

A Phylogenetic Mixture Model for the Evolution of Gene Expression

Kevin H. Eng,*† Héctor Corrada Bravo,‡ and Sündüz Keleş*†

*Department of Statistics, University of Wisconsin-Madison; †Department of Biostatistics and Medical Informatics, University of Wisconsin-Madison; and ‡Department of Computer Sciences, University of Wisconsin-Madison

Microarray platforms are used increasingly to make comparative inferences through genome-wide surveys of gene expression. Although recent studies focus on describing the evidence for natural selection using estimates of the within- and between-taxa mutational variances, these methods do not explicitly or flexibly account for predicted nonindependence due to phylogenetic associations between measurements. In the interest of parsing the effects of selection: we introduce a mixture model for the comparative analysis of variation in gene expression across multiple taxa. This class of models isolates the phylogenetic signal from the nonphylogenetic and the heritable signal from the nonheritable while measuring the proper amount of correction. As a result, the mixture model resolves outstanding differences between existing models, relates different ways to estimate the across taxa variance, and induces a likelihood ratio test for selection. We investigate by simulation and application the feasibility and utility of estimation of the required parameters and the power of the proposed test. We illustrate analysis under this mixture model with a gene duplication family data set.

Introduction

The availability of gene expression data en masse admits a genomic resolution comparative expression experiment that measures many homologous gene transcripts simultaneously across many taxa in the interest of determining which genes are likely to undergo selective forces (Rifkin et al. 2003; Nuzhdin et al. 2004; Whitehead and Crawford 2006). Through such an experiment, the investigator may determine the relative strengths of neutral drift and natural selection forces on gene expression traits (Fay and Wittkopp 2008) at the single gene level while isolating whole groups of genes that act together and that might have a common evolutionary history. These investigators propose the use of the variance within and between taxa to determine the strength and form of hypothesized selection forces. The expression of each gene is a single, continuously valued trait, and, as in the usual comparative experiment, the analysis is potentially obfuscated by the evolutionary dependence common to the taxa (Felsenstein 1985; Harvey and Pagel 1991; Martins and Garland 1991; Purvis and Garland 1993; Garland et al. 2005; Rohlf 2006).

To account for this dependence, we may examine the structured form of the phylogenetic covariance matrix defined between taxa. The investigator typically considers the evolutionary relationship evidenced by a phylogenetic tree estimated from molecular sequence characters, but, for model-based comparative analyses, we wish to translate these trees into covariance matrices. Under the assumption of a Brownian motion process underlying the evolution of the trait, we may construct a phylogenetic covariance from a known tree (Felsenstein 1988). For general phylogenetic covariance matrices, Martins and Housworth (2002) suggested an eigenvector decomposition to identify variance with specific tree shapes. In Corrada Bravo et al. (2008), we developed a new algorithm for estimating a tree and

its matching Brownian motion covariance directly from observed continuous trait data. As opposed to methods like neighbor joining (Saitou 1987), this method globally optimizes a projection criterion over all possible tree topologies using proven efficient methods for combinatorial optimization. For expression from gene duplication families, Gu (2004) and Oakley et al. (2005) both reparameterize the mutational rates on each branch of a known tree covariance allowing it to better fit the phylogeny information. Of particular note, the addition of an error component allows these covariances to extend to a model for the entire experiment with a single covariance matrix (Ives et al. 2007; Felsenstein 2008).

Practically, linear models model both dependence and error by implicitly assuming a covariance structure that decomposes the observed or experimental variance. Such decompositions are especially desirable because they correspond to known structures in the experiment. Lynch (1991) defines a mixed-effects model across multiple traits, capturing the phylogenetic structure in a relationship matrix G and covariance between traits as a series of single parameter variance components. Although adapting Lynch's model for biological replicate data, Christman et al. (1997) extend a memetic, due to Cheverud et al. (1985), where the trait value (T) is separated into a phylogenetic component (P), a specific value (S), and a random error component (E), namely $T = P + S + E$. This decomposition leads the authors to conclude that Lynch's model isolates heritable effects ($P + S$) from noise (E) but fails to separate them from one another (P from S). Housworth et al. (2004) reformulate Lynch's model to address this deficiency by emphasizing a parameter that indirectly estimates the degree of phylogenetic signal in the sample. Guo et al. (2007) fit three types of Bayesian-flavored mixed-effects models each parameterizing an increasing amount of phylogenetic signal, finding that modeling the degree of signal present yields better models.

The importance of determining the amount of phylogenetic signal in a sample cannot be understated. If there is a phylogenetic signal, the comparative analysis ought to find and remove the extra variation. If no signal can be detected, then corrective methods will overzealously bias the final estimates. Permutation tests at the level of tree

Key words: gene expression, likelihood ratio test, mixture model, natural selection, phylogenetic correction.

E-mail: keles@stat.wisc.edu.

Mol. Biol. Evol. 26(10):2363–2372. 2009
doi:10.1093/molbev/msp149
Advance Access publication July 14, 2009

© 2009 The Authors

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/2.0/uk/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

estimation offer a way of testing for the presence of a signal or not (Blomberg et al. 2003). Pagel (1999) introduced λ as a measure of the strength of the signal and developed a likelihood ratio test (LRT) for its presence. A continuous estimate carries more information than a dichotomous hypothesis test and should indicate a strong signal: we ought to apply an appropriate phylogenetic correction.

Our goal in this paper was to integrate a framework for studying selection forces into phylogenetic, variance-decomposing models in a gene expression context. With respect to tests of selection, Rifkin et al. (2003) proposed the use of the estimated mean squares to model expected variation between and within taxa. In this framework, evidence of deviation from expectation under neutrality is evidence of the effect of natural selection. Nuzhdin et al. (2004) revised this idea using nested random effects in an analysis of variance (ANOVA) model and proposing the numerator and denominator of the standard uncorrected F ratio to be estimates of the between- and within-taxa variance. In particular, they give forms of the tests that distinguish between purifying and adaptive selection. Whitehead and Crawford (2006) continue the use of plain mean square estimates, adding a test for stabilizing selection.

In this article, we present a mixture model for the covariance in order to resolve predecessor models' inability to separate phylogenetic effects from nonphylogenetic ones. In such a model, the necessary degree of correction is freely estimated so the investigator may draw inferences on parameters unconfounded by dependence. We discuss the convergence of existing models by demonstrating the relationships between their assumptions on the covariance; the mixture formulation covers a continuum of models set between independent contrasts and the class of phylogenetic mixed-effects models. We describe the main assumptions and implications of the model from the practical analysis point of view, illustrating its effect with a simulation study and demonstrating its use in the study of gene family evolution in *Saccharomyces cerevisiae* (Oakley et al. 2005).

Methods

We are interested in modeling the variance of gene expression measurements, $Y_{gr} = (Y_{gr1}, \dots, Y_{grT})$, which are made across the T taxa of a known phylogenetic tree for $g = 1, \dots, G$ many genes with $r = 1, \dots, n$ many microarray replicates. Through a Brownian motion process (Felsenstein 1985), a rooted, bifurcating, phylogenetic tree has a well-understood translation into the covariance matrix of a multivariate normal random vector.

Pagel (1999) and Freckleton et al. (2002) introduce the parameter λ as a measure of the strength of phylogenetic correlation or the "loss of history," which induces a covariance matrix $V(\lambda)$. In defining $V(\lambda)$ to be a phylogenetic covariance matrix whose off-diagonals are multiplied by λ , the authors implicitly assume that opposing the

phylogenetic structure V_0 is a specific, nonphylogenetic structure Λ_0 :

$$\begin{aligned} V(\lambda) &= [\lambda J_T + (1 - \lambda)I_T] \circ V_0 \\ &= \lambda V_0 + (1 - \lambda)\Lambda_0. \end{aligned} \quad (1)$$

Here, J_T is a $T \times T$ matrix of ones, I_T is the identity matrix of the same dimension, and \circ is the elementwise (Hadamard) product. We define Λ_0 to be the diagonal matrix with the same main diagonal as V_0 and assume that $0 \leq \lambda \leq 1$.

Adding the experimental error to this variance, we might model the variability of this set of measurements as

$$\text{Var}(Y_{gr}) = \tau^2 a \lambda V_0 + \tau^2 a (1 - \lambda) \Lambda_0 + \sigma^2 I_T \quad (2)$$

$$\begin{aligned} &= \frac{\tau^2 \lambda}{\tau^2 + \sigma^2} a (\tau^2 + \sigma^2) V_0 \\ &\quad + \frac{\tau^2 (1 - \lambda)}{\tau^2 + \sigma^2} a (\tau^2 + \sigma^2) \Lambda_0 \\ &\quad + \frac{\sigma^2}{\tau^2 + \sigma^2} (\tau^2 + \sigma^2) I_T \end{aligned} \quad (3)$$

$$= p_1 (\kappa_1 V_0) + p_2 (\kappa_1 \Lambda_0) + p_3 (\kappa_2 I_T), \quad (4)$$

where τ^2 is the rate of variation in the expression trait across taxa, σ^2 is the rate of variation within taxa, and a is a nuisance scale factor accounting for the difference in units between sequence-based trees (typically, the expected number of sequence substitutions) and the log ratio scale of the gene expression measurements. This point is discussed further in the following sections. The proportions p_1 , p_2 , and p_3 are constrained to sum to 1, and κ_1 and κ_2 are measurements of total variation on the sequence scale and the expression scale.

This variance has two important interpretations. The relative rate interpretation (eq. 2) expresses the variance in terms of rates of mutation so that sequence-based models and expression-based models of evolution may be compared. The second mixture model interpretation (eq. 4) is that p_1 , p_2 , and p_3 represent the proportion of the observed variance attributable to certain archetypical signals: the phylogenetic history, nonphylogenetic variation, and within-taxa variation, respectively. One should note that this is precisely the desired decomposition of the variance into $T = P + S + E$ components from Christman et al. (1997).

Under this second interpretation, formalized in the following section, these parameters are estimable and evidence for selection forces can be evaluated. It may appear that a density with this mixture covariance is not identifiable because different sets of $(\mathbf{p}, \mathbf{\kappa})$ yield the same marginal covariance. In the mixture model, those densities are not the same; that is, two identifiable mixtures may produce the same marginal covariance.

A Phylogenetic Mixture Model

We suppose that Y_{gr} follows a mixture distribution with variance equation 4 and some mean vector $\boldsymbol{\mu}$.

The mixture model probability density function of Y_{gr} is given by

$$\begin{aligned} \Pr(Y_{gr}; \mathbf{p}, \boldsymbol{\mu}, \boldsymbol{\kappa}, V_0, \Lambda_0, I_0) = & p_1 f(Y_{gr} | \boldsymbol{\mu}, \kappa_1 V_0) \\ & + p_2 f(Y_{gr} | \boldsymbol{\mu}, \kappa_1 \Lambda_0) \\ & + p_3 f(Y_{gr} | \boldsymbol{\mu}, \kappa_2 I_0), \end{aligned} \quad (5)$$

where $f(y|\boldsymbol{\mu}, \boldsymbol{\Sigma})$ is a normal density with mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$ and $p_1 + p_2 + p_3 = 1$, $\kappa_1, \kappa_2 > 0$. This mixture model supposes that the observed variation has three sources, the correlated phylogenetic signal, the uncorrelated nonphylogenetic signal, and residual experimental variance each represented by the three distributions.

These component distributions may be interpreted as particular archetypical scenarios. If the data show phylogenetic signal (a particular type of nonindependence), then we believe that they come from the $f(Y_{gr} | \boldsymbol{\mu}, \kappa_1 V_0)$ component. If the data were independent but not identically distributed (each has its own specific variance), then $f(Y_{gr} | \boldsymbol{\mu}, \kappa_1 \Lambda_0)$ is the correct model. If the taxa were truly independent and identically distributed (i.i.d.) noise, then $f(Y_{gr} | \boldsymbol{\mu}, \kappa_2 I_0)$ takes precedence. Mixing proportions p_1 , p_2 , and p_3 represent the relative strengths or the probabilities of each component.

Because V_0 and Λ_0 are sequence-based estimates, they require a different scale (κ_1) than the expression log ratio-based error term (κ_2). Recall that V_0 is a tree-structured covariance, Λ_0 is a diagonal matrix with the same diagonal entries (same specific variance but no covariance), and we typically assume that I_0 is the identity matrix of size T .

This model can be fit using the Expectation Maximization algorithm (Dempster et al. 1977) outlined in the Appendix, where the strategy is to find the maximum-likelihood estimates ($\mathbf{p}, \boldsymbol{\kappa}$) and to transform them into $(\tau^2, \sigma^2, \lambda, a)$. Some technical details (Everitt and Hand 1981) require that V_0 and Λ_0 are identifiable that the off-diagonal entries of V_0 are not too small or that V_0 is a reasonable tree estimate, a case readily checked by the investigator.

Testing Selection Hypotheses

We consider the evidence in favor of natural selection forces characterized by the variance between and within taxa (Rifkin et al. 2003) and the ratio between them, the F ratio. Nuzhdin et al. (2004) identify genes with both variance estimates low as undergoing stabilizing selection, genes with low F ratios may be undergoing balancing selection, and genes with large F ratios may undergo adaptive divergence. Whitehead and Crawford (2006) add the constraint that genes undergoing adaptive divergence ought to favor a particular direction, that is, correlate with an additional environmental covariate.

The variance estimates in these studies vary: the first article uses the mean squared error for the variance within taxa and the mutational variance scaled by time for the variance between taxa. The second uses the variance of a nesting factor (species) and the nested factor (line). The last uses the variance among the population

means and the variance within populations. Using these ANOVA, sums of squares implicitly assumes the following variances,

$$\begin{aligned} \text{Var}(Y_{gr}) = & \sigma_{\text{error}}^2 I_T \\ & \text{(Rifkin et al. 2003)} \end{aligned} \quad (6)$$

$$\begin{aligned} = & \sigma_{\text{species}}^2 I_T + \sigma_{\text{line (species)}}^2 I_T \\ & \text{(Nuzhdin et al. 2004)} \end{aligned} \quad (7)$$

$$\begin{aligned} = & \sigma_{\text{pop.}}^2 I_T + \sigma_{\text{individual (pop.)}}^2 I_T \\ & \text{(Whitehead and Crawford 2006).} \end{aligned} \quad (8)$$

Although these studies do consider phylogenetic corrections at other points in their analysis, their ANOVA mean square estimates for the variances are uncorrected for possible phylogenetic dependence (they use diagonal I_T).

In our mixture parametrization, τ^2 is the gene-specific between-taxa variance (numerator of the F ratio) and σ^2 is the gene-specific within-taxa variance (denominator of the F ratio). Because σ^2 is interpreted as the rate of mutation in the expression trait, the relative sizes of τ^2 and σ^2 imply the following different evolutionary scenarios. When $\tau^2 = \sigma^2$, the signal is consistent with a Brownian motion process evolving along the given tree, representing the neutral drift null hypothesis. If $\tau^2 < \sigma^2$, there is less expression divergence than predicted by sequence divergence, suggesting that the gene may be undergoing balancing selection. Inversely, $\tau^2 > \sigma^2$ favors directional selection because the observed divergence is larger than expectation. We relax the requirement that the residuals must also show correlation with environmental covariates, that is, that they show a particular direction as well. If τ^2 and σ^2 are both “small,” then we conclude that there is evidence of purifying or stabilizing selection. Because it is not clear what constitutes an unusually small variance, we do not consider testing stabilizing selection hypotheses at this time (see Discussion section).

For relative rate type models, a neutral model variance supposes that the divergence given by a sequence-based tree directly predicts the divergence in the expression trait up to a mutation rate constant, that is, it assumes that $\sigma^2 = \tau^2$ or that

$$\text{Var}(Y_{gr}) = 2\sigma^2(p_1 a V_0 + p_2 a \Lambda_0 + p_3 I_T), \quad (9)$$

where $p_1 = \lambda/2$, $p_2 = (1 - \lambda)/2$, and $p_3 = 1/2$. This variance may be used as the null model for a LRT, which compares the log likelihoods of the model fit under the general mixture variance (eq. 4) and the model under the neutral variance (Hulslenbeck and Rannala 1997). Because the within- and between-taxa variance estimates can take several possible values, we can test for evidence of each of these types of selection using a single omnibus test. If the test is not significant, we cannot reject the neutral drift model, but if it is significant, we must look at the estimates of τ^2 and σ^2 to determine the type of selection evidenced.

To conduct the test, compare the LRT statistic versus its asymptotic distribution, where $ll(\sigma^2, \tau^2, \lambda)$ is the log

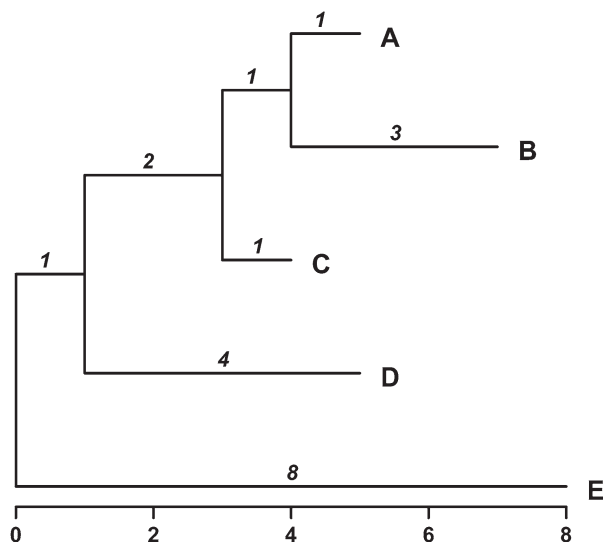


FIG. 1.—Example tree for simulation study.

likelihood of the unrestricted model and $l_0(\sigma^2, \lambda)$ is the log likelihood of the model when $\sigma^2 = \tau^2$:

$$LR = -2\{l(\hat{\sigma}^2, \hat{\tau}^2, \hat{\lambda}) - l_0(\hat{\sigma}^2, \hat{\lambda})\} \sim \chi^2_1, \quad (10)$$

for unrestricted maximum-likelihood estimates $(\hat{\sigma}^2, \hat{\tau}^2, \hat{\lambda})$ and estimates $(\hat{\sigma}^2, \hat{\lambda})$ under the null model. An algorithm for computing both likelihoods is given in the Appendix.

Results

Simulation: Need for Corrections

Although it is well accepted that phylogenetic corrections are necessary in comparative studies, we construct the following simulation study to illustrate the cost of failing to correct a phylogenetic signal on the statistical evolutionary hypotheses posited above. Suppose that V_0 is the following tree-structured matrix with corresponding tree in figure 1, that is, the main diagonal entries are the specific variances for each taxa (total branch length) and the off-diagonals are the covariances between taxa (shared branch lengths). For example, with reference to the branch lengths in figure 1, $\text{Var}(\text{Taxa A}) = 1 + 2 + 1 + 1 = 5 = [V_0]_{11}$; $\text{Var}(\text{Taxa B}) = 1 + 2 + 1 + 3 = 7 = [V_0]_{22}$; and $\text{Cov}(\text{Taxa A, Taxa B}) = 1 + 2 + 1 = 4 = [V_0]_{12}$.

$$V_0 = \begin{pmatrix} 5 & 4 & 3 & 1 & 0 \\ 4 & 7 & 3 & 1 & 0 \\ 3 & 3 & 7 & 1 & 0 \\ 1 & 1 & 1 & 5 & 0 \\ 0 & 0 & 0 & 0 & 8 \end{pmatrix}.$$

Under the mixture model proposed above, we define the selection hypotheses in table 1. We construct an artificial array of 350 genes (50 genes undergo each type of selection) and an artificial experiment where each gene is measured 500 times (5 taxa in V_0 above, 100 individuals). λ is generated by sampling a Uniform(0,1) random variable once for each of the 350 genes. Each gene has replicates drawn from one of three sources

Table 1
Simulation Design

Hypothesis	τ^2	σ^2	Number of Genes	Plot Color
Neutral drift	1.00	1.00	50	Black
Balancing selection, weak	1.00	5.00	50	Light red
Balancing selection, strong	1.00	10.00	50	Dark red
Directional selection, weak	5.00	1.00	50	Light blue
Directional selection, strong	10.00	1.00	50	Dark blue
Stabilizing selection, weak	0.10	0.10	50	Light green
Stabilizing selection, strong	0.05	0.05	50	Dark green
Total			350	

with probabilities $p_1 = \frac{\tau^2 \lambda}{\tau^2 + \sigma^2}$, $p_2 = \frac{\tau^2(1-\lambda)}{\tau^2 + \sigma^2}$, and $p_3 = \frac{\sigma^2}{\tau^2 + \sigma^2}$. The details of the data generation are given below.

$$X_{gr} \sim \text{Multinomial}(p_1, p_2, p_3) \quad (11)$$

$$Y_{gr} | \{X_{gr} = 1\} \sim \mathcal{N}(0, (\tau^2 + \sigma^2)V_0) \quad (12)$$

$$Y_{gr} | \{X_{gr} = 2\} \sim \mathcal{N}(0, (\tau^2 + \sigma^2)\Lambda_0) \quad (13)$$

$$Y_{gr} | \{X_{gr} = 3\} \sim \mathcal{N}(0, (\tau^2 + \sigma^2)I_0). \quad (14)$$

Note that we choose a large number of individuals (arrays) to illustrate this problem clearly; if the problem exists for a large number of individuals, then it ought to exist for a small number of individuals. Also, all the methods presented operate gene by gene so that these conclusions scale to any sized experiment.

We are most interested in considering the adequacy of methods based on ANOVA sums of squares. For each gene g , the mean square approach generates estimates from the usual one-way ANOVA table and considers significant effects using an F -ratio test. Because there are $T = 5$ rows in V_0 , and we select $n = 100$ replicates, the proper reference for this test is the F distribution with 4 and 495 degrees of freedom.

The resulting simulated data are analyzed in figure 2, which plots the logged values of τ^2 and σ^2 under the scenarios tabled above. The seven versions of the variance-based hypotheses are color coded: the black points represent a neutral drift null scenario, the two shades of blue are genes undergoing strong and weak directional selection; two shades of red, balancing selection and two shades of green, stabilizing selection. Two gray lines indicate the level 0.05 two-sided thresholds for the F -ratio test. Points above the upper threshold show evidence of directional selection. Points below the lower threshold show evidence of balancing selection. We do not implement the corresponding stabilizing selection tests.

The top two panels illustrate the same data when V_0 captures the true correlation between taxa. The true values of τ^2 and σ^2 are the same for every gene in the same group, so the spread of points represents sampling variability (and to some extent the effect of λ). The plot on the left (fig. 2a) shows the ANOVA estimates and on the right (fig. 2b)

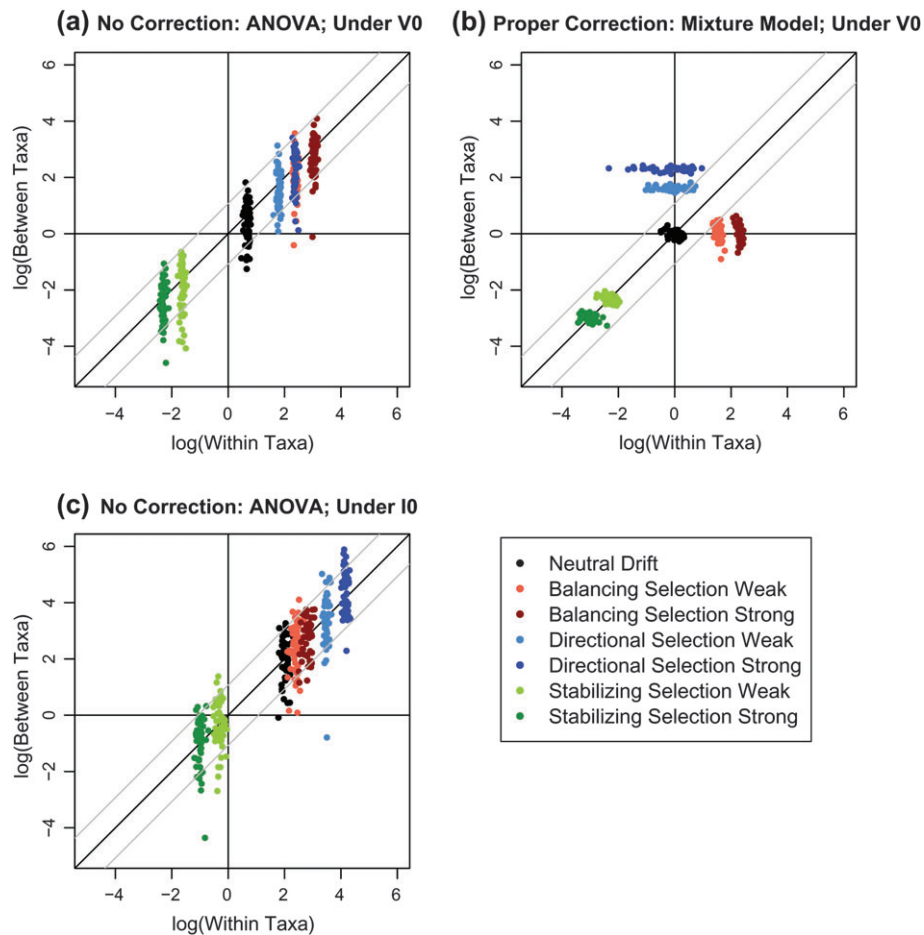


FIG. 2.—Simulation example. Simulation under ideal settings for selection hypotheses defined in the text with a tree-structured covariance (V_0) and a nonphylogenetic covariance (I_0) shows that the ANOVA estimators do not discriminate between the hypotheses. The gray lines identify tests of selection: the corresponding two-sided F -test thresholds at $\alpha = 0.05$ for $F_{4,95}$. Panel (a) shows the ANOVA estimates of data generated under V_0 ; panel (b) shows the mixture model estimates under V_0 ; and panel (c) illustrates that the ANOVA estimates are inflated even under i.i.d. characters (the primary ANOVA assumption).

shows the estimates from the mixture model. Intuitively, both variance estimation procedures partition the total observed variance into within- and between-taxa parts. Because we assume the mixture model is the true generating model, we can see that the ANOVA estimates tend to over estimate σ^2 and make up for the excess by increasing the variance in the estimate of τ^2 . In a joint bias–variance tradeoff, the ANOVA estimate trades low variance in the σ^2 estimate for bias and higher overall error in the τ^2 estimate.

The left-hand plots (panels 2a and c) employ the mean square estimates for the between- and within-taxa variances. In panel 2a, because all the groups of genes in each class of hypotheses are centered about the identity line, it is clear that choosing genes using their F ratio is not specific for the directional alternative or the balancing alternative; some genes from each group fall above or below the corresponding threshold. This pattern holds even in the lower left plot (2c) where we assume the data really are i.i.d.

In addition to problems with the hypothesis tests, the estimates of σ^2 and τ^2 appear overestimated when we do not account for the dependence structure. In figure 2a and

c, the neutral drift cluster and the stabilizing selection clusters (which are and should be centered on the identity line) appear biased much farther up the identity line than they should.

Next, we contrast these observations with the estimates from with the mixture model (panel 2b). The plot shows what we would ideally like to see: all the genes clearly separate based on the true values of their parameters. The effects are clearly separated implying that there are a sufficient number of replicates to identify all the effects. Note that this scenario represents artificially ideal conditions: a large number of observations, good separation, each gene class has the same true parameter. The point is that the mean square estimates do not behave as expected under this optimal setting and we would not expect them to do so under more realistic experimental conditions. In practice, we might expect each gene to have a different set of parameters (τ^2 and σ^2) and the groups to overlap significantly. Furthermore, the proportion of genes undergoing natural selection may alter the plot significantly, the plot will depend on the proportion of genes under each type of selection and their relative strengths.

Simulation: Calibration

As we discussed in the Methods section, the mixture model relies on a V_0 matrix that captures the phylogenetic relationship between the taxa. Additionally, tests of selection hypotheses require a model that preserves the relationship between τ^2 and σ^2 . Because estimates of the phylogenetic tree are typically obtained from sequence information (or similar independent sources), there is no reason to believe that it is of the appropriate scale for expression level data. If the given covariance structure is scaled too small, then estimates of τ^2 will be artificially large; likewise, if the given covariance is too large, τ^2 will be too small. The mixture model variance includes the parameter a to account for this scale (eq. 2).

We need to emphasize that the mixture assumption, made in the Methods section, is necessary to obtain an identifiable estimate of a . Had we assumed a marginal mixed-effects model (Lynch 1991; Martins and Hansen 1997; Guo et al. 2007) with the same variance (eq. 2), the scale parameter and the variance would only be estimable as $a\tau^2$. Practically, the investigator would have to assume some value of a in order to conduct selection tests, but this would create an uncorrectable bias in the testing framework.

Figure 3 summarizes the scaling problem and this correction using the same set of seven hypotheses from the simulation section. For these plots, we assume that we know V_0 , the same as in the last section, but that data are generated from scaled versions of V_0 . In the left panels (3a and d), the true $V_{1/100} = V_0/100$ is 100 times smaller than the given V_0 . In the middle (3b and e), the scale is correct. In the right panels (3c and f), the true $V_{100} = 100V_0$ is 100 times larger than the given V_0 .

The top row of figure 3 demonstrates the effect of the wrong-sized covariance by fitting the mixture with V_0 given. Note that estimates are drawn uniformly downward in panel 3a but pushed upward in panel 3c. For reference, panel 3b is the same plot from figure 2b. Fewer points appear in the latter plot because estimates may be unobtainable when this scaling is too far off. The bottom row of figure 3 shows the effect of estimating nuisance scale a for large and small true values.

It makes sense that the procedure fails for a very small because this case corresponds to the scenario where the heritable component is weak, that is, there is very little signal. At present, this case can be identified by observing an unusually large proportion of genes for which $\hat{\lambda} = 0$ because very small a forces λ to shrink even if the signal is present.

Simulation: Tests

Based on the data presented in the next section, we choose more realistic strong and weak presentations of balancing and directional selection forces for studying the LRT. As we mentioned in the Methods section, each hypothesis depends on the ratio τ^2/σ^2 assuming that $\tau^2 = \sigma^2$ represents the null hypothesis. We set $n = 5, 10$ for small sample microarray studies and $n = 15, 30, 50$ to check that the asymptotic distribution is correctly chosen. For each

hypothesis, we draw a simulation data set by selecting three nuisance parameters (σ^2, a, λ),

$$\sigma^2 \sim \chi_1^2, \quad (15)$$

$$a = a'|a' > 1, \quad a' \sim \chi_1^2 \quad (16)$$

$$\lambda \sim \text{Uniform}(0, 1), \quad (17)$$

and computing the LRT. We repeat this procedure 10,000 times, tabulating the proportion of tests with P values less than 0.05 in table 2. Repeating the procedure 50 times allows us to compute the simulation error (in all cases < 0.005). Note that we consider $a > 1$ because the calibration simulation indicates that when $a < 1$, the available signal is hard to estimate.

Although the LRT has reasonable power only at moderate sample sizes ($n = 15$), it should be noted that it is an omnibus test in the sense that it tests for any deviation from $\sigma^2 = \tau^2$. It is likely that a more powerful test can be built when it is of interest to test for either balancing or directional selection. One could also conduct these tests separately (as in Whitehead and Crawford 2006), but that procedure would raise concerns about multiple testing.

Housworth et al. (2004) observed that the small number of replicates available in comparative experiments may not reach the statistical power necessary to make strong inferences. For gene expression data, we observe elsewhere (Eng et al. 2008) that this sort of per gene analysis may also suffer from low power but propose that clustering together genes with similar a covariance structure may generate additional power. That is, if we believe that several genes evolved in concert and are willing to draw selection inferences on a whole group (i.e., a each member of a group undergoes the same selection force versus a single gene under a unique force), then we may employ genes as identical replicates in order to increase the power of the test. For example, if we have $n = 5$ replicates, then the powers tabulated are quite poor. If we are able to suppose that $g = 2$ genes (or experiments in the example in the next section) have the same covariance structure, then there are effectively $n = 5 \times 2 = 10$ available replicates so the power nearly triples for the balancing and stabilizing selection hypotheses.

Gene Family Data Example

The application of phylogenetic techniques to gene duplication families in the yeast *S. cerevisiae* supposes that the members of these families have sequences that are linked to a common ancestor sequence, the target of a duplication event, and that the expression level of these descendent sequences is itself a trait subject to evolutionary forces (Thornton and DeSalle 2000; Gu 2004; Oakley et al. 2005). In such an analysis, the members of these families constitute the taxa of interest and the models developed in Gu (2004) and Oakley et al. (2005) test how well a sequence-derived covariance matrix matches the predicted history of the expression trait. Because there is good reason to expect a phylogenetic structure between the genes, we will re-analyze the data set presented in Oakley et al. (2005) to illustrate the mixture model.

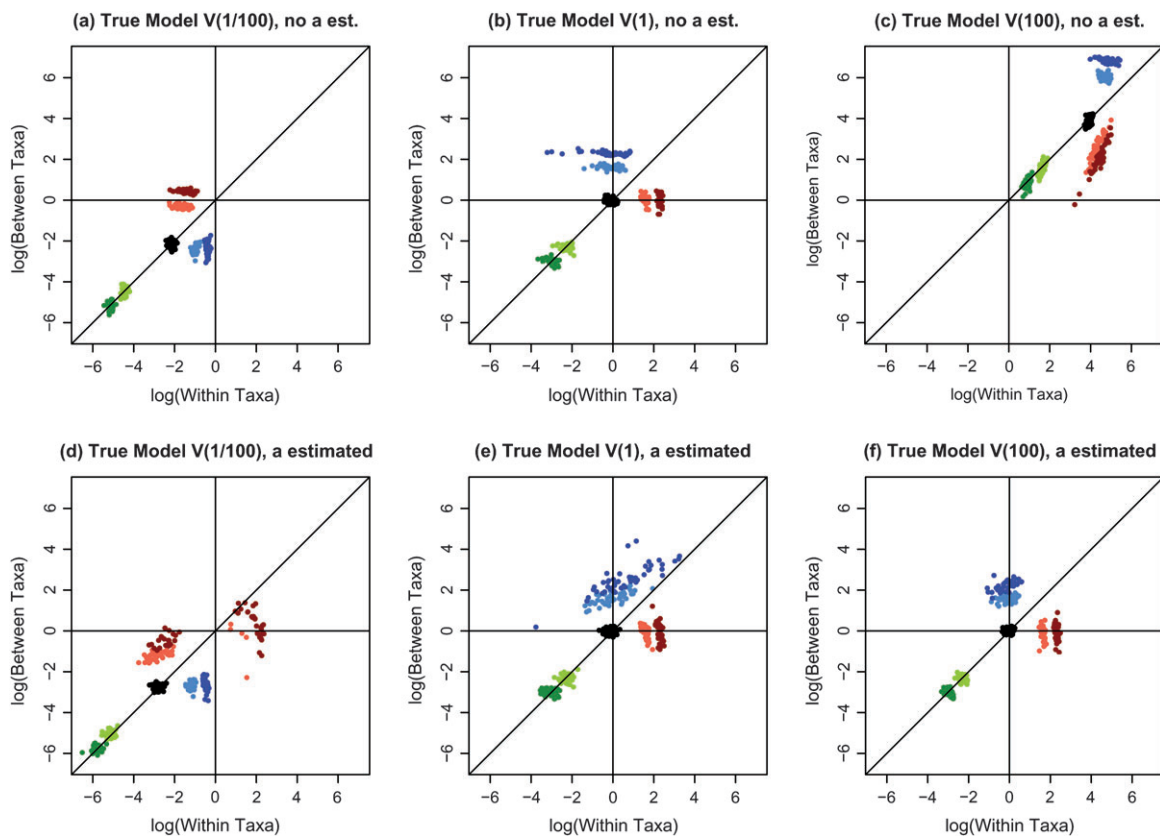


FIG. 3.—Calibration problem/solution. Estimation of σ^2 and τ^2 is sensitive to mis-specifying the scale of the phylogenetic covariance matrix, a . When a is not accounted for (panels *a–c*), estimates are shrunk for a small (panel *a*). When a is big, estimates are too big (panel *c*). The ideal pattern appears in the top center (panel *b*). Simultaneously estimating a fixes the problem (panels *e* and *f*) for all but the smallest case (panel *d*).

Using Gu's procedure (2004) for searching the proteome to identify 10 large gene families (between 7 and 18 genes each), Oakley et al. (2005) process expression arrays from 19 experiments from the Stanford Microarray Database (<http://genome-www5.stanford.edu>) and compute maximum-likelihood phylogenetic trees for each family. Each experiment represents a different experimental condition, so we may draw inferences about the evidence of selection under particular conditions. There are 19 experiments each of which contains some of the 10 gene families for a total of 169 family-specific measurements. Each experiment is a separate data set where g corresponds to a gene family, t a single transcript in the gene family, and r an array in the experiment. The data analyzed

Table 2
Simulated Power for LRT

Selection Hypothesis	τ^2/σ^2	Estimated Power				
		$n = 5$	$n = 10$	$n = 15$	$n = 30$	$n = 50$
Neutral drift (null)	$1/1$	0.017	0.034	0.038	0.049	0.050
Balancing, weak	$1/2$	0.006	0.079	0.178	0.374	0.573
Balancing, strong	$1/5$	0.002	0.223	0.517	0.884	0.981
Directional, weak	$2/1$	0.034	0.091	0.121	0.336	0.549
Directional, strong	$5/1$	0.043	0.167	0.278	0.648	0.845

NOTE.—The neutral drift null hypothesis row corresponds to the level of the test. The four selection hypotheses rows are estimated statistical power at level 0.05. All powers were estimated with 10,000 replicates and 50 simulation replicates (simulation errors < 0.005).

are available from the supplementary materials from Oakley et al. (2005) (<http://www.lifesci.ucsb.edu/eemb/labs/oakley/pubs/MBE2005data/>).

First, we consider the application of an ANOVA model, which assumes that the residuals from its fit will be i.i.d. For each gene family in table 3, the maximum residual correlation between all pairs of taxa over all replicates in all experiments demonstrates that the residuals are frequently

Table 3
Diagnostics for ANOVA Residuals

Gene Family	Maximum Residual Correlation	Levene's Test (P Value)
ABC Transporters	0.40	0.0093
ADP Ribosylation	0.56	0.0751
Alpha Glucosidases	0.71	0.1139
DUP	0.84	0.3998
GTP Binding	0.51	<0.0001
HSP DnaK	0.78	<0.0001
Hexose Transport	0.93	<0.0001
Kinases	0.43	0.0001
Permeases	0.60	<0.0001
Putative Helicases	0.75	0.0626

NOTE.—The ANOVA method for estimating the mutational variances assumes that the residuals will be i.i.d. The maximum residual correlation between pairs of taxa over all replicates in all experiments demonstrates that the residuals are frequently not independent (8 of 10 have correlation greater than 0.50) and Levene's test for the homogeneity of variances shows that the identically distributed assumption holds for only 4 of the 10 families.

Table 4
Yeast Gene Family Data

Gene Family (Number of Taxa)	Number of Experiments	$\lambda > 0.5$	Tests at Level 0.05	
			Directional	Balancing
ABC Transporters (8)	17	4	1	0
ADP Ribosylation (7)	17	7	0	0
Alpha Glucosidases (6)	19	4	1	0
DUP (10)	13	8	0	0
GTP Binding (11)	17	7	0	1
HSP DnaK (10)	16	1	0	1
Hexose Transport (18)	14	8	0	7
Kinases (7)	16	8	2	1
Permeases (17)	12	5	0	1
Putative Helicases (11)	11	2	0	4

NOTE.—Gene family data analyzed under the mixture model. A small number of experiments show strong phylogenetic signal ($\lambda > 0.5$), whereas the number of experiments with level 0.05 significant ratios τ^2/σ^2 large (directional) or small (balancing) are tabulated above.

not independent (8 of 10 families have correlation greater than 0.50 in at least one pair of taxa) and Levene's test for the homogeneity of variances rejects the identically distributed assumption for 6 of the 10 families. These observations reinforce the need for an adjustment to account for the violation of the i.i.d. assumptions.

The same 10 gene families appear in table 4, which lists the number of experiments in which the gene family was measured, the number of these experiments that show some evidence of phylogenetic signal ($\lambda > 0.5$) and the number of experiments that may have had significant selection LRTs at level 0.05. The tests are split into balancing and directional selection hypotheses. Tables of the experiments with significant tests and their P values are available in the supplementary materials.

The plot in figure 4 shows the ANOVA estimates and the mixture model corrected estimates on log scale (τ^2 is between taxa and σ^2 is within taxa). We have enlarged and colored points with a significant LRT at level 0.05 to illustrate their positions on these plots. As in the simulation plots, points about the identity line favor neutral expecta-

tions and points significantly distant from the line favor selection hypotheses. Not every extreme point is consistently highlighted because each test has a different sample size (number of genes in the family).

The ANOVA estimates appear to have a strong trend where τ^2 is smaller than expected, reflecting the tendency of the ANOVA estimate to favor σ^2 at the cost of shrinking τ^2 to zero if necessary (we saw this same pattern in fig. 2). The mixture estimates are more in line with neutral expectations.

The third panel (fig. 4c) subtracts the estimated common variance estimate under the neutral model; this is the scale that the LRT considers for significant deviations from the common variance. We see that most of the points with nonsignificant tests appear near the origin (the unrestricted maximum-likelihood estimates are close to the neutral variance estimate) and points further from the origin have significant test statistics.

Concordant with the finding in Oakley et al. (2005) that most families have a “nonphylogenetic” model in different experimental conditions (117 of 152), a large proportion of experiments corrected with the mixture model show weak phylogenetic signal, $\lambda < 0.5$ (115 of 169). This raises some questions about how to interpret the results because Whitehead and Crawford (2006) only defined selection scenarios for τ^2 and σ^2 supposing that $\lambda = 1$. We do find, however, that the Hexose Transport gene family appears to show strong phylogenetic signal in 8 of 14 experiments versus 12 of 14 in Oakley et al.'s analysis (2005). This family is also strongly represented in the balancing selection list (5 of 7 significant experiments).

As an illustration for combining experiments, the family of heat shock proteins has 10 genes and the data sets examined in Oakley et al. (2005) contain two separate heat shock-related experiments with $n = 2$ and $n = 7$ arrays each. When analyzed separately, neither experiment has a significant LRT. But, if we suppose that the experiments may be treated as replicates, there are $n = 9$ arrays available and the LRT is significant at level 0.05 ($P = 0.037$) with ratio 0.286 indicating balancing selection.

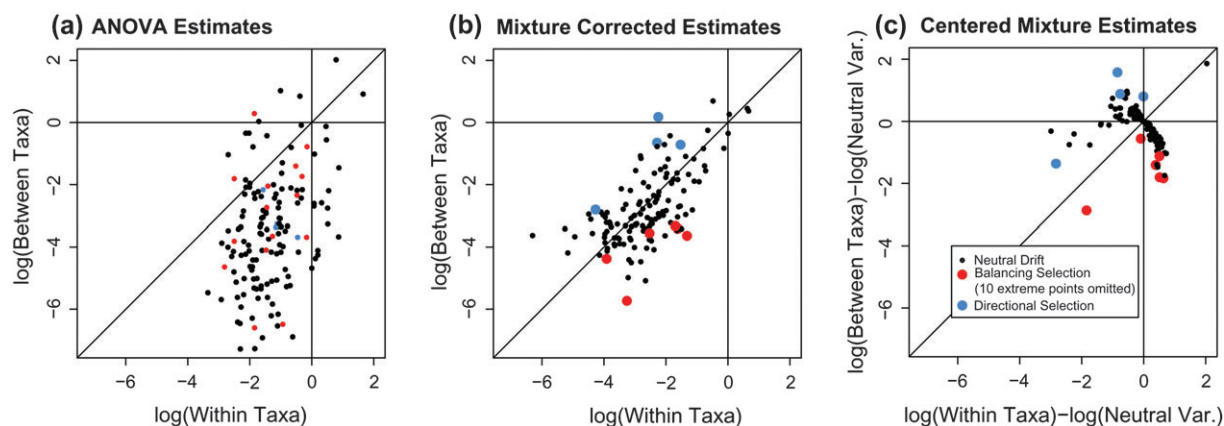


FIG. 4.—ANOVA and mixture model estimates for data from Oakley et al. (2005). Uncorrected ANOVA estimates (a) show a marked trend toward small between-taxa variances (τ^2), whereas corrected estimates (b) fit more neutral expectation. The effect of the common variance in the neutral model is removed in panel (c). The ANOVA estimates show the same low variance pattern in the σ^2 estimate as in figure 2. Ten extreme points are omitted from panels (b) and (c) to make the scales comparable. Points with significant tests at level 0.05 are enlarged and colored red for balancing and blue for directional selection evidence.

Discussion

We have investigated the estimation of the variance of expression traits within and between the tips of a given phylogenetic tree, demonstrating a problem with current analyses and proposing a solution. The model we developed anticipates the use of microarray platforms to make general inferences about the strength of evidence for natural selection forces in expression traits. As described previously, mutational variability/mean square type selection inference relies on estimating τ^2 and σ^2 . We have paid less attention to the frameworks put forward in Gu (2004) and Oakley et al. (2005) that define natural selection on the basis of a likely history parameterized in the form of V_0 . This model is also different from the Bayesian type Guo et al. (2007) model, which allows one of some, all, or no phylogenetic signal; we emphasize the transformative role of λ , τ^2 , and σ^2 in the sense that they also find a tree-structured covariance matrix consistent with observed data. Furthermore, likelihood ratio-based hypothesis testing is straightforward with these parameters.

Because this model only considers likelihood-based decompositions of the variance, we can augment it with the application of standard statistical linear model theory to accommodate much more complicated experiments. In time course expression experiments, this form of linear model may account for the correlation over time (Guo et al. 2007) and also approximate gene associations by clustering genes with similar mean and variance effects together (Eng et al. 2008). For comparison, Oakley et al. (2005) corrected for correlated adjacent time points by using the first-order differences, whereas Gu (2004) found the effect negligible. It is not unbelievable that more complex factors like expression under various conditions/treatments across taxa will be of interest and the model we have presented may serve as a useful component in that analysis.

Supplementary Material

Supplementary file 1 including tables of significant gene families from the example analysis are available at *Molecular Biology and Evolution* online (<http://mbe.oxfordjournals.org/>).

Acknowledgments

We thank Carol Lee (University of Wisconsin-Madison) and Anthony Ives (University of Wisconsin-Madison) for their feedback on earlier drafts of this article and Andrew Whitehead (Louisiana State University) for discussion on his experiments. This work was supported by National Science Foundation grant DMS 0804597 (S.K.), a PhRMA Foundation Research Starter Grant (S.K.), and the National Institute of Health grant HG03747 (S.K.).

Appendix

EM Algorithm for V_0 and Λ_0 known up to scale a for a single gene.

For all replicates $r = 1, \dots, R$ of a single gene, let C be the class variable taking values 1, 2, and 3 for components $\mathcal{N}(\mu, \kappa_1 V_0)$, $\mathcal{N}(\mu, \kappa_1 \Lambda_0)$, and $\mathcal{N}(\mu, \kappa_2 I_T)$, respectively. Suppose T is the rank of V_0 (the number of taxa). The algo-

rithm stops when the observed data log likelihood (eq. 5) increases by less than 0.001. Because the estimate μ does not depend on C , its estimate is $\hat{\mu} = \frac{1}{R} \sum_{r=1}^R (Z'Z)^{-1} Z'Y_r$ for design matrix Z (e.g., for balanced replicates, these are the group means).

1. E-Step

$$\begin{aligned} \hat{C}_{ri} &= E(C_r = i | Y_r, \hat{\mu}, \hat{\kappa}_1^{(t)}, \hat{\kappa}_2^{(t)}, \hat{p}_i^{(t)}), \\ &= \frac{P(Y_r - \hat{\mu} | C_r = i, \hat{\kappa}_1^{(t)}, \hat{\kappa}_2^{(t)}) \hat{p}_i^{(t)}}{\sum_{i'=1}^3 P(Y_r - \hat{\mu} | C_r = i', \hat{\kappa}_1^{(t)}, \hat{\kappa}_2^{(t)}) \hat{p}_{i'}^{(t)}}. \end{aligned}$$

2. M-step

$$\begin{aligned} \hat{p}_i^{(t+1)} &= \frac{\sum_r \hat{C}_{ri}}{R}, \\ \hat{\kappa}_1^{(t+1)} &= \frac{\frac{1}{T} \sum_{r=1}^R [\hat{C}_{r1}(Y_r - \hat{\mu})' V_0^{-1} (Y_r - \hat{\mu}) + \hat{C}_{r2}(Y_r - \hat{\mu})' \Lambda_0^{-1} (Y_r - \hat{\mu})]}{\sum_{r=1}^R \hat{C}_{r1} + \hat{C}_{r2}}, \\ \hat{\kappa}_2^{(t+1)} &= \frac{\frac{1}{T} \sum_{r=1}^R [\hat{C}_{r3}(Y_r - \hat{\mu})' (Y_r - \hat{\mu})]}{\sum_{r=1}^R \hat{C}_{r3}}. \end{aligned}$$

The required estimates from this algorithm are converted back to the original parametrization:

$$\begin{aligned} \hat{\tau}^2 &= \hat{\kappa}_2(1 - \hat{p}_3), \\ \hat{\sigma}^2 &= \hat{\kappa}_2(\hat{p}_3), \\ \hat{\lambda} &= \frac{\hat{p}_1}{1 - \hat{p}_3}, \\ \hat{a} &= \frac{\hat{\kappa}_1}{\hat{\kappa}_2}. \end{aligned}$$

Null Model EM Algorithm for LRT

The previous EM estimates ($\tau^2, \sigma^2, \lambda, a$) under the unrestricted full model to compute the LRT, we compare the log likelihood of that model with the log likelihood of a restricted model (neutral drift). Supposing that $\tau^2 = \sigma^2$, the mixing proportions become $(p_1, p_2, p_3) = \left(\frac{\lambda}{2}, \frac{1-\lambda}{2}, \frac{1}{2}\right)$. The component densities are components $\mathcal{N}(\mu, \kappa_1 V_0)$, $\mathcal{N}(\mu, \kappa_1 \Lambda_0)$, and $\mathcal{N}(\mu, \kappa_2 I_T)$ where $\tau^2 = \sigma^2 = \kappa_2/2$ and $a = \kappa_1/\kappa_2$.

1. E-Step

$$\begin{aligned} \hat{C}_{ri} &= E(C_r = i | Y_r, \hat{\mu}, \hat{\kappa}_1^{(t)}, \hat{\kappa}_2^{(t)}, \hat{p}_i^{(t)}), \\ &= \frac{P(Y_r - \hat{\mu} | C_r = i, \hat{\kappa}_1^{(t)}, \hat{\kappa}_2^{(t)}) \hat{p}_i^{(t)}}{\sum_{i'=1}^3 P(Y_r - \hat{\mu} | C_r = i', \hat{\kappa}_1^{(t)}, \hat{\kappa}_2^{(t)}) \hat{p}_{i'}^{(t)}}. \end{aligned}$$

2. M-step

$$\begin{aligned} \hat{p}_1^{(t+1)} &= \frac{\sum_r \hat{C}_{r1}}{2 \sum_r \hat{C}_{r1} + \hat{C}_{r2}}, \\ \hat{p}_2^{(t+1)} &= \frac{\sum_r \hat{C}_{r2}}{2 \sum_r \hat{C}_{r1} + \hat{C}_{r2}}, \\ \hat{\kappa}_1^{(t+1)} &= \frac{\frac{1}{T} \sum_{r=1}^R [\hat{C}_{r1}(Y_r - \hat{\mu})' V_0^{-1} (Y_r - \hat{\mu}) + \hat{C}_{r2}(Y_r - \hat{\mu})' \Lambda_0^{-1} (Y_r - \hat{\mu})]}{\sum_{r=1}^R \hat{C}_{r1} + \hat{C}_{r2}}, \\ \hat{\kappa}_2^{(t+1)} &= \frac{\frac{1}{T} \sum_{r=1}^R [\hat{C}_{r3}(Y_r - \hat{\mu})' (Y_r - \hat{\mu})]}{\sum_{r=1}^R \hat{C}_{r3}}. \end{aligned}$$

Again, we obtain the required estimates by converting back to the original parametrization:

$$\hat{\sigma}^2 = \hat{\kappa}_2/2,$$

$$\hat{\lambda} = 2\hat{\rho}_1,$$

$$\hat{a} = \frac{\hat{\kappa}_1}{\hat{\kappa}_2}.$$

Literature Cited

- Blomberg SP, Garland TG, Ives AR. 2003. Testing for phylogenetic signal in comparative data: behavioral traits are more labile. *Evolution* 57:717–745.
- Cheverud JM, Dow MM, Leutenegger W. 1985. The quantitative assessment of phylogenetic constraints in comparative analyses: sexual dimorphism in body weight among primates. *Evolution* 39:1335–1351.
- Christman MC, Jernigan RW, Culver D. 1997. A comparison of two models for estimating phylogenetic effect on trait variation. *Evolution* 51:262–266.
- Corrada Bravo H, Wright S, Eng KH, Keleş S, Wahba G. 2009. Machine learning for Tracking. In: van Dyk D, Welling M, editors. Estimating tree structured covariance matrices via mixed-integer programming. Proceedings of the twelfth international conference on artificial intelligence and statistics (AISTATS). JMLR: W&CP 5, Clearwater Beach, Florida, April 16–18, 2009, pp. 41–48.
- Dempster AP, Laird NM, Rubin DB. 1977. Maximum likelihood from incomplete data via the EM algorithm. *J R Stat Soc.* 39:185–197.
- Eng KH, Keleş S, Wahba G. 2008. A linear mixed effects clustering model for multi-species time course gene expression data. Madison, WI: University of Wisconsin-Madison, Department of Statistics. Report No.: 1143.
- Everitt BS, Hand DJ. 1981. Finite mixture distributions. London: Chapman and Hall.
- Fay JC, Wittkopp PJ. 2008. Evaluating the role of natural selection in the evolution of gene regulation. *Heredity* 100:191–199.
- Felsenstein J. 1985. Phylogenies and the comparative method. *Am Nat.* 125:1–15.
- Felsenstein J. 1988. Phylogenies and the quantitative characters. *Annu Rev Ecol Syst.* 19:447–471.
- Felsenstein J. 2008. Comparative methods with sampling error and within-species variation: contrasts revisited and revised. *Am Nat.* 171:713–725.
- Freckleton RP, Harvey PH, Pagel M. 2002. Phylogenetic analysis and comparative data: a test and review of evidence. *Am Nat.* 160:712–726.
- Garland T Jr, Bennett AF, Rezende EL. 2005. Phylogenetic approaches in comparative physiology. *J Exp Biol.* 208:3015–3035.
- Gu X. 2004. Statistical framework for phylogenomic analysis of gene family expression profiles. *Genetics* 167:531–542.
- Guo H, Weiss RE, Gu X, Suchard M. 2007. Time squared: repeated measures on phylogenies. *Mol Biol Evol.* 24:352–362.
- Harvey PH, Pagel MD. 1991. The comparative method in evolutionary biology. Oxford: Oxford University Press.
- Housworth EA, Martins EP, Lynch M. 2004. The phylogenetic mixed model. *Am Nat.* 163:84–96.
- Hulsenbeck JP, Rannala B. 1997. Phylogenetic methods come of age: testing hypotheses in an evolutionary context. *Science* 276:227–232.
- Ives AR, Midford PE, Garland T Jr. 2007. Within-species variation and measurement error in phylogenetic comparative methods. *Syst Biol.* 56:252–270.
- Lynch M. 1991. Methods for the analysis of comparative data in evolutionary biology. *Evolution* 45:1065–1080.
- Martins EP, Garland T Jr. 1991. Phylogenetic analyses of the correlated evolution of continuous characters: a simulation study. *Evolution* 45:534–557.
- Martins EP, Hansen TF. 1997. Phylogenies and the comparative method: a general approach to incorporating phylogenetic information into the analysis of interspecific data. *Am Nat.* 149:646–667.
- Martins EP, Housworth EA. 2002. Phylogeny shape and the phylogenetic comparative method. *Syst Biol.* 51:873–880.
- Nuzhdin SV, Wayne ML, Harmon K, McIntyre LM. 2004. Common pattern of evolution of gene expression level and protein sequence in *Drosophila*. *Mol Biol Evol.* 21:1308–1317.
- Oakley TH, Gu Z, Abouheif E, Patel NH, Li WH. 2005. Comparative methods for the analysis of gene-expression evolution: an example using yeast functional genomic data. *Mol Biol Evol.* 22:40–50. Available from: <http://www.lifesci.ucsb.edu/eemb/labs/oakley/pubs/MBE2005data/>.
- Pagel M. 1999. Inferring the historical patterns of biological evolution. *Nature* 401:877–884.
- Purvis A, Garland T Jr. 1993. Polytomies in comparative analyses of continuous characters. *Syst Biol.* 42:596–575.
- Rifkin SA, Kim J, White KP. 2003. Evolution of gene expression in the *Drosophila melanogaster* subgroup. *Nat Genet.* 33:138–144.
- Rohlf FJ. 2006. A comment on phylogenetic correction. *Evolution* 60:1509–1515.
- Saitou N. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol.* 4:406–425.
- Thornton JW, DeSalle R. 2000. Gene family evolution and homology: genomics meets phylogenetics. *Annu Rev Genomics Hum Genet.* 1:41–73.
- Whitehead A, Crawford DL. 2006. Neutral and adaptive variation in gene expression. *Proc Natl Acad Sci USA.*

Yi-Ju Li, Associate Editor

Accepted July 4, 2009