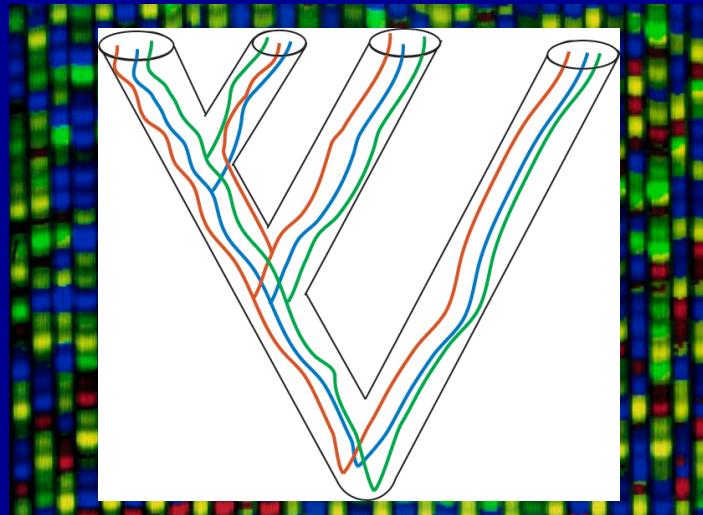
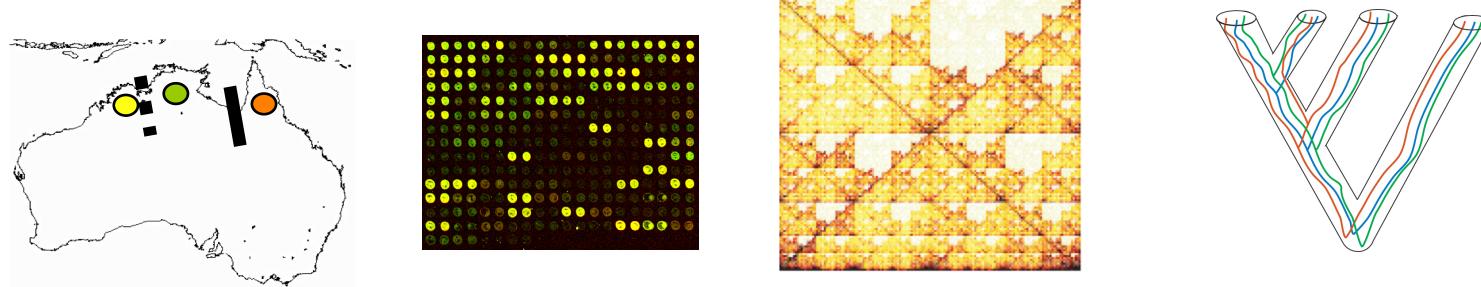


The continuum between phylogeography and phylogenetics



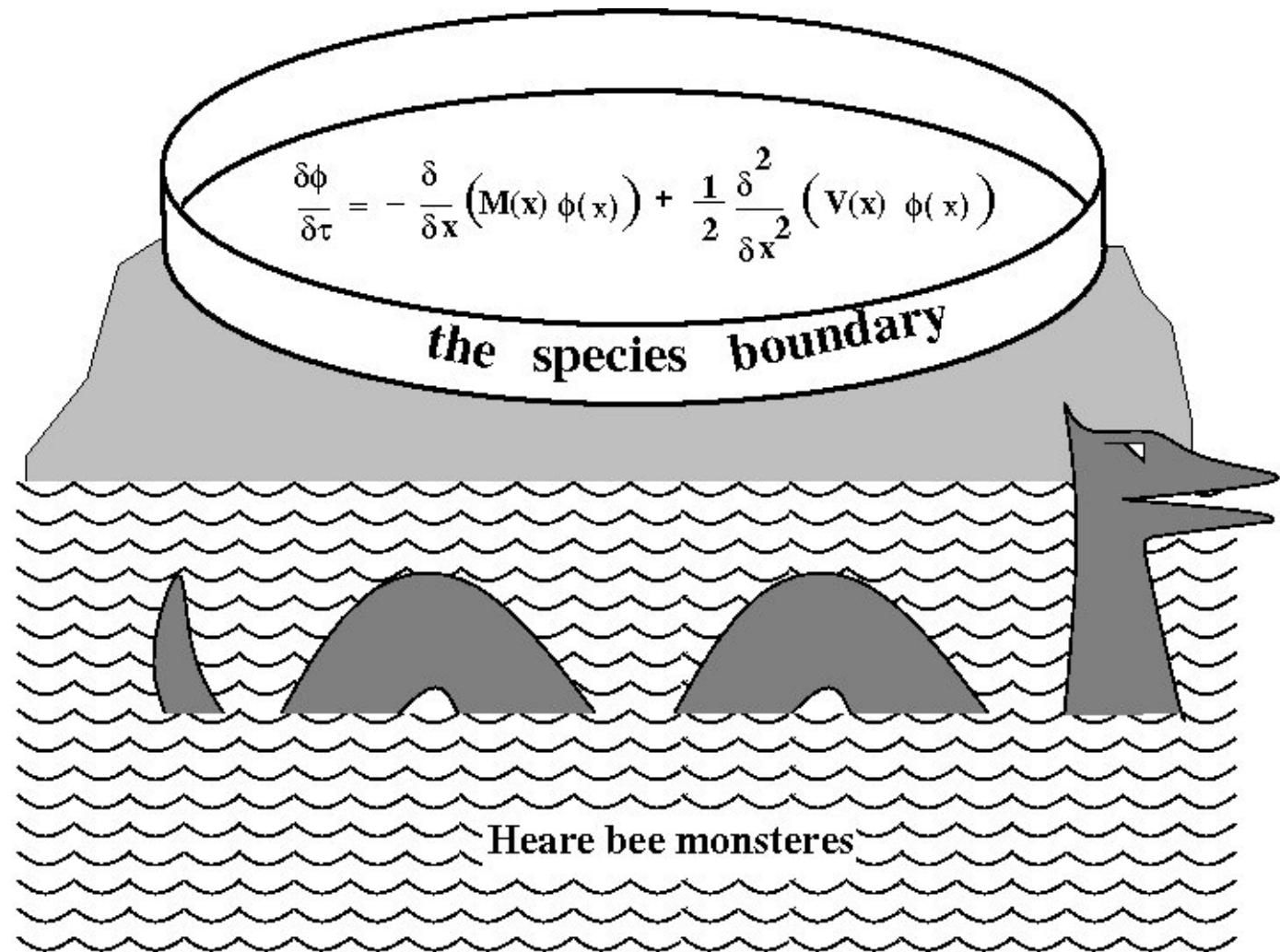
Scott V. Edwards
Department of Organismic and Evolutionary Biology
Harvard University
Cambridge, MA USA
<https://edwards.oeb.harvard.edu/>

Edwards lab – birds, genomics and evolution



March 2018

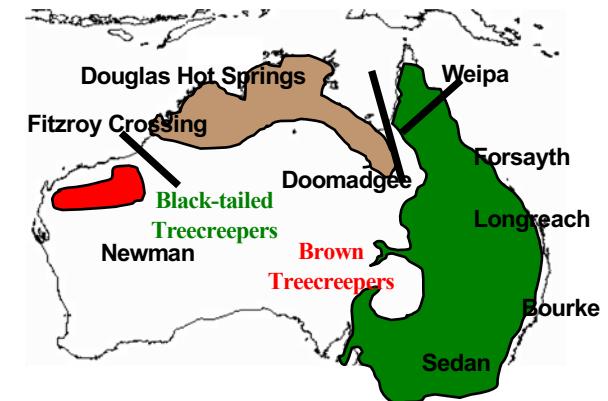
A 1980s conundrum in evolutionary biology



Talk overview

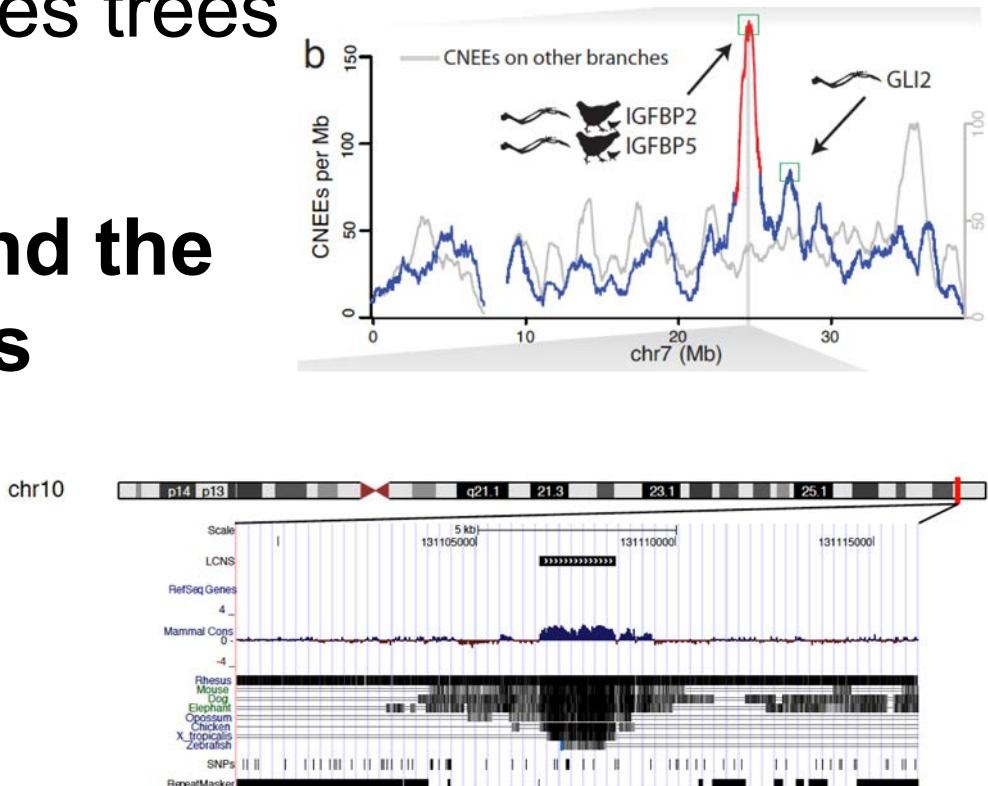
Part I: The emerging continuum between phylogeography and phylogenetics

- Multilocus analyses of gene flow, phylogeography and species trees



Part II: Macroevolution and the origin of phenotypic traits

- Regulatory or protein-coding innovation?
- Genomics of convergence



Case studies



Babblers



Honeyeaters



Flycatchers



Treecreepers

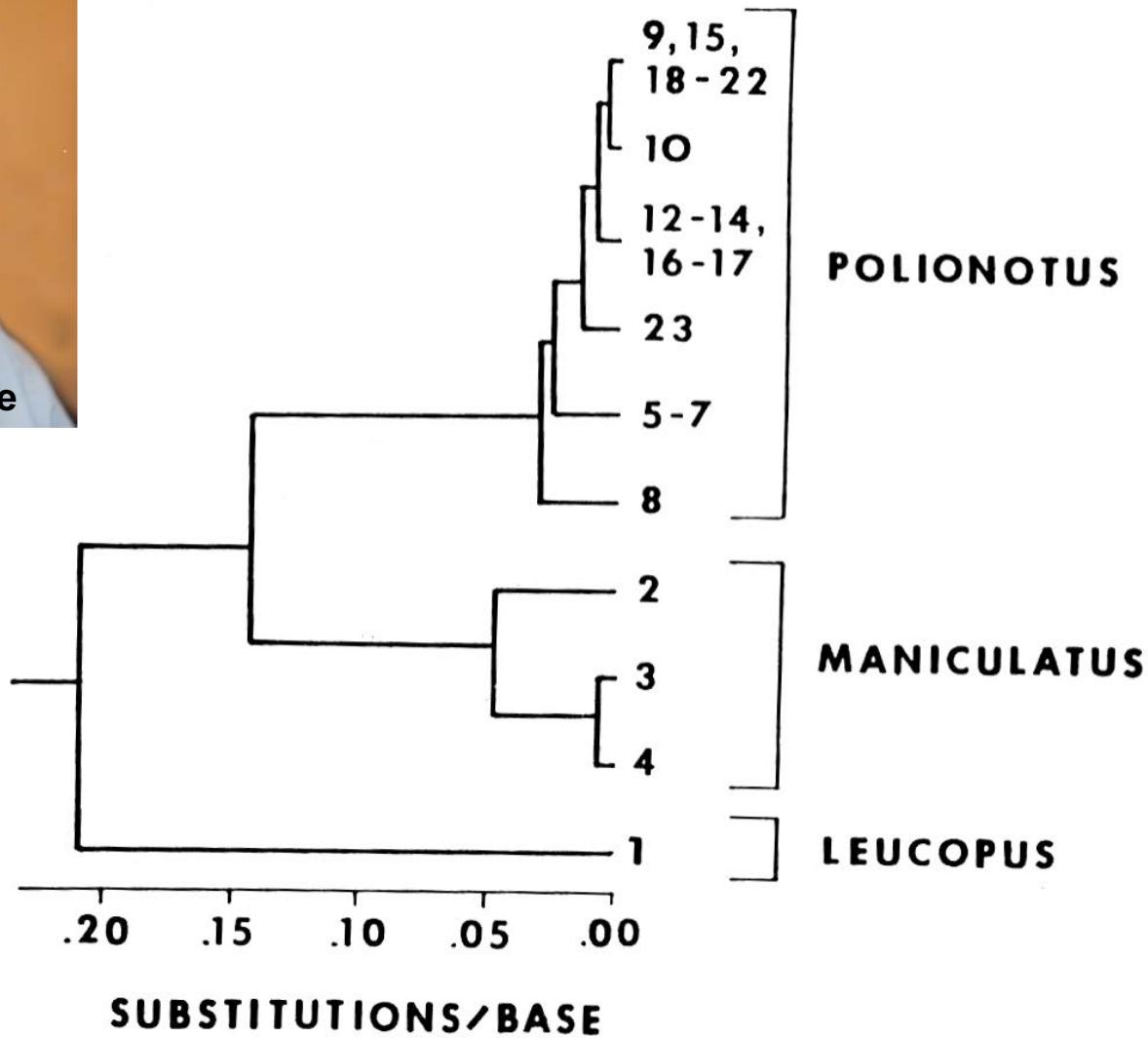


Fairy Wrens

The first ‘gene tree’, 1979



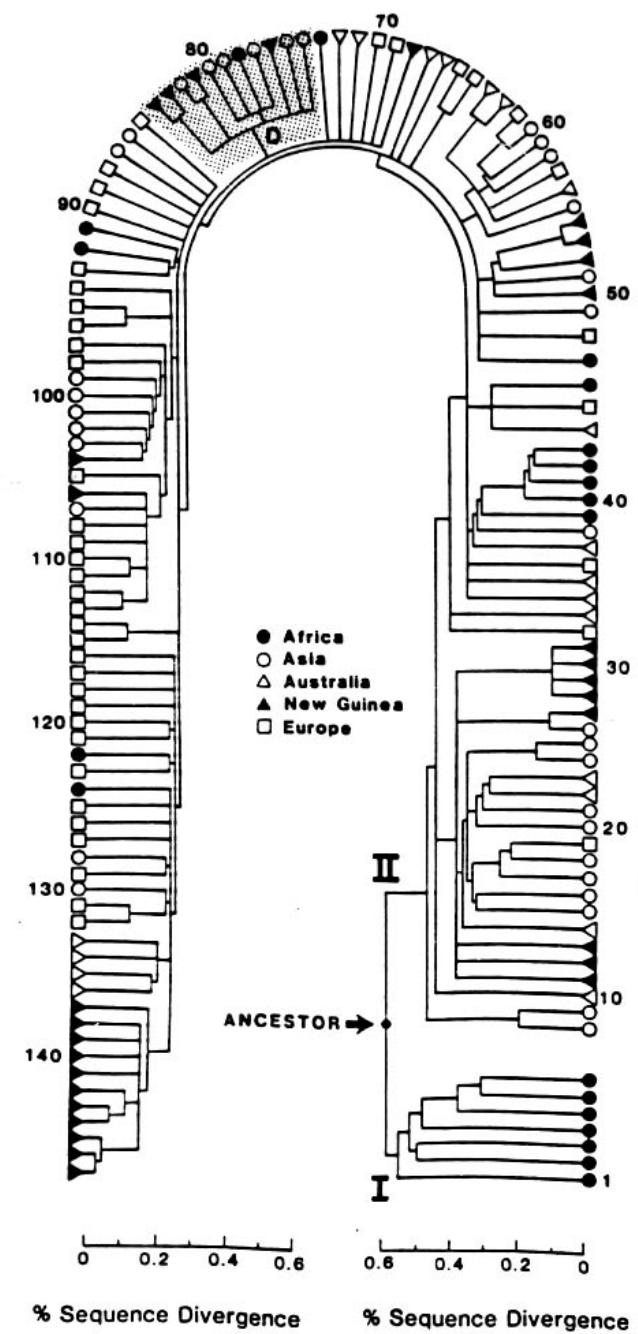
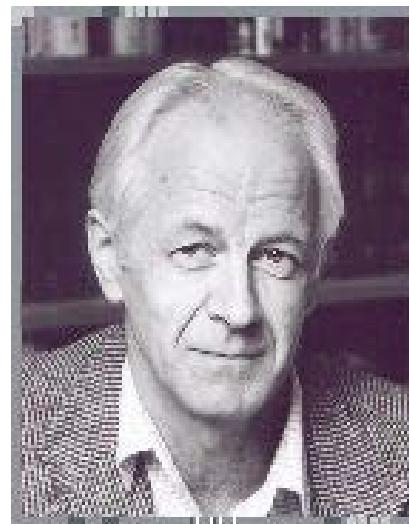
J. C. AVISE, R. A. LANSMAN AND R. O. SHADE



Mitochondrial DNA and human evolution

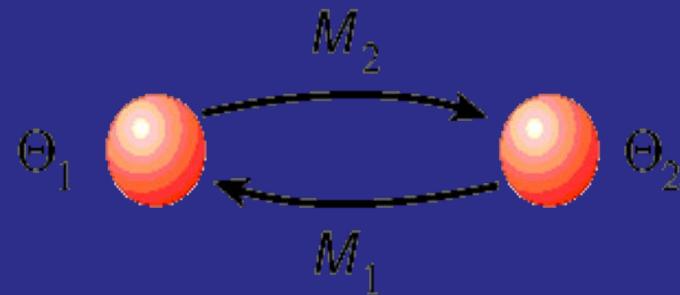
Rebecca L. Cann*, Mark Stoneking & Allan C. Wilson

Department of Biochemistry, University of California, Berkeley, California 94720, USA



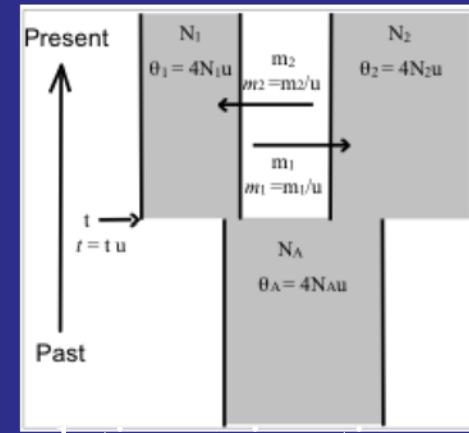
Multilocus models in phylogeography

Population size: $\theta = 4N\mu$ Divergence time: $\tau = \mu t$ Gene flow: $M = m/\mu$



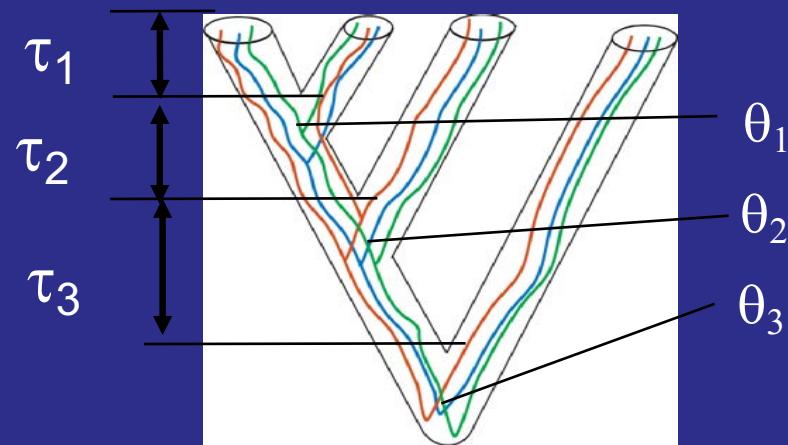
equilibrium migration model

MIGRATE: Beerli 2006 *Bioinformatics*



isolation-migration model

IM: Hey and Nielsen 2004 *Genetics*

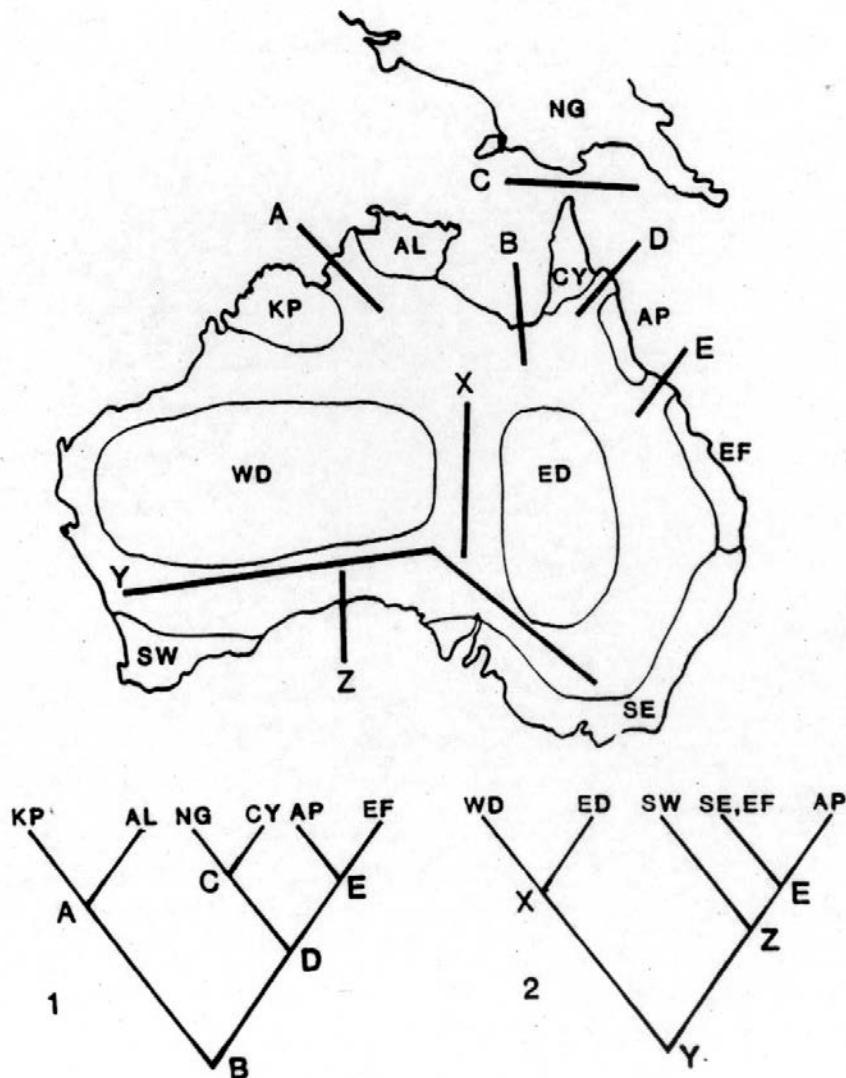


pure isolation
(phylogeny)

BEST: Liu and Pearl. 2007. *Syst. Biol*

BEAST*: Heled and Drummond. 2010. *Mol. Biol. Evol.*

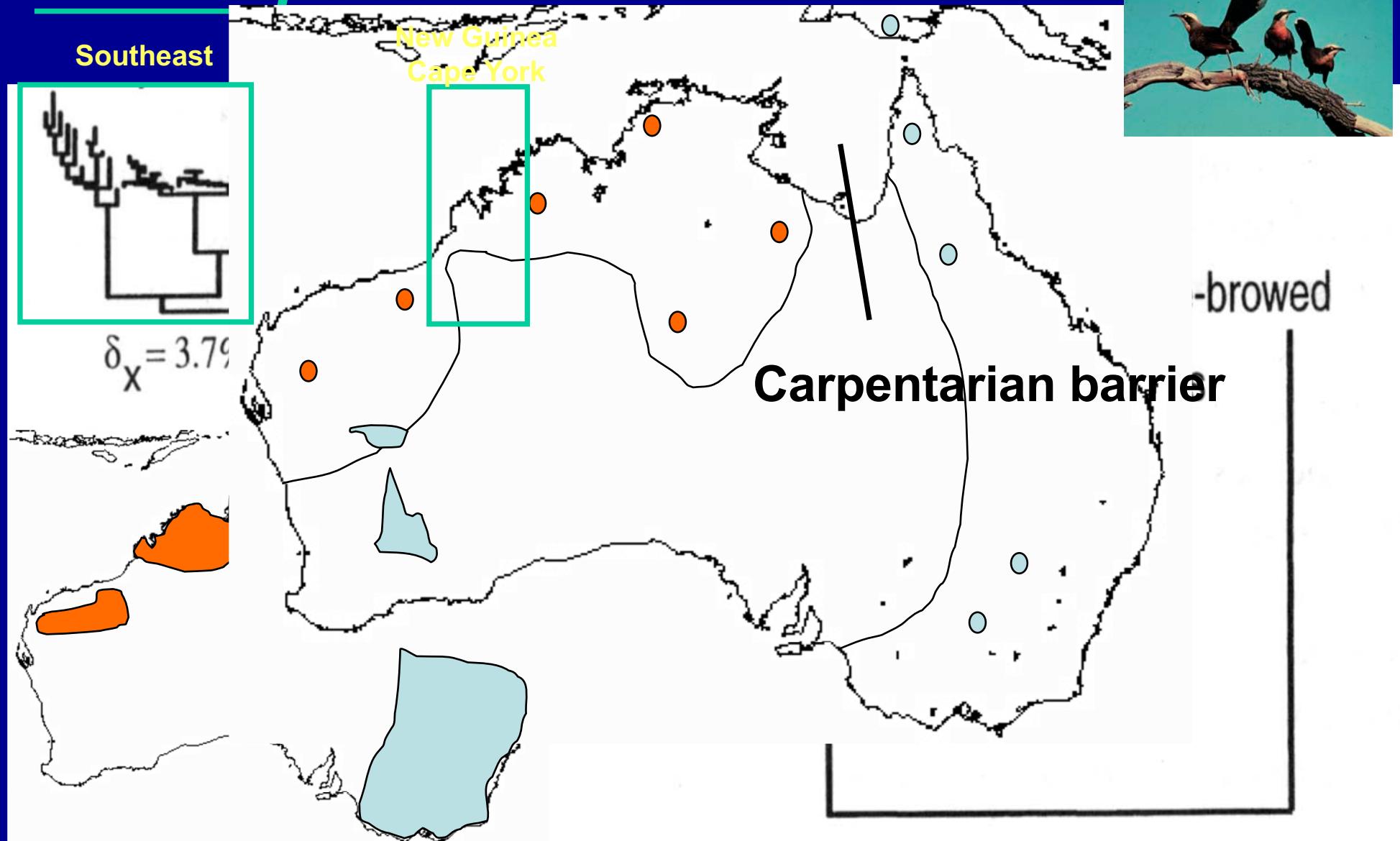
Carpentarian barrier (B) is deepest split in area cladograms of Australian biota



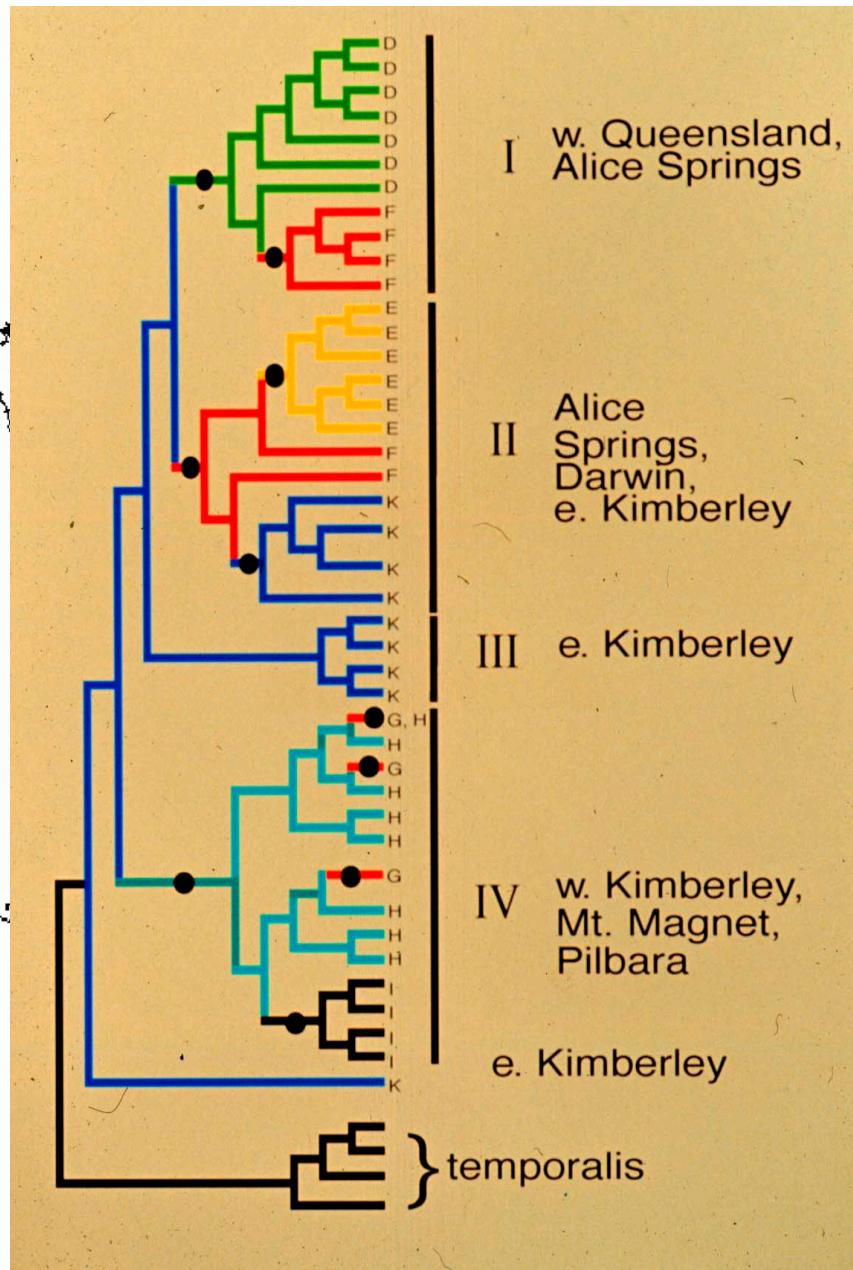
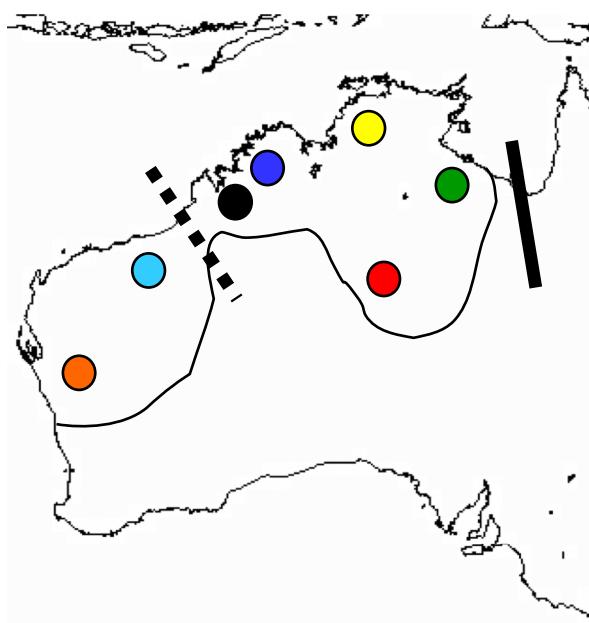
Warm welcome in the outback...



Mitochondrial genetic ties of orange-eared babbles *temporalis* *rubeculus*



Gene flow erodes population monophyly



● Migration event
reconstructed by
parsimony

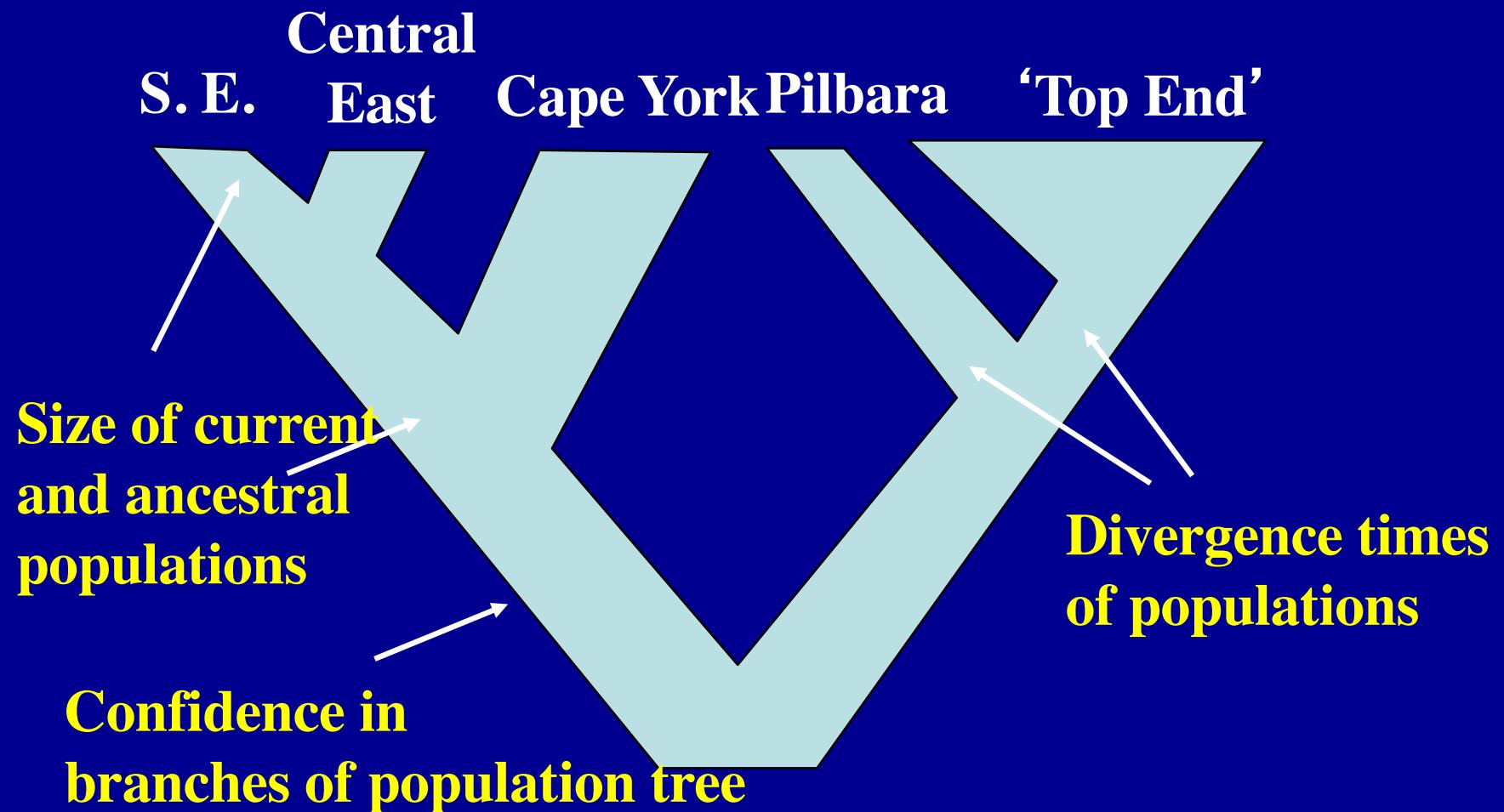
Bayesian approach: Advantages for estimating gene flow

- Parsimony approach
 - Assumes complete certainty of gene tree
 - Statistical testing of different hypotheses of gene flow is cumbersome
 - Confidence limits reflect only uncertainty in s
- ML approach (MIGRATE)
 - Integrates the two sources of uncertainty
 - Accommodates uncertainty in trees by integrating over all trees
 - Accommodates the range of Nm given these trees
 - Hypothesis testing easier



Photo: Graeme Chapman

Hypothetical species tree for babblers



Treecreepers (*Climacteris*)

Black-tailed treecreeper

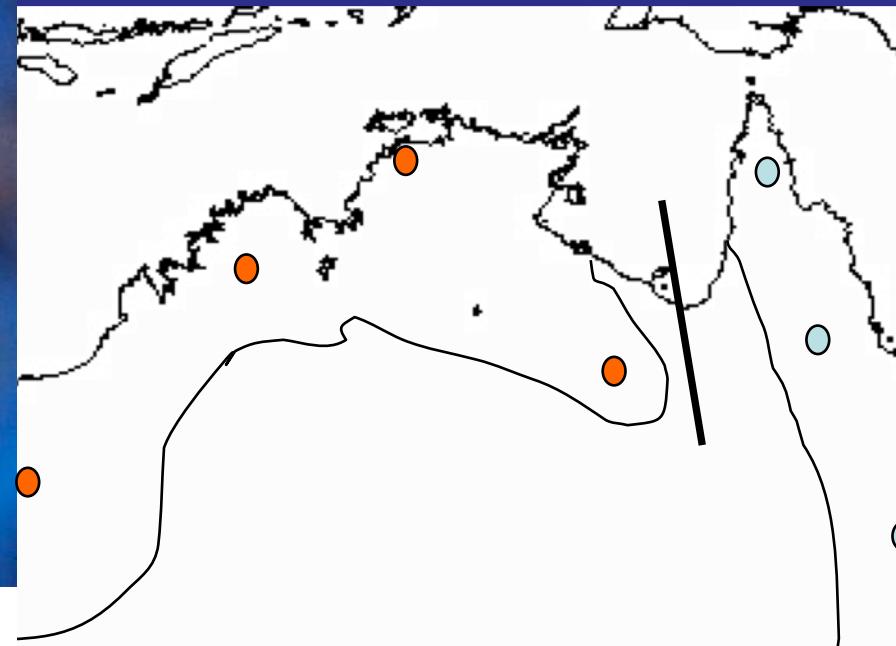


C. melanura

Brown treecreeper



C. picumnus



12 gene trees in Australian treecreepers

Brown Treecreeper
(eastern) lineages

Black-tailed (western)
Treecreeper lineages

AL3

AL5

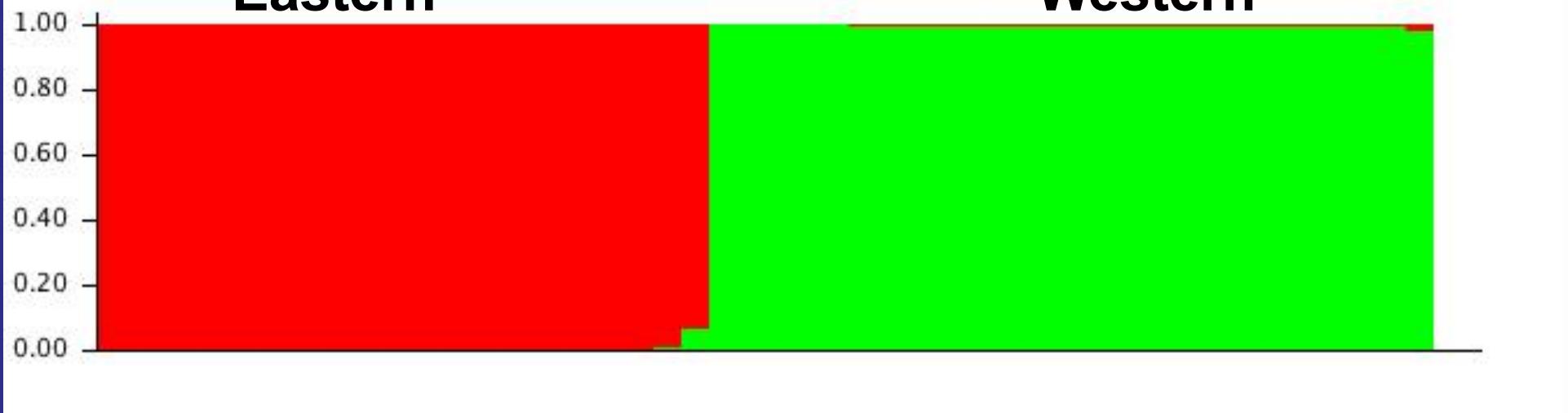
AL7

AL14

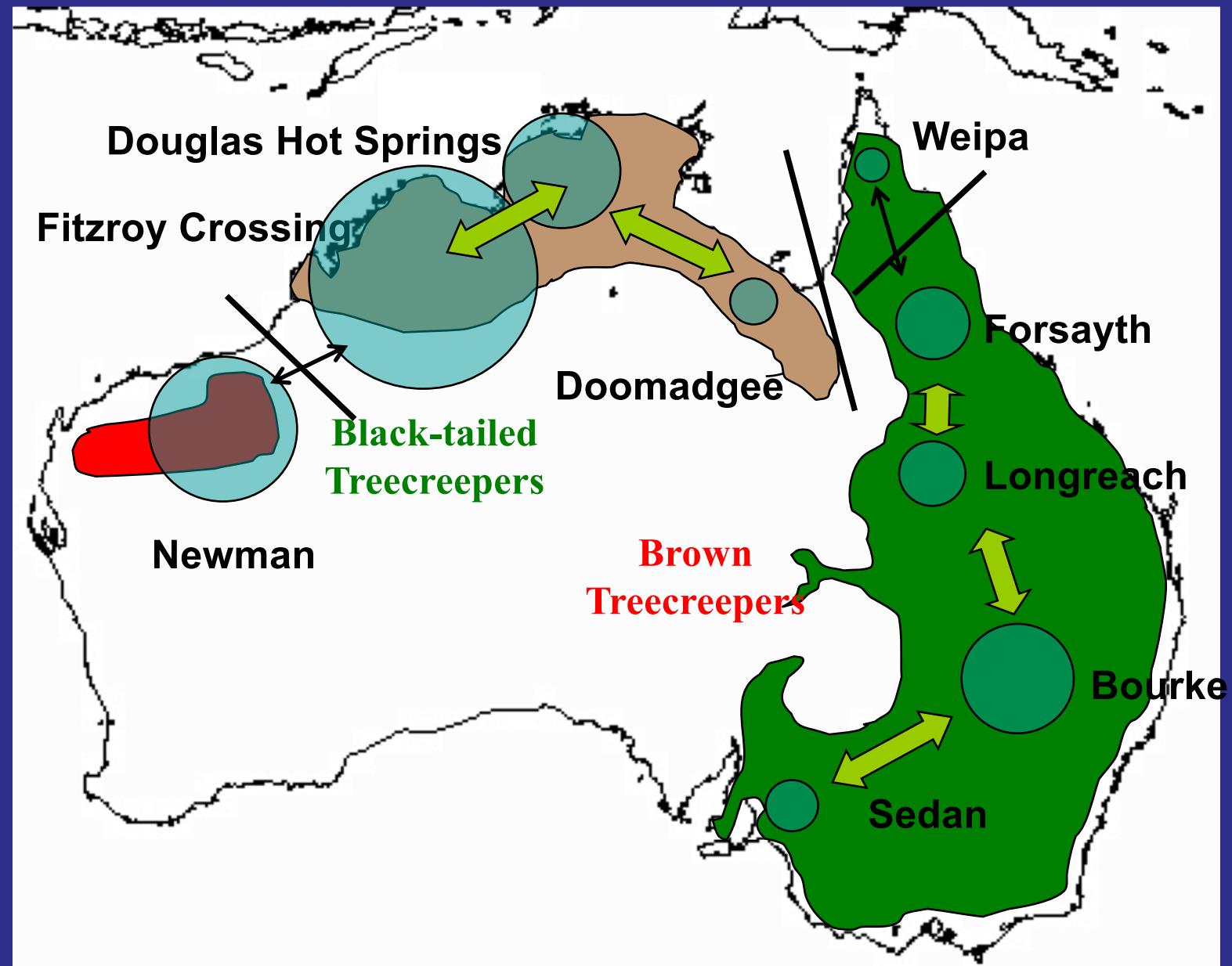
AL16

Eastern

Western

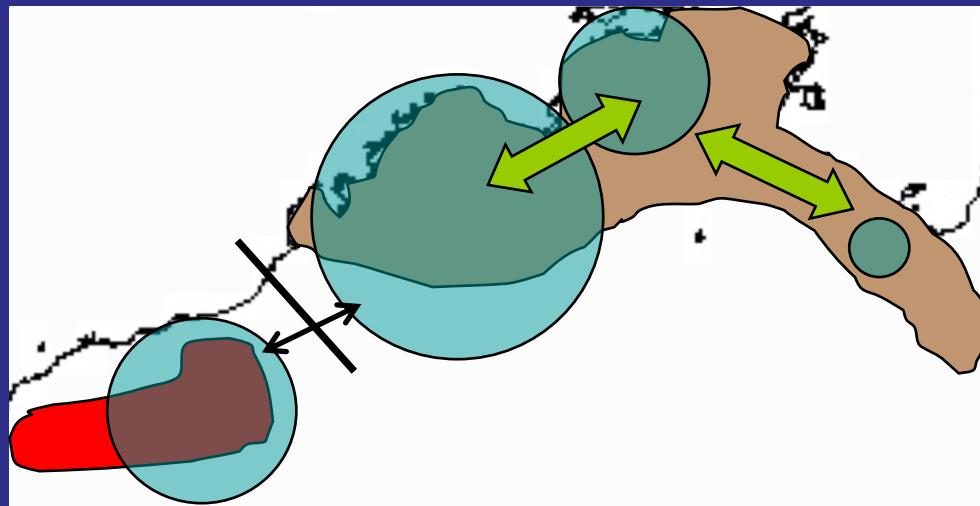


Treecreeper populations are connected but variable in size (MIGRATE)



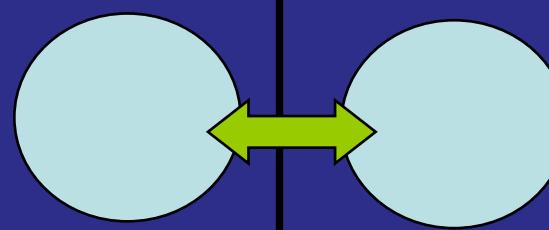
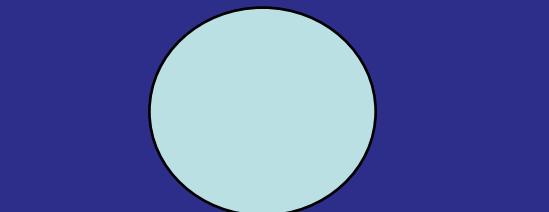
Rotzel, Edwards and Beerli, unpubl. data

Using Bayes Factors to evaluate phylogeographic models

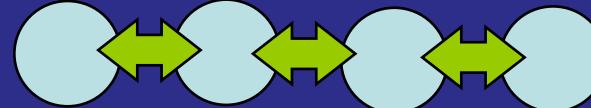


$2(mL_1 - mL_2)$

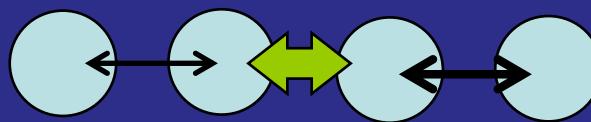
-98.60



-83.14



-54.98

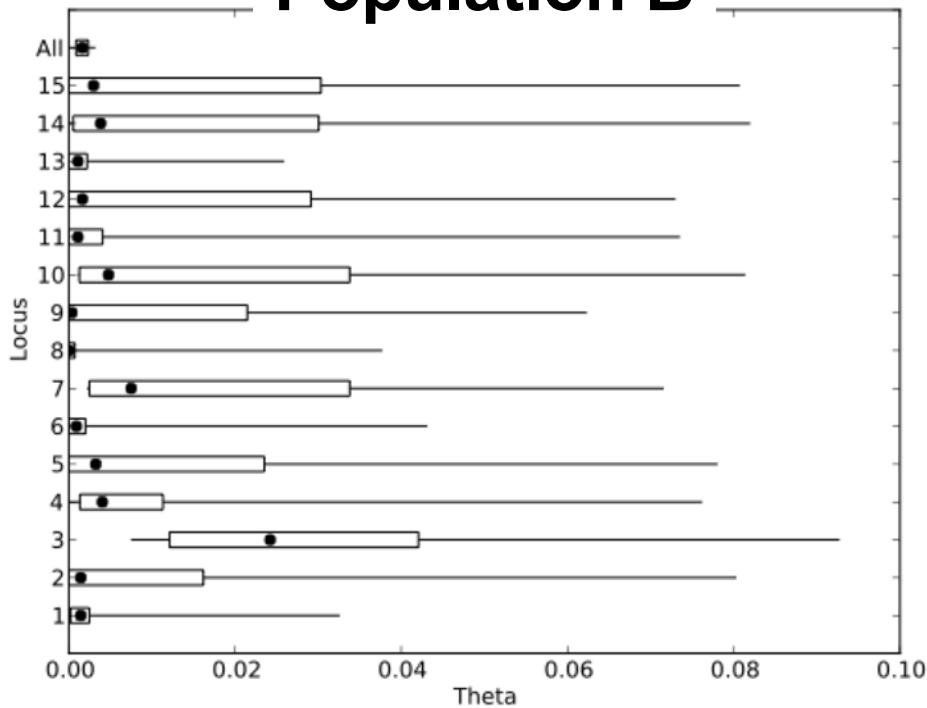


0.00

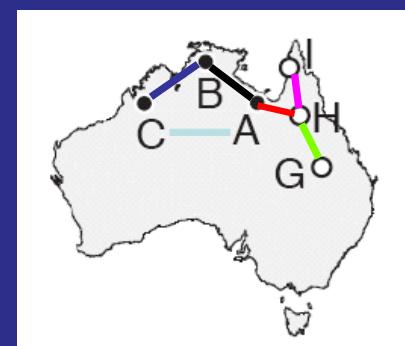
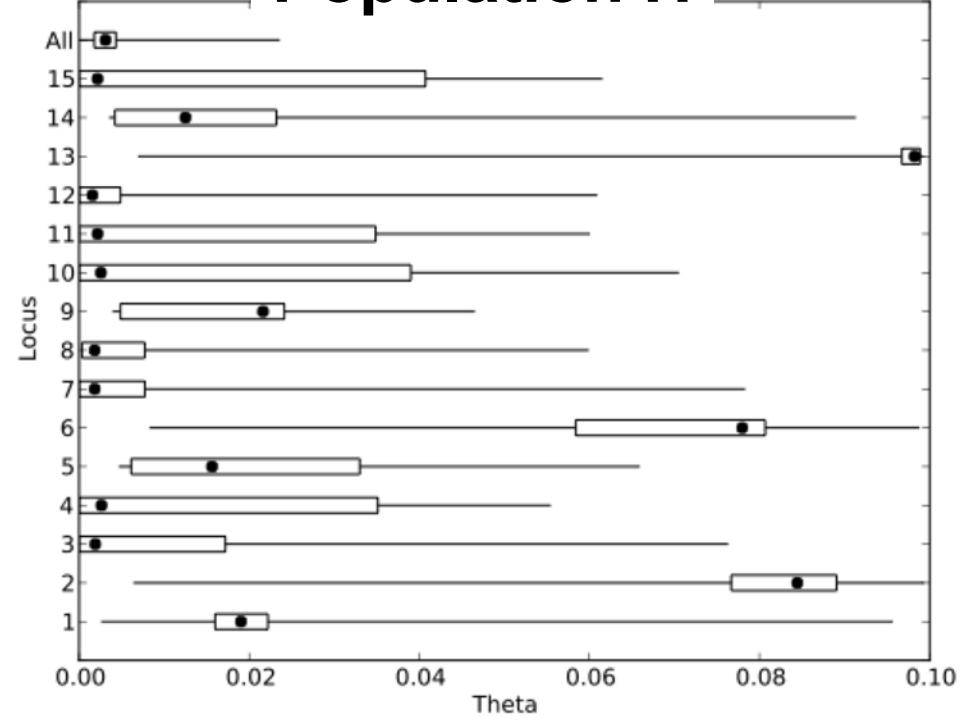
mL = marginal likelihood

Higher precision estimates of demographic parameters (θ) with more loci

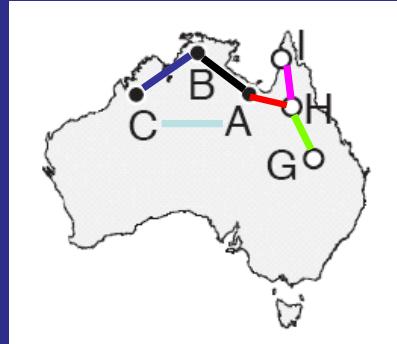
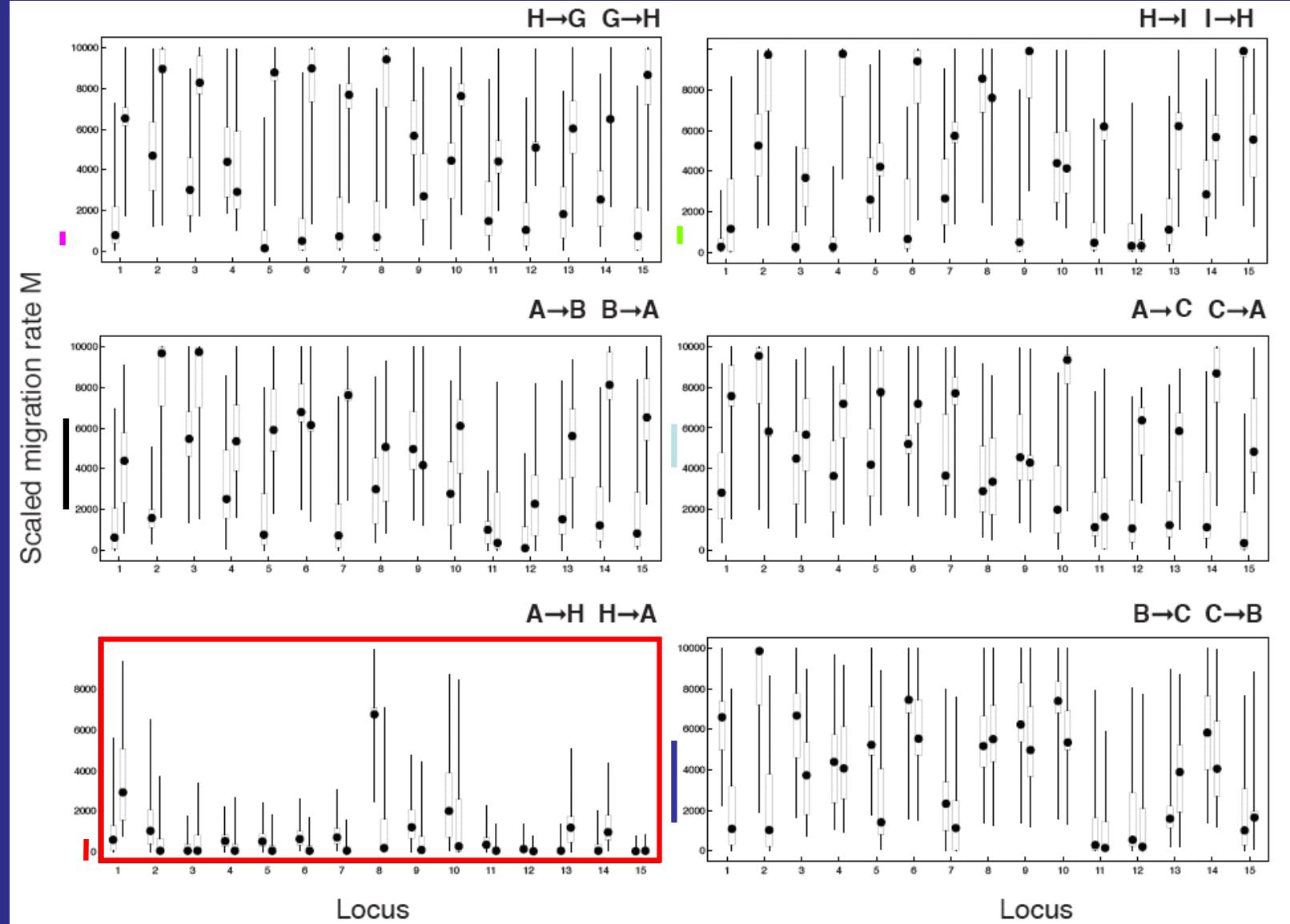
Population B



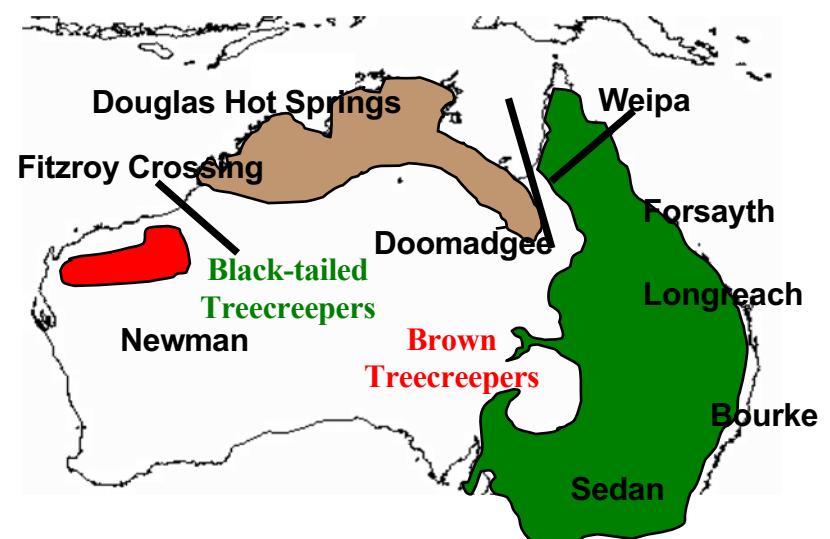
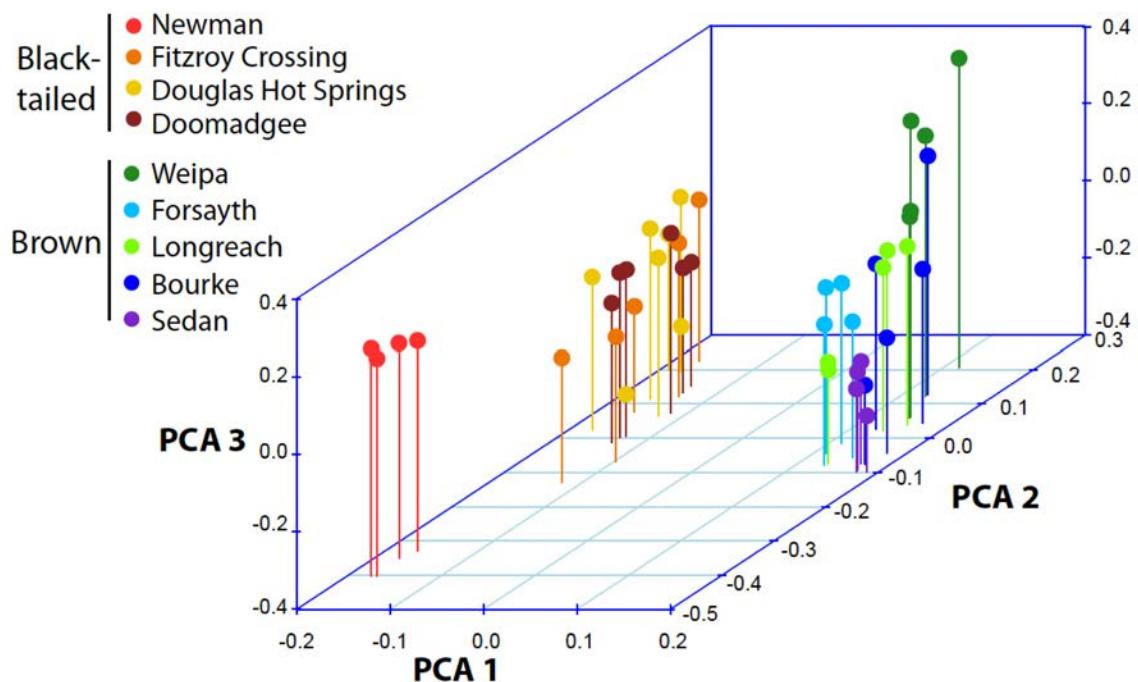
Population H



Multilocus estimates of migration rate



Correspondence between PCA and geography



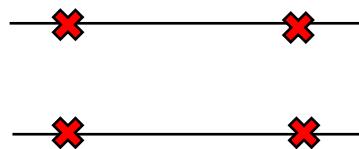
Rad-seq: powerful markers for phylogeography



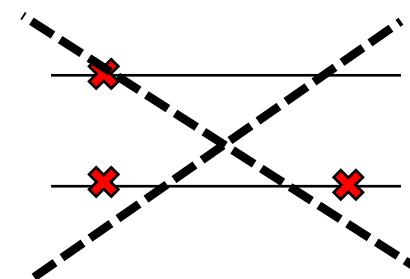
Caveat: Rad-seq assumes homozygous restriction sites

Arnold et al. 2013. *Mol. Ecol.* 22:3179-3190

Locus with homozygous restriction sites

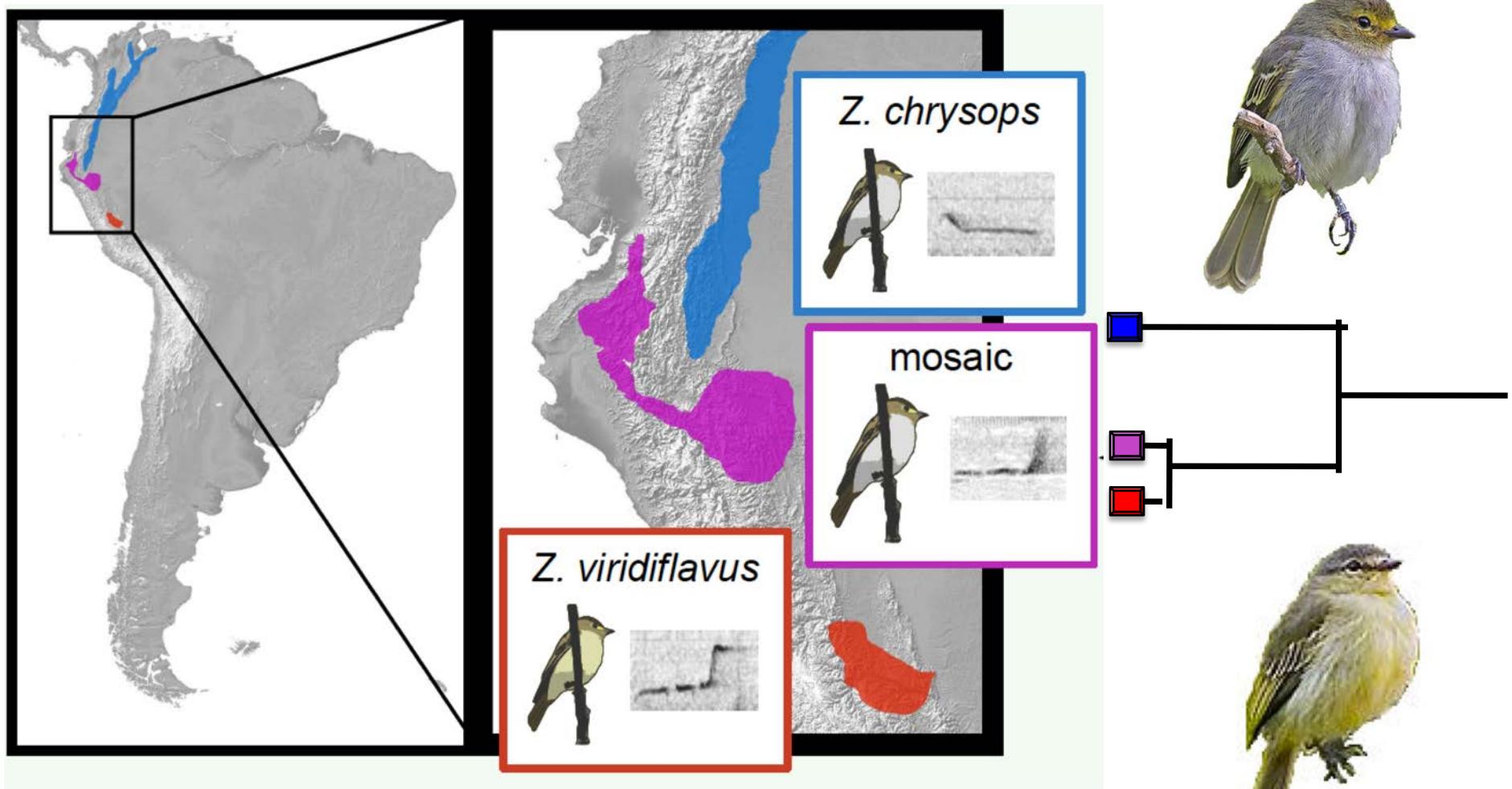


Loci with heterozygous restriction sites
– not sampled



Geography and mtDNA relationships among *Zimmerius* flycatchers

Zimmerius chrysops



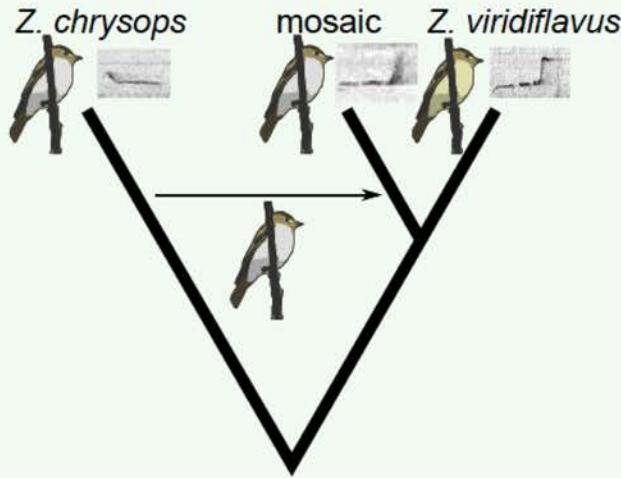
maximum parsimony; ND2 (mtDNA)

Rheindt et al. (2008c; Mol. Phylogenet. Evol.)

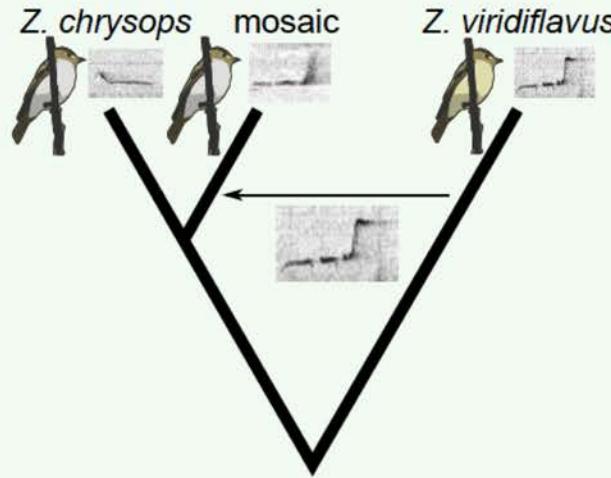
Zimmerius viridiflavus

Hypotheses for the origin of the ‘mosaic’ lineage

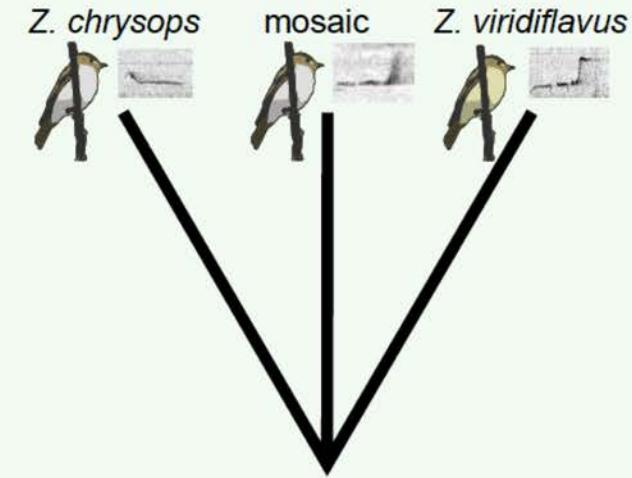
Hypothesis 1



Hypothesis 2



Hypothesis 3

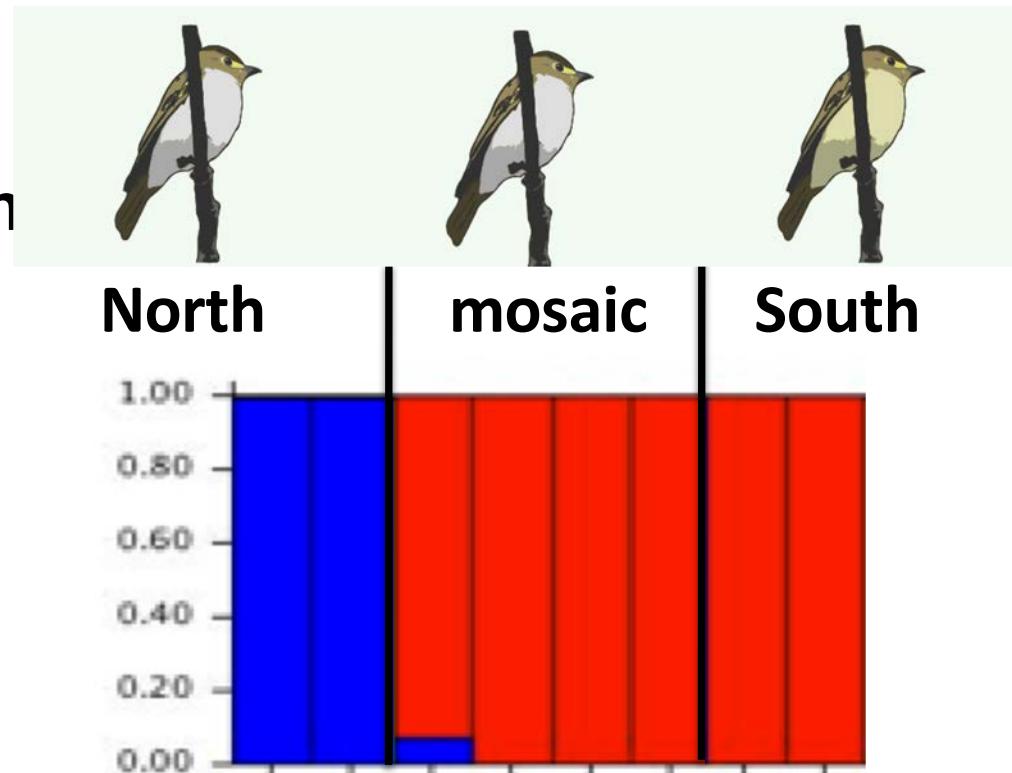


Flow of ‘plumage’ genes
Flow of ‘song’ genes
Inheritance of mosaic traits
into mosaic populations
from ancestral populations

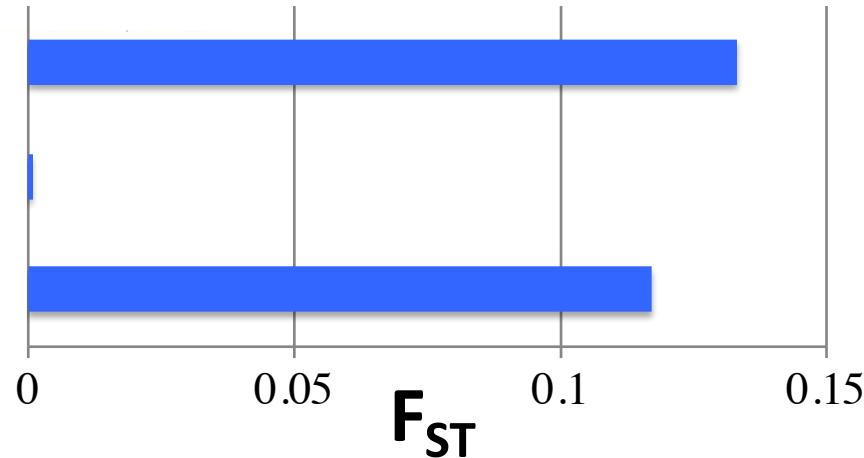
Population genomics of *Zimmerius* Flycatchers

STRUCTURE
plot based on
2120 SNPs
shared
among all
individuals

K=2



North vs South
mosaic vs South
mosaic vs North



SNAPP tree: A species tree made from SNP genotypes

Bryant et al. 2012 *MBE*

0 = homozygote A/A 1=heterozygote A/G 2 = homozygote G/G

<i>chrysops</i>	SI	II	III
	northern_individual2		0100000000001001000000002000
	northern_individual3		0100000000001000000101000000
<i>mosaic</i>	northern_individual1		0100000000000000100100200000
	mosaic_individual11		1100000001101000000000000000
	mosaic_individual9		01010001000000000000100200000
	mosaic_individual10		0100000000001010000001000000
	mosaic_individual8		011001100000000000000000000000
	mosaic_individual12		010000100000000000000000000000
	southern_individual5		010000001000110000000000000000
	southern_individual4		010000000001200000001010100
	gracilipes_individual6		010000000002001000010000002
	acer_individual7		0100100000002000011000000002

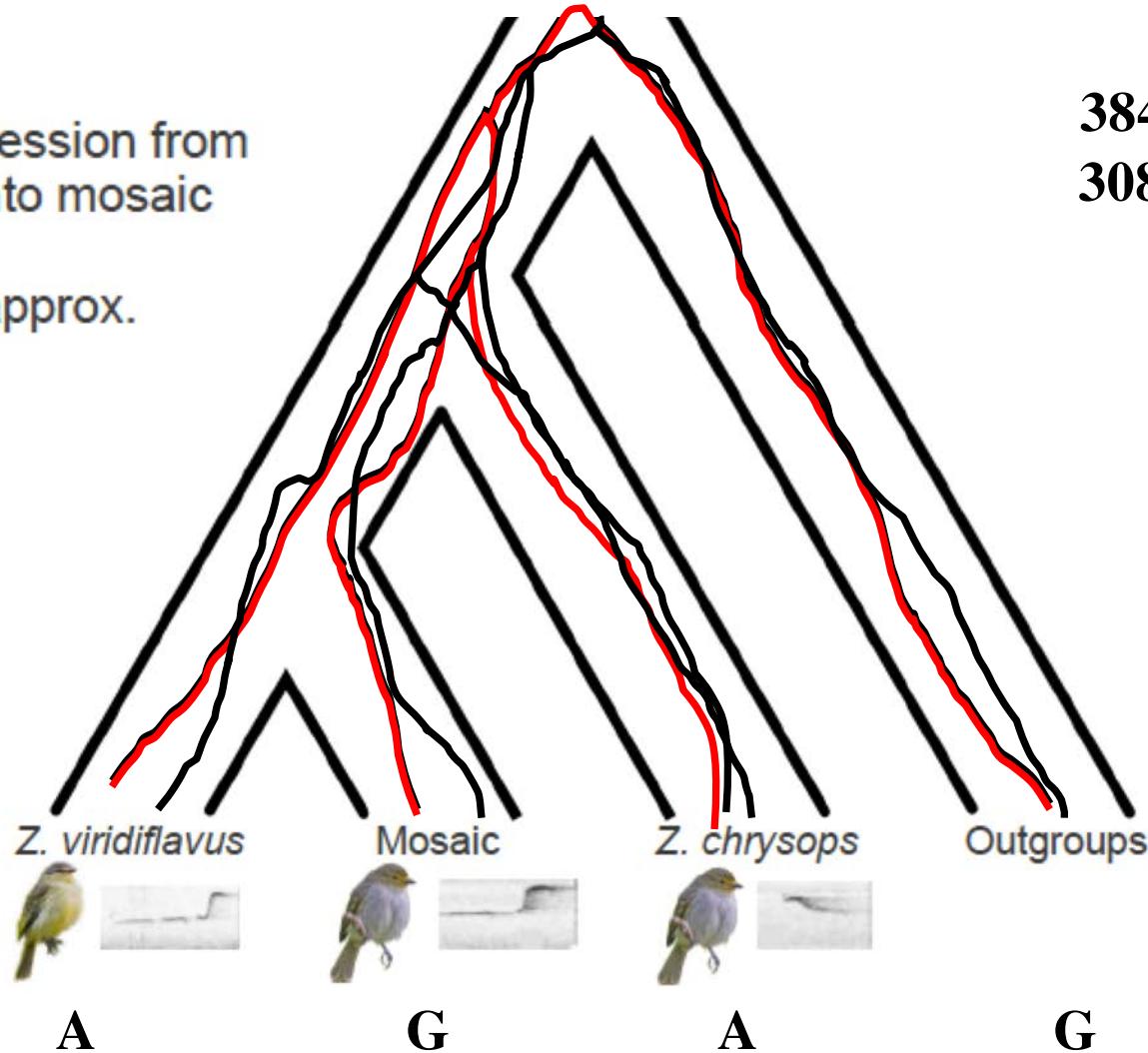
Genome-wide test of introgression in *Zimmerius* flycatchers

$$D = \frac{\text{Num. ABBA} - \text{Num. BABA}}{\text{Num SNPs}}$$

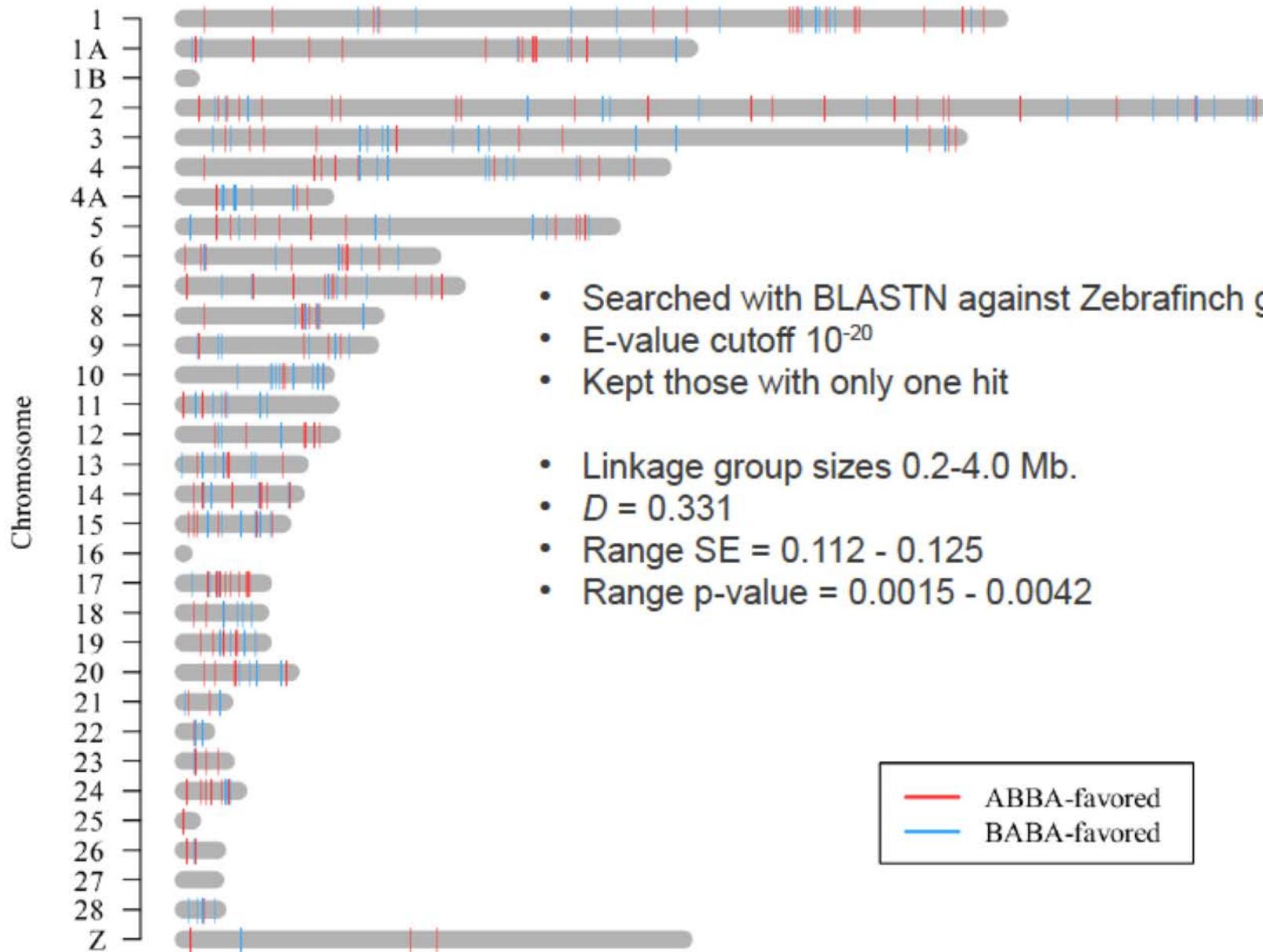
- $D > 0$: introgression from *Z. chrysops* into mosaic population
- (D assumed approx. Normal)

384 ABBA-like SNPs
308 BABA-like SNPs

$$D = 0.110$$



Introgressing SNPs are underrepresented on the Z chromosome



Introgressing SNPs are enriched for several cell functions

GO.ID	Term	Annotated	Significant	Corrected P-value
GO:0031253	cell projection membrane	53	7	0.0018632
GO:0042995	cell projection	485	17	0.0107408
GO:0005886	plasma membrane	1407	32	0.0107408
GO:0016020	membrane	3985	65	0.0107408
	plasma membrane			
GO:0044459	part	736	21	0.0107408
	neuron projection membrane			
GO:0032589		8	3	0.012056
GO:0071944	cell periphery	1460	32	0.012056
	cell projection part			
GO:0044463		208	10	0.0137
	leading edge membrane			
GO:0031256		27	4	0.025573333
	membrane			
GO:0044425	part	2740	47	0.04932



Conclusions: Flycatchers

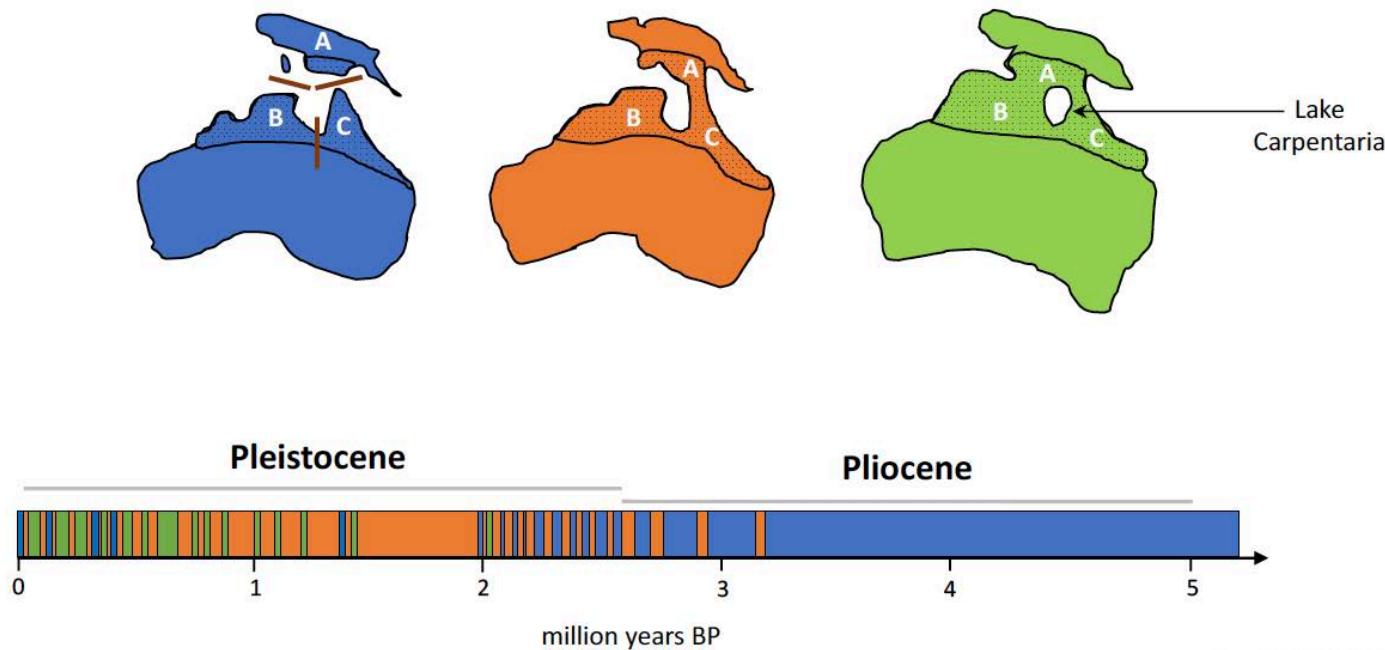
- Vocalizations are a better guide to genomic affinities than morphology
- Small portions of the genome can flow between species, even when most of the genome is distinct
- Locations of introgressing loci suggest a role for carotenoid metabolism in plumage differences

Whole-genome phylogeography of a widespread Australian honeyeater

John Burley, L. Joseph, A. Drew, N. Backström and S. Edwards



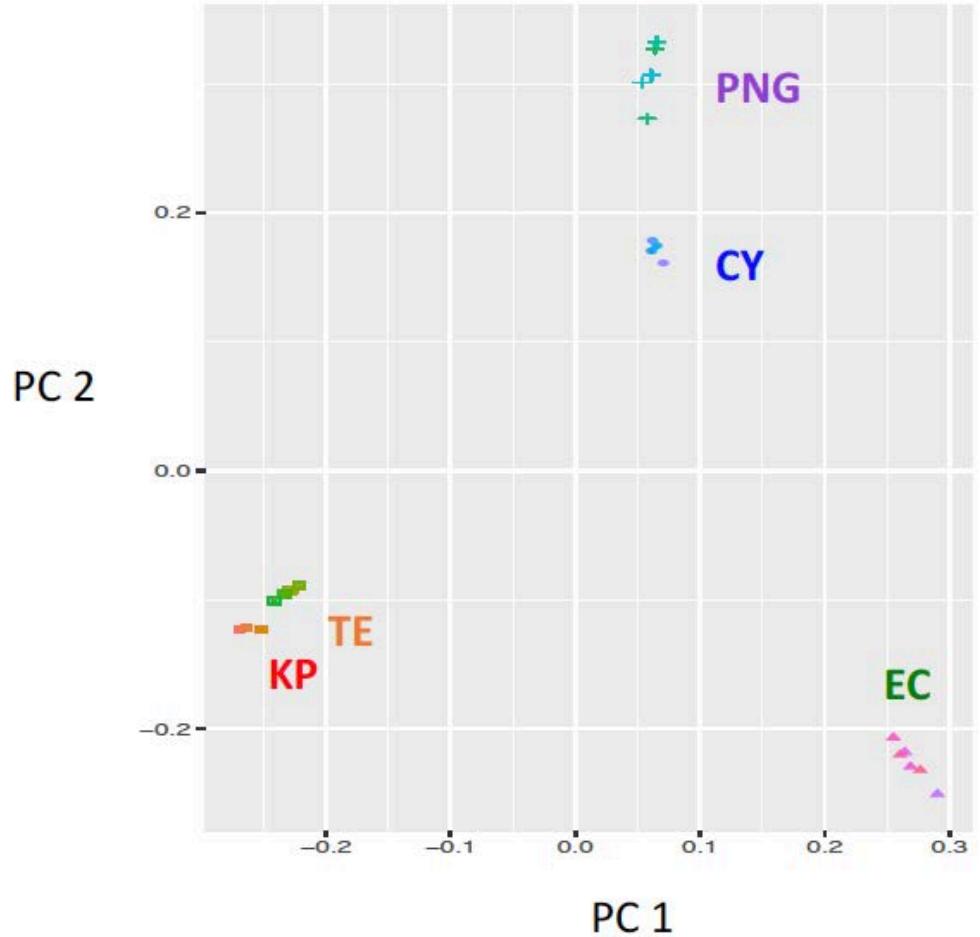
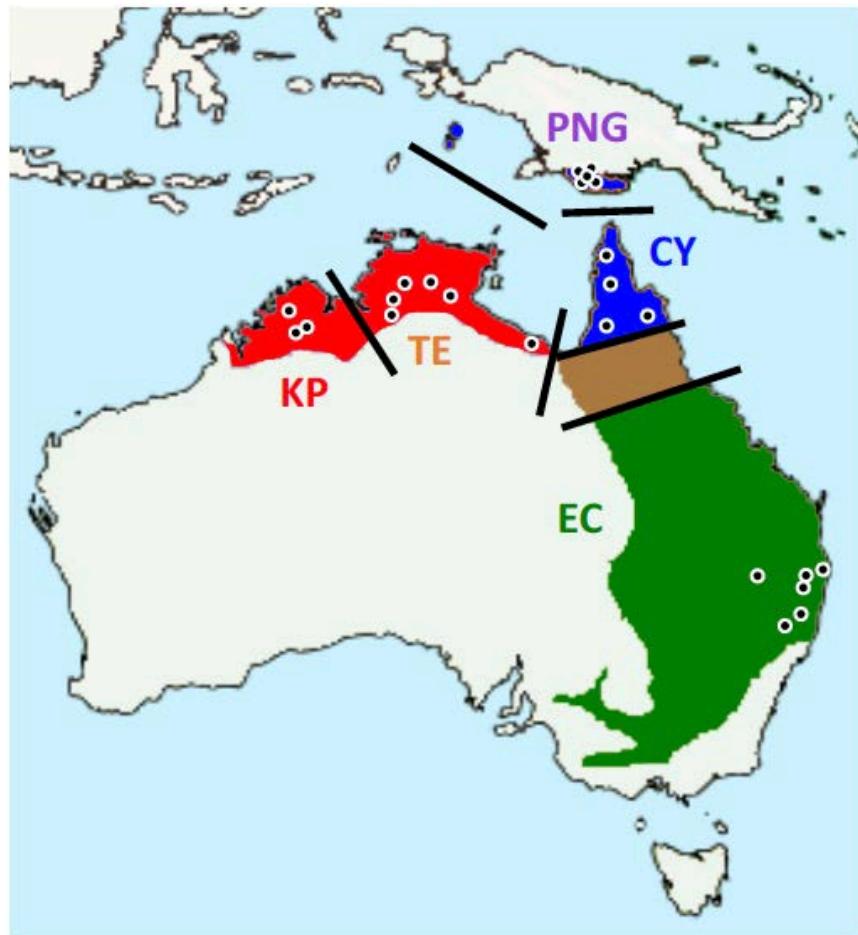
Blue-faced honeyeater
Entomyzon cyanotis



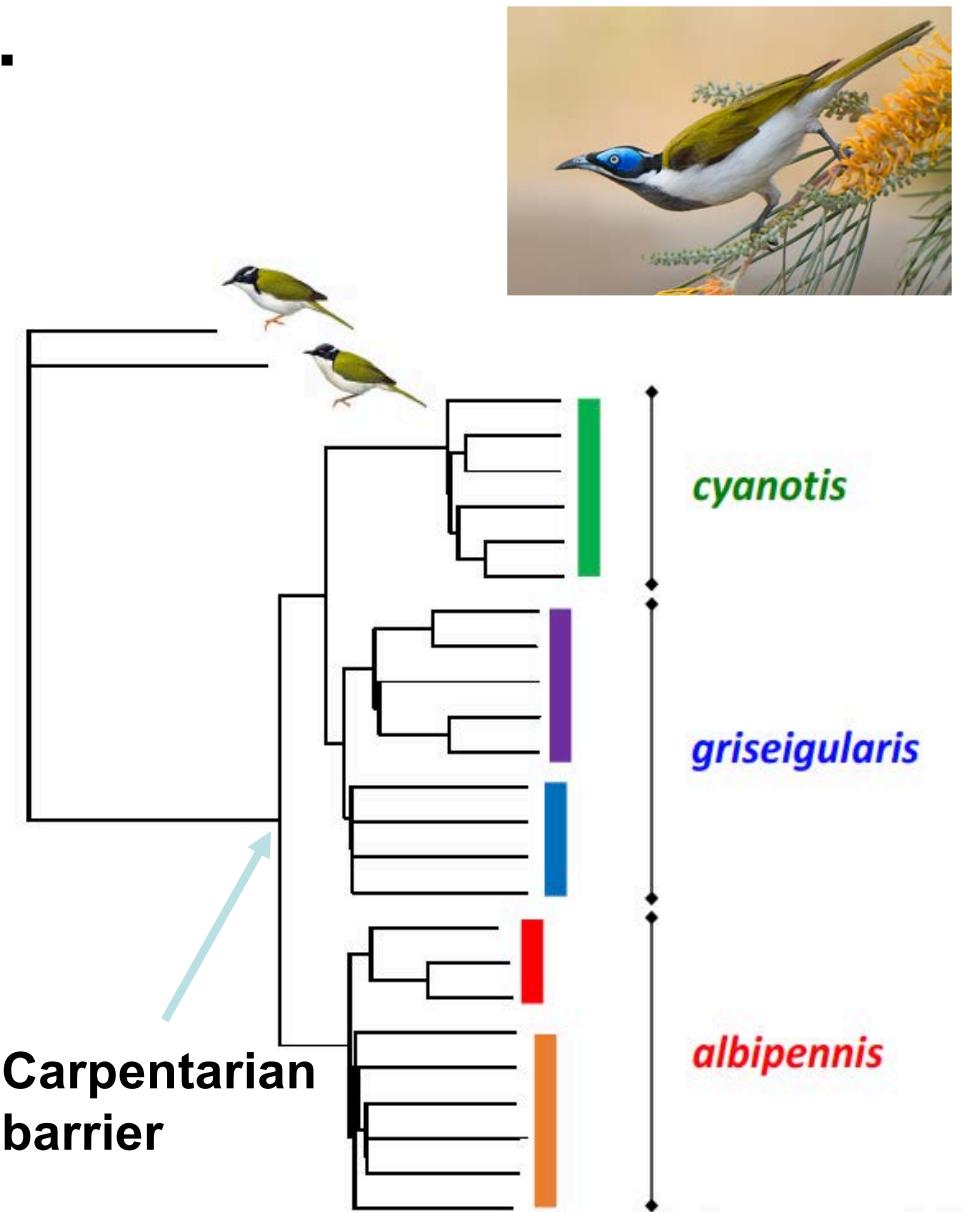
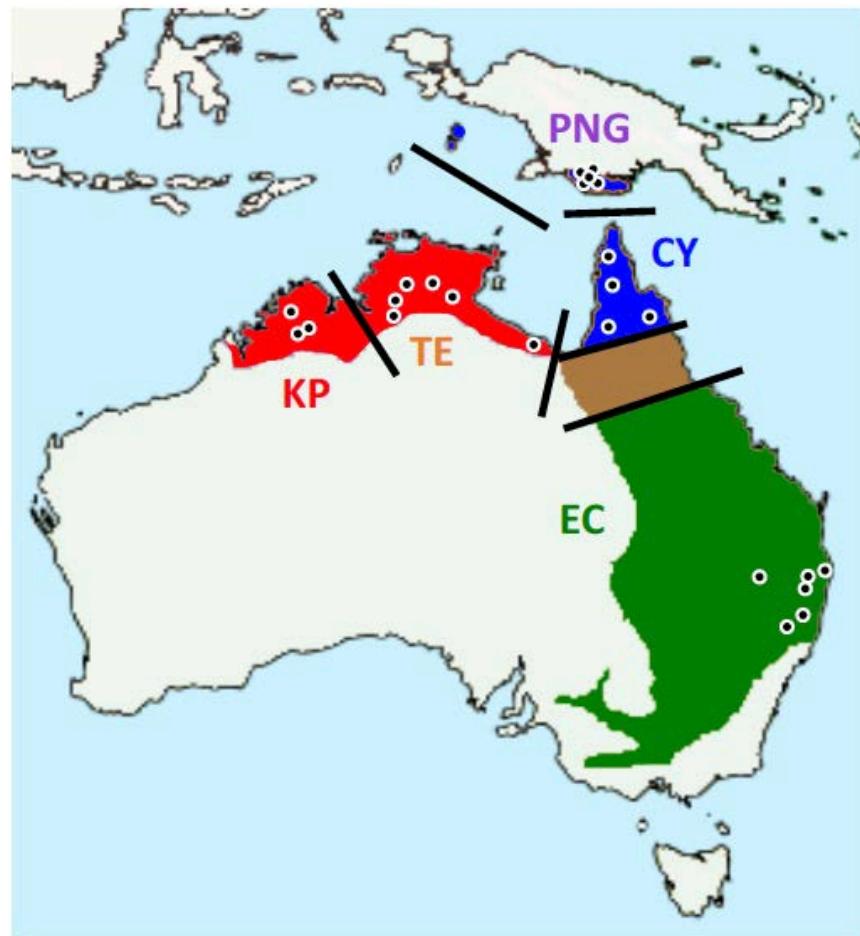
Naish et al. (2009) *Nature*
Kearns et al. (2011) *Mol. Ecol.*

- 24 samples
- + 2 outgroups
- 64X reference genome
- 7-12X population resequencing
- GATK variant calling
- ANGSD (Korneliussen et al. 2014. *BMC Genomics*)

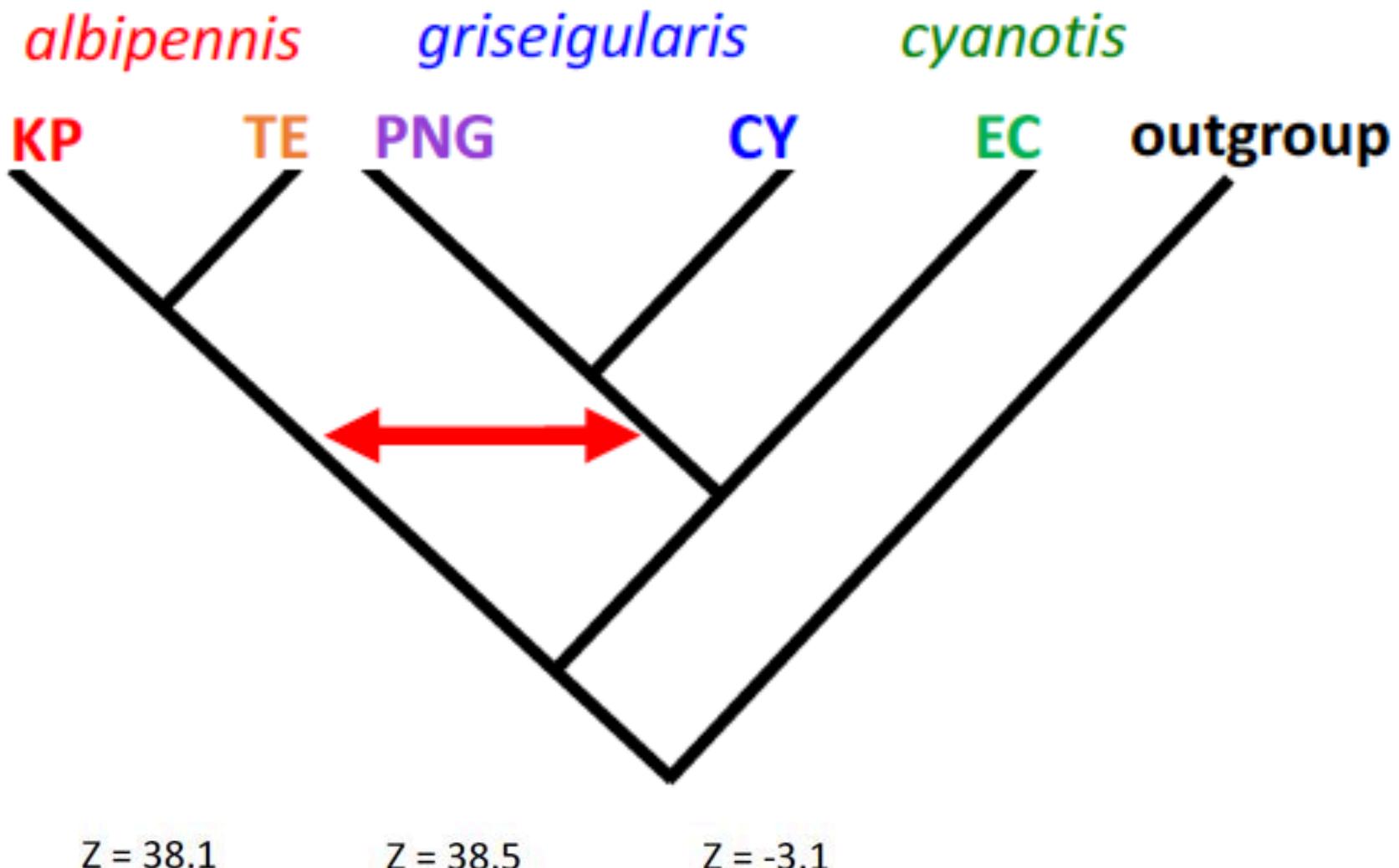
PCA of ~6 M SNPs shows close correspondence with geography



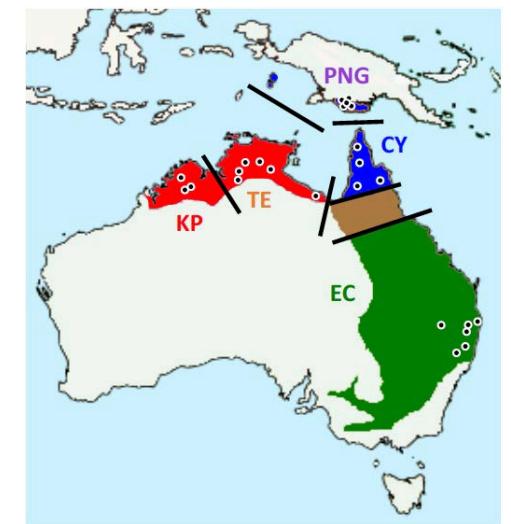
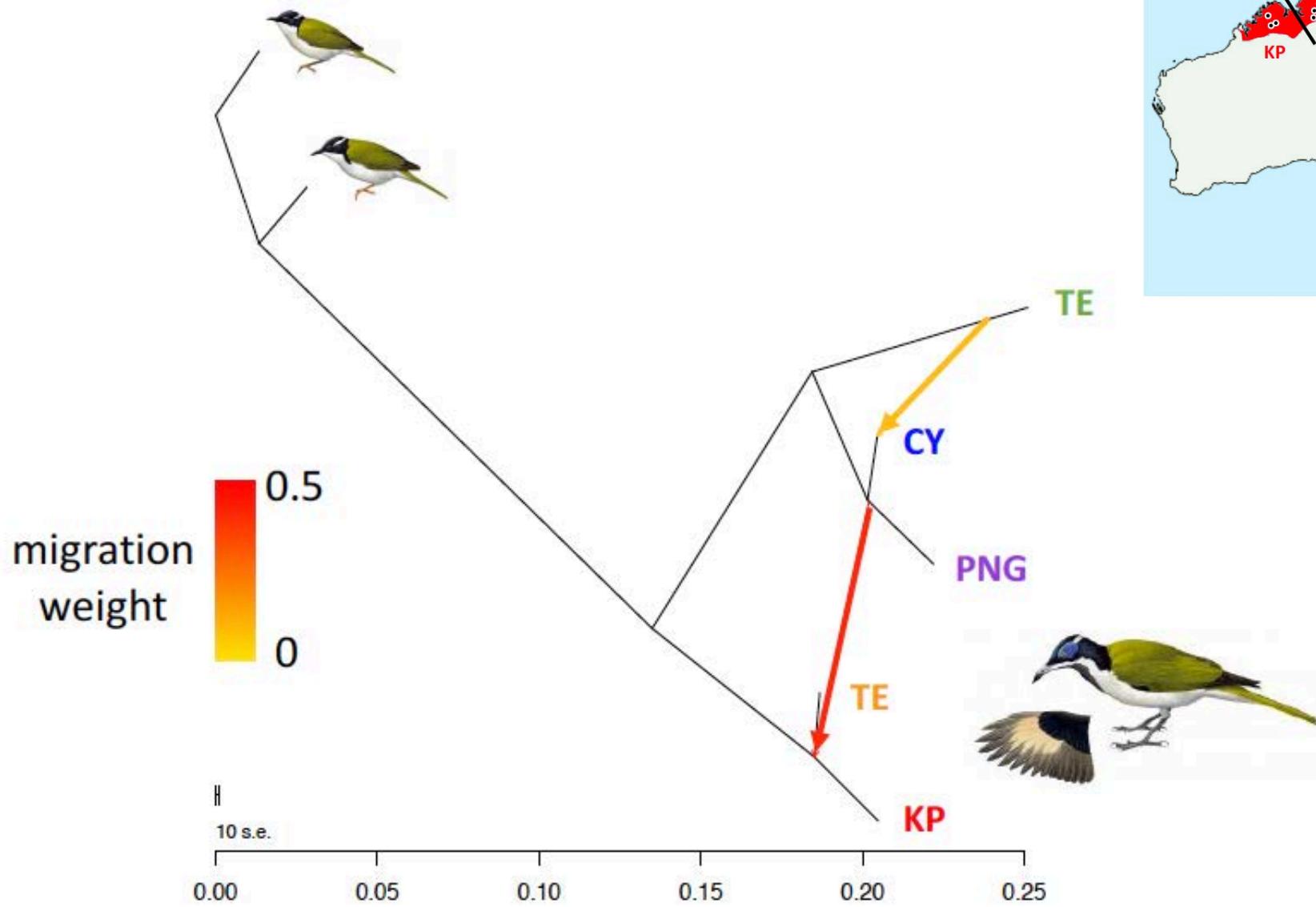
Carpentarian and other barriers cause divergence...



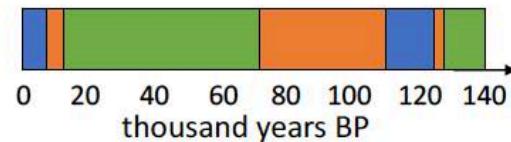
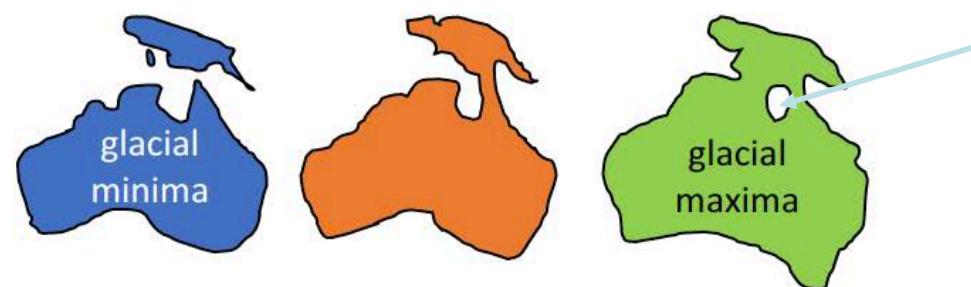
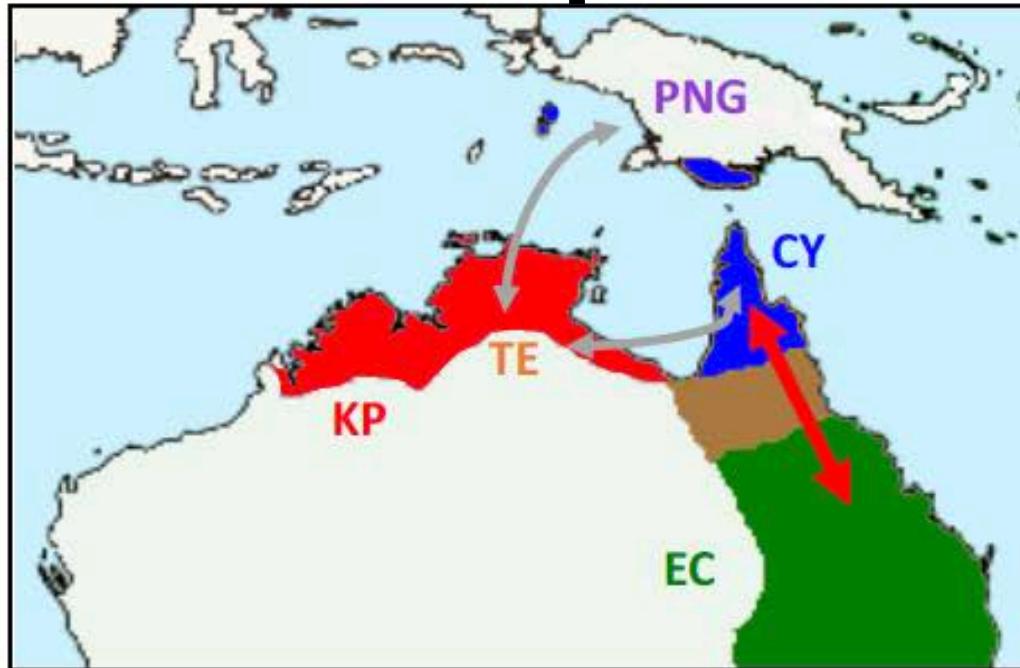
...but ABBA-BABA tests reveal signals of past introgression



TreeMix analysis also suggests introgression



Phylogeographic shadows of Lake Carpentaria



Australo-Papuan Fairy Wrens - Maluridae



Splendid Fairywren (*Malurus*)



Superb Fairywren (*Malurus*)



Striated Grasswren (*Amytornis*)



Southern Emu-wren (*Stipiturus*)

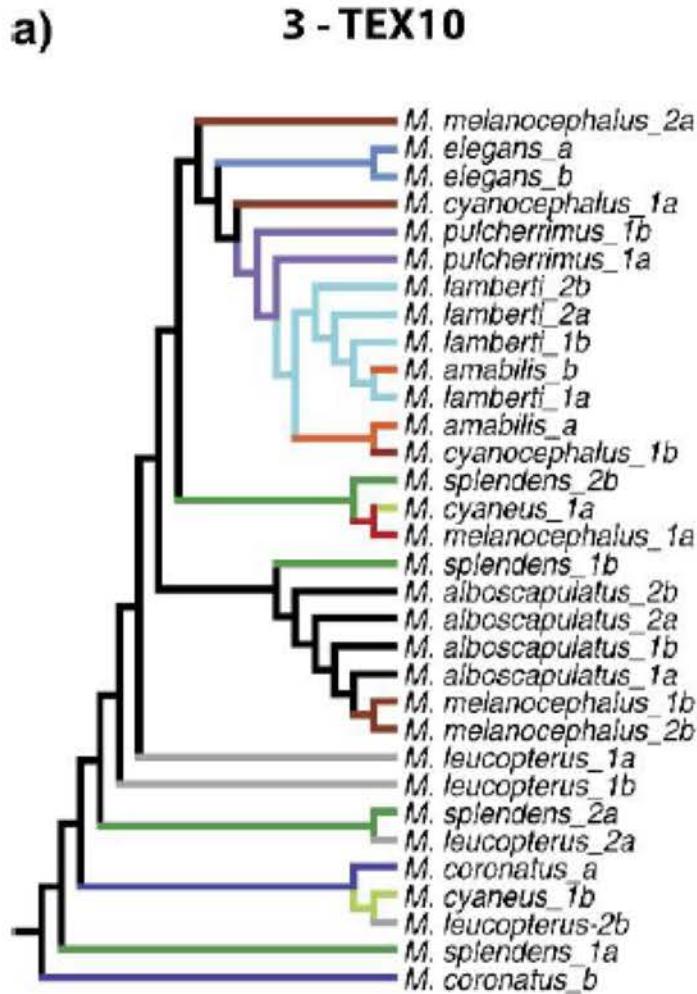
BROAD SAMPLING WITHIN FAIRY WRENS

No	Species Name	Common Name	Depository	Specimen No.	State
1	<i>Amytornis ballarae</i>	Kalkadoon Grasswren	ANWC	41740	WA
2	<i>Amytornis barbatus</i>	Grey Grasswren	ANWC	41789	WA
3	<i>Amytornis dorotheae</i>	Carpentarian Grasswren	ANWC		NT
4	<i>Amytornis goyderi</i>	Eyrean Grasswren	ANWC		SA
5	<i>Amytornis housei</i>	Black Grasswren			PNG
6	<i>Amytornis merrotsyi</i>	Short-tailed Grasswren			QLD
7	<i>Amytornis purnelli</i>	Dusky Grasswren			QLD
8	<i>Amytornis textilis</i>	Thick-tailed Grasswren			QLD
9	<i>Amytornis striatus</i>				QLD
10	<i>Clytorhynchus insignis</i>				QLD
11	<i>Malurus amabilis</i>				SA
12	<i>Malurus cyaneus</i>				PNG
13	<i>Stipiturus malachurus</i>				WA
14	<i>Stipiturus malaceus</i>				SA
15	<i>Stipiturus mallee</i>				PNG
16	<i>Stipiturus ruficeps</i>				QLD
17	<i>Stipiturus malachurus</i>	Emu-wren	ANWC	20947	QLD
18	<i>Stipiturus malaceus</i>	Red-capped Fairy wren	ANWC	29906	QLD
19	<i>Stipiturus mallee</i>	Variegated Fairy wren	ANWC	28009	QLD
20	<i>Stipiturus ruficeps</i>	Spotted Fairy wren	ANWC	31655	NSW
21	<i>Stipiturus malachurus</i>	Purple-crowned Fairy wren	BMNH	60807	NT
22	<i>Stipiturus malaceus</i>	Southern Emu-wren	ANWC	20748	SA
23	<i>Stipiturus mallee</i>	Mallee Emu-wren	ANWC	40418	SA
24	<i>Stipiturus ruficeps</i>	Rufous-crowned Emu-wren	ANWC	39914	QLD
25	<i>Greygone olivacea</i>	White-throated Greygone	MCZ	336023	NSW

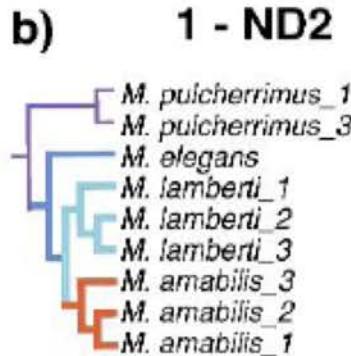
26 SPECIES 51 INDIVIDUALS
18 LOCI

High diversity of gene trees in the noncoding genome of Fairy Wrens

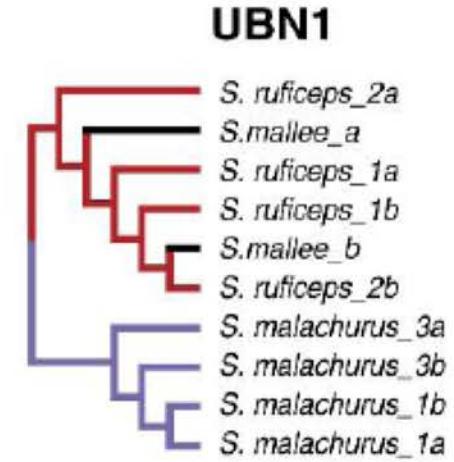
a)



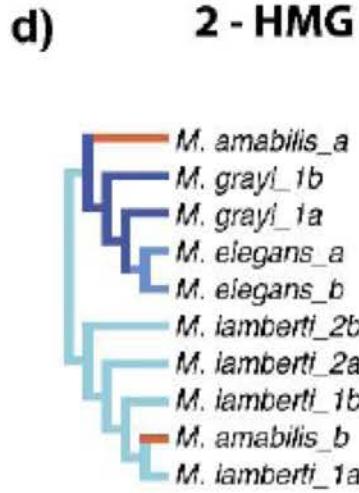
b)



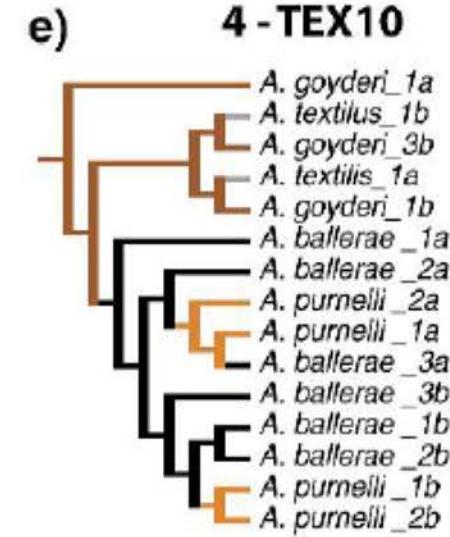
c)



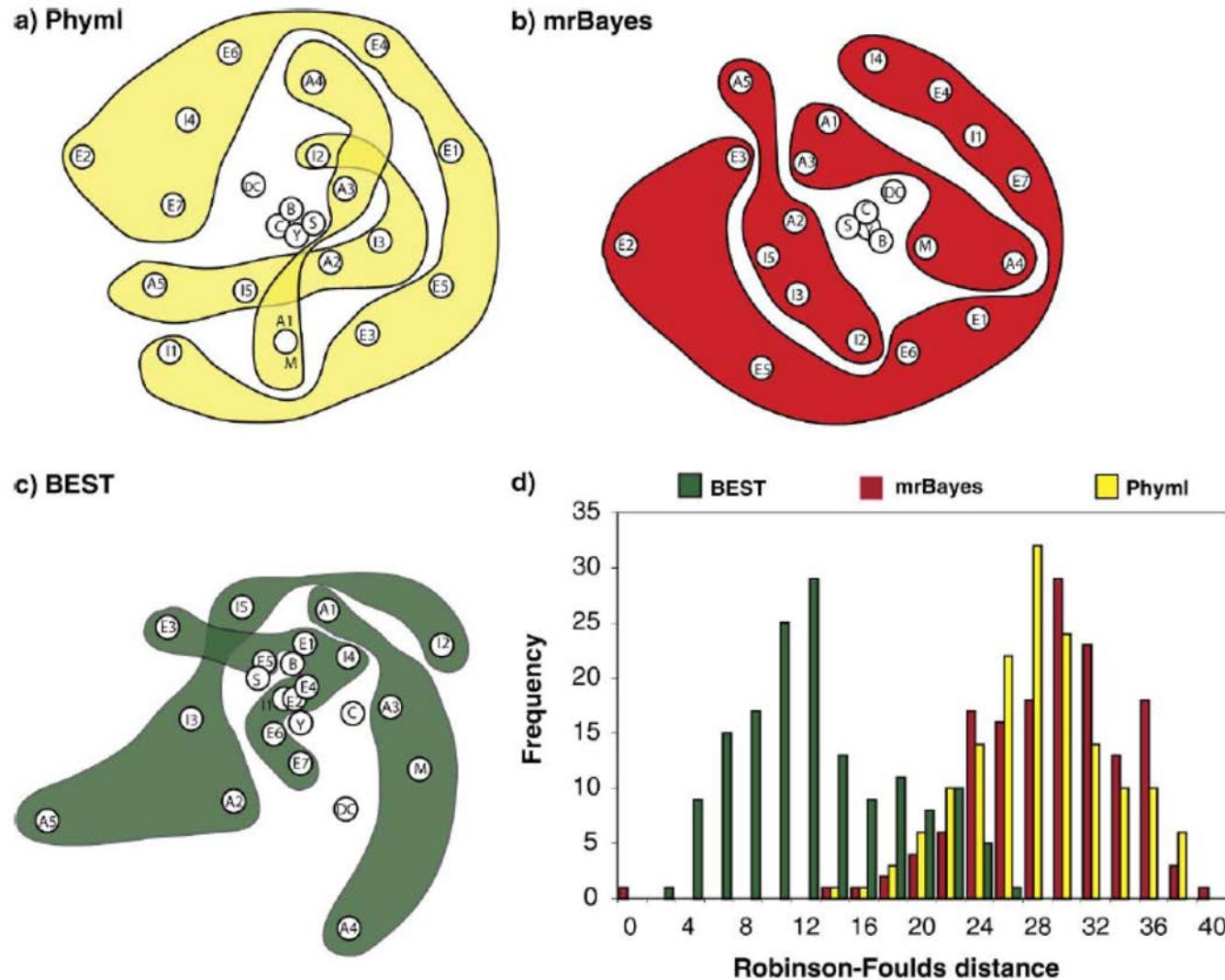
d)



e)

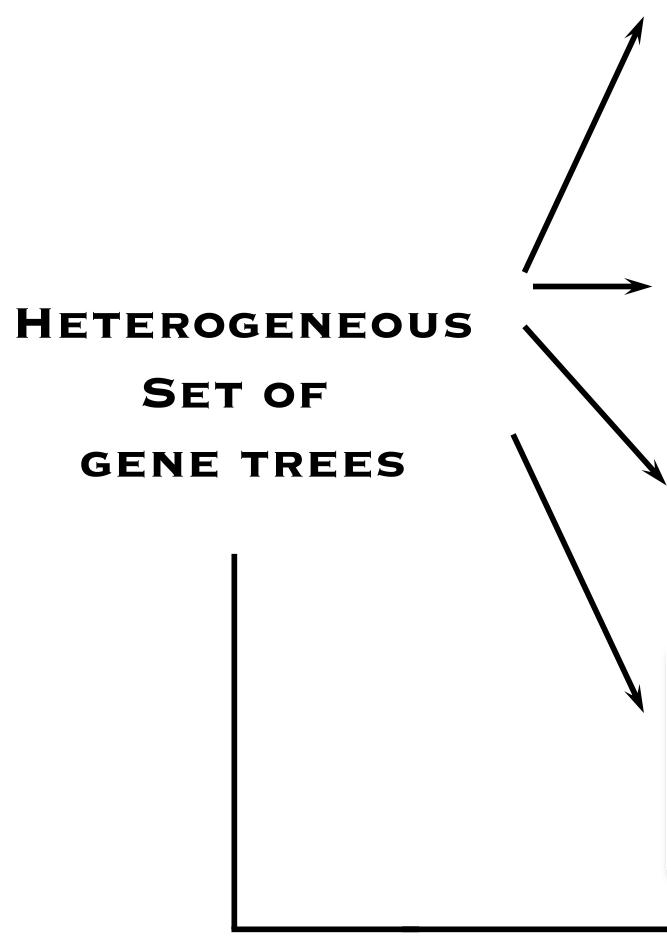


Heterogeneity in Fairy Wren gene trees across loci

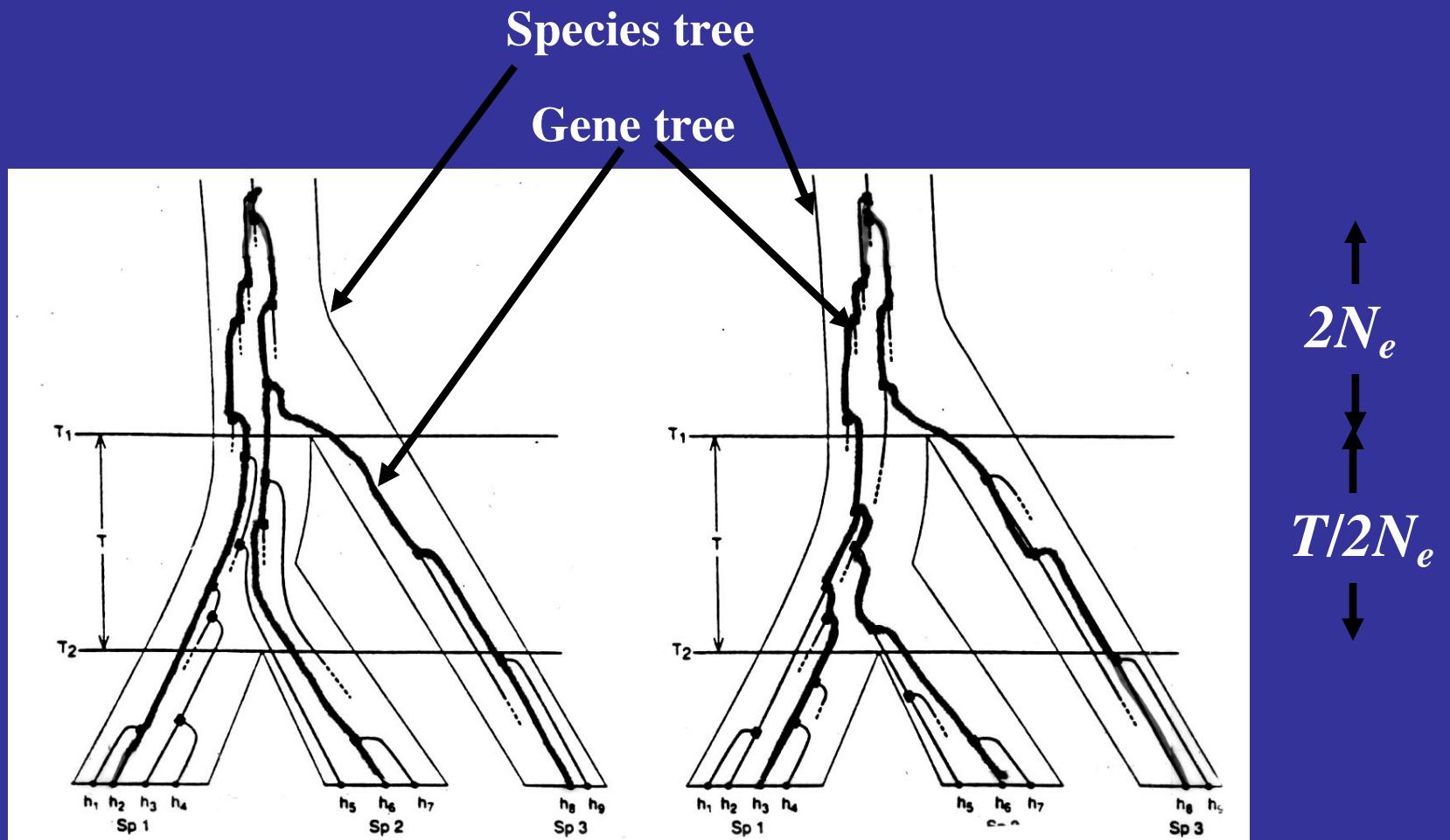


Tree Set Visualization (by multidimensional scaling): Hillis et al 2005. *Syst. Biol.* 54:471-482.
Concaterpillar: Leigh, J. W., et al.. 2008. *Syst. Biol.* 57: 104-115.

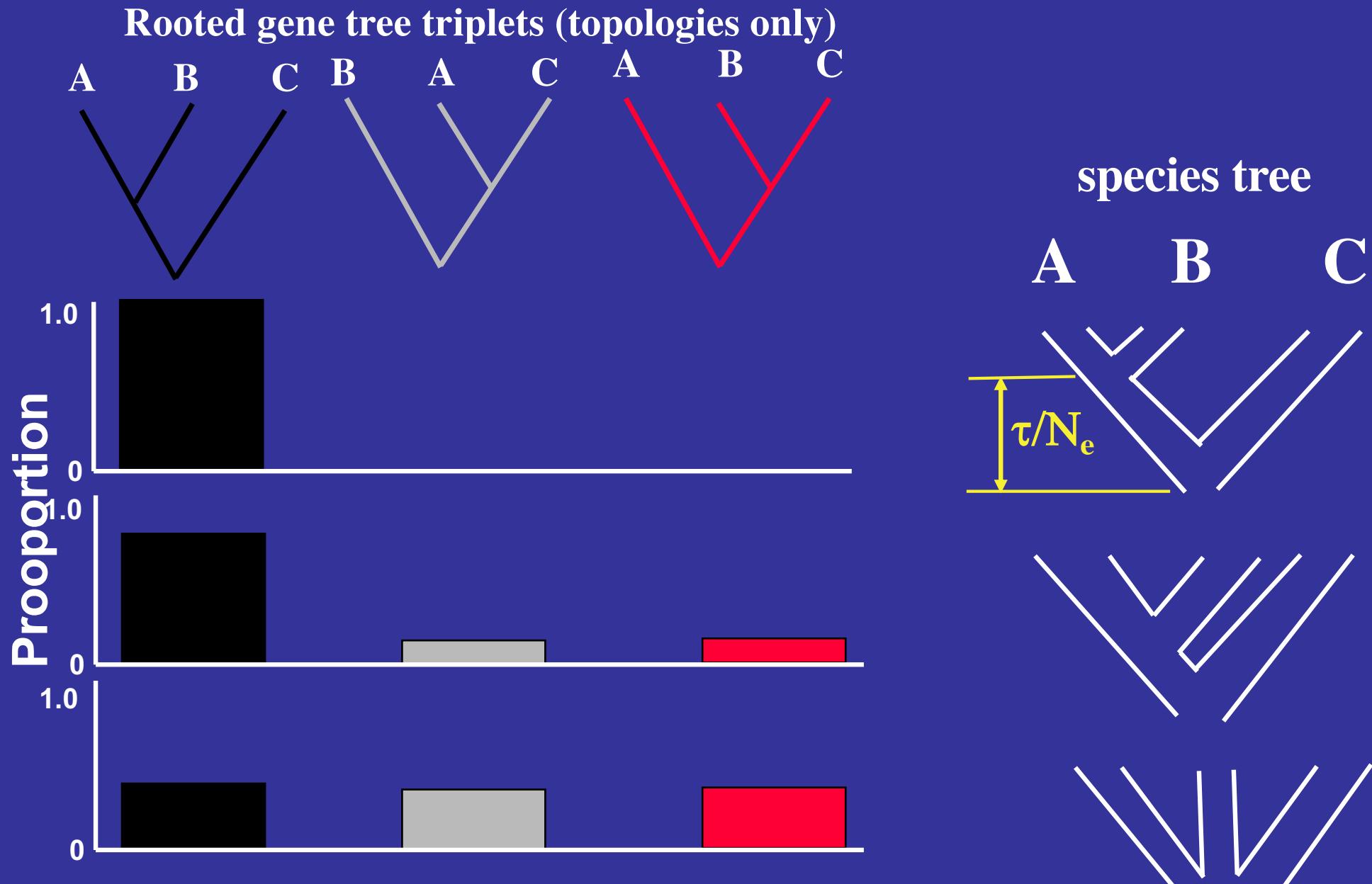
Exploring incomplete lineage sorting through species tree methods



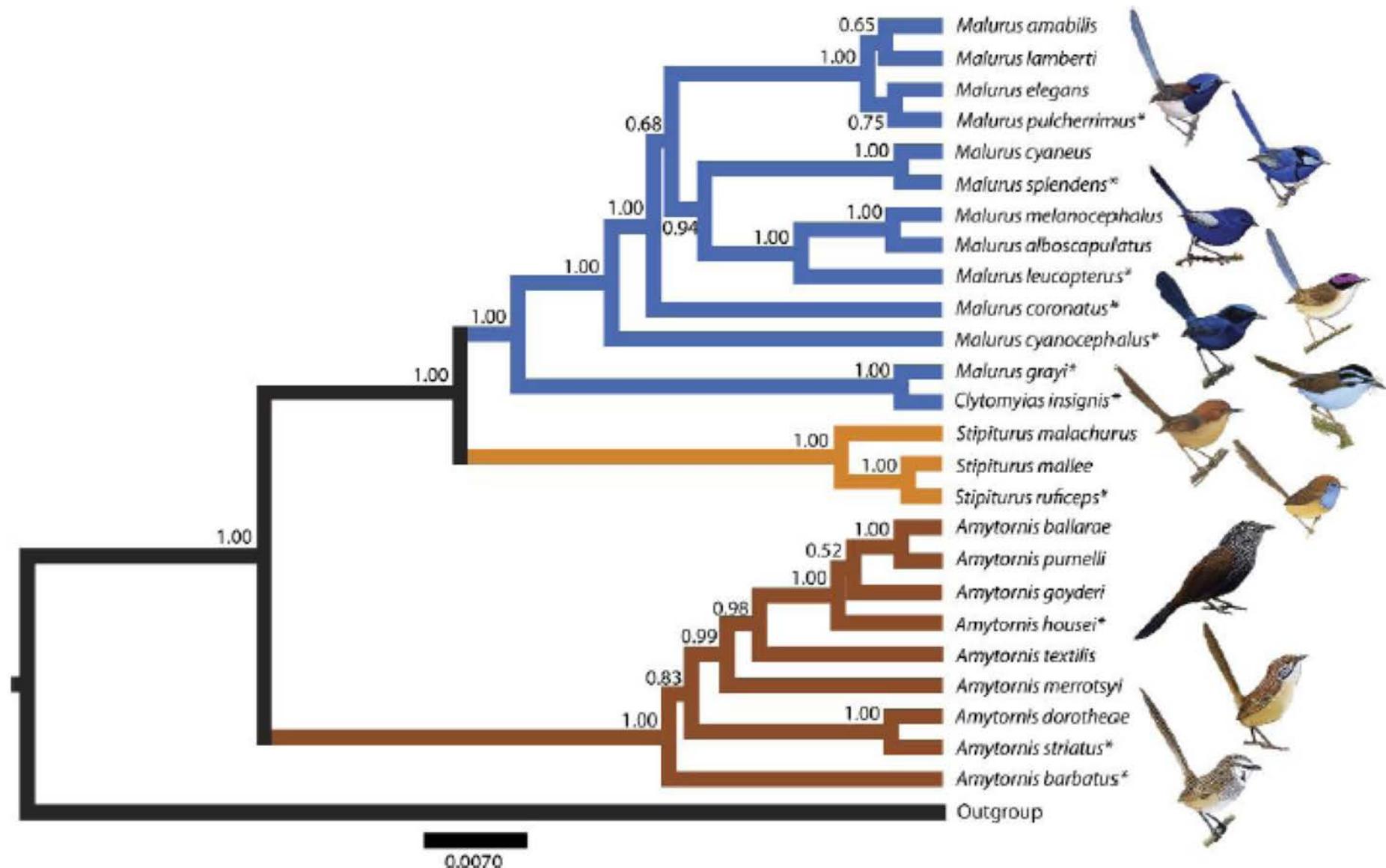
Stochastic gene tree variation during rapid radiations



Maximum (pseudo) likelihood method for species trees



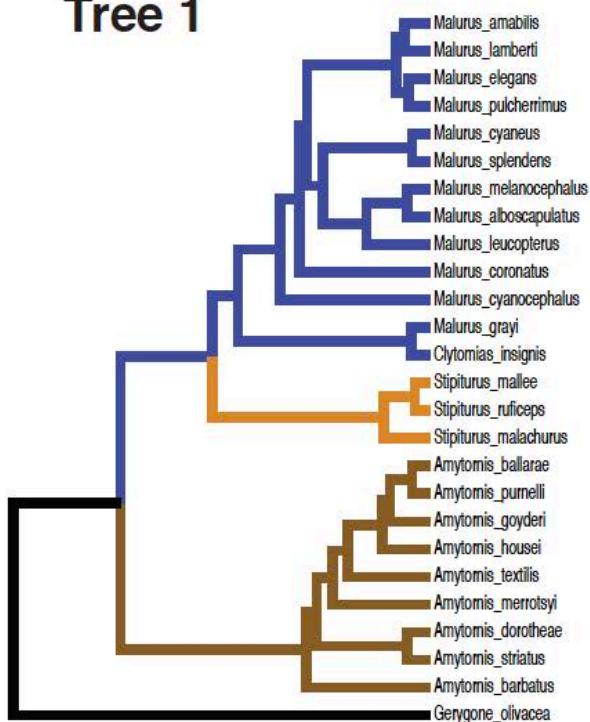
Multiple genes resolve the phylogenetic tree for Fairy Wrens



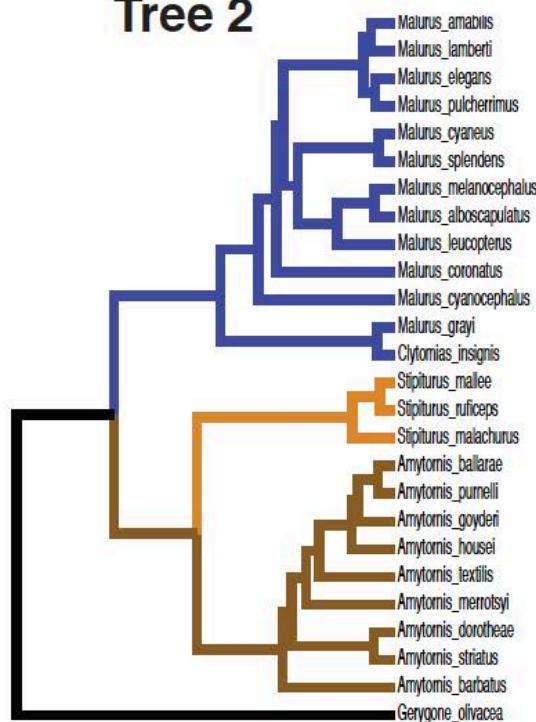
Lee, Joseph, & Edwards 2012. *Syst. Biol.* 61:253-271.

Testing tree topologies under the multispecies coalescent model using MP-EST

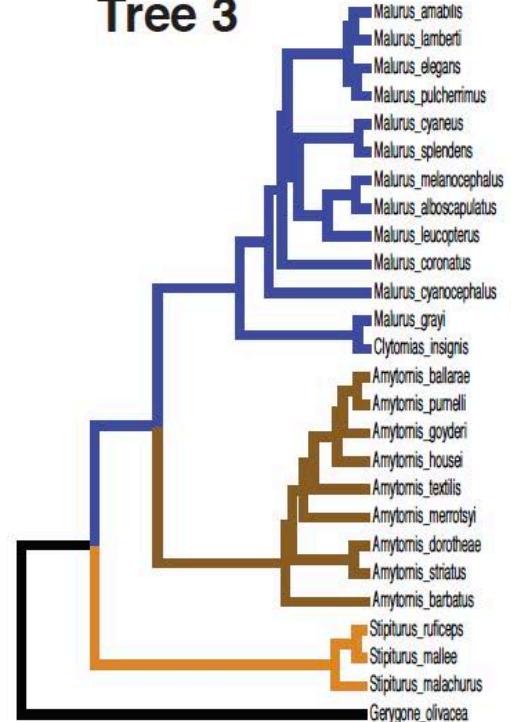
Tree 1



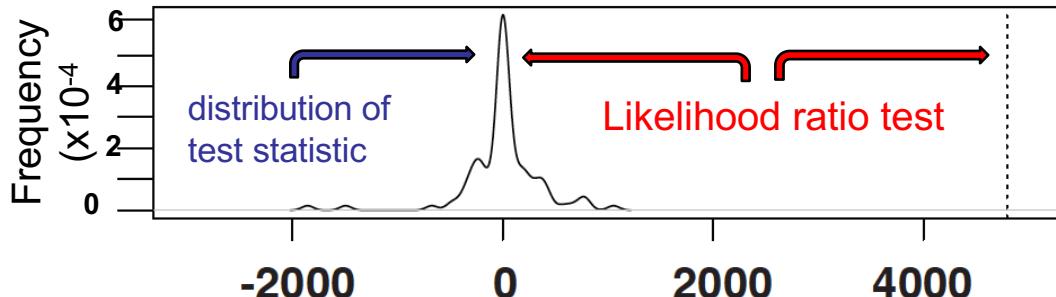
Tree 2



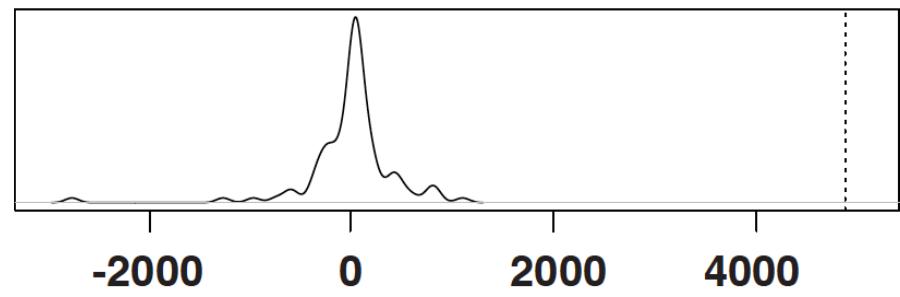
Tree 3



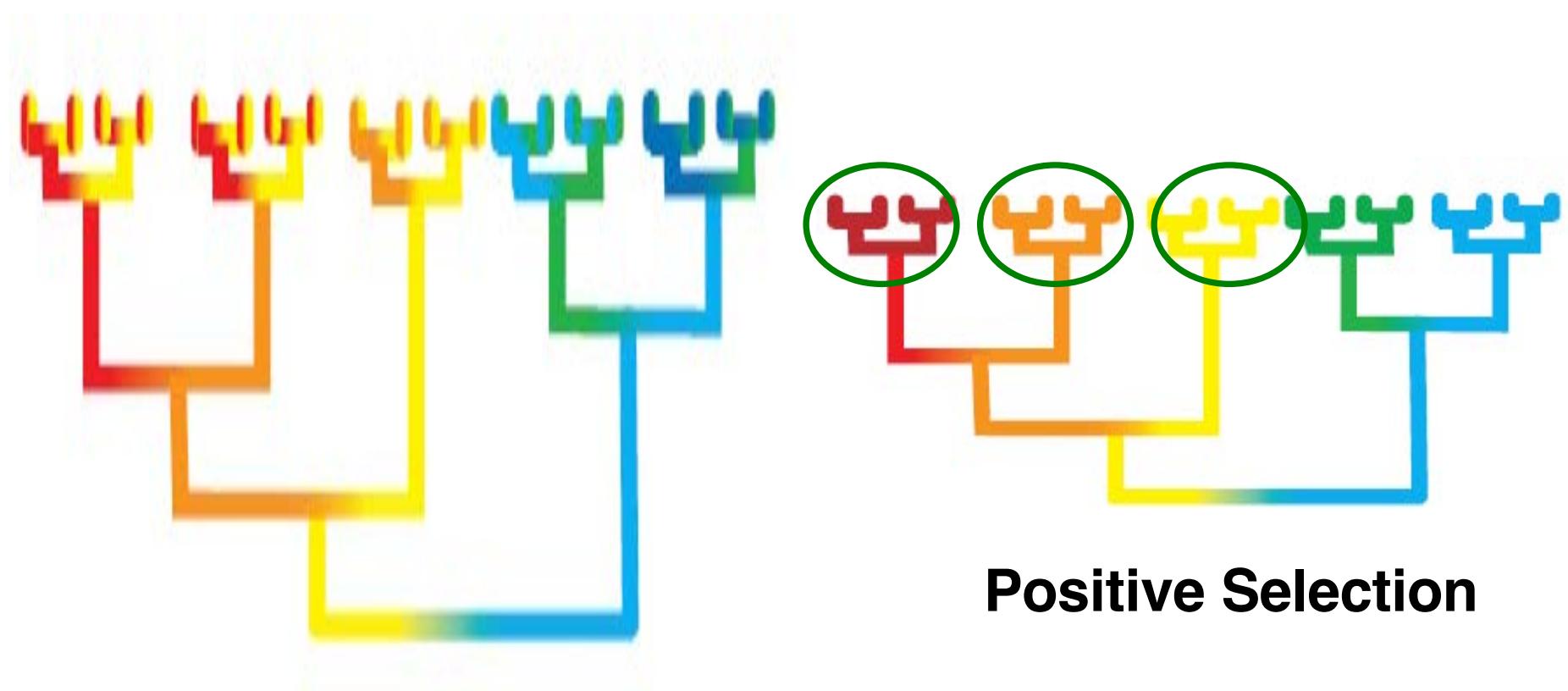
Tree 1 vs. Tree 2



Tree 1 vs. Tree 3



Phylogenetic manifestations of natural selection



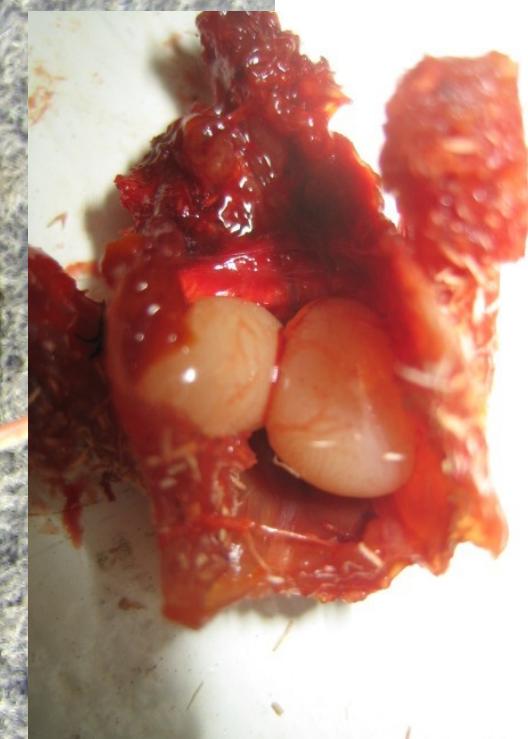
Neutral evolution

Positive Selection

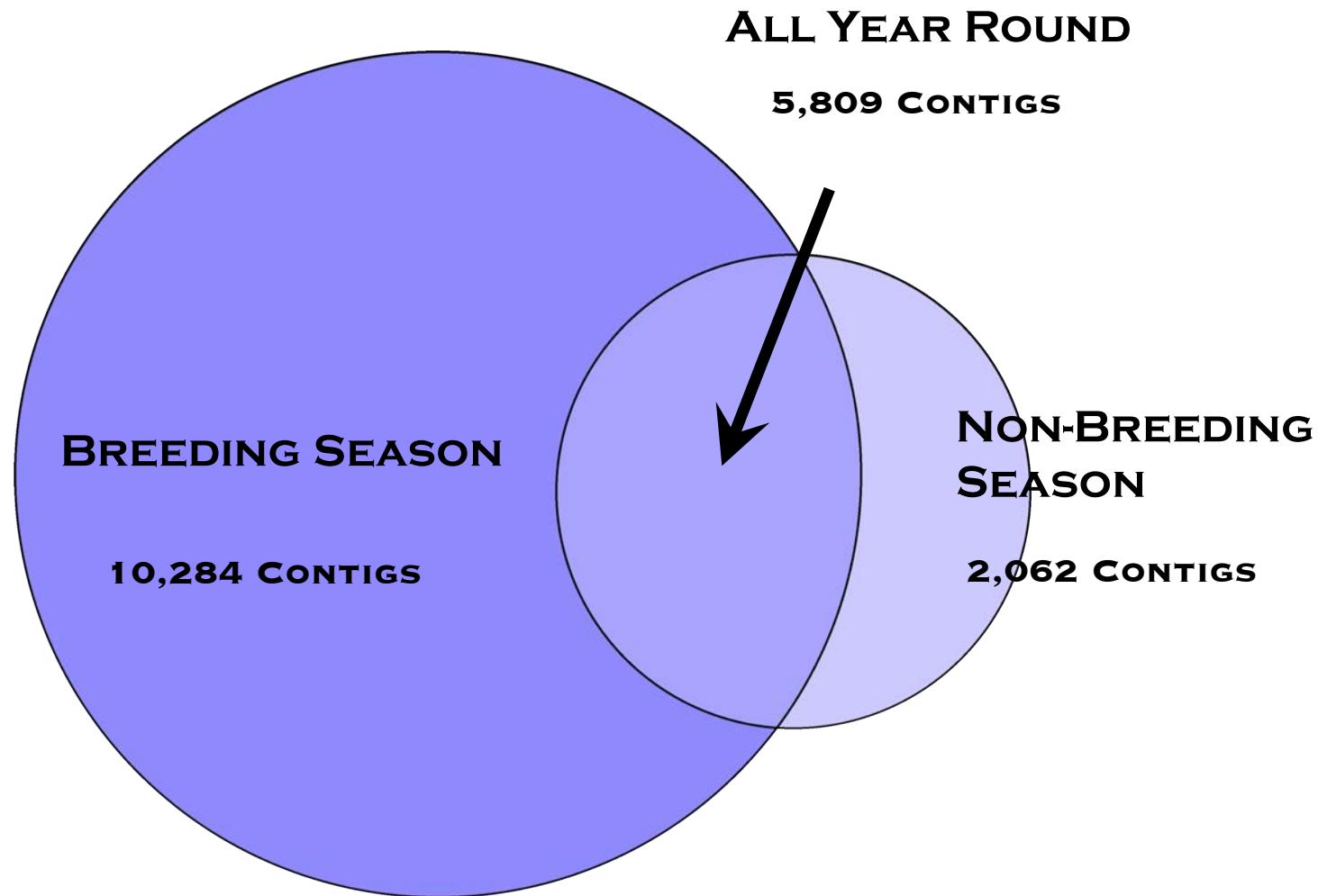
Extraordinary Sperm Production In Fairy Wrens



	Domestic chicken	Guinea fowl	White-winged fairy-wren
Body mass (g)	2300	1850	7.6
Sperm/day ($\times 10^6$)	2000	70	646 (164–1974)
Sperm/h ($\times 10^6$)	83.3	3.0	26.9 (6.8–82.3)
Sperm/day/g of body mass ($\times 10^6$)	0.87	0.04	85.0 (19.5–259.7)
Source	de Reviers & Williams 1981	Brillard & de Reviers 1981	Tuttle et al. 1996



Transcriptomics: Higher diversity of genes in breeding vs. non-breeding season testes

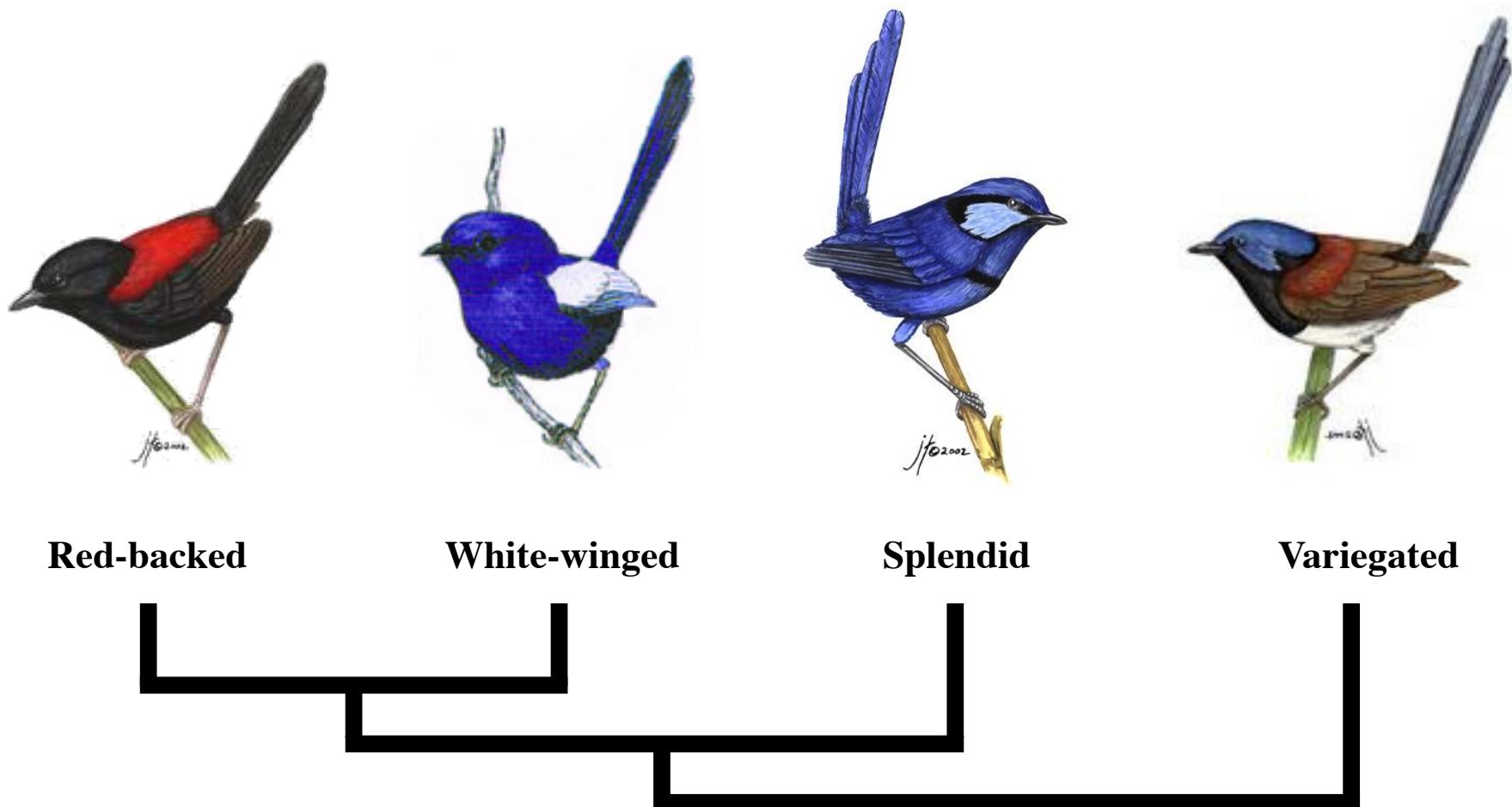


20 REPRODUCTIVE GENES SURVEYED

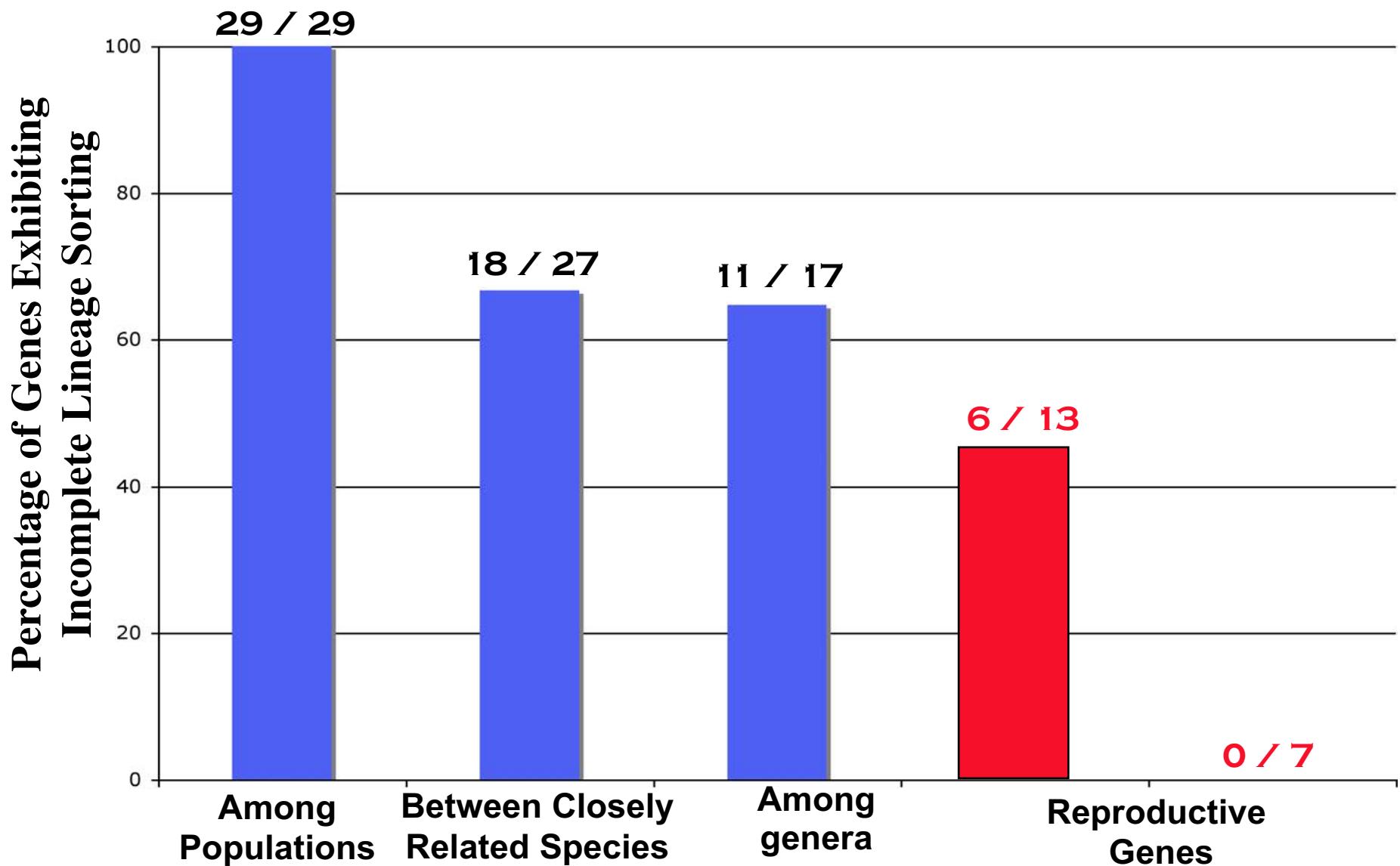
Contig	Gene Name	Function	Coverage
722	Kazal-type Serine Proteinase Inhibitor	Protection of viable spermatozoa from the proteinase	125.583
5992	Beta Defensin (6e-05)	Microbicidal peptides active against many bacteria, fungi, viruses	97.053
1948	Proacrosin	Mediating binding between acrosomal membrane and IPVL	91.025
1175	Tubulin Polymerization Promoting Protein	Promoting tubulin assemblies and cell proliferation	86.585
10816	Outer Dense Fiber Protein	Consists of sperm tail	64.102
5083			60.349
7574	Chromosome 10 Open Reading Frame 122		59.409
7761	Heat Shock Protein 40kd		42.143
5235	Creatine Kinase B-Variant	Catalyses the conversion of creatine	34.376
2098	Family with Sequence Similarity 46 Member D	Cancer/Testis Antigens	32.419
1895	Chromosome 20 Open Reading Frame 85		30.557
6856	Ribonuclease / angiogenin Inhibitor 1	Rnase Inhibitor	29.209
2639	Cytochrome c Oxidase Subunit V _i c		24.909
5768			19.38
6676	Myeloid Leukemia Factor 1		18.959
17202			13.465
13352			9.971
5085			9.316
7629			6.759
16688			6.496

Red=genes showing d_n/d_s ratio > 1

Phylogenetic survey of reproductive genes across Fairy wren species

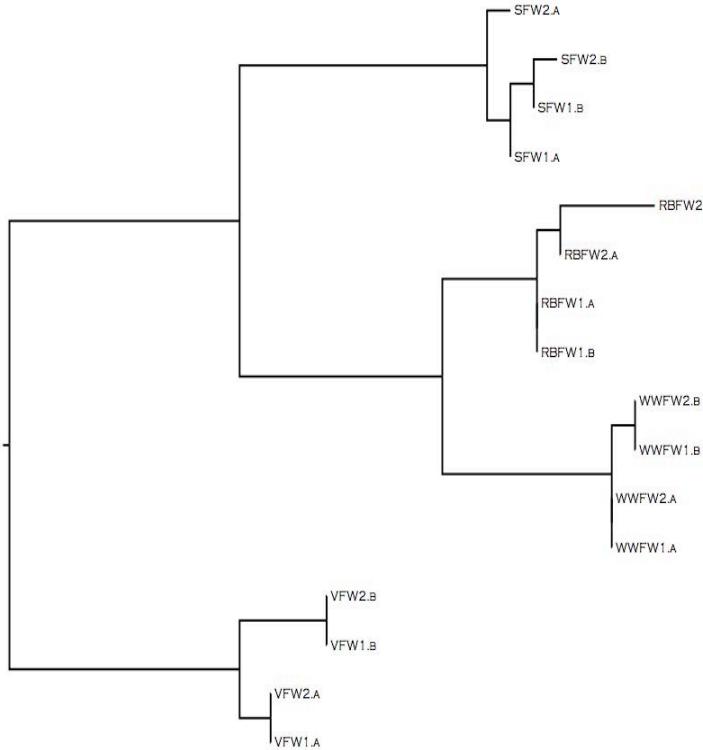


Incomplete lineage sorting is common in Fairy Wrens

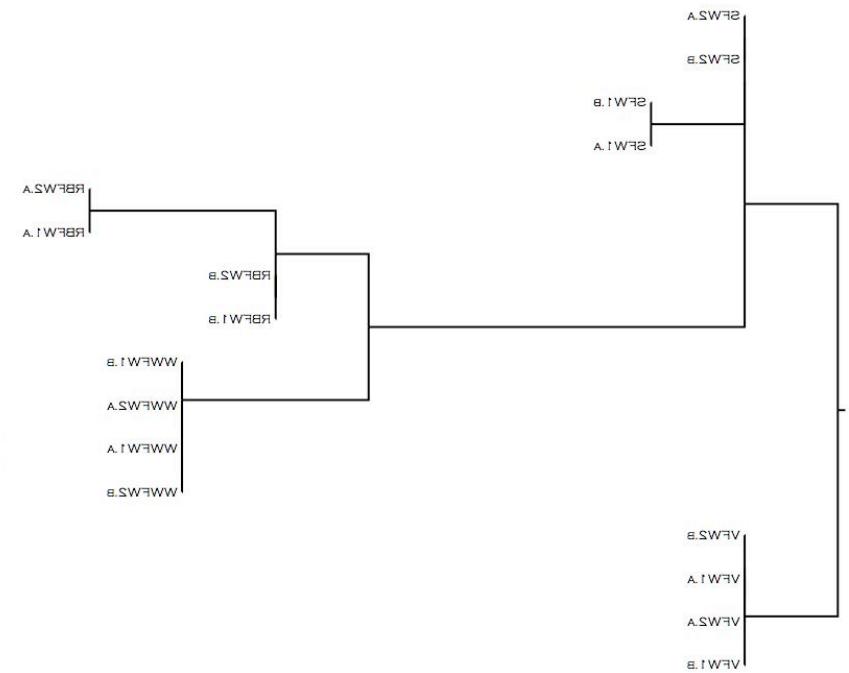


Examples of genes showing reciprocal monophyly among Fairy Wren species

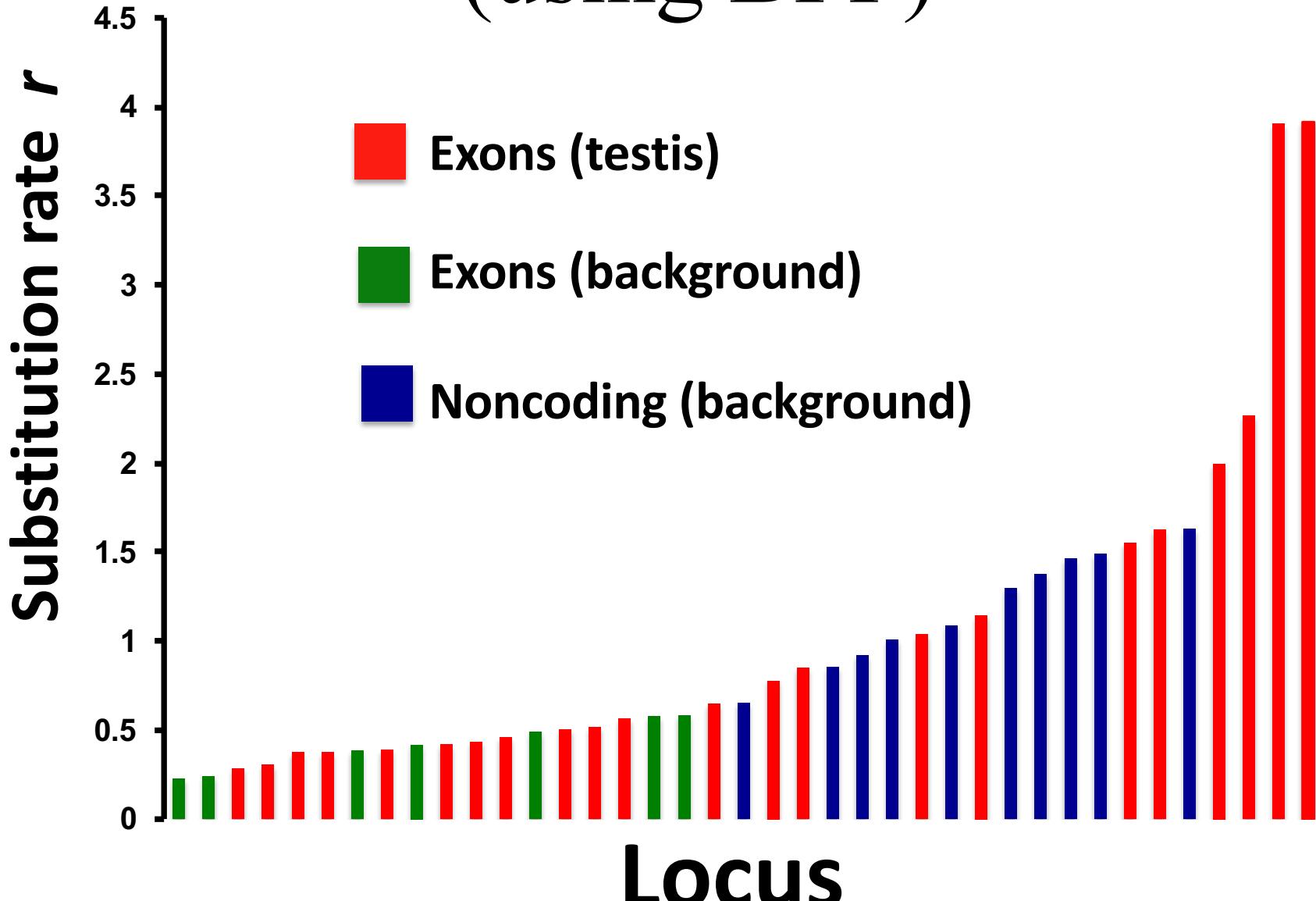
Proacrosin



Outer Dense Fiber Protein

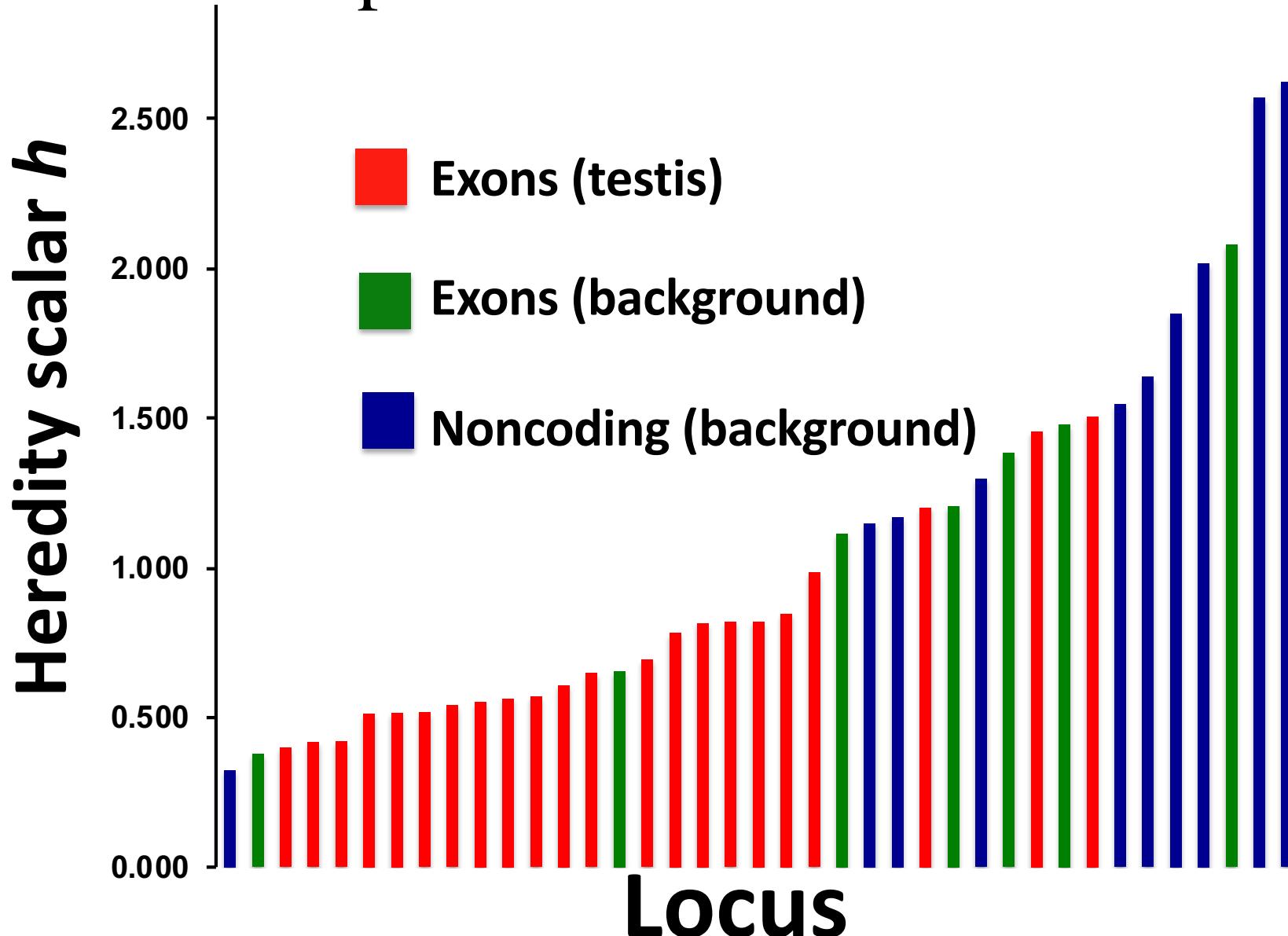


Rapid evolution of testis transcripts (using BPP)



BPP: Yang Z, Rannala B, 2010. *PNAS* 107: 9264–9269.

Low inheritance scalar of testis-transcripts: evidence for selection?



BPP: Yang Z, Rannala B, 2010. PNAS 107: 9264–9269.