PHYLOGEOGRAPHY

History of Biogeography

outline

WHAT IS PHYLOGEOGRAPHY

MARKERS

MODELS

AVISE’S HYPOTHESIS and COROLLARIES

POPULATION GENETICS, PHYLOGEOGRAPHY

Questions

- How is genetic variation distributed in time and space?
- What factors account for the spatial and temporal distribution of genetic lineages?

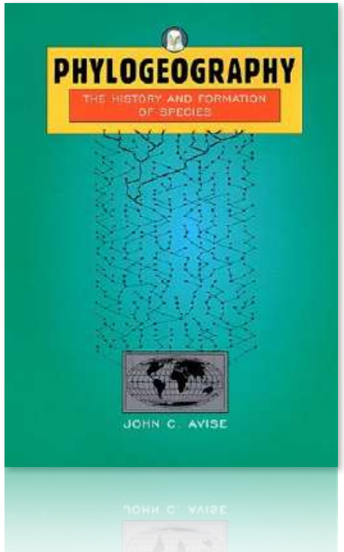
Phylogeography: introduction

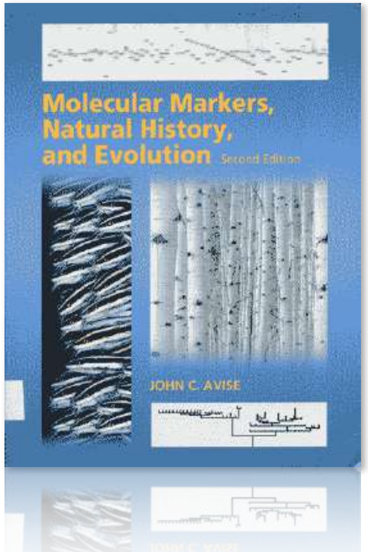
What is phylogeography?

The are the benefits of phylogeography?understanding genetic structure

How do we do it?

historical and present-day genetic data





Phylogeography: introduction

What is phylogeography?

A field of study concerned with the principles and processes governing the ***geographic distribution of genealogical lineages***, especially those within and among closely related species.

Avise 2000

Understanding genetic structure

Phylogeography

provide a means of examining the *history of genetic exchange* among populations, with the potential to distinguish *biogeographic patterns of genetic variation* caused by *gene flow* from those caused by *common ancestry*.

Schaal 1998

Phylogeography: introduction

Population genetics

Within populations
Shallow timescale

Phylogeography: introduction

Population genetics

Within populations
Shallow timescale

Phylogenetics

Between species
Deep timescale

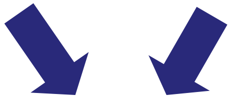
Phylogeography: introduction

Population genetics

Within populations
Shallow timescale

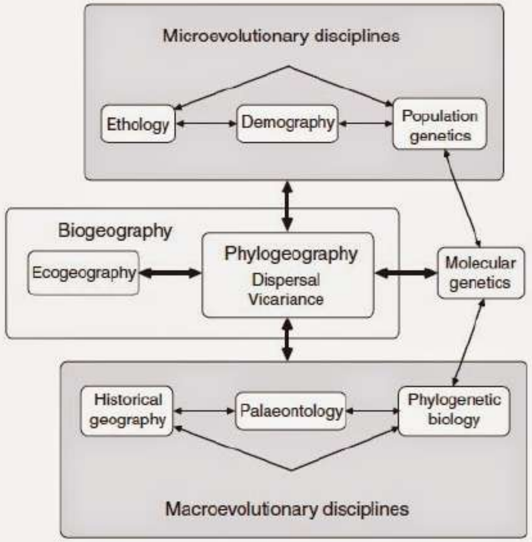
Phylogenetics

Between species
Deep timescale



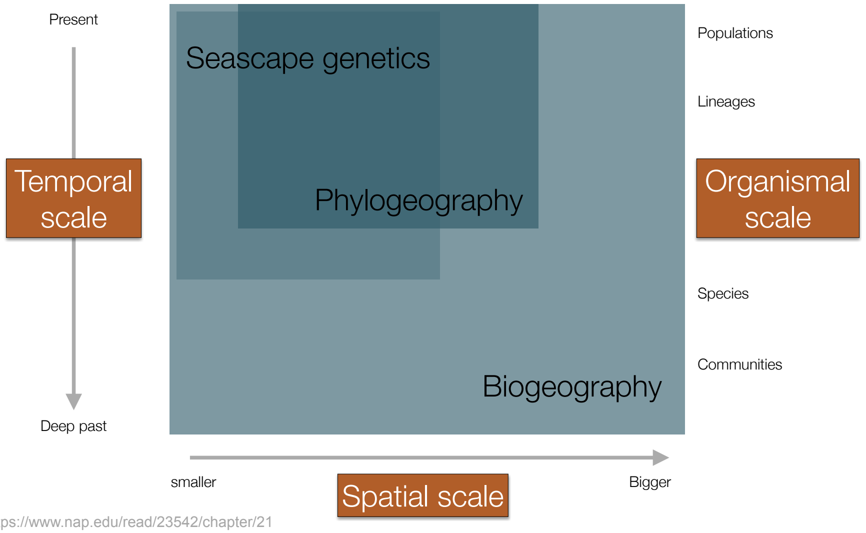
Phylogeography

Between populations and species
Medium timescale
Plus geography



Avise 2000

Phylogeography



Present

Temporal scale

Deep past

Seascape genetics

Phylogeography

Biogeography

Populations

Lineages

Organismal scale

Species

Communities

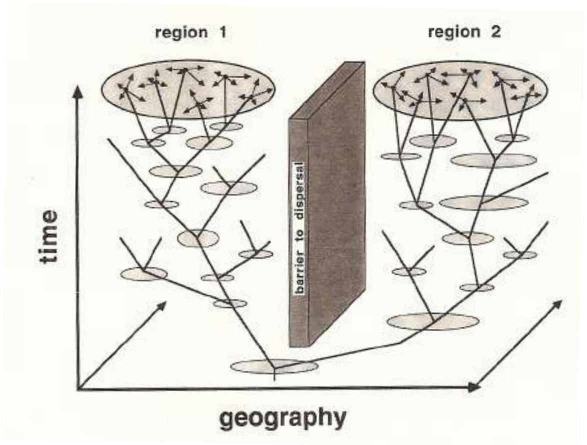
smaller

Spatial scale

Bigger

https://www.nap.edu/read/23542/chapter/21

Phylogeography: introduction



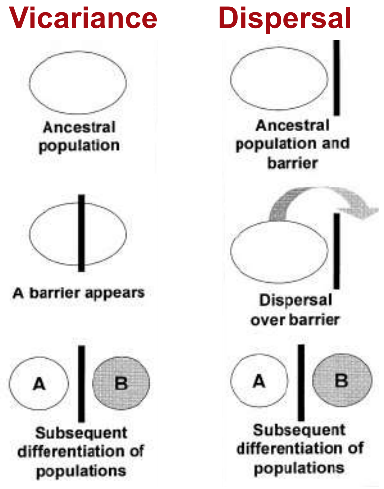
Gene genealogies of interest are mapped in space and time.

Avise 2000

Phylogeography: introduction

Goal: To understand the factors contributing to the formation of population (or species-level) genetic structure.

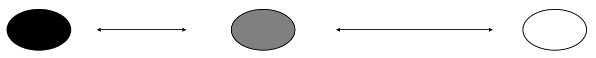
Can evaluate alternative historical scenarios that account for current spatial patterns.



Determinants of genetic structure

Historical relationships and contemporary gene flow influence spatial patterns of genetic variation.

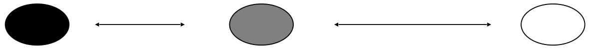
Similarity due to gene flow



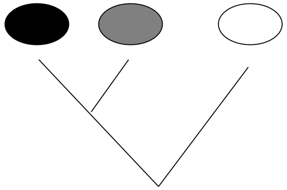
Determinants of genetic structure

Historical relationships and contemporary gene flow influence spatial patterns of genetic variation.

Similarity due to gene flow



Similarity due to historical relationships



Understanding genetic structure

Within species, *genetic exchange* has been emphasized as the cause of similarity.

Understanding genetic structure

Traditionally:

Within species, *genetic exchange* has been emphasized as the cause of similarity.

Among species, *historical relationship* has been emphasized as the cause of similarity.

Understanding genetic structure

We can tease apart
contemporary forces of genetic exchange
from
historical relationships

Understanding genetic structure

Phylogeography uses principles from

population genetics

and

phylogenetics

Understanding genetic structure

Phylogeography uses principles from **population genetics** and **phylogenetics.**

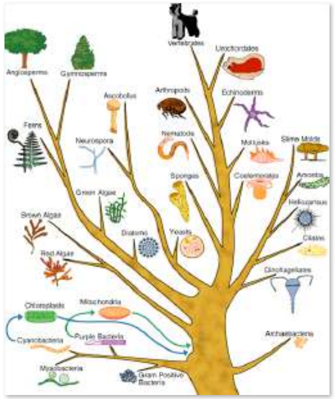
However…

Understanding genetic structure

Does not assume an equilibrium between genetic drift and gene flow

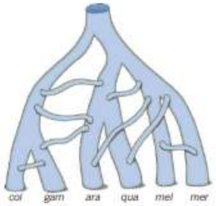
Assumes reticulation

Understanding genetic structure



Evolution can be displayed in a bifurcating pattern

There are exceptions…. Lateral gene transfer



More web than tree. The pattern of evolution seen in the An. gambiae species complex resembles a network more than a tree. This type of evolutionary network is referred to as reticulate evolution (10).

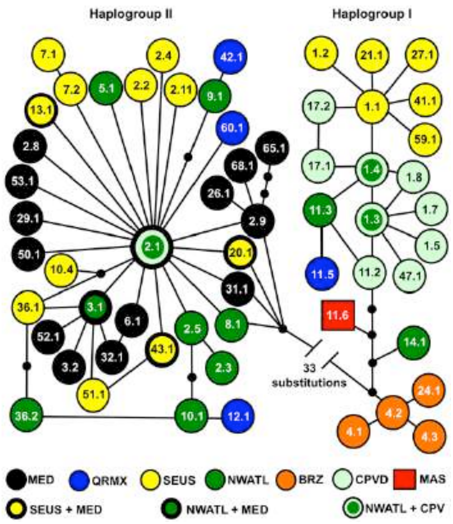
Understanding genetic structure

Does not assume an equilibrium between genetic drift and gene flow

Assumes reticulation

Relies on the use of **gene genealogies** [phylogeographic methods provide historical information that individuals or populations cannot]

Genetic variation



http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0085956

Genetic variation

We need significant genetic variation at the appropriate level
(i.e. among the populations or taxonomic units under investigation)

Genetic variation

We need significant genetic variation at the appropriate level
(i.e. among the populations or taxonomic units under investigation)

Mitochondrial DNA
great for animals
low rates of molecular evolution in plants, intramolecular recombination

Genetic variation

We need significant genetic variation at the appropriate level
(i.e. among the populations or taxonomic units under investigation)

Mitochondrial DNA
great for animals
low rates of molecular evolution in plants, intramolecular recombination

Chloroplast DNA
faster rates of molecular evolution but still low (and variable among species)
no recombination
chloroplast capture can occur (e.g. in oaks, sunflowers, poplar)

Organelle genomes are essentially a single locus - inferences based on multiple loci give more accurate estimations of the population history of a species

---

Genetic variation

We need significant genetic variation at the appropriate level
(i.e. among the populations or taxonomic units under investigation)

Mitochondrial DNA
great for animals
low rates of molecular evolution in plants, intramolecular recombination

Chloroplast DNA
faster rates of molecular evolution but still low (and variable among species)
no recombination
chloroplast capture can occur (e.g. in oaks, sunflowers, poplar)

Organelle genomes are essentially a single locus - inferences based on multiple loci give more accurate estimations of the population history of a species

Nuclear DNA
problems of intra-allelic recombination, heterozygosity and gene families

---

Phylogeography: early years

Phylogeography was essentially *descriptive*:

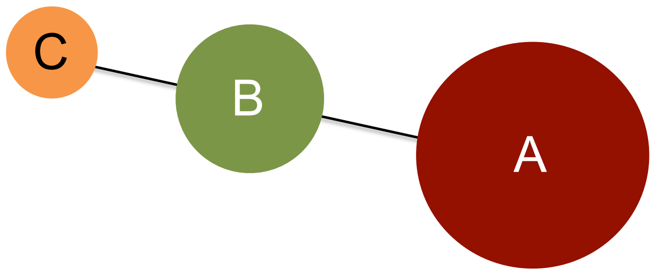
Plot haplotypes on map

Classify pattern of phylogeographic structure

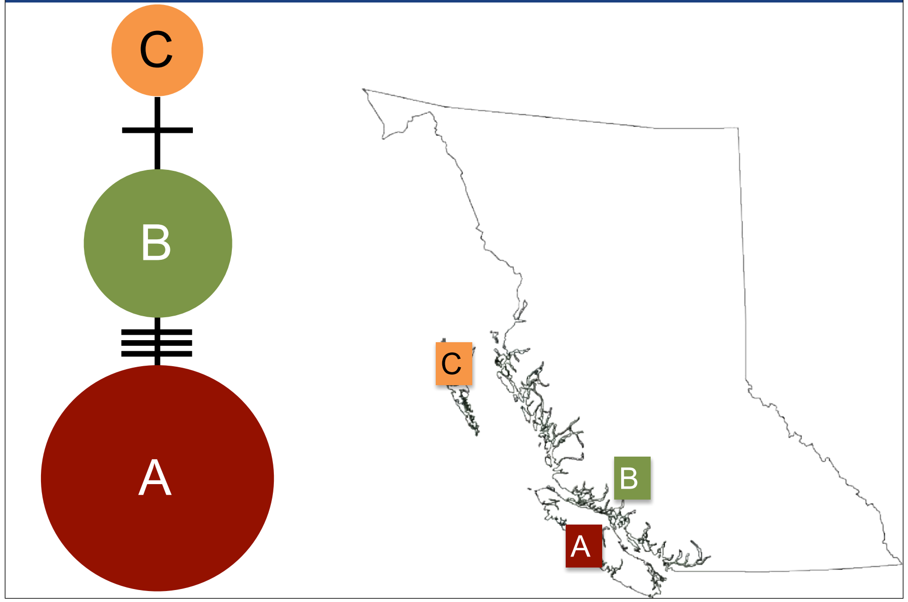
Consider historical explanations

Look for concordance among different species (comparative phylogeography)

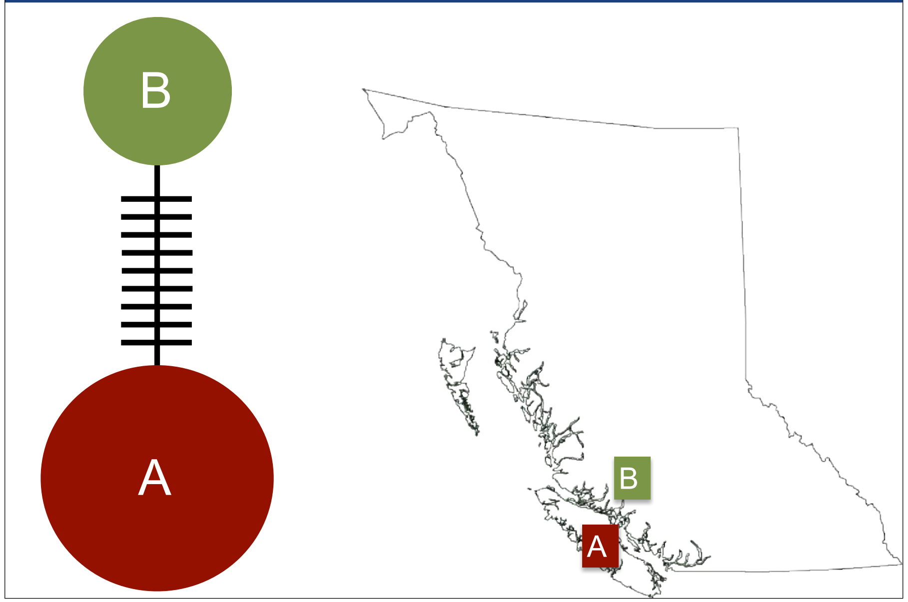
Haplotype network



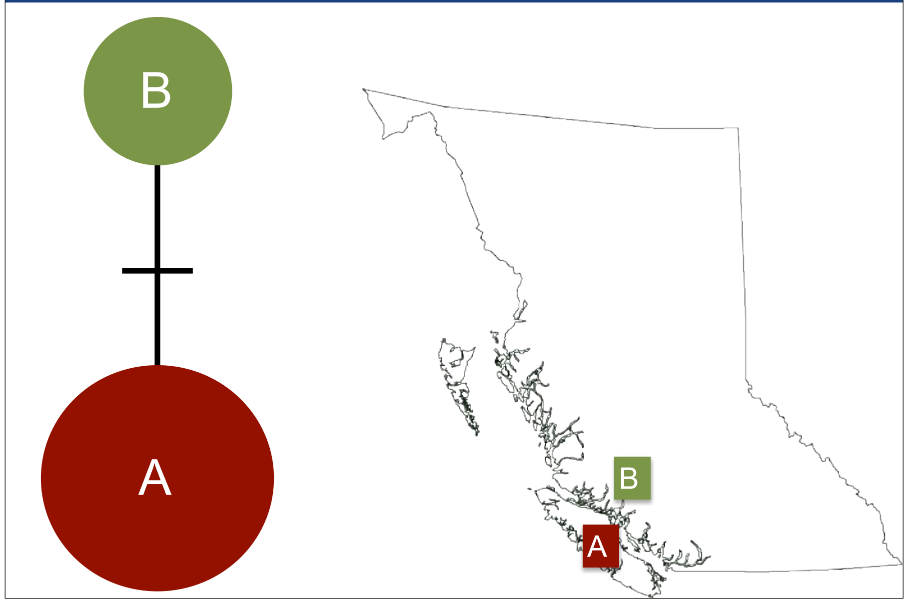
Where did C originate from?



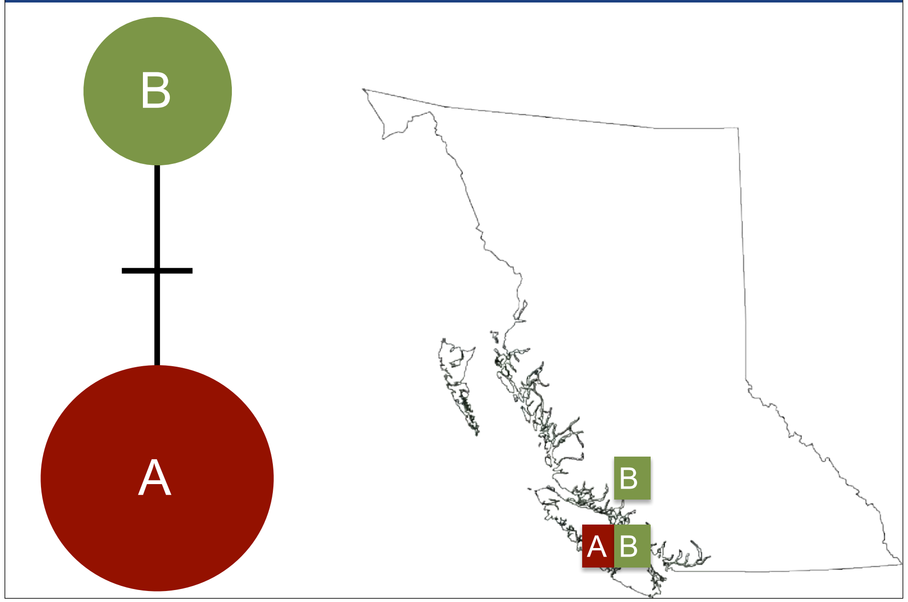
Deep divergence



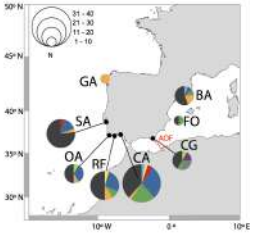
Shallow divergence

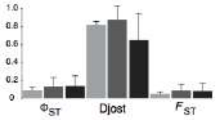


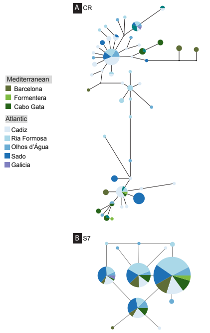
Gene flow



Phylogeography: early years







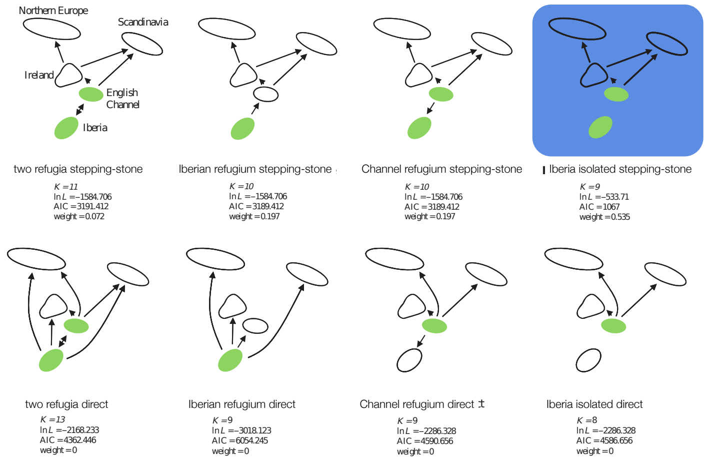
			Mitochondrial D-loop						First intron of S7 gene				
Region	Locations	Code	*N*	NH	NP	Haplotype diversity ± s.d.	Nucleotide diversity ± s.d	PS	*N*	NH	Gene diversity ± s.d.	Nucleotide diversity ± s.d	Observed Heterozygosity
M	Barcelona	BA	16	9	3	0.92 ± 0.04	0.0520 ± 0.0274	38	15	5	0.65 ± 0.07	0.0017 ± 0. 0013	0.53
	Formentera	FO	2	2	1				3	1			
	Cabo de Gata	CG	12	6	2	0.88 ± 0.06	0.0167 ± 0.0098	10	13	4	0.64 ± 0.07	0.0015 ± 0.0013	0.46
A	Cadiz	CA	39	19	13	0.88 ± 0.04	0.0151 ± 0.0084	25	22	6	0.72 ± 0.05	0.0019 ± 0.0015	0.41
	Ria Formosa	RF	23	13	8	0.94 ± 0.03	0.0111 ± 0.0066	17	40	6	0.62 ± 0.04	0.0016 ± 0.0013	0.45
	Olhos de Água	OA	12	7	6	0.91 ± 0.08	0.0170 ± 0.0100	18	12	5	0.78 ± 0.07	0.0023 ± 0.0017	0.42
	Sado	SA	26	7	4	0.72 ± 0.08	0.0145 ± 0.0083	12	30	4	0.73 ± 0.03	0.0019 ± 0.0014	0.40
	Galicia	GA	1	1	1				1	1			

Notes.
N, number of individuals per location; NH, haplotype richness; NP, number of private haplotypes; PS, number of polymorphic sites.

Hypoteses testing

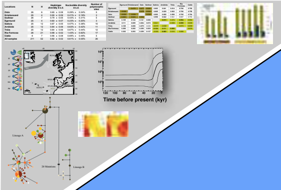
Stepping-stone

Direct



Provan & Maggs 2011. PRSL

The idea



Questions
Hypotheses

Hypothesis:
Oceanic fronts may act as a barrier to dispersal for littoral fish irrespective of the life-history traits.

Predictions:
(1) populations separated by oceanic fronts belong to genetically differentiated subunits;
(2) the degree of connectivity between populations is related to early-life-history traits and dispersal capabilities.

Testing by:
(1) plotting standardized species Fst
(2) estimating pairwise Fst;
(3) mapping posterior probabilities of population membership and
(4) mapping spatial location of genetic discontinuities for different species.

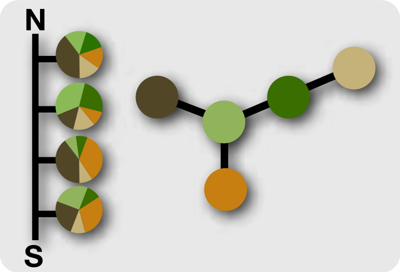
Galarza (2009). PNAS

Expectations in phylogeographic data

How do we reconstruct the **origin**, **dispersal**, and **extinction** of taxa?

Phylogeographic data

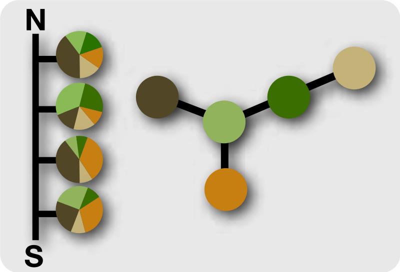
Null model



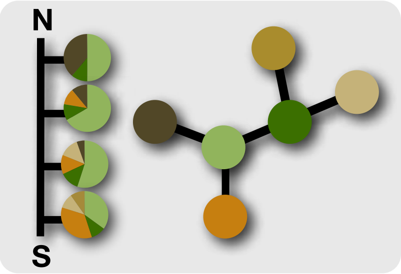
Maggs et al. (2008). *Ecology.*

Phylogeographic data

Null model



Latitudinal cline

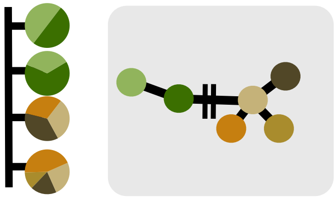


Maggs et al. (2008). *Ecology.*

Phylogeographic data: expectations

genealogical
concordance

no
geographic
contact



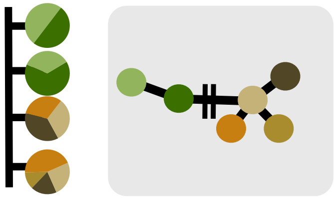
Maggs et al. (2008). *Ecology.*

Phylogeographic data: expectations

genealogical
concordance

no genealogical
concordance

no
geographic
contact



?

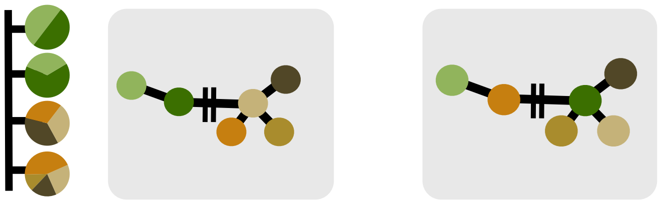
Maggs et al. (2008). *Ecology.*

Phylogeographic data: expectations

genealogical concordance

no genealogical concordance

no geographic contact



Maggs et al. (2008). *Ecology*.

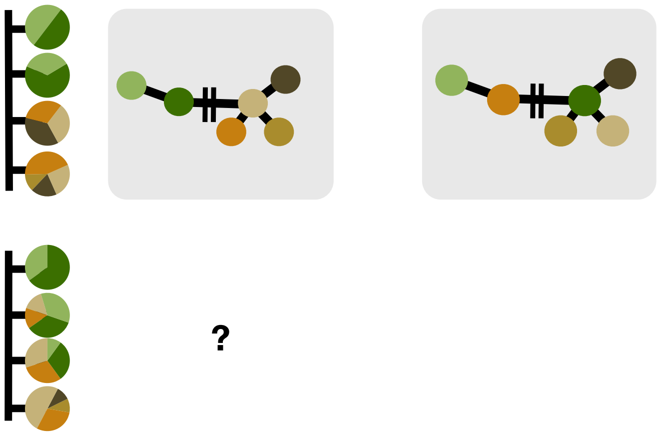
Phylogeographic data: expectations

genealogical concordance

no genealogical concordance

no geographic contact

geographic contact



?

Maggs et al. (2008). *Ecology*.

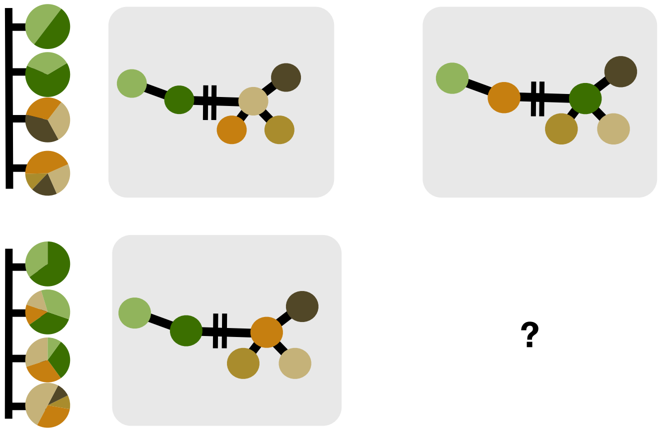
Phylogeographic data: expectations

genealogical concordance

no genealogical concordance

no geographic contact

geographic contact



?

Maggs et al. (2008). *Ecology*.

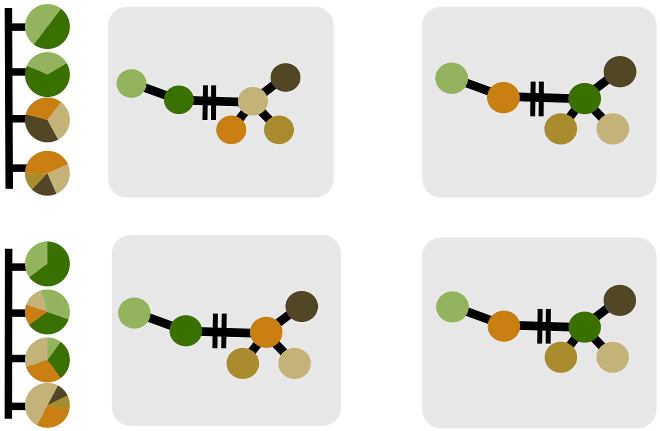
Phylogeographic data: expectations

genealogical concordance

no genealogical concordance

no geographic contact

geographic contact



Maggs et al. (2008). *Ecology*.

Avise phylogeography

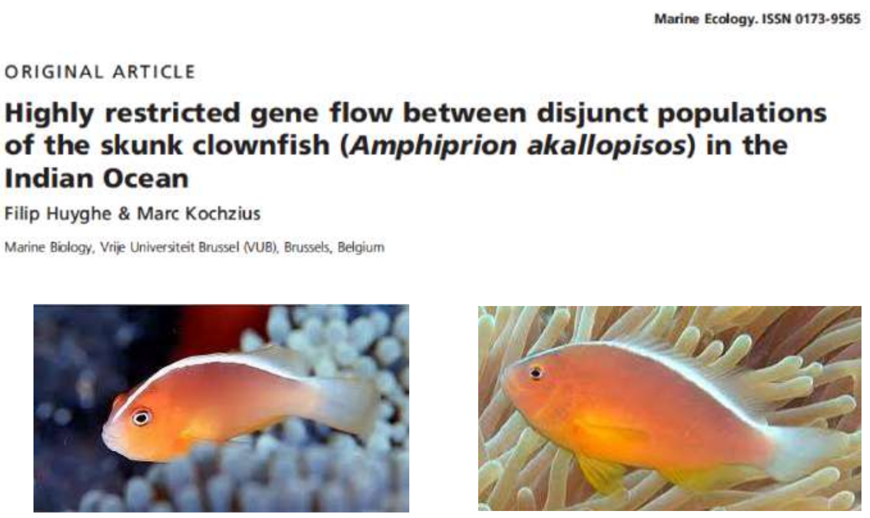
Phylogeographic hypotheses and corollaries

(Corollary:
1.something that naturally follows;
2.proposition that follows with little or no proof required from one already proven;
3. A deduction or an inference;
4. A natural consequence or effect; a result)

Hypothesis 1

Most species are composed of
geographic populations
whose members occupy recognisable matrilineal branches.

Populations of most species display
significant phylogeographic structure
supported by mtDNA data.



Marine Ecology. ISSN 0173-9565

ORIGINAL ARTICLE

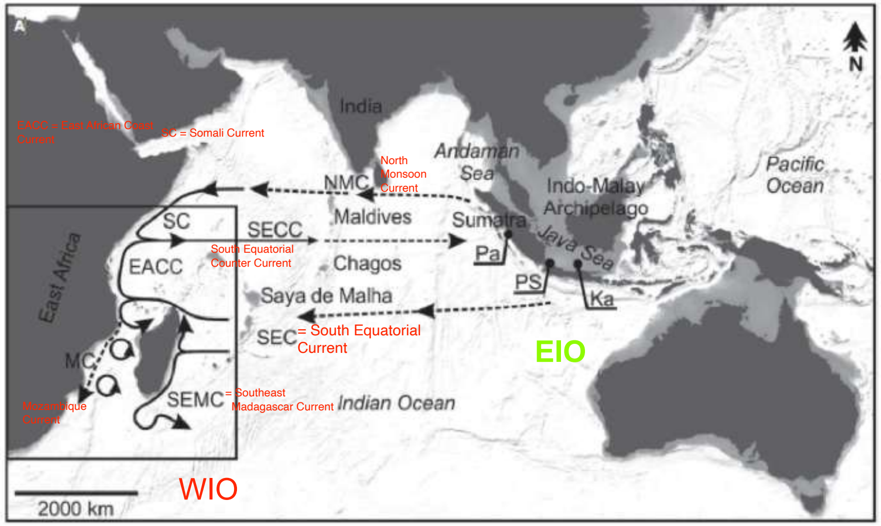
Highly restricted gene flow between disjunct populations of the skunk clownfish (*Amphiprion akallopisos*) in the Indian Ocean

Filip Huyghe & Marc Kochzius

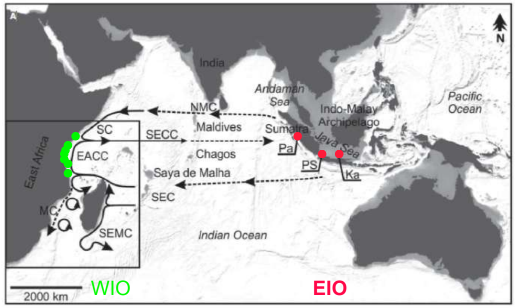
Marine Biology, Vrije Universiteit Brussel (VUB), Brussels, Belgium

---

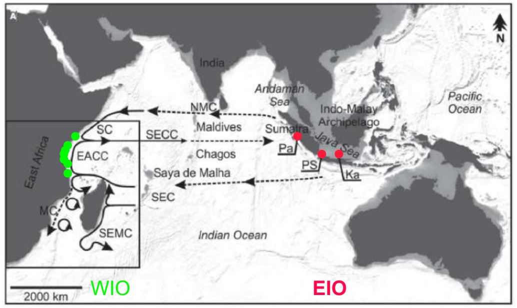
Hypothesis 1

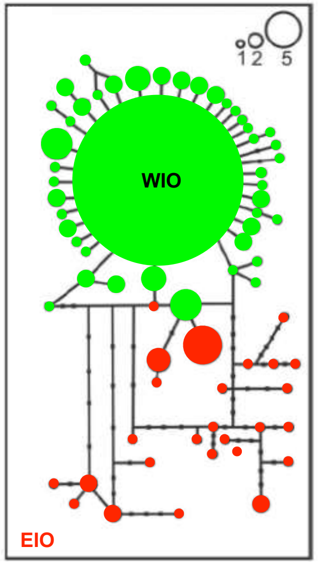


Hypothesis 1

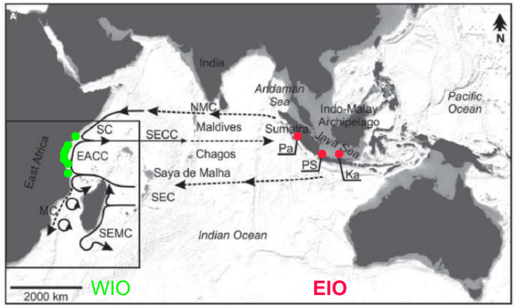


Hypothesis 1

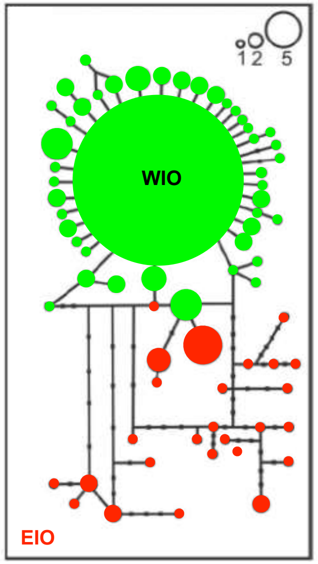




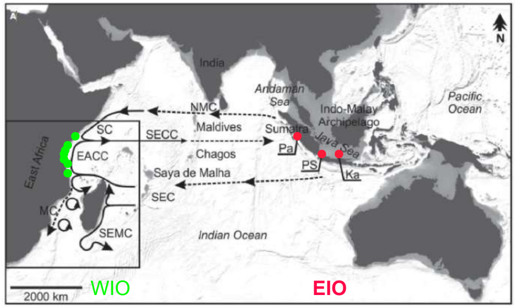
Hypothesis 1



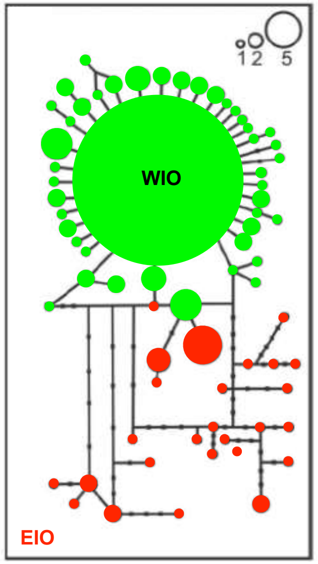
None of the identified 69 haplotypes was shared between the WIO and EIO.



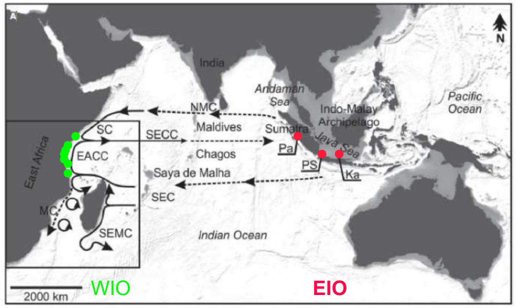
Hypothesis 1

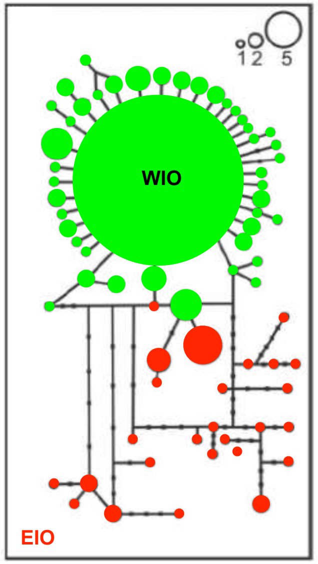


None of the identified 69 haplotypes was shared between the WIO and EIO.
Haplotype as well as nucleotide diversity was considerably higher in the EIO than in the WIO.



Hypothesis 1





None of the identified 69 haplotypes was shared between the WIO and EIO.

Haplotype as well as nucleotide diversity was considerably higher in the EIO than in the WIO.

Support for EIO as the geographical origin of the species

Hypothesis 1

Given
the **large distance**
between the disjunct
populations
the **short pelagic larval**
duration, long-distance
dispersal is rather
unlikely.

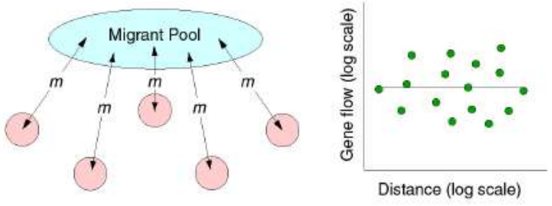
Hypothesis 1

Given
the **large distance**
between the disjunct
populations
the **short pelagic larval**
duration, long-distance
dispersal is rather
unlikely.

Alternatively a **stepping-stone model** involving islands in the Central Indian Ocean is a more likely scenario for colonization of the WIO.

Hypothesis 1

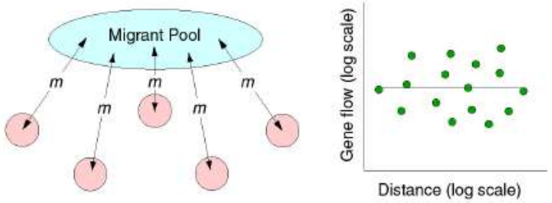
Island Model



Wright, S. 1931. Evolution in Mendelian populations. Genetics 16:97-159.

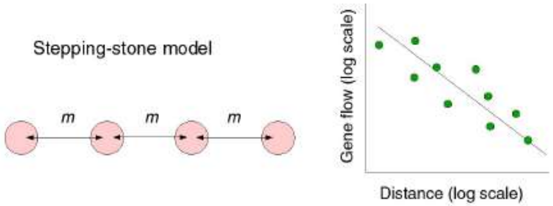
Hypothesis 1

Island Model



Wright, S. 1931. Evolution in Mendelian populations. Genetics 16:97-159..

Stepping-stone model

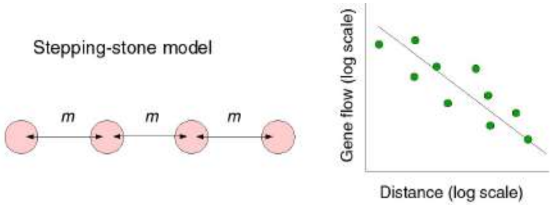


Kimura, M., and W. H. Weiss. 1964. The stepping stone model of genetic structure and the decrease of genetic correlation with distance. Genetics 49:561-576.

Hypothesis 1

Given
the **large distance** between the disjunct populations
the **short pelagic larval** duration, long-distance dispersal is rather unlikely.

Stepping-stone model



Kimura, M., and W. H. Weiss. 1964. The stepping stone model of genetic structure and the decrease of genetic correlation with distance. Genetics 49:561-576.

Alternatively a **stepping-stone model** involving islands in the Central Indian Ocean is a more likely scenario for colonization of the WIO.

PHYLOGEOGRAPHY

Hypothesis 3

Major phylogeographic units within a species reflect
long-term historical barriers
to gene flow

Hypothesis 3

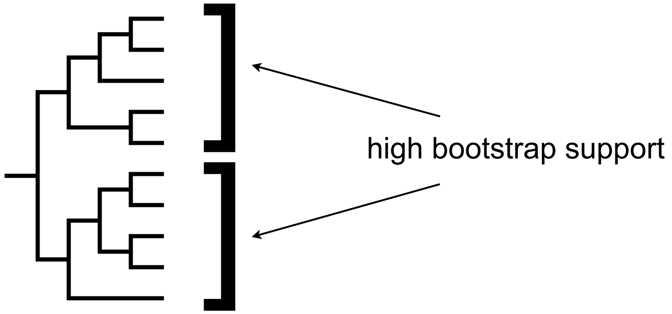
This hypothesis includes **four corollaries** that represent different aspects of genealogical concordance.

Intraspecific monophyletic groups distinguished by large genealogical gaps usually arise from long-term extrinsic (biogeographic) barriers to gene flow.

Major phylogeographic units within a species reflect long-term historical barriers to gene flow.

Hypothesis 3: C1

Species “a”
Gene “1”



Concordance across sequence characters within a gene. (Agreement across characters within a gene)

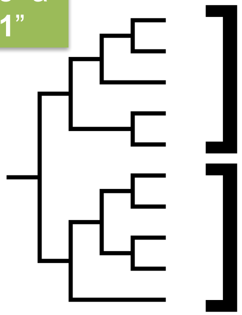
A

Every deep phylogenetic split in the intraspecific gene tree is supported by multiple diagnostic characters within mtDNA. If this is not the case, such matrilineal separations would not be evident in data analysis (not supported by bootstrap).

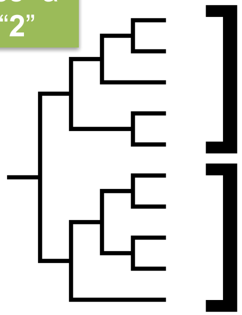
Avise, J.C. (2000) Phylogeography: the history and formation of species. Harvard University Press, Cambridge, MA.

Hypothesis 3: C2

Species “a”
Gene “**1**”



Species “a”
Gene “**2**”



Concordance in significant genealogical partitions across *multiple genes* within a species. (Agreement across genes)

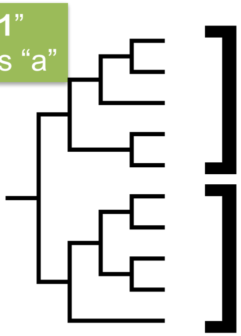
B

Empirical examples show general agreement between deep phylogeographic topologies in multiple gene trees within the species. These deep branch separations characterize the same sets of geographic populations.

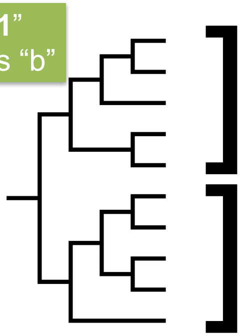
Avise, J.C. (2000) Phylogeography: the history and formation of species. Harvard University Press, Cambridge, MA.

Hypothesis 3: C3

Gene “**1**”
Species “a”



Gene “**1**”
Species “b”



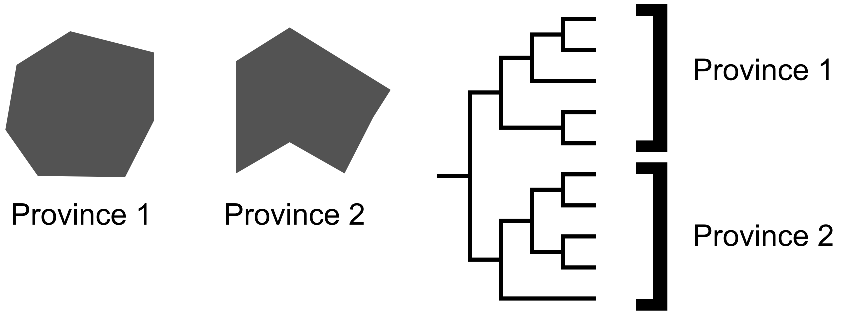
Concordance in the geography of gene-tree partitions across multiple co-distributed species. (Agreement across co-distributed species)

C

Several co-distributed species with comparable natural histories or habitat requirements proved to be phylogeographicaly structured in similar fashion. In particular, divergent branches in the intraspecific gene trees might map consistently to the same geographic regions.

Avise, J.C. (2000) Phylogeography: the history and formation of species. Harvard University Press, Cambridge, MA.

Hypothesis 3: C4



D

Concordance of gene-tree partitions with spatial boundaries between traditionally recognized biogeographic provinces. (Agreement with other biogeographic data)
An emerging generality from molecular phylogeography studies is that deeply separated phylogroups at the intraspecific level often are confined to biogeographic provinces or subprovinces as identified from traditional faunal lists.

Avise, J.C. (2000) Phylogeography: the history and formation of species. Harvard University Press, Cambridge, MA.

Conceptual approaches

Classical population genetics
based on allele frequencies
alleles are either the same or different, no degrees of similarity

---

Conceptual approaches

Classical population genetics
based on allele frequencies
alleles are either the same or different, no degrees of similarity

SUMMARY STATISTICS
(descriptive, exploratory statistics)

How many alleles are there per locus per population?
What are allelic frequencies?
What is the proportion of heterozygous individuals?
What would be the proportion of heterozygous individuals in HW equilibrium?
What are the distances between populations?
Are distinct loci in linkage equilibrium?
How many positions vary in a nucleotide alignment?
How many different haplotypes are there?
How different are haplotypes?
What is the average difference between two sequences in a population?
Is most of the variation within or between populations?

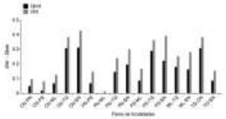
Conceptual approaches

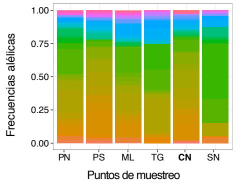
Classical population genetics
based on allele frequencies
alleles are either the same or different, no degrees of similarity

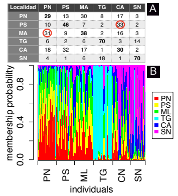
Conceptual approaches

Classical population genetics
based on allele frequencies
alleles are either the same or different, no degrees of similarity

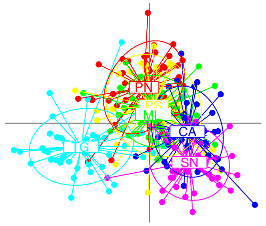
Localidad	Código	Longitud	Latitud	N	AT	Nap	Ar (35)	H_o	H_e	Gis
Portugal - Norte	PN	-8.8	40.7	45	136	15	15.4	0.77	0.86	0.10
Portugal - Sur	PS	-8.4	37.1	43	133	11	15.6	0.83	0.88	0.06
Malaga	ML	-4.3	36.6	46	131	14	15.1	0.76	0.88	0.14
Tánger	TG	-5.9	35.9	46	112	11	12.6	0.69	0.80	0.14
Islas Canarias	CA	-15.0	28.3	42	142	22	16.9	0.79	0.89	0.11
Senegal	SN	-17.6	14.8	37	135	19	16.7	0.77	0.88	0.12







Localidad	PN	PS	MA	TG	CA	SN
PN	**29**	13	30	8	17	3
PS	10	**46**	7	2	33	2
MA	31	9	**38**	2	16	3
TG	6	2	6	**70**	3	14
CA	18	32	17	1	**30**	2
SN	4	1	6	18	1	**70**

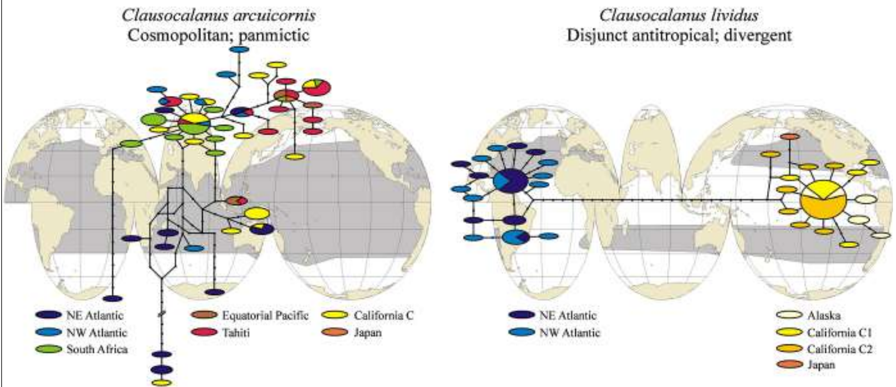


Conceptual approaches

Phylogeography
based on gene genealogies
both the frequencies of sequence haplotypes and their relationship is considered

Conceptual approaches

Phylogeography
based on gene genealogies
both the frequencies of sequence haplotypes and their relationship is considered



Blanco-Bercial et al. 2011

History of Biogeography

outline

WHAT IS PHYLOGEOGRAPHY

MARKERS

MODELS

AVISE'S HYPOTHESIS and COROLLARIES