二、研究計畫中英文摘要：請就本計畫要點作一概述，並依本計畫性質自訂關鍵詞。

（一）計畫中文摘要。（五百字以內）

自迴歸條件異質變異數模型對物種性狀演化率之建模與探討

本計劃提出以自迴歸條件異質變異數模型對生物性狀『演化變動幅度』進行統計建模。其模型考量演化過程中祖先跟後代的變動率是否存在關聯性，目前提供蒙地卡羅法搭配吉柏斯的貝氏抽樣方法來進行參數估計。預計將開發 R 套件 *phylorates* 供社群使用者使用。本提案目前所開發的程式碼可由以下網址取得 `http://tonyjhwueng.info/phygarchrate`。

關鍵詞: 自迴歸條件異質變異數模型, 演化率、種系比較方法、布朗尼運動、性狀演化

（一）計畫英文摘要。（五百字以內）

Over the last decades, there has been increased attention on studying the rate of evolution. Various statistical models were developed for estimating the rate parameters for a group of related species that evolved along the phylogenetic tree. Current existing models assume that the rate parameters were independent, this may not cover a scenario where the rate of evolution was correlated to their past or not. In this work, we propose the use of the generalized autoregressive conditional heteroskedastic (GARCH) model for modeling the rate of evolution by concerning that the rates between two successive generations (ancestor-descendant) are time-dependent and correlated along the tree, A Markov chain Monte Carlo blended with Gibb's sampling algorithm is developed for parameter estimation and inference. An R package Called *phylorates: Investigate serially autocorrelated rate evolution using the Generalized AutoRegressive Conditional Heteroskedastic (GARCH) process* Scripts and relevant files developed for this project can be accessed in `http://tonyjhwueng.info/phygarchrate`.

Keywords: Generalized autoregressive conditional heteroskedasticity model; evolutionary rates; phylogenetic comparative method, Brownian motion, trait evolution.

請概述執行本計畫可能產生對社會、經濟、學術發展等面向的預期影響性(一百五十字內)。研究演化率爲演化生物學領域中常被關切注意的議題，目前文獻上對演化率的研究多爲對演化率參數本身假設爲常數或隨機之變數，或是與其它性狀有所關聯。本提案考慮演化率的動態關係可以使用自迴歸異質變異數模型來描述，目前文獻並無類似模型來分析並解釋實際資料上演化率變化情形，因此本提案提出建立新模型以期對該領域學術上發展具有增值效果，主要是可提供給該社群的使用人員有更多選擇，並使用更合適的模型分析他們的資料。

三、研究計畫：

(一) 研究計畫之背景。請詳述本研究計畫所要探討或解決的問題、研究原創性、重要性、預期影響性及國內外有關本計畫之研究情況、重要參考文獻之評述等。如爲連續性計畫應説明上年度研究進度。

1 Introduction

On the Application of the GARCH Model for Phylogenetic Rate of Trait Evolution

In evolution, studying heritable characteristics of the biological population is helpful to understand the diversity among species on our planet Earth (Pettay et al., 2005). Macroevolution is expected to occur when selection acts on a trait that has a heritable basis of phenotypic variation. During the evolutionary process speciation results in new species, and the comparison of traits (e.g. height, weight, size $\cdots$, etc) among a group of related species can be made by studying the speed of changes of the characteristics over successive generations. In particular, the rate of evolution is assessed on a lineage. While one subgroup of species evolved at a faster rate and resulted in a larger variation in the trait, the other subgroup of species evolved at a relatively lower rate and yielded a moderate variation of the trait.

When evolutionary processes such as natural selection (including sexual selection) and genetic drift act on this variation, resulting in certain characteristics become more common or rare within a population(Scott-Phillips et al., 2014). For instance, Darwin's finches are a group of about 18 species of dull-colored passerine birds in the Galápagos islands (Soons et al., 2010). They are well known for their remarkable diversity in beak form/size and function which are highly adapted to different food sources. (Podos and Nowicki, 2004). The smallest is the warbler finch and the largest is the vegetarian finch. The birds vary in size from 10 cm to 16 cm and weigh between 8 and 38 grams. Another example is angiosperms which successfully survive and thrive on our planet, the evolution of fruits shall be the most important feature as the fruits not only comprise a food source for other species but protect seeds and contribute to dispersal seed (Xiang et al., 2017). The survival and thriving of fruits require adaptation to an environment. For instance, while dragon fruit (*Selenicereus*) endures temperatures up to 40°C (104°F) for survival, watermelon (*Citrullus lanatus*) needs temperatures higher than about 25°C (77°F) to thrive. Studying the reproducible via the rate of evolution to the resistance of temperature would help us shed light on the evolution of angiosperms themselves and also understand their ecological implications.

The rate of evolution is a measurement of the change in an evolutionary lineage over time and can be defined as the ratio of the character displacement over a certain time interval.

(Thorne et al., 1998) defined the rate of change between two samples using three quantities : (i) the proportional difference between the sample means, $d = \bar{y}_1 - \bar{y}_2$; (ii) the pooled standard deviation of the samples, (e.g., $s_p = \sqrt{s_p^2}$, where $sp^2 = ((n_1 - 1)s_1^2 + (n_2 - 1)s_2^2)/(n_1 + n_2^2)$; and (iii) the time interval between the samples, $I = t_2 - t_1$, counted or estimated in generations. Usually, the log-transformed traits are used as the logging is necessary because of the geometric normality of biological variation (comparative data are log-transformed. Finally, the rate of change in standard deviations per generation (Haldane) is calculated as: $H_{(\log I)} = D/I$, where $D = d/s_p$.

For simplicity, we adopt the following definition. Suppose that a character has been measured at two times, t_1 and t_2 where t_1 and t_2 are expressed as the time before the present in millions of years. The time length between the two samples can be written as: $dt = t_1 - t_2$, which is 1 million years if $t_1 = 6.27$ and $t_2 = 5.27$. The average value of the character is defined as x_1 in the earlier sample and x_2 in the later sample. Let $y_{t_i} = \log(x_{t_i}), i = 1, 2$ be the natural logarithm taking on x_{t_1} and x_{t_2}. Then the evolutionary rate (σ) can be defined in Eq. (1).

$$\sigma = \frac{y_{t_2} - y_{t_1}}{t_2 - t_1}. \tag{1}$$

The rate σ is measured in *darwins* which is a change in the character by a factor of one million years. A classical case study can be dated to the horse teeth evolution (MacFadden, 2005). Rates calculated over different intervals are not directly comparable, which is why rates must be indexed by the log (here base 10) of the interval.

If evolutionary events (e.g., nucleotide substitutions) occur independently, then the number of evolutionary events that occur on a branch existing from time 0 to tome T and have rate $\sigma(t)$ at time t follows a Poisson distribution with mean $b(T) = \int_0^T \sigma(t)dt$ where $b(T)$ is referred a branch length. A main difficulty is that $\sigma(t)$ can not be directly observed (see Figure 1)

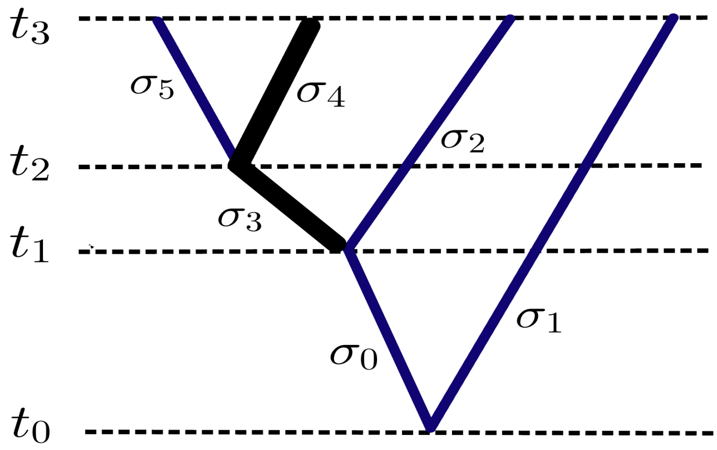

Figure 1: Rates and node times on an example tree. re-graph from (Thorne et al., 1998) The rate of branch i will be denoted σ_i. $\sigma(t)$ is the rate at time t and is used to compute an amount of evolutionary events $b(T) = \int_0^T \sigma(t)dt$ that occur on a branch existing from time 0 to to time T. When $\sigma(t) = \sigma, T_{i-1} \leq t \leq T_i, b(T) = \sigma(T_i - T_{i-1})$.

Multiply the time difference $t_2 - t_1$ on both sides of Equation (1) yields the character difference within a time unit $y_{t_2} - y_{t_1} = \sigma(t_2 - t_1)$. Conceptually consider that the character change occurred in an infinitesimal time and denote the character displacement by $dy_t = \lim_{t_1 \to t_2} y_{t_2} - y_{t_1}$, then one has the differential equation

$$dy_t = \sigma dt.$$

Given $y(0) = y_0$, the solution is $y_t = y_0 + \sigma t$ which shows that y_t is increased with time t. Unfortunately, this may not be an appropriate model to describe character change in the evolutionary sense. Instead, one may consider the variation of the character change adopts a certain dynamic. For instance, if considering the variation of the character change is proportional to time, then a stochastic variable W_t can be introduced where W_t is a Wiener process with independent increment and $dW_t = W_{t+dt} - W_t \sim \mathcal{N}(0, dt)$. Thus, the dynamic of the trait variable y_t $(= \log x_t)$ follows a Brownian motion in Eq. (2)

$$dy_t = \sigma dW_t. \tag{2}$$

Given $y(0) = y_0$, integrate both side of Equation (2), one has

$$y_t = y_0 + \sigma W_t \text{ where } W_t \sim \mathcal{N}(0, t). \tag{3}$$

Equation (3) describes the dynamic of the trait variable y_t at time t. A simulation of the trajectories of y_t using Brownian motion under two different rates is shown in Figure 2.

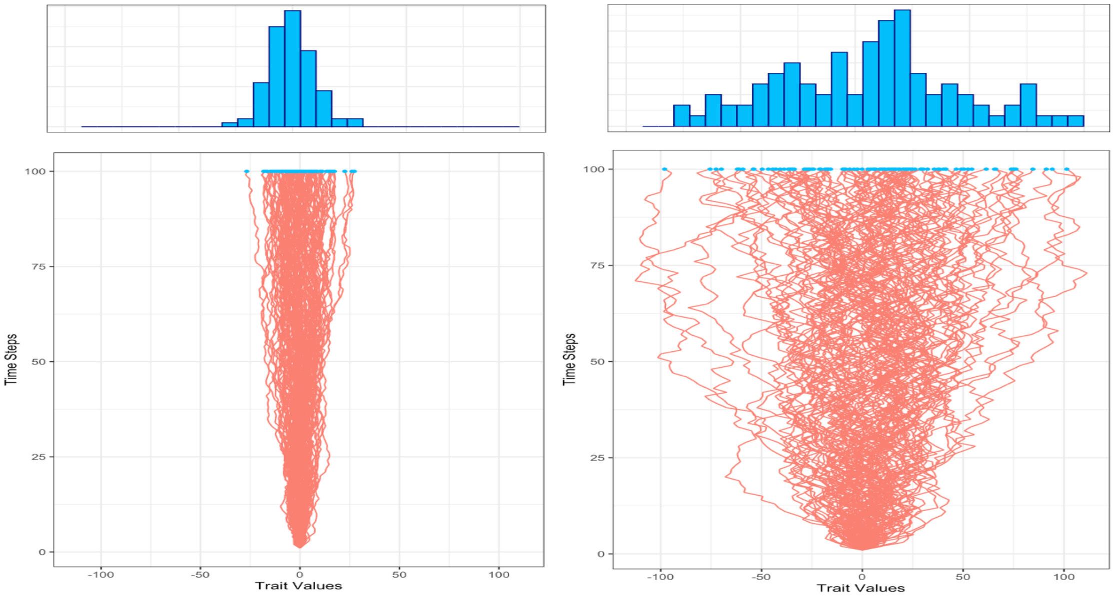

Figure 2: Trajectories of Brownian motion. One hundred trajectories are generated using rate $\sigma = 1$ (bottom left), and $\sigma = 4$ (bottom right), respectively. The sample path on each trajectory is composed of 100 samples drawn from the Brownian motion. The histograms are plotted using the endpoints of the trajectories.

It can be seen straightforwardly that while a smaller rate ($\sigma = 1$) yields a narrower range of the character value, the larger rate ($\sigma = 4$) yields a wider range of character value.

Phylogenetic comparative methods (PCMs) are statistical methods that are frequently applied to analyze empirical datasets from ecology and evolution. Many PCMs are able to estimate the rate of evolution of the phenotype for a group of related but different species and then the rate is used to interpret how diverse the phenotype/genotype change. Various works have been proposed to infer the rate of evolution from the maximum likelihood approach (Caetano and Harmon, 2017; Martin et al., 2022; Revell, 2021) to Bayesian approach (May and Moore, 2020; Eastman et al., 2011) introduce a Bayesian approach for identifying complex patterns in the evolution of continuous traits. The method (auteur) uses reversible-jump Markov chain Monte Carlo sampling to more fully characterize the complexity of trait evolution, considering models that range in complexity from those with a single global rate to potential ones in which each branch in the tree has its own independent rate.

The model in Eq. (2) is applied to estimate the common rate for a group of related species(Felsenstein, 1985). More models were created based on Eq. 2 and involved more sophisticated scenarios such as multiple rates on different clades (OMeara et al., 2006) and multivariate rates for multiple traits (Adams, 2014) Maddison et al. (2007) (Uyeda et al., 2011) Gingerich (2001) Kratsch and McHardy (2014) Kühnert et al. (2016) Uyeda and Harmon (2014).

1.1 Trait Evolution with Time *Uncorrelated* Rate

For a group of related species, estimation of the rate of evolution can be conducted by the phylogenetic comparative methods (PCMs)(Uyeda et al., 2018; Garamszegi, 2014; Cornwell and Nakagawa, 2017). PCMs are statistical models that assume trait variables y_t follow a certain random process along a rooted phylogenetic tree $\mathcal{T}$ with known topology and branch lengths (evolutionary time). The most commonly used continuous random process for trait evolution is the Brownian motion model as shown in Eq. (2).

A Brownian motion variable y_t with a single rate parameter σ is a continuous random process with state space $\mathcal{R}$ with $y(0) = y_0$, stationary increment (i.e. $y_t - y_s$ has the same distribution with y_{t-s}), independent increment (i.e. $y_{t_i} - y_{t_{i-1}}, i = 1, 2, \cdots, n$) are independent, normal distribution. Let $y_0 = \mu$ be the ancestral value for a trait of a species. Then the trait variable y_t at time t can be defined as $y_t = \mu + \sigma W_t$, where $W_t \sim \mathcal{N}(0, t)$ is a random variable with variance t. Hence y_t is a Brownian motion variable with mean μ and variance $\sigma^2 t$.

With the assumption that the evolution of a species change over time, the rates of evolution can be either modeled as constants or stochastic variables along times (branch lengths). However, the constant rates BM model may not be well addressed the evolution in many scenarios. Models that expand from the BM model have been developed where the rate parameter is a function of time (i.e. $\sigma = \sigma_t$). The trait variable y_t adopts the following dynamics

$$y_t = \mu + \sigma_t W_t, \tag{4}$$

where σ_t can be constant (i.e. $\sigma_t = \sigma, t > 0$) (Felsenstein, 1985), piecewise constant (i.e. $\sigma_t = \sigma_i, \gamma_{i-1} \leq t \leq \gamma_i$ where γ_i and γ_{i-1} are the successive time regimes) (OMeara et al., 2006) or a random variable modeled by another pertinent process (i.e. $\sigma_t \sim \mathcal{D}(t)$ where $\mathcal{D}(t)$ is a distribution function for the stochastic variable) (Jhwueng and Maroulas, 2016; Jhwueng, 2020). Those models have been applied to study the trait evolution (Hunt, 2012; Rabosky et al., 2013; Landis and Schraiber, 2017; Soul and Wright, 2021; Chira et al., 2018). Martin et al. (2022) developed a trait evolution model that allows the rate to change gradually and stochastically across a clade, allowing for flexible modeling of early/late bursts of trait evolution.

Trajectories of the trait variable y_t shown in Eq. (4) simulated under a continuous random process evolved along a tree of four taxa using a constant rate vs. two rates are shown in Figure 3.

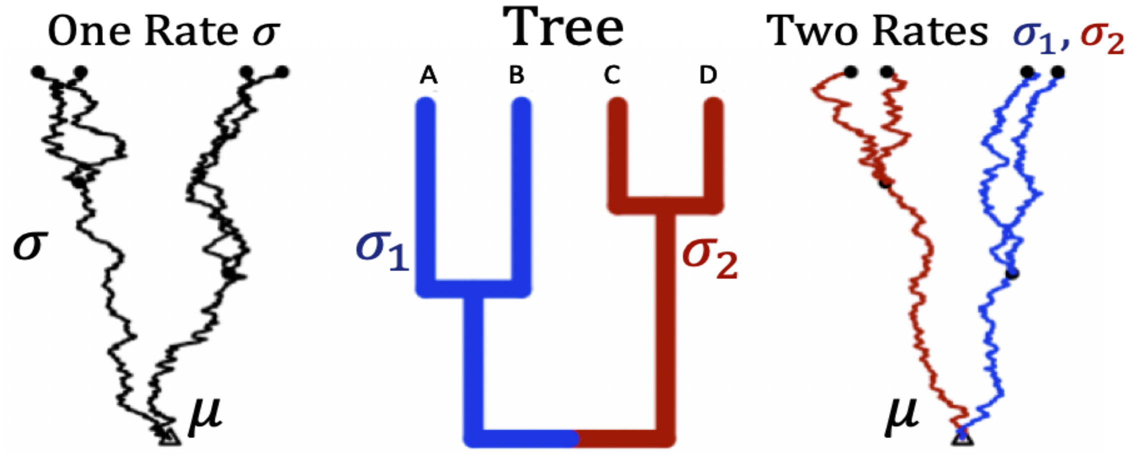

Figure 3: Trajectories of trait evolution along a rooted phylogenetic tree. Middle panel: four taxa rooted tree of 4 tips A, B, C, D. Left panel: a set of four dependent trajectories along the tree using a single rate ($\sigma_t = \sigma$ on all branches). Right panel: a set of four dependent trajectories along the tree using two rates ($\sigma_t = \sigma_1$ on the blue branches, and $\sigma_t = \sigma_2$ on the red branches).

The models have been broadly applied in many studies. For instance, in the morphology evolution of the world's largest flowers (*Raffkesianeae*: up to 1 meter in diameter), Davis et al. (2007) found that the enormous flowers evolved from ancestors with tiny flowers. The increment of the flower diameter (cm) along a stem lineage is approximately 189.1 mm $/2.4$ mm ≈ 78.79 folds in size in a period of ca. 46 million years and a likelihood ratio test rejects the hypothesis of a single rate of flower size evolution in the entire Euphorbiaceae-plus-Rafflesiaceae clade. In the study of the evolution of plant genome size, Beaulieu et al. (2012) found that the woody lineages had a rate of stochastic motion that was nearly five times slower than the rate of herbaceous lineages.

Although the rate models have been developed under the existing framework, a scenario that considers the rate as a time-correlated stochastic variable could provide an extra feature and be useful for studying the rate evolution. The time series model has been well recognized in many fields (see the application in finance (Tsay, 2005) and in engineering (Pham and Yang, 2010). In this proposal, we aim to answer a couple of research questions

Research Question

- **Are those evolutionary rates of biological traits statistically non-independent?**
- or **shall they be believed to be serially autocorrelated?**

Sakamoto and Venditti (2018) investigated whether or not the evolutionary rates of such biological traits are statistically non-independent by verifying testing whether the rate contains phylogenetic signal (statistics to measure how closely related biological taxa share similar traits.) The commonly used statistics are Pagel's λ statistic measures the strength of trait heritability from the ancestor (Pagel, 1999); Bloomberg's K statistic which measures the relatedness of species in a clade when compared to randomly selected species from the same tree (Blomberg et al., 2003). And for the trait, Ackerly (2009) developed methods to detect the trait(not rates) relatedness.

Consider that implementation of the time-correlated rate evolution σ_t could possibly provide an alternative to reveal embedded information in the species evolution, in this work, we intend to use the model $y_t = \mu + \sigma_t W_t$ in Eq. (4) under the framework of correlated rate evolution ($\sigma_t = g(\sigma_s|\Theta)$ for $s < t$ where Θ is a parameter vector) to model the trait evolution for phylogenetic comparative analysis. In particular, we use the generalized autoregressive conditional heteroskedastic (GARCH) time series model that has been broadly applied in econometrics for modeling the rate parameter (Bollerslev, 1986; Engle, 2001). We give a brief description of the GARCH model and its variants in the following section. Later on, we will implement the GARCH for comparative analysis.

2 Methods

2.1 The GARCH Model

In econometrics, the autoregressive conditional heteroskedasticity (ARCH) model is a statistical model for time series data that describes the variance of the current error term or innovation as a function of the actual sizes of the previous time periods' error terms; often the variance is related to the squares of the previous innovations. The ARCH model is appropriate when the error variance in a time series follows an autoregressive (AR) model; if an autoregressive moving average (ARMA) model is assumed for the error variance, the model is a generalized autoregressive conditional heteroskedasticity (GARCH) model. Hence, the GARCH models are used when the variance of the error term is not constant. That is, the error term is heteroskedastic which has an irregular pattern in a statistical model.

In literature, the generalized autoregressive conditional heteroskedasticity (GARCH) model is commonly applied for estimating stochastic volatility (Tsay, 2005). The assumption of the randomness of the variance process varies with the variance. GARCH is very commonly used in analyzing time-series data where the variance error is believed to be serially autocorrelated and where the variance of the error term is assumed to follow an autoregressive moving average process. Suppose that a time series y_t with respect to an information filtration $\mathscr{F}_t$ can be written as

$$y_t = \mu + \epsilon_t, \text{ where } \epsilon_t = \sigma_t z_t \tag{5}$$

where μ is a constant, and $z_t|\mathscr{F}_t \sim \mathscr{D}(0,1)$ where $\mathscr{D}(0,1)$ denotes some distribution with zero mean and unit variance.

Therefore, the GARCH models differ in their specification for σ_t^2 (Nelson, 1990; Zakoian, 1994; Hajiramezanali and Amindavar, 2012; Duan, 1997). The classical formula in the GARCH regression model is obtained by assuming an autoregressive moving average equation on an observable variable y_t, the conditional variance being expressed as a linear function of past squared innovation and of its past values. Inspired by Engle (1982), the first GARCH (p,q) model developed by Bollerslev (1986) which posed the assumption that $\epsilon_t|\psi_t \sim \mathcal{N}(0, \sigma_t)$ where ψ_t is the information set (σ-field) of all information through time t. The representation is written as $\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i \sigma_{t-i}^2$ where $\omega > 0$ is the minimum variance allowed, $\alpha_i \geq 0, i = 1, \cdots, q$, and $\beta_i \geq 0, i = 0, 1, 2, \cdots, p$.

In literature, many augmented GARCH models are developed (Li et al., 2018), and they were in particular useful to capture pertinent processes to datasets with their own features (long-term, mean reverting, fluctuation). In particular, the GARCH $(1,1)$ specification has been found to be adequate in most applications. Despite of being serially uncorrelated, the series ϵ_t does not need to be serially independent. The GARCH $(1,1)$ process is described as follows. We say that $\epsilon_t \sim$ GARCH if we can write $\epsilon_t = \sigma_t z_t$, where z_t is standard Gaussian and σ_t has a series dependent relationship shown in Eq. (6)

$$\sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2, \tag{6}$$

where $\omega > 0, \alpha \geq 0, \beta \geq 0$.

In the following section, we implement the GARCH to model the rate of evolution along the tree.

2.2 Trait Evolution with Time *Correlated* Rate

We describe our model in two steps. In 2.2.1, the connection between the GARCH and trait evolution is defined where In 2.2.2, the incorporation of the tree to the model associated the tree branch length is provided.

2.2.1 Step 1: GARCH rate for trait evolution

Let the trait variable y_t adopt a pertinent process $y_t = \mu + \epsilon_t$, where μ is the expected ancestral value and ϵ_t is a zero-mean white noise. To implement the GARCH $(1,1)$ model for studying the rate of evolution, initially one can consider the rate σ_t adopt the GARCH model shown in Eq. (6). The illustration of the time-correlated model for the rate of trait evolution is shown in Figure 4.

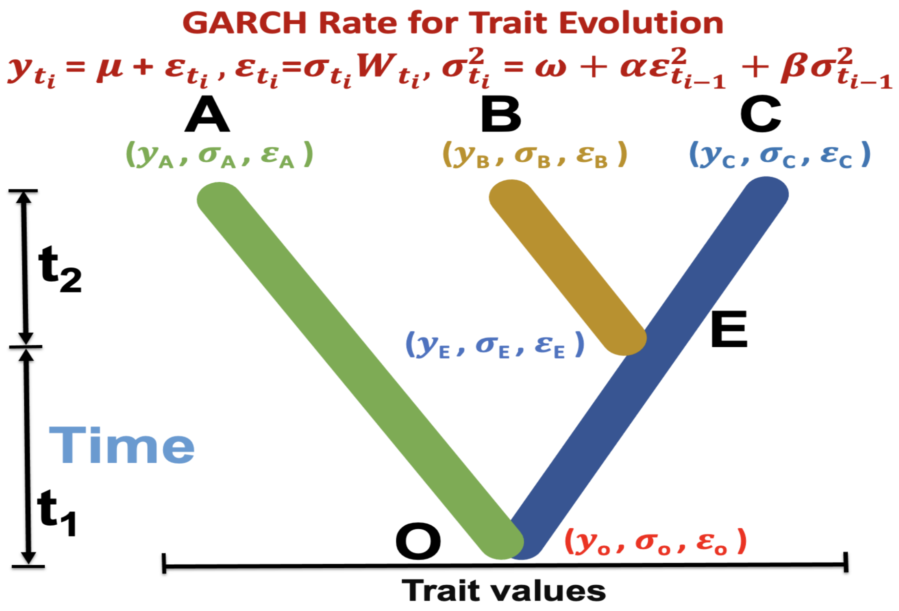

Figure 4: Preliminary scenario for modeling rate of trait evolution via directly using GARCH $(1,1)$ process along a three taxa tree A, B, C.

From Figure 4, the evolution started at the root at node O with unknown ancestral status $(y_0, \sigma_0, \epsilon_0)$. At node E, speciation occurred when two species B and C were formed and evolved to the tip.

The unknown ancestral status $(y_E, \sigma_E, \epsilon_E)$. For instance, the rate at the node E can be modeled by the GARCH model conditioned on its ancestor $(y_0, \sigma_0, \epsilon_0)$. And then the tip status $(y_B, \sigma_B, \epsilon_B)$ at node B can be realized by the GARCH model for the rate of evolution conditioned on its ancestor $(y_E, \sigma_E, \epsilon_E)$. The full scenario for the trait evolution can be realized using the tree traversal algorithm (Pfaff, 1998).

When focusing on studying the rate parameter, concerning that the study of interest is to specify the distribution of σ_t on internal nodes and tips given posterior of parameter, the model for the trait evolution that implements the GARCH model for the rate is described in the following.

For modeling trait evolution, consider the diffusion process $y_t = \mu + \epsilon_t = \mu + \sigma_t z_t$. Instead of setting $z_t \sim \mathscr{D}(0, 1)$ (unit variation) in the conventional GARCH framework, we replace z_t by a Wiener process W_t shown in Eq. (7)

$$z_t := W_t \sim \mathcal{N}(0, t). \tag{7}$$

Next, we denote anc and des as the ancestral and the descendant status at the time t, the relationship between anc and des can be written as

$$y_{\text{des}} = \mu_{\text{des}} + \epsilon_{\text{des}} \text{ where } \epsilon_{\text{des}} = \sigma_{\text{des}} W_{t_{\text{anc,des}}}, \tag{8}$$

where $W_{t_{\text{anc,des}}} \sim \mathcal{N}(0, t_{\text{anc,des}})$ denotes the Wiener process of time length $t_{\text{anc,des}}$ for an ancestor evolved to its descendant.

By considering time series model in Eq. (6), the GARCH rate model for trait evolution is shown in Eq. (9)

$$\sigma^2_{\text{des}} = \omega + \alpha\epsilon^2_{\text{anc}} + \beta\sigma^2_{\text{anc}}, \tag{9}$$

where the subscripts anc and des are the nodes of a successive ancestor-descendant with time length $t_{\text{anc,des}}$.

When in Eq. (9) the $\sigma^2_{\text{des}} = \omega$ is a constant, the model in Eq. (8) reduce to the Brownian motion model $y_{\text{des}} = \mu_{\text{des}} + \omega W_{\text{anc,des}}$. In that case, where setting α and β to zero would yield the same likelihood as the Brownian motion model.

Notice that Hansen et al. (2021) proposed regression models for studying the rate of evolution σ_t of the response trait y_t via building up a relationship with another predicted trait variable x_t. Their models assume that the predictor follows the Brownian motion (i.e., $x_t = x_0 + \sigma_x W^x_t$) or geometric Brownian motion predictor (i.e., $x_t = x_0 \exp\left((\mu - \frac{\sigma_x^2}{2})t + \sigma_x W^x_t\right)$. And the model can be written in a system of stochastic differential equations as $dy_t = \sigma_t dW^y_t$, where $\sigma^2_t = \omega + bx_t$. Notice that the σ_t in Hansen et al. (2021) model is *not self–correlated*, rather it is correlated to another predictor x_t. Therefore our new model is distinguishable from Hansen's model in the sense of the autocorrelated relationship of the σ_t.

2.2.2 Step 2: modeling GARCH rate on tree

For trait evolution, we incorporate the branch lengths to the modified GARCH(1,1) model in Eq. (10)

$$\sigma^2_{\text{des}} = \omega + \frac{\alpha}{l_{\text{des}}}\sigma^2_{\text{anc}} W^2_{l_{des}} + \frac{\beta}{l_{\text{des}}}\sigma^2_{\text{anc}}, \kappa \geq 0, \alpha > 0, \tag{10}$$

where ω is the minimum variance allowed, we use the estimator from the ridge regression via `RRphylo` (Castiglione et al., 2018).

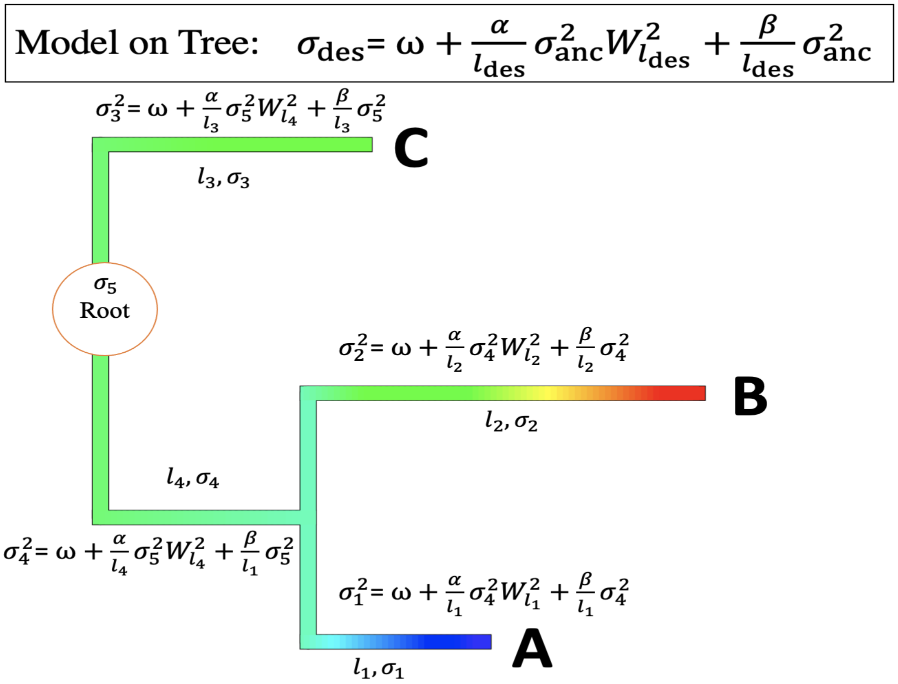

Figure 5: GARCH $(1, 1)$ for the rate of evolution on the tree

The Model Likelihood

To perform model inference on parameter estimation, given trait vector $\boldsymbol{Y} = (y_1, y_2, \cdots, y_n)^t$ on the tip, and a phylogenetic tree $\mathcal{T}$ with known branch length set $\mathscr{B} = \{b_0, b_1, b_2, \cdots, b_m\}$ where $b_0 = 0$ is the root node branch length, and each branch $b_k, k \geq 1$ represents as the time from t_{k-1} to t_k. Use $\boldsymbol{Y}$ to get the rate estimates at the root σ_0. We are interested in the joint distribution $L(\Theta|\boldsymbol{Y}, \boldsymbol{\sigma}^2, \boldsymbol{\epsilon}_{\mathscr{B}}, \mathcal{T})$ where Θ is the parameter vector (i.e. $\Theta = (\omega, \alpha, \beta)$ for the GARCH $(1, 1)$. The log-likelihood on a branch is a 1 dimensional GARCH $(1, 1)$ which has likelihood

$$\log L(\Theta|\boldsymbol{Y}, \sigma^2_{b_k}, \epsilon_{b_k}, \mathcal{T}) = \frac{1}{2}\left(-\log(2\pi) - \log(\sigma^2_{b_k}) - \frac{\epsilon^2_{b_k}}{\sigma^2_{b_k}}\right). \tag{11}$$

Hence, the full log likelihood given the tree and trait can be written as

$$\log L(\Theta|\boldsymbol{Y}, \boldsymbol{\sigma}^2, \boldsymbol{\epsilon}_{\mathscr{B}}, \mathcal{T}) = \log L(\Theta|\epsilon_0, \sigma_0^2, \mathcal{T}) + \sum_{b_k \in \mathscr{B}} \log L(\Theta|\boldsymbol{Y}, \sigma^2_{b_k}, \epsilon_{b_k}, \mathcal{T}), \tag{12}$$

where ϵ_{b_0} is the residual at the root and the product operator $b_k \in \mathscr{B}$ follows a tree topology.

Once the likelihood is determined, parameter estimation through maximum likelihood estimation (Zivot, 2017; Francq and Zakoian, 2007) or Bayesian approach (Ardia and Hoogerheide, 2010) is a feasible option. As shown in Eq. (12) the whole likelihood takes the tree as a input where on each branch a single likelihood is determined by Eq. (11), we adopt the Bayesian approach as shown the Algorithm 1.

Algorithm 1 Metropolis Hasting Sampling Algorithm for GARCH Rate Model of Trait Evolution.

```
Require: Trait data at the tip $\boldsymbol{Y}$, tree $\mathcal{T}$ with branch length $\mathscr{B} = \{b_k\}$, initial parameters values $\Theta_0$,
    a prior distribution $\pi$.
Ensure: MCMC samples $\Theta_j, j = 1, 2, \cdots, L$ from the posterior distribution.
Initial Step:
Use $\boldsymbol{Y}$ and $\mathcal{T}$ to get the rate estimates at the root $\sigma_0$.
Set $\epsilon_0 = 0$.
Set start value $\Theta_0 = (\omega_0, \alpha_0, \beta_0)$.
Compute $\ell_0 = \log L(\Theta_0 | \epsilon_0, \sigma_0^2, \mathcal{T})$.
for j=1:L do
    Draw proposal for $\Theta_1 = (\omega, \alpha, \beta)$.
    ##Apply tree traversal algorithm to simulate samples at internal node.
    for $b_k \in \mathscr{B}$ do
        Use $\Theta_1, \epsilon_{b_{k-1}}, \sigma_{b_{k-1}}$ to compute $\sigma_{b_k} = \omega + \frac{\alpha}{b_k}\epsilon^2_{b_{k-1}} + \frac{\beta}{b_k}\sigma_{b_{k-1}}$.
        Use $\sigma_{b_k}$ and $b_k$ to draw sample $\epsilon_{b_k} \sim \mathcal{N}(0, \sigma^2_{b_k} t^2_{b_k})$.
    end for
    Compute the likelihood $\ell_1$ in Eq. (12).
    Set $\delta_j = \min\{1, \exp(\ell_1 + \log \text{pr}(\Theta_1) - \ell_0 - \log \text{pr}(\Theta_0))\}$.
    Draw $u$ from $\mathcal{U}[0, 1]$.
    if $u < \delta_j$ (accept proposal)
        Set $\Theta_j = \Theta_1$, $\ell_0 = \ell_1$.
    else
        Set $\Theta_j = \Theta_{j-1}$.
end for
return $\Theta_j, j = 1, 2, \cdots, L$.
```

3 Ongoing Work

3.1 Simulation

We assess the model performance using simulation. Four types of trees: star tree, balanced tree, left tree, coalescent tree, and birth-death tree each of size $16, 32, 64, 128$ are used for assessment. Initial samples are drawn from an independent normal distribution The initial estimate for σ at the root is estimated by the Brownian motion model using the `R` package `geiger`. For each type of tree, for each data, and for each prior set, a million replicates of the sample will be generated to assess the posterior sample distribution.

Currently, the estimation of one parameter under MCMC seems to work quite well and the estimation of two parameters under MCMC still works well, too. However, the estimation of three parameters simutaneously under MCMC seems not ok because ω is hard to estimate. Instead, we fixed ω to the root value μ estimated from the Brownian motion model, so the focus parameters would be the α and β. Given the trait dataset parameter of ω is estimated fixed by the Brownian motion model, then $\hat{\omega}$ is directly used for the next step of estimation for the other two parameters. The trajectories for the parameters α and β for a taxa size of 128 using a balanced tree are shown in Figure 6.

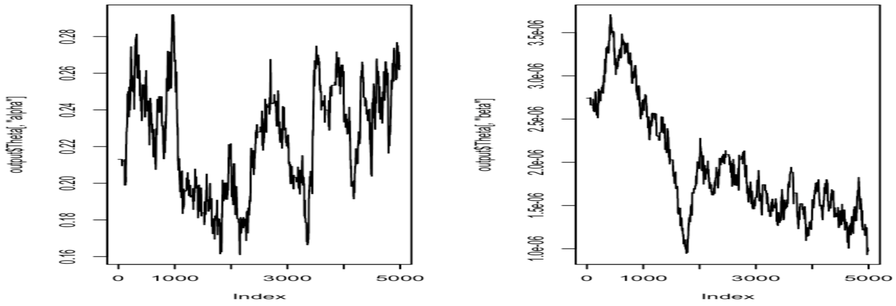

Figure 6: The trajectories of the phygarch model for 128 taxa under a balanced tree.

The posterior distributions of the parameters α and β for a taxa size of 128 using a balanced tree are shown in Figure 7.

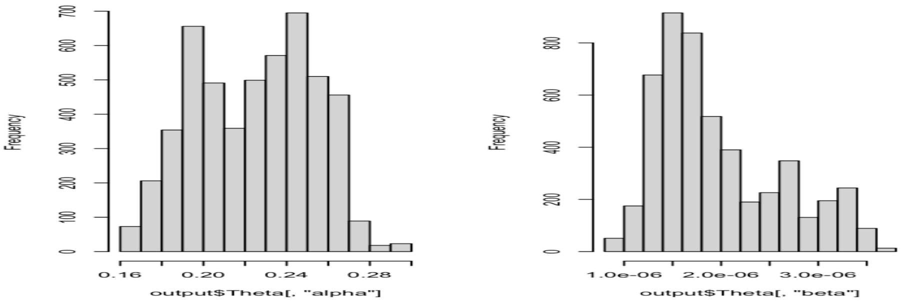

Figure 7: The posterior distribution of parameter α and β using a 128 taxa balanced tree.

3.1.1 Testing and Model comparison

Testing: To answer the research question: are those evolutionary rates of biological traits statistically non-independent ? or shall they be believed to be serially autocorrelated ? We are interested in testing assumptions of the form testing the nullity of

$$H_0 : \alpha = 0 \text{ or } H_0 : \beta = 0. \tag{13}$$

Conventionally, this is for instance the case when a GARCH $(0, 1)$ (or a GARCH $(1, 0)$ is tested against a GARCH $(1, 1)$). In practice, the most widely used test for a simple hypothesis is the so-called t–ratio defined, in the case of Eq. (13), by $t = \hat{\alpha}/\text{sd}(\hat{\alpha})$ or $t = \hat{\beta}/\text{sd}(\hat{\beta})$ with standard notations. We expect to explore further via simulation.

Model comparison: The proposed models here can also be compared with the popular Brownian motion (OMeara et al., 2006) and Ornstein Uhlenbeck model (Hansen, 1997) using the Bayes factor (Lee and Wagenmakers, 2014; Kass and Raftery, 1995).

3.2 Empirical Analysis

We do not have good datasets yet to validate the new models and distinguish our models from other well known models. And we plan to search the data that contains certain signals of the correlated rate (β) conditioned on the ancestor.

4 Appendix

4.1 Code and Scripts

以下連結爲此計畫目前所開發的程式，所有的相關資料可以由以下連結取得 `https://tonyjhwueng.info/phygarchrate`。

1. Figure 1: `https://tonyjhwueng.info/phygarchrate/thorne.jpg`
2. Figure 2: `https://tonyjhwueng.info/phygarchrate/twosigbmv3.html`
3. Figure 3: `https://tonyjhwueng.info/phygarchrate/4spetreemap.html`
4. Figure 4: `https://tonyjhwueng.info/phygarchrate/3taxaTraitdemo.png`
5. Figure 5: `https://tonyjhwueng.info/phygarchrate/garchrate.pptx`
6. Figures 6, 7: `https://tonyjhwueng.info/phygarchrate/phygarch_sims_balanced.html`

References

Ackerly, D. (2009). Conservatism and diversification of plant functional traits: evolutionary rates versus phylogenetic signal. *Proceedings of the National Academy of Sciences 106*(supplement_2), 19699–19706.

Adams, D. C. (2014). A method for assessing phylogenetic least squares models for shape and other high-dimensional multivariate data. *Evolution 68*(9), 2675–2688.

Ardia, D. and L. F. Hoogerheide (2010). Bayesian estimation of the garch (1, 1) model with student-t innovations. *The R Journal 2*(2), 41–47.

Beaulieu, J., D.-C. Jhwueng, C. Boettiger, and B. O'Meara (2012). Modeling stabilizing selection: expanding the Ornstein-Uhlenbeck model of adaptive evolution. *Evolution 66*(8), 2369–2383.

Blomberg, S. P., T. Garland Jr, and A. R. Ives (2003). Testing for phylogenetic signal in comparative data: behavioral traits are more labile. *Evolution 57*(4), 717–745.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics 31*(3), 307–327.

Caetano, D. S. and L. J. Harmon (2017). ratematrix: an r package for studying evolutionary integration among several traits on phylogenetic trees. *Methods in Ecology and Evolution 8*(12), 1920–1927.

Castiglione, S., G. Tesone, M. Piccolo, M. Melchionna, A. Mondanaro, C. Serio, M. Di Febbraro, and P. Raia (2018). A new method for testing evolutionary rate variation and shifts in phenotypic evolution. *Methods in Ecology and Evolution 9*(4), 974–983.

Chira, A.-M., C. Cooney, J. A. Bright, E. Capp, E. Hughes, C. Moody, L. Nouri, Z. Varley, and G. Thomas (2018). Correlates of rate heterogeneity in avian ecomorphological traits. *Ecology letters 21*(10), 1505–1514.

Cornwell, W. and S. Nakagawa (2017). Phylogenetic comparative methods. *Current Biology 27*(9), R333–R336.

Davis, C. C., M. Latvis, D. L. Nickrent, K. J. Wurdack, and D. A. Baum (2007). Floral gigantism in rafflesiaceae. *Science 315*(5820), 1812–1812.

Duan, J.-C. (1997). Augmented garch (p, q) process and its diffusion limit. *Journal of Econometrics 79*(1), 97–127.

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: International Journal of Organic Evolution 65*(12), 3578–3589.

Engle, R. (2001). Garch 101: The use of arch/garch models in applied econometrics. *Journal of economic perspectives 15*(4), 157–168.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica: Journal of the econometric society*, 987–1007.

Felsenstein, J. (1985). Phylogeny and the comparative method. *America Naturalist 125*(1), 1–15.

Francq, C. and J.-M. Zakoian (2007). Quasi-maximum likelihood estimation in garch processes when some coefficients are equal to zero. *Stochastic Processes and their Applications 117*(9), 1265–1284.

Garamszegi, L. Z. (2014). *Modern phylogenetic comparative methods and their application in evolutionary biology: concepts and practice*. Springer.

Gingerich, P. D. (2001). Rates of evolution on the time scale of the evolutionary process. *Microevolution rate, pattern, process*, 127–144.

Hajiramezanali, M. and H. Amindavar (2012). Maneuvering target tracking based on sde driven by garch volatility. In *2012 IEEE Statistical Signal Processing Workshop (SSP)*, pp. 764–767. IEEE.

Hansen, T. (1997). Stabilizing selection and the comparative analysis of adaptation. *Evolution 51*, 1341–1351.

Hansen, T. F., G. H. Bolstad, and M. Tsuboi (2021). Analyzing disparity and rates of morphological evolution with model-based phylogenetic comparative methods. *Systematic Biology*.

Hunt, G. (2012). Measuring rates of phenotypic evolution and the inseparability of tempo and mode. *Paleobiology 38*(3), 351–373.

Jhwueng, D.-C. (2020). Modeling rate of adaptive trait evolution using cox–ingersoll–ross process: An approximate bayesian computation approach. *Computational Statistics & Data Analysis 145*, 106924.

Jhwueng, D.-C. and V. Maroulas (2016). Adaptive trait evolution in random environment. *Journal of Applied Statistics 43*(12), 2310–2324.

Kass, R. and A. Raftery (1995). Bayes factors, in journal of the american statistical association.

Kratsch, C. and A. C. McHardy (2014). Ridgerace: ridge regression for continuous ancestral character estimation on phylogenetic trees. *Bioinformatics 30*(17), i527–i533.

Kühnert, D., T. Stadler, T. G. Vaughan, and A. J. Drummond (2016). Phylodynamics with migration: a computational framework to quantify population structure from genomic data. *Molecular biology and evolution 33*(8), 2102–2116.

Landis, M. J. and J. G. Schraiber (2017). Pulsed evolution shaped modern vertebrate body sizes. *Proceedings of the National Academy of Sciences 114*(50), 13224–13229.

Lee, M. D. and E.-J. Wagenmakers (2014). *Bayesian cognitive modeling: A practical course*. Cambridge university press.

Li, M., F. Mercurio, and S. Resnick (2018). The garch linear sde: Explicit formulas and the pricing of a quanto cds. *Available at SSRN 3188272*.

MacFadden, B. J. (2005). Fossil horses–evidence for evolution. *Science 307*(5716), 1728–1730.

Maddison, W. P., P. E. Midford, and S. P. Otto (2007). Estimating a binary character's effect on speciation and extinction. *Systematic biology 56*(5), 701–710.

Martin, B. S., G. S. Bradburd, L. J. Harmon, and M. G. Weber (2022). Modeling the evolution of rates of continuous trait evolution. *bioRxiv*.

May, M. R. and B. R. Moore (2020). A bayesian approach for inferring the impact of a discrete character on rates of continuous-character evolution in the presence of background-rate variation. *Systematic biology 69*(3), 530–544.

Nelson, D. B. (1990). Arch models as diffusion approximations. *Journal of econometrics 45*(1-2), 7–38.

OMeara, B., C. Ané, M. Sanderson, and P. Wainwright (2006). Testing different rates of continuous trait evolution using likelihood. *Evolution 60*, 922–933.

Pagel, M. (1999). Inferring the historical patterns of biological evolution. *Nature 401*(6756), 877.

Pettay, J. E., L. E. Kruuk, J. Jokela, and V. Lummaa (2005). Heritability and genetic constraints of life-history trait evolution in preindustrial humans. *Proceedings of the National Academy of Sciences 102*(8), 2838–2843.

Pfaff, B. (1998). An introduction to binary search trees and balanced trees. *Libavl Binary Search Tree Library 1*, 19–20.

Pham, H. T. and B.-S. Yang (2010). Estimation and forecasting of machine health condition using arma/garch model. *Mechanical Systems and Signal Processing 24*(2), 546–558.

Podos, J. and S. Nowicki (2004). Beaks, adaptation, and vocal evolution in darwin's finches. *Bioscience 54*(6), 501–510.

Rabosky, D. L., F. Santini, J. Eastman, S. A. Smith, B. Sidlauskas, J. Chang, and M. E. Alfaro (2013). Rates of speciation and morphological evolution are correlated across the largest vertebrate radiation. *Nature communications 4*(1), 1–8.

Revell, L. J. (2021). A variable-rate quantitative trait evolution model using penalized-likelihood. *PeerJ 9*, e11997.

Sakamoto, M. and C. Venditti (2018). Phylogenetic non-independence in rates of trait evolution. *Biology letters 14*(10), 20180502.

Scott-Phillips, T. C., K. N. Laland, D. M. Shuker, T. E. Dickins, and S. A. West (2014). The niche construction perspective: a critical appraisal. *Evolution 68*(5), 1231–1243.

Soons, J., A. Herrel, A. Genbrugge, P. Aerts, J. Podos, D. Adriaens, Y. De Witte, P. Jacobs, and J. Dirckx (2010). Mechanical stress, fracture risk and beak evolution in darwin's ground finches (geospiza). *Philosophical Transactions of the Royal Society B: Biological Sciences 365*(1543), 1093–1098.

Soul, L. C. and D. F. Wright (2021). *Phylogenetic Comparative Methods: A User's Guide for Paleontologists*. Cambridge University Press.

Thorne, J. L., H. Kishino, and I. S. Painter (1998). Estimating the rate of evolution of the rate of molecular evolution. *Molecular biology and evolution 15*(12), 1647–1657.

Tsay, R. S. (2005). *Analysis of financial time series*, Volume 543. John wiley & sons.

Uyeda, J. C., T. F. Hansen, S. J. Arnold, and J. Pienaar (2011). The million-year wait for macroevolutionary bursts. *Proceedings of the National Academy of Sciences 108*(38), 15908–15913.

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. *Systematic biology 63*(6), 902–918.

Uyeda, J. C., R. Zenil-Ferguson, and M. W. Pennell (2018). Rethinking phylogenetic comparative methods. *Systematic Biology 67*(6), 1091–1109.

Xiang, Y., C.-H. Huang, Y. Hu, J. Wen, S. Li, T. Yi, H. Chen, J. Xiang, and H. Ma (2017). Evolution of rosaceae fruit types based on nuclear phylogeny in the context of geological times and genome duplication. *Molecular biology and evolution 34*(2), 262–281.

Zakoian, J.-M. (1994). Threshold heteroskedastic models. *Journal of Economic Dynamics and control 18*(5), 931–955.

Zivot, E. (2017). *Introduction to computational finance and financial econometrics*. Chapman & Hall Crc.