UC Irvine

UC Irvine Previously Published Works

Title

Drift and Directional Selection Are the Evolutionary Forces Driving Gene Expression Divergence in Eye and Brain Tissue of Heliconius Butterflies

Permalink

https://escholarship.org/uc/item/5501q3zh

Journal

Genetics, 213(2)

ISSN

0016-6731

Authors

Catalán, Ana
Briscoe, Adriana D
Höhna, Sebastian

Publication Date

2019-10-01

DOI

10.1534/genetics.119.302493

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Drift and Directional Selection Are the Evolutionary Forces Driving Gene Expression Divergence in Eye and Brain Tissue of *Heliconius* Butterflies

Ana Catalán,*,†,1 Adriana D. Briscoe,‡ and Sebastian Höhna†,§,,1**

*Department of Evolutionary Biology, Evolutionary Biology Centre (EBC), Uppsala University, 75236, Sweden, †Division of Evolutionary Biology, Ludwig-Maximilians-Universität München, Planegg-Martinsried 82152, Germany, ‡Department of Ecology and Evolutionary Biology, University of California, Irvine, California 92697, §Department of Earth and Environmental Sciences, Paleontology and Geobiology, and **GeoBio-Center, Ludwig-Maximilians-Universität München, 80333 Munich, Germany

ABSTRACT Investigating gene expression evolution over micro- and macroevolutionary timescales will expand our understanding of the role of gene expression in adaptation and speciation. In this study, we characterized the evolutionary forces acting on gene expression levels in eye and brain tissue of five *Heliconius* butterflies with divergence times of ~5–12 MYA. We developed and applied Brownian motion (BM) and Ornstein–Uhlenbeck (OU) models to identify genes whose expression levels are evolving through drift, stabilizing selection, or a lineage-specific shift. We found that 81% of the genes evolve under genetic drift. When testing for branch-specific shifts in gene expression, we detected 368 (16%) shift events. Genes showing a shift toward upregulation have significantly lower gene expression variance than those genes showing a shift leading toward downregulation. We hypothesize that directional selection is acting in shifts causing upregulation, since transcription is costly. We further uncovered through simulations that parameter estimation of OU models is biased when using small phylogenies and only becomes reliable with phylogenies having ≥ 50 taxa. Therefore, we developed a new statistical test based on BM to identify highly conserved genes (*i.e.*, evolving under strong stabilizing selection), which comprised 3% of the orthoclusters. In conclusion, we found that drift is the dominant evolutionary force driving gene expression evolution in eye and brain tissue in *Heliconius*. Nevertheless, the higher proportion of genes evolving under directional than under stabilizing selection might reflect species-specific selective pressures on vision and the brain that are necessary to fulfill species-specific requirements.

KEYWORDS Brownian motion; natural selection; stabilizing selection; Ornstein–Uhlenbeck; RevBayes

Species and populations diverge through the accumulation of genetic changes that affect coding or non-coding genomic regions that Genetic variation affecting gene expression has the potential of changing gene expression patterns in a spatiotemporal manner by changing gene expression profiles in specific organs and cell types at particular developmental stages (Carroll 2005; Signor and Nuzhdin 2018). This spatiotemporal attribute of gene expression might enable evolutionary change in a compartmentalized manner, allowing for change where it is required but also allowing for the needed processes to remain conserved. Phenotypic diversity caused by changes in gene expression encompasses a great variety of traits, including changes affecting an organism's coloration (Nadeau 2016), size, and shape (Ahi *et al.* 2017), as well as sensory perception and behavior, among other phenotypes (Lee *et al.* 2000; Wanner *et al.* 2007). Even though major advances have been made in linking gene expression variation to a phenotype (Catalán *et al.* 2016; Glaser-Schmitt and Parsch 2018), discerning the evolutionary forces that shape gene expression level variation among closely related species is an area that needs further research.

To understand the evolutionary forces acting on gene expression it is necessary to model within- and between-species

Copyright © 2019 Catalán *et al.*
doi: https://doi.org/10.1534/genetics.119.302493
Manuscript received July 8, 2019; accepted for publication August 24, 2019; published Early Online August 29, 2019.
Available freely online through the author-supported open access option.
This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
Supplemental material available at FigShare: https://doi.org/10.25386/genetics.9736253.
[1]Corresponding authors: LMU Biozentrum Department Biologie II, Ludwig-Maximilians-Universität München, Großhaderner Str. 2, Planegg-Martinsried, Munich 82152, Germany. E-mail: ana.catalan@gmail.com; and sebastian.hoehna@gmail.com

gene expression variance. Neutral gene expression divergence between species leads to gene expression differences through divergence alone. Thus, neutral changes in gene expression provide a null hypothesis to detect deviations from the expected neutral gene expression divergence. A linear relationship between divergence time and gene expression variance has been proposed for closely related species, assuming a clock-like (*i.e.*, constant through time) rate of gene expression divergence (Khaitovich *et al.* 2004, 2005a). Another approach to study the evolutionary forces shaping gene expression evolution, which is motivated by statistical phylogenetics, is fitting Brownian motion (BM) models. BM models are often used to describe the rate of change of continuous traits through time taking into account the known phylogeny of the taxa of interest (Felsenstein 1985). Thus, in a BM context, the parameter σ^2 is often described as the volatility parameter that determines the rate at which a trait's value diffuses away from its current state (Bedford and Hartl 2009). Fitting BM models to investigate gene expression evolution has shown that stabilizing selection and evolution through drift can be readily characterized (Kalinka *et al.* 2010; Wong *et al.* 2015).

Ornstein–Uhlenbeck (OU) models have also been used to study continuous trait evolution in a phylogenetic context (Hansen 1997; Butler and King 2004). OU models, an extension to BM models, include two extra parameters, α and θ. As in a BM context, σ^2 is the rate at which a trait changes through time and the parameter α is the force pulling back the diffused trait to an optimum state. This is analogous to stabilizing selection pulling back a trait to its optimum value after having experienced a departure from it. θ is described as the trait's optimum state at a particular time point toward which the pull of α is aimed (Hansen 1997; Butler and King 2004). OU models offer a useful framework to generate hypotheses about the evolutionary forces acting on transcriptome levels, whether it is drift, stabilizing selection, or directional selection (Bedford and Hartl 2009; Rohlfs and Nielsen 2015; Wong *et al.* 2015; Chen *et al.* 2017; Stern and Crandall 2018).

In this study, we used five closely related species of *Heliconius* butterflies to explore the evolutionary forces shaping gene expression variation in combined eye and brain tissue. *Heliconius charithonia*, *H. sara*, *H. erato*, *H. melpomene*, and *H. doris* (Figure 1) belong to four of the seven distinct *Heliconius* phylogenetic clades with divergence times ranging from 5.5 to 11.8 MYA. Beside showing great diversity in wing color patterns (Kronforst and Papa 2015), *Heliconius* butterflies also show diversity in life history traits (Salcedo 2010; Merrill *et al.* 2015), mating systems (Beltrán *et al.* 2007; Walters *et al.* 2012), and behavior (Mendoza-Cuenca and Macías-Ordóñez 2005; Merrill *et al.* 2019). Since *Heliconius* butterflies are diurnal species, visual stimuli provide key sources of information about the environment. For example, flowers and oviposition sites, potential mates, or predators are all targets of interest to butterflies in which the first line of perception is visual (Finkbeiner *et al.* 2014, 2017). After

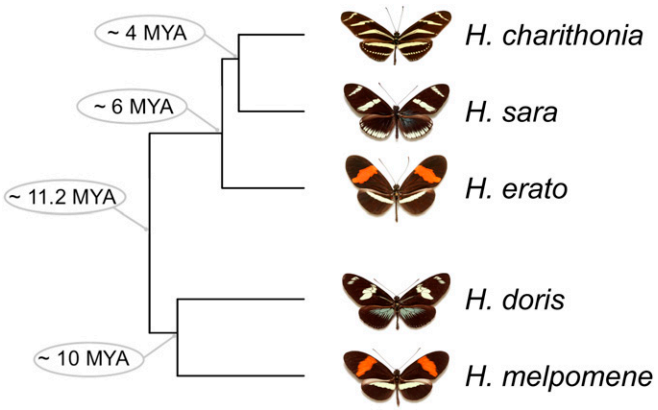

Figure 1 Phylogenetic relationship of the *Heliconius* species used in this study showing divergence times at each node (Kozak *et al.* 2015).

visual cues are detected by the visual system, the detected information travels to the brain where it is processed, and its output can result in a specific behavior or physiological response. Thus, the brain's processing and output together with the visual system have the potential of being finely tuned according to a species' life history. In the case of *Heliconius* butterflies, high diversity of adult compound eye retinal mosaics (between sexes and species) has been described (McCulloch *et al.* 2016, 2017), as well as species-specific differences in brain morphology (Montgomery *et al.* 2016). Which evolutionary forces are shaping adult eye and brain expression in *Heliconius* is one question we seek to investigate, and in that way, gain an understanding of the potential role of interspecies gene expression differences in speciation and adaptation.

Therefore, in this study we investigated which evolutionary forces are driving gene expression variation in combined eye and brain tissue. More specifically, we aimed to identify if expression variation in individual genes is evolving, for example, through drift, stabilizing selection, or directional selection. To this end, we generated a set of orthoclusters shared among our five butterfly species together with expression data for each gene in each orthocluster. We characterized the selective forces acting on gene expression levels, thereby revealing the fraction of the transcriptome evolving under drift, directional selection, and stabilizing selection.

Materials and Methods

Data set, samples, and RNA-sequencing

De novo transcriptome assemblies for *H. charithonia*, *H. sara*, and *H. doris* were downloaded from Dryad with data identifier doi: 10.5061/dryad.ds21fv5 (Catalán *et al.* 2018). Transcriptomes from *H. erato* and *H. melpomene* were downloaded from Dryad with data identifier doi: 10.5061/dryad.8d724 (Smith *et al.* 2016). RNA-sequencing (RNA-seq) data for all the species were downloaded from ArrayExpress: E-MTAB-6810 (Catalán *et al.* 2018). The RNA-seq data were generated as follows: pupae were obtained from The

Table 1 Summary of the models implemented in this work

Model	Description	Parameters
Equal species means	All species have the same mean gene expression level	μ: global mean gene expression level
Unequal species means	All species have their own independent mean gene expression level	μ_i: mean gene expression level per species
Brownian motion	Random drift of the species' mean gene expression level along the phylogeny	σ^2: rate of drift
Brownian motion with shift	Random drift with one branch having a different rate (directional selection)	σ^2_B: rate of drift background branch σ^2_F: rate of drift foreground branch
Ornstein–Uhlenbeck	Stabilizing selection of the species mean gene expression level evolving along the phylogeny	σ^2: rate of drift α: strength of selection θ: optimal gene expression level
Ornstein–Uhlenbeck with shift	Directional selection due to a shift in optimal gene expression level	σ^2_B: rate of drift background branch σ^2_F: rate of drift foreground branch. θ_B: optimal gene expression level background branch θ_F: optimal gene expression level foreground branch α: strength of selection

Butterfly Farm, Costa Rica Entomological Supply and allowed to eclose under controlled laboratory conditions.

Biological replicates for females (F) and males (M) for each species were generated: *H. charithonia* (F = 6, M = 6), *H. sara* (F = 5, M = 6), *H. erato* (F = 3, M = 3), *H. doris* (F = 6, M = 6), and *H. melpomene* (F = 4, M = 4). Three days after eclosion, the butterflies were flash frozen at −80° until RNA extraction. Combined eye and brain tissue was dissected from the same individual by removing the antennae, palps, and proboscis from the head capsule. RNA was extracted using TRIzol (Thermo Fisher Scientific, Waltham, MA) following the manufacturer's instructions. RNA was purified using a NucleoSpin RNA II kit (Macherey-Nagel, Bethlehem, PA). Purified RNA was quantified using a Qubit 2.0 Fluorometer (Thermo Fisher Scientific) and RNA integrity was checked using an Agilent Bioanalyzer 2100 (Agilent Technologies, Santa Clara, CA), with RNA integrity number (RIN) ranging from 8 to 10. RNA-seq libraries were prepared for each individual using a TruSeq RNA Sample Preparation Kit v2 (Illumina, San Diego, CA). Double-stranded cDNA libraries were quantified, quality checked, normalized, and pooled for sequencing according to their concentrations. Pooled libraries were run on a 2% agarose gel to size select fragments of ~240–600 bp. cDNA was recovered using a Geneclean III kit (MP Biomedical, Santa Ana, CA) and purified using Agencourt AMPure XP beads (Beckman Coulter, Brea, CA). Sequencing was conducted at the University of California, Irvine Genomics High-Throughput Facility using a HiSeq 2500 (Illumina), paired end 100-cycle sequence run.

Reads were mapped to their corresponding transcriptome using Bowtie (Langmead *et al.* 2009), and FPKM (fragments per kilobase per million mapped read) values were calculated for each species and used for downstream analysis (Catalán *et al.* 2018). Annotation files for each transcriptome were downloaded from doi: 10.5061/ dryad.ds21fv5 (Catalán *et al.* 2018) with the exception of the transcriptome annotation for *H. erato*, which was newly annotated. For the annotation of the *H. erato* transcriptome, TransDecoder (version 5.0.2) was used to identify candidate coding regions. The predicted coding sequences were utilized to identify orthologous hits in the UniProt, FlyBase, and Pfam databases using blastp (2.2.30) (Altschul *et al.* 1990; Chintapalli *et al.* 2007; Punta *et al.* 2012) (Appendix I).

Orthology assessment

Orthology across species was determined by using the Unrooted Phylogenetic Orthology (UPhO) pipeline (Ballesteros and Hormiga 2016). UPhO uses an all-species pairwise blastp search and a Markov clustering algorithm (MCL) (version 1.0.0) (Enright *et al.* 2002) to cluster sequences according to sequence similarity. Clustered sequences were aligned with MAFFT (version 7.3.05) (Katoh and Toh 2008) and curated after alignment with trimAl (version 1.3) (Capella-Gutiérrez *et al.* 2009). Phylogenetic inference for each sequence cluster was done using RAxML (version 8.2.10) (Stamatakis 2006) and orthology was assessed for each generated tree using the UPhO algorithm. A matrix with $\log_2$ FPKM values was generated for each orthocluster, which was used to analyze and model gene expression variance (Appendix I).

Modeling gene expression evolution

To study the forces driving gene expression evolution, we implemented a set of six different statistical models (Table 1). Each one models the mean species gene expression level (between-species variance) and the gene expression levels of individual samples per species (within-species variance). How these mean species gene expression levels evolve, or not, along the phylogeny and over time, is specific and central to each model. We estimated the parameters of each model and performed Bayesian model selection using Bayes factors to establish which model describes the observed data best, and thus which process is most likely to drive gene expression evolution in the five *Heliconius* species of our study (see below).

The simplest model of gene expression assumes that all species have the exact same mean gene expression level. In this case, we only have one parameter, μ, which defines the

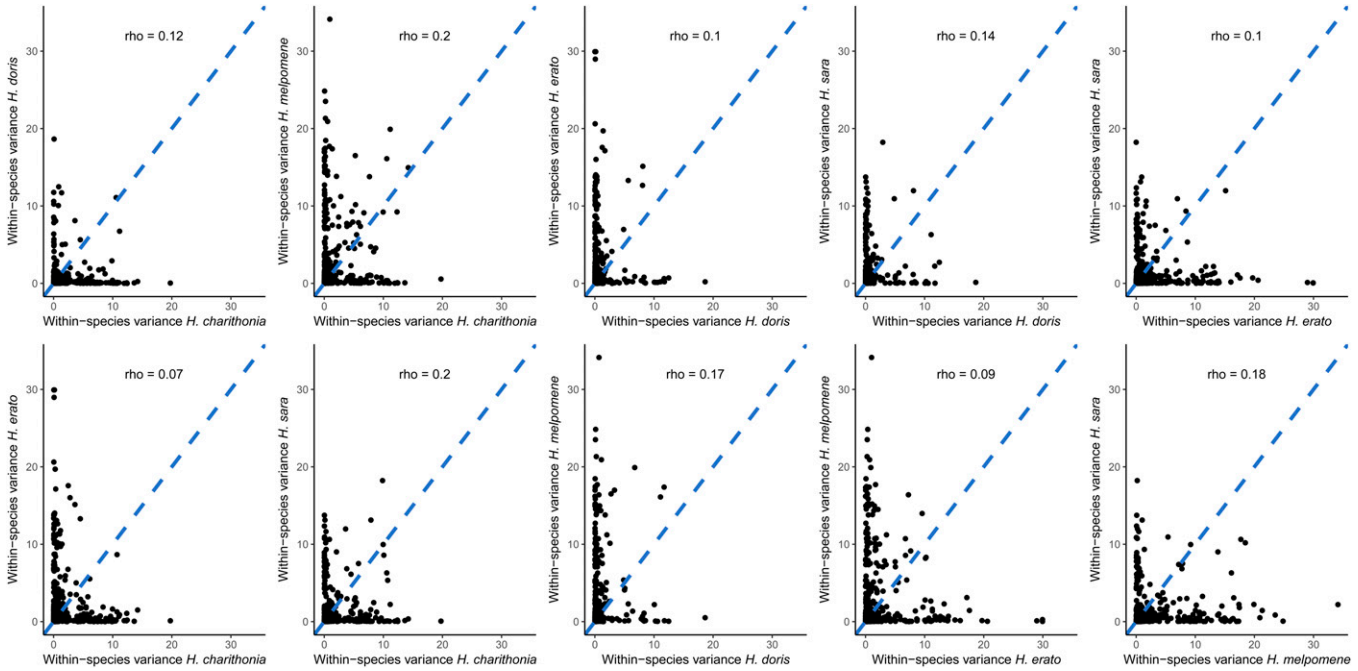

Figure 2 Pairwise correlation between five *Heliconius* species and their respective per gene expression variances. The correlation strength between per gene expression variances was estimated by calculating Pearson's ρ correlation coefficient.

mean gene expression level of all species. The expression level X_{ij} of individual i from species j is modeled using a normal distribution with $X_{ij} \sim \text{Norm}(\mu, \delta_i^2)$. We chose a uniform prior distribution between -20 and $+20$ for the mean gene expression parameter μ. Note that we assume that every species has its own gene expression variance parameter δ_i^2 (see below). This model assumes that there is no evolution of gene expression levels, *i.e.*, gene expression levels are completely conserved among species.

The second model that we implemented was a model where each species has its own gene expression mean μ_i. Thus, we modeled the gene expression level X_{ij} of gene i from species j using a normal distribution with $X_{ij} \sim \text{Norm}(\mu_i, \delta_i^2)$. In this model, each species has a different mean gene expression level, but these gene expression levels do not evolve under an evolutionary model; they are intrinsically different without any mechanistic reason (no phylogenetic signal). As with the first model, we assumed a uniform prior distribution between -20 and $+20$ for each mean gene expression level μ_i.

The third model that we implemented was a phylogenetic BM model (Felsenstein 1985). We assume that any gene expression value at the root of the phylogeny is equally probable. Then, the mean gene expression levels μ evolve along the lineages of the phylogeny. The BM model specifies that the focal variable, μ in our case, is drawn from a normal distribution centered around the value of the ancestor, μ_A. The amount of change, *i.e.*, the rate of random drift, is defined by the parameter σ^2. We assumed a log-uniform prior distribution between 10E−5 and 10E5 for the drift parameter σ^2. Thus, the mean gene expression levels μ_i for the species of the phylogeny are distributed according to a multivariate normal distribution where the covariance structure is defined by the phylogeny (Felsenstein 1985). This means that more closely related species are expected to have a more similar mean gene expression level because they share more evolutionary history (*i.e.*, they are more recently diverged), which is modeled by the covariance structure. Such a phylogenetic model of gene expression evolution has been applied previously by Bedford and Hartl (2009). Importantly, the BM model only defines how the mean gene expression levels evolve but does not allow for any sample variance of the individuals of a species. Therefore, we extended the standard phylogenetic BM model to allow for within-species sample variance, where again the expression level X_{ij} of individual i from species j is normally distributed with $X_{ij} \sim \text{Norm}(\mu_i, \delta_i^2)$ where δ_i^2 is the within-species variance parameter (Ives *et al.* 2007; Rohlfs and Nielsen 2015). This extension to allow for within-species variance was developed for all phylogenetic models (BM, BM with shift, OU, and OU with shift).

The fourth model that we implemented was a phylogenetic BM model with branch-specific rates of evolution, thus detecting directional selection. The mean gene expression level evolves under a BM model (*i.e.*, random drift) where the rate of evolution for branch k is given by σ_k^2 (O'Meara *et al.* 2006; Eastman *et al.* 2011). Thus, a branch with a higher rate of evolution σ_k^2 signifies more change in gene expression levels than under a constant rate random drift model (the BM model). Directional selection can therefore be detected by inferring an elevated estimate of σ_k^2 compared with the background rate of drift σ^2. Specifically, we applied a background

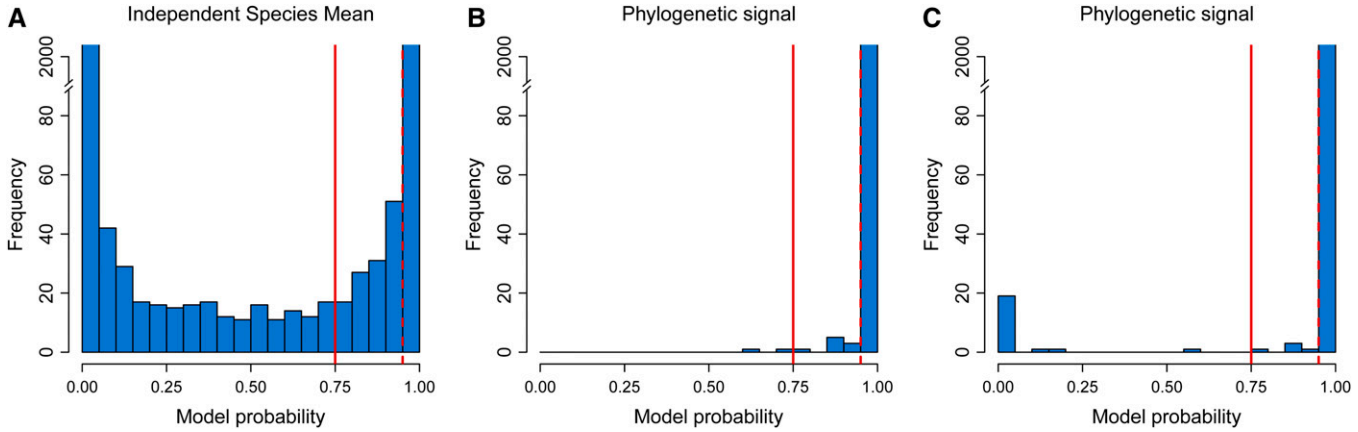

Figure 3 Testing for a phylogenetic signal in gene expression levels of *Heliconius* using BM. Significance is shown at model probability > 0.75 (solid red line, Bayes factor > 3, positive support) and model probability > 0.95 (dashed red line, Bayes factor > 20, strong support). (A) Shows the comparison between the two nonphylogenetic models (identical *vs.* independent species mean). (B) Shows the model probability of the BM model compared with the independent species mean model. (C) Shows the model probability of the BM model compared with the identical species mean model. BM, Brownian motion.

rate of drift σ^2_B to all branches except the chosen foreground branch, which received its own rate of drift σ^2_F.

The fifth model we implemented was a phylogenetic OU process (Hansen 1997). Similar to BM, the OU process models the evolution of the mean gene expression level per species along a phylogenetic tree. However, unlike BM, the mean expression level diffuses with rate σ^2 and is attracted with strength α to an optimum level θ. Thus, the OU process has an expected variance of $\sigma^2/2\alpha$ that is independent of time, *i.e.*, does not increase with increasing time but instead stabilizes. The variance becomes small if either the strength of selection is large or the rate of drift is small. This is, in fact, a major problem for OU models, which cannot distinguish if attraction (or selection) is strong or diffusion is weak (Ho and Ané 2014; Cooper *et al.* 2016).

The sixth model we implemented was an OU process with a branch-specific shift in both the rate of drift σ^2 and the optimum gene expression level θ (Rohlfs *et al.* 2014; Uyeda and Harmon 2014). Thus, this branch-specific OU model is analogous to the branch-specific BM model, allowing for directional selection in an OU framework. Specifically, we tested if there was significant support for the chosen foreground branch, which received its own rate of drift σ^2_F and optimum θ_B to be different from the background rate of drift σ^2_B and optimum θ_B. We used the same prior distributions as before, and assumed that both parameters for the background and foreground branches are drawn from the identical prior distribution. This model has, in total, five free additional parameters along with the five nuisance parameters (the within-species variances). Thus, we expect that this model is more prone to be overparameterized for our data set with five species. Nevertheless, our Bayesian approach for parameter estimation and model selection integrates over parameter uncertainty, and penalizes extra parameters by integrating over the prior distribution.

Parameter estimation and model selection

We estimated parameters for our different models in a Bayesian statistical framework. Thus, we approximated the posterior distribution of the model parameters using Markov chain Monte Carlo (MCMC) sampling (Metropolis *et al.* 1953; Hastings 1970). We ran a separate MCMC analysis for each model and gene, 2393 analyses per model. Every model parameter was updated twice per MCMC iteration where the order of parameter updates was chosen randomly. We applied the same settings of the MCMC algorithm for each model. First, a burn-in phase of the MCMC algorithm was run for 2000 iterations with auto-tuning every 100 iterations. Then, the actual MCMC simulation was run for 50,000 iterations with sampling 10 iterations, yielding 5000 samples from the posterior distribution (Höhna *et al.* 2017).

Model selection was performed using marginal likelihoods. Marginal likelihoods are the probability of the data for a specific model integrated over all possible parameter values. From the marginal likelihood we can then compute Bayes factors and model probabilities (*i.e.*, weights of a model being the true model generating the data given a set of candidate models). We approximated the marginal likelihoods using stepping-stone sampling (Fan *et al.* 2011). The stepping-stone algorithm implemented in RevBayes consisted of 128 MCMC runs, where each MCMC run had the likelihood function raised to the power of β computed by the quantiles of a β probability distribution (Höhna *et al.* 2017).

Data availability

The five different models that we used in our study are implemented in Bayesian phylogenetic inference software RevBayes v1.0.8 (Höhna *et al.* 2016). For efficient computation, we implemented the restricted maximum likelihood algorithm for BM (Felsenstein 1985) and OU models (FitzJohn 2012; Freckleton 2012). The source code and compiled

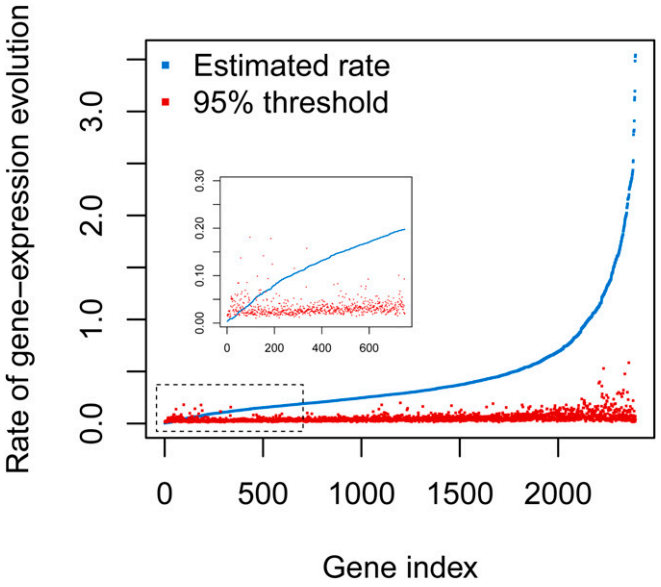

Figure 4 Posterior mean estimates of the rate of gene expression change (σ^2 in blue) and the 95% threshold computed (red) using Monte Carlo simulations. The genes were sorted by an ascending estimate of σ^2. Inset: close-up of genes whose σ^2 is not significantly bigger than zero.

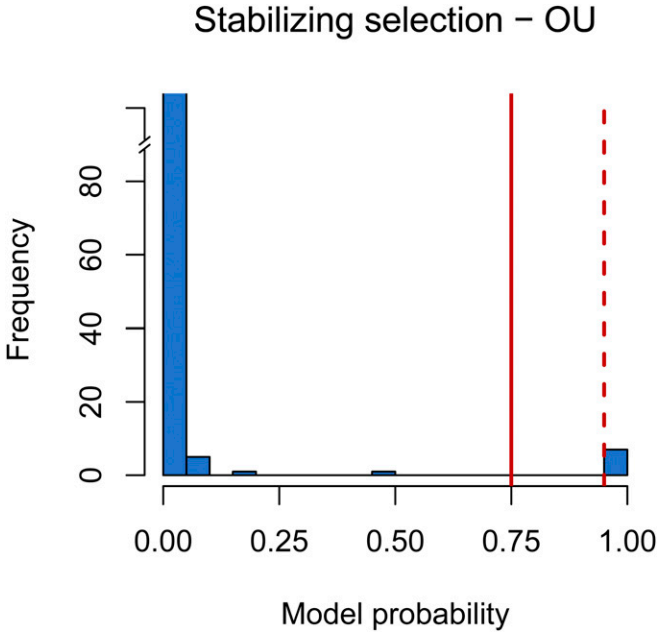

Figure 5 Model probability when testing model suitability when fitting an OU model for the assessment of stabilizing selection. Significance is shown at model probability > 0.75 (solid red, Bayes factor > 3, positive support) and model probability > 0.95 (solid red, Bayes factor > 20, strong support). There are only seven genes with significant support for stabilizing selection through an OU model. OU, Ornstein–Uhlenbeck.

executables of RevBayes are available from https://github.com/revbayes/revbayes, and tutorials about the analyses are available from https://revbayes.github.io/tutorials/. Supplemental material available at FigShare: https://doi.org/10.25386/genetics.9736253.

Results

One of our main objectives in this study was to detect the evolutionary forces acting on gene expression by identifying deviations of gene expression variation from the variation expected from phylogeny alone. With this aim, we used transcriptomic data from combined eye and brain tissue of five *Heliconius* species (*H. charithonia*, *H. sara*, *H. erato*, *H. doris*, and *H. melpomene*) (Figure 1). From orthologous genes present and expressed in the five species, we built an expression matrix using $\log_2$-transformed FPKM values, which formed a data set of 2373 orthologous expressed genes (Appendix I). We used this data set to investigate the evolutionary forces shaping gene expression variation across species.

Testing for equality of within-species variance in gene expression levels

Previous models that assess the evolutionary forces acting on gene expression levels (Warnefors and Eyre-Walker 2012; Rohlfs *et al.* 2014) assume equal gene expression variance across species. Assuming equality of variance when it is not the case can lead to high type I error rates (Gastwirth *et al.* 2009). To test for equality of variance of gene expression across our five species, we calculated the within-species variance in gene expression. We calculated the within-species gene expression variance by calculating the gene expression variance from the mean, as measured in FKPMs, for each gene (as we have 6–12 biological replicates per species) and tested for a correlation of gene expression variance between every species pair (Figure 2).

In Figure 2, each black dot represents the value of the variance for each gene of the *Heliconius* species presented on either of the axes. From this pairwise assessment, we found no significant correlation among all possible pairs, with Pearsons's ρ values ranging from 0.07 to 0.2 (Figure 2). Since gene expression variance across species is heterogeneous, and hence not correlated across species, we treated per gene expression variance as a random variable when fitting BM and OU models.

Testing for a phylogenetic signal

First, we started checking whether we could detect a phylogenetic signal from our gene expression data by using BM models. To test for this, we used two nonphylogenetic models where either all species had identical mean gene expression levels (model 1) or all species had their own independent mean gene expression levels (model 2). For each gene, we computed the probability of the BM model having produced the observed data, *i.e.*, a high probability means that it is more probable that the gene expression levels evolved under a BM model, whereas a low probability means that it is more probable that the gene expression levels evolved under a nonphylogenetic model (model 1 and model 2). A model probability of >0.75 corresponds to a Bayes factor of >3 (positive support) and a model probability of >0.95 corresponds to a Bayes factor of >20 (strong support) (Figure 3).

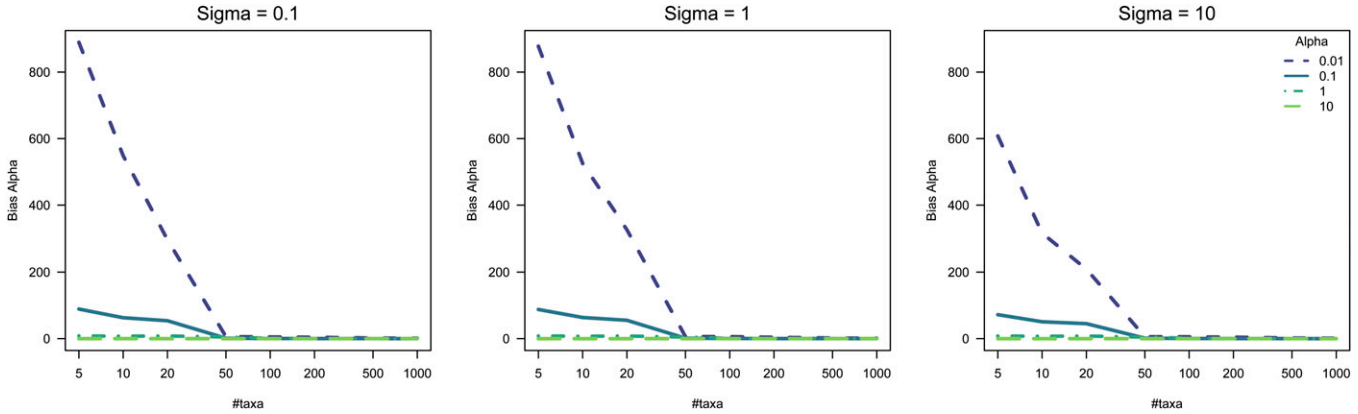

Figure 6 Simulation study for the assessment of parameter estimation bias under an OU model. The relative bias in estimates of the attraction/selection parameter (α) through 1000 simulations under σ values ranging from 0.1 to 10, and α values ranging from 0.01 to 10. Simulations were performed for phylogenies with sizes ranging from 5 to 1000 taxa. OU, Ornstein–Uhlenbeck.

Our results show that for the majority of gene expression levels (2369 out of 2393) a phylogenetic signal can be recovered. Since RNA-seq data have high sensitivity to experimental and environmental noise, gene expression levels are prone to strong stochastic changes. The identification of a phylogenetic signal in most genes shows that BM models are suited to the investigation of the evolutionary forces acting on our gene expression data set.

Testing for conserved gene expression levels

The next question we explored was how prevalent conserved gene expression levels are in combined eye and brain tissue of *Heliconius*. This question could be answered with our previous results by computing how often model 1, with identical species means, was recovered (Figure 2C). However, our model selection procedure relied on computing marginal likelihoods, which are intrinsically sensitive to the choice of prior distribution (Berger 1990; Kass and Raftery 1995; Sinharay and Stern 2002). Therefore, we additionally performed a sensitivity analysis of $\sigma^2 = 0$ using Monte Carlo simulation as follows (Goldman 1993). First, we estimated the posterior distribution of all parameters under the identical species mean model (the only parameters were the per gene expression variances), then we used 1000 parameter samples from the posterior distribution to simulate gene expression data sets (*e.g.*, a data set consisting of a single gene with five species and 6–12 individuals per species) under the identical species mean model, yielding 1000 simulated data sets per gene in total. Then, for every gene of the 2393 genes, we estimated σ^2 for each simulated data set as well as the original data set, which amounted to a total of 2,395,393 MCMC analyses. Finally, we calculated if the mean posterior estimate of the empirical data set was $>$95% of the mean posterior estimates of the simulated data sets. In cases where the mean posterior estimate of σ^2 was not larger than the mean estimate of 95% of the simulated data sets, we concluded that these genes are highly conserved (Figure 4).

By using the described approach, we uncovered a set of 83 orthoclusters whose rates of gene expression evolution (Figure 4) and gene expression levels (Supplemental Material, Figure S1) across species are highly conserved. A σ^2 not significantly different from zero can be caused by stabilizing selection hindering gene expression divergence, resulting in more similar gene expression patterns across different *Heliconius* species than expected.

Testing for stabilizing selection acting on gene expression levels

Subsequently, we moved forward into implementing an OU to investigate the strength of stabilizing selection (Bedford and Hartl 2009; Beaulieu *et al.* 2012; Rohlfs *et al.* 2014). OU models include two extra parameters, α and θ. In a BM context, if σ^2 is the rate at which a trait changes through time, α is then described as a force pulling back the diffused trait to an optimum state (θ).

We estimated the marginal likelihood for each gene under a BM model and an OU model. Then, we computed the probability (*i.e.*, support) of an OU model over a BM model using the marginal likelihoods. Our results show very low support for stabilizing selection under an OU model (Figure 5). When the marginal likelihoods were examined, in 99.7% of the cases a BM model explained our expression data better than an OU model.

Testing the power to estimate stabilizing selection

Our results indicate that very few genes in *Heliconius* combined eye and brain tissue are evolving under stabilizing selection. It has previously been discussed that when working with small phylogenies ($<$10 species) there is a lack of power for parameter estimation when using an OU model (Beaulieu *et al.* 2012; Rohlfs *et al.* 2014), but no simulation studies have been done. By simulating data under an OU model using phylogenies with varying numbers of taxa, we were able to show how parameter estimation is biased. The attraction parameter α could only be estimated closely to the true values

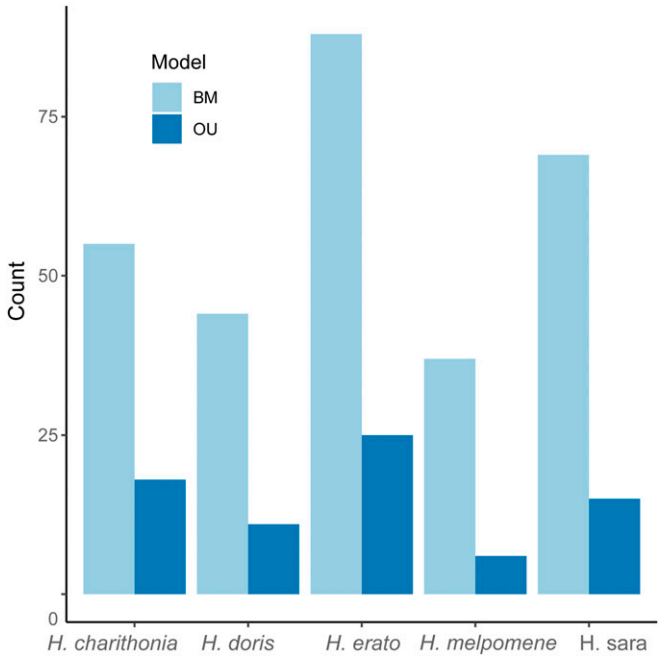

Figure 7 Bar plot showing branch-specific shifts on gene expression levels in *Heliconius*. Bars in light blue show branch shifts identified by BM models and dark blue bars show branch shifts identified by OU models. BM, Brownian motion; OU, Ornstein–Uhlenbeck.

used for the simulations when the phylogenies contained ≥50 taxa. Thus, we can observe that the bias observed for parameter estimation drops considerably when the number of taxa composing the phylogeny reaches 50 (Figure 6). This observation holds as well for the estimation of σ^2 under a range of σ values (Figures S2 and S3). Our simulation study shows that attention needs be paid when applying OU models to assess gene expression evolution for phylogenies containing <50 taxa.

Detection of branch-specific shifts in gene expression

To reveal genes whose gene expression patterns have putatively been shaped by directional selection, we tested for branch-specific shifts in evolutionary rates along the *Heliconius* tree. To explore branch-specific shifts in gene expression, we first used a BM model to test for the evolutionary rate (σ^2) of a focal branch being different from the background rate (*i.e.*, the rest of the branches in the phylogenetic tree) and assessed significance by applying Bayes factors (Figure 7 and Figure S4). Second, we also tested branch-specific shifts through an OU model and tested for a branch-specific shift in gene expression level optimum (θ_F) *vs.* the rest of the tree's θ_B.

With a BM approach, we were able to detect a total of 322 branch-specific shifts when considering only tip branches (Figure 7). We found 112 branch-specific shifts in the *H. erato* linage, 70 in *H. sara*, 67 in *H. charithonia*, 44 in *H. doris*, and 29 in *H. melpomene* (Figure 7 and Figure S5). *H. charithonia*, *H. sara*, and *H. doris* had more shifts toward a downregulation, although only in *H. charithonia* and *H. sara* was this difference significant (sign test, *H. charithonia:* *P*-value 6.738e−05 and *H. sara*: *P*-value 1.653e−06). In *H. erato* and *H. melpomene*, more upregulated genes were causing a branch-specific shift, although no significant difference was found.

When implementing an OU model, we recovered a total of 75 genes showing a branch-specific shift in gene expression optimum (Figure 7). From these genes, 55 also show a branch-specific shift when implementing a BM model and 20 genes show uniquely a gene expression-level shift in optimum when using an OU model (Figure S6).

Next, we assessed gene expression variance of all the genes identified as having a branch-specific shift in gene expression through BM and OU models. When we plotted the distribution of the gene expression variance, we found that upregulated genes have a significantly lower variance when compared to genes with a gene expression shift toward downregulation (Figure 8).

Discussion

Gene expression evolution through genetic drift

Our study on the evolutionary forces acting on gene expression in combined eye and brain tissue of *Heliconius* butterflies reveals that drift is the dominant evolutionary force driving gene expression divergence (81% of the transcriptome). According to neutral expectations, phenotypic changes are expected to accumulate as a function of time, by drift and mutation alone (Lande 1976). Following this rationale, the change of transcriptomic levels through drift should reflect the divergence history of the taxa of interest. From our BM analysis, we show that in most of the gene expression levels in combined eye and brain, a phylogenetic signal can be recovered (Figure 3). Nevertheless, we have to keep in mind that a phylogenetic signal can also be recovered even if other evolutionary forces are acting on the transcriptome, such as stabilizing selection or directional selection (Harmon *et al.* 2010). There are other evolutionary scenarios, beside drift, which resemble a random walk as modeled by BM. For example, directional selection that varies in strength and direction in a random fashion from one generation to the next can also be described as having a BM behavior. Similarly, for strong stabilizing selection, when the trait's optimum changes randomly, It can also be described by BM. For example, in a study investigating pulsed evolution in vertebrates, the authors included BM to model incremental phenotypic change, where the trait of interest followed a wandering optimum (Landis and Schraiber 2017). Thus, drift, randomly varying selection, and varying stabilizing selection can be modeled by BM.

Evolutionary rates of gene expression, which have been investigated at both population and species level, show that the proportion of the type of evolutionary force acting on transcriptomic levels is not constant across taxa (Whitehead and Crawford 2006a; Landis and Schraiber 2017; Nourmohammad *et al.* 2017; Stern and Crandall

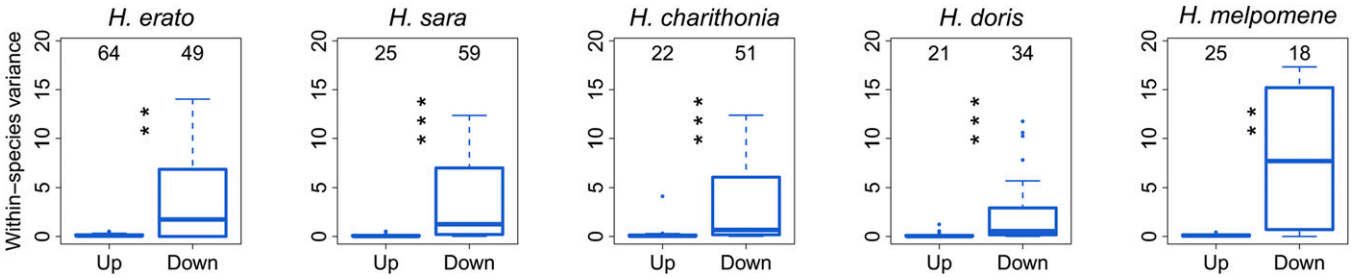

Figure 8 The gene expression variances for all the genes showing a shift toward an up- and a downregulation are depicted as box plots for each species. Numbers above the box plots show the total number of genes identified with a BM and an OU model. Wilcoxon test: * $P < 0.05$, ** $P < 0.01$, and *** $P < 0.001$. BM, Brownian motion; OU, Ornstein–Uhlenbeck.

2018). For example, when examining the evolutionary forces acting on gene expression levels in several fish populations, the authors reported that the dominant force driving expression changes was genetic drift (Whitehead and Crawford 2006b). Comparably, in studies concerning primates, genetic drift was the main force driving gene expression evolution (Khaitovich *et al.* 2005b; Chaix *et al.* 2008). The proportion of gene expression levels evolving by drift depends on the strength of natural selection acting on the interrogated transcriptome. For example, in a comparison between different organ types in mammals, gonad gene expression showed the lowest phylogenetic signal when compared to other organs like cerebellum or heart (Brawand *et al.* 2011). In *Heliconius* butterflies, other organs would need to be tested to get a more global understanding of how gene expression is evolving in the whole organism.

We further explored our expression data by comparing the expected gene expression divergence under a BM model to the observed data. Consequently, we simulated expression levels for 10,000 genes along the known *Heliconius* phylogeny and computed the mean of the pairwise species differences. Similarly, we computed the mean pairwise differences of the observed gene expression data. Alternatively, we can also derive the expected divergence in gene expression levels between two species over time under BM. Both species evolve under random drift and, thus, their gene expression values are normally distributed with variance $\sigma^2 \times T$, where T is the time since the most recent common ancestor of the two species. Therefore, the difference in gene expression levels between the two species is normally distributed with variance $2 \times \sigma^2 \times T$. Since we are only interested in the absolute value of the gene expression difference, we use a truncated normal distribution instead. From this truncated normal distribution with mean zero and variance $2 \times \sigma^2 \times T$, we compute the expected gene expression difference through time (Figure 9). For the empirical data, we estimate σ^2 using a sum of squares approach. We find that our observed gene expression data have a close fit to the simulated data (Figure 9).

Gene expression evolution through stabilizing selection

Studies done in *Drosophila* and mammals have shown that stabilizing selection is the main evolutionary force driving gene expression evolution (Rifkin *et al.* 2003; Lemos *et al.* 2005; Rohlfs and Nielsen 2015). In contrast to these studies, in *Heliconius*, we discovered that only 3% of gene expression levels are either highly conserved (Figure 4) or are evolving through stabilizing selection (Figure 5). Factors such as tissue type, gene functionality turnover, or epistatic levels have the potential to influence the degree of stabilizing selection acting on the transcriptome (Larracuente *et al.* 2008; Kalinka *et al.* 2010; Romero *et al.* 2012). Additionally, in groups that have experienced an adaptive radiation, such as *Heliconius* butterflies (Kozak *et al.* 2015), and have thus recently experienced an elevated rate of trait evolution, directional selection might be more recurrent than stabilizing selection. Factors such as the evolutionary history, the topology of the phylogenetic tree (*e.g.*, the depth of the phylogeny), and the type of continuous trait being studied will determine if a model describing drift or stabilizing selection describes the data best (Fay and Wittkopp 2008).

OU models are suitable models to study the force of stabilizing selection acting on a phenotype since the α parameter simulates the strength of selection keeping a trait close to an optimum (Beaulieu *et al.* 2012), as several studies exemplify (Kalinka *et al.* 2010; Brawand *et al.* 2011; Stern and Crandall 2018). When we applied an OU model to identify stabilizing selection on gene expression, we detected parameter estimation biases as shown by our simulation study (Figure 6). For small phylogenies, accurate parameter estimation is challenging since statistical power is weak with small sample sizes (Rohlfs *et al.* 2014), and parameters like α and σ^2 tend to be overestimated (Beaulieu *et al.* 2012). Specifically, it is very challenging with small phylogenies to distinguish between conserved gene expression levels due to low values of drift (*i.e.*, no change) *vs.* high values of directional selection (*i.e.*, drift is removed due to selection) (Appendix II). Therefore, we propose that for small phylogenies, testing for $\sigma^2 = 0$ under a BM framework and assessing for significance by applying Monte Carlo simulations is a better approach to uncover stabilizing selection. From our estimates on the rate of gene expression evolution (σ^2 ranging from $\sim$0 to 9) and using Monte Carlos simulations to test for a σ^2 significantly different from zero, we show that for 88% of the data we have a false discovery rate $\leq$ 5%. Thus, the

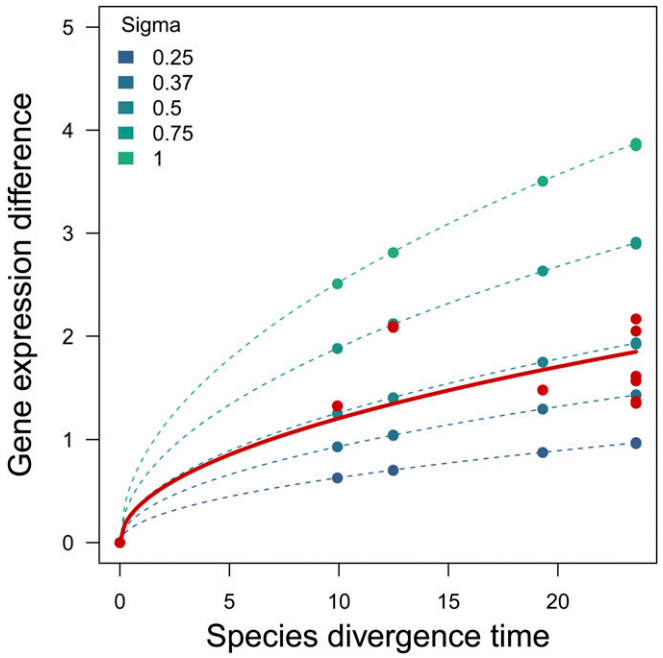

Figure 9 Between-species gene expression divergence plotted as a function of divergence time according to the *Heliconius* phylogeny. Red: σ^2 from gene expression levels observed in *Heliconius*. Blue: simulated gene expression divergence under random drift with different values of σ.

likelihood that our Monte Carlo simulation approach for assessing conserved genes is reporting a false positive is very low for larger values of the rate of evolution (Appendix II).

When using this approach, we identified 83 genes with conserved gene expression levels across species. These genes might be involved in maintaining conserved processes that are essential for the function of eye and brain tissue in *Heliconius*. For example, from the top 10 genes with the most conserved gene expression levels, we found the transcription factor *bobby sox* (*bbx*) (Group_674, Appendix I). BBX belongs to the high-mobility box domain superfamily, which is involved in transcription, replication, and chromatin remodeling (Chintapalli *et al.* 2007). BBX has also been found to have orthologs in flies, humans, and mice (Nitta *et al.* 2015), suggesting high essentiality of *bbx* expression. Another highly conserved orthocluster (Group_977, Appendix I) was annotated as *glaikit* (Chintapalli *et al.* 2007), which is known to be essential for the formation of epithelial polarity and nervous system development (Dunlop *et al.* 2004).

To complement our BM approach, we further explored our data by simulating 10,000 genes under an OU model under a range of σ^2 and α values, and computed the mean differences between pairs of species. We observed a reasonably good fit to our data (Figure 10), but it is worth pointing out that a steeper change in gene expression divergence is observed between closely related species when compared to the calculated expected divergence. This suggests that different evolutionary scenarios might explain the data better at different depths of the phylogeny, but could also be a characteristic signal of our data set. Consequently, adding more species, including those that are closely related, could not only improve OU parameter estimation but could also help disentangle the evolutionary forces acting on gene expression divergence, particularly between closely related species. Additional to the implementation of BM and OU models to individual genes (as we have done in this study), investigating how whole-gene networks or groups of coexpressed genes are evolving following BM/OU models along a phylogeny will increase the power to detect deviations to neutral expectation in small phylogenies (Schraiber *et al.* 2013). This approach was implemented in yeast where genes belonging to the same gene pathway were jointly analyzed, resulting in the identification of pathways with constrained and accelerated gene expression evolution, even when using small phylogenies. Such an approach could be explored with our data set once we have identified the genes that are evolving together because they form part of the same pathway.

Gene expression evolution through genetic directional selection

To reveal branch-specific shifts in gene expression levels, we applied BM and OU models allowing for branch-specific shifts at the tips of the phylogeny. Using this approach, we found that 16% of the genes show a branch-specific shift, toward either up- or downregulation, with increased expression levels showing lower variance than expected (Figure 8, Figure S5, and Appendix II for false discovery rate analysis). The direction of a gene expression shift might be influenced by its mode of regulation. For example, in yeast, it was found that regulatory mutations affecting *trans*-regulatory factors were more likely to cause an increase in gene expression. Conversely, mutations in *cis*-regulatory elements were found to be skewed toward a decrease in expression (Metzger *et al.* 2016). On the other hand, it has been observed that mutations affecting gene expression are more prone to cause a downregulation as opposed to those mutations causing an upregulation, which tend to be less frequent and to cause bigger expression changes (Khaitovich *et al.* 2005b). Because random mutations increase entropy, there is a higher chance that a mutation in a regulatory region will decrease or disrupt the binding site of a transcription factor causing downregulation (Chaix *et al.* 2008). Investigating the mutation dynamics affecting gene expression variation in *Heliconius* will help us understand how mutational variance is linked to changes in gene expression (Hodgins-Davis *et al.* 2015).

Overall, if directional selection is causing a branch-specific shift in gene expression one would expect to see low within-species variance, whereas if the shift is caused by a relaxation of purifying selection or balancing selection, greater within-species variance would be expected. When we looked at the degree of variability between genes showing a shift toward a higher or a lower expression level, we observed that downregulated genes have significantly higher variance than genes showing upregulation (Figure 8). From this observation, we hypothesize that relaxation of purifying selection might be driving the shifts causing downregulation on gene

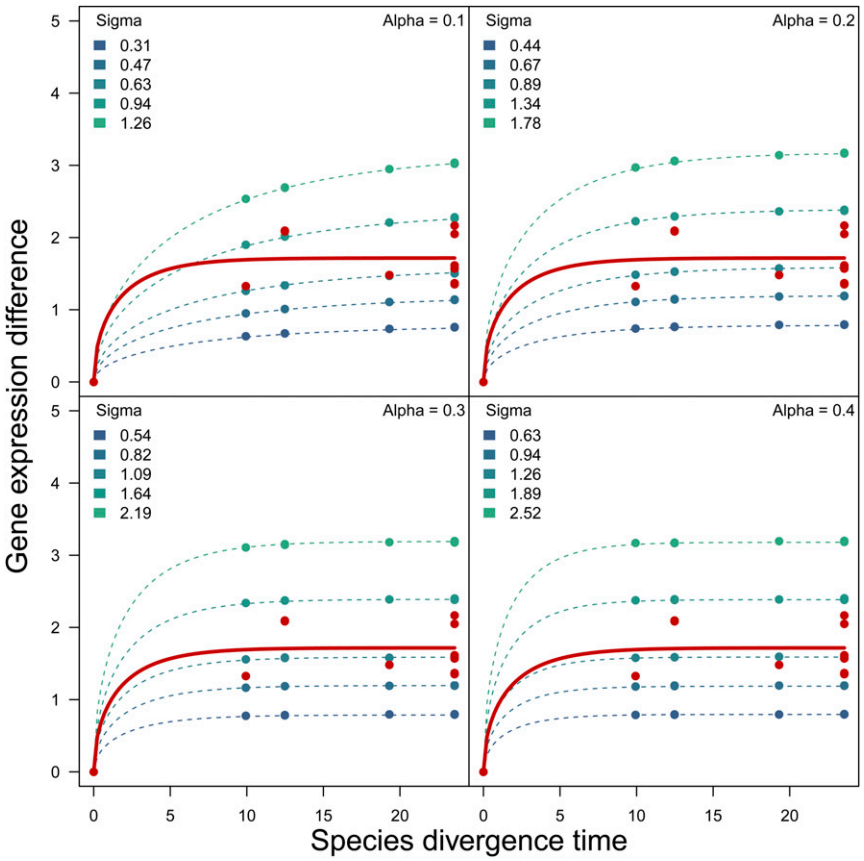

Figure 10 Between-species gene expression divergence plotted as a function of divergence time according to the *Heliconius* phylogeny. Red: σ^2 from gene expression levels observed in *Heliconius*. Blue: simulated gene expression divergence under different values of σ. Each panel shows estimates for a different value of α.

expression, a pattern that could eventually lead to a loss of expression. However, balancing selection or experimental noise could also lead to elevated within-species variance. Because of the cost of gene expression, it is expected that only those genes that are essential and have fitness effects will continue to be expressed, whereas genes that are not will eventually stop being transcribed (Stern and Crandall 2018). However, a shift toward downregulation does not always have to be a consequence of relaxed purifying selection. For example, in the orthocluster with identifier Group_449_-clean_0, a sevenfold lower expression shift was detected in the branch leading to *H. doris* (Figure S7) and significantly less variance than was expected transcriptome-wide (Fisher's exact test, P-value < 0.001). Directional selection favoring downregulation of gene expression can occur in a scenario where fine-tuning of expression levels is necessary for optimal cell or tissue function (Cayirlioglu *et al.* 2008; Catalán *et al.* 2016).

On the other hand, genes showing a branch-specific shift toward upregulation have significantly lower variances when compared to expression level shifts toward downregulation (Figure 8). This observed pattern could be a result of directional selection acting on gene expression levels leading to a reduction in the variation observed in gene expression. It is possible that to achieve an increase in gene expression levels, the selective forces leading to upregulation would have to be sufficiently strong to result in a greater investment of energy allocated to transcription costs (Wagner 2005; Lang *et al.* 2009). Some of the genes having the most extreme branch shifts in expression, either toward a higher or a lower expression level, are involved in enzymatic activity (Appendix I). Enzymes support biochemical and physiological processes, helping the optimization of tissue function (Wagner and Altenberg 1996; Feller and Gerday 1997). Thus, optimal enzymatic activity might be a key factor for species-specific brain and eye function, which in turn might be optimized for the species-specific life history and ecological niche. An approach to further test for positive selection acting on the genes showing a branch shift would be to take a population genetic approach and identify selective sweeps (Fraser *et al.* 2010; Catalán *et al.* 2012, 2016). This approach would provide a second line of evidence for positive selection acting on the transcriptome but will require demographic studies for each species (Kim and Stephan 2002).

A factor possibly influencing the proportion of transcriptome levels found to be evolving through drift, or stabilizing or directional selection is the methodology used for orthology assessment. In our analysis of gene expression variation, we assessed variation in orthoclusters where an orthologous hit was found for each of our five *Heliconius* species. Additionally, because *de novo* transcriptome assemblies are prone to form chimeric transcripts, we used strict filtering criteria when

assessing for orthology (see *Materials and Methods*). Genes with fast-evolving protein rates—to the point that orthology assessment becomes challenging—might also show gene expression shifts, which would not be detected in our experimental design. For example, orthology assessment for genes showing sex-biased gene expression, which tend to have higher evolutionary rates than unbiased genes, might require an alternative method. In fact, from the orthoclusters that we identified in this study, only two included genes with sex-biased expression (Catalán *et al.* 2018).

With this work, we have generated a set of candidate genes that are putatively evolving through directional selection, and that have the potential to be involved in the processes of adaptation and speciation. To test the role of these genes in such processes, functional validation will be necessary to gain deeper insight in the evolutionary consequences of gene expression shifts. Techniques like *in situ* hybridization, RNA interference, and clustered regularly interspaced short palindromic repeats/Cas9 are tools that can be used to shed light on the functionality of these genes. Particularly interesting could be those genes whose gene expression levels have shifted to the degree of showing absence of expression (Figure S8). The evolution of gain and loss of gene expression across a phylogeny requires a suitable theoretical framework that should be explored carefully, since such events have the potential to cause big phenotypic shifts.

Acknowledgments

We thank Aline Rangel Olguín for technical assistance with the orthocluster analysis. This work was partially supported by National Science Foundation grant IOS-1656260 to A.D.B., a Knut and Alice Wallenberg Foundation grant to Jochen Wolf, and the Deutsche Forschungsgemeinschaft (DFG) Emmy Noether-Program awarded to S.H. (Award HO 6201/1-1).

Literature Cited

Ahi, E. P., F. Richter, and K. M. Sefc, 2017 A gene expression study of ornamental fin shape in Neolamprologus brichardi, an African cichlid species. Sci. Rep. 7: 17398. https://doi.org/10.1038/s41598-017-17778-0

Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman, 1990 Basic local alignment search tool. J. Mol. Biol. 215: 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2

Ballesteros, J. A., and G. Hormiga, 2016 A new orthology assessment method for phylogenomic data: unrooted phylogenetic orthology. Mol. Biol. Evol. 33: 2117–2134 (erratum: Mol. Biol. Evol. 33: 2481). https://doi.org/10.1093/molbev/msw069

Beaulieu, J. M., D. C. Jhwueng, C. Boettiger, and B. C. O'Meara, 2012 Modeling stabilizing selection: expanding the Ornstein-Uhlenbeck model of adaptive evolution. Evolution (N. Y.) 66: 2369–2383. https://doi.org/10.1111/j.1558–5646.2012.01619.x

Bedford, T., and D. L. Hartl, 2009 Optimization of gene expression by natural selection. Proc. Natl. Acad. Sci. USA 106: 1133–1138. https://doi.org/10.1073/pnas.0812009106

Beltrán, M., C. Jiggins, A. Brower, E. Bermingham, and J. Mallet, 2007 Do pollen feeding and pupal-mating have a single origin in *Heliconius*? Inferences from multilocus sequence data. Biol. J. Linn. Soc. Lond. 92: 221–239. https://doi.org/10.1111/j.1095-8312.2007.00830.x

Berger, J. O., 1990 Robust Bayesian analysis: sensitivity to the prior. J. Stat. Plan. Inference 25: 303–328. https://doi.org/10.1016/0378-3758(90)90079-A

Brawand, D., M. Soumillon, A. Necsulea, P. Julien, G. Csárdi *et al.*, 2011 The evolution of gene expression levels in mammalian organs. Nature 478: 343–348. https://doi.org/10.1038/nature10532

Butler, M. A., and A. A. King, 2004 Phylogenetic comparative analysis: a modeling approach for adaptive evolution. Am. Nat. 164: 683–695. https://doi.org/10.1086/426002

Capella-Gutiérrez, S., J. M. Silla-Martínez, and T. Gabaldón, 2009 trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics 25: 1972–1973. https://doi.org/10.1093/bioinformatics/btp348

Carroll, S. B., 2005 Evolution at two levels: on genes and form. PLoS Biol. 3: e245. https://doi.org/10.1371/journal.pbio.0030245

Catalán, A., S. Hutter, and J. Parsch, 2012 Population and sex differences in *Drosophila melanogaster* brain gene expression. BMC Genomics 13: 654. https://doi.org/10.1186/1471-2164-13-654

Catalán, A., A. Glaser-Schmitt, E. Argyridou, P. Duchen, and J. Parsch, 2016 An indel polymorphism in the MtnA 3′ untranslated region is associated with gene expression variation and local adaptation in *Drosophila melanogaster*. PLoS Genet. 12: 1–24. https://doi.org/10.1371/journal.pgen.1005987

Catalán, A., A. Macias-Muñoz, and A. D. Briscoe, 2018 Evolution of sex-biased gene expression and dosage compensation in the eye and brain of *Heliconius* butterflies. Mol. Biol. Evol. 35: 2120–2134. https://doi.org/10.1093/molbev/msy111

Cayirlioglu, P., I. G. Kadow, X. Zhan, K. Okamura, G. S. Suh, *et al.*, 2008 Hybrid neurons in a microRNA mutant are putative evolutionary intermediates in insect CO2 sensory systems. Science 319: 1256–1260. https://doi.org/10.1126/science.1149483

Chaix, R., M. Somel, D. P. Kreil, P. Khaitovich, and G. A. Lunter, 2008 Evolution of primate gene expression: drift and corrective sweeps? Genetics 180: 1379–1389. https://doi.org/10.1534/genetics.108.089623

Chen, J., R. Swofford, J. Johnson, B. B. Cummings, N. Rogel *et al.*, 2017 A quantitative model for characterizing the evolutionary history of mammalian gene expression. bioRxiv. Available at: https://doi.org/10.1101/229096. doi: 10.1101/229096https://doi.org/10.1101/229096

Chintapalli, V. R., J. Wang, and J. A. T. Dow, 2007 Using FlyAtlas to identify better *Drosophila melanogaster* models of human disease. Nat. Genet. 39: 715–720. https://doi.org/10.1038/ng2049

Cooper, N., G. H. Thomas, C. Venditti, A. Meade, and R. P. Freckleton, 2016 A cautionary note on the use of Ornstein Uhlenbeck models in macroevolutionary studies. Biol. J. Linn. Soc. Lond. 118: 64–77. https://doi.org/10.1111/bij.12701

Dunlop, J., X. Morin, M. Corominas, F. Serras, and G. Tear, 2004 *glaikit* is essential for the formation of epithelial polarity and neuronal development. Curr. Biol. 14: 2039–2045. https://doi.org/10.1016/j.cub.2004.10.048

Eastman, J. M., M. E. Alfaro, P. Joyce, A. L. Hipp, and L. J. Harmon, 2011 A Novel comparative method for identifying shifts in the rate of character evolution on trees. Evolution (N. Y.) 65: 3578–3589. https://doi.org/10.1111/j.1558–5646.2011.01401.x

Enright, A. J., S. Van Dongen, and C. A. Ouzounis, 2002 An efficient algorithm for large-scale detection of protein families. Nucleic Acids Res. 30: 1575–1584. https://doi.org/10.1093/nar/30.7.1575

Fan, Y., R. Wu, M. H. Chen, L. Kuo, and P. O. Lewis, 2011 Choosing among partition models in Bayesian phylogenetics. Mol. Biol. Evol. 28: 523–532. https://doi.org/10.1093/molbev/msq224

Fay, J. C., and P. J. Wittkopp, 2008 Evaluating the role of natural selection in the evolution of gene regulation. Heredity (Edinb) 100: 191–199. https://doi.org/10.1038/sj.hdy.6801000

Feller, G., and C. Gerday, 1997 Psychrophilic enzymes: molecular basis of cold adaptation. Cell. Mol. Life Sci. 53: 830–841. https://doi.org/10.1007/s000180050103

Felsenstein, J., 1985 Phylogenies and the comparative method. Am. Nat. 125: 1–15. https://doi.org/10.1086/284325

Finkbeiner, S. D., A. D. Briscoe, and R. D. Reed, 2014 Warning signals are seductive: relative contributions of color and pattern to predator avoidance and mate attraction in *Heliconius* butterflies. Evolution 68: 3410–3420. https://doi.org/10.1111/evo.12524

Finkbeiner, S. D., D. A. Fishman, D. Osorio, and A. D. Briscoe, 2017 Ultraviolet and yellow reflectance but not fluorescence is important for visual discrimination of conspecifics by *Heliconius erato*. J. Exp. Biol. 220: 1267–1276. https://doi.org/10.1242/jeb.153593

FitzJohn, R. G., 2012 Diversitree: comparative phylogenetic analyses of diversification in R. Methods Ecol. Evol. 3: 1084–1092. https://doi.org/10.1111/j.2041-210X.2012.00234.x

Fraser, H. B., A. M. Moses, and E. E. Schadt, 2010 Evidence for widespread adaptive evolution of gene expression in budding yeast. Proc. Natl. Acad. Sci. USA 107: 2977–2982. https://doi.org/10.1073/pnas.0912245107

Freckleton, R. P., 2012 Fast likelihood calculations for comparative analyses. Methods Ecol. Evol. 3: 940–947. https://doi.org/10.1111/j.2041-210X.2012.00220.x

Gastwirth, J. L., Y. R. Gel, and W. Miao, 2009 The impact of Levene's test of equality of variances on statistical theory and practice. Stat. Sci. 24: 343–360. https://doi.org/10.1214/09-STS301

Glaser-Schmitt, A., and J. Parsch, 2018 Functional characterization of adaptive variation within a cis-regulatory element influencing Drosophila melanogaster growth. PLoS Biol. 16: e2004538. https://doi.org/10.1371/journal.pbio.2004538

Goldman, N., 1993 Simple diagnostic statistical tests of models for DNA substitution. J. Mol. Evol. 37: 650–661. https://doi.org/10.1007/BF00166252

Hansen, T. F., 1997 Stabilizing selection and the comparative analysis of adaptation. Evolution (N. Y.) 51: 1341. https://doi.org/10.2307/2411186

Harmon, L. J., 2018 *Phylogenetic Comparative Methods: Learning from Trees*. CreateSpace Independent Publishing Platform, Scotts Valley, CA.

Harmon, L. J., J. B. Losos, T. Jonathan Davies, R. G. Gillespie, J. L. Gittleman *et al.*, 2010 Early bursts of body size and shape evolution are rare in comparative data. Evolution (N. Y.) 64: 2385–2396. https://doi.org/10.1111/j.1558–5646.2010.01025.x

Hastings, W. K., 1970 Monte Carlo sampling methods using Markov chains and their applications. Biometrika 57: 97–109. https://doi.org/10.1093/biomet/57.1.97

Ho, L. S. T., and C. Ané, 2014 Intrinsic inference difficulties for trait evolution with Ornstein-Uhlenbeck models. Methods Ecol. Evol. 5: 1133–1146. https://doi.org/10.1111/2041-210X.12285

Hodgins-Davis, A., D. P. Rice, J. P. Townsend, and J. Novembre, 2015 Gene expression evolves under a house-of-cards model of stabilizing selection. Mol. Biol. Evol. 32: 2130–2140. https://doi.org/10.1093/molbev/msv094

Höhna, S., M. J. Landis, T. A. Heath, B. Boussau, N. Lartillot *et al.*, 2016 RevBayes: Bayesian phylogenetic inference using graphical models and an interactive model-specification language. Syst. Biol. 65: 726–736. https://doi.org/10.1093/sysbio/syw021

Höhna, S., M. J. Landis, and T. A. Heath, 2017 Phylogenetic inference using RevBayes. Curr. Protoc. Bioinformatics 57: 6.16.1–6.16.34. https://doi.org/10.1002/cpbi.22

Ives, A. R., P. E. Midford, and T. Garland, 2007 Within-species variation and measurement error in phylogenetic comparative methods. Syst. Biol. 56: 252–270. https://doi.org/10.1080/10635150701313830

Kalinka, A. T., K. M. Varga, D. T. Gerrard, S. Preibisch, D. L. Corcoran *et al.*, 2010 Gene expression divergence recapitulates the developmental hourglass model. Nature 468: 811–814. https://doi.org/10.1038/nature09634

Kass, K. E., and A. E. Raftery, 1995 Bayes factors. J. Am. Stat. Assoc. 90: 773–795. https://doi.org/10.1080/01621459.1995.10476572

Katoh, K., and H. Toh, 2008 Recent developments in the MAFFT multiple sequence alignment program. Brief. Bioinform. 9: 286–298. https://doi.org/10.1093/bib/bbn013

Khaitovich, P., G. Weiss, M. Lachmann, I. Hellmann, W. Enard *et al.*, 2004 A neutral model of transcriptome evolution. PLoS Biol. 2: e132. https://doi.org/10.1371/journal.pbio.0020132

Khaitovich, P., I. Hellmann, W. Enard, K. Nowick, M. Leinweber *et al.*, 2005a Parallel patterns of evolution in the genomes and transcriptomes of humans and chimpanzees. Science 309: 1850–1854. https://doi.org/10.1126/science.1108296

Khaitovich, P., S. Pääbo, and G. Weiss, 2005b Toward a neutral evolutionary model of gene expression. Genetics 170: 929–939. https://doi.org/10.1534/genetics.104.037135

Kim, Y., and W. Stephan, 2002 Detecting a local signature of genetic hitchhiking along a recombining chromosome. Genetics 160: 765–777.

Kozak, K. M., N. Wahlberg, A. F. E. E. Neild, K. K. Dasmahapatra, J. Mallet *et al.*, 2015 Multilocus species trees show the recent adaptive radiation of the mimetic *Heliconius* butterflies. Syst. Biol. 64: 505–524. https://doi.org/10.1093/sysbio/syv007

Kronforst, M. R., and R. Papa, 2015 The functional basis of wing patterning in *Heliconius* butterflies: the molecules behind mimicry. Genetics 200: 1–19. https://doi.org/10.1534/genetics.114.172387

Lande, R., 1976 Natural selection and random genetic drift in phenotypic evolution. Evolution 30: 314–334. https://doi.org/10.1111/j.1558-5646.1976.tb00911.x

Landis, M. J., and J. G. Schraiber, 2017 Pulsed evolution shaped modern vertebrate body sizes. Proc. Natl. Acad. Sci. USA 114: 13224–13229. https://doi.org/10.1073/pnas.1710920114

Lang, G. I., A. W. Murray, and D. Botstein, 2009 The cost of gene expression underlies a fitness trade-off in yeast. Proc. Natl. Acad. Sci. USA 106: 5755–5760. https://doi.org/10.1073/pnas.0901620106

Langmead, B., C. Trapnell, M. Pop, and S. L. Salzberg, 2009 Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 10: R25. https://doi.org/10.1186/gb-2009–10–3-r25

Larracuente, A. M., T. B. Sackton, A. J. Greenberg, A. Wong, N. D. Singh *et al.*, 2008 Evolution of protein-coding genes in *Drosophila*. Trends Genet. 24: 114–123. https://doi.org/10.1016/j.tig.2007.12.001

Lee, G., M. Foss, S. F. Goodwin, T. Carlo, B. J. Taylor *et al.*, 2000 Spatial, temporal, and sexually dimorphic expression patterns of the fruitless gene in the *Drosophila* central nervous system. J. Neurobiol. 43: 404–426. https://doi.org/10.1002/1097-4695(20000615)43:4<404::AID-NEU8>3.0.CO;2-D

Lemos, B., C. D. Meiklejohn, M. Cáceres, D. L. Hartl, and S. Url, 2005 Rates of divergence in gene expression profiles of primates, mice, and flies: stabilizing selection and variability among functional categories. Evolution 59: 126–137. https://doi.org/10.1111/j.0014-3820.2005.tb00900.x

McCulloch, K. J., D. Osorio, and A. D. Briscoe, 2016 Sexual dimorphism in the compound eye of Heliconius erato: a nymphalid

butterfly with at least five spectral classes of photoreceptor. J. Exp. Biol. 219: 2377–2387. https://doi.org/https://10.1242/jeb.136523
McCulloch, K. J., F. Yuan, Y. Zhen, M. L. Aardema, G. Smith *et al.*, 2017 Sexual dimorphism and retinal mosaic diversification following the evolution of a violet receptor in butterflies. Mol. Biol. Evol. 34: 2271–2284. https://doi.org/10.1093/molbev/msx163
Mendoza-Cuenca, L., and R. Macías-Ordóñez, 2005 Foraging polymorphism in *Heliconius charitonia* (Lepidoptera: Nymphalidae): morphological constraints and behavioural compensation. J. Trop. Ecol. 21: 407–415. https://doi.org/10.1017/S0266467405002385
Merrill, R. M., K. K. Dasmahapatra, J. W. Davey, D. D. Dell'Aglio, J. J. Hanly *et al.*, 2015 The diversification of *Heliconius* butterflies: what have we learned in 150 years? J. Evol. Biol. 28: 1417–1438. https://doi.org/10.1111/jeb.12672
Merrill, R. M., P. Rastas, S. H. Martin, M. C. Melo, S. Barker *et al.*, 2019 Genetic dissection of assortative mating behavior. PLoS Biol. 17: 1–21. https://doi.org/10.1371/journal.pbio.2005902
Metropolis, N., A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, 1953 Equation of state calculations by fast computing machines. J. Chem. Phys. 21: 1087–1092. https://doi.org/10.1063/1.1699114
Metzger, B. P. H., F. Duveau, D. C. Yuan, S. Tryban, B. Yang *et al.*, 2016 Contrasting frequencies and effects of *cis*- and *trans*-regulatory mutations affecting gene expression. Mol. Biol. Evol. 33: 1131–1146. https://doi.org/10.1093/molbev/msw011
Montgomery, S. H., R. M. Merrill, and S. R. Ott, 2016 Brain composition in *Heliconius* butterflies, posteclosion growth and experience-dependent neuropil plasticity. J. Comp. Neurol. 524: 1747–1769. https://doi.org/10.1002/cne.23993
Nadeau, N. J., 2016 Genes controlling mimetic colour pattern variation in butterflies. Curr. Opin. Insect Sci. 17: 24–31. https://doi.org/10.1016/j.cois.2016.05.013
Nitta, K. R., A. Jolma, Y. Yin, E. Morgunova, T. Kivioja *et al.*, 2015 Conservation of transcription factor binding specificities across 600 million years of bilateria evolution. Elife 2015: 1–20. https://doi.org/10.7554/eLife.04837
Nourmohammad, A., J. Rambeau, T. Held, V. Kovacova, J. Berg *et al.*, 2017 Adaptive evolution of gene expression in *Drosophila*. Cell Rep. 20: 1385–1395. https://doi.org/10.1016/j.celrep.2017.07.033
O'Meara, B. C., C. Ané, M. J. Sanderson, and P. C. Wainwright, 2006 Testing for different rates of continuous trait evolution using likelihood. Evolution 60: 922–933. https://doi.org/10.1111/j.0014–3820.2006.tb01171.x
Punta, M., P. Coggill, R. Eberhardt, J. Mistry, J. Tate *et al.*, 2012 The Pfam protein families databases. Nucleic Acids Res. 30: 1–12. https://doi.org/10.1093/nar/gkp985
Rifkin, S. A., J. Kim, and K. P. White, 2003 Evolution of gene expression in the *Drosophila melanogaster* subgroup. Nat. Genet. 33: 138–144. https://doi.org/10.1038/ng1086
Rohlfs, R. V., and R. Nielsen, 2015 Phylogenetic ANOVA: the expression variance and evolution model for quantitative trait evolution. Syst. Biol. 64: 695–708. https://doi.org/10.1093/sysbio/syv042
Rohlfs, R. V., P. Harrigan, and R. Nielsen, 2014 Modeling gene expression evolution with an extended Ornstein-Uhlenbeck process accounting for within-species variation. Mol. Biol. Evol. 31: 201–211. https://doi.org/10.1093/molbev/mst190
Romero, I. G., I. Ruvinsky, and Y. Gilad, 2012 Comparative studies of gene expression and the evolution of gene regulation. Nat. Rev. Genet. 13: 505–516. https://doi.org/10.1038/nrg3229
Salcedo, C., 2010 Environmental elements involved in communal roosting in *Heliconius* butterflies (Lepidoptera: Nymphalidae). Environ. Entomol. 39: 907–911. https://doi.org/10.1603/EN09340
Schraiber, J. G., Y. Mostovoy, T. Y. Hsu, and R. B. Brem, 2013 Inferring evolutionary histories of pathway regulation from transcriptional profiling data. PLoS Comput. Biol. 9: e1003255. https://doi.org/10.1371/journal.pcbi.1003255
Signor, S. A., and S. V. Nuzhdin, 2018 The evolution of gene expression in cis and trans. Trends Genet. 34: 532–544. https://doi.org/10.1016/j.tig.2018.03.007
Sinharay, S., and H. S. Stern, 2002 On the sensitivity of Bayes factors to the prior distributions. Am. Stat. 56: 196–201. https://doi.org/10.1198/000313002137
Smith, G., A. Macias-Muñoz, and A. D. Briscoe, 2016 Gene duplication and gene expression changes play a role in the evolution of candidate pollen feeding genes in *Heliconius* butterflies. Genome Biol. Evol. 8: 2581–2596. https://doi.org/10.1093/gbe/evw180
Stamatakis, A., 2006 RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22: 2688–2690. https://doi.org/10.1093/bioinformatics/btl446
Stern, D. B., and K. A. Crandall, 2018 The evolution of gene expression underlying vision loss in cave animals. Mol. Biol. Evol. 35: 2005–2014. https://doi.org/10.1093/molbev/msy106
Uyeda, J. C., and L. J. Harmon, 2014 A novel Bayesian method for inferring and interpreting the dynamics of adaptive landscapes from phylogenetic comparative data. Syst. Biol. 63: 902–918. https://doi.org/10.1093/sysbio/syu057
Wagner, A., 2005 Energy constraints on the evolution of gene expression. Mol. Biol. Evol. 22: 1365–1374. https://doi.org/10.1093/molbev/msi126
Wagner, G. P., and L. Altenberg, 1996 Complex adaptations and the evolution of evolvability. Evolution 50: 967–976.
Walters, J. R., C. Stafford, T. J. Hardcastle, and C. D. Jiggins, 2012 Evaluating female remating rates in light of spermatophore degradation in *Heliconius* butterflies: pupal-mating monandry *vs.* adult-mating polyandry. Ecol. Entomol. 37: 257–268. https://doi.org/10.1111/j.1365-2311.2012.01360.x
Wanner, K. W., A. R. Anderson, S. C. Trowell, D. A. Theilmann, H. M. Robertson *et al.*, 2007 Female-biased expression of odourant receptor genes in the adult antennae of the silkworm, *Bombyx mori*. Insect Mol. Biol. 16: 107–119. https://doi.org/10.1111/j.1365-2583.2007.00708.x
Warnefors, M., and A. Eyre-Walker, 2012 A selection index for gene expression evolution and its application to the divergence between humans and chimpanzees. PLoS One 7: e34935. https://doi.org/10.1371/journal.pone.0034935
Whitehead, A., and D. L. Crawford, 2006a Variation within and among species in gene expression: raw material for evolution. Mol. Ecol. 15: 1197–1211. https://doi.org/10.1111/j.1365-294X.2006.02868.x
Whitehead, A., and D. L. Crawford, 2006b Neutral and adaptive variation in gene expression. Proc. Natl. Acad. Sci. USA 103: 5425–5430. https://doi.org/10.1073/pnas.0507648103
Wong, E. S., D. Thybert, B. M. Schmitt, K. Stefflova, D. T. Odom *et al.*, 2015 Decoupling of evolutionary changes in transcription factor binding and gene expression in mammals. Genome Res. 25: 167–178. https://doi.org/10.1101/gr.177840.114

Communicating editor: P. Wittkopp