# Phylogenetic Inference of Character Evolution I: Discretely Coded Traits

Rich Glor University of Kansas March 11, 2014

## **Research Interests**

Hypothe

- Herpetology
- Adaptive radiation
- Species diversification
- Speciation









# Outline

- 1. Insights from studies of character evolution
- 2. Requirements for studies of character evolution
- 3. Data exploration
  - Visualizing trait data on phylogenetic trees (R tutorial)
- 4. Trait evolution
  - ML and Bayesian methods for estimating ancestral states, rates, and models of evolution (R & BayesTraits tutorials)
  - Stochastic character mapping (R & SIMMAP tutorials)
  - Correlations among characters (BayesTraits tutorial)

# Outline

- 1. Insights from studies of character evolution
- 2. Requirements for studies of character evolution
- 3. Data exploration
  - Visualizing trait data on phylogenetic trees (R tutorial)
- 4. Trait evolution
  - ML and Bayesian methods for estimating ancestral states, rates, and models of evolution (R & BayesTraits tutorials)
  - Stochastic character mapping (R & SIMMAP tutorials)
  - Correlations among characters (BayesTraits tutorial)





Lunch

Coffee

## Character evolution road map

	Discrete	Continuous					
Trait evolution: ancestral reconstruction, patterns of trait, rates of trait evolution, biogeographic analyses	Glor: Tuesday Landis: Tuesday	Mahler: Wednesday Wainwright: Thursday Price: Thursday					
rates of trait evolution, biogeographic analyses Trait interactions: trait correlations, impact of traits on one another	Glor: Tuesday Price: Thursday	Mahler: Wednesday Wainwright: Thursday Price: Thursday					

Impact of traits on taxonomic diversification

Moore: Thursday

#### ACTA CHEMICA SCANDINAVICA 17 (1963) S9-S16

**Chemical Paleogenetics** 

Molecular "Restoration Studies" of Extinct Forms of Life

LINUS PAULING and EMILE ZUCKERKANDL\*

Division of Chemistry and Chemical Engineering, California Institute

- Trait evolution
  - Reconstruction of ancestral traits
    - Proteins

#### of Technology, Pasadena, California, USA\*\* S 10 PAULING AND ZUCKERKANDL Time Present X (a)Time (a) (b) (b) (b) (c) (c)(c)

Fig. 1. Different evolutionary relationships of the amino-acid residues  $\rho$  and  $\sigma$  found at corresponding sites in the homologous polypeptide chains C, D, and E. See text.

(d)

(c)



- Trait evolution
  - Reconstruction of ancestral traits
    - Proteins

#### **Recreating a Functional Ancestral Archosaur Visual Pigment**

Belinda S. W. Chang,\* Karolina Jönsson,\* Manija A. Kazmi,\* Michael J. Donoghue,† Thomas P. Sakmar\*



Chang et al. 2002. MBE.

- Trait evolution
  - Reconstruction of ancestral traits
    - Proteins





Chang et al. 2002. MBE.

- Trait evolution
  - Reconstruction of ancestral traits
    - Proteins



Copyright © 1995 by the Genetics Society of America

#### A New Method of Inference of Ancestral Nucleotide and Amino Acid Sequences

Ziheng Yang, Sudhir Kumar and Masatoshi Nei

Institute of Molecular Evolutionary Genetics and Department of Biology, The Pennsylvania State University, University Park, Pennsylvania 16802

> Manuscript received May 8, 1995 Accepted for publication September 14, 1995

### Trait evolution

- Reconstruction of ancestral traits
  - Proteins

TABLE 3

Maximum-likelihood reconstructions of amino acids at interior nodes of the tree

		_		_			_	_		_							_		_	_							_			_		_		_	_	_	_	_	_	_	_
						1							2					3						4							5				б						7
Site	2	3	4	5	8	0	1	4	5	6	7	8	0	1	3	7	9	1	2	3	4	7	8	1	3	5	б	7	8	9	0	2	б	9	1	2	3	6	7	8	1
Langur	I	F	Е	R	L	R	т	К	L	G	L	D	Y	к	v	N	V	$\mathbf{L}$	A	K	W	G	Y	Е	т	Y	N	Ρ	G	D	E	т	I	I	S	R	Y	N	N	G	Р
Baboon	I	F	Е	R	L	R	т	R	$\mathbf{L}$	G	$\mathbf{L}$	D	Y	R	Ι	N	V	$\mathbf{L}$	А	К	W	D	Y	Q	т	Y	Ν	Ρ	G	D	Q	т	I	I	S	Н	Y	Ν	D	G	Ρ
Human	v	F	Е	R	L	R	т	R	L	G	М	D	Y	R	Ι	N	М	L	А	K	W	G	Y	R	т	Y	Ν	А	G	D	R	т	I	I	S	R	Y	N	D	G	Ρ
Rat	т	Y	Е	R	F	R	т	R	Ν	G	М	s	Y	Y	v	D	v	$\mathbf{L}$	A	Q	н	Ν	Y	Q	R	Y	D	Ρ	G	D	Q	т	I	Ι	S	R	Y	N	D	G	Ρ
Cow	V	F	Е	R	L	R	т	к	$\mathbf{L}$	G	$\mathbf{L}$	D	Y	K	v	Ν	L	$\mathbf{L}$	т	K	W	S	Y	Κ	т	Y	Ν	Ρ	s	s	Е	т	I	I	s	К	W	N	D	G	Ρ
Horse	v	F	s	К	L	н	K	А	Q	Е	М	D	F	G	Y	N	v	М	А	Е	Y	Ν	F	R	F	G	К	N	А	Ν	G	s	L	L	Ν	К	W	K	D	N	R
Node 7	v	F	R	R	T.	R	т	R	L	G	м	р	v	R	i	N	v	L		ĸ	w	2	Y	0	т	Y	N	P	G	р	0	т	т	т	s	R	v	N	р	G	P
Node 8	v	F	E	R	ī	R	T	R	L.	G	м	Đ	v	R	i	N	v	L	Ä	ĸ	w	G	Ŷ	õ	т	ÿ	N	P	G	D	õ	т	ī	Ŧ	s	R	v	N	D	G	P
Node 9	Ť	F	Е	R	ī.	R	Ť	R	L	Ğ	L	Ð	v	R	i	N	v	L		к	w	G	Ŷ	õ	T	Ŷ	N	P	G	Б	õ	т	ī	ī	s	R	÷	N	D	ā	P
Node 10	v	F	E	R	ī.	R	T	-	L	G	м	Б	v	R	i	N	v	L	А	ĸ	w	2	v	2	Ŧ	÷	N	P	a	2	*	T	ĩ	Ŧ	s	k	w	N	Б	G	- P
noue re	•	•	-	-	-		-	Ē	-	-		-	-	Ü		~*	•	-		-		•	-	•	-	-		-	э	•	0	-	-	-		~		-	-	•	•
																						1					1							1							1
	7							8							9							0					1							2							3
Site	2	3	4	5	6	8	9	0	2	3	5	6	7	8	0	1	2	3	4	8	9	1	2	4	6	7	0	3	4	5	7	8	9	1	2	3	5	6	7	9	0
Langur	G	A	v	D	А	Н	I	S	s	A	$\mathbf{L}$	Q	N	Ŋ	А	D	A	V	A	R	v	s	D	Q	I	R	v	R	Ν	Н	Q	N	к	v	s	Q	v	К	G	G	v
Baboon	G	А	V	Ν	А	Н	I	S	Ν	А	L	Q	D	N	т	D	A	V	А	R	v	S	D	Q	I	R	V	R	Ν	Н	Q	Ν	R	V	S	Q	V	Q	G	G	V
Human	G	A	V	Ν	А	Н	L	S	S	А	$\mathbf{L}$	Q	D	N	А	D	Α	V	А	R	V	R	D	Q	I	R	V	R	Ν	R	Q	Ν	R	V	R	Q	V	Q	G	G	V
Rat	R	А	К	Ν	А	G	I	Ρ	s	А	L	Q	D	D	т	Q	А	I	Q	R	v	R	D	Q	Ι	R	V	Q	R	Н	Κ	Ν	R	L	S	G	I	R	Ν	G	V
Cow	Ν	А	v	D	G	н	V	S	S	Е	М	Е	N	D	А	К	А	V	А	К	I	s	Е	Q	I	т	V	К	s	Н	R	D	Н	v	S	S	V	Ε	G	т	L
Horse	S	S	s	Ν	A	N	I	М	S	К	$\mathbf{L}$	D	Е	N	D	D	D	I	S	R	v	R	D	K	М	S	К	V	K	н	K	D	К	L	s	Е	$\mathbf{L}$	А	s	N	L
Node 7	G	А	v	N	λ	н	I	s	s	А	L	Q	D	N	А	D	λ	v	А	R	v	R	D	Q	I	R	v	R	N	н	a	N	R	v	s	0	v	0	G	G	v
Node 8	G	A	v	N	λ	н	I	s	s	A	L	Q	D	N	A	D	A	v	A	R	v	R	D	Q	I	R	v	R	N	н	ō	N	R	v	s	Q	v	Q	G	G	v
Node 9	G	A	v	N	A	н	I	s	s	A	L	Q	D	N	A	D	A	v	A	R	v	s	D	Q	I	R	v	R	N	н	Q	N	R	v	s	Q	v	Q	G	G	v
Node 10	?	A	v	N	A	н	I	s	s	e	L	e	D	N	A	D	A	v	A	R	v	R	D	Q	I	?	v	r	N	н	?	đ	R	v	s	q	v	q	G	?	1

The posterior probability for the reconstructed amino acid at an interior node is calculated according to (4) and is indicated using different typefaces: upper case bold for 0.9-1.0, upper case for 0.8-0.9, lower case bold 0.7-0.8, lower case 0.5-0.7, and question mark (?) for <0.5. Only variable sites in the sequence data are shown.



- Trait evolution
  - Reconstruction of ancestral traits
    - Proteins

### Inferring the palaeoenvironment of ancient bacteria on the basis of resurrected proteins

Eric A. Gaucher  $^1$  , J. Michael Thomson  $^{2}\star$  , Michelle F. Burgan  $^3$  & Steven A. Benner  $^{1,2,3}$ 

<sup>1</sup>NASA Astrobiology Institute, <sup>2</sup>Department of Anatomy and Cell Biology, College of Medicine, and <sup>3</sup>Department of Chemistry, University of Florida, Gainesville, Florida 32611-7200, USA

#### CHAPTER 2

Ancestral sequence reconstruction as a tool to understand natural history and guide synthetic biology: realizing and extending the vision of Zuckerkandl and Pauling

Eric A. Gaucher



Figure 1 The two unrooted universal trees used to reconstruct ancestral bacterial sequences. Archaea and Eukarya serve to provide a node within the bacterial subtree from which ancient sequences can be inferred. Thermophilic lineages are highlighted in bold. Aquificaceae subfamily not shown. a, Maximum likelihood topology used to reconstruct

the stem elongation factors from bacteria (ML-stem), or most recent common ancestor of bacteria, and the ancestral sequence for mesophilic lineages only (ML-meso). **b**, Alternative topology used to reconstruct the stem elongation factors from bacteria (Alt-stem).

- Trait evolution
  - Reconstruction of ancestral traits
    - Proteins

### Inferring the palaeoenvironment of ancient bacteria on the basis of resurrected proteins

Eric A. Gaucher  $^1$  , J. Michael Thomson  $^{2*}$  , Michelle F. Burgan  $^3$  & Steven A. Benner  $^{1,2,3}$ 

<sup>1</sup>NASA Astrobiology Institute, <sup>2</sup>Department of Anatomy and Cell Biology, College of Medicine, and <sup>3</sup>Department of Chemistry, University of Florida, Gainesville, Florida 32611-7200, USA

\* Present address: Department of Cell and Developmental Biology, University of North Carolina, Chapel Hill, North Carolina 27599, USA



**Figure 3** GDP-binding assay to test thermostability of ancestral and modern EF proteins. The amount of tritium-labelled GDP bound at 0 °C was subtracted from all other temperature values for a given protein. Shown is the relative amount of GDP bound compared with the amount bound at the optimal temperature for each protein.

### Gaucher et al. 2003. Nature

- Trait evolution
  - Reconstruction of ancestral traits
    - Behavior



### Female Responses to Ancestral Advertisement Calls in Túngara Frogs

Michael J. Ryan and A. Stanley Rand



- Trait evolution
  - Reconstruction of ancestral traits
    - Behavior



Fig. 2. Phonotactic responses of female túngara frogs (P. pustulosus) to conspecific (Con.) versus heterospecific (Het.) calls (8).

### Female Responses to Ancestral Advertisement Calls in Túngara Frogs

Michael J. Ryan and A. Stanley Rand

Phylogenetic techniques were used to estimate and reconstruct advertisement calls at ancestral nodes. These calls were used to investigate the degree of preference of female



Fig. 3. Phonotactic responses of female túngara frogs (*P. pustulosus*) to heterospecific calls versus no response, which includes no phonotaxis and a response to white noise (8). In only four of the 280 tests did females exhibit phonotaxis to noise.

Ryan & Rand. 1995. Science (see also Losos 1999 and Ryan & Rand 1999 Animal Behavior)

# What Questions Can You Ask?

- Testing trait correlation
  - Adaptation and natural selection







What happens when we replay the tape of evolution?

# What Questions Can You Ask?

- Testing trait correlation
  - Adaptation and natural selection







What happens when we replay the tape of evolution?

- Trait evolution
  - Patterns of trait evolution

OPEN O ACCESS Freely available online

🤎 PLoS one

### **Repeated Origin and Loss of Adhesive Toepads in Geckos**

Tony Gamble<sup>1,2</sup>, Eli Greenbaum<sup>3#</sup>, Todd R. Jackman<sup>3</sup>, Anthony P. Russell<sup>4</sup>, Aaron M. Bauer<sup>3</sup>\*

1 Department of Genetics, Cell Biology and Development, University of Minnesota, Minnesota, Minnesota, United States of America, 2 Bell Museum of Natural History, University of Minnesota, St. Paul, Minnesota, United States of America, 3 Department of Biology, Villanova University, Villanova, Pennsylvania, United States of America, 4 Department of Biological Sciences, University Department of Calgary, Calgary, Canada







Figure 1. Gecko phylogeny and the evolution of adhesive toepads. Maximum likelihood tree showing phylogenetic relationships among gecko genera. Toepad traits, including the presence of adhesive toepads, toepad shape and the presence of paraphalanges, are illustrated by colored squares on the tips of the branches (squares with two colors indicate polymorphism within the clade). Rectangles at internal nodes represent ancestral presence or absence probabilities of adhesive toepads interned using the 6-parameter binary-state speciation and extinction (BISSE) model. Details for lettered clades are presented in Table 1. Representative images illustrate a variety of gecko toepad morphologies. Single digits from representative gecko species illustrating the morphological diversity of paraphalangeal elements (in gray with stippling) are shown on the right. Clades enclosed in gray boxes are shown in greater detail in Figures 3 and 4. doi:10.1371/journal.pone.0039429.001

### Gamble et al. 2012. PLoS ONE

Origin and Loss of Adhesive Toepads in Geckos



Jønsson et al. 2008. Biology Letters

### Trait evolution

Patterns of trait evolution

**BIOINFORMATICS DISCOVERY NOTE** 

Vol. 28 no. 5 2012, pages 603–606 doi:10.1093/bioinformatics/bts008

**Phylogenetics** 

Advance Access publication January 11, 2012

# The evolution of nitrogen fixation in cyanobacteria

Natasha Latysheva<sup>1</sup>, Vivien L. Junker<sup>1</sup>, William J. Palmer<sup>1,2</sup>, Geoffrey A. Codd<sup>3</sup> and Daniel Barker<sup>1,\*</sup>

<sup>1</sup>Centre for Evolution, Genes and Genomics, School of Biology, University of St Andrews, St Andrews, Fife, KY16 9TH, UK, <sup>2</sup>Department of Genetics, University of Cambridge, Cambridge, CB2 3EH, UK and <sup>3</sup>Division of Molecular Microbiology, College of Life Sciences, University of Dundee, Dundee, DD1 4HN, UK Associate Editor: David Posada



Fig. 1. Rooted phylogeny of cyanobacterial species. Bootstrap support is 100% for all nodes except where shown (vertical percentages). Pie charts represent the posterior probability of the presence (black) and absence (white) of nitrogen fixation according to the Asymm.2 model of trait evolution. Grey crosses on branches indicate the most probable locations of trait loss according to stochastic mapping, and grey bars indicate the most probable sites of gain through horizontal transfer (frequency in mappings  $\geq$ 12%, excluding change and reversal on the same branch; frequencies shown as horizontal percentages).

### Latysheva et al. 2012. Bioinformatics





Figure 3 | Bayesian DNA gene tree of the kalilkrein toxin family. Multiple support values are given at key nodes in the following order: Bayesian DNA, ML DNA, Bayesian aa, ML aa; x indicates no support for the node in that analysis. Tips of the tree coloured in red indicate Toxicoferan sequences sourced from the venom gland and those coloured in blue indicate the ones sourced from non-venom gland tissues ('physiological' non-toxins). Pie charts represent the bpp of ancestral state reconstructions at that node: red = venom, blue = non-venom. The numbered codes for each sequence presented in the genetree represent GenBank GI accession numbers.

### Casewell et al. 2012. Nature Communications

### Trait evolution

Patterns of trait evolution

Ecology, 93(8) Supplement, 2012, pp. S151-S166 © 2012 by the Ecological Society of America

### Niche evolution across spatial scales: climate and habitat specialization in California Lasthenia (Asteraceae)

NANCY C. Emery,<sup>1,2,3</sup> Elisabeth J. Forrestel,<sup>1,4</sup> Ginger Jui, Michael S. Park,<sup>1</sup> Bruce G. Baldwin,<sup>1</sup> and David D. Ackerly<sup>1</sup>

<sup>1</sup>Department of Integrative Biology and Jepson Herbarium, University of California, Berkeley, California 94720 USA <sup>2</sup>Department of Biological Sciences and Botany and Plant Pathology, Purdue University, West Lafayette, Indiana 47907-2054 USA



Fig. 2 A representative ancestral state reconstruction of habitat affinities in *Lasthenia* using Lagrange (Ree and Smith 2008). The presented phylogeny was randomly selected from all trees containing the most commonly observed habitat transition rates (Table 3). Circled numbers represent independent transitions from terrestrial habitat into either vernal pool habitat or terrestrial/ aquatic habitat (see Fig. 1).



### Emery et al. 2012. Ecology

### Spiny frogs (Paini) illuminate the history of the Himalayan region and Southeast Asia

Jing Che<sup>a,b</sup>, Wei-Wei Zhou<sup>a</sup>, Jian-Sheng Hu<sup>a,c</sup>, Fang Yan<sup>a</sup>, Theodore J. Papenfuss<sup>b</sup>, David B. Wake<sup>b,1</sup>, and Ya-Ping Zhang<sup>a,d,1</sup>

<sup>a</sup>State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming 650223, People's Republic of China; <sup>b</sup>Department of Integrative Biology, Museum of Vertebrate Zoology, University of California, Berkeley, CA 94720; and <sup>c</sup>College of Life Sciences and <sup>d</sup>Laboratory for Conservation and Utilization of Bioresources, Yunnan University, Kunming 650091, People's Republic of China

Contributed by David B. Wake, June 15, 2010 (sent for review May 6, 2010)

Asian frogs of the tribe Paini (Anura: Dicroglossidae) range across frogs belonging to the tribe Paini, ranoid frogs endemic to Asia—







### Bizarre frogs rafted atop ancient continents

Well-muscled spiny frogs offer clues to evolution of continents, research finds

### Che et al. 2010. PNAS

- Trait evolution
  - Patterns of trait evolution

VOL. 154, NO. 2 THE AMERICAN NATURALIST AUGUST 1999

Using Phylogenies to Test Macroevolutionary Hypotheses of Trait Evolution in Cranes (Gruinae)



Arne Ø. Mooers,<sup>1,2,\*</sup> Steven M. Vamosi,<sup>1</sup> and Dolph Schluter<sup>1</sup>



Observed Change E. Free

**Figure 1:** Characterization of five models for the evolution of quantitative traits. Branch lengths represent the expected amount of change occurring along that lineage. *A*, Gradual model, in which change is correlated with time. This mirrors the inferred phylogenetic tree. *B*, Genetic distance model, in which each branch is proportional to the amount of genetic change that has occurred along it in the gene(s) used to make the tree. *C*, Speciational change model, in which change is correlated with speciation events. *D*, Nonhistorical model, in which there is no phylogenetic component to trait evolution, but each tip is equally divergent from all others. *E*, Free model, in which the branches represent the best fit of the observed data to the topology, rather than a priori hypotheses.

### Mooers et al. 1999. American Naturalist

- Trait evolution
  - Trait interactions



nature

# Pollinator shifts drive increasingly long nectar spurs in columbine flowers

Justen B. Whittall<sup>1</sup>† & Scott A. Hodges<sup>1</sup>





Whittall & Hodges. Nature. 2007

Vol 447 7 June 2007 dok10.1038/nature05857

Figure 3 | Phylogenetic analysis of pollination syndrome evolution in

- Trait evolution
  - Trait interactions



nature

### Pollinator shifts drive increasingly long nectar spurs in columbine flowers

Vol 447 7 June 2007 dok10.1038/nature05857

Justen B. Whittall<sup>1</sup>† & Scott A. Hodges<sup>1</sup>



Impact of traits on taxonomic diversification

VOL. 180, NO. 1 THE AMERICAN NATURALIST JULY 2012

Did Pollination Shifts Drive Diversification in Southern African *Gladiolus*? Evaluating the Model of Pollinator-Driven Speciation

Luis M. Valente,<sup>1,2,\*</sup> John C. Manning,<sup>3</sup> Peter Goldblatt,<sup>4</sup> and Pablo Vargas<sup>1</sup>



### Valente et al. 2012. American Naturalist

#### 88 The American Naturalist





Figure 2: Evolution of pollination systems in *Gladiolus*. The tree shown is the majority rule consensus from the MrBayes (Ronquist and Huelsenbeck 2003) analysis. Symbols above the nodes represent Bayesian posterior probabilities/maximum likelihood bootstrap values. Lozenges represent posterior probabilities above 95% and bootstrap values above 80%. Triangles represent posterior probabilities above 90% and bootstrap values above 60%. Pie charts at selected nodes summarize the results of the maximum likelihood character optimization analyses in Mesquite (Maddison and Maddison 2009), conducted for 1,000 highly probable Bayesian trees. Each chart shows the percentage of trees for which a given pollination system was reconstructed as ancestral for that node. Branches are colored according to a parsimony mapping of ancestral pollination systems conducted in Mesquite onto the tree shown in the figure. The crown nodes of the Cape radiation clade and the Mediterranean basin lineage are indicated by arrows.

### Valente et al. 2012. American Naturalist

### Trophic evolution and mammal diversification

### Tempo of trophic evolution and its impact on mammalian diversification

Samantha A. Price<sup>a,1,2</sup>, Samantha S. B. Hopkins<sup>b,1</sup>, Kathleen K. Smith<sup>c</sup>, and V. Louise Roth<sup>c</sup>

\*Department of Evolution and Ecology, University of California, Davis, CA 95616; \*Department of Geological Sciences, Clark Honors College, University of Oregon, Eugene, OR 97403; and \*Department of Biology, Duke University, Durham NC 27708–0338

Edited by David B. Wake, University of California, Berkeley, CA, and approved March 20, 2012 (received for review October 20, 2011)



**Fig. 2.** Plot of the posterior probability density of the parameter estimates of the all rates-free model (which was the best-fitting model from the maximum-likelihood analysis) analyzed using Bayesian MCMC methods on 10 of the 100 replicate phylogenies. (A) Depiction of the net diversification rate (speciation minus extinction) for all three trophic strategies and (B) the rate of transition between the trophic strategies.



## Coral reefs and fish diversification



Price et al. 2011. Ecology Letters

## Coral reefs and fish diversification







Figure 3 Model-averaged rates of ecomorphological evolution for coral reef (black bars) and non-coral reef (grey bars) labrids. Data includes all PC axes of the complete dataset (a) and the dataset which excludes the novel trophic strategies (b). Rates of evolution are the fastest on coral reefs for all morphological axes, however, in a model selection framework only PC1, PC4 and PC7 have strong support ( $\Delta$ AICc > 2) for the two-rate Brownian model in the complete dataset. Data are model-averaged estimates of the Brownian motion rate parameters with standard errors from a set of 500 character maps of reef/non-reef habitat.

### Price et al. 2011. Ecology Letters

### Ecological opportunity and anole diversification



ORIGINAL ARTICLE

doi:10.1111/j.1558-5646.2010.01026.x

### ECOLOGICAL OPPORTUNITY AND THE RATE OF MORPHOLOGICAL EVOLUTION IN THE DIVERSIFICATION OF GREATER ANTILLEAN ANOLES

D. Luke Mahler,<sup>1,2</sup> Liam J. Revell,<sup>3</sup> Richard E. Glor,<sup>4</sup> and Jonathan B. Losos<sup>1</sup>

<sup>1</sup>Department of Organismic and Evolutionary Biology, and Museum of Comparative Zoology, Harvard University, Cambridge, Massachusetts 02138

<sup>2</sup>E-mail: Imahler@oeb.harvard.edu

<sup>3</sup>National Evolutionary Synthesis Center, Durham, North Carolina 27705

<sup>4</sup>Department of Biology, University of Rochester, Rochester, New York 14627



# Outline

1. Insights from studies of character evolution

### 2. Requirements for studies of character evolution

- 3. Data exploration
  - Visualizing trait data on phylogenetic trees (R tutorial)
- 4. Trait evolution
  - ML and Bayesian methods for estimating ancestral states, rates, and models of evolution (R & BayesTraits tutorials)
  - Stochastic character mapping (R & SIMMAP tutorials)
  - Correlations among characters (BayesTraits tutorial)

### Why is it Necessary to Use the Phylogeny?

Vol. 125, No. 1

γ

The American Naturalist

January 1985

#### PHYLOGENIES AND THE COMPARATIVE METHOD

#### JOSEPH FELSENSTEIN

Department of Genetics SK-50, University of Washington, Seattle, Washington 98195

Submitted November 30, 1983; Accepted May 23, 1984



х

## Why is it Necessary to Use the Phylogeny?



х

# **Phylogenetic Signal**

- "tendency for related species to resemble each other more than they resemble species drawn at random from the [phylogenetic] tree" Blomberg & Garland 2002
- "Phylogenetic signal is a measure of the statistical dependence among species' trait values due to their phylogenetic relationships." Revell et al. 2008



# **Phylogenetic Signal**

- "tendency for related species to resemble each other more than they resemble species drawn at random from the [phylogenetic] tree" Blomberg & Garland 2002
- "Phylogenetic signal is a measure of the statistical dependence among species' trait values due to their phylogenetic relationships." Revell et al. 2008


# Phylogenetic Signal

- "tendency for related species to resemble each other more than they resemble species drawn at random from the [phylogenetic] tree" Blomberg & Garland 2002
- "Phylogenetic signal is a measure of the statistical dependence among species' trait values due to their phylogenetic relationships." Revell et al. 2008
- On the one hand...
  - Phylogenetic signal suggests that phylogenetic analyses are necessary.
- On the other hand...
  - Lack of phylogenetic signal means that the history of traits cannot be inferred.

## What Do You Need to Investigate Trait Evolution?

- 1. Phylogeny for the taxa of interest
- 2. Trait values for the taxa of interest
- 3. A model to describe trait evolution

## What Do You Need to Investigate Trait Evolution?

1. Phylogeny for the taxa of interest



## Phylogenetic Trees for Comparative Analyses

• Preferably trees with meaningful branch lengths and sets of trees that account for uncertainty in branch length and topology.



## What Do You Need to Investigate Trait Evolution?

1. Phylogeny for the taxa of interest

2. Trait values for the taxa of interest



- Discretely-coded traits
  - Intrinsically discrete traits



- Discretely-coded traits
  - Intrinsically discrete traits





Wings



No wings



Aquatic



Terrestrial

- Discretely-coded traits
  - Discretize continuous traits



Fig. 1. Discrete approximation to the gamma distribution  $G(\alpha,\beta)$ , with  $\alpha = \beta = \frac{1}{2}$ . Four categories are used to approximate the continuous distribution, with equal probability for each category. The three boundaries are 0.1015, 0.4549, and 1.3233, which are the percentage points corresponding to  $p = \frac{1}{4}, \frac{2}{4}, \frac{3}{4}$ . The means of the four categories are 0.0334, 0.2519, 0.8203, 2.8944. The medians are 0.0247, 0.2389, 0.7870, 2.3535, and these are scaled to get 0.0291, 0.2807, 0.9248, and 2.7654, so that the mean of the discrete distribution is one.

Yang 1994

- Discretely-coded traits
  - Discretize continuous traits



Pit Viper











Pupil shape in snakes

Brischoux et al. 2010

Oxyuranus scutellatus Oxyuranus microlepidotu

# Outline

- 1. Insights from studies of character evolution
- 2. Requirements for studies of character evolution

#### 3. Data exploration

• Visualizing trait data on phylogenetic trees (R tutorial)

#### 4. Trait evolution

- ML and Bayesian methods for estimating ancestral states, rates, and models of evolution (R & BayesTraits tutorials)
- Stochastic character mapping (R & SIMMAP tutorials)
- Correlations among characters (BayesTraits tutorial)

JOURNAL<sup>of</sup> OOLOGICAL SYSTEMATICS

© 2011 Blackwell Verlag GmbH

Accepted on 21 October 2011 J Zool Syst Evol Res doi: 10.1111/j.1439-0469.2011.00646.x

<sup>1</sup>Department of Biology, University of Central Florida, Orlando, FL, USA; <sup>2</sup>Department of Ecology and Evolutionary Biology, Corson Hall, Cornell University, Ithaca, NY, USA

The serpent and the egg: unidirectional evolution of reproductive mode in vipers?

Allyson M. Fenwick<sup>1</sup>, Harry W. Greene<sup>2</sup> and Christopher L. Parkinson<sup>1</sup>



GA1469 [RM] © www.visualphotos.com

#### Viviparous





511317 [RM] @ www.visualphotos Oviparous



Table 1. Maximum likelihood models tested. All models have some support under AIC, optimal model is bold. Parameter values are averages taken over the sample of 600 trees including standard deviations. Eggs symbolize rates under oviparity; snakes symbolize viviparity

Model			Character state			
	Parameters	LnL	0- <i>2</i> ~	e~-0	AIC	ΔΑΙΟ
Variable rates Equal rates Dollo	2: $q_{\text{ovip to vivip}} \neq q_{\text{vivip to ovip}}$ 1: $q_{\text{ovip to vivip}} = q_{\text{vivip to ovip}}$ 1: $q_{\text{vivip to ovip}} = 0$	$\begin{array}{r} -49.30 \pm 0.83 \\ -53.42 \pm 0.84 \\ -54.21 \pm 1.45 \end{array}$	$\begin{array}{r} \textbf{0.0340} \ \pm \ \textbf{0.0037} \\ \textbf{0.0136} \ \pm \ \textbf{0.0014} \\ \textbf{0.0487} \ \pm \ \textbf{0.0057} \end{array}$	$\begin{array}{c} \textbf{0.00323} \ \pm \ \textbf{0.000} \\ \textbf{0.0136} \ \pm \ \textbf{0.0014} \\ \textbf{0} \end{array}$	<b>102.6</b> 108.8 110.4	<b>0</b> 6.244 7.824

AIC, Akaike information criterion.

Table 2. BiSSE models tested. Model name includes number of parameters for that model. Model 5d has some support under AIC, optimal model is bold. Parameters not mentioned in models were allowed to vary independently of each other. Parameter values are harmonic means taken over the sample of 600 trees. Eggs symbolize rates under oviparity; snakes symbolize viviparity

			Speci rate	ation $(\lambda)$	Extinctio	n rate (µ)	Charact	ter state n rate (q)		
Model	Parameters	LnL	٢	~	٢	~~	0-2-	Em-0	AIC	ΔΑΙΟ
6	All rates variable	646.7	0.113	0.196	2.33e-6	3.62e-7	0.0282	0.0031	-1283	19.13
5a	Speciation rates equal ( $\lambda_{ovip} = \lambda_{vivip}$ )	647.8	0.193	0.193	3.874e-6	6.350e-6	6.770e-7	0.0102	-1286	16.39
5b	Extinction rates equal ( $\mu_{ovip} = \mu_{vivip}$ )	647.8	0.093	0.193	4.925e-6	4.925e-6	1.355e-6	0.0104	-1286	16.45
5c	Character state transition rates equal $(a_{\text{min}} + a_{\text{min}}) = a_{\text{min}} + a_{\text{min}}$	646.7	0.113	0.197	2.635e-6	1.350e-6	0.0031	0.0031	-1283	18.65
5d	Dollo transition rates $(q_{vivin to ovin} = 0)$	654.0	0.122	0.196	1.039e-6	1.042e-6	0.0416	1.000e-7	-1298	4.00
4a	$\lambda_{\text{ovin}} = \lambda_{\text{vivin}}, \mu_{\text{ovin}} = \mu_{\text{vivin}}$	646.7	0.196	0.196	2.493e-6	2.493e-6	0.0282	0.0031	-1285	16.65
3a	$\lambda_{\rm ovip} = \lambda_{\rm vivip},  \mu_{\rm ovip} = \mu_{\rm vivip},$	646.7	0.197	0.197	1.159e-6	1.159e-6	0.003	0.003	-1287	14.65
3b	$q_{\text{ovip to vivip}} = q_{\text{vivip to ovip}}$ $\lambda_{\text{ovip}} = \lambda_{\text{vivip}}, \mu_{\text{ovip}} = \mu_{\text{vivip}}, q_{\text{vivip to ovip}} = 0$	654.0	0.196	0.196	1.328e-6	1.328e-6	0.042	1.000e-7	-1302	0

AIC, Akaike information criterion.



#### Visualizing body size in mammals



**Figure 1.** Joint reconstruction of divergence times and adult body mass (in grams) under the covariant model. Horizontal bars indicate 95% credibility intervals for node ages. Dark- and light-shaded disks indicate 95% credibility intervals for ancestral body masses. In the case of extant taxa for which body size was missing, credibility intervals are reconstructed as for internal nodes.

## Visualizing morphological data from PCA



#### Egg laying strategies in darter fish



**Fig. 1** Time-calibrated maximum-clade credibility phylogeny containing 245 of 248 recognized darter species inferred from a threegene DNA sequence data set (see Near *et al.* 2011). Species that were included in our analysis and thus for which we have egg-deposition mode data are marked with a black circle.

Kelly et al. 2012. J. Evol. Biology

#### Kelly et al. 2012. J. Evol. Biology



## Egg laying strategies in darter fish

# Diet in mammals



Price et al. 2012. PNAS



#### ELEVATED RATES OF MORPHOLOGICAL AND FUNCTIONAL DIVERSIFICATION IN REEF-DWELLING HAEMULID FISHES

doi:10.1111/j.1558-5646.2012.01773.x

Samantha A. Price,<sup>1,2,3</sup> Jose J. Tavera,<sup>3,4</sup> Thomas J. Near,<sup>5</sup> and Peter. C. Wainwright<sup>1</sup> <sup>1</sup>Department of Evolution and Ecology, University of California Davis, Davis, California 95616 <sup>2</sup>E-mail: saprice@ucdavis.edu

<sup>4</sup>Centro de Investigaciones Biológicas del Noroeste, S.C., La Paz, B.C.S., México <sup>5</sup>Department of Ecology & Evolutionary Biology, Yale University, New Haven, Connecticut 06520



## Lunch break

# Outline

- 1. Insights from studies of character evolution
- 2. Requirements for studies of character evolution
- 3. Data exploration
  - Visualizing trait data on phylogenetic trees (R tutorial)

#### 4. Trait evolution

- ML and Bayesian methods for estimating ancestral states, rates, and models of evolution (R & BayesTraits tutorials)
- Stochastic character mapping (R & SIMMAP tutorials)
- Correlations among characters (BayesTraits tutorial)

## What Do You Need to Investigate Trait Evolution?

- 1. Phylogeny for the taxa of interest
- 2. Trait values for the taxa of interest
- 3. A model to describe trait evolution

## Parsimony

The Fitch-Wagner down-pass/up-pass method





## **Alternative Parsimony Models**

 Unordered (Fitch parsimony) Ordered (Wagner parsimony) 0 - 1 - 2 - 3 - 4 Irreversible (Camin-Sokal parsimony)  $0 \rightarrow 1 \rightarrow 2 \rightarrow 3 \rightarrow 4$ Once only Site Site absent Dollo present May happen more than once • Step matrix

## Limitations of Parsimony

- Doesn't take branch lengths into account (limited to one change per branch, regardless of how long)
- Performs poorly with rapidly evolving traits, tending to push divergence toward the tips of the tree
- Underestimates variance of rate parameters and ancestral reconstructions and doesn't provide estimates of error

ANIMAL BEHAVIOUR, 1999, 58, 1319-1324

			Article No. anbe.1999.1261, available online at http://www.idealibrary.com on IDELL®		
1999	POINTS OF VIEW	665	Æ		
Syst. Biol. 48(3):665–674, 1999			COMMENTARIES		
Some Limitations of Ancestral Character-State Reconstruction When Testing Evolutionary Hypotheses			Uncertainty in the reconstruction of ancestral character states and limitations on the use of phylogenetic comparative methods		
	CLIFFORD W. CUNNINGHAM				
Zoology Department, Duke University, Durham, North Carolina 27708-0325, USA; E-mail: cliff@duke.edu		nail: cliff@duke.edu	JONATHAN B. LOSOS Department of Biology, Washington University, St Louis		
			(Received 14 January 1999; initial acceptance 24 May 1999; final acceptance 19 July 1999; MS. number: AS-1236)		

## Maximum Likelihood Analysis of Trait Evolution

No. 6

Detecting correlated evolution on phylogenies: a general method for the comparative analysis of discrete characters

#### MARK PAGEL

School of Mathematical Sciences, Queen Mary and Westfield College, University of London, London El 4NS, U.K.

Pagel 1994. Proc. Royal Soc.

#### **EVOLUTION**

INTERNATIONAL JOURNAL OF ORGANIC EVOLUTION

PUBLISHED BY

December 1997

THE SOCIETY FOR THE STUDY OF EVOLUTION

Vol. 51

Evolution, 51(6), 1997, pp. 1699-1711

#### LIKELIHOOD OF ANCESTOR STATES IN ADAPTIVE RADIATION

DOLPH SCHLUTER,<sup>1,2</sup> TREVOR PRICE,<sup>3</sup> ARNE Ø. MOOERS<sup>1,4</sup> AND DONALD LUDWIG<sup>1,5</sup> <sup>1</sup>Department of Zoology and Centre for Biodiversity Research, University of British Columbia, Vancouver, British Columbia V6T 1Z4, Canada <sup>2</sup>E-mail: schluter@zoology.ubc.ca <sup>3</sup>Biology Department O-116, University of California—San Diego, La Jolla, California 92093 E-mail: tprice@ucsd.edu <sup>4</sup>E-mail: mooers@bio.uva.nl <sup>5</sup>E-mail: ludwig@zoology.ubc.ca

Schluter et al. 1997



#### Maximum Likelihood Analysis of Trait Evolution

Detecting correlated evolution on phylogenies: a general method for the comparative analysis of discrete characters

MARK PAGEL

School of Mathematical Sciences, Queen Mary and Westfield College, University of London, London E1 4NS, U.K.

- 1. The Markov model for evolution of discretely coded traits
- 2. Using maximum likelihood to estimate rates of character evolution and ancestral character states

$$P_{ij}(t + dt) = P_{ii}(t) q_{ij} dt + P_{ij}(t) (1 - q_{ji}) dt,$$
(1)

$$\begin{aligned} \boldsymbol{P}(t+dt) &= \boldsymbol{P}(t) \left( \boldsymbol{I} + \boldsymbol{Q} dt \right), \end{aligned} \tag{2} \\ \begin{bmatrix} 1 - P_{01} \left( t + dt \right) & P_{01} \left( t + dt \right) \\ P_{10} \left( t + dt \right) & 1 - P_{10} \left( t + dt \right) \end{bmatrix} \\ &= \begin{bmatrix} P_{00}(t) & P_{01}(t) \\ P_{10}(t) & P_{11}(t) \end{bmatrix} \begin{bmatrix} (1 - q_{01}) dt & q_{01} dt \\ q_{10} dt & (1 - q_{10}) dt \end{bmatrix}, \end{aligned}$$

$$\begin{split} L(I) &= \prod_{s=0}^{X \text{ and } Y} \sum_{s=0}^{1} \sum_{s=0}^{1} \sum_{s=0}^{1} \sum_{s=0}^{1} \sum_{s=0}^{1} P(s8, s9, t8) \\ &\times P(s3, s8, t3) \ P(s7, s8, t7) \ P(s5, s7, t5) \\ &\times P(s4, s7, t4) \ P(s6, s9, t6) \\ &\times P(s1, s6, t1) \ P(s2, s6, t2), \end{split}$$

$$L(I) = \prod_{s=0}^{X \text{ and } Y} \sum_{s=0}^{1} \left[ \left( \sum_{s=0}^{1} P(s8, s9, t8) P(s3, s8, t3) \right) \right] \\ \times \sum_{s=0}^{1} P(s7, s8, t7) P(s5, s7, t5) P(s4, s7, t4) \\ \times \sum_{s=0}^{1} P(s6, s9, t6) P(s1, s6, t1) \\ \times P(s2, s6, t2) \right].$$
(6)

for branches leading to the tips of the tree because u will be a constant. For all other branches both u and v can vary from 0 to 1 so four possibilities arise in each variable. Then the likelihood for this tree will be given by,

$$\begin{split} L(I) &= \prod_{s=0}^{X \text{ and } Y} \sum_{s=0}^{1} \sum_{s=0}^{1} \sum_{s=0}^{1} \sum_{s=0}^{1} \sum_{s=0}^{1} P(s8, s9, t8) \\ &\times P(s3, s8, t3) \ P(s7, s8, t7) \ P(s5, s7, t5) \\ &\times P(s4, s7, t4) \ P(s6, s9, t6) \\ &\times P(s1, s6, t1) \ P(s2, s6, t2), \end{split}$$

where the summations are from 0 to 1 at each node. This likelihood is a function of four parameters, two for the X variable and two for the Y variable. One then searches the likelihood surface to find the maximum likelihood solution to L(I). This is the set of four parameters that makes L(I) largest.

Although the likelihood in equation (5) is found over all possible assignments of the ancestral character states, this is not strictly necessary. One may have prior information or beliefs about the ancestral values at the root of the tree. In such instances, the root of the tree

Proc. R. Soc. Lond. B (1994)

Felsenstein (1993) encountered this problem in connection with estimating tree topologies and developed his 'pruning' algorithm (1983; pp. 251–252) to reduce the computational effort. Felsenstein's pruning algorithm can be applied in this instance (I refer the reader to his article for further details) by recognizing that the four nested summations in (5) can be distributed along the product terms:

$$\begin{split} L(I) &= \prod_{s_{0}=0}^{X \text{ and } Y} \sum_{s_{0}=0}^{1} \left[ \left( \sum_{s_{0}=0}^{1} P(s8, s9, t8) \ P(s3, s8, t3) \right) \\ &\times \sum_{s_{0}=0}^{1} P(s7, s8, t7) \ P(s5, s7, t5) \ P(s4, s7, t4) \right) \\ &\times \sum_{s_{0}=0}^{1} P(s6, s9, t6) \ P(s1, s6, t1) \\ &\times P(s2, s6, t2) \right]. \end{split}$$
(6)

No expression of the form P(u, v, t) can appear to the left of a summation over u. Applying the pruning algorithm makes the computing task linear rather than exponential in n: that is, proportional to 4(n-1).

for branches leading to the tips of the tree because u will be a constant. For all other branches both u and v can vary from 0 to 1 so four possibilities arise in each variable. Then the likelihood for this tree will be given



where the summations are from 0 to 1 at each node. This likelihood is a function of four parameters, two for the X variable and two for the Y variable. One then searches the likelihood surface to find the maximum likelihood solution to L(I). This is the set of four parameters that makes L(I) largest.

Although the likelihood in equation (5) is found over all possible assignments of the ancestral character states, this is not strictly necessary. One may have prior information or beliefs about the ancestral values at the root of the tree. In such instances, the root of the tree

Proc. R. Soc. Lond. B (1994)

Felsenstein (1993) encountered this problem in connection with estimating tree topologies and developed his 'pruning' algorithm (1983; pp. 251–252) to reduce the computational effort. Felsenstein's pruning algorithm can be applied in this instance (I refer the reader to his article for further details) by recognizing that the four nested summations in (5) can be distributed along the product terms:



No expression of the form P(u, v, t) can appear to the left of a summation over u. Applying the pruning algorithm makes the computing task linear rather than exponential in n: that is, proportional to 4(n-1).

## Advantages of Maximum Likelihood

- Incorporates branch length information and permits multiple changes along a single branch
- Superior performance with rapidly evolving traits
- Provides estimates of variance and error



## The Markov Model

- 1. Probability of change at a point in time along any branch of the tree depends only on the character state at that time, not on prior character states
- 2. Transitions along each branch are independent of changes elsewhere on the tree
- 3. Rates of change are constant throughout time and along branches
- 4. Rates are estimated only from the tree and the data available for extant species, not on prior knowledge or belief



## Parameters for Likelihood Evaluation

 $\mathbf{q}_{01}$  = rate at which character changes from 0 to 1 over a short interval *dt* 

 $\mathbf{q}_{10}$  = rate at which character changes from 1 to 0 over a short interval *dt* 

 $t_i$  = branch lengths

 $s_i = ancestors$ 

 $X_N$  = set of possible ancestral reconstructions



## Parameters for Likelihood Evaluation



#### Likelihood Calculation

 $\mathbf{q}_{01}$  = rate at which character changes from 0 to 1 over a short interval *dt* 

 $\mathbf{q}_{10}$  = rate at which character changes from 1 to 0 over a short interval *dt* 

 $t_i$  = branch lengths

 $s_i = ancestors$ 

$$L(I) = \sum_{s=0}^{1} \sum_{s=0}^{1} \sum_{s=0}^{1} \sum_{s=0}^{1} \sum_{s=0}^{1} \sum_{s=0}^{1} P(s8, s9, t8)$$

x P(s3,s8,t3)P(s7,s8,t7)P(s5,s7,t5) x P(s4,s7,t4)P(s6,s9,t6) x P(s1,s6,t1)P(s2,s6,t2)



(5)


n	# Calculations
5	16
10	512
50	5.63E+14
100	6.3383E+29

#### Felsenstein's Pruning Algorithm

$$L(I) = \sum_{s \neq 0} \sum_{s \neq 0} \sum_{s \neq 0} P(s8, s9, t8)$$

$$x P(s3, s8, t3) P(s7, s8, t7) P(s5, s7, t5) = (5)$$

$$x P(s4, s7, t4) P(s6, s9, t6)$$

$$x P(s1, s6, t1) P(s2, s6, t2)$$

$$L(I) = \sum_{s \neq 0}^{1} \left( \sum_{s \neq 0}^{1} P(s8, s9, t8) P(s3, s8, t3) \right)^{(5)}$$

$$x \sum_{s \neq 0}^{1} P(s7, s8, t7) P(s5, s7, t5) P(s4, s7, t4)$$

$$x \sum_{s \neq 0}^{1} P(s6, s9, t6) P(s1, s6, t1) \left[ x P(s2, s6, t2) \right]$$

$$(6)$$



$$L(I) = \sum_{s=0}^{1} [(\sum_{s=0}^{1} P(s8, s9, t8)) P(s3, s8, t3)]$$
  

$$x \sum_{s=0}^{1} P(s7, s8, t7) P(s5, s7, t5) P(s4, s7, t4))$$
  

$$x \sum_{s=0}^{1} P(s6, s9, t6) P(s1, s6, t1) [$$
  

$$x P(s2, s6, t2)]$$

 $2^{n-1}$  calculations v. 2 (n-1)

n	# Calculations	# Calculations
5	16	8
10	512	18
50	5.63E+14	98
100	6.3383E+29	198



#### The Transition Matrix

• Exponentiate the rate matrix (Q) times the length of the interval to obtain the probabilities over longer intervals

$$\mathbf{Q} = \begin{pmatrix} 0 & 1 \\ -q & q \\ q & -q \end{pmatrix}$$
$$\mathbf{P}(t) = \mathbf{e}^{-\mathbf{Q}t} = \begin{bmatrix} P & P \\ P & P \end{bmatrix}$$

## Probability of Change

#### $P_{ij}(t+dt) = \text{probability of change from } i \text{ to } j \text{ over}$ interval t+dt =

Probability that *i* remains *i* over *t* Probability that *i* goes to *j* over *t* 

$$P_{ii}(t)q_{ij}dt + P_{ij}(t)(1-q_{ji})dt$$

Rate for *i* to *j* over *t* 

Rate for Staying *j* 

	0 (yellow)	1 (blue)
0 (yellow)	0.9	0.1
1 (blue)	0.1	0.9



	0 (yellow)	1 (blue)
0 (yellow)	0.9	0.1
1 (blue)	0.1	0.9

$$(S_7) \rightarrow (4) = 0.10$$
  
 $(S_7) \rightarrow (5) = 0.10$   
 $(S_7) = 0.10 \times 0.10 = 0.01$ 



	0 (yellow)	1 (blue)
0 (yellow)	0.9	0.1
1 (blue)	0.1	0.9

$$(57) \rightarrow (4) = 0.90$$
  
 $(57) \rightarrow (5) = 0.90$   
 $(57) = 0.90 \times 0.90 = 0.81$ 



	0 (yellow)	1 (blue)
0 (yellow)	0.9	0.1
1 (blue)	0.1	0.9



$$\begin{array}{c} \hline S_8 \\ \longrightarrow \\ \hline 0.01 \\ \hline 0.81 \\ \hline 0.90 \\ x \\ \hline 0.01 \\ \hline 0.81 \\ \hline 0.091 \\ \hline 0.001 \\ \hline 0.81 \\ \hline 0.091 \\ \hline 0.001 \\$$

	0 (yellow)	1 (blue)
0 (yellow)	0.9	0.1
1 (blue)	0.1	0.9



$$S_{8} \longrightarrow 0.01 \ 0.81 = 0.10 \times 0.01 + 0.90 \times 0.81 = 0.001 + 0.729 = 0.730$$
  

$$S_{8} \longrightarrow 3 = 0.10$$
  

$$S_{8} = 0.730 \times 0.10 = 0.073$$

	0 (yellow)	1 (blue)
0 (yellow)	0.9	0.1
1 (blue)	0.1	0.9



$$\begin{array}{c} \hline \mathbf{S}_{6} & \longrightarrow & 1 \\ \hline \mathbf{S}_{6} & \longrightarrow & 1 \\ \hline \mathbf{S}_{6} & \longrightarrow & 1 \\ \hline \mathbf{S}_{6} & = & 0.90 \\ \hline \mathbf{S}_{6} & = & 0.81 \end{array}$$

	0 (yellow)	1 (blue)
0 (yellow)	0.9	0.1
1 (blue)	0.1	0.9



$$S_{6} \rightarrow 2 = 0.10$$
  
 $S_{6} \rightarrow 2 = 0.10$   
 $S_{6} \rightarrow 2 = 0.10$   
 $S_{6} = 0.10 \times 0.10 = 0.01$ 

	0 (yellow)	1 (blue)	
0 (yellow)	0.9	0.1	$t_1$ $t_2$ $s_6$ 0.81 0.0
1 (blue)	0.1	0.9	
			$t_6 \setminus$



5

 $t_5$ 

 $(S_8)$ 

0.81

 $t_7$ 

0.073

 $(S_9)$ 

 $(S_{7})$ 

 $t_4$ 

0.01

 $t_3$ 

0.0585

0.081

 $t_8$ 

0.0049



#### Interpreting Ancestral Reconstructions



	0 (yellow)	1 (blue)
0 (yellow)	0.9	0.1
1 (blue)	0.1	0.9

$$(S_9) = 0.0802 \times 0.730 = 0.0585$$

 $s_9$  = 0.0657 x 0.0747 = 0.0049

#### $\sum \pi L_0$

lnL = ln(0.0585 x 0.5 + 0.0049 x 0.5) = ln(0.029 + 0.002) = ln(0.031) = -3.459



#### Interpreting Ancestral Reconstructions



 $L(q_{01}, q_{10}) = \sum P(s_1, s_2 \dots s_N)$ 

Obtain rate parameter(s) with highest likelihood among all possible reconstructions for a given tree topology and branch lengths.

# Justification for the Markov Model & Some Potential Problems

- Although not necessarily the most realistic models possible, Markov processes are "simple ways to represent unpredictability." - Schluter et al. 1997
- Assumption of rate constancy cited by Schluter et al. 1997 as "the most glaring weakness," but other problems exist as well
  - Good news: new methods permit alternative models
- Assumes that the tree topology and branch lengths are known with certainty
  - More good news: Bayesian methods to the rescue!
- Assumption that traits do not drive diversification.
  - BiSSE model and extensions

## **Another Important Limitation**

- May not be able to accurately estimate multiple rates (e.g., q01 v. q10)
- "Limiting the number of parameters is more crucial when estimating ancestor states than when estimating phylogenetic trees because here we are interested in a single character. In contrast, likelihood estimates of trees use information on many characters (i.e., base pairs) simultaneously."
- Often assume a single rate



## Examples

- Marginal likelihoods and rates
- Ancestral reconstruction
- Rate comparison and variation

## **Bayesian Analysis of Trait Evolution**

Syst. Biol. 53(5):673–684, 2004 Copyright © Society of Systematic Biologists ISSN: 1063-5157 print / 1076-836X online DOI: 10.1080/10635150490522232

#### **Bayesian Estimation of Ancestral Character States on Phylogenies**

MARK PAGEL, ANDREW MEADE, AND DANIEL BARKER

School of Animal and Microbial Sciences, University of Reading, Whiteknights, Reading RG6 6AJ England; E-mail: m.pagel@rdg.ac.uk (M.P.)

Pagel et al. 2004. Systematic Biology

## **Bayesian Analysis of Trait Evolution**



## **Bayes Theorem for Tree Reconstruction**

- T = tree
- **s** = aligned sequences



## **Bayes Theorem for Rate Coefficients**

- **Q** = set of rate coefficients
- **D** = dataset of trait values for species in our tree



#### Accounting for Phylogenetic Uncertainty

- **Q** = set of rate coefficients
- **D** = dataset of trait values for species in our tree



## Accounting for Phylogenetic Uncertainty

- **Q** = set of rate coefficients
- **D** = dataset of trait values for species in our tree



#### The Ancestral States

 $s_{ij} = j$  is the character state at node *i*  j = 0 or 1 for a binary trait  $p(s_{ij} | D)_{i \in T} =$  Posterior probability of *j* at node *i* given that node *i* appears in the tree ( $i \in T$ )

$$p(s_{ij}|D)_{i\in T} = \frac{\int_{T}}{\int_{T}}$$
(5)

#### The Ancestral States

This equation only works when the node is present.

This is problem because we're going to end up overestimating  $p(s_{ij}|D)$ .

$$p(s_{ij}|D)_{i\in T} = \frac{\int_{T}}{\int_{T}}$$
(5)

#### The Ancestral States

This equation only works when the node is present.

This is problem because we're going to end up overestimating  $p(s_{ij}|D)$ .

$$p(s_{ij}|D)_{i\in T} = \frac{\int_{T}}{\int_{T}}$$
(5)  
$$p(s_{ij}|D) = \frac{\int_{T}}{\int_{T}}$$
(6)

## **Bayesian MCMC for Character Reconstruction**

- 1. Start with random parameters and tree sampled from posterior distribution of Bayesian phylogenetic analysis.
- 2. At each generation propose a new combination of rate parameters and a new tree from the posterior of our previous phylogenetic analyses

## **Bayesian MCMC for Character Reconstruction**

- 1. Start with random parameters and tree sampled from posterior distribution of Bayesian phylogenetic analysis.
- 2. At each generation propose a new combination of rate parameters and a new tree from the posterior of our previous phylogenetic analyses
- 3. If new tree has higher posterior probability, accept the move
- 4. If the new tree has a lower probability, accept with probability new/current
- 5. Save parameters every n generation



#### The Acceptance Ratio

$$R = \frac{Pr(T_1 \mid D)}{Pr(T_2 \mid D)}$$

$$Pr(T_{1} | D) = \frac{Pr(T_{1}) Pr(D | T_{1})}{\sum_{T} Pr(T_{i}) Pr(D | T_{i})} \qquad Pr(T_{2} | D) = \frac{Pr(T_{2}) Pr(D | T_{2})}{\sum_{T} Pr(T_{i}) Pr(D | T_{i})}$$
$$Pr(T_{1}) Pr(D | T_{1})$$
$$R = \frac{Pr(T_{1}) Pr(D | T_{1})}{Pr(T_{2}) Pr(D | T_{2})} \qquad R = \frac{Pr(T_{1}) Pr(D | T_{1})}{Pr(T_{2}) Pr(D | T_{2})}$$

# A Worked Example

- The evolution of pancreatic ribonuclease
- Shift from G to D at the base of the ruminants
- Previous analyses support this scenario, but with considerable uncertainty







#### Priors for Bayesian Character Reconstruction

• "Priors are the soft underbelly of Bayesian analyses." - BayesTraits manual



#### Priors for Bayesian Character Reconstruction

- "Priors are the soft underbelly of Bayesian analyses." BayesTraits manual
- Because we're concerned with only a single character, the priors can have a stronger impact than they might in a dataset with hundreds or thousands of nucleotides.
   Small sample
   Large sample


### Priors for Bayesian Character Reconstruction

- "Priors are the soft underbelly of Bayesian analyses." BayesTraits manual
- Because we're concerned with only a single character, the priors can have a stronger impact than they might in a dataset with hundreds or thousands of nucleotides.
- In Pagel et al.'s worked example, the uniform or uninformative prior produces poor log-likelihood scores and rate posteriors with large variances and means.

TABLE 2. Posterior distributions derived from three different prior distributions. All priors are beta distributed. The uniform prior assumes that the two parameters of the beta distribution are  $\alpha = \beta = 1$ ; the 'maximum likelihood' prior bases the beta distribution on the mean and variance across trees of the maximum likelihood values of the rate parameters; the 'likelihood surface' prior bases the beta distribution on the mean and variance of the likelihood surface of the rate parameters on the consensus tree.

	Average $\pm$ SD of posterior distribution			
Prior	q <sub>GD</sub>	q <sub>DG</sub>	Log-likelihood	
Uniform q <sub>GD</sub> , q <sub>DG</sub> : 0–100	$11.47\pm6.28$	$13.28\pm9.13$	$-16.63 \pm 1.50$	
Maximum likelihood $q_{GD}$ : beta(5.82 ± 0.75) $q_{DG}$ : beta(6.63 ± 1.17)	$5.82 \pm 0.90$	$6.63 \pm 1.01$	$-15.22 \pm 0.59$	
Likelihood surface $q_{GD}$ , $q_{DG}$ : beta(8.2 ± 5.39)	$7.42\pm3.48$	8.05 ± 3.77	$-15.82 \pm 0.89$	

### Priors for Bayesian Character Reconstruction

- First calculate the rate parameter of the model on each tree in the dataset using maximum likelihood and derive a prior that emulates this distribution
- The "empirical Bayes estimator"

TABLE 2. Posterior distributions derived from three different prior distributions. All priors are beta distributed. The uniform prior assumes that the two parameters of the beta distribution are  $\alpha = \beta = 1$ ; the 'maximum likelihood' prior bases the beta distribution on the mean and variance across trees of the maximum likelihood values of the rate parameters; the 'likelihood surface' prior bases the beta distribution on the mean and variance of the likelihood surface of the rate parameters on the consensus tree.

	Average $\pm$ SD of posterior distribution			
Prior	$q_{GD}$	q <sub>DG</sub>	Log-likelihood	
Uniform q <sub>GD</sub> , q <sub>DG</sub> : 0-100	$11.47 \pm 6.28$	$13.28\pm9.13$	$-16.63 \pm 1.50$	
Maximum likelihood	$5.82 \pm 0.90$	$6.63 \pm 1.01$	$-15.22 \pm 0.59$	
$q_{GD}$ : beta(5.82 ± 0.75)				
$q_{DG}$ : beta(6.63 ± 1.17)				
Likelihood surface q <sub>GD</sub> ,	$7.42 \pm 3.48$	$8.05 \pm 3.77$	$-15.82 \pm 0.89$	
$q_{DG}$ : beta(8.2 ± 5.39)				

# Fine Tuning Bayesian Inference of Character Evolution

- Obtaining a suitable estimate of the prior
- Ensuring proper mixing during MCMC
- Assessing burn-in
- Assessing variation among independent runs

# Fine Tuning Bayesian Inference of Character Evolution: MCMCMC?

- "Our (unpublished) experience is that MCMCMC is of limited value in a phylogenetic context. Swapping of states is rare before convergence when the chains might be in different regions of the universe; and yet it is in these parts of the runs when exchanging information could be most valuable."
- "By comparison, some number of independent MCMC chains begun from random starting points, requires the same computing power as one MCMCMC run with the same number of chains, but each chain can be used for inference. This makes the MCMC procedure, other things equal, more efficient. If all of the independent runs converge to the same region of the tree space, this provides evidence that the chains have explored the tree space effectively."



### Within Versus Between Tree Variance

TABLE 3. Estimates of the between- and within-tree components of variance in ancestral state reconstructions. The MS between has 499 df and the MS within has 50,000 - 499 or 49,501 df (see text).

Measure	Mean square between trees	Mean square within trees	F = MSB/MSW	<i>P</i> -value
q <sub>GD</sub>	34.2	11.9	2.87	< 0.001
q <sub>DG</sub>	65.7	13.7	4.81	< 0.001
Log-likelihood	32.85	0.45	73.35	< 0.0001

### BayesTraits

- Investigating trait evolution in a plant radiation
- Estimating forward and reverse rate parameters
- Ancestral character state reconstruction

### **Stochastic Character Mapping**

Syst. Biol. 52(2):131-158, 2003 DOI: 10.1080/10635150390192780

#### Stochastic Mapping of Morphological Characters

JOHN P. HUELSENBECK,<sup>1</sup> RASMUS NIELSEN,<sup>2</sup> AND JONATHAN P. BOLLBACK<sup>1</sup>

<sup>1</sup>Section of Ecology, Behavior and Evolution, Division of Biology, University of California–San Diego, La Jolla, California 92093-0116, USA <sup>2</sup>Department of Biometrics, Cornell University, 439 Warren Hall, Ithaca, New York 14853-7801, USA

Abstract.— Many questions in evolutionary biology are best addressed by comparing traits in different species. Often such studies involve mapping characters on phylogenetic trees. Mapping characters on trees allows the nature, number, and timing of the transformations to be identified. The parsimony method is the only method available for mapping morphological characters on phylogenies. Although the parsimony method often makes reasonable reconstructions of the history of a character, it has a number of limitations. These limitations include the inability to consider more than a single change along a branch on a tree and the uncoupling of evolutionary time from amount of character change. We extended a method described by Nielsen (2002, Syst. Biol. 51:729–739) to the mapping of morphological characters under continuous-time Markov models and demonstrate here the utility of the method for mapping characters on trees and for identifying character correlation. [Bayesian estimation; character correlation; character mapping; Markov chain Monte Carlo.]

#### **BMC Bioinformatics**

Bio Med Central

**Open Access** 

Software

Syst. Biol. 51(5):729-739, 2002 DOI: 10.1080/10635150290102393

#### Mapping Mutations on Phylogenies

#### RASMUS NIELSEN

Department of Biological Statistics, Cornell University, 439 Warren Hall, Ithaca, New York 14853-7801, USA; E-mail: rn28@cornell.edu

Abstract.—Mapping of mutations on a phylogeny has been a commonly used analytical tool in phylogenetics and molecular evolution. However, the common approaches for mapping mutations based on parsimony have lacked a solid statistical foundation. Here, I present a Bayesian method for mapping mutations on a phylogeny. I illustrate some of the common problems associated with using parsimony and suggest instead that inferences in molecular evolution can be made on the basis of the posterior distribution of the mappings of mutations. A method for simulating a mapping from the posterior distribution of mappings is also presented, and the utility of the method is illustrated on two previously published data sets. Applications include a method for testing for variation in the substitution rate along the sequence and a method for testing whether the  $d_N/d_s$  ratio varies among lineages in the phylogeny. [Data augmentation; DNA sequence evolution; Markov chain Monte Carlo; phylogenetics.] SIMMAP: Stochastic character mapping of discrete traits on phylogenies Jonathan P Bollback\*

Address: Bioinformatics Center, University of Copenhagen, Universitetsparken 15, Building 10, 2100 Copenhagen Ø, Denmark Email: Jonathan P Bollback\* - bollback@binf.ku.dk \* Corresponding author

### **Stochastic Character Mapping**

- Advantages over parsimony
  - 1. (most importantly) only one change per branch
  - 2. underestimates variance by only considering the most parsimonious reconstruction
  - 3. no framework for accommodating uncertainty in the phylogenetic reconstruction
  - Bollback 2006

### Evolution of venom proteins in reptiles



Casewell et al. 2012. Nature Communications

### **Evolution of Pharyngognathy**





### Painting branches



**Figure 2.** Phylogenetic relationships, habitat reconstruction, and distribution of body size for *Varanus* species. The phylogeny depicted is the maximum clade credibility tree resulting from Bayesian inference on mtDNA sequences for 37 species. Nodes are supported by at least 0.99 posterior probabilities unless otherwise noted. The history of habitat mapped onto this tree is a single stochastic character map given the observed states for species and this phylogeny. Colors on branches indicate inferred habitat state and colored boxes next to species names represent species habitat states; orange is terrestrial, green is arboreal, and dark gray is rock-dwelling. Species values for adult SVL are based on data for measurements on preserved adult specimens (see Materials and Methods for details).

doi:10.1111/j.1558-5646.2011.01335.x

### **EVOLUTION OF EXTREME BODY SIZE DISPARITY IN MONITOR LIZARDS (VARANUS)**

David C. Collar<sup>1,2,3</sup>, James A. Schulte II<sup>4,5</sup>, and Jonathan B. Losos<sup>1,6</sup>



Figure 4. Boxplots for species' size-corrected morphological traits in each of the three habitat categories. Whiskers are standard, extending to the 9th and 91st percentile of the distributions. Labeled points are species values falling outside of this range. Colors correspond to habitat states: orange is terrestrial, green is arboreal, and gray is rock-dwelling.

### Stochastic mapping

- The basic algorithm
  - 1. Calculate conditional likelihood for each state at each node
  - 2. Simulate ancestral states at internal nodes by sampling from posterior
  - Simulate substitution history by sampling from posterior conditional on reconstructions from step 2 and observed states at the tips of the tree.
    Waiting times from exponential distribution with rate from diagonal of Q matrix conditioned on current state

## Stochastic mapping in R

- make.simmap function
- Primitive functionality relative to SIMMAP, but easier to learn with
- Line by line through make.simmap