mvMORPH: an R package for the fitting of multivariate evolutionary models to morphometric data
Using phylogeny to describe evolutionary dynamic:

tempo of phenotypic evolution
We can use Maximum Likelihood to estimate phenotypic rate of evolution.

- Is disparity related to the age of the clade?
- We need a model of trait evolution

Modelling approach: univariate models
Phenotypic Evolution on Phylogeny

- Evolution of the phenotypic variance between species
- *e.g.* Brownian motion (BM)

Brownian rate: rate of variance accumulation through time
Modelling approach: univariate models

✅ *Filling the VCV matrix* or *transforming the tree*

Different evolutionary rates between clades

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1.0</td>
<td>0.28</td>
<td>0.28</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>B</td>
<td>0.28</td>
<td>1.0</td>
<td>0.42</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>C</td>
<td>0.28</td>
<td>0.42</td>
<td>1.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>D</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>2.0</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>E</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.5</td>
<td>2.0</td>
<td>0.85</td>
</tr>
<tr>
<td>F</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.5</td>
<td>0.85</td>
<td>2.0</td>
</tr>
</tbody>
</table>
Modelling approach: univariate models

✓ **Filling the VCV matrix/ or transforming the tree**

Different evolutionary rates between clades

\[
V = \begin{bmatrix}
1 & 0.28 & 0.28 & 0 & 0 & 0 \\
0.28 & 1 & 0.42 & 0 & 0 & 0 \\
0.28 & 0.42 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 2 & 0.5 & 0.5 \\
0 & 0 & 0 & 0.5 & 2 & 0.85 \\
0 & 0 & 0 & 0.5 & 0.85 & 2
\end{bmatrix}
\]

Phenotypic trait vector

Expectation

Evolutionary vcv matrix

\[
\log(L) = \log \left[ \frac{\exp \left\{ -\frac{1}{2} [X - E(X)]' (V)^{-1} [X - E(X)] \right\}}{\sqrt{2\pi}^N |V|} \right]
\]
Macroevolution

Multivariate evolution?
Dealing with multivariate data: phenotypic disparity and integration

Traits interactions in a hypothesis testing framework

Evolutionary rates are expressed in $n \times n$ dimensions matrices
Dealing with multivariate data: phenotypic disparity and integration

Selection on multiple traits allows making predictive assumptions on the way the traits may covary through time. Evolutionary rates are expressed in $n \times n$ dimensions matrices.

Evolutionary rates matrix describe at least three aspects:

- The size: the total amount of phenotypic variance in traits (evolutionary rates)
- Phenotypic covariances/correlations (describe the phenotypic integration)
- Orientation of covariances (Does the evolutionary rates matrix for different clades share the same eigen-structure?)

Changes in shape, size, and orientation are illustrated in the diagrams.
Dealing with multivariate data: kroenecker product of rate and vcv matrix

Multivariate models

$R_1$}

\[
\begin{bmatrix}
1 & 0 \\
0 & 1
\end{bmatrix}
\]

$C_1$

\[
\begin{bmatrix}
\begin{array}{cccccc}
A & B & C & D & E & F \\
A & 1 & 0.28 & 0.28 & 0 & 0 & 0 \\
B & 0.28 & 1 & 0.42 & 0 & 0 & 0 \\
C & 0.28 & 0.42 & 1 & 0 & 0 & 0 \\
D & 0 & 0 & 0 & 0 & 0 & 0 \\
E & 0 & 0 & 0 & 0 & 0 & 0 \\
F & 0 & 0 & 0 & 0 & 0 & 0 \\
\end{array}
\end{bmatrix}
\]

\[
\begin{bmatrix}
\begin{array}{cccccc}
A & B & C & D & E & F \\
A & 0 & 0 & 0 & 0 & 0 & 0 \\
B & 0 & 0 & 0 & 0 & 0 & 0 \\
C & 0 & 0 & 0 & 0 & 0 & 0 \\
D & 0 & 0 & 0 & 0 & 0 & 0 \\
E & 0 & 0 & 0 & 0 & 0 & 0 \\
F & 0 & 0 & 0 & 0 & 0 & 0 \\
\end{array}
\end{bmatrix}
\]
**Multivariate models**

*Dealing with multivariate data: kroenecker product of rate and vcv matrix*

\[
\begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 2 & 0.5 & 0.5 \\
0 & 0 & 0 & 0.5 & 2 & 0.85 \\
0 & 0 & 0 & 0.5 & 0.85 & 2
\end{bmatrix}
\times
\begin{bmatrix}
2 & 0 \\
0 & 2
\end{bmatrix}
= \begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 2 & 0.5 & 0.5 \\
0 & 0 & 0 & 0.5 & 2 & 0.85 \\
0 & 0 & 0 & 0.5 & 0.85 & 2
\end{bmatrix}
\]
Multivariate models

Dealing with multivariate data: kroenecker product of rate and vcv matrix

\[ V = C_1 + C_2 = \]

\[
\begin{pmatrix}
1 & 0.28 & 0.28 & 0 & 0 & 0 \\
0.28 & 1 & 0.42 & 0 & 0 & 0 \\
0.28 & 0.42 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 2 & 0.5 & 0.5 \\
0 & 0 & 0 & 0.5 & 2 & 0.85 \\
0 & 0 & 0 & 0.5 & 0.85 & 2
\end{pmatrix}
\]
Multivariate models

Dealing with multivariate data: phenotypic disparity and integration

Evolutionary traits interactions in a hypothesis testing framework

\[
\begin{bmatrix}
5 & 2.8 \\
2.8 & 3
\end{bmatrix}
\approx \text{correlation: 0.72}
\]

\[
\begin{bmatrix}
8 & 1.5 \\
1.5 & 6
\end{bmatrix}
\approx \text{correlation: 0.21}
\]
**Dealing with multivariate data: phenotypic disparity and integration**

Evolutionary traits interactions in a hypothesis testing framework

More variance on traits for generalists species
And lower evolutionary correlations: disparate clade
Dealing with multivariate data: phenotypic disparity and integration

Evolutionary traits interactions in a hypothesis testing framework

Higher evolutionary covariance between traits: correlative selection? Functional constraints?

≈ correlation: 0.72

≈ correlation: 0.21
Dealing with multivariate data: phenotypic disparity and integration

Selection on multiple traits allows making predictive assumptions on the way the traits may covary through time. Evolutionary rates are expressed in $n \times n$ dimensions matrices.

Evolutionary rates matrix describe at least three aspects:

- The size: the total amount of phenotypic variance in traits (evolutionary rates)
- Phenotypic covariances/correlations (describe the phenotypic integration)
- Orientation of covariances (Does the evolutionary rates matrix for different clades share the same eigen-structure?)

Example: evolution toward three optima with OU process.

Multivariate models

*Dealing with multivariate data: phenotypic disparity and integration*

-E.g. we could expect that environnementally driven changes in evolutionary modes is accompanied by a change in phenotypic integration of functional traits.

*Exemple:* diversification associated to the intrusion of a competitive specie followed by a functional diversification

*But... few tools allows such a modeling approach*
Dealing with paleotrees: Development of models of evolutionary shifts

Paleo-trees allows estimating changes in evolutionary rates (tempo), but also changes in modes! No local or global minimum on ultrametric trees
mvMORPH: an R package for the fitting of multivariate evolutionary models and other things...

mvMORPH: an R package motivated by the need for tools to study the evolution of phenotypic traits with fossil data, and to study multivariate evolution of disparity and phenotypic integration in macroevolutionary studies.
mvMORPH

R packages for dealing with multivariate data

- Some packages availables on CRAN

+ multivariate Early-Burst Model, models of changes in evolutionary modes through times, possibility of testing for differences in evolutionary rates between traits...
mvMORPH: an R package for the fitting of multivariate evolutionary models and other things...

- Allows the fitting of 14 evolutionary models by maximum likelihood, mostly dedicated to multivariate data, but also on models dedicated to paleo-trees

<table>
<thead>
<tr>
<th>Function</th>
<th>Model</th>
<th>Short description of the associated evolutionary model</th>
</tr>
</thead>
<tbody>
<tr>
<td>mvBM</td>
<td>BM1</td>
<td>One BM rate per trait</td>
</tr>
<tr>
<td></td>
<td>BM1-constrained</td>
<td>Constrained to a unique BM rate common to all traits</td>
</tr>
<tr>
<td></td>
<td>BMM</td>
<td>One BM rate per selective regime and per trait</td>
</tr>
<tr>
<td></td>
<td>BMM-constrained</td>
<td>Constrained to one BM rate per selective regime common to all traits</td>
</tr>
<tr>
<td>mvOU</td>
<td>OU1</td>
<td>OU process with a unique adaptive optimum per trait</td>
</tr>
<tr>
<td></td>
<td>OU1-constrained</td>
<td>OU process constrained to evolve independently on each traits</td>
</tr>
<tr>
<td></td>
<td>OUM</td>
<td>OU process with multiple adaptive optima per trait</td>
</tr>
<tr>
<td></td>
<td>OUM-constrained</td>
<td>OU process constrained to evolve independently on each traits (with multiple optima)</td>
</tr>
<tr>
<td>mvEB</td>
<td>EB-ACDC</td>
<td>EB model or decelerating model of evolutionary rates.</td>
</tr>
<tr>
<td>mvSHIFT</td>
<td>ER</td>
<td>OU process that collapses to a BM drift at a given point in time</td>
</tr>
<tr>
<td></td>
<td>RR</td>
<td>OU process that switches to an independent BM process at a given point in time</td>
</tr>
<tr>
<td></td>
<td>EC</td>
<td>BM process pulled to an OU process at a given point in time</td>
</tr>
<tr>
<td></td>
<td>RC</td>
<td>Independent BM process switch to an OU process</td>
</tr>
<tr>
<td></td>
<td>SR</td>
<td>Change in BM rates at a given point in time</td>
</tr>
</tbody>
</table>

Current models in mvMORPH 1.0.2 available on CRAN
Main drawback

- Dealing with huge dataset (uncertainty due to rounding errors, and computational time (number of arithmetic operations \(O^3\) ) for matrix inversion)
- Number of parameters

But…

- Allows SIMMAP mapping, integrating uncertainty, working with paleotrees
- Building of variances-covariances matrix using faster C-code

Using Phylogenetic independent contrast (PIC) instead?

Main advantage:

- Does not need the computation (and inversion) of a variance-covariance matrix
- Time linear algorithm, need a transformed tree.
Performances

✓ Using Phylogenetic independent contrast (PIC) instead?

\[ a2 = \text{Contraste } (B-C) = \frac{(x2-x3)}{0.6} \]
\[ a1 = \text{Contraste } (A-a2) = \frac{(((x2+x3)/2)-x1)}{0.4} \]

Main disadvantage:

✓ Some tree transformation don’t works for paleo-trees

✓ Complex models that need a conception matrix are not easily handled
mvMORPH: an R package for the fitting of multivariate evolutionary models and other things...

Forthcoming: mvMORPH 1.0.3

- Additional models (multiple optimums in shifts,…)
- MANOVA parameterizable with the various models of mvMORPH
- S3 class for various functions of the package (loglik, summary, simulate…)
- Simulations functions with fitted models parameters
- Fast likelihood algorithms (three points structure) – Ho & Ané 2014 (Syst. Biol.)
- Combined calculation of the inverse and determinant of phylogenetic evolutionary matrix
  Involved in likelihood estimation with a LDL Cholesky factorizations (actually in C but maybe also done in Fortran for performances?)

The goal:
Tools for working with huge paleotrees!! (e.g. Raia et al. (2013) Proc. Roy. Soc., paleotree with 1100 taxa!); and multivariate datasets…
Thank you very much for your attention

mvMORPH: Multivariate Comparative Tools for Fitting Evolutionary Models to Morphometric Data

Fits multivariate (Brownian Motion, Early Burst, ACDC, Ornstein-Uhlenbeck and Shifts) models of continuous traits evolution on trees

Version: 1.0.2
Depends: R (≥ 2.9.1), phytools, ape, corpcor, subplex
Published: 2014-04-24
Author: Julien Clavel, with contributions from Aaron King, and Emmanuel Paradis
Maintainer: Julien Clavel <julien.clavel at hotmail.fr>
License: GPL-2 | GPL-3 [expanded from: GPL (≥ 2.0)]
NeedsCompilation: yes
CRAN checks: mvMORPH results

Downloads:

Reference manual: mvMORPH.pdf
Package source: mvMORPH_1.0.2.tar.gz
Windows binaries: r-devel: mvMORPH_1.0.2.zip, r-release: mvMORPH_1.0.2.zip, r-oldrel: mvMORPH_1.0.2.zip
OS X Snow Leopard binaries: r-release: mvMORPH_1.0.2.tgz, r-oldrel: mvMORPH_1.0.2.tgz
OS X Mavericks binaries: r-release: mvMORPH_1.0.2.tgz
Old sources: mvMORPH archive

Laboratoire de Géologie de Lyon
Terre Planètes Environnement

ENS DE LYON