Optimizing Variable Selection in Phylogenetic Eigenvector Regression for Trait Correlation Analysis

Zheng-Lin Chen and Deng-Ke Niu*

MOE Key Laboratory for Biodiversity Science and Ecological Engineering and Beijing Key Laboratory of Gene Resource and Molecular Development, College of Life Sciences, Beijing Normal University, Beijing, 100875, China

*Correspondence author:
Deng-Ke Niu, College of Life Sciences, Beijing Normal University, Beijing, 100875, China.
dkniu@bnu.edu.cn, dengkeniu@hotmail.com

Abstract

Phylogenetic autocorrelation refers to the similarity in traits among closely related species due to shared evolutionary history, violating the assumption of independence in conventional statistical analyses. Phylogenetic eigenvector regression (PVR) addresses this issue by incorporating phylogenetic eigenvectors derived from phylogenetic trees to control for shared ancestry and isolate trait correlations. In this study, we investigated how different selections of dependent and independent variables affect eigenvector identification and trait correlation results. We used simulated data to analyze both fixed-balanced and random phylogenetic trees, which allowed us to record internal node data. This enabled us to determine the true correlation between traits using traditional Pearson and Spearman rank correlations, providing a gold standard for comparison. We found that the eigenvectors selected differed significantly when switching between dependent variables (X_1 and X_2), which led to inconsistent PVR correlation results. We assessed three criteria—R^2, Pagel's λ, and AIC—to determine which dependent variable produced more accurate results. Our results indicate that, under scenarios with significant phylogenetic autocorrelation and larger sample sizes (128 and 1024 species), the criterion of maximizing R^2 consistently provided the more accurate model. This was true across multiple evolutionary models, tree structures, and regression estimators, including L2 (OLS), MM, and L1. However, in smaller samples (16 species), Pagel's λ was more effective as a selection criterion. These findings emphasize the importance of using appropriate criteria, such as R^2 or λ, for selecting dependent variables to enhance the accuracy and consistency of evolutionary inferences. Additionally, our preliminary findings suggest that similar issues of variable selection may affect other spatial statistical tools, such as spatial eigenvector mapping, conditional autoregressive models, simultaneous autoregressive models, and generalized additive models. These insights extend beyond phylogenetic analyses to other contexts where spatial autocorrelation exists or coexists with phylogenetic autocorrelation. Future work is needed to address how to appropriately select dependent and independent variables in these contexts, which remains a key challenge for understanding evolutionary and ecological relationships.

Keywords

phylogenetic autocorrelation; phylogenetic eigenvectors regression (PVR); dependent and independent variable selection; evolutionary correlation analysis; robust regression; phylogenetic signal; spatial autocorrelation.

Introduction

Understanding the correlations between traits in evolutionary biology is crucial for elucidating evolutionary relationships and processes. Due to their common evolutionary history, closely related species share more similar traits. This phenomenon, called phylogenetic autocorrelation, violates the assumption of independence in conventional statistical analyses, potentially leading to inaccurate conclusions when phylogenetic relationships are ignored (Cheverud et al. 1985, Felsenstein 1985, Garamszegi 2014). To address the issue of interdependence among samples, previous researchers have proposed various methods, such as phylogenetic independent contrasts (PIC), phylogenetic generalized least squares (PGLS), and phylogenetic eigenvector regression (PVR) (Felsenstein 1985, Grafen 1989, Diniz−Filho et al. 1998, Pagel 1999, Keitt et al. 2002, Dormann et al. 2007, Peres-Neto and Legendre 2010, Garamszegi 2014, O'Meara 2016, Cornwallis and Griffin 2024).

Among these methods, PVR stands out as a flexible approach by incorporating phylogenetic eigenvectors as covariates, which represent the shared evolutionary history of species to control for phylogenetic relationships and isolate trait correlations (Diniz−Filho et al. 1998, Diniz−Filho and Tôrres 2002). PVR can be traced back to autoregressive models proposed by Cheverud (1985), which decompose studied traits into components influenced by phylogeny (P) and components reflecting independent evolution (S) (Harvey et al. 1995, Westoby et al. 1995, Desdevises et al. 2003). PVR entails performing principal coordinate analysis on a double-centred phylogenetic matrix ($n \times n$), generating $n - 1$ orthogonal phylogenetic eigenvectors. These eigenvectors represent species' positions on the phylogenetic tree and can be incorporated as covariates into models to account for phylogenetic relationships, allowing for partial regression between traits while controlling for shared phylogenetic history (Desdevises et al. 2003, Safi and Pettorelli 2010, Diniz−Filho et al. 2011, Peres-Neto et al. 2012).

A key challenge, however, lies in the careful selection of phylogenetic eigenvectors. Including all eigenvectors as covariates may lead to overfitting, where the model excessively attributes variation in the dependent variable to phylogenetic effects, thus obscuring meaningful trait correlations (Rohlf 2001). On the other hand, excluding relevant phylogenetic eigenvectors can result in the loss of crucial phylogenetic information. Fortunately, research has demonstrated that not all phylogenetic eigenvectors are necessary for robust analysis, as many contribute minimally to the overall model. By carefully selecting a subset of phylogenetic eigenvectors, it is often possible to capture sufficient phylogenetic information and avoid overfitting (Peres-Neto et al. 2005, Griffith and Peres-Neto 2006, Diniz−Filho et al. 2009, Diniz−Filho et al. 2012). These selected eigenvectors help to isolate the effects of phylogeny from other sources of variation, ultimately providing a clearer understanding of trait relationships that are independent of phylogenetic influence (Diniz−Filho and Nabout 2009, Diniz−Filho et al. 2009, Burns et al. 2015, Divíšek et al. 2018, Diniz−Filho et al. 2019, LeRoy et al. 2020, Gao et al. 2023, Lanuza et al. 2023).

Effectively selecting phylogenetic eigenvectors can significantly enhance the performance of PVR, improving its overall efficacy. However, an equally critical issue in eigenvector selection concerns the designation of dependent and independent variables. We are particularly concerned that swapping these variables may lead to different eigenvector selections, which could substantially affect the correlation analysis. As a result, traits may appear correlated under one configuration but uncorrelated—or even inversely correlated—under another, leading to mutually exclusive results.

We constructed simulated data sets (X_1 and X_2) to select phylogenetic eigenvectors, comparing differences between different selections of independent and dependent variables and evaluating their influence on accurately predicting trait correlations. Our analysis revealed conflicting correlation results due to different selections of dependent and independent variables exhibiting a significant occurrence. To accurately depict evolutionary correlations, we established a gold standard based on traits' changes per unit length along the phylogenetic tree branches. Comparing models based on R^2, Pagel's λ (Pagel 1999), and Akaike information criterion (AIC) values, we found that larger R^2 models provided more accurate predictions. After determining variables based on R^2, we compared eigenvector-selection methods for their effects on PVR model fit.

In statistical modeling, traditional methods such as ordinary least squares (OLS) regression rely on assumptions like the normality of errors and the absence of influential outliers. When these assumptions are violated, model performance can be compromised. Robust statistics provides methods that are less sensitive to such deviations, using estimators like M-estimators, L1 estimators, S-estimators, and MM-estimators (Huber 1964, Yu and Yao 2017). These robust estimators adjust model weights to reduce the influence of outliers, improving model stability and accuracy. When applied to phylogenetic data, robust phylogenetic regression can integrate robust estimators into the phylogenetic regression framework, such as PGLS and PIC (Adams et al. 2024). These methods account for the shared evolutionary history of species while reducing the impact of data irregularities. In this paper, we integrated robust estimators into PVR analysis and examined the effects of variable selection on the performance of robust PVR analysis.

Materials and Methods

Study Workflow Overview

This study involves a detailed comparative analysis of PVR with multiple simulation scenarios and regression methods, as well as the addition of robust phylogenetic estimators.

To clarify the complexity of the process, we present a flowchart (Fig. 1) outlining the workflow used in our analysis, consisting of the following key steps:

1. Tree construction: Both random trees and fixed-balanced trees were generated and used in the analysis.
2. Simulation scenarios: We performed simulations under four different scenarios, including combinations of Brownian motion (BM) and normal trait evolution models.
3. Determination of eigenvectors: A subset of eigenvectors was selected based on three methods before performing the robust regression analysis. Notably, eigenvectors selected using the minimizing AIC (mAIC) method were not carried forward to subsequent analysis because the large number of selected eigenvectors resulted in a high degree of multicollinearity with the independent variables. This decision ensured that only eigenvectors that did not induce multicollinearity were used in further analysis.
4. Regression estimators: To address the issue of robustness in the analysis, five robust regression estimators, as described by Adams et al. (Adams et al. 2024), were applied to the PVR models. These included M, L1, S, MM, and the classical **L2 estimator, which corresponds to OLS regression**.
5. Criteria performance in variables identification: The identification of appropriate dependent and independent variables was crucial to obtaining the most suitable combination of eigenvectors. Three criteria, R^2, AIC, and Pagel's λ, were used to evaluate the model performance under different scenarios.

The flowchart (Fig. 1) illustrates the entire process from tree generation to the final evaluation of model performance. For a detailed explanation of each step, please refer to the subsequent subsections in the Materials and Methods.

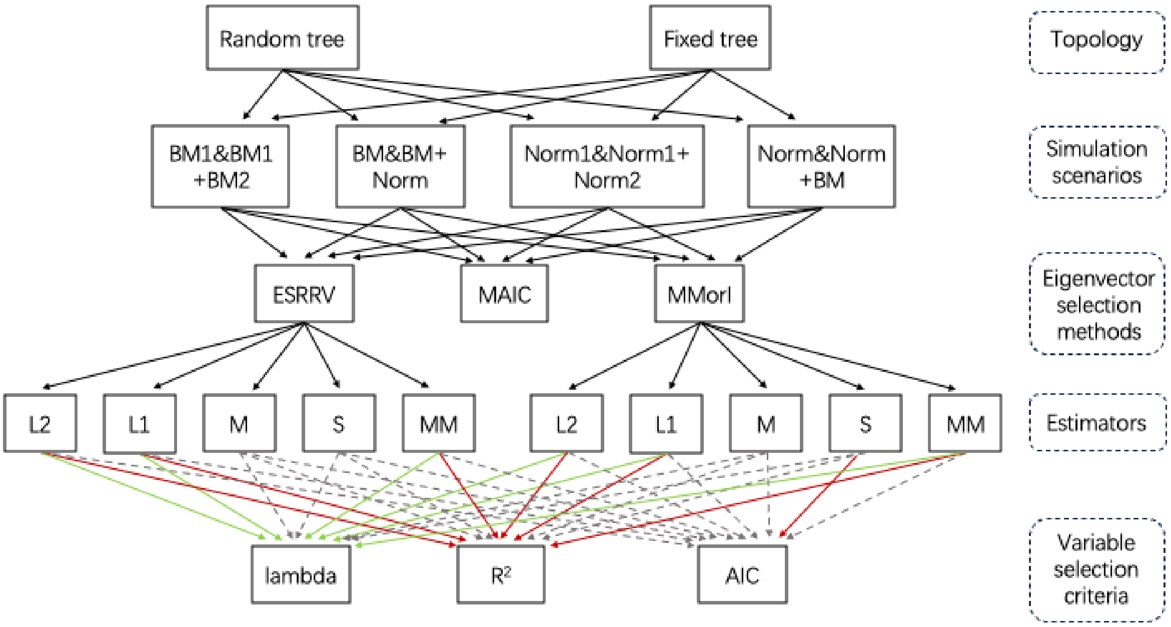
Random tree
Fixed tree
Topology
BM1&BM1 +BM2
BM&BM+ Norm
Norm1&Norm1+ Norm2
Norm&Norm +BM
Simulation scenarios
ESRRV
MAIC
MMorI
Eigenvector selection methods
L2
L1
M
S
MM
L2
L1
M
S
MM
Estimators
lambda
R²
AIC
Variable selection criteria

Figure 1. Workflow for the present study. The analysis was conducted using two tree topologies, followed by simulations under four distinct evolutionary scenarios. A subset of eigenvectors was selected for each round of simulation using three methods. Eigenvectors selected using the mAIC criterion were not used in subsequent analyses because the large number of eigenvectors led to high multicollinearity with the independent variables. Each simulation and its selected eigenvectors were analyzed using five regression estimators, including the classical OLS (i.e., L2) method and four robust phylogenetic estimators. A final outcome of the study was the identification of a selection criterion for dependent and independent variables, which ultimately determined the most appropriate combination of eigenvectors. Model performance was assessed using R^2, AIC, and Pagel's λ. The flowchart summarizes the overall process, providing a visual representation of the detailed steps described in the Materials and Methods section.

Simulations of Phylogenetic Trees

In PVR, the R^2 value is influenced by the eigenvalues of the selected eigenvectors, which reflect the structure of the phylogenetic tree (Rohlf 2001, Diniz−Filho et al. 2012). Given this dependency, changes in the shape of the phylogeny can significantly affect the R^2 value, complicating the interpretation of model performance. To address this issue, we used a fixed phylogeny across all simulations to ensure consistency in tree topology. We selected a balanced tree for the fixed group, as it provides a unique structure for 128 species, simplifying comparisons between evolutionary models compared to the numerous possible non-balanced trees. To evaluate how tree size impacts the efficiency of PVR in capturing phylogenetic signals and the scalability of our methods, we conducted additional simulations using phylogenetic trees with 16 and 1024 species.

Although the fixed-balanced tree offers theoretical advantages, real evolutionary data exhibit a wide variety of tree topologies, and balanced trees are relatively rare. To account for this variability, we repeated our analysis using random phylogenetic trees, where a new phylogenetic tree was generated for each iteration, resulting in different topologies across repetitions. This approach enabled us to test the robustness of our findings under more diverse and realistic tree shapes, ensuring that our results are applicable to a broader range of evolutionary models and tree topologies. All random phylogenetic trees contained 128 species.

We used the *stree* function from the *ape* package (Version 5.8) (Paradis and Schliep 2019) and the *paste.tree* function from the *phytools* package (Version 2.3-0) (Revell 2024) to construct balanced trees, and the *rtree* function from the *ape* package to construct random trees.

Simulated Data

Using the simulated phylogenetic trees described above, we modeled the evolution of two traits across these tree structures, X_1 and X_2, where X_2 was defined as X_1 plus a noise term (ε). The variance of ε was varied to manipulate the correlation strength between traits, inspired by Revell (2010). Four evolutionary scenarios were modeled to explore different aspects of this relationship:

1. **BM1 & BM1+BM2**: Trait X_1 evolves according to one Brownian motion model (variance = 1) on the phylogenetic tree, while X_2 evolves as X_1 plus a noise term ε, where ε evolves under another Brownian motion model with a variance of d. Again, the values of d range from 10^{-4} to 1024.
2. **BM & BM+Norm**: Trait X_1 evolves along the phylogenetic tree following a Brownian motion model (variance = 1), while X_2 evolves as X_1 plus a noise term ε, where ε follows a normal distribution with mean 0 and variance d. Values of d used are 10^{-4}, 0.01, 1, 4, 16, 64, 256, and 1024.
3. **Norm & Norm+BM**: Trait X_1 evolves under a normal distribution (mean = 0, variance =1), while X_2 evolves as X_1 plus a noise term ε, where ε evolves along the phylogenetic tree following a Brownian motion model with a variance of d. The same range of d values is utilized: 10^{-4}, 0.01, 1, 4, 16, 64, 256, and 1024.
4. **Norm1 & Norm1+Norm2**: Trait X_1 follows a normal distribution with a mean of 0 and a variance of 1, while X_2 evolves as X_1 plus a noise term ε, where ε follows a normal distribution with a mean of 0 and a variance of d. The values used for d are consistent with the previous scenarios. This scenario serves as a control with no phylogenetic influence, highlighting the phylogenetic effects present in the first three scenarios, where trait evolution is shaped by phylogenetic autocorrelation.

For each value of variance (d) within each scenario, we conducted 1,000 repeated simulations per tree or set of random trees, resulting in 8,000 simulations per scenario. Given that we have four scenarios in total, this led to a cumulative 32,000 simulations per tree type.

Our study aimed to capture realistic situations where evolutionary pressures differentially influence traits, as well as scenarios where traits become effectively uncorrelated due to high noise variance. This approach allowed us to explore not only varying correlation strengths but also the differential impact of phylogenetic influences on each trait. For instance, while traits like body size in mammals often exhibit strong phylogenetic conservation, other traits, such as dietary preference, may be more influenced by local environmental factors and display weaker phylogenetic signals.

We used the *rTraitCont* function from the *ape* package to generate trait data under the BM evolutionary model and the *rnorm* function from the *stats* package to generate data under a normal distribution.

Phylogenetic Eigenvectors Selection Methods

We initiated our analysis with Cheverud's autoregressive model (Cheverud et al. 1985), defined by the equation $T = P + S$. In this model, T represents the trait being analyzed, P accounts for the influence of phylogenetic relationships, and S encompasses influences from other traits. Utilizing the PVR method, we performed principal coordinate analysis on a double-centered phylogenetic matrix ($n \times n$), resulting in $n - 1$ phylogenetic eigenvectors. We then selected k eigenvectors (where typically $k < n - 1$) by using the following model as the underlying principle to assess their relevance with the dependent variable:

$$Y = V\beta + \varepsilon,$$

where Y is the dependent variable representing the trait, V is the set of selected phylogenetic eigenvectors, β is the vector of regression coefficients, and ε represents the residual component S, accounting for additional explanatory variables not included in V. In this study, we aim to determine which of the two variables, X_1 or X_2, serves better as the dependent variable in the regression. Consequently, each variable was used as Y in turn, leading to the identification of corresponding sets of eigenvectors, denoted as V_{X_1} and V_{X_2}, respectively.

The specific methods for eigenvector selection used in this study are based on approaches detailed in Diniz−Filho and Nabout (2009) and Diniz−Filho et al. (2012).

Diniz-Filho et al. (2012) proposed two principal approaches for selecting phylogenetic eigenvectors: sequential and non-sequential. In the sequential approach, the $n - 1$ phylogenetic eigenvectors derived from the matrix decomposition are sorted by their eigenvalues from largest to smallest (EV_1, EV_2, ..., EV_{n-1}). These eigenvectors are then incrementally added to the model in descending order of importance until Moran's I for the residuals, ε, reaches zero, indicating the absence of significant phylogenetic autocorrelation. Moran's I measures the degree of autocorrelation, ranging from −1 (total negative autocorrelation) to +1 (total positive autocorrelation), with 0 representing no autocorrelation.

The non-sequential approach includes three distinct methods. The first method, known as ESRRV (Evaluating Significant Relationships between phylogenetic eigenvectors and the response variable), uses regression models to identify which eigenvectors are relevant to the trait being studied (Diniz−Filho and Nabout 2009, Diniz−Filho et al. 2012). In this study, we start by performing normality tests on both the eigenvectors and the response variable. Depending on the results, we use either Pearson correlation (if the data is normally distributed) or Spearman rank correlation (for non-normal data).

The second and third methods use a strategy called "stepwise regression," specifically employing what is known as "forward selection." In this approach, we start with an empty model (i.e., no phylogenetic eigenvectors included), and then, one by one, we add the eigenvectors that best improve the model’s fit. We continue adding eigenvectors until no more significant improvement can be achieved. This process is called "forward selection" because we gradually build up the model by adding variables rather than removing them.

The difference between the second and third methods lies in the criteria used to determine whether adding an eigenvector improves the model. The second method, called mAIC, adds variables in a way that aims to minimize the AIC value, a measure of model quality that balances goodness of fit and complexity. The third method, called mMorI, adds eigenvectors until the residuals of the model have minimal phylogenetic autocorrelation, as indicated by Moran's I, a statistic that measures the degree of autocorrelation. The difference between the second and third methods lies in the criteria used to determine whether adding an eigenvector improves the model. The second method, called mAIC, seeks to minimize the AIC value. The third method, called mMorI, adds eigenvectors until the model's residuals show minimal phylogenetic autocorrelation, as measured by Moran's I.

In essence, this study exclusively uses the three non-sequential methods (ESRRV, mAIC, and mMorI) to select the most relevant phylogenetic eigenvectors, avoiding redundancy. Compared to the sequential approach, the mMorI method offers more flexibility, encompassing the logic of sequential eigenvector addition while allowing for greater adaptability by exploring different combinations and orders of eigenvector inclusion. Consequently, mMorI supersedes the need for the sequential approach, providing more effective control of phylogenetic signals and enhancing model robustness.

Once selected, these phylogenetic eigenvectors are integrated as covariates into a simple linear regression model, formulated as $y = \sum_{i=1}^{k} a_i * ev_i + b * x + \varepsilon$, where ev_i represents the set of selected eigenvectors, y is the dependent variable trait, x is the independent variable trait, and ε is the residual term.

Moran's I was originally designed as a metric for evaluating spatial autocorrelation, capable of measuring both positive and negative correlations. In its typical application, Moran's I aims to quantify the clustering of similar values, with the minimization process continuing until a suitable *p*-value (e.g., $p > 0.05$) is reached, indicating no significant autocorrelation remains. However, in the context of phylogenetic autocorrelation, we primarily observe positive correlations, as related species tend to have similar trait values.

Negative phylogenetic autocorrelation, while possible, is far less common. Therefore, to better suit our requirements, we modified the eigenvector selection procedure in the PVR package. Specifically, we altered the iteration process so that it terminates upon encountering a negative value of Moran's I, as negative correlations are not typical in our dataset. If no negative value appears, the iteration proceeds until a suitable *p*-value is reached.

To accomplish this, we rewrote the relevant code in the PVR package for selecting phylogenetic eigenvectors by minimizing Moran's I, resulting in a standalone function named MoransI. This custom script is now used for Moran's I method of eigenvector selection in PVR, while all other computations, such as eigenvector decomposition, still utilize the original PVR package. For the ESRRV method, we created the ESRRV function, and for the mAIC method, we employed the stepAIC function from the MASS package (Version 7.3-55) (Venables and Ripley 2002).

Robust Estimators Used in PVR Analysis

In this study, we employed several robust estimators to improve the performance of our PVR analysis, addressing potential violations of model assumptions, such as the presence of outliers and deviations from normality.

The L2 estimator, used in traditional OLS regression, minimizes the sum of squared residuals. However, it is highly sensitive to outliers, as squaring large residuals gives them a disproportionate influence on the model's results.

The L1 estimator, also known as the least absolute deviation, minimizes the sum of absolute residuals, making it more robust than L2 by reducing the influence of outliers. It is particularly useful when outliers or deviations from normality are expected, as it diminishes the impact of extreme values.

M-estimators generalize the maximum likelihood estimation framework and allow for different loss functions to down-weight the effect of outliers. They combine the strengths of both L1 and L2 estimators, treating small residuals quadratically (like L2) and large residuals linearly (like L1).

S-estimators focus on minimizing the dispersion of the residuals to achieve robustness. These estimators aim to find regression coefficients that minimize a robust measure of scale, such as the median absolute deviation. With a high breakdown point, S-estimators can handle a substantial proportion of outliers without losing efficiency.

MM-estimators further refine M- and S-estimators by combining their strengths: they start with an S-estimator for initial robustness and use an M-estimator for final coefficient estimation. This method balances robustness and efficiency, making it well-suited for datasets with outliers and deviations from typical statistical assumptions.

The mathematical formulas for the robust estimators, L1, M, S, and MM, are widely available in statistical references and online resources, such as Wikipedia. For comparative purposes, we applied these robust estimators alongside the traditional L2 estimator. All estimators were implemented using R packages, with M-estimators via the *rlm* function from the MASS package (Version 7.3-55) (Venables and Ripley 2002), L1-estimators via *lad* function from L1PACK package (Version 0.41-245) (Phillips 2002), and S- and MM-estimators via the *lmRob* function from the *robust* package (Version 0.7-4) (Gervini and Yohai 2002).

Gold Standard for Trait Correlations in Each Simulation

In our study, we simulated the evolution of traits X_1 and X_2 along a phylogenetic tree, recording trait values at each node. We then calculated the changes in these traits per unit branch length, represented as $\Delta X_1/L$ and $\Delta X_2/L$, where L is the length of the branch.

These changes are independent from one branch to another, which allows us to use conventional correlation tests to analyze the relationships between $\Delta X_1/L$ and $\Delta X_2/L$, similar to the approach used in PIC.

For each simulation, we tested the normality of $\Delta X_1/L$ and $\Delta X_2/L$. As shown in Tables S1-S3, approximately 90% of cases involving the fixed balanced tree (with 16, 128, and 1024 species) showed normally distributed data, and the frequency of normality increased slightly after applying the Yeo-Johnson transformation, which effectively accommodates both positive and negative values, making it suitable for our dataset where trait changes may include negative values. Given that normality was achieved for the majority of the data, the Pearson correlation was used as the gold standard for the simulations with the three fixed-balanced trees. However, for the random tree simulations (with 128 species), none of the cases showed normal distributions for both $\Delta X_1/L$ and $\Delta X_2/L$, and the Yeo-Johnson transformation resulted in normally distributed data in less than 1% of cases (Table S4). This difference is likely due to the balanced structure of fixed trees, which promotes more consistent trait evolution and a higher likelihood of normality, compared to the uneven and unpredictable nature of random trees (nearly all trees are unbalanced), which results in skewed and heteroscedastic trait distributions. As a result, Spearman rank correlation was employed to examine these correlations.

By assessing the relationship between $\Delta X_1/L$ and $\Delta X_2/L$, we effectively captured the true evolutionary relationship between the traits, X_1 and X_2. This approach enabled us to evaluate the accuracy of various models by comparing their predicted trait correlations with the actual evolutionary relationships observed in the simulations.

It is important to note that, in our simulations, internal node data were directly recorded. This allowed us to perform statistical analyses without the need for model-based estimations of node data, as would be required for empirical datasets. In contrast, empirical datasets often lack direct access to internal node data and must rely on estimation methods, such as PIC. While PIC is a useful method for estimating internal node values, it assumes that traits evolve according to a Brownian motion model. It is also worth noting that PIC is a special case of PGLS under perfect Brownian motion, as described by Rohlf (2001). Our simulations, however, did not rely on this assumption, enabling more flexible and accurate statistical analyses, even in cases where trait evolution deviated from a Brownian motion model.

The Yeo-Johnson transformation (Version 8.23.0) (Guerrero 1993) was implemented using the *BoxCox.lambda* function from the *forecast* package, and the normality test was conducted using the *shapiro.test* function from the *stats* package.

Results

When ordinary least squares linear regression models were used to assess the relationship between X_1 and X_2, we use $OLS_{X_1 \sim X_2 + V_{X_1}}$ to represent the model with X_1 as the dependent variable, X_2 as the independent variable, and V_{X_1} as the covariates. Similarly, $OLS_{X_2 \sim X_1 + V_{X_2}}$ denotes the model with X_2 as the dependent variable, X_1 as the independent variable, and V_{X_2} as the covariates. Unless otherwise specified, the results presented in this section are based on phylogenetic trees with 128 species. A dedicated subsection will later compare the findings across trees with 16, 128, and 1024 species.

Collinearity of Independent Variables of Phylogenetic Eigenvectors

First, we assess the collinearity of independent variables when different methods of phylogenetic eigenvector selection are applied. Our results showed that, under a fixed balanced phylogenetic tree, the variance inflation factor (VIF) for independent variables was consistently higher when phylogenetic eigenvectors were selected using the mAIC method, compared to when they were selected using the ESRRV or mMorI methods, regardless of the direction of the regression ($X_1 \sim X_2$ or $X_2 \sim X_1$) (Tables 1 and S5). In the BM1 & BM1+BM2 and BM & BM+Norm scenarios, VIF values as high as six were observed, indicating severe multicollinearity that could compromise model stability and affect regression results. Similarly, under random phylogenetic trees (Tables S6 and S7), the mAIC method also showed higher multicollinearity compared to the other two methods. Therefore, due to the frequently high multicollinearity associated with the mAIC method, we excluded it from further comparative analyses.

Table 1. Variance inflation factor (VIF) of X_2 under a fixed balanced phylogenetic tree

Simulation scenarios	Eigenvector selection methods	Noise term variances							
		10^{-4}	0.01	1	4	16	64	258	1024
BM1 & BM1+BM2	ESRRV	2.6	2.5	2.1	1.9	1.8	1.8	1.8	1.8
	mAIC	6.3	6.3	4.0	3.4	3.2	3.1	3.1	3.1
	mMorI	3.7	3.7	2.8	2.5	2.4	2.4	2.4	2.4
BM & BM+Norm	ESRRV	2.6	2.5	1.6	1.2	1.1	1.0	1.1	1.0
	mAIC	6.3	6.1	2.7	2.2	2.0	2.0	2.0	2.0
	mMorI	3.7	3.6	1.9	1.3	1.1	1.1	1.0	1.1
Norm & Norm+BM	ESRRV	1.3	1.3	1.2	1.1	1.1	1.1	1.1	1.1
	mAIC	3.1	3.1	1.9	1.6	1.5	1.5	1.5	1.5
	mMorI	1.0	1.0	1.1	1.1	1.1	1.1	1.1	1.2
Norm1 & Norm1+Norm2	ESRRV	1.3	1.3	1.2	1.1	1.1	1.1	1.1	1.1
	mAIC	3.1	3.1	2.0	1.7	1.5	1.5	1.5	1.5
	mMorI	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0

The regression model used in the analyses of this table is specified as $X_1 \sim X_2$, where X_1 is the dependent variable, and X_2 is the independent variable. Higher VIF values indicate greater multicollinearity, which may affect the stability of the regression model. Please refer to the Materials and Methods section for details regarding the three phylogenetic eigenvector selection methods and the four simulation scenarios.

Impact of Variable Designation on Eigenvector Selection

The designation of dependent and independent variables had a substantial impact on the selection of phylogenetic eigenvectors. Each bar in Figure 2 corresponds to an 8,000 simulation set, showing how often the sets of selected eigenvectors differed when the dependent and independent variables were interchanged. Overall, the discrepancy rate between eigenvector selections was high, with approximately 80% of the cases showing different sets of eigenvectors following the variable swap. The discrepancy rates were consistent across different tree types, with 80.6% for the fixed balanced tree and 83.4% for the random trees. However, there was a slightly larger difference between the methods, with an 86.3% discrepancy for the ESRRV method and 77.6% for the mMorI method. These results demonstrate how the selection of dependent and independent variables influences the identification of phylogenetic eigenvectors.

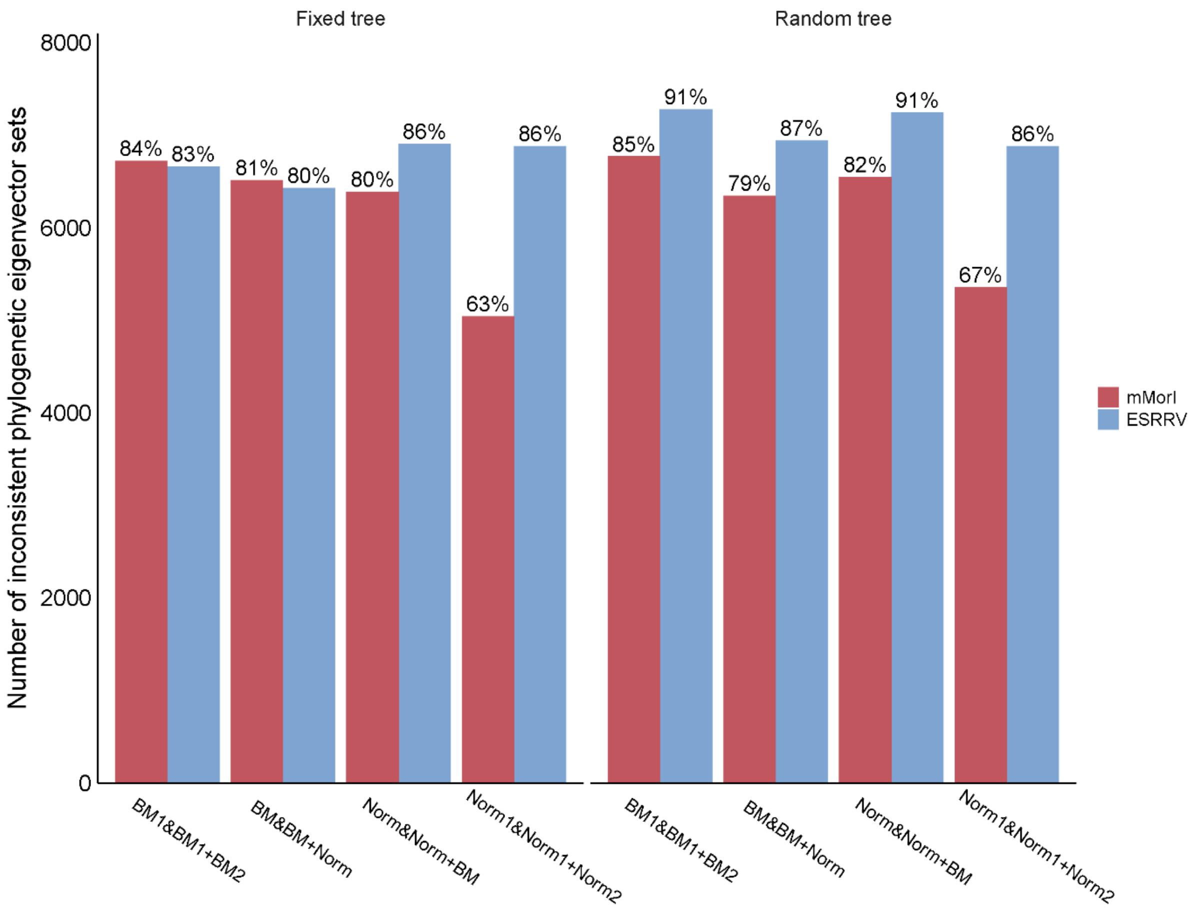

Figure 2. Dependence of phylogenetic eigenvector selection on variable exchange in the case of fixed balanced tree and random trees. This figure illustrates the count of divergent selections of phylogenetic eigenvector sets resulting from interchanging dependent and independent variables (denoted as V_{X_1} and V_{X_2}). Each scenario encompasses 8000 simulations. Please refer to the Materials and Methods section for details regarding the two phylogenetic eigenvector selection methods (ESRRV and mMorI) and the four simulation scenarios ("BM1 & BM1+BM2," "BM & BM+Norm," "Norm & Norm+BM," and "Norm1 & Norm1+Norm2").

Divergent PVR Results Due to Different Eigenvector Sets

In both the fixed balanced tree and random tree scenarios, the interchanging dependent and independent variables resulted in differing sets of selected phylogenetic eigenvectors, leading to conflicting correlation outcomes in our simulations (Fig. 3). The M estimator consistently exhibited the highest number of inconsistencies. In contrast, the L2, L1, and

MM estimators showed fewer inconsistencies, with relatively similar performance. The S estimator displayed intermediate results. Notably, the M estimator resulted in 35-39% of correlation outcomes being conflicting for both the ESRRV and mMorI selection methods. Even for the L2 estimator, which had the fewest conflicting correlation outcomes, there were still 11% conflicting outcomes for ESRRV and 8-9% for mMorI out of 64,000 simulations, highlighting the significant impact of variable selection on correlation analysis. Since the L2 estimator corresponds to OLS, our results demonstrate that using robust estimators in PVR does not mitigate the influence of dependent and independent variable selection on the results. In fact, the influence is even greater, suggesting that in robust PVR, more careful selection of the dependent variable is necessary.

The distribution of divergent outcomes across different scenarios can be seen in Figs. S1-S5, which shows the varying results for each estimator under different conditions.

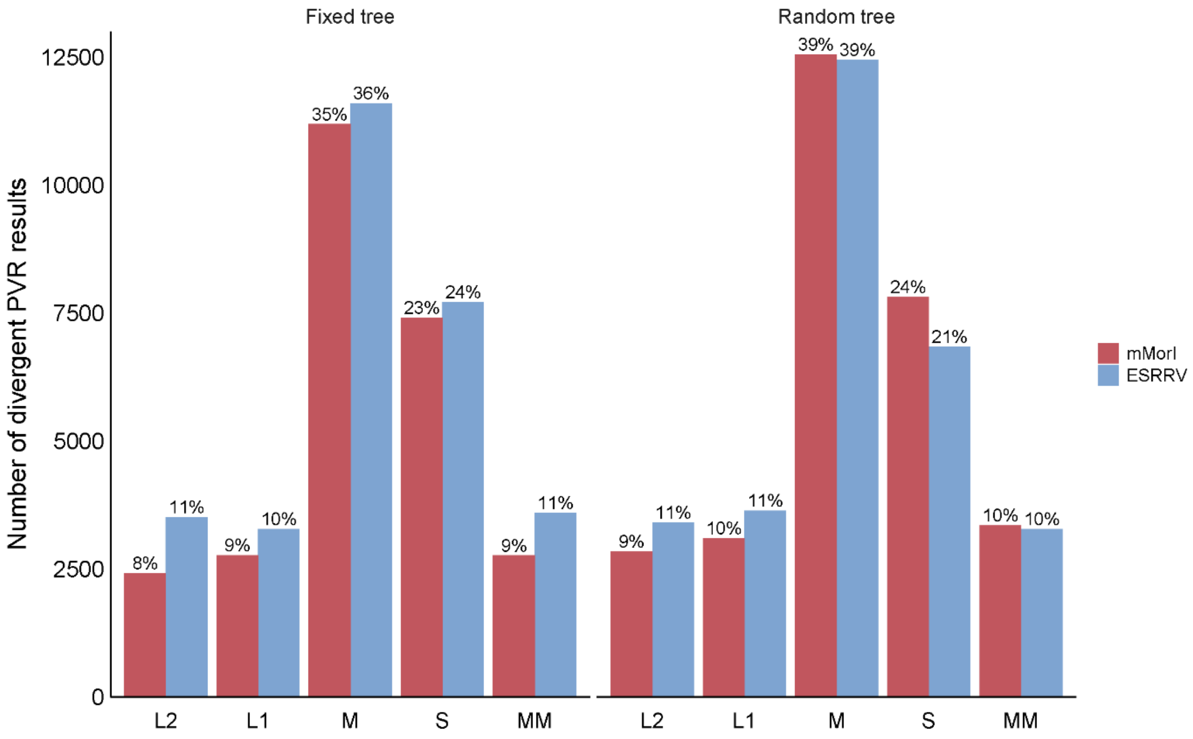

Figure 3. Divergent phylogenetic eigenvector regression outcomes due to variable swap. This figure illustrates the number of divergent phylogenetic eigenvector regression outcomes when dependent

and independent variables are swapped. The analysis includes five robust regression estimators (L1, L2, M, S, and MM), two phylogenetic eigenvector selection methods (ESRRV and mMorI), and two tree types (fixed balanced and random trees, 32,000 simulations per tree type). Please refer to the Materials and Methods section for details regarding the phylogenetic eigenvector selection methods and the robust regression estimators.

Impact of Eigenvector Set Selection on Correlation Divergence

In this section, we provide a more detailed evaluation of the divergent correlation outcomes, specifically focusing on the correlation coefficients for the L2 estimator, which corresponds to OLS. We investigated whether the different sets of eigenvectors (V_{X_1} and V_{X_2}) resulting from swapping the dependent and independent variables could lead to differences in partial correlation coefficients between X_1 and X_2. Specifically, we looked at how the partial correlation between X_1 and X_2 changed when we swapped the eigenvectors selected for X_1 with those selected for X_2. This approach provides a more nuanced evaluation of the variability introduced by these different eigenvector sets across various scenarios.

Our results indicate that the differences in partial correlation coefficients ($|\mathrm{pcor}(X_1, X_2 + V_{X_1}) - \mathrm{pcor}(X_2, X_1 + V_{X_2}|$) were significantly larger in datasets with strong phylogenetic structure, particularly in the 'BM1 & BM1+BM2' scenario (Fig. 4). In contrast, in scenarios with minimal phylogenetic autocorrelation, such as 'Norm1 & Norm1+Norm2,' the differences were negligible. These results provide a more in-depth view of the impact of eigenvector selection on correlation analysis when using L2, highlighting that both the choice of dependent and independent variables and the selection of phylogenetic eigenvectors can substantially influence correlation estimates when phylogenetic autocorrelation is present.

This section supplements the findings presented in Fig. 3, where we observed that the designation of dependent variables led to differing sets of selected phylogenetic eigenvectors, ultimately resulting in inconsistent correlation outcomes in our simulations.

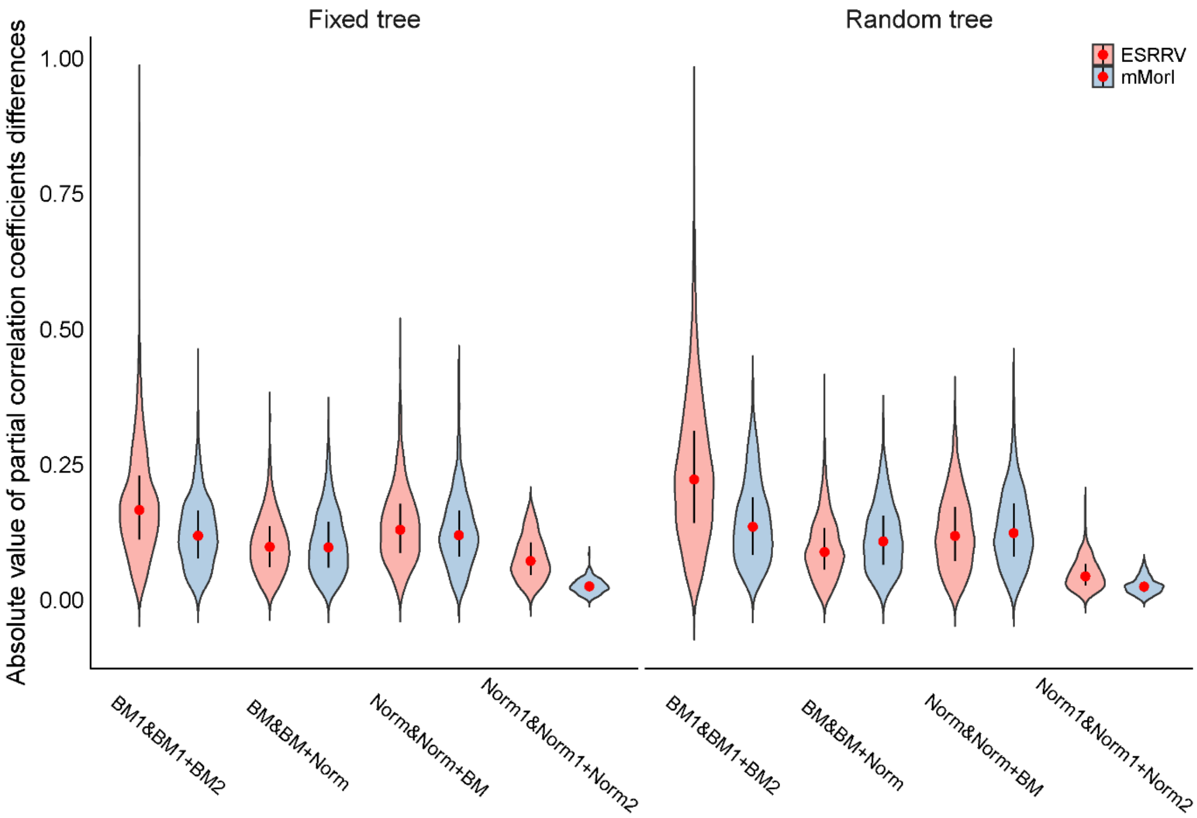

Figure 4. Distributions of differences in partial correlation coefficients between X_1 and X_2. The violin plot illustrates the absolute differences in partial correlation coefficients (|pcor(X_1, $X_2 + V_{X_1}$) - pcor(X_2, $X_1 + V_{X_2}$)|) for cases with conflicting PVR results. The left half represents results from the fixed tree case, while the right half shows the random tree case. The blue sections correspond to the mMorI selection method, and the red sections represent the ESRRV method. Red dots indicate the median values, and the vertical lines within the violins show the range of error bars. These distributions highlight the variability introduced by swapping dependent and independent variables across different eigenvector sets. Please refer to the Materials and Methods section for details regarding the two phylogenetic eigenvector selection methods (ESRRV and mMorI) and the four simulation scenarios ("BM1 & BM1+BM2," "BM & BM+Norm," "Norm & Norm+BM," and "Norm1 & Norm1+Norm2").

Accuracy of PVR Outcomes Derived from Variable Swapping via L2 Estimator

To establish a baseline for model evaluation, we used the Gold Standard for Trait Correlations in Each Simulation (see Materials and Methods for further details), determining the true relationships between traits X_1 and X_2 by examining the relationship between the changes per unit length ($\Delta X_1/L$ and $\Delta X_2/L$) for each branch of the phylogenetic trees using traditional correlation analysis methods. This provided a standard for evaluating the accuracy of PVR results.

Next, we evaluated the accuracy of the PVR models by comparing the results obtained using the L2 estimator with the Gold Standard. Specifically, we compared the outcomes of two models: $OLS_{X_1 \sim X_2 + V_{X_1}}$ and $OLS_{X_2 \sim X_1 + V_{X_2}}$, where the dependent and independent variables were swapped. The analysis focused on the frequency of outcomes where both models yielded consistent and correct results, consistent but incorrect results, and cases where only one model provided an accurate correlation result.

Our findings show no substantial difference in performance between the ESRRV and mMorI eigenvector selection methods. From Fig. 5, it can be observed that in most cases (blue portion), both X_1 and X_2, when treated as the dependent variable, yielded correct results. In some cases (yellow portion), only one of the variables as the dependent variable produced correct results, while the other resulted in incorrect outcomes. Additionally, in a portion of the simulations (red portion), neither choice of dependent variable produced correct results, indicating that there is still room for improvement in the current methods. Across all four simulation scenarios, the proportion of cases where only one model was correct represented nearly 10% of the total simulations for both methods. The distribution of 'only one correct' outcomes revealed that more inconsistent results were concentrated in the 'BM1 & BM1+BM2,' 'BM & BM+Norm,' and 'Norm & Norm+BM' scenarios, while the

'Norm1 & Norm1+Norm2' scenario exhibited fewer such cases (Fig. 3). In this last scenario, trait evolution was logically less influenced by phylogeny, as the Norm model represents no phylogenetic influence. By contrast, the first three scenarios required careful handling of phylogenetic autocorrelation. The occurrence of these 'only one correct' results underscores the importance of careful selection of the dependent variable in PVR analysis.

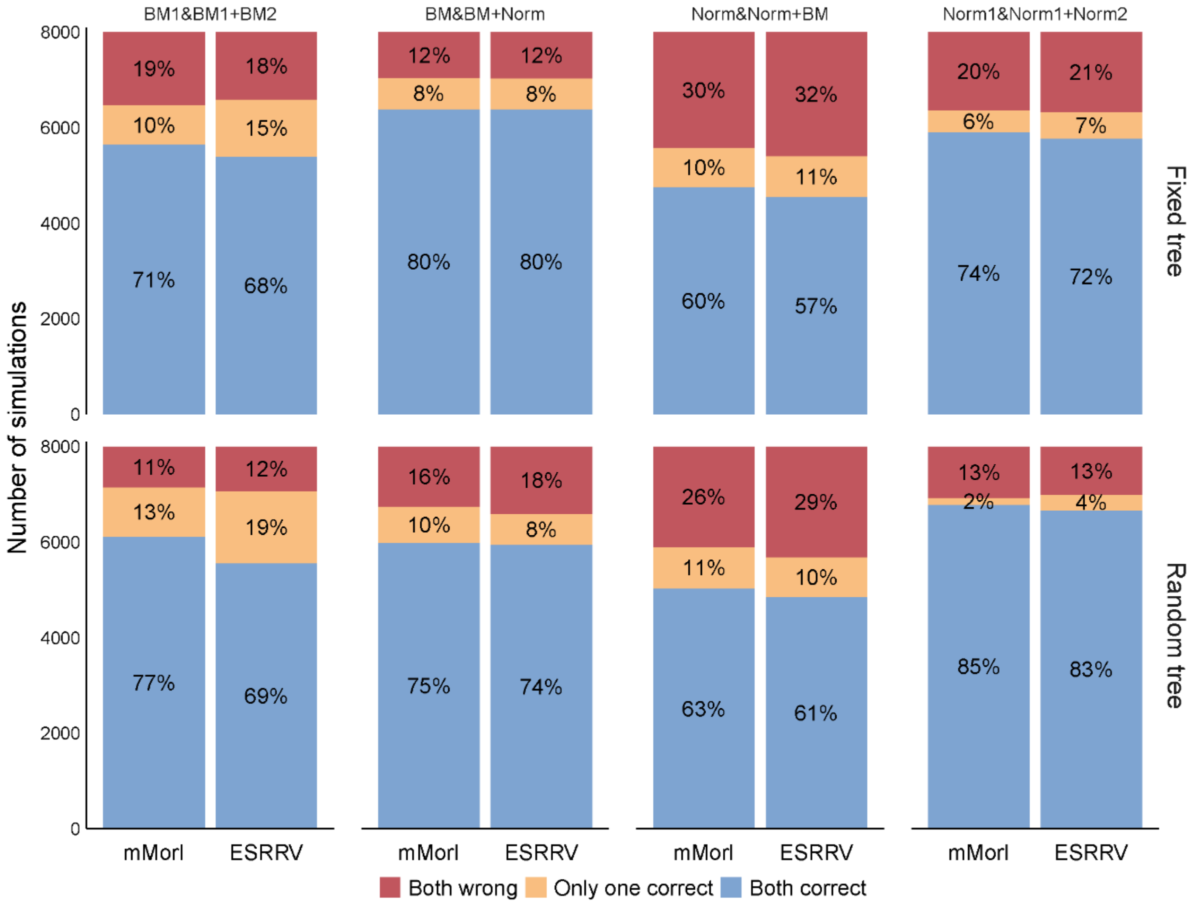

Figure 5. Outcomes when either X_1 or X_2 is treated as the dependent variable in PVR analysis of various simulations. This figure shows the distribution of outcomes from simulations using the ESRRV and mMorI selection methods. The blue sections represent cases where both choices of the dependent variable (X_1 or X_2) yielded correct results, the yellow sections indicate cases where only one choice produced correct results, and the red sections reflect instances where neither choice produced correct results. Please refer to the Materials and Methods section for details regarding the two phylogenetic eigenvector selection methods (ESRRV and mMorI) and the four simulation

scenarios ("BM1 & BM1+BM2," "BM & BM+Norm," "Norm & Norm+BM," and "Norm1 & Norm1+Norm2").

Evaluating Dependent Variable Selection Criteria in PVR with L2 Estimator

Accurately identifying the model that best reflects the relationships between traits is critical, especially when swapping dependent and independent variables can result in conflicting outcomes. The "Gold Standard for Trait Correlations" allowed us to determine trait relationships in our simulations by directly recording internal node data, but this is not possible in empirical datasets, where internal node data are typically unavailable. In light of these limitations, we propose three criteria based on R^2, phylogenetic signal λ, and AIC, using the L2 estimator (i.e., OLS) in PVR analyses. These criteria provide a systematic way to guide variable selection, thereby improving the robustness of PVR analysis, particularly when a "gold standard" is unavailable for empirical datasets.

1. **R^2-Criterion**: This criterion prioritizes models based on the magnitude of R^2, reflecting the proportion of variance explained by the independent variable(s). Higher R^2 values suggest a better model fit.
2. **λ-Criterion**: This criterion prioritizes dependent variables based on the strength of the phylogenetic signal, measured using Pagel′s λ. A higher λ indicates a stronger correlation between regression residuals and phylogenetic structure (Pagel 1997).
3. **AIC-Criterion**: This criterion prioritizes models based on the relative magnitude of AIC, which balances model fit and complexity. Lower AIC values indicate models that achieve a better balance between goodness of fit and model simplicity.

The results of applying these criteria are summarized in Table 2, where 'H' and 'L' indicate whether models with higher (H) or lower (L) values for each criterion aligned better with the "Gold Standard" for trait correlations.

***R^2* Criterion**: In most scenarios ("BM1 & BM1+BM2," "BM & BM+Norm," and "Norm & Norm+BM"), higher R^2 values (H) aligned better with the "Gold Standard," indicating that maximizing R^2 generally improved model performance. However, in the "Norm1 & Norm1+Norm2" scenario, models with lower R^2 values (L) performed better, likely due to the weaker phylogenetic structure in this scenario.

λ Criterion: Models with stronger phylogenetic signals (H) typically performed better. However, in some cases, such as the random tree scenario for "BM1 & BM1+BM2," models with lower λ values (L) aligned better with the "Gold Standard," suggesting that overestimating phylogenetic signal can lead to poorer performance.

AIC Criterion: Results for AIC were mixed. In some scenarios, lower AIC values (L) corresponded to better model performance, but in others, higher AIC values (H) produced more accurate models. This suggests that while AIC is generally useful, its reliability in selecting dependent variables may vary based on the evolutionary scenario and eigenvector selection method.

Overall, R^2 emerged as the most reliable criterion, especially in scenarios with strong phylogenetic autocorrelation. For the "Norm1 & Norm1+Norm2" scenario, no consistent patterns emerged for all three criteria, making it difficult to provide reliable recommendations for dependent variable selection. In cases where the phylogenetic signal is negligible or absent, such as in this scenario, we suggest that phylogenetic comparative methods like PVR may not be appropriate. Instead, traditional methods that do not rely on phylogenetic relationships may yield more accurate results.

Table 2. Performance of dependent variable selection criteria in PVR using the L2 estimator

Eigenvector selection methods	Phylogenetic trees	Simulation scenarios	Criteria		
			R^2	λ	AIC

ESRRV	Fixed tree	BM1 & BM1+BM2	H	/	/
		BM & BM+Norm	H	H	L
		Norm & Norm+BM	H	H	H
		Norm1 & Norm1+Norm2	L	H	H
	Random tree	BM1 & BM1+BM2	H	/	H
		BM & BM+Norm	H	H	L
		Norm & Norm+BM	H	H	H
		Norm1 & Norm1+Norm2	L	H	/
mMorI	Fixed tree	BM1 & BM1+BM2	H	H	/
		BM & BM+Norm	H	H	L
		Norm & Norm+BM	H	H	H
		Norm1 & Norm1+Norm2	/	/	/
	Random tree	BM1 & BM1+BM2	H	L	/
		BM & BM+Norm	H	H	L
		Norm & Norm+BM	H	H	H
		Norm1 & Norm1+Norm2	L	/	L

This table evaluates the effectiveness of three criteria—R^2, λ, and AIC—in selecting the dependent variable for PVR models. 'H' indicates that models with higher values for the criterion performed better, 'L' indicates that models with lower values performed better, and '/' denotes no significant difference. Model performance was evaluated by conducting Chi-squared tests on the number of correct and incorrect correlation outcomes for each model. Detailed Chi-squared test results can be found in Tables S8-S13. Please refer to the Materials and Methods section for details regarding the phylogenetic eigenvector selection methods, the phylogenetic trees, and the simulation scenarios.

Evaluating Dependent Variable Selection Criteria with Various Estimators

In the previous sections, we analyzed the performance of different variable selection criteria (R^2, λ, and AIC) using the L2 estimator under various scenarios, trees, and phylogenetic eigenvector selection methods. Now, we introduce another layer of complexity by incorporating robust estimators (L1, M, S, and MM). To streamline the presentation of results, we combined the first three simulation scenarios ("BM1 & BM1+BM2," "BM & BM+Norm," and "Norm & Norm+BM"), which are logically expected to have strong phylogenetic autocorrelation, into a unified evaluation.

The performance in these scenarios was summarized using indicators like H1, H2, and H3: 'H1' represents cases where higher criterion values outperformed in one scenario, with no significant difference in the remaining two; 'H2' indicates that higher values outperformed in two scenarios; and 'H3' indicates higher criterion values consistently outperformed across all three scenarios. Conflicting results, where both higher and lower criterion values showed significant performance in different scenarios, are indicated by 'C' (Table 3).

To facilitate a comparison of how each criterion performs under L2 versus other robust estimators (L1, M, S, and MM), we included L2 results in Table 3. This allows for a direct assessment of each criterion's effectiveness across different phylogenetic and evolutionary contexts. Specifically, Table 3 highlights the performance of the three selection criteria under different estimators, tree types, and eigenvector selection methods.

R^2 Criterion: Generally performed better than λ and AIC, especially with the L2 and MM estimators. For instance, models with higher R^2 values consistently outperformed in the scenarios with strong phylogenetic autocorrelation ("BM1 & BM1+BM2," "BM & BM+Norm," and "Norm & Norm+BM") for both fixed and random trees (denoted by 'H3'). Under the L1 estimator, R^2 still generally performed well but was slightly less consistent compared to L2 and MM, as seen with an H2 outcome in the mMorI method for the fixed tree. Under the S estimator, R^2 performed less effectively, with multiple instances of mixed ('M') results, indicating inconsistent performance across different scenarios. Under the M estimator, the R^2 criterion did not perform strongly, showing mixed results (M) across all four cases (Table 3).

λ Criterion: Performed moderately well across scenarios but did not dominate any specific estimator or tree type. It showed stronger results under L2, L1, and MM estimators but was inconsistent with M and S estimators, where no clear advantage was observed.

AIC Criterion: Showed unusual behavior with the M estimator, where higher AIC values were consistently associated with better model performance (e.g., 'H2' or 'H3' in three rows). The S estimator also showed cases where higher AIC values led to better results under the mMorI selection method. In contrast, AIC did not significantly differentiate models with L2, L1, and MM estimators, as evidenced by more mixed ('M') results.

Table 3. Summary of the performance of three criteria on different estimators.

Eigenvector selection methods	Phylogenetic trees	Criteria	Robust estimators				
			L2	L1	M	S	MM
ESRRV	fixed tree	R^2	H3	H3	C	H2	H3
		λ	H2	H3	C	H2	H3
		AIC	C	C	H2	C	C
	random tree	R^2	H3	H3	C	C	H3
		λ	H2	H2	C	C	H2
		AIC	C	C	H3	H1	C
mMorI	fixed tree	R^2	H3	H2	C	C	H3
		λ	H3	H1	C	C	H2
		AIC	C	H1	C	H2	C
	random tree	R^2	H3	H3	C	C	H3
		λ	C	H2	C	C	H2
		AIC	C	C	H2	H2	C

In the table, 'C' denotes conflicting cases across the three scenarios with strong phylogenetic autocorrelation—'BM1 & BM1+BM2,' 'BM & BM+Norm,' and 'Norm & Norm+BM'—where models with higher criteria values performed significantly better in some instances, while in others, models with lower criteria values performed significantly better. 'H' indicates that models with higher corresponding criteria values significantly outperformed others, with the number following 'H' specifying how many of the three scenarios showed this pattern (e.g., 'H1' indicates one scenario, 'H3' indicates all three scenarios). The significance test results for each scenario are detailed in Tables S14–S17. Please refer to the Materials and Methods section for details regarding the phylogenetic eigenvector selection methods, the phylogenetic trees, and the simulation scenarios.

Impact of Sample Size on the Performance of Dependent Variable Selection Criteria

To examine how the number of species (sample size) influences the selection of dependent and independent variables, we analyzed fixed-balanced trees containing 16, 128, and 1024 species across four simulation scenarios ("BM1 & BM1+BM2," "BM & BM+Norm," "Norm & Norm+BM," and "Norm1 & Norm1+Norm2"). We evaluated the effect of swapping dependent and independent variables on the PVR results under these different sample sizes and assessed the performance of the three variable selection criteria (R^2, λ, and AIC).

The mMorI eigenvector selection method, due to its iterative nature, has a time complexity of $O(n^2)$, making it computationally intensive as the number of species increases. When sample sizes reached 1024 species, the time required for a single simulation using mMorI exceeded 7 hours (7h 04m per simulation) in our laboratory's computational environment. Given that our study required conducting a total of 8,000 simulations for each scenario, the cumulative time required for the mMorI method with large trees became prohibitive. Table 4 presents the time consumption for a single simulation under different sample sizes, from which it is evident that the ESRRV method demonstrates a clear time advantage when handling large datasets. As such, we primarily focused on the ESRRV method when analyzing the effect of sample size on dependent and independent variable selection, which has a time complexity of O(n) and is more efficient for large datasets.

Table 4. Time consumption per single simulation across different sample sizes

Eigenvector selection methods	Sample sizes		
	16	128	1024
ESRRV	0.04s	0.14s	1s
mMorI	0.07s	6.45s	7h 04m

The simulations were performed using a single thread of an AMD EPYC 7702 64-core processor. While absolute times may vary depending on the computational environment, this table highlights the relative time differences between different sample sizes and eigenvector selection methods, with the mMorI method being particularly time-intensive for larger trees. Please refer to the Materials and Methods section for details regarding the two phylogenetic eigenvector selection methods.

In order to streamline our analysis and reduce workload, we did not perform all estimators across the various sample sizes. Based on the results from previous sections, which demonstrated that the L2 estimator (i.e., OLS) generally performed well across different scenarios, we chose to use the L2 estimator for this comparison of sample size effects. Furthermore, to minimize computational burden, we restricted our analysis to the fixed tree case. This approach allowed us to focus on how sample size impacts the performance of the three variable selection criteria (R^2, λ, and AIC) without unnecessarily increasing the computational burden.

Our results indicate that the R^2 criterion performs well with larger sample sizes (128 and 1024), particularly in scenarios with strong phylogenetic autocorrelation. However, with a sample size of 16, no significant differences were observed between models with higher and lower R^2 values under the "BM & BM+Norm" scenario (Table 5). The λ criterion tends to perform better with smaller sample sizes, but it underperforms with larger sample sizes, particularly in the "BM1 & BM1+BM2" scenario. Lastly, the AIC criterion showed conflicting results across sample sizes. In the "BM & BM+Norm" scenario, models with lower AIC values performed better, whereas in the "Norm & Norm+BM" scenario, models with higher AIC values outperformed.

Table 5. Performance of the three criteria under different sample sizes for the ESRRV method.

Species number	Simulation scenarios	Variable selection criteria		
		R^2	λ	AIC
16	BM1 & BM1+BM2	H	H	/
	BM & BM+Norm	/	H	L
	Norm & Norm+BM	H	H	H
	Norm1 & Norm1+Norm2	L	/	/
128	BM1 & BM1+BM2	H	/	/
	BM & BM+Norm	H	H	L
	Norm & Norm+BM	H	H	H
	Norm1 & Norm1+Norm2	L	H	H
1024	BM1 & BM1+BM2	H	/	/
	BM & BM+Norm	H	H	L
	Norm & Norm+BM	H	H	H
	Norm1 & Norm1+Norm2	L	/	/

'H' indicates that models with higher values for the criterion performed better, 'L' indicates that models with lower values performed better, and '/' denotes no significant difference. Model performance was evaluated by conducting Chi-squared tests on the number of correct and incorrect correlation outcomes for each model. Detailed Chi-squared test results can be found in Tables S8-S10 and S18-S20. Please refer to the Materials and Methods section for details regarding the simulation scenarios.

Comparison of Eigenvector Selection Methods in PVR Analysis

Using R^2 as our selection criterion, we effectively selected the dependent variable from the two traits X_1 and X_2. We then assessed the performance of two phylogenetic eigenvector selection methods, ESRRV and mMorI, within the framework of PVR analysis. These analyses were conducted on both fixed and random trees containing 128 species, utilizing the L2 estimator.

Our results indicate that for fixed trees and across different simulation scenarios, significant differences between the ESRRV and mMorI methods were found in only 5 out of the 32 cases (four scenarios, each with eight variance levels), using $p = 0.05$ as the significance threshold. These differences were concentrated in specific variance gradients (particularly 4, 16, and 64). Among these cases, mMorI outperformed ESRRV in 4, while ESRRV was superior in only 1. When the variance of the noise term was either very small or very large, no significant differences between the two methods were observed (Fig. 6A).

In the random tree case, we observed that at very low noise variances (0.0001, 0.01, and 1), there were no significant differences between the mMorI and ESRRV methods (Fig. 6B). However, as the noise variance increased, particularly for variance values of 4, 16, and 64, the mMorI method outperformed the ESRRV method in several instances. No significant differences were found in the scenario with the weakest phylogenetic autocorrelation, 'Norm1 & Norm1+Norm2,' regardless of the noise variance.

In terms of accuracy, the mMorI method showed superior performance overall when compared to ESRRV (Fig. 6). However, considering the time consumption (Table 4), the ESRRV method may be more suitable for analyses involving large sample sizes.

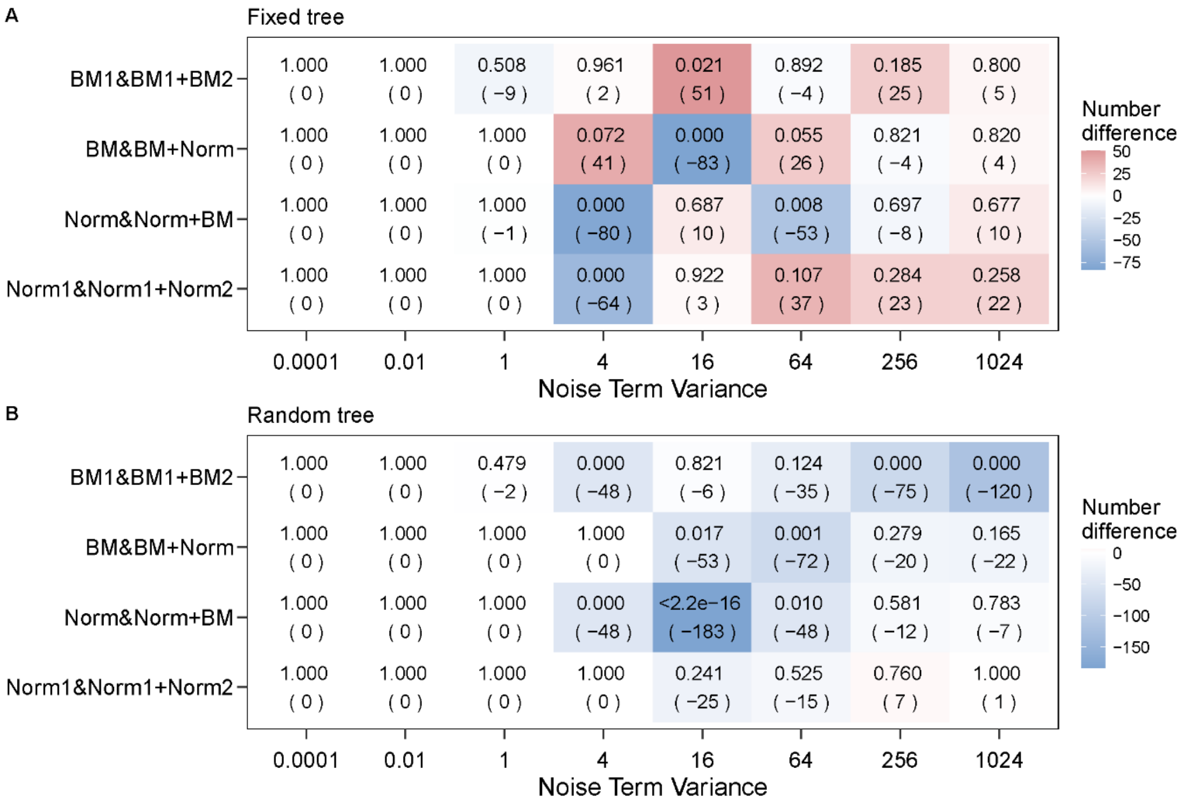

Figure 6. Significance heatmap for the differences between ESRRV and mMorI methods in their correct detections of trait correlations in the fixed tree (A) and random trees (B) with 128 species using the L2 estimator. Each number above the cell represents the *p*-value of the Chi-square test, and the number below represents the difference in correct detections between the ESRRV and mMorI methods. Red indicates better performance by ESRRV, while blue indicates better performance by mMorI. Please refer to the Materials and Methods section for further details on the two phylogenetic eigenvector selection methods (ESRRV and mMorI) and the four simulation scenarios ("BM1 & BM1+BM2," "BM & BM+Norm," "Norm & Norm+BM," and "Norm1 & Norm1+Norm2").

Discussion

Model-Based vs. Eigenvector-Based Methods in Phylogenetic Comparative Analysis

Although model-based methods like PGLS and PIC are prone to certain limitations, particularly in cases where singular evolutionary events distort associations between traits (Uyeda et al. 2018), they remain more widely adopted phylogenetic comparative methods addressing phylogenetic autocorrelation compared to the eigenvector-based approach, PVR. (Garamszegi 2014). The primary advantage of PIC over PVR is its simplicity and intuitiveness, making it sufficient for certain evolutionary scenarios. Previous studies have suggested that PGLS tends to be more efficient than PVR (Freckleton et al. 2011). With small sample sizes, PVR often performs worse due to the statistical cost of estimating parameters, as each eigenvector included in the model consumes one degree of freedom. In larger samples, PVR exhibits another weakness, where residual phylogenetic signal persists after dimensionality reduction, which can lead to higher Type I error rates.

Despite these limitations, we believe that PVR still has a role in phylogenetic comparative analysis. Since the advent of the genomic era, studies commonly involve hundreds or even thousands of species, making the limitations of PVR for small datasets less of a concern. Moreover, the challenges related to residual phylogenetic signals in larger datasets and the sensitivity of PVR to eigenvector selection can be addressed by optimizing algorithms and improving both dimensionality reduction and the selection of phylogenetic eigenvectors. Diniz-Filho et al. (2012) demonstrated how these enhancements can significantly improve the performance of PVR in large-scale phylogenetic analyses. While PGLS can be considered a developed method, PVR remains a developing method with relatively fewer studies, offering opportunities for further refinement. The main focus of this study is to improve the accuracy of PVR by optimizing eigenvector selection by switching the roles of dependent and independent variables. Additionally, our results, presented in Tables 1 and S5-S7, show that the selected eigenvectors exhibit significantly reduced collinearity with independent variables when using the ESRRV and mMorI methods compared to the mAIC method. Furthermore, in accurately detecting trait correlations across various evolutionary scenarios, the mMorI method generally outperformed ESRRV (Fig. 6).

However, for large sample sizes, the time required by mMorI is significantly greater than that for ESRRV (Table 4).

While PGLS and other model-based methods are favored for their flexibility and robustness in evaluating phylogenetic patterns (Freckleton et al. 2011, O'Meara and Beaulieu 2016), PVR offers valuable insights through its eigenvector decomposition, particularly when the phylogenetic signal is complex, or the evolutionary model is uncertain. Above all, the evolutionary models we currently understand are unlikely to capture all patterns of evolution. Even if future research develops numerous evolutionary models capable of accounting for nearly every possible scenario, distinguishing subtle differences between these models using PGLS would remain cumbersome, as each model would require a separate reanalysis.

The utility of eigenanalysis, which generates phylogenetic eigenvectors, extends beyond PVR and contributes more broadly to the field. Unlike PGLS, where phylogenetic information is embedded directly into the model through assumptions about residual distributions (Symonds and Blomberg 2014), eigenanalysis produces modular eigenvectors that capture phylogenetic structure. These eigenvectors can be transferred and utilized in various models beyond PVR, including random forests, support vector machines, and neural networks, such as in studies that impute missing trait data within a phylogenetic context (Debastiani et al. 2021).

Future research may reveal that PGLS and PVR each have their strengths, excelling in certain evolutionary contexts while underperforming in others. A complementary relationship between PVR and PGLS could emerge, where each method is optimal for different scenarios.

Asymmetry in OLS Regression: Challenges in Assigning Variables

When studying the relationship between two variables, two traditional approaches are correlation analysis and regression analysis. Correlation analysis describes the strength and direction of the relationship between two variables symmetrically, meaning the variables X and Y are interchangeable. It measures the degree of association without implying causation, making it suitable for preliminary investigations where both variables are sampled without systematic control. Its simplicity and minimal assumptions are advantageous, but it lacks the ability to identify cause-and-effect relationships or make predictions.

Regression analysis, on the other hand, provides more detailed information and involves a clear distinction between the independent variable (X) and the dependent variable (Y). In OLS regression, one of the most commonly used regression methods due to its simplicity and wide applicability, the independent variable is assumed to be fixed and free of measurement error (McArdle 1988, Osborne and Waters 2002, Casson and Farmer 2014, Jarantow et al. 2023), which makes OLS inherently asymmetrical—X and Y cannot be freely interchanged. The purpose of regression is to explain or predict the variation in Y based on X, which generally requires establishing a logical causal direction. However, in many studies, causal relationships are unclear, and measurement errors in X can be unavoidable. These issues complicate the use of OLS, requiring caution in interpreting its results.

These challenges also exist in phylogenetic comparative analyses that control for phylogenetic autocorrelation. However, suitable methods for controlling phylogenetic effects while studying trait relationships are more limited. The most commonly used are PIC and PGLS, with PVR also being a notable method. PGLS was developed from OLS, inheriting its asymmetry issues. As with OLS, PGLS requires a distinction between independent and dependent variables. Moreover, most studies using PIC employ it in regression contexts, where they are similarly subject to asymmetry problems. Since PVR also relies on the foundation of OLS, it suffers from the same asymmetry. One reviewer rightly pointed out that this asymmetry becomes particularly pronounced when R^2 is low, which can make interpretation challenging.

In PGLS, our previous findings suggest that, in the absence of observational error, the trait with a stronger phylogenetic signal should generally be treated as the dependent variable to more effectively control for phylogenetic autocorrelation (Chen et al. 2023). However, if practical experience indicates that this trait has smaller measurement errors, it should be treated as the independent variable instead. This highlights the complexities in the selection of dependent and independent variables when asymmetry and measurement errors are present.

PVR offers a distinct advantage over PGLS by extracting phylogenetic information into separate eigenvectors, which serve as additional variables in the analysis. The present study indicates that by testing different combinations of dependent and independent variables, PVR can effectively identify the optimal set of eigenvectors to control phylogenetic autocorrelation. Once the best set of eigenvectors is determined, the subsequent regression analysis becomes less sensitive to the choice of dependent or independent variables, allowing researchers to designate the logically inferred causal trait as the independent variable without compromising the model's validity. In cases where there is no obvious causal relationship between traits (e.g., two physiological or genomic traits), the trait with less measurement error can be chosen as the independent variable.

To address the limitations of OLS regression concerning measurement error in independent variables, previous researchers have developed several regression methods, such as orthogonal regression, errors-in-variables regression, and reduced major axis (RMA) regression. These methods account for errors in both variables. Similarly, in the field of phylogenetic comparative analyses, it is important to develop methods that not only control for phylogenetic autocorrelation but also consider measurement errors in both variables.

However, it is crucial to recognize that methods like RMA have their limitations. According to Kilmer and Rodríguez (2017), RMA does not account for covariance between variables, which can lead to the overestimation of slopes. While RMA addresses some limitations of OLS, it is not without challenges, particularly when dealing with natural variability beyond measurement error. Accordingly, OLS-based phylogenetic comparative methods, like PGLS and PVR, are still acceptable, especially when measurement errors are relatively small and do not significantly impact conclusions.

Furthermore, corphylo (correlations among multiple traits with phylogenetic signal) offers a framework for analyzing correlations among traits, including phylogenetic signals (Zheng et al. 2009). However, this method does not specifically handle regression-related concerns, such as distinguishing between dependent and independent variables. Although many evolutionary analyses focus primarily on determining whether there is a significant correlation between traits rather than on the specific influence of one on the other, corphylo has not yet been widely adopted in this capacity.

Robustness of R^2 Criterion in PVR Across Phylogenetic Tree Structures

Our results indicate that selecting appropriate dependent and independent variables significantly enhances the reliability of PVR analysis, particularly when guided by model selection criteria such as R^2. As Rohlf (2001) and Diniz-Filho et al. (2012) have noted, the R^2 value in PVR is closely tied to the eigenvalues of selected eigenvectors, which reflect the phylogenetic tree's structure. This study incorporated both fixed balanced and random trees, with random trees offering a more realistic representation of potential evolutionary scenarios. Although the performance of PVR and its associated R^2 metric depends on the underlying evolutionary model, differences between random and fixed balanced trees had minimal impact on the overall results, underscoring the robustness of this approach.

Explanation of Unexpected Results for R^2 and AIC in PVR Analysis

In some scenarios, we observed that models with lower R^2 values and higher AIC scores produced more accurate correlation results when evaluating the outcomes of PVR analyses using the L2 estimator (i.e., OLS, Table 2). This counterintuitive finding likely reflects the nature of the phylogenetic signal and the trait evolution model. In scenarios where the phylogenetic signal was weaker, such as in the "Norm1 & Norm1+Norm2" scenario, higher R^2 values may have overestimated model fit due to the lack of a strong phylogenetic structure. Similarly, models with lower AIC values may have been overfitting the noise in the data, leading to less accurate predictions.

Notably, when using the M estimator, models with higher AIC scores generally performed better (Table S16). This pattern was also observed with other robust estimators (Tables S14–S17) and was not limited to scenarios with weaker phylogenetic autocorrelation, 'Norm1 & Norm1+Norm2.' Interestingly, these outcomes were most frequently seen in scenarios like 'BM & BM+Norm,' and even occasionally appeared in the 'BM1 & BM1+BM2' scenario, which had the strongest phylogenetic signal. These findings suggest that the relationship between model selection criteria and phylogenetic structure is complex. In certain cases, simpler models with higher AIC scores may yield more accurate representations of trait relationships, even when the phylogenetic autocorrelation is strong. This highlights the importance of considering the evolutionary context and the nature of the data when interpreting traditional model selection criteria like R^2 and AIC, particularly in phylogenetic comparative analyses.

Challenges of Variable Swapping in Methods Addressing Spatial Autocorrelation

Beyond phylogenetic autocorrelation, spatial autocorrelation also exists, which refers to the phenomenon where observations located closer in space tend to be more similar to each other than those further apart. This effect can be observed in a range of ecological and evolutionary datasets, necessitating specialized statistical methods that account for spatial dependence to avoid biased results (Doreian 1980, Dow et al. 1982, Legendre 1993, He et al. 2003, Cottenie 2005, Diniz-Filho and Bini 2005, Dale and Fortin 2009). While the principal focus of this study has been on the impacts of variable selection within PVR, our preliminary investigations into other spatial statistical techniques—such as spatial eigenvector mapping, conditional autoregressive models, simultaneous autoregressive models, and trend-surface generalized additive model—have also indicated similar discrepancies when dependent and independent variables are swapped. These findings are documented in Supplementary Table S21, which summarize the initial results and suggest that the issue of variable swapping leading to discrepancies in model outcomes may be widespread across various statistical models. Detailed analysis of these results will be pursued in future work to elucidate further the implications for a broader range of ecological and evolutionary research methodologies. This anticipates a significant expansion of our understanding of how methodological choices impact the reliability and interpretability of statistical models in these fields.

Future Directions and Methodological Considerations

While distinguishing between $X_1{\sim}X_2$ and $X_2{\sim}X_1$, R^2 consistently demonstrated good performance across several evolutionary contexts. In light of suggestions from a reviewer, future work may benefit from exploring the use of phylogenetic signal-representation (PSR) curves to assess non-stationary evolutionary patterns, as described by Diniz-Filho et al. (2019). PSR curves could provide additional insights by offering an alternative to traditional phylogenetic signal measures like Pagel's λ and Blomberg's K, particularly in cases where more detailed assessments of non-stationarity and complex evolutionary dynamics are required.

Robust phylogenetic regression, as proposed by Adams et al. (2024), offers a promising approach to address some of the limitations inherent in traditional phylogenetic comparative methods, such as sensitivity to outliers and deviations from normality. In our study, we incorporated robust phylogenetic regression into PVR to investigate the performance of different criteria (e.g., R^2, λ, and AIC) for selecting dependent variables. Our findings demonstrate that using the L2 and MM estimators, the R^2 criterion was effective in identifying the appropriate dependent variable, suggesting that R^2 is a reliable standard in these contexts. With the L1 estimator, R^2 also performed reasonably well, albeit with some inconsistencies. By contrast, under the M and S estimators, none of the criteria proved consistently effective for selecting dependent variables, highlighting the challenges of using robust PVR with these estimators. Future studies are needed to develop more effective standards for variable selection under these conditions.

In some cases, it becomes necessary to account for multiple sources of autocorrelation simultaneously, such as controlling for both spatial and phylogenetic autocorrelation. This introduces additional complexity in selecting variables and eigenvectors, as balancing multiple forms of autocorrelation must be tailored to the specific goals of the analysis. Safi and Pettorelli (2010) discuss how to balance phylogenetic, spatial, and environmental components in such cases, providing useful insights into eigenvector selection strategies when dealing with multiple forms of autocorrelation. Although our study focuses primarily on phylogenetic relationships, their approach highlights the importance of considering different sources of autocorrelation depending on the study's objectives, which may be relevant in future studies requiring the management of multiple or complex sources of autocorrelation.

Data availability

Supplementary Tables S1-S20 and Figures S1-S5 are included in Supplementary_tables_figures.docx. The raw data for the four simulation scenarios are provided in the file Raw_data.zip. The scripts used to generate the results presented in the main text, including phylogenetic tree construction, trait data simulation, regression model fitting, and statistical analysis, are available in Main_code.zip. The relevant codes for evaluating spatial statistical methods are included in Spatial_methods.zip. Additionally, standalone functions for eigenvector selection (MoransI.R) and regression (ESRRV.R) are available separately. All these files and .R scripts can be found in the Dryad data repository: https://datadryad.org/xxxxx.

Acknowledgments

We are deeply grateful to the reviewers for their valuable comments and suggestions, which significantly improved the quality of our work. In particular, we would like to thank Prof. J. A. F. Diniz-Filho for his insightful feedback. We also appreciate the contributions from the anonymous reviewer, whose detailed suggestions greatly helped us refine our analysis and presentation. This study was partially funded by the National Natural Science Foundation of China (grant number 31671321). To enhance language and readability, this article has undergone refinement by ChatGPT 4o. Subsequently, the authors conducted a thorough review and made necessary edits, taking full responsibility for the publication's content.

Author Contributions

DKN and ZLC conceived the research hypothesis; ZLC performed the simulations and analyzed the data. ZLC drafted the manuscript, and DKN revised the manuscript. Both authors agree with the final version and the submission.

References

Adams R., Cain Z., Assis R., DeGiorgio M. 2024. Robust phylogenetic regression. Syst. Biol., 73:140-157.

Burns J.H., Anacker B.L., Strauss S.Y., Burke D.J. 2015. Soil microbial community variation correlates most strongly with plant species identity, followed by soil chemistry, spatial location and plant genus. AoB PLANTS, 7.

Casson R.J., Farmer L.D. 2014. Understanding and checking the assumptions of linear regression: a primer for medical researchers. Clinical & Experimental Ophthalmology, 42:590-596.

Chen Z.-L., Guo H.-J., Niu D.-K. 2023. Dependent variable selection in phylogenetic generalized least squares regression analysis under Pagel's lambda model. bioRxiv:2023.2005.2021.541623.

Cheverud J.M., Dow M.M., Leutenegger W. 1985. The quantitative assessment of phylogenetic constraints in comparative analyses: sexual dimorphism in body weight among primates. Evolution, 39:1335-1351.

Cornwallis C.K., Griffin A.S. 2024. A guided tour of phylogenetic comparative methods for studying trait evolution. Annu. Rev. Ecol. Evol. Syst., 55:181-204.

Cottenie K. 2005. Integrating environmental and spatial processes in ecological community dynamics. Ecol. Lett., 8:1175-1182.

Dale M.R.T., Fortin M.-J. 2009. Spatial autocorrelation and statistical tests: Some solutions. JABES, 14:188-206.

Desdevises Y., Legendre P., Azouzi L., Morand S. 2003. Quantifying phylogenetically structured environmental variation. Evolution, 57:2647-2652.

Diniz-Filho J.A.F., Bini L.M. 2005. Modelling geographical patterns in species richness using eigenvector-based spatial filters. Glob. Ecol. Biogeogr., 14:177-185.

Diniz−Filho J.A.F., Bini L.M., Rangel T.F., Morales-Castilla I., Olalla-Tárraga M.Á., Rodríguez M.Á., Hawkins B.A. 2012. On the selection of phylogenetic eigenvectors for ecological analyses. Ecography 35:239-249.

Diniz−Filho J.A.F., Cianciaruso M.V., Rangel T.F., Bini L.M. 2011. Eigenvector estimation of phylogenetic and functional diversity. Funct. Ecol., 25:735-744.

Diniz−Filho J.A.F., de Sant'Ana C.E.R., Bini L.M. 1998. An eigenvector method for estimating phylogenetic inertia. Evolution, 52:1247-1262.

Diniz−Filho J.A.F., Jardim L., Mondanaro A., Raia P. 2019. Multiple components of phylogenetic non-stationarity in the evolution of brain size in fossil hominins. Evol. Biol., 46:47-59.

Diniz−Filho J.A.F., Nabout J.C. 2009. Modeling body size evolution in Felidae under alternative phylogenetic hypotheses. Genet Mol Biol, 32:170-176.

Diniz−Filho J.A.F., Rodríguez M.Á., Bini L.M., Olalla-Tarraga M.Á., Cardillo M., Nabout J.C., Hortal J., Hawkins B.A. 2009. Climate history, human impacts and global body size of Carnivora (Mammalia: Eutheria) at multiple evolutionary scales. J. Biogeogr., 36:2222-2236.

Diniz−Filho J.A.F., Tôrres N.M. 2002. Phylogenetic comparative methods and the geographic range size – body size relationship in new world terrestrial carnivora. Evol. Ecol., 16:351-367.

Divíšek J., Chytrý M., Beckage B., Gotelli N.J., Lososová Z., Pyšek P., Richardson D.M., Molofsky J. 2018. Similarity of introduced plant species to native ones facilitates naturalization, but differences enhance invasion success. Nat. Commun., 9:4631.

Doreian P. 1980. Linear models with spatially distributed data: spatial disturbances or spatial effects? Soc. Methods Res., 9:29-60.

Dormann C.F., McPherson J.M., Araújo M.B., Bivand R., Bolliger J., Carl G., Davies R.G., Hirzel A., Jetz W., Kissling W.D., Kühn I., Ohlemüller R., Peres-Neto P.R., Reineking B., Schröder B., Schurr F.M., Wilson R. 2007. Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. Ecography, 30:609-628.

Dow M.M., Burton M.L., White D.R. 1982. Network autocorrelation: A simulation study of a foundational problem in regression and survey research. Social Networks, 4:169-200.

Felsenstein J. 1985. Phylogenies and the comparative method. Am. Nat., 125:1-15.

Freckleton R.P., Cooper N., Jetz W. 2011. Comparative methods as a statistical fix: the dangers of ignoring an evolutionary model. Amer. Natur., 178:E10-E17.

Gao X., Dai J., Tao Z., Shahzad K., Wang H. 2023. Autumn phenology of tree species in China is associated more with climate than with spring phenology and phylogeny. Front. Plant Sci., 14:1040758.

Garamszegi L.Z. 2014. Modern Phylogenetic Comparative Methods and Their Application in Evolutionary Biology: Concepts and Practice. Berlin, Springer.

Gervini D., Yohai V.J. 2002. A class of robust and fully efficient regression estimators. The Annals of Statistics, 30:583-616.

Grafen A. 1989. The phylogenetic regression. Philos. Trans. R. Soc. Lond. B Biol. Sci., 326:119-157.

Griffith D.A., Peres-Neto P.R. 2006. Spatial modeling in ecology: the flexibility of eigenfunction spatial analyses. Ecology, 87:2603-2613.

Guerrero V.M. 1993. Time-series analysis supported by power transformations. Journal of Forecasting, 12:37-48.

Harvey P.H., Read A.F., Nee S. 1995. Why ecologists need to be phylogenetically challenged. J Ecol, 83:535-536.

He F., Zhou J., Zhu H. 2003. Autologistic regression model for the distribution of vegetation. JABES, 8:205-222.

Huber P.J. 1964. Robust estimation of a location parameter. Ann. Math. Statist., 35:73-101, 129.

Jarantow S.W., Pisors E.D., Chiu M.L. 2023. Introduction to the use of linear and nonlinear regression analysis in quantitative biological assays. Current Protocols, 3:e801.

Keitt T.H., Bjørnstad O.N., Dixon P.M., Citron-Pousty S. 2002. Accounting for spatial pattern when modeling organism-environment interactions. Ecography, 25:616-625.

Kilmer J.T., Rodríguez R.L. 2017. Ordinary least squares regression is indicated for studies of allometry. J. Evol. Biol., 30:4-12.

Lanuza J.B., Rader R., Stavert J., Kendall L.K., Saunders M.E., Bartomeus I. 2023. Covariation among reproductive traits in flowering plants shapes their interactions with pollinators. Funct. Ecol., 37:2072-2084.

Legendre P. 1993. Spatial autocorrelation: trouble or new paradigm? Ecology, 74:1659-1673.

LeRoy C.J., Hipp A.L., Lueders K., Follstad Shah J.J., Kominoski J.S., Ardón M., Dodds W.K., Gessner M.O., Griffiths N.A., Lecerf A., Manning D.W.P., Sinsabaugh R.L., Webster J.R. 2020. Plant phylogenetic history explains in-stream decomposition at a global scale. J Ecol, 108:17-35.

McArdle B.H. 1988. The structural relationship: regression in biology. Canadian journal of zoology, 66:2329-2339.

O'Meara B. 2016. Phylogenetic Comparative Method. In: Kliman RM editor. Encyclopedia of Evolutionary Biology. Oxford, Academic Press, p. 254-256.

O’Meara B.C., Beaulieu J.M. 2016. Past, future, and present of state-dependent models of diversification. Am. J. Bot., 103:792-795.

Osborne J.W., Waters E. 2002. Four assumptions of multiple regression that researchers should always test”, Practical Assessment, Research, and Evaluation 8(1): 2. doi: . Practical Assessment, Research, and Evaluation, 8:2.

Pagel M. 1997. Inferring evolutionary processes from phylogenies. Zool. Scr., 26:331-348.

Pagel M. 1999. Inferring the historical patterns of biological evolution. Nature, 401:877-884.

Paradis E., Schliep K. 2019. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics, 35:526-528.

Peres-Neto P.R., Jackson D.A., Somers K.M. 2005. How many principal components? stopping rules for determining the number of non-trivial axes revisited. Comput. Statist. Data Anal., 49:974-997.

Peres-Neto P.R., Legendre P. 2010. Estimating and controlling for spatial structure in the study of ecological communities. Glob. Ecol. Biogeogr., 19:174-184.

Peres-Neto P.R., Leibold M.A., Dray S. 2012. Assessing the effects of spatial contingency and environmental filtering on metacommunity phylogenetics. Ecology, 93:S14-S30.

Phillips R.F. 2002. Least absolute deviations estimation via the EM algorithm. Statistics and Computing, 12:281-285.

Revell L.J. 2010. Phylogenetic signal and linear regression on species data. Methods Ecol. Evol., 1:319-329.

Revell L.J. 2024. phytools 2.0: an updated R ecosystem for phylogenetic comparative methods (and other things). PeerJ, 12:e16505.

Rohlf F.J. 2001. Comparative methods for the analysis of continuous variables: geometric interpretations. Evolution, 55:2143-2160.

Safi K., Pettorelli N. 2010. Phylogenetic, spatial and environmental components of extinction risk in carnivores. Glob. Ecol. Biogeogr., 19:352-362.

Symonds M.R.E., Blomberg S.P. 2014. A primer on phylogenetic generalised least squares. In: Garamszegi LZ editor. Modern Phylogenetic Comparative Methods and Their Application in Evolutionary Biology: Concepts and Practice. Berlin, Heidelberg, Springer Berlin Heidelberg, p. 105-130.

Uyeda J.C., Zenil-Ferguson R., Pennell M.W. 2018. Rethinking phylogenetic comparative methods. Syst. Biol., 67:1091-1109.

Venables W.N., Ripley B.D. 2002. Modern Applied Statistics with S. 4th ed., Springer.

Westoby M., Leishman M.R., Lord J.M. 1995. On Misinterpreting the `Phylogenetic Correction'. J Ecol, 83:531-534.

Yu C., Yao W. 2017. Robust linear regression: A review and comparison. Commun. Stat. Simul. Comput., 46:6261-6282.

Zheng L., Ives A.R., Garland T., Larget B.R., Yu Y., Cao K. 2009. New multivariate tests for phylogenetic signal and trait correlations applied to ecophysiological phenotypes of nine Manglietia species. Funct. Ecol., 23:1059-1069.