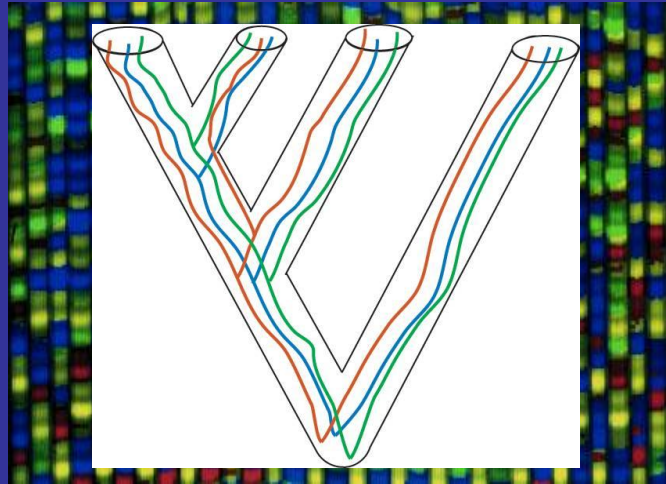


The phylogeography-phylogenetics continuum



Scott V. Edwards

Department of Organismic and Evolutionary Biology

Harvard University

Cambridge, MA USA

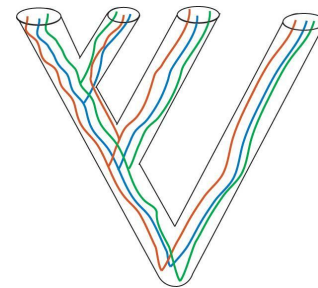
<http://www.oeb.harvard.edu/faculty/edwards>

The phylogeography-phylogenetics continuum

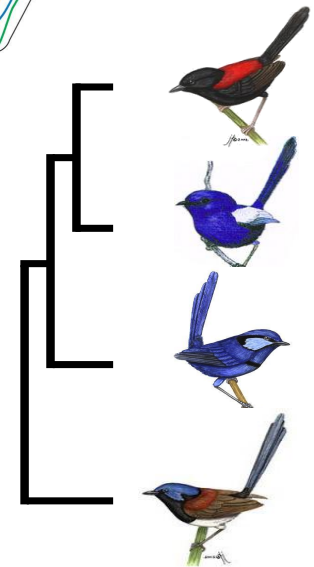
- **Multilocus phylogeography** across the Carpentarian barrier



- **Species trees:** just another phylogeographic model



- How **natural selection** can create novel patterns in gene trees



Concordance of geographic ranges of Australian songbirds



Brown
Treecreeper

+



Black-tailed
Treecreeper

=



Grey-crowned
Babbler



Black-throated
Finch

+



Long-tailed
Finch

=



Red-backed
Wren



Eastern Spinebill

+



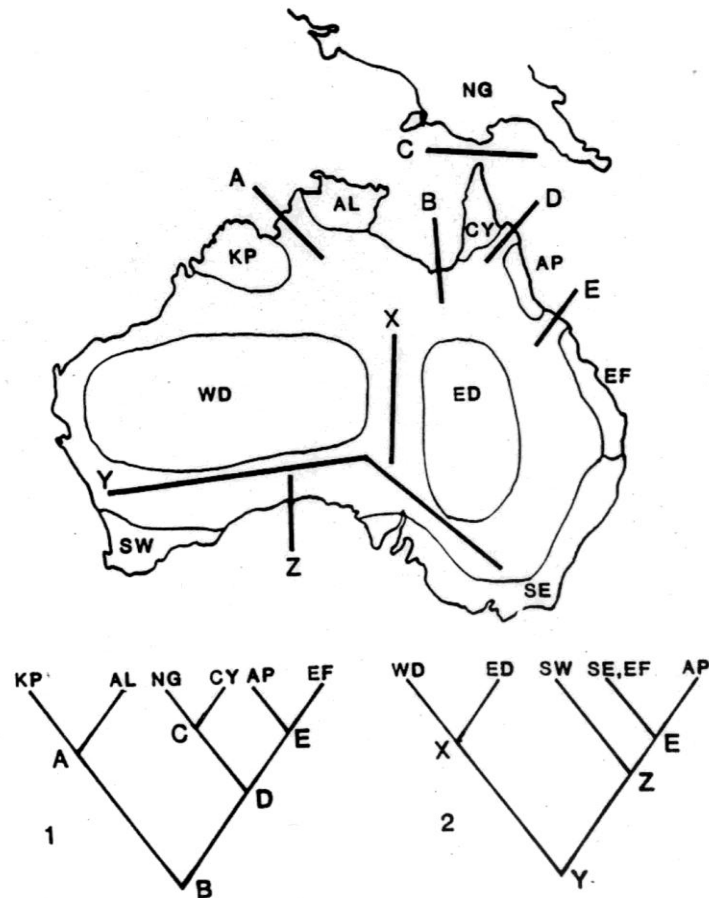
Western Spinebill

=

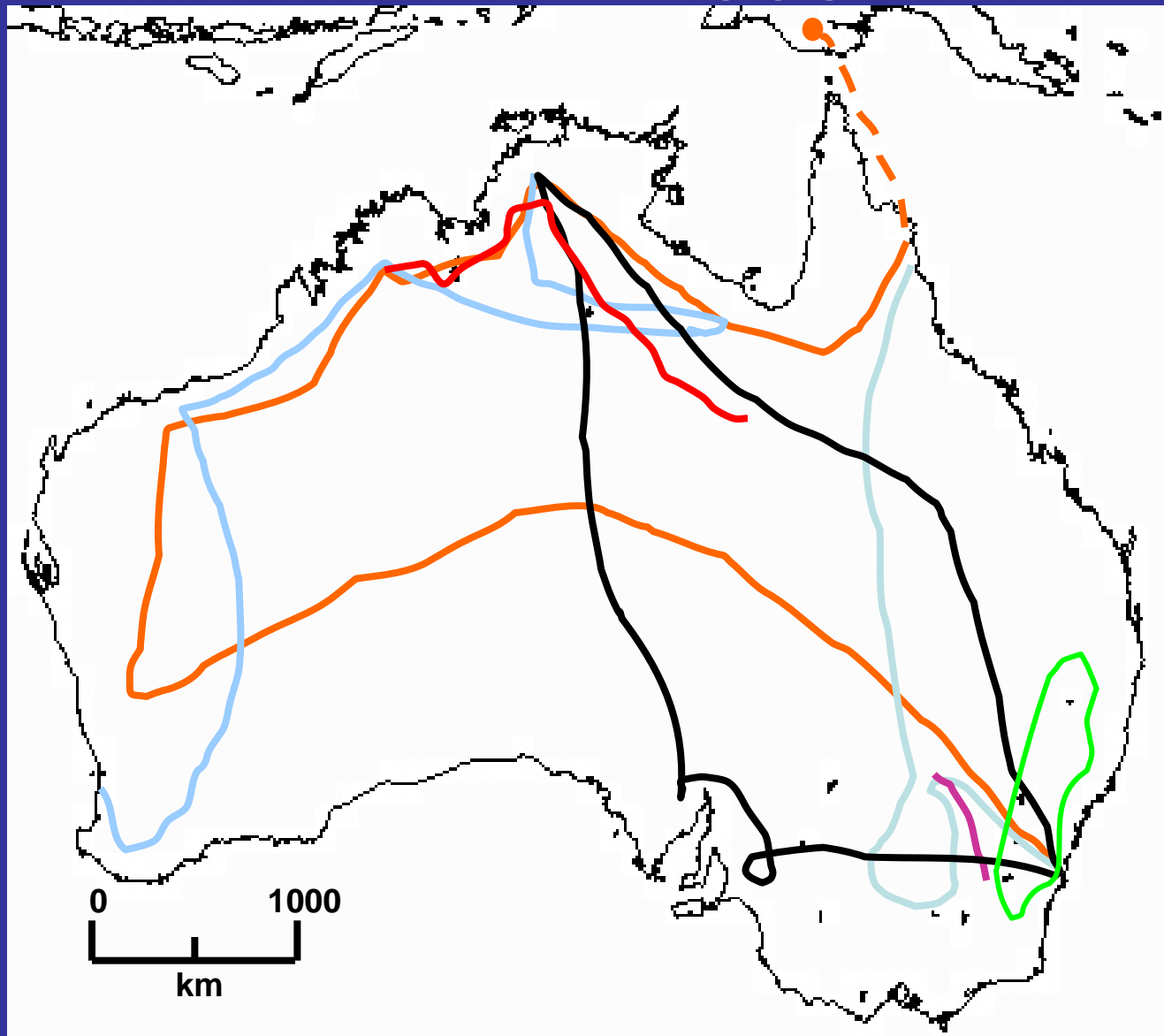


Golden Whistler

Carpentarian barrier (B) is deepest split in area cladograms



Australia expeditions, 1987 - 2005



- 1987
- 1990
- 1996
- 1997
- 2002a
- 2002b
- 2005

Those welcoming Aussies...



Case studies



Grassfinches (*Poephila*)
W. Bryan Jennings



Red-backed Wren
(*Malurus*
***melanocephalus*)**
June Lee



Treecreepers (*Climacteris*)
Nancy Rotzel



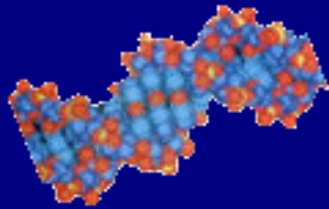
Zebra Finches (*Taenopygia guttata*)
Chris Balakrishnan

Anonymous loci: advantages over microsatell

- Mutational scale directly comparable to mtDNA
- Mutational homoplasy is minimal
- Gene trees easily constructed
- Diversities easily compared across species

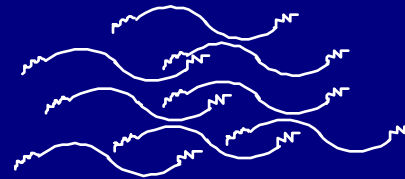
Obtaining anonymous loci

Genomic DNA

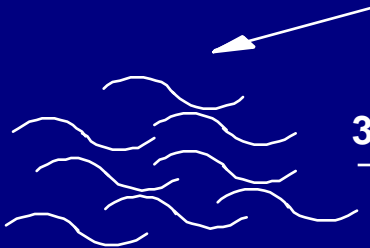


1. sonication

1-2 kb sheared fragments



2. repair ragged ends; make blunt

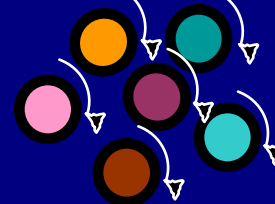


3. clone blunt fragments

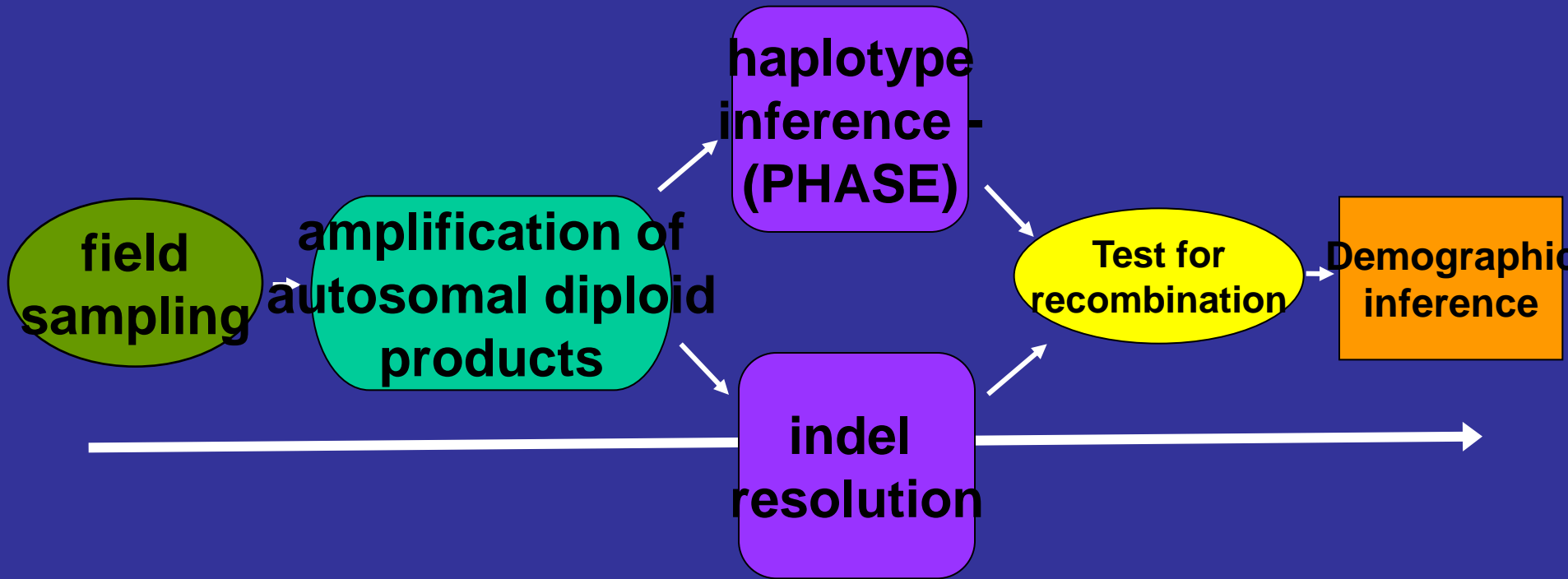
1-2 kb insert

pUC 18 vector

4. sequence random clones



Pipeline for multilocus data analysis

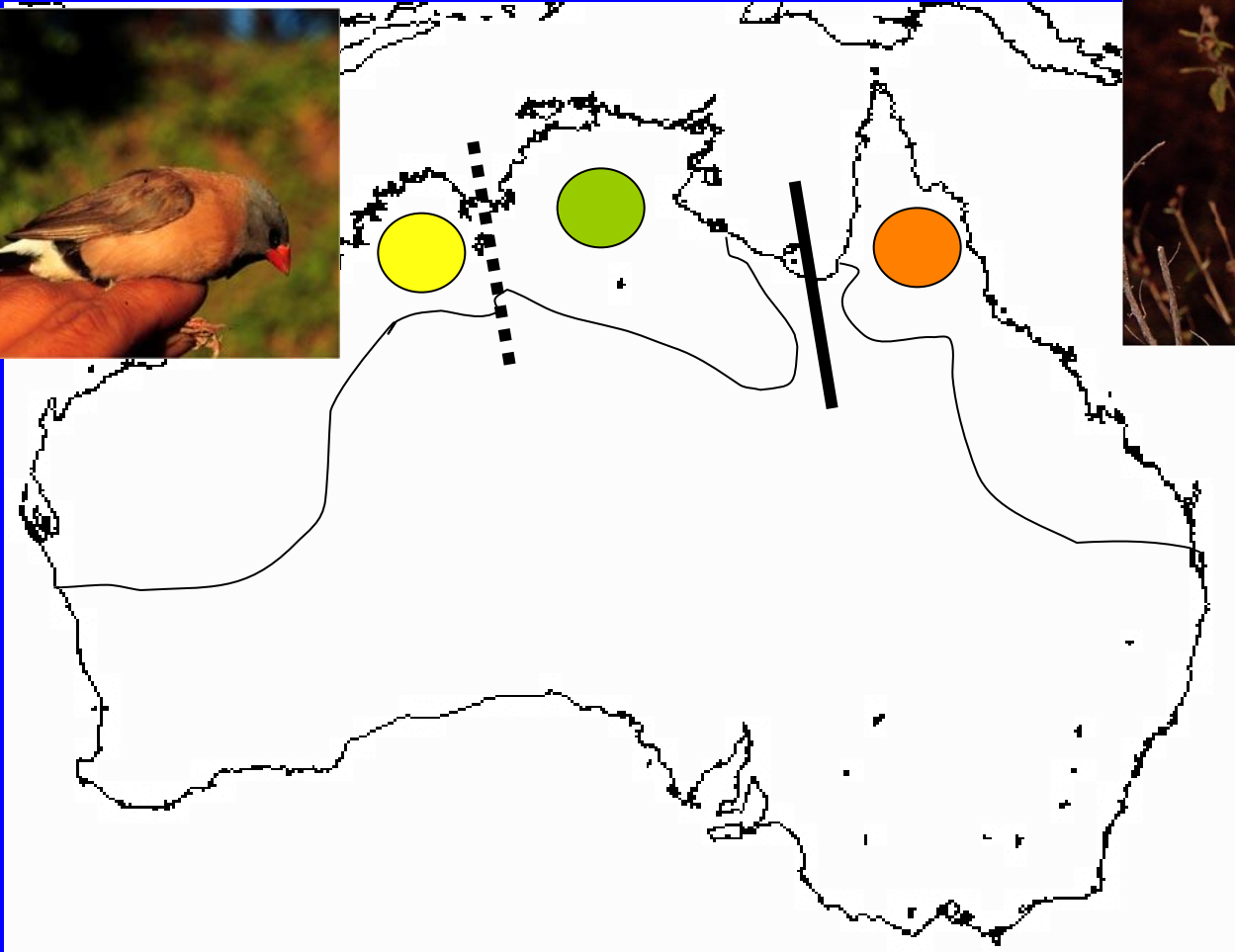


East-west pairs -- grassfinches (*Poephila*)

P. acuticauda

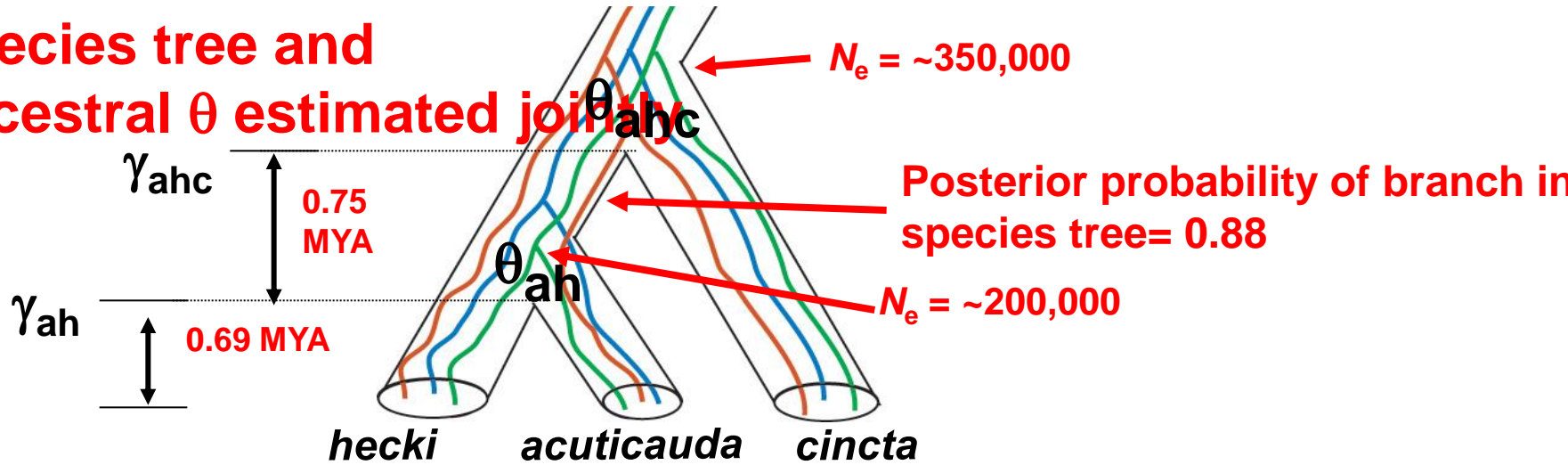
P. hecki

P. cincta



Species tree of *Poephila* grassfinches

Species tree and
ancestral θ estimated



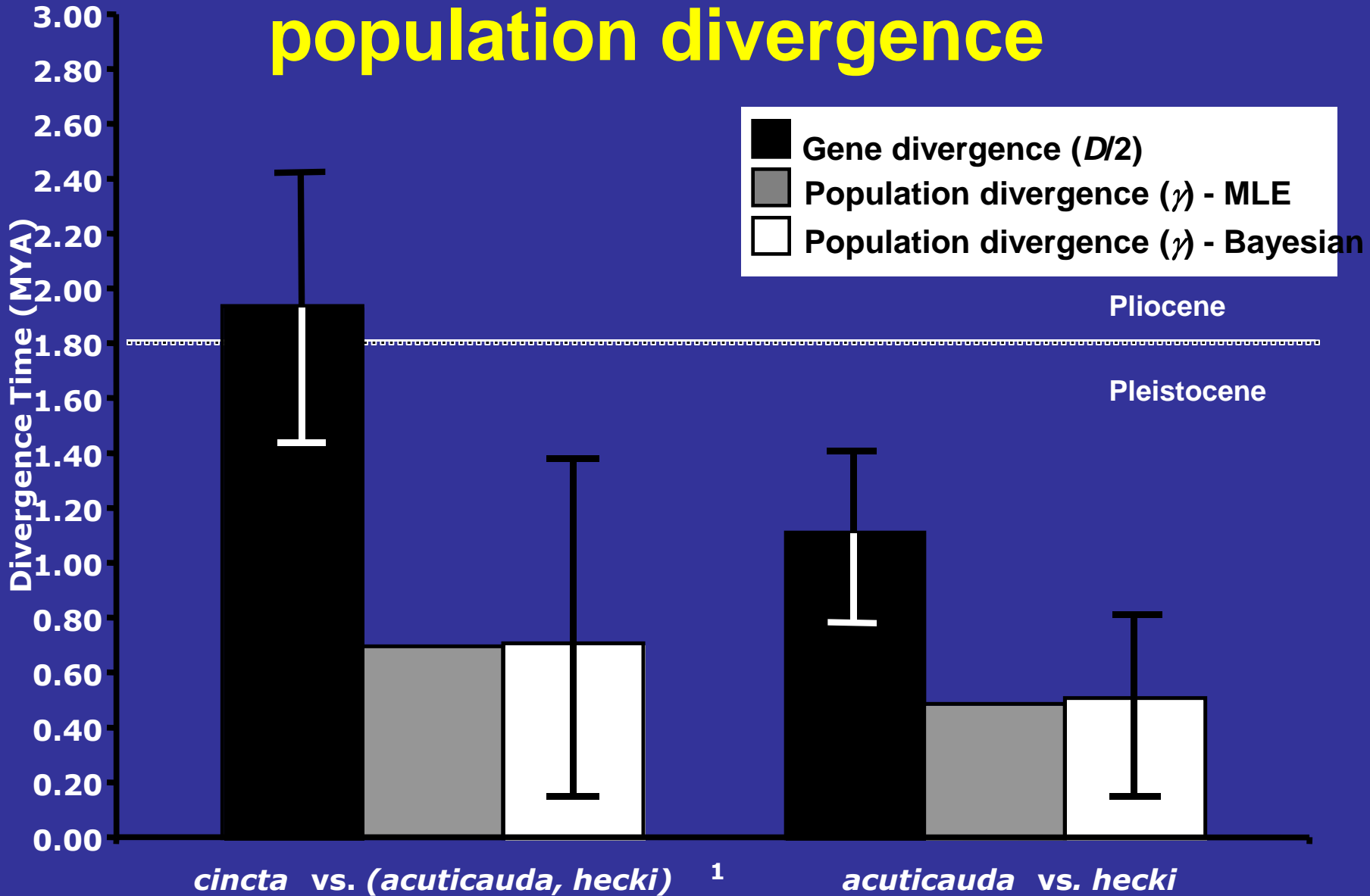
Long-tailed Finch



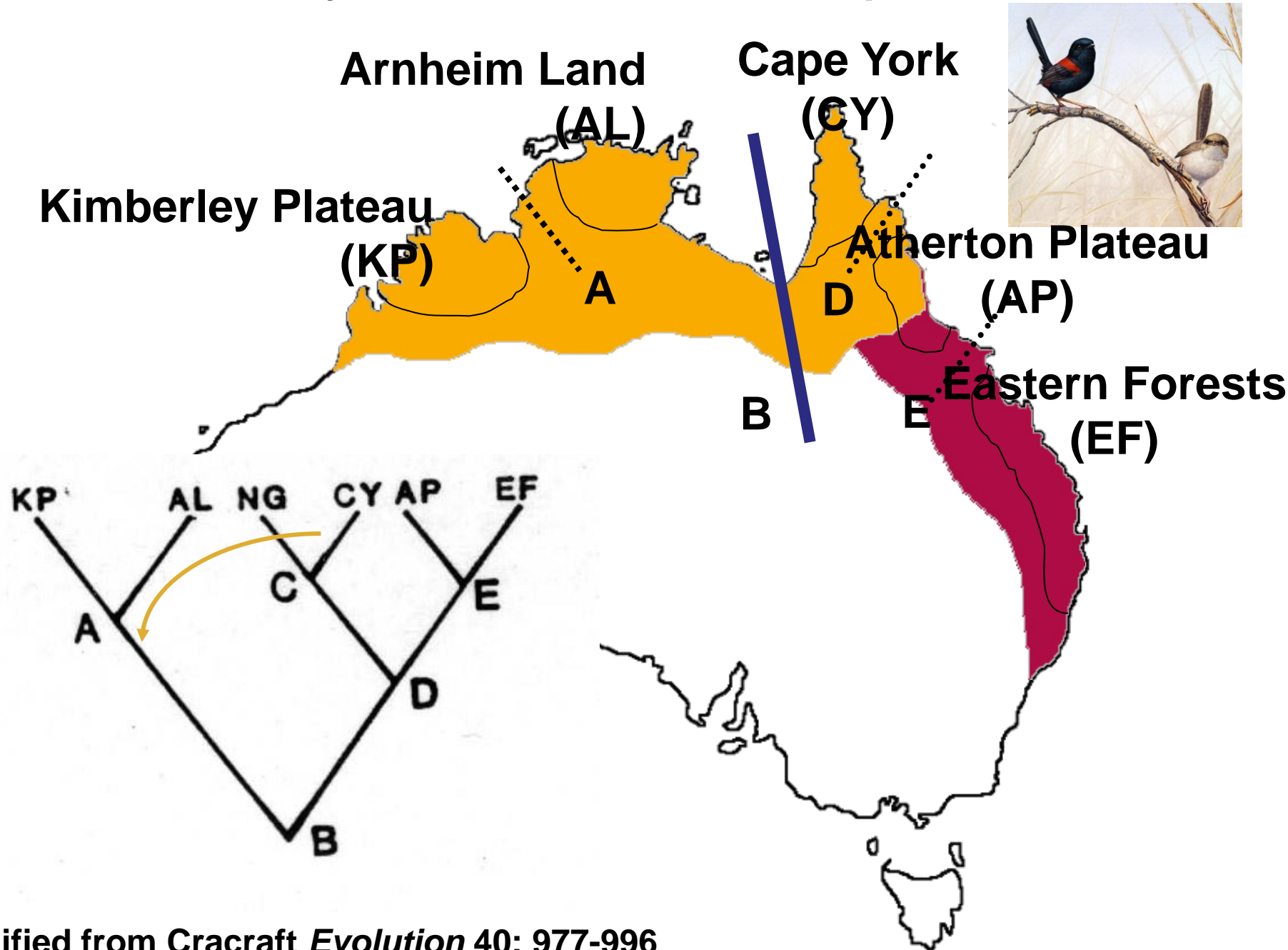
Black-throated Finch



Gene divergence substantially predates population divergence



Red-backed Fairy wren - discordant subspecies boundaries



Modified from Cracraft *Evolution* 40: 977-996

Population Assignment - Structure

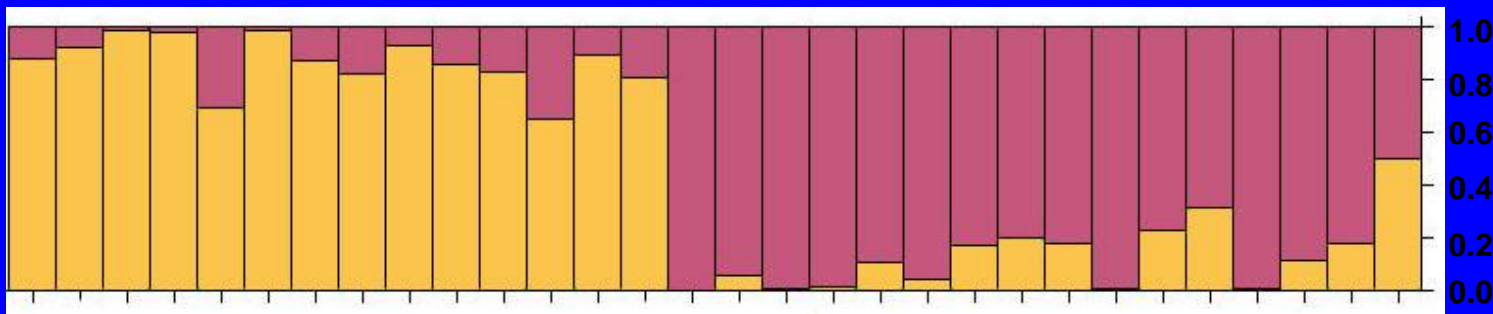
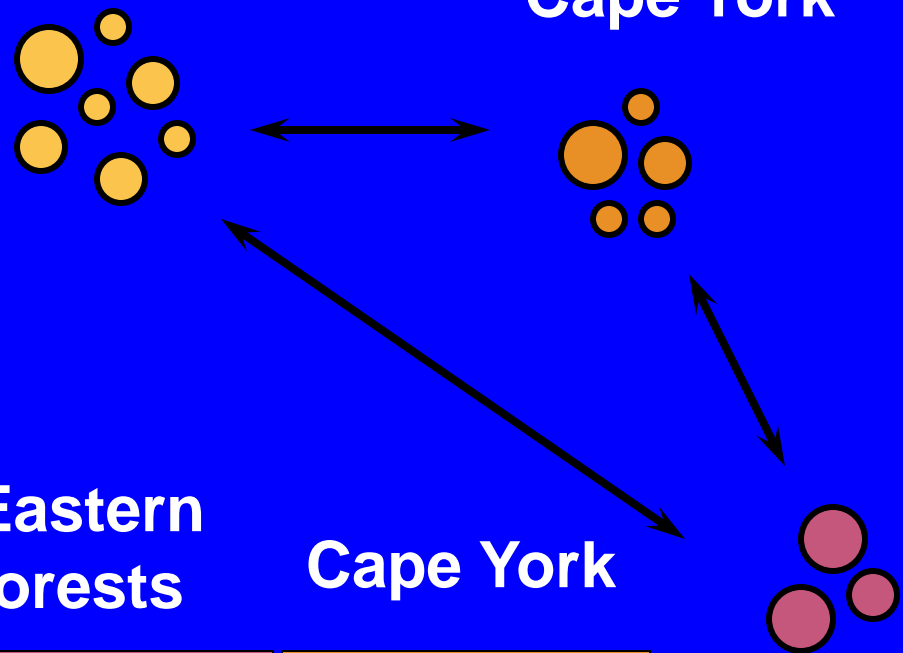
| K | Ln P(D) | Var[LnP(D)] |
|---|----------|-------------|
| 1 | -22842.8 | 805.2 |
| 1 | -22847.2 | 813.6 |
| 1 | -22844.8 | 809.7 |
| 2 | -22062.8 | 2148.2 |
| 2 | -22077.3 | 2176.7 |
| 2 | -22060.1 | 2135.9 |
| 3 | -28951.7 | 14967.9 |
| 4 | -26786.3 | 10266.9 |
| 5 | -32129.0 | 22148.8 |

Kimberley

**Eastern
forests**

Cape York

Cape York



**Eastern
forests**

Treecreepers (*Climacteris*)

Black-tailed treecreeper

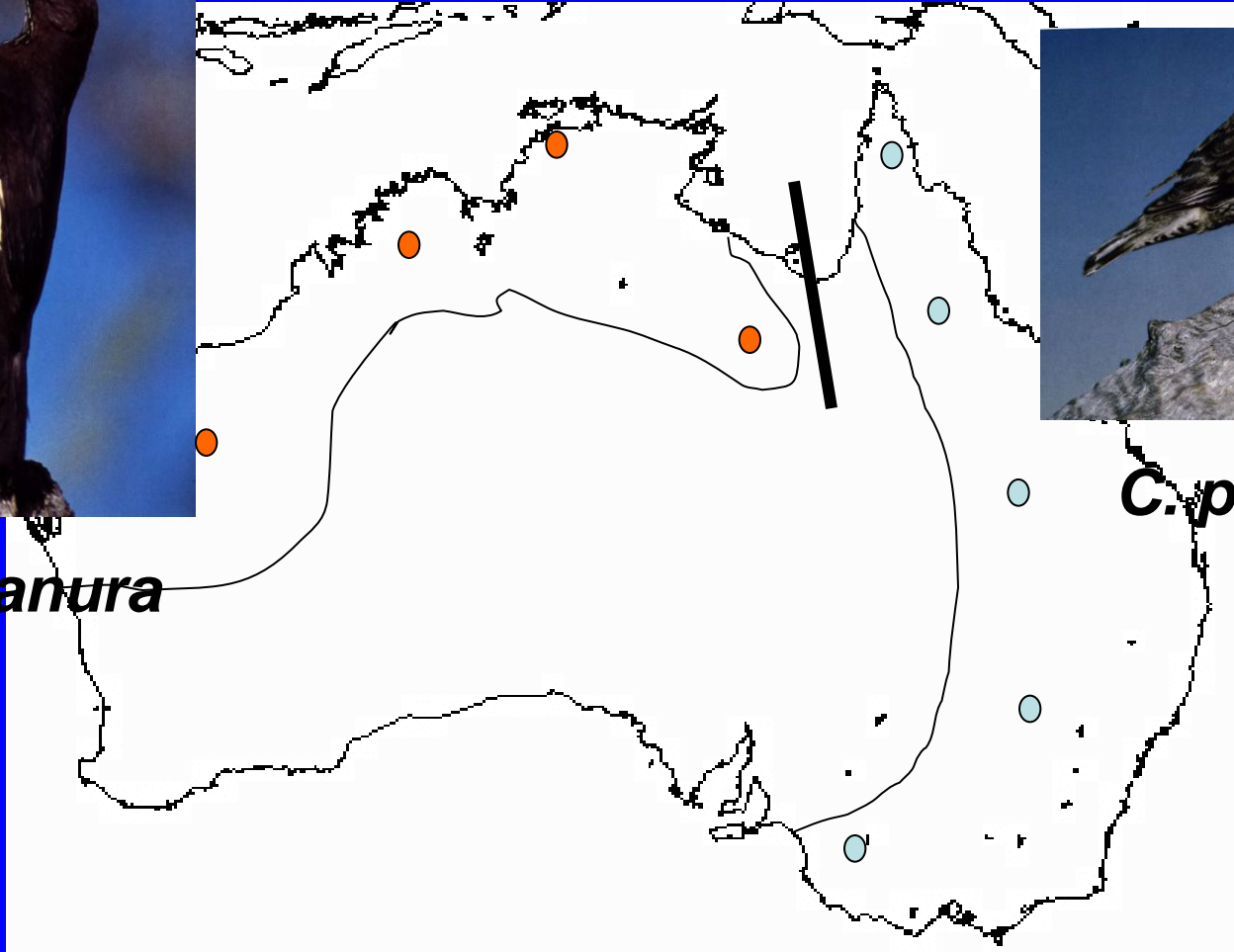


Brown treecreeper



C. melanura

C. picumnus



12 gene trees in Australian treecreepers

Brown Treecreeper
(eastern) lineages

Black-tailed (western)
Treecreeper lineages

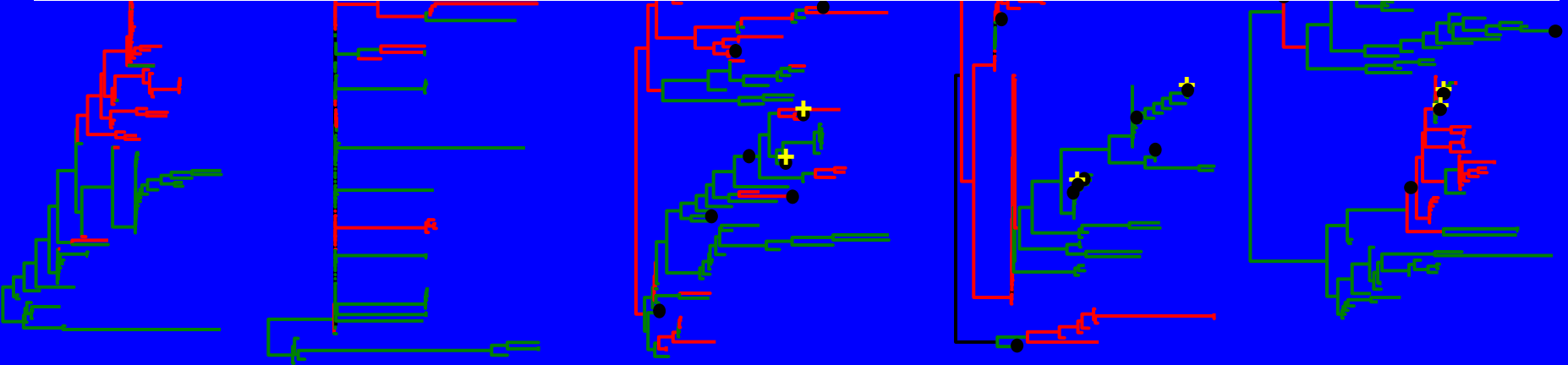
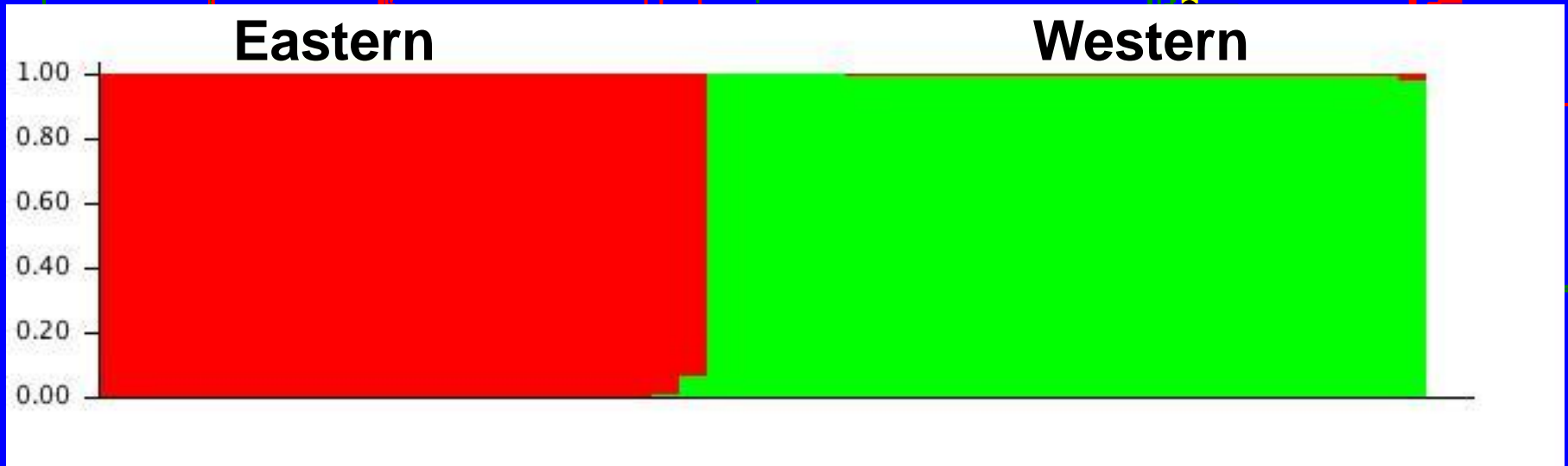
AL3

AL5

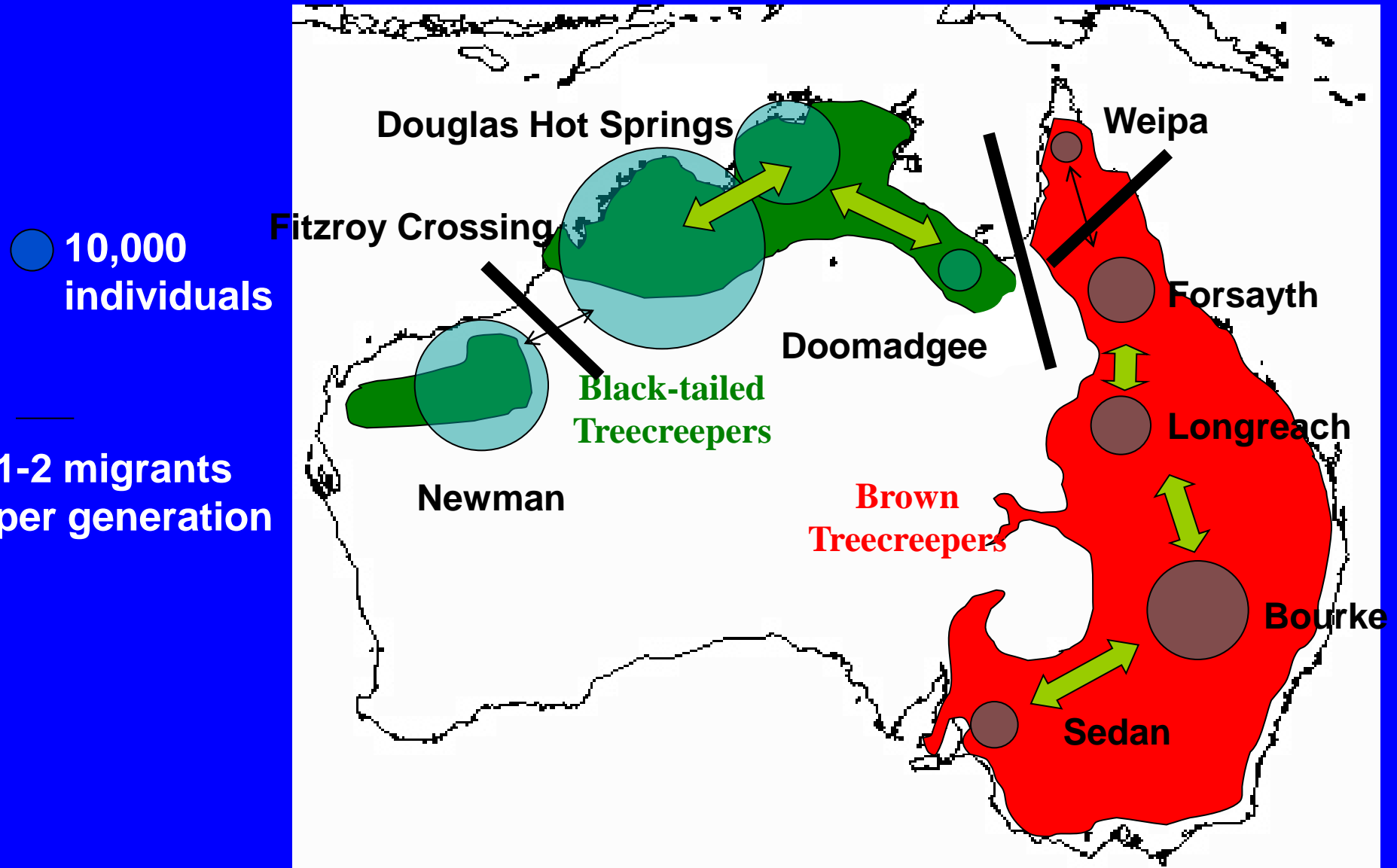
AL7

AL14

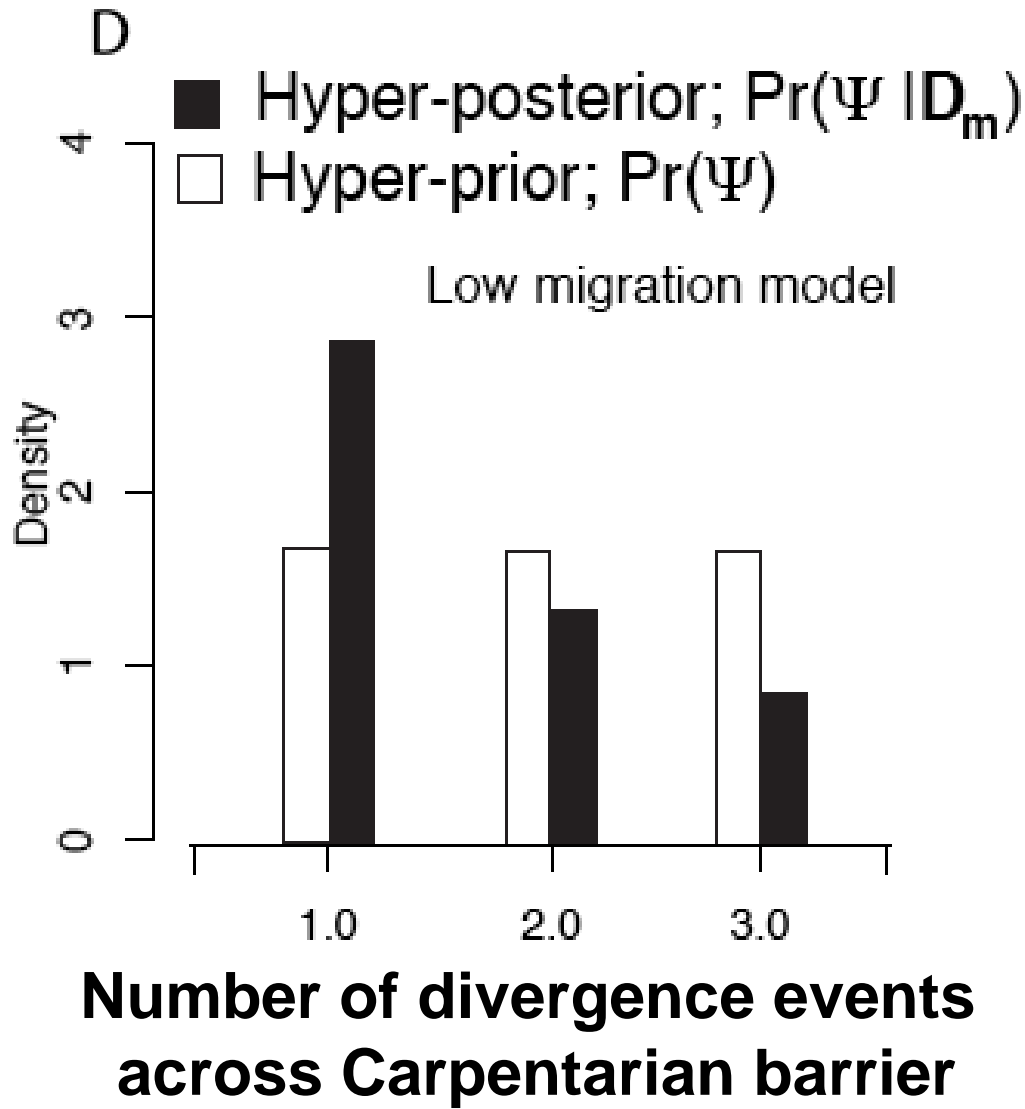
AL16



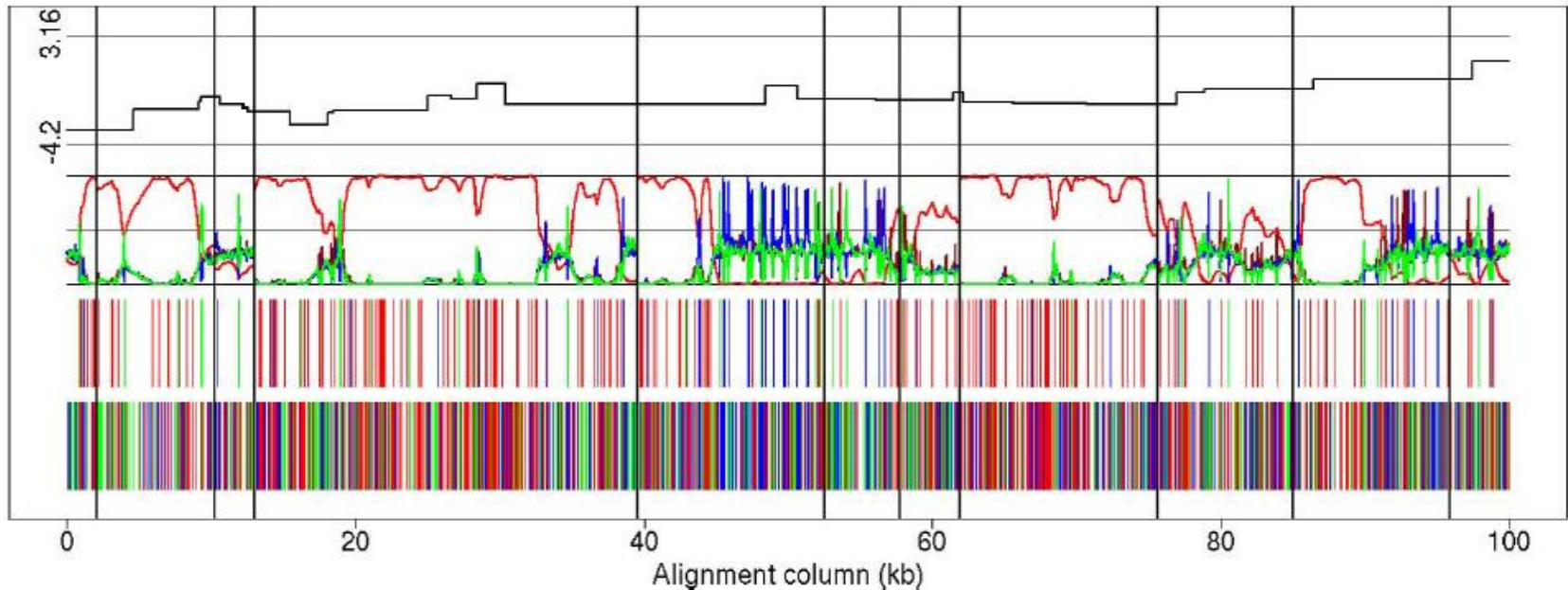
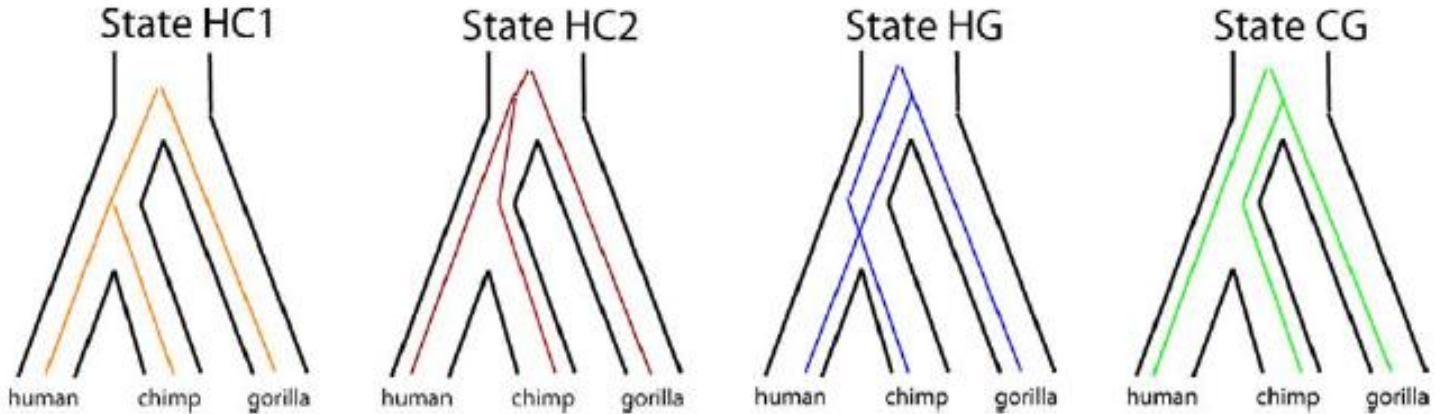
Treecreeper populations are connected but variable in size (MIGRATE)



Test of contemporaneous divergence across Carpentarian barrier using msBayes

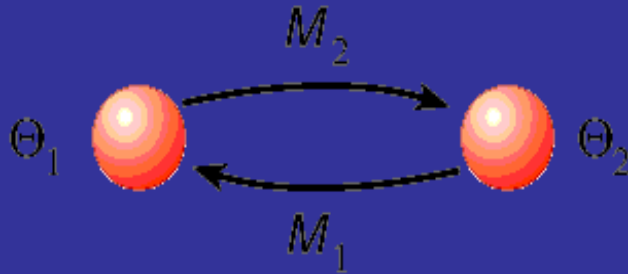


Regions of fixed differences between primate species



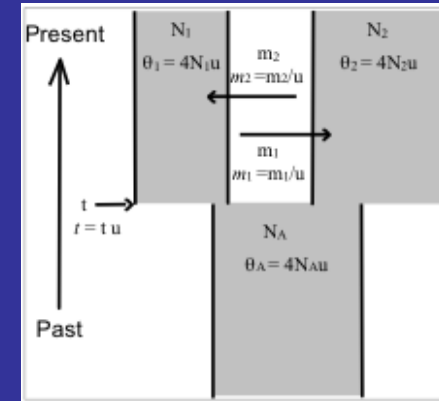
Alternative models of population history

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



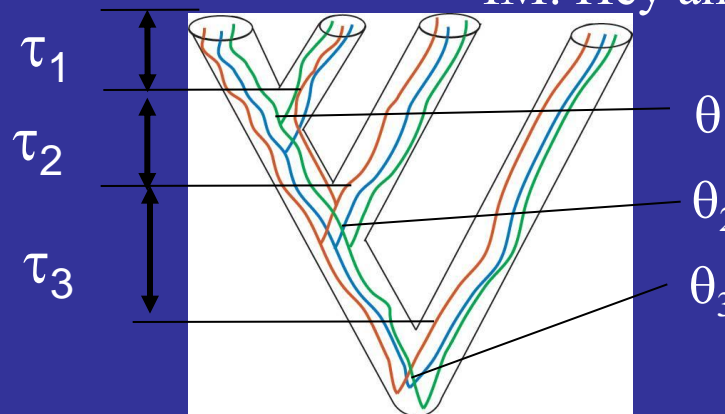
equilibrium migration model

MIGRATE: Beerli 2006 *Bioinformatics*



isolation-migration model

IM: Hey and Nielsen 2004 *Genetics*

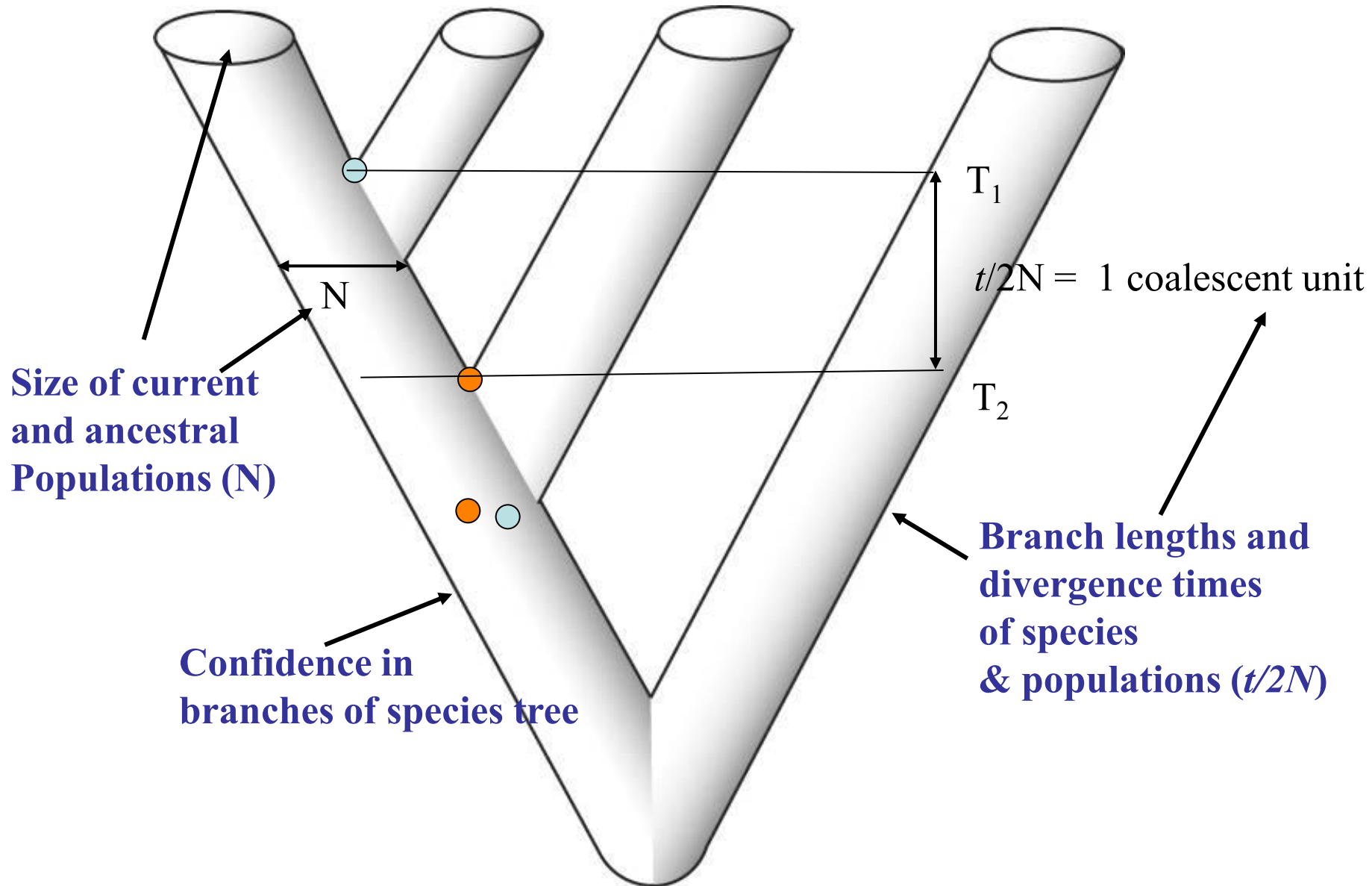


pure isolation
(phylogeny)

MCMCcoal: Rannala + Yang 2003 *Genetics*

BEST: Liu and Pearl, *Syst. Biol.* *In press*

Properties of species trees



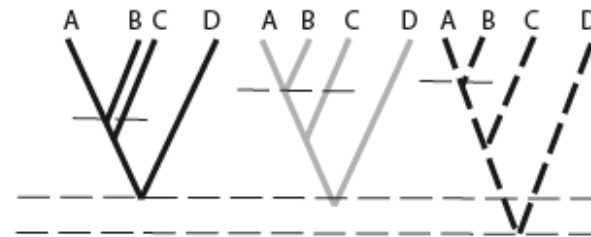
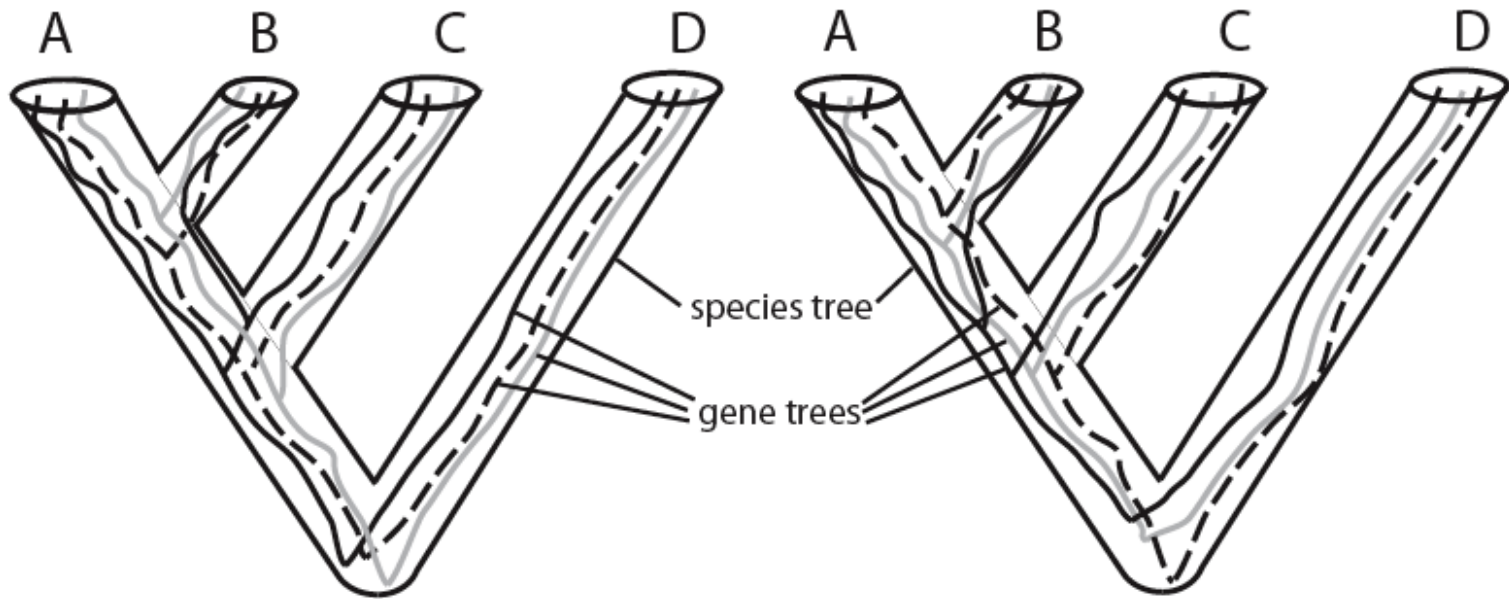
Justifications for species trees

- ***Current paradigm:*** gene trees & concatenation
- ***Philosophical justification:*** what are we estimating and how do we combine data in systematics?
- ***Improved model:*** better describes patterns in sequence data and the relationship between gene and species trees
- ***Performance:*** outperforms concatenation in some areas of tree space and gives more accurate confidence than does concatenation

Deep coalescence vs. branch length heterogeneity

Deep coalescence

Branch length heterogeneity



Species tree approaches span all major statistical methods

Bayesian

BEST

Liu & Pearl. 2007. *Syst. Biol.* 56: 504-514

BEAST*

Heled & Drummond. 2010. *Mol. Biol. Evol.* 27: 570-580.

Maximum Likelihood

MP-EST

Liu, et al. 2010. *BMC Evol. Biol.* (in press)

Mixed methods

STAR/STEAC

Liu, et al. 2009. *Syst. Biol.* 58: 468-477

Parsimony

Deep coalescence

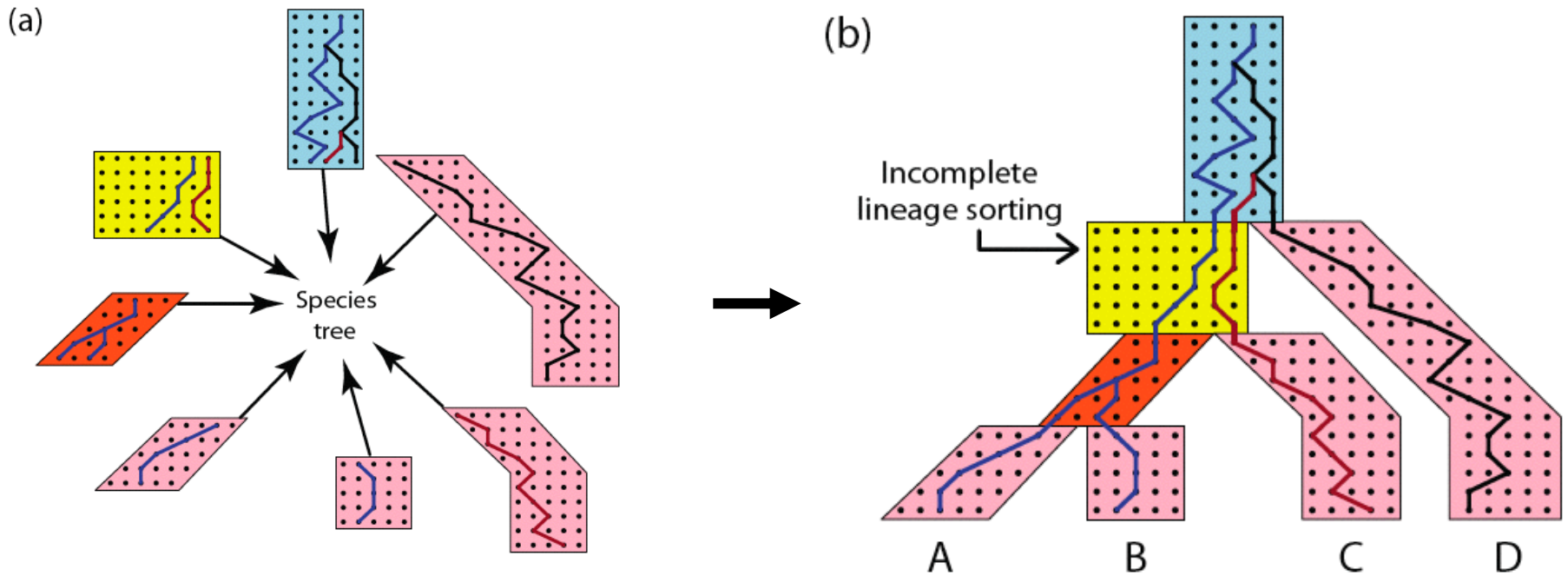
Maddison & Knowles. 2006. *Syst. Biol.* 55: 21-30.

Summary statistics

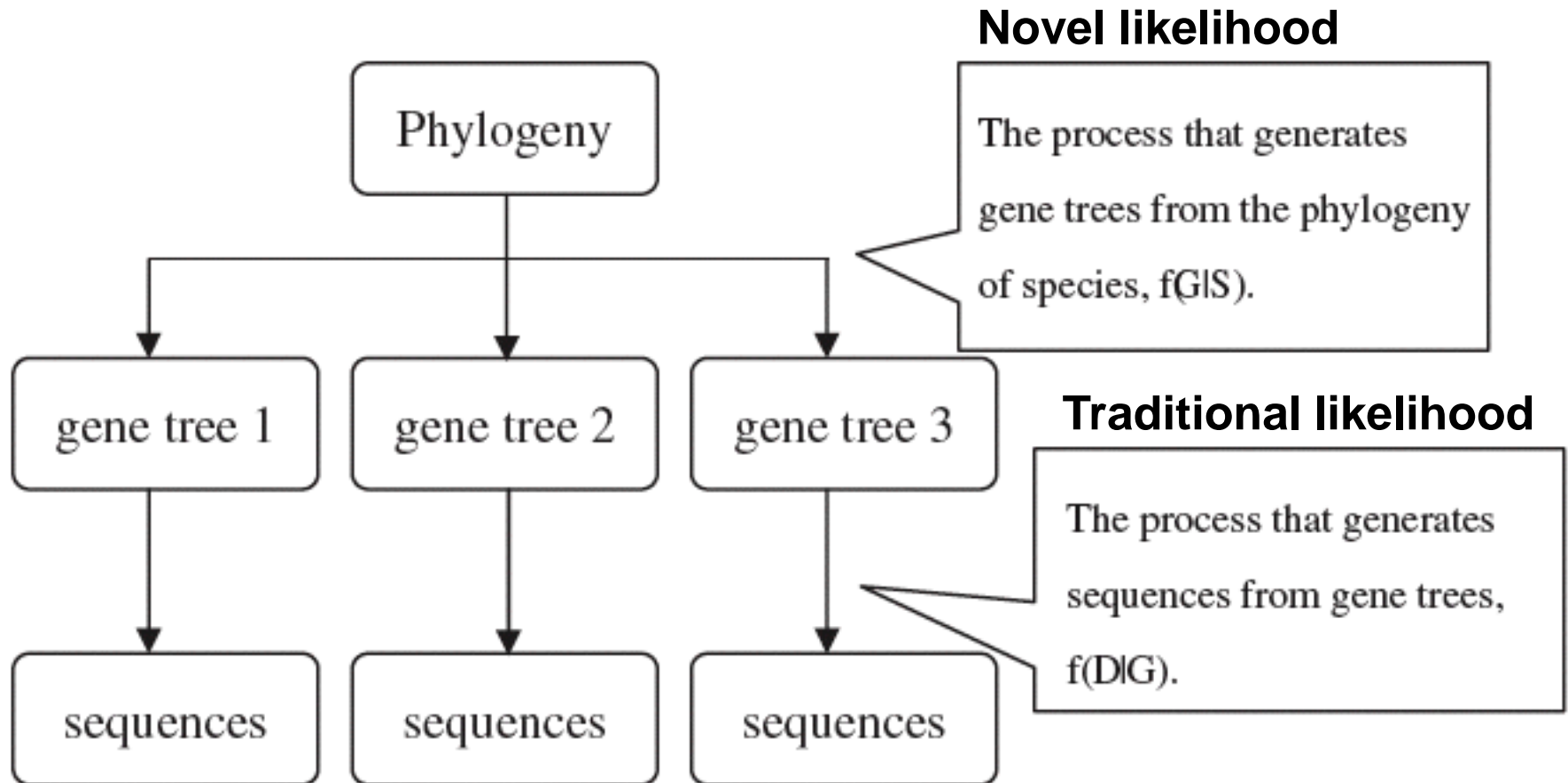
Maximum Tree

Liu, Yu and Pearl. 2009. *J Math Biol.* 60:95-106.

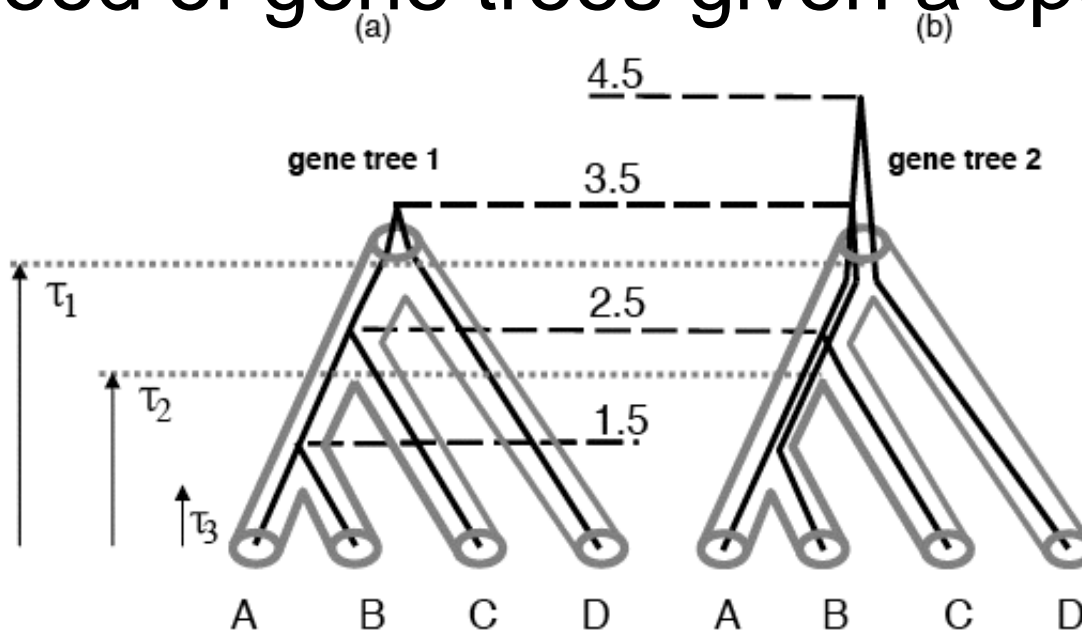
The multispecies coalescent



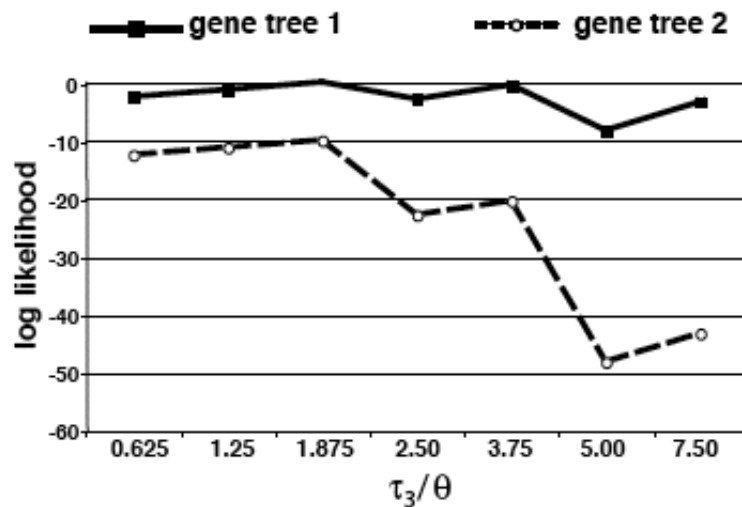
A tale of two likelihoods



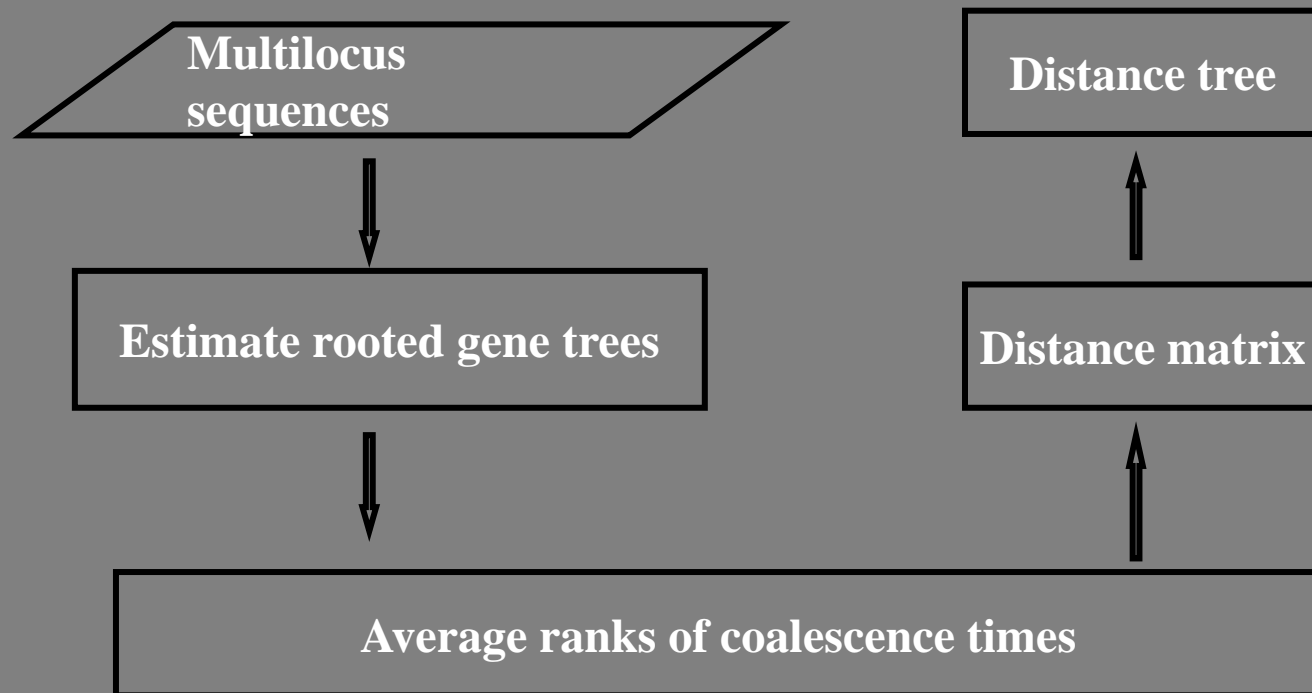
Likelihood of gene trees given a species tree



(c)

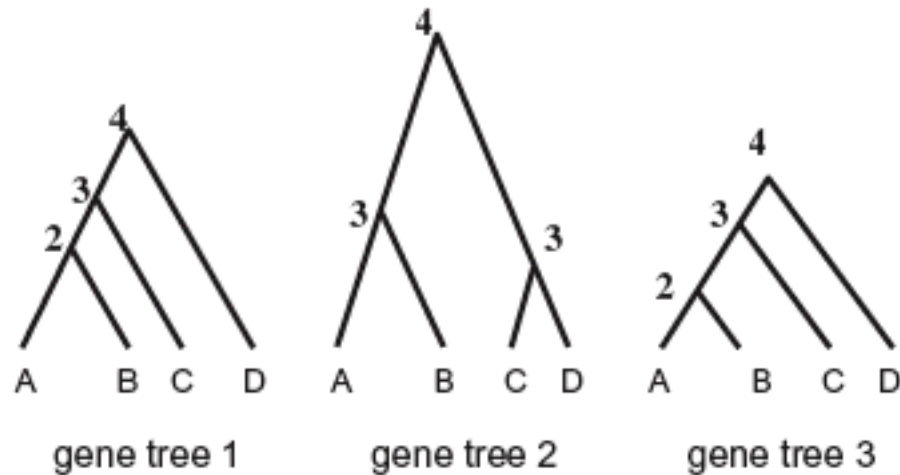


Species Trees from Average Ranks of Coalescence Times (STAR)



Calculating a STAR tree

(a)



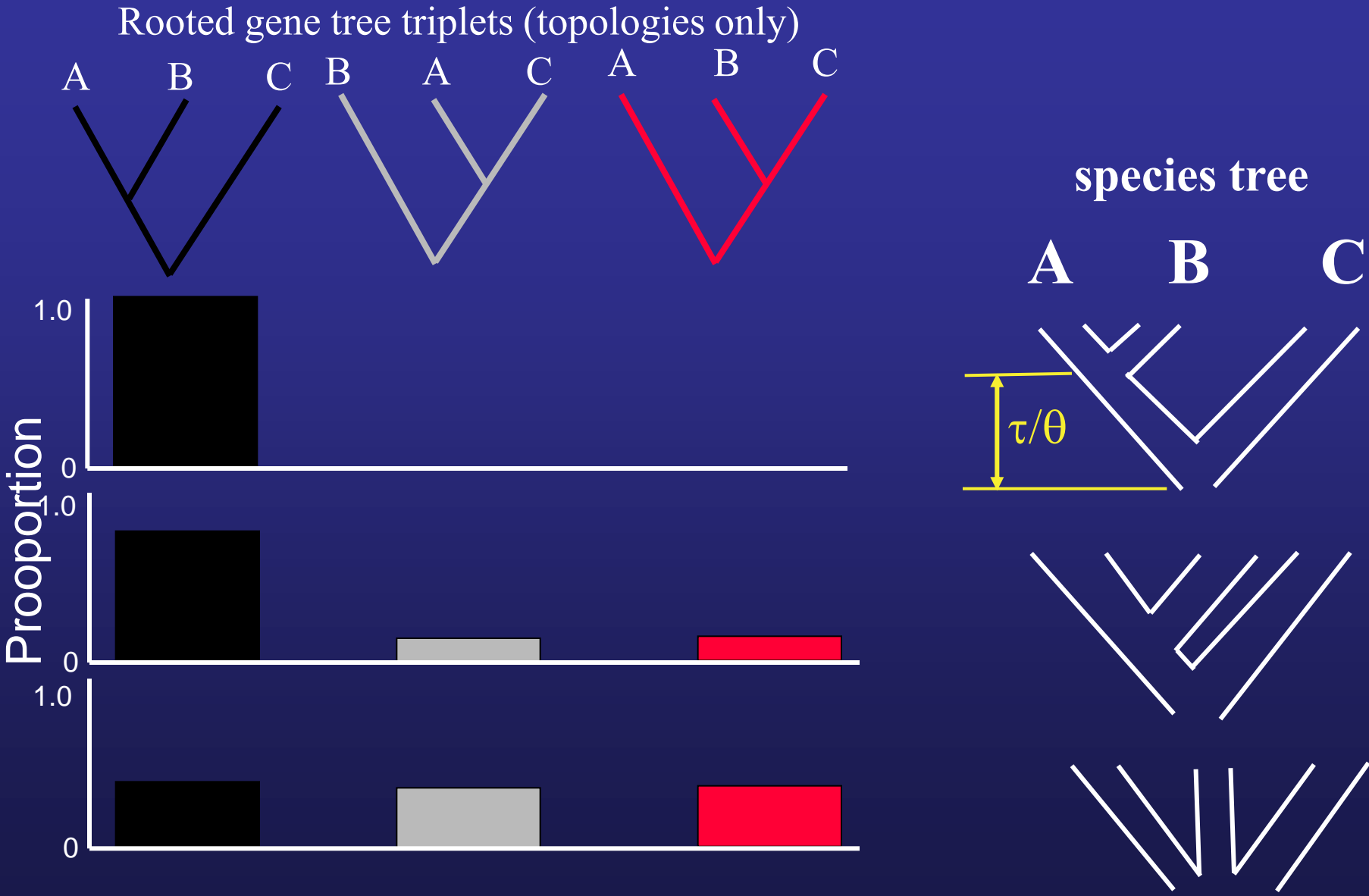
(c)

| | A | B | C | D |
|---|------|------|------|------|
| A | -- | 4.66 | 6.66 | 8 |
| B | 4.66 | -- | 6.66 | 8 |
| C | 6.66 | 6.66 | -- | 7.34 |
| D | 8 | 8 | 7.34 | -- |



topology of the STAR tree

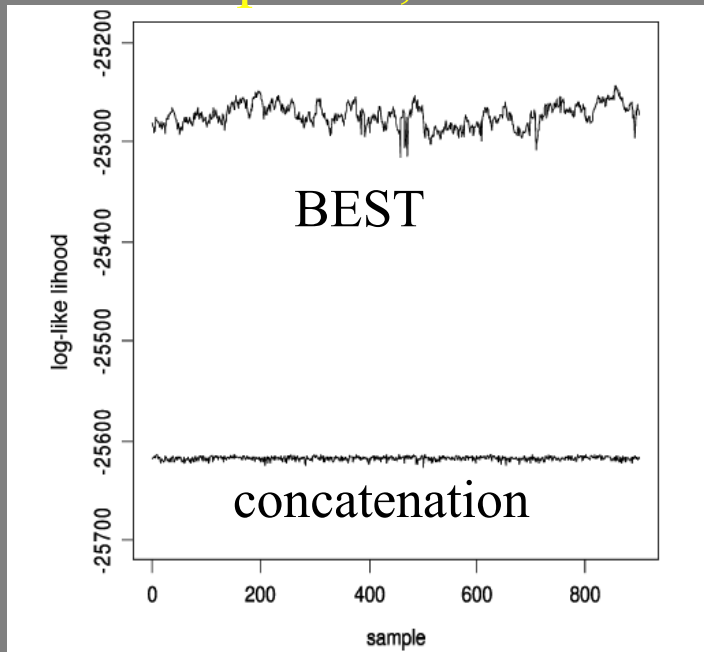
Maximum (pseudo) likelihood method for species trees



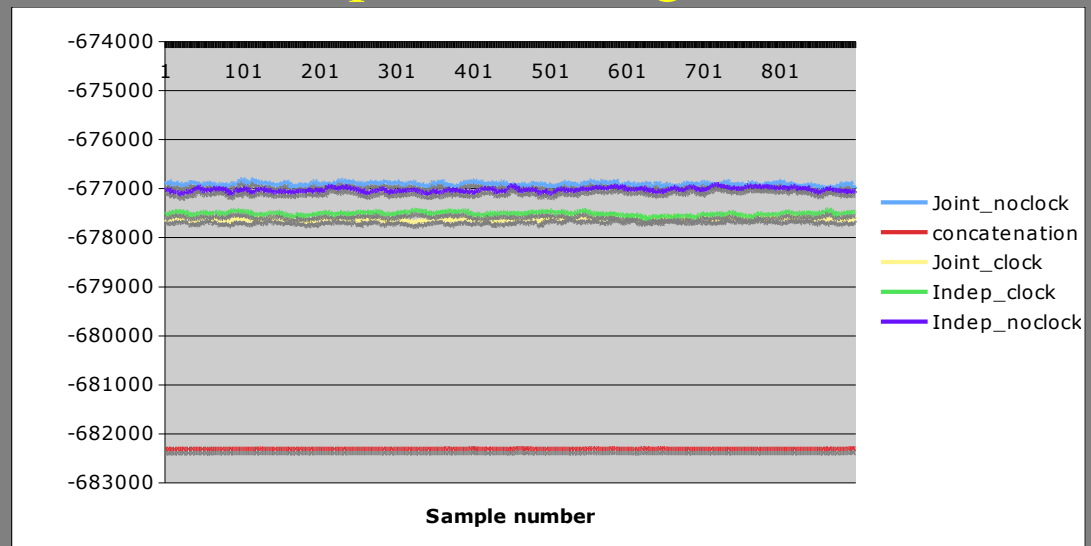
Is the coalescent model a better fit to the data than concatenation?

Yes

Australian grassfinches
4 species, 30 loci



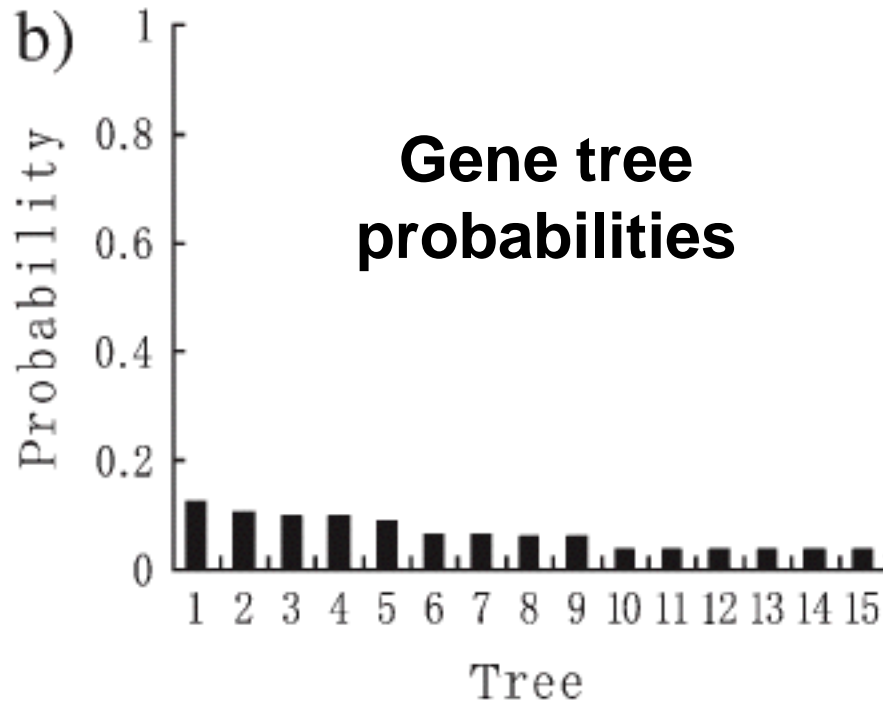
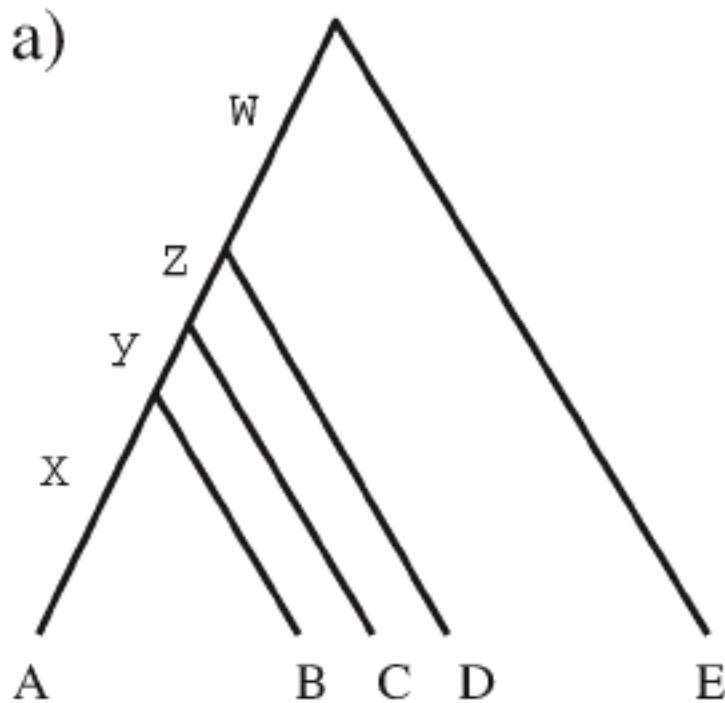
Yeast data
8 species, 106 genes



Liu & Pearl, D. K. (2007) *Syst. Biol.* 56, 504-14; Edwards, Liu and Pearl. 2007. *PNAS* 104:5936
Jennings, W. B. & Edwards, S. V. (2005) *Evolution* 59, 2033-2047; Rokas et al. 2003. *Nature* 425: 798-804.

The anomaly zone

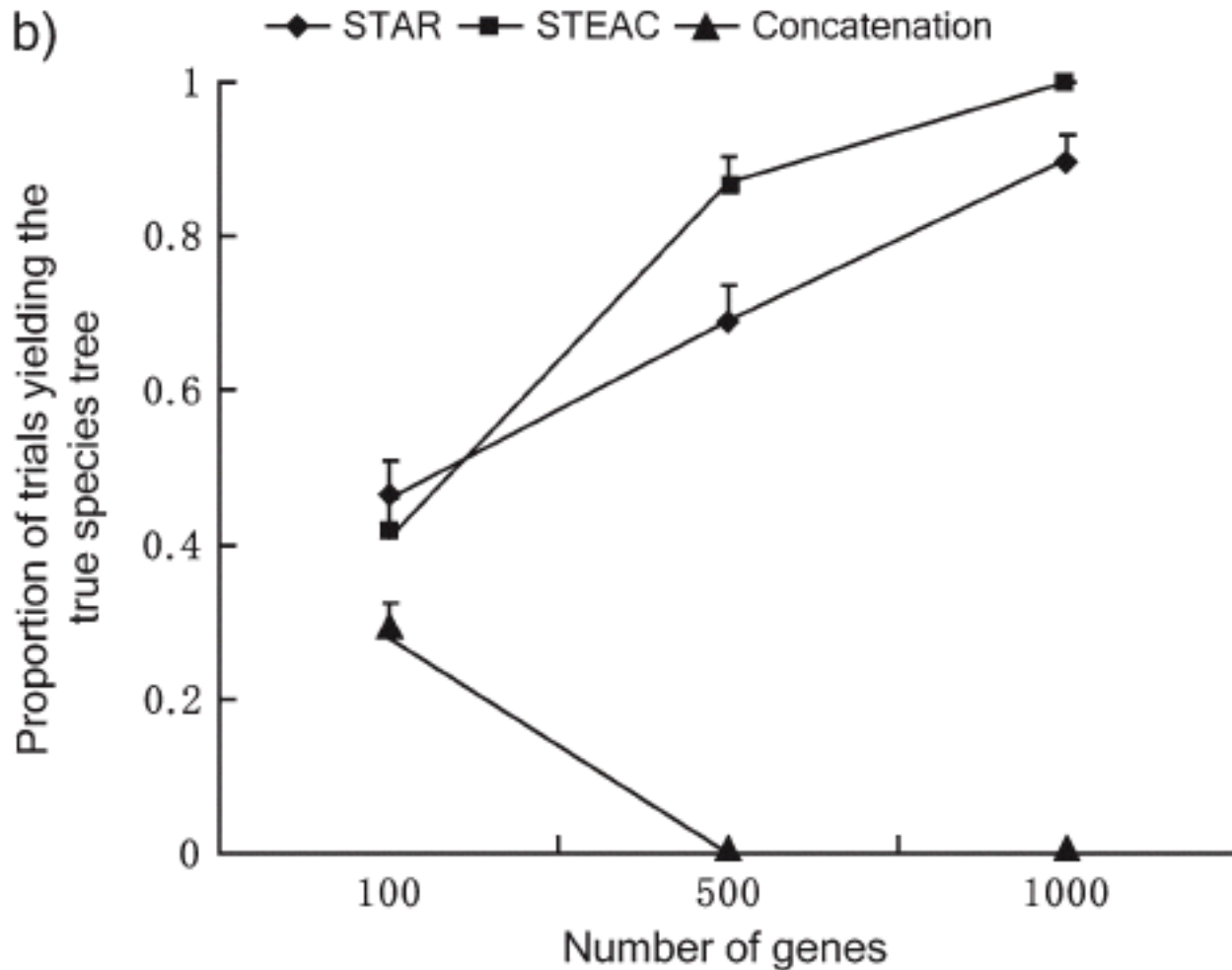
Species tree



Gene trees



Phylogenetic analysis in the anomaly zone



Australo-Papuan Fairy Wrens - Maluridae



Splendid Fairywren (*Malurus*)



Superb Fairywren (*Malurus*)



Striated Grasswren (*Amytornis*)



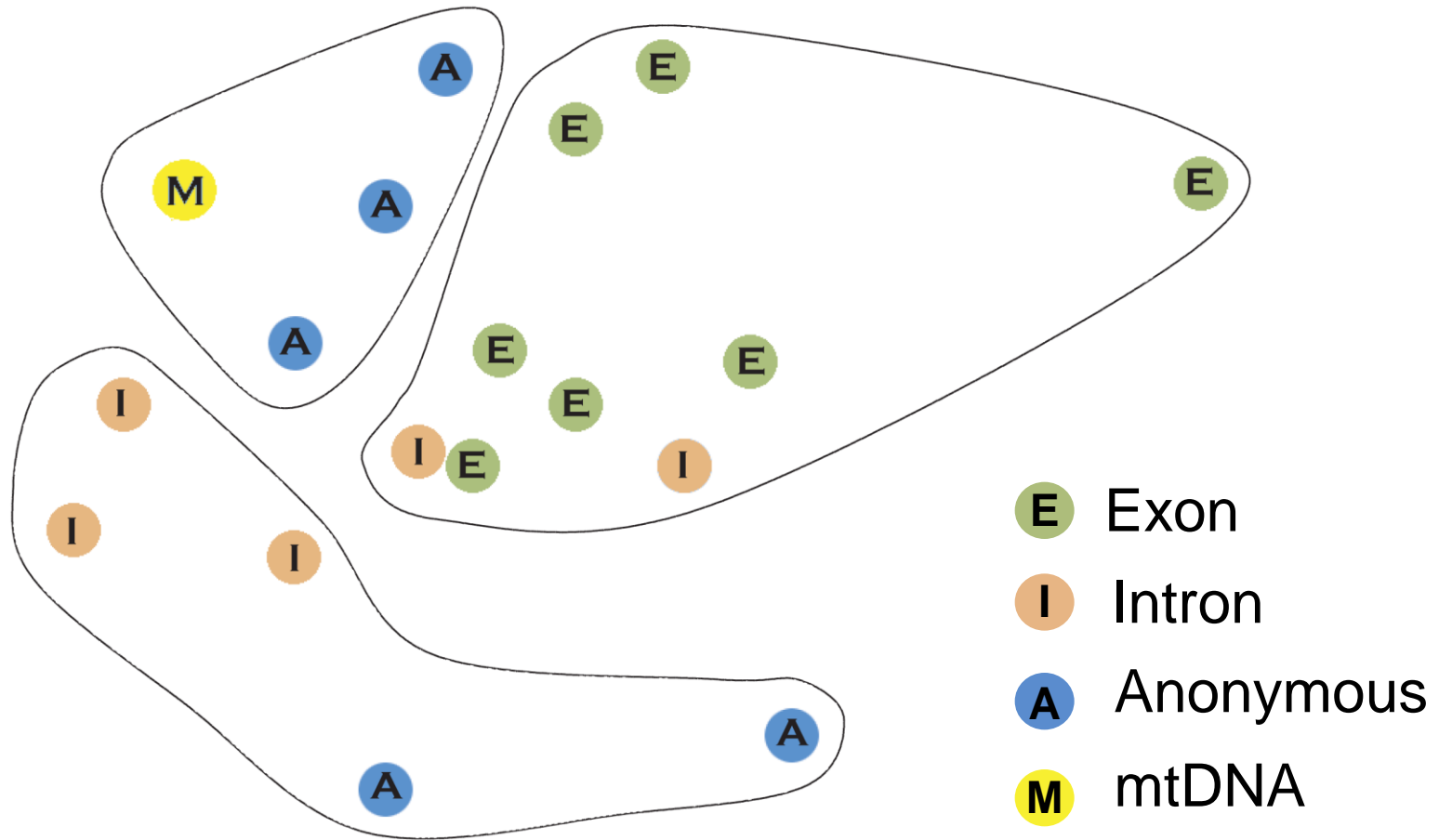
Southern Emu-wren (*Stipiturus*)

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 | 41788 | WA |
| 3 | <i>Amytornis dorotheae</i> | Carpentarian Grasswren | ANWC | | WA |
| 4 | <i>Amytornis goyderi</i> | Eyrean Grasswren | ANWC | | WA |
| 5 | <i>Amytornis housei</i> | Black Grasswren | | | WA |
| 6 | <i>Amytornis merrotsyi</i> | Short-tailed Grasswren | | | WA |
| 7 | <i>Amytornis pumelli</i> | Dusky Grasswren | | | WA |
| 8 | <i>Amytornis textiles</i> | Thicket Grasswren | | | WA |
| 9 | <i>Amytornis striatus</i> | White-throated Grasswren | | | WA |
| 10 | <i>Clytomias insignis</i> | White-winged Grasswren | | | WA |
| 11 | <i>Malurus amabilis</i> | Yellowthroat | | | WA |
| 12 | <i>Malurus melanocephalus</i> | Black-headed Yellowthroat | | | WA |
| 13 | <i>Malurus melanocephalus</i> | Black-headed Yellowthroat | | | WA |
| | | | | | WA |
| | | | | | SA |
| | | | | 12171 | PNG |
| | | | | 7564 | PNG |
| | | | KU | 7082 | PNG |
| | | | ANWC | 20947 | QLD |
| | | Fairy wren | ANWC | 29906 | QLD |
| | | Spotted Fairy wren | ANWC | 28009 | QLD |
| 21 | | Variiegated Fairy wren | ANWC | 31655 | NSW |
| 22 | | Purple-crowned Fairy wren | BMNHC | 60807 | NT |
| 23 | <i>Stipiturus</i> | Southern Emu-wren | ANWC | 20748 | SA |
| 24 | <i>Stipiturus mallee</i> | Mallee Emu-wren | ANWC | 40418 | SA |
| 25 | <i>Stipiturus ruficeps</i> | Rufous-crowned Emu-wren | ANWC | 39914 | QLD |
| 26 | <i>Greygone olivacea</i> | White-throated Greygone | MCZ | 336023 | NSW |

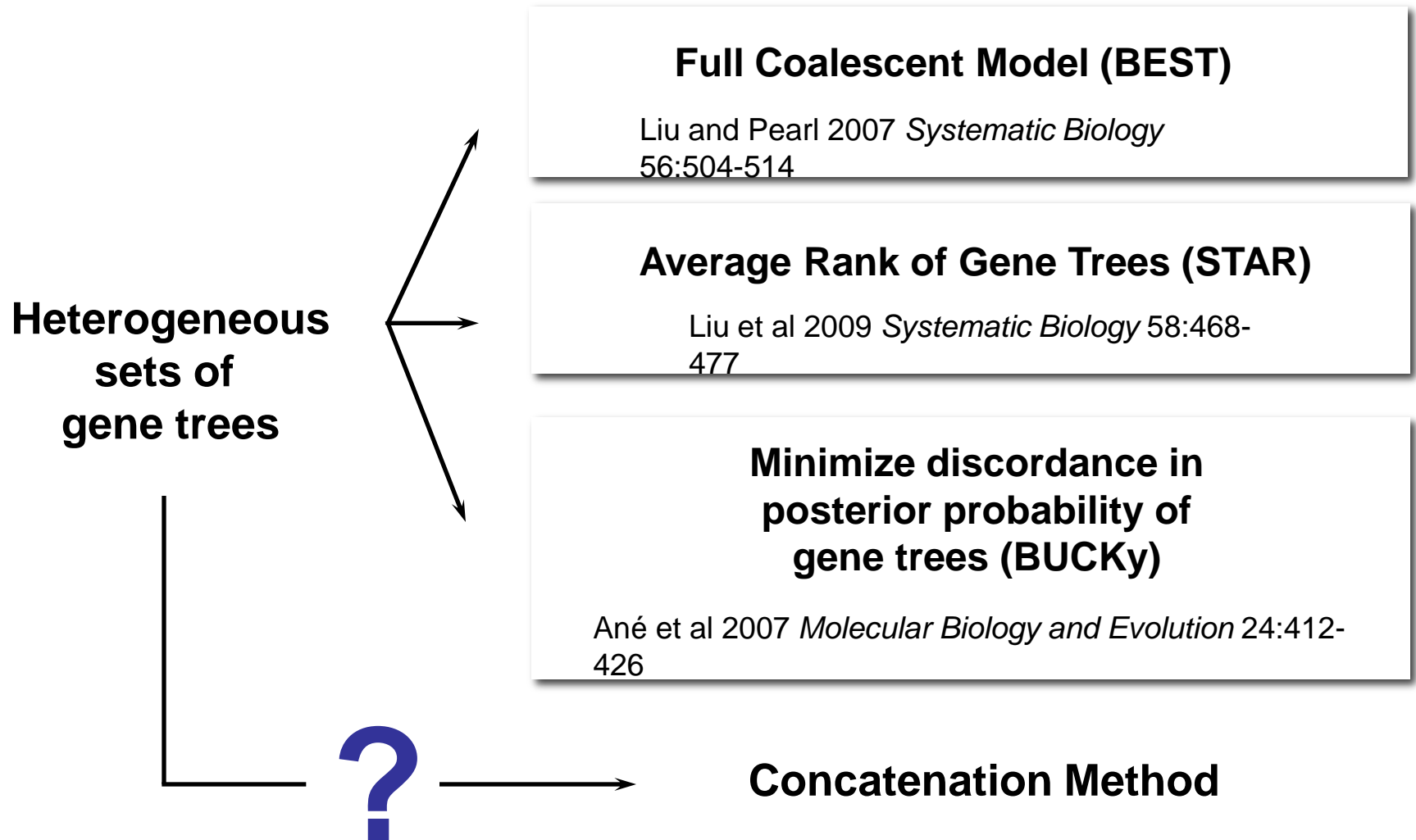
26 Species 51 Individuals
18 Loci

Heterogeneity in Fairy Wren Gene Trees Across Loci

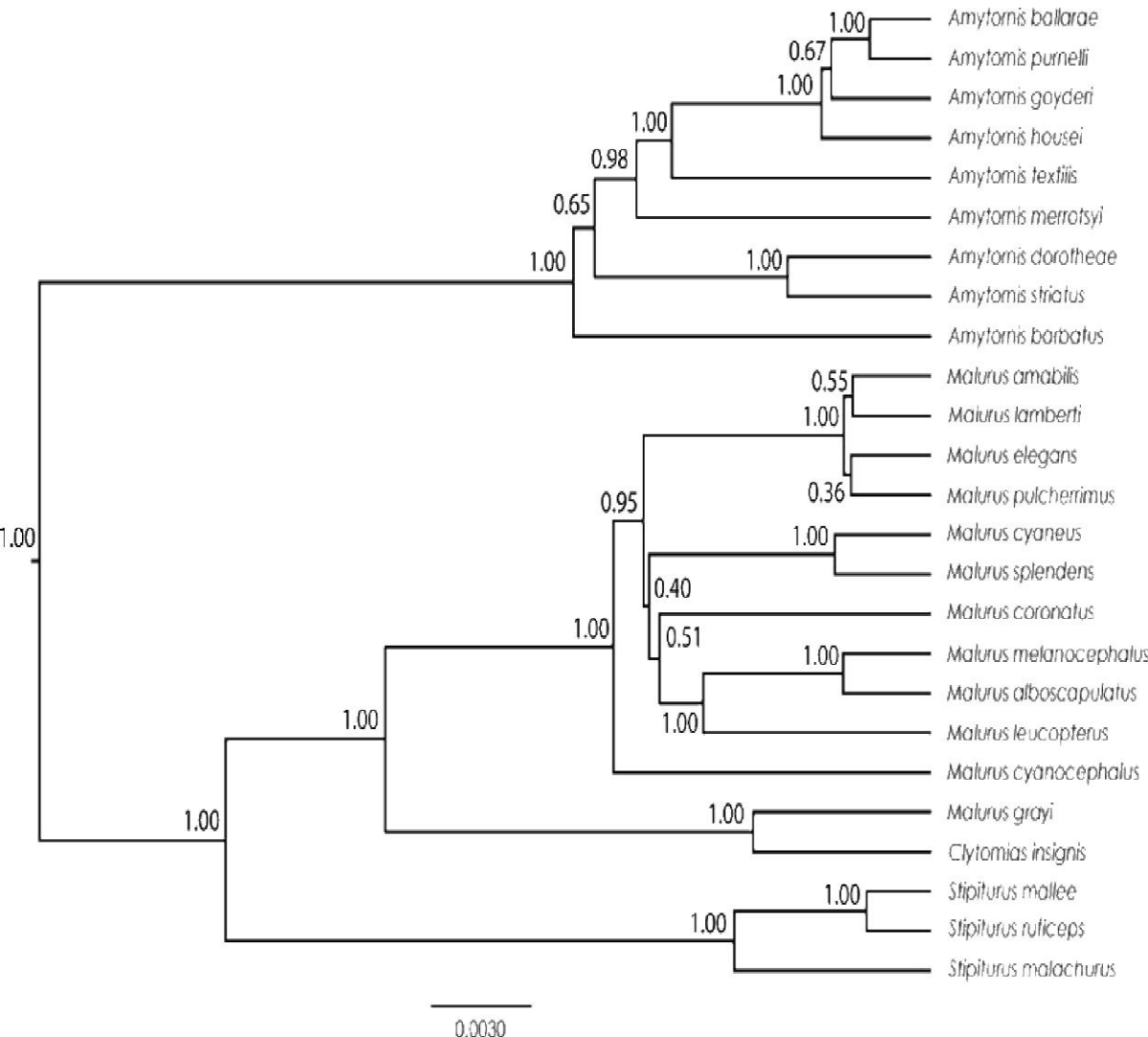


Tree Set Visualization (by multidimensional scaling): Hillis et al 2005. *Syst. Biol.* 54:471-482.
Concatenation: Leigh, J. W., et al.. 2008. *Syst. Biol.* 57: 104-115.

Exploring Incomplete Lineage sorting Through Species Tree Methods



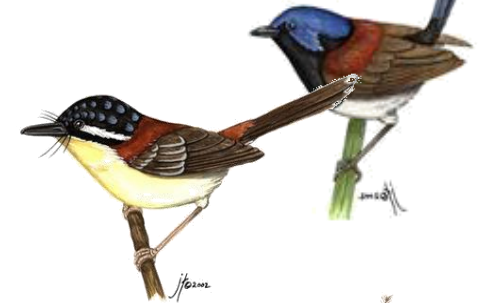
A Species Tree For Fairy Wrens (BEST Method)



GRASSWRENS

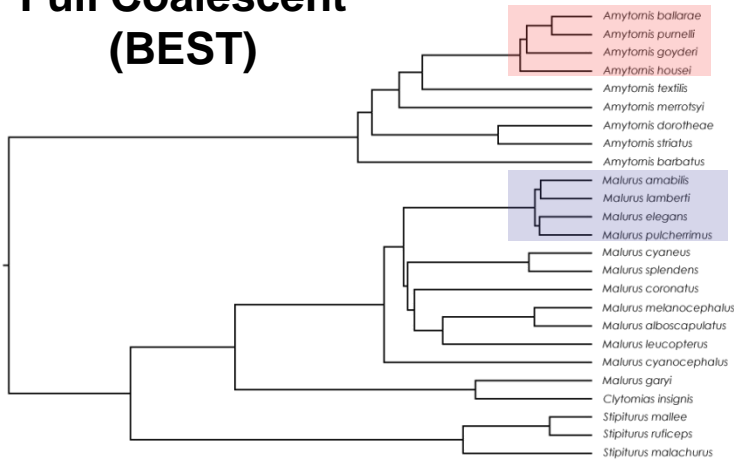
FAIRYWRENS

EMU WRENS

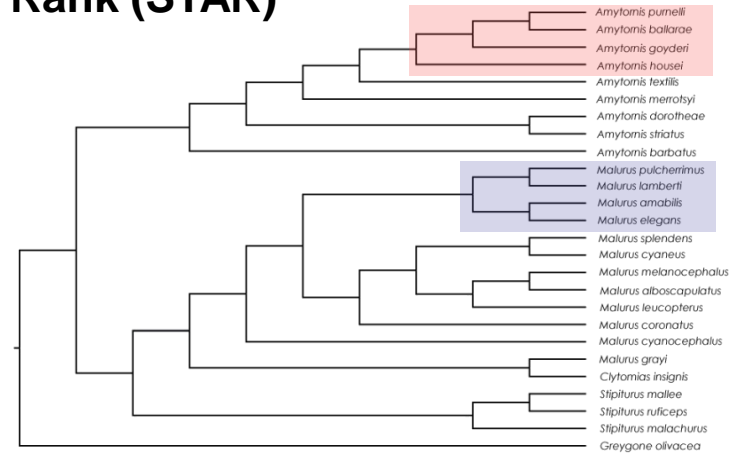


Fairy Wren Species Trees: High Consistency among Methods

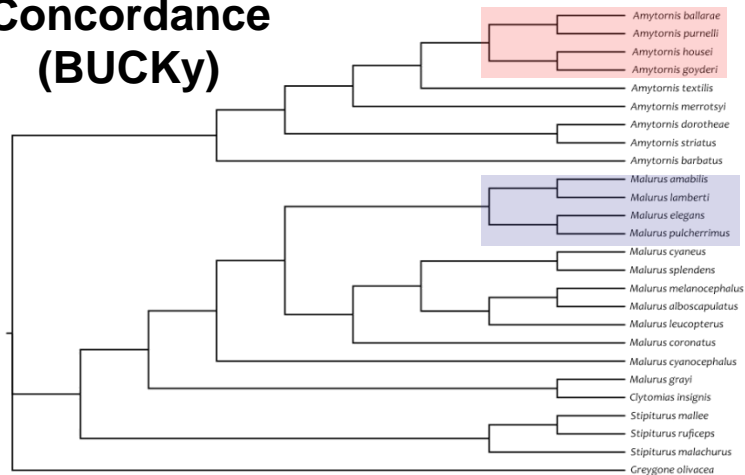
Full Coalescent (BEST)



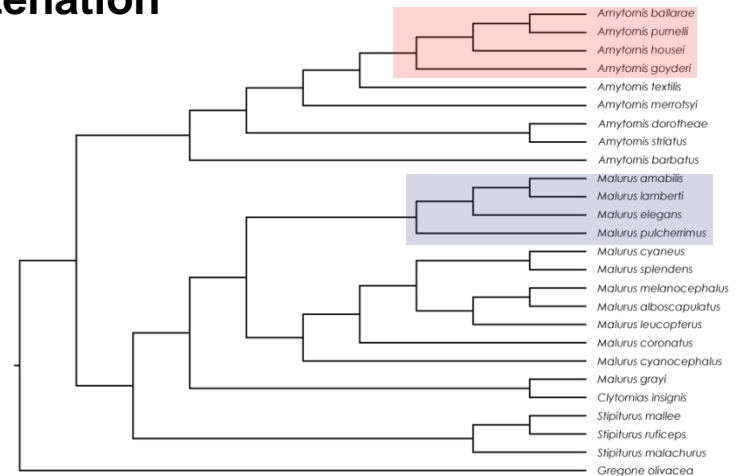
Average Rank (STAR)



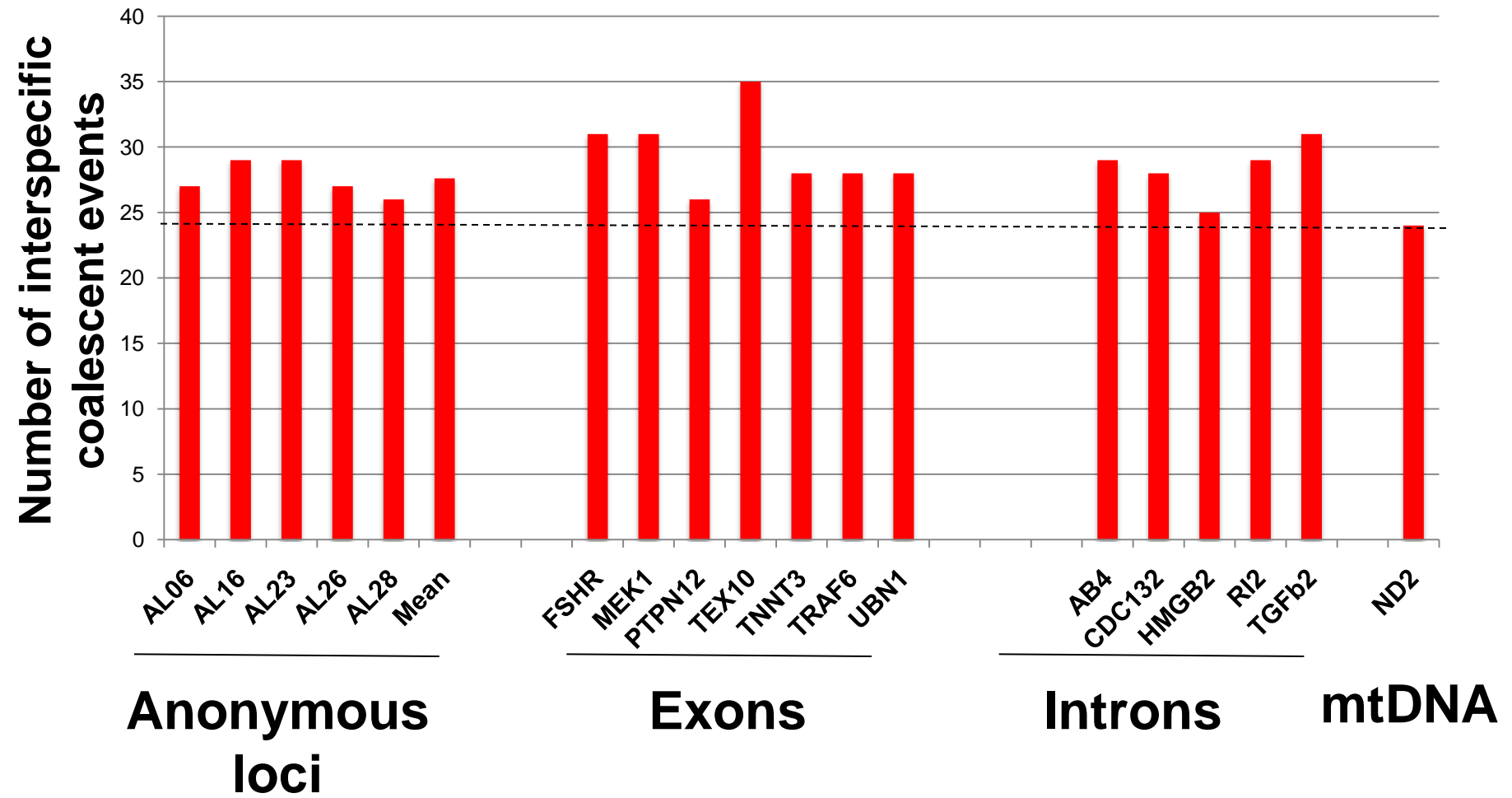
Gene Tree Concordance (BUCKy)



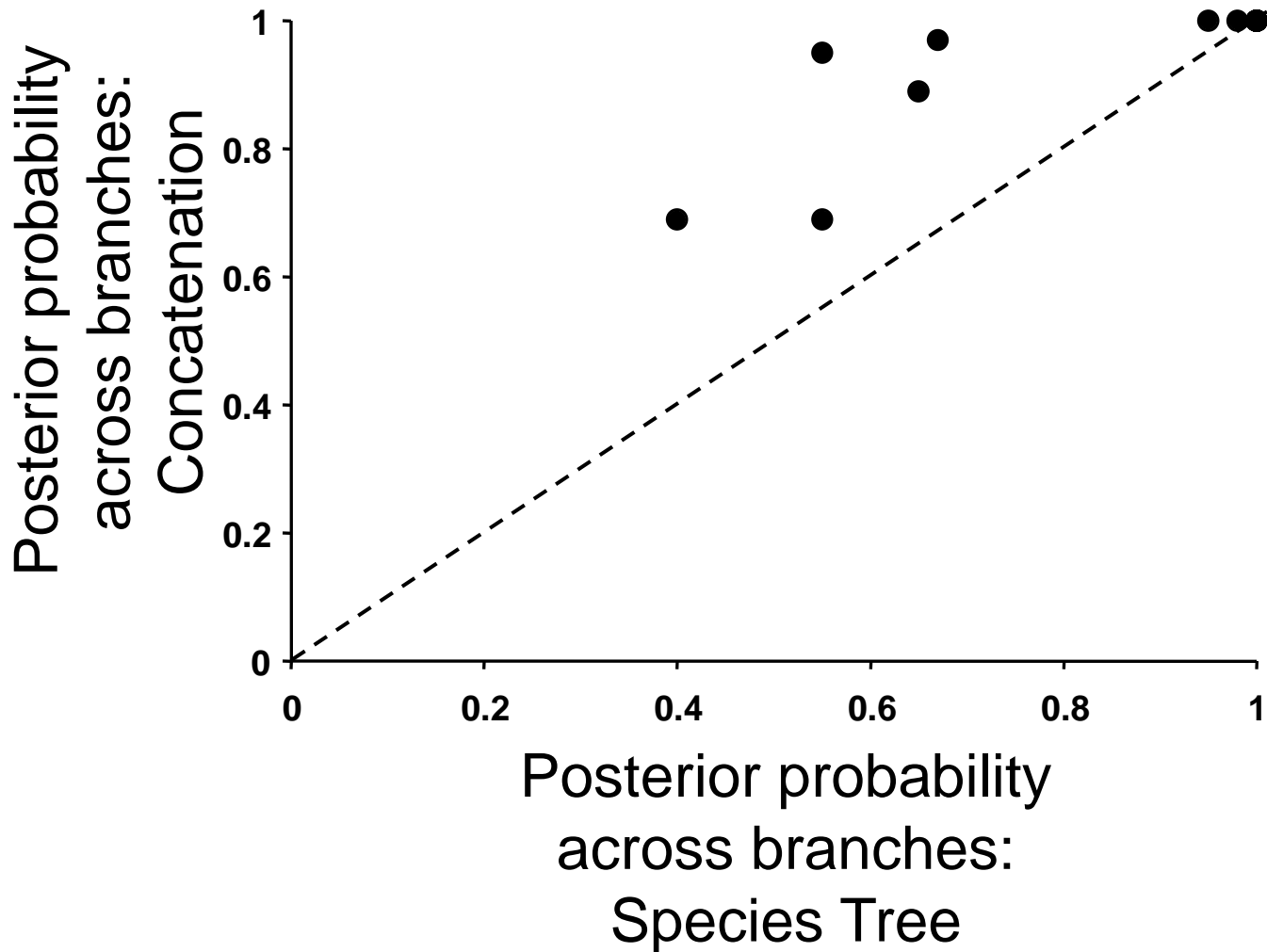
Concatenation



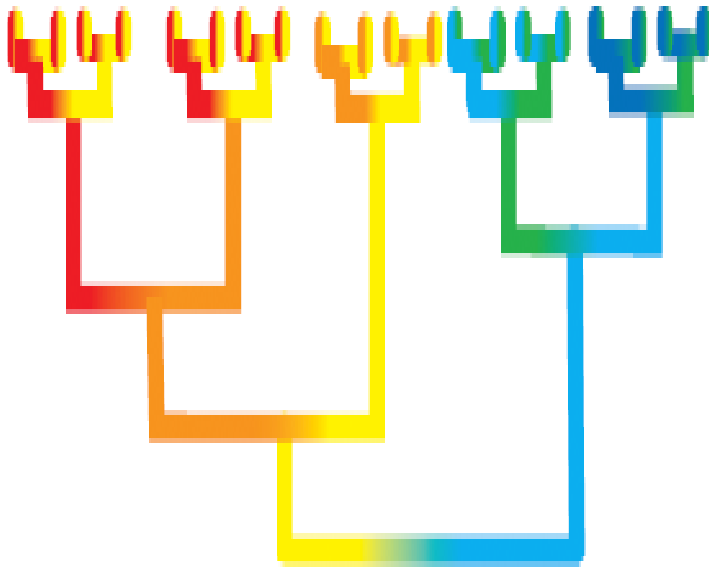
Incomplete lineage sorting is common in fairy wrens



Posterior probabilities under concatenation are uniformly higher than under species tree analysis



Phylogenetic Manifestations of Neutrality and Selection

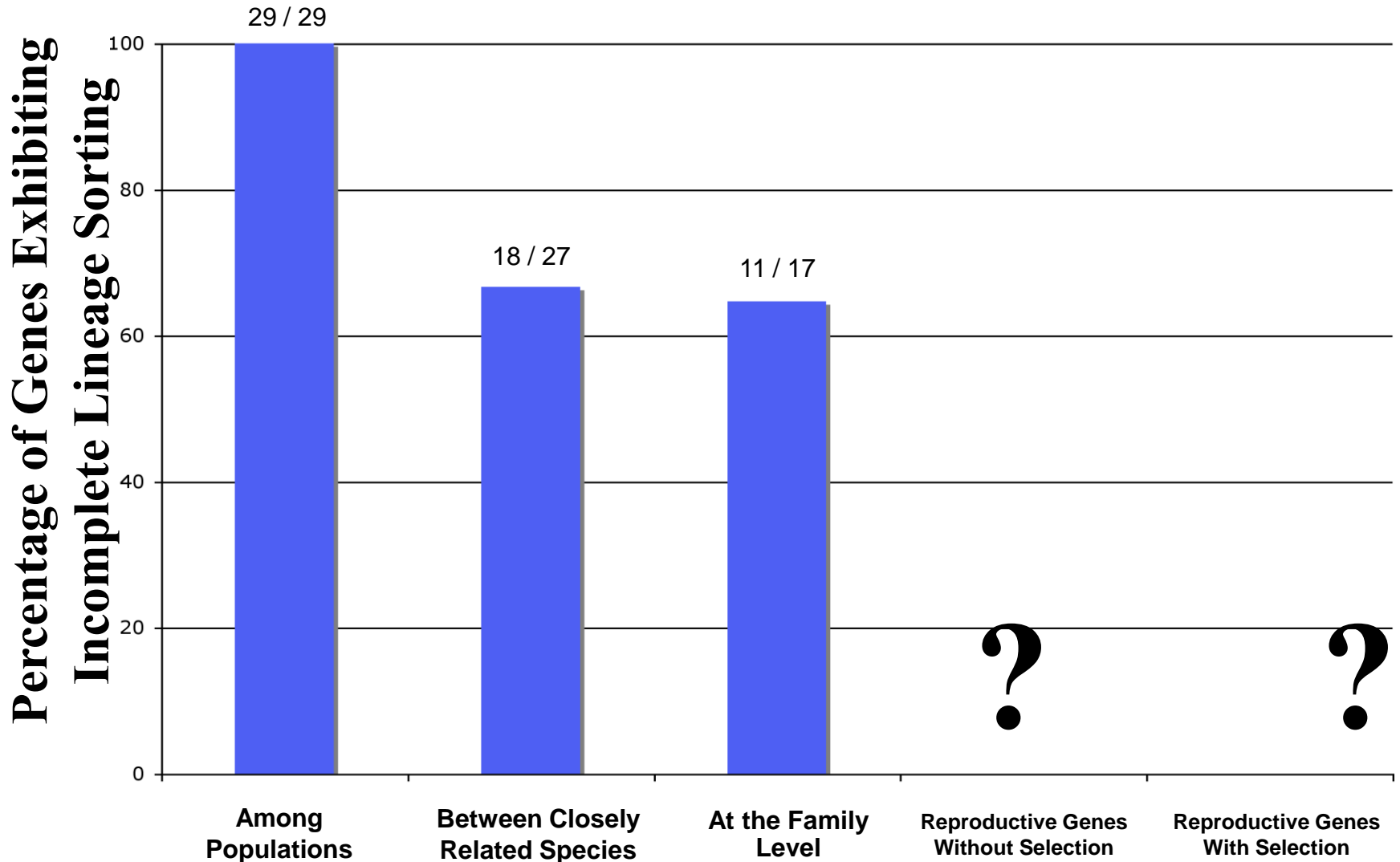


Neutral Evolution



Positive Selection

incomplete lineage sorting is Common In Fairy Wrens



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 Reviere & Williams 1981 | Brillard & de Reviere 1981 | Tuttle et al. 1996 |

Outreach in the outback

Movie
Camera

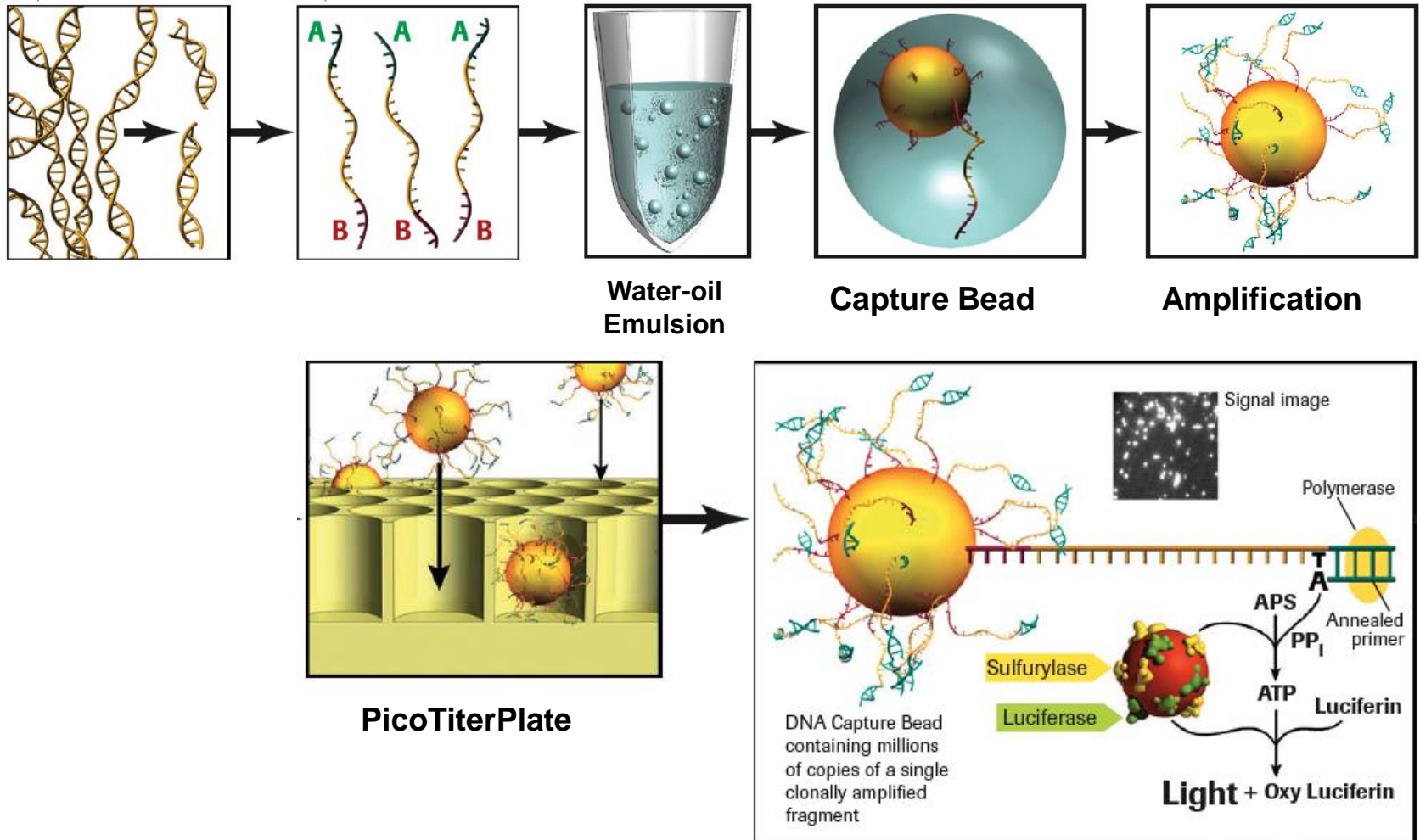
Grad Student

Postdoc

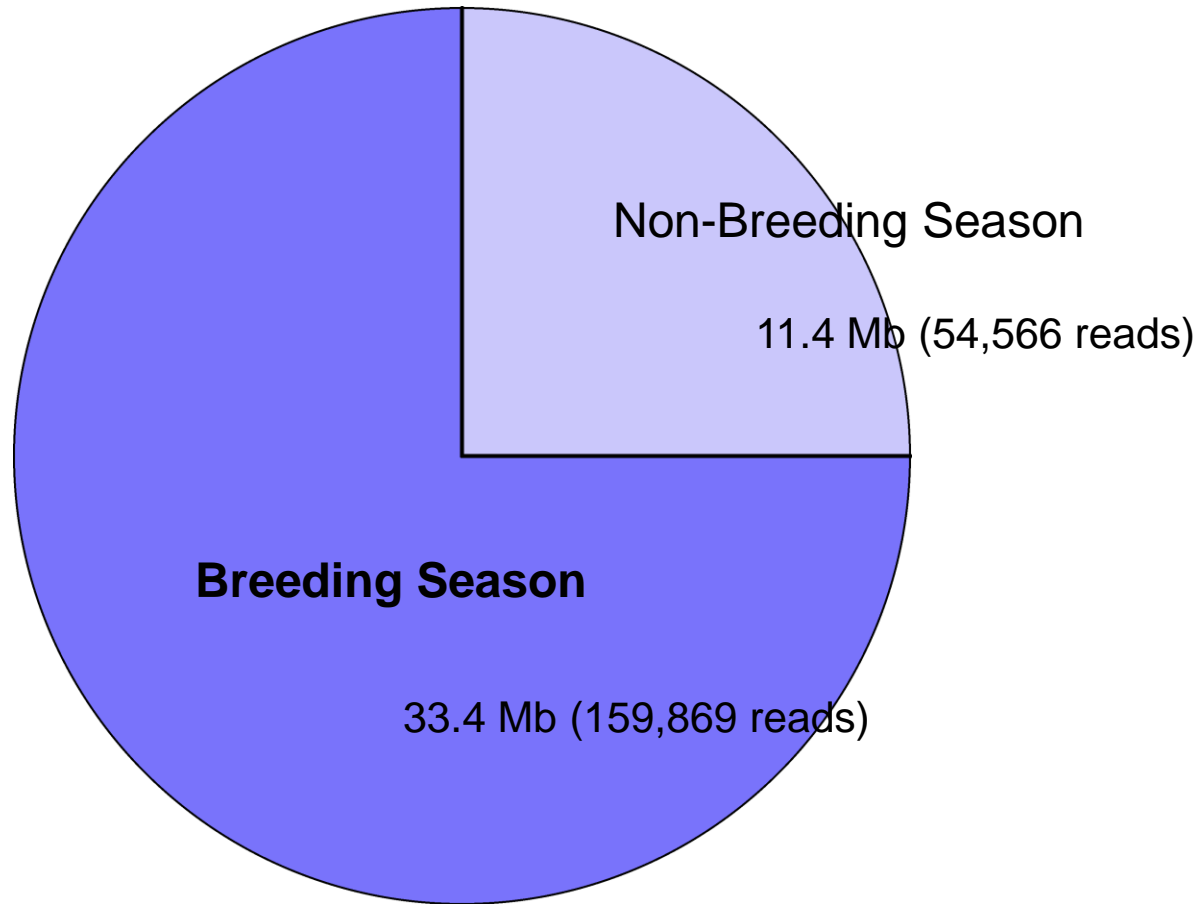




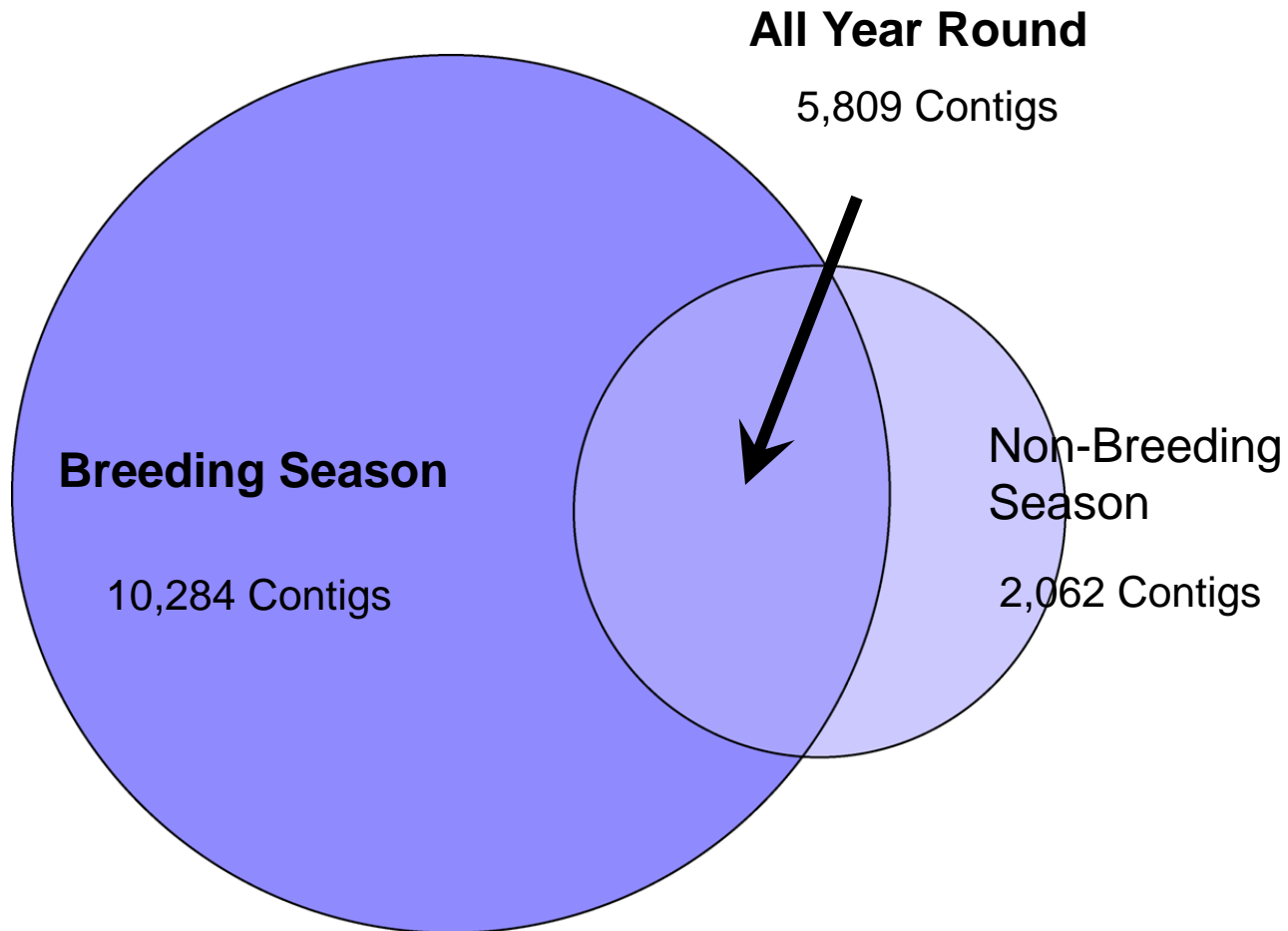
Massively Parallel Sequencing Of testis Transcripts in Breeding And NonBreeding Seasons



454 Sequencing (Raw Data)



Higher Diversity of Contigs In Breeding vs. Non-Breeding Season Testis

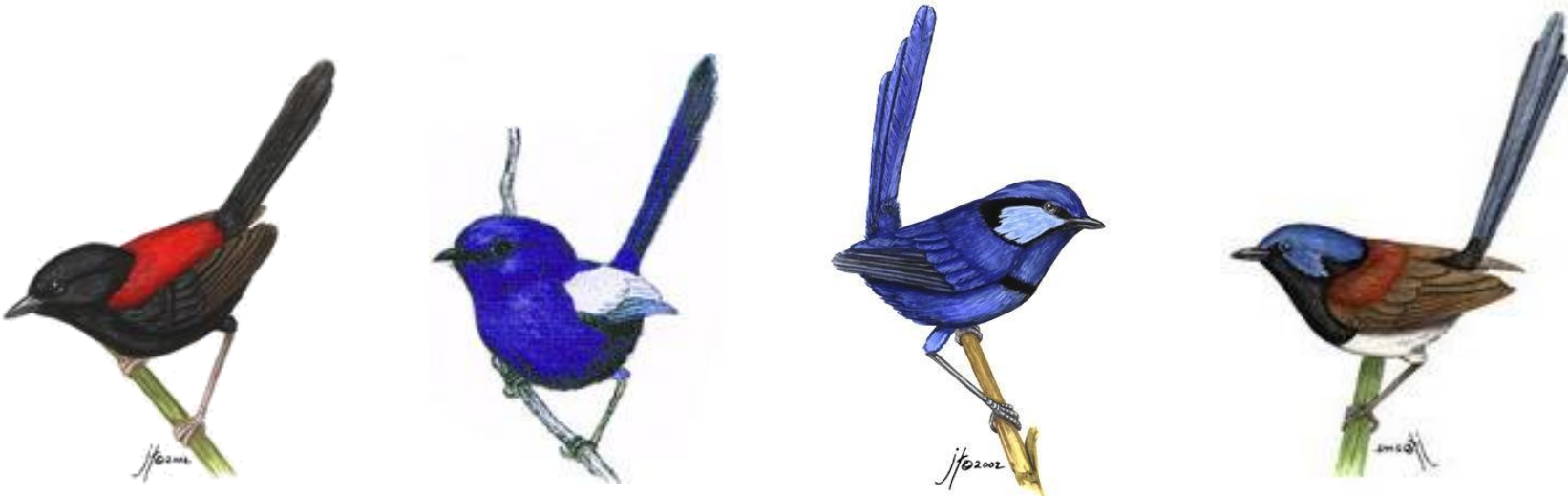


20 Reproductive Genes surveyed

| 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 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 Vic | | 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



Red-backed

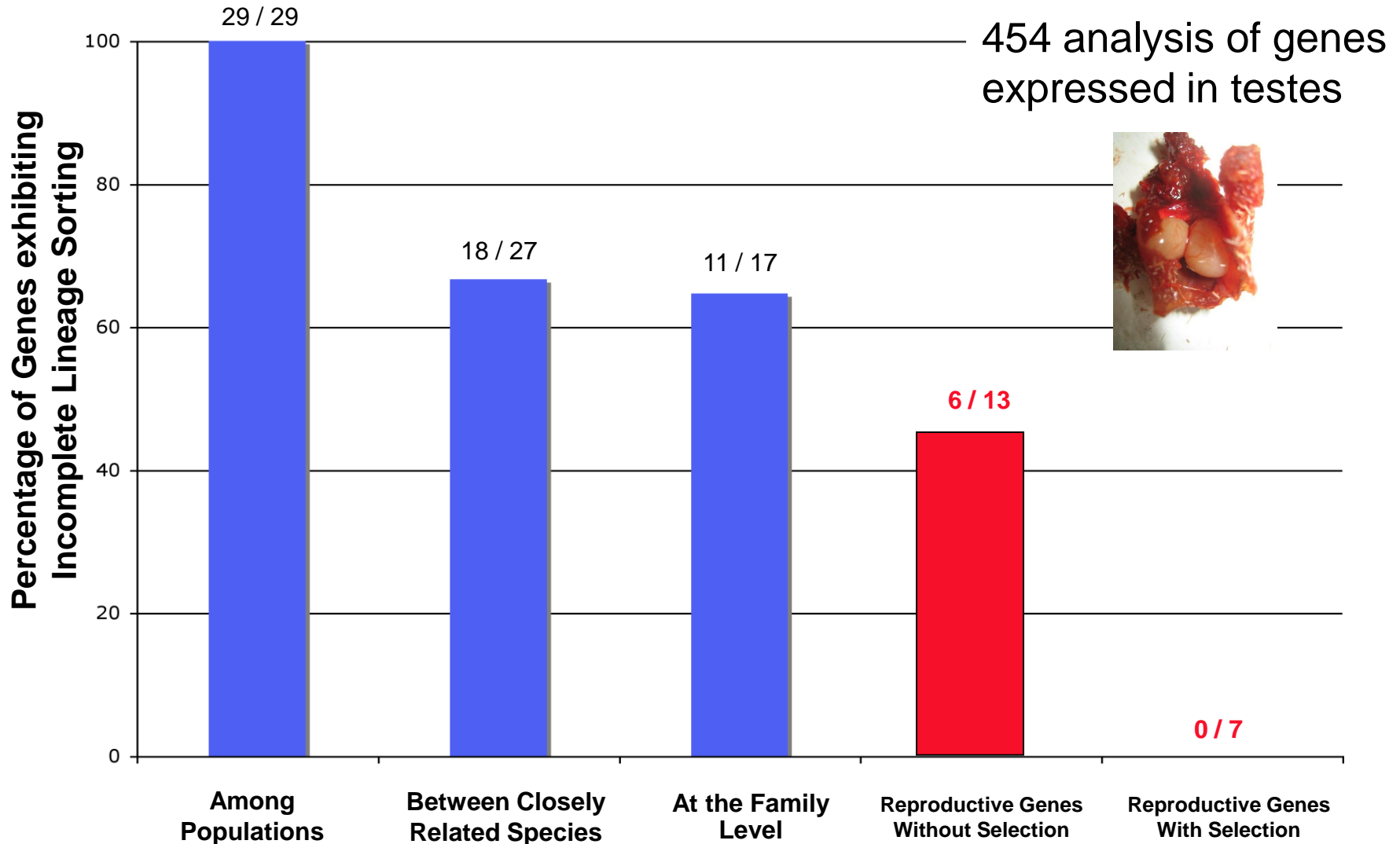
White-winged

Splendid

Variegated

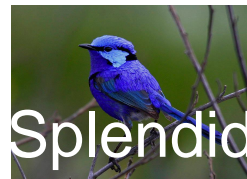
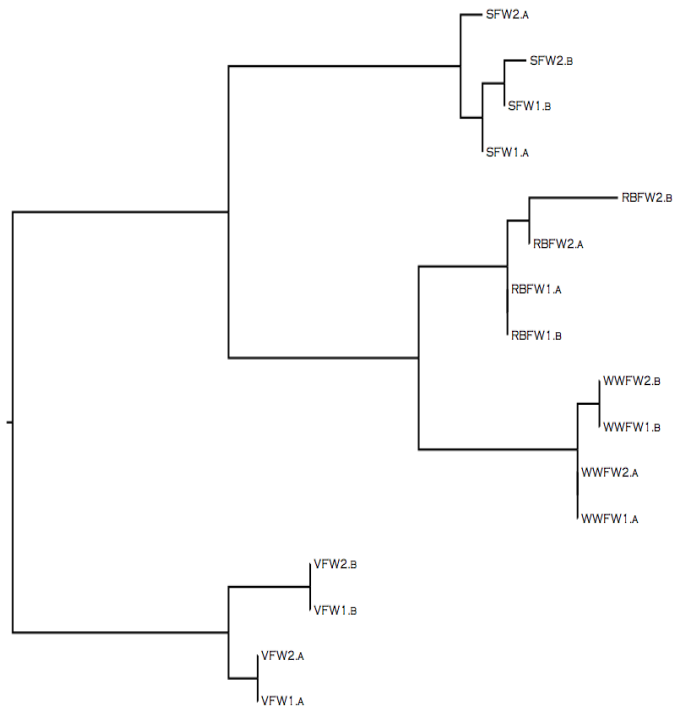


Directional Selection reduces incomplete lineage sorting In Reproductive Genes

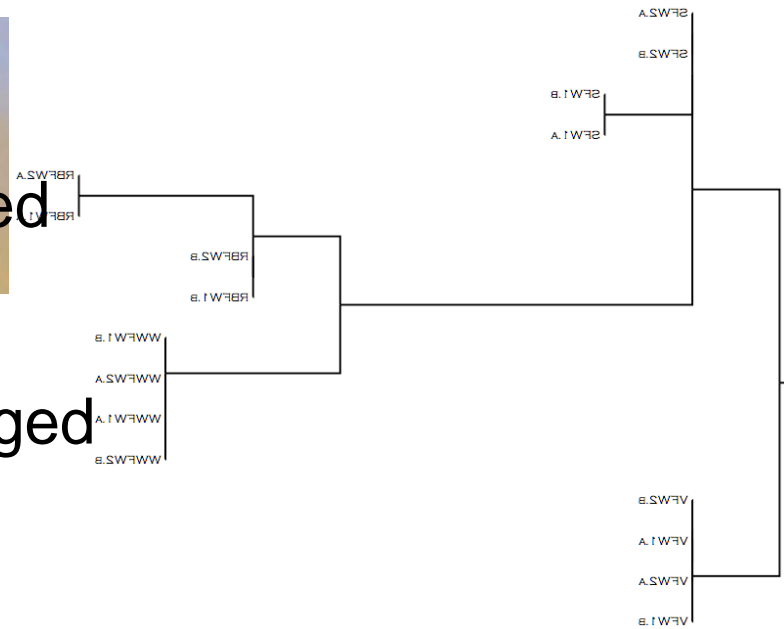


Examples of genes showing Reciprocal monophyly among Fairy Wren species

Proacrosin



Outer Dense Fiber Protein



Conclusions

- Multilocus phylogeography captures major breaks within species
- Species trees can be thought of as an extension of phylogeography
- New methods incorporate various aspects of coalescent process into phylogenetics
- Natural selection can speed up rate of monophyly

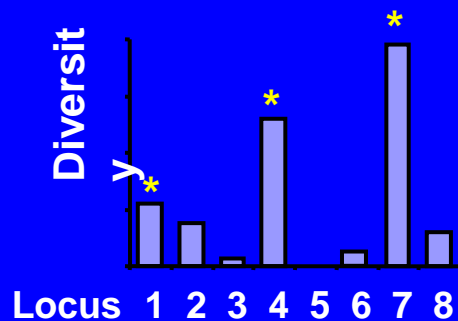


Thanks to ...

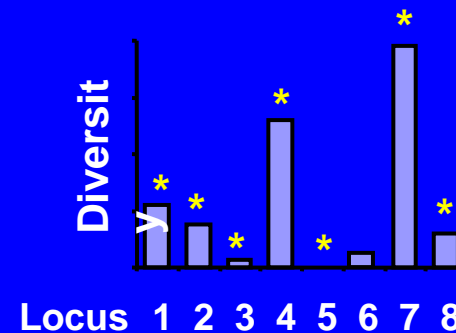
- Bryan Jennings, June Lee, Nancy Rotzel
- Liang Liu, Santiago Castillo, Dennis Pearl, June Lee, Leo Joseph
- National Science Foundation

Ascertainment bias and sampling strategy for SNPs

Microsatellites



SNPs



* Locus chosen for phylogeographic survey

