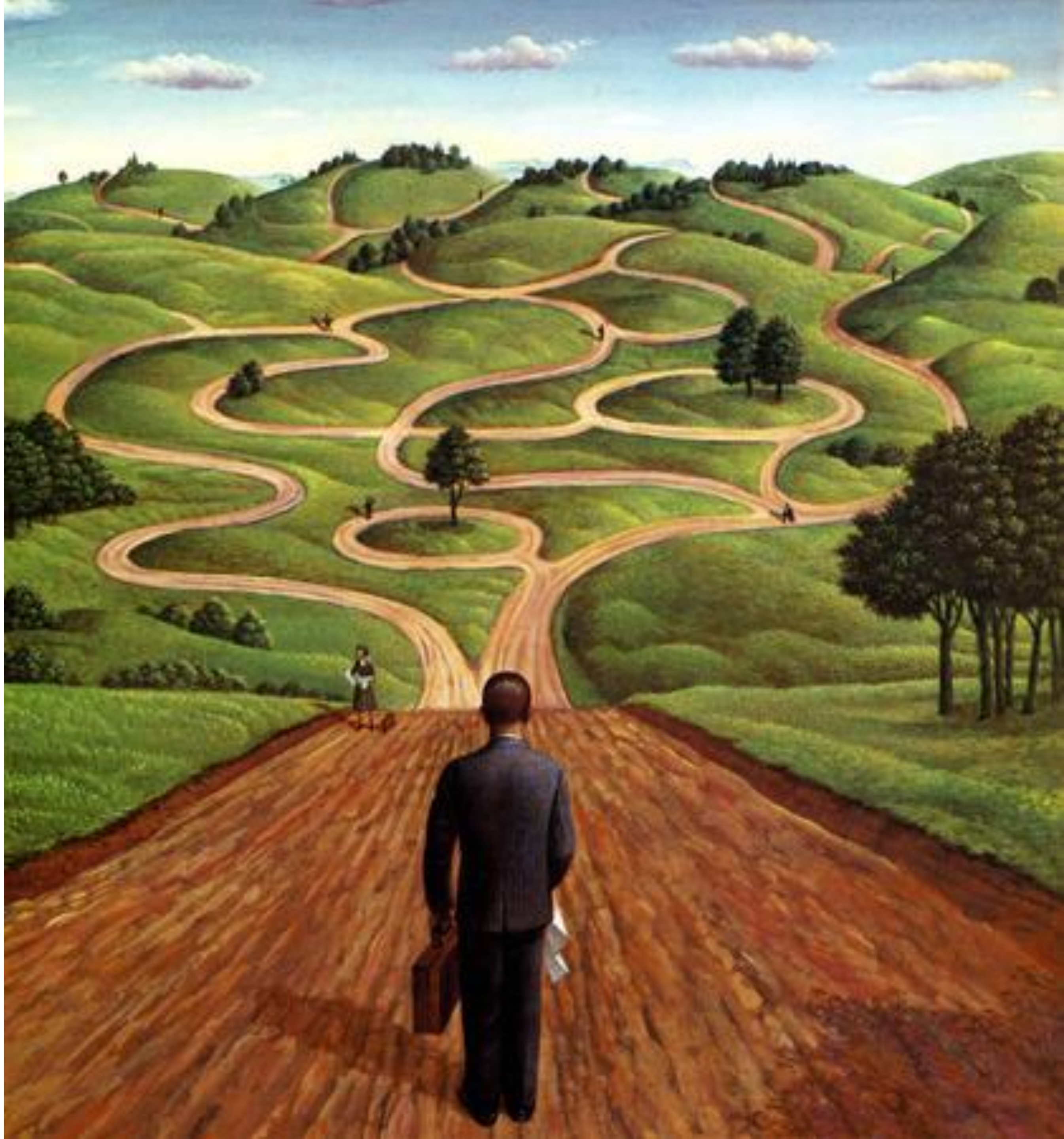
Paths through morphospace

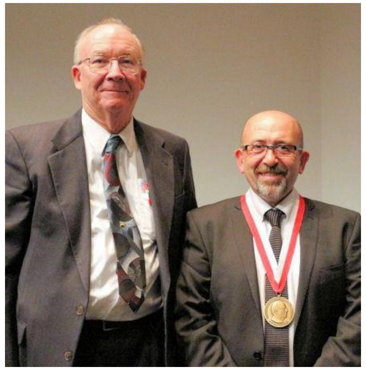
Morphometrics, phylogenetics, and evolution

P. David Polly
Indiana University
pdpolly@pollylab.org
https://pollylab.org/

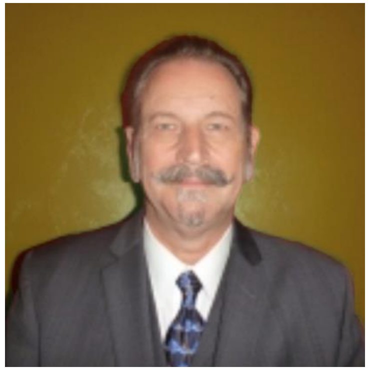


Previous Rohlf Medalists


Bookstein


O'Higgins


Hallgrimsson

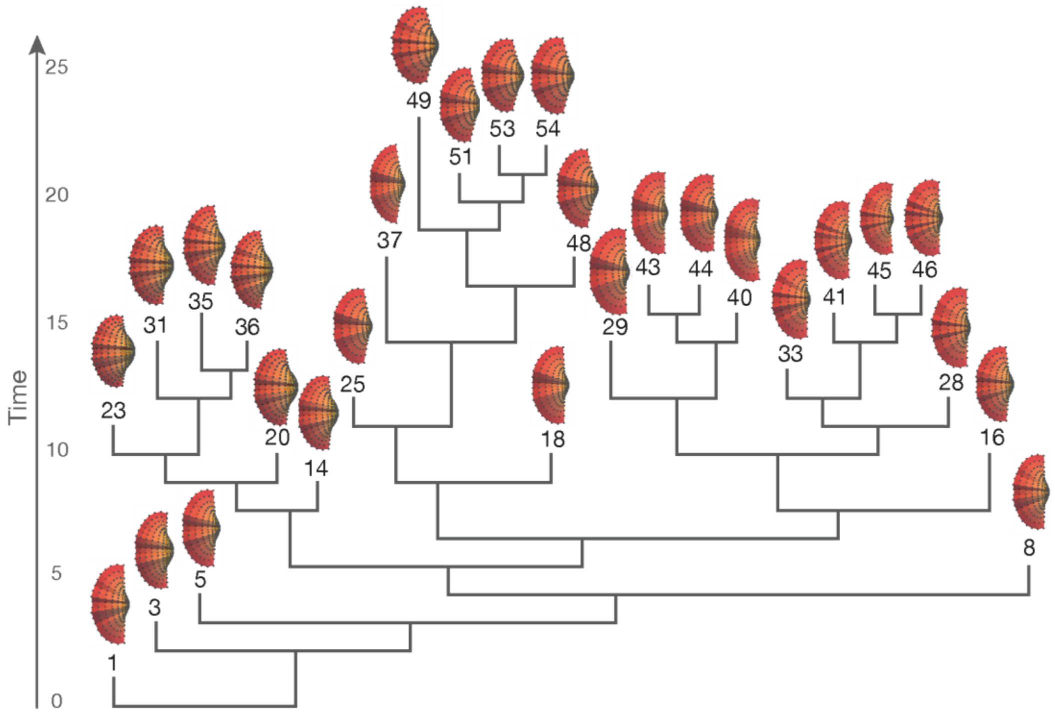

Slice

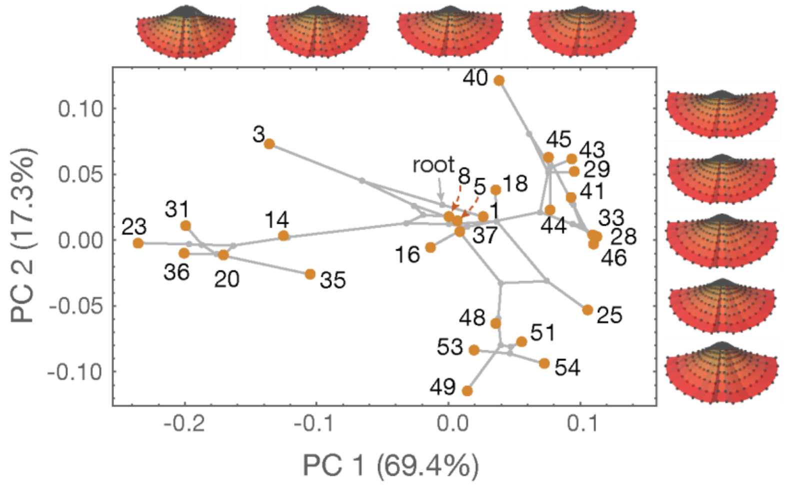

Adams


Richtsmeier


Mitteröcker

Phylogeny and morphometrics





Polly PD & Motz GJ, 2017, *Paleontological Society Papers*, 22: 71-99.

Is it always possible to estimate…

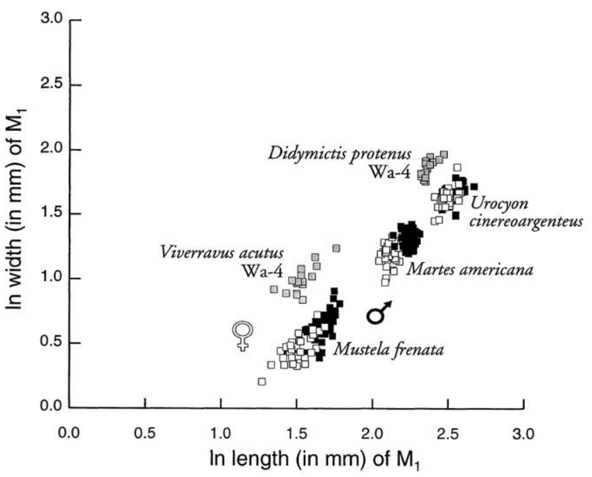
Thomas Hart Benton, *High Plains* (1953)

- evolutionary transitions?
- ancestor shape reconstructions?
- phenotypic change on a phylogeny?
- rates or modes of evolution?
- adaptive landscapes?

In what ways do evolutionary paths map onto mathematical paths?

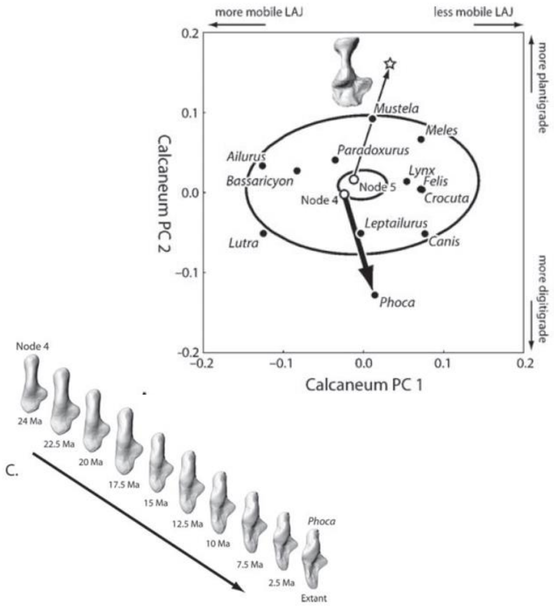
We estimate paths in nearly every morphometric study

Taxonomic differences



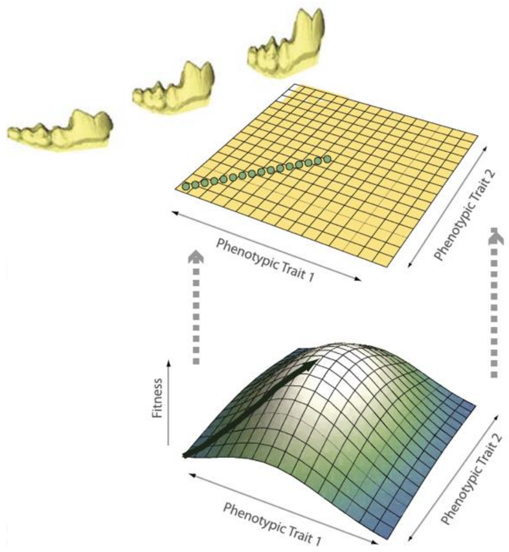
Polly PD, 1998. *Contr. Mus. Paleo. U. Mich.*, 30: 1-53.

Evolutionary trajectories



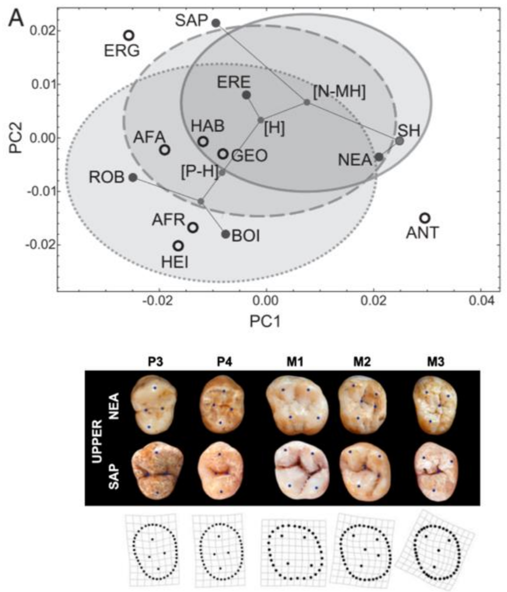
Polly PD, 2008. *Mammalian Evolutionary Morphology*, 167-198.

Modeling and simulation



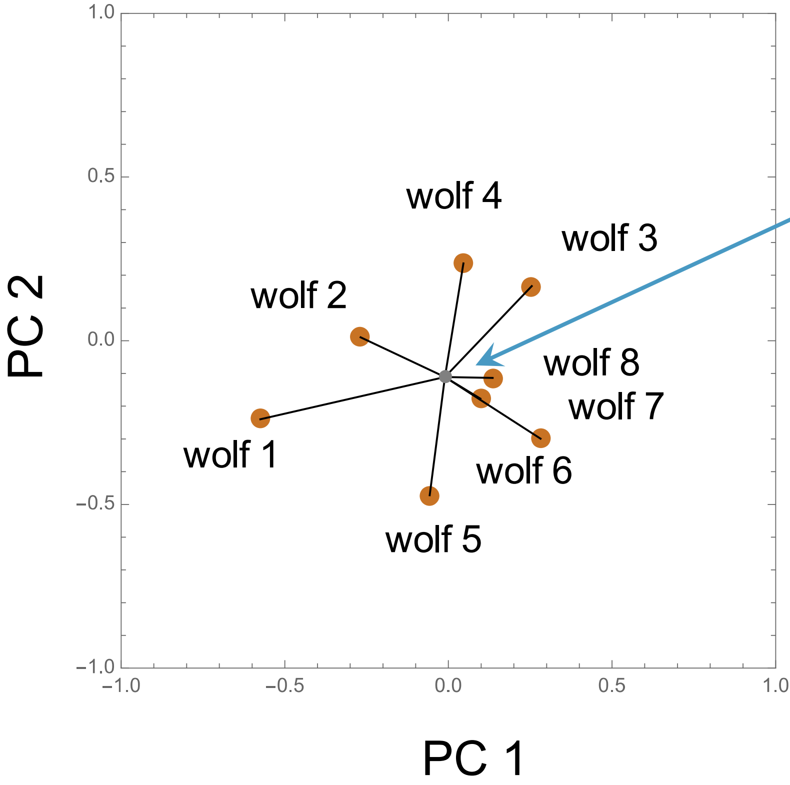
Polly PD, 2008. *Evol. Biol.*, 35: 85-96.

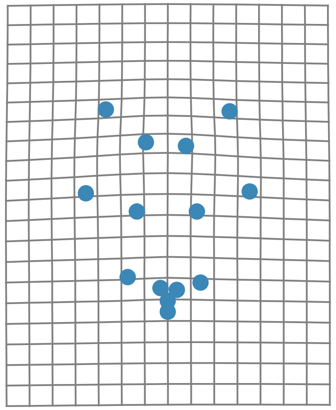
Ancestor reconstruction



Gomez-Robles et al. 2013. *PNAS*, 110: 18196-18201.

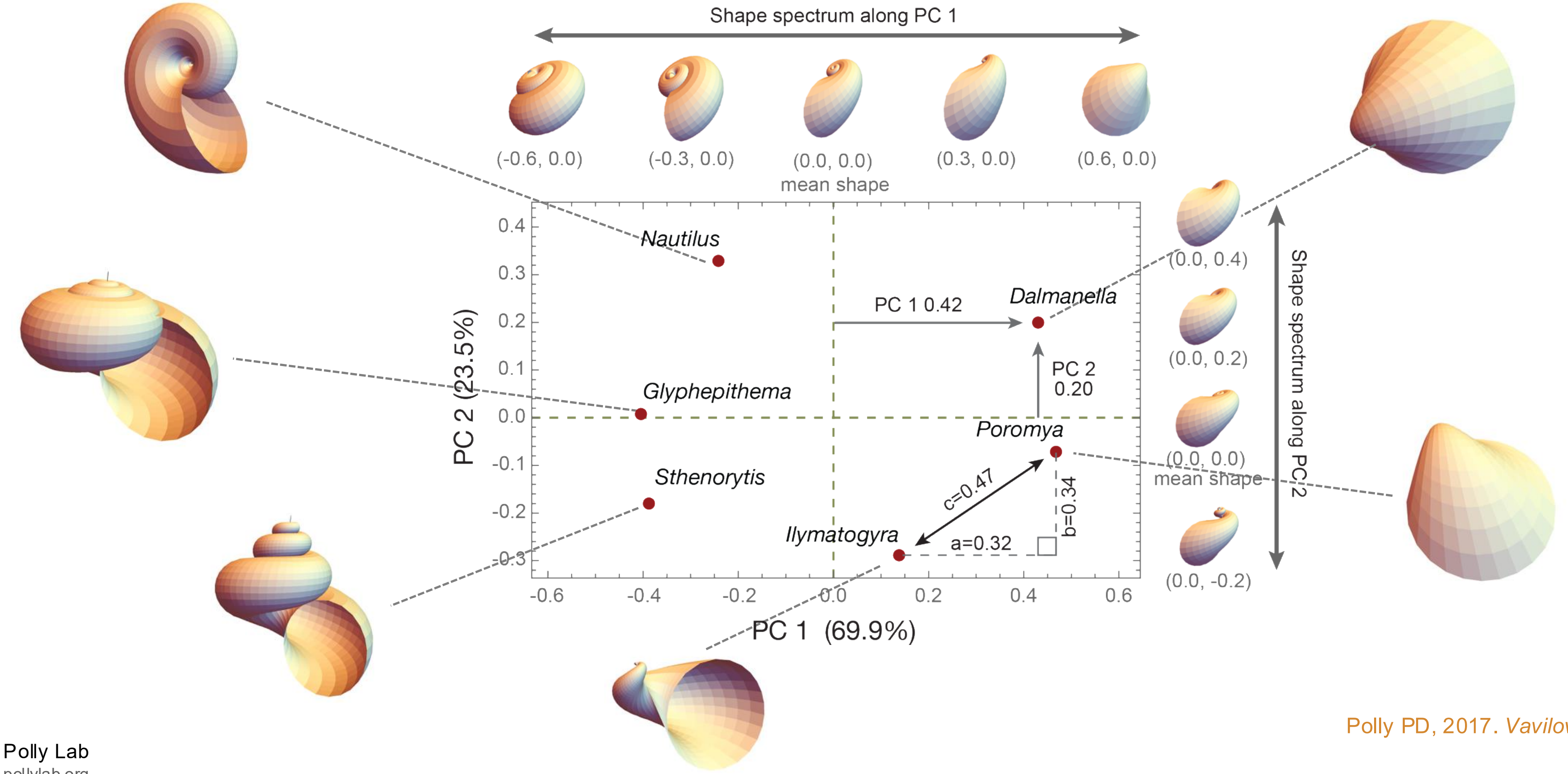
Even a population means involves paths…





Mean wolf

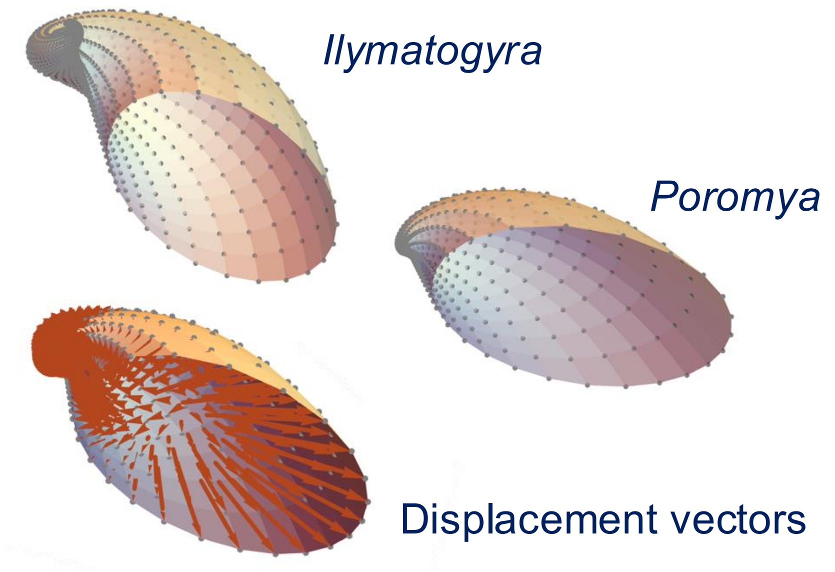
A typical geometric morphometric morphospace…



Polly PD, 2017. *Vavilov Journal* **21:** 452-461

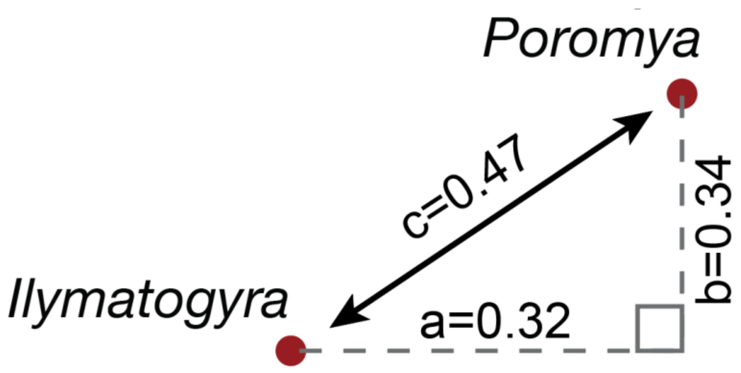
Distances between objects are preserved

Landmark space



=

PCA space



Procrustes distance = sum of displacements between homologous landmarks

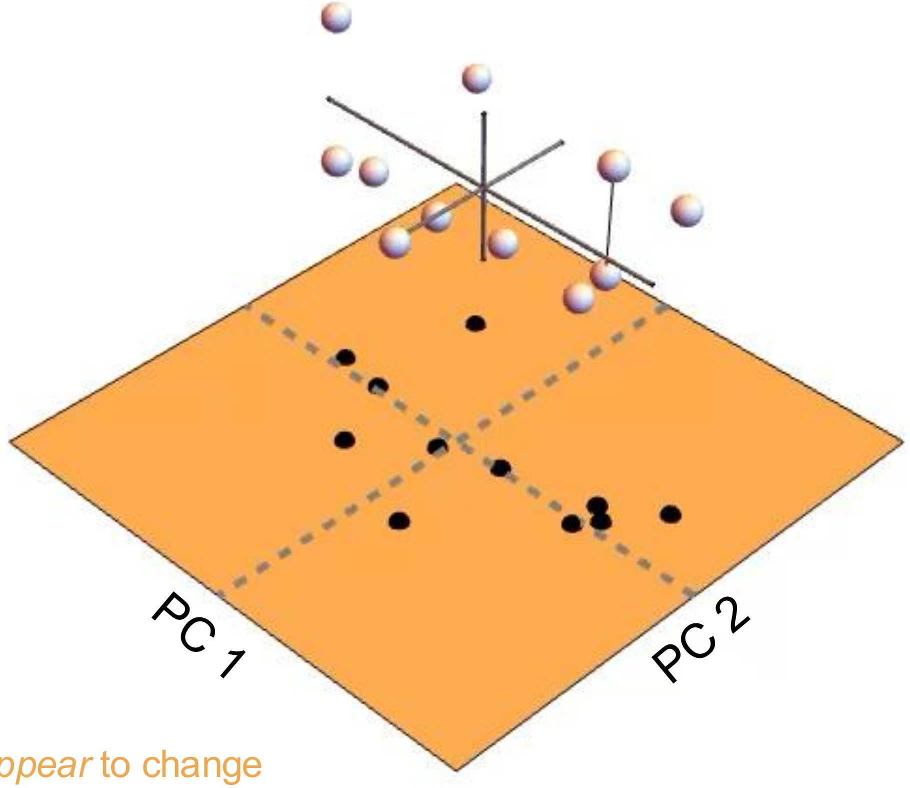
Procrustes distance = full multivariate distance between objects in morphospace

Distances between objects are preserved

Geometric morphometric morphospaces are strongly multivariate

- distances between objects in the full space are constant
- PC axes ***are*** sample dependent
- operations on a small number of PCs ***are*** sample dependent
- patterns off interest may not be represented on first PC axes
- operations carried out in full multivariate morphospace ***are not*** sample-dependent

Relationships are constant in multidimensional space

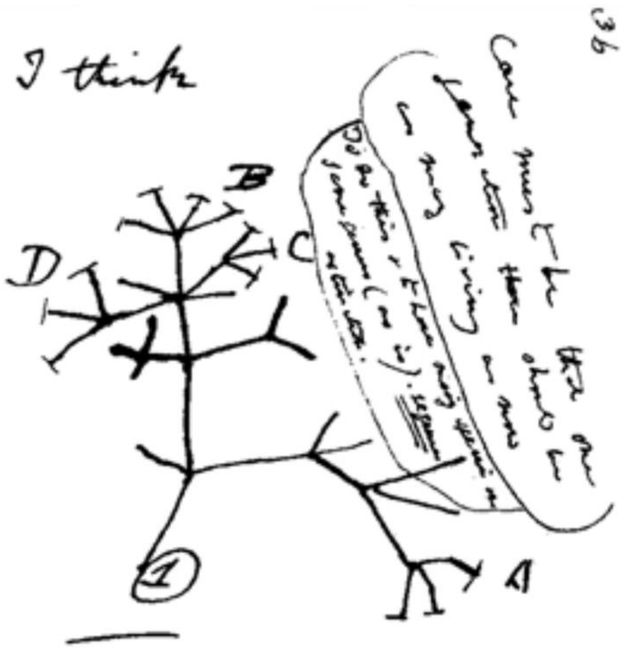


Addition of new specimen simply changes coordinate system, not the relationship between shapes

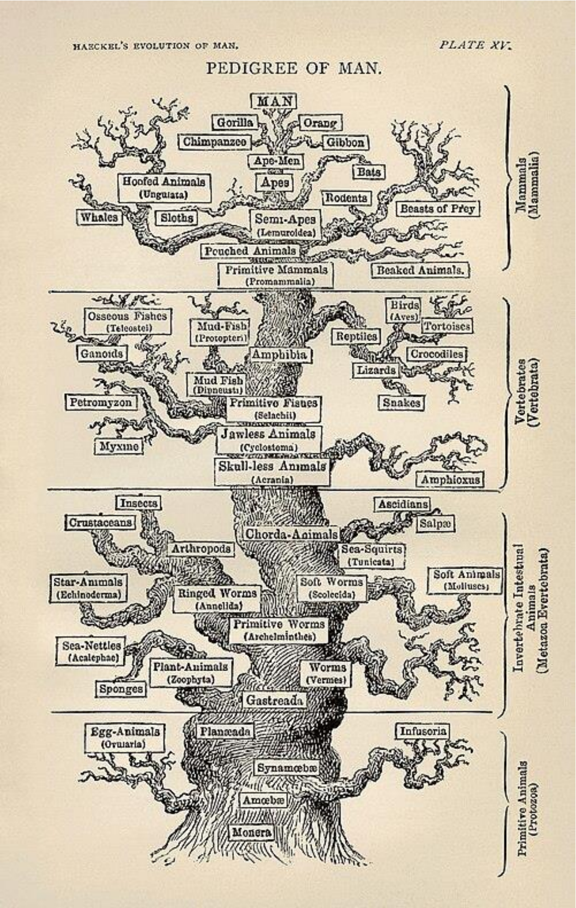
Relationships *appear* to change on two-dimensional projection

Phylogenetics

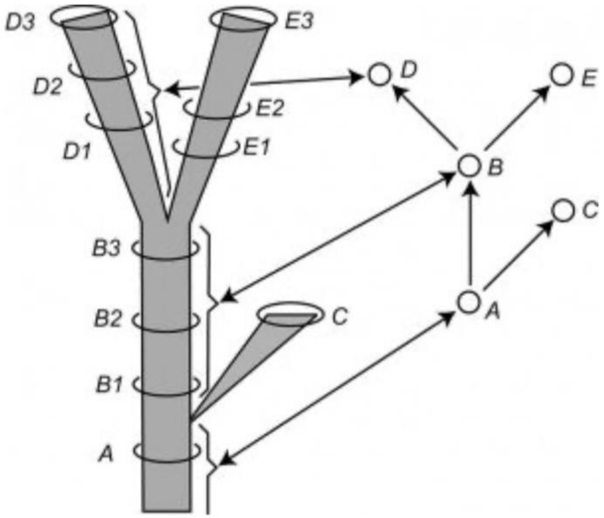
(and its interaction with high dimensional spaces)



Darwin, 1838



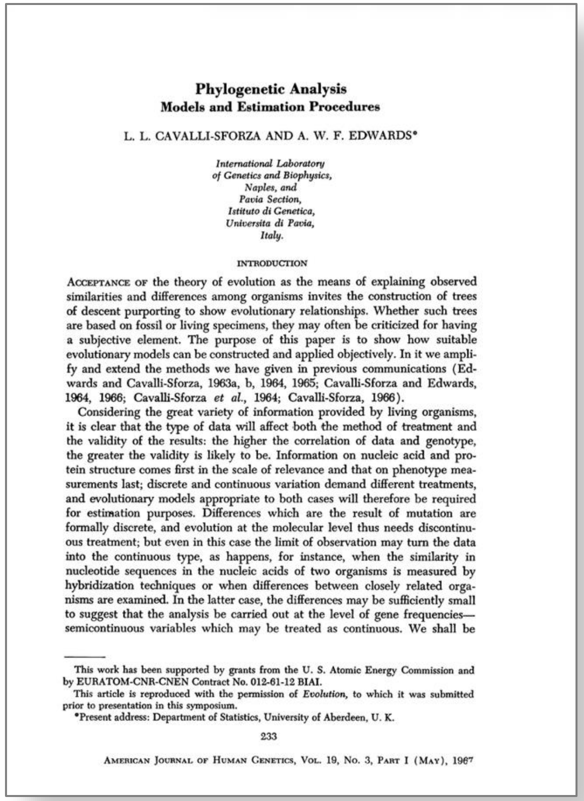
Haeckel, 1879



Hennig, 1966

Computational phylogeny reconstruction for morphology

Gene frequencies & likelihood
1965



Phylogenetic Analysis
Models and Estimation Procedures

L. L. CAVALLI-SFORZA AND A. W. F. EDWARDS*

International Laboratory of Genetics and Biophysics, Naples, and Pavia Section, Istituto di Genetica, Universita di Pavia, Italy.

INTRODUCTION

ACCEPTANCE OF the theory of evolution as the means of explaining observed similarities and differences among organisms invites the construction of trees of descent purporting to show evolutionary relationships. Whether such trees are based on fossil or living specimens, they may often be criticized for having a subjective element. The purpose of this paper is to show how suitable evolutionary models can be constructed and applied objectively. In it we amplify and extend the methods we have given in previous communications (Edwards and Cavalli-Sforza, 1963a, b, 1964, 1965; Cavalli-Sforza and Edwards, 1964, 1966; Cavalli-Sforza *et al.*, 1964; Cavalli-Sforza, 1966).

Considering the great variety of information provided by living organisms, it is clear that the type of data will affect both the method of treatment and the validity of the results: the higher the correlation of data and genotype, the greater the validity is likely to be. Information on nucleic acid and protein structure comes first in the scale of relevance and that on phenotype measurements last; discrete and continuous variation demand different treatments, and evolutionary models appropriate to both cases will therefore be required for estimation purposes. Differences which are the result of mutation are formally discrete, and evolution at the molecular level thus needs discontinuous treatment; but even in this case the limit of observation may turn the data into the continuous type, as happens, for instance, when the similarity in nucleotide sequences in the nucleic acids of two organisms is measured by hybridization techniques or when differences between closely related organisms are examined. In the latter case, the differences may be sufficiently small to suggest that the analysis be carried out at the level of gene frequencies—semicontinuous variables which may be treated as continuous. We shall be

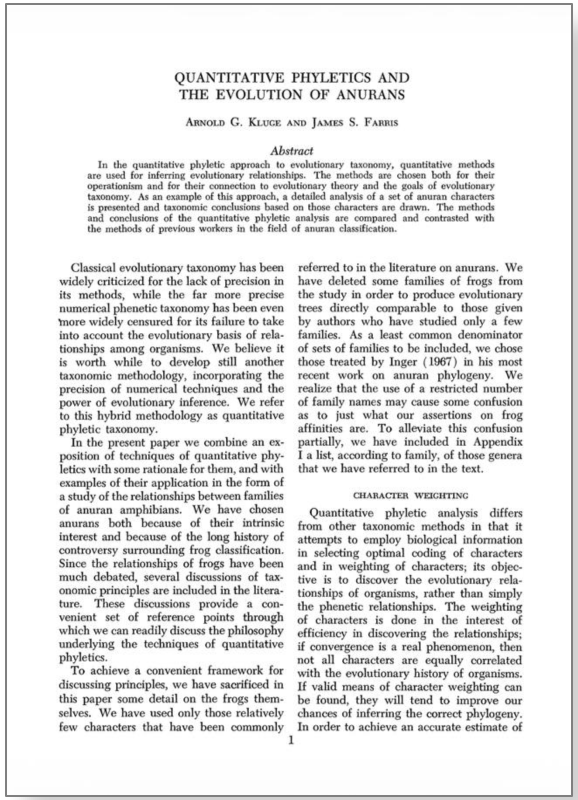
This work has been supported by grants from the U. S. Atomic Energy Commission and by EURATOM-CNR-CNEN Contract No. 012-61-12 BIAI.

This article is reproduced with the permission of *Evolution*, to which it was submitted prior to presentation in this symposium.

*Present address: Department of Statistics, University of Aberdeen, U. K.

Cavalli-Sforza LL, Edwards AW, 1967. Phylogenetic analysis. Models and estimation procedures. *Am J Human Genetics*, 19: 233-257.

Discrete traits & parsimony
1969



QUANTITATIVE PHYLETICS AND THE EVOLUTION OF ANURANS

ARNOLD G. KLUGE AND JAMES S. FARRIS

Abstract

In the quantitative phyletic approach to evolutionary taxonomy, quantitative methods are used for inferring evolutionary relationships. The methods are chosen both for their operationism and for their connection to evolutionary theory and the goals of evolutionary taxonomy. As an example of this approach, a detailed analysis of a set of anuran characters is presented and taxonomic conclusions based on those characters are drawn. The methods and conclusions of the quantitative phyletic analysis are compared and contrasted with the methods of previous workers in the field of anuran classification.

Classical evolutionary taxonomy has been widely criticized for the lack of precision in its methods, while the far more precise numerical phenetic taxonomy has been even more widely censured for its failure to take into account the evolutionary basis of relationships among organisms. We believe it is worth while to develop still another taxonomic methodology, incorporating the precision of numerical techniques and the power of evolutionary inference. We refer to this hybrid methodology as quantitative phyletic taxonomy.

In the present paper we combine an exposition of techniques of quantitative phyletics with some rationale for them, and with examples of their application in the form of a study of the relationships between families of anuran amphibians. We have chosen anurans both because of their intrinsic interest and because of the long history of controversy surrounding frog classification. Since the relationships of frogs have been much debated, several discussions of taxonomic principles are included in the literature. These discussions provide a convenient set of reference points through which we can readily discuss the philosophy underlying the techniques of quantitative phyletics.

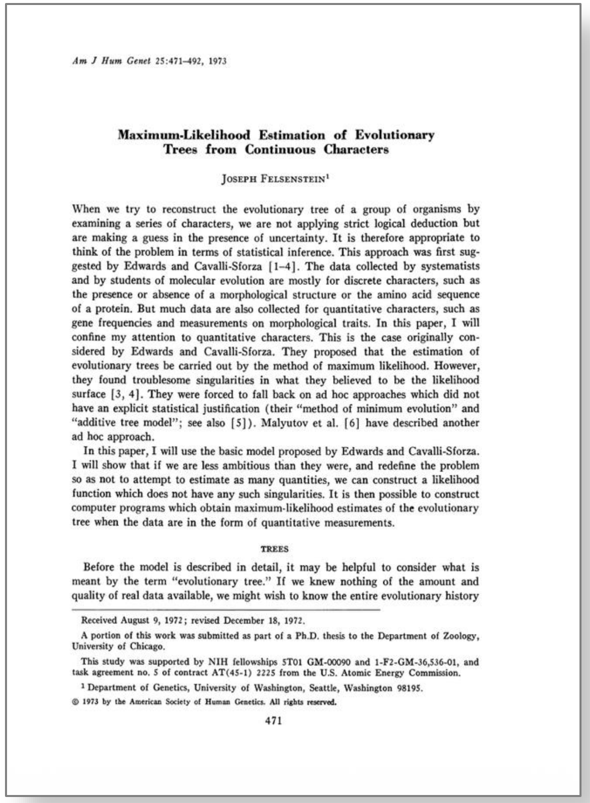
To achieve a convenient framework for discussing principles, we have sacrificed in this paper some detail on the frogs themselves. We have used only those relatively few characters that have been commonly referred to in the literature on anurans. We have deleted some families of frogs from the study in order to produce evolutionary trees directly comparable to those given by authors who have studied only a few families. As a least common denominator of sets of families to be included, we chose those treated by Inger (1967) in his most recent work on anuran phylogeny. We realize that the use of a restricted number of family names may cause some confusion as to just what our assertions on frog affinities are. To alleviate this confusion partially, we have included in Appendix I a list, according to family, of those genera that we have referred to in the text.

CHARACTER WEIGHTING

Quantitative phyletic analysis differs from other taxonomic methods in that it attempts to employ biological information in selecting optimal coding of characters and in weighting of characters; its objective is to discover the evolutionary relationships of organisms, rather than simply the phenetic relationships. The weighting of characters is done in the interest of efficiency in discovering the relationships; if convergence is a real phenomenon, then not all characters are equally correlated with the evolutionary history of organisms. If valid means of character weighting can be found, they will tend to improve our chances of inferring the correct phylogeny. In order to achieve an accurate estimate of

Kluge AG, Farris JS. 1969. Quantitative phyletics and the evolution of anurans. *Systematic Zooology*: 1-32.

Continuous traits & likelihood
1973



Am J Hum Genet 25:471–492, 1973

Maximum-Likelihood Estimation of Evolutionary Trees from Continuous Characters

JOSEPH FELSENSTEIN[1]

When we try to reconstruct the evolutionary tree of a group of organisms by examining a series of characters, we are not applying strict logical deduction but are making a guess in the presence of uncertainty. It is therefore appropriate to think of the problem in terms of statistical inference. This approach was first suggested by Edwards and Cavalli-Sforza [1–4]. The data collected by systematists and by students of molecular evolution are mostly for discrete characters, such as the presence or absence of a morphological structure or the amino acid sequence of a protein. But much data are also collected for quantitative characters, such as gene frequencies and measurements on morphological traits. In this paper, I will confine my attention to quantitative characters. This is the case originally considered by Edwards and Cavalli-Sforza. They proposed that the estimation of evolutionary trees be carried out by the method of maximum likelihood. However, they found troublesome singularities in what they believed to be the likelihood surface [3, 4]. They were forced to fall back on ad hoc approaches which did not have an explicit statistical justification (their "method of minimum evolution" and "additive tree model"; see also [5]). Malyutov et al. [6] have described another ad hoc approach.

In this paper, I will use the basic model proposed by Edwards and Cavalli-Sforza. I will show that if we are less ambitious than they were, and redefine the problem so as not to attempt to estimate as many quantities, we can construct a likelihood function which does not have any such singularities. It is then possible to construct computer programs which obtain maximum-likelihood estimates of the evolutionary tree when the data are in the form of quantitative measurements.

TREES

Before the model is described in detail, it may be helpful to consider what is meant by the term "evolutionary tree." If we knew nothing of the amount and quality of real data available, we might wish to know the entire evolutionary history

Received August 9, 1972; revised December 18, 1972.

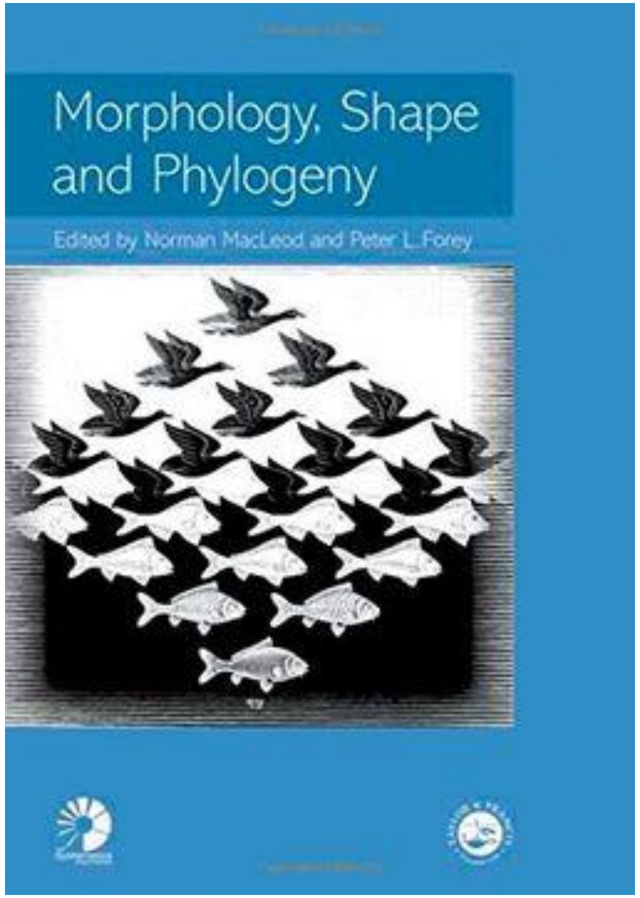
A portion of this work was submitted as part of a Ph.D. thesis to the Department of Zoology, University of Chicago.

This study was supported by NIH fellowships 5T01 GM-00090 and 1-F2-GM-36,536-01, and task agreement no. 5 of contract AT(45-1) 2225 from the U.S. Atomic Energy Commission.

[1] Department of Genetics, University of Washington, Seattle, Washington 98195.

Felsenstein J. 1973. Maximum-likelihood estimation of evolutionary trees from continuous characters. *Am J Human Genetics*, 25: 471-492.

Integrating phylogenetics and geometric morphometrics



2002

Felsenstein Polly
Rohlf Cole Pagel Swiderski Rae Reid Purvis

Webster Humphries Forey MacLeod Bookstein Sidwell

Systematics Association Annual Meeting
August 1999
Glasgow, Scotland

Phylomorphospace

The first(?) of geometric morphometrics and statistical phylogenetics

Chapter 9

Geometric morphometrics and phylogeny

F. James Rohlf

ABSTRACT

This chapter reviews some of the important properties of geometric morphometric shape variables and discusses the advantages and limitations of the use of such data in studies of phylogeny. A method for fitting morphometric data to a phylogeny (i.e., estimating ancestral states of the shape variables) is presented using the squared-change parsimony estimation criterion. These results are then used to illustrate shape change along a phylogeny as a deformation of the shape of any other node on the tree (e.g., the estimated root of the tree). In addition, a method to estimate the digitized image of an ancestor is given that uses averages of unwarped images. An example dataset with 18 wing landmarks for 11 species of mosquitoes is used to illustrate the methods.

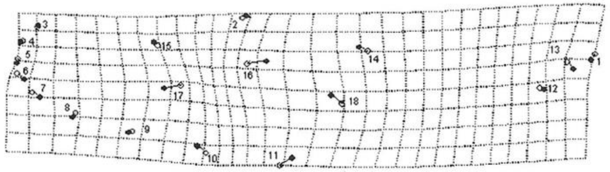


Figure 9.3 Visualization of the estimated shape for ancestor of the Aedini (*Aedes*, *Psorphora*, and *Mansonia*, see Figure 9.2). The grid shows the thin-plate spline transformation from the starting form (●) to the estimated configuration (○). The estimated shape at the root of the tree was used as the starting form.

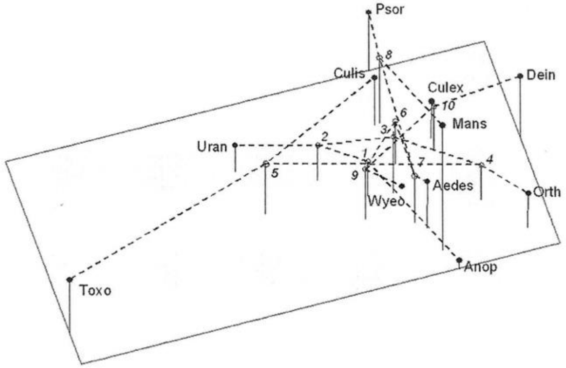


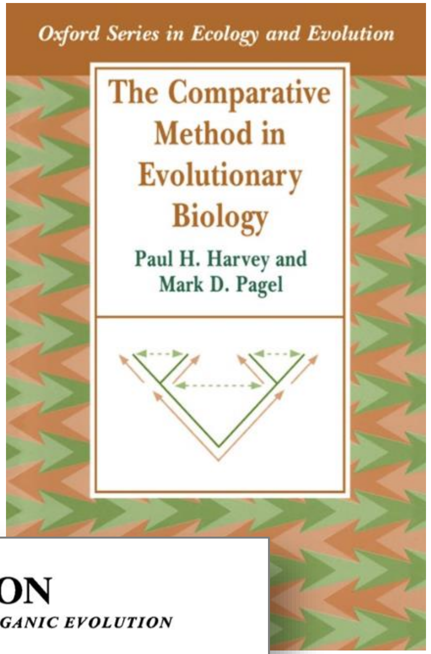
Figure 9.4 Ordination from a non-metric multidimensional scaling analysis of a matrix of distances between all species (●) and estimated internal nodes (○). Phylogenetic tree from Figure 9.2 superimposed using broken lines to link points. OTU codes are given in Table 9.1. Internal nodes numbered in a preorder traversal of the tree beginning with the root. Stress = 0.145. matrix correlation = 0.978.

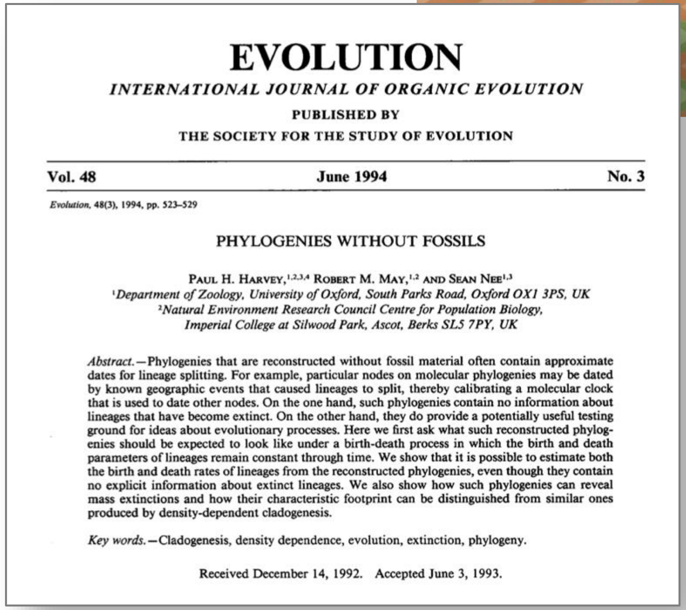
Rohlf FJ, 2002, Geometric morphometrics and Phylogeny, pp. 175-193 in *Morphology, Shape, and Phylogenetics.*

Statistical phylogenetics

Probabilistic, likelihood, and Bayesian models for relationships between characters and trees

- phylogeny reconstruction
- ancestral state (shape) reconstruction
- estimation of evolutionary rates
- model selection for evolutionary “modes”
- hypothesis testing for rate and “mode” shifts
- testing for evolutionary convergence
- testing for multiple adaptive peaks
- performance trade-off modeling





EVOLUTION
INTERNATIONAL JOURNAL OF ORGANIC EVOLUTION
PUBLISHED BY
THE SOCIETY FOR THE STUDY OF EVOLUTION

Vol. 48 June 1994 No. 3

Evolution, 48(3), 1994, pp. 523–529

PHYLOGENIES WITHOUT FOSSILS

Paul H. Harvey,[1,2,3,4] Robert M. May,[1,2] and Sean Nee[1,3]
[1]*Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, UK*
[2]*Natural Environment Research Council Centre for Population Biology, Imperial College at Silwood Park, Ascot, Berks SL5 7PY, UK*

Abstract.—Phylogenies that are reconstructed without fossil material often contain approximate dates for lineage splitting. For example, particular nodes on molecular phylogenies may be dated by known geographic events that caused lineages to split, thereby calibrating a molecular clock that is used to date other nodes. On the one hand, such phylogenies contain no information about lineages that have become extinct. On the other hand, they do provide a potentially useful testing ground for ideas about evolutionary processes. Here we first ask what such reconstructed phylogenies should be expected to look like under a birth-death process in which the birth and death parameters of lineages remain constant through time. We show that it is possible to estimate both the birth and death rates of lineages from the reconstructed phylogenies, even though they contain no explicit information about extinct lineages. We also show how such phylogenies can reveal mass extinctions and how their characteristic footprint can be distinguished from similar ones produced by density-dependent cladogenesis.

Key words.—Cladogenesis, density dependence, evolution, extinction, phylogeny.

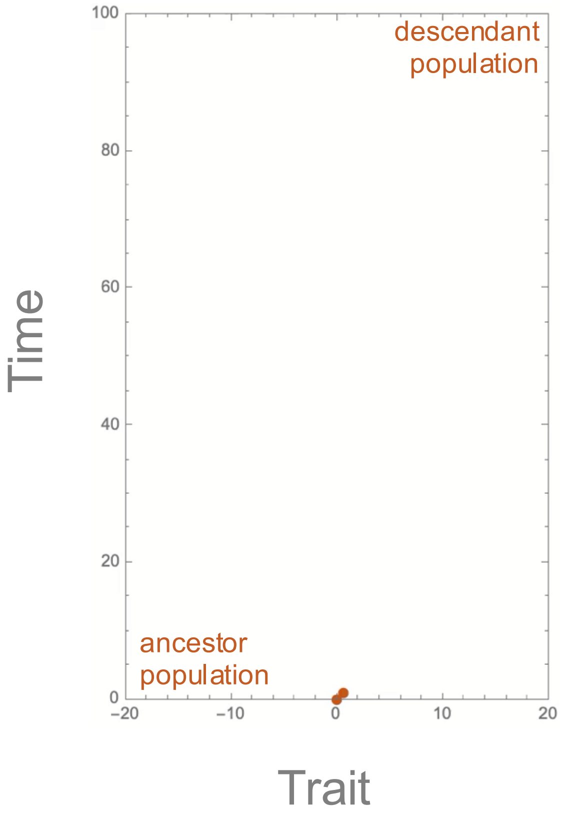
Received December 14, 1992. Accepted June 3, 1993.

Brownian motion is the null model…

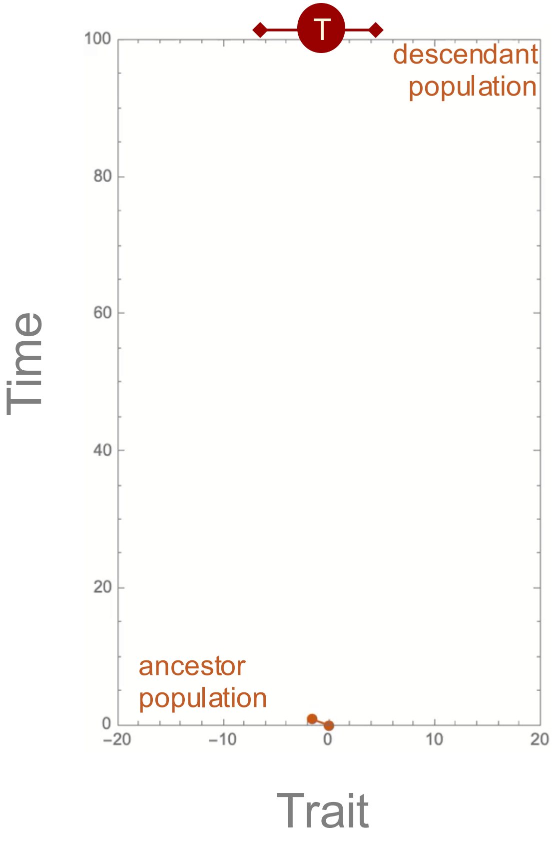
…most other models are juxtaposed to it

evolution at each step is random in direction and magnitude

Brownian motion
(genetic or selectional drift)
in a single lineage



Ornstein-Uhlenbeck
(stabilizing selection)
in a single lineage



evolution at each step depends on distance from target *T*

Nearly all phylogenetic statistical are based on…

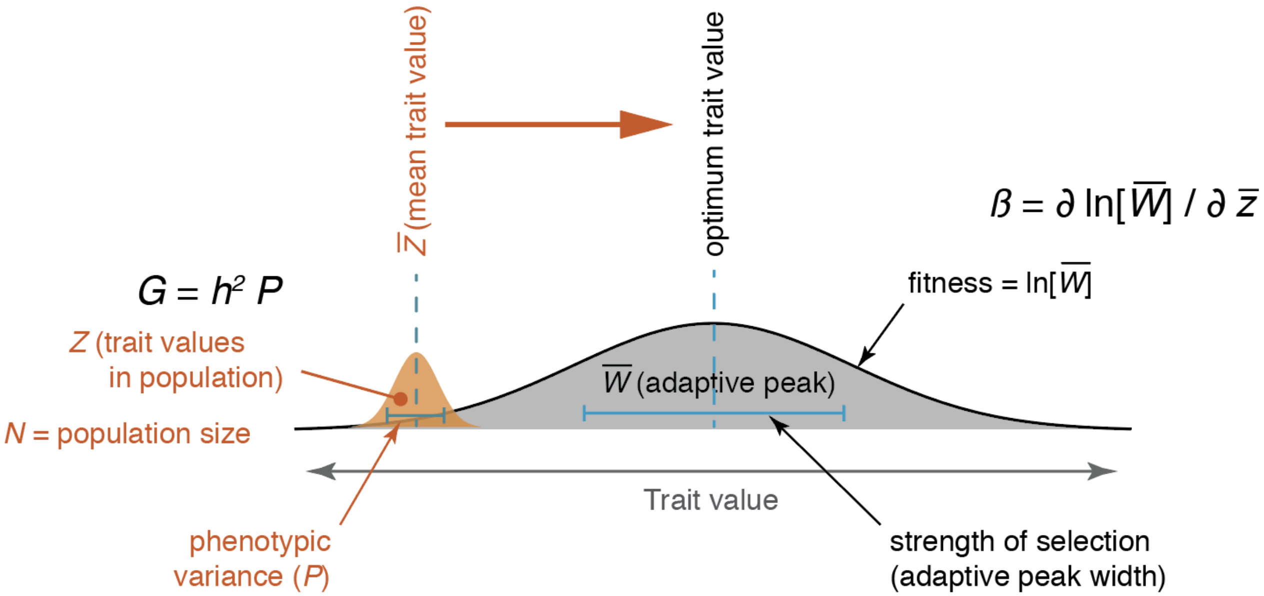
Lande’s adaptive peak model for quantitative traits

change in mean trait value

selection component

drift component

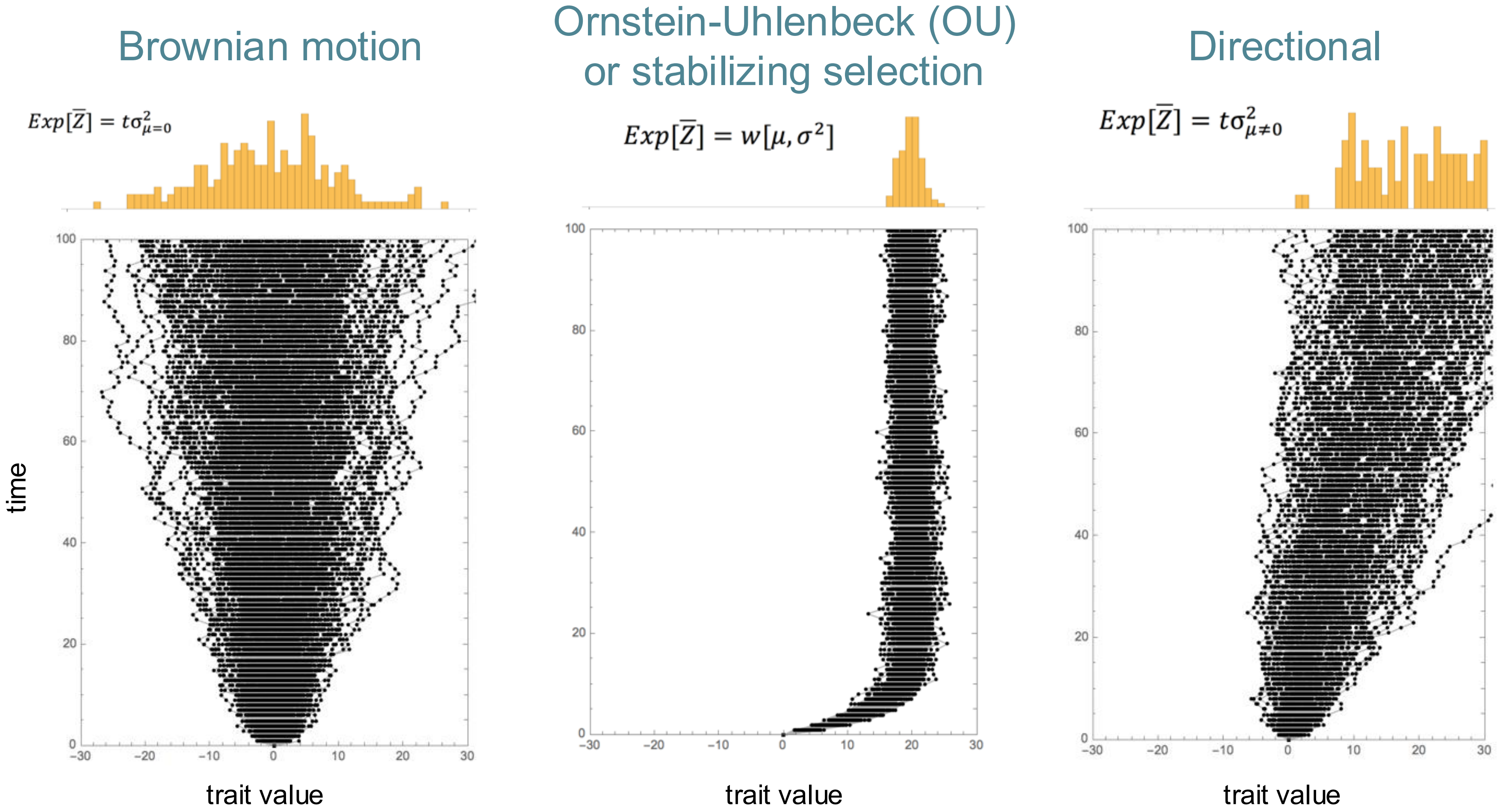
$$\Delta \bar{Z} = \beta G + G/N$$



Lande R, 1976. Natural selection and random genetic drift in phenotypic evolution. *Evolution*, 30: 314-334.

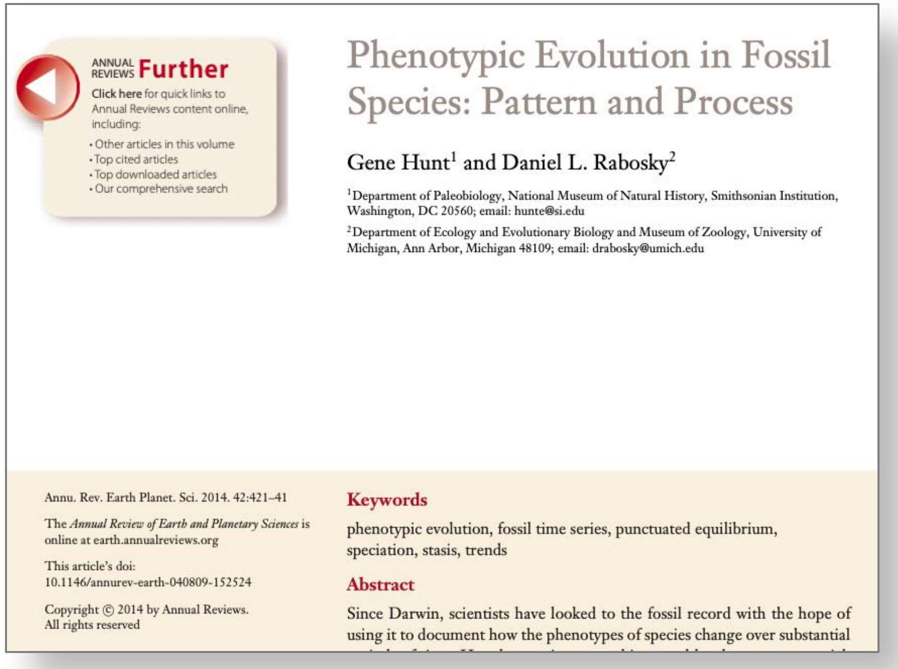
Likelihood of an evolutionary outcome depends on the model

Best model can be selected based on observed outcomes



Example: model selection in a paleontological lineage

Gene Hunt



ANNUAL REVIEWS Further

Click here for quick links to Annual Reviews content online, including:

- Other articles in this volume
- Top cited articles
- Top downloaded articles
- Our comprehensive search

Phenotypic Evolution in Fossil Species: Pattern and Process

Gene Hunt[1] and Daniel L. Rabosky[2]

[1]Department of Paleobiology, National Museum of Natural History, Smithsonian Institution, Washington, DC 20560; email: hunte@si.edu

[2]Department of Ecology and Evolutionary Biology and Museum of Zoology, University of Michigan, Ann Arbor, Michigan 48109; email: drabosky@umich.edu

Annu. Rev. Earth Planet. Sci. 2014. 42:421–41

The *Annual Review of Earth and Planetary Sciences* is online at earth.annualreviews.org

This article's doi: 10.1146/annurev-earth-040809-152524

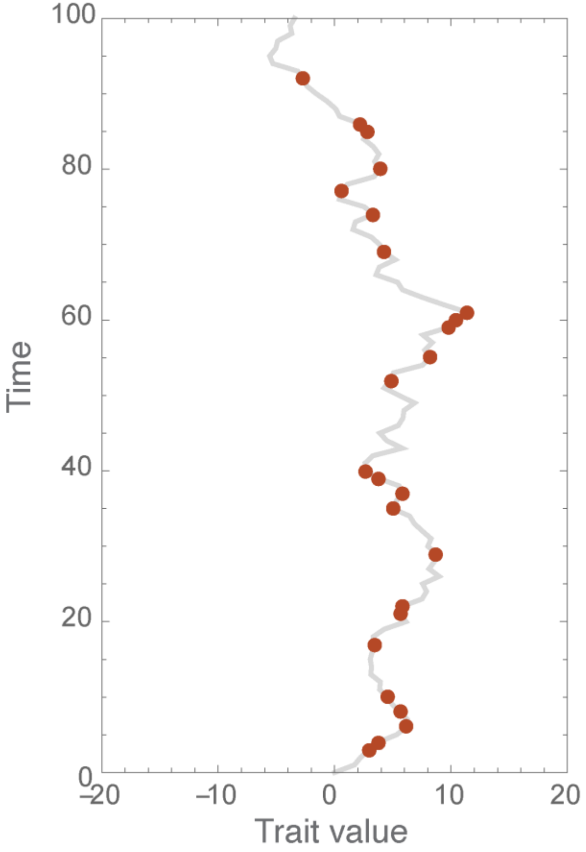
Keywords

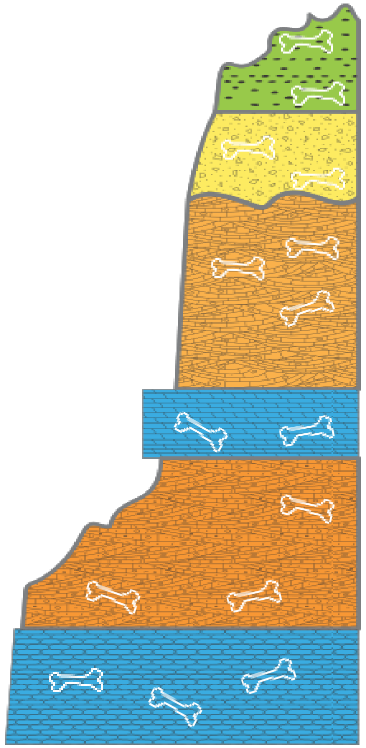
phenotypic evolution, fossil time series, punctuated equilibrium, speciation, stasis, trends

Abstract

Since Darwin, scientists have looked to the fossil record with the hope of using it to document how the phenotypes of species change over substantial

Best model	URW
N	39
$\hat{\mu}_{step}$	0.166
$\hat{\sigma}^2_{step}$	0.034
$\hat{\theta}$	1.384
$\hat{\omega}$	0.004
AIC_c URW	-124.270
AIC_c GRW	-122.860
AIC_c Stasis	-96.080
wt URW	0.670
wt GRW	0.330
wt Stasis	5.064×10^{-7}

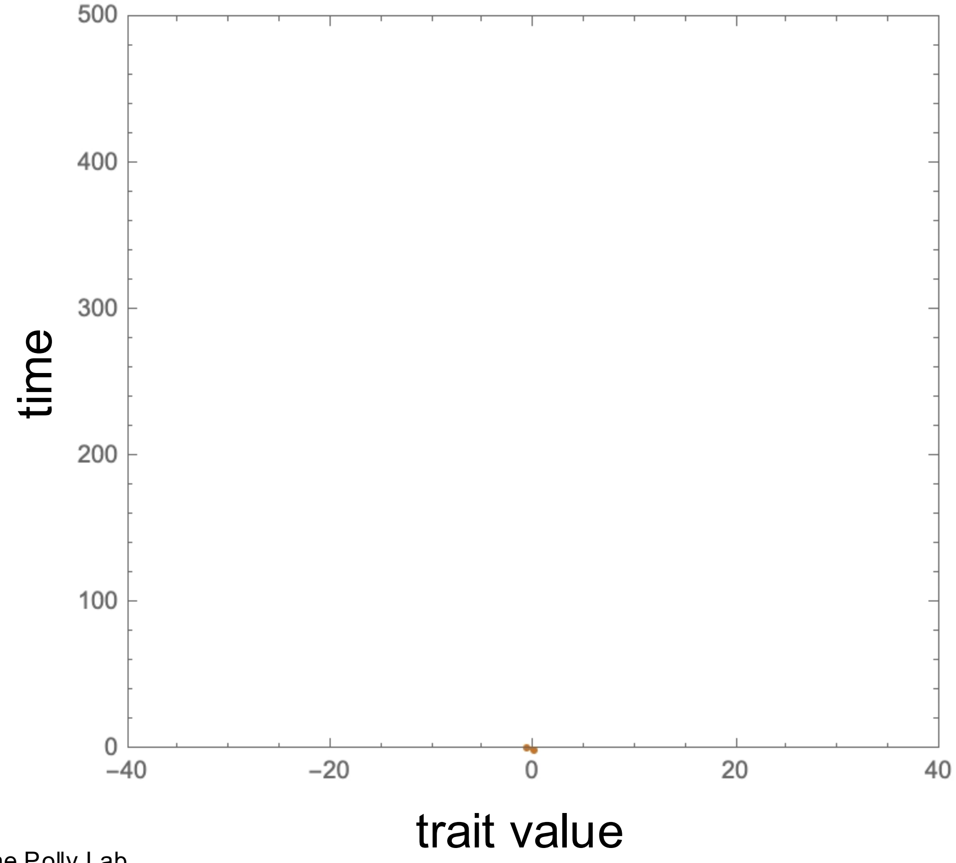




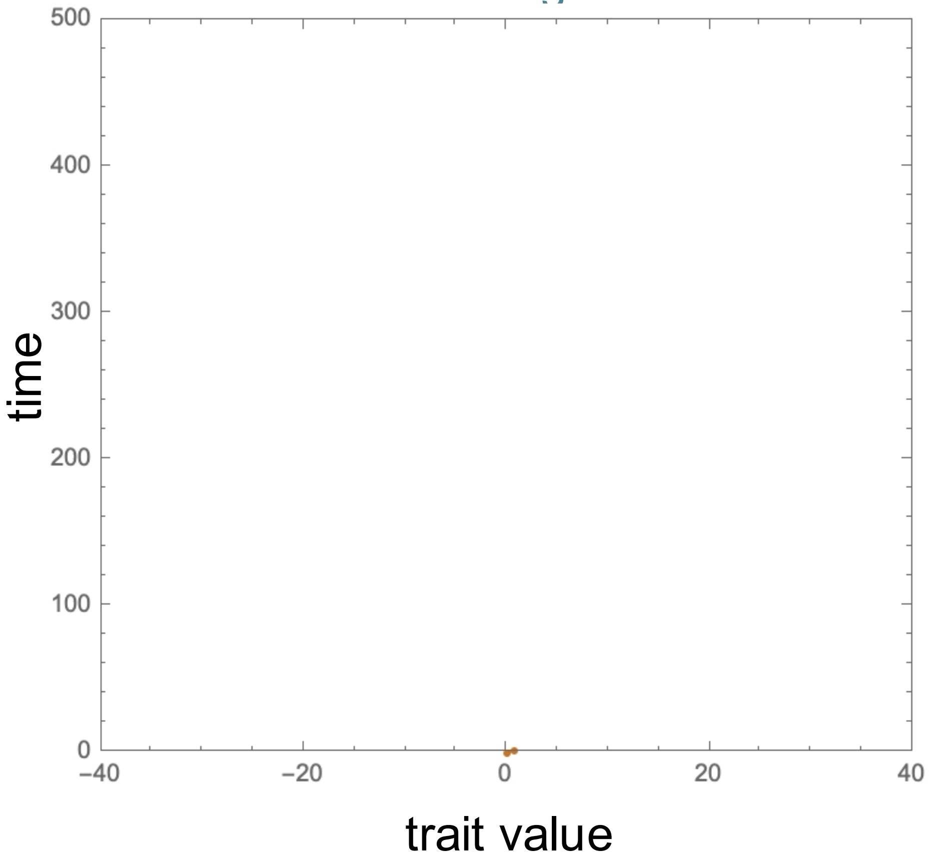
Statistical models of evolution on phylogenetic trees

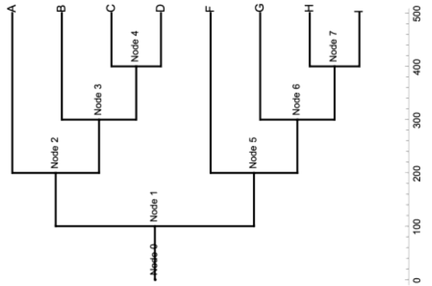
Same idea, but tip values covary by shared ancestry

Brownian motion



Ornstein-Uhlenbeck (OU)
or stabilizing selection

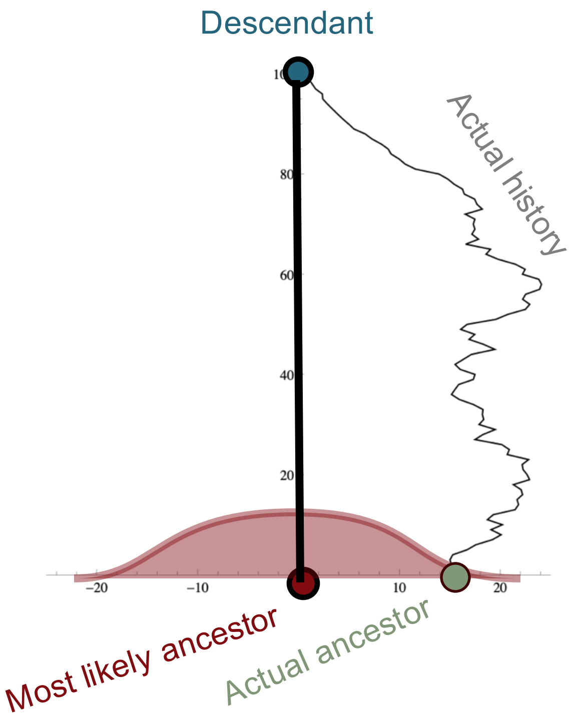




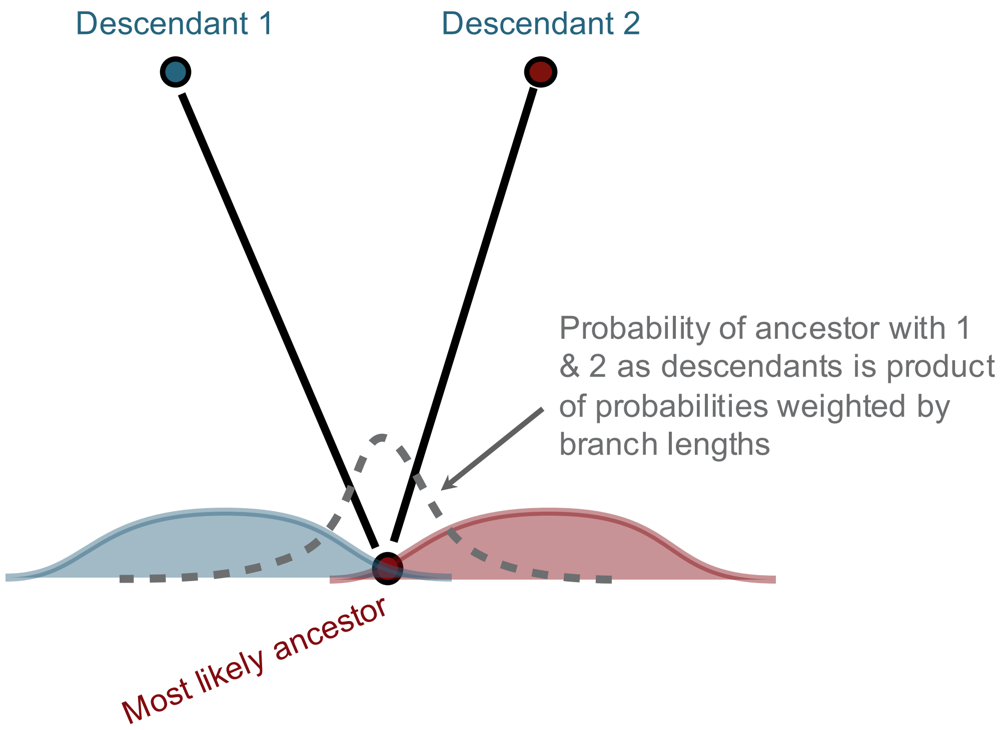
Ancestor reconstruction using a BM model

For one descendant

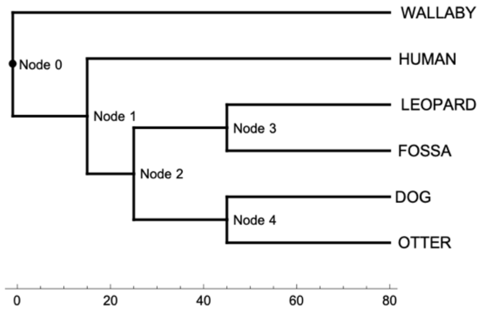
- For one tip, the most likely ancestor value is the same
- Variance in likelihood is proportional to time elapse
- Most likely path is a straight line

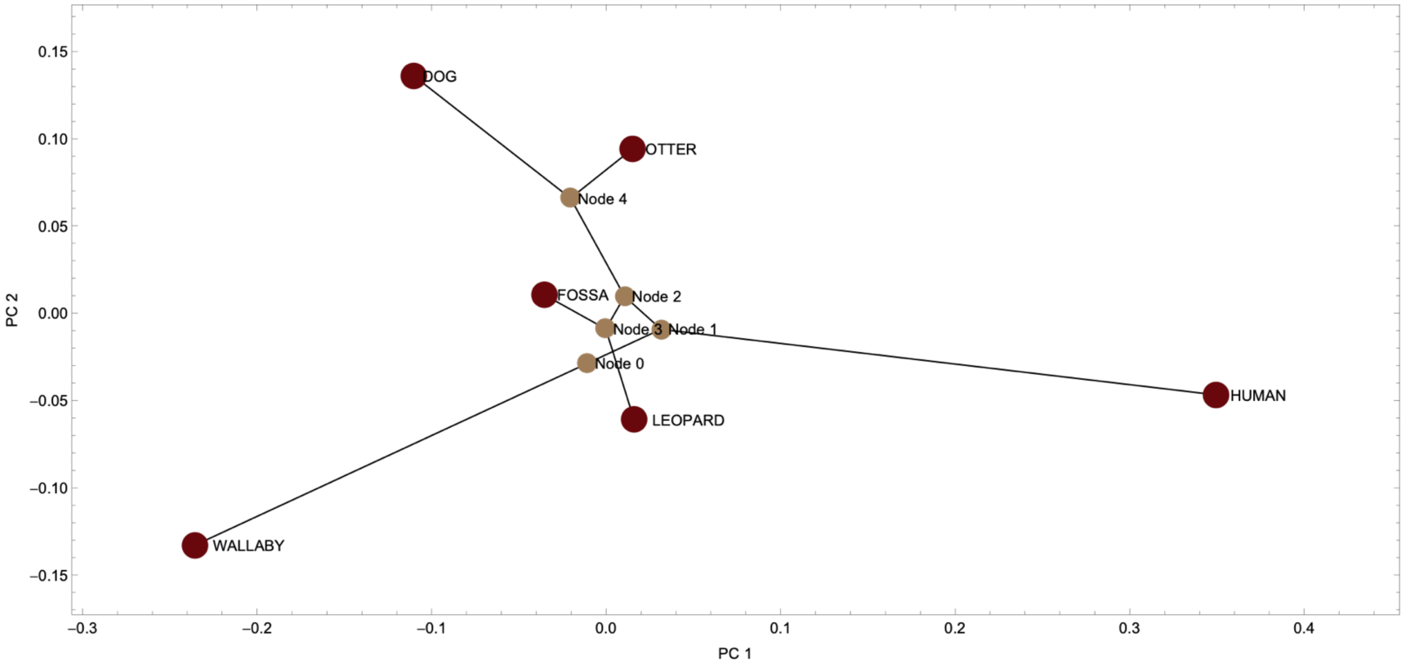


For two or more descendants



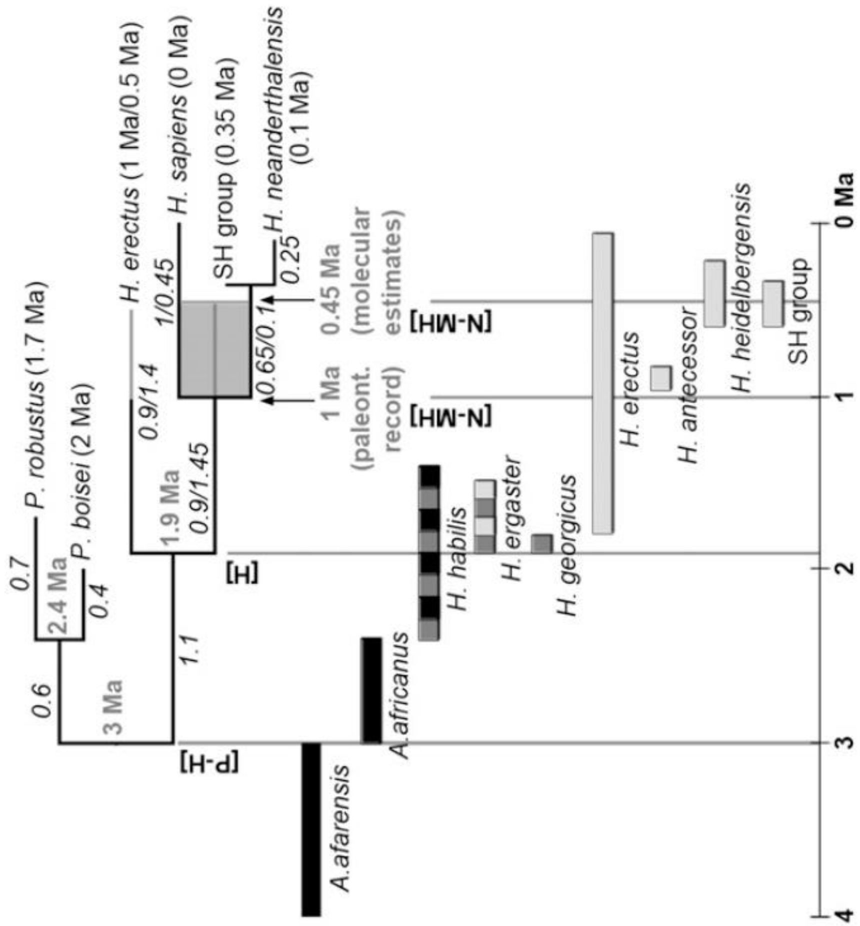
Phylomorphospaces are based on these models

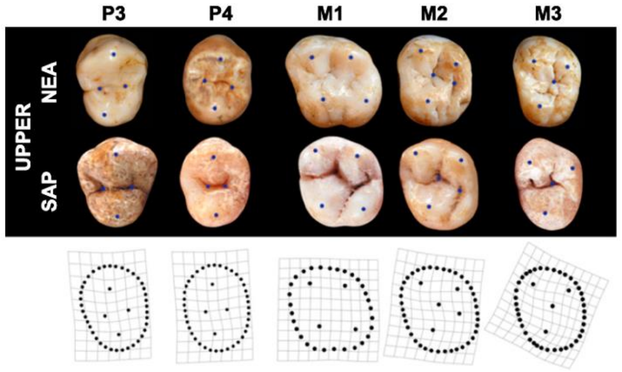


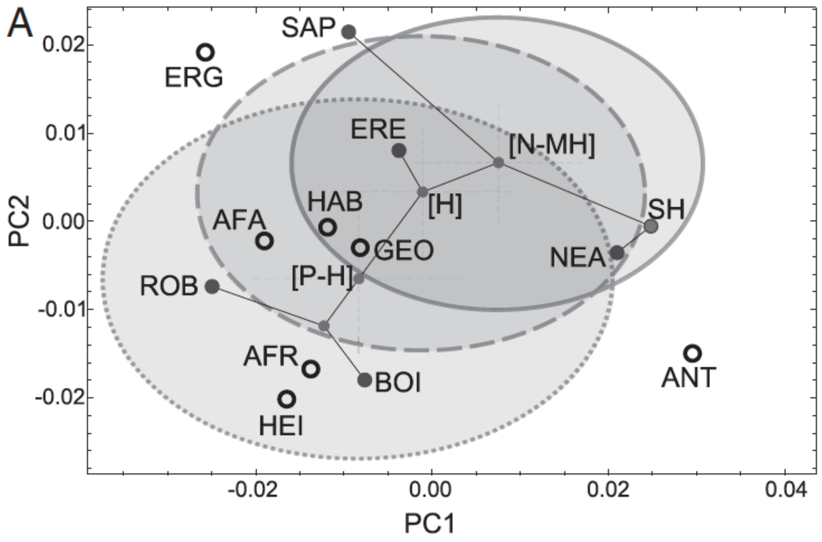


No known hominin species matches the expected dental morphology of the last common ancestor of Neanderthals and modern humans

Aida Gómez-Robles[a,b,1], José María Bermúdez de Castro[c], Juan-Luis Arsuaga[d], Eudald Carbonell[e], and P. David Polly[f]







Example: *ProbabilitiesOfShapesAsAncestors[proc, labels, tree]*

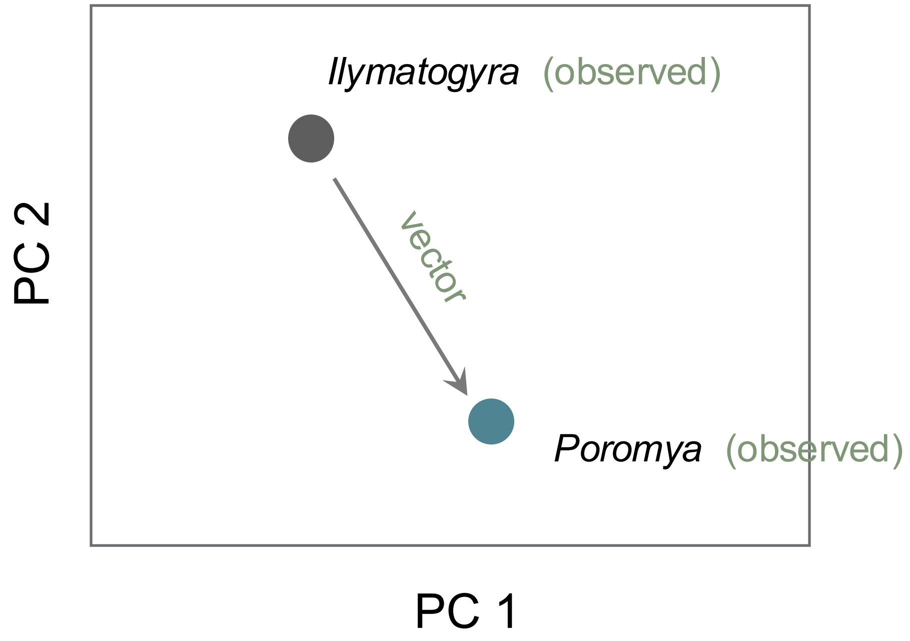
	Node 0	Node 1	Node 2	Node 3	Node 4	Node 5
A.africanus	**0.8920**	**0.7500**	0.0249	0.0130	0.0002	0.0000
H.habilis	**0.6460**	**0.9110**	**0.2520**	**0.6490**	**0.2460**	0.0009
H.ergaster	0.0392	0.0022	0.0000	0.0002	0.0001	0.0000
H.georgicus	**0.3120**	0.0099	0.0000	0.0001	0.0000	0.0000
H.antecessor	0.0109	0.0008	0.0000	0.0001	0.0000	0.0000
H.heidelbergensis	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000

Gómez-Robles et a.. 2013. *PNAS*, 110: 18196-18201.

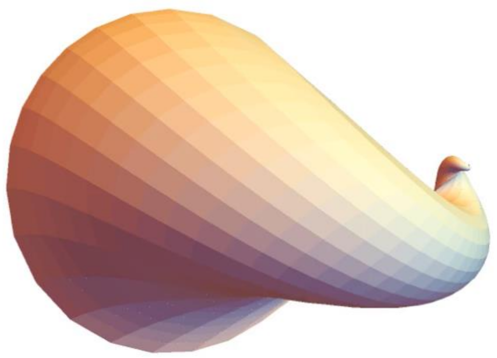
Paths through morphospace

Do straight lines represent realistic biological paths?

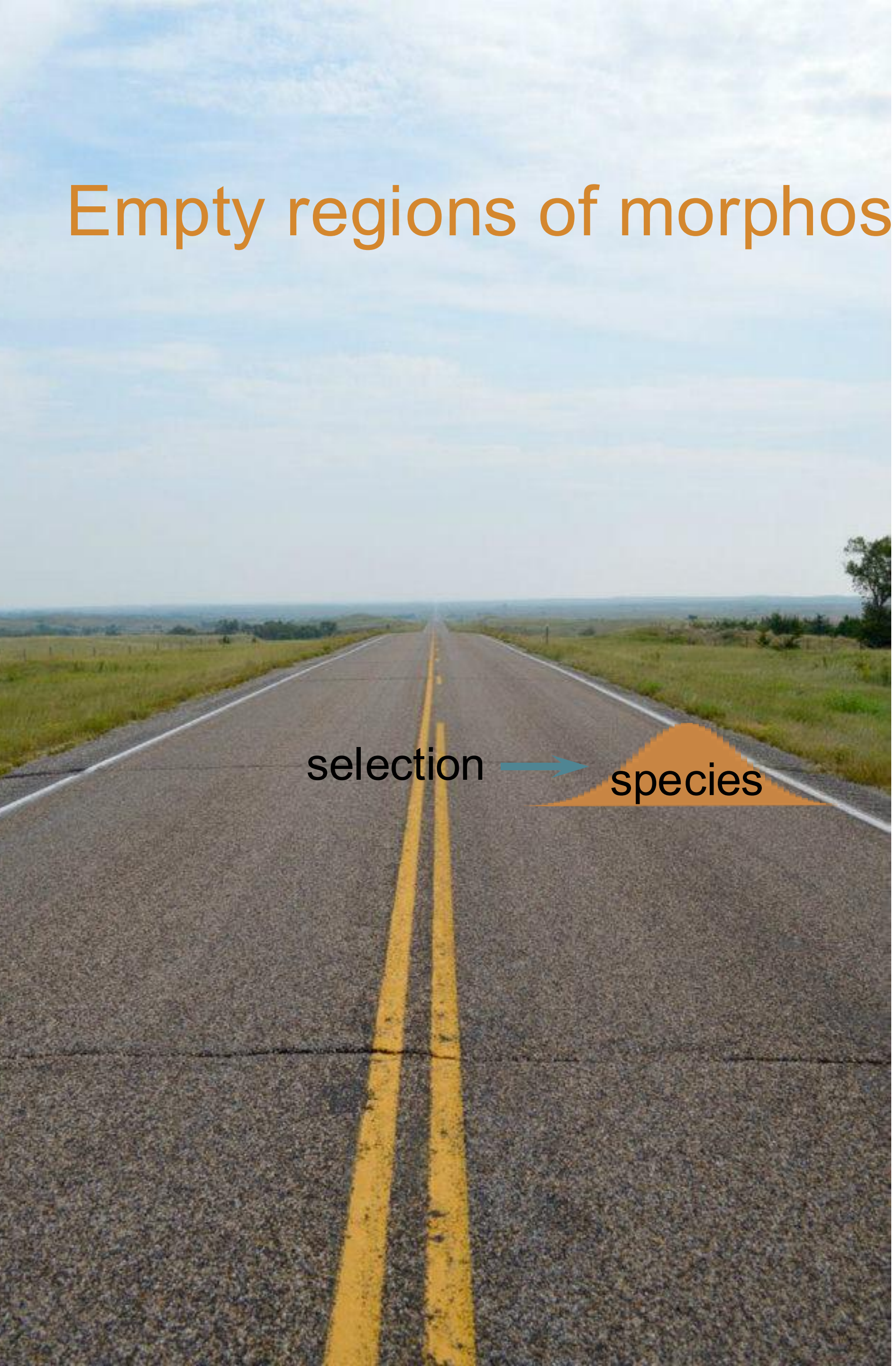
morphospace



morphologies along path



Empty regions of morphospace….

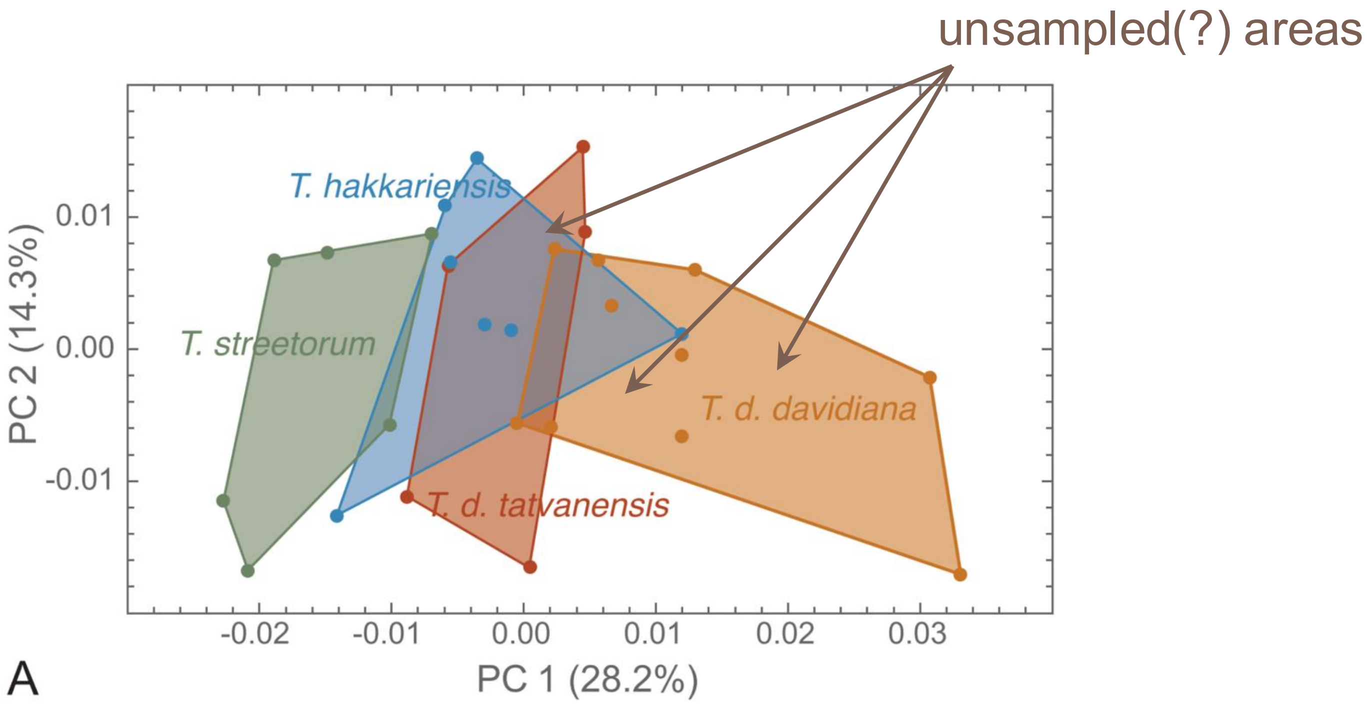


… do not always follow straight paths

Empty areas of morphospace – Type 1

Ordinary gaps

Talpa hakkariensis

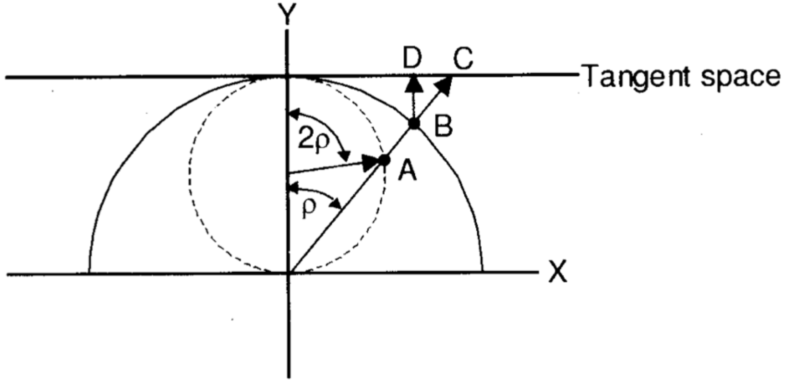


Gündüz et al., 2023, *Zool J Linn Soc*, 199: 567-593

Empty areas of morphospace – Type 2

"Kendall curvature"

Mathematical topology constraints



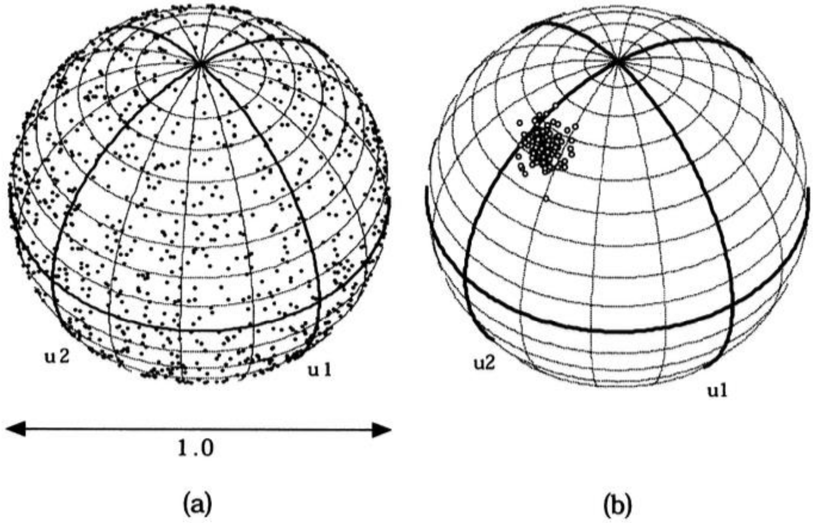


FIGURE 1. A view of Kendall's shape space for triangles showing (a) the mapping of 2,000 random triangles generated by the independent, normal displacement of triplets of points from the origin and (b) the mapping of 110 triangles describing the overall shape of gorilla scapulae. The north pole in both plots corresponds to an equilateral triangle. The south pole corresponds to the reflection of the triangle at the north pole. The equator of each sphere corresponds to triangles with collinear vertices. Longitude is defined with respect to Bookstein's (1996a) linearized Procrustes estimates of uniform shape differences, u1 and u2.

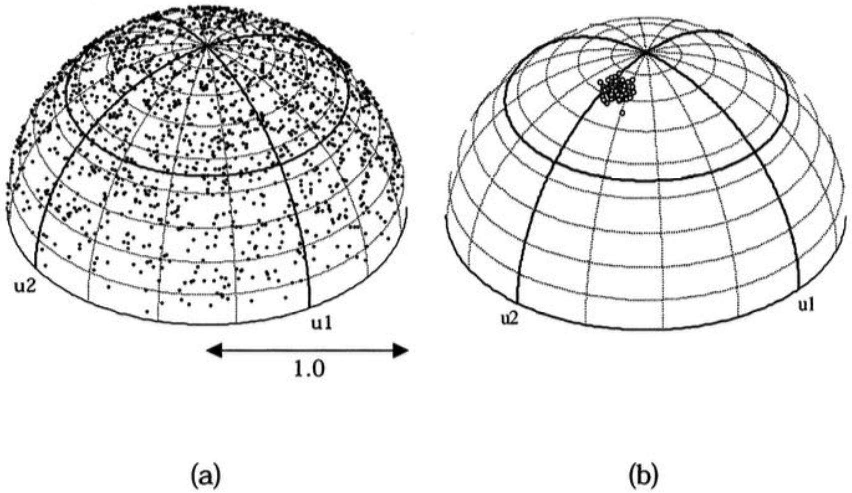


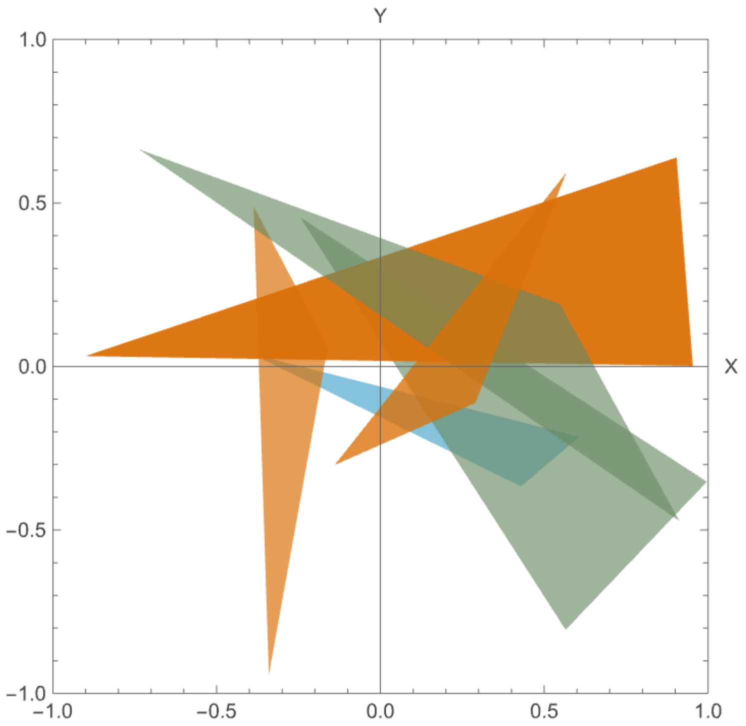
FIGURE 2. A view of the hemisphere of Procrustes-superimposed triangles showing (a) the mapping of the 2,000 random triangles from Figure 1a and (b) the mapping of the 110 gorilla scapulae from Figure 1b. Both sets of data were fit by using an equilateral triangle (north pole) as the reference configuration. The heavy latitudinal line on each hemisphere corresponds to triangles with collinear vertices that map to the equators of the spheres in Figure 1. The equator of the hemisphere corresponds to the reflection of the equilateral triangle at the north pole. Longitude is defined with respect to Bookstein's (1996a) linearized Procrustes estimates of uniform shape differences, u1 and u2.

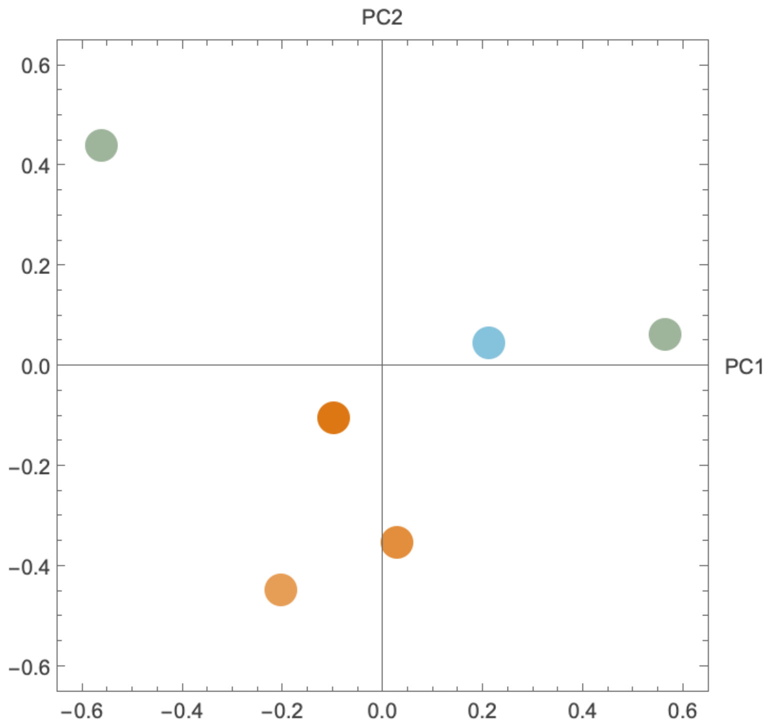
Slice, 2001, "Landmark coordinates aligned by Procrustes analysis do not lie in Kendall's shape space", *Syst Biol* 50: 141-149

Empty areas of morphospace – Type 2

“Kendall curvature”

Mathematical topology constraints

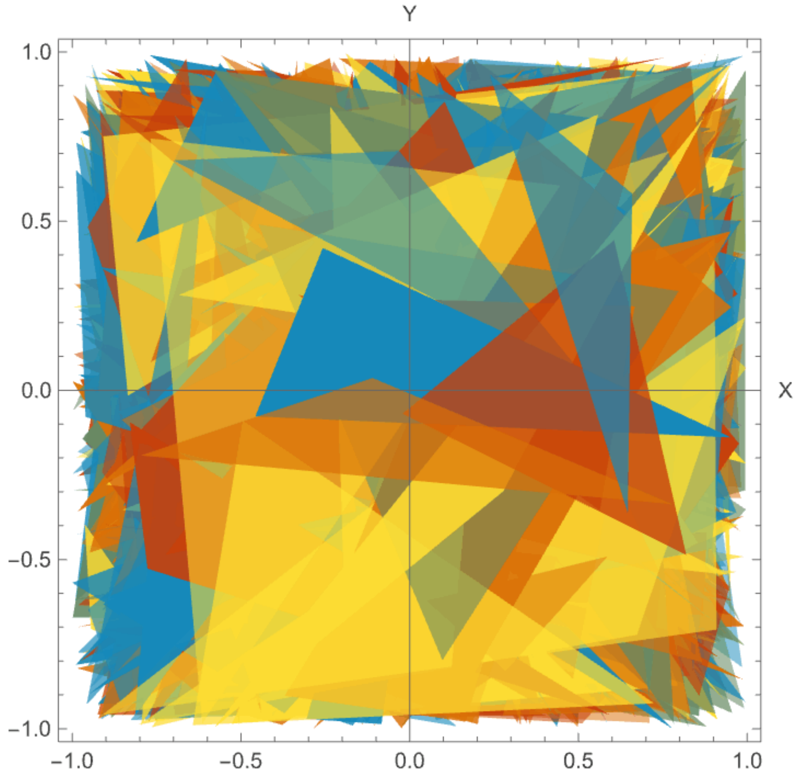


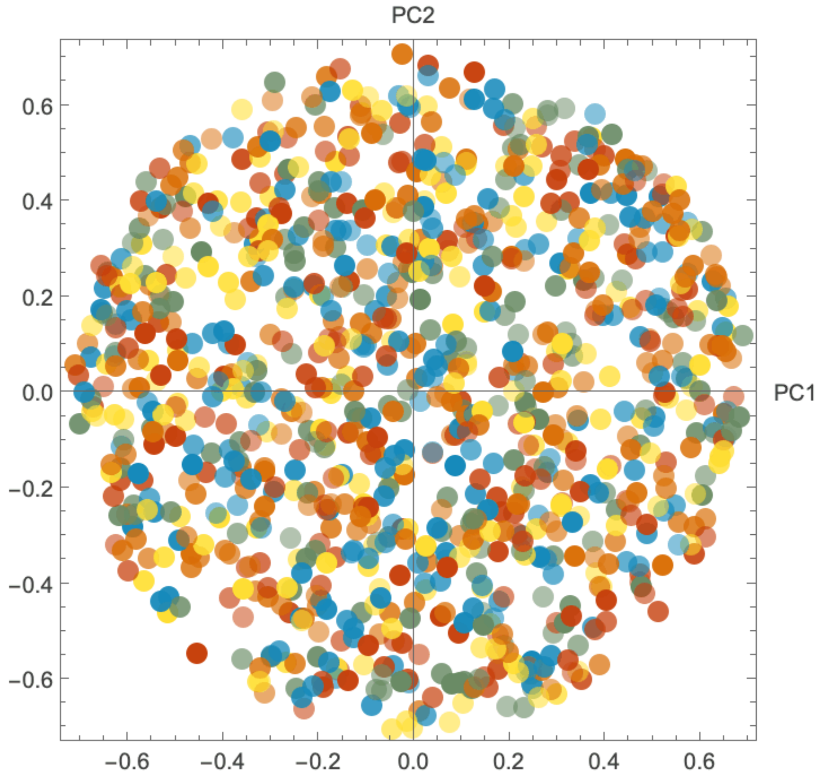


Empty areas of morphospace – Type 2

"Kendall curvature"

Mathematical topology constraints

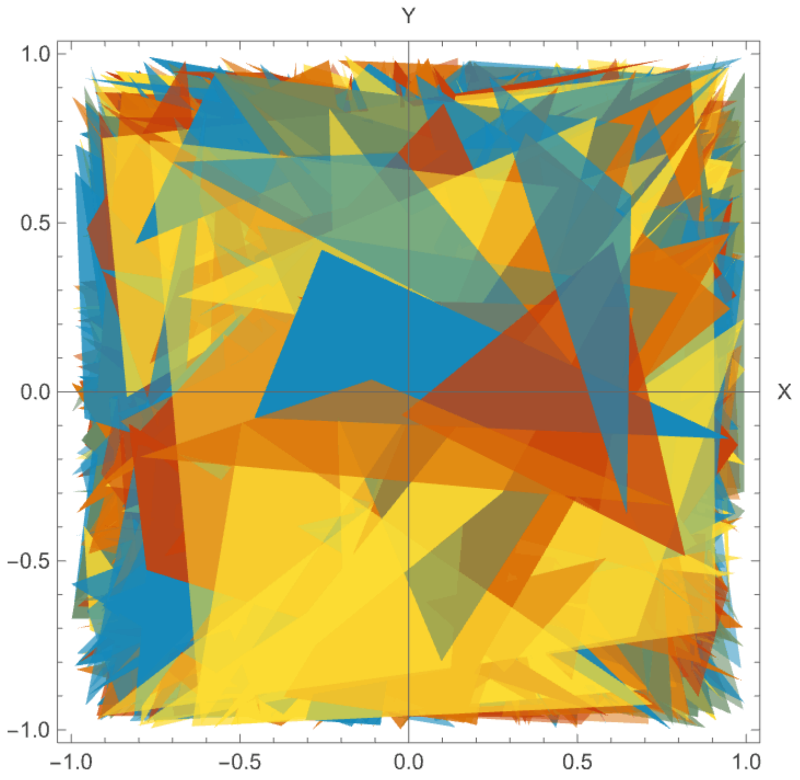


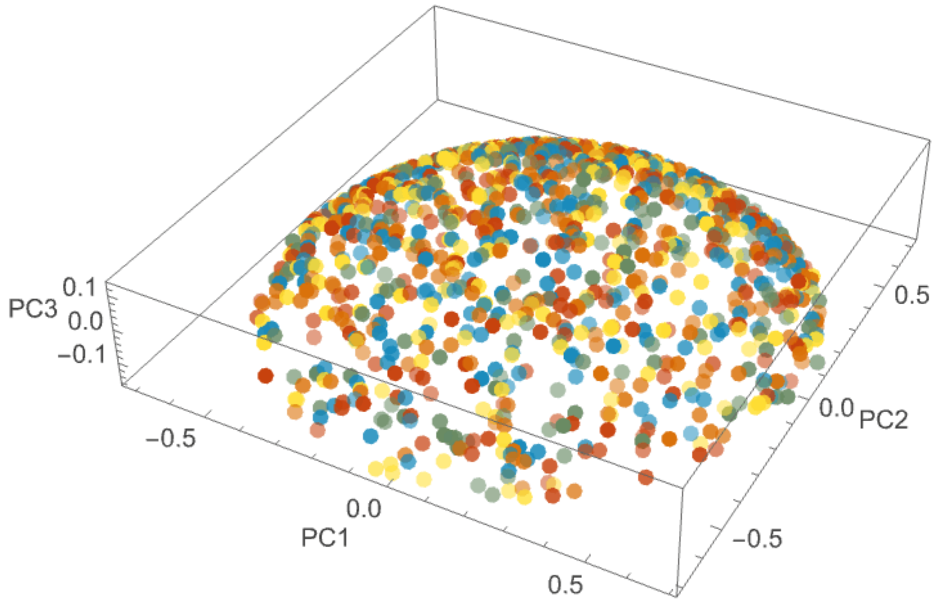


Empty areas of morphospace – Type 2

“Kendall curvature”

Mathematical topology constraints





Empty areas of morphospace – Type 3

“Bookstein bends”

Mathematical interactions between shape gradients and ”major axes”

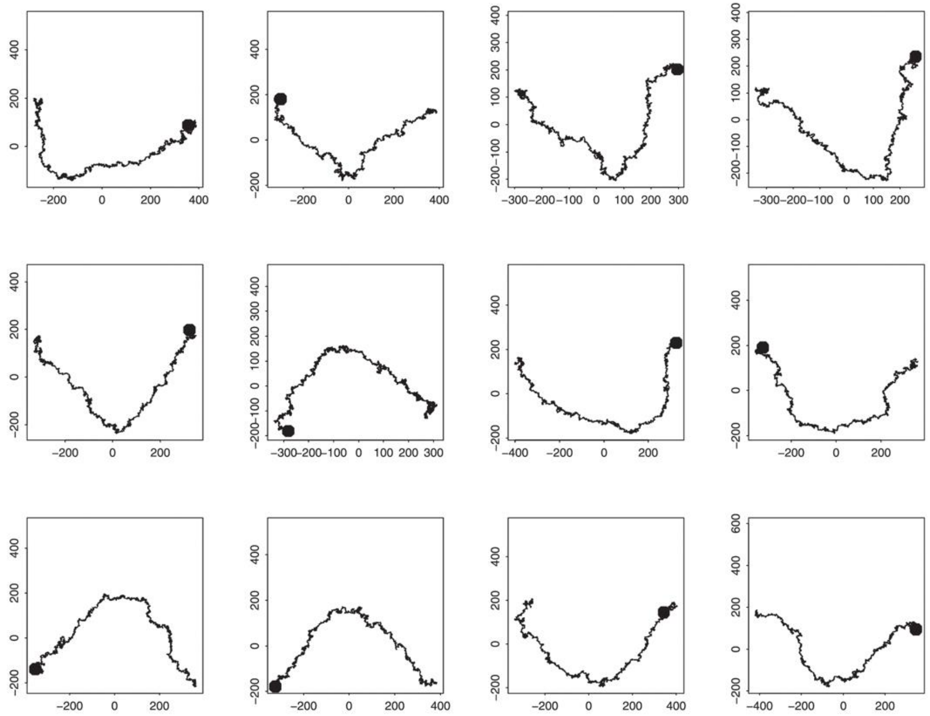


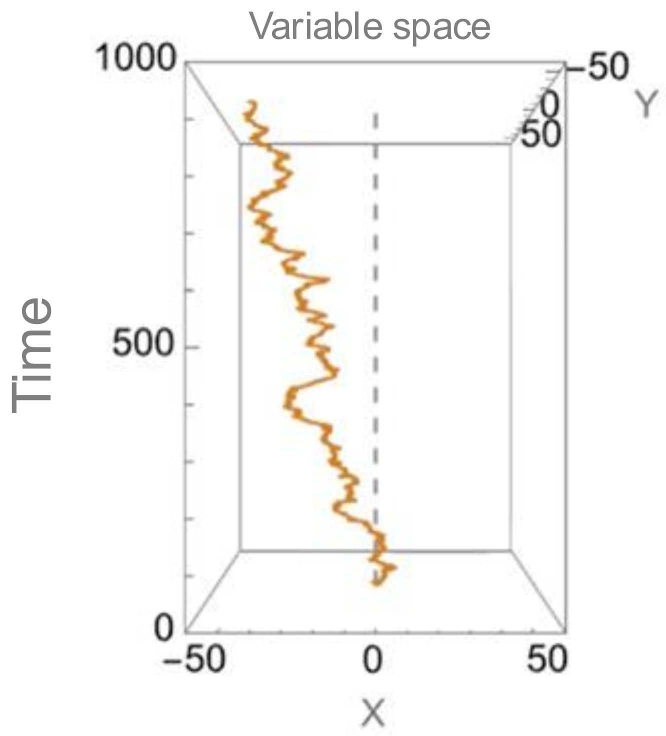
FIGURE 6. First two principal components of twelve isotropic 50-dimensional random walks (Kendall’s model for diffusion in shape space). These each had a 50-dimensional sphere of directions for PC_1 to choose from, and given PC_1 a 49-dimensional sphere of directions for PC_2; yet the resulting principal component analyses all look more or less the same. All walks are of length 10,000. Axes are as in Figure 5.

Bookstein, 2013, “Random walk as a null model”, *Paleobiology* 39: 52-74.

Empty areas of morphospace – Type 3

"Bookstein bends"

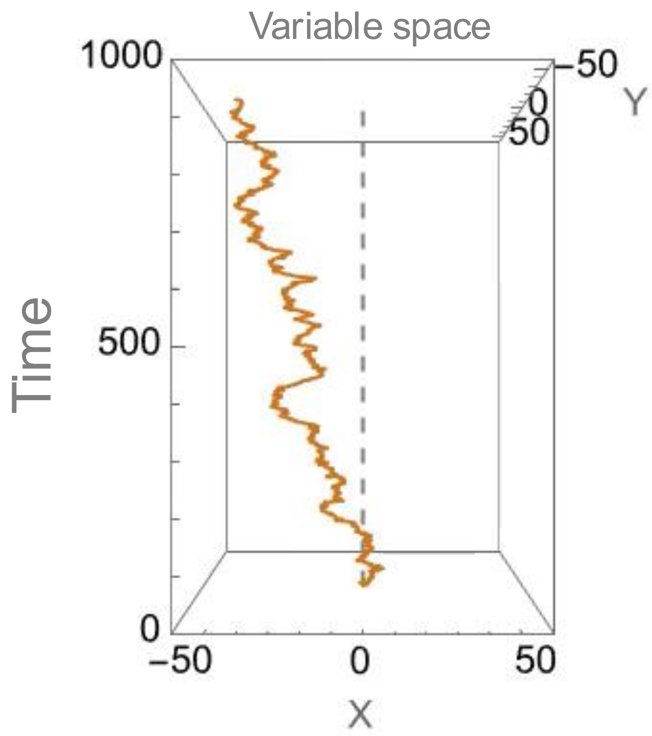
Evolutionary Random Walk
(2 of 56 variables)



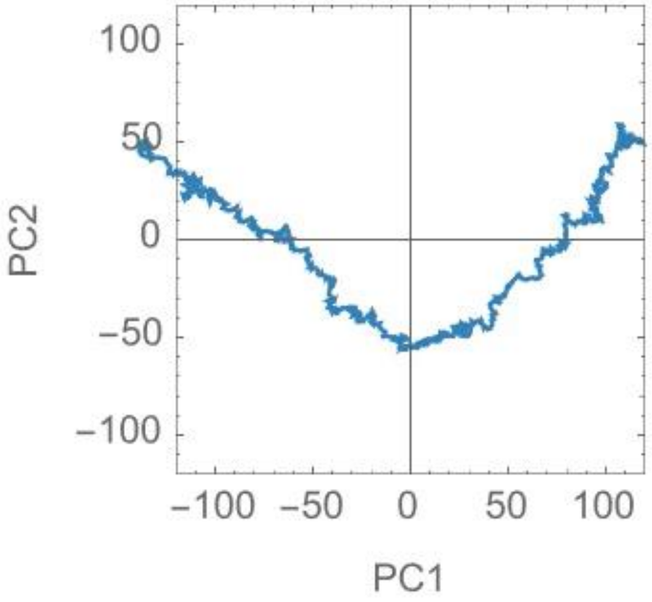
Empty areas of morphospace – Type 3

“Bookstein bends”

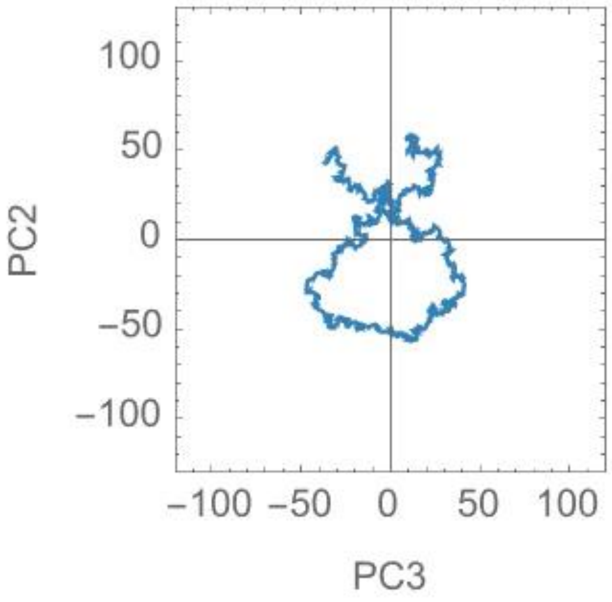
Evolutionary Random Walk
(2 of 56 variables)



PC morphospace
(1st two dimensions)

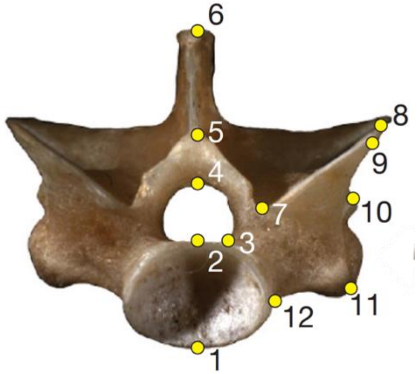


PC morphospace
(2nd two dimensions)

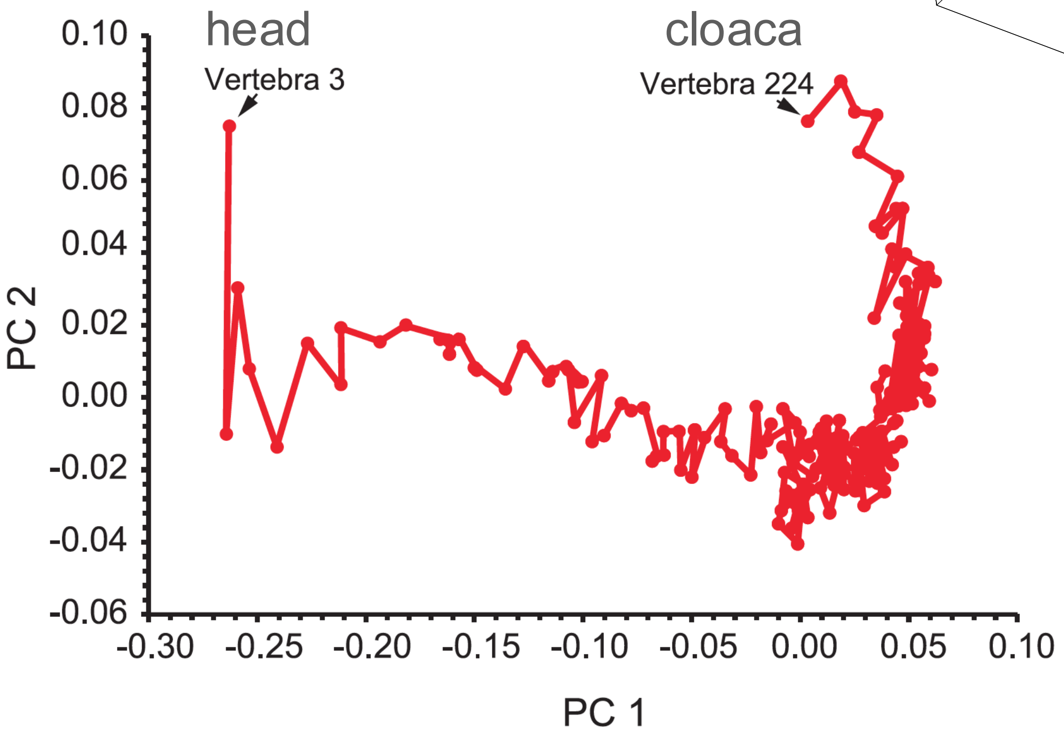


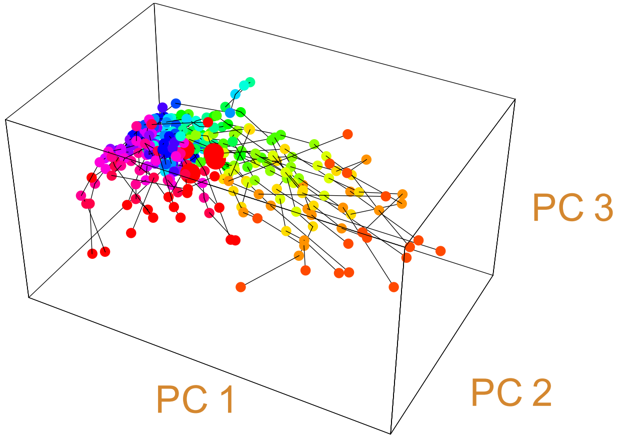
Example Type 3 : morphological series

Jason J. Head – snake vertebral regionalization



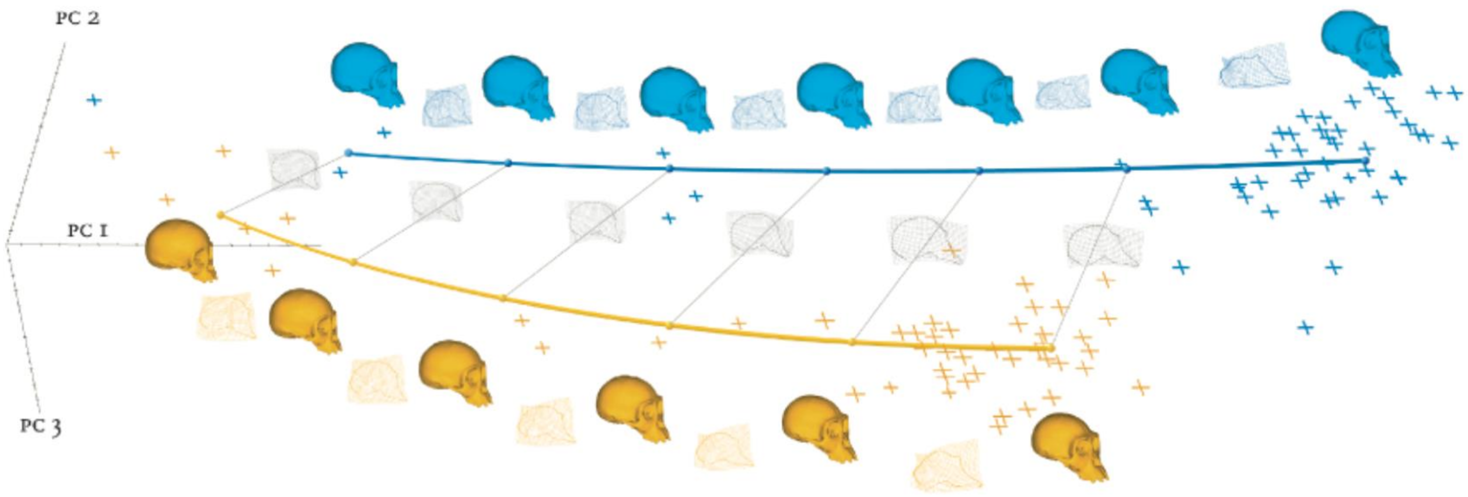
Head & Polly, 2015, *Nature* **520:** 86-89





Example Type 3: ontogenetic series

Philipp Mitteröcker –growth and allometry differences in *chimps*



Mitteroecker et al., 2005. *Evolution & Development,* 7: 244-258.

Example Type 3: paleontologic study of group origins

Timothy B. Rowe – phylogenetic origin of Mammalia

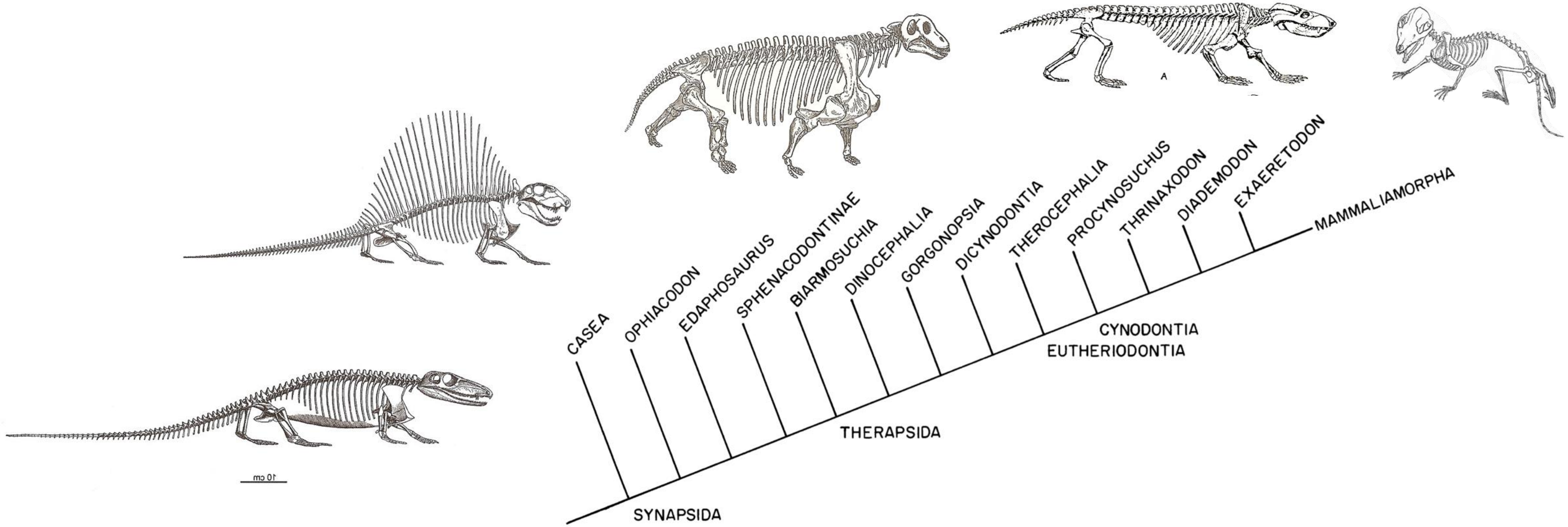
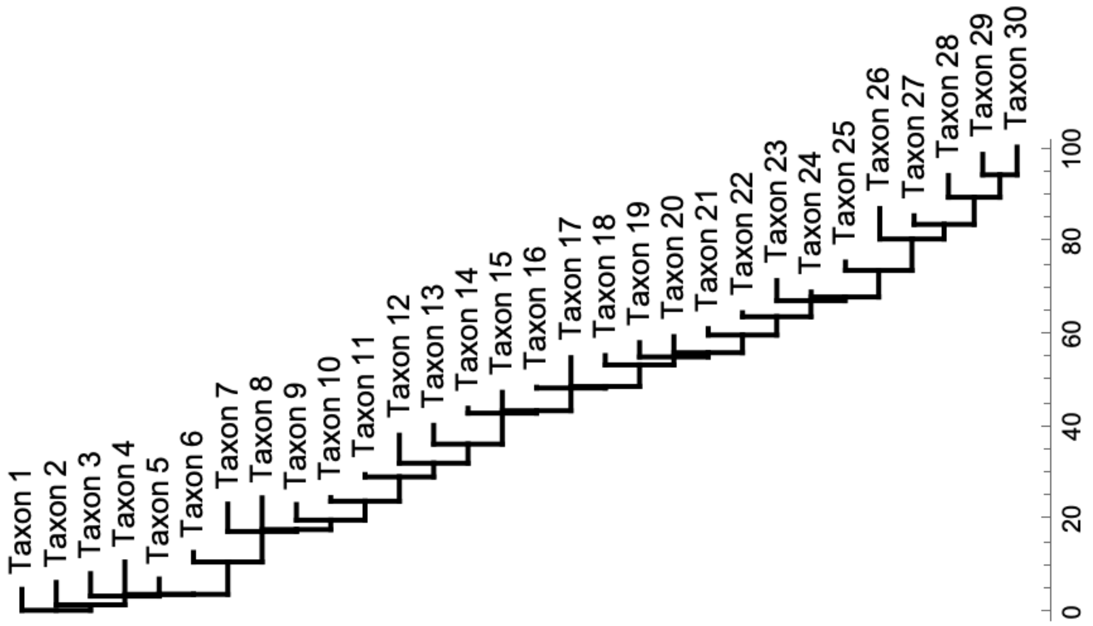
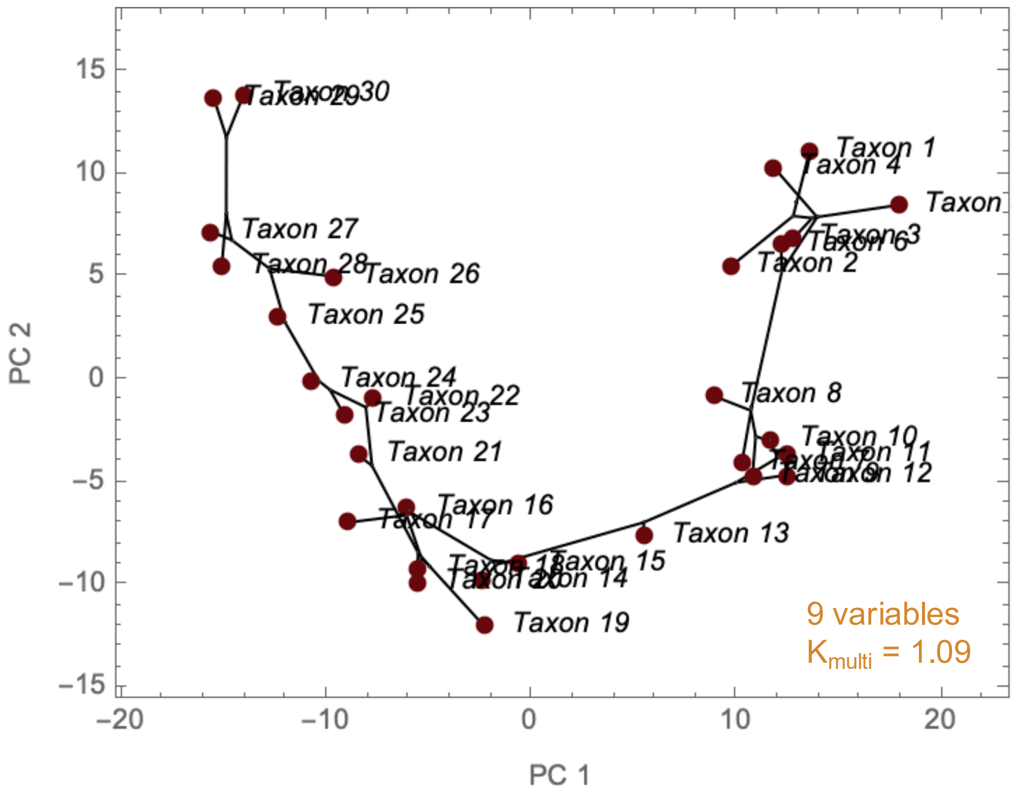


FIGURE 2. Phylogeny of higher systematic categories of Synapsida. This hypothesis depicts the consecutive outgroups used to determine ancestral states for the terminal taxa in this analysis. Character data for this hypothesis are discussed in Rowe (1986a) and Gauthier et al. (1988a).

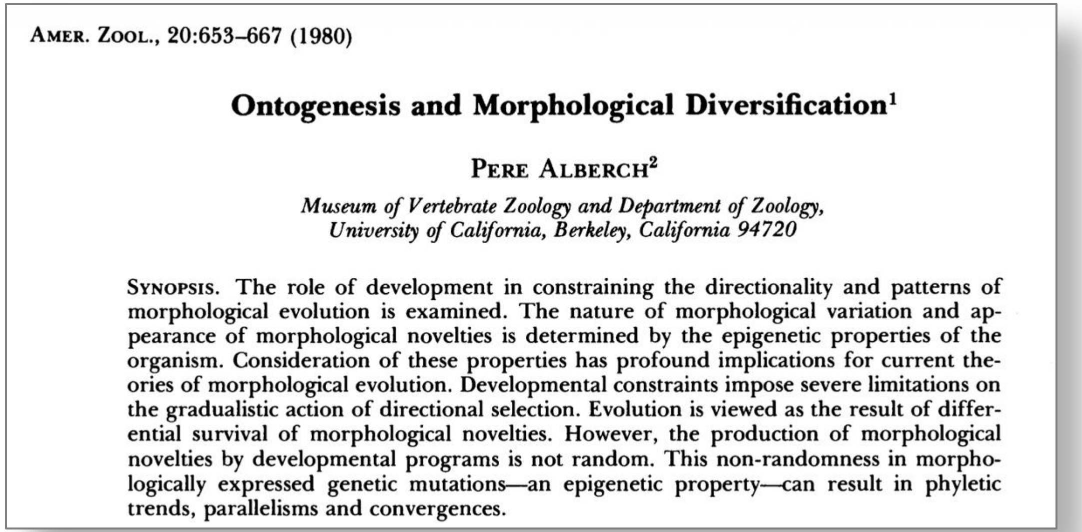
Pectinate trees likely to create "Bookstein bends"





Empty areas of morphospace – Type 4

“Constraint chasms”



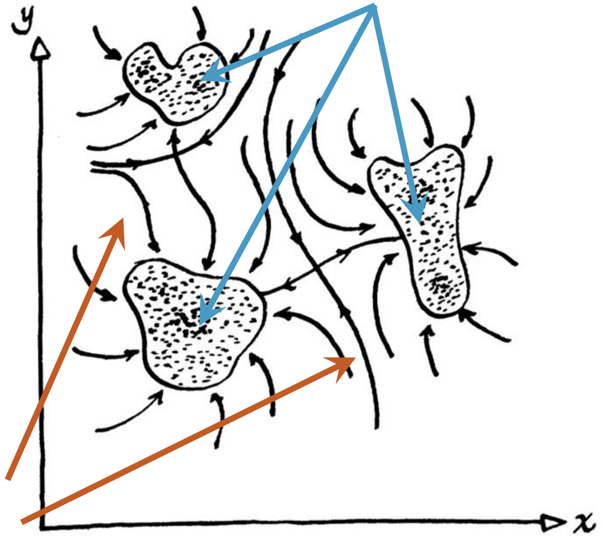
AMER. ZOOL., 20:653–667 (1980)

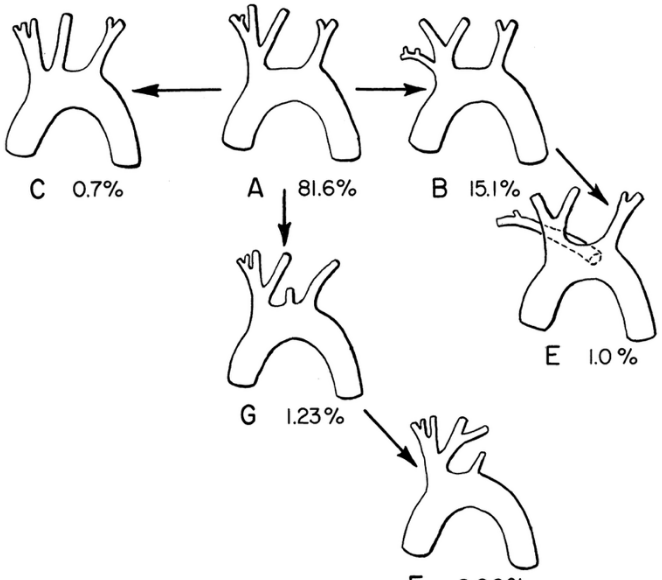
Ontogenesis and Morphological Diversification[1]

PERE ALBERCH[2]

Museum of Vertebrate Zoology and Department of Zoology, University of California, Berkeley, California 94720

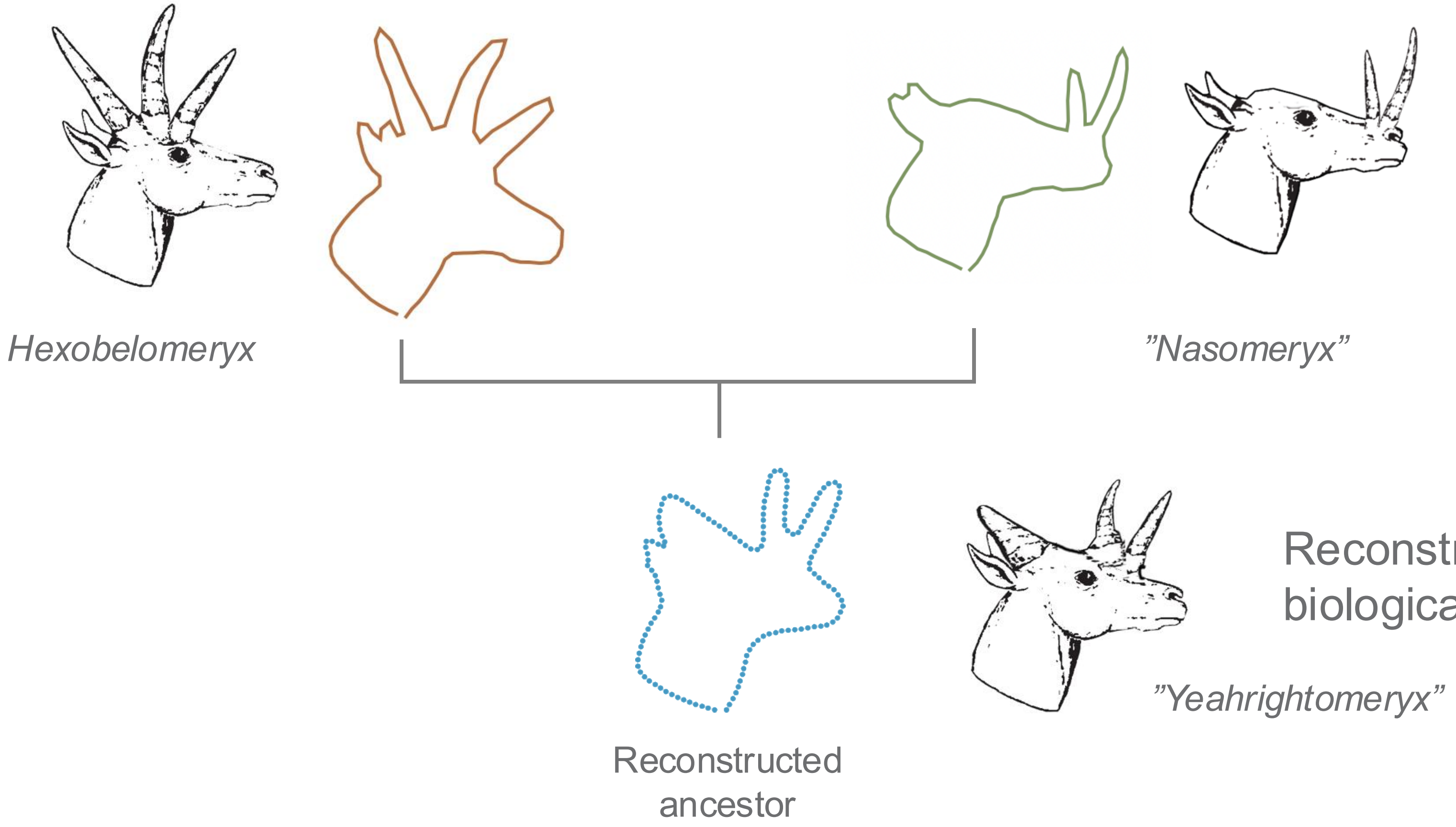
SYNOPSIS. The role of development in constraining the directionality and patterns of morphological evolution is examined. The nature of morphological variation and appearance of morphological novelties is determined by the epigenetic properties of the organism. Consideration of these properties has profound implications for current theories of morphological evolution. Developmental constraints impose severe limitations on the gradualistic action of directional selection. Evolution is viewed as the result of differential survival of morphological novelties. However, the production of morphological novelties by developmental programs is not random. This non-randomness in morphologically expressed genetic mutations—an epigenetic property—can result in phyletic trends, parallelisms and convergences.





Empty areas of morphospace – Type 4

“Constraint chasms”

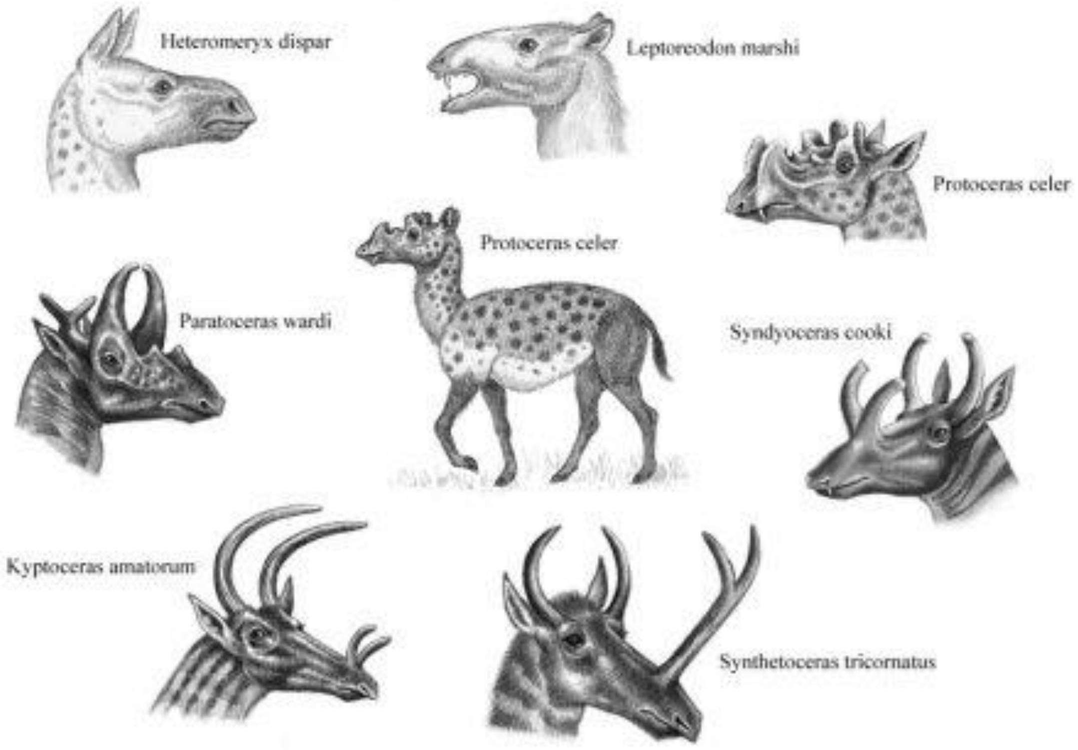


Reconstruction is biologically implausible

Poll, 2017. *Vavilov Journal of Genetics and Selection*, **21:** 452-461

Empty areas of morphospace – Type 4

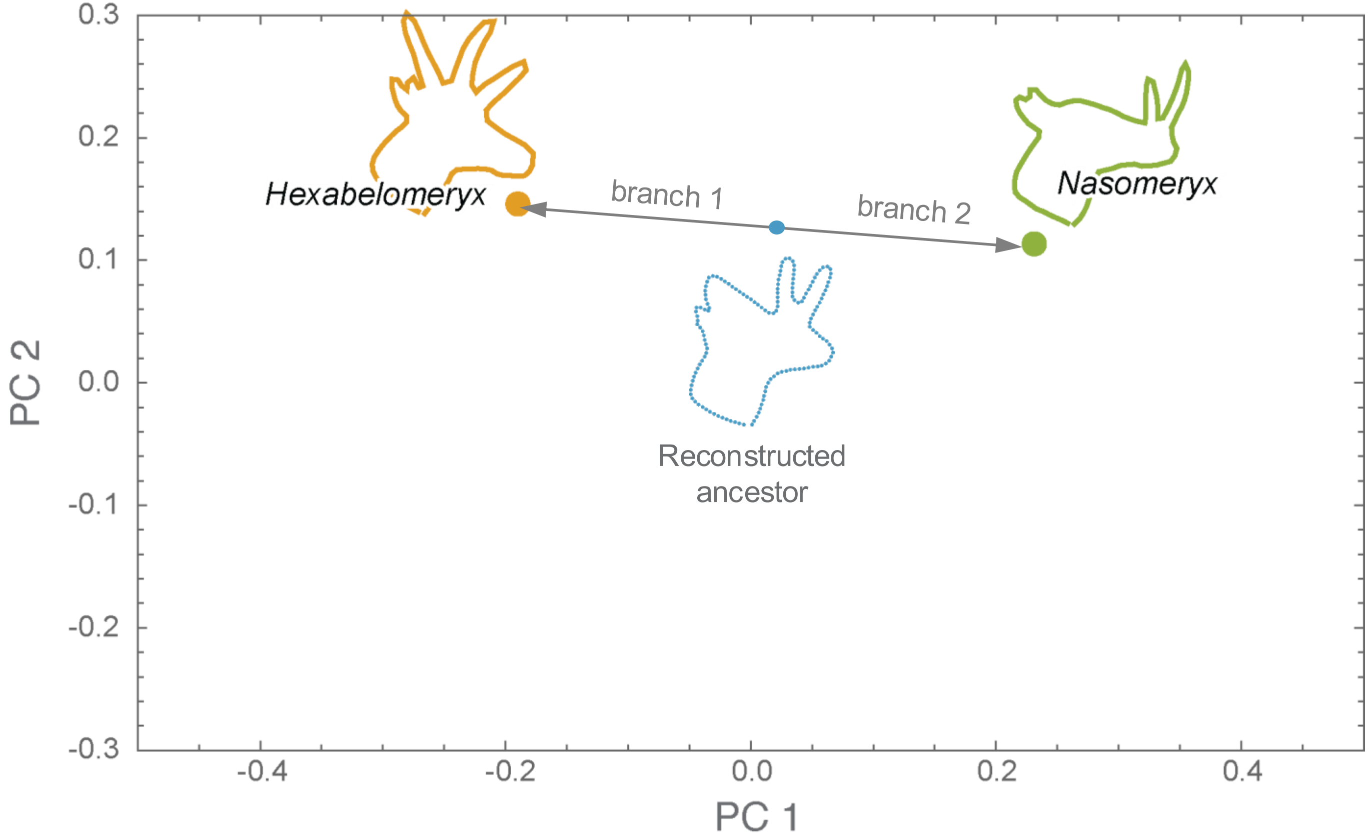
“Constraint chasms”



ferrebeekeeper

Empty areas of morphospace – Type 4

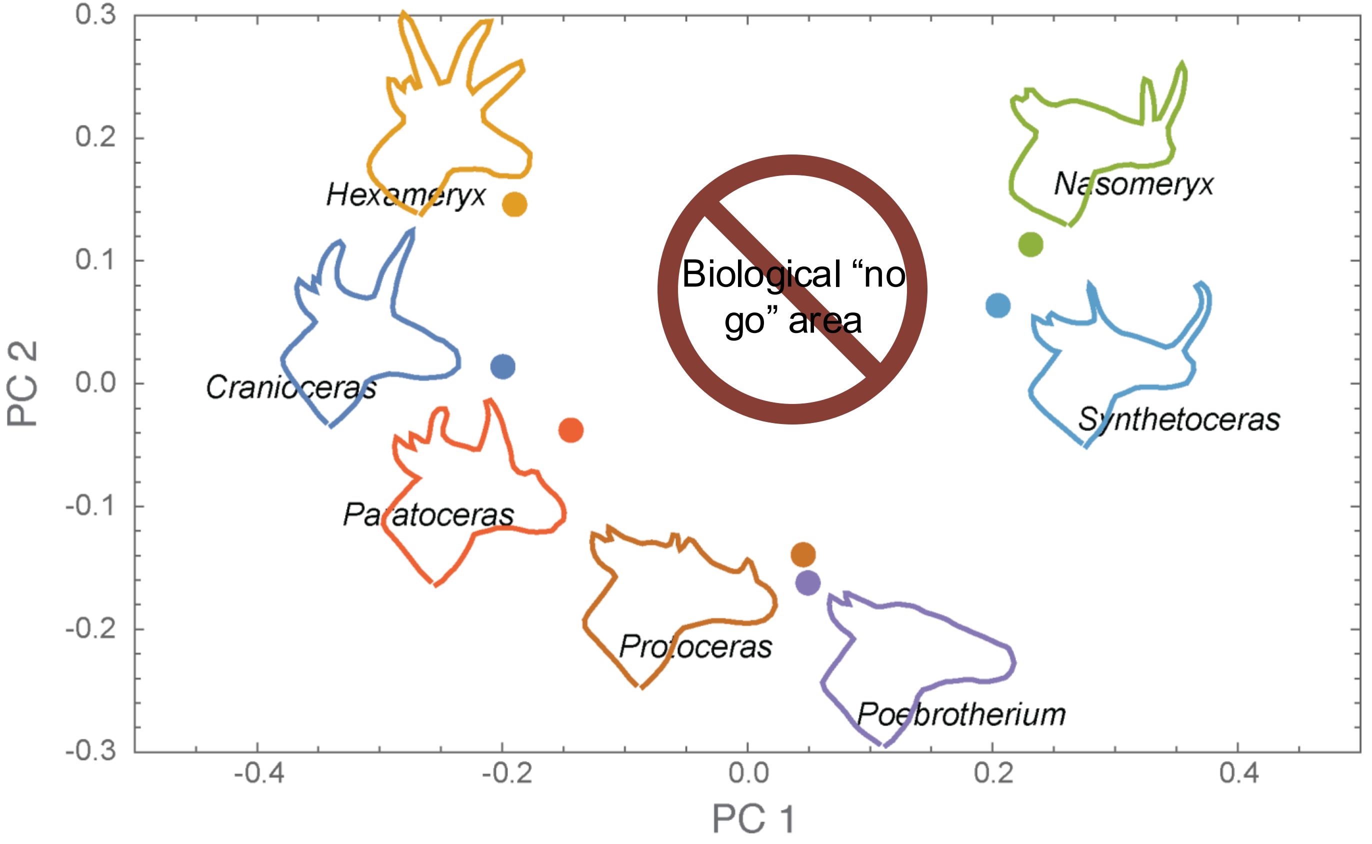
“Constraint chasms”



The Polly Lab
www.pollylab.org

Empty areas of morphospace – Type 4

"Constraint chasms"



Example Type 4: mammal teeth

Isaac Salazar-Ciudad

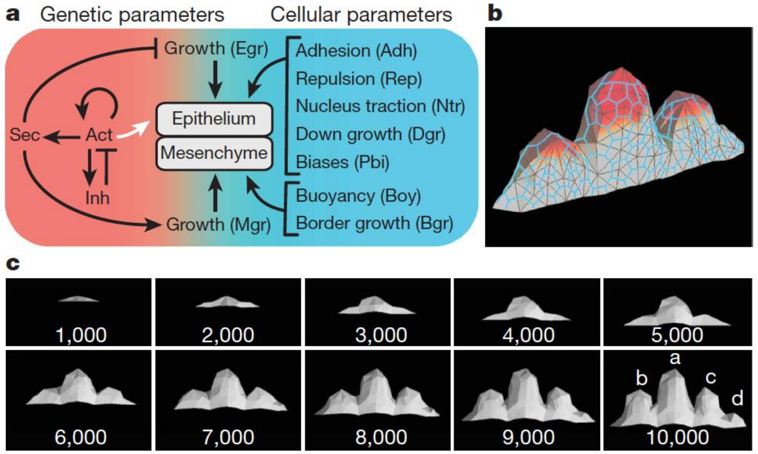


Figure 1 | A model integrating gene networks and tissue mechanics. a, Nine

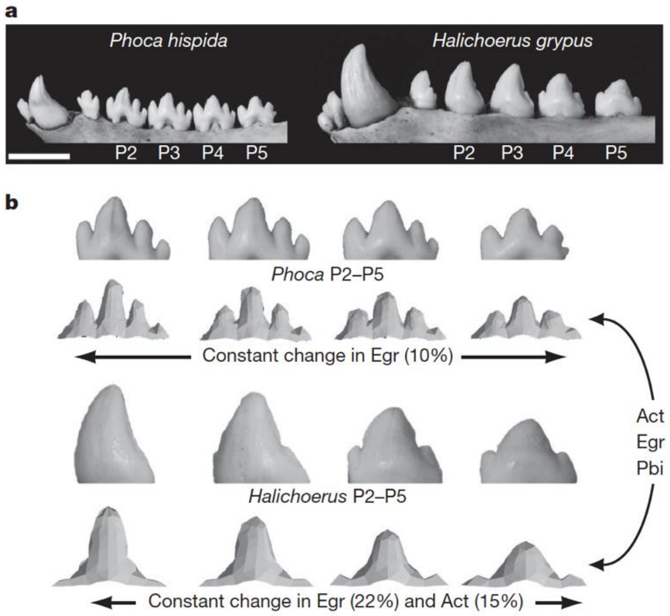
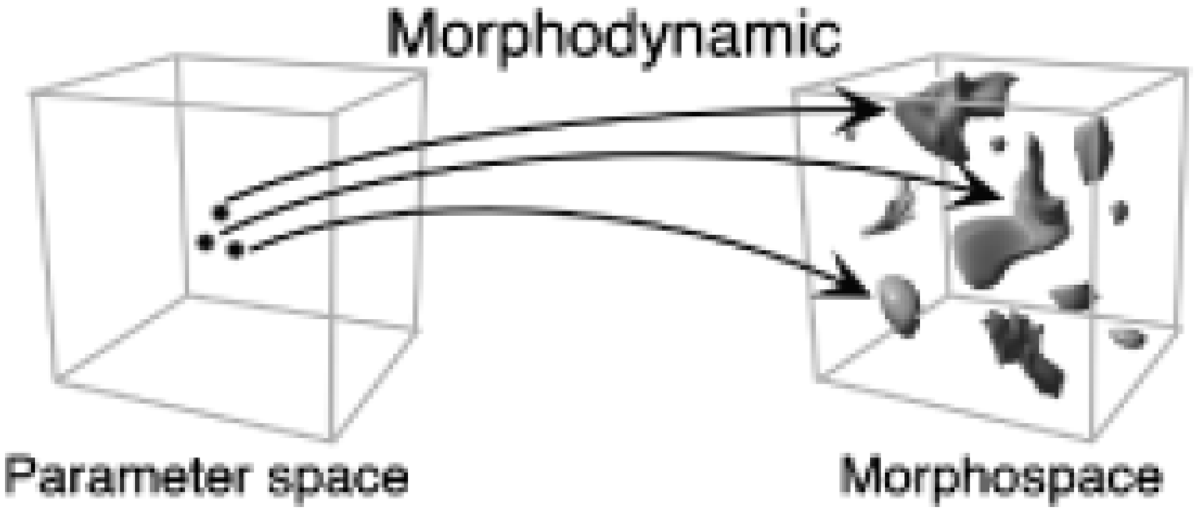


Figure 4 | Serial tooth-to-tooth variation implicates a cellular parameter.



Salazar-Ciudad I, Jernvall J, 2004, *Evolution and Development*, **6:** 6-16; 2010, *Nature*, **464**: 583–586.

Empty spaces summary

Type 1 – ordinary gaps

- everything good

Type 2 – Kendall curvature

- usually correctable with projection to Euclidean space
- usually negligible for biological data sets (e.g., phylogenetic ones)
- can be an issue with:
  - non-biological objects (e.g. stone tools)
  - simulated shapes
- testable by comparing full and projected Procrustes distances

Type 3 – Bookstein bends

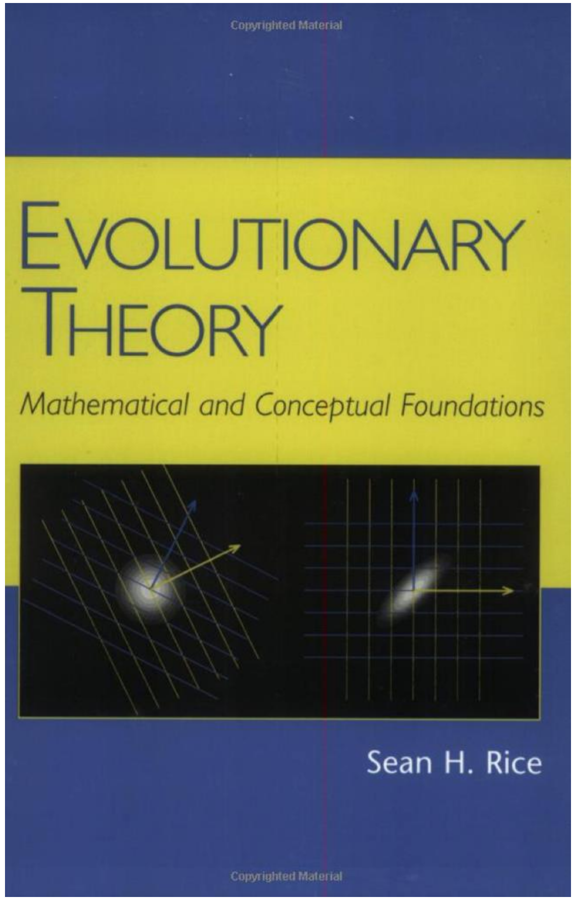
- fairly common
- can easily be overinterpreted on low PCs
- statistics and model fitting may thwarted on low PCs
- may misguide ancestor reconstruction
- using full morphospace (all PCs) should alleviate statistical issues

Type 4 – Constraint chasms

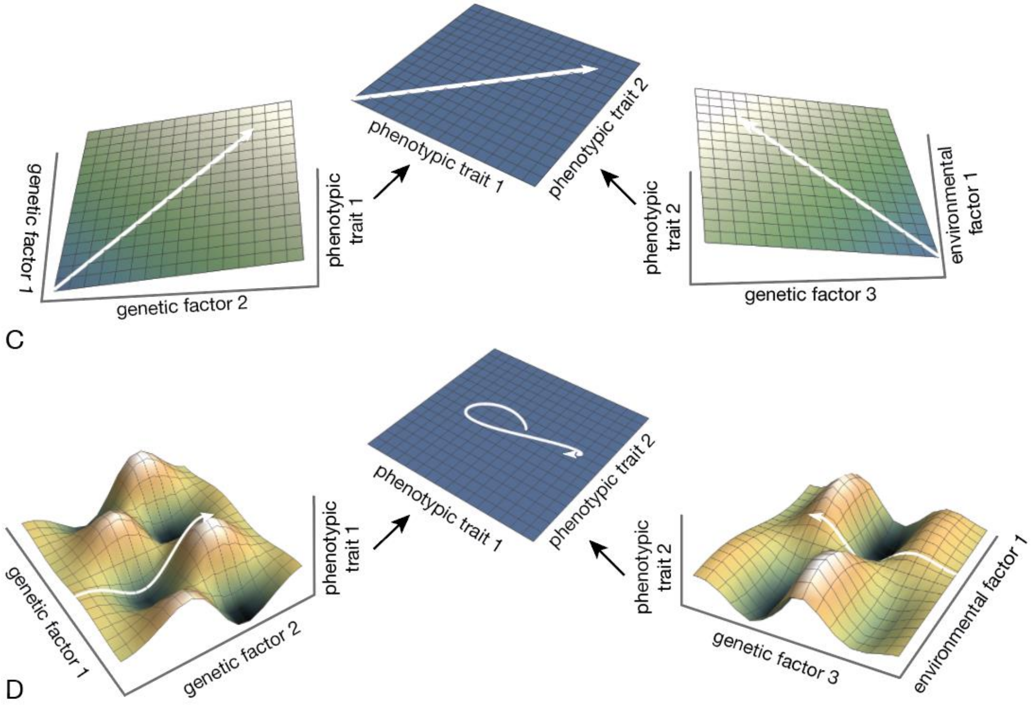
- biological, not mathematical in nature
- caused by non-linear relationship between biological and mathematical paths
- may thwart ancestor reconstruction and model fitting (but not statistics?)

Phenotypic landscape theory to the rescue?

Rice S, 2004. *Evolutionary Theory.* Sinauer Associates.



Genotype-phenotype map combined with morphometrics could be used to reconstruct biologically realistic trajectories



Polly PD, 2008. Evolutionary Biology, 35: 83-96

Review of morphospace concepts

The Concept of Morphospaces in Evolutionary and Developmental Biology: Mathematics and Metaphors

Philipp Mitteroecker
Department of Theoretical Biology
University of Vienna
Vienna, Austria
philipp.mitteroecker@univie.ac.at

Simon M. Huttegger
Department of Logic and Philosophy of Science
University of California, Irvine
Irvine, CA, USA
shuttegg@uci.edu

Abstract

Formal spaces have become commonplace conceptual and computational tools in a large array of scientific disciplines, including both the natural and the social sciences. Morphological spaces (morphospaces) are spaces describing and relating organismal phenotypes. They play a central role in morphometrics, the statistical description of biological forms, but also underlie the notion of adaptive landscapes that drives many theoretical considerations in evolutionary biology. We briefly review the topological and geometrical properties of the most common morphospaces in the biological literature. In contemporary geometric morphometrics, the notion of a morphospace is based on the Euclidean tangent space to Kendall's shape space, which is a Riemannian manifold. Many more classical morphospaces, such as Raup's space of coiled shells, lack these metric properties, e.g., due to incommensurably scaled variables, so that these morphospaces typically are affine vector spaces. Other notions of a morphospace, like Thomas and Reif's (1993) skeleton space, may not give rise to a quantitative measure of similarity at all. Such spaces can often be characterized in terms of topological or pretopological spaces.

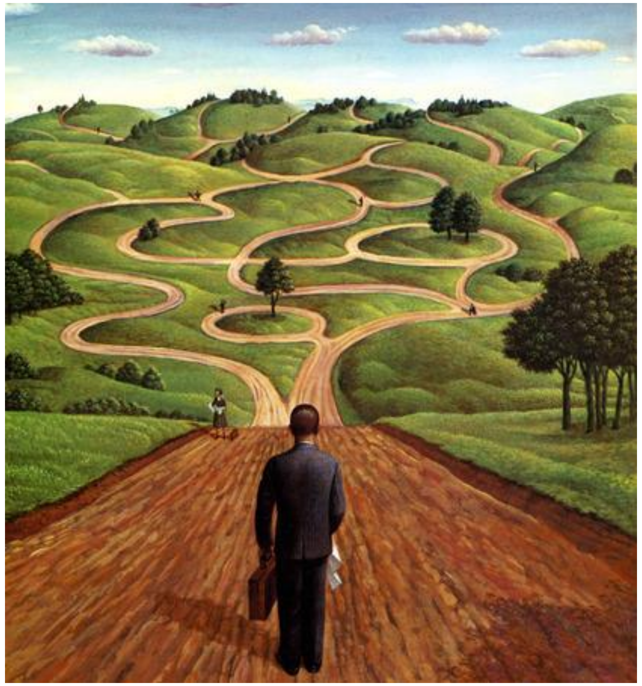
The typical language of theoretical and evolutionary biology, comprising statements about the "distance" among phenotypes in an according space or about different "directions" of evolution, is not warranted for all types of morphospaces. Graphical visualizations of morphospaces or adaptive landscapes may tempt the reader to apply "Euclidean intuitions" to a morphospace, whatever its actual topology might be. We discuss the limits of metaphors such as the developmental hourglass and adaptive landscapes that ensue from the geometric properties of the underlying morphospace.

Keywords

adaptive landscapes, affine space, developmental hourglass, morphometrics, phenetic space, sequence space, shape space, skeleton space, theoretical morphology, topology

January 30, 2009; accepted September 5, 2009
Biological Theory 4(1) 2009, 54–67. ©2009 Konrad Lorenz Institute for Evolution and Cognition Research

Thinking forward…



Are there better ways for evaluating “empty” spaces?

- tests for “Bookstein bends” vs. “Constraint chasms”

Are there better ways for managing high-dimensionality?

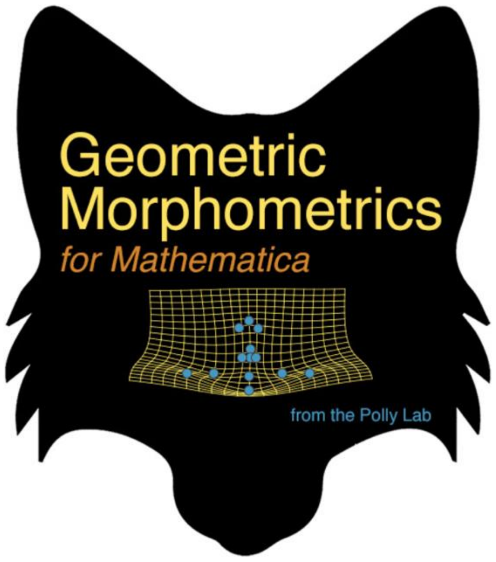
- Trade-off between analysis analysis in full vs reduced dimensionality (captures full patterns but over parameterized?)

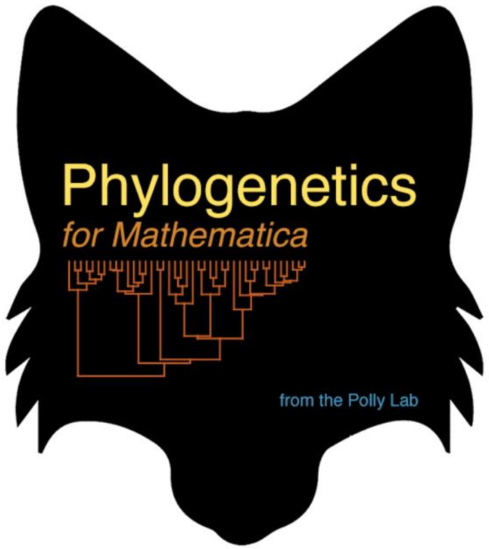
Can easy-to-use tools be developed?

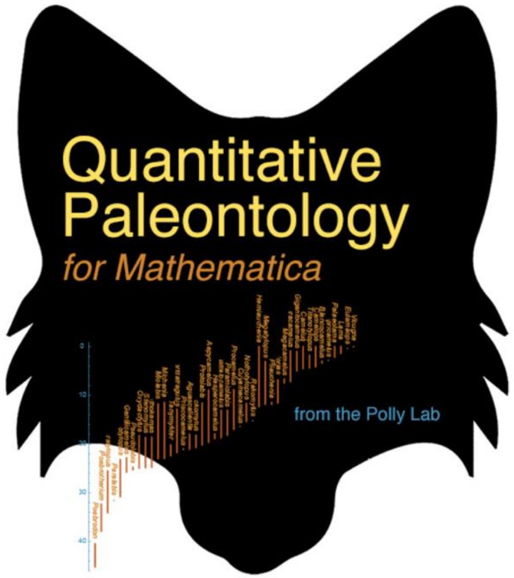
- tests for effects of data pattern on evolutionary models?

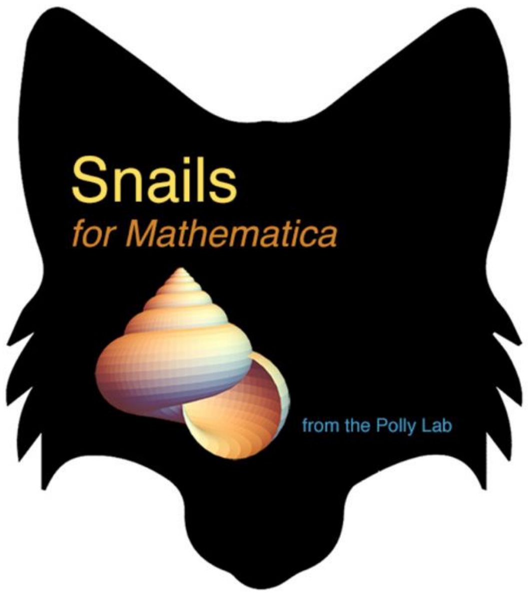
Software for Mathematica©

https://github.com/pdpolly





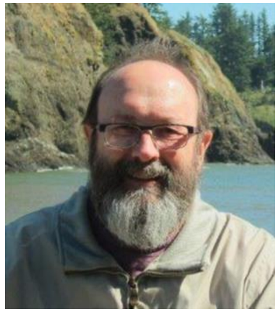




Jason Head, Cambridge

Andrea Cardini
Modena

Radha Caumul
Burty David School, Mauritius



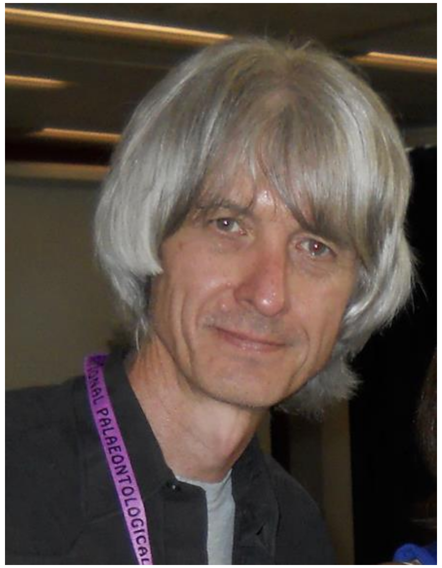
Joe Felsenstein
University of Washington

Philip Gingerich
U Michigan

Michelle Lawing
Texas A&M

Anjali Goswami
NHM London

Emília Martins
Arizona State University



Norman MacLeod, Nanjing U

Steve Le Comber
Queen Mary, U London
(deceased)

Aida Gómez Robles
Universty College London

Katrina Jones
Bristol University

Miriam Zelditch
U Michigan