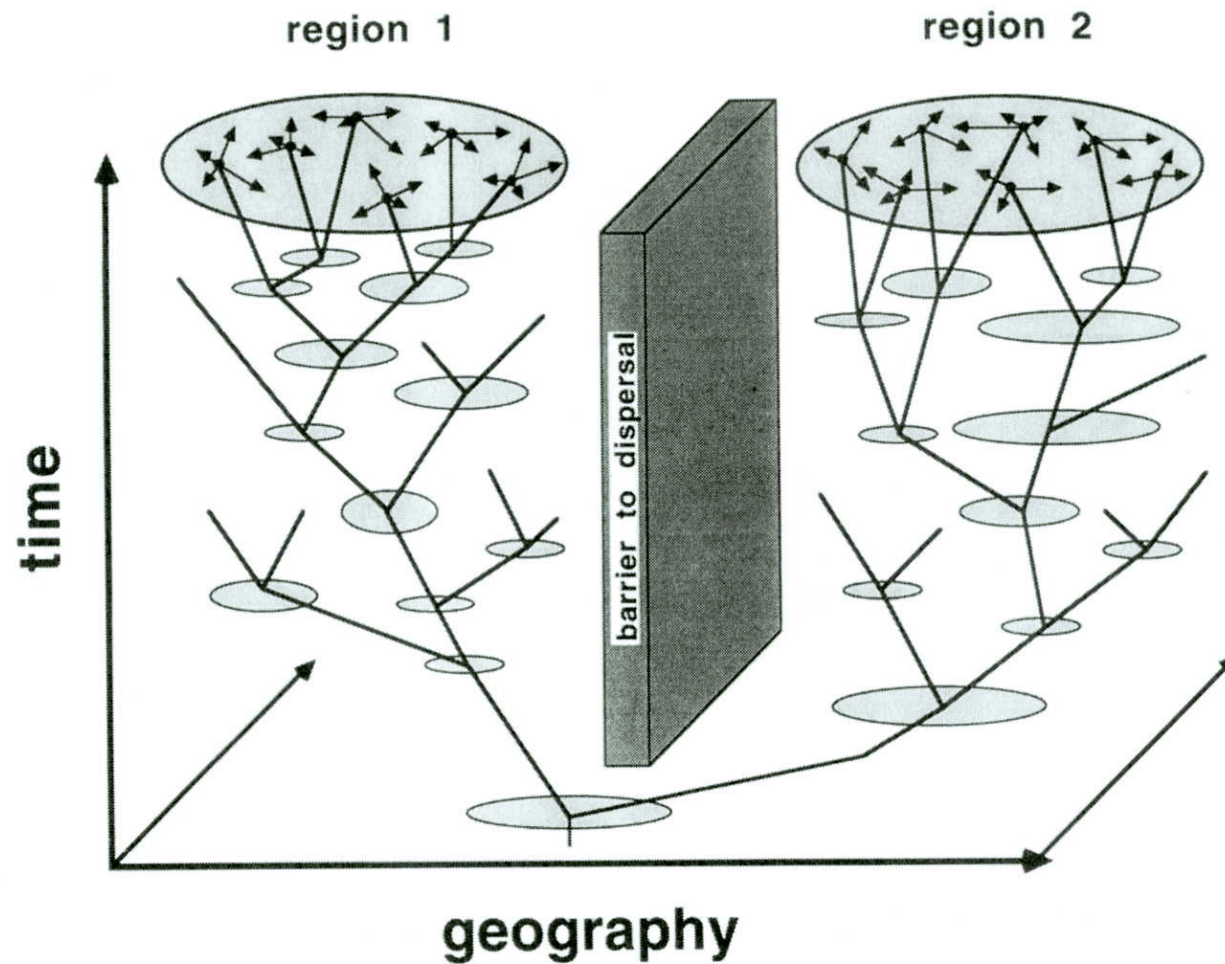


FILOGEOGRAFIA

O QUE É A FILOGEOGRAFIA

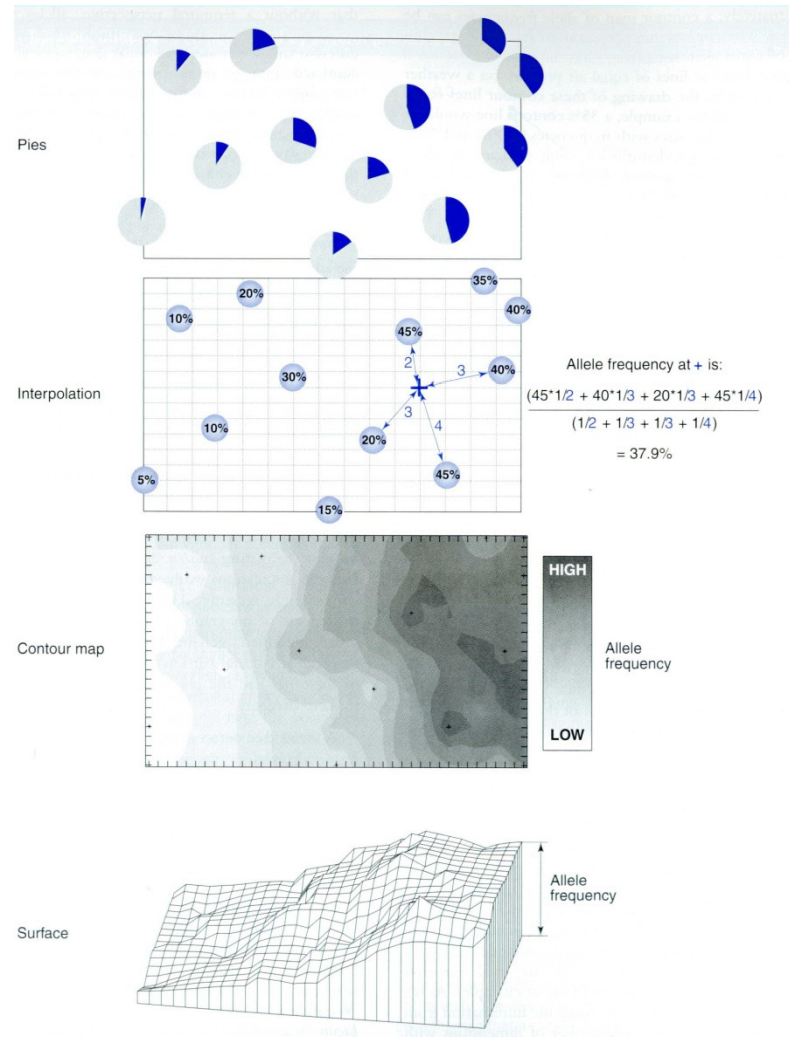


ANTES DA FILOGEOGRAFIA

Mesmo antes da palavra ou disciplina “FILOGEOGRAFIA”, os investigadores já estudavam a relação entre a variação genética e a geografia

Alguns métodos já bastante “antigos” (mas que continuam a usar-se) para dados alélicos incluem:

1. A geração de mapas de frequências alélicas por interpolação e os mapas sintéticos de Cavalli-Sforza

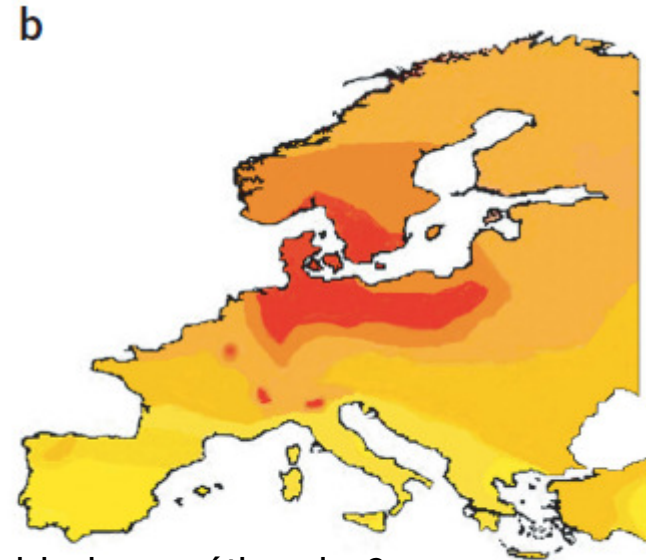


Gene-culture coevolution between cattle milk protein genes and human lactase genes

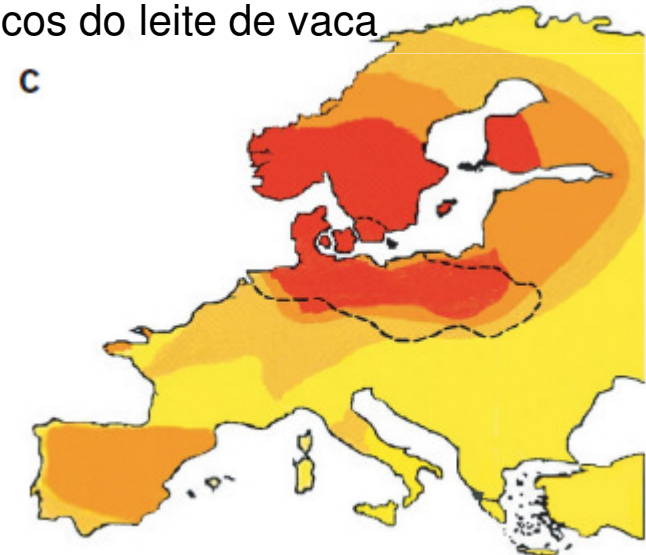
Albano Beja-Pereira^{1,2}, Gordon Luikart¹, Phillip R England¹, Daniel G Bradley³, Oliver C Jann⁴, Giorgio Bertorelle⁵, Andrew T Chamberlain⁶, Telmo P Nunes⁷, Stoitcho Metodiev⁸, Nuno Ferrand^{2,9} & Georg Erhardt⁴

Milk from domestic cows has been a valuable food source for over 8,000 years, especially in lactose-tolerant human societies that exploit dairy breeds. We studied geographic patterns of variation in genes encoding the six most important milk proteins in 70 native European cattle breeds. We found substantial geographic coincidence between high diversity in cattle milk genes, locations of the European Neolithic cattle farming sites (>5,000 years ago) and present-day lactose tolerance in Europeans. This suggests a gene-culture coevolution between cattle and humans.

Coincidência geográfica entre a diversidade genética encontrada em 6 proteínas do leite de vaca com os principais centros de “produção de leite” conhecidos durante o Neolítico e a actual tolerância à lactose em Europeus

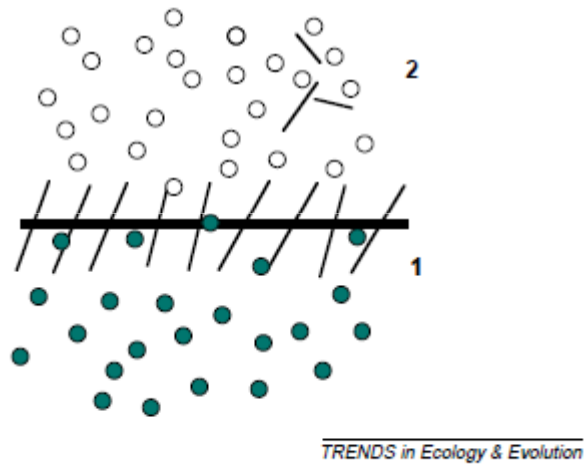


Diversidade genética de 6 genes proteicos do leite de vaca



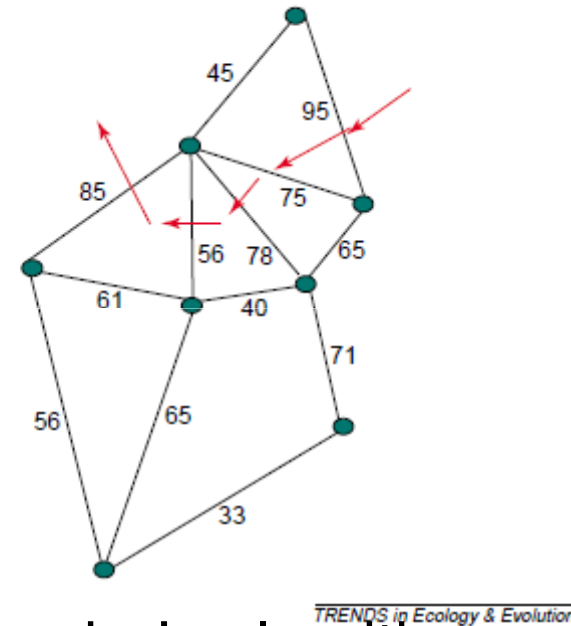
Distribuição geográfica actual do alelo que confere tolerância à lactose em humanos

2. Detecção de descontinuidades genéticas: Wombling e Monmonier's algoritmo



wombling

Detecta alterações no declive das frequências alélicas interpoladas

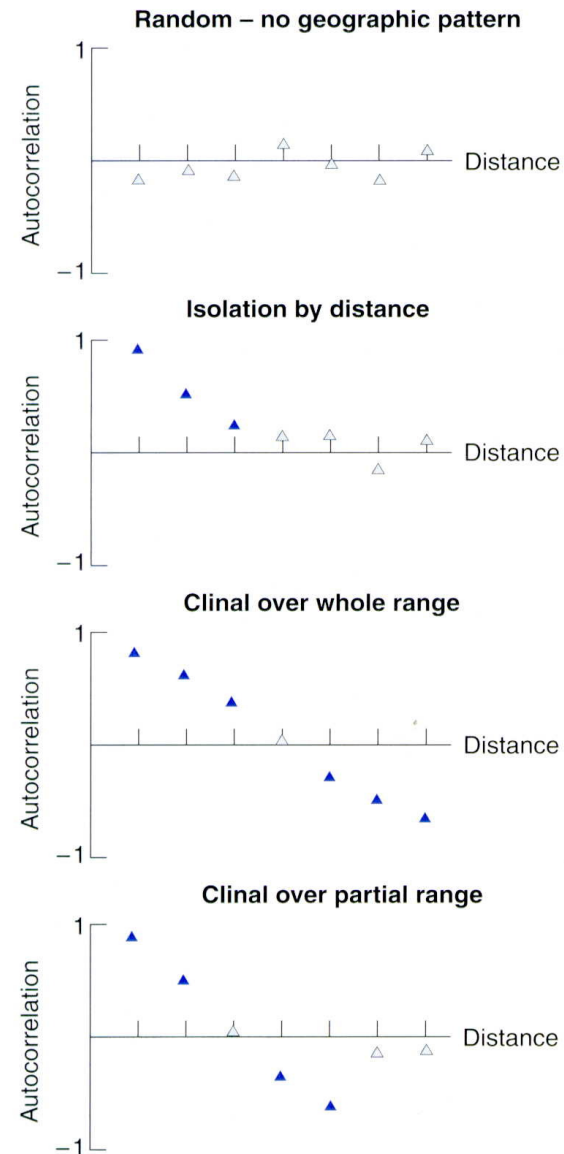


Monmonier's algorithm

Detecta fronteiras entre os vértices de um gráfico de valores. Isto é feito encontrando o percurso que exhibe a distância mais longa entre os vértices. O algoritmo procura a maior distância entre dois vértices vizinhos. Uma fronteira poderá resultar de um decréscimo abrupto desses valores.

3. Autocorrelação espacial

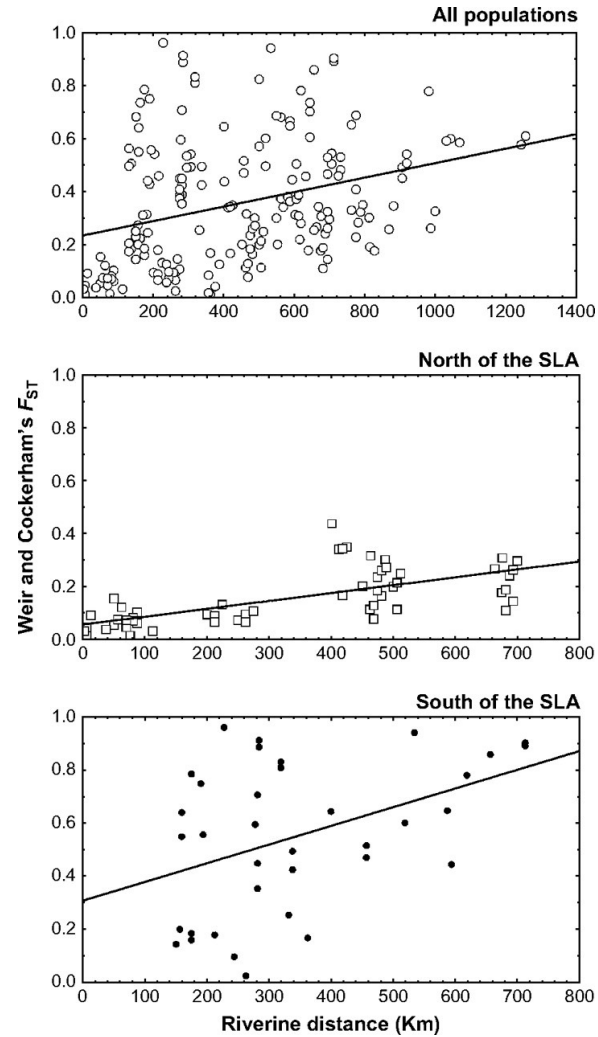
Correlação das frequências alélicas entre pares de localidades e sua variação com a distância geográfica



4. Teste de Mantel

Investiga a significância estatística entre duas matrizes: a matriz de distâncias genéticas e a matriz das distâncias geográficas

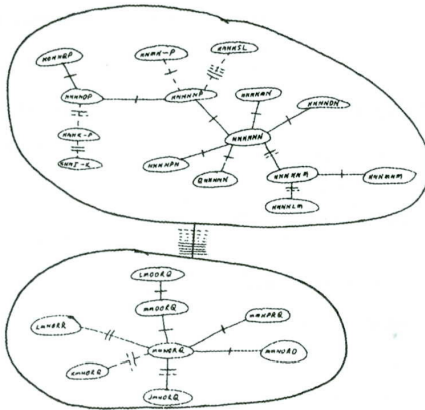
Scatterplots of protein multilocus F_{ST} estimates versus riverine distance for pairwise combination of populations.



Antunes A et al. *J Hered* 2006;97:193-205

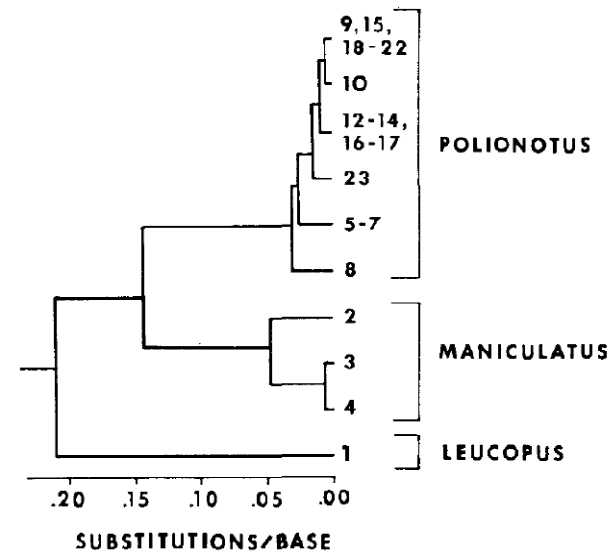
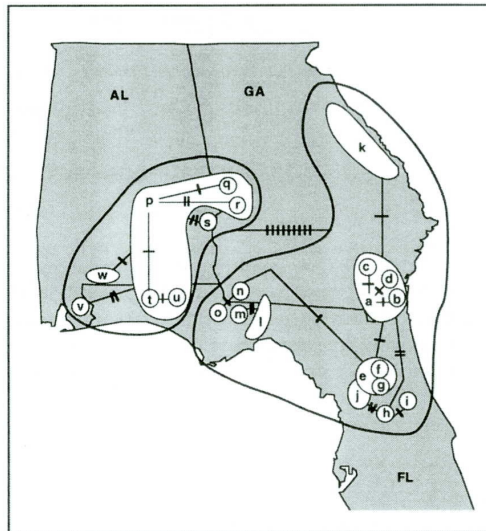
O NASCIMENTO DA FILOGEOGRAFIA

Estimated genetic distance (mean ± standard error) in base substitutions per nucleotide between Peromyscus populations at various stages of evolutionary divergence



Comparison	Mean distance between collections	Mean ± S.E. base substitutions per nucleotide	
		4-base enzyme	6-base enzymes
Within geographic locality			
<i>P. leucopus</i>	500 feet	—	0.000
<i>P. maniculatus</i>	1/4 mile	—	<0.010
<i>P. polionotus</i>	1 mile	0.004 ± 0.002	0.002 ± 0.001
Between geographic localities			
<i>P. maniculatus</i>	1000 miles	0.042	0.031 ± 0.013
<i>P. polionotus</i>	80-500 miles	0.006 ± 0.001	0.015 ± 0.001
Sibling species			
<i>P. maniculatus</i> vs. <i>P. polionotus</i>		0.166 ± 0.006	0.132 ± 0.006
Nonsibling species			
<i>P. leucopus</i> vs. <i>P. maniculatus</i> and <i>P. polionotus</i>		—	0.214 ± 0.018

**John Avise
(1979)
revisited**



FILOGEOGRAFIA: O BAPTISMO

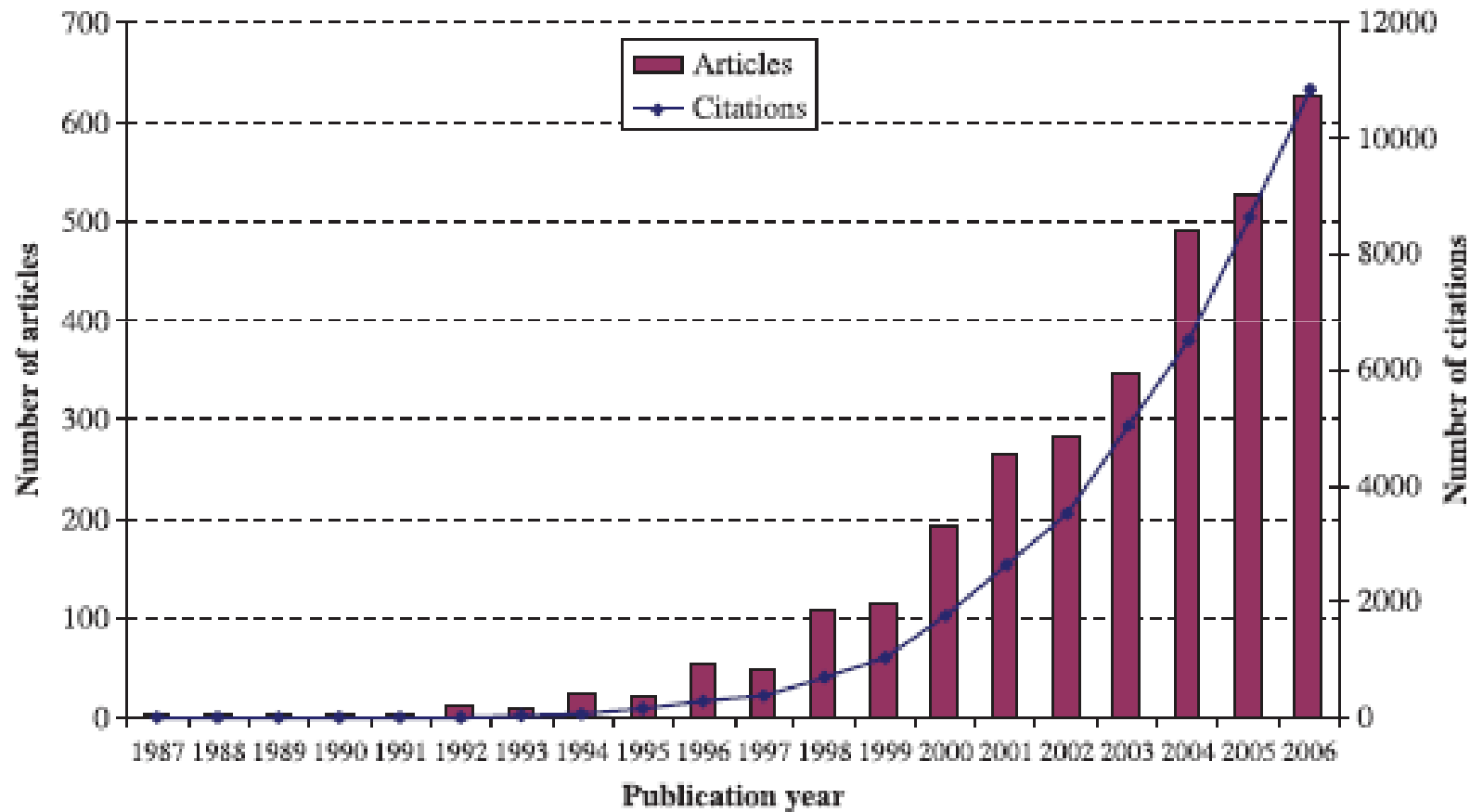
Avise JC *et al* (1987). Intraspecific **phylogeography**: the mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology and Systematics* 18: 489-522.

i) A maior parte das espécies é constituída por populações geograficamente distintas cujos membros ocupam diferentes ramos numa árvore filogenética intraespecífica

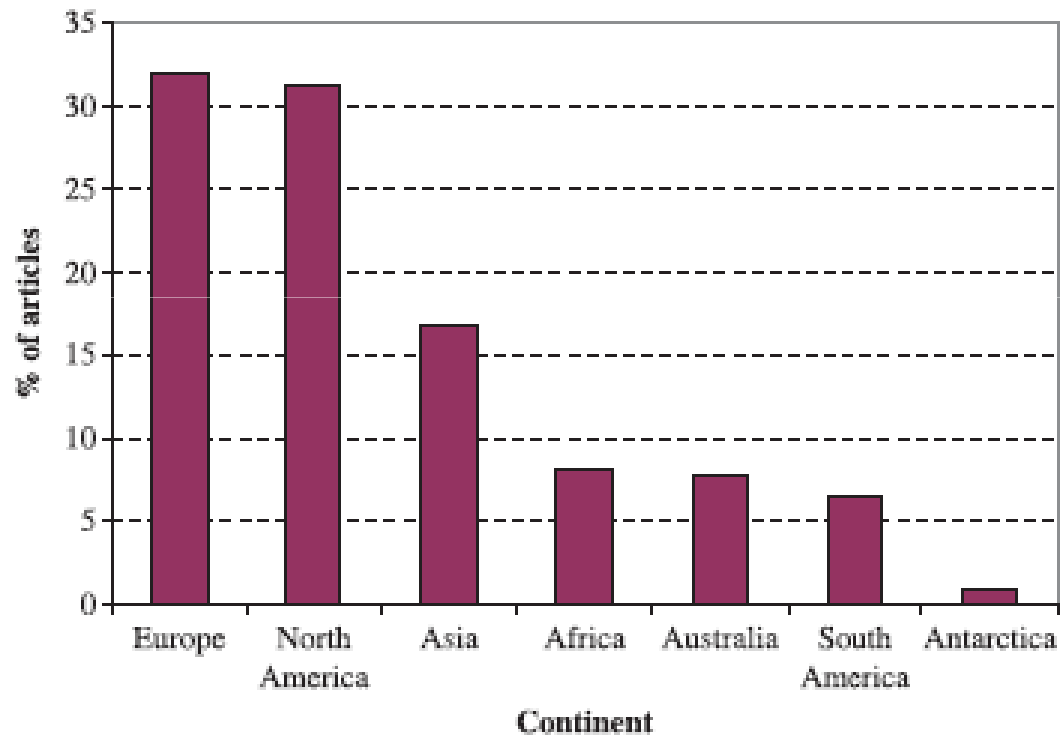
ii) Espécies com elevada capacidade de dispersão não têm estruturas filogeográficas bem marcadas, como, por exemplo, insectos voadores, aves, peixes marinhos, ou ainda aquelas, como presumivelmente a espécie humana, que se caracterizam por uma expansão recente a partir de um único refúgio

iii) Os grupos monofiléticos que se caracterizam por grandes descontinuidades filogenéticas são geralmente o resultado da existência de barreiras extrínsecas antigas ao fluxo génico

A EVOLUÇÃO DOS ESTUDOS FILOGEOGRÁFICOS ENTRE 1987 E 2006



A EVOLUÇÃO DOS ESTUDOS FILOGEOGRÁFICOS ENTRE 1987 E 2006



A EVOLUÇÃO DOS ESTUDOS FILOGEOGRÁFICOS ENTRE 1987 E 2006

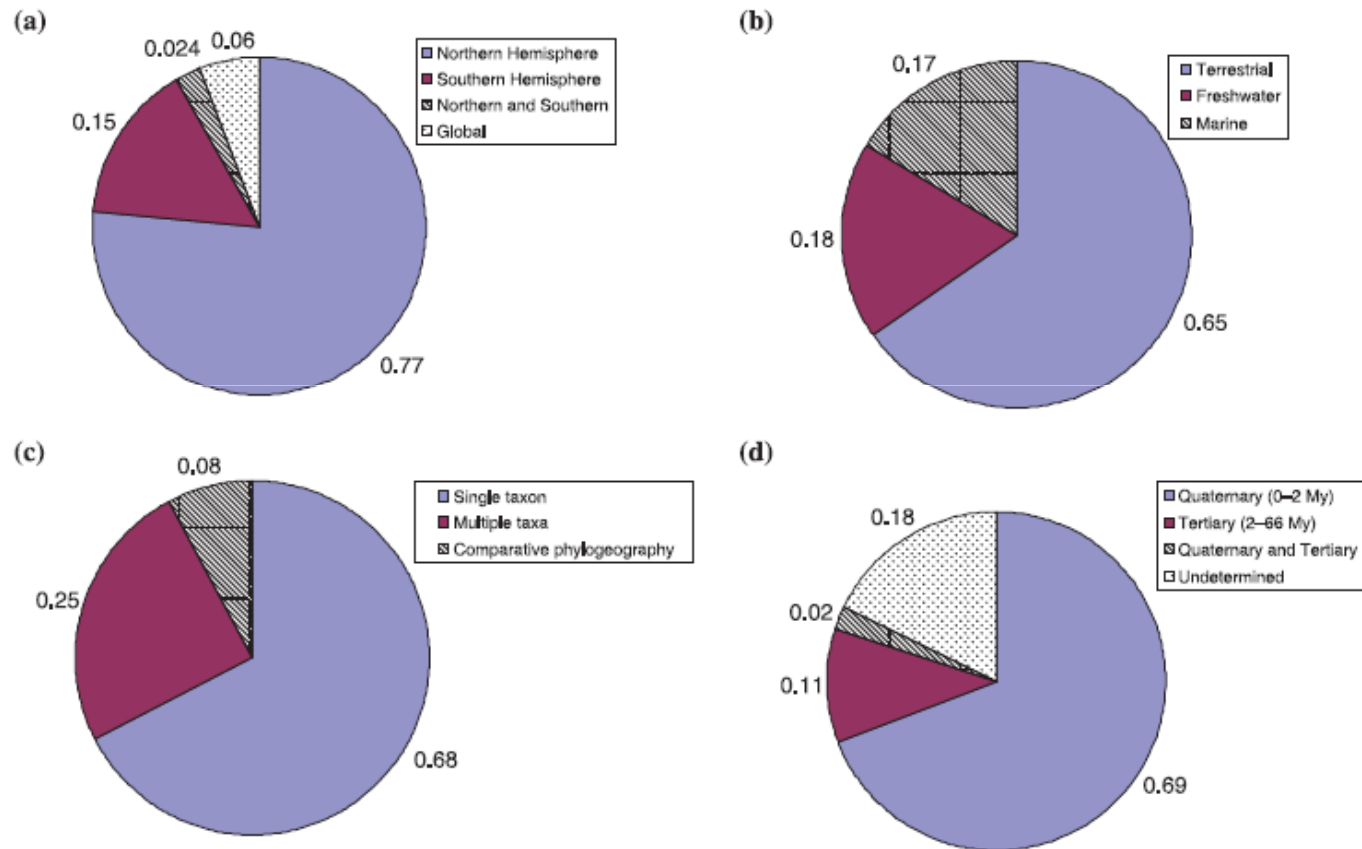
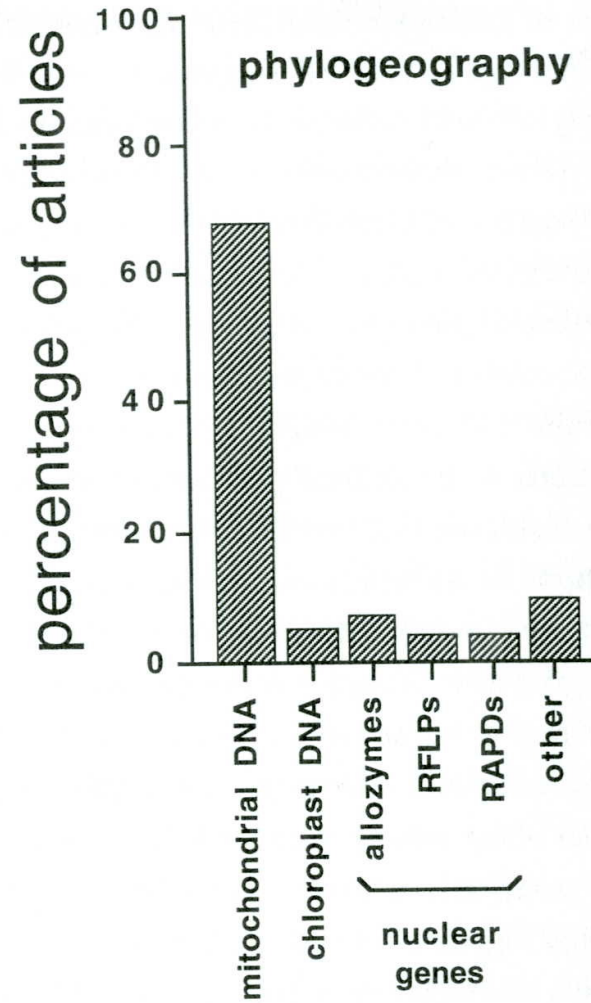


Fig. 2 Proportion of articles in phylogeography published between 1987 and 2006 according to (a) hemisphere or global scale, (b) type of environment, (c) nature of study and (d) timescale (see text for details).

A POPULARIDADE DO mtDNA



FILOGEOGRAFIA

Durante muitos anos, as análises filogeográficas usando dados de sequencias baseavam-se em meras inferências de observação, carecendo de suporte estatístico.

In 1998, Templeton desenvolveu um método - Nested Clade Analyses (NCA)

Molecular Ecology (1998) 7, 381–397

Nested clade analyses of phylogeographic data: testing hypotheses about gene flow and population history

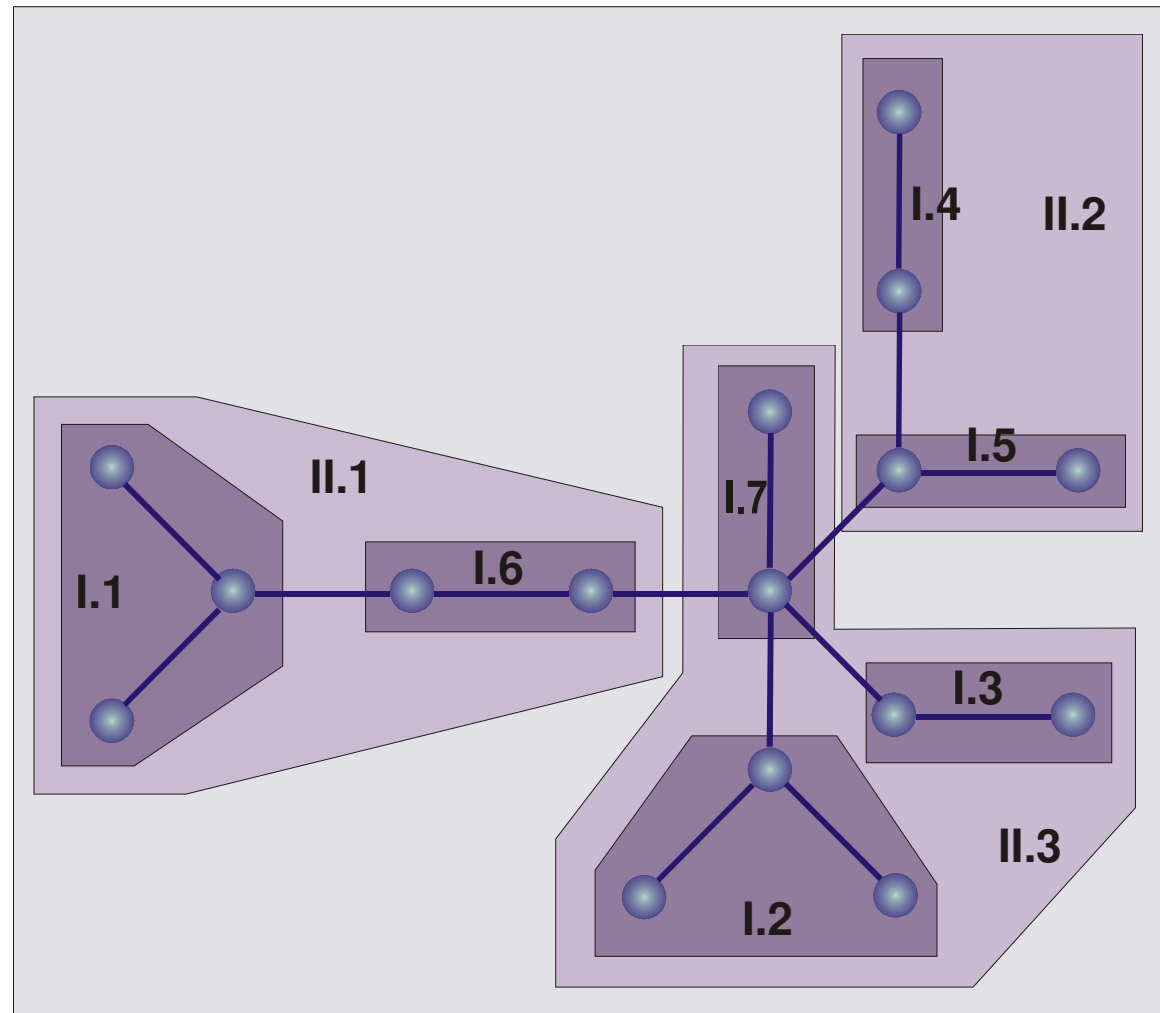
ALAN R. TEMPLETON

Department of Biology, Washington University, St Louis, Missouri 63130-4899, USA

O DESENVOLVIMENTO DA ANÁLISE CLADÍSTICA HIERARQUIZADA (NCA ou NCPA) I

1. A construção de uma rede de haplótipos
2. A existência de uma idade diferente para os haplótipos: os externos (*tip clades*), ou mais recentes, e os internos (*internal clades*), ou mais antigos
3. A construção de clados progressivamente mais antigos
4. Um teste estatístico de associação entre estes clados e a geografia

O DESENVOLVIMENTO DA ANÁLISE CLADÍSTICA HIERARQUIZADA (NCA ou NCPA) II



A ANÁLISE ESTATÍSTICA DA NCA ou NCPA I

D_c – distância entre os membros de um clado e o seu centro geográfico

D_n – distância entre os membros de um clado hierarquizado e o seu centro geográfico

D_c (int) – D_c (tip)

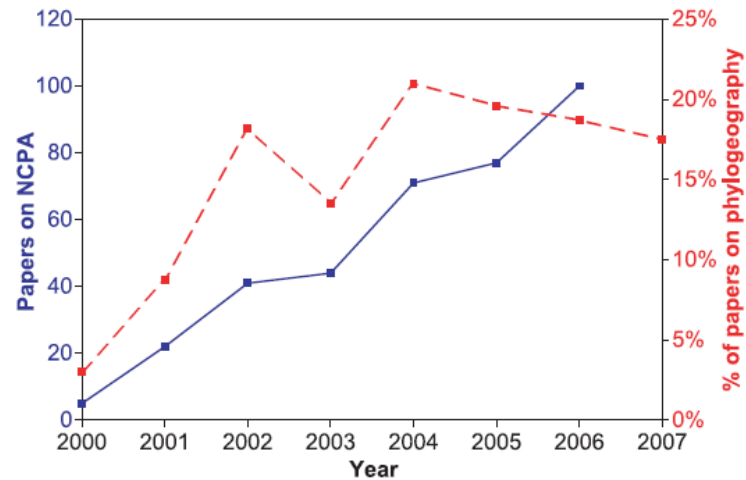
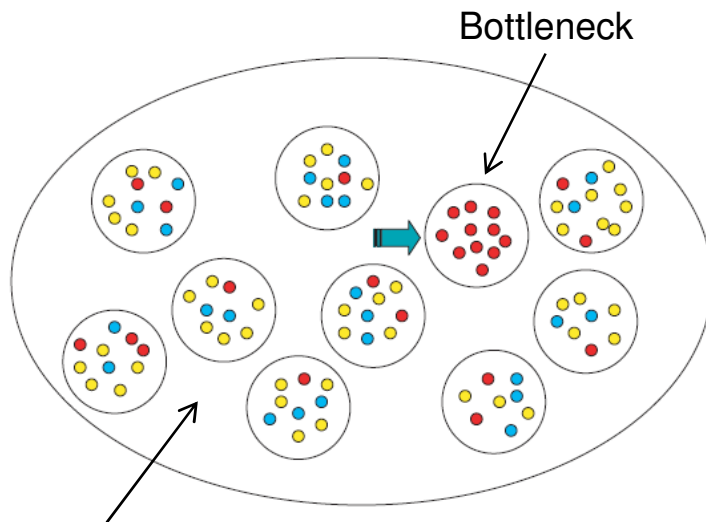
D_n (int) – D_n (tip)

A aplicação de permutações que testam a H_0 de não associação entre a distribuição dos haplótipos e a geografia

A chave interpretativa

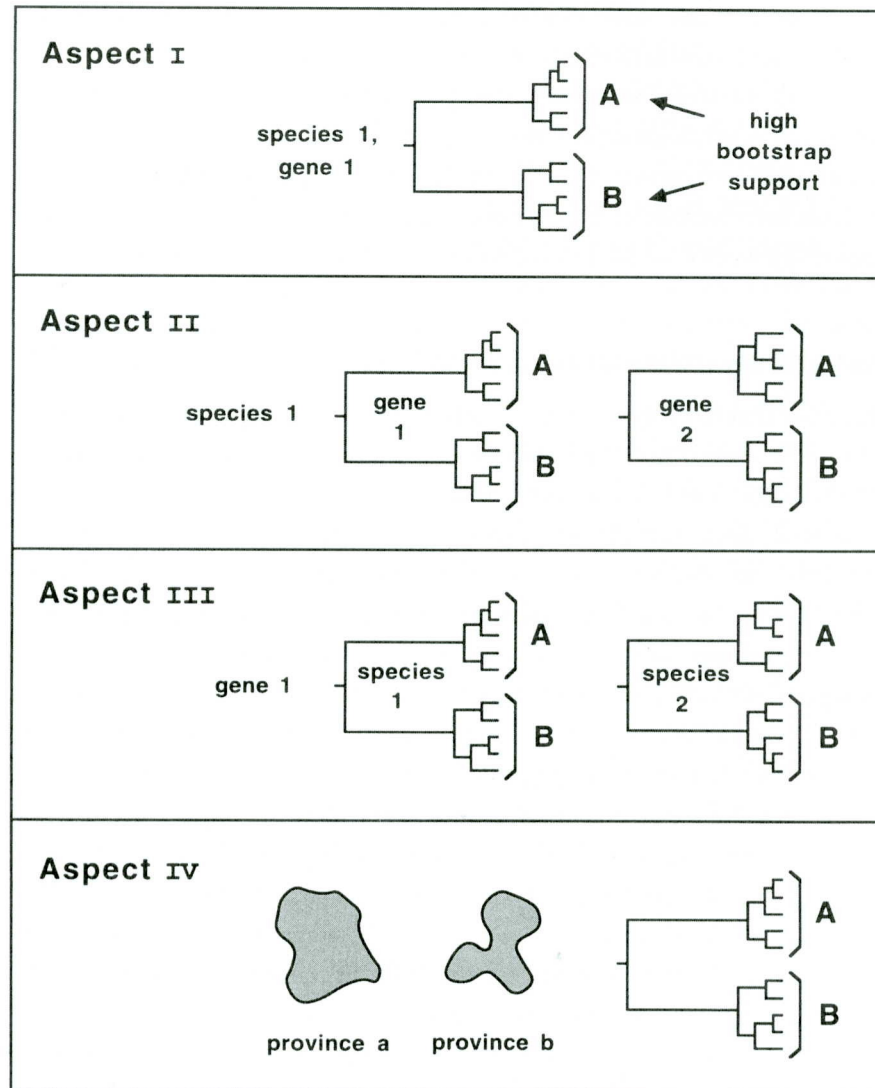
Limitações e críticas à NCA ou NCPA

- ❖ Falta de rigor estatístico; excesso de falsos positivos (fragmentação e expansão)
- ❖ Baixo poder para detectar expansão populacional em situações extremas
- ❖ Tendência para detectar isolamento pela distância
- ❖ Baseia-se numa só genealogia, ignorando a estocasticidade das genealogias
- ❖ O teste de permutações são feitas ao nível do indivíduo e não das populações
- ❖ Chave de interpretação



Estrutura espacial significativa, possivelmente com expansão (caso o alelo vermelho seja derivado)

A CONCORDÂNCIA GENEALÓGICA



Múltiplos caracteres genéticos diagnósticos formam clados suportados estatisticamente

Concordância entre múltiplos loci independentes

Factores biogeográficos históricos afectam múltiplas espécies

Concordância entre dados genéticos e evidências biogeográficas tradicionais

A DISCORDÂNCIA GENEALÓGICA

- ❖ Polimorfismo ancestral
- ❖ Recombinação
- ❖ Selecção natural
- ❖ Introgessão

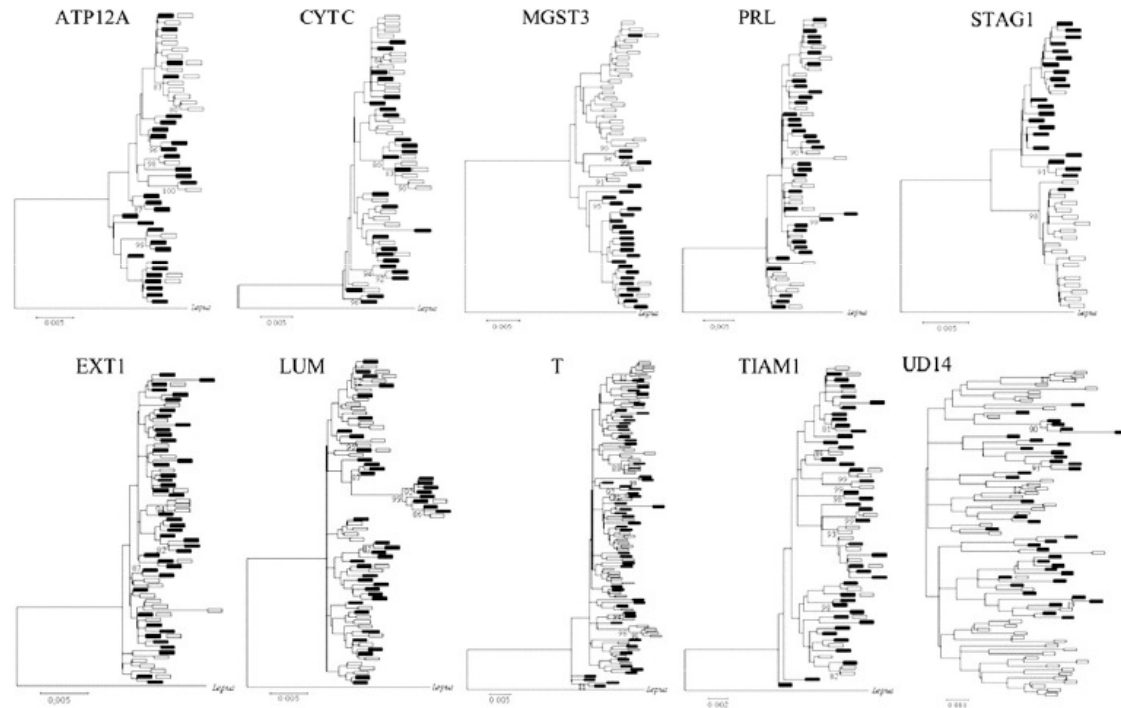


FIGURE 3.—Haplotype variation illustrated as neighbor-joining trees for each of the 10 autosomal loci. Because recombination events were detected for all loci in our data set, the trees represent overall similarity between the haplotypes. Portions of each gene may have different genealogies. Haplotypes found in *O. c. algirus* and *O. c. cuniculus* are followed by a solid and an open box, respectively, and shared haplotypes by both solid and open boxes. *Lepus granatensis* was used as an outgroup except in UD14. Bootstrap values ≥ 80 are shown next to branches.

FILOGEOGRAFIA E ESPECIAÇÃO

Os problemas associados ao estudo de uma única genealogia

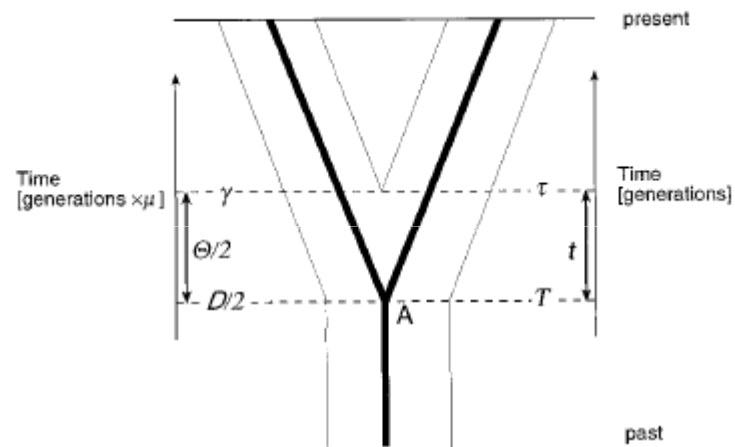
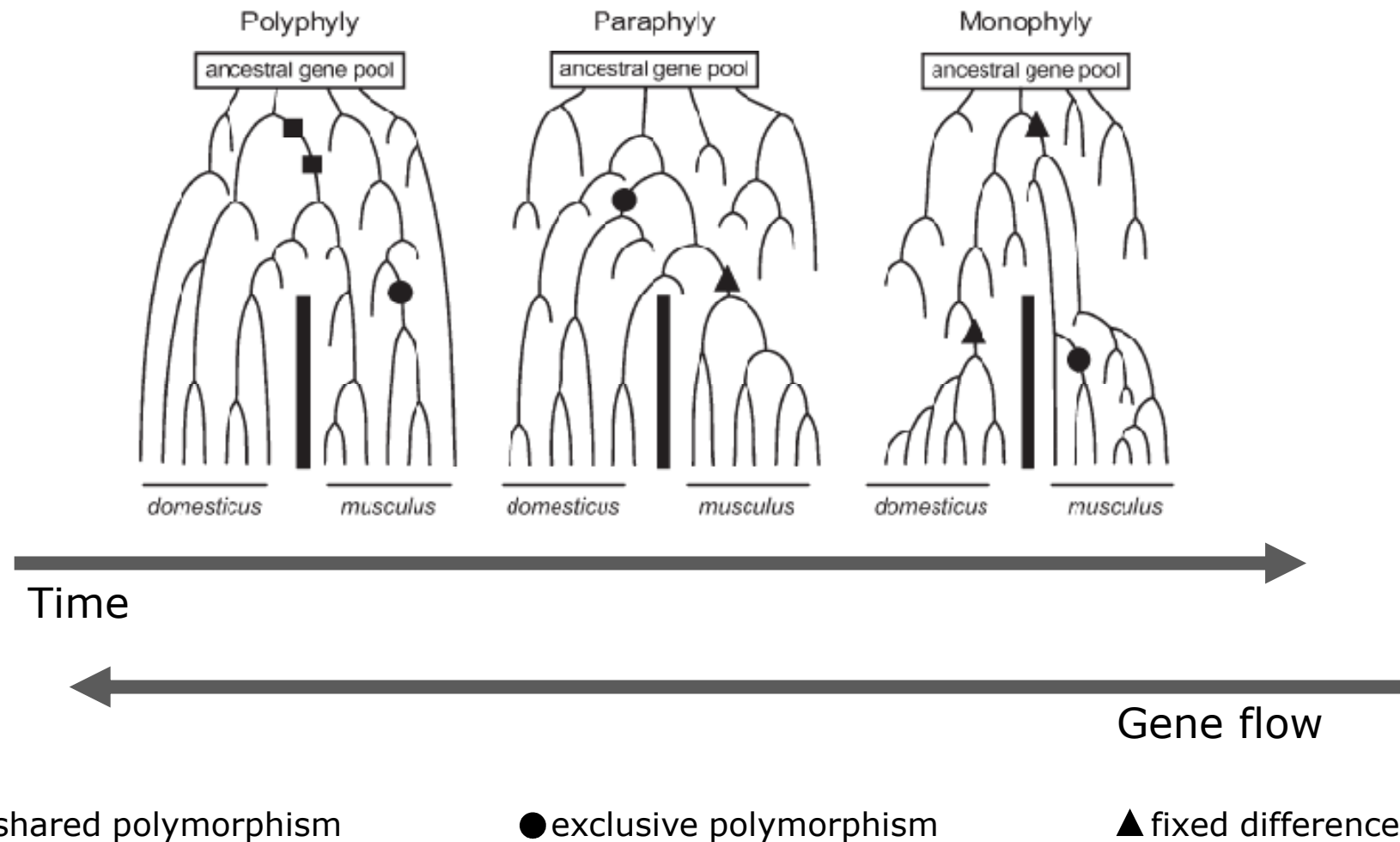
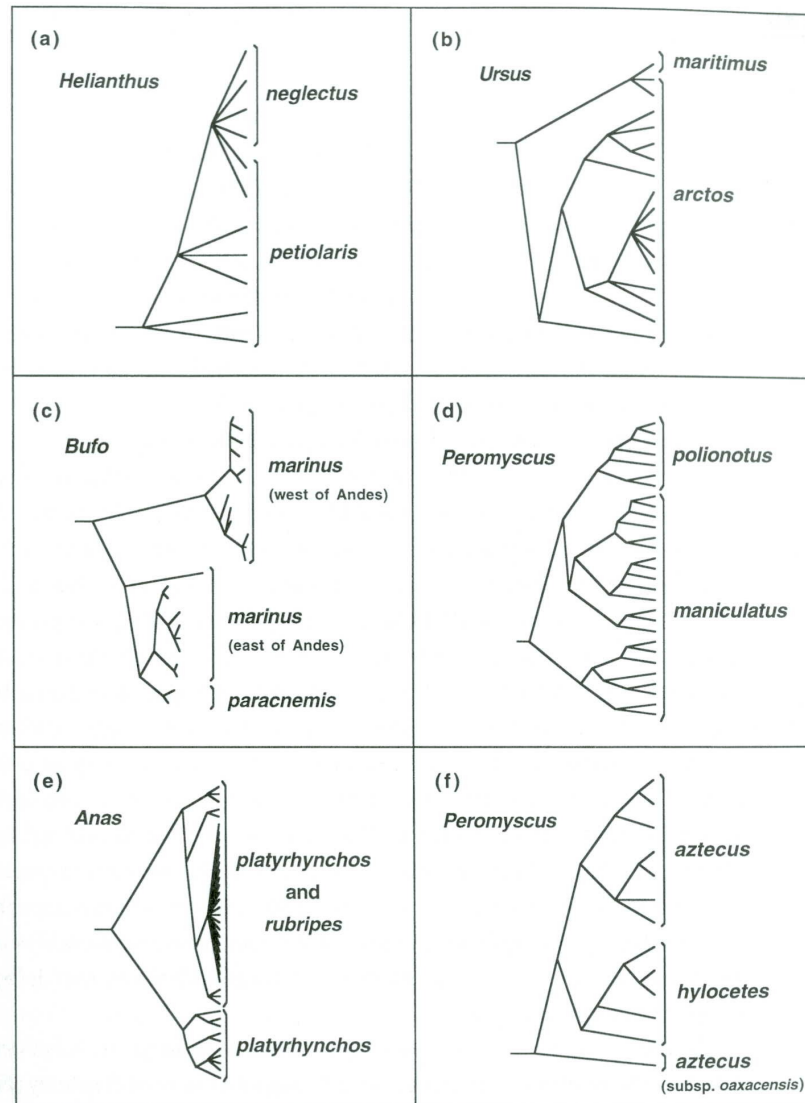


FIG. 1. Schematic diagram showing a single ancestral population splitting into two descendent populations, and associated variables. T is the time in generations since the gene divergence occurred, τ is population divergence, γ is the population divergence τ multiplied by the mutation rate μ , and D is the gene divergence time T scaled by μ . t is the difference between τ and T , and has an expected value of $2N_e$ generations, or $\Theta/2$ mutational events, where $\Theta = 4N_e\mu$. The bold A marks the most recent common ancestor (MRCA) to which alleles in the two descendent species coalesce, not the MRCA of all the alleles in the ancestral species.

FILOGEOGRAFIA E ESPECIAÇÃO I



FILOGEOGRAFIA E ESPECIAÇÃO II



A FILOGEOGRAFIA ESTATÍSTICA

Molecular Ecology (2002) 11, 2623–2635

Statistical phylogeography

L. LACEY KNOWLES and WAYNE P. MADDISON

Molecular Ecology (2009) 18, 1034–1047

INVITED REVIEW

Statistical inferences in phylogeography

RASMUS NIELSEN*† and MARK A. BEAUMONT‡

Statistical Phylogeography

L. Lacey Knowles

Department of Ecology and Evolutionary Biology, Museum of Zoology, University of Michigan,
Ann Arbor, Michigan 48109-1079; email: knowlesl@umich.edu

THE STUDY OF STRUCTURED
POPULATIONS — NEW HOPE FOR A
DIFFICULT AND DIVIDED SCIENCE

Jody Hey and Carlos A. Machado†*

Molecular Ecology (2010) 19, 436–446

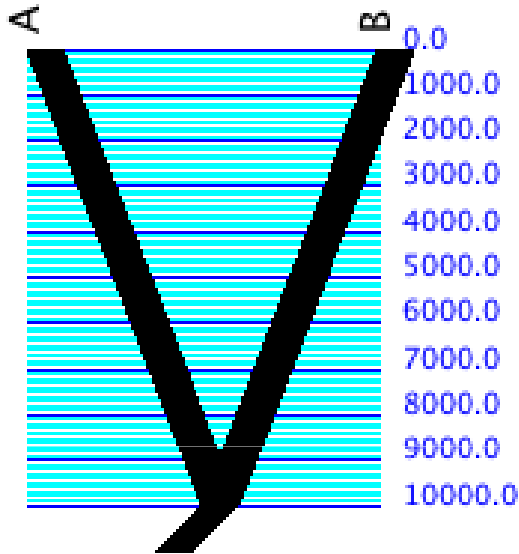
NEWS AND VIEWS

REPLY

In defence of model-based inference in phylogeography

MARK A. BEAUMONT,¹ RASMUS NIELSEN,² CHRISTIAN ROBERT,³ JODY HEY,⁴ OSCAR GAGGIOTTI,⁵ LACEY KNOWLES,⁶ ARNAUD ESTOUP,⁷ MAHESH PANCHAL,⁸ JUKKA CORANDER,⁹ MIKE HICKERSON,¹⁰ SCOTT A. SISSON,¹¹ NELSON FAGUNDES,¹² LOUNÈS CHIKHI,¹³ PETER BEERLI,¹⁴ RENAUD VITALIS,¹⁵ JEAN-MARIE CORNUET,⁷ JOHN HUELSENBECK,² MATTHIEU FOLL,^{16,17} ZIHENG YANG,¹⁸ FRANCOIS ROUSSET,¹⁹ DAVID BALDING²⁰ and LAURENT EXCOFFIER^{16,17}

A FILOGEOGRAFIA ESTATÍSTICA



Testar hipóteses comparando os nossos dados com cenários hipotéticos – Qual o cenário que melhor explica os nossos dados?

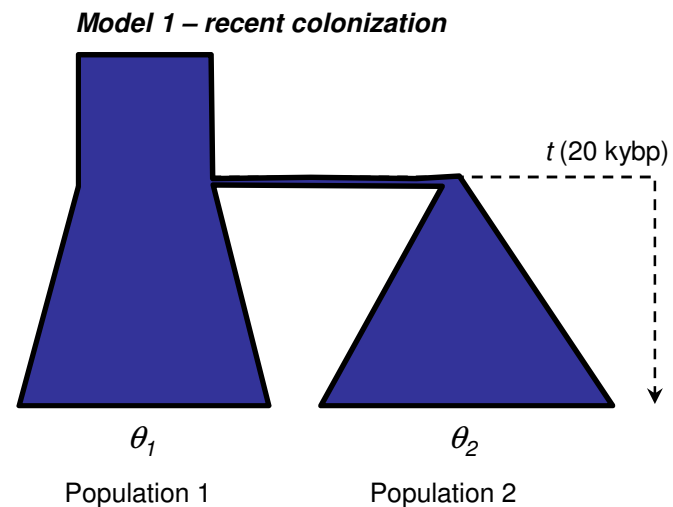
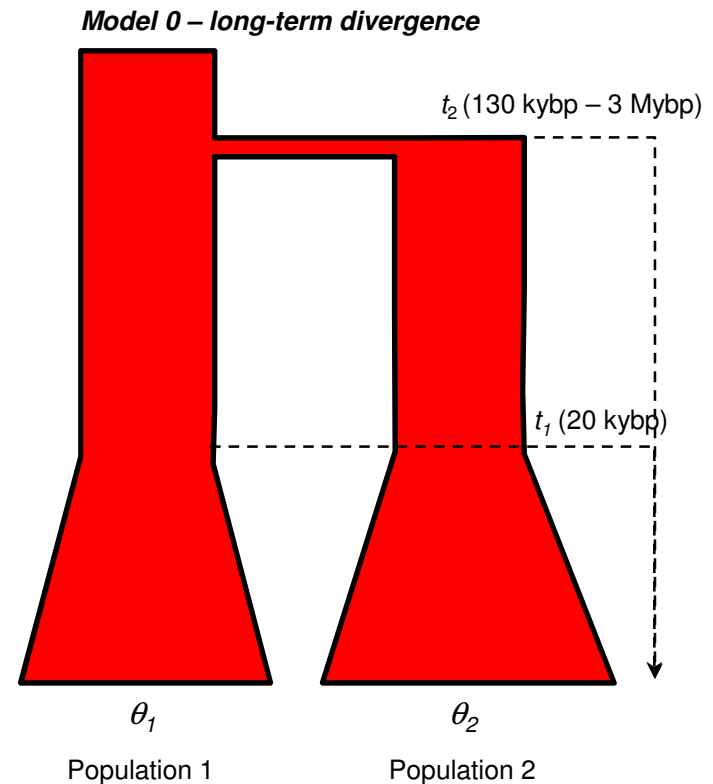
- ❖ Vamos simular várias genealogias usando o modelo coalescente de acordo com o cenário hipotético histórico que queremos testar (assumindo um N_e ou divergência ou número de sítios variáveis ou outros parâmetros, que é estimado a partir dos dados).
- ❖ Vamos comparar os valores obtidos através de simulações coalescentes (medidos através de uma estatística) com a mesma estatística obtida para os nossos dados.

ms (program for generating samples under neutral models)

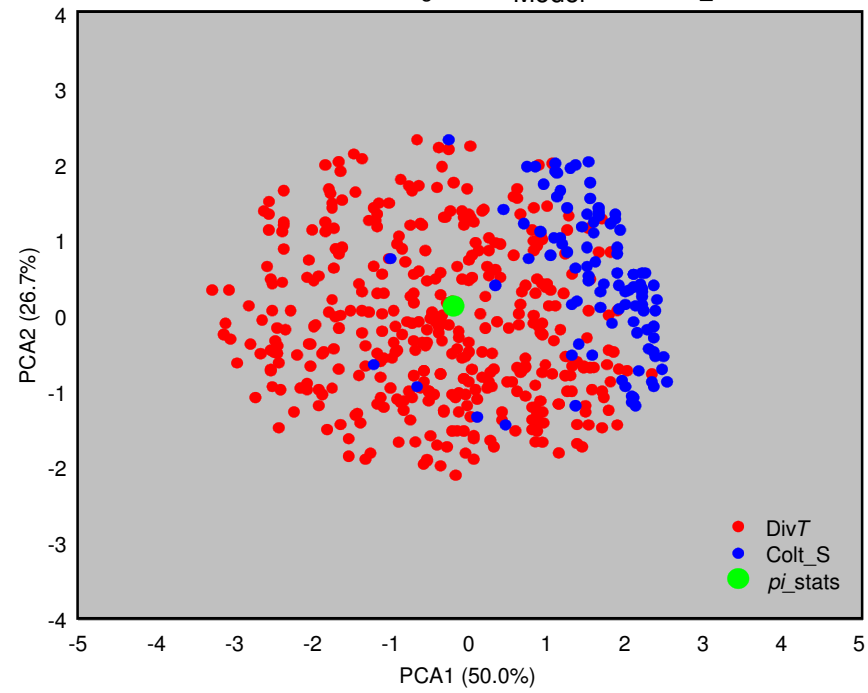
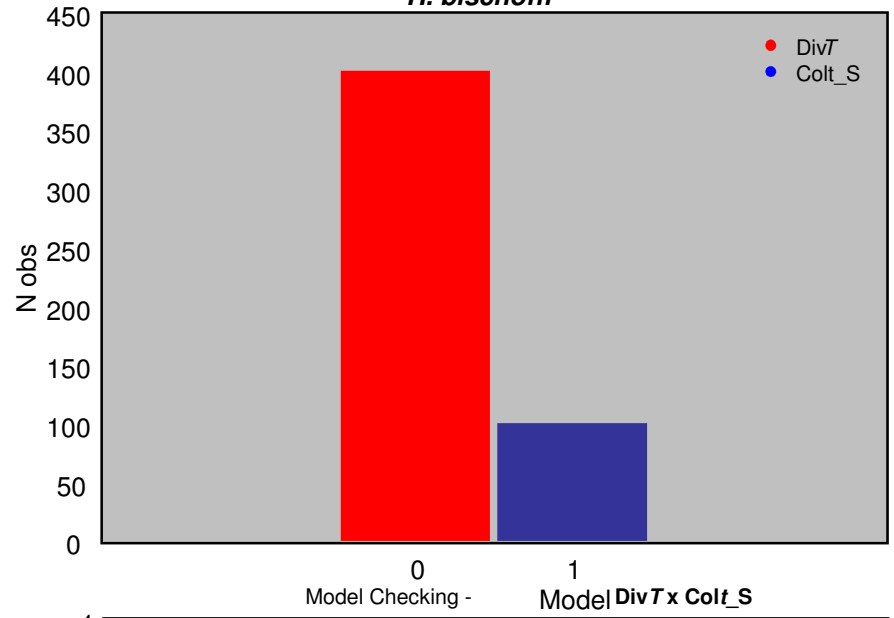
Hudson, R. R. (2002) *Bioinformatics*.

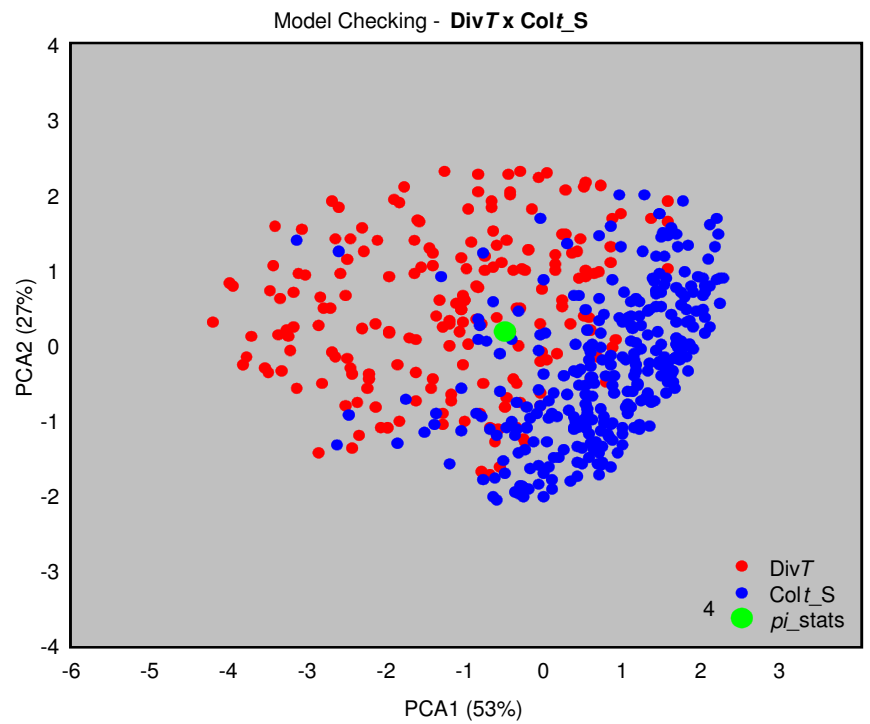
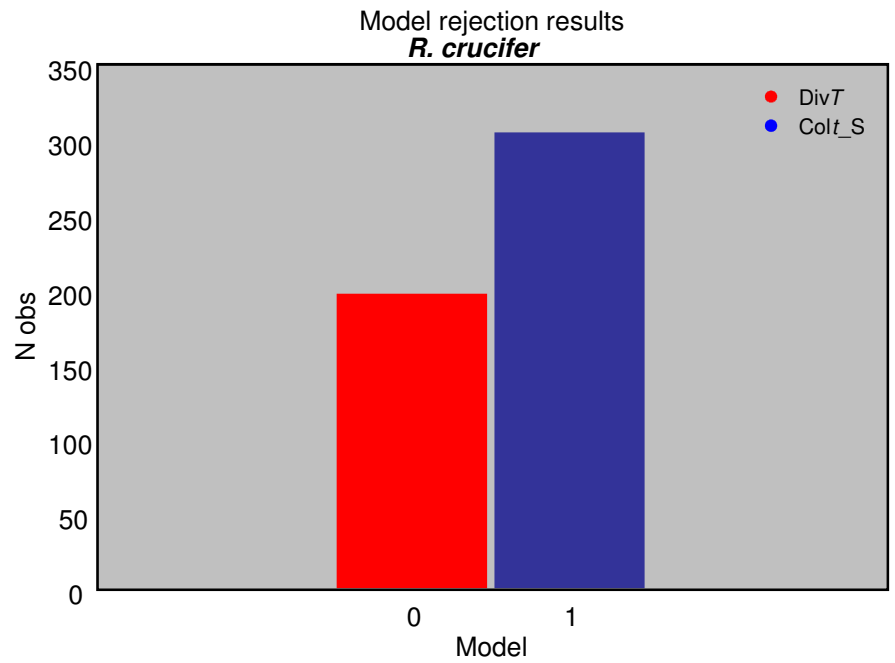
❖ Modelos para testar hipóteses filogeográficas alternativas, que incorporam a divergência temporal e outros parâmetros adicionais, como p. ex o Theta (estimado a partir dos nossos dados)

❖ Depois de definido os modelos, vamos comparar os valores obtidos (para determinados parâmetros ou estatísticas) através de simulações coalescentes de acordo com o cenário hipotético que queremos testar com os valores para os mesmos parâmetros ou estatística obtidos para os nossos dados.



Model rejection results
H. bischoffi

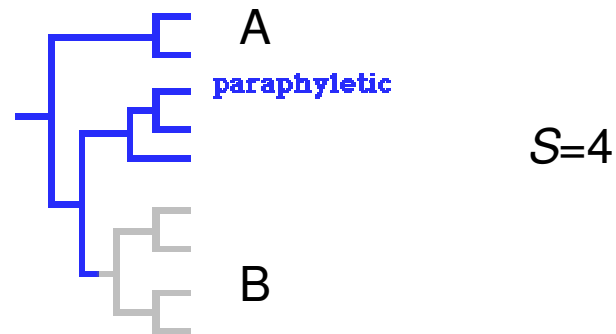




A FILOGEOGRAFIA ESTATÍSTICA

MESQUITE

Maddison & Maddison.
<http://mesquiteproject.org>.

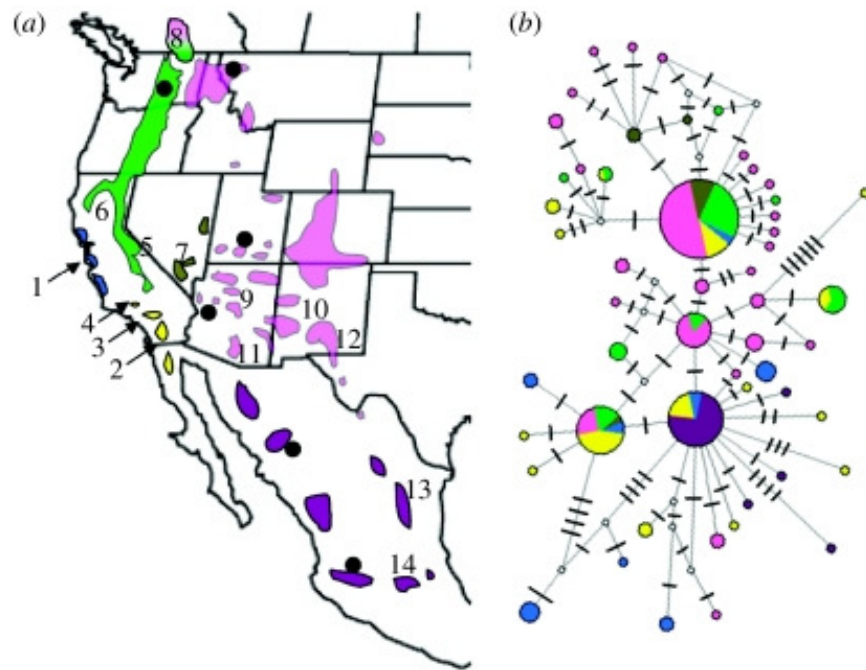


Vamos estimar Slatkin & Maddison's estatística s - discordância entre a árvore do gene e a subdivisão populacional.

P. ex. Se as populações já divergiram há relativamente muito tempo, o único processo responsável pelo padrão observado é a migração. Então S pode ser interpretado como o número mínimo de eventos migratórios entre as populações

Qual a probabilidade de obter $S = 4$, se as populações divergiram à 10 mil anos (assumindo total isolamento desde a divergência), tendo em consideração o efectivo populacional das populações?

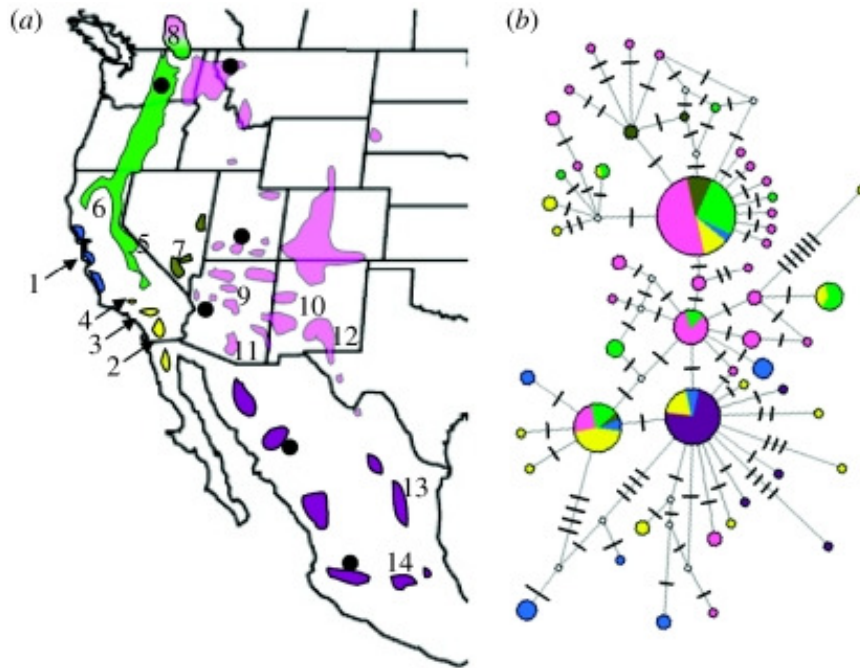
A FILOGEOGRAFIA ESTATÍSTICA



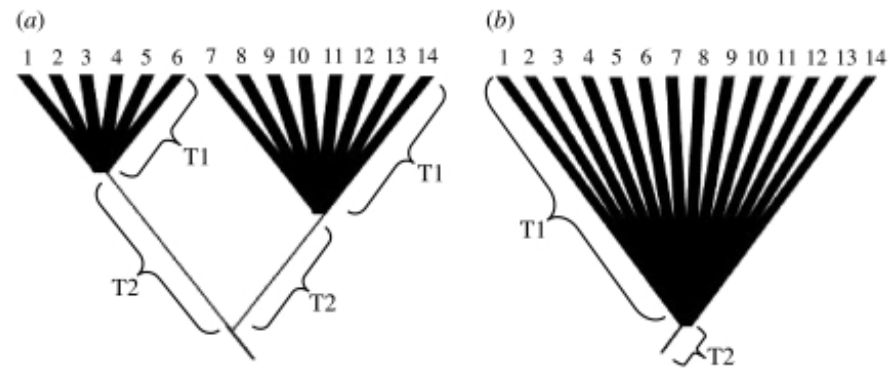
Sitta pygmaea)

Spellman & Klicka 2006. *Proc. R. Soc. B*

A FILOGEOGRAFIA ESTATÍSTICA



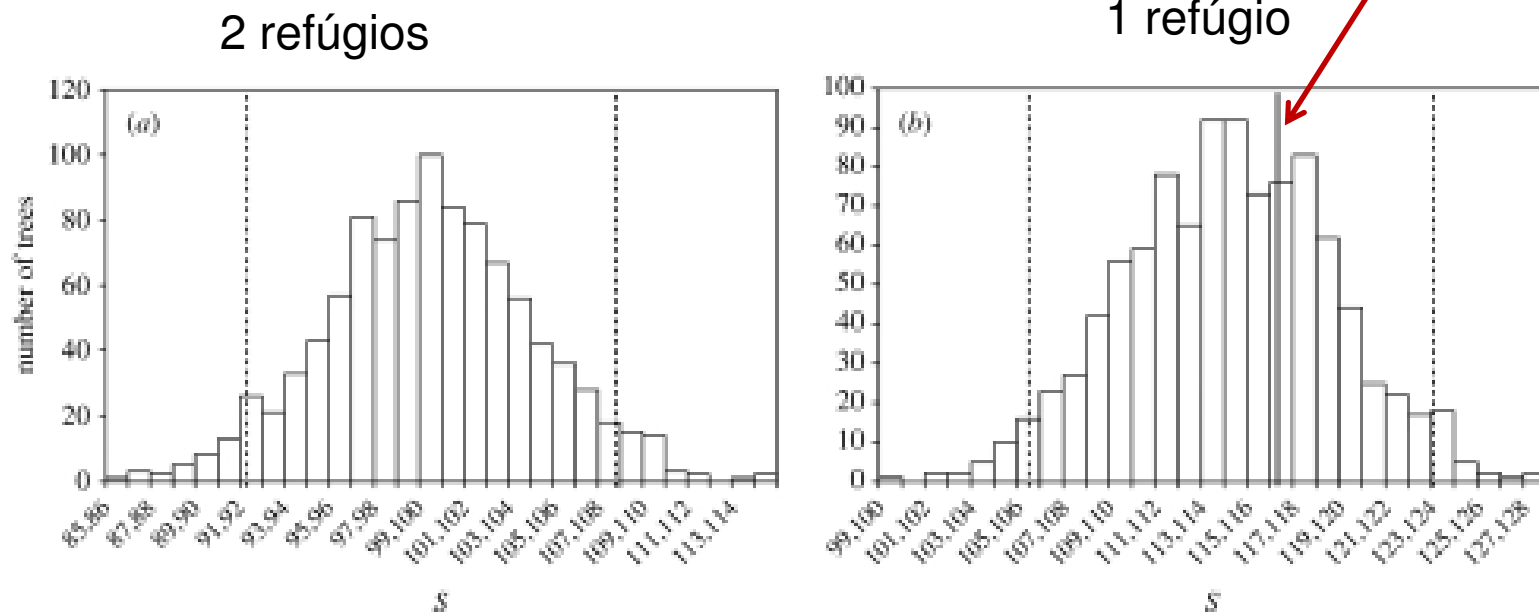
Sitta pygmaea



Hipótese 1 – Dois refúgios pleistocénicos durante as glaciações

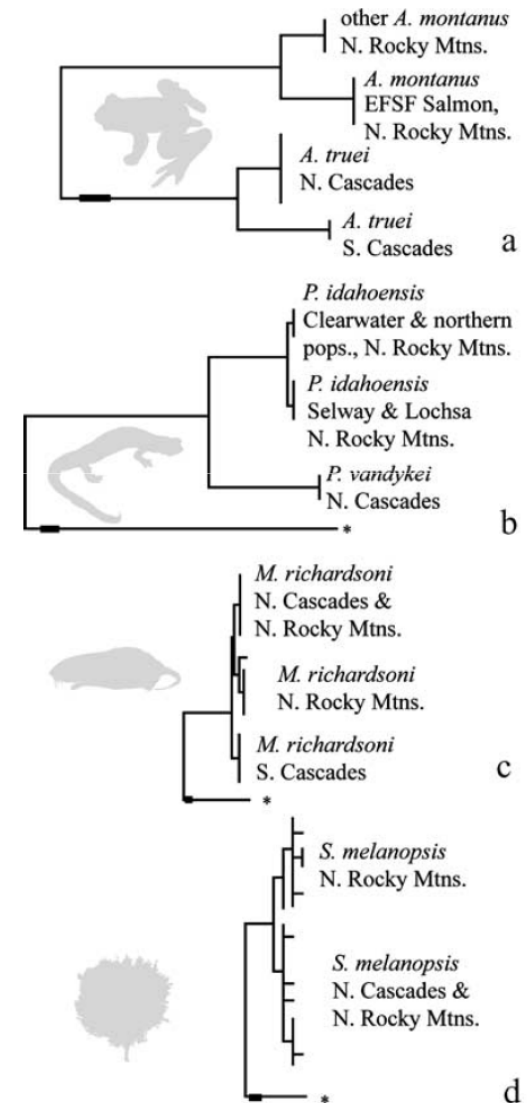
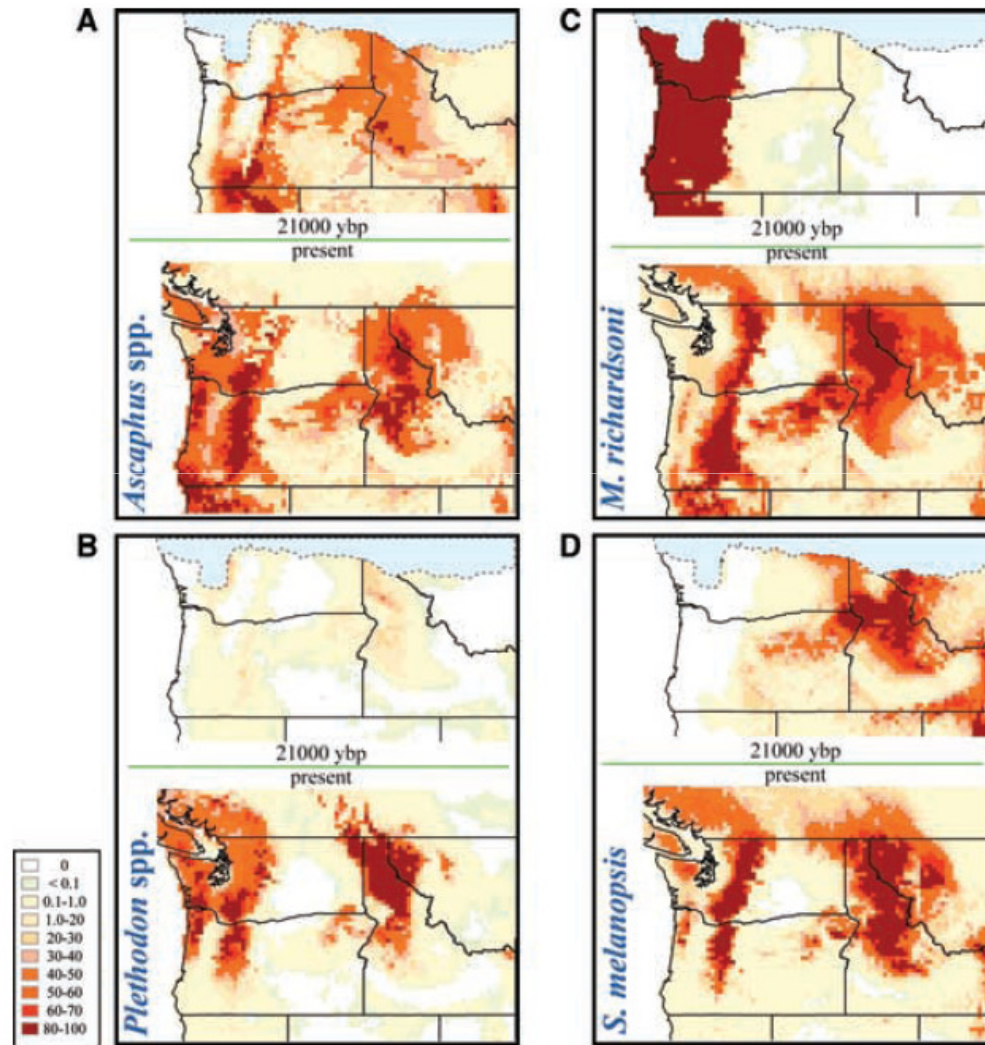
Hipótese 2 – Um refúgio pleistocénico durante as glaciações

A FILOGEOGRAFIA ESTATÍSTICA



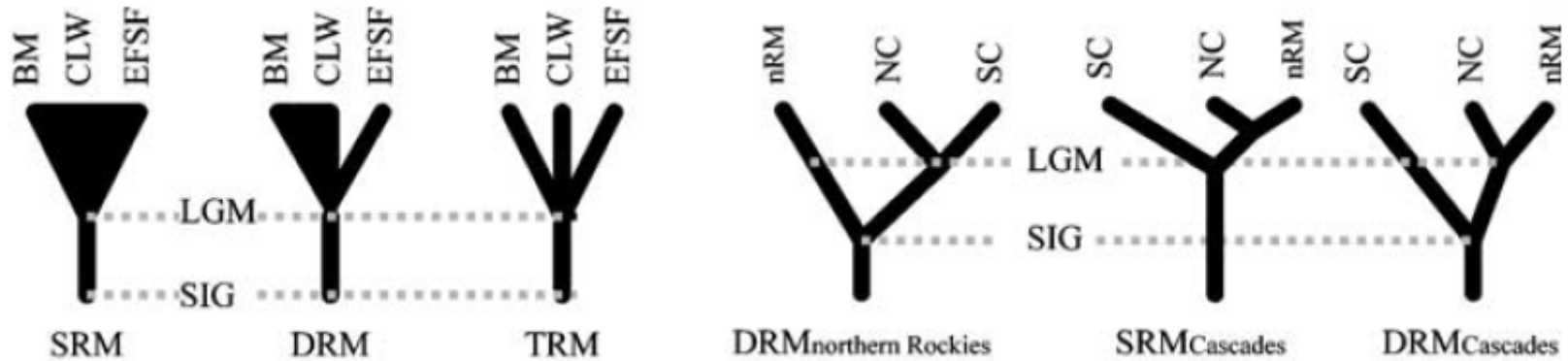
A hipóteses de dois refúgios (talvez a mais provável) é rejeitada

A FILOGEOGRAFIA ESTATÍSTICA



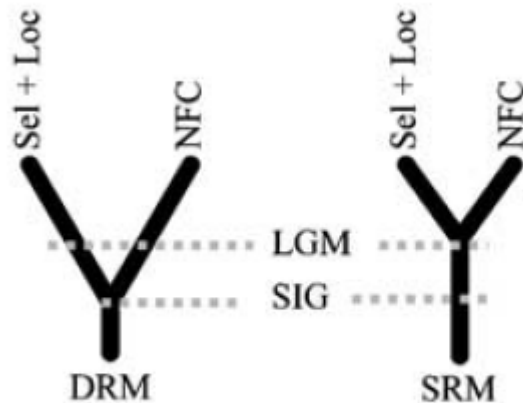
Modelos de nicho ecológico e de paleodistribuição

A FILOGEOGRAFIA ESTATÍSTICA

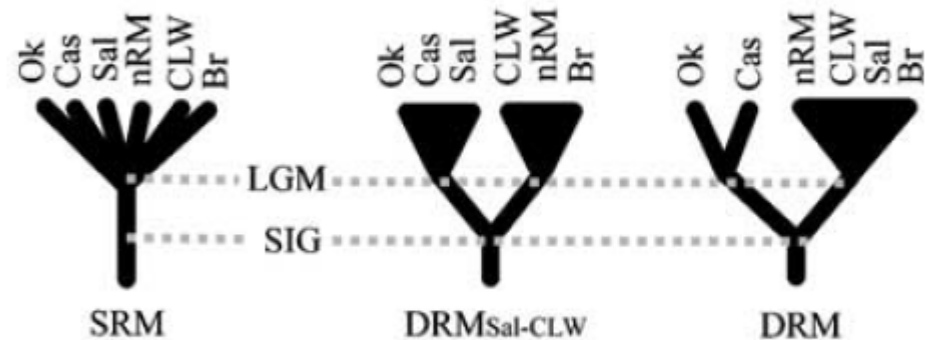


(A) *Ascapus montanus*

(C) *Microtus richardsoni*



(B) *Plethodon idahoensis*

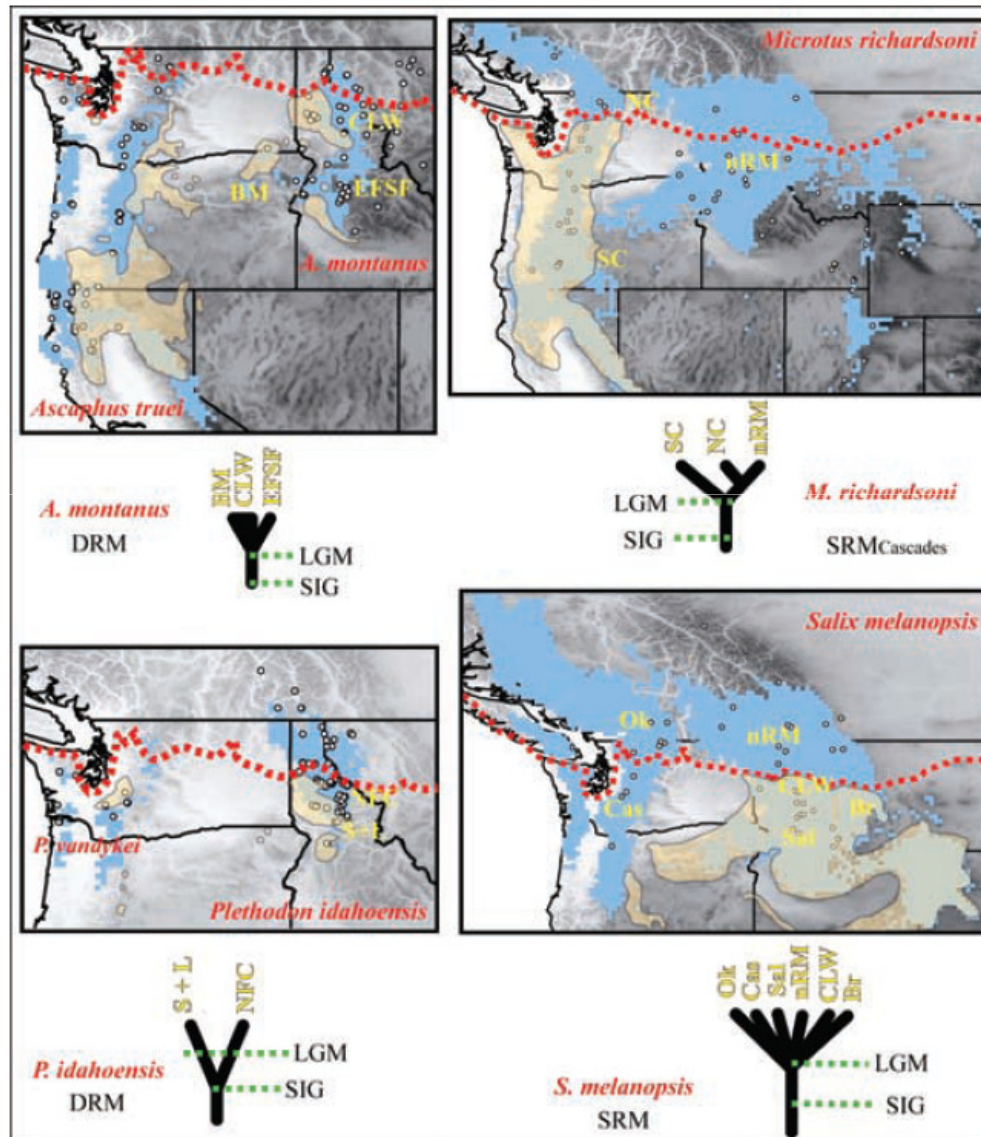


(D) *Salix melanopsis*

A FILOGEOGRAFIA ESTATÍSTICA

Taxa	H_o	Test stat	$N_{e\text{low}}$	N_e	$N_{e\text{high}}$
<i>A. montanus</i>	SRM	RECIPROCAL MONO ((CLW+BM)(EFSF))	<0.01	<0.001	<0.01
	DRM	RECIPROCAL MONO ((CLW+BM)(EFSF))	0.26	0.156	0.09
	TRM	RECIPROCAL MONO ((CLW)(BM)(EFSF))	<0.01	<0.001	<0.01
<i>P. idahoensis</i>	SRM	RECIPROCAL MONO ((SEL+LOC)(NFC))	<0.01	<0.001	<0.01
	SRM _{Bottleneck}	RECIPROCAL MONO ((SEL+LOC)(NFC))	<0.01	<0.001	<0.01
	DRM	RECIPROCAL MONO ((SEL+LOC)(NFC))	<0.01	<0.001	<0.01
	DRM _{Bottleneck}	RECIPROCAL MONO ((SEL+LOC)(NFC))	<0.01	<0.001	<0.01
<i>M. richardsoni</i>	SRM _{Cascades}	RECIPROCAL MONO ((SC)(NRM+NC))	0.61	0.390	0.29
		DIST = 0.02856 subs./site	0.08	0.113	0.13
	DRM _{Cascades}	RECIPROCAL MONO ((SC)(NRM+NC))	1.0	1.0	1.0
		DIST = 0.02856 subs./site	<0.01	<0.01	<0.01
	DRM _{northernRockies}	RECIPROCAL MONO ((SC)(NRM+NC))	<0.01	<0.01	<0.01
<i>S. melanopsis</i>	DRM	DIST = 0.00239 subs./site	<0.01	<0.01	<0.01
	DRM _{Sal-CLW}	DIST = 0.00239 subs./site	0.01	0.024	0.01
	SRM	DIST = 0.00239 subs./site	0.33	0.651	0.52

A FILOGEOGRAFIA ESTATÍSTICA



A FILOGEOGRAFIA ESTATÍSTICA

BMC Bioinformatics

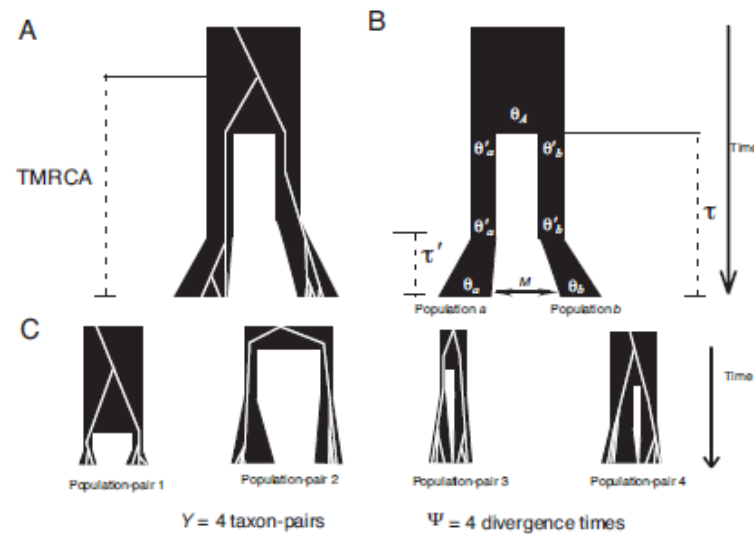


Software

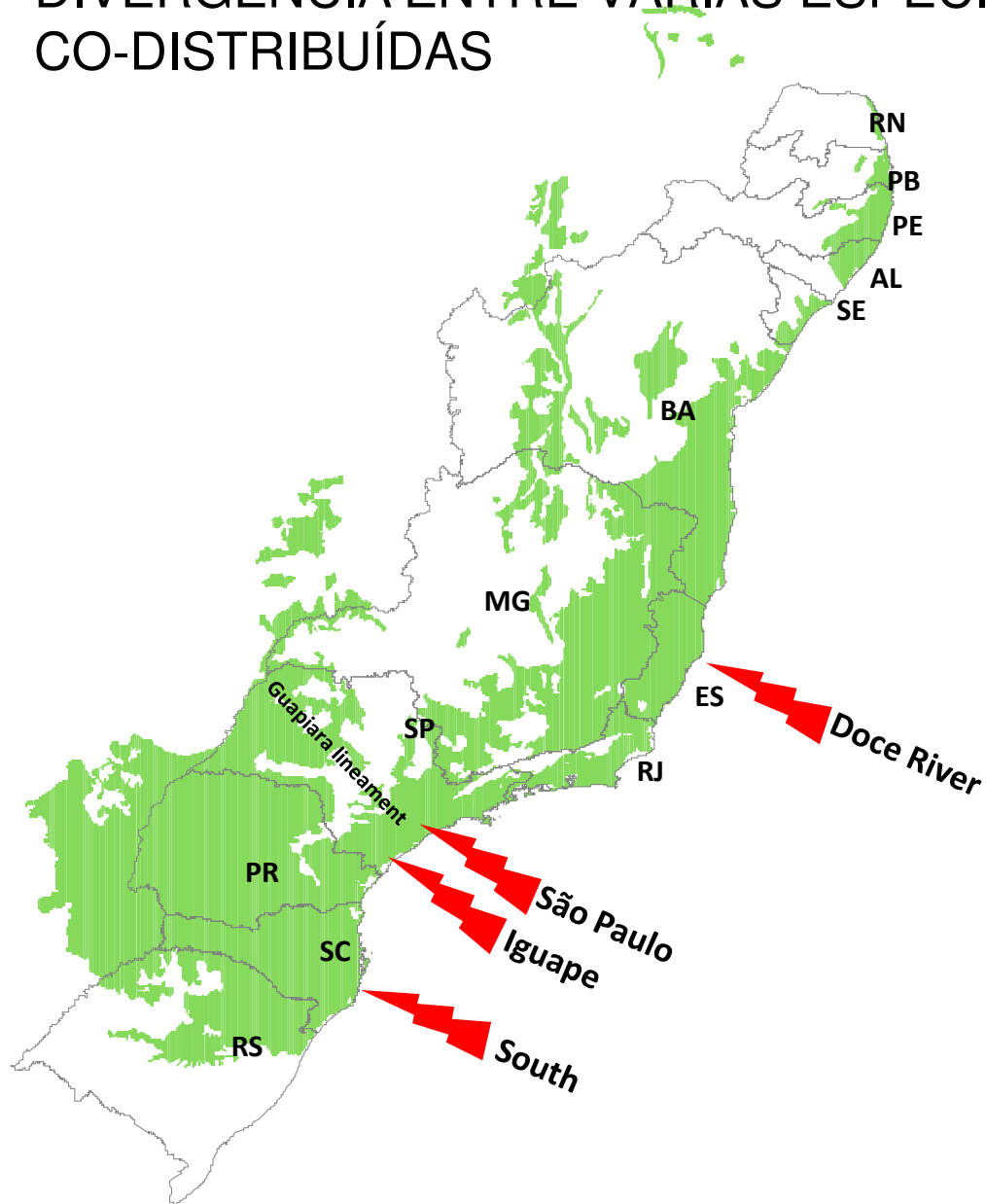
Open Access

msBayes: Pipeline for testing comparative phylogeographic histories using hierarchical approximate Bayesian computation

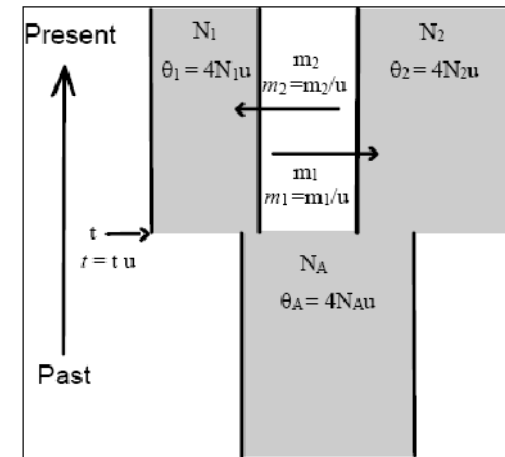
Michael J Hickerson^{*1}, Eli Stahl² and Naoki Takebayashi³



CONGRUÊNCIA TEMPORAL DA DIVERGÊNCIA ENTRE VÁRIAS ESPÉCIES CO-DISTRIBUÍDAS

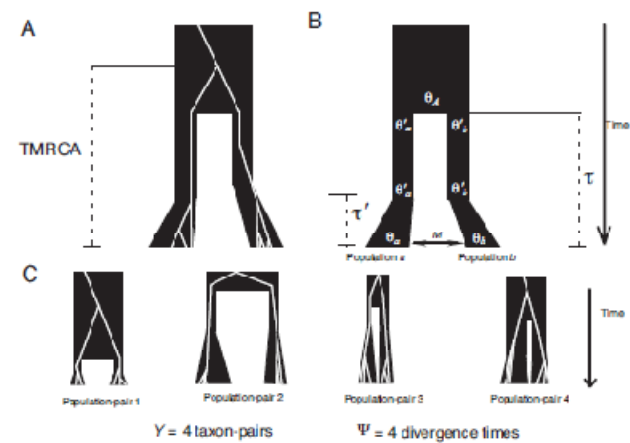


Isolation with Migration



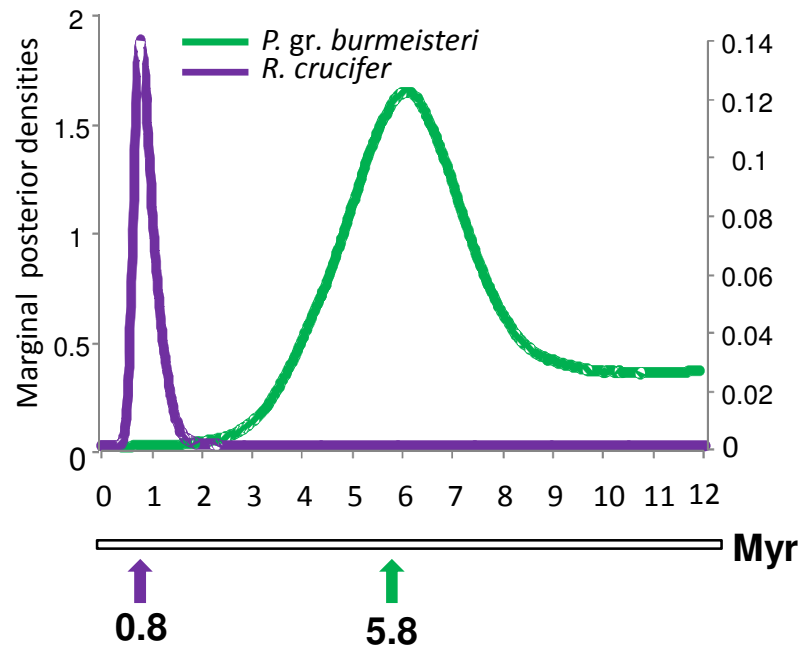
Hey J. Nielsen R. 2004. *Genetics*

msBayes: hierarchical ABC

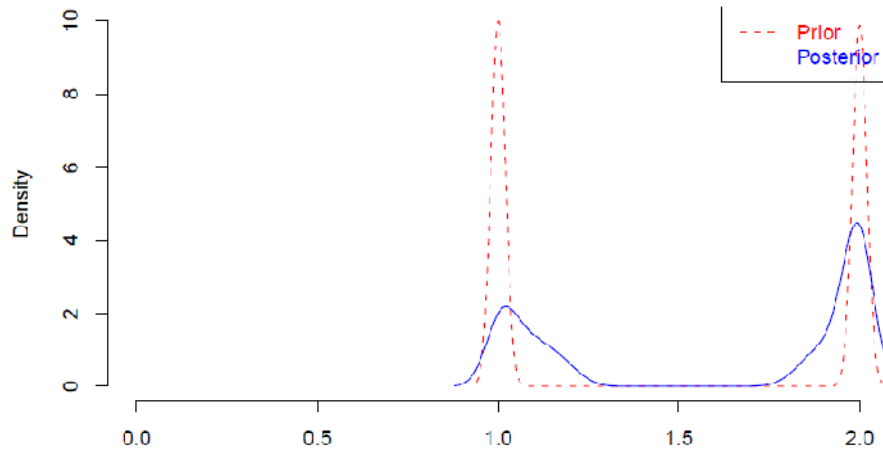


Hickerson MJ. et al. 2007. *Bioinformatics*

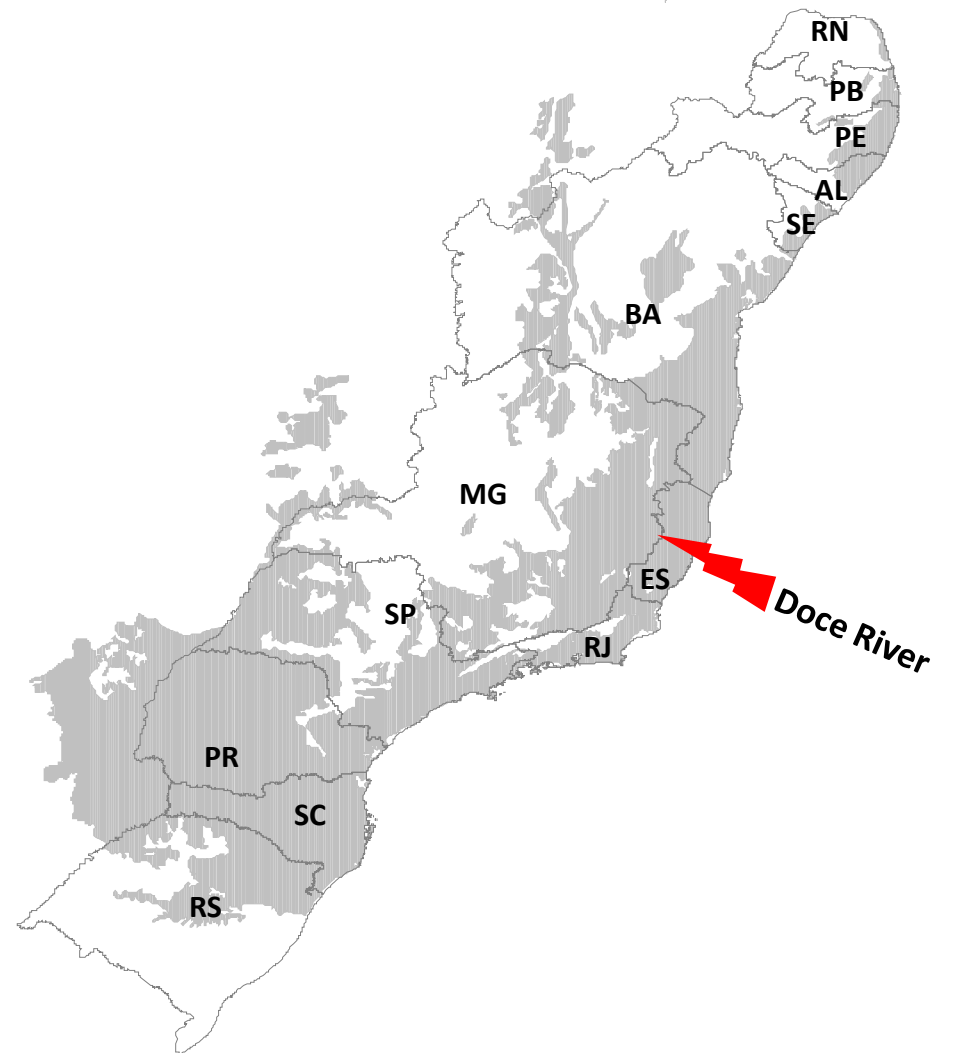
$t = T * \mu$ (time since divergence)

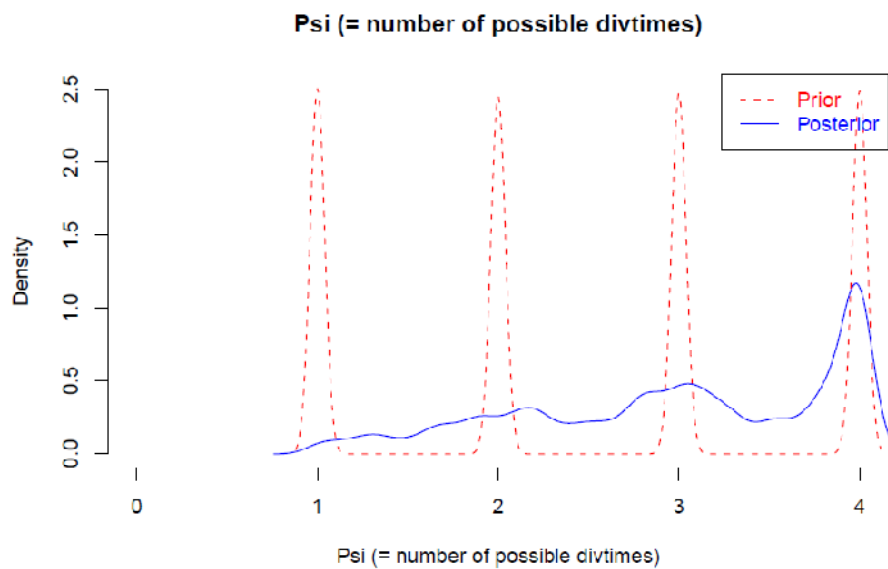
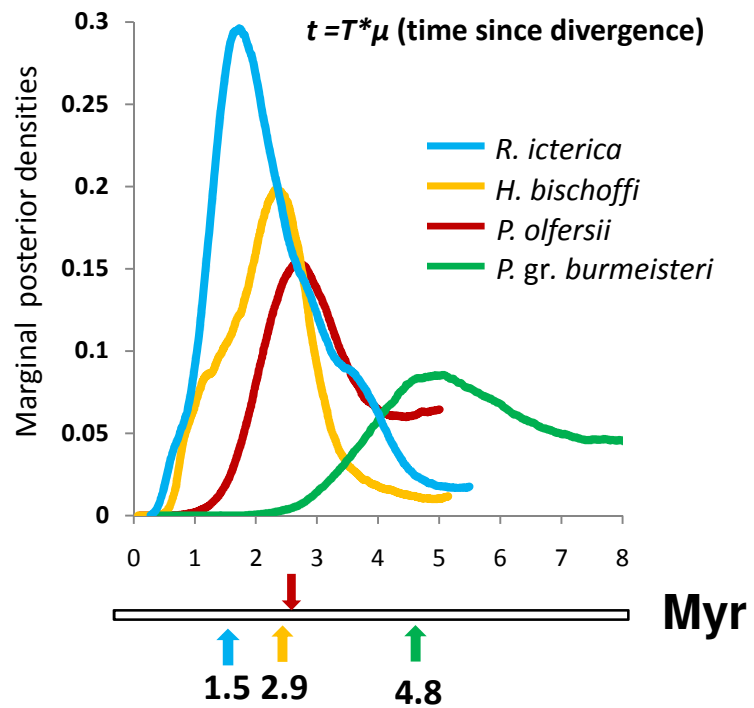


Psi (= number of possible divtimes)

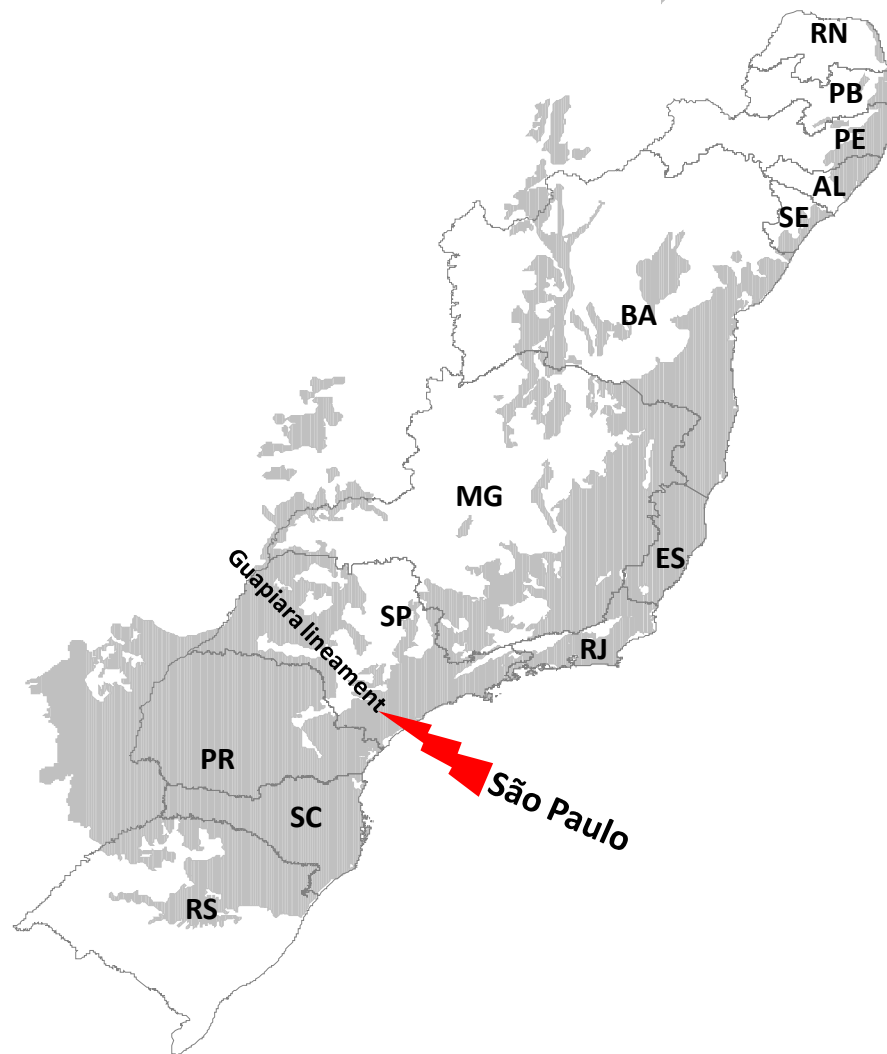


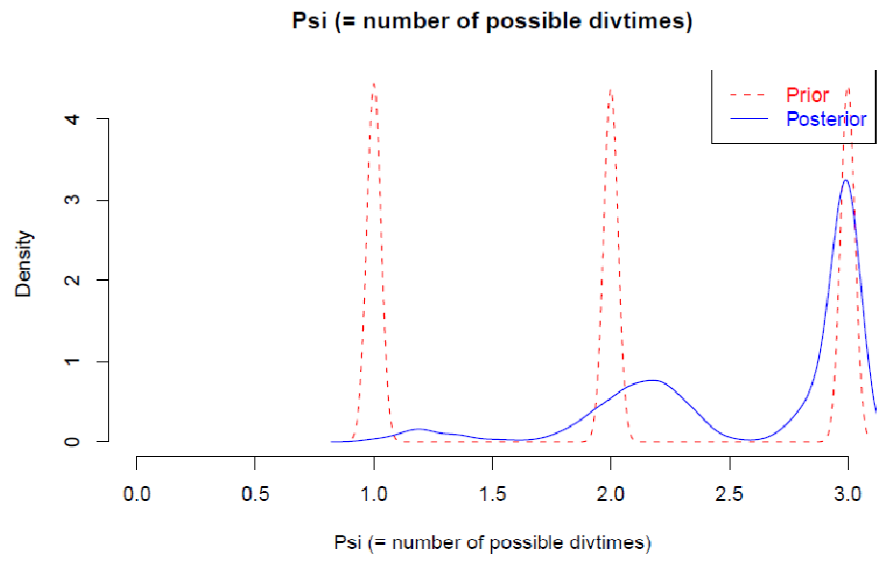
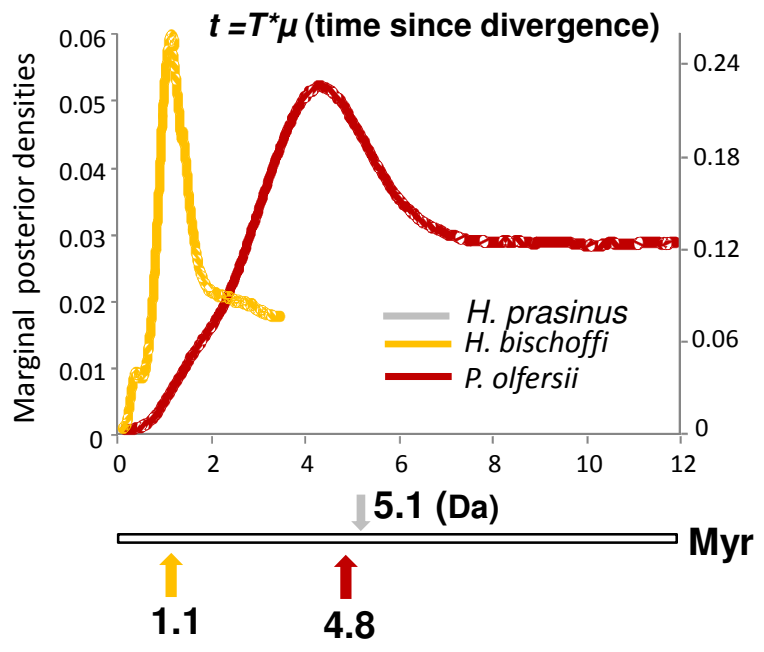
Doce River





São Paulo



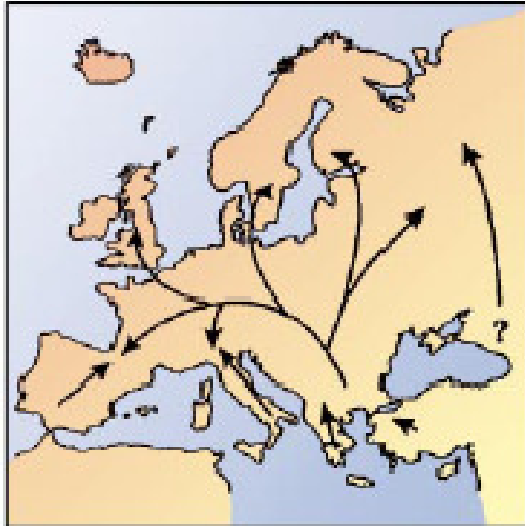


South

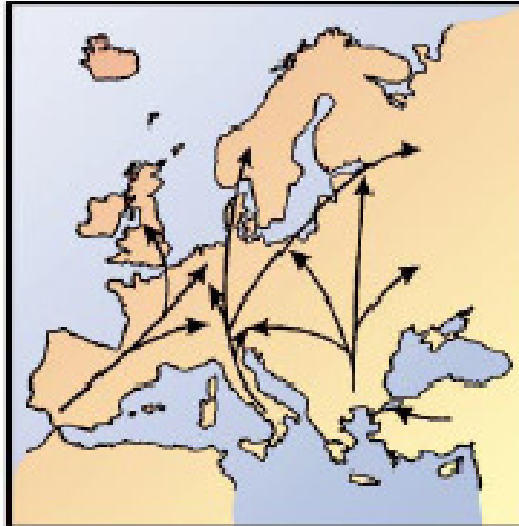


OS CENÁRIOS FILOGEOGRÁFICOS NA EUROPA

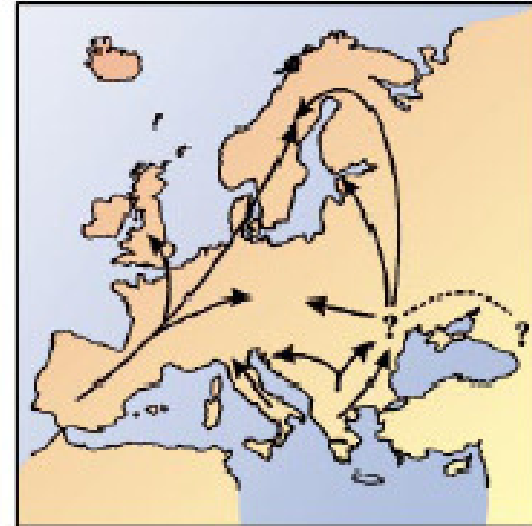
Gafanhoto



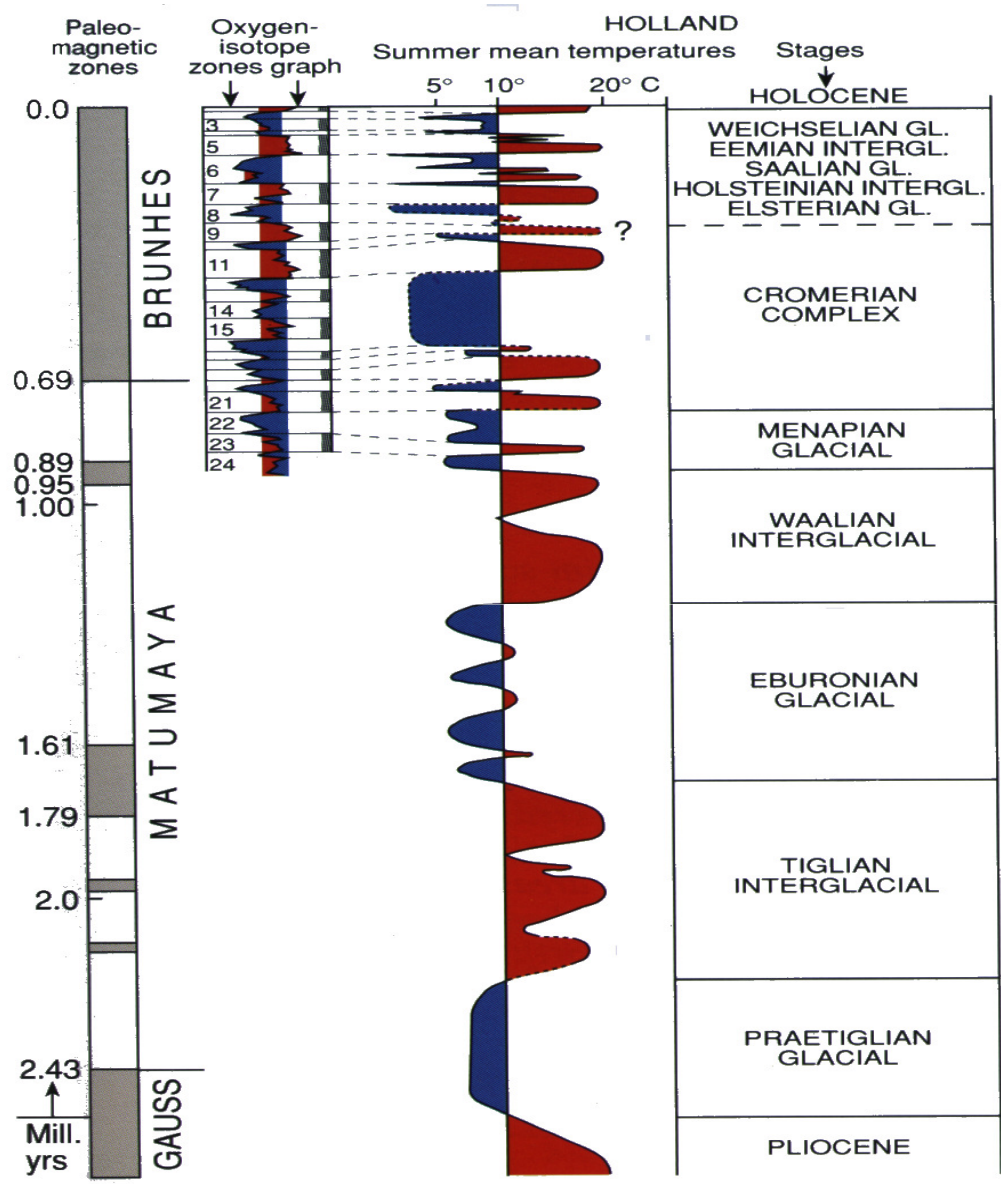
Ouriço-cacheiro

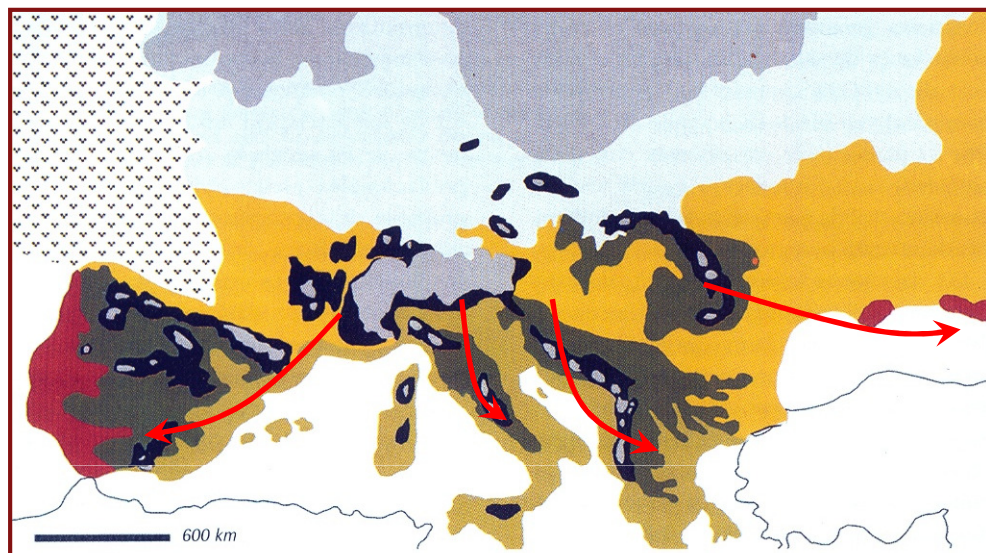


Urso-pardo



Cenários deduzidos através de estudos genéticos

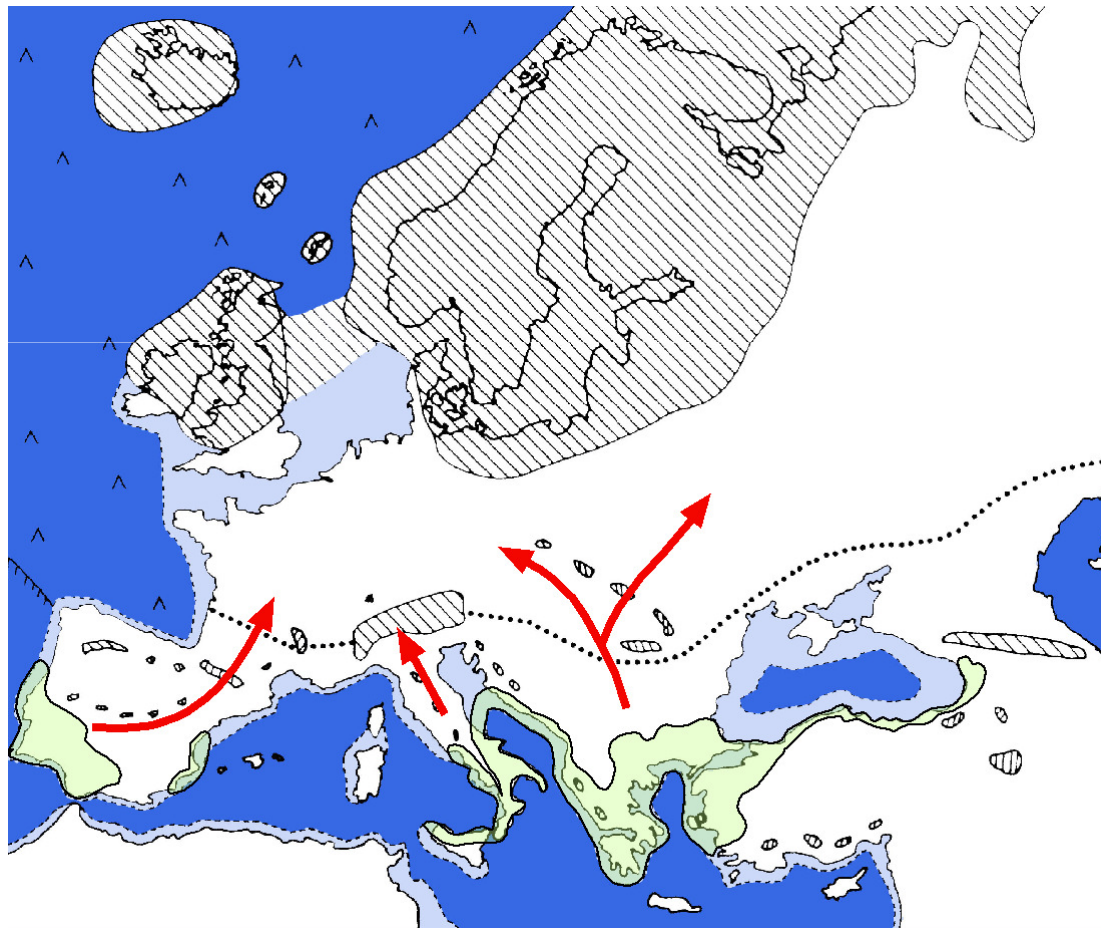




- Bosques Caducifólios
- Estepe Meridional
- Bosques Coníferas
- Estepe
- Vegetação Alpina
- Glaciares ou Gelo Permanente
- Tundra Ártica



Rotas de expansão pós-glaciares



Lack of phylogeography in European mammals before the last glaciation

Michael Hofreiter^{*,†,‡}, David Serre^{*,†,‡}, Nadin Rohland^{*}, Gernot Rabeder[§], Doris Nagel[§], Nicholas Conard[¶], Susanne Münzel[¶], and Svante Pääbo^{*}

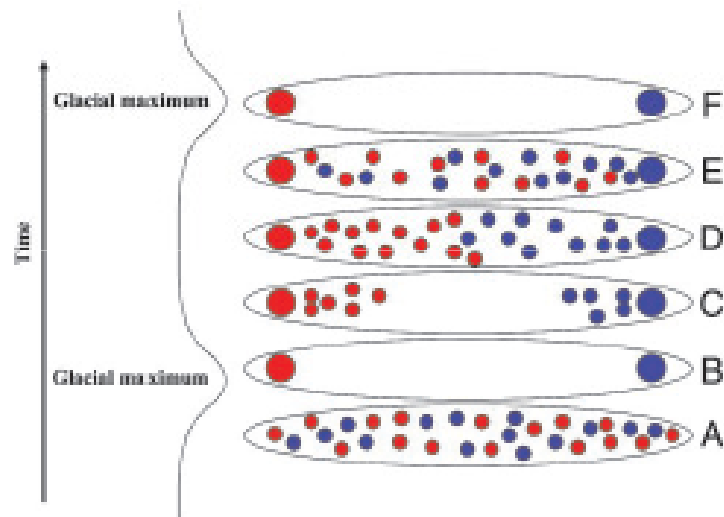


Fig. 5. Schematic figure showing the possible effects of glacial cycles on phylogeographic mtDNA patterns. (A) A population that lacks phylogeographic structure is shown shortly before a glacial maximum. (B) During a glacial maximum only individuals in the refugia survive. By drift, different mtDNA types become fixed in the two refugia. (C) After the glaciation, recolonization from the refugia occurs. (D) Individuals from the two clades meet to form a “hybrid zone.” (E) Migration eventually erases the phylogeographic pattern for most of the population’s range, but the higher population density slows migration into the refugia. (F) During a subsequent glacial maximum the refugia are likely to remain distinct with respect to mtDNA clades.

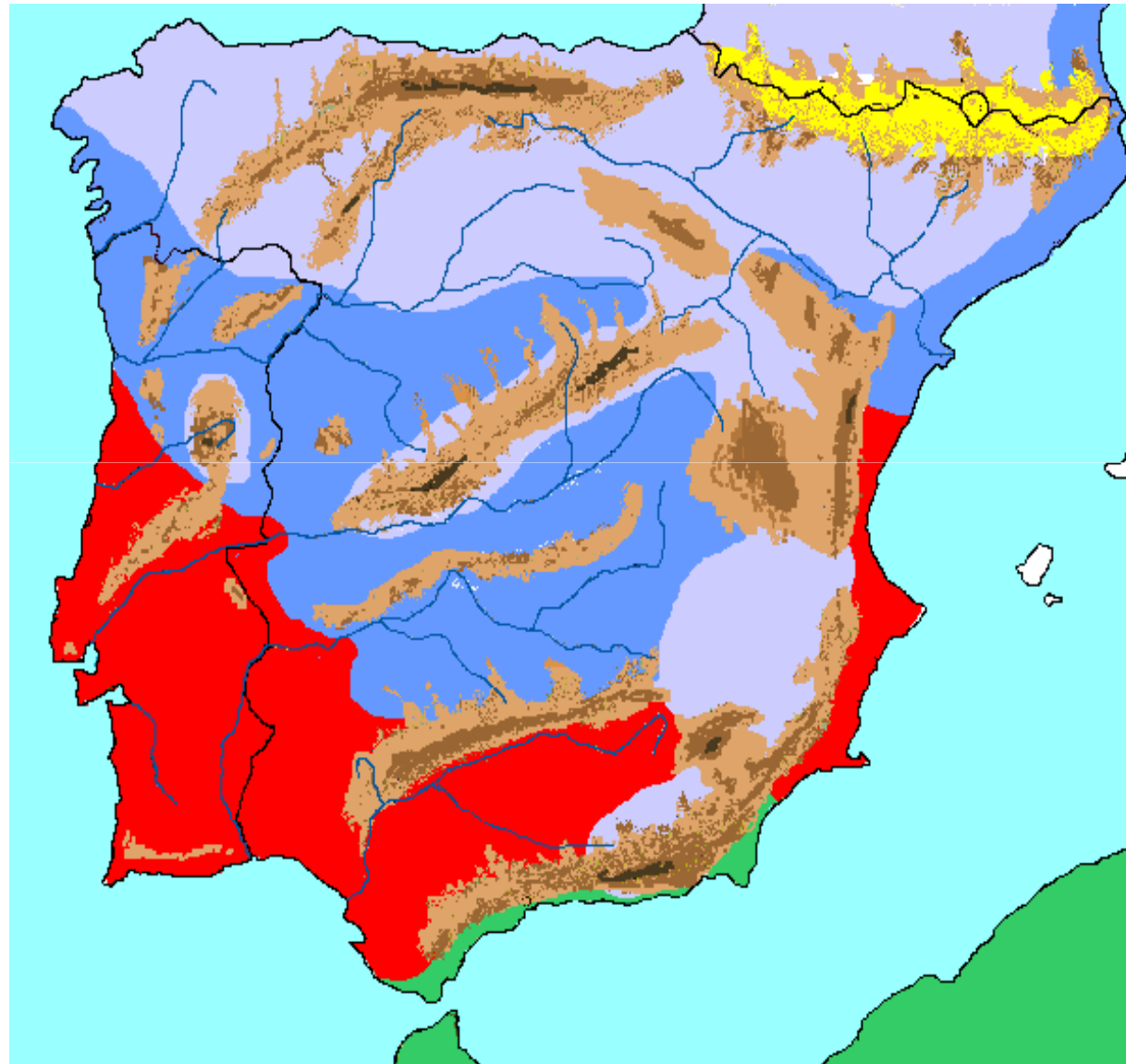
MtDNA sequencição de fósseis de mamíferos extintos

Ausência de estrutura filogeográfica no passado, em oposição à forte estrutura no presente

“Weak phylogeography”

(os padrões observados actualmente poderão ser de transição)

Península Ibérica há 18-20 milhares de anos



 **Mountains**

 **Rivers**

 **Glacier ice**

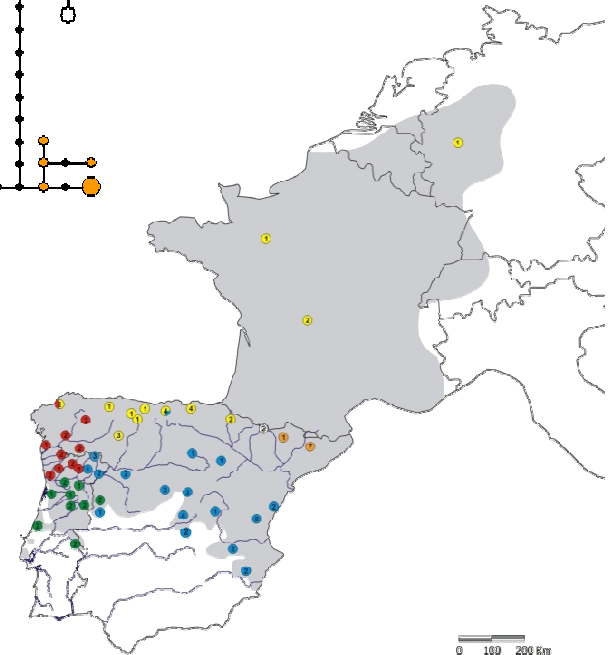
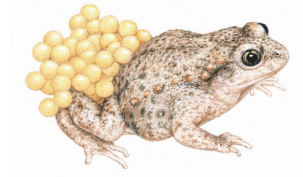
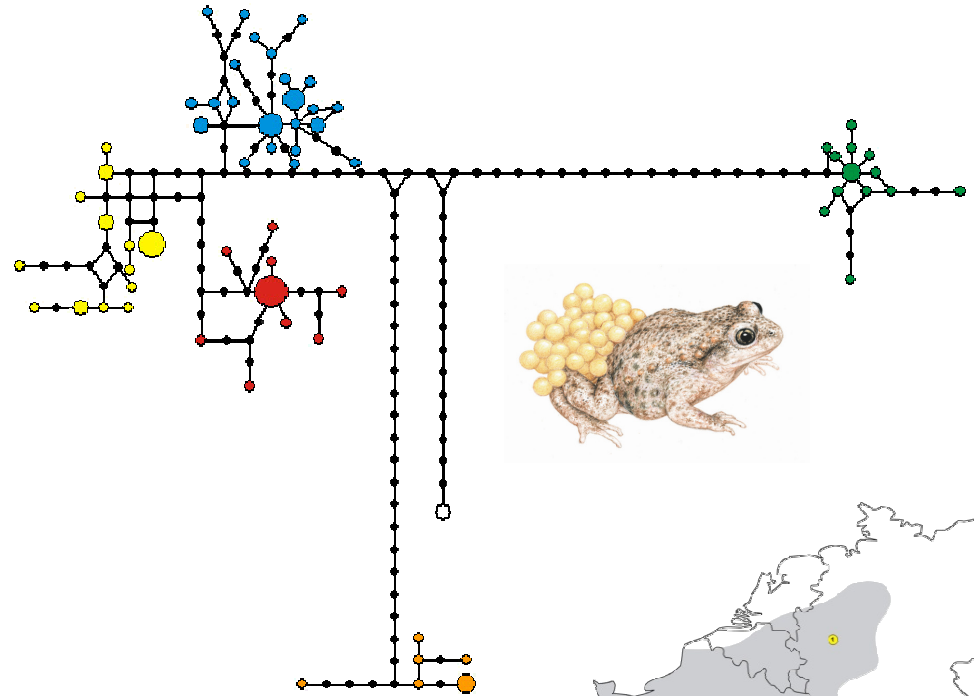
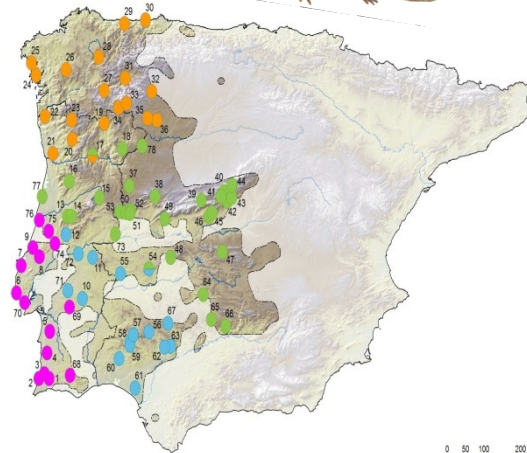
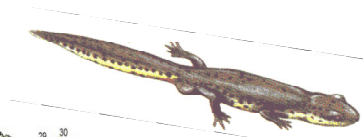
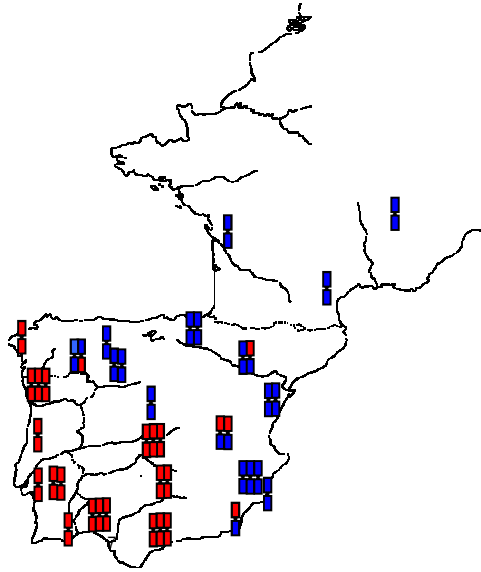
 **Tundra**

 **Boreal forest**

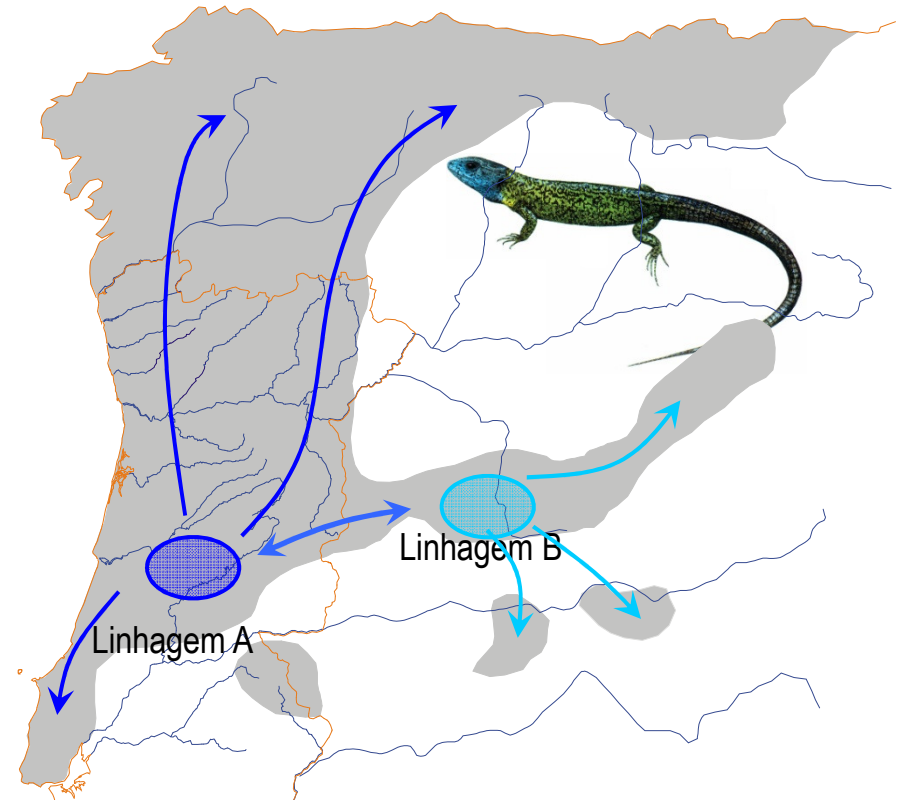
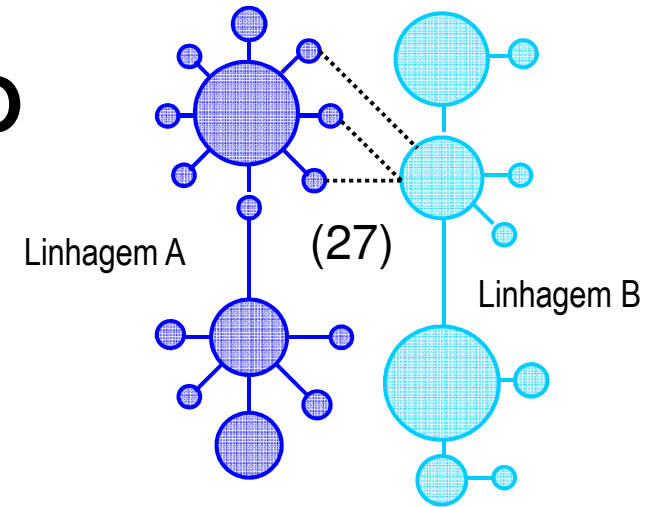
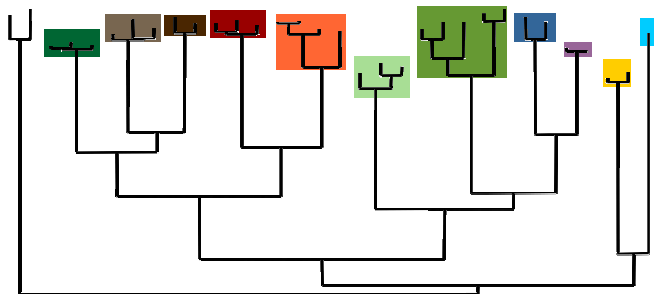
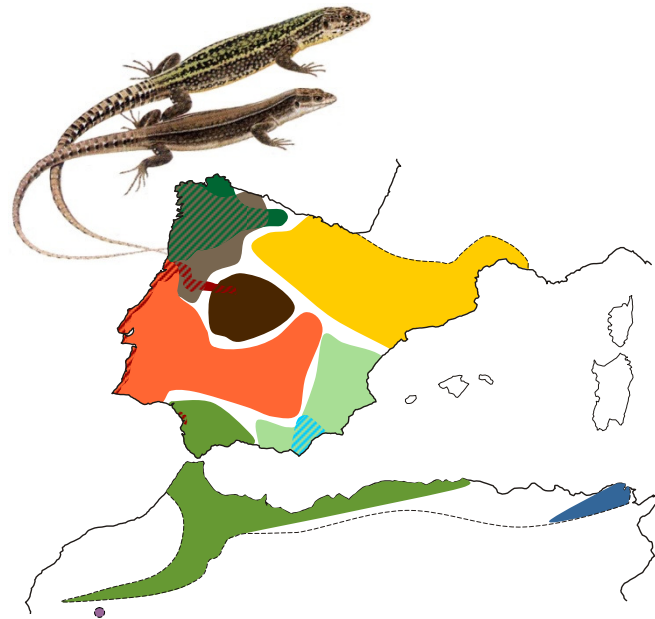
 **Temperated mixed forest**

 **Mediterranean vegetation**

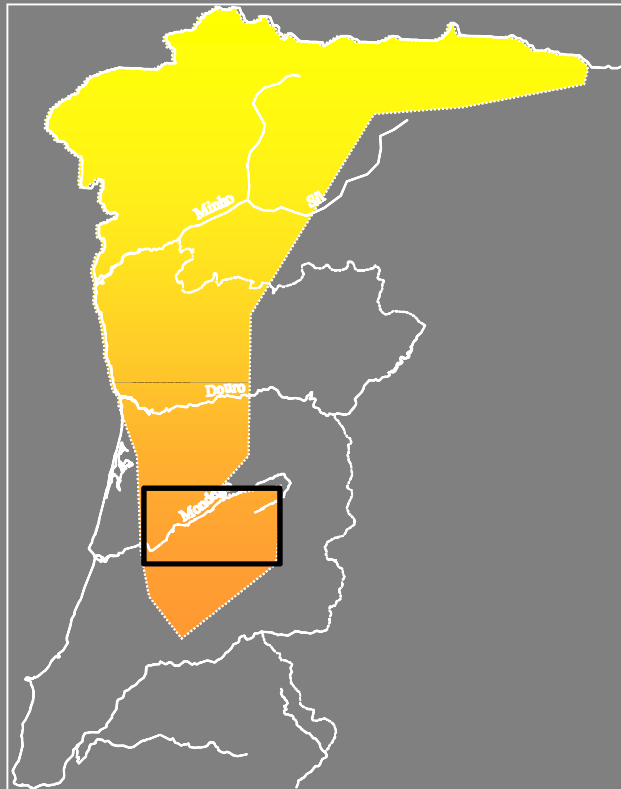
REFÚGIO DENTRO DO REFÚGIO



REFÚGIO DENTRO DO REFÚGIO



SALAMANDRA-LUSITÂNICA



1. Glaciação;

Refúgios

2. Pós-glaciação;

Expansão Sul-Norte e formação de uma área de contacto secundário

3. Distribuição actual.