



Hybridisation

Joana Meier

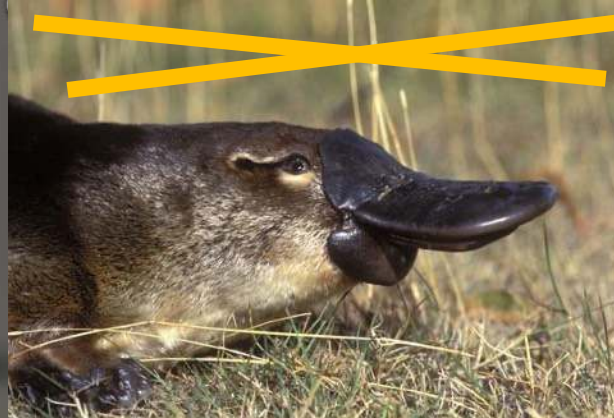
Tree of Life Programme, Wellcome Sanger Institute

THE
ROYAL
SOCIETY

The
Branco Weiss
Fellowship
Society in Science



UNIVERSITY OF
CAMBRIDGE

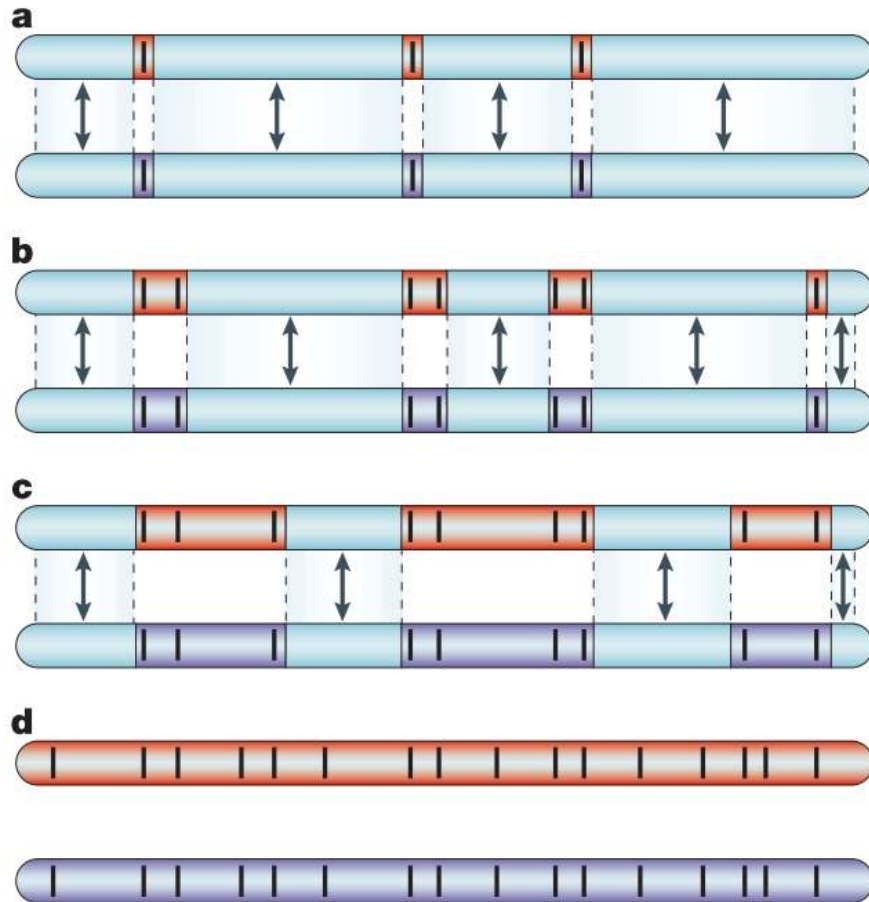


Outline

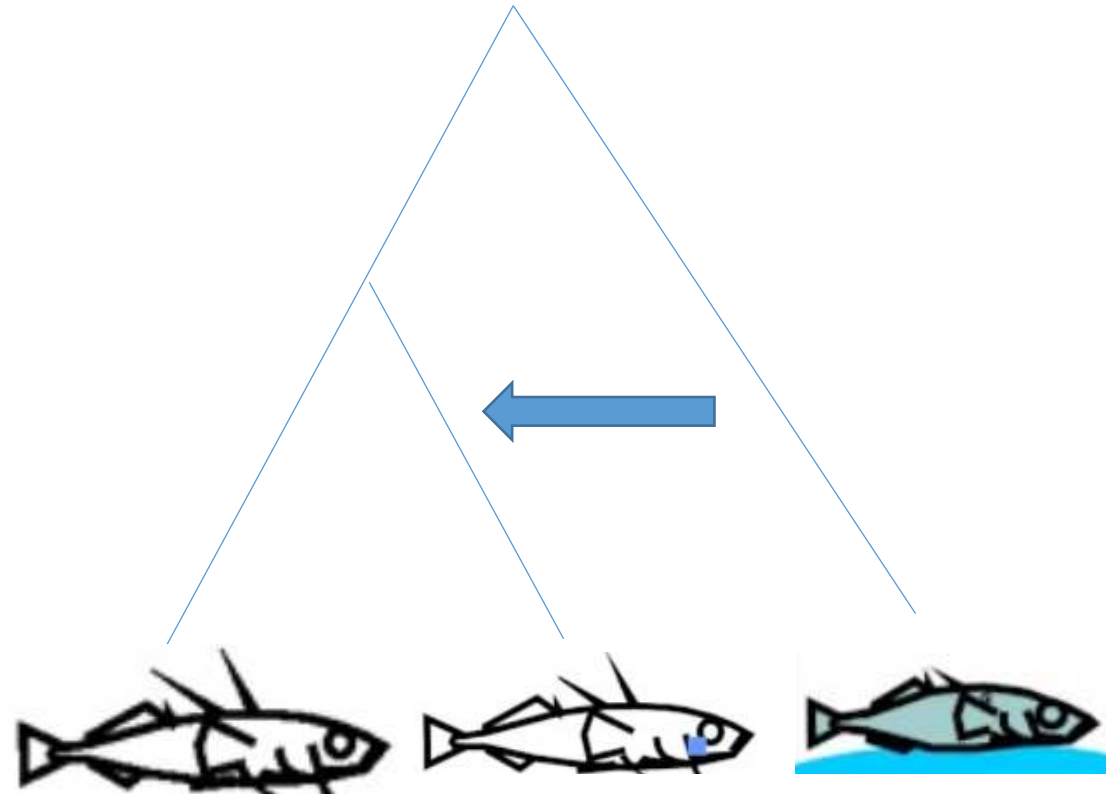
- What is hybridisation and introgression?
 - Defining terms
 - How common is it?
- Evolutionary consequences of hybridisation
 - From genome swamping to hybrid speciation
- Detecting hybridisation
 - Time scale matters (ongoing versus ancestral)
- An example of my work on hybridisation

The dual role of hybridisation in speciation and adaptation

Gene flow between sister taxa
counteracts speciation

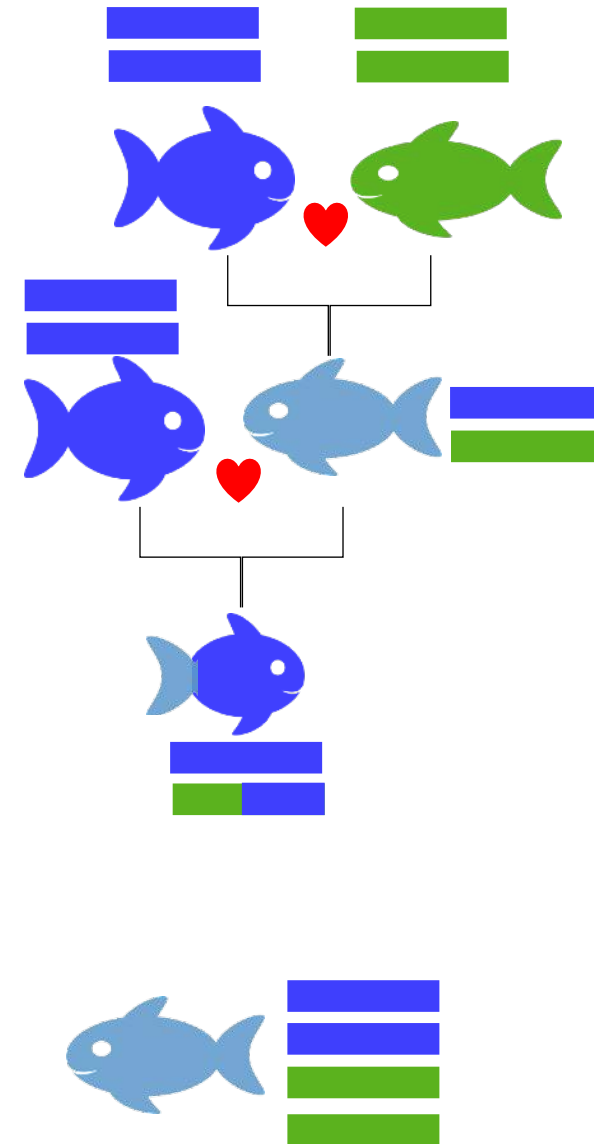


Hybridisation can introduce
new genetic variation



Key terminology

- **Hybridisation:** sexual reproduction between individuals from different species or genetically divergent populations
- **Back-crossing:** hybrid individuals mating with «pure» individuals and thus leading to gene transfer between species
- **Admixture:** the genetic pattern of mixed ancestry in hybrid individuals
- **Introgression:** movement of gene variants across lineage boundaries through hybridisation and backcrossing
- **Gene flow:** gene exchange between closely related populations
- **Allopolyploid** : organism with multiple sets of chromosomes from different parental lineages due to hybridisation



History

Historically, zoologists considered hybridisation between species as a rare process or as a sign of incomplete speciation (Dobzhansky 1937; Mayr 1942). Botanists recognized that hybridisation may sometimes generate novel traits and even species (Lotsy, 1916) (Anderson 1949; Anderson and Stebbins 1954).

“The rarity of allopolyploids constitutes the greatest known difference between the evolutionary patterns in the two kingdoms” (Dobzhansky, 1937, p. 219)

„the majority (if not all) of organisms have an evolutionary history (...) that includes genetic exchange events“ (Arnold, 2009, p. 183).

The majority of arctic plants are allopolyploid hybrids



Hybridisation is widespread but not common

Table 1. Hybridization in the wild, including rates of hybridization per species^a

Taxon	No. species sampled	No. species hybridizing with at least one other	% species hybridizing with at least one other	Approx. rate of hybridization per individual
UK vascular plants (Pteridophyta, Angiospermae)	539 (sampled from 2290 species known in the UK)	135	25.0%	Variable, but mostly <1%
Flora of Concord, Massachusetts (Angiospermae)	838	>27	>3.2%	Unknown
European butterflies (Lepidoptera: Rhopalocera)	379	47	12.4%	Mostly <0.1%
Swallowtail butterflies (Lepidoptera: Papilionidae; <i>Papilio</i>)	216	14–32	6.5–14.8%	Some hybrids common in restricted areas
Passion flower butterflies (Lepidoptera: Heliconiina)	73	19	26.0% (34.8% in genus <i>Heliconius</i>)	Mostly <0.1%

Widespread

Not common

Hybridisation is widespread but not common

Table 1. Hybridization in the wild, including rates of hybridization per species^a

Taxon	No. species sampled	No. species hybridizing with at least one other	% species hybridizing with at least one other	Approx. rate of hybridization per individual	Notes and references
Native grouse of Britain (Tetraonidae)	4	4	100.0%	Mostly <0.1%	[83]



All grouse species of Britain hybridise with other species, but only <0.1% individuals are visible hybrids

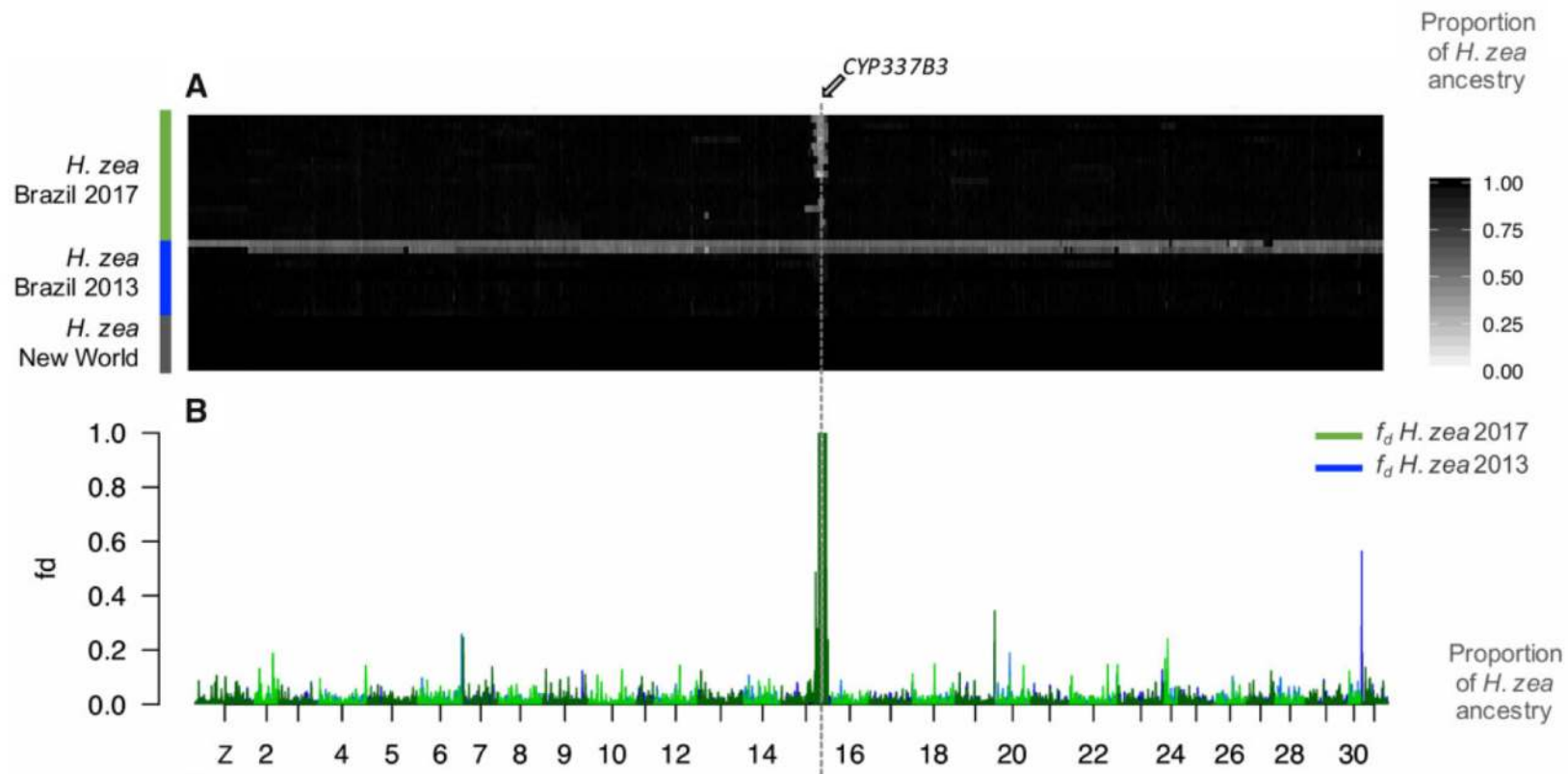
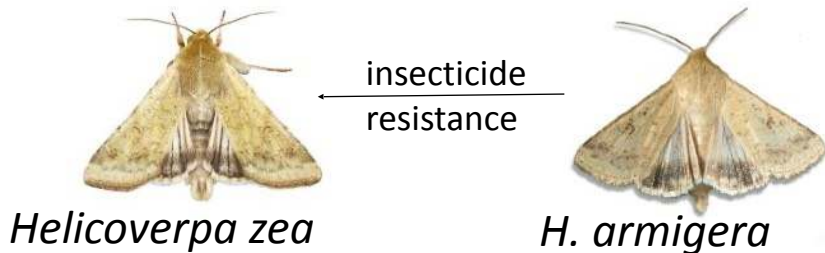


Hybridisation does not necessarily lead to introgression



Even if hybrids are very rare, introgression can be highly relevant, e.g. if beneficial allele introgresses

Insecticide resistance



High altitude adaptation in Tibetans



Admixture between Denisovans and humans in Asia 50k-30k years ago led to 1-6% Denisovan ancestry in southeast Asians.

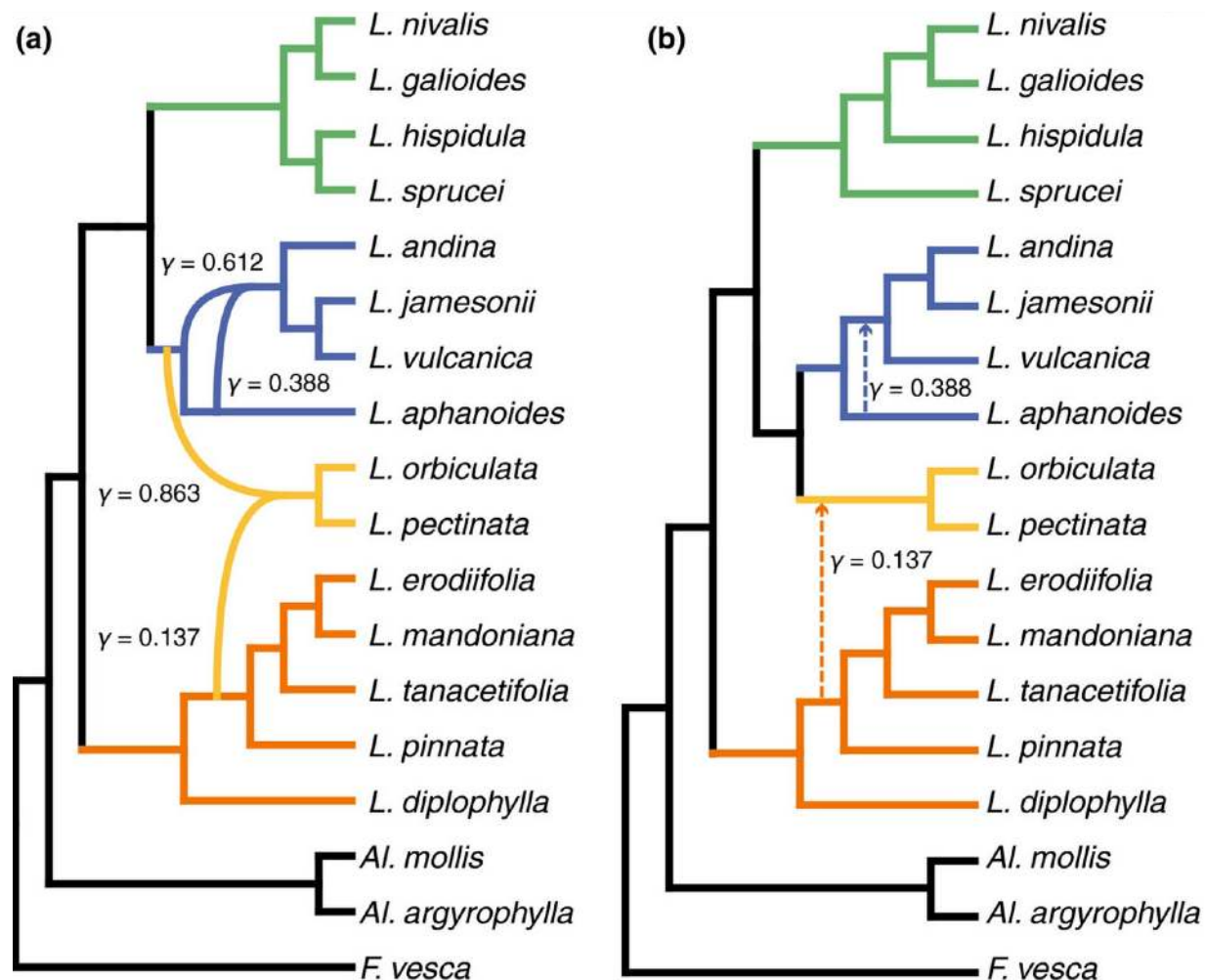
EPAS1 introgression from Denisovans increases the efficiency in oxygen transport in Tibetans, without straining the heart

Because of introgression, the tree of life is actually a network of life

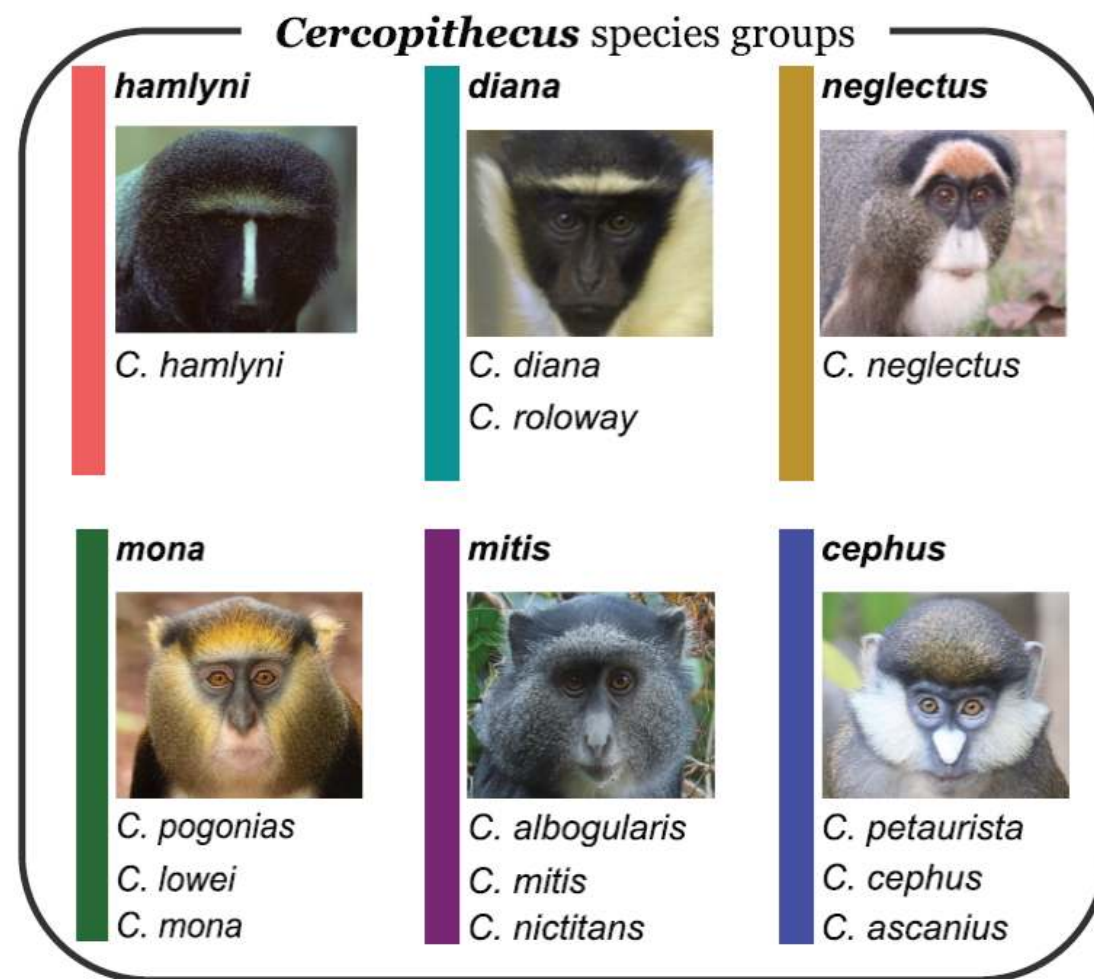
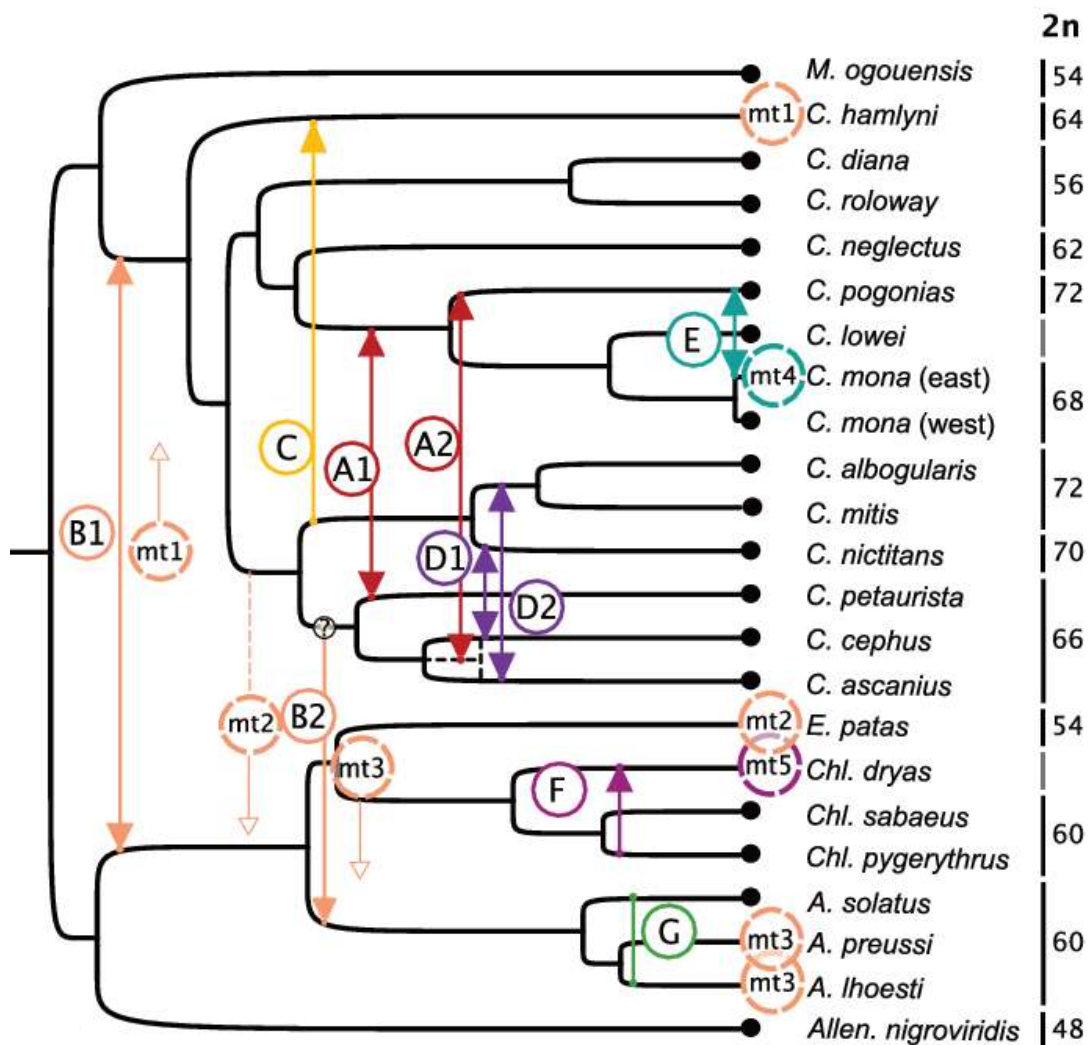
«With whole-genome sequencing we can study gene flow in the past and assess levels of reticulation but the quest for the 'correct tree' is flawed»

Phylogenies are inaccurate

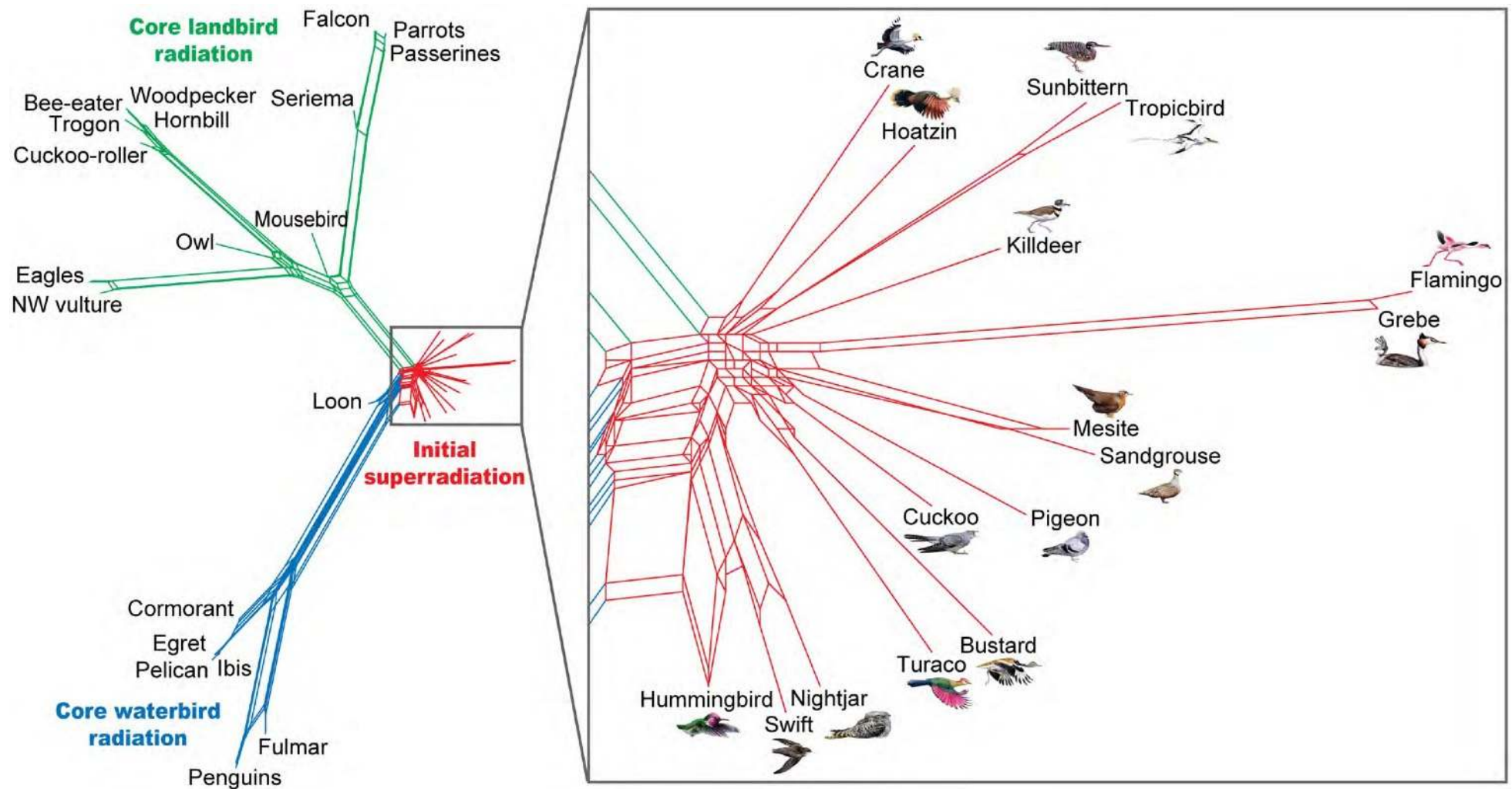
Lachemilla (Rosaceae)



The same is true for animals (e.g. guenons)

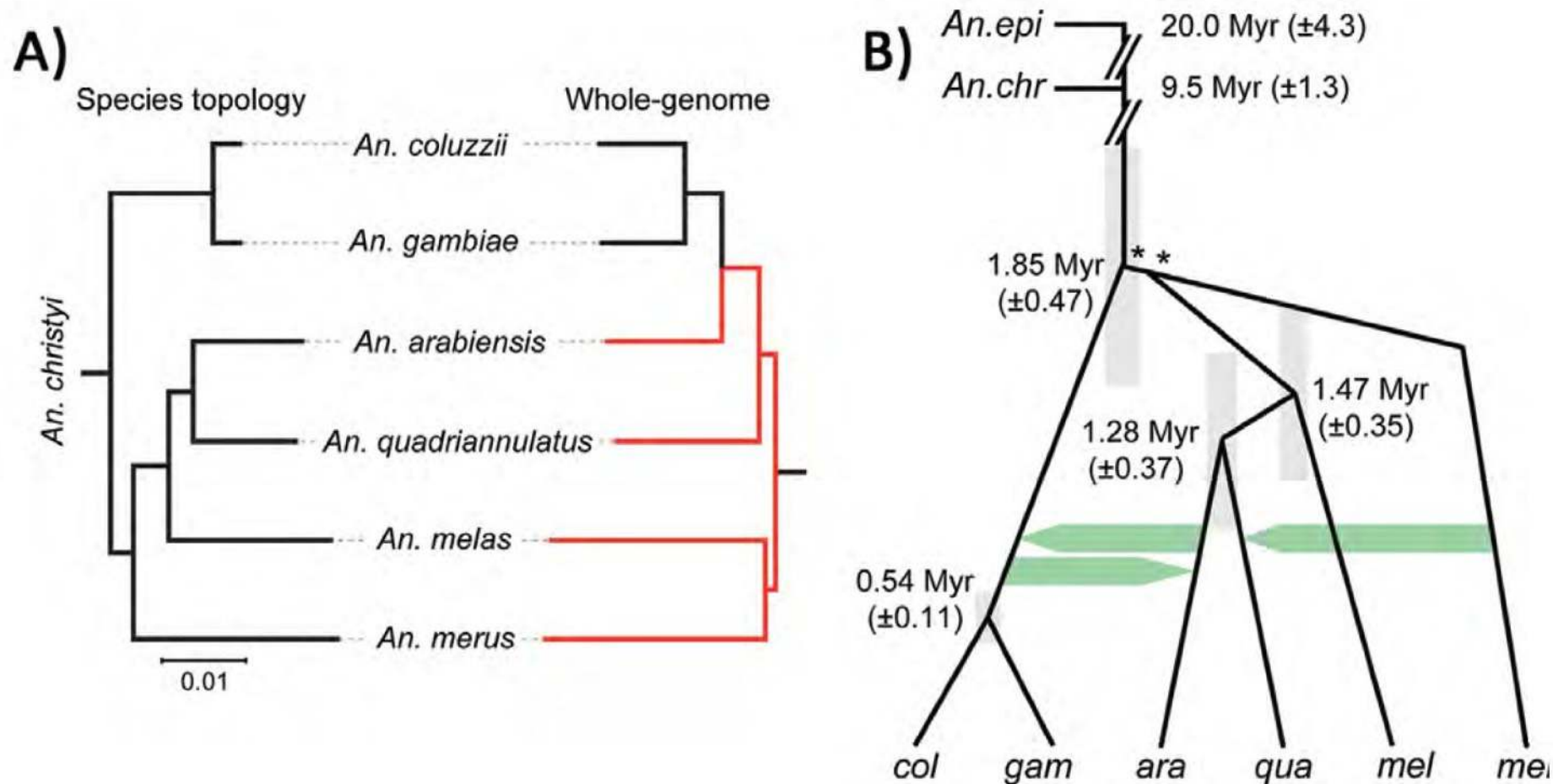


Phylogenies as networks

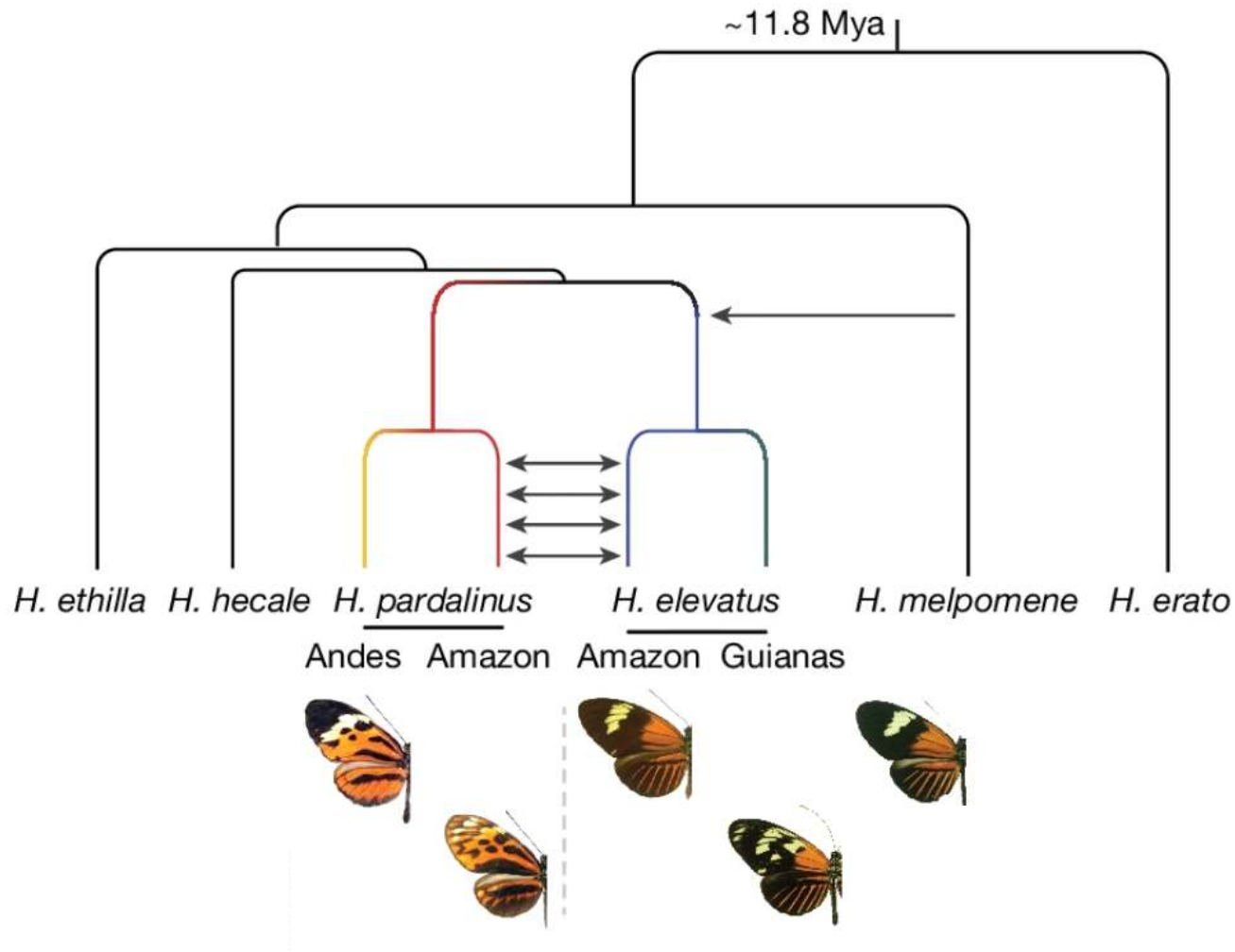


The bird radiation

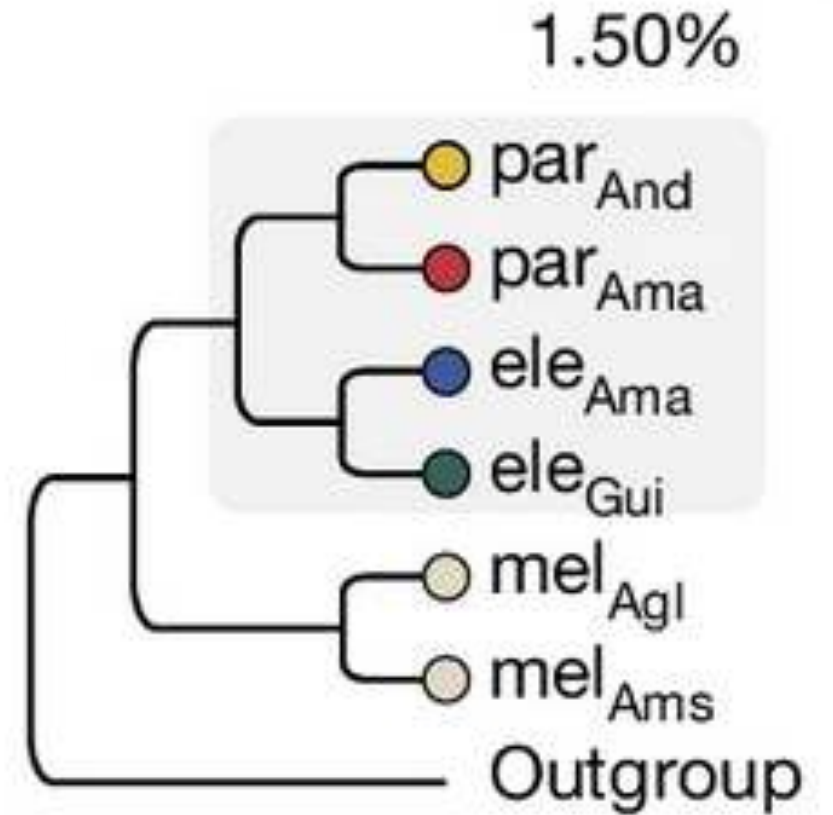
Introgression can affect large amounts of the genome: In *Anopheles* most parts of the genome do not represent the species tree



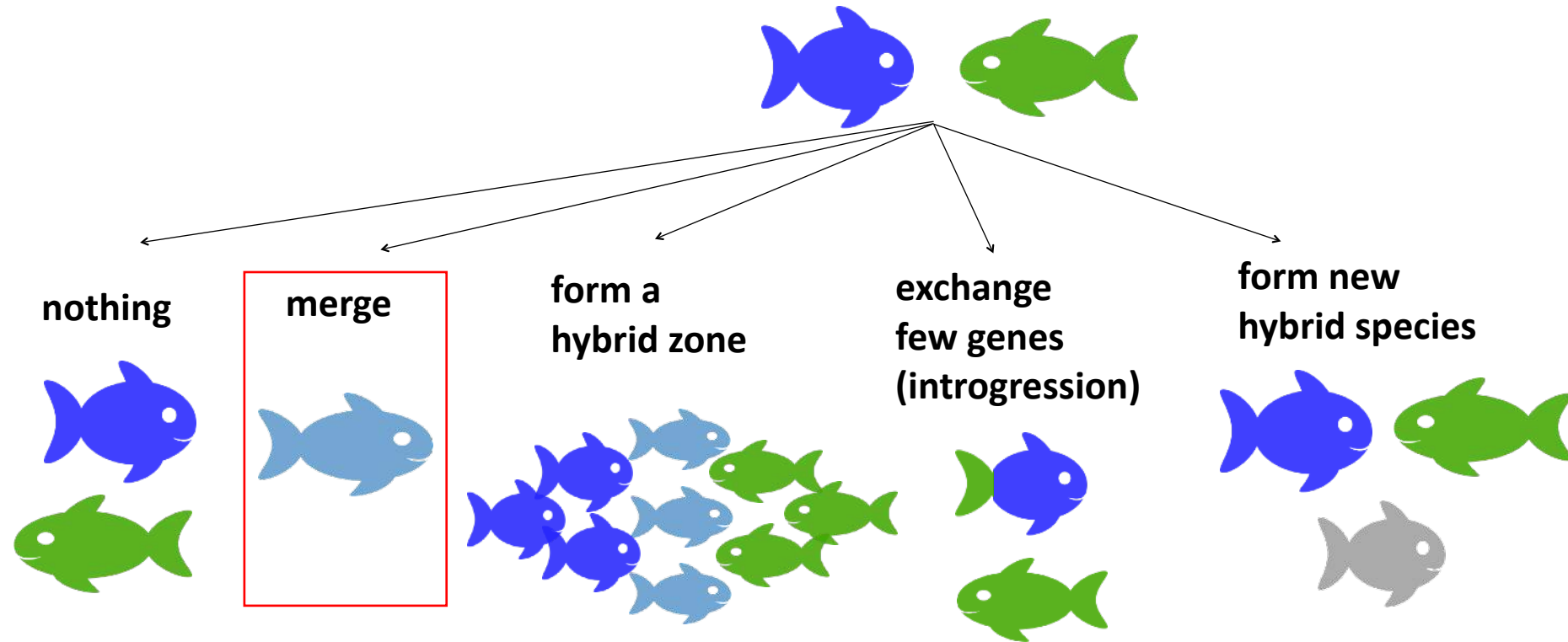
Heliconius elevatus and *H. pardalinus*



Only 1.5% of the phylogenies of different genomic windows show the species topology



Evolutionary consequences of hybridisation



Local adaptation can be lost and species can fuse

Introduced mallards threaten local duck species through hybridisation



THE GLOBAL FERAL MALLARD THREAT



Photo © H. Stiver



Photo © H. Stiver



Photo © H. Stiver



Photo by B. Zaun



Photo by C.D. Roderick



U.S. Fish and Wildlife Service photo

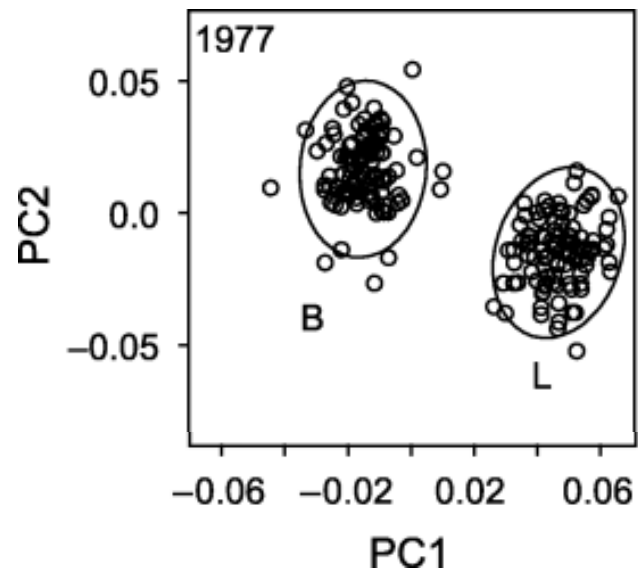


Photo © M. Epstein

Worldwide, native ducks are vulnerable to invasion and hybridization with feral Mallards, particularly when already in jeopardy from habitat loss and introduced mammalian predators. The Mallard occurs naturally in the Northern Hemisphere, but because of its popularity, people have transported it to areas where it was not previously found. Mallard introductions have caused conflicts with native duck populations throughout the world. In New Zealand, descendants of game-farm Mallards, estimated to number 4 to 5 million, are responsible for the near extinction of the New Zealand Grey Duck. In South Africa, Mallards introduced by farmers and waterfowl collectors now threaten both the African Black Duck and Yellow-billed Duck.

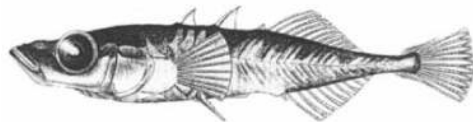
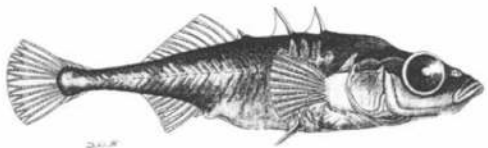


Stickleback despeciation / speciation reversal



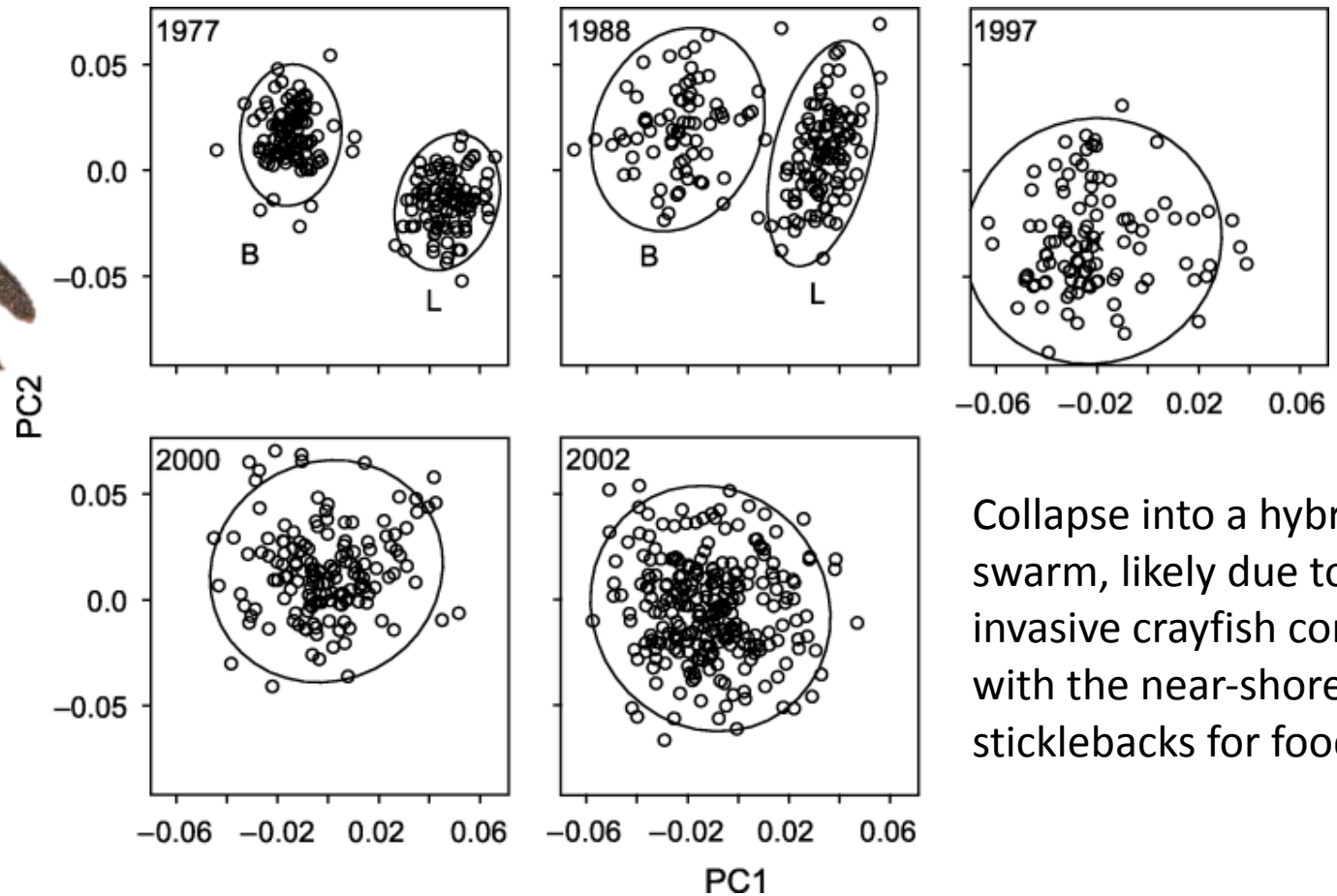
benthic

limnetic



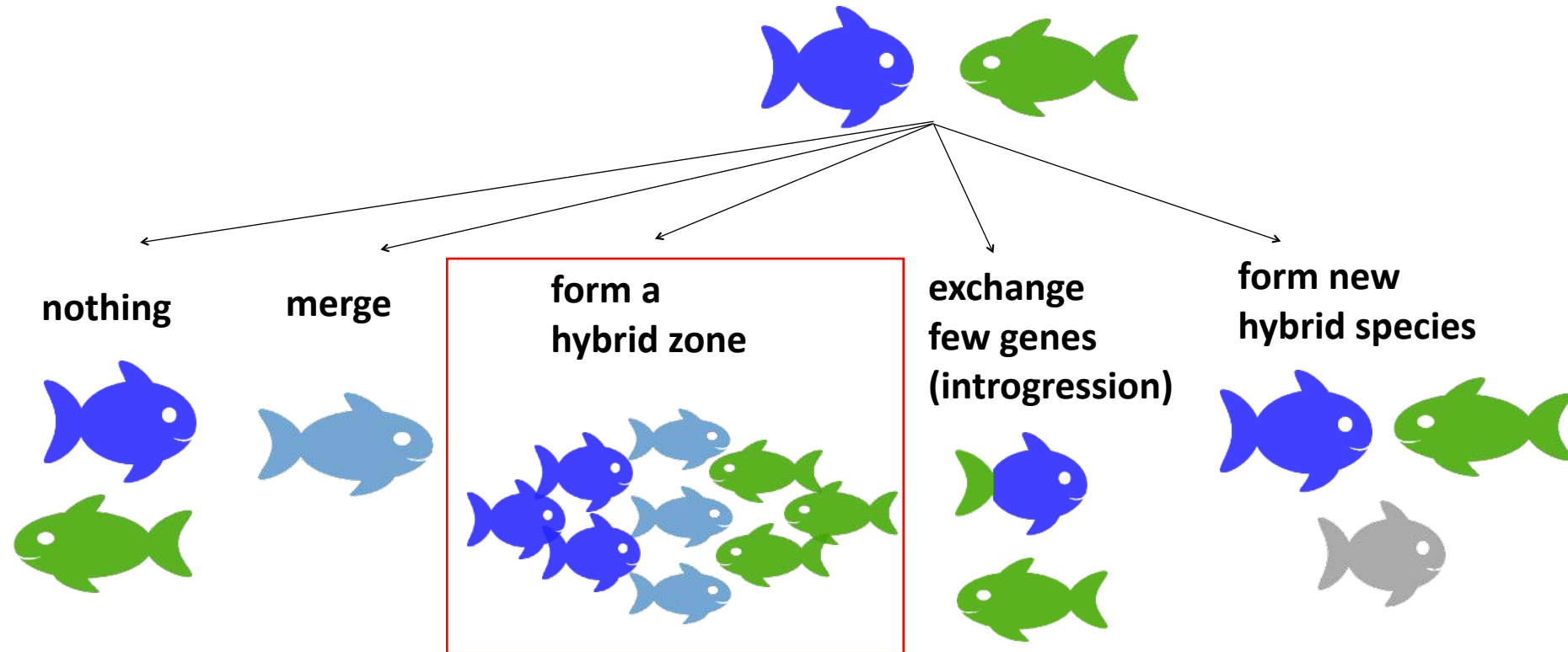
Stickleback despeciation / speciation reversal

Morphological variation got more and more lost

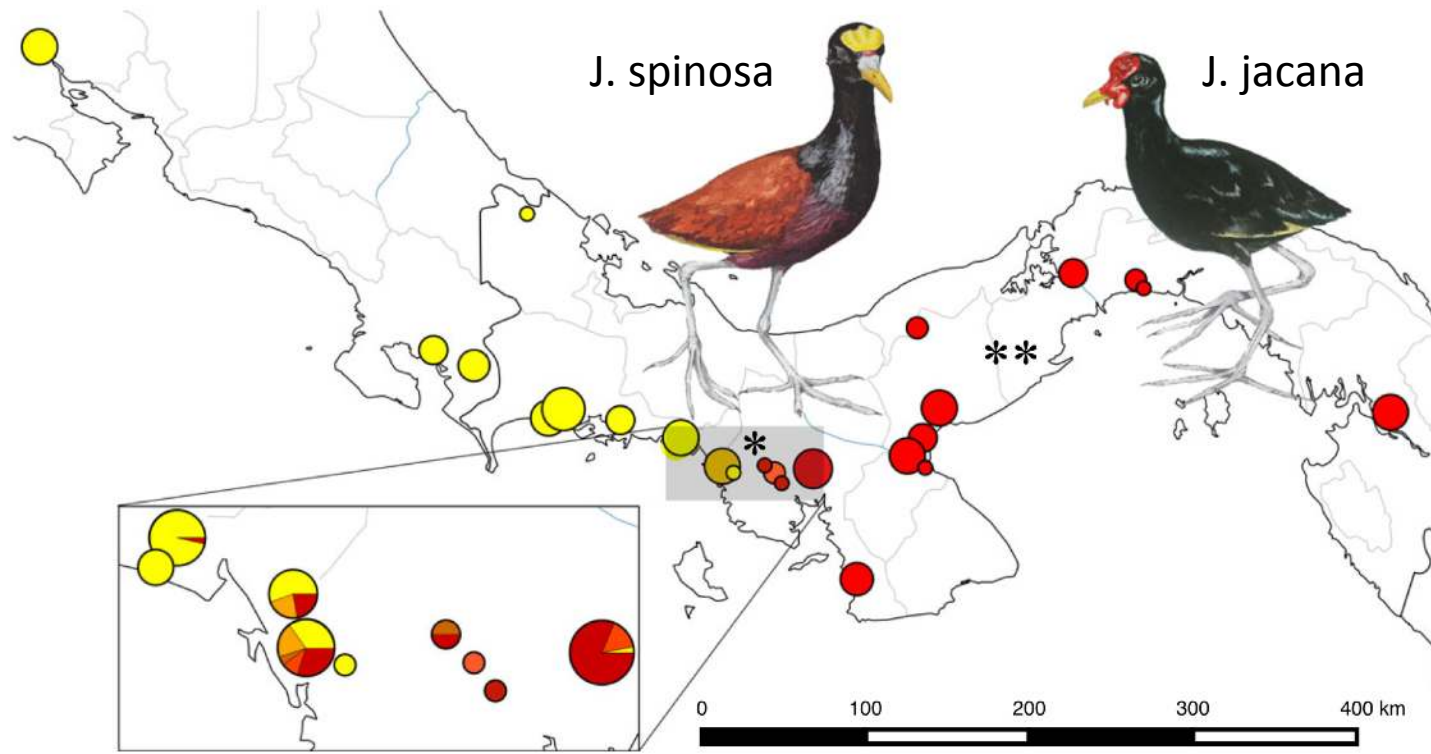


Collapse into a hybrid swarm, likely due to the invasive crayfish competing with the near-shore sticklebacks for food.

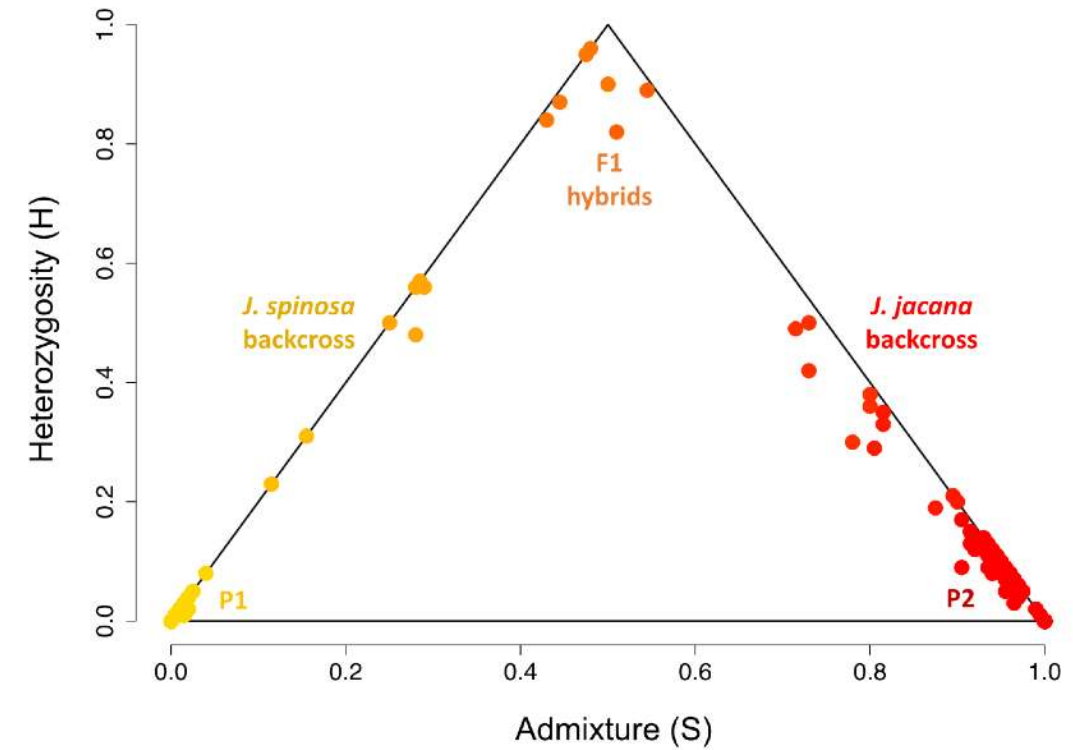
Evolutionary consequences of hybridisation



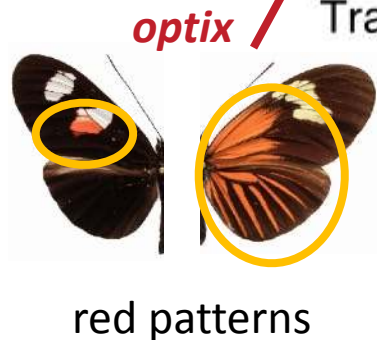
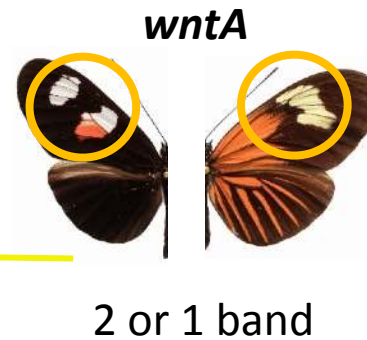
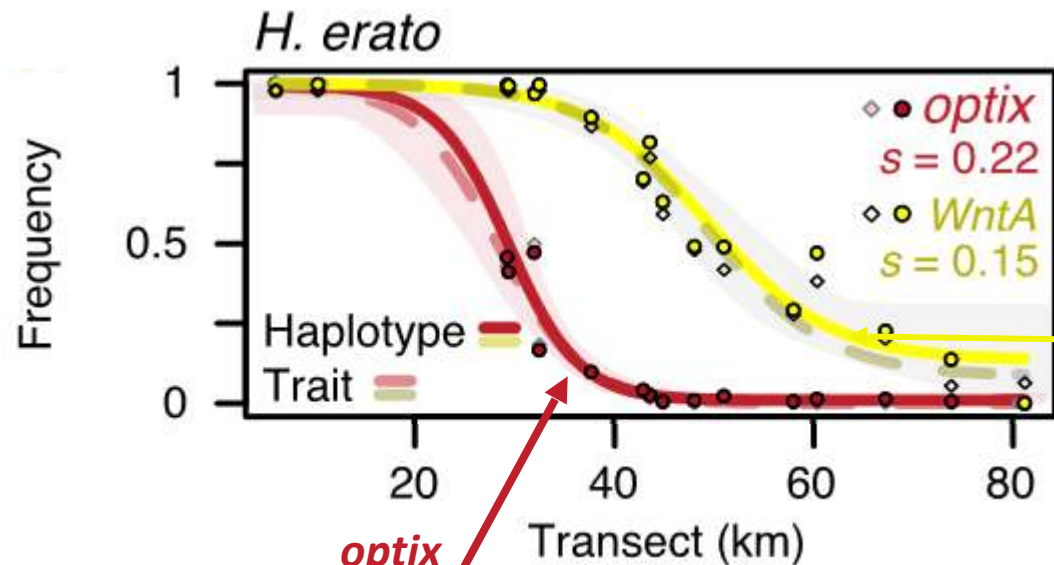
Jacana hybrid zone in Panama



Triangle plot shows
assymmetric introgression



Hybrid zones allow us to test for barriers to reproductive isolation and estimate selection



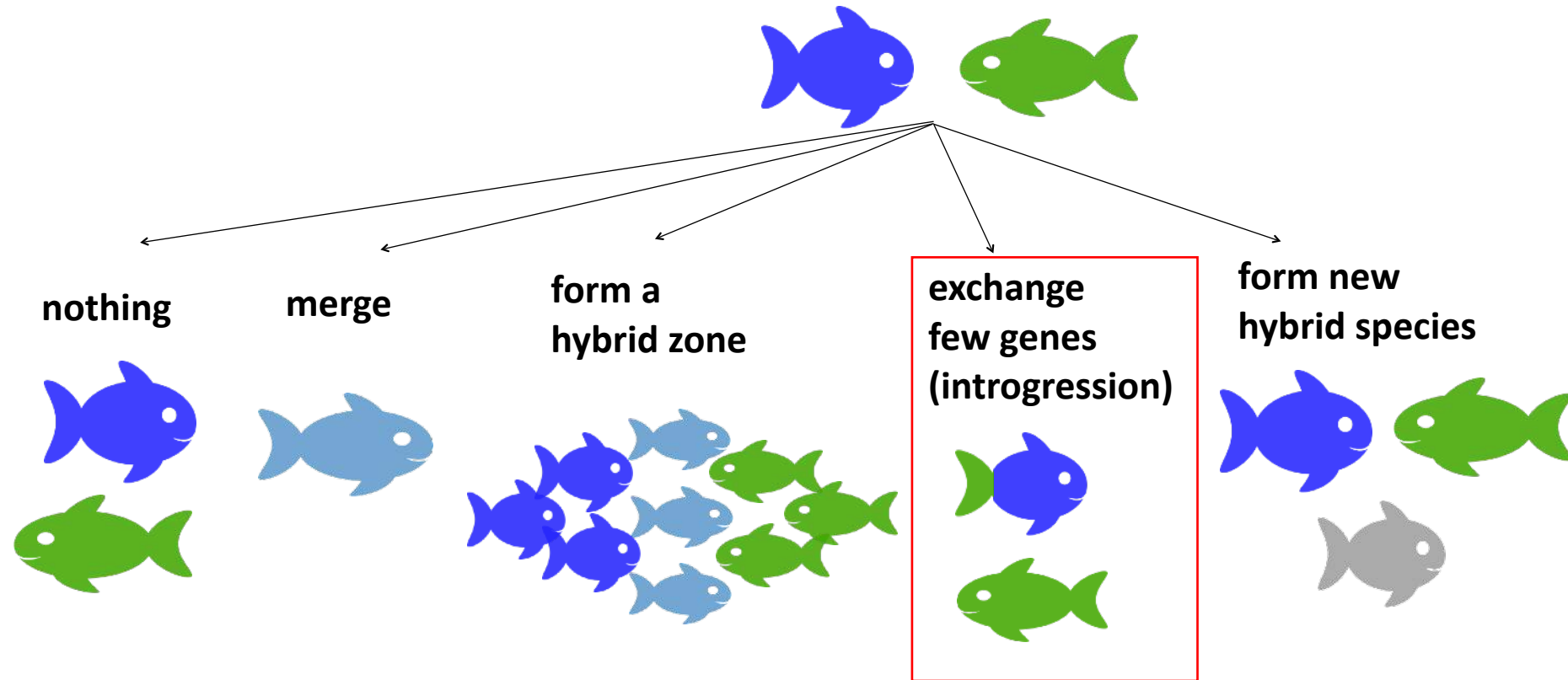
$$\sigma \sim w \sqrt{s/12}$$

Dispersal rate

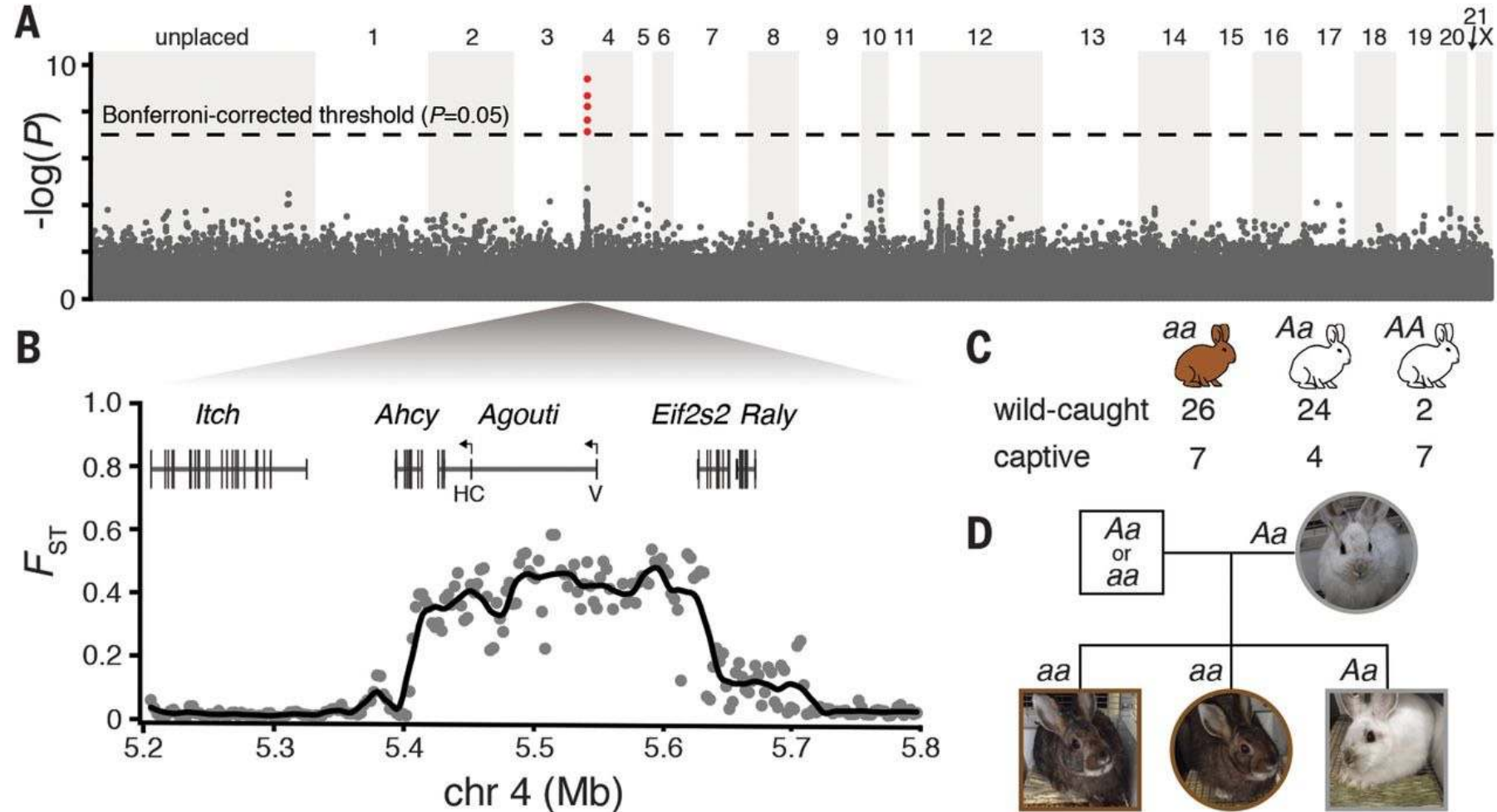
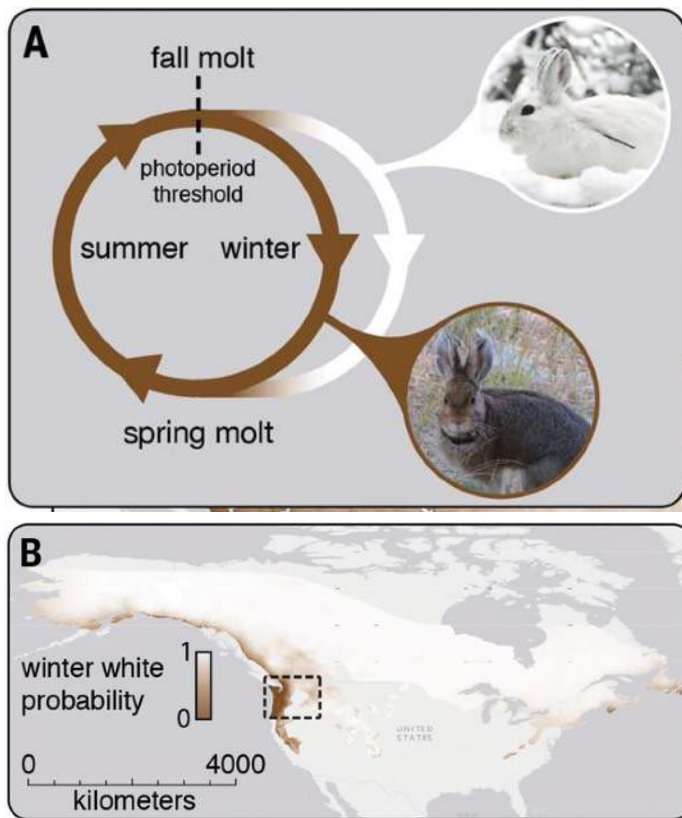
Cline width

Selection coefficient

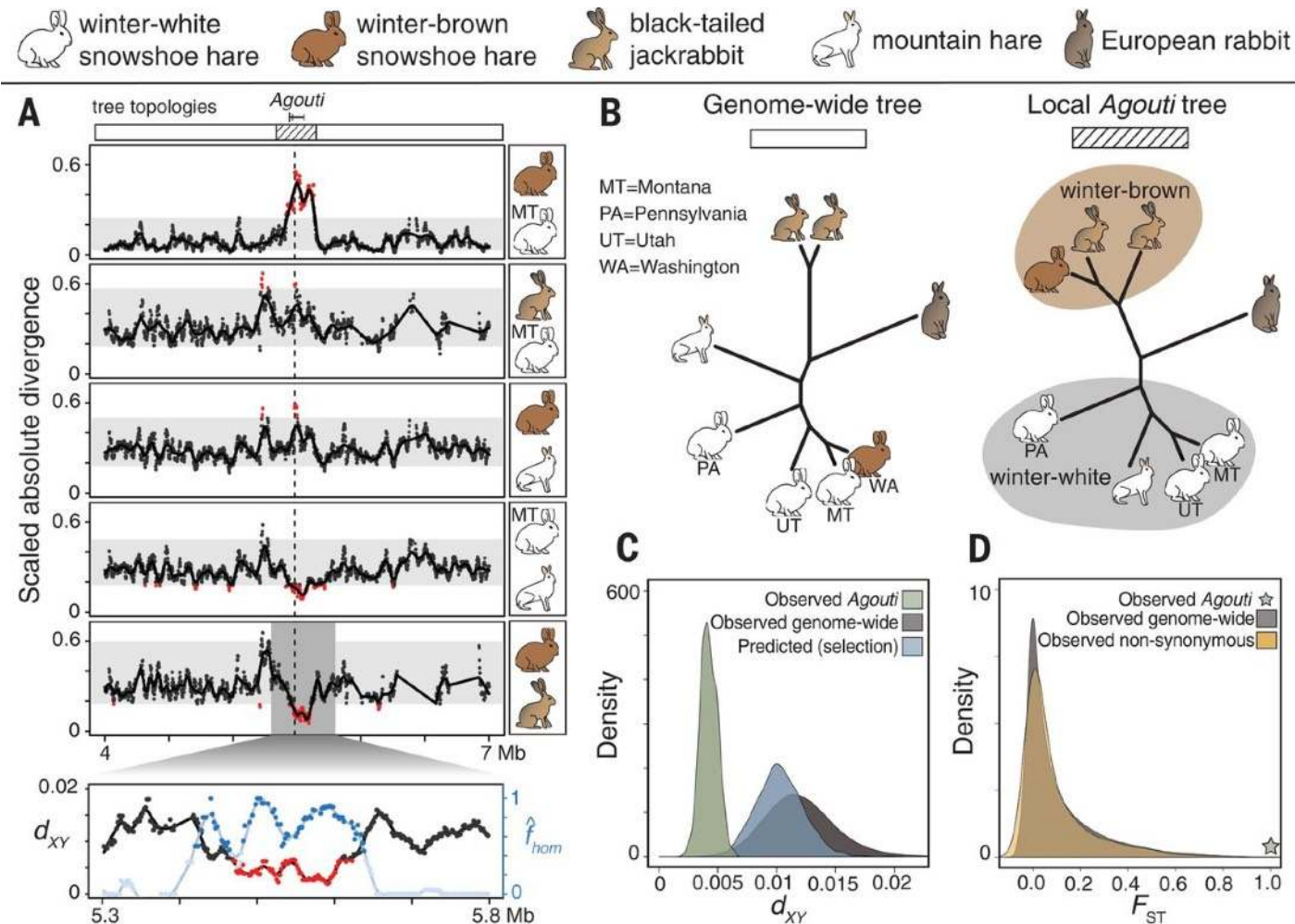
Evolutionary consequences of hybridisation



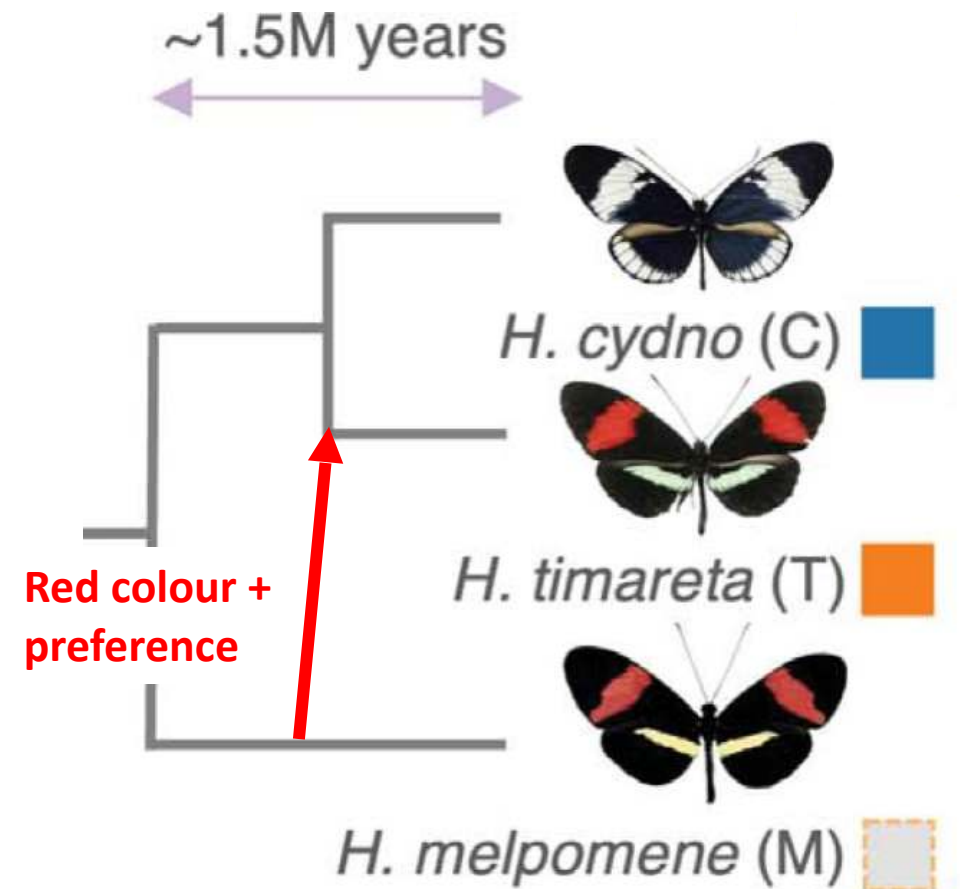
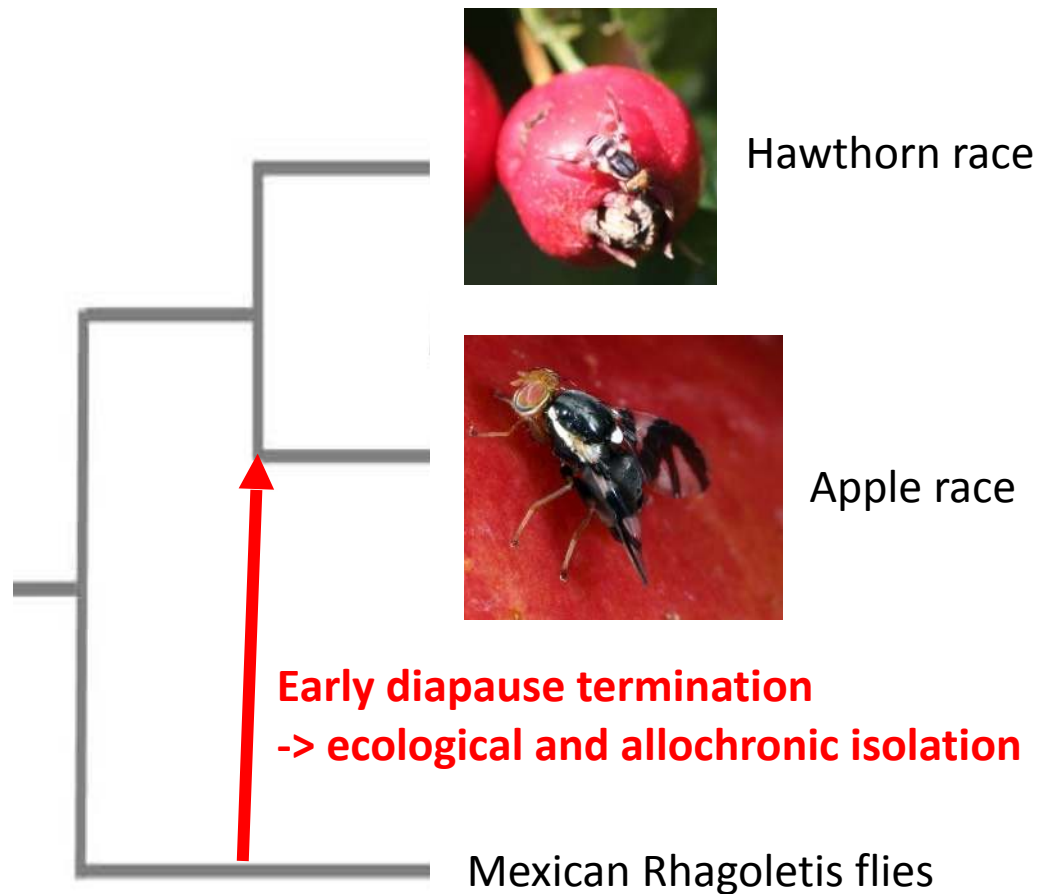
Example of adaptation to climate change: Adaptive introgression in snowshoe hares



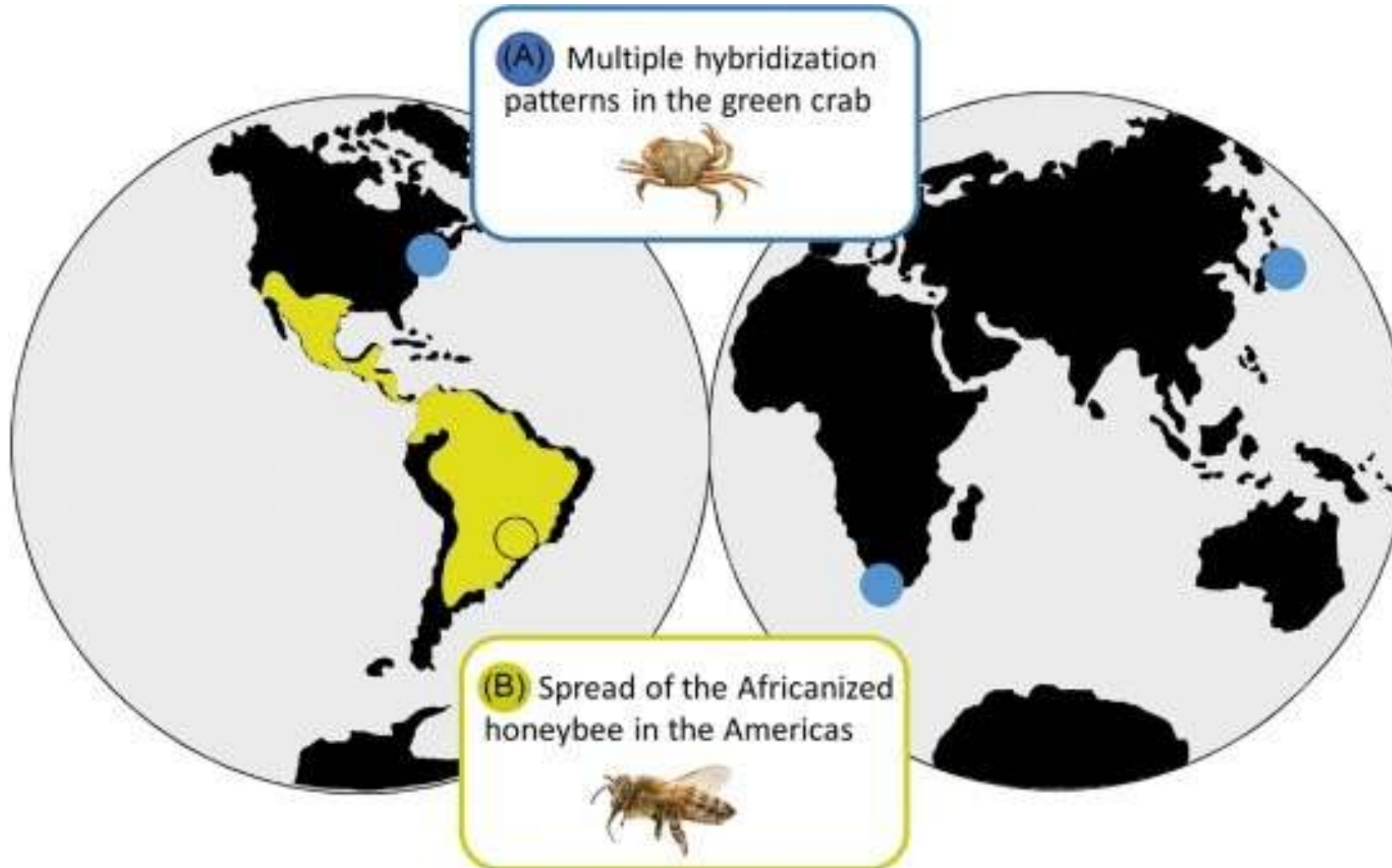
Example of adaptation to climate change: Adaptive introgression in snowshoe hares



Introgression of key traits contributes to speciation



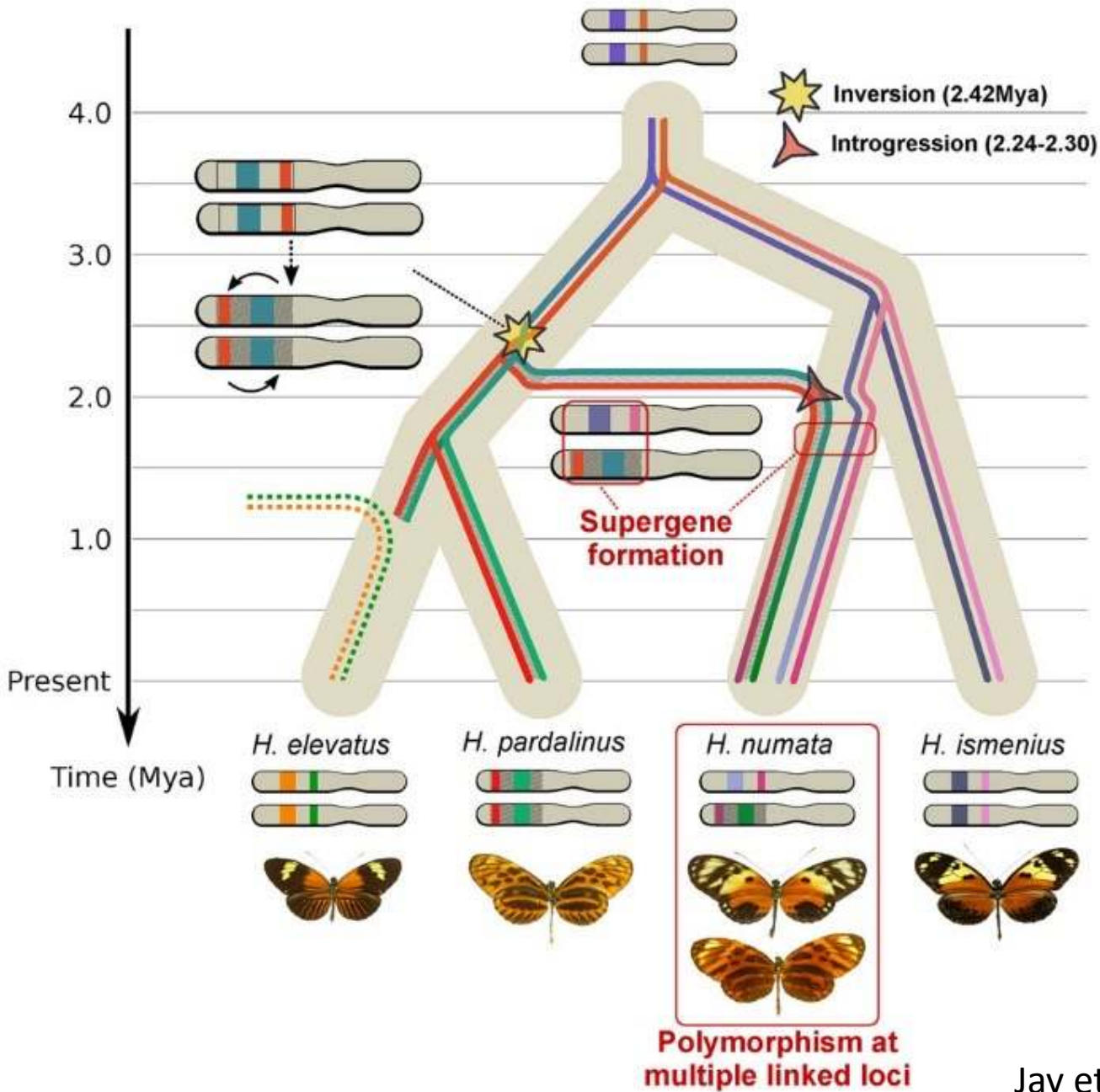
Hybridisation can facilitate invasiveness



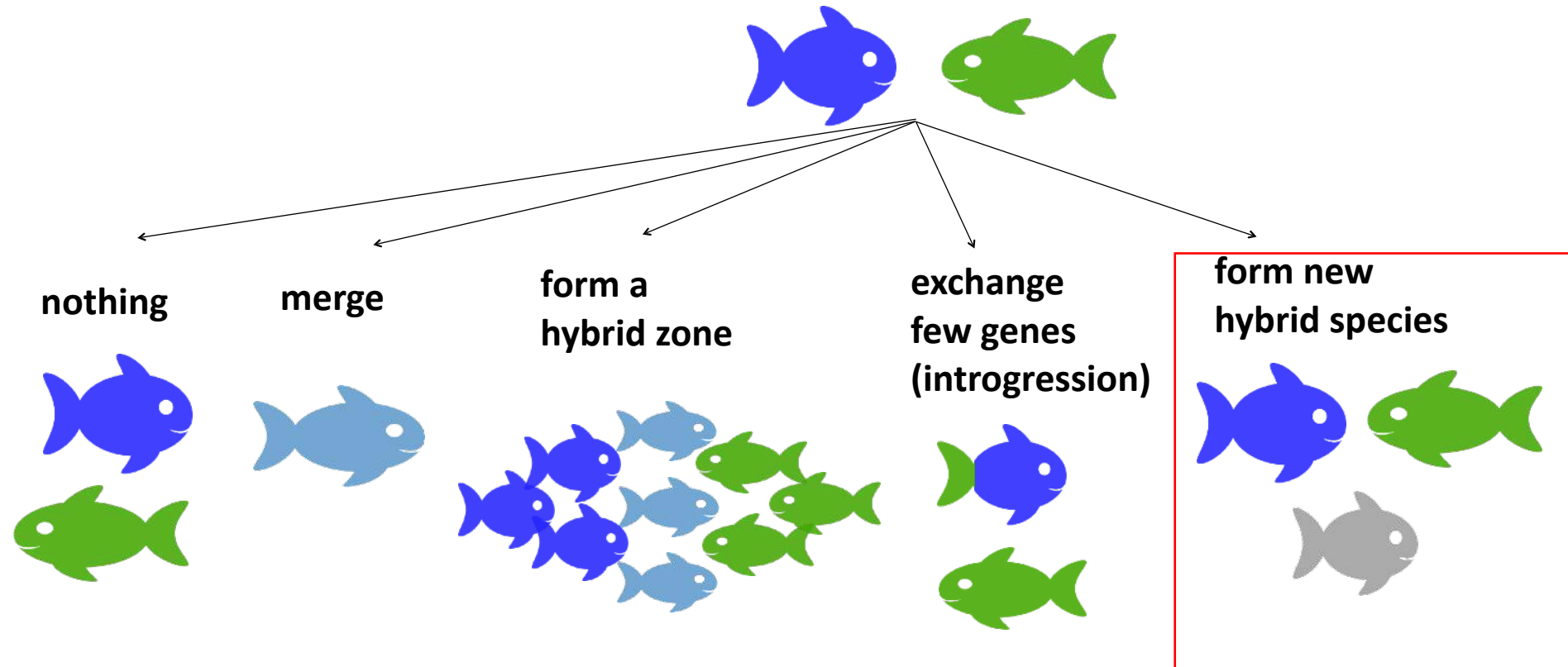
Supergene formation through introgression



Melinaea species *Heliconius numata*



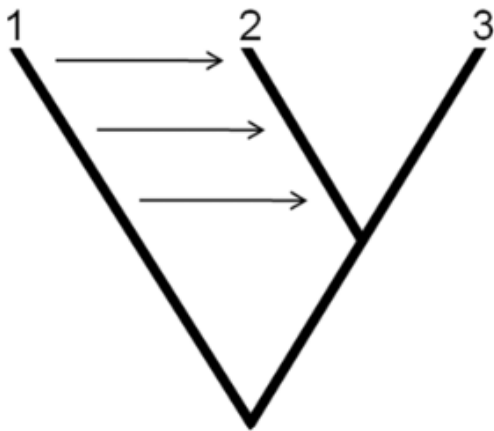
Evolutionary consequences of hybridisation



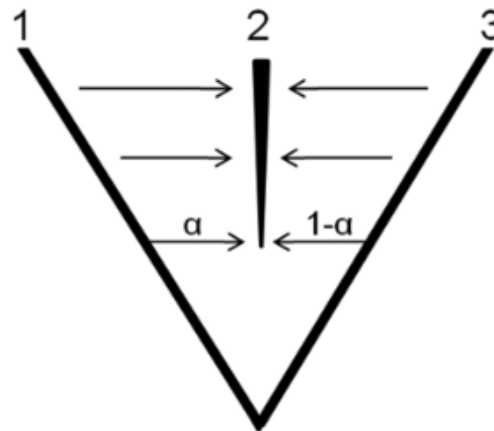
Homoploid hybrid speciation

- Hybridisation must precede speciation
- Hybrid species must be reproductively isolated from both parental lineages
- Hybridisation must be causal for reproductive isolation

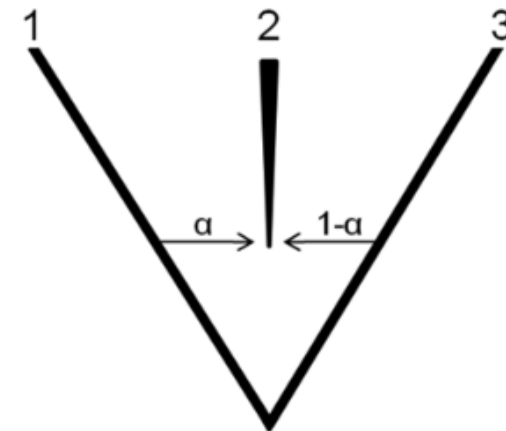
A Secondary gene flow



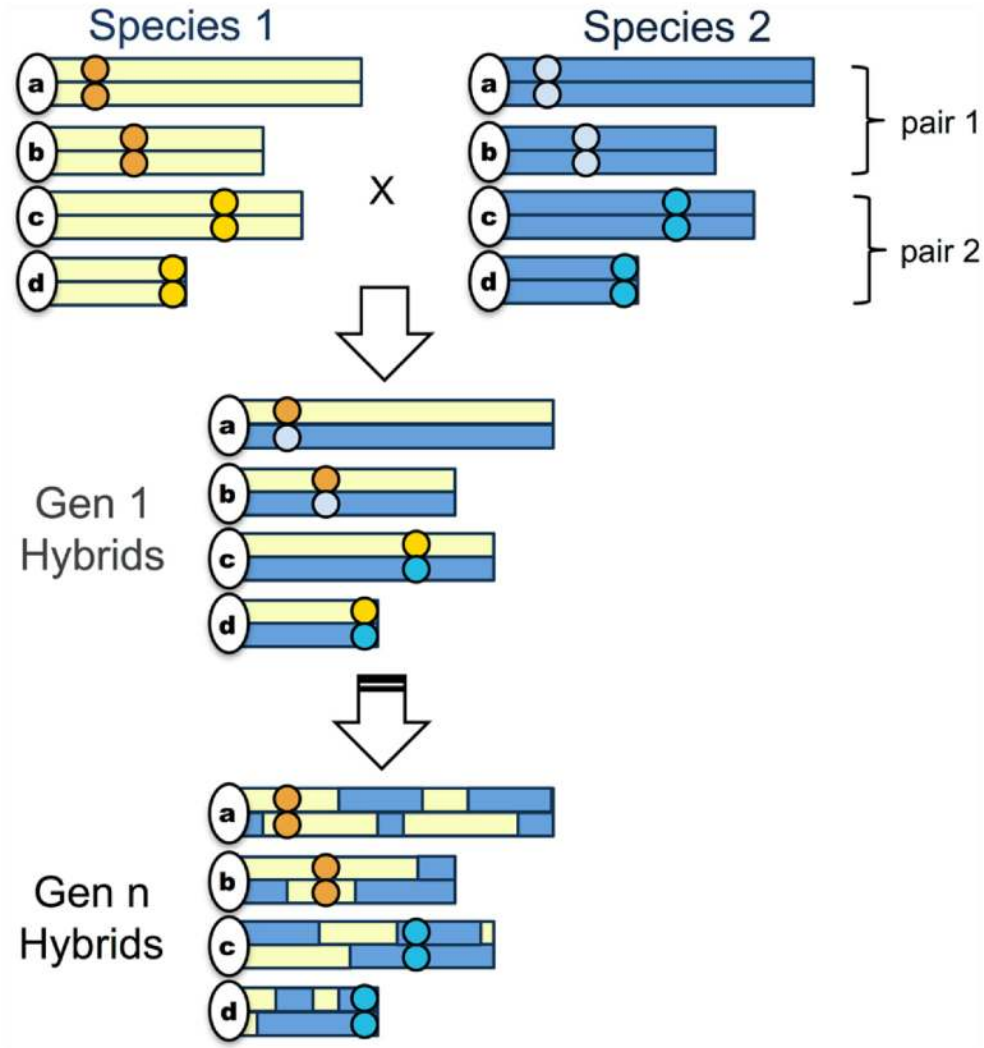
B Hybrid swarms



C Hybrid speciation

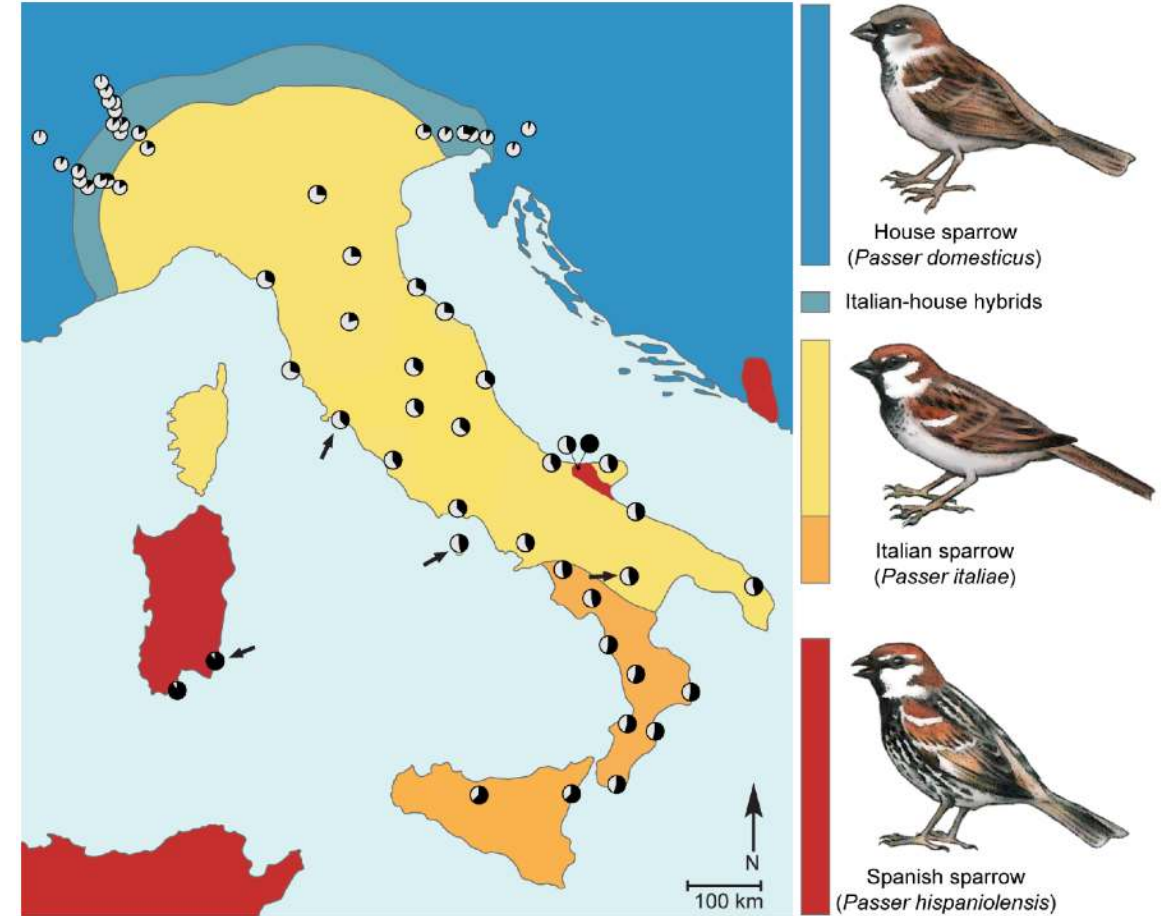


Sorting of incompatibilites



Schumer et al. 2015

Italian sparrow: Putative hybrid species



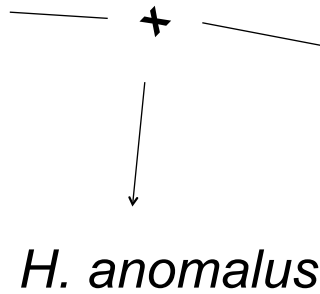
Hermansen et al. 2014

Hybrid sunflower species: homoploid hybrids

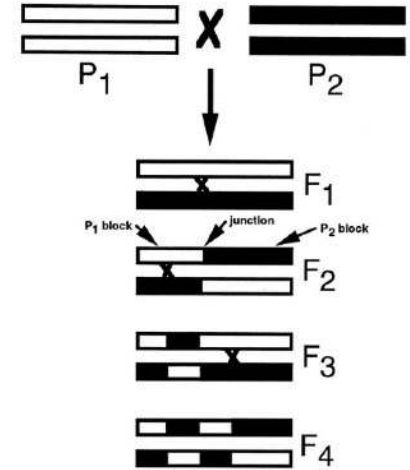
H. annuus



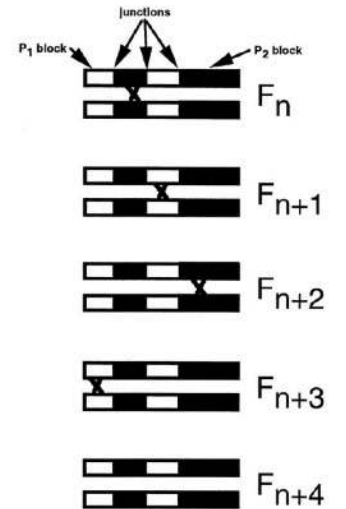
H. petiolaris



a

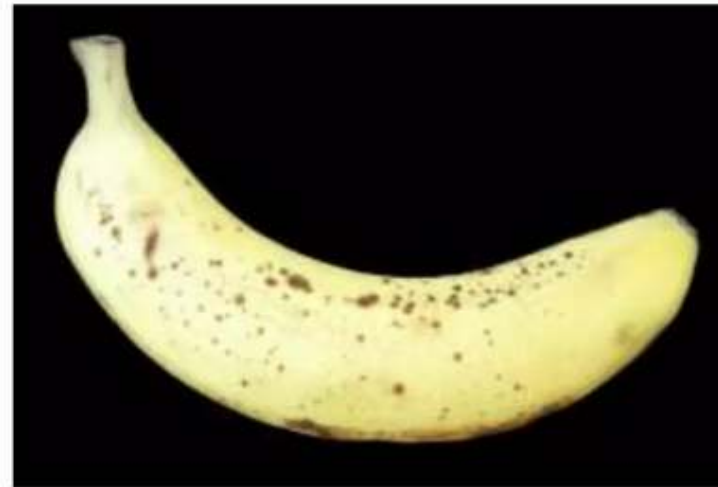
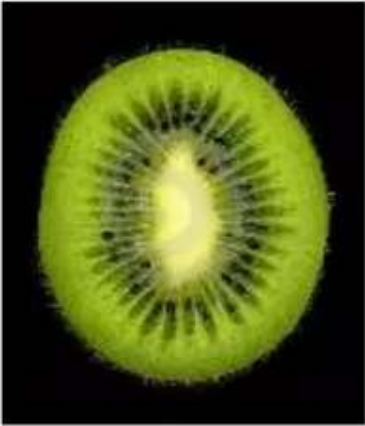


b



Ungerer et al. 1998 PNAS

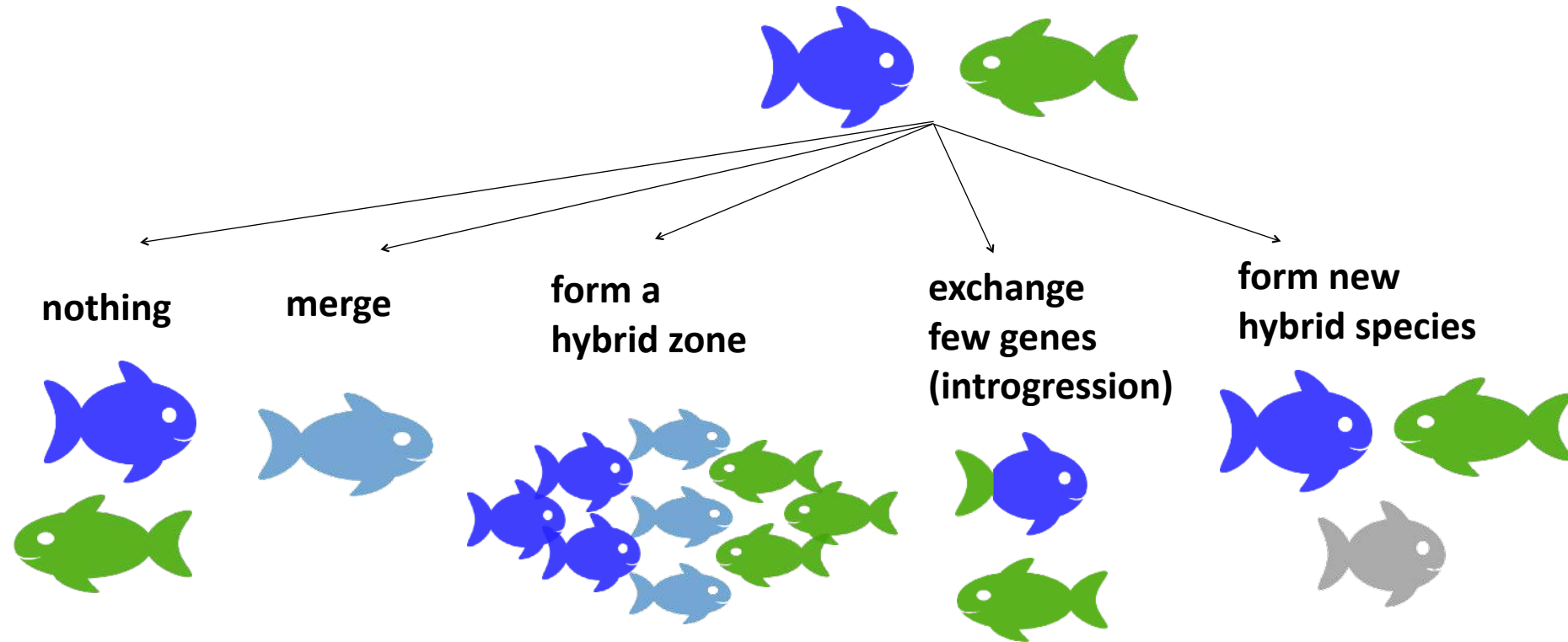
Allopolyploid hybrid species



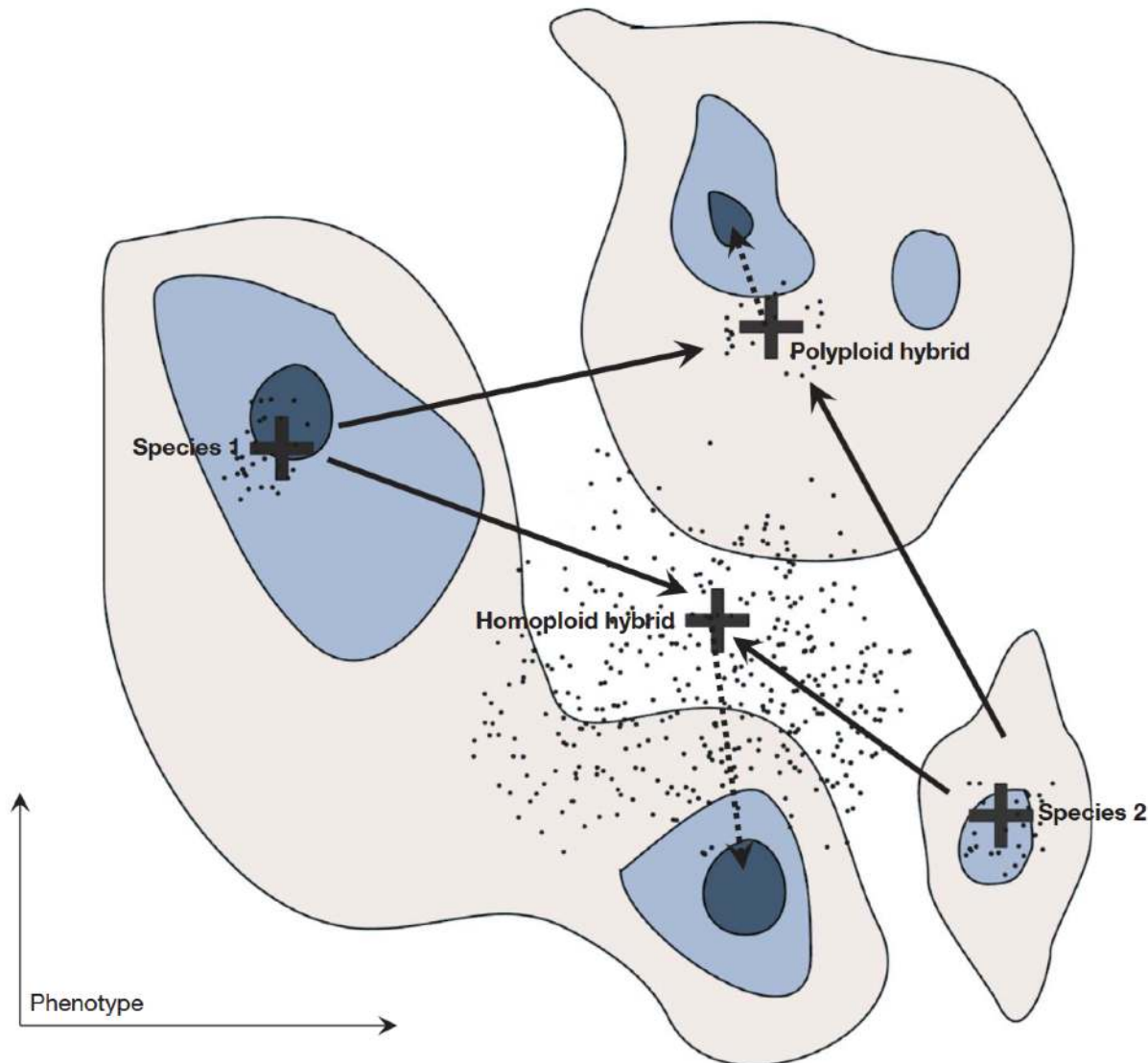
Old genetic variants, often derived from hybridisation, contributed to fast speciation and adaptation



What determines the evolutionary consequences?



Fitness of hybrids is key

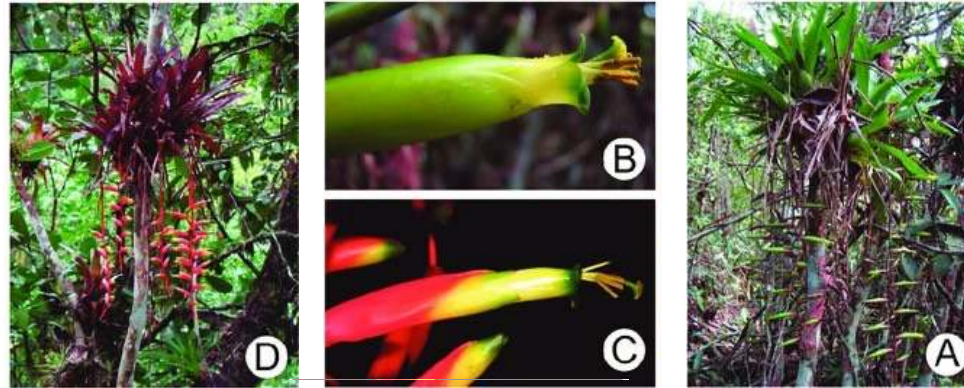


Hybrid speciation Fitness optima ('adaptive peaks') are coloured blue. Adaptive landscapes are not rigid, but are readily distorted by environmental or biotic changes, including evolutionary change. Mean phenotypes of species and their hybrids are shown as crosses, and offspring distributions as dots.

The time scale matters

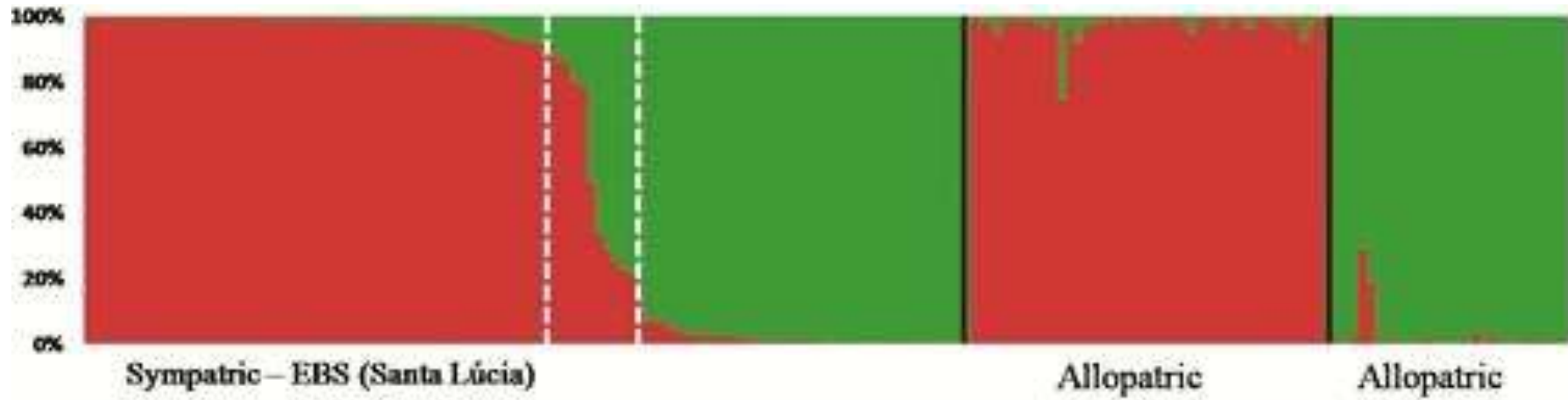
- Methods that only work with recent/ongoing hybridisation:
 - Hybrid zone study: cline analyses
 - Triangle plot
 - Clustering methods such as ADMIXTURE or STRUCTURE
 - F_{ST} and D_{XY} to infer regions with low gene flow as barriers to reproduction
 - Haplotype assignment
- Methods that also work with ancient hybridisation:
 - Demographic modeling (e.g. Fastsimcoal, gIMble)
 - Phylogeny-based methods such as D statistics
 - TWISST (Martin et al. 2019)
 - MSCi (multispecies coalescent-with-introgression, in BPP: Flouri et al. 2020)
 - QuIBL (Edelman et al. 2019)
 - Network methods such as Phylonet, AIM, GPhoCS

Hybrid zones: Clustering methods to identify hybrids and infer ancestry proportions: Bromeliads

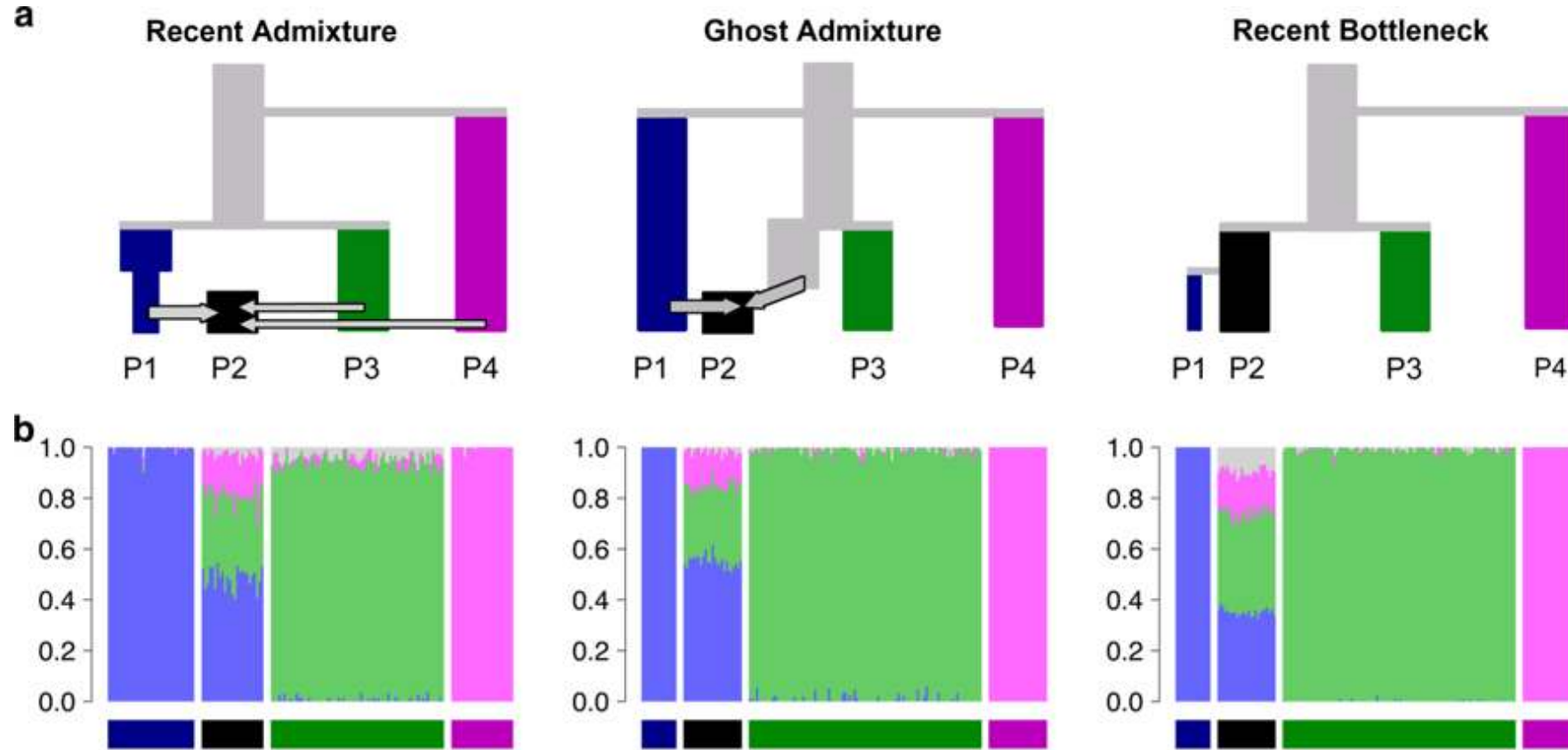


Vriesia simplex

Vriesia scalaris

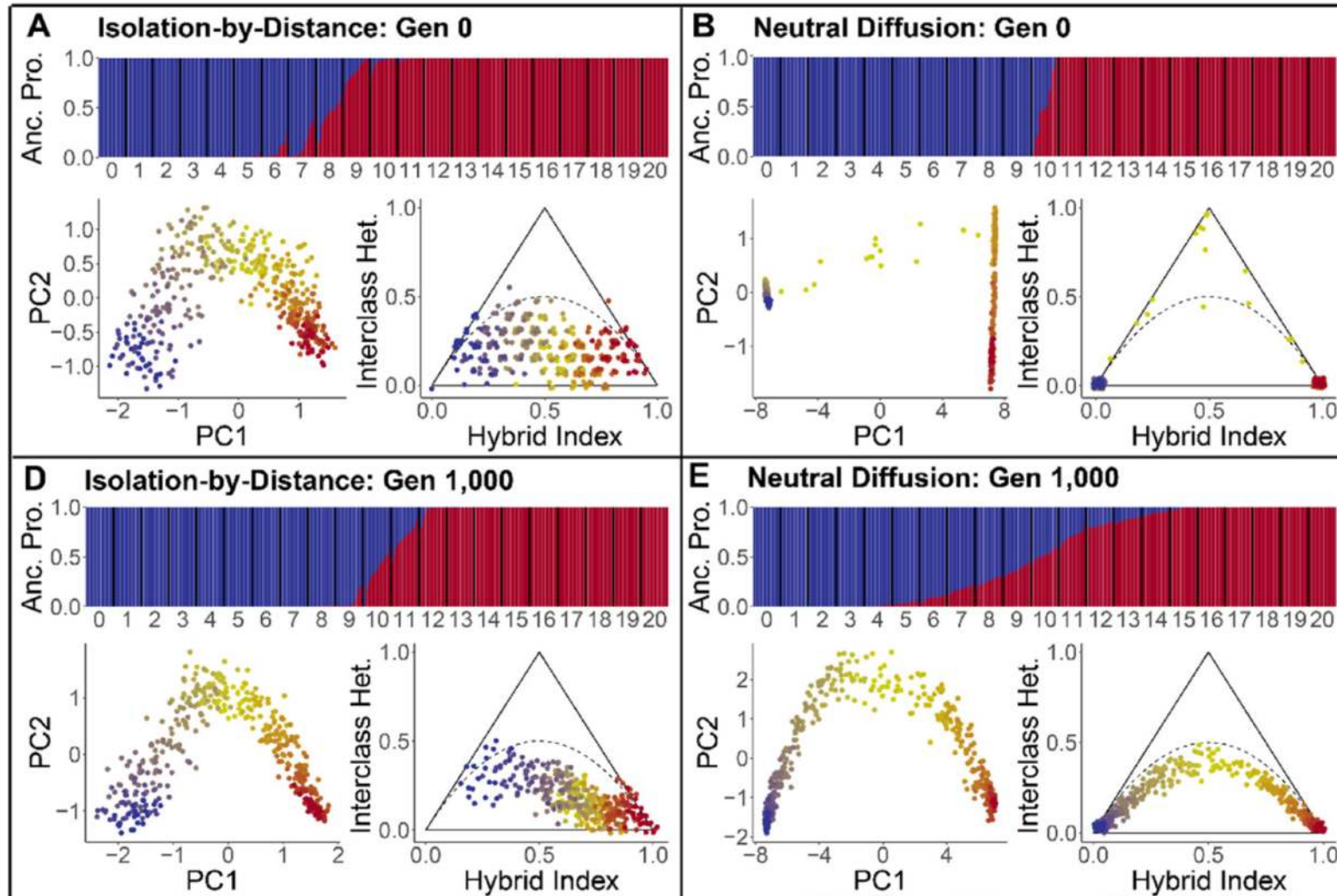


Very different evolutionary histories can give the same ADMIXTURE plot -> badMixture

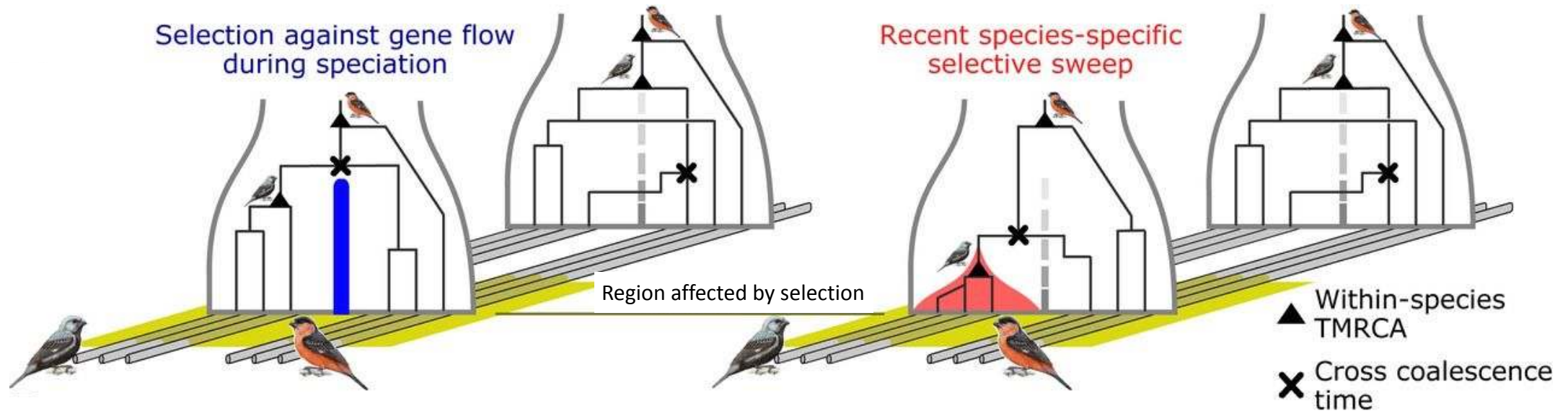


Very important: Use STRUCTURE/ADMIXTURE plots only to get a first overview of genetic structure. Putative cases of hybridization need to be followed up with other analyses such as D statistics and/or demographic modeling.

Isolation by distance can mimic hybridisation in STRUCTURE plots -> use triangle plots



Gene flow creates gene tree – species tree discordance

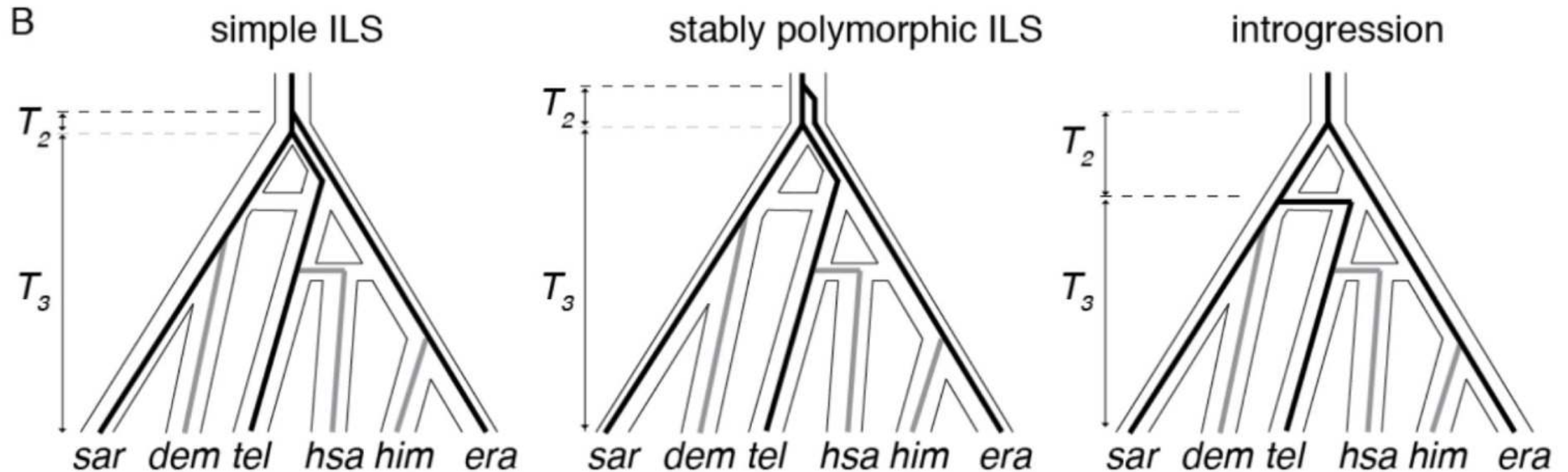


Genomic regions with selection against gene flow or recent selective sweeps show less ILS

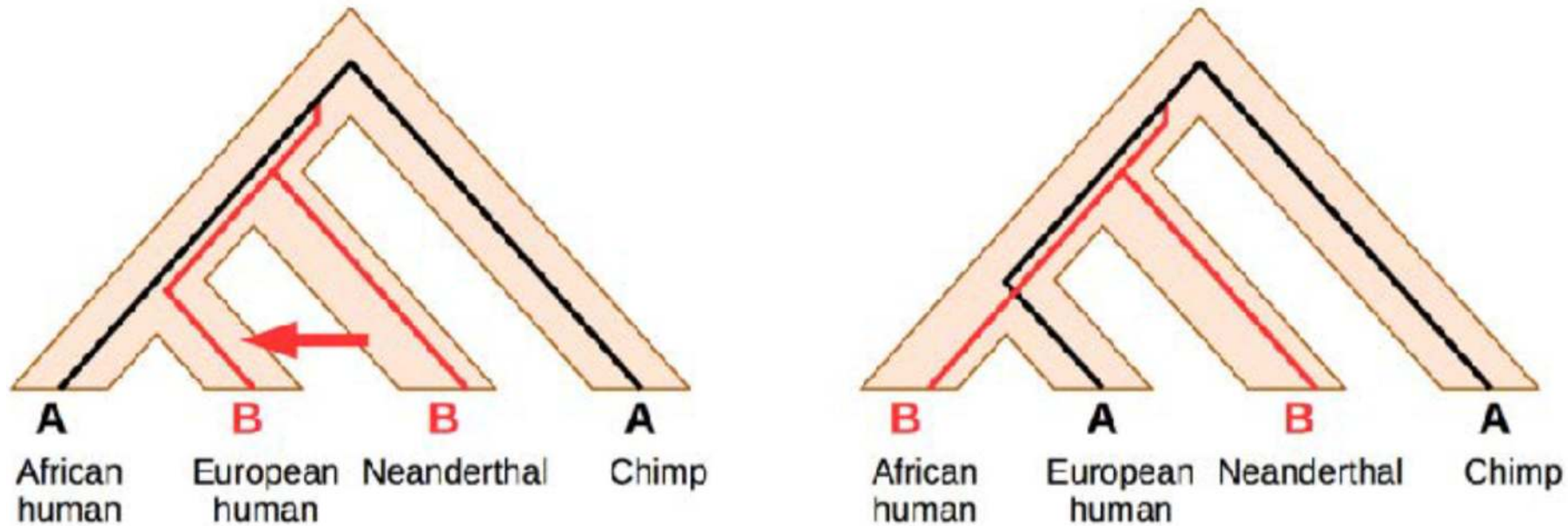
-> Selection can affect levels of ILS

Methods that use branch lengths

e.g. QuIBL (Edelman et al. 2019)



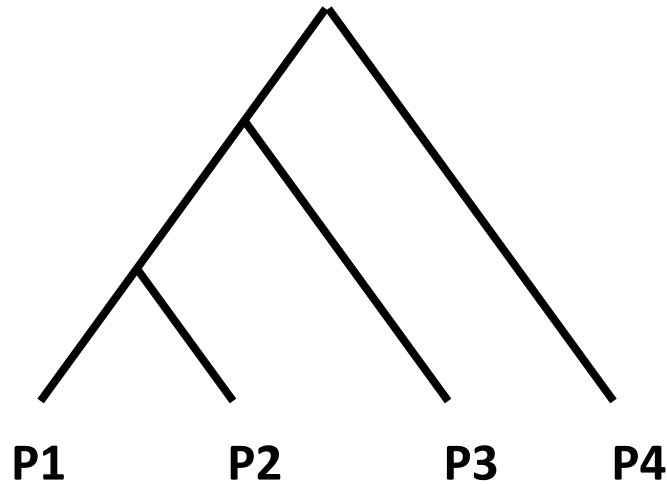
Patterson's D statistics to identify hybridisation



$$D(P_1, P_2, P_3, O) = \frac{\sum C_{ABBA}(i) - C_{BABA}(i)}{\sum C_{ABBA}(i) + C_{BABA}(i)} \quad (1)$$

Patterson's D statistics to identify hybridisation

- Also called ABBA-BABA test: $D = (ABBA - BABA) / (BABA + ABBA)$

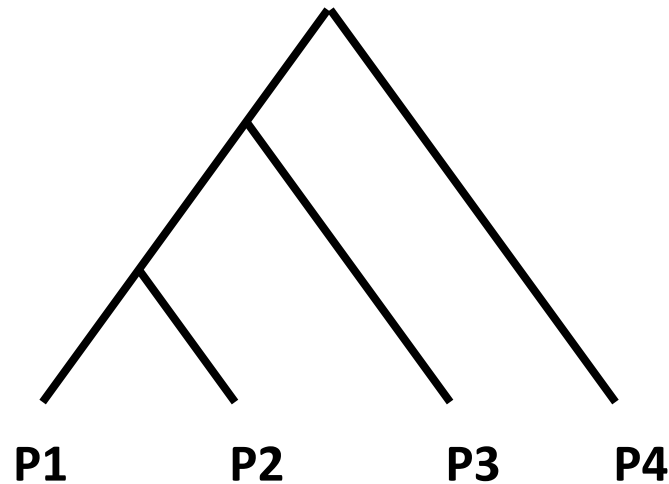


$$D = \frac{ABBA - BABA}{ABBA + BABA}$$

	P1	P2	P3	P4	
Pos 1	T	T	A	A	Concordant SNPs
Pos 20	G	G	G	T	
Pos 23	C	T	T	T	
Pos 35	C	A	C	A	Discordant SNPs
Pos 41	A	T	T	A	

Patterson's D statistics to identify hybridisation

- Also called ABBA-BABA test: $D = (ABBA - BABA) / (BABA + ABBA)$

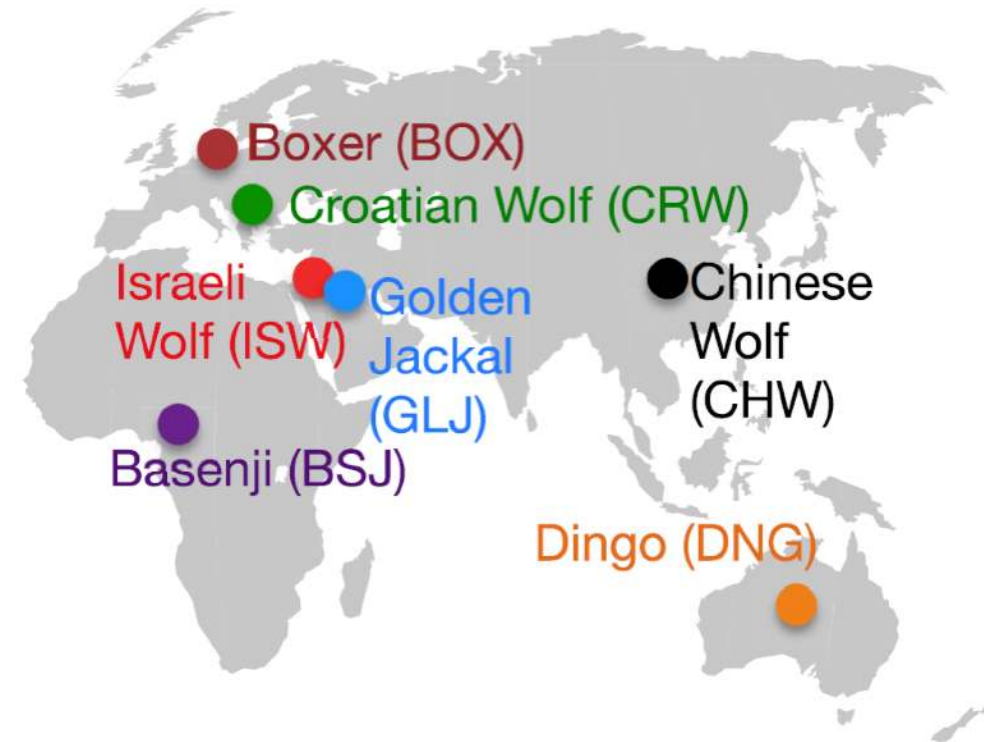
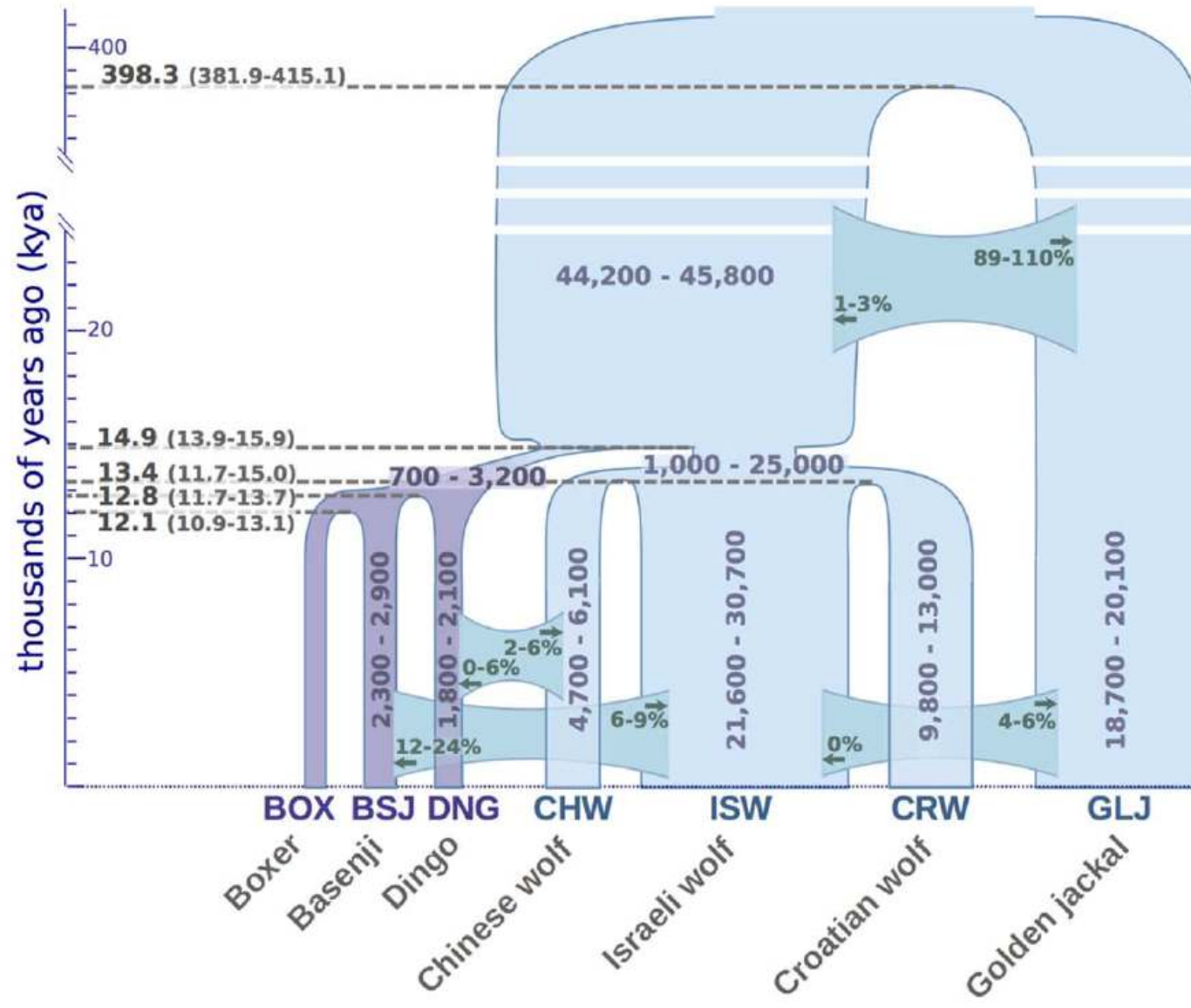


$$D = \frac{ABBA - BABA}{ABBA + BABA}$$

Pos 1	T	T	A	A	} Concordant SNPs
Pos 20	G	G	G	T	
Pos 23	C	T	T	T	
Pos 35	C	A	C	A	← BABA
Pos 41	A	T	T	A	← ABBA
Pos 44	C	G	G	C	← ABBA

$$D = \frac{2 - 1}{2 + 1} = \frac{1}{3}$$

Network methods, e.g. GPhoCS



Examples from my work

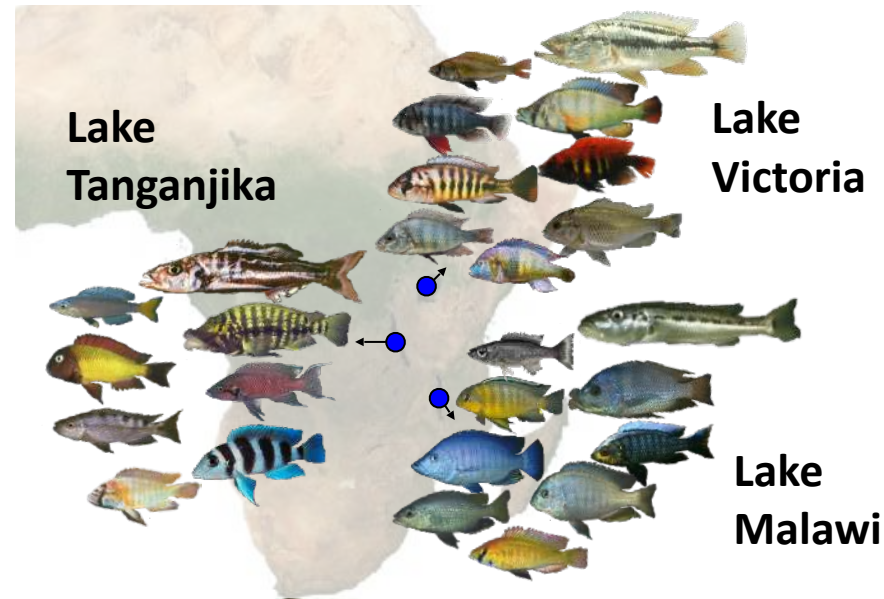
Some lineages diversify again and again

Anolis lizards on the Caribbean islands and the mainland



Huie et al., 2021

Cichlid fishes in East Africa



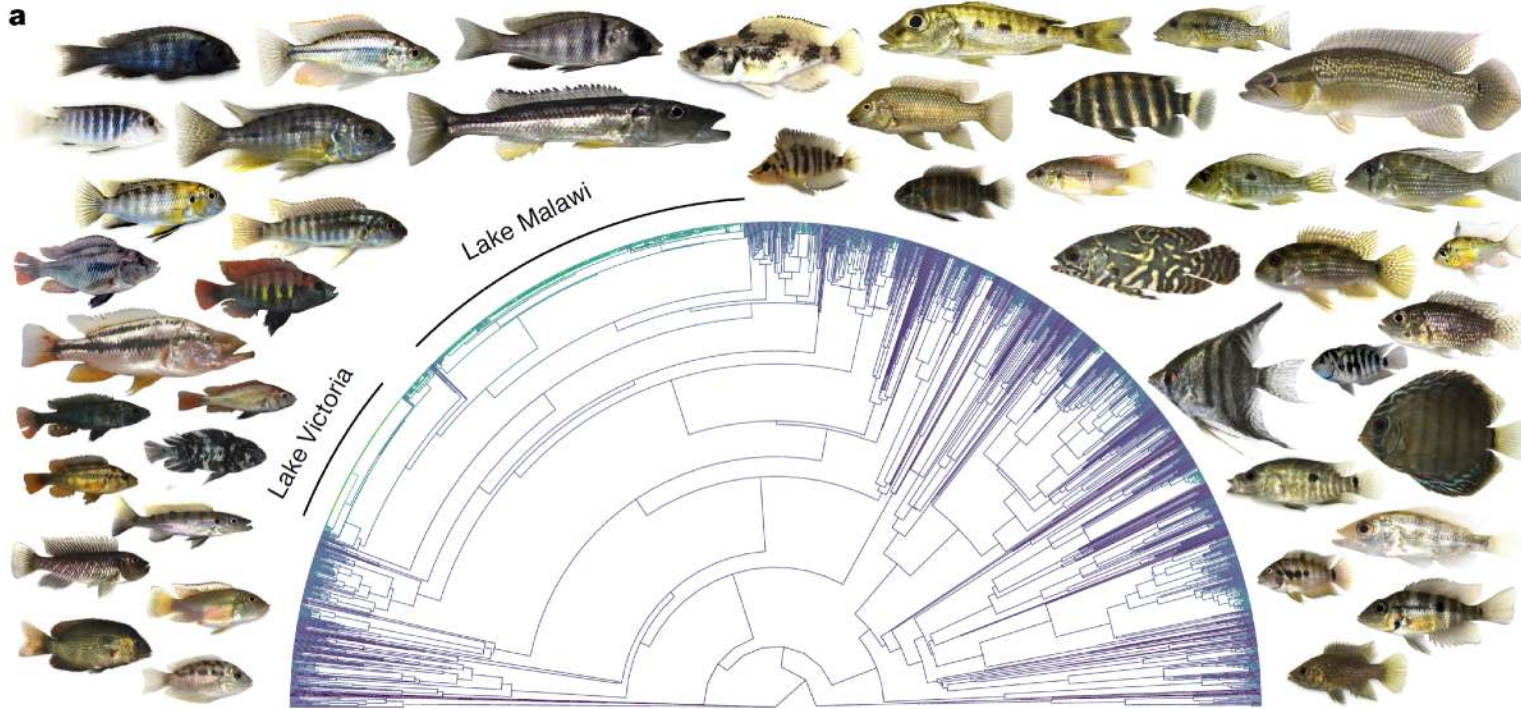
Brawand et al., 2014

Hawaiian lobeliads



Givnish et al. 2008

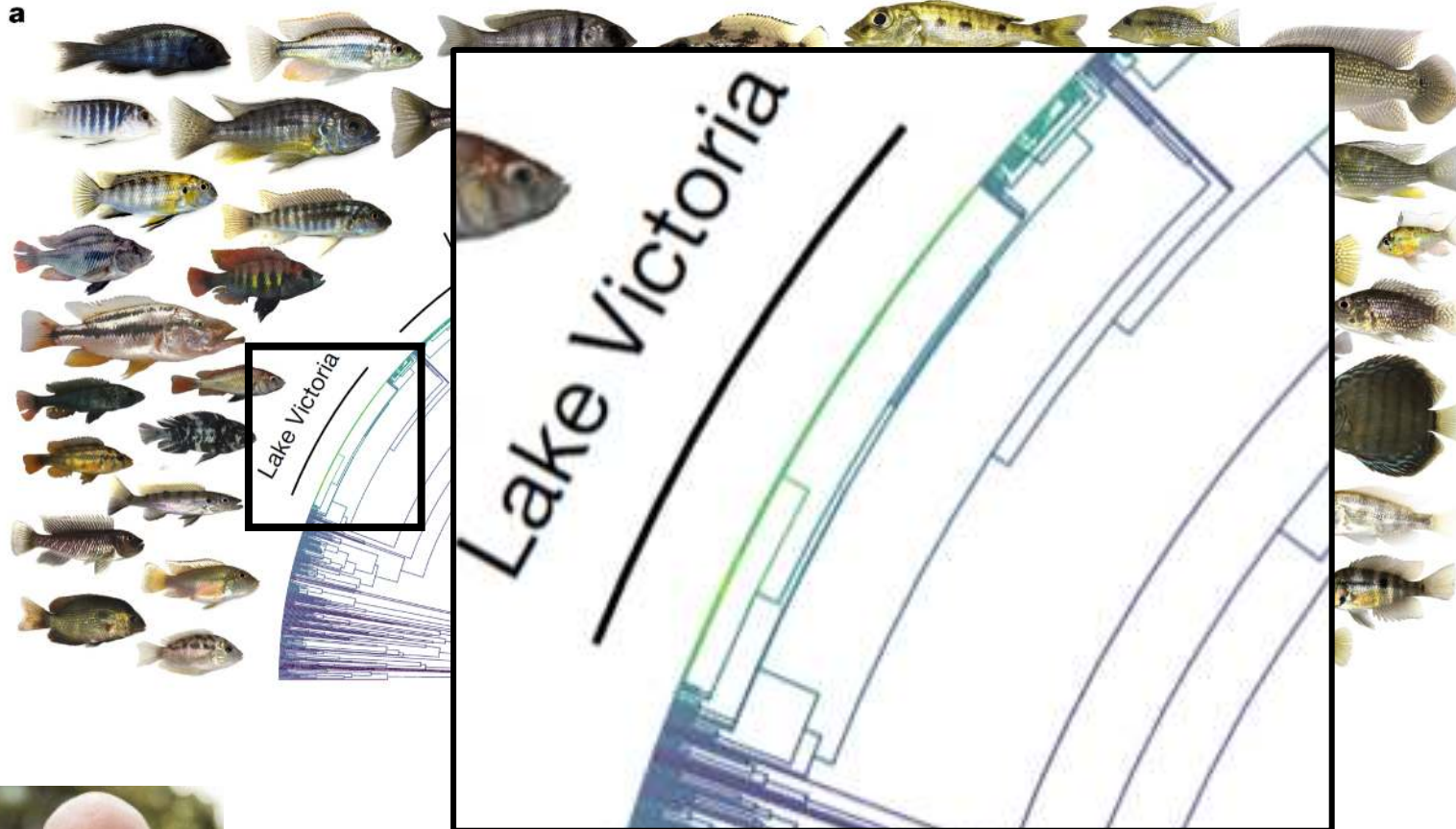
Extrinsic and intrinsic factors together best predict adaptive radiation



**Matt
McGee**
(Monash U)

All described 1,712 species
McGee et al., 2020, Nature

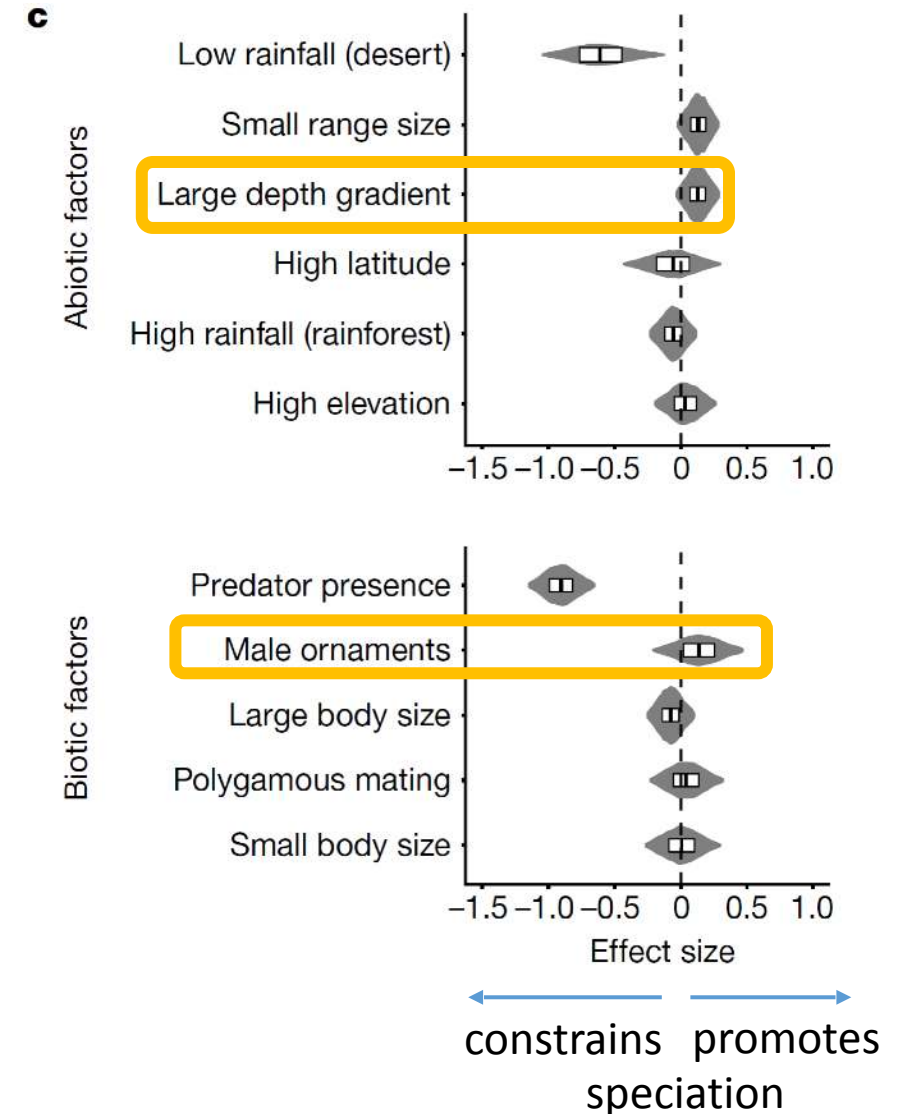
Extrinsic and intrinsic factors together best predict adaptive radiation



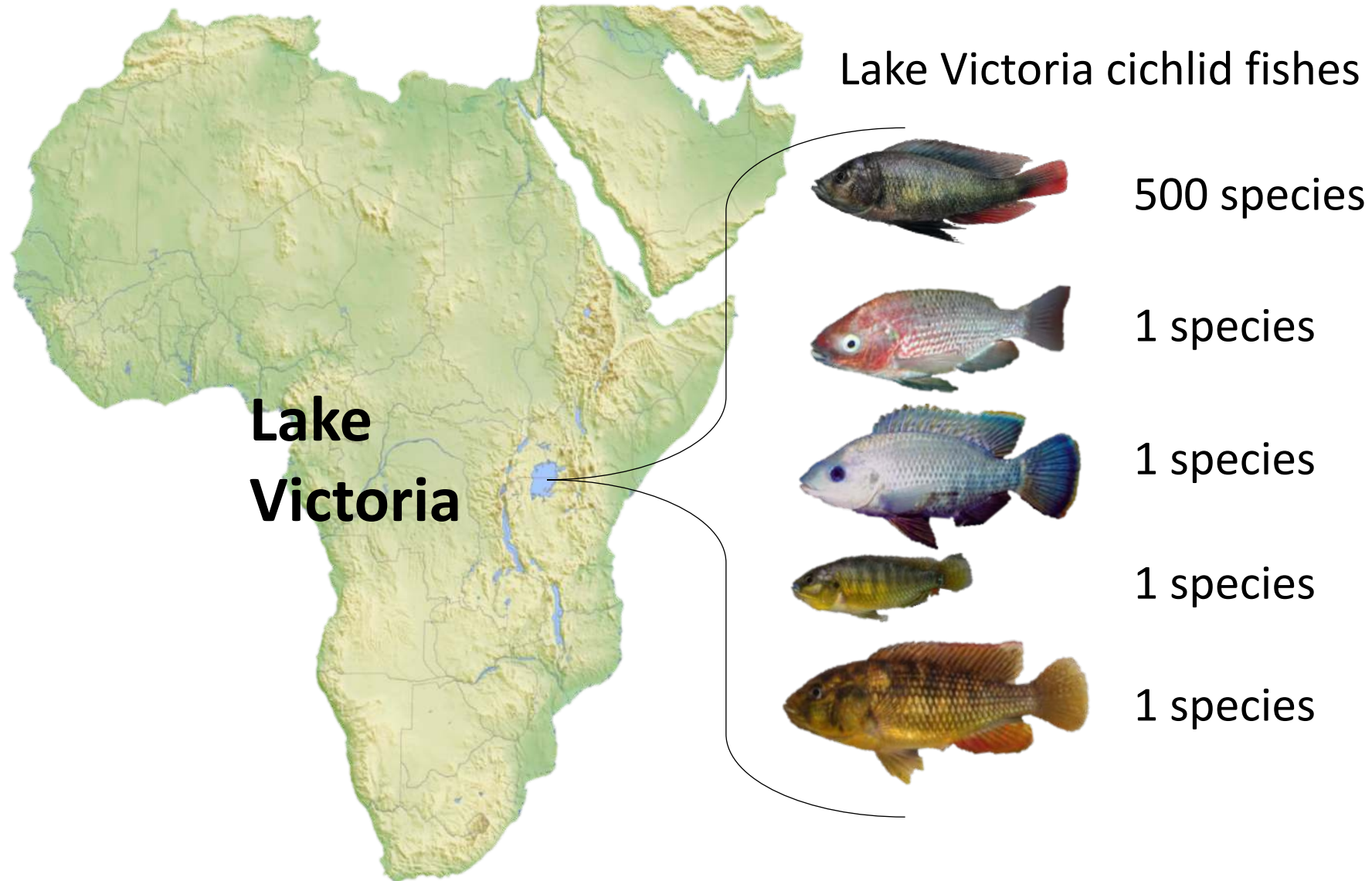
Matt McGee
(Monash U)

All described 1,712 species
McGee et al., 2020, Nature

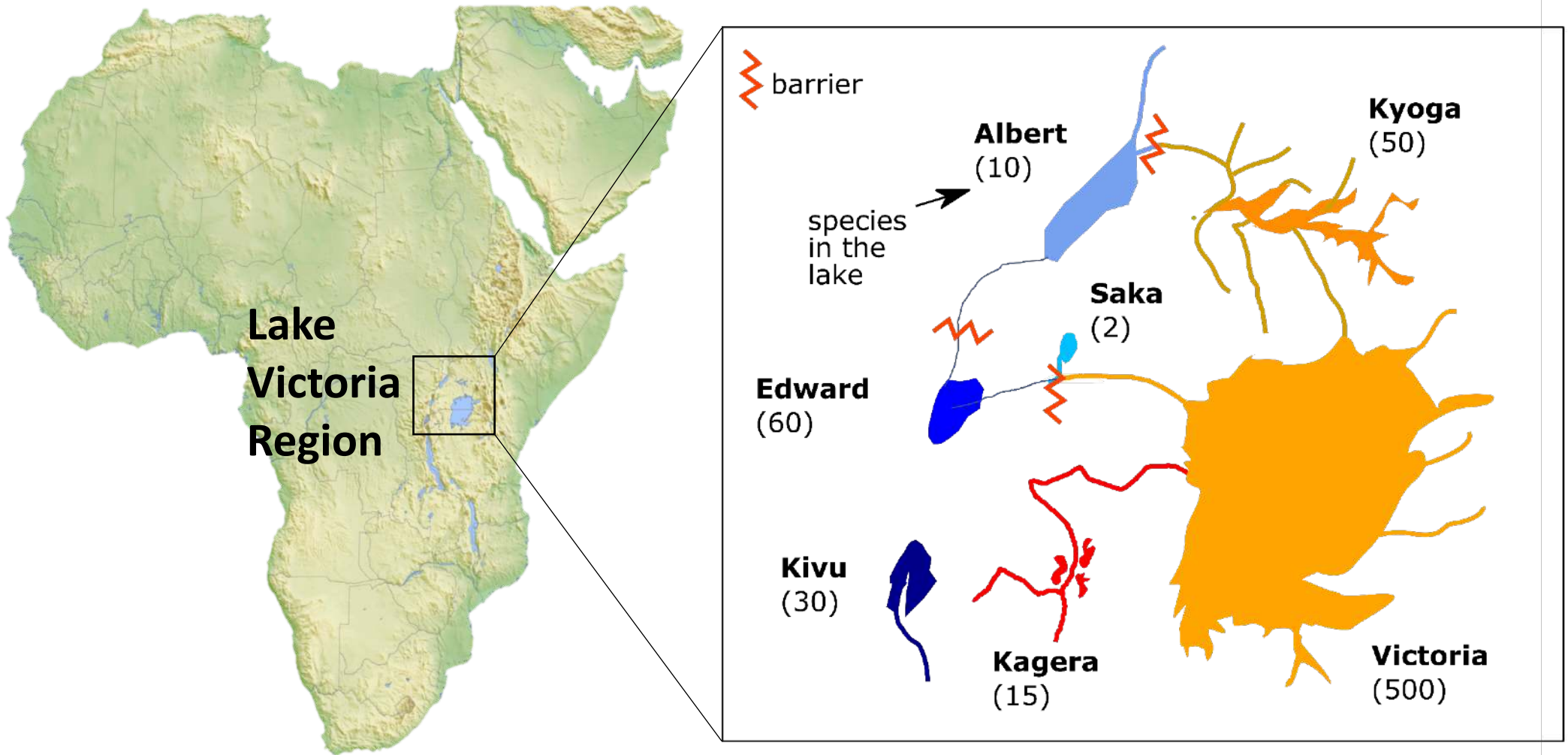
Factors affecting speciation rates



Cichlid lineages vary massively in species richness

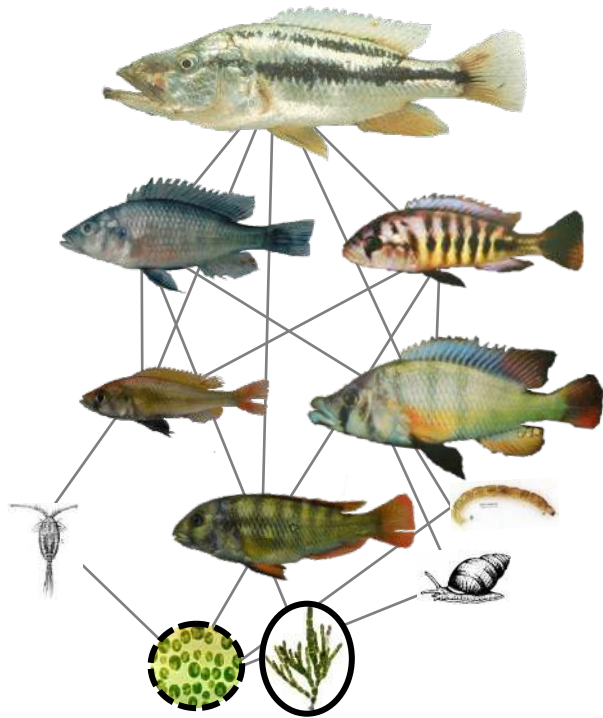


The same lineage diversified in all major lakes in the Lake Victoria Region in only 150,000 years into 700 species

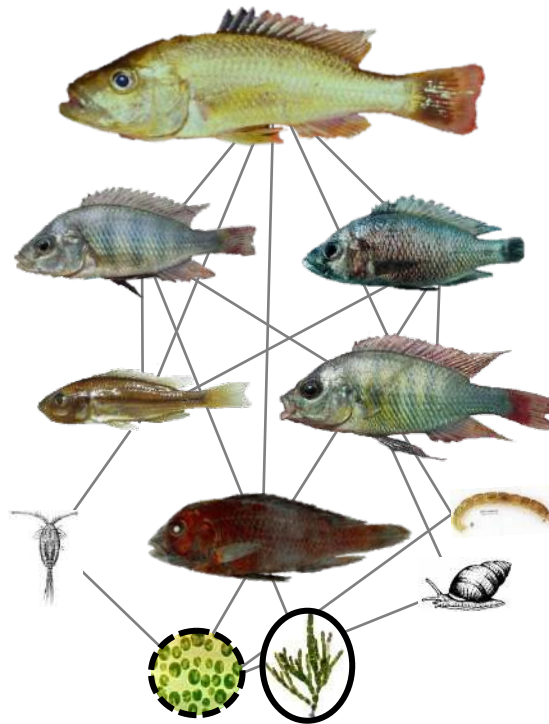


In each lake, the cichlids form species assemblages with similar ecomorphs

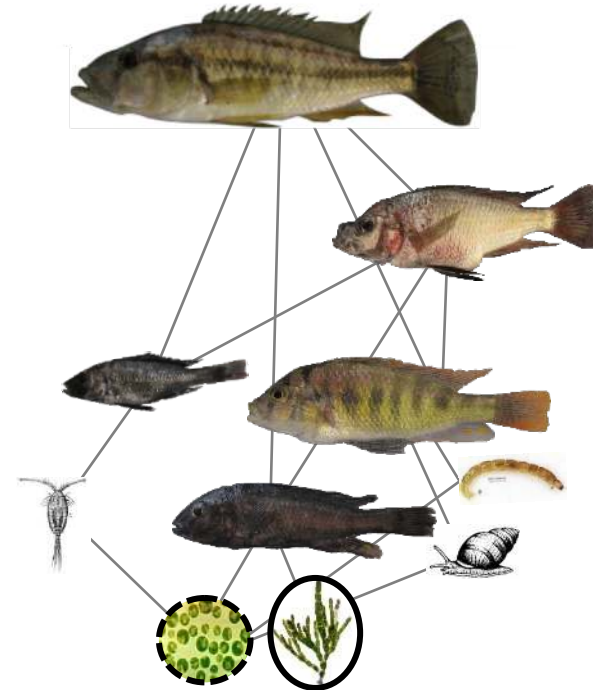
Lake Victoria



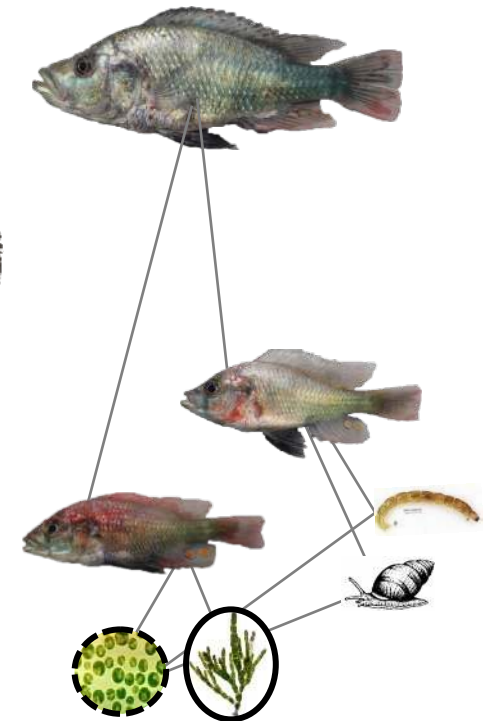
Lake Edward



Lake Kivu



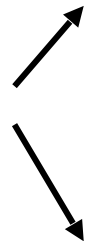
Kagera Lakes



Requirements for rapid adaptive radiation

Hybridisation

- boost genetic variation
- introduce incompatibilities



- Rapid adaptation to different niches for stable coexistence
- Reproductive isolation barriers that avoid merging of nascent species

Chromosomal rearrangements

- decrease hybrid fitness
- link co-adapted genes





**Ole
Seehausen**
(EAWAG,
U Bern,
Switzerland)



**David
Marques**
(Natural History
Museum Basel,
Switzerland)



**Laurent
Excoffier**
(U Bern,
Switzerland)



**Salome
Mwaiko**
(EAWAG,
Switzerland)



**Mary
Kishe**
(TAFIRI,
Tansania)



**Katie
Wagner**
(U Wyoming,
USA)



**Matt
McGee**
(Monash U,
Australia)

u^b

u^b

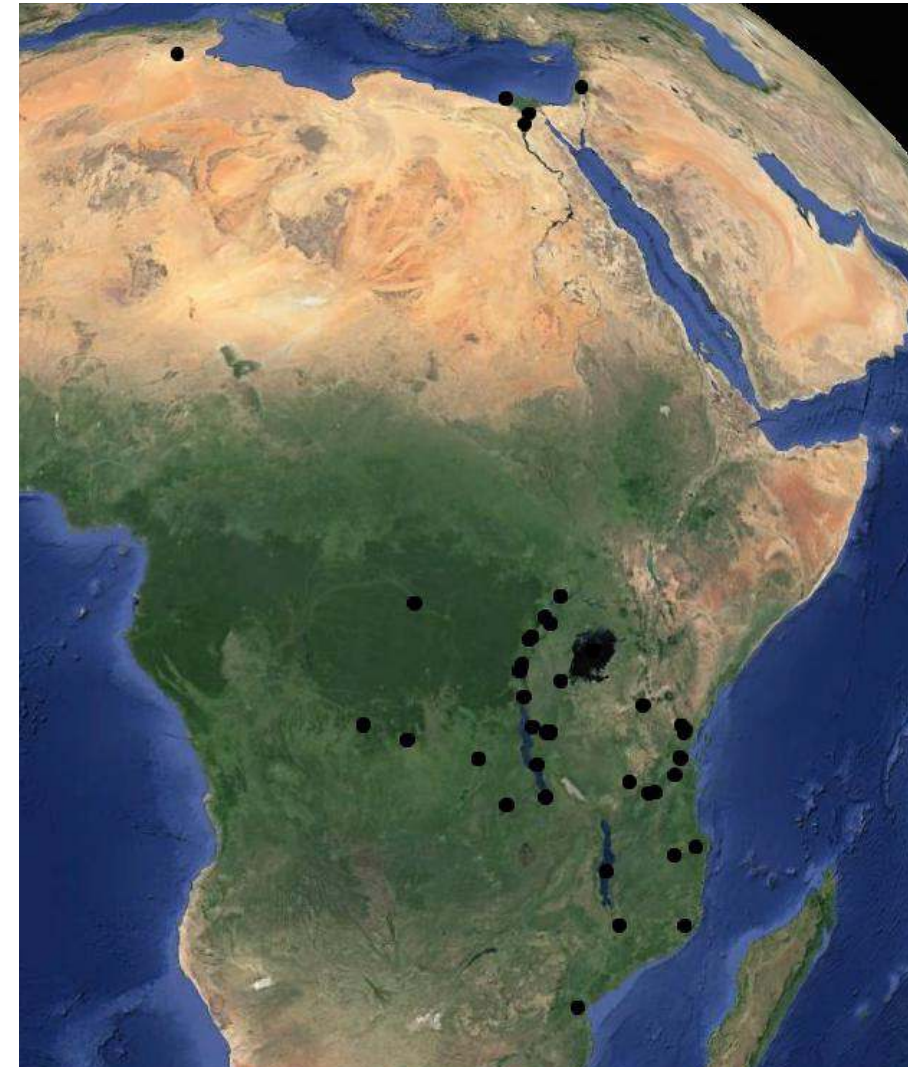
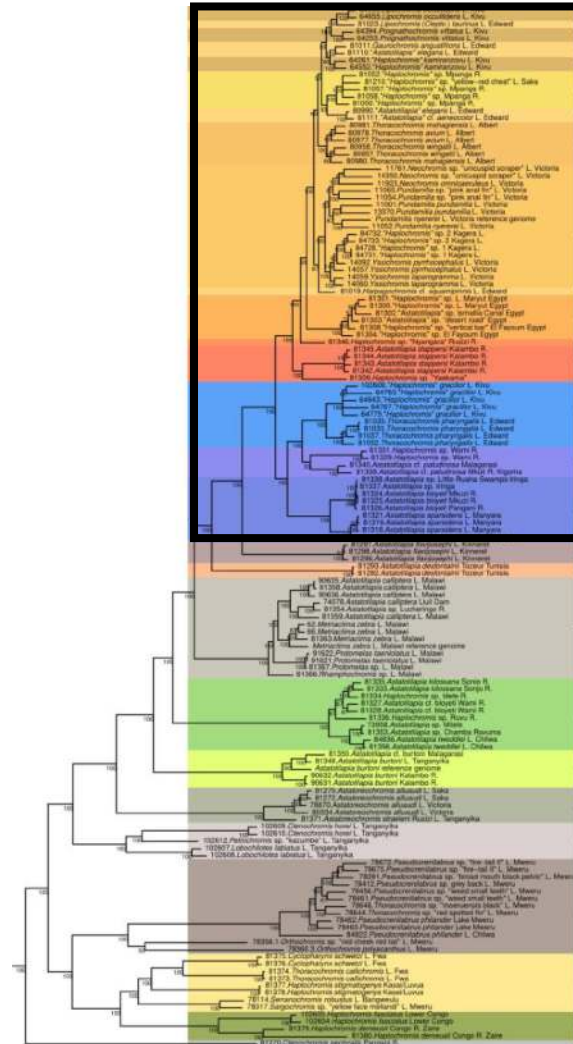
**UNIVERSITÄT
BERN**

eawag
aquatic research ooo



Identifying the closest relatives of the radiations

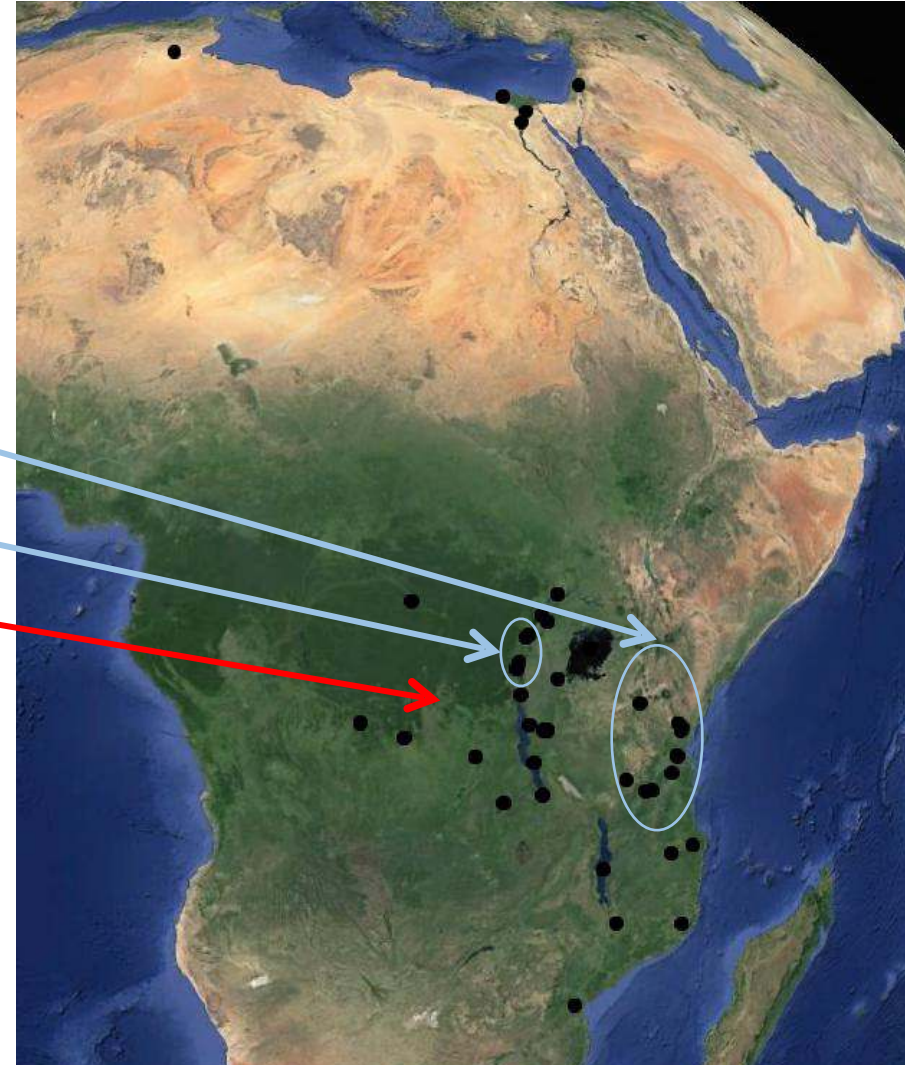
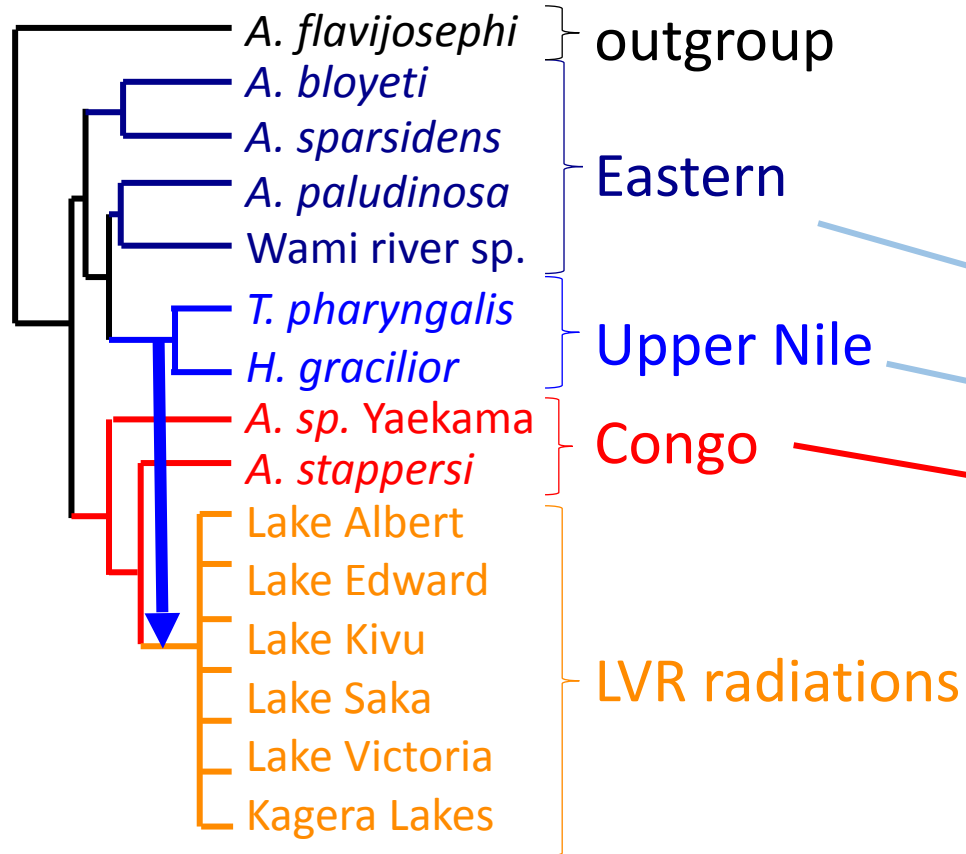
(RAD-sequencing of >100 cichlid species)



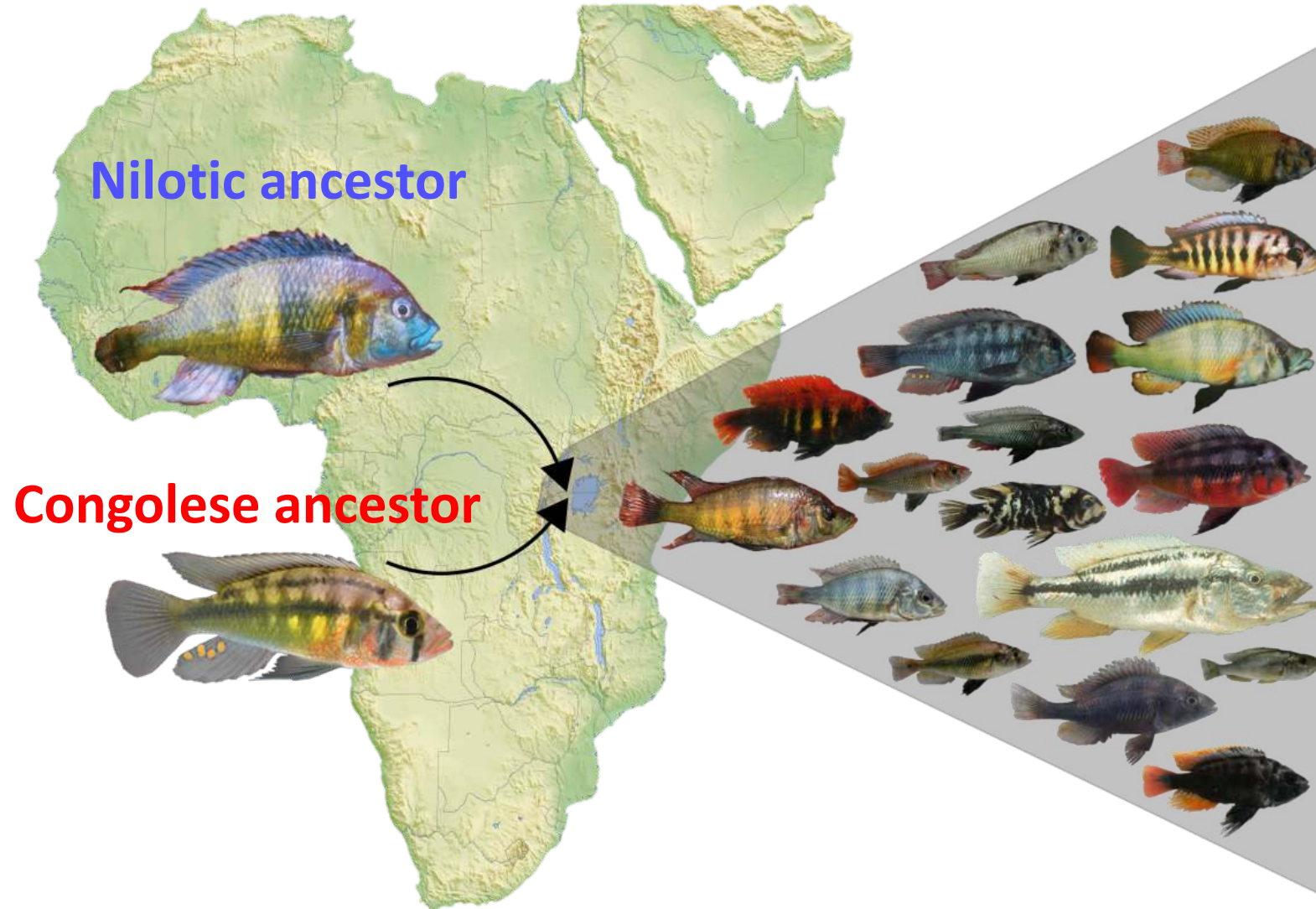
Meier *et al.* (2017) *Nat Commun*

Identifying the closest relatives of the radiations

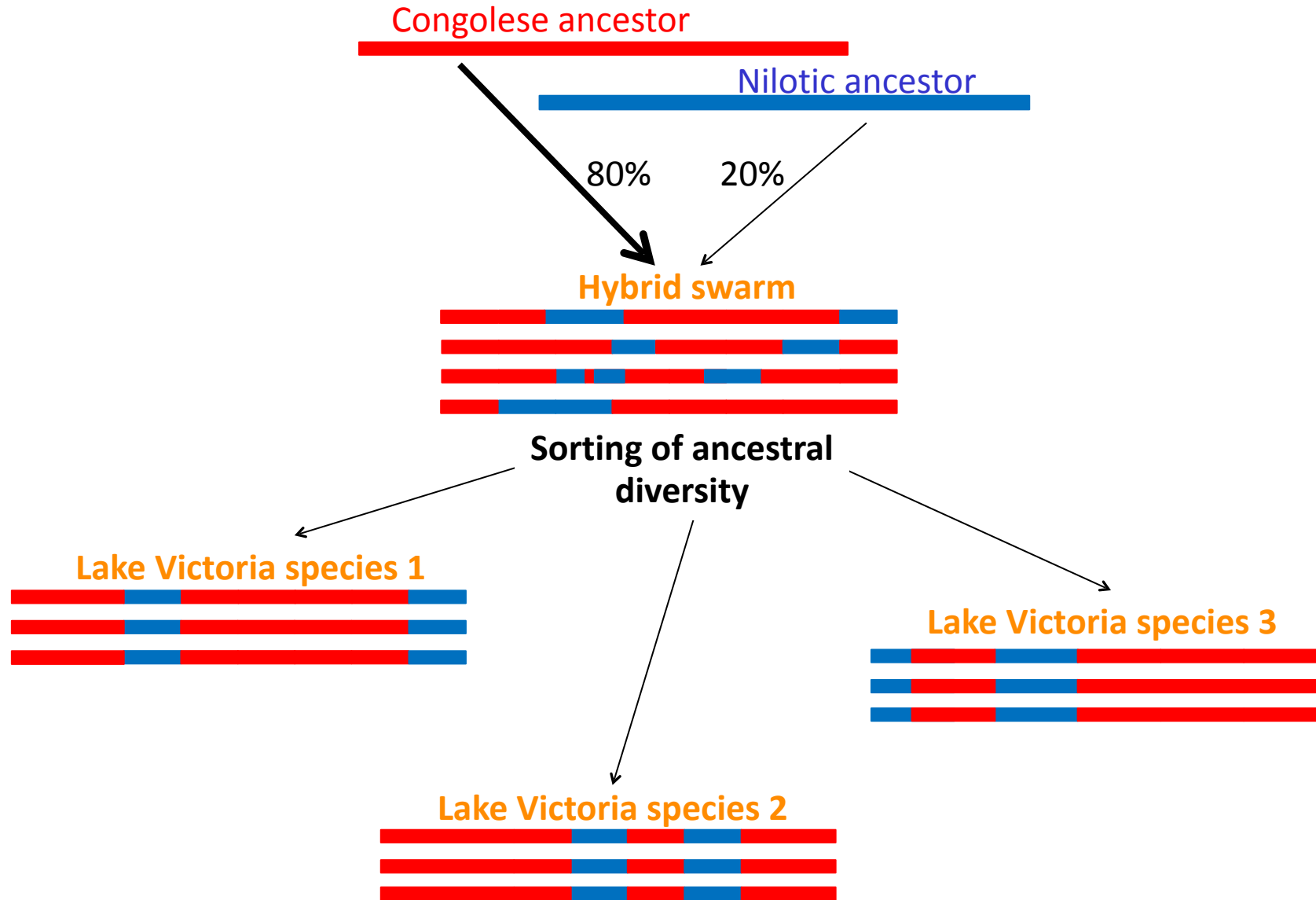
(RAD-sequencing of >100 cichlid species)



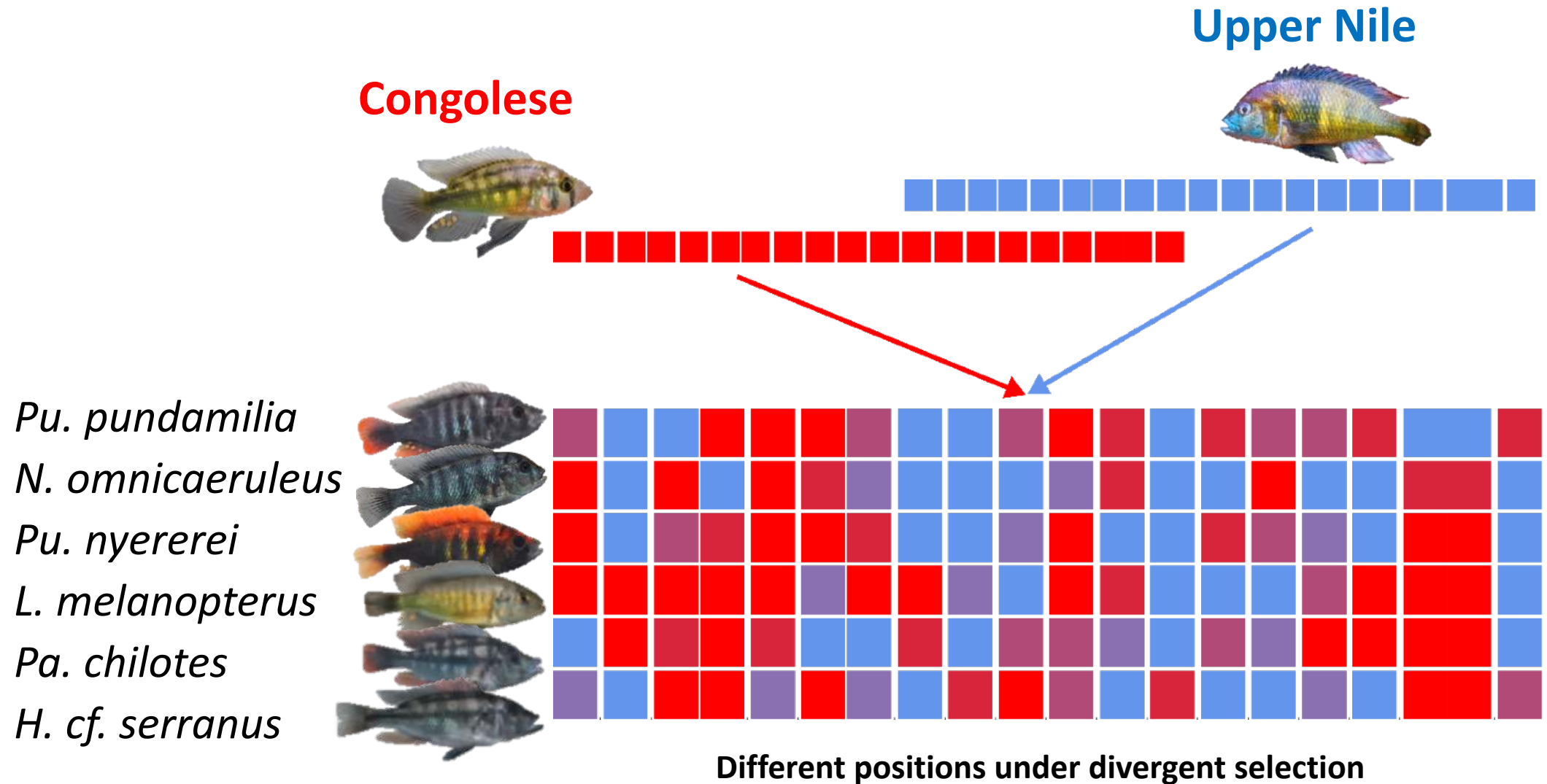
Hybrid origin of the lineage that speciated rapidly



Sorting of the parental lineage variation into different species



At sites under divergent selection, different species are different genomic mosaics of the parental lineages



Admixture provided key genetic variation

LWS opsin gene
(red-sensitive)

e.g. clear shallow water algivores



haplotype
class I

Congo



96

e.g. turbid deep water detritivores



haplotype
class II

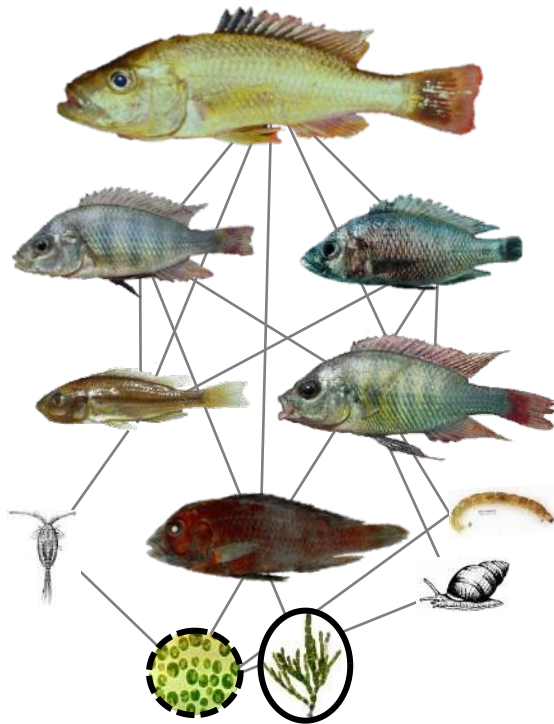
Upper Nile



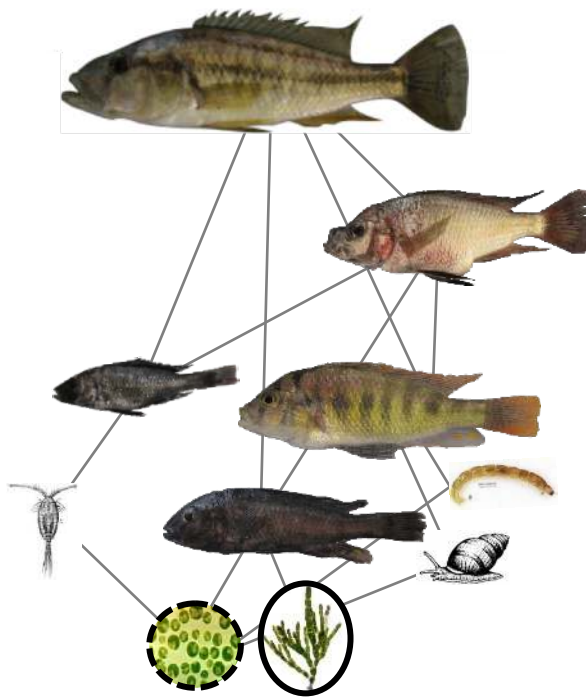
63

How did this admixture variation of 150 kya contribute to the origin of the much younger radiations after the Pleistocene dry period?

Lake Edward

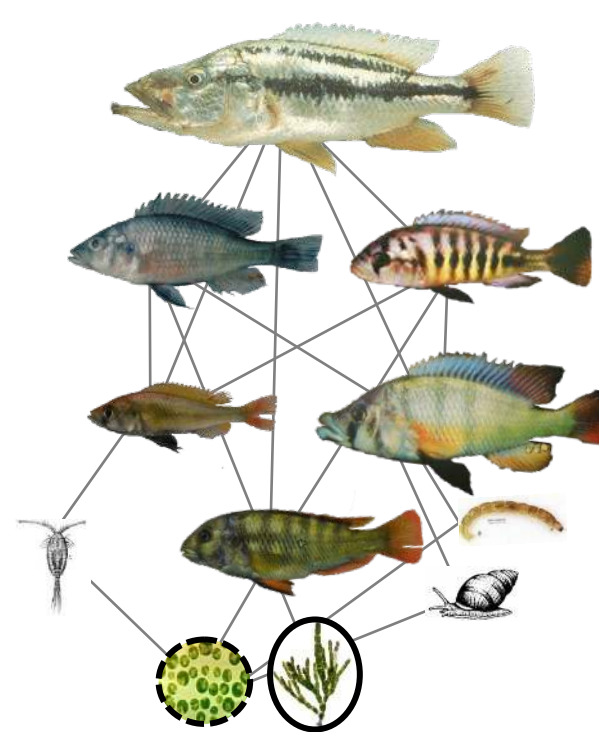


Lake Kivu

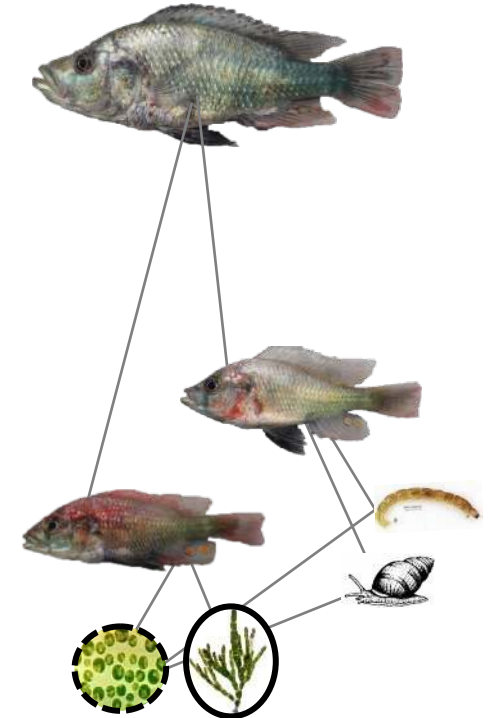


Less than 16,000 years old!

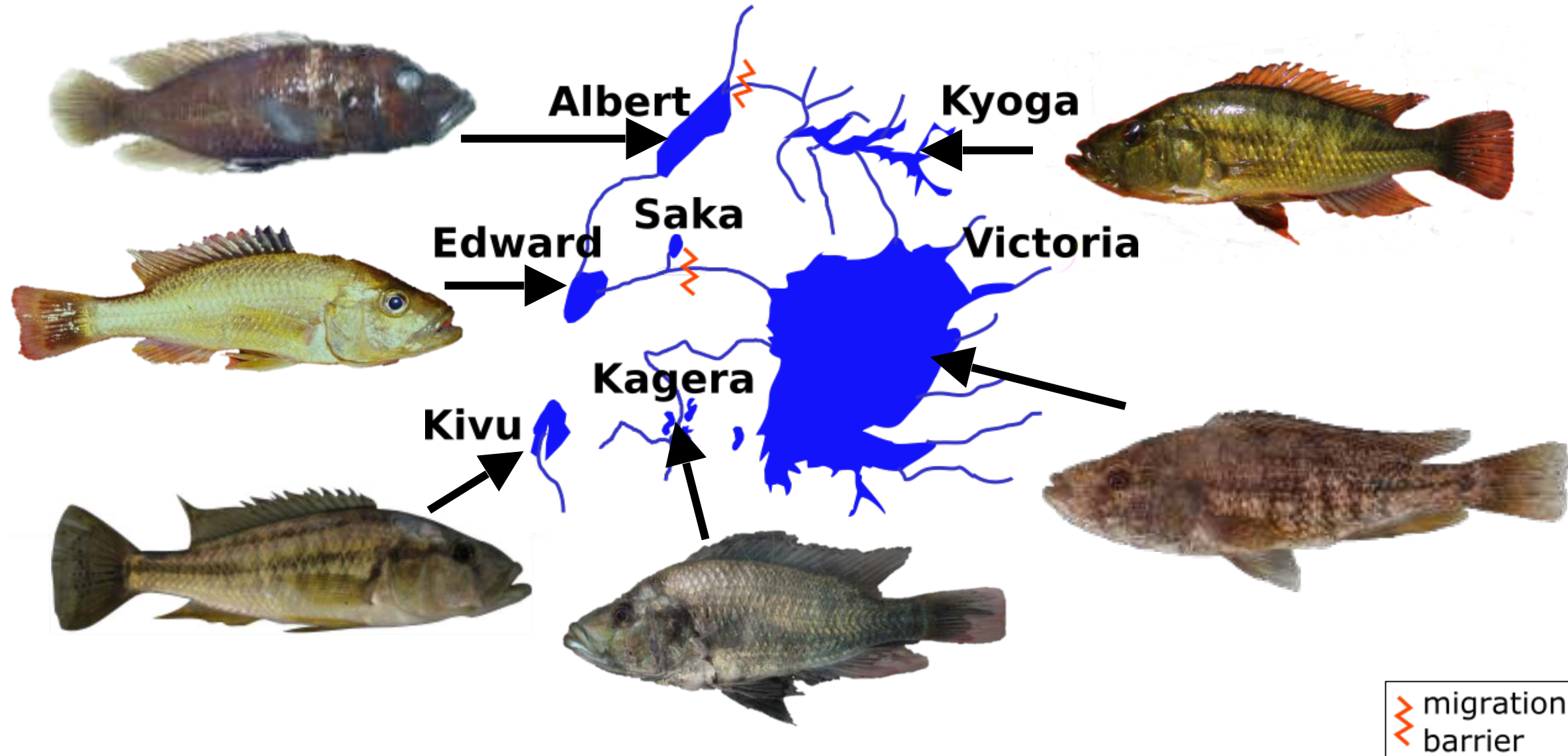
Lake Victoria



Kagera Lakes

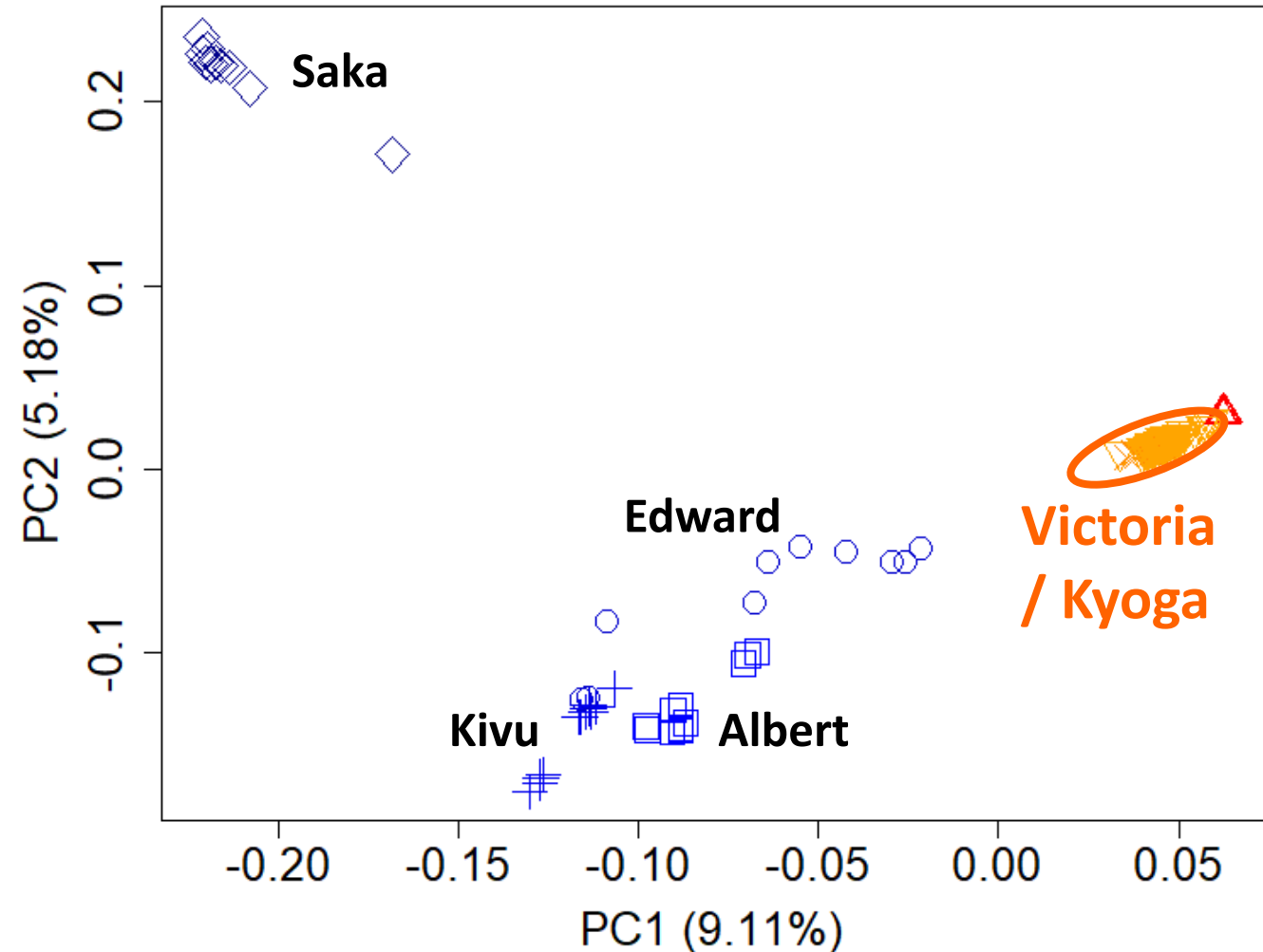
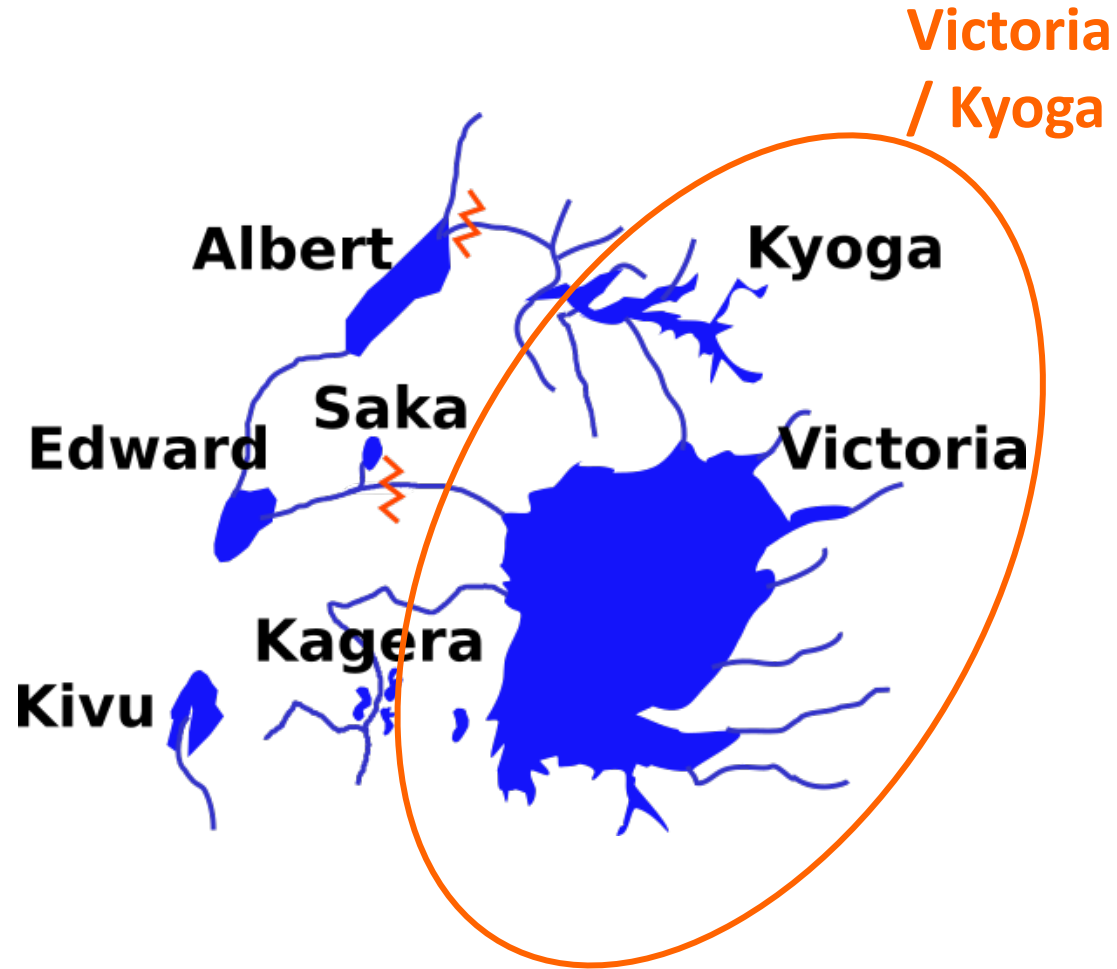


Predators across different lake radiations are extremely similar in morphology



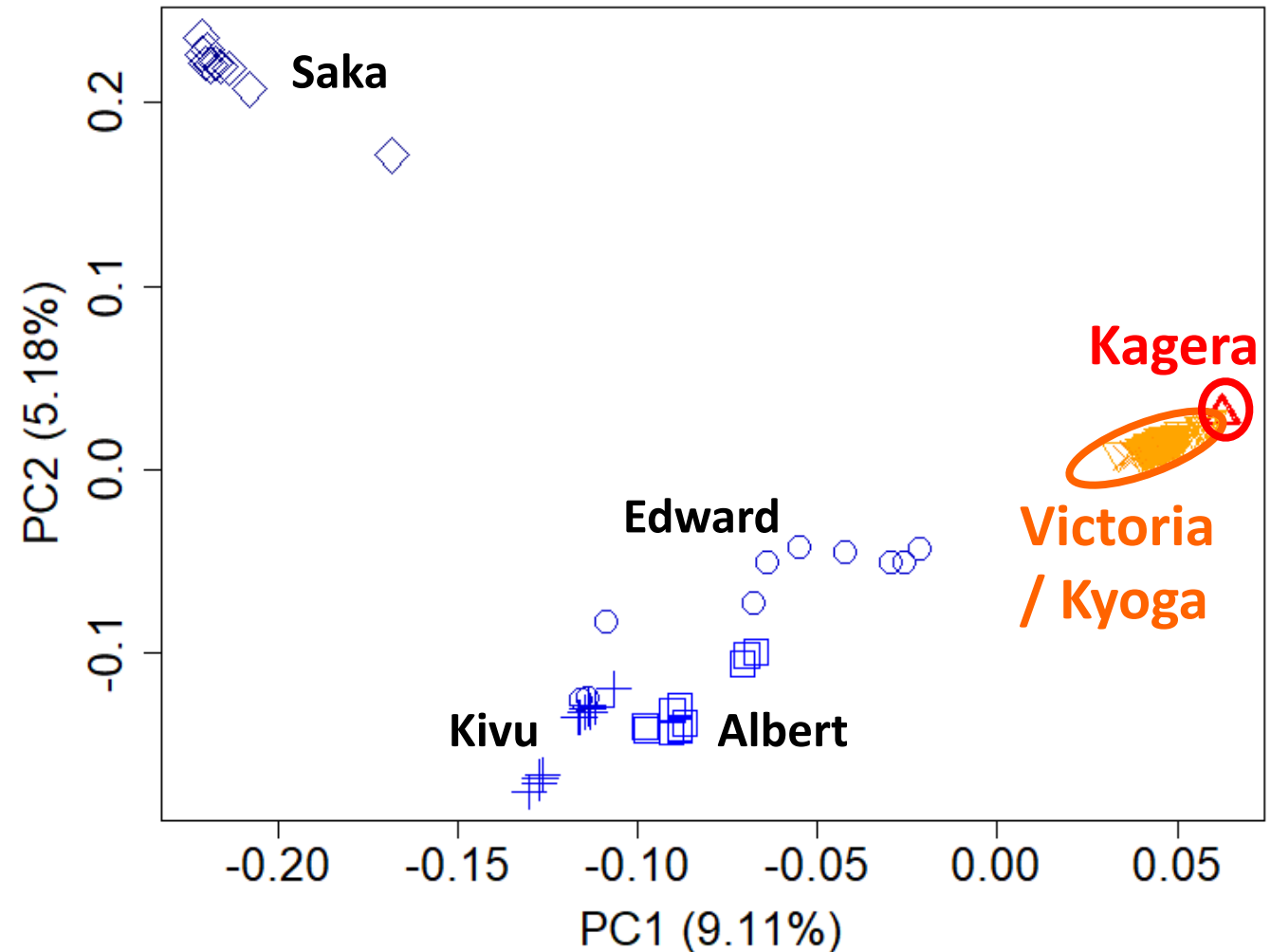
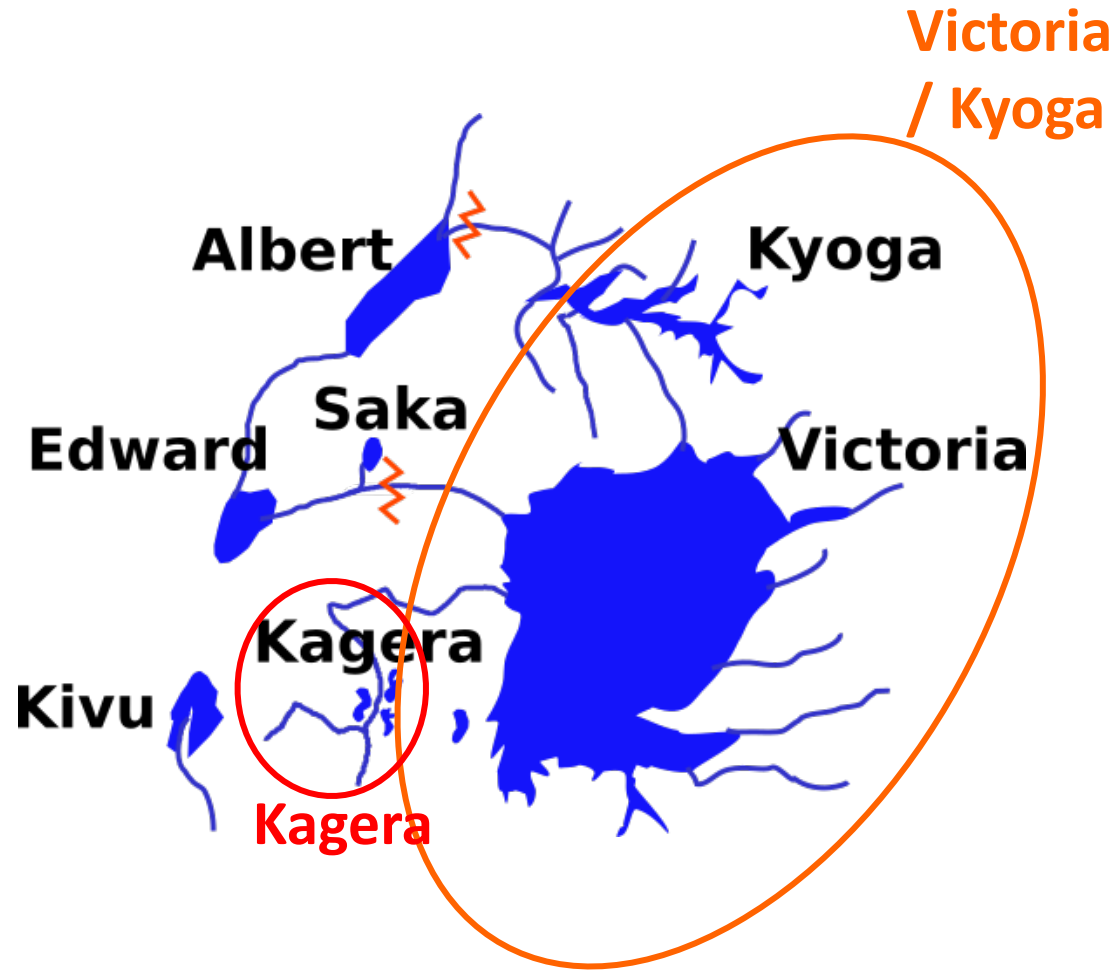
PCA on whole-genomes separates Lake Victoria/Kyoga cichlids from others

(152 genomes, 1.6M LD-pruned SNPs)



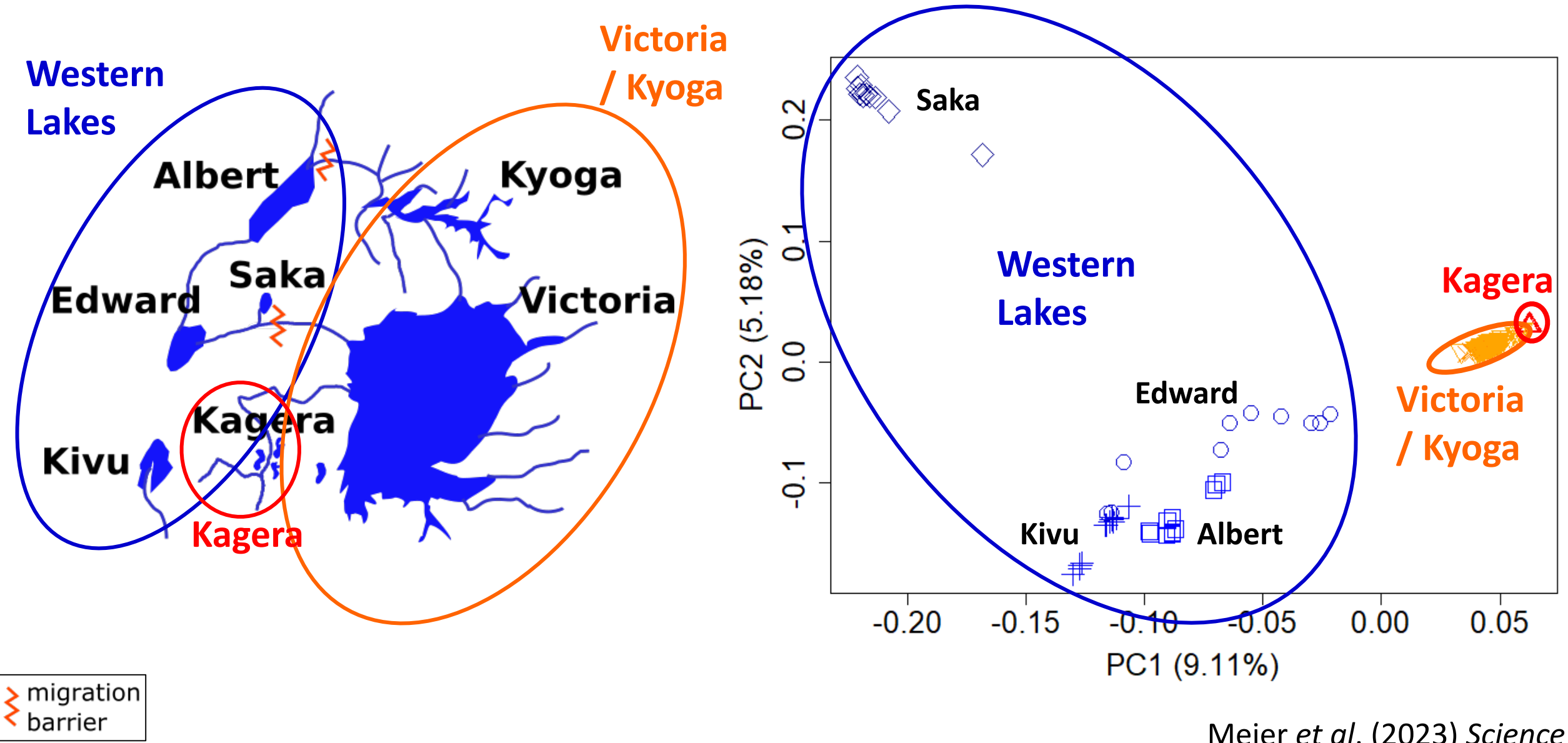
PCA on whole-genomes separates Lake Victoria/Kyoga cichlids from others

(152 genomes, 1.6M LD-pruned SNPs)



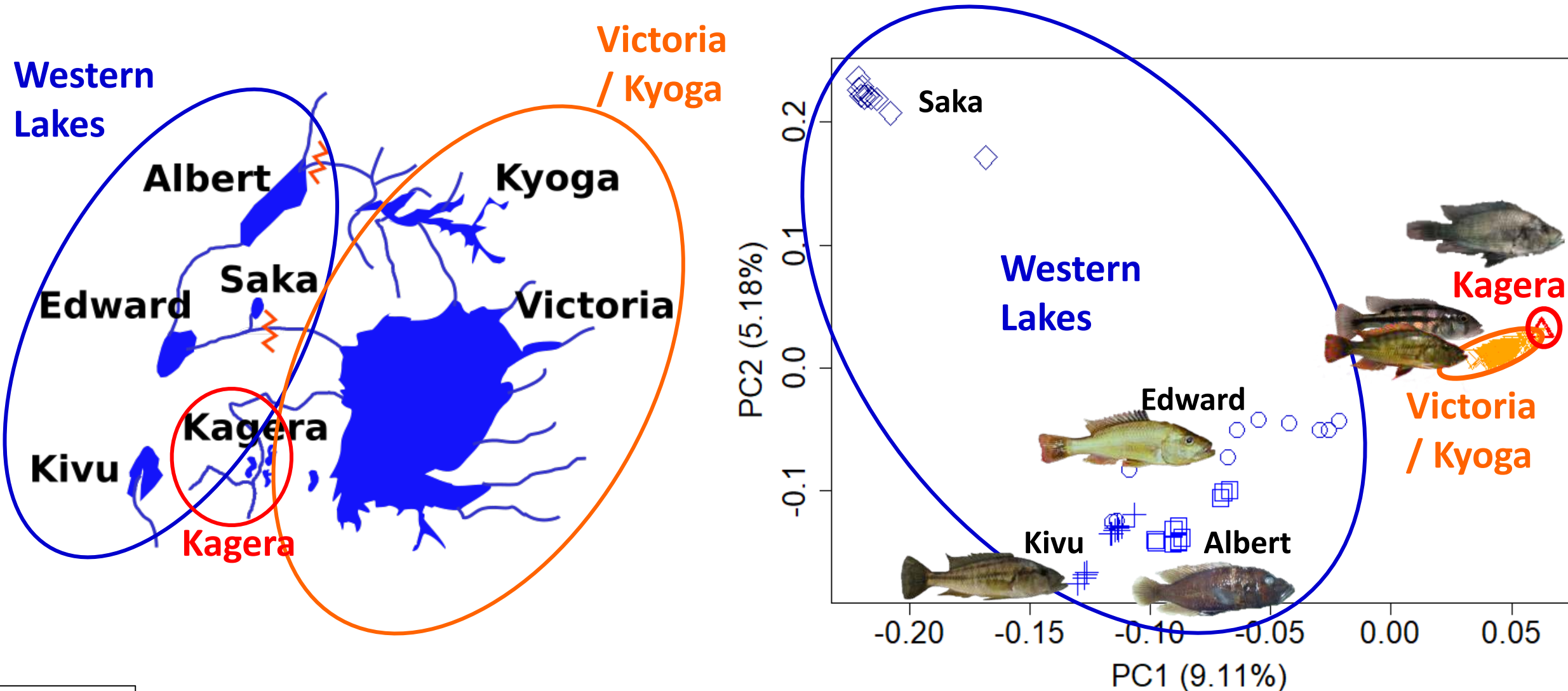
PCA on whole-genomes separates Lake Victoria/Kyoga cichlids from others

(152 genomes, 1.6M LD-pruned SNPs)



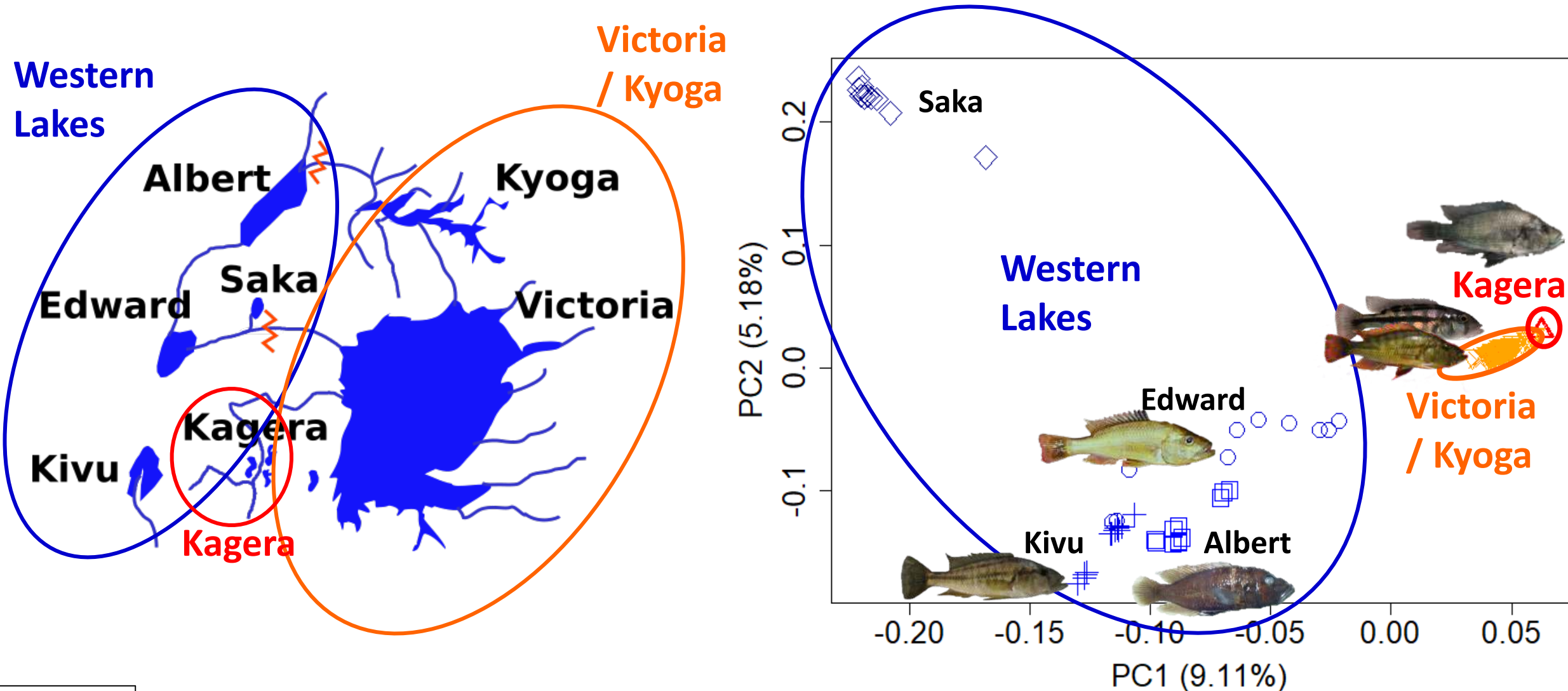
PCA on whole-genomes separates Lake Victoria/Kyoga cichlids from others

(152 genomes, 1.6M LD-pruned SNPs)

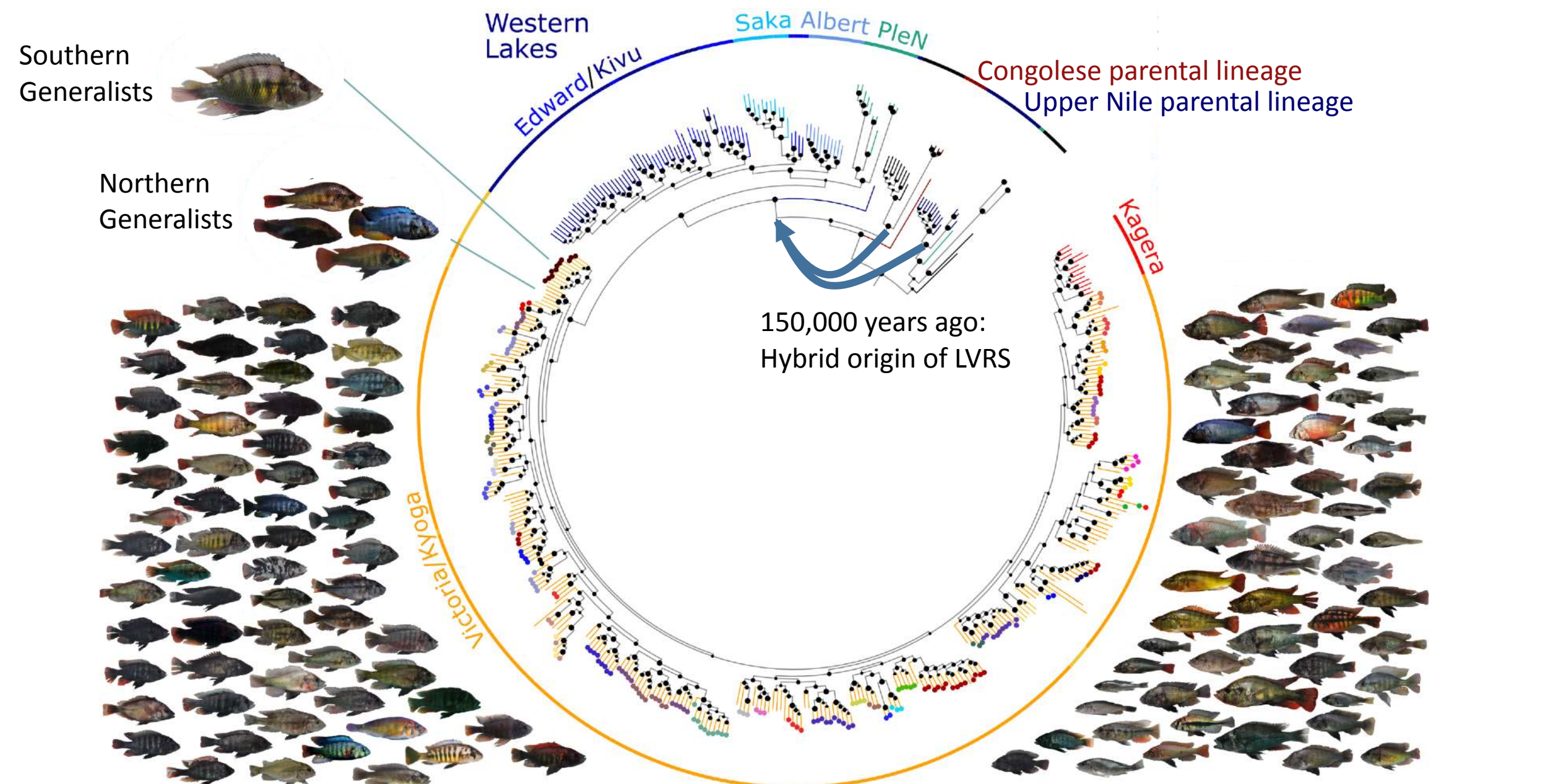


PCA on whole-genomes separates Lake Victoria/Kyoga cichlids from others

(152 genomes, 1.6M LD-pruned SNPs)

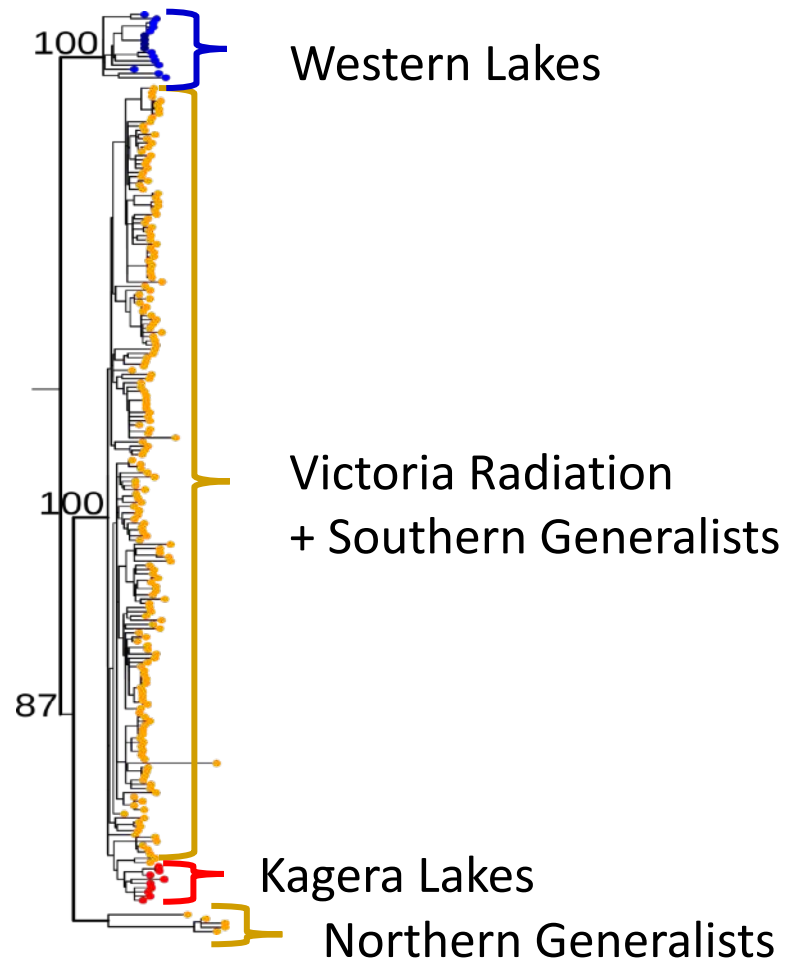


Phylogeny of 474 genomes reveals two lineages of generalists as closest relatives

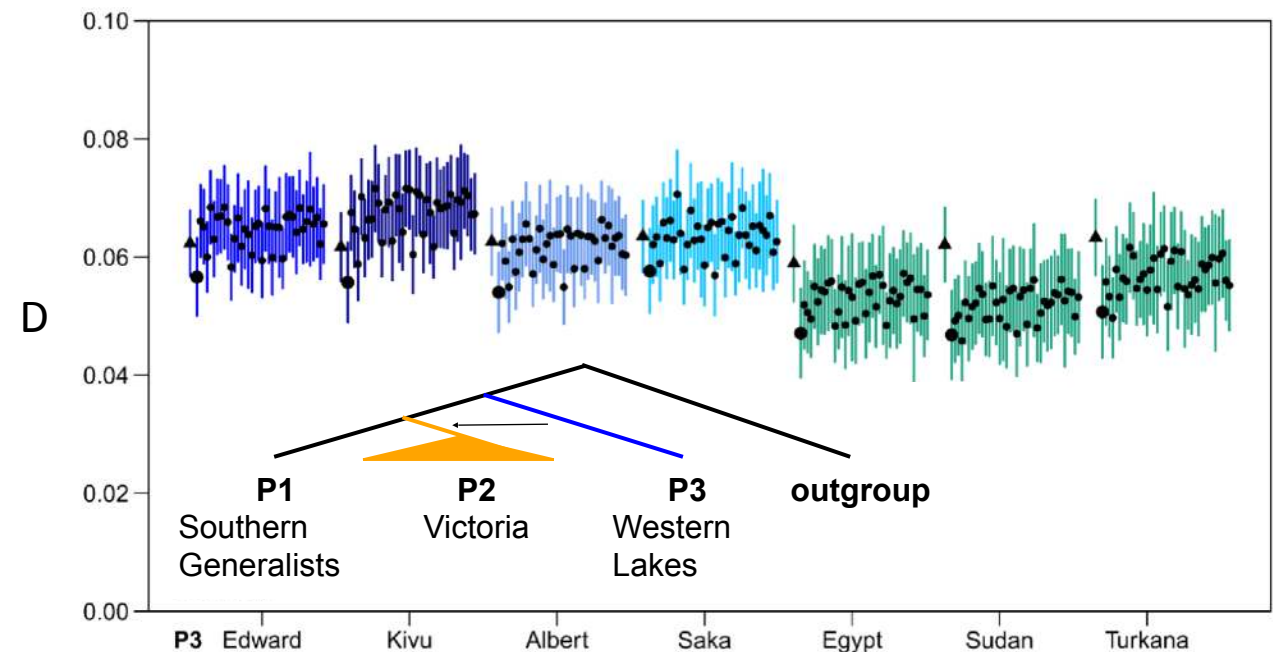


Multiple lines of evidence for another hybrid origin

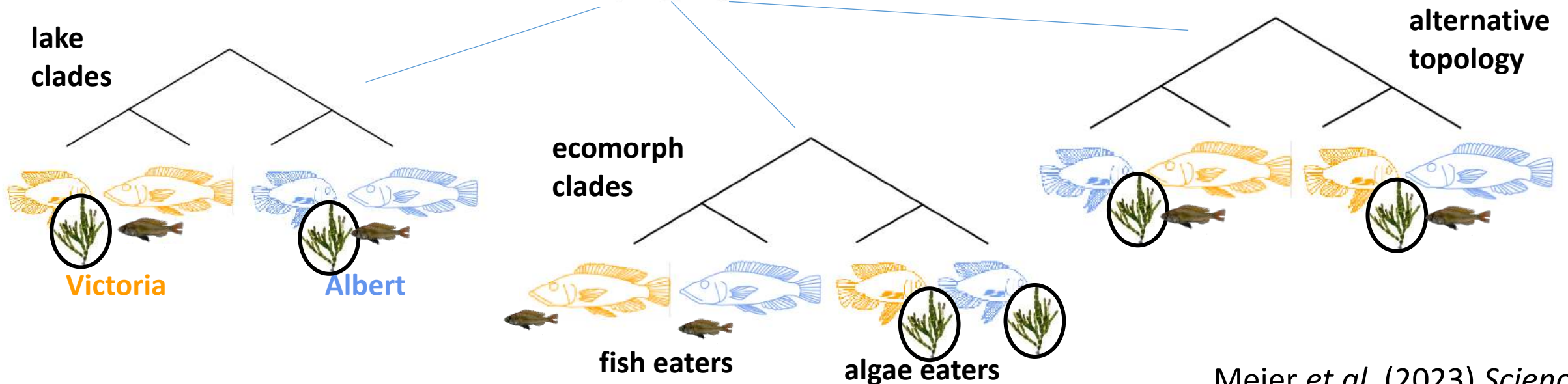
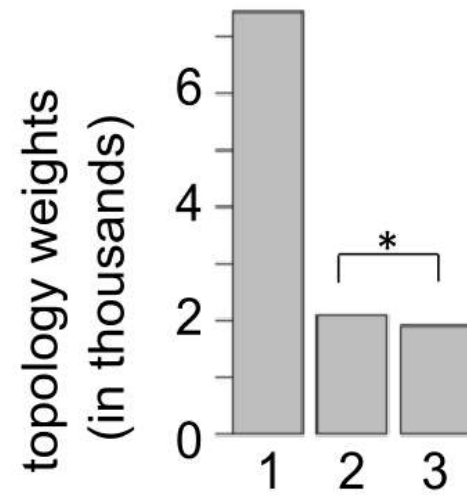
Deep mitochondrial divergence
not reflected in the nuclear phylogeny



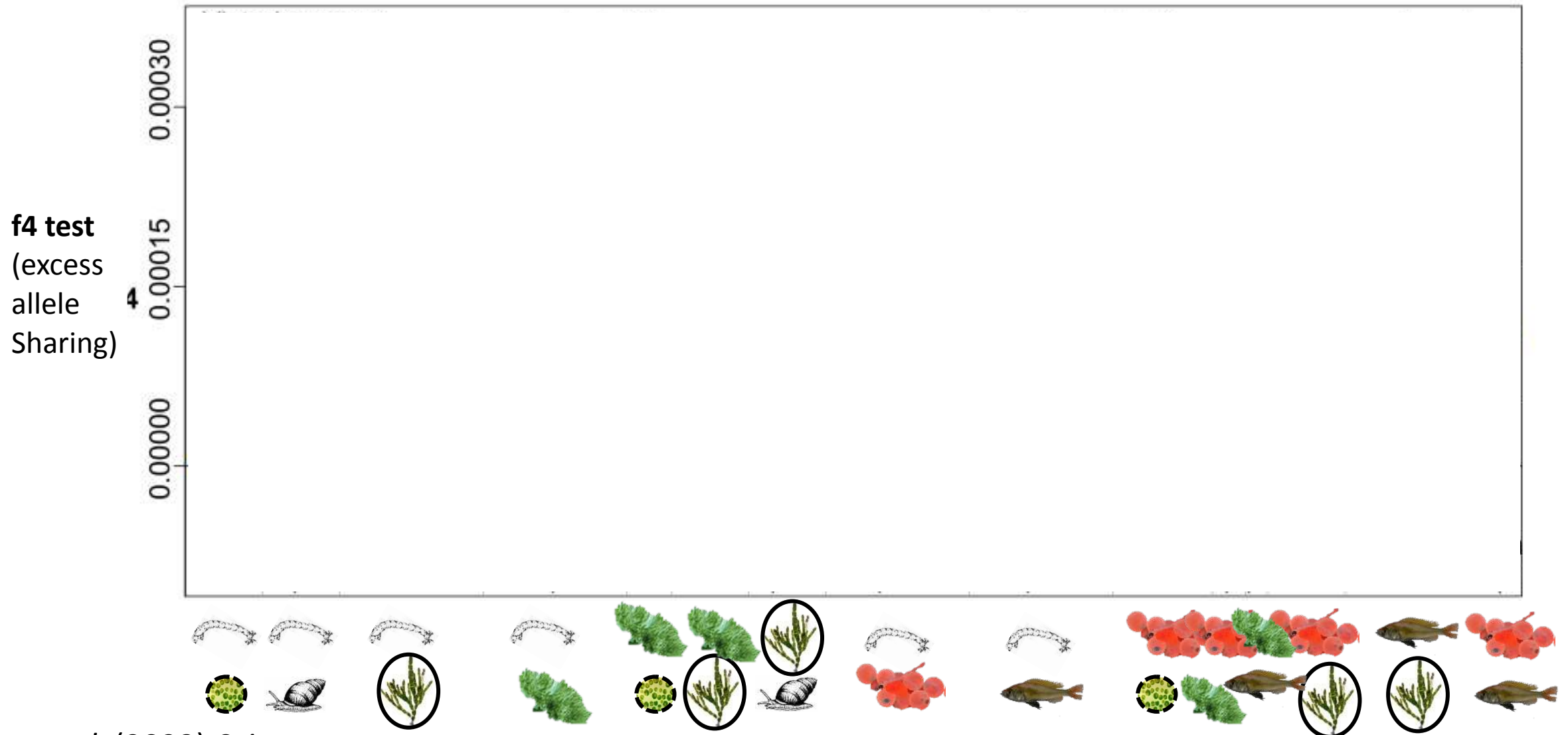
Tests of admixture reveal an additional
ancestry contribution of a relative of
the Western Lake radiations



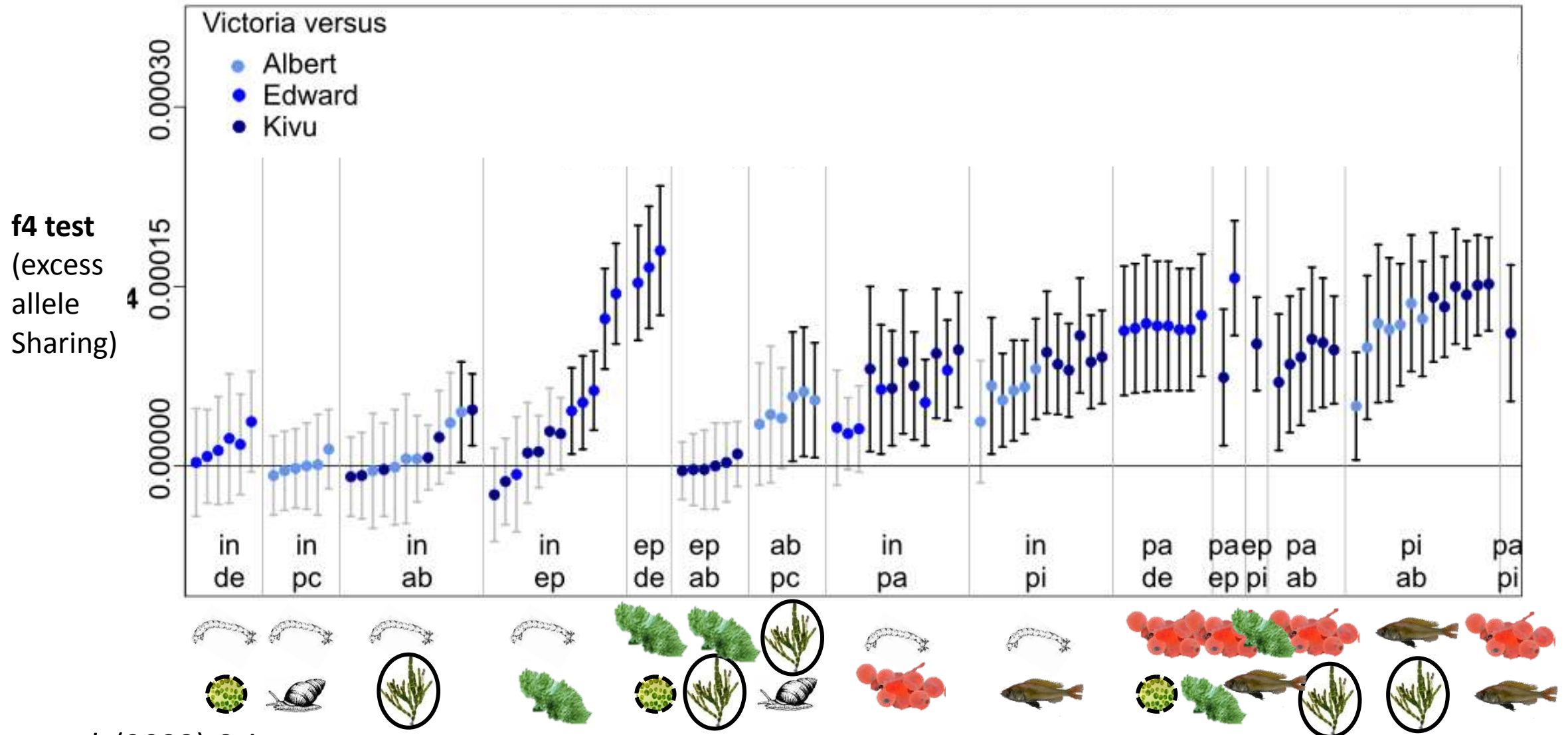
Slight excess allele sharing between ecologically similar species of different lake radiations



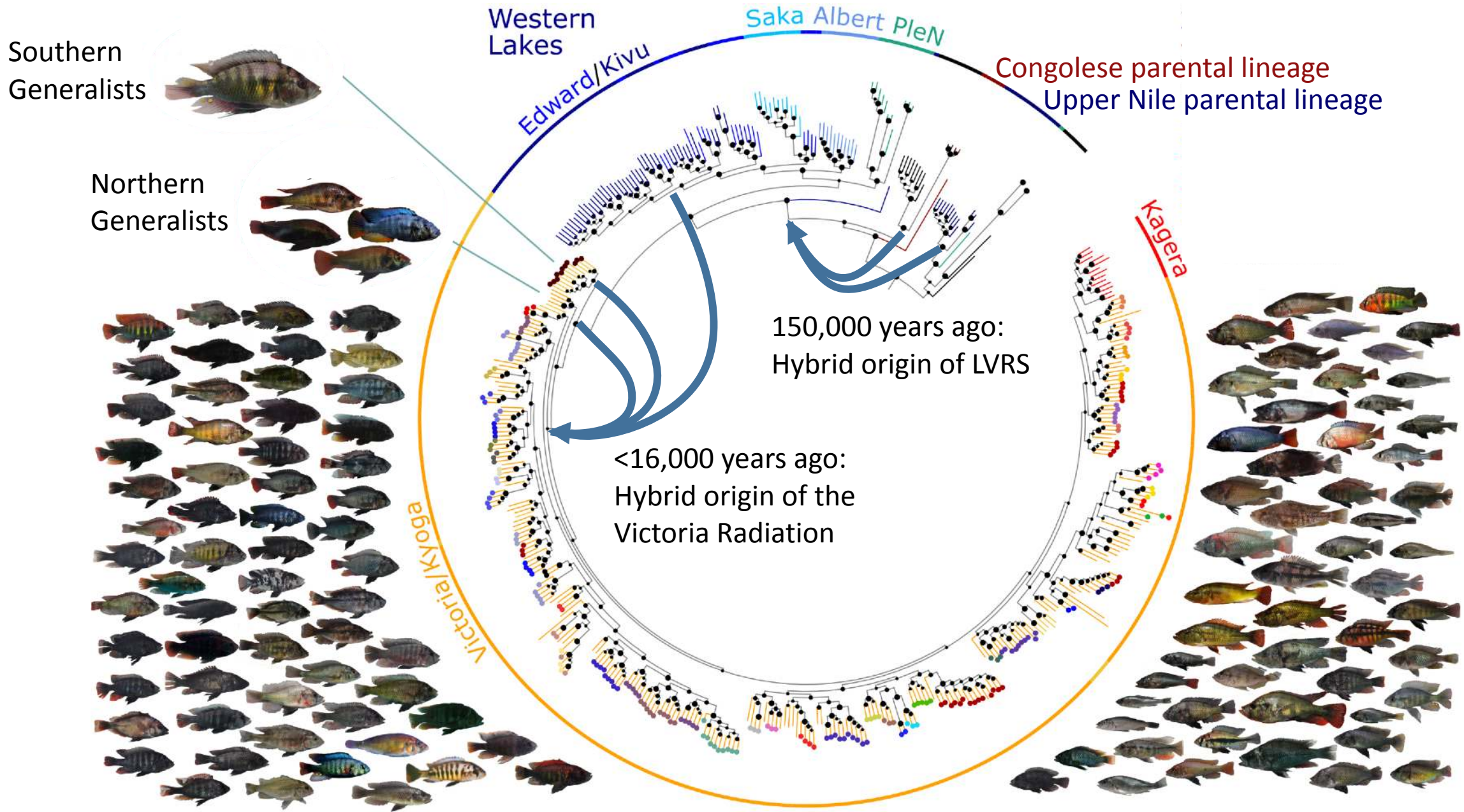
Victoria ecomorphs share alleles with the corresponding ecomorphs of the other radiations



Victoria ecomorphs share alleles with the corresponding ecomorphs of the other radiations



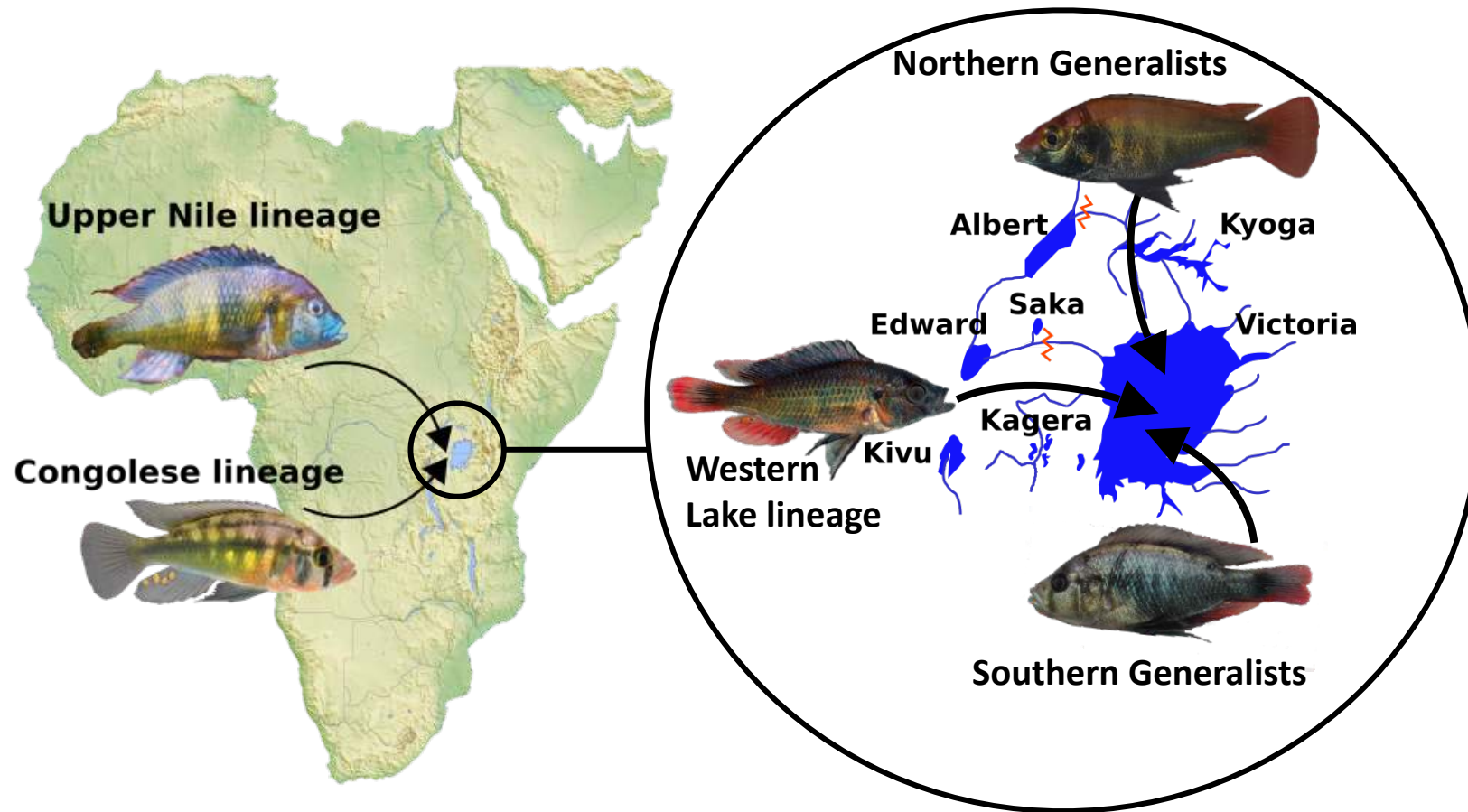
Hybrid origin at the base of the Victoria Radiation



Repeated cycles of admixture and diversification

1. level: ~150,000 years ago:
Hybrid origin of the entire LVRS

2. level: ~16,000 years ago:
Hybrid origin of the Victoria Radiation



Hybrid origin of pelagic dwarf predators

pelagic
dwarf predators



Hybrid origin of pelagic dwarf predators

large predators



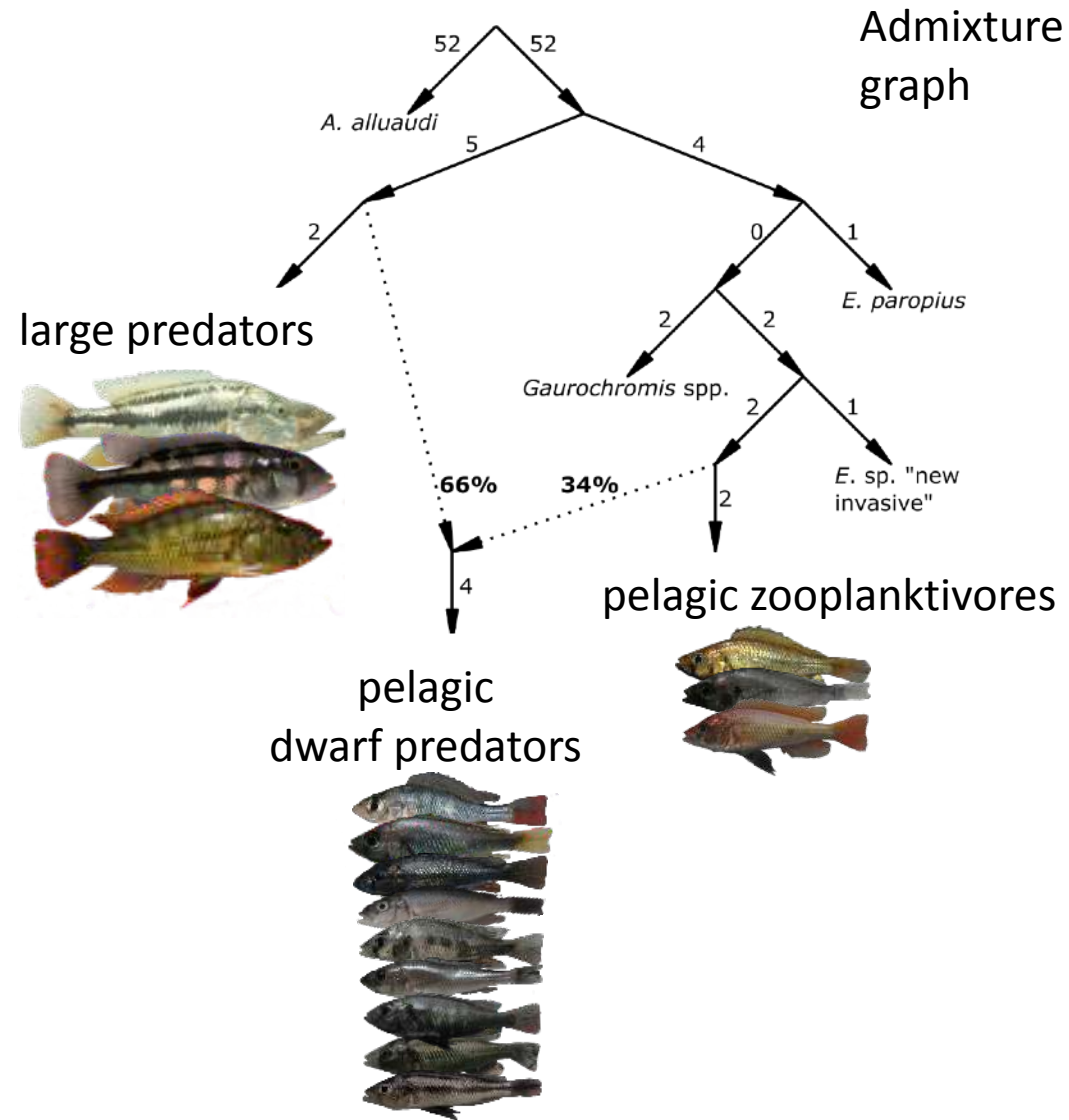
pelagic zooplanktivores



pelagic
dwarf predators

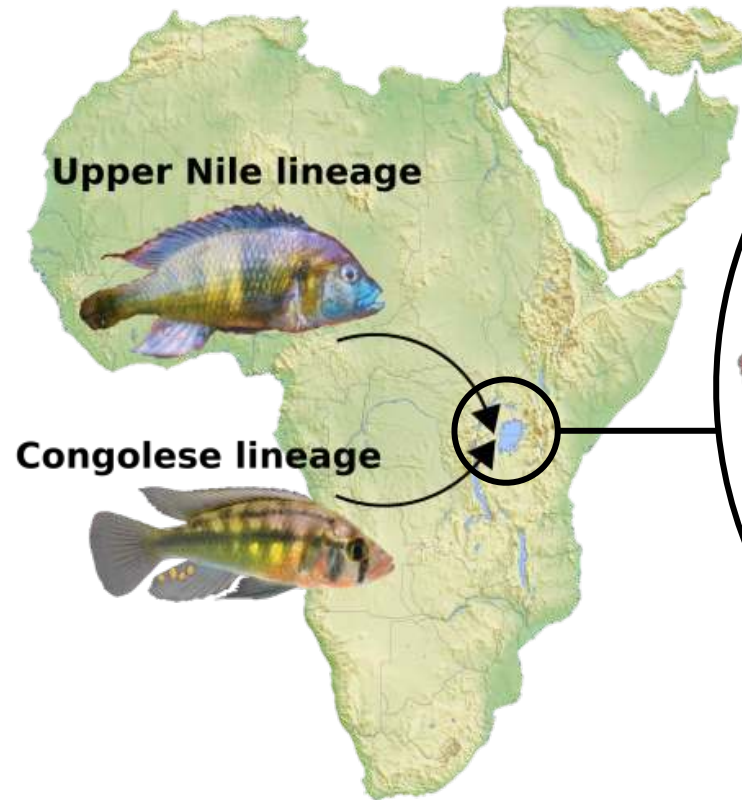


Hybrid origin of pelagic dwarf predators

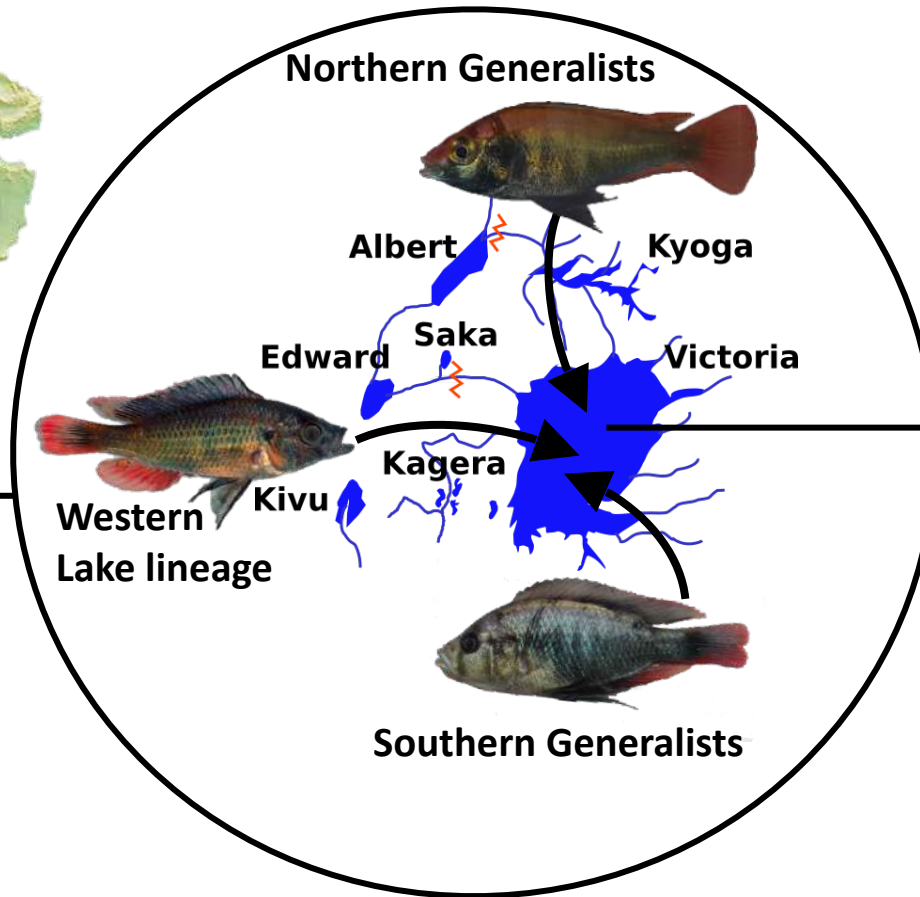


Repeated cycles of admixture and diversification

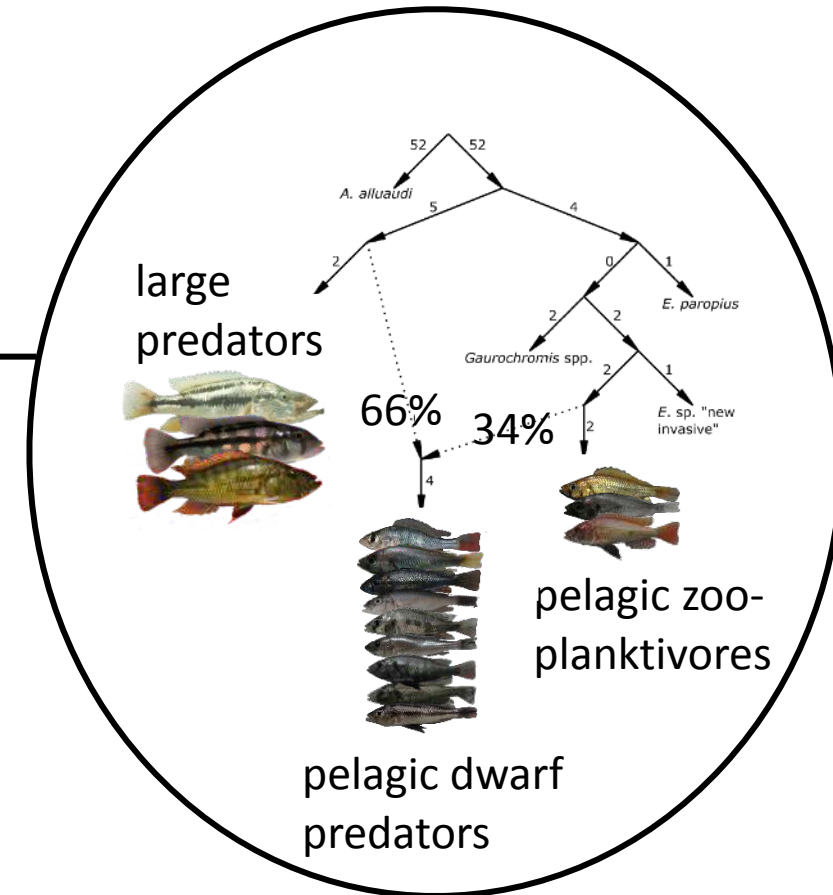
1. level: ~150,000 years ago:
Hybrid origin of the entire LVRS



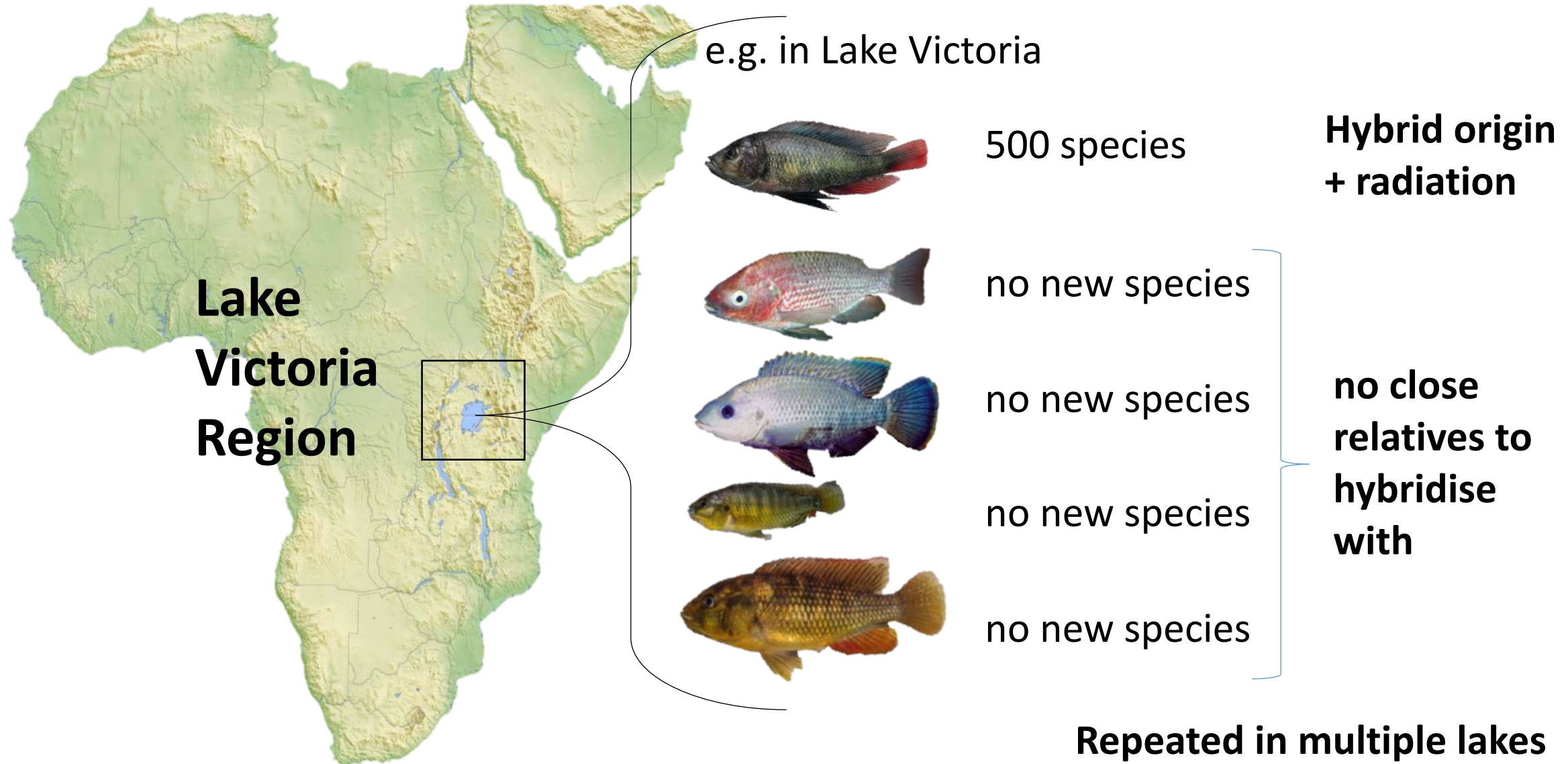
2. level: ~16,000 years ago:
Hybrid origin of the Victoria Radiation



3. level: <16,000 years ago:
Hybrid origin of dwarf predators
and other species



Did the other lineages not hybridize?



Did the other lineages not hybridize?



**Lake
Victoria
Region**

e.g. in Lake Victoria



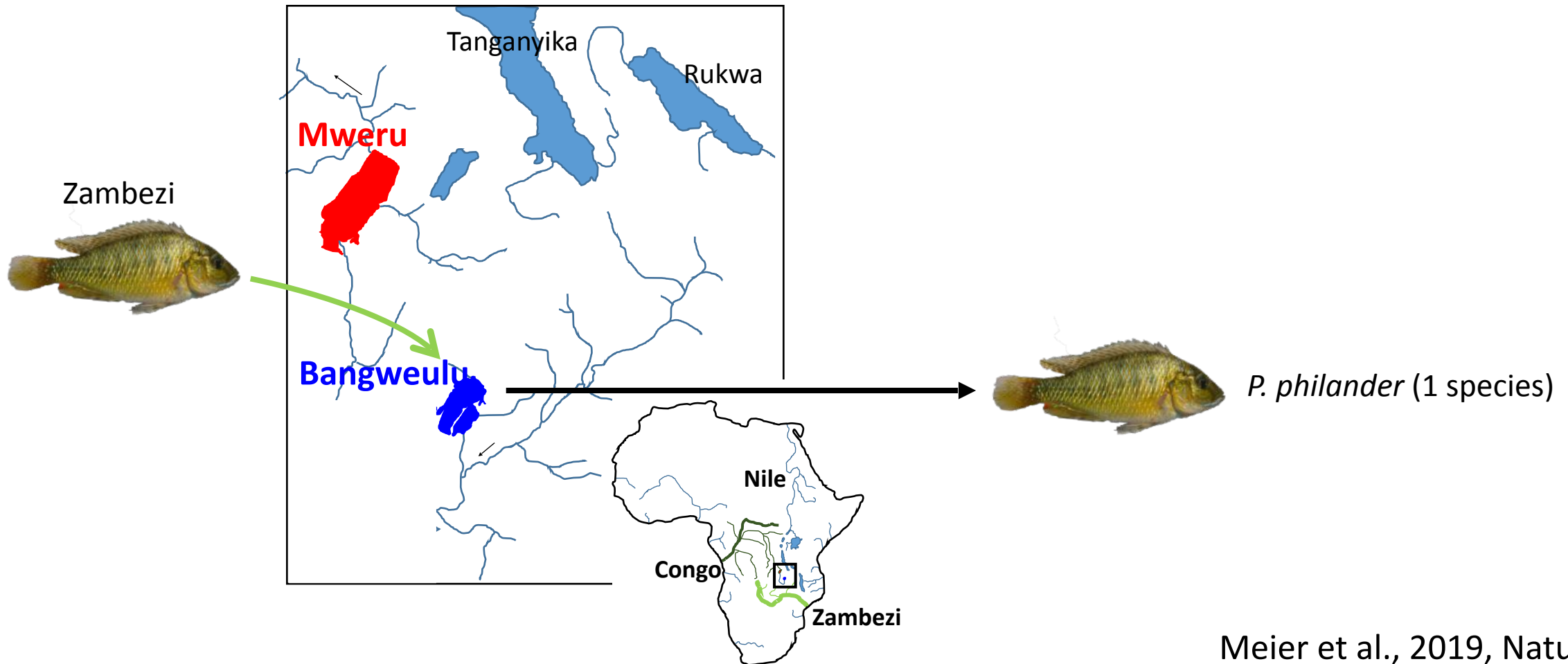
500 species (LVRS) **Hybrid origin
+ radiation**

Pseudocrenilabrus

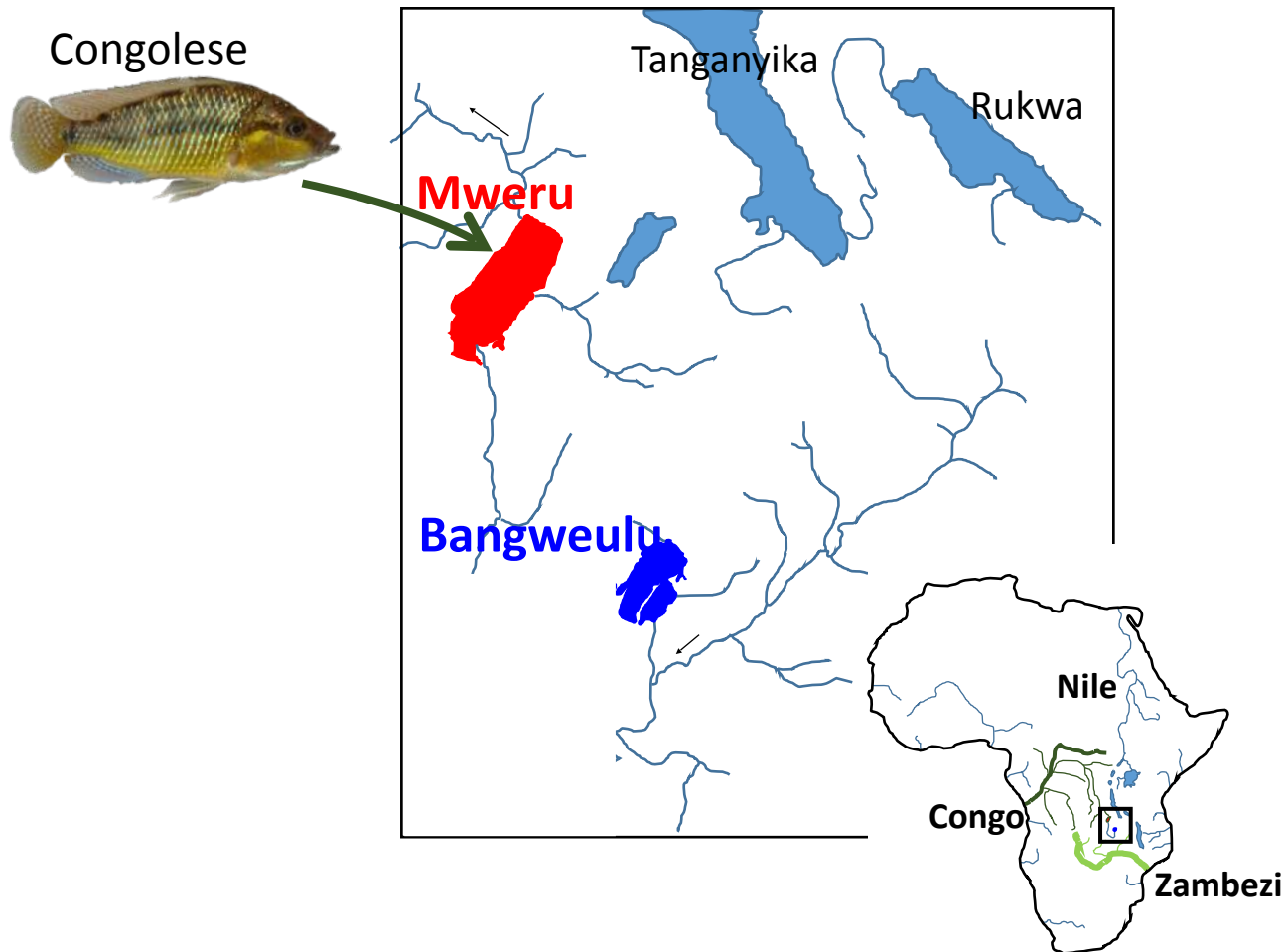


no new species
in 5 lakes

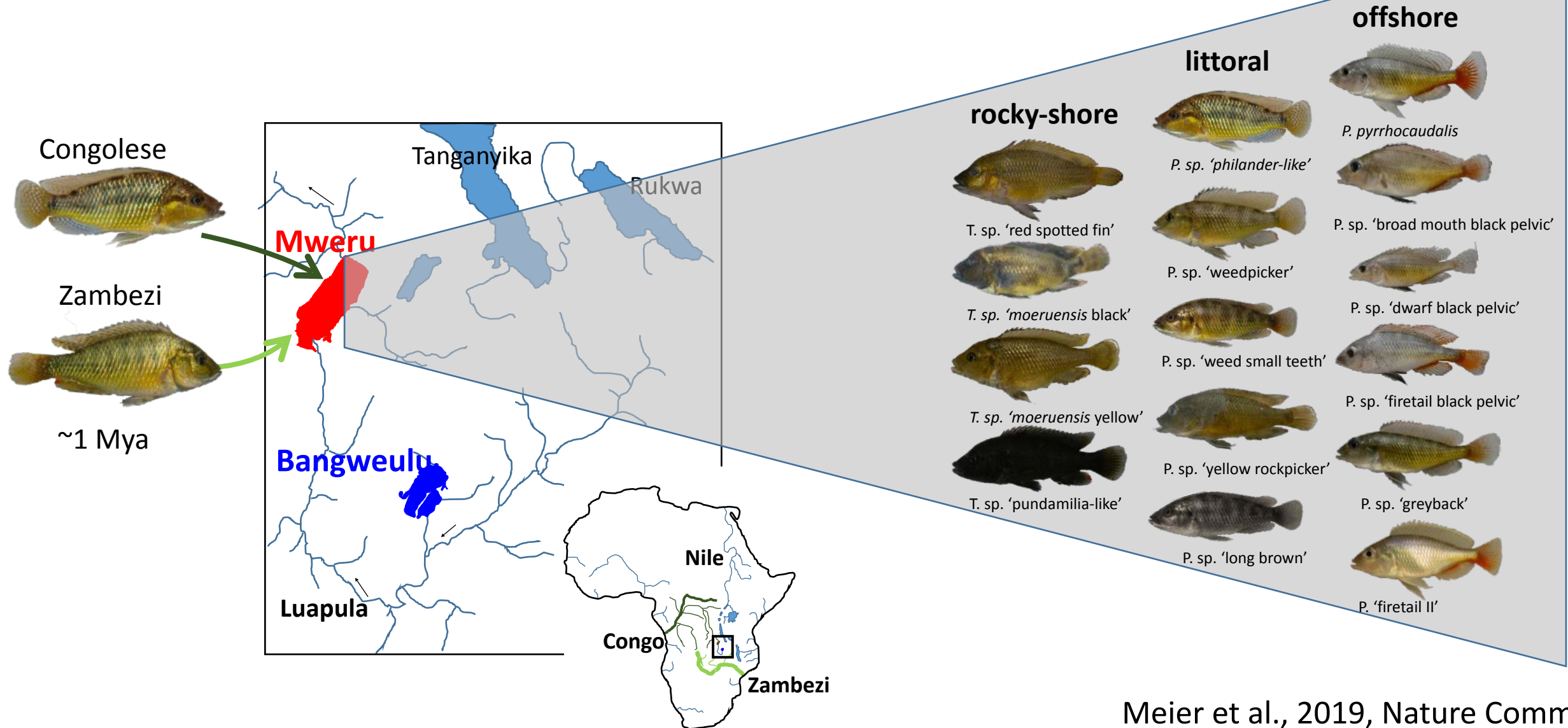
Lake Bangweulu contains a single generalistic Zambezi *Pseudocrenilabrus* species



Lake Mweru was colonized by Zambezian and Congolese *Pseudocrenilabrus* lineages and radiated into 15 species



Lake Mweru was colonized by Zambezian and Congolese *Pseudocrenilabrus* lineages and radiated into 15 species



In Lake Mweru ~40 new species evolved from hybrid origin between Zambezian and Congolese lineages

Congolese lineages

Pseudocrenilabrus

Serranochromis
lineage

Orthochromis
kalungwishiensis

Orthochromis stormsi

Orthochromis sp.

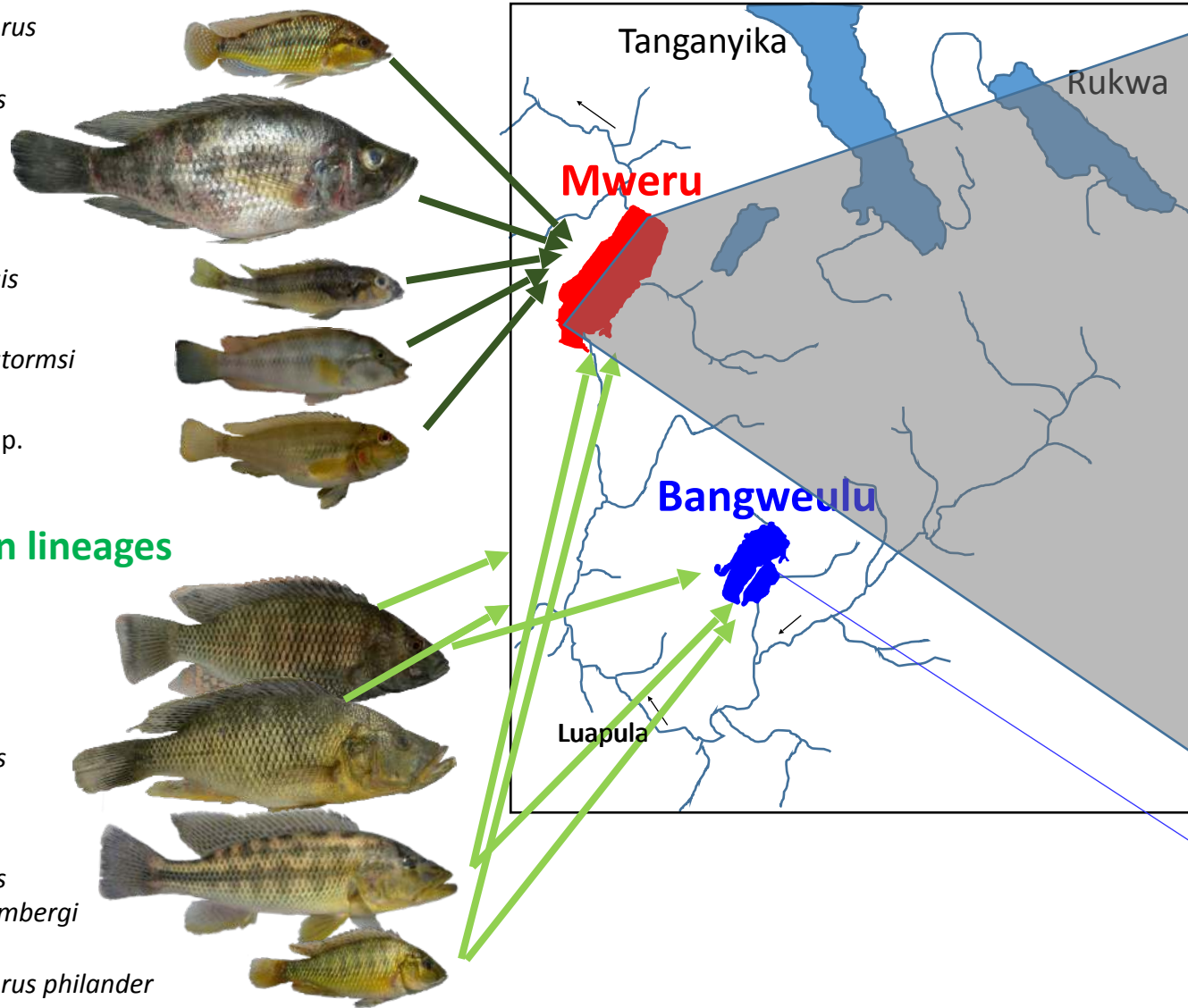
Zambezan lineages

Sargochromis
mellandi

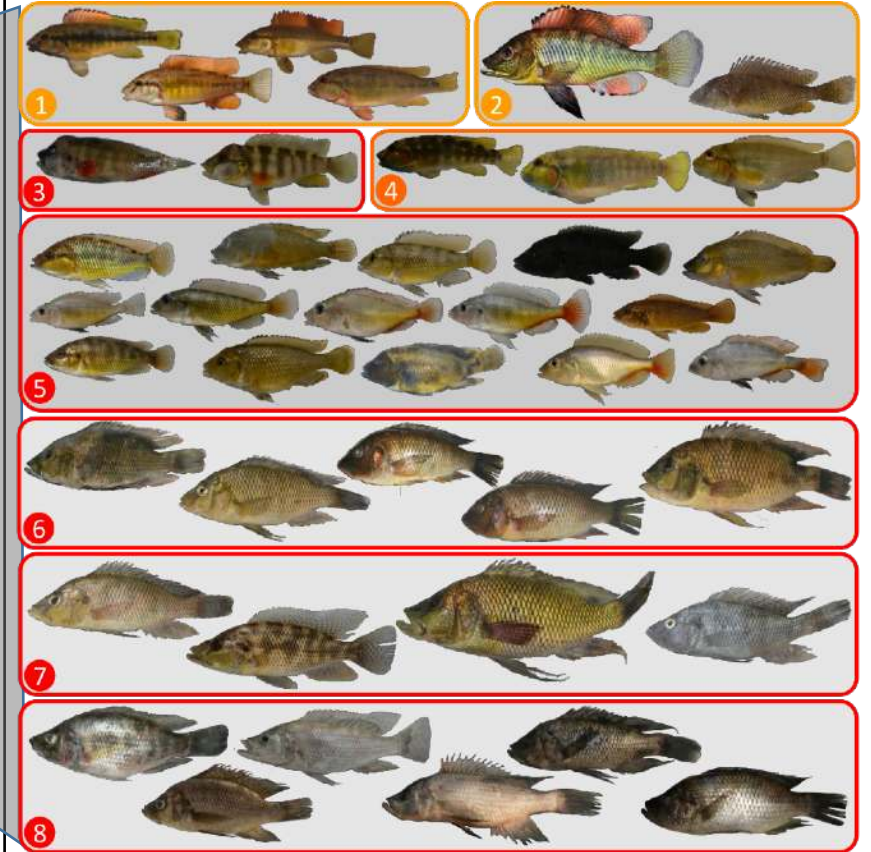
Serranochromis
altus &
angusticeps

Serranochromis
robustus & *thumbergi*

Pseudocrenilabrus philander



~40 species evolved in 8 radiations



no adaptive radiation

Slow evolution through new mutations



Boost of diversity through hybridisation

