Optimized Phylogenetic Eigenvector Regression Provides Robust and Flexible Trait Correlation Analysis under Diverse Evolutionary Conditions

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 comparative methods are essential tools for analyzing trait correlations among species while controlling for evolutionary dependencies. Phylogenetic eigenvector regression (PVR) uses eigenvectors (EVs) derived from the phylogenetic tree as covariates to remove shared evolutionary history. Although conceptually simple and independent of predefined evolutionary models, PVR has traditionally been viewed as inferior to phylogenetic generalized least squares (PGLS), which explicitly incorporates evolutionary models. Here, we systematically optimized the PVR framework by integrating robust statistical estimators (L1, L2, M, and MM) and developing a theoretically justified approach to selecting EVs, particularly by using the union of EVs from both traits rather than a single trait alone. Using extensive simulations, we compared optimized PVR methods to established approaches, including PGLS and phylogenetic independent contrasts (PIC), across a range of evolutionary scenarios, spanning strong, weak, and asymmetric phylogenetic signals as well as non-stationary shifts; and across multiple tree structures (balanced and ladder-shaped), branch-length scalings ($\rho = 0.1, 1, 2$), and sample sizes (128 and 16 species). Our results challenge the prevailing assumption that PGLS universally outperforms PVR. While PGLS and standard PIC generally excelled when dependent traits exhibited strong phylogenetic signals, optimized PVR methods—particularly PVR-MM—demonstrated superior robustness under scenarios involving mixed or weak phylogenetic signals, as well as abrupt, non-stationary evolutionary shifts. Even the simplest optimized PVR variant (PVR-L2) exhibited inherent robustness, highlighting a fundamental advantage of phylogenetic eigenvectors in controlling evolutionary structure. Furthermore, PVR provides greater methodological flexibility compared to model-based methods like PGLS, as it is not constrained by predefined evolutionary models and can readily extend to advanced machine-learning approaches. Lastly, unlike PGLS, PVR effectively addresses the directional asymmetry inherited from ordinary least squares regression, allowing researchers to select predictors based on biological reasoning or measurement accuracy without compromising model validity. Overall, this study highlights that optimized PVR is not merely a competitive alternative to traditional methods, but also a flexible and robust framework that adapts well to a wide range of evolutionary scenarios and analytical needs.

KEYWORDS: phylogenetic comparative methods; phylogenetic eigenvector regression (PVR); robust regression; evolutionary shift; phylogenetic autocorrelation; phylogenetic signal; dependent and independent variable asymmetry; trait correlation.

INTRODUCTION

Phylogenetic comparative methods correct for non-independence among species due to shared ancestry (Felsenstein 1985, Garamszegi 2014, Cornwallis and Griffin 2024, Dewar et al. 2025). Among these, phylogenetic generalized least squares (PGLS) is widely regarded as the standard method, modeling trait covariation using predefined evolutionary models such as Brownian motion (BM) or Ornstein-Uhlenbeck (OU) (Grafen 1989, Pagel 1999, Garamszegi 2014). By contrast, phylogenetic eigenvector regression (PVR) uses phylogenetic eigenvectors (EVs)—orthogonal axes extracted from the phylogenetic tree—to statistically isolate trait correlations from phylogenetic structure (Diniz-Filho et al. 1998, Diniz-Filho and Tôrres 2002, Diniz-Filho et al. 2014). These EVs 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).

Despite its intuitive appeal and flexibility, PVR has been considered less reliable than PGLS in controlling phylogenetic non-independence (Pennell and Harmon 2013). This conventional view—articulated in foundational comparative studies (Laurin 2010, Freckleton et al. 2011)—suggests that PVR lacks the explicit evolutionary modeling necessary for producing robust estimates, and may suffer from inflated Type I error rates particularly when broken-stick criterion was used to select EVs. Moreover, PVR has been viewed as a purely descriptive tool rather than one suitable for detecting adaptation or making evolutionary inferences (Hansen 2014). As a result, PVR has been underutilized, despite its potential for broader applications.

A key challenge to PVR lies in the selection of EVs. Including all EVs 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 EVs can result in the loss of crucial phylogenetic information. Fortunately, researches have demonstrated that not all EVs are necessary for robust analysis, as many contribute minimally to the overall model. By carefully selecting a subset of EVs, 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. 2012a). These selected EVs 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).

Although EVs are derived solely from the phylogenetic tree, the selection of which EVs to include in a regression model is inherently trait-specific. Different biological traits emphasize different components of the tree's structure, leading to divergent EV sets depending on which trait is analyzed. This divergence is not trivial: when EVs are selected separately based on different traits, subsequent regression analyses may yield conflicting correlation results. Rather than focusing on distinctions between dependent and independent variables, which are often ambiguous in biological systems, we addressed this issue by treating the two traits symmetrically. Inspired by Cheverud's (1985) decomposition of trait covariation into phylogenetic and specific components (T = P + S), we hypothesized that using the union of EVs selected from both traits (*EV_U*) would provide a more stable and comprehensive representation of the phylogenetic structure relevant to either trait. Subsequent simulation analyses confirmed this hypothesis, showing that *EV_U* led to more accurate estimates of trait correlations across a wide range of evolutionary conditions.

In addition to improving EV selection, we explored a second optimization strategy for PVR: enhancing its robustness to model violations and outliers through the integration of robust regression estimators. Traditional ordinary least squares (OLS) regression, as used in standard PVR, assumes normally distributed errors and is sensitive to extreme values. Robust regression techniques—such as L1, M-, S-, and MM-estimators—have been developed to mitigate these limitations by down-weighting the influence of outliers (Huber 1964, Yu and Yao 2017). Robust regression estimators have increasingly been adapted for use in phylogenetic comparative methods (Slater and Pennell 2014, Adams et al. 2024). In particular, Adams et al. (2024) demonstrated that the MM-estimator applied to phylogenetic independent contrasts (PIC), PIC-MM, not only outperformed other robust variants of PIC and PGLS (e.g., PIC-L1, PIC-M, PGLS-L1), but also consistently surpassed standard implementations such as PIC-L2 and PGLS-L2—especially under non-stationary evolutionary conditions.

This study revisits and systematically improves the PVR framework through two methodological enhancements: (1) using *EV_U*, the full union of eigenvectors selected separately from X_1 and X_2, including their shared components, to more comprehensively account for overlapping and trait-specific phylogenetic signals; (2) integrating robust

regression estimators to improve model resilience to outliers and deviations from normality. These optimized variants are then benchmarked against established methods such as PGLS and PIC across a wide range of evolutionary conditions. Contrary to the long-standing view that PGLS consistently outperforms alternative approaches (Laurin 2010, Freckleton et al. 2011, Pennell and Harmon 2013, Hansen 2014), our findings demonstrate that optimized PVR methods—particularly those incorporating robust estimators—can match or exceed PGLS in accuracy, primarily under non-stationary or weak signal scenarios. These results highlight the context-dependent nature of methodological performance and suggest that comparative researchers should move beyond one-size-fits-all approaches when selecting analytical tools.

MATERIALS AND METHODS

Simulation of Phylogenetic Trees

We simulated five fully bifurcating phylogenetic trees to evaluate the performance of PVR, PGLS, and PIC under diverse evolutionary conditions. (1) A 128-species balanced tree with uniformly scaled branch lengths ($\rho = 1$) was used as the baseline structure. (2) To assess the effects of branch-length scaling (Grafen 1989), we constructed two additional 128-species balanced trees: one with compressed ancestral and elongated terminal branches ($\rho = 0.1$), and one with elongated ancestral and shortened terminal branches ($\rho = 2$). (3) To evaluate the impact of taxonomic sampling, we included a 16-species balanced tree with uniformly scaled branch lengths ($\rho = 1$). (4) To examine performance under topological asymmetry, we simulated a 128-species ladder-shaped (comb-shaped) tree with $\rho = 1$. This tree represents extreme imbalance, with each bifurcation producing one descendant lineage that does not split further—modeling prolonged unbranched evolutionary trajectories.

All balanced and ladder-shaped phylogenetic trees were constructed using the *stree* function from the R package ape (Version 5.8) (Paradis and Schliep 2019) and the *paste.tree* function from the R package phytools (Version 2.3-0) (Revell 2024).

Simulated Trait Generation

We simulated the evolution of two traits (X_1 and X_2) along each of the five phylogenetic trees, following the simulation logic proposed in Revell (2010) and Adams et al. (2024). In all scenarios, X_2 was modeled as:

$$X_2 = \beta X_1 + \varepsilon$$

where β is a scalar coefficient controlling the trait association, and ε is an independently simulated noise term. Unless otherwise noted, β was set to 1 to introduce a correlation between the traits. The magnitude and structure of ε were varied to simulate different evolutionary dynamics. We tested eight levels of ε magnitude: 10^{-4}, 0.01, 1, 4, 16, 64, 256, and 1024. Each setting was repeated 1,000 times per tree, totaling 8,000 simulations per scenario and 40,000 simulations per tree type.

Five simulation scenarios were defined. (1) **BM1 & BM1+BM2**, where X_1 evolves under Brownian motion (BM) with unit variance. ε also evolves under a separate BM process with the variable magnitude. Both X_1 and X_2 thus exhibit strong phylogenetic autocorrelation. (2) **BM & BM+Norm**, where X_1 evolves under BM (variance = 1). ε is drawn from an i.i.d. normal distribution with varying magnitude. This setup retains strong phylogenetic signal in X_1, but the phylogenetic signal in X_2 is diluted by non-phylogenetic noise. (3) **Norm & Norm+BM**, where X_1 follows a normal distribution (mean = 0, variance = 1) without phylogenetic structure. ε evolves under BM. X_2 thus acquires some phylogenetic structure from ε, while X_1 remains unstructured. (4) **Norm1 & Norm1+Norm2**, where Both X_1 and ε follow i.i.d. normal distributions. Neither trait is expected to carry phylogenetic signal. This scenario serves as a null benchmark, though some simulations may exhibit weak incidental phylogenetic structure due to noise. (5) **Shift**, where β was set to 0 to eliminate inherent correlation between X_1 and X_2. Both X_1 and ε evolved independently under a BM model with unit variance. To simulate non-stationary evolutionary change, a synchronous, instantaneous shift was introduced along a single internal branch near the root. In the ladder-shaped tree, this shift was applied to the branch leading to the lineage that continues branching. The shift simultaneously displaced the values of X_1 and ε by a specified magnitude s, which was varied across eight values: 10^{-4}, 0.01, 1, 4, 16, 64, 256, and 1024. The case $s = 1$ served as the baseline (no shift), while other values captured greater degrees of evolutionary displacement without introducing causal association between traits.

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

Building on Cheverud's (1985) autoregressive framework (T = P + S), where T denotes trait values, P represents the phylogenetic component, and S accounts for other influences, PVR applies principal coordinate analysis to a double-centered phylogenetic distance matrix to derive $n - 1$ EVs, from which a relevant subset is selected as covariates in regression models.

EVs were selected separately for each trait (X_1 or X_2) by fitting the models: $X_1 = EV_X_1 * \beta + \varepsilon$, and $X_2 = EV_X_2 * \beta + \varepsilon$, where β is the regression coefficient vector and ε is the residual (S component). We adopted three non-sequential EV selection methods, adapted from references (Diniz-Filho and Nabout 2009, Diniz-Filho et al. 2012a). (1) In the ESRRV approach, EVs significantly correlated with the dependent variable were identified via Pearson or Spearman correlations, depending on normality. (2) The mAIC method applied forward stepwise regression to sequentially add EVs that minimized the Akaike Information Criterion (AIC), with the selection process terminating when no further improvement could be achieved. (3) The mMorI approach used forward selection to include EVs until the residuals minimized Moran's I, a statistic ranging from –1 to +1 that quantifies the strength of phylogenetic autocorrelation, with 0 indicating no detectable autocorrelation.

To simplify interpretation and improve computational performance, we made a practical adjustment to the mMorI algorithm. Originally developed to detect spatial autocorrelation, Moran's I is capable of capturing both positive and negative structure by quantifying the clustering of similar values. In conventional applications, variable selection continues until the residuals reach non-significance (e.g., $p > 0.05$), indicating that spatial autocorrelation has been adequately removed. However, in phylogenetic contexts, trait values among closely related species tend to be positively correlated—negative phylogenetic autocorrelation is theoretically possible but biologically rare. Therefore, rather than relying solely on *p*-values, we modified the mMorI procedure to terminate the eigenvector selection process once a negative value of Moran's I is observed in the residuals. If no negative value was reached, selection proceeded until a non-significant *p*-value threshold was met.

To accomplish this, we rewrote the relevant code in the PVR package for selecting EVs by minimizing Moran's I, resulting in a standalone function named *MoransI*. This custom script is now used for Moran's I method of EV selection in PVR, while all other computations, such as EV decomposition, still utilize the original *PVR* package (Diniz-Filho et al. 1998, Diniz-Filho et al. 2012a, Diniz-Filho et al. 2012b). 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).

Accounting for Phylogenetic Autocorrelation in Trait Regression Models

Following EV selection, we implemented four approaches to account for phylogenetic autocorrelation. (1) In the ***EV_X₁*-corrected regression**, the EVs selected for trait X_1 were included as covariates in the model $X_1 \sim X_2 + EV_X_1$. (2) In the ***EV_X₂*-corrected regression**, the EVs from trait X_2 were used as covariates in the model $X_2 \sim X_1 + EV_X_2$. It is important to note that the relationship between X_1 and X_2 derived from $X_2 \sim X_1 + EV_X_2$ is equivalent to that from $X_1 \sim X_2 + EV_X_2$, since both models adjust for phylogenetic autocorrelation in trait X_2, and the direction of the relationship does not influence the correlation or slope between the two traits. (3) In the ***EV_U*-corrected regression**, the union of EVs selected from both traits (denoted *EV_U*) was incorporated as covariates in the model $X_2 \sim X_1 + EV_U$. (4) Finally, in the **residual correlation analysis**, each trait was regressed on its own EVs (X_1 on EV_X_1 and X_2 on EV_X_2), and the Pearson correlation between the resulting residuals was calculated. To avoid inflated type I error due to overfitting, degrees of freedom were adjusted to n – k – 1, where n is the number of species and k is the number of eigenvectors in *EV_U* (Diniz-Filho and Tôrres 2002).

Robust Estimators Used in PVR Analysis

To address potential violations of model assumptions—such as outliers and non-normal residuals—we incorporated several robust regression estimators into our PVR analysis. The **L2** estimator, equivalent to ordinary least squares (OLS), represents the conventional regression method in robust regression contexts. It minimizes the sum of squared residuals but is highly sensitive to outliers due to its quadratic loss function, which disproportionately amplifies large errors. The **L1** estimator, or least absolute deviation, minimizes the sum of absolute residuals. Compared to L2, it is less affected by extreme values and performs better when outliers are present. **M**-estimators generalize the likelihood framework and down-weight the influence of large residuals. They behave like L2 for small errors and like L1 for large errors, striking a balance between efficiency and robustness. **S**-estimators minimize a robust measure of residual spread—typically based on median absolute deviation—offering a high breakdown point and strong resistance to outliers. **MM**-estimators combine S- and M-estimators, using the former for initial robustness and the latter for efficient final estimation. This hybrid approach ensures both high robustness and statistical efficiency, making it well-suited for noisy or heterogeneous datasets. 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

Trait evolution for X_1 and X_2 was simulated along phylogenetic trees, with values recorded at both tips and internal nodes. This design enabled the calculation of standardized evolutionary changes per branch, expressed as $\Delta X_1/L$ and $\Delta X_2/L$, where L denotes is branch length. These changes are independent from one branch to another, which allows us to use conventional correlation tests to evaluate the relationship between trait changes across the tree.

Our approach, while conceptually similar to PIC, avoids the need for model-based ancestral state reconstruction by directly recording internal node values during simulations. In contrast, PIC relies on reconstructing these states under a BM assumption, which may not hold in many empirical datasets. Notably, PIC is mathematically equivalent to PGLS only when the BM assumption is satisfied (Rohlf 2001, Blomberg et al. 2012). Because our simulations are not constrained by such assumptions, they provide a more accurate and flexible benchmark for evaluating comparative methods, including PIC, PGLS, and optimized PVR.

To ensure methodological consistency across datasets and facilitate fair comparisons, we uniformly applied Spearman's rank correlation to assess the association between $\Delta X_1/L$ and $\Delta X_2/L$. This nonparametric method is robust to non-normality and outliers, and avoids assumptions such as linearity or homoscedasticity.

RESULTS

To evaluate regression outcomes under different EV sets, we applied two models: OLS($X_1 \sim X_2 + EV_X_1$), where X_1 is the dependent variable; and OLS($X_2 \sim X_1 + EV_X_2$), where X_2 is the dependent variable. In each model, the corresponding eigenvectors (EV_X_1 or EV_X_2) were included as covariates to account for phylogenetic structure. Unless otherwise specified, all results in this section are based on 128-species balanced phylogenetic trees with uniformly scaled branch lengths.

Eigenvector-Set Size and Exclusion of mAIC

To assess model complexity and mitigate the risk of overfitting, we first examined the number of EVs selected by each method, ESRRV, mAIC, and mMorI. These methods were applied separately to traits X_1 and X_2, and the resulting EVs were used as covariates in regression models. As shown in Figure 1, the mAIC method produced substantially larger EV sets, with median counts of 57 for EV_X_1, 56 for EV_X_2, and a striking 123 for their union (*EV_U*), which is close to the total number of species (N = 128) in the dataset. Such dimensionality severely compromises degrees of freedom and increases the risk of overfitting. In contrast, ESRRV and mMorI consistently produced much smaller EV sets, preserving model simplicity and statistical power. As a result, mAIC was excluded from subsequent comparative analyses, and only ESRRV and mMorI were retained for downstream evaluations.

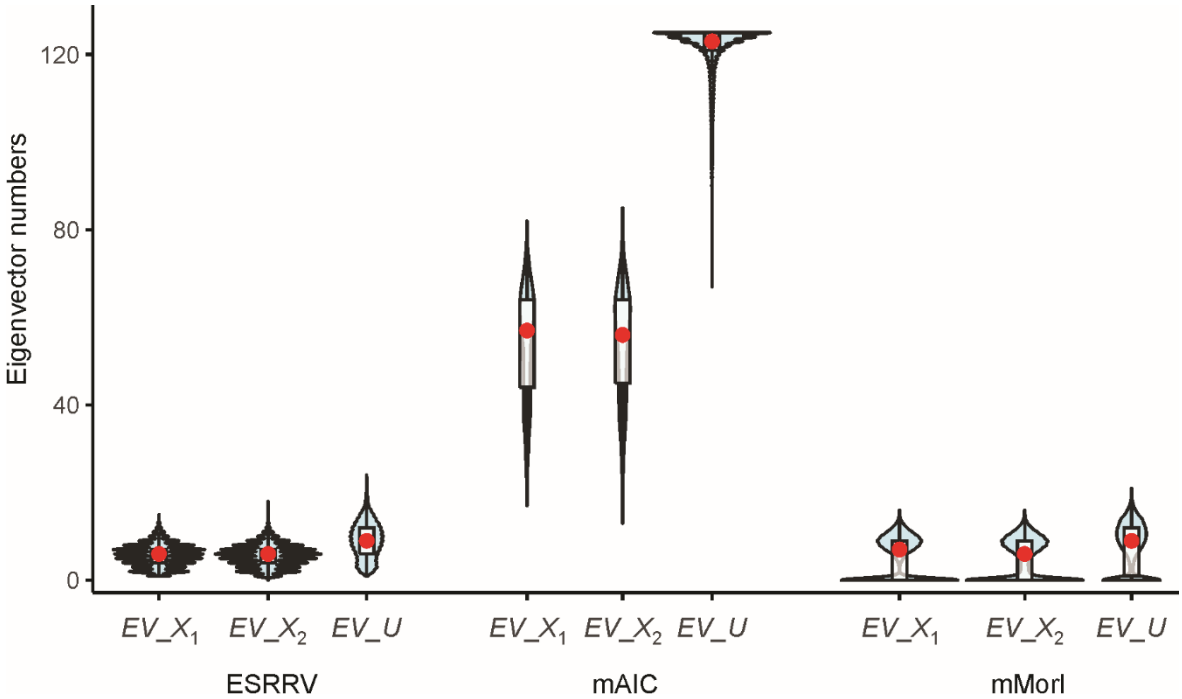

Figure 1. Number of phylogenetic eigenvectors (EVs) selected by each method across trait sets. For each method, EVs were selected separately for trait X_1 and trait X_2. The mAIC method produced notably larger sets, with median counts reaching 57, 56, and 123 for EV_X_1, EV_X_2, and their union *EV_U*, respectively. By contrast, ESRRV and mMorI consistently selected fewer EVs, with union set medians remaining below 10. The total number of species in the dataset was 128. See MATERIALS AND METHODS for details regarding the three EV selection methods, ESRRV, mAIC, and mMorI. The data used to plot this figure are available in Supplementary Table S1.

Impact of Variable Designation on Eigenvector Selection

The designation of dependent and independent variables significantly influences the results of EV selection. As shown in Table 1, the discrepancy rate between EV selections was notably high, with an average of 77% across all methods and simulation scenarios. This demonstrates that relying on EVs selected from only one trait in regression models—without accounting for discrepancies between the two sets—can introduce considerable risk of bias, potentially compromising the robustness and reliability of the analysis.

Table 1. Inconsistency rates between phylogenetic eigenvector (EV) sets selected from traits X_1 and X_2.

Simulation scenarios	ESRRV	mMorI
BM1 & BM1+BM2	83%	84%
BM & BM+Norm	80%	81%
Norm & Norm+BM	86%	76%
Norm1 & Norm1+Norm2	86%	4%
Shift	94%	100%

Values represent the percentage of inconsistent selections across 8,000 simulations per scenario. Please refer to the MATERIALS AND METHODS for details regarding the EV selection methods (ESRRV and mMorI) and the five simulation scenarios.

Divergent PVR Results Due to Different Eigenvector Sets

To assess how discrepancies between EV_X_1 and EV_X_2 affect PVR outcomes, we compared two regression models: OLS($X_1 \sim X_2 + EV_X_1$) and OLS($X_2 \sim X_1 + EV_X_2$), applied across various evolutionary scenarios and EV selection methods. Divergence was defined by either a change in statistical significance ($p < 0.05$ vs. $p \geq 0.05$) or a reversal in slope direction (≥ 0 vs. < 0). As summarized in Table 2, divergence rates varied from 0.13% to 30.7%, depending on the scenario and EV selection method. While mMorI generally produced more consistent results, ESRRV showed greater divergence in several scenarios, most notably in the Shift scenario (30.7%). To quantify these discrepancies, we computed Cohen's Kappa values for each combination (Landis and Koch 1977). Asterisks in the table indicate the degree of disagreement based on Kappa statistics: more stars denote greater inconsistency. These results highlight that the inconsistencies in EV selection can substantially affect

regression outcomes—especially under non-stationary evolutionary conditions—underscoring the need for trait-agnostic strategies such as *EV_U*.

Table 2. PVR model divergence due to differences in EV sets derived from X_1 and X_2.

Simulation scenarios	ESRRV	mMorI
BM1 & BM1+BM2	19.8% **	11.4% **
BM & BM+Norm	9.18% *	18.2% **
Norm & Norm+BM	11.5% **	10.8% **
Norm1 & Norm1+Norm2	6.44% *	0.13% *
Shift	30.7% ***	14.9% ***

Values represent the percentage of divergent regression results across 8,000 simulations per scenario. Divergence was defined as differences in statistical significance ($p < 0.05$ vs. $p \geq 0.05$) or slope direction (≥ 0 vs. < 0) between the two models. Asterisks indicate the degree of inconsistency between regression results based on Cohen's Kappa values (reversed scale): *, $\kappa \geq 0.8$; **, $0.6 \leq \kappa < 0.8$; ***, $0.2 \leq \kappa < 0.4$. See Materials and Methods for details on phylogenetic eigenvector (EV) selection methods (ESRRV and mMorI) and simulation scenarios.

Divergence in Correlation Coefficients Due to EV Set Differences

To quantify the sensitivity of PVR outcomes to differences in EV selection, we examined the absolute differences in partial correlation coefficients between OLS($X_1 \sim X_2 + EV_X_1$) and OLS($X_2 \sim X_1 + EV_X_2$) across five simulation scenarios. To this end, we computed the absolute difference in partial correlation coefficients between the two regression models: $|\text{pcor}(X_1, X_2 + EV_X_1) - \text{pcor}(X_2, X_1 + EV_X_2|$.

As shown in Figure 2, coefficient discrepancies were consistently observed across all five evolutionary scenarios, with many differences exceeding 10% in absolute value. This suggests that EV selection based on variable designation can substantially affect not just the significance of results, but also the estimated strength of trait associations. Notably, in the Norm1 & Norm1+Norm2 scenario—characterized by minimal phylogenetic signal—coefficient differences were markedly smaller, particularly when EVs were selected using the mMorI method. These results reinforce the methodological risks associated with trait-specific EV selection and motivate the use of more stable, trait-agnostic alternatives.

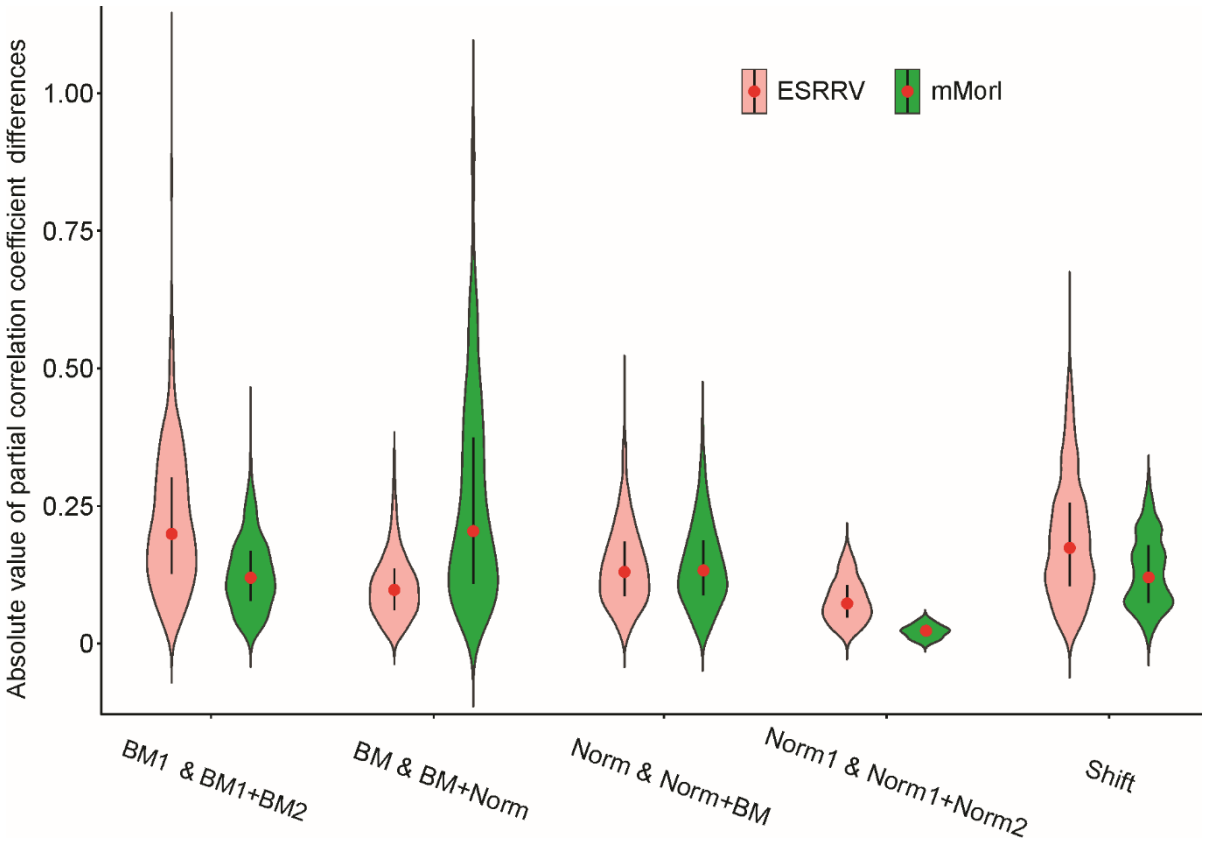

Figure 2. Distributions of differences in partial correlation coefficients between X_1 and X_2, comparing PVR models using EV_X_1 versus EV_X_2 as covariates. The violin plot illustrates the absolute differences in partial correlation coefficients ($|\text{pcor}(X_1, X_2 + EV_X_1) - \text{pcor}(X_2, X_1 + EV_X_2|$) for cases with conflicting PVR results. Red dots indicate the median values, and the vertical lines within the violins show the range of error bars. Paired t-tests and Wilcoxon signed-rank tests both showed that the differences in partial correlation coefficients between using EV_X_1 and EV_X_2 as covariates were highly significant across all scenarios (all p-values $< 10^{-23}$). See MATERIALS AND METHODS for details regarding the EV selection methods (ESRRV and mMorI) and the simulation scenarios (BM1 & BM1+BM2, BM & BM+Norm, Norm & Norm+BM, and Norm1 & Norm1+Norm2, and Shift). The data used to plot this figure are available in Supplementary Table S2.

Theoretical Rationale for Using the Union of Trait Eigenvectors

Our empirical findings demonstrated that regression outcomes based on EV sets selected from different traits (EV_X_1 vs. EV_X_2) often diverged—both in terms of statistical significance and estimated coefficients. This variability underscores a key methodological issue: the conventional approach of selecting EVs based on a single trait may not adequately capture the phylogenetic structure of the traits.

To clarify this point, we revisited the trait decomposition framework T = P + S, originally proposed by Cheverud (1985), where where P represents the phylogenetic component and S the specific or residual variation. Our derivation (see Appendix 1) shows that when two traits share overlapping but non-identical phylogenetic structure, relying solely on trait-specific EVs may lead to incomplete or biased control of phylogenetic effects. In contrast, using the union of EVs selected from both traits (denoted *EV_U*) can offer more stable and comprehensive adjustment. This insight also aligns with critiques of PGLS by Westoby et al. (2023), who emphasized that PGLS typically accounts for phylogenetic structure in the response variable but not in the predictor—potentially undermining model reliability when both traits exhibit phylogenetic signal.

This theoretical insight also inspired an alternative strategy for jointly controlling the phylogenetic non-independence of both traits: a two-step residual regression approach. In this framework, each trait is first regressed against its own EVs to remove trait-specific phylogenetic signals, and then the resulting residuals are used in a standard regression to assess the trait–trait relationship. In the following sections, we compare this residual-based method with models using *EV_U* and EVs derived independently from each trait across five evolutionary scenarios.

Performance of Eigenvector Selection Methods and Regression Approaches

In addition to evaluating different regression strategies, we compared the performance of two eigenvector selection methods—ESRRV and mMorI—to examine how the choice of EV selection influences regression outcomes. As shown in Figure 3, mMorI (solid lines) generally outperformed ESRRV (dashed lines), particularly in terms of regression accuracy, except in one case—when using EV_X_2 under the BM & BM+Norm scenario.

When comparing regression approaches under a fixed EV selection method, we found that models using *EV_U* consistently achieved the highest accuracy across five evolutionary scenarios when mMorI was used (red solid lines). With ESRRV, *EV_U* was still preferred in scenarios such as BM1 & BM1+BM2, BM & BM+Norm, and Shift, where it was nearly the top performer (red dashed lines), though not always the best. Notably, under conditions of minimal phylogenetic autocorrelation (Norm1 & Norm1+Norm2), the union method performed the worst among all tested approaches.

We also assessed a two-step residual regression strategy, where each trait was regressed on its own EVs before analyzing the correlation between residuals. However, this approach yielded lower accuracy overall and was therefore not selected for further analyses.

Based on these findings, we chose the mMorI method combined with *EV_U* as the representative PVR configuration in subsequent comparisons, which included integration with robust estimators and benchmarking against PIC and PGLS methods.

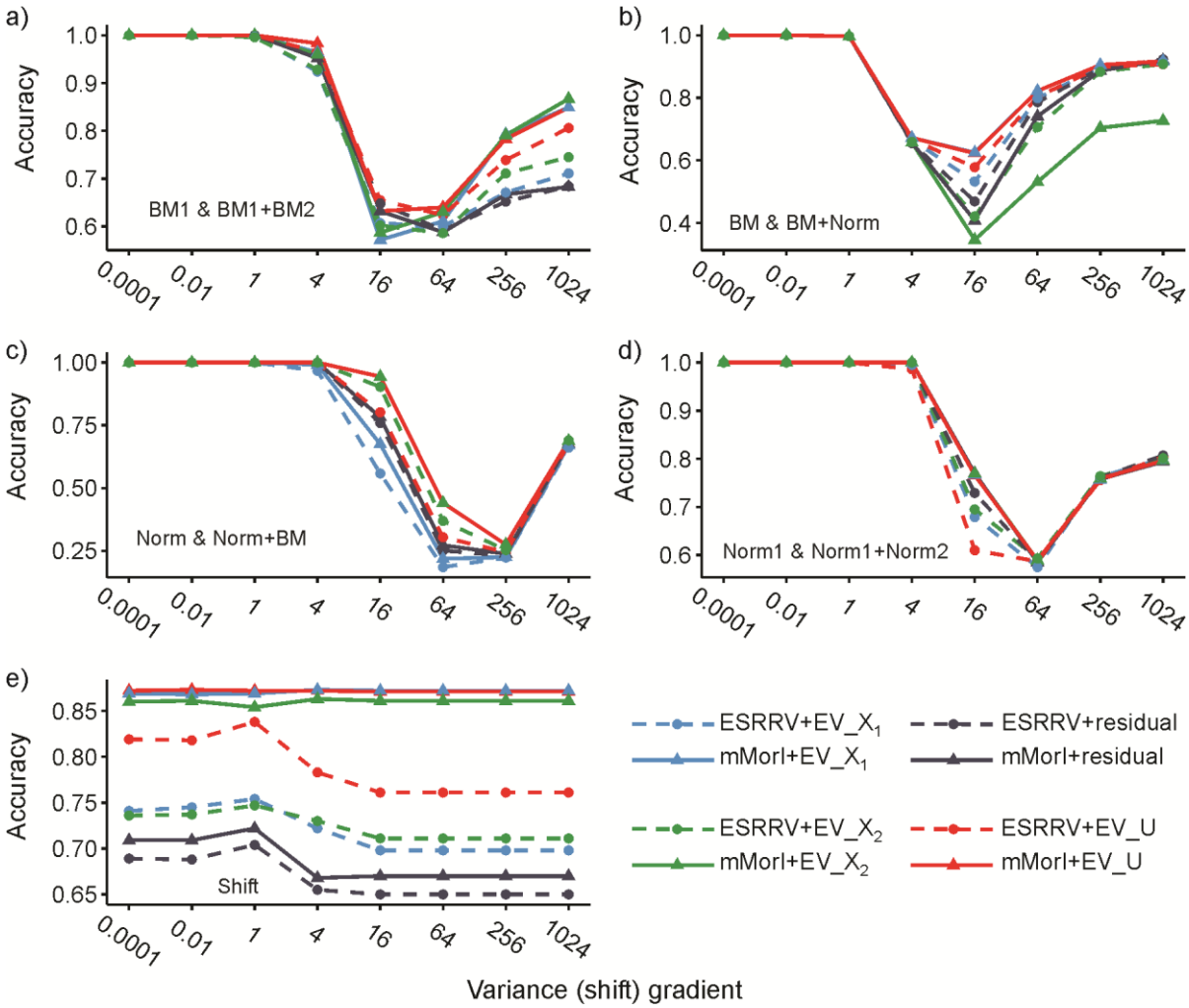

Figure 3. Performance across different phylogenetic eigenvector (EV) selection methods and regression approaches. This figure shows the performance of different EV selection methods (ESRRV and mMorI) combined with four regression strategies: (1) using EV_X_1 as covariates; (2) using EV_X_2 as covariates; (3) using the union set of EV_X_1 and EV_X_2, *EV_U*, as covariates; and (4) regressing the residuals of X_1 and X_2, each obtained from separate regressions against their respective EVs. Simulations were conducted on 128-species balanced trees ($\rho = 1$). The comparisons were performed across five simulation scenarios: a) BM1 & BM1+BM2; b) BM & BM+Norm; c) Norm & Norm+BM; d) Norm1 & Norm1+Norm2; and e) Shift. See MATERIALS AND METHODS for details regarding the EV

selection methods (ESRRV and mMorI) and the definitions of each simulation scenario. The data used to plot this figure are available in Supplementary Table S3.

Performance of Robust Estimators in PVR Models

We evaluated the performance of five regression estimators—L1, L2, M, MM, and S—within the optimized PVR framework (using mMorI and EV_U), focusing on the accuracy of regression outcomes across five evolutionary scenarios (Fig. 4). The S estimator generally exhibited the lowest accuracy, ranking last in three scenarios and performing suboptimally in the others, particularly under the Shift condition. In contrast, L1, L2, M, and MM produced broadly similar results, with each showing modest advantages under certain conditions. No single estimator consistently outperformed the others across all scenarios. Based on these findings, we excluded S from further analysis and retained the remaining four estimators for integration into the PVR framework in subsequent comparisons with PIC and PGLS.

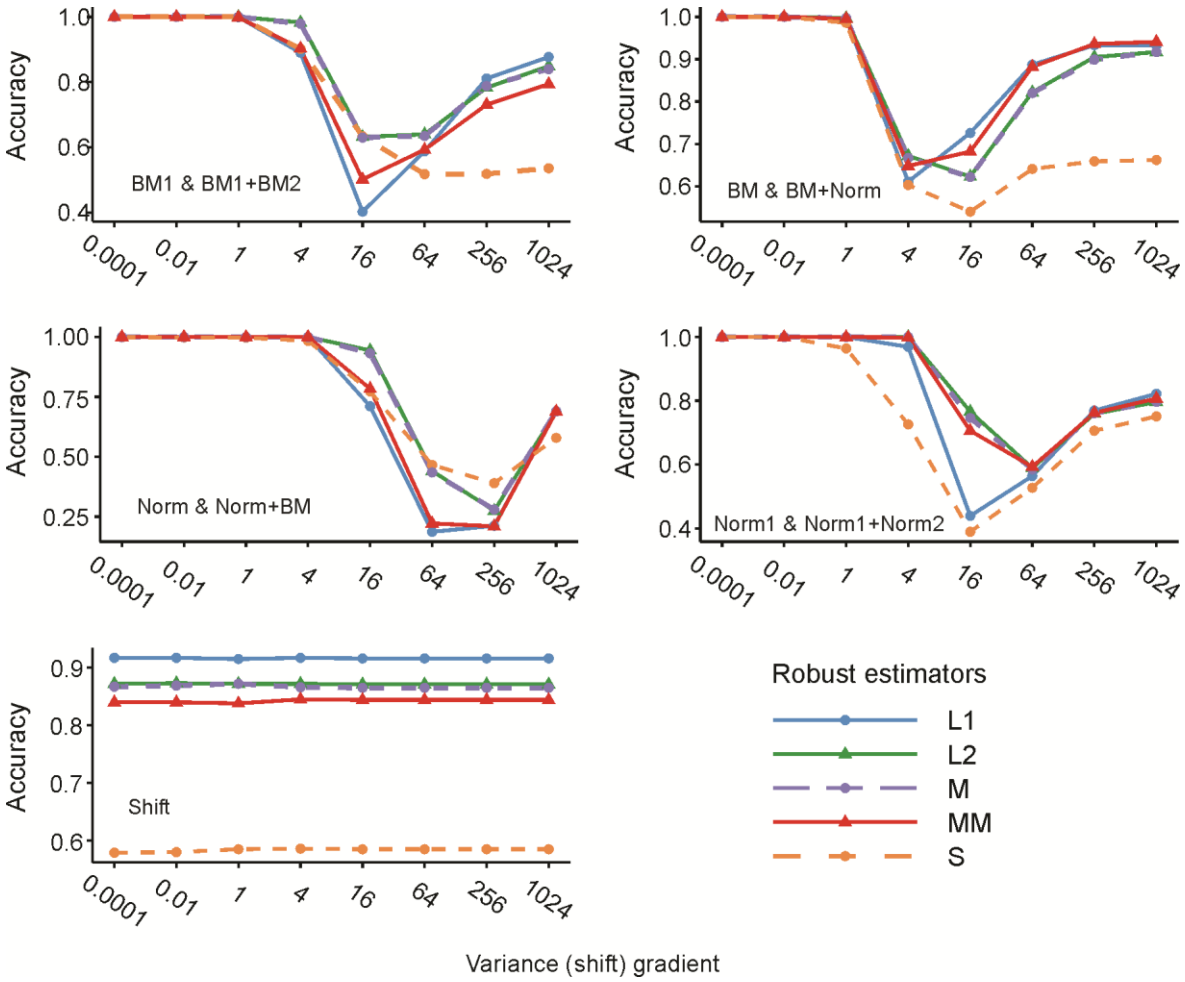

Figure 4. Comparison of robust estimator performance in phylogenetic eigenvector regression (PVR). This figure compares the performance of five robust estimators (L1, L2,

M, MM, and S) integrated into the optimized PVR framework (selected based on mMorI and *EV_U*). The five evolutionary scenarios—a) BM1 & BM1+BM2, b) BM & BM+Norm, c) Norm & Norm+BM, d) Norm1 & Norm1+Norm2, and e) Shift—represent different combinations of evolutionary models and relationships between X_1 and X_2. See MATERIALS AND METHODS for details on the phylogenetic EV selection method (mMorI), the definitions of robust estimators, the construction of *EV_U*, and the definitions of each simulation scenario. The data used to plot this figure are available in Supplementary Table S4.

Comparative Evaluation of Optimized PVR, PGLS, and PIC

Overall research setup.—Having integrated four estimators (L1, L2, M, and MM) into the PVR framework using mMorI and *EV_U*, we proceeded to compare these optimized PVR approaches with PGLS and PIC. For PGLS, the most appropriate model for each dataset was selected from five candidate evolutionary models—BM, Pagel's λ, OU fixed, OU random, and EB—via AIC minimization (hereafter referred to as PGLS-best). Following Adams et al. (2024), who demonstrated that PIC-MM outperformed other PIC variants (e.g., PIC-L1, PIC-M) and various robust-estimator PGLS approaches under non-stationary evolutionary scenarios, we included PIC-MM and PIC-L2 (representing standard PIC) in our analyses to compare their performance with optimized PVR across different evolutionary contexts.

To comprehensively assess the performance of the optimized PVR framework relative to PIC and PGLS, we conducted evaluations under five distinct evolutionary scenarios, applied consistently across all tree types. These scenarios were tested on multiple phylogenetic structures, beginning with 128-species balanced trees under three different branch-length scaling conditions: (1) uniformly scaled branches ($\rho = 1$), (2) compressed ancestral and elongated terminal branches ($\rho = 0.1$), and (3) elongated ancestral branches ($\rho = 2$). We then extended the analyses to two additional tree types representing key structural contrasts: a reduced sampling scenario using 16-species balanced trees ($\rho = 1$) and a topologically imbalanced scenario using 128-species ladder-shaped trees ($\rho = 1$).

128-species balanced trees with uniformly scaled branches ($\rho = 1$).—Across five evolutionary scenarios differing in the strength and structure of phylogenetic autocorrelation, method accuracy was plotted against increasing variance (shift) gradient values (Fig. 5). At low variance, all methods—including optimized PVR variants (solid lines) and benchmarks

PIC-L2, PIC-MM, and PGLS-best (dashed lines)—performed similarly. However, performance diverged significantly as variance (shift) increased.

Under BM1 & BM1+BM2, both X_1 and X_2 evolved under strong phylogenetic signal, with PGLS-best and PIC-L2 consistently outperforming all PVR variants (Fig. 5a). Under BM & BM+Norm, which features substantial differences in phylogenetic signal between traits, PGLS-best exhibited significant accuracy loss, while PVR-MM and PVR-L1 performed slightly better than the other methods (Fig. 5b). The opposite pattern emerged under Norm & Norm+BM, where PGLS-best and PIC-L2 again dominated, while PVR-MM and PVR-L1 performed most poorly (Fig. 5c). In Norm1 & Norm1+Norm2, all methods performed similarly, with the exception of PVR-L1, which exhibited a noticeable drop in accuracy at one variance level (Fig. 5d). Among the remaining methods, the three other PVR variants and PGLS-best demonstrated marginally higher accuracy than the PIC-based methods.

The most pronounced contrasts emerged in the Shift scenario, modeling non-stationary evolutionary change (Fig. 5e). Whereas PGLS-best and PIC-L2 suffered dramatic accuracy reductions as shift magnitude increased, PVR methods and PIC-MM maintained consistently high performance, demonstrating robustness against abrupt, branch-specific evolutionary changes. Surprisingly, PVR-L2, without explicit robust estimators, performed comparably to PIC-MM, suggesting inherent robustness conferred by phylogenetic eigenvectors themselves. Among PVR variants, while L2 and M estimators showed slightly lower accuracy, L1 and MM—together with PIC-MM—consistently maintained higher accuracy across different shift magnitudes.

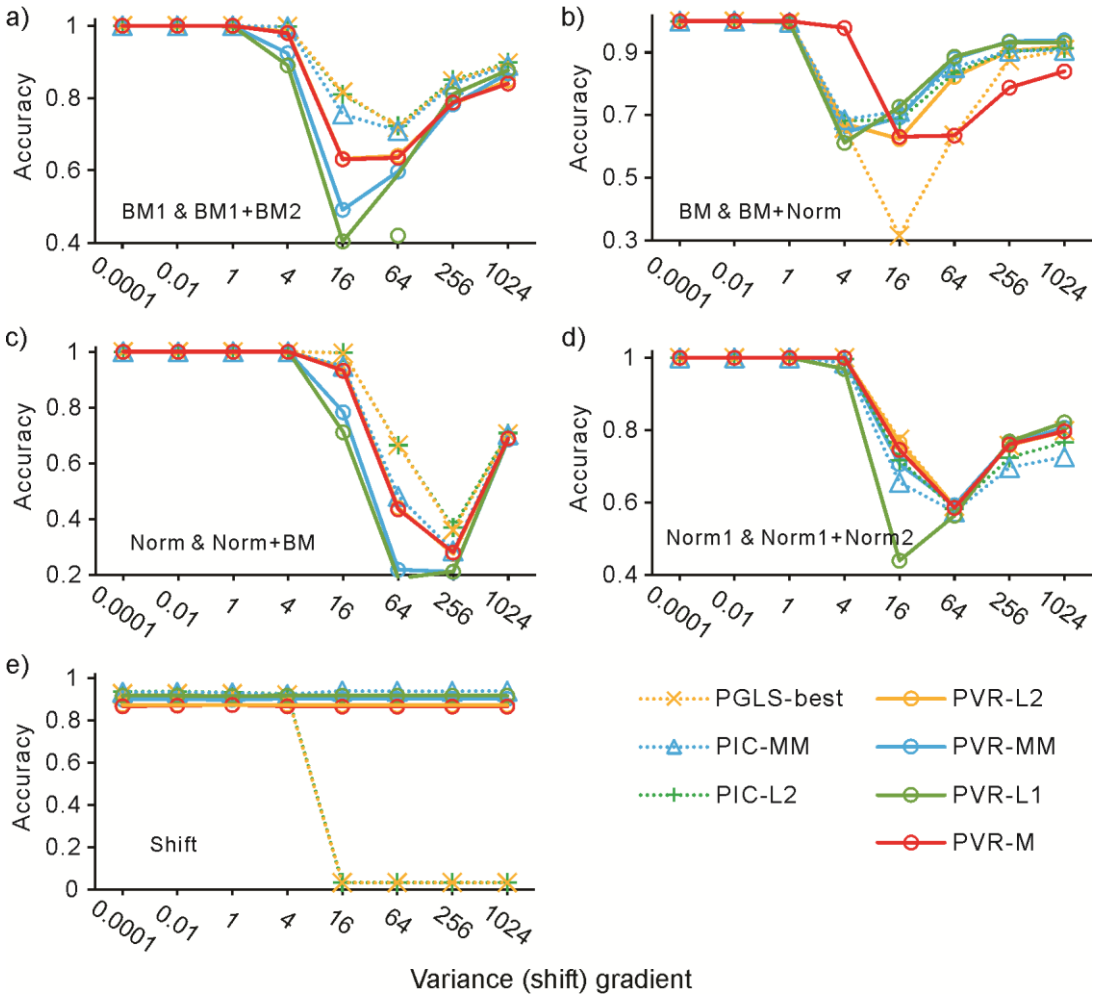

Figure 5. Comparative accuracy of PVR, PGLS, and PIC methods on 128-species balanced phylogenetic trees (ρ = 1). The comparisons are performed under five evolutionary scenarios: a) BM1 & BM1+BM2, both traits exhibit strong phylogenetic signal; b) BM & BM+Norm and c) Norm & Norm+BM, large disparity in phylogenetic signal strength between traits; d) Norm1 & Norm1+Norm2, both traits lack phylogenetic structure; e) Shift, a synchronous trait shift was introduced on a single internal branch, while the two traits evolved independently. See MATERIALS AND METHODS for details on the definitions of each simulation scenario. The data used to plot this figure are available in Supplementary Table S5.

128-species balanced trees with compressed ancestral branches (ρ = 0.1).—With compressed ancestral branches, methods exhibited closely aligned performance across stationary scenarios, except for PVR-L1 (Fig. 6a-d). Under BM & BM+Norm, PVR-L1 was distinctly superior; however, it underperformed in other stationary scenarios (BM1 & BM1+BM2, Norm & Norm+BM, and Norm1 & Norm1+Norm2). PVR-L2, PVR-M, and

PVR-MM consistently matched PIC-L2, PIC-MM, and PGLS-best. In the Shift scenario, PGLS-best and PIC-L2 again showed accuracy declines, while PIC-MM and all PVR methods remained robust (Fig. 6e).

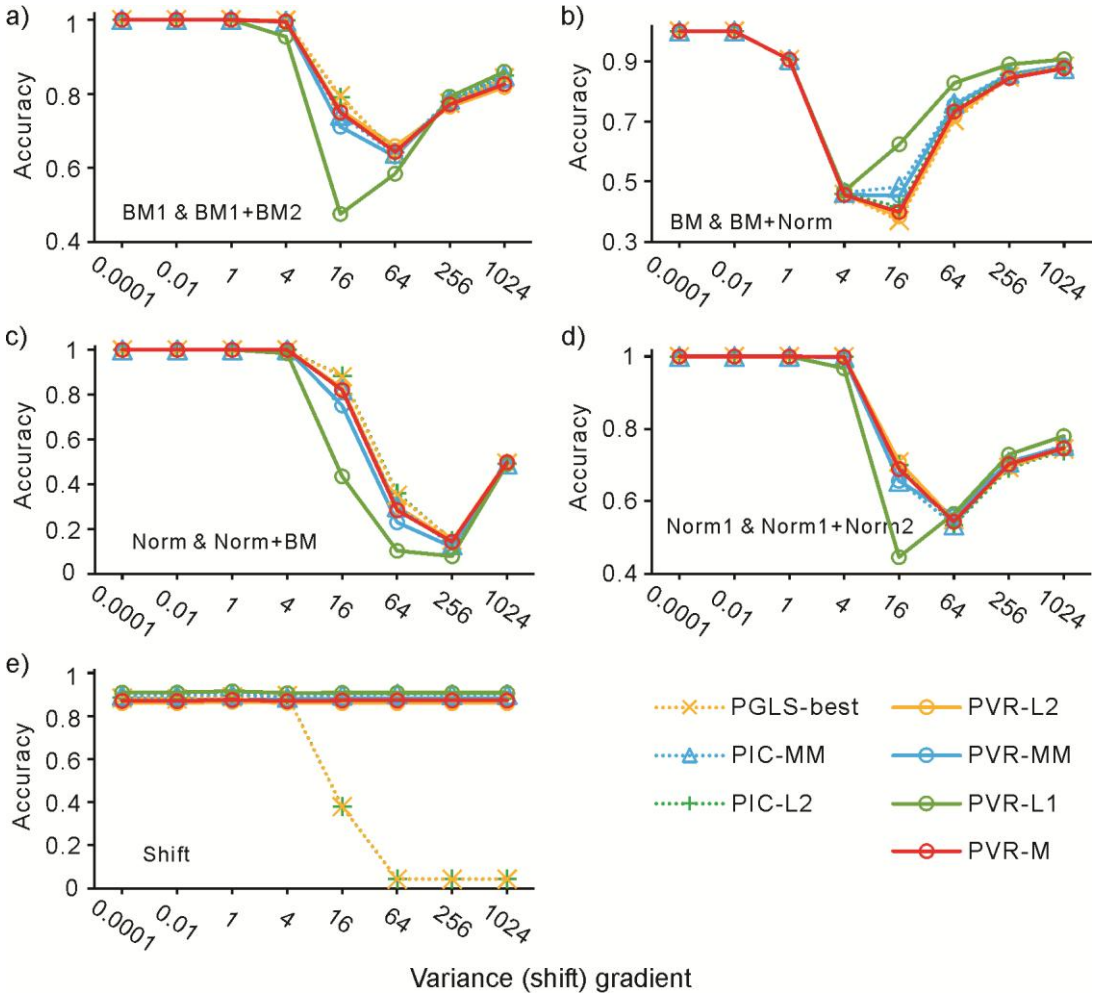

Figure 6. Performance of PVR, PGLS, and PIC on 128-species balanced phylogenetic trees ($\rho = 0.1$). The comparisons are performed under five evolutionary scenarios: a) BM1 & BM1+BM2; b) BM & BM+Norm; c) Norm & Norm+BM; d) Norm1 & Norm1+Norm2; e) Shift. See MATERIALS AND METHODS for details on the definitions of each simulation scenario. The data used to plot this figure are available in Supplementary Table S6.

128-species balanced trees with elongated ancestral branches ($\rho = 2$).—When deep ancestral branches were elongated, accuracy patterns diverged notably from the other branch-length scenarios (Fig. 7). In BM1 & BM1+BM2, PGLS-best, PIC-L2, and PIC-MM substantially outperformed PVR methods. Under BM & BM+Norm, PGLS-best accuracy declined dramatically at intermediate variance, whereas other methods remained stable.

Interestingly, in Norm & Norm+BM with elongated ancestral branches (ρ = 2), all methods achieved perfect (100%) accuracy (Fig. 7c). Analysis of raw simulation data indicated strongly positive correlations between traits (Supplementary Table S7). With elongated ancestral branches and extremely short terminal branches, BM—which scales proportionally with branch length—contributed minimally to trait differences along the short terminal branches. In contrast, the independent normal (Norm) component, being independent of branch length, dominated these short branches. Consequently, the minimal terminal BM contributions made X_2 nearly collinear with X_1, effectively eliminating methodological performance differences.

In the Shift scenario, PGLS-best and PIC-L2 experienced marked accuracy reductions at high shift magnitudes (Fig. 7c). PIC-MM and PVR methods remained relatively robust. PIC-MM generally maintained a slight advantage at this larger ρ values (Fig.7), whereas PVR-L1 performed slightly better at smaller ρ (see Figure 6 and Supplementary Table S6).

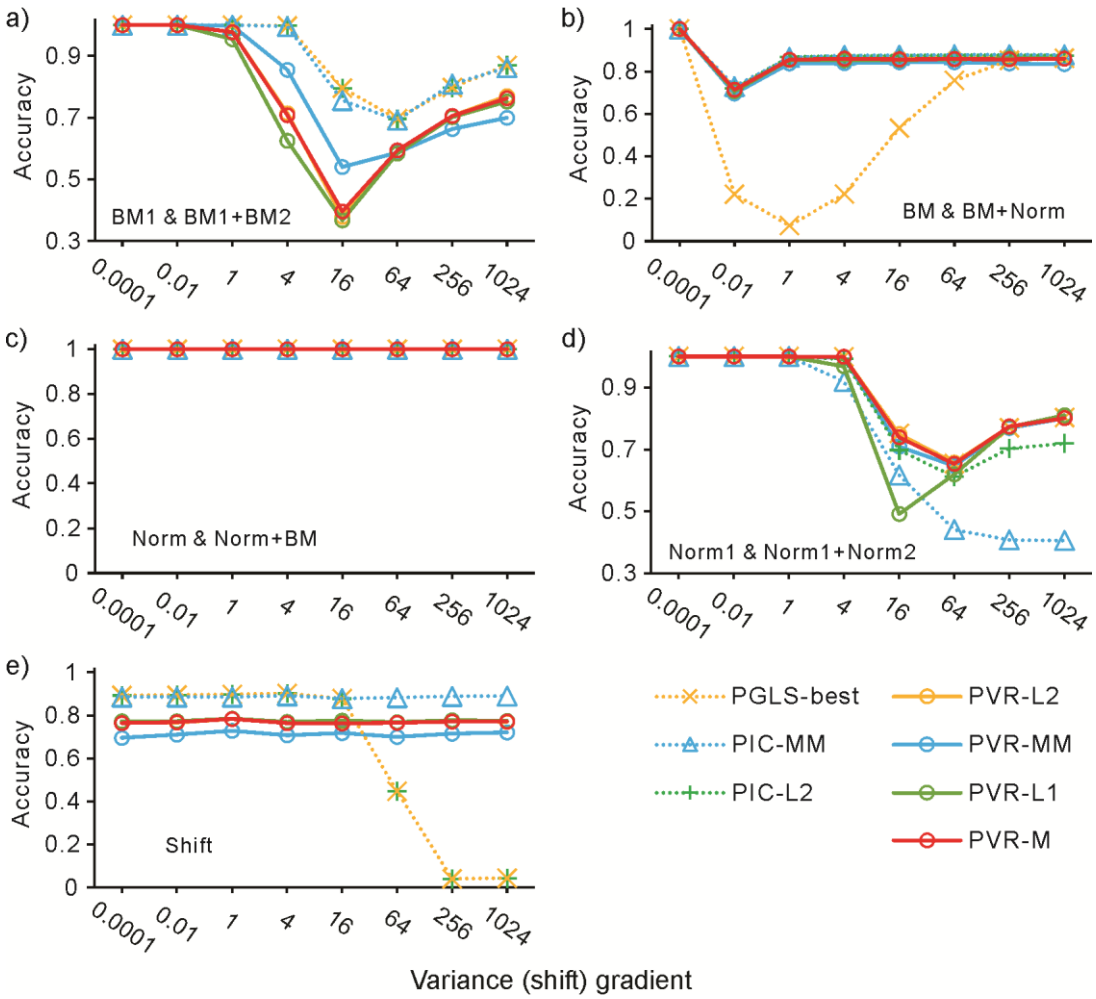

Figure 7. Performance of PVR, PGLS, and PIC on 128-species balanced phylogenetic trees (ρ = 2). The comparisons are performed under five evolutionary scenarios: a) BM1 &

BM1+BM2; b) BM & BM+Norm; c) Norm & Norm+BM; d) Norm1 & Norm1+Norm2; e) Shift. See MATERIALS AND METHODS for details on the definitions of each simulation scenario. The data used to plot this figure are available in Supplementary Table S8.

16-species balanced trees with uniformly scaled branches.— The overall performance patterns under reduced taxonomic sampling were broadly consistent with those observed under 128-species balanced trees (Fig. 8). Although minor differences existed in the exact performance ranks of certain methods, the scenario-dependent relative advantages among methods remained unchanged.

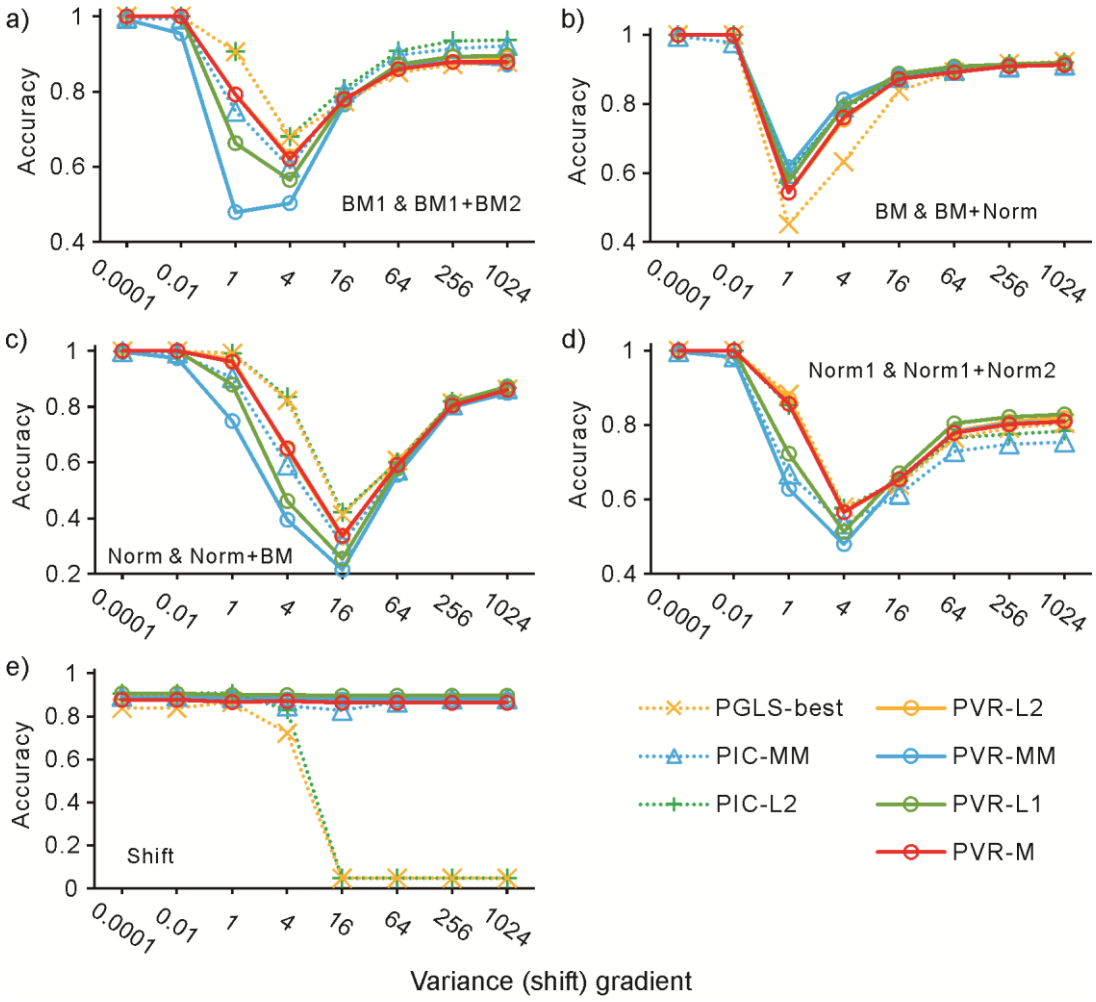

Figure 8. Performance of PVR, PGLS, and PIC on 16-species balanced trees with uniformly scaled branches (ρ = 1). The comparisons are performed under five evolutionary scenarios: a) BM1 & BM1+BM2; b) BM & BM+Norm; c) Norm & Norm+BM; d) Norm1 & Norm1+Norm2; e) Shift. See MATERIALS AND METHODS for details on the definitions of each simulation scenario. The data used to plot this figure are available in Supplementary Table S9.

128-species Ladder-Shaped Trees (ρ = 1)—During simulation analyses using ladder-shaped phylogenies, we observed that the mMorI method frequently failed to select a sufficient number of eigenvectors, often returning none at all. Quantitatively, the median number of eigenvectors in the union set (*EV_U*) was only 1. In contrast, the mAIC method produced excessively large EV sets, with a median *EV_U* count of 80—raising concerns about model overfitting. The ESRRV method yielded a more moderate and balanced result (median *EV_U* = 31), effectively supporting phylogenetic correction. Accordingly, ESRRV was adopted for all subsequent analyses involving ladder-shaped trees.

Given these patterns, we further examined the underlying reasons why mMorI underperformed in ladder-shaped trees. Cumulative eigenvalue distributions differed notably between balanced and ladder-shaped trees (Fig. 9). The extreme topological imbalance of ladder-shaped trees distorts the distribution of phylogenetic signal, resulting in smaller disparities among eigenvalues. This in turn hinders the identification of informative eigenvectors—likely explaining why mMorI failed to select sufficient EVs under these conditions. These findings highlight the importance of tailoring eigenvector selection strategies to tree topology and signal structure.

Figure 9. Contribution of phylogenetic eigenvectors across different tree structures. Both the balanced and ladder-shaped trees shown here are 128-species phylogenies with uniformly scaled branches (ρ = 1). The data used to plot this figure are available in Supplementary Table S10.

We next evaluated the comparative performance of PVR, PGLS, and PIC methods under ladder-shaped tree conditions (Fig. 10). While overall trends resembled those observed under balanced trees, important distinctions emerged in the Shift scenario. PVR-L1, PVR-L2, and

PVR-M showed notable accuracy declines at higher shift magnitudes—similar to PGLS-best and PIC-L2—whereas PIC-MM and PVR-MM maintained robustness, reaffirming their superior performance under non-stationary conditions.

Collectively, these results illustrate that method performance varies markedly depending on tree structure and evolutionary scenario. No single method consistently outperformed the others across all conditions, highlighting the need for context-specific model selection in phylogenetic comparative analysis.

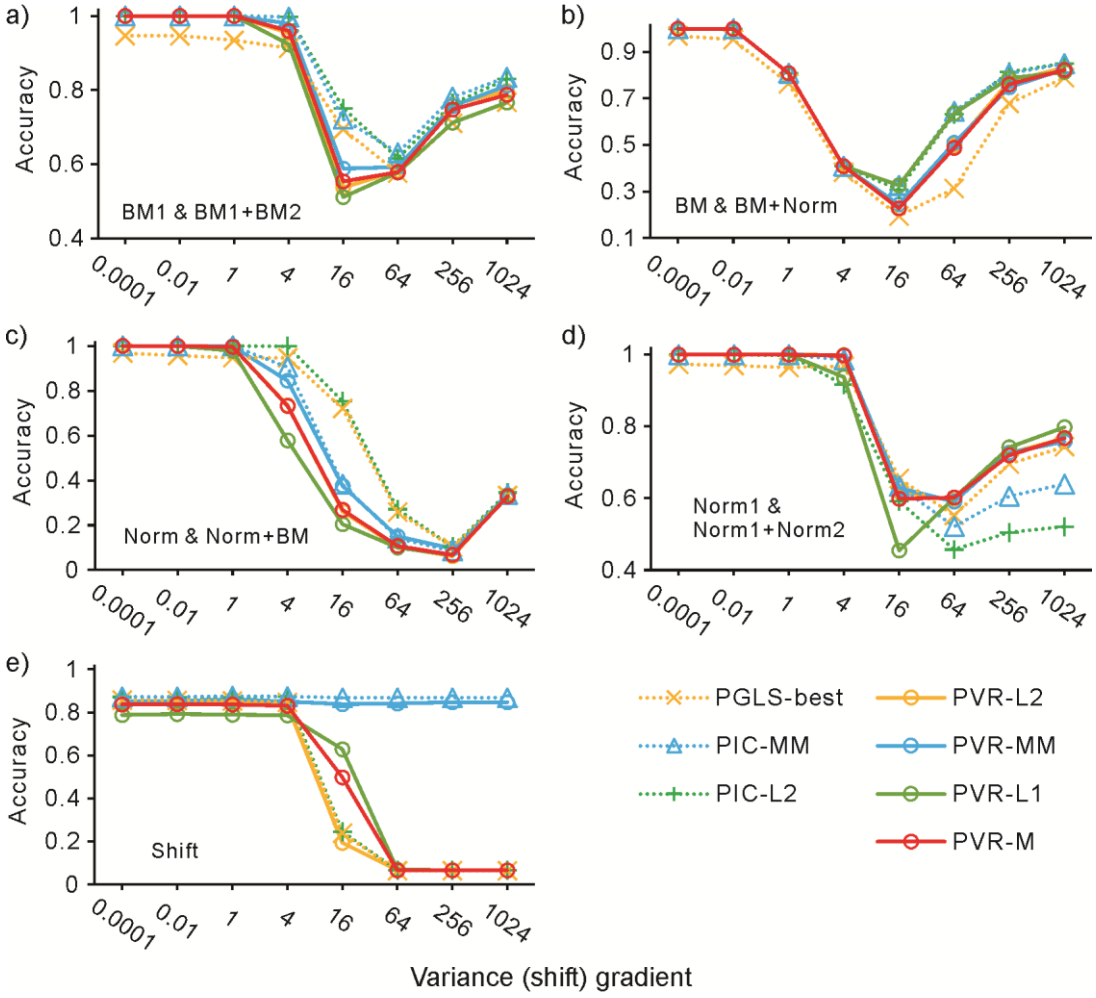

Figure 10. Performance of PVR, PGLS, and PIC on on 128-species ladder-shaped phylogenetic trees (ρ = 1). The comparisons are performed under five evolutionary scenarios: a) BM1 & BM1+BM2; b) BM & BM+Norm; c) Norm & Norm+BM; d) Norm1 & Norm1+Norm2; e) Shift. This figure differs from Figures 5-8 in that all PVR methods use ESRRV for eigenvector selection, rather than mMorI. See MATERIALS AND METHODS for details on the definitions of each simulation scenario and the two EV selection methods. The data used to plot this figure are available in Supplementary Table S11.

DISCUSSION

PVR was originally proposed as a flexible alternative to model trait–trait correlations while accounting for phylogenetic structure through eigenvectors derived from phylogenetic distance matrices (Diniz-Filho et al. 1998). However, this conventional view (Laurin 2010, Freckleton et al. 2011, Pennell and Harmon 2013, Hansen 2014) holds that PVR is less reliable in controlling for phylogenetic non-independence and produces less robust estimates than likelihood-based approaches, particularly PGLS.

In this study, we revisited the PVR framework and introduced several key optimizations. First, we evaluated how the selection of phylogenetic eigenvectors varies depending on which trait is treated as the dependent variable. Our analysis revealed that the sets of eigenvectors selected based on X_1 and X_2 often differ substantially (Table 1), and when these trait-specific eigenvectors are used as covariates in regression models, they can lead to markedly different correlation estimates (Table 2 and Fig. 2). Both theoretical derivation and simulation-based validation demonstrated that using the union of eigenvectors selected from both traits (*EV_U*) better control for phylogenetic independence than using those selected from a single trait (Appendix 1 and Fig.3).

Building on this foundation, we further integrated robust estimators into the PVR framework to improve its resilience to noise and data irregularities. The resulting variants—PVR-L1, PVR-L2, PVR-M, and PVR-MM—were then systematically compared against established methods including PGLS-best, PIC-L2 (standard PIC), and PIC-MM across a wide range of evolutionary scenarios.

Our findings challenge the conventional view that PGLS universally outperforms PVR. While PGLS-best and PIC-L2 excelled when dependent traits exhibited strong phylogenetic signal, optimized PVR methods—particularly PVR-MM—demonstrated superior robustness in scenarios characterized by weak or asymmetric phylogenetic signal and non-stationary evolutionary shifts. These results underscore the context-dependent nature of method performance and suggest that researchers should move away from universally applying PGLS, instead tailoring model choice to the evolutionary properties of their data. The optimized PVR framework, especially when coupled with robust estimators, offers a valuable and versatile analytical alternative for phylogenetic comparative studies.

PVR Robustness Under Non-Stationary Evolutionary Shift

One of the most challenging conditions for phylogenetic models is non-stationary evolutionary change. Under such shift scenarios, our results showed that the optimized PVR framework—especially PVR-MM and even PVR-L2—maintained high and stable accuracy, closely matching PIC-MM across most tree structures and shift magnitudes. In contrast, PGLS-best and PIC-L2 suffered pronounced accuracy declines under increasing shift intensity.

Remarkably, PVR-L2, which does not incorporate any explicit robust estimator, performed comparably to PIC-MM—one of the most robust models available (Adams et al. 2024)— under balanced tree conditions. This suggests that the use of phylogenetic eigenvectors may confer inherent robustness, likely due to their capacity to absorb structured variation across the tree. Among PVR variants, L1 and MM estimators provided additional but modest gains, especially under large shifts and in most tree conditions. However, under ladder-shaped phylogenies, only PVR-MM and PIC-MM retained robustness, highlighting their stability in particularly asymmetric topologies.

Theoretically, this robustness stems from PVR's model-agnostic design: it does not rely on global stationarity assumptions but instead decomposes phylogenetic structure via orthogonal eigenvectors, with some exhibiting greater sensitivity to variation in specific regions of the tree. This multiscale sensitivity allows PVR to flexibly accommodate localized shifts in evolutionary rates (Diniz-Filho et al. 2012b, Gouveia et al. 2013, Diniz-Filho et al. 2014, Diniz-Filho et al. 2015, Diniz-Filho et al. 2019).

Together, these empirical and theoretical insights reinforce the view that PVR is not only a competitive alternative to conventional phylogenetic comparative methods under stationary models, but also a particularly robust tool for modeling evolutionary nonstationarity—especially when paired with robust estimators such as MM.

PVR Advantage Under Weak Phylogenetic Signal

In addition to its robustness under non-stationary evolutionary shifts, the PVR framework also exhibited a notable advantage in scenarios characterized by weak phylogenetic signal. In scenarios with weak phylogenetic signal, such as Norm1 & Norm1+Norm2, PVR methods, particularly L2 and M, showed a clear advantage.

We found that the phylogenetic signal (λ) in X_1 and X_2 under this scenario was significantly greater than zero in a proportion ranging from 0.5% to 6% (depending on the structures of phylogenetic trees), with the maximum value reaching 1.13. While such signals

are generally weak, their presence indicates that assessing the performance of phylogenetic methods in these contexts is not without value. Note that although λ is typically bounded between 0 and 1, values slightly above 1 can occur due to maximum likelihood estimation and are not uncommon in empirical datasets.

When the phylogenetic signal is weak, the coefficients of eigenvectors as covariates tend to be small. Since these coefficients are estimated freely based on the data, without constraint to a specific range, they provide modeling flexibility even when phylogenetic effects are minimal. Therefore, PVR methods may be more effective than PIC and PGLS in handling cases with weak phylogenetic signals.

Comparison with Two-Step Residual Regression Approaches

In addition to comparing EV sets derived from individual traits and their union (*EV_U*), we also evaluated a two-step residual regression approach as an alternative means of controlling phylogenetic non-independence. In this method, each trait is first regressed against its own set of EVs to remove trait-specific phylogenetic signals. The resulting residuals are then used in a standard regression to assess the relationship between traits. This residual-based strategy offers a symmetric structure. However, our results showed that this method consistently underperformed relative to the *EV_U*-based regression approach across all evolutionary scenarios (Fig. 4). While the residual method did not exhibit clear advantages over trait-specific EV regressions, it showed notably lower accuracy in certain scenarios—particularly under shift conditions.

Theoretically, this outcome is not unexpected. When two traits share overlapping phylogenetic structure—as often occurs due to shared ancestry—independently removing phylogenetic signal from each trait can result in the unintended removal of shared components that may underlie genuine biological associations. As a result, residual-based approaches may reduce correlation accuracy or introduce bias under certain conditions.

In contrast, the *EV_U* approach directly incorporates the phylogenetic information from both traits within a single model, without presuming a causal direction or requiring sequential correction. This unified strategy helps preserve meaningful shared signal while effectively controlling for phylogenetic effects. Our simulation results support this reasoning: *EV_U* consistently yielded higher accuracy across diverse evolutionary scenarios, reinforcing its utility as the preferred method in our optimized PVR framework.

Flexibility of Eigenvector-Based Approaches in Phylogenetic Modeling

Unlike PGLS and other model-based approaches that rely on predefined evolutionary models (Freckleton et al. 2011, O'Meara and Beaulieu 2016), PVR extracts phylogenetic structure through eigenvector decomposition, offering greater modeling flexibility. This feature is particularly valuable when the true evolutionary process deviates from canonical models (e.g., BM, OU or EB), or when model uncertainty is high (Diniz-Filho et al. 2014). Given the limited repertoire of evolutionary models currently available, it is unlikely that all real-world evolutionary patterns can be adequately captured by likelihood-based methods. Even if model libraries were expanded in the future, distinguishing among them would require repeated model fitting and comparison, rendering the process both cumbersome and prone to error.

PVR's modular use of phylogenetic eigenvectors offers a key advantage: unlike PGLS, where phylogenetic information is embedded in model-specific assumptions, PVR expresses it through separable components, enabling seamless integration with machine learning algorithms such as random forests and neural networks (Penone et al. 2014, Debastiani et al. 2021). This portability enables applications beyond traditional regression, including trait imputation and classification in a phylogenetic context.

Third, PVR offers a structural advantage in handling directional asymmetry in trait relationships. Unlike correlation analysis, which treats both variables symmetrically, OLS requires a clear distinction between the independent variable (X) and the dependent variable (Y). This distinction is built into the model's structure: X is assumed to be fixed and measured without error, while all residual variation is attributed to Y (McArdle 1988, Osborne and Waters 2002, Casson and Farmer 2014, Jarantow et al. 2023). Moreover, OLS implicitly encodes a causal direction—predicting changes in Y as a function of X. As a result, OLS is inherently asymmetrical: reversing X and Y leads to a different model, different estimates, and potentially different conclusions. PGLS inherited the asymmetry from OLS regression, requiring a clear assignment of dependent and independent variables. However, when the two traits differ in the strength of their phylogenetic signal, assigning the trait with weaker signal as the dependent variable can distort the estimated relationship between them (Chen et al. 2023). In situations where the trait that is either logically considered causal or measured with less error happens to exhibit a stronger phylogenetic signal, researchers face a dilemma: treating it as the independent variable aligns with

biological reasoning or data reliability, but doing so may compromise phylogenetic model performance in PGLS.

PVR, by contrast, avoids this conflict. By including the union of eigenvectors selected from both traits as covariates, PVR allows either trait to be designated as the independent variable without affecting the accuracy of the model. This design enables researchers to select predictors based on causal logic or measurement quality, without being constrained by the phylogenetic properties of the traits—removing a key asymmetry-related limitation present in OLS-derived methods like PGLS.

Dependence of PGLS-best Performance on Dependent Variable Selection

From Figures 5-8 and Figure 10, we observed a consistent pattern in the relative performance of PGLS-best compared to other methods across the first four evolutionary scenarios. In BM1 & BM1+BM2, PGLS-best generally performed the best or was close to the best; in BM & BM+Norm, it was the worst or close to the worst; in Norm & Norm+BM, it reversed again to be the best or close to the best; and in Norm1 & Norm1+Norm2, its performance was average. In BM1 & BM1+BM2 and Norm & Norm+BM, X_2 exhibited stronger phylogenetic signal than X_1, while in BM & BM+Norm, X_2 had a weaker phylogenetic signal compared to X_1. In Norm1 & Norm1+Norm2, there was little to no difference in the phylogenetic signals of the two traits. In this study, X_2 was treated as the dependent variable and X_1 as the independent variable. It seems that the performance of PGLS-best depends on whether the trait with the stronger phylogenetic signal is chosen as the dependent variable, which is the focus of our unpublished manuscript (Chen et al. 2023). As the study has faced multiple rejections in various journals and still requires further refinement, it should not be considered a mature benchmark method for evaluating PVR performance. However, it can be speculated that, if such optimization were performed, the relative advantage of PGLS-best might become even more pronounced.

Broader Implications: Variable Asymmetry Beyond Phylogenetic Models

While this study focuses primarily on phylogenetic comparative methods, similar challenges may arise in other contexts involving structured data—particularly those with spatial autocorrelation. Spatial autocorrelation refers to the tendency for geographically proximate observations to exhibit greater similarity than those farther apart. It is common in ecological and evolutionary datasets and requires statistical models that account for spatial structure to

avoid biased estimates (Doreian 1980, Dow et al. 1982, Legendre 1993, Keitt et al. 2002, He et al. 2003, Cottenie 2005, Diniz-Filho and Bini 2005, Dormann et al. 2007, Dale and Fortin 2009, Peres-Neto and Legendre 2010).

Preliminary analyses suggest that the asymmetry introduced by variable designation—previously discussed in the context of PGLS and PVR—also occurs in spatial models. In methods such as spatial eigenvector mapping, conditional autoregressive (CAR) models, simultaneous autoregressive (SAR) models, and trend-surface generalized additive models (GAMs), we observed outcome discrepancies when the dependent and independent variables were swapped. These findings, summarized in Supplementary Table S12, indicate that variable swapping may affect model performance across a wide range of autocorrelation-adjusted frameworks. A full analysis of these effects is beyond the scope of this study, but future work will explore their implications more systematically, with potential relevance to both spatial and ecological inference.

In many real-world applications, multiple forms of autocorrelation—such as spatial and phylogenetic structure—must be addressed simultaneously. This introduces additional complexity in both variable designation and eigenvector selection. As discussed by Safi and Pettorelli (2010), properly balancing spatial, phylogenetic, and environmental components depends on the specific analytical goals. Although our current study emphasizes phylogenetic relationships, their work underscores the importance of aligning methodological choices with the biological structure of the data. Future extensions of PVR and related models may benefit from more flexible integration of multiple autocorrelation sources to improve model reliability across complex ecological scenarios.

DATA AVAILABILITY

Supplementary Tables S1-S12 are included in Supplementary_tables.xlsx. The raw data for the five 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 ESRRV.R) are available separately. All these files and .R scripts can be found in the Dryad data repository: http://datadryad.org/stash/share/dLL0-zHGm_Wyrli08AJ9P4pR-yFycPeYOqktVZhEUVA.

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 supported in part by the National Natural Science Foundation of China (grant number 31671321), with leftover resources used for occasional research expenses. 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.

Blomberg S.P., Lefevre J.G., Wells J.A., Waterhouse M. 2012. Independent contrasts and PGLS regression estimators are equivalent. Syst. Biol., 61:382-391.

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:plv030.

Casson R.J., Farmer L.D. 2014. Understanding and checking the assumptions of linear regression: a primer for medical researchers. Clin. Exp. Ophthalmol., 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.

Debastiani V.J., Bastazini V.A.G., Pillar V.D. 2021. Using phylogenetic information to impute missing functional trait values in ecological databases. Ecol. Inform., 63:101315.

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

Dewar A.E., Belcher L.J., West S.A. 2025. A phylogenetic approach to comparative genomics. Nat. Rev. Genet., 10.1038/s41576-024-00803-0.

Diniz-Filho J.A.F., Alves D.M.C.C., Villalobos F., Sakamoto M., Brusatte S.L., Bini L.M. 2015. Phylogenetic eigenvectors and nonstationarity in the evolution of theropod dinosaur skulls. J. Evol. Biol., 28:1410-1416.

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. 2012a. On the selection of phylogenetic eigenvectors for ecological analyses. Ecography 35:239-249.

Diniz-Filho J.A.F., Bini M., Sakamoto M., Brusatte S. 2014. Phylogenetic eigenvector regression in paleobiology. Rev. Bras. Paleontol., 17:107-122.

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., Rangel T.F., Santos T., Mauricio Bini L. 2012b. Exploring patterns of interspecific variation in quantitative traits using sequential phylogenetic eigenvector regressions. Evolution, 66:1079-1090.

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. Amer. Natur., 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. Ann. Statist., 30:583-616.

Gouveia S.F., Hortal J., Cassemiro F.A.S., Rangel T.F., Diniz-Filho J.A.F. 2013. Nonstationary effects of productivity, seasonality, and historical climate changes on global amphibian diversity. Ecography, 36:104-113.

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.

Hansen T.F. 2014. Use and misuse of comparative methods in the study of adaptation. In: Garamszegi LZ editor. Modern Phylogenetic Comparative Methods and Their Application in Evolutionary Biology: Concepts and Practice. Berlin, Heidelberg, Springer Berlin Heidelberg, p. 351-379.

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.

Jarantow S.W., Pisors E.D., Chiu M.L. 2023. Introduction to the use of linear and nonlinear regression analysis in quantitative biological assays. Curr. Protoc., 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.

Landis J.R., Koch G.G. 1977. The measurement of observer agreement for categorical data. Biometrics, 33:159-174.

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.

Laurin M. 2010. Assessment of the relative merits of a few methods to detect evolutionary trends. Syst. Biol., 59:689-704.

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. Can. J. Zool., 66:2329-2339.

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. Pract. Assess. Res. Eval., 8:2.

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.

Pennell M.W., Harmon L.J. 2013. An integrative view of phylogenetic comparative methods: connections to population genetics, community ecology, and paleobiology. Ann NY Acad Sci, 1289:90-105.

Penone C., Davidson A.D., Shoemaker K.T., Di Marco M., Rondinini C., Brooks T.M., Young B.E., Graham C.H., Costa G.C. 2014. Imputation of missing data in life-history trait datasets: which approach performs the best? Methods Ecol. Evol., 5:961-970.

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.

Slater G.J., Pennell M.W. 2014. Robust regression and posterior predictive simulation increase power to detect early bursts of trait evolution. Syst. Biol., 63:293-308.

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

Westoby M., Yates L., Holland B., Halliwell B. 2023. Phylogenetically conservative trait correlation: Quantification and interpretation. J Ecol, 111:2105-2117.

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