Trait-dependent dispersal models for phylogenetic biogeography, in the R package *BioGeoBEARS*

Nicholas J. Matzke[*,1,2]

**nick.matzke@anu.edu.au*

[1] Discovery Early Career Researcher Award (DECRA) Fellow; Division of Ecology, Evolution, and Genetics; Research School of Biology; The Australian National University, Canberra

[2] Work began at: National Institute of Mathematical and Biological Science (NIMBioS) and Department of Ecology & Evolutionary Biology, University of Tennessee, Knoxville

(Data provided by collaborator Jeremy J. Kirchman)

Twitter and photos are OK! @NickJMatzke

QR code: CV, website, R code!

Abstract

Organism traits must be important in historical biogeography. In particular, rates of dispersal (both range-expansion dispersal, and jump dispersal leading to founder-event speciation) must depend to some degree on traits such as flight and its loss, and seed dispersal mechanisms and the dispersal abilities of animals that transport seeds.

However, to date no probabilistic historical biogeographical models have been available that allow geographic range and traits to co-evolve on the phylogeny, with traits influencing dispersal ability. In purely continuous-time Markov models, adding a trait is just a matter of doubling the size of the rate matrix; however, biogeographical models also include a much more complex discrete-time model describing how geographic range can change during cladogenesis. Traits might also influence this process.

I present an addition to the R package *BioGeoBEARS* that enables an evolving discrete trait to influence dispersal ability for both anagenetic and cladogenetic range change. This model can be freely combined with models adding jump dispersal (e.g., DEC+J), distance as a predictor of dispersal (+x models, with dispersal rate multiplied by distance^x), and other variants. I test the model against simulations and datasets where large evolutionary changes in dispersal ability are highly likely (e.g., Pacific rails, which have repeatedly lost flight).

BioGeoBEARS trait + dispersal model

Build a trait-based dispersal model

1. Trait flight/flightlessness trait evolves on tree
   Parameters of model:
   - *t12* -> Rate of flight loss
   - *t21* -> Rate of flight gain
   - *m2* -> Multiplier on dispersal prob. while flightless

	FLY	NOT
FLY	-	t12
NOT	t21	-

2. Cladogenetic parameters of model:
   jt12 -> Chance of "jumping" from flight-> flightless during founder-event speciation

Study system: "typical" rails of the Pacific

Adolf Meyer (1898). The Birds of Celebes and the neighbouring islands. https://en.wikipedia.org/wiki/Snoring_rail

G. pacificus, from Tahiti, "known only from a painting made by Georg Forster on James Cook's second voyage."

Data from: Kirchman (2012)

Collaborator: Jeremy J. Kirchman, Curator of Birds, New York State Museum

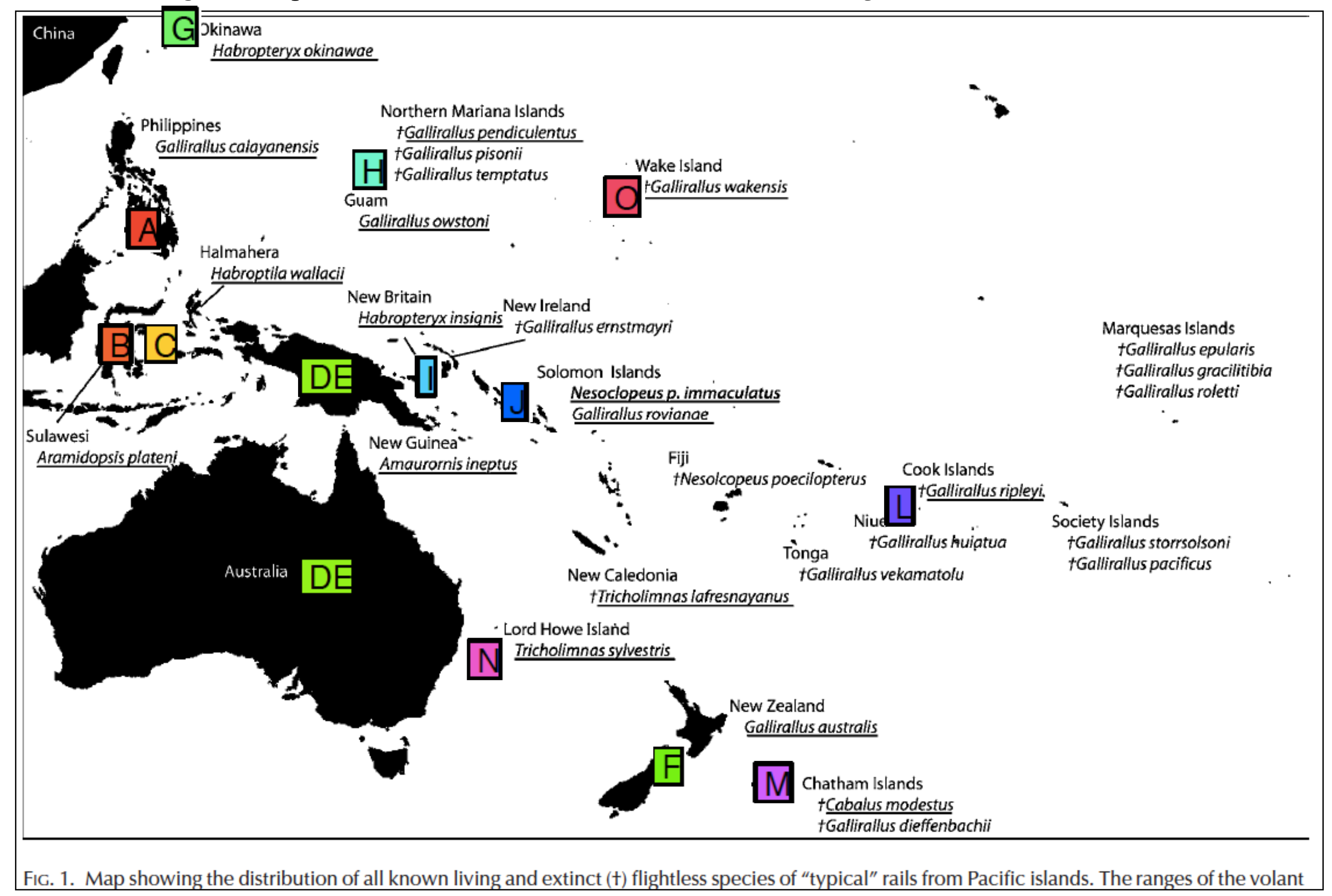


FIG. 1. Map showing the distribution of all known living and extinct (†) flightless species of "typical" rails from Pacific islands. The ranges of the volant

Source: Kirchman, Jeremy J. (2012). Speciation of Flightless Rails on Islands: A DNA-Based Phylogeny of the Typical Rails of the Pacific, *The Auk*, 129(1):56-69.

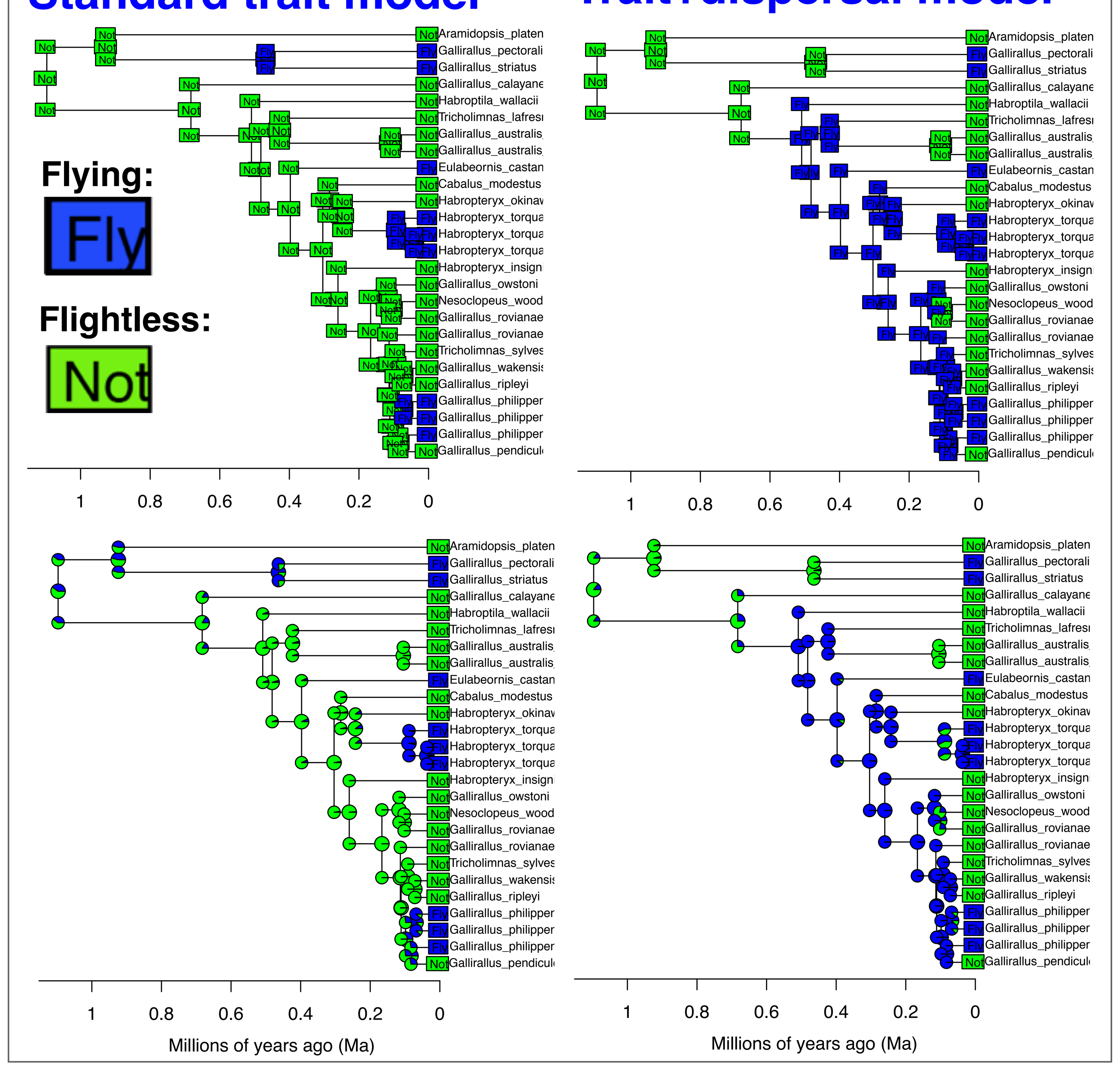
Statistical Model Choice

	Base model	rate of range expansion d	rate of range loss e	weight of founder-event speciation j	rate of flight -> flightlessness t_{12}	rate of flightlessness -> flight t_{21}	dispersal multiplier when trait=1 (flight) m_1	dispersal multiplier when trait=2 (flightless) m_2	rate of switch to flightlessness during founder event jt_{12}	rate of switch to flight during founder event jt_{21}	number of free parameters np	log-likelihood LnL	delta AICc dAICc	model weight mw
		Black: free parameter that is estimated. Gray: fixed parameter.												
DEC: "Dispersal-Extinction-Cladogenesis"(of *Lagrange*, Ree & Smith 2008)	DEC	3.84	1.94	0	3.31	0.94	1	0	0	0	4	-119.6	37.6	0.0%
	DEC	0.18	0.00	0.28	5.18	1.65	1	0	0	0	5	-103.9	9.4	0.9%
	DEC	3.81	2.01	0	2.58	0	1	0	0	0	3	-120.3	36.2	0.0%
	DEC	0.15	0.00	0.16	2.84	0	1	0	0	0	4	-107.5	13.5	0.1%
	DEC	2.50	1.69	0	3.60	1.28	1	0	1.00	0	5	-119.4	40.3	0.0%
	DEC	0.18	0.00	0.82	4.37	1.12	1	0	0.75	0	6	-104.2	13.3	0.1%
	DEC	3.31	1.98	0	2.57	0	1	0	1.00	0	4	-120.2	38.9	0.0%
	DEC	0.15	1E-12	0.16	2.84	0	1	0	0.00	0	5	-107.5	16.6	0.0%
Likelihood version of Ronquist (1997)'s Dispersal-Vicariance Analysis	DIVALIKE	1.99	0.82	0	5.16	1.48	1	0	0	0	4	-115.2	28.9	0.0%
	DIVALIKE	0.24	0.00	0.22	5.73	1.71	1	0	0	0	5	-104.1	9.8	0.7%
	DIVALIKE	2.68	1.50	0	2.60	0	1	0	0	0	3	-117.7	31.0	0.0%
	DIVALIKE	0.16	0.00	0.10	2.61	0	1	0	0	0	4	-109.3	17.1	0.0%
	DIVALIKE	2.35	0.90	0	4.38	1.11	1	0	0.91	0	5	-115.3	32.2	0.0%
	DIVALIKE	0.24	0.00	0.49	5.49	1.49	1	0	0.62	0	6	-104.1	13.2	0.1%
	DIVALIKE	2.56	1.45	0	2.70	0	1	0	0.58	0	4	-117.7	33.8	0.0%
	DIVALIKE	0.17	1E-12	0.64	2.81	0	1	0	0.87	0	5	-109.3	20.1	0.0%
(no vicariance)	BAYAREALIKE	6.40	3.70	0	4.13	2.37	1	0	0	0	4	-129.7	57.8	0.0%
	BAYAREALIKE	5.34	3.70	0.95	4.14	2.38	1	0	0	0	5	-122.3	46.2	0.0%
	BAYAREALIKE	3.32	3.91	0	2.67	0	1	0	0	0	3	-130.8	57.2	0.0%
	BAYAREALIKE	3.34	3.91	0.78	2.67	0	1	0	0	0	4	-123.0	44.4	0.0%
	BAYAREALIKE	4.36	3.46	0	5.66	2.45	1	0	0.98	0	5	-129.9	61.4	0.0%
	BAYAREALIKE	0.09	0.16	0.35	9.24	3.01	1	0	0.42	0	6	-97.51	0.0	96.5%
	BAYAREALIKE	3.36	2.66	0	2.61	0	1	0	0.00	0	4	-132.5	63.5	0.0%
	BAYAREALIKE	0.07	1E-12	0.14	2.59	0	1	0	0.14	0	5	-103.4	8.4	1.5%

Best Model under these constraints

Standard trait model

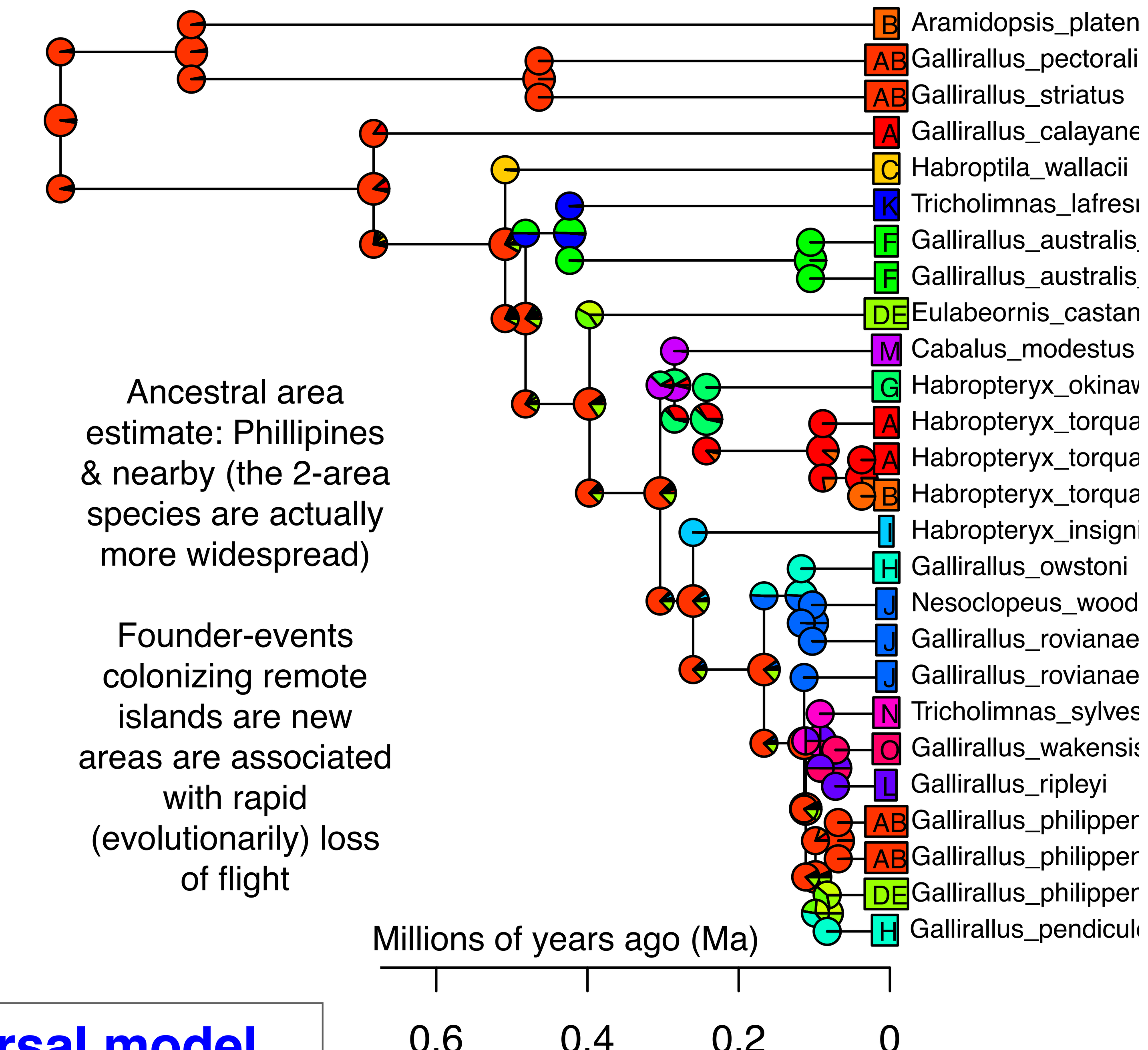
Trait+dispersal model



Results: Ancestral geography

Ancestral area estimate: Phillipines & nearby (the 2-area species are actually more widespread)

Founder-events colonizing remote islands are new areas are associated with rapid (evolutionarily) loss of flight



Conclusions

Linked trait+dispersal model allows statistical inferences not possible with either individually. Although this case is "obvious," in many datasets, the influence of traits on dispersal may be unknown.

- Flightlessness is more frequent at the tips, so models that allow reversal are favored. Regardless, the trait inference is improved (not perfect, see root).
- Inferring *m2* is difficult, but fixing to 0 is feasible and nearly as good (~1 LnL).
- Founder-event speciation is a key process, as is founder+flight loss.
- The base model matters *a lot*. This *must* be part of the model choice!
- These models ignore lineage extinction, but ClaSSE models infeasible.
- Models with more areas will be slower, but *BioGeoBEARS* is parallelized.

Check out *BioGeoBEARS*, updated code at PhyloWiki: http://phylo.wikidot.com/

Acknowledgements: *BioGeoBEARS* was developed during support by: NSF Award 0919124, and by the National Institute for Mathematical and Biological Synthesis (NIMBioS), an Institute sponsored by the National Science Foundation through NSF Award #DBI-1300426 with additional support from The University of Tennessee, Knoxville. NJM was also supported by the Australian Research Council's Discovery Early Career Researcher Award #DE150101773, and by The Australian National University.

ANU Research School of Biology. Poster at Early-and Mid-Career Conference. R. N. Robertson Lecture Theatre, Bldg. 46. Feb. 11th, 2016, 4:45 pm.

References & more info: Use QR code or go to: *http://phylo.wikidot.com/nicholas-j-matzke*