An empirical bayes phylogenetic comparative approach to study the phenotypic evolution from high-dimensional data

Paola Montoya Valencia*[1], Hélène Morlon[2], and Julien Clavel[1]

[1]University Lyon, UMR CNRS 5023 LEHNA, F-69622, Villeurbanne, France – CNRS : UMR5023 – France

[2]Institut de biologie de l'ENS Paris – Département de Biologie - ENS Paris, Institut National de la Santé et de la Recherche Médicale, Centre National de la Recherche Scientifique – 46, Rue d'Ulm75005 Paris, France

Résumé

The fast increase of multivariate phylogenetic comparative methods (PCMs) and data across species is extending our capacity to understand the evolution of complex phenotypes. However, high dimensional datasets such as those from geometric morphometric studies – where the number of traits is often higher than the number of species – are still challenging for PCMs. Here, we propose a new regularized approach using an empirical bayes framework, to fit evolutionary models on high-dimensional datasets. Using simulations, we tested the accuracy and time performances of the proposed approach to estimate multiple parameters from commonly used evolutionary models such as Pagel's lambda or Brownian motion, while comparing them to current PCMs. The proposed approach exhibits high accuracy in parameter estimation (e.g., Pagel's lambda) in a wide range of simulations and conditions, including a dimensionality ranging from equal number of traits and species up to 30 times more traits than species. The estimated traits covariance matrices remain close to the simulated (true) one as the number of traits increases relative to the number of species, with similar performances to recently proposed approaches that can handle high-dimensional datasets. However, our approach shows a considerable reduction in computing time (up to 200 times faster), providing an accurate and fast alternative to deal with increasingly large datasets (e.g., from 3D geometric morphometrics) along with a promising framework for extending current evolutionary models.

Mots-Clés: multivariate phylogenetic comparative methods, regularization, evolutionary models

---

*Intervenant