**Jointly Modeling Discrete and Continuous Traits**

**as a Means of Discovering Hidden Rate Heterogeneity**

James D. Boyko1\*, Brian O’Meara2, Jeremy M. Beaulieu1

*1Department of Biological Sciences, University of Arkansas, Fayetteville, AR, 72701, USA*

*2Department of Ecology and Evolutionary Biology, University of Tennessee, Knoxville, TN, USA*

*\*To whom correspondence should be addressed (jboyko@uark.edu).*

**Running head**: **Jointly Modeling Discrete and Continuous Traits**

**Abstract**

Rates of evolution have changed throughout the history of life and produced the rich diversity of morphology, behaviour, and ecology that characterizes the biodiversity we see today. As such, models favoring variation in rates and states has fast become the rule, as opposed to the rare exception in empirical applications of comparative methods. The challenge then becomes how to model this heterogeneous process with the limitation that, for most comparative biologists, our data comes exclusively from extant species. Our model, which we call hOUwie, uses hidden Markov models as a statistically rigorous way to learn about variation in the evolutionary process. hOUwie is composed of two processes: one describes the evolution of a discrete character (such as whether a mammal is a carnivore, herbivore, or omnivore) and the other the evolution of a continuous character (such as mammalian body size). To model the evolution of our continuous character we have chosen to use an Ornstein-Uhlenbeck (OU) model. In evolutionary biology, we use it to model the change in traits through time, which is the product of two antagonistic forces continuously pulling trait values towards some optimal value and randomly moving away from the optimum. The dynamics of these forces can also vary by so-called “regimes”. Under a typical OU model, these underlying regimes are taken as a given and are defined a priori leaving no room for inference about the evolution of the regimes themselves. To resolve this problem, we will model the evolution of these regimes as a discrete character via the aforementioned hidden Markov model. One of the most important properties we will utilize is the inclusion of hidden states. Hidden states give us a way to learn about the variation in the evolutionary process without the necessity of it being linked to one of our explanatory variables.

**Introduction**

**Methods**

The hOUwie model

The hOUwie model is composed of two processes: one describes the evolution of a discrete character and the other the evolution of a continuous character. To model the evolution of a single continuous character we have chosen to use an Ornstein-Uhlenbeck (OU) model (Hansen 1997; Butler and King 2004; Hansen et al. 2008; Beaulieu et al. 2012; Ho and Ané 2014). This model combines the stochastic evolution of a trait through time with a deterministic component which models the tendency for a trait to evolve towards an adaptive optimum. In this model, a trait () is pulled towards an optimum at a rate scaled by the parameter , while the optimum itself (which may change through time) is denoted by the parameter . is piecewise constant on intervals and takes values in a finite set. This represents the set of “selective regimes”, “regimes”, or Simpson’s “adaptive zones” based on narrational preference (Cressler et al. 2015). Additionally, random deviations are introduced by Gaussian white noise , which is distributed as a normal random variable with mean zero and variance one. The magnitude of these deviations is scaled by the noise intensity . has been interpreted as the strength of selection (Simpson 1953, Lande 1976, 1980) and has been referred to as genetic drift (Lande 1976, Hansen 1997). This latter interpretation has been criticized on the grounds that the stochasticity modeled by an OU on a macroevolutionary time scale is likely not a consequence of genetic drift (Hansen 2012). Formally, the OU process is an Itô diffusion satisfying:

Here, we use the set of extensions introduced by Beaulieu et al. (2012) and implemented in the R-package OUwie. This allows for multiple primary optima () in which both the pull strength () and the noise intensity () can vary across the phylogeny. The derivation of this model is given in Beaulieu et al. (2012). However, the algorithm used to calculate the likelihood described in Beaulieu et al. (2012) involves matrix inversion - a computationally costly procedure. Therefore, we implement a linear-time computation of the likelihood of Gaussian trait models following Ho and Ané (2014). To do this we first transform the phylogeny such that its variance covariance matrix,, is 3-point structured. We can write the variance covariance matrix of the untransformed phylogeny as , where following Beaulieu et al. (2012) and Ho and Ané (2014),

where, is the distance from the root to the beginning of the selective regime () for the number of selective regimes along the path from the root to the last common ancestor of and ( or from the root to the terminal tip (). Our transformed phylogeny now has a variance covariance matrix and diagonal matrix . We can then apply the function three.point.compute from the R-package phylolm (Ho et al. 2016) to calculate the quadratic quantities and determinant of . The likelihood of our continuous trait is given by

where is the number of tips in the phylogeny, is the continuous trait value of each species, and is the expected value of each species given the continuous trait model – calculated following equation (11) of Beaulieu et al. (2012).

Most previous phylogenetic comparative models of an OU process have assumed that the intervals and regimes are known, though the optimum trait values associated with each regime are not (but see, Uyeda et al. 2018; May and Moore 2020). This leaves no room for inference about the regimes themselves and how they change through time, nor the possibility that the evolution of the continuous character could influence change in the regimes or vice versa. To resolve this problem, we will model the evolution of these regimes as a discrete character. For this, we have assumed that regime change follows a hidden Markov model (Felsenstein and Churchill 1996; Yang 1994; Beaulieu et al. 2013). Hidden Markov models have a hierarchical structure that can be broken down into two components: a “state-dependent process” and an unobserved “parameter process” (Zucchini et al. 2017). Under an HMM, observations are generated by a given state-dependent process, which in turn depends on the state of the parameter process. In other words, the observed data are the product of several processes occurring in different parts of a phylogeny and the parameter process is way of linking them. It is initially unknown what the parameter process corresponds to biologically, hence the moniker “hidden” state. Nevertheless, the information for detecting hidden states comes from the differences in how the observed states change. If the transitions between observed states of different lineages are more adequately described by several Markov processes rather than a single process, there will be information to detect hidden states (Boyko and Beaulieu 2021).

In comparative biology, for characters that take on discrete states the standard “state-dependent process” is typically a continuous-time Markov chain with finite state-space (CTMC-FS) (O’Meara 2012). In part, the common use of this model is because the likelihood of discrete data can be calculated efficiently using conditional likelihoods at nodes (Felsenstein 1981). However, our goal is to model the discrete and continuous characters as following the same underlying regime painting. Therefore, we calculate the probability of a discrete character () as the being equivalent to the probability of an underlying regime structure (). This means, unlike the implementations of software programs like corHMM (Beaulieu et al. 2013) which marginalize over all possible path reconstructions, we are interested in calculating the probability of a specific path (Steel and Penny 2000). For a branch of length and a discrete model with instantaneous rate, , of going from state to state we can calculate the probability of a transition occurring at time as and the probability of not changing for the remaining , as .