– SUPPLEMENTARY MATERIAL –

In 1), we present the results of the simulations for traits with low ($\rho = 0.2$) average correlations and with or without phylogenetic structure in the predictor variable (Fig. S1 and S2 respectively), and simulations with higher correlations between the traits ($\rho = 0.8$) with and without phylogenetic structure in the predictor variable (Fig. S3 and S4 respectively). In 2), we display the three first phylogenetic principal components of the phyllostomid bats dataset under the various feeding modes groupings described in the main text (Figs. S5-6). Phylogenetic PCs were obtained using penalized-likelihood with the *mvgls.pca* function in mvMORPH (PL estimate of Pagel's $\lambda = 0.45$; see Clavel et al. 2019). In 3), we provide additional results for univariate analyses conducted separately on the three first pPC axes of the bats mandible dataset. In 4), we present the general linear hypothesis testing using contrasts coding and test specific hypotheses about the dietary grouping schemes on the phyllostomid bats mandible dataset. In 5) we briefly explain how the datasets with main differences along specific axes of the multidimensional space were simulated, and finally in 6) we provide a step by step derivation of the first and second order derivative of the log-likelihood presented in Appendix 2.

1) Simulations results for type I and II errors (statistical power and error) on datasets simulated with low ($\boldsymbol{\rho = 0.2}$) and high ($\boldsymbol{\rho = 0.8}$) average correlations.

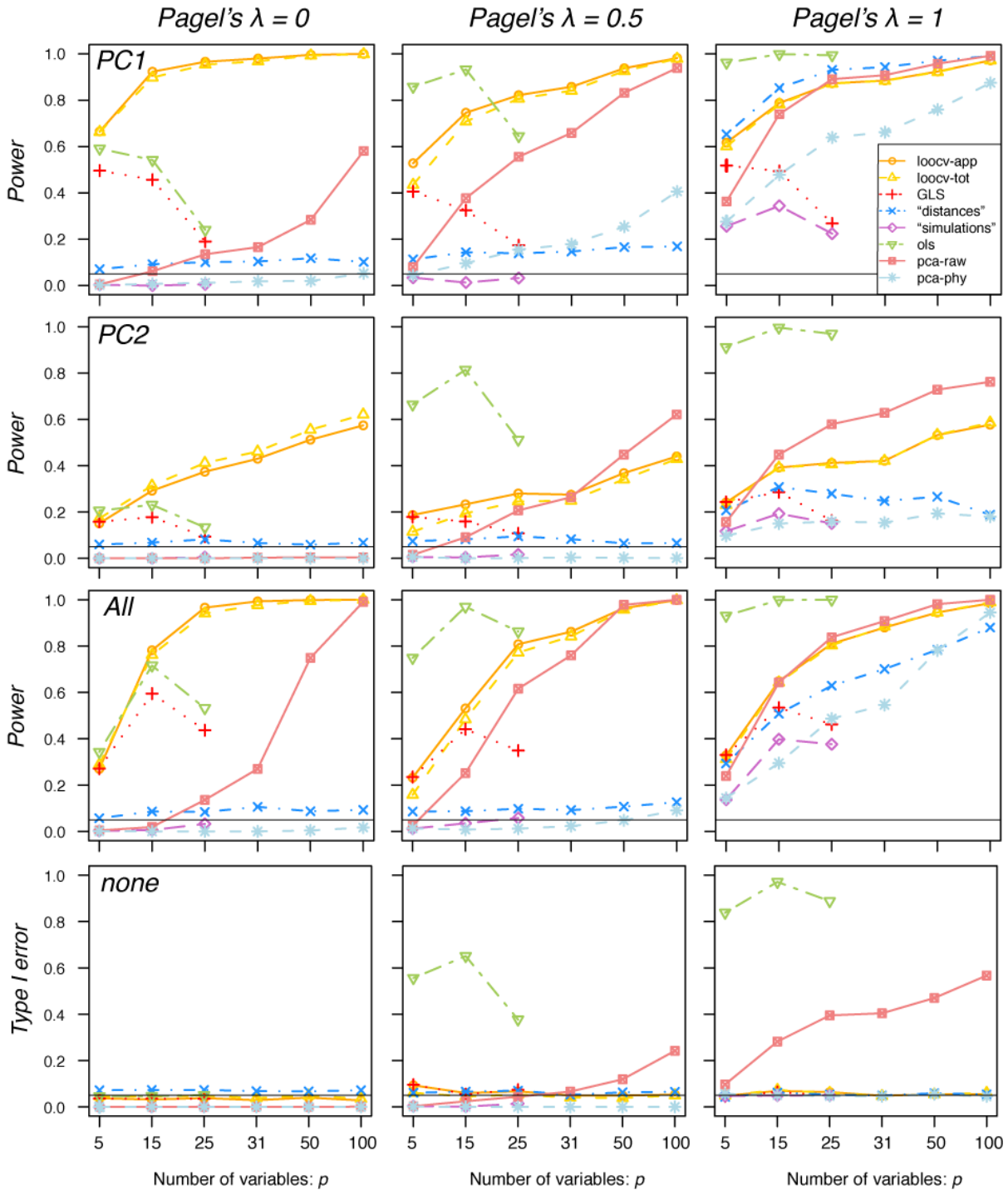

Figure S1. *Comparison of the statistical performances (statistical power and error) for the various MANOVA approaches with a phylogenetically clustered binary predictor variable and an average correlation between the dependent variables of* ***0.2****. In the top row the differences between groups were simulated along the first axis of variance (PC1), along the second axis (PC2) on the second row, and across all the axes (All) on the third row (see the main text). There were no differences simulated between groups (none) on the last row (for type I error testing). We can note that the regularized phylogenetic MANOVAs based on penalized likelihood ("loocv-1" and "loocv-2") show good performances in terms of statistical power in all of the scenarios and for various level of phylogenetic signal (Pagel's λ), and type I error to the nominal rate. The conventional OLS (i.e. non-phylogenetic) MANOVA shows increased type I error with increasing phylogenetic signal.*

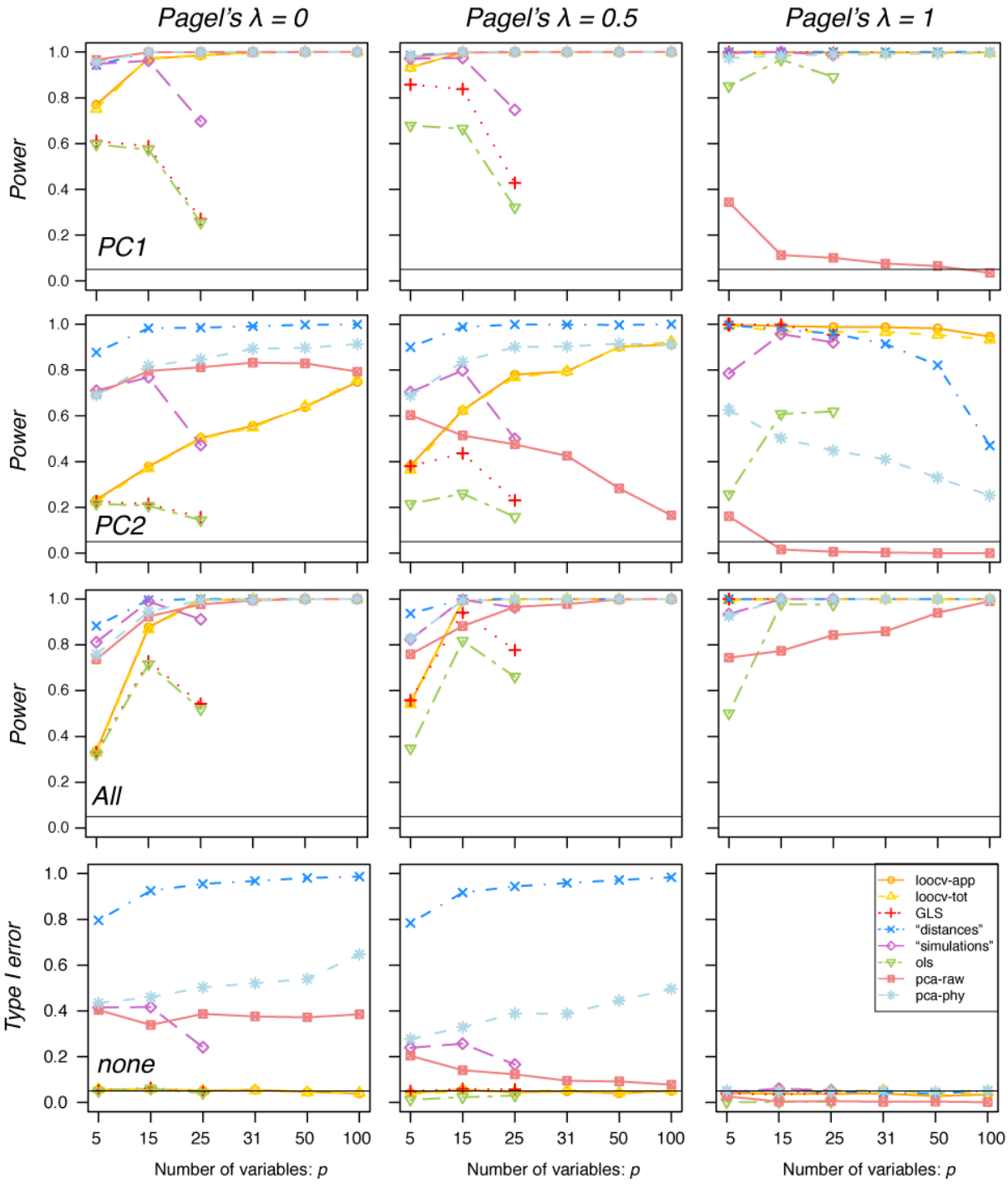

Figure S2. *Comparison of the statistical performances (statistical power and error) for the various MANOVA approaches with a non-phylogenetically distributed binary predictor and an average correlation between the dependent variables of* ***0.2****. In the top row the differences between groups were simulated along the first axis of variance (PC1), along the second axis (PC2) on the second row, and across all the axes (All) on the third row (see text). There were no differences simulated between groups (none) on the last row (for type I error testing). We can note that the regularized phylogenetic MANOVAs based on penalized likelihood ("loocv-1" and "loocv-2") show good performances in terms of statistical power in all of the scenarios and for various level of phylogenetic signal (Pagel's λ), and type I error to the nominal rate. In contrast, methods that are restricted to Brownian motion (distance-based, simulations-based, and PCA-based MANOVA) show high type I error with decreasing phylogenetic signal (or equivalently increased random noise; see text).*

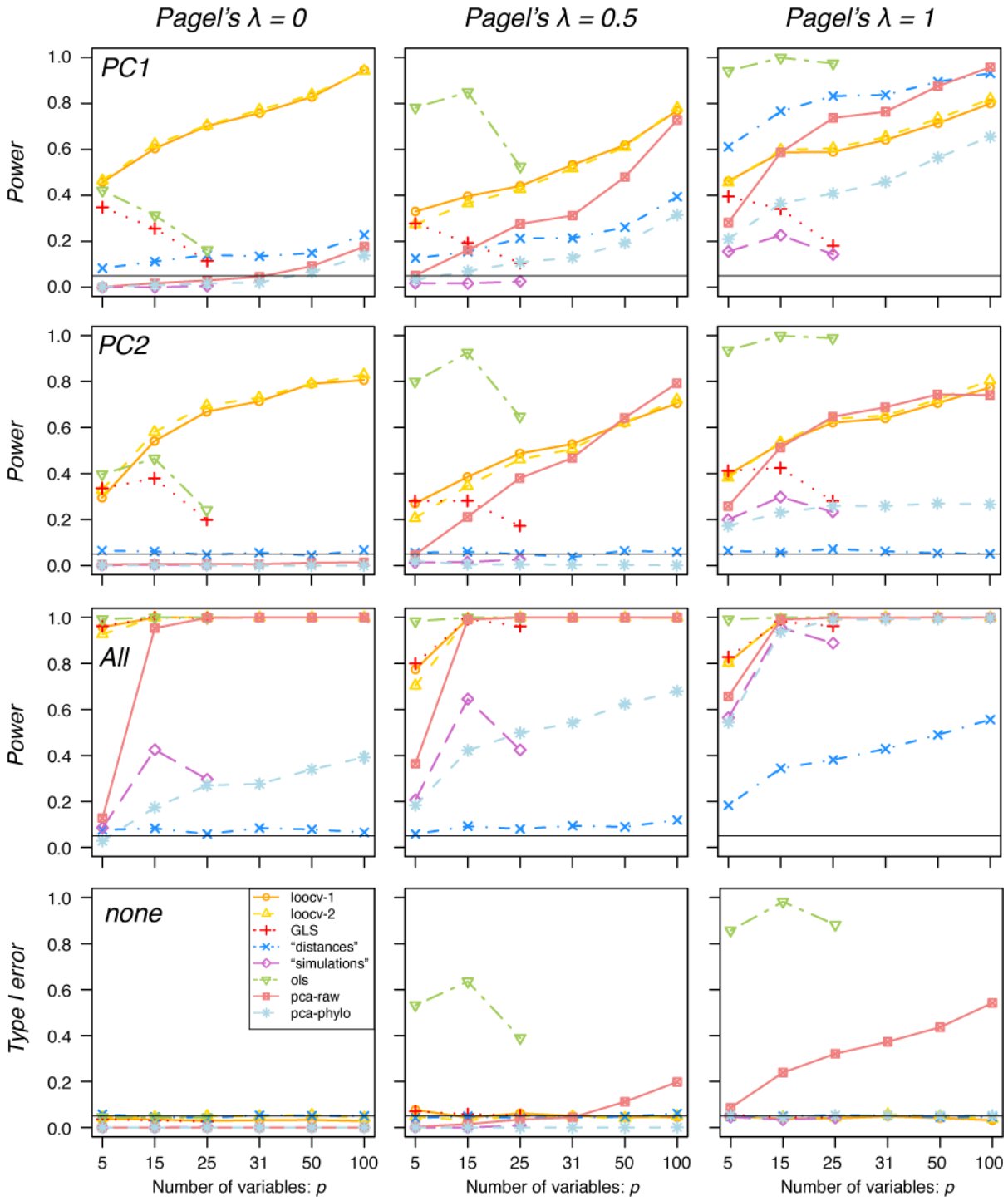

Figure S3. *Comparison of the statistical performances (statistical power and error) for the various MANOVA approaches with a phylogenetically clustered binary predictor variable and an average correlation between the dependent variables of* ***0.8****. In the top row the differences between groups were simulated along the first axis of variance (PC1), along the second axis (PC2) on the second row, and across all the axes (All) on the third row (see text). There were no differences simulated between groups (none) on the last row (for type I error testing). We can note that the regularized phylogenetic MANOVAs based on penalized likelihood ("loocv-1" and "loocv-2") show good performances in terms of statistical power in all of the scenarios and for various level of phylogenetic signal (Pagel's λ), and type I error to the nominal rate. The conventional OLS (i.e. non-phylogenetic) MANOVA shows increased type I error with increasing phylogenetic signal.*

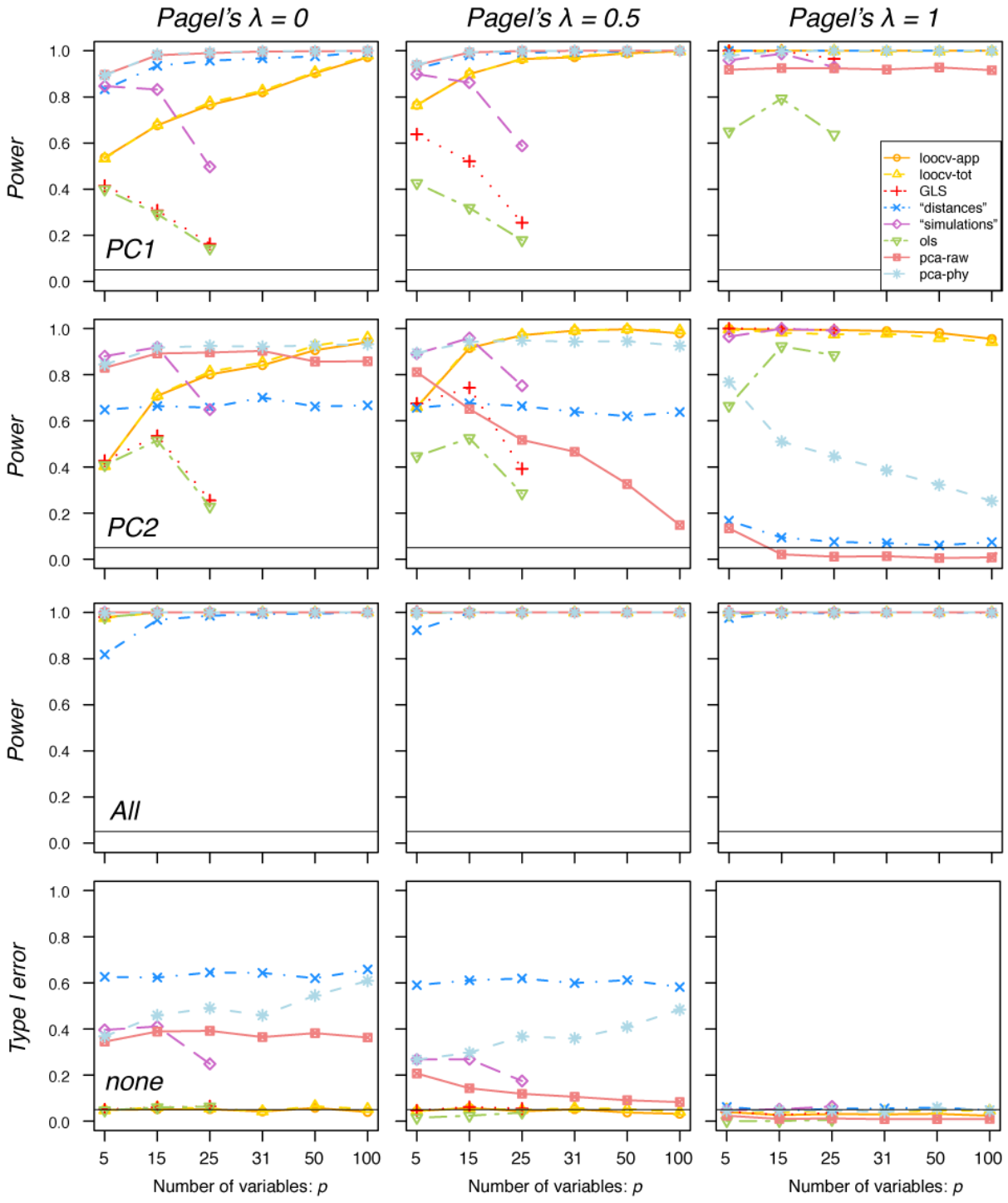

Figure S4. *Comparison of the statistical performances (statistical power and error) for the various MANOVA approaches with a non-phylogenetically distributed binary predictor and an average correlation between the dependent variables of* ***0.8****. In the top row the differences between groups were simulated along the first axis of variance (PC1), along the second axis (PC2) on the second row, and across all the axes (All) on the third row (see text). There were no differences simulated between groups (none) on the last row (for type I error testing). We can note that the regularized phylogenetic MANOVAs based on penalized likelihood ("loocv-1" and "loocv-2") show good performances in terms of statistical power in all of the scenarios and for various level of phylogenetic signal (Pagel's λ), and type I error to the nominal rate. In contrast, methods that are restricted to Brownian motion (distance-based, simulations-based, and PCA-based MANOVA) show high type I error with decreasing phylogenetic signal (or equivalently increased random noise; see text).*

2) Plots of the PL-pPCA axes for the Phyllostomid bats datasets along with their diet and feeding modes.

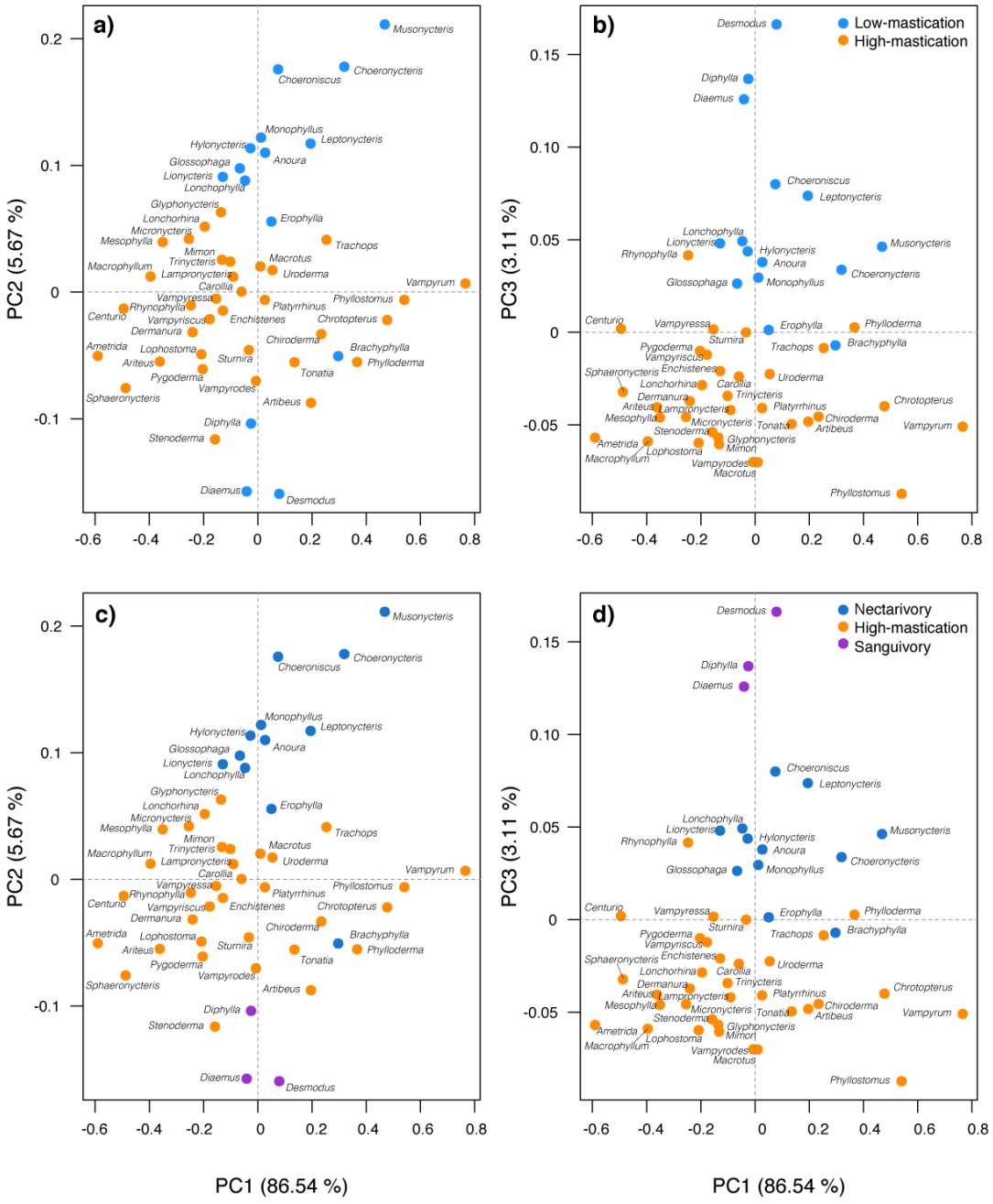

Figure S5. *Plot of the three first PL-pPC axes on the Phyllostomid bats dataset. In a) and b) the species are distinguished by a colour coding depending on whether their feeding mode involves a considerable amount of mastication ("high-mastication") or little or no mastication ("low-mastication"). In c) and d), the "low-mastication" species are further distinguished depending on whether they have a sanguivores or nectarivores diet. The proposed feeding modes and diets are described in Monteiro & Nogueira (2011). Sanguivores and nectarivores species are well separated from the species with high-mastication on the second and third PL-pPC axes.*

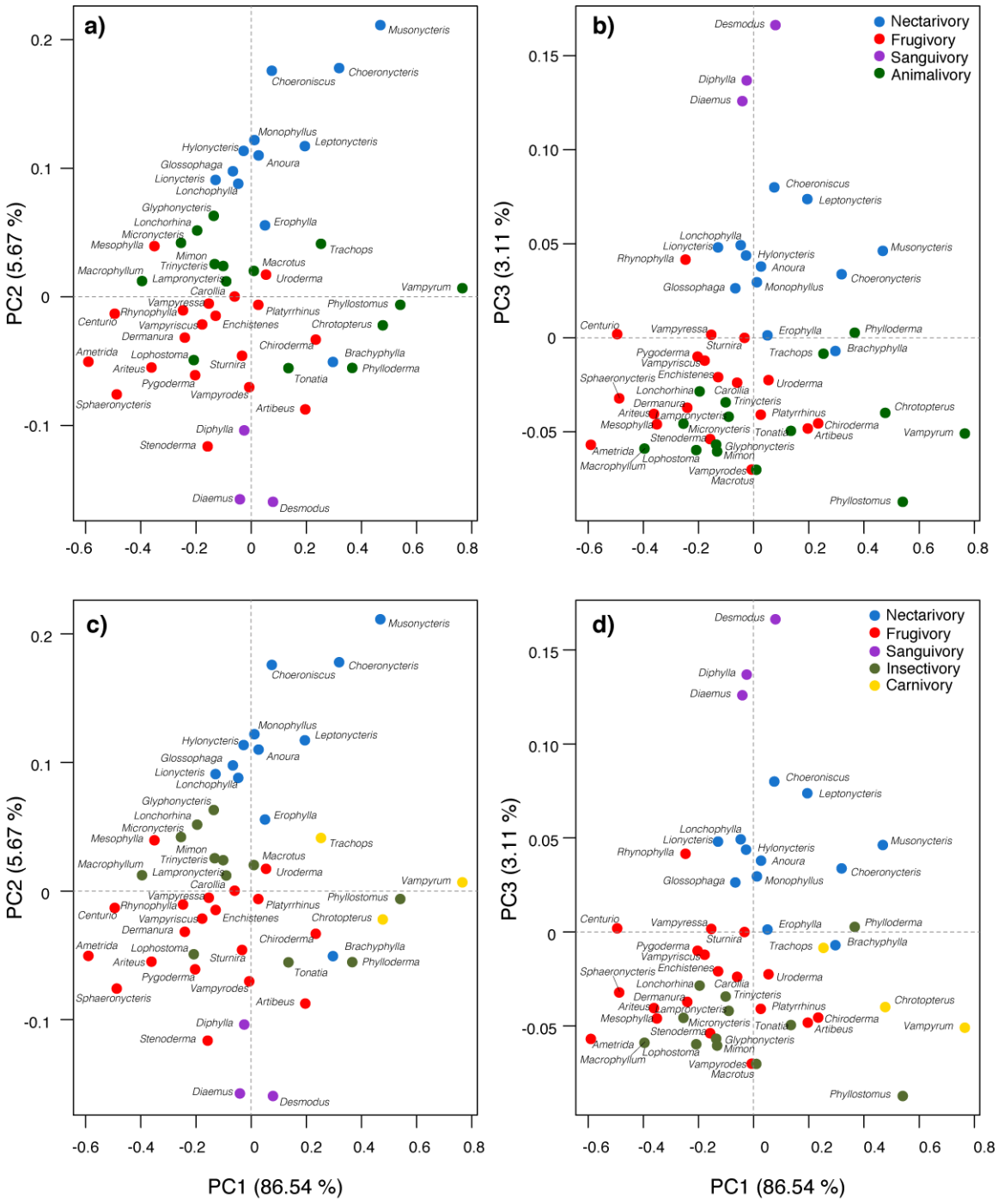

Figure S6. *Plot of the three first PL-pPC axes on the Phyllostomid bats dataset. In a) and b) the species are distinguished by a colour coding depending on their feeding mode: frugivores, animalivores, nectarivores, or sanguivores. In c) and d), the "animalivores" species are further distinguished depending on whether they are specialized in insectivory or carnivory. The proposed feeding modes and diets are described in Monteiro & Nogueira (2011). We can notice that the carnivore's species are the only one that can be separated from the other species on the first PC axis.*

3) Results of univariate analysis of variance on the pPCs axes of the phyllostomid bats dataset.

To further explore (and test) on which axes the differences detected by the penalized likelihood approach are located, we performed univariate phylogenetic ANOVAs on each of the three first separate phylogenetic PC axes (Table S1). For each grouping configuration, we first transformed the phylogenetic tree according to the Pagel's λ values estimated by the PL-MANOVA procedure (Table 2a). Next, we computed the pPC axes using the *mvgls.pca* function in mvMORPH under a BM process. Finally, we performed a phylogenetic ANOVA on each of the three first pPC axes using the *gls* function from "nlme" package, with the covariance structure provided by the *corPagel* function in "ape" (Paradis et al. 2004). No significant differences were detected on the first PC axis, with the exception of the five groups scenario (Table S1, see also Figures S5-6). Differences occurred on pPC2 (except for the 2-groups scenario, see also Figure 2 in the main text and Figure S5) and pPC3 (for all scenarios).

Table S1. *Results of univariate pANOVAs on the first pPC axes for the various dietary grouping scenarios considered for the phyllostomid bats dataset.*

	pANOVA - Tests significance (p-values)			
	Grouping 1	Grouping 2	Grouping 3	Grouping 4
pPC1*	0.5331	0.7739	0.3552	**0.0006**
pPC2*	0.0983	**<0.0001**	**<0.0001**	**<0.0001**
pPC3*	**<0.0001**	**<0.0001**	**<0.0001**	**<0.0001**

*The phylogenetic tree was transformed to the Pagel's λ values estimated by the PL approach for each scenario before computing the pPCs and conducting the pANOVAs

4) General linear hypothesis testing using multivariate contrasts and application to the bats dataset

When the multivariate test on the GLS model $\boldsymbol{Y} = \boldsymbol{XB} + \boldsymbol{\Xi}$ is significant, we can use general linear hypothesis testing to test more specific hypotheses of the form:

$$H_0: \boldsymbol{LB} = \boldsymbol{\Theta}$$

Where $\boldsymbol{B}$ is the $q \times p$ matrix of parameters estimated while performing the GLS fit, $\boldsymbol{L}$ is a $k \times q$ matrix of full row rank, $rank(\boldsymbol{L}) = k \leq q$, called the contrast coding matrix, and $\boldsymbol{\Theta}$ is a $k \times p$ constant matrix, usually full of zeros, called the right-hand-side (rhs) matrix, that together specify the hypotheses to be tested using *k* linear combinations of the parameters $\boldsymbol{B}$ (see details in Fox 2015; and also Rencher 2002 - p. 180-183).

The multivariate tests of such general hypotheses (e.g., using the Wilks's lambda, Pillai's trace, Lawley-Hotelling, or Roy's largest root tests) are based on the eigenvalues of the $p \times p$ matrix $\boldsymbol{E}^{-1}\boldsymbol{H}$ described in the main text with the hypothesis matrix $\boldsymbol{H}$ given by (see Rao et al. 1993; Timm 2002):

$$\boldsymbol{H} = (\boldsymbol{B}^T\boldsymbol{L}^T - \boldsymbol{\Theta}^T)[\boldsymbol{L}(\boldsymbol{X}^T\boldsymbol{C}^{-1}\boldsymbol{X})^{-1}\boldsymbol{L}^T]^{-1}(\boldsymbol{LB} - \boldsymbol{\Theta})$$

Where $\boldsymbol{C}$ is the phylogenetic variance-covariance matrix estimated during the GLS fit. When the significance of the test is assessed using the permutation procedure (e.g. in the PL test, cf main text), we need the design matrix of the null (or reduced) model $\boldsymbol{X_0}$, which is given by $\boldsymbol{X_0} = \boldsymbol{X}(\boldsymbol{I} - \boldsymbol{L}^T(\boldsymbol{L}^T)^+)$ where $\boldsymbol{A}^+$ stands for the pseudoinverse of a given matrix $\boldsymbol{A}$, and $\boldsymbol{I}$ is the $q \times q$ identity matrix (Winkler et al. 2014). The test statistic distribution is then obtained as described in the main text.

We illustrate the general linear hypothesis testing on the bat dataset. The overall MANOVA test in the 5 groups scenario rejected the null hypothesis ($H_0: \boldsymbol{XB} = \boldsymbol{1B}_{null}$). We test three specific hypotheses about morphological difference between groups: ***1)*** difference

between the sanguivorous and the nectivorous species, ***2)*** between the frugivorous and the animalivorous, and ***3)*** between the carnivorous and the insectivorous. These three tests correspond to the following linear hypotheses:

$$0\boldsymbol{B}_{Frugivory} + 0\boldsymbol{B}_{Insectivory} + 1\boldsymbol{B}_{Nectarivory} + 0\boldsymbol{B}_{Carnivory} - 1\boldsymbol{B}_{Sanguivory} = \boldsymbol{\Theta}_1$$

$$1\boldsymbol{B}_{Frugivory} - 0.5\boldsymbol{B}_{Insectivory} + 0\boldsymbol{B}_{Nectarivory} - 0.5\boldsymbol{B}_{Carnivory} + 0\boldsymbol{B}_{Sanguivory} = \boldsymbol{\Theta}_2$$

$$0\boldsymbol{B}_{Frugivory} + 1\boldsymbol{B}_{Insectivory} + 0\boldsymbol{B}_{Nectarivory} - 1\boldsymbol{B}_{Carnivory} + 0\boldsymbol{B}_{Sanguivory} = \boldsymbol{\Theta}_3$$

Where $\boldsymbol{B}_{diet}$ is the row of $\boldsymbol{B}$ that corresponds to the specified diet category, and $\boldsymbol{\Theta}_1, \boldsymbol{\Theta}_2$ and $\boldsymbol{\Theta}_3$ are taken to be vectors of length *p* full of zeros. The corresponding $\boldsymbol{L}$ contrast coding matrices used to test each of the three hypotheses are therefore:

$$\boldsymbol{L}_1 = [0 \quad 0 \quad 1 \quad 0 \quad -1]$$

$$\boldsymbol{L}_2 = [1 \quad -0.5 \quad 0 \quad -0.5 \quad 0]$$

$$\boldsymbol{L}_3 = [0 \quad 1 \quad 0 \quad -1 \quad 0]$$

We ran the *manova.gls* function with the contrast coding matrices $\boldsymbol{L}$ defined above and right-hand-side null matrices $\boldsymbol{\Theta}$ with the Wilks statistic (9999 permutations) on the bat dataset. We used Holm and Bonferroni multiple testing corrections (Holm 1979) on the estimated *p-values* using the *p.adjust* function in the R package "stats". Results are presented in Table (S2) and show that all the tests rejected the null hypothesis, suggesting that the various dietary strategies are associated to well differentiated morphologies.

Table S2. *Results from linear hypothesis tests (MANOVA on contrasts variables) on the bats dataset (9999 permutations).*

	Linear hypothesis			
	Wilks λ	p-value	p-value (Holm)	p-value (Bonferonni)
Animalivorous vs. Frugivorous	0.0521	**0.0001**	**0.0003**	**0.0003**
Insectivorous vs. Carnivorous	0.1331	**0.0033**	**0.0033**	**0.0099**
Nectivorous vs. Sanguivorous	0.0279	**0.0001**	**0.0003**	**0.0003**

5) Simulating differences between groups in the multidimensional space.

To simulate scenarios where the differences between the binary groups are located on specific dimensions of the multivariate space (i.e., $\boldsymbol{\mu}_2 \neq \boldsymbol{\mu}_1$ on PC1, PC2, or all the PCs axes), we first rotated the simulated datasets $\boldsymbol{Y}$ to its principal component axes:

$$\widetilde{\boldsymbol{Y}} = \boldsymbol{YU}$$

Where $\boldsymbol{R} = \boldsymbol{U}^T\boldsymbol{DU}$ is an eigen decomposition of the trait's covariance matrix $\boldsymbol{R}$. In practice we used the simulated covariance matrix $\boldsymbol{R}$ (as explained in the main text) rather than the empirical (sample) covariance because it corresponds to the population estimate, and because the later estimate is singular in high-dimensions. Differences between groups were then introduced by shifting part of the values from the j^{th} column (or columns) of $\widetilde{\boldsymbol{Y}}$ corresponding to the desired PCs dimensions (e.g., PC1, PC2 and all the PCs in our experiment):

$$\widehat{\boldsymbol{Y}}_j = \widetilde{\boldsymbol{Y}}_j + \boldsymbol{X}\beta$$

Where $\boldsymbol{X}$ is a design matrix with dummy coding specifying the row of $\widetilde{\boldsymbol{Y}}_j$ that correspond to each of the two groups, and $\beta = \{\boldsymbol{\mu}_1; \boldsymbol{\mu}_2\}$ is a vector with the two groups mean ($\boldsymbol{\mu}_1 = 0$ and $\boldsymbol{\mu}_2 \neq \boldsymbol{\mu}_1$)**.** The effect-size chosen for the shifts from $\boldsymbol{\mu}_1$ to $\boldsymbol{\mu}_2$ are described in the main text. To preserve the rank of the eigenvalues of the original data on the transformed dataset, we need to standardize first each transformed PCs axis to the variance of the original PCs:

$$\widehat{\boldsymbol{Y}}_j = \widehat{\boldsymbol{Y}}_j \times \sqrt{Var(\widetilde{\boldsymbol{Y}}_j)/Var(\widehat{\boldsymbol{Y}}_j)}$$

We then center the *j* transformed PCs on the values of the original PCs:

$$\widehat{\boldsymbol{Y}}_j = \widehat{\boldsymbol{Y}}_j - E[\widehat{\boldsymbol{Y}}_j] + E[\widetilde{\boldsymbol{Y}}_j]$$

And we finally transform back the transformed and standardized PCs to the original space using the eigenvector matrix $\boldsymbol{U}$:

$$\boldsymbol{Y}^* = \widehat{\boldsymbol{Y}}\boldsymbol{U}^T$$

The values of $\boldsymbol{Y}^*$ show differences located on the desired PCs axes and are used in the downstream simulations. Note that, when the traits were simulated with a phylogenetic signal, the mean and variance of the respective PCs are computed using the generalized least squares estimates given by:

$$E[\boldsymbol{Y}] = (\mathbf{1}^T \boldsymbol{C}^{-1} \mathbf{1})^{-1} \mathbf{1}^T \boldsymbol{C}^{-1} \boldsymbol{Y}$$

and,

$$Var(\boldsymbol{Y}) = diag((\boldsymbol{Y} - E[\boldsymbol{Y}])^T \boldsymbol{C}^{-1} (\boldsymbol{Y} - E[\boldsymbol{Y}]))/(n-1)$$

Where $\boldsymbol{Y}$ is either $\widehat{\boldsymbol{Y}}_j$ or $\widetilde{\boldsymbol{Y}}_j$, n is the number of species, $\boldsymbol{C}$ is the variance covariance matrix of the evolutionary process, and $\mathbf{1}$ is a column vector of one.

6) Step by step calculation of the first and second order derivative of the log-likelihood in Appendix 2

It is shown in the Appendix 2, that the cross-validated log-likelihood (Equation A2 in Appendix 1) can be computed using the following equation:

$$\mathcal{L}_{CV} \propto -\frac{p\log|\boldsymbol{C}| - p\log|\boldsymbol{X}^T \boldsymbol{C}^{-1} \boldsymbol{X}|}{2}$$
$$-\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{p}\left[\frac{(n-q)}{n}\log\left[(1-\gamma)d_j^{(-i)} + \gamma t_j\right] + \frac{\left(z_j^{(i)}\right)^2}{(1-\gamma)d_j^{(-i)} + \gamma t_j}\right]$$

To differentiate this function with respect to the regularization parameter γ, we make use of the sum/difference rule $(f \pm g)' = f' \pm g'$ to differentiate the summands separately. We start by using the chain rule $[\log(u(x))]' = \frac{1}{u(x)} \times u'(x)$ for the term $\log\left[(1-\gamma)d_j^{(-i)} + \gamma t_j\right]$ and we factor out the constant terms in $\frac{\left(z_j^{(i)}\right)^2}{(1-\gamma)d_j^{(-i)} + \gamma t_j}$ and $c = \frac{(n-q)}{n}$ using the constant factor rule $(a \times f)' = a \times f'$. We can now apply the reciprocal rule $[\frac{1}{u(x)}]' = -\frac{u'(x)}{u(x)^2}$ which leads to:

$$\frac{\partial \mathcal{L}_{CV}}{\partial \gamma} = -\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{p}\left[-\frac{\frac{\partial}{\partial \gamma}\left[(1-\gamma)d_j^{(-i)} + \gamma t_j\right]}{\left((1-\gamma)d_j^{(-i)} + \gamma t_j\right)^2} \times \left(z_j^{(i)}\right)^2 + \frac{1}{(1-\gamma)d_j^{(-i)} + \gamma t_j} \times \frac{\partial}{\partial \gamma}\left[(1-\gamma)d_j^{(-i)} + \gamma t_j\right] \times c\right]$$

We reuse the sum rule and the standard derivative rules $x' = 1$ and $a' = 0$ (for some constant a), for solving $\frac{\partial}{\partial \gamma}\left[(1-\gamma)d_j^{(-i)} + \gamma t_j\right]$:

$$\frac{\partial}{\partial \gamma}\left[(1-\gamma)d_j^{(-i)} + \gamma t_j\right] = \frac{\partial}{\partial \gamma}\left[(1-\gamma)d_j^{(-i)}\right] + \frac{\partial}{\partial \gamma}\left[\gamma t_j\right] = d_j^{(-i)}\frac{\partial}{\partial \gamma}[(1-\gamma)] + t_j\frac{\partial}{\partial \gamma}[\gamma]$$
$$= -d_j^{(-i)} + t_j$$

We then substitute back into the previous equation

$$\frac{\partial \mathcal{L}_{CV}}{\partial \gamma} = -\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{p}\left[-\frac{\left(-d_j^{(-i)} + t_j\right)\left(z_j^{(i)}\right)^2}{\left((1-\gamma)d_j^{(-i)} + \gamma t_j\right)^2} + \frac{(-d_j^{(-i)} + t_j) \times c}{(1-\gamma)d_j^{(-i)} + \gamma t_j}\right]$$
$$= -\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{p}\left[\frac{\left(t_j - d_j^{(-i)}\right) \times c}{(1-\gamma)d_j^{(-i)} + \gamma t_j} - \frac{\left(t_j - d_j^{(-i)}\right)\left(z_j^{(i)}\right)^2}{\left((1-\gamma)d_j^{(-i)} + \gamma t_j\right)^2}\right]$$

After simplifications we obtain the first derivative as follow:

$$\frac{\partial \mathcal{L}_{CV}}{\partial \gamma} = -\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{p}\frac{\left(t_j - d_j^{(-i)}\right)\left(\gamma t c + d_j^{(-i)}c - d_j^{(-i)}\gamma c - \left(z_j^{(i)}\right)^2\right)}{\left(d_j^{(-i)} - d_j^{(-i)}\gamma + \gamma t_j\right)^2}$$

We can again use the constant factor rule to the above equation to obtain the closed form solution to the second order derivative:

$$\frac{\partial^2 \mathcal{L}_{CV}}{\partial \gamma^2} = -\frac{1}{2n}\sum_{i=1}^{n}\sum_{j=1}^{p}\left(t_j - d_j^{(-i)}\right)\frac{\partial}{\partial \gamma}\left[\frac{\left(\gamma t c + d_j^{(-i)}c - d_j^{(-i)}\gamma c - \left(z_j^{(i)}\right)^2\right)}{\left(d_j^{(-i)} - d_j^{(-i)}\gamma + \gamma t_j\right)^2}\right]$$

We then use the quotient rule $\left[\frac{u(x)}{v(x)}\right]' = \frac{u'(x)v(x) - u(x)v'(x)}{v(x)^2}$ to solve the derivative:

$$\frac{\partial^2 \mathcal{L}_{CV}}{\partial \gamma^2} = -\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{p}\left(t_j - d_j^{(-i)}\right) \times$$

$$\frac{\frac{\partial}{\partial\gamma}\left[c\gamma t + c d_j^{(-i)} - c d_j^{(-i)}\gamma - \left(z_j^{(i)}\right)^2\right] \times \left(d_j^{(-i)} - d_j^{(-i)}\gamma + \gamma t_j\right)^2 - \left(ct\gamma + c d_j^{(-i)} - c d_j^{(-i)}\gamma - \left(z_j^{(i)}\right)^2\right) \times \frac{\partial}{\partial\gamma}\left[\left(d_j^{(-i)} - d_j^{(-i)}\gamma + \gamma t_j\right)^2\right]}{\left(\left(d_j^{(-i)} - d_j^{(-i)}\gamma + \gamma t_j\right)^2\right)^2} =$$

$$-\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{p}\left(t_j - d_j^{(-i)}\right) \times$$

$$\frac{\frac{\partial}{\partial\gamma}\left[ct + 0 - c d_j^{(-i)} - 0\right] \times \left(d_j^{(-i)} - d_j^{(-i)}\gamma + \gamma t_j\right)^2 - \left(ct\gamma + c d_j^{(-i)} - c d_j^{(-i)}\gamma - \left(z_j^{(i)}\right)^2\right) \times \frac{\partial}{\partial\gamma}\left[\left(d_j^{(-i)} - d_j^{(-i)}\gamma + \gamma t_j\right)^2\right]}{\left(\left(d_j^{(-i)} - d_j^{(-i)}\gamma + \gamma t_j\right)^2\right)^2}$$

Finally, we need the power rule $[u(x)^n]' = n \times u(x)^{n-1} \times u'(x)$ to differentiate the term $\frac{\partial}{\partial\gamma}\left[\left(d_j^{(-i)} - d_j^{(-i)}\gamma + \gamma t_j\right)^2\right]$ in the above equation:

$$\frac{\partial}{\partial\gamma}\left[\left(d_j^{(-i)} - d_j^{(-i)}\gamma + \gamma t_j\right)^2\right] = 2\left(d_j^{(-i)} - d_j^{(-i)}\gamma + \gamma t_j\right) \times \frac{\partial}{\partial\gamma}\left[\left(d_j^{(-i)} - d_j^{(-i)}\gamma + \gamma t_j\right)\right]$$

Again, we have $\frac{\partial}{\partial\gamma}\left[\left(d_j^{(-i)} - d_j^{(-i)}\gamma + \gamma t_j\right)\right] = -d_j^{(-i)} + t_j = t_j - d_j^{(-i)}$. Substituting we obtain:

$$\frac{\partial^2 \mathcal{L}_{CV}}{\partial \gamma^2} =$$

$$-\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{p}\frac{\left(t_j - d_j^{(-i)}\right)\left(\left(ct_j - c d_j^{(-i)}\right)\left(d_j^{(-i)} - d_j^{(-i)}\gamma + \gamma t_j\right)^2 - 2\left(t_j - d_j^{(-i)}\right)\left(\left(d_j^{(-i)} - d_j^{(-i)}\gamma + \gamma t_j\right)\right)\left(ct\gamma + c d_j^{(-i)} - c d_j^{(-i)}\gamma - \left(z_j^{(i)}\right)^2\right)\right)}{\left(\left(d_j^{(-i)} - d_j^{(-i)}\gamma + \gamma t_j\right)^2\right)^2}$$

After simplifications we can rewrite the above equation in a more compact form:

$$\frac{\partial^2 \mathcal{L}_{CV}}{\partial \gamma^2} = -\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{p}\frac{\left(t_j - d_j^{(-i)}\right)^2\left(ct\gamma + c d_j^{(-i)} - c d_j^{(-i)}\gamma - 2\left(z_j^{(i)}\right)^2\right)}{\left(d_j^{(-i)} - d_j^{(-i)}\gamma + \gamma t_j\right)^3}$$

References:

Anderson M., Braak C.T. 2003. Permutation tests for multi-factorial analysis of variance. Journal of Statistical Computation and Simulation. 73:85–113.

Fox J. 2015. Applied Regression Analysis and Generalized Linear Models. SAGE Publications.

Freedman D., Lane D. 1983. A nonstochastic interpretation of reported significance levels. Journal of Business & Economic Statistics. 1:292–298.

Holm S. 1979. A simple sequentially rejective multiple test procedure. Scandinavian Journal of Statistics. 6:65–70.

Monteiro L.R., Nogueira M.R. 2011. Evolutionary patterns and processes in the radiation of phyllostomid bats. BMC Evolutionary Biology. 11:1–23.

Paradis E., Claude J., Strimmer K. 2004. APE: Analysis of Phylogenetics and Evolutions in R language. Bioinformatics. 20:289–290.

Rao J.N.K., Brajendra C.S., Yue K. 1993. Generalized Least Squares F test in regression analysis with two-stage cluster samples. Journal of the American Statistical Association. 88:1388–1391.

Rencher A.C. 2002. Methods of Multivariate Analysis. New York: John Wiley & Sons.

Timm N.H. 2002. Applied Multivariate Analysis. Springer-Verlag New York.

Winkler A.M., Ridgway G.R., Webster M.A., Smith S.M., Nichols T.E. 2014. Permutation inference for the general linear model. NeuroImage. 92:381–397.