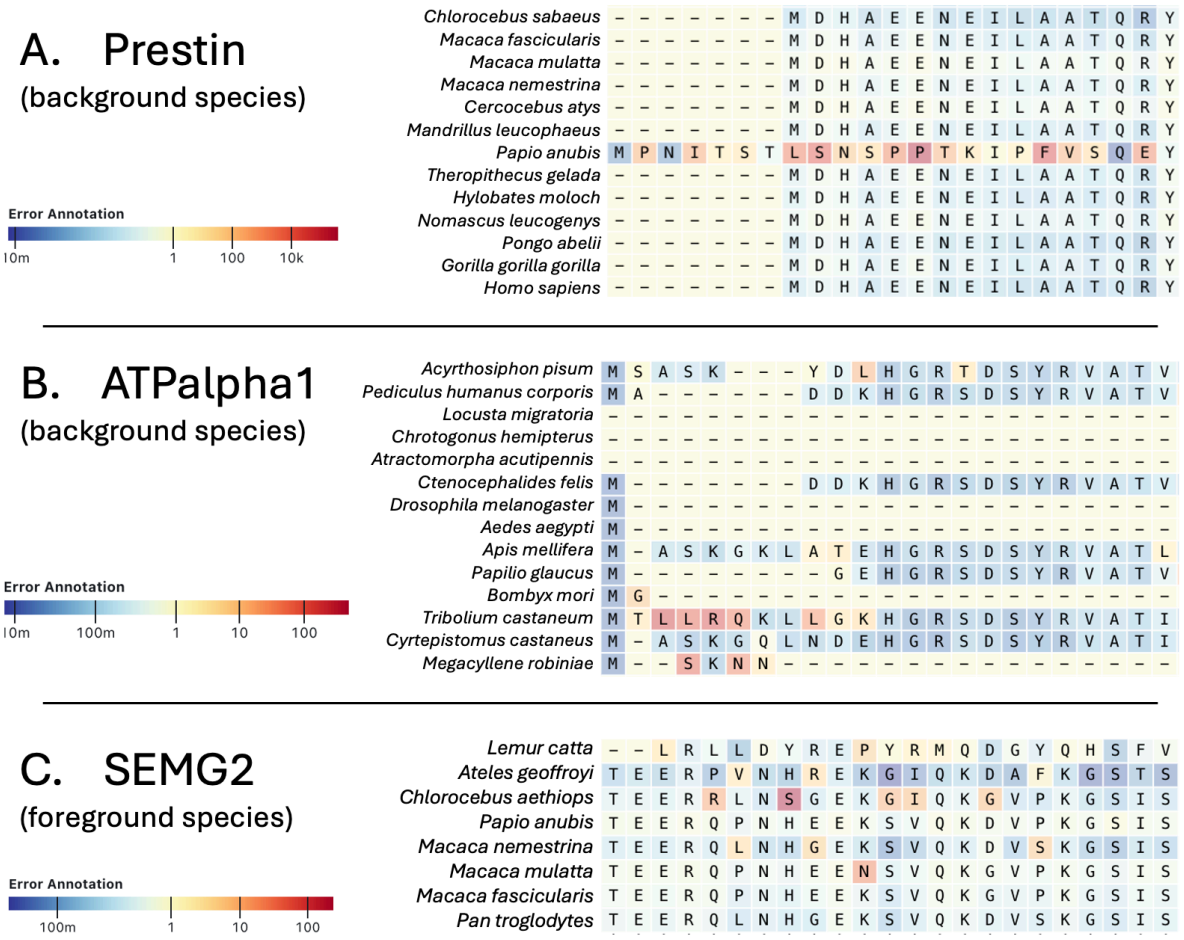

Figure S3. Residual alignment errors detected by BUSTED-PH in three positive-control genes. Prestin and ATPalpha1 (A, B) contain localized misaligned regions on background lineages, flagged by elevated error annotation values. SEMG2 (C) shows more scattered individual foreground codons inferred as highly unlikely substitutions. These examples reflect the residual errors present in the empirical alignments used for BUSTED-PH analyses.

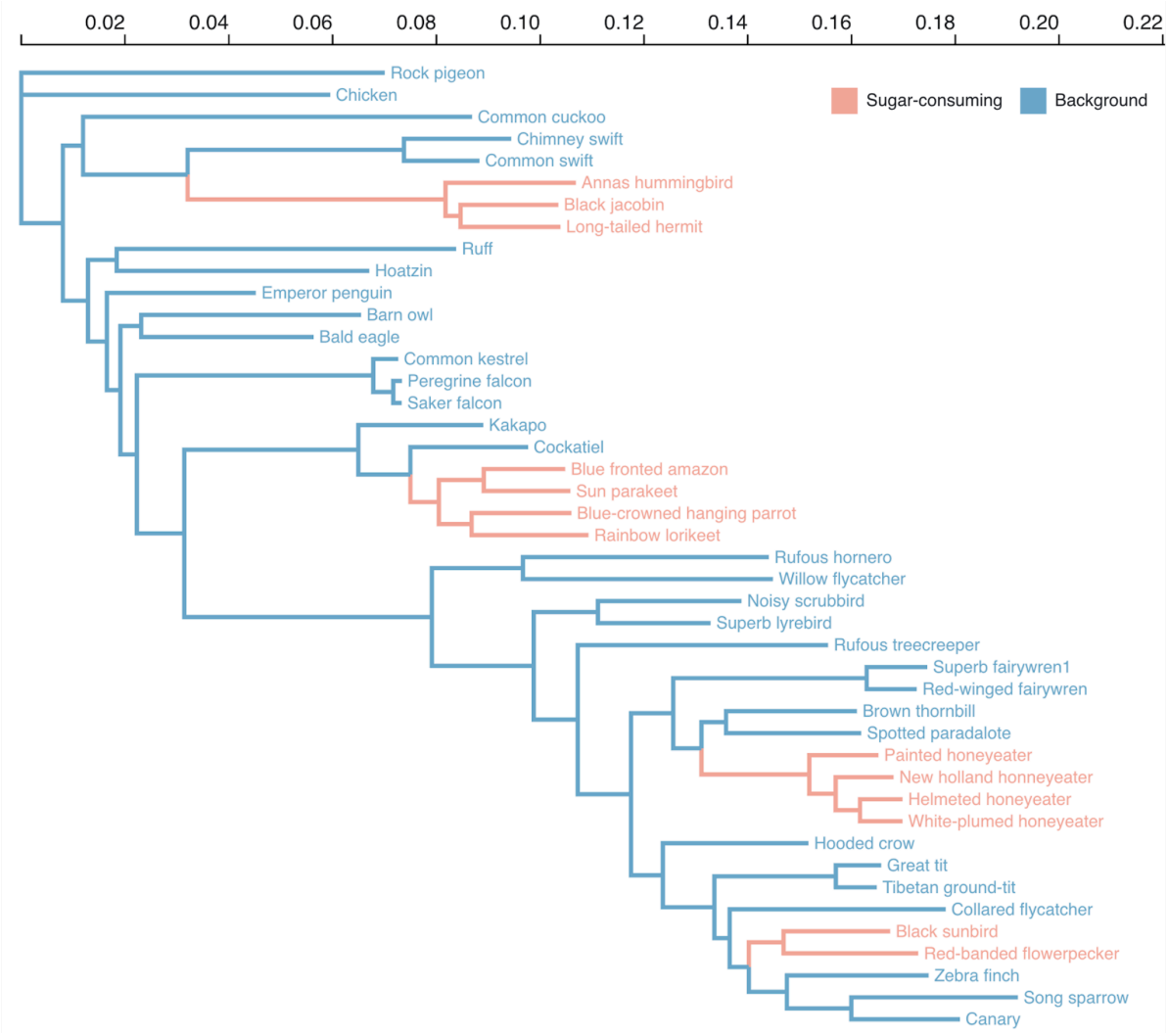

Figure S4. Tree used for power simulations (Table 3), from Osipova et al. (2024).

Study	Organisms	Trait	No. of Genes tested
Ludington et al. (2023)	Snakes	Marine adaptation	8654
Dong et al. (2025)	Locusts	Migratory behavior	6874
Cicconardi et al. (2025)	Heliconiini butterflies	Social behavior	3393
Dhakal et al. (2025)	Green plants	Chromosome number	54 gene families
Ruesink-Bueno et al. (2024)	Birds	Passerine vs nonpasserine birds	1; DAA1
Onetto et al. (2025)	Yeast	Wine fermentation	2036
Armitage et al. (2025)	Symbiotic cyanobacteria	Azolla symbionts vs free-living bacteria	3520
Barkdull et al. (2023)	Ants	worker caste polymorphism and reproductive capacity	32792
Singh et al. (2025)	Bees	Diet breadth	251
Berger et al. (2025)	Copepods	Oil sac used to sequester lipids	13640
Yusuf et al. (2023)	Ray-finned fishes	Viviparity vs oviparity	1044
Mulhair et al. (2023)	Lepidoptera	Nocturnal vs diurnal eye genes	3
Kopania et al. (2025)	Murine rodents	Relative testes mass	11775
Donaldson et al. (2025)	Ants	Aggression and nestmate discrimination	16233

Table S1. Studies that have implemented BUSTED-PH to test for lineage or trait-associated positive selection (excluding self-citations).

Gene	FG p-value	BG p-value	DIFF p-value	ω_+	P[ω_+]	Ranking
CDH23	0.0000	0.5000	0.0000	9.43	0.01	34.36
ACTBL2	0.0000	0.5000	0.0000	4.42	0.05	31.94
ACE	0.0000	0.5000	0.0000	7.35	0.02	30.91
COL6A6	0.0000	0.0538	0.0000	6.35	0.02	30.11
SLC26A5	0.0000	0.1677	0.0000	11.55	0.01	28.81
ITPRID1	0.0000	0.2605	0.0000	20.10	0.04	28.77
DSP	0.0000	0.1315	0.0000	2.19	0.05	27.60
RUVBL1	0.0001	0.5000	0.0000	7.65	0.02	27.58
KNG1	0.0000	0.3269	0.0000	4.09	0.17	27.27
BCLAF3	0.0001	0.5000	0.0000	4.03	0.14	26.68
LOXHD1	0.0000	0.0919	0.0000	3.52	0.03	26.64
OR10X1	0.0000	0.0930	0.0000	29.63	0.07	26.11
ITGAD	0.0000	0.1020	0.0000	3.87	0.08	25.93
MAGED4B	0.0000	0.1633	0.0000	7.35	0.03	25.87
KARS1	0.0004	0.5000	0.0000	7.29	0.02	25.58
FFAR3	0.0004	0.5000	0.0000	5.78	0.05	25.52
EPC2	0.0005	0.5000	0.0000	28.53	0.01	25.33
HSD17B13	0.0005	0.5000	0.0000	4.73	0.09	25.25
SRRM1	0.0006	0.5000	0.0000	42.52	0.00	25.23
GPR42	0.0015	0.5000	0.0000	6.84	0.04	24.22
ZFHX4	0.0017	0.5000	0.0000	10.61	0.00	24.10
PHGDH	0.0017	0.5000	0.0000	2.59	0.06	24.10
ZNF783	0.0021	0.5000	0.0000	3.02	0.11	23.87
ZNF778	0.0005	0.1199	0.0000	15.35	0.08	23.72
TMEM63C	0.0027	0.5000	0.0000	14.17	0.01	23.66
TMPRSS11D	0.0004	0.5000	0.0000	1.93	0.31	23.65
SAMHD1	0.0027	0.5000	0.0000	9.04	0.02	23.64
AP2B1	0.0015	0.2130	0.0000	18.38	0.00	23.40
TMC1	0.0036	0.5000	0.0000	8.49	0.02	23.34
F11	0.0020	0.2437	0.0000	11.51	0.02	23.23
PCDHGB1	0.0051	0.5000	0.0000	20.57	0.01	23.01
DNAH10	0.0063	0.5000	0.0000	7.80	0.00	22.79
GUCY2C	0.0070	0.5000	0.0000	2.92	0.06	22.69
NDST4	0.0071	0.5000	0.0000	8.39	0.01	22.68
ADAMTS6	0.0073	0.5000	0.0000	11.20	0.01	22.64
PRPF18	0.0080	0.5000	0.0000	2.51	0.07	22.55
DPP4	0.0095	0.5000	0.0000	4.29	0.05	22.38

TNN	0.0079	0.3816	0.0000	4.50	0.02	22.30
RIBC1	0.0106	0.5000	0.0000	11.04	0.03	22.28
TRPV2	0.0024	0.1115	0.0000	5.97	0.02	22.26
MOV10	0.0117	0.5000	0.0000	5.93	0.01	22.18
CYP4X1	0.0078	0.3337	0.0000	4.73	0.08	22.17
DNHD1	0.0125	0.5000	0.0000	4.28	0.02	22.11
ADCY5	0.0137	0.5000	0.0000	8.82	0.01	22.02
ZFP30	0.0147	0.5000	0.0000	31.30	0.02	21.95
SF3B1	0.0150	0.5000	0.0000	27.22	0.00	21.93
H1-4	0.0096	0.3064	0.0000	5.38	0.03	21.89
ITGA4	0.0152	0.4661	0.0000	15.67	0.00	21.85
PHLDB1	0.0134	0.3458	0.0000	5.69	0.01	21.67
KRT83	0.0162	0.4042	0.0000	20.01	0.01	21.63
FRK	0.0201	0.5000	0.0000	4.82	0.05	21.63
UGT8	0.0240	0.5000	0.0000	5.29	0.03	21.46
MYH1	0.0004	0.3594	0.0000	4.57	0.02	21.43
NR1I3	0.0310	0.5000	0.0000	3.69	0.08	21.20
KRTAP5-3	0.0000	0.1628	0.0000	449.77	0.07	21.12
DNAH17	0.0368	0.5000	0.0000	1.58	0.04	21.03
CFAP57	0.0373	0.5000	0.0000	5.40	0.01	21.02
PALM	0.0077	0.1036	0.0000	3.38	0.06	21.02
SDK1	0.0230	0.2859	0.0000	4.05	0.01	20.94
ACE2	0.0113	0.1311	0.0000	3.51	0.08	20.87
CYP2S1	0.0273	0.5000	0.0000	8.83	0.01	20.86
POLR2B	0.0370	0.4201	0.0000	6.94	0.00	20.85
PARP14	0.0000	0.5000	0.0007	7.16	0.04	20.85
GPR108	0.0442	0.5000	0.0000	6.21	0.02	20.85
TAX1BP1	0.0445	0.5000	0.0000	6.67	0.01	20.84
GRIN2A	0.0017	0.5000	0.0000	8.15	0.01	20.81
MYH3	0.0468	0.5000	0.0000	4.31	0.01	20.79
ATF6	0.0399	0.3425	0.0000	5.40	0.01	20.57
CAMSAP2	0.0353	0.2796	0.0000	8.45	0.01	20.49
BPIFA1	0.0122	0.0919	0.0000	4.65	0.11	20.44
ITIH3	0.0249	0.1781	0.0000	5.21	0.04	20.39
TIGD6	0.0203	0.1381	0.0000	13.17	0.02	20.34
COL4A2	0.0149	0.0898	0.0000	2.73	0.06	20.21
ALOXE3	0.0268	0.5000	0.0000	6.87	0.01	20.19
PCDHA11	0.0030	0.4851	0.0000	14.02	0.01	20.14
MAP3K15	0.0147	0.0759	0.0000	3.53	0.08	20.06
AGRN	0.0244	0.1477	0.0000	20.95	0.00	19.89

USP8	0.0004	0.5000	0.0000	6.26	0.02	19.74
MMP2	0.0342	0.3059	0.0000	3.34	0.02	19.48
TMEM132D	0.0261	0.0658	0.0000	6.75	0.01	19.35
KLHL10	0.0007	0.4917	0.0000	7.29	0.01	19.23
MLKL	0.0000	0.2701	0.0002	5.17	0.12	19.16
TBX4	0.0025	0.5000	0.0000	8.20	0.01	19.04
SYNPO2	0.0285	0.1422	0.0000	13.43	0.01	18.74
LCE1D	0.0419	0.5000	0.0000	3.94	0.12	18.41
ZNF599	0.0149	0.5000	0.0000	5.35	0.04	18.33
MYT1L	0.0135	0.5000	0.0000	16.58	0.00	18.25
FTHL17	0.0224	0.3379	0.0000	2.32	0.20	18.05
LCE1C	0.0103	0.1761	0.0000	50.06	0.03	18.02
SIGLECL1	0.0012	0.1966	0.0000	9.86	0.07	17.95
LMNTD2	0.0328	0.5000	0.0000	4.06	0.06	17.85
MAP1B	0.0042	0.5000	0.0000	4.41	0.01	17.85
CRH	0.0212	0.0597	0.0000	8.14	0.04	17.80
FBXO43	0.0000	0.5000	0.0020	4.73	0.05	17.77
FREM2	0.0002	0.5000	0.0001	3.94	0.01	17.75
PNMA6A	0.0002	0.1385	0.0000	13.04	0.02	17.69
RFWD3	0.0096	0.1440	0.0000	7.74	0.02	17.55
PSKH2	0.0005	0.1591	0.0000	18.21	0.03	17.35
CEP164	0.0228	0.5000	0.0000	3.17	0.07	17.25
MMP12	0.0288	0.4992	0.0000	2.23	0.26	17.13

Table S2. List of top genes from BUSTED-PH showing evidence for trait-associated EDS. From left to write: Gene ID name, FG, BG, and Difference between FG and BG p-values, omega values for positive selection class, proportion of sites with positive selection omega value, and ranking determined by Ranking score = $-\log_{10}(\max(10^{-8}, pFG)) - \log_{10}(\max(10^{-8}, pDIFF)) + \log_{10}(\max(10^{-8}, pBG))$.