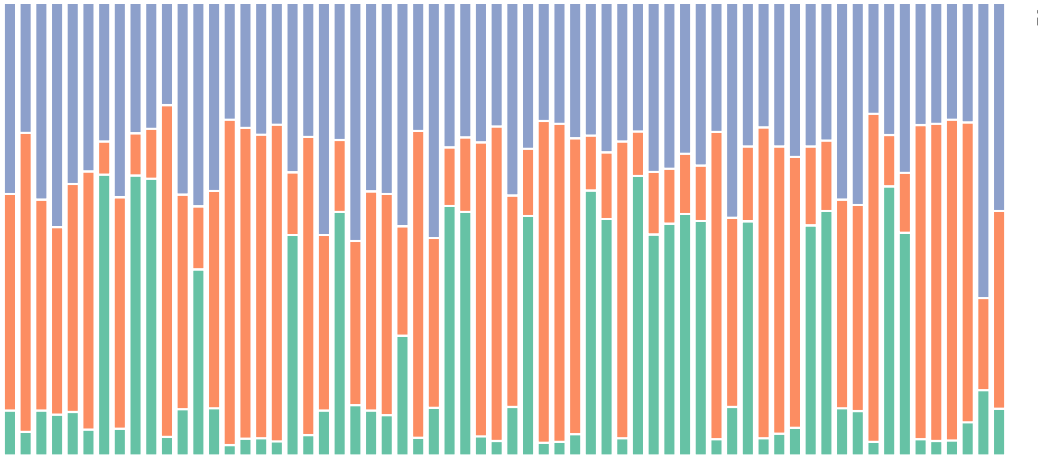
Identifying Models of Trait-Mediated Community Assembly

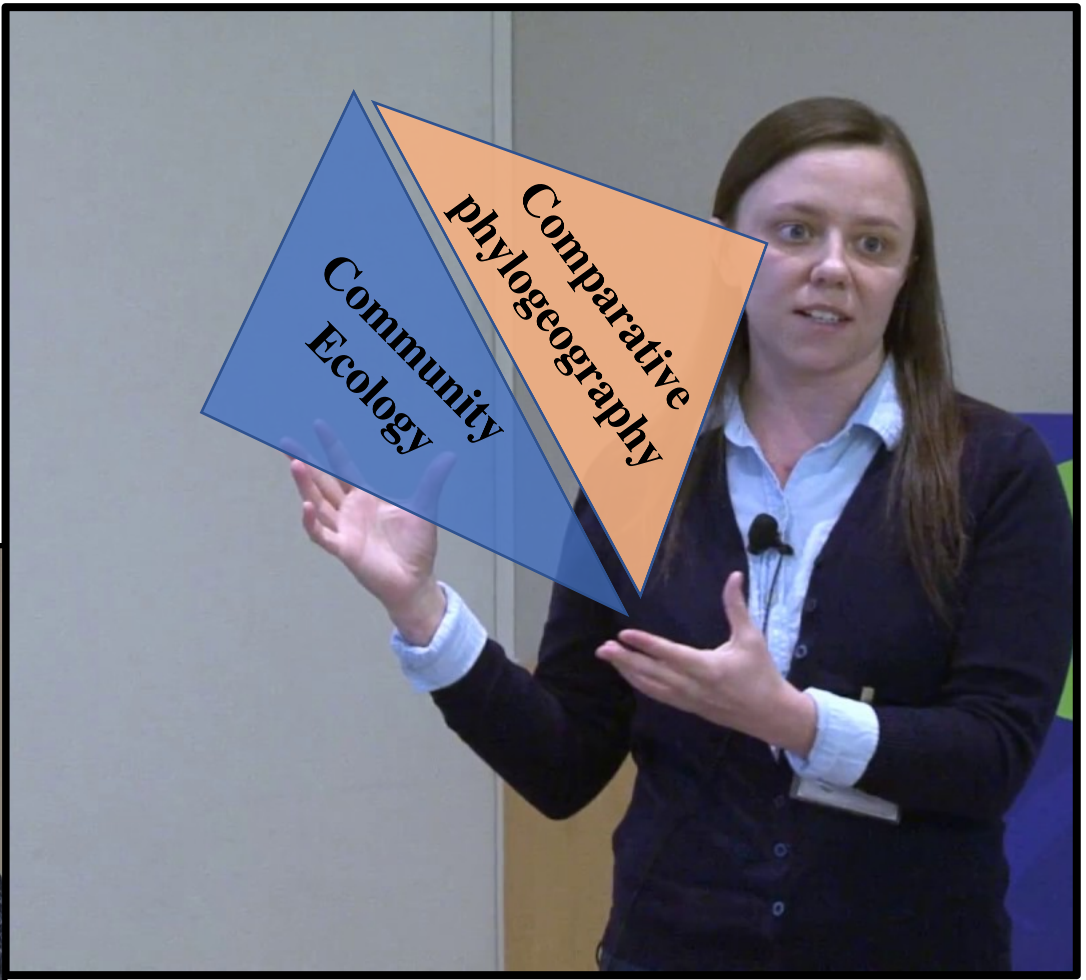
using random forests and approximate Bayesian computation

Megan Ruffley, Katie Peterson,
Bob Week, David C. Tank,
Luke J. Harmon



Society of Systematic Biologists
University of Idaho
SSE
SOCIETY for the STUDY of EVOLUTION
iDiv
German Centre for Integrative Biodiversity Research
SANTA FE INSTITUTE
NSF
THE OHIO STATE UNIVERSITY
BOTANICAL SOCIETY OF AMERICA
FOR BIO
the City College of New York
UiO : Life Science
University of Oslo
ibest
Institure for Bioinformatics and Evolutionary Studies
ID
University of Idaho Stillinger Herbarium

- Megan Ruffley (she/her)
- University of Idaho
- Institute for Bioinformatics and Evolutionary Studies (**Ibest**)
- Bioinformatics and Computation Biology
- Plants
- Roller Derby




David C. Tank


Jack Sullivan

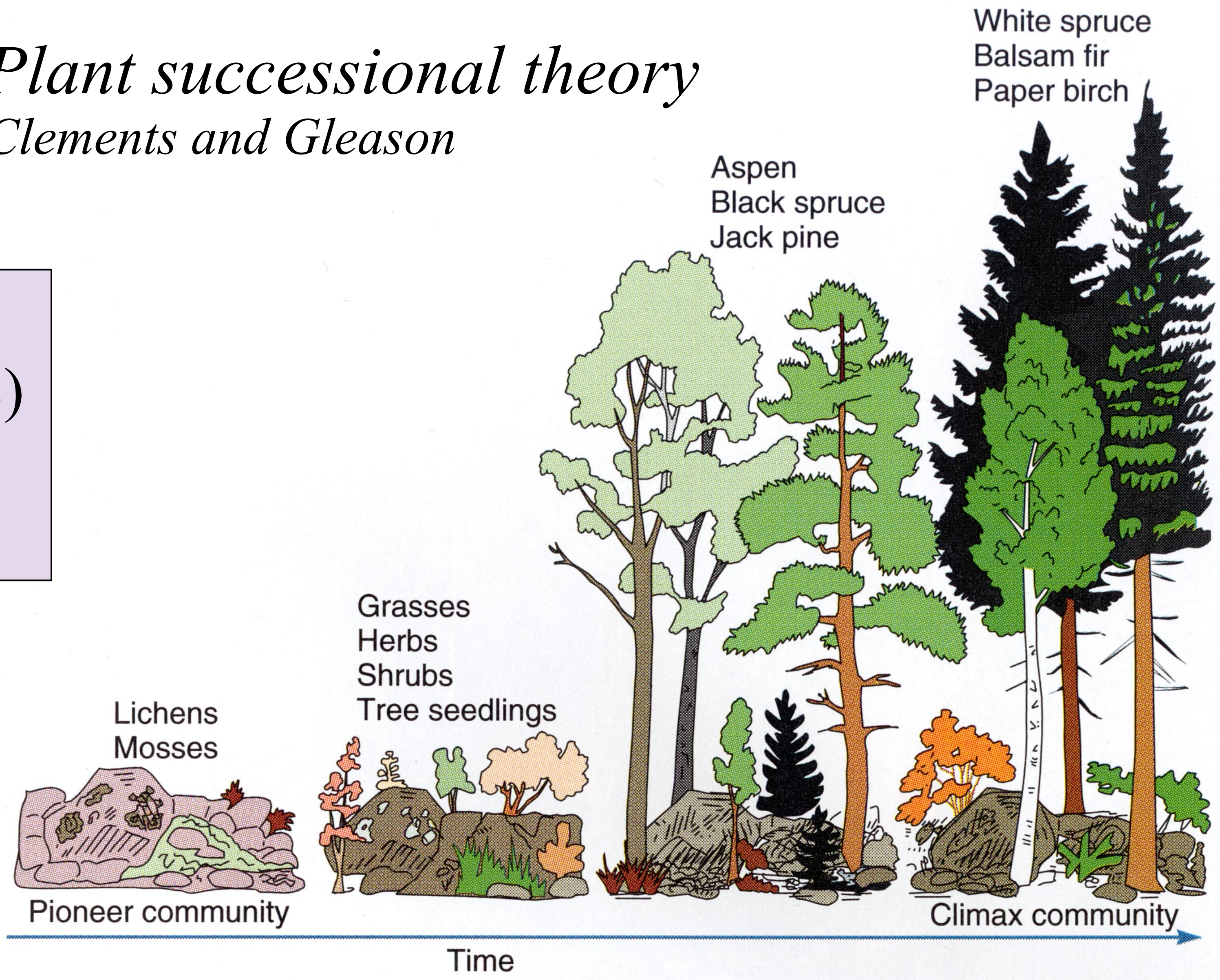

Luke J. Harmon

Interested in identifying neutral and non-neutral processes of assembly

"The processes that determine the patterns of the number and composition of co-occurring species have been central to community ecology for decades…" Chase 2003.

Plant successional theory
Clements and Gleason

Are communities a result of ***deterministic*** (Clements 1938) or ***stochastic*** (Gleason 1927) processes?



Frederic Clements (1916)

"Succession is deterministic"

-stable 'climax' community that is in perfect balance with the climate

"Communities are super-organisms"

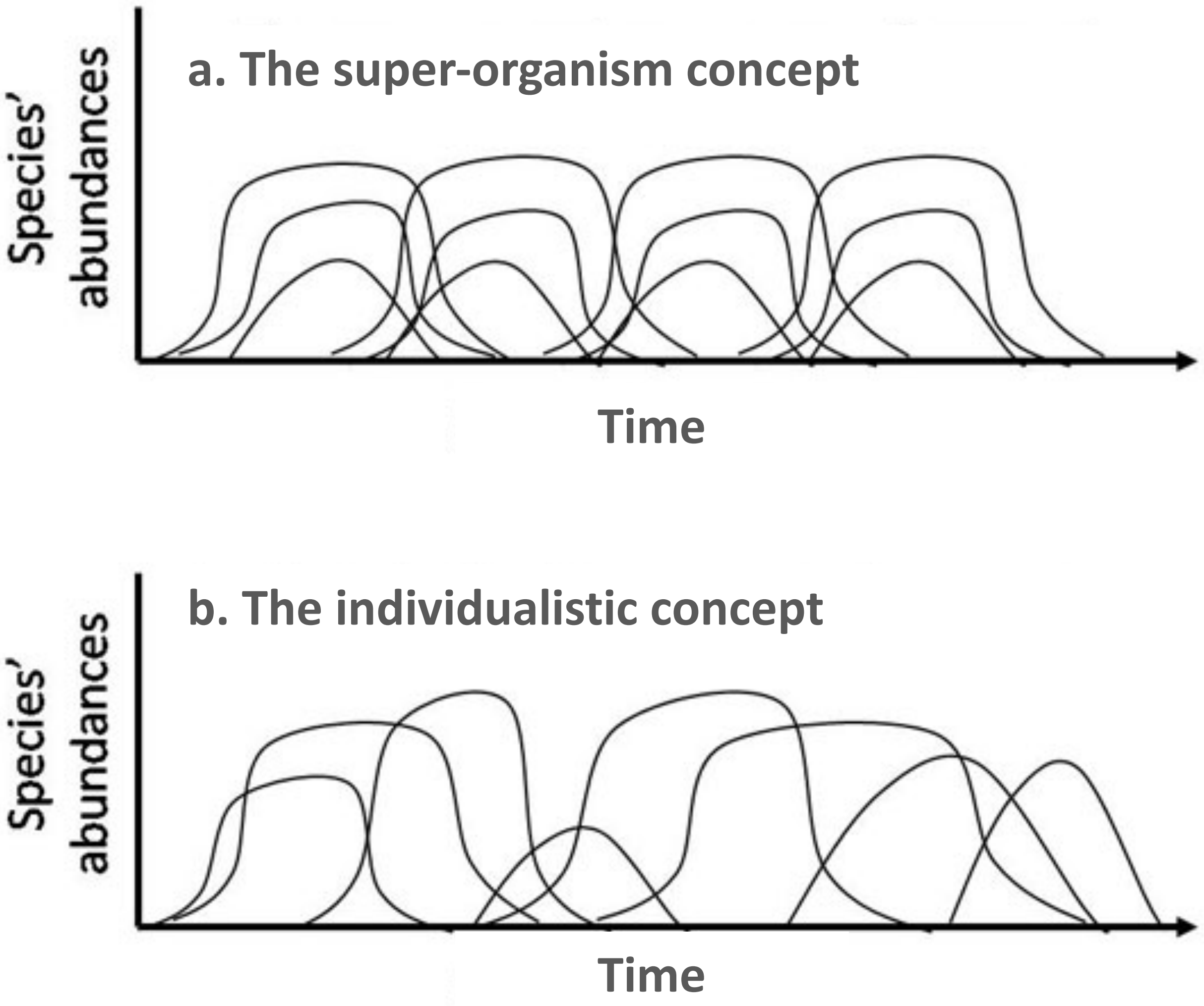
-Communities are dependent on the species interactions among members

Henry Gleason (1917)

- "Each species is an independent entity"

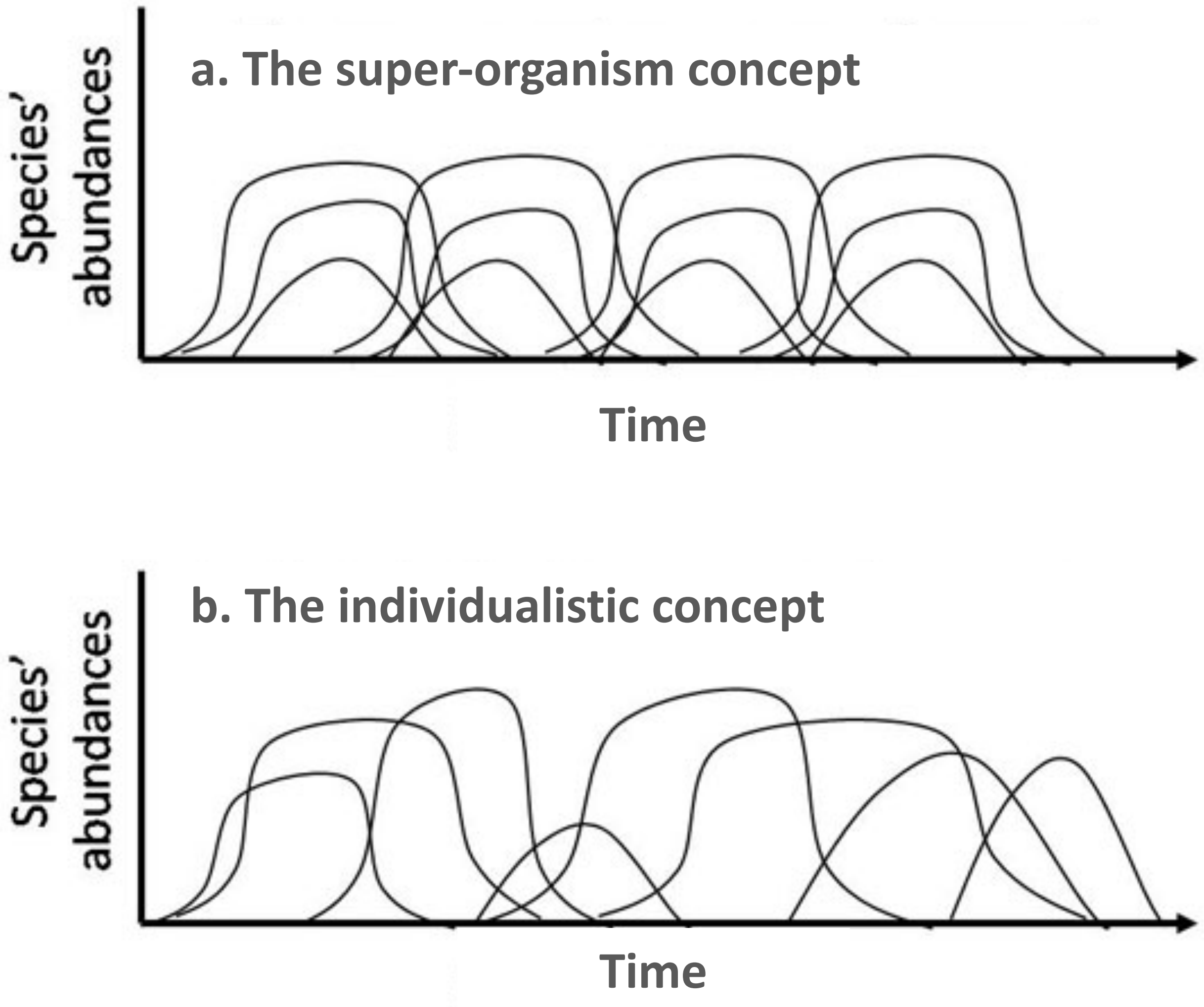
    -distribution of each species is due to genetics, physiology, life history, and relations to abiotic and biotic factors

    -communities are just a consequence of independent interactions among species and their environment



(A) Groups of species are tightly associated, and are supplanted by other groups of tightly associated species. (B) Individual species independently respond to environmental conditions. Each curve on the graphs represents the abundance of a single species.

Emery, S. (2010)



(A) Groups of species are tightly associated, and are supplanted by other groups of tightly associated species. (B) Individual species independently respond to environmental conditions. Each curve on the graphs represents the abundance of a single species.

**note on phylogeographic concordance

Emery, S. (2010)

Assembly “rules”

- Jared Diamond 1975 – the community is a product of the species pool, the environment, and the interactions amongst species.

- Problem was, if all of these things seems similar, why were communities behaving differently? **cough** (gleason)

Assembly "rules"

- The whole debate begins again in a slightly different context.

- 1.) 1:1 match between species and environment, even if the timing of species entering a community is different
  - (Neill 1975, Tilman et al. 1986, Sommer 1991, Law & Morton 1996)

- 2.) depends on what species get there first (multiple state equilibria)
  - (Luh and Pimm 1993, Law 1999, Robinson and Dickerson 1988, Drake 1991, Samuels & Drake 1997)

Assembly "rules"

- The whole debate begins again in a slightly different context.

- 1.) 1:1 match between species and envi… he timing of species entering a community i…
  - (Neill 1975, Til… Morton 1996)

Evidence has been found for both arguments

- 2.) d… species get there first (multiple state equilibria)
  - (L… Pimm 1993, Law 1999, Robinson and Dickerson 1988, Drake 1991, Samuels & Drake 1997)

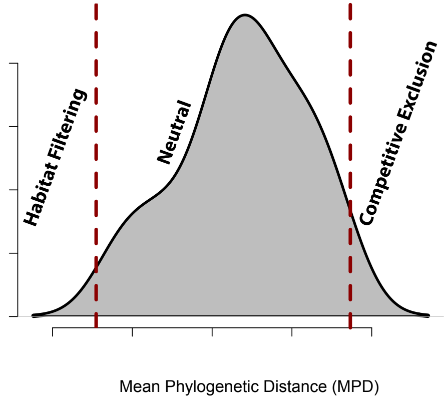
Measure Community Structure

- Species/genus ratios
  - Moreau 1948, Simberloff 1970
- Abundance distributions
  - Simpson 1949, Magurran 1988
- Higher Taxonomic diversity metrics
  - Faith 1992, Gotelli & Colwell 2001
- Functional traits information
  - Macarthur & Levins 1967; Weiher *et al*. 1999; McGill *et al*. 2006

Measure Community Structure

- Species/genus ratios
  - Moreau 1948, Simberloff 1970
- Abundance distributions
  - Simpson 1949, Magurran 1988
- Higher Taxonomic diversity metrics
  - Faith 1992, Gotelli & Colwell 2001
- Functional traits information
  - Macarthur & Levins 1967; Weiher *et al*. 1999; McGill *et al*. 2006

Null Models in Community Ecology (Gotelli & Graves 1996)



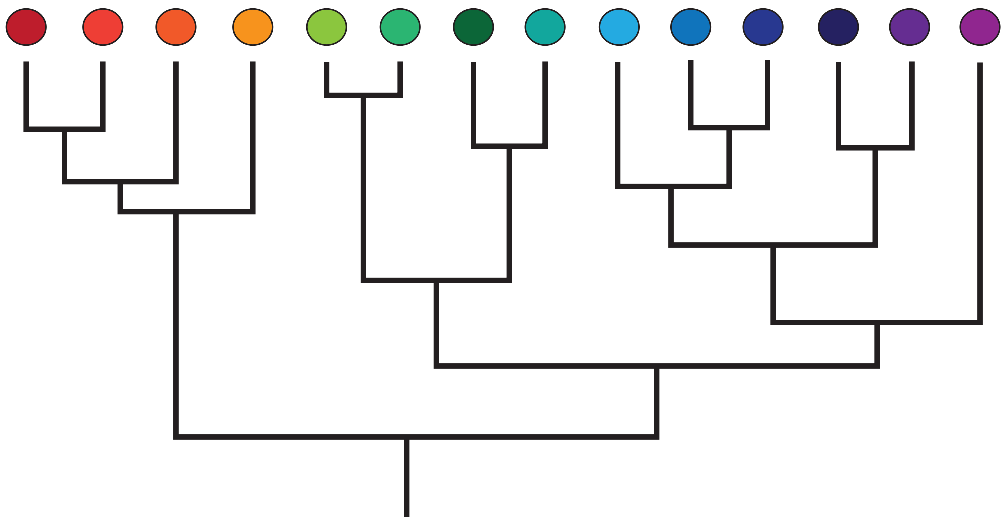
Phylogenetic Community Ecology

- We should use phylogenies in community ecology!
  - Brooks & McLennan 1991, Losos 1996, Thompson et al. 2001
- The phylogeny encapsulates information about the functional traits important for assembly, as they are "phylogenetic conserved"
  - **assumption

Webb et al. 2002

Phylogenetic Community Ecology

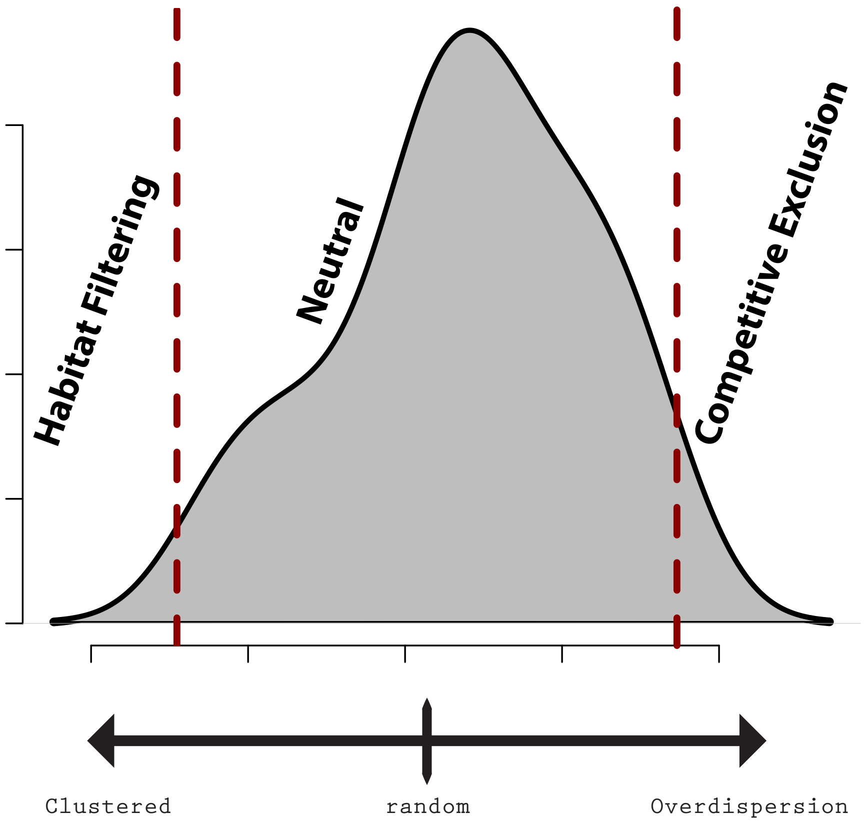
- We should use phylogenies in community ecology!
  - Brooks & McLennan 1991, Losos 1996, Thompson et al. 2001
- The phylogeny encapsulates information about the functional traits important for assembly, as they are "phylogenetic conserved"
  - **assumption



Webb et al. 2002

Phylogenetic Community Ecology

- Is the distribution of species in a community nonrandom with respect to phylogeny?

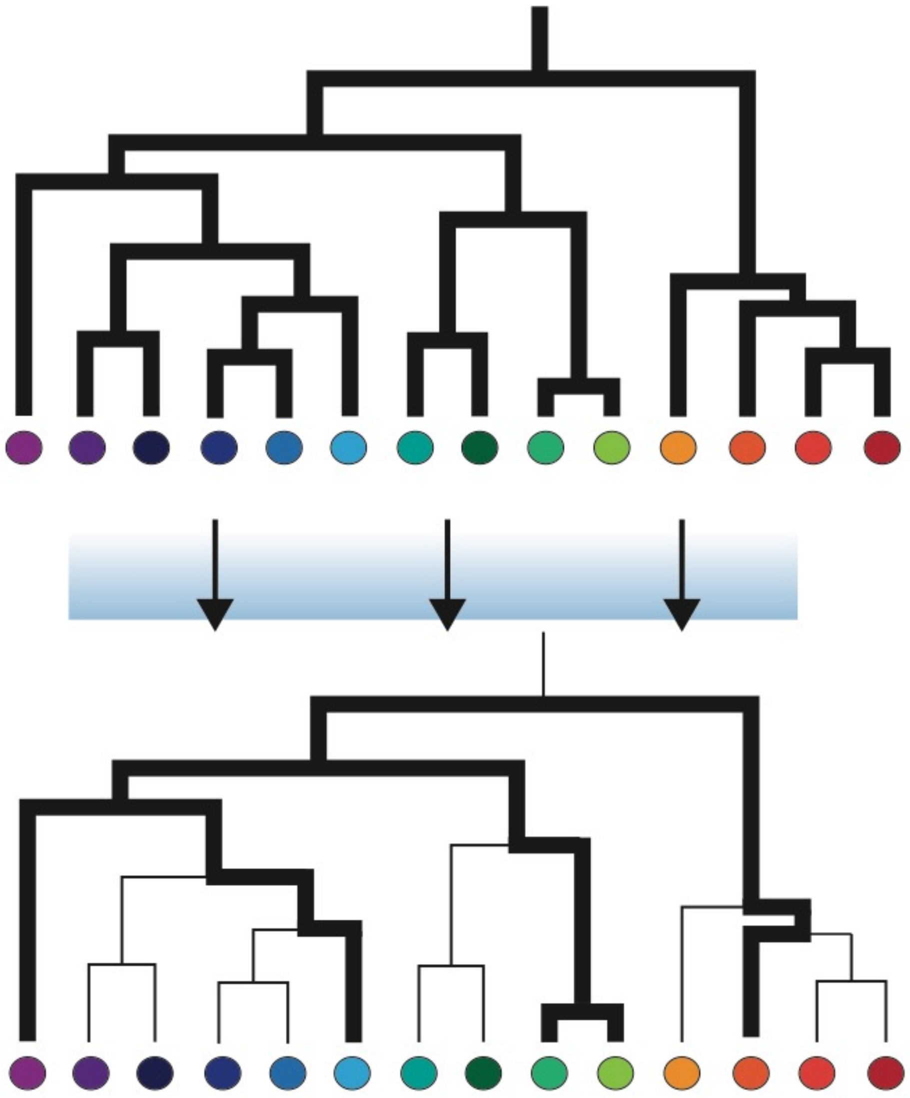


Regional Species Pool

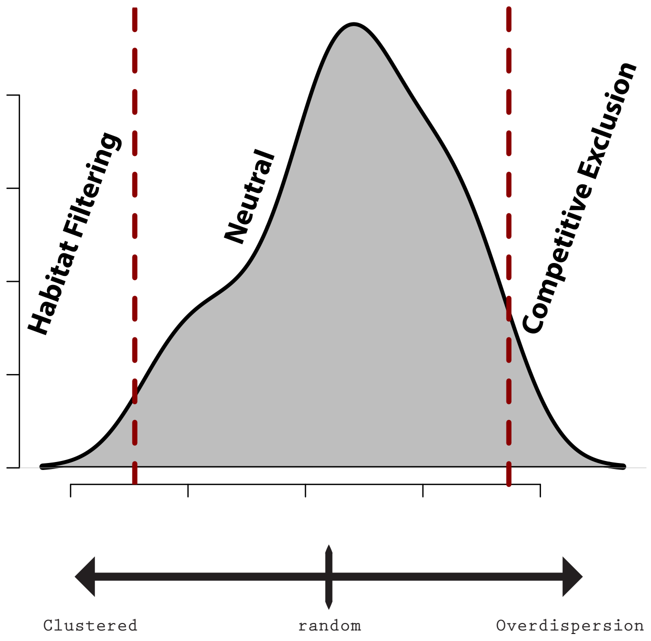
Environmental Filter

Competitive Exclusion

Local Species Pool



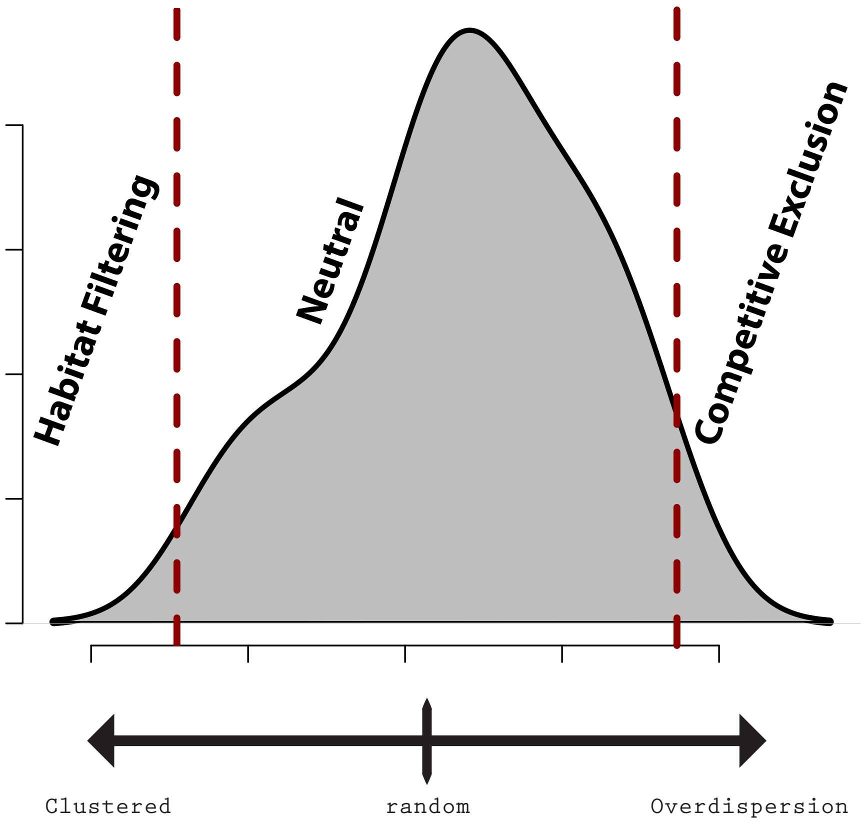
Phylogenetic dispersion metrics: Mean Phylogenetic Distance (MPD)



Non-neutral Assembly Models

- Is the distribution of species in a community nonrandom with respect to phylogeny?

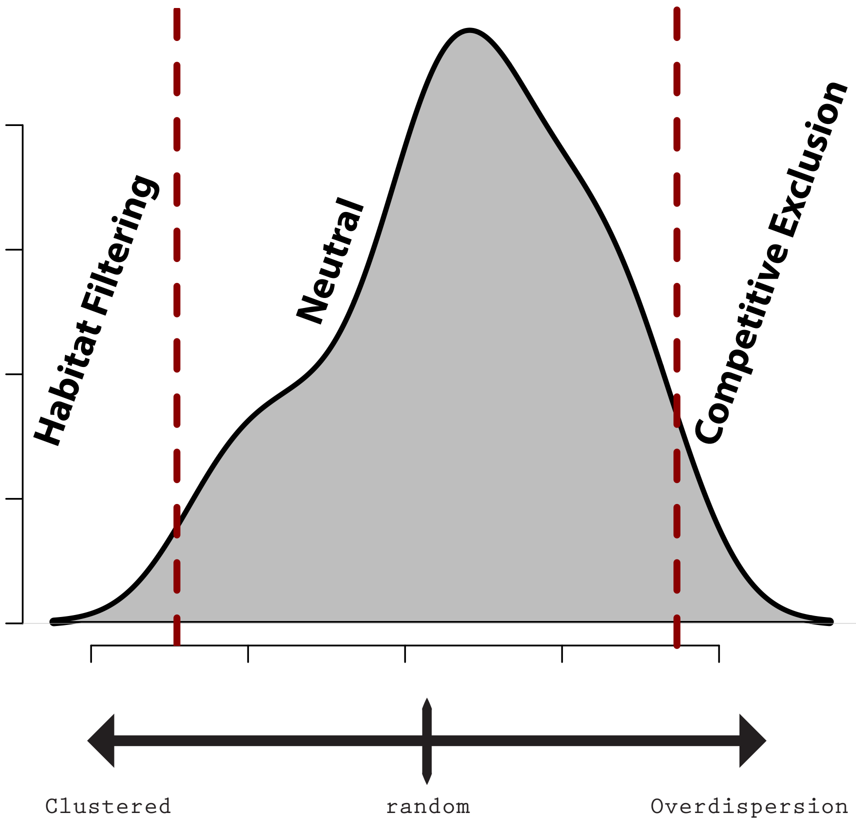
Habitat Filtering – species fail to establish in community due to incompatibility with relevant environmental factors (Bazzaz 1991)



Non-neutral Assembly Models

- Is the distribution of species in a community nonrandom with respect to phylogeny?

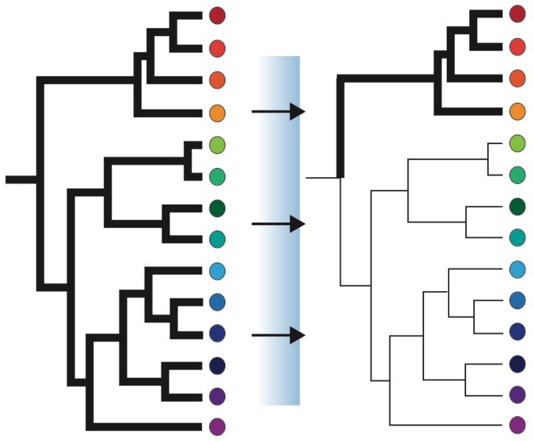
Habitat Filtering – species fail to establish in community due to incompatibility with relevant environmental factors (Bazzaz 1991)

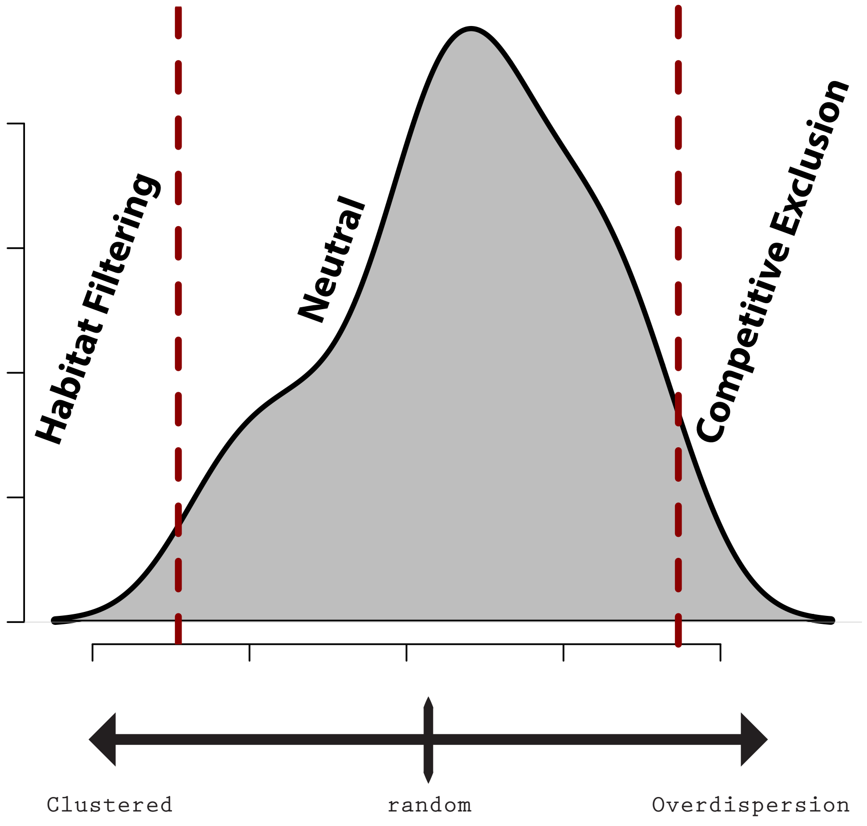


Competitive Exclusion – species fail to establish in a community when they cannot out-compete another species for their desired niche (Macarthur & Levins 1967)

Traits are phylogenetically conserved

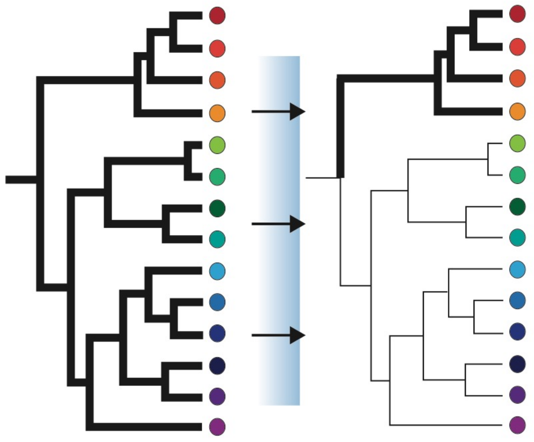
Habitat Filtering =
phylogenetic underdispersion



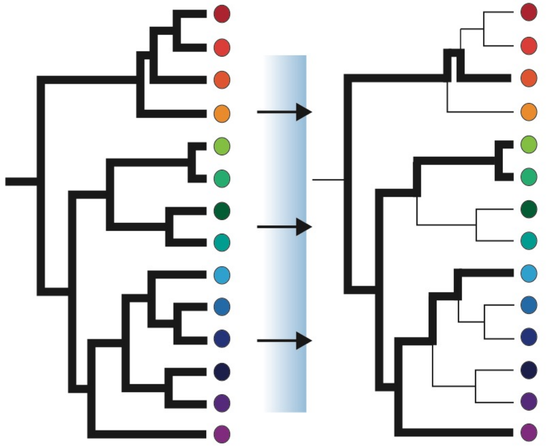


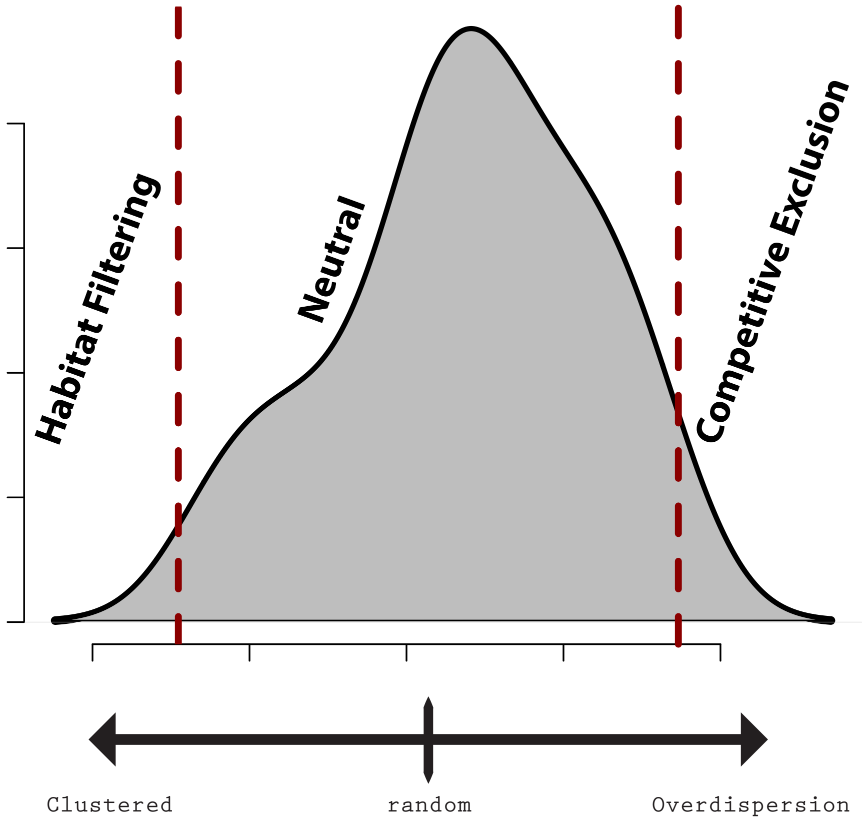
Traits are phylogenetically conserved

Habitat Filtering =
phylogenetic underdispersion



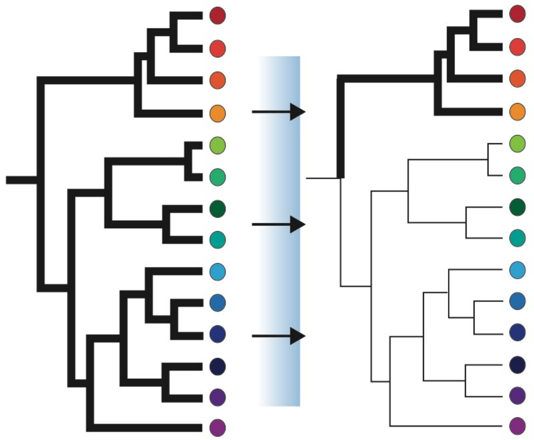
Competitive Exclusion =
phylogenetic overdispersion



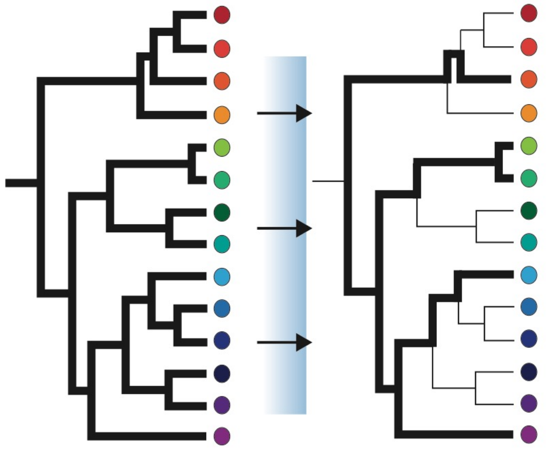


Traits are phylogenetically conserved

Habitat Filtering =
phylogenetic underdispersion



Competitive Exclusion =
phylogenetic overdispersion



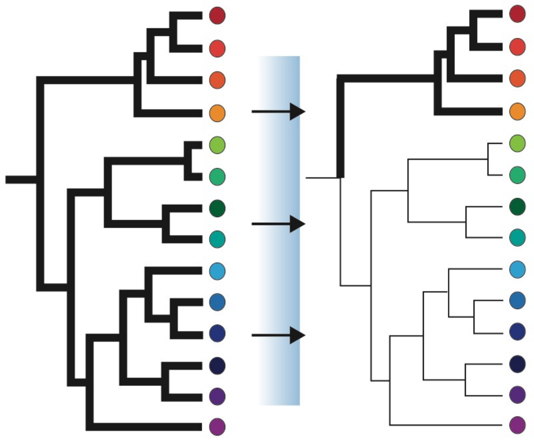
Traits are phylogenetically convergent

Habitat Filtering =
phylogenetic overdispersion

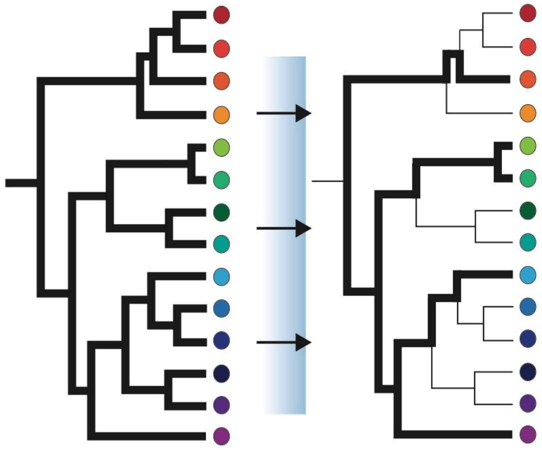
Competitive Exclusion =
phylogenetic underdispersion

Traits are phylogenetically conserved

Habitat Filtering = phylogenetic underdispersion

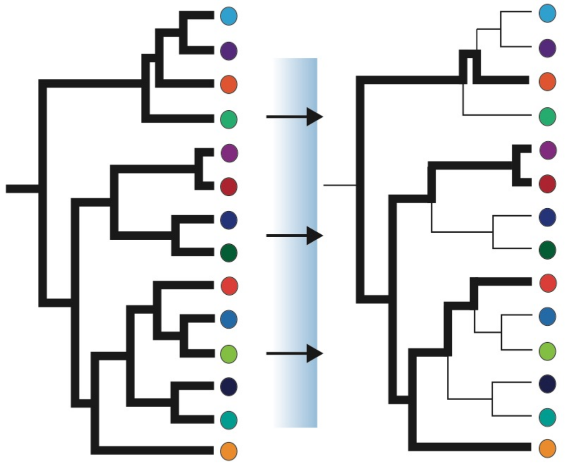


Competitive Exclusion = phylogenetic overdispersion

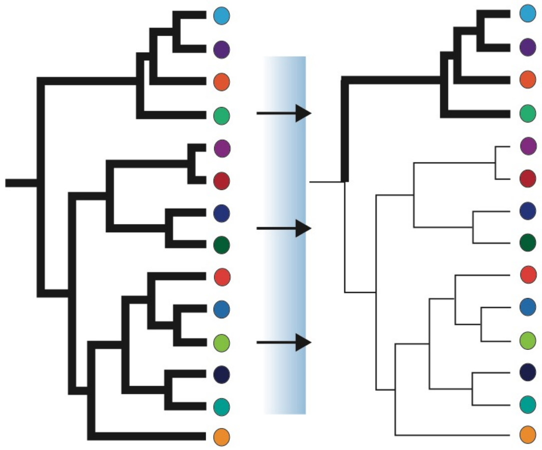


Traits are phylogenetically convergent

Habitat Filtering = phylogenetic overdispersion



Competitive Exclusion = phylogenetic underdispersion



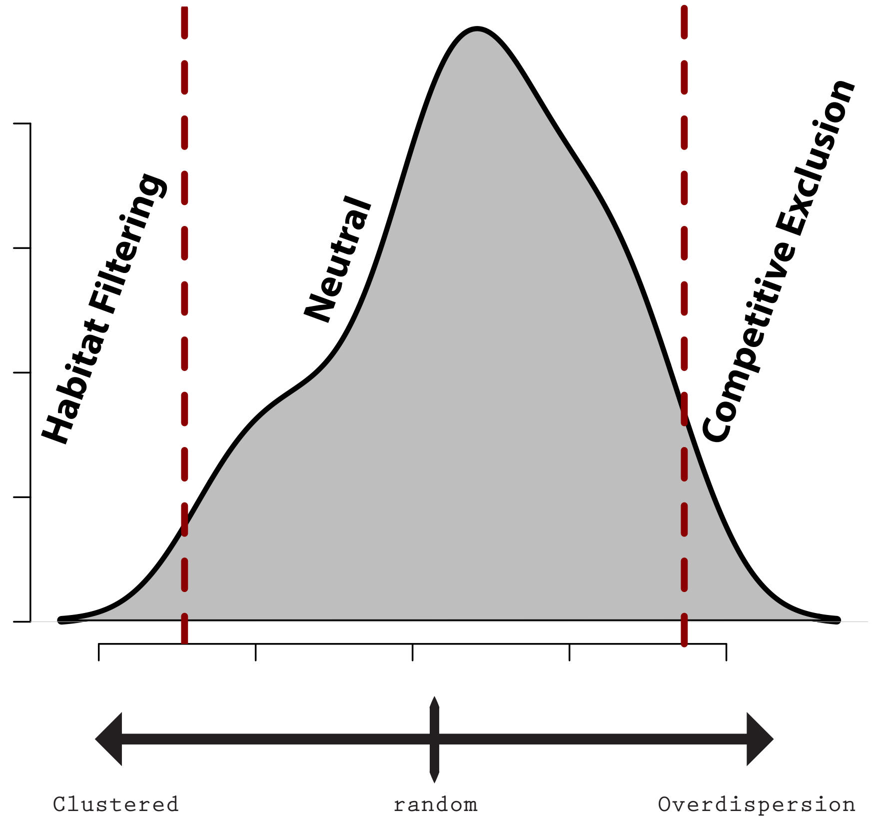
Community Assembly Inference; dispersion metrics

- Phylogenetic Community Structure

  Grandcolas 1998, Webb et al. 2000, 2002, Kraft et al. 2007

- Trait Community Structure

  Weiher et al. 1999, McGill et al. 2006, Cornwell et al. 2006, Kraft et al. 2007



Community Assembly Inference; dispersion metrics

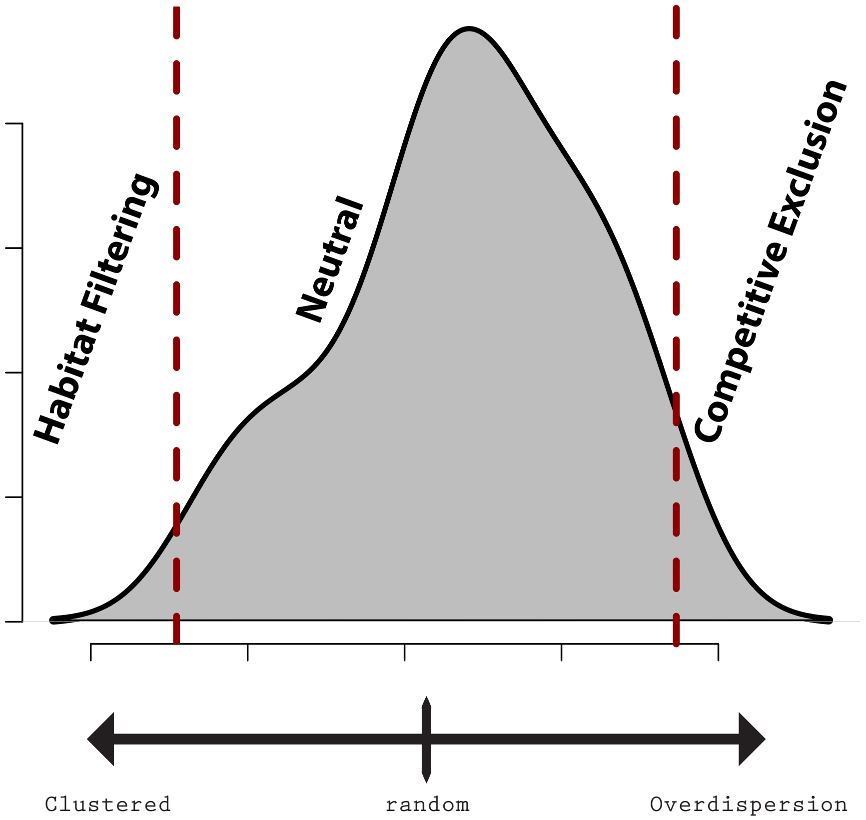
- Phylogenetic Community Structure

  Grandcolas 1998, Webb et al. 2000, 2002, Kraft et al. 2007

- Trait Community Structure

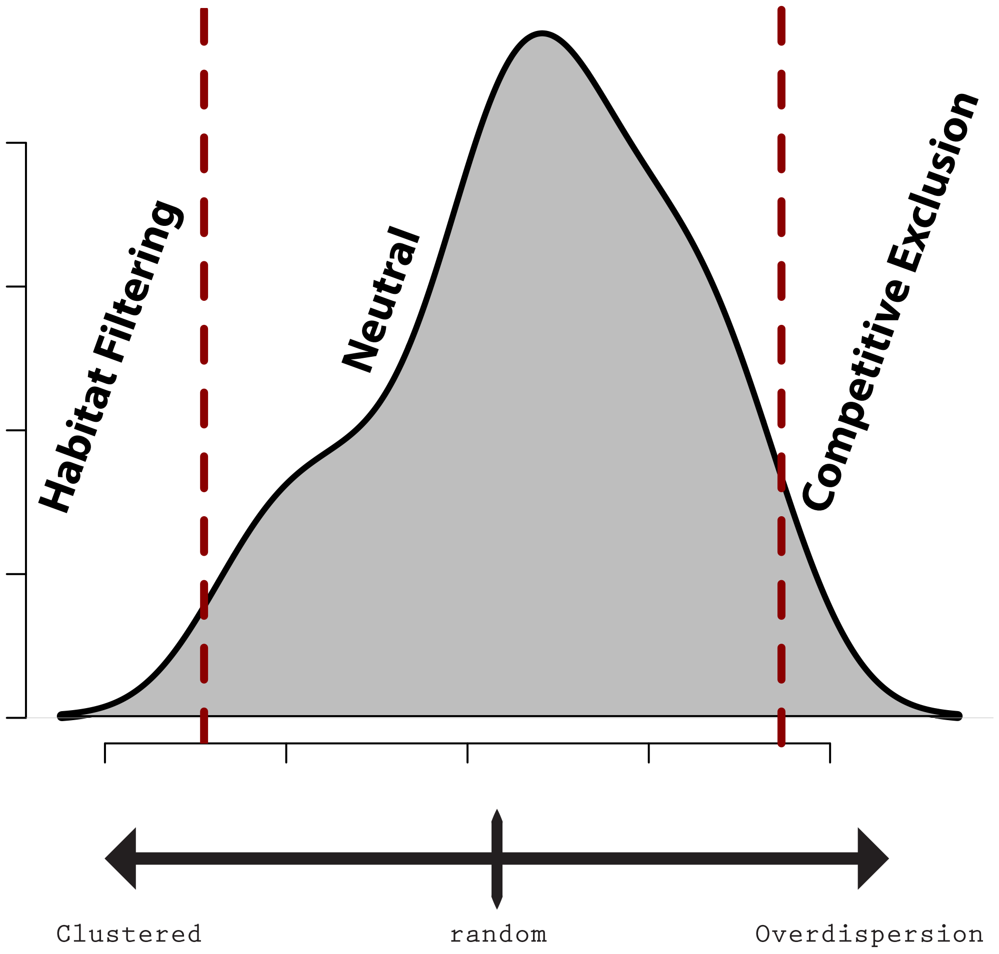
  Weiher et al. 1999, McGill et al. 2006, Cornwell et al. 2006, Kraft et al. 2007

Weiher and Keddy 1995, Kraft et al. 2007, 2010, Cavender-Bares et al. 2009; Kemble 2009; Mayfield and Levine 2010; Gerhold et al. 2015



Current Community Assembly Inference Methods

- Relying on often-violated assumptions
- Statistical significance != Biological significance
- Cannot simultaneously compare model support
- Not widely comparable across communities



CAMI development objectives

1. *Implement a model-based inference procedure*

CAMI development objectives

1. *Implement a model-based inference procedure*
    - *Compare support for neutral and non-neutral models simultaneously*

CAMI development objectives

1. *Implement a model-based inference procedure*
   - *Compare support for neutral and non-neutral models simultaneously*

   ***Cannot write down a likelihood, so we will need to use an approximate approach based on simulations*

CAMI development objectives

1. *Implement a model-based inference procedure*
   - *Compare support for neutral and non-neutral models simultaneously*

   ***Cannot write down a likelihood, so we will need to use an approximate approach based on simulations*

2. *Parameterize the strength of the non-neutral assembly processes*

CAMI development objectives

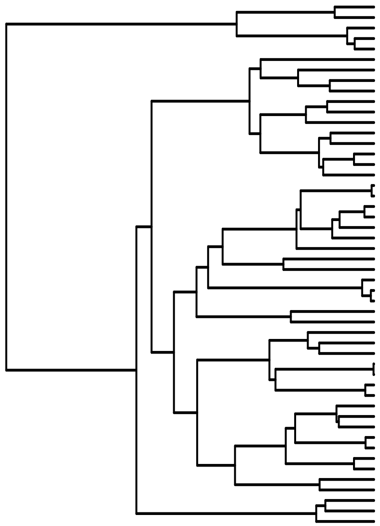
1. ***Implement a model-based inference procedure***
   - *Compare support for neutral and non-neutral models simultaneously*

   ***Cannot write down a likelihood, so we will need to use an approximate approach based on simulations*

2. *Parameterize the strength of the non-neutral assembly processes*

Objective 1: Implement a model-based inference procedure

1.1 Regional community phylogeny: N, λ, μ

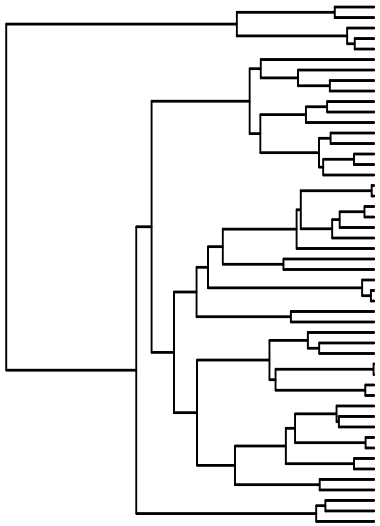


N	regional species pool
λ	speciation rate
μ	extinction rate
σ^2	rate of character change
α	strength of constraints (OU only)
n	local species pool
t_E	effect of environmental filtering
t_C	effect of competitive exclusions

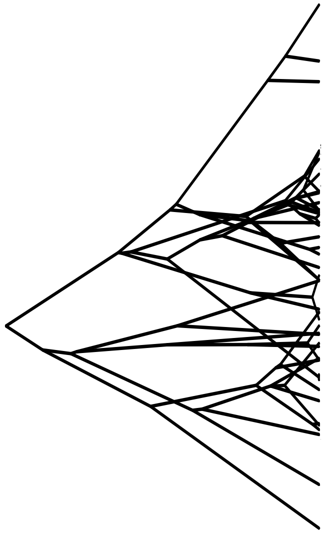
Ruffley *et al.* 2019 *Ecol. & Evol. Accepted*

Objective 1: Implement a model-based inference procedure

1.1 Regional community phylogeny: N, λ, μ



1.2 Trait evolution: σ^2, α



N	regional species pool
λ	speciation rate
μ	extinction rate
σ^2	rate of character change
α	strength of constraints (OU only)
n	local species pool
t_E	effect of environmental filtering
t_C	effect of competitive exclusions

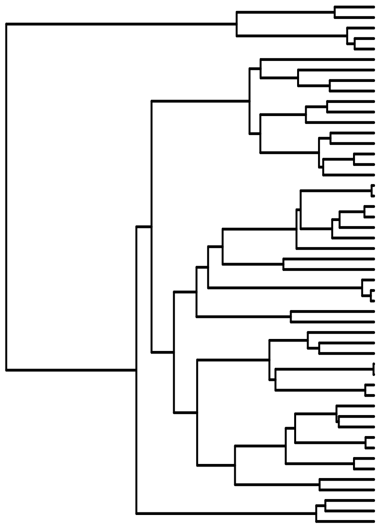
Ruffley *et al.* 2019 *Ecol. & Evol. Accepted*

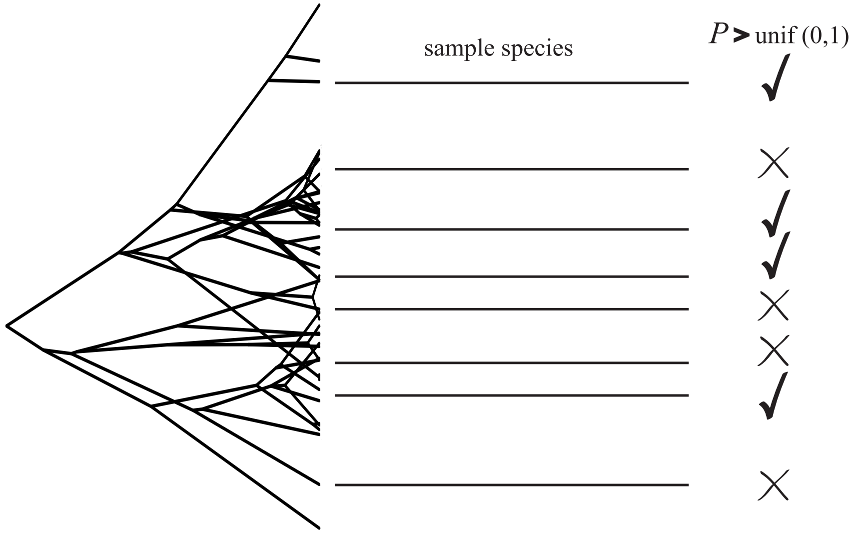
Objective 1: Implement a model-based inference procedure

1.1 Regional community phylogeny: N, λ, μ

1.2 Trait evolution: σ^2, α

1.3 Local community assembly: n, t_E, t_C





N	regional species pool
λ	speciation rate
μ	extinction rate
σ^2	rate of character change
α	strength of constraints (OU only)
n	local species pool
t_E	effect of environmental filtering
t_C	effect of competitive exclusions

- Neutral
- Environmental Filtering
- Competitive Exclusion

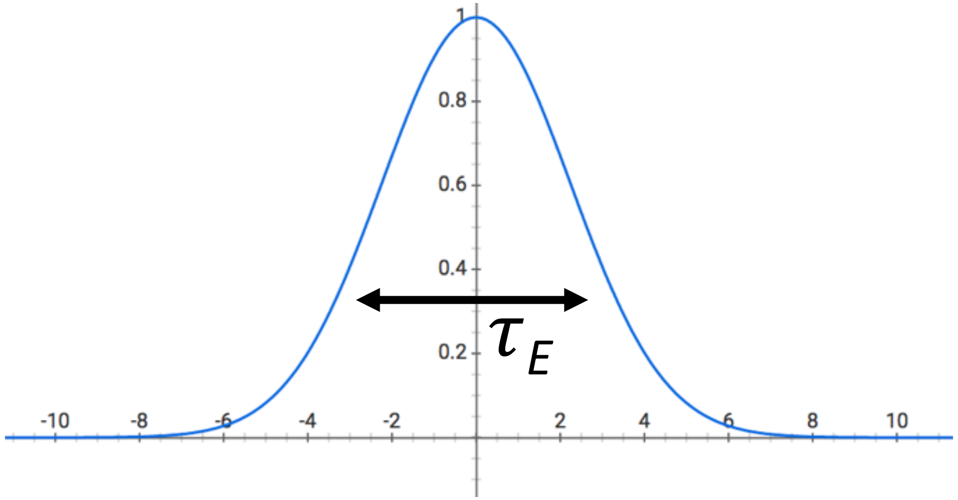
Ruffley *et al.* 2019 *Ecol. & Evol. Accepted*

Objective 2: Parameterize the strength of the non-neutral assembly processes

Objective 2: Parameterize the strength of the non-neutral assembly processes

Environmental Filtering

$$P = e^{-\frac{(Zi-Zj)^2}{\tau}}$$

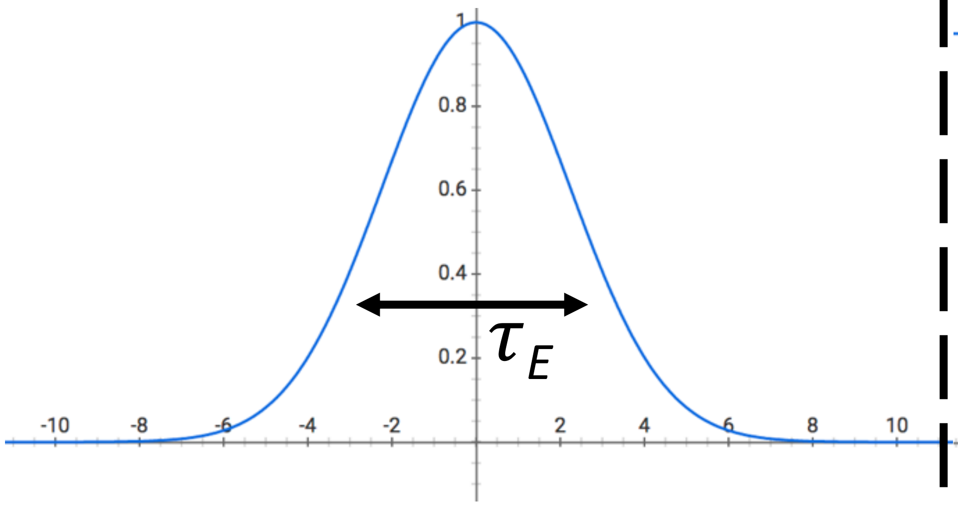


Ruffley *et al.* 2019 *Ecol. & Evol. Accepted*

Objective 2: Parameterize the strength of the non-neutral assembly processes

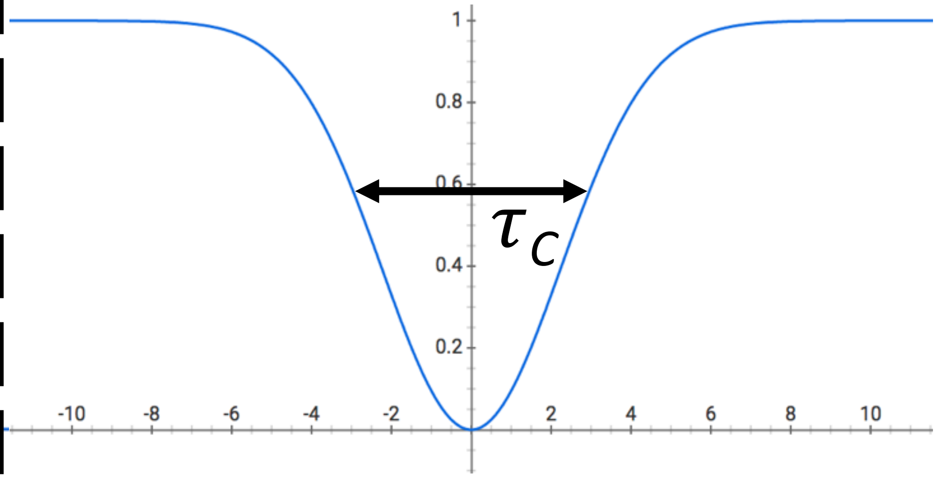
Environmental Filtering

$$P = e^{-\frac{(Zi-Zj)^2}{\tau}}$$



Competitive Exclusion

$$P = 1 - e^{-\frac{(Zi-Zj)^2}{\tau}}$$



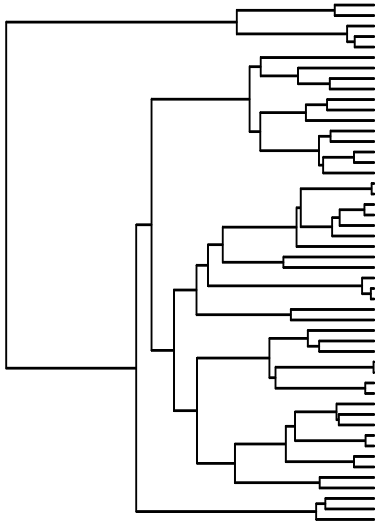
Ruffley *et al.* 2019 *Ecol. & Evol. Accepted*

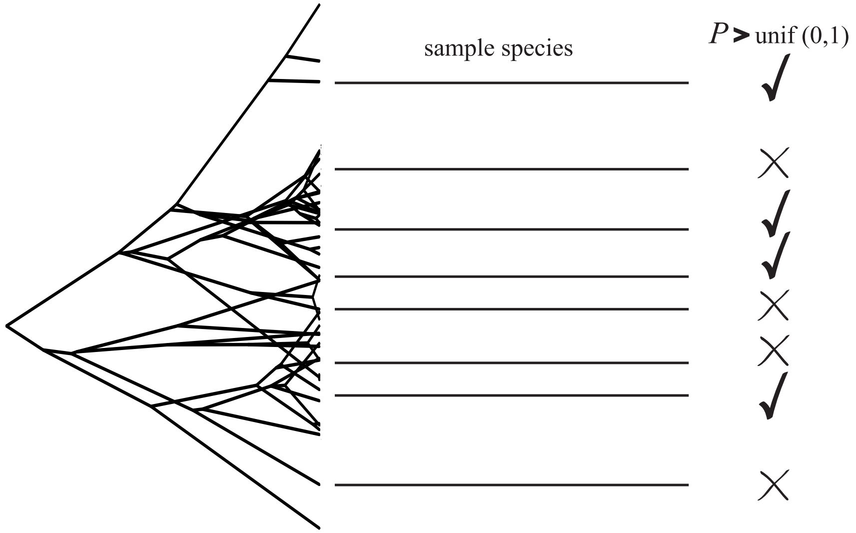
Objective 1: Implement a model-based inference procedure

1.1 Regional community phylogeny: N, λ, μ

1.2 Trait evolution: σ^2, α

1.3 Local community assembly: n, t_E, t_C



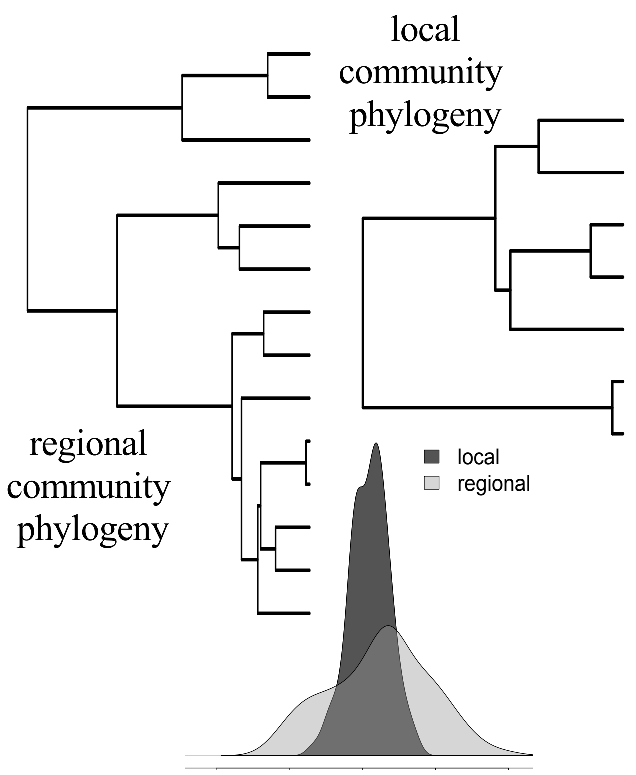


N	regional species pool
λ	speciation rate
μ	extinction rate
σ^2	rate of character change
α	strength of constraints (OU only)
n	local species pool
t_E	effect of environmental filtering
t_C	effect of competitive exclusions

- Neutral
- Environmental Filtering
- Competitive Exclusion

Ruffley *et al.* 2019 *Ecol. & Evol. Accepted*

Objective 1: Implement a model-based inference procedure



Simulation #	*SS.1*	*SS.2*	*SS.3*	*…..*	*SS.30*	*Model*
1	29.3201204	0.047795881	2.202369339	0.024039567	-0.551807157	**Neutral**
2	58.01018705	0.05683677	4.296006276	0.014001377	-0.141285853	**Neutral**
3	28.35028897	0.057748823	2.565336481	0.02472052	0.543737366	**Neutral**
4	54.4519145	0.058161893	4.462506196	0.006577184	-0.297716652	**Neutral**
5	29.52415799	0.0449925	1.332280042	-0.027307941	-0.424597011	**Neutral**
6	13.8002455	0.043015337	3.827085961	0.052868925	0.00038215	**Filtering**
7	68.51513963	0.05221236	3.773466742	0.006122832	-0.373019012	**Filtering**
8	6.409247373	0.067539483	1.27370363	0.056550081	-0.419361409	**Filtering**
9	7.869379132	0.044893904	1.621625795	0.005944114	-0.225941885	**Filtering**
10	56.74215972	0.042376631	3.870139564	-0.010551762	1.612683176	**Filtering**
11	32.21489305	0.039088122	4.436192932	-0.005017495	-0.608370405	**Competition**
12	35.31406132	0.055897014	1.802319164	0.00665565	0.575928644	**Competition**
…	15.17747744	0.052302964	2.833102254	-0.007315425	-0.343133625	**Competition**
10,000	176.8646799	0.046759559	3.404803136	0.003965272	-0.492954288	**Competition**

Ruffley *et al.* 2019 *Ecol. & Evol. Accepted*

Objective 1: Implement a model-based inference procedure

Model-selection using approximate approaches

- randomForests (*Breiman 2001, Breiman and Cutler 2007*)**
- Approximate Bayesian Computation (ABC; *Beaumont, et al. 2002, Csilléry, et al. 2012)*

Simulation #	*SS.1*	*SS.2*	*SS.3*	*…..*	*SS.30*	*Model*
1	29.3201204	0.047795881	2.202369339	0.024039567	-0.551807157	Neutral
2	58.01018705	0.05683677	4.296006276	0.014001377	-0.141285853	Neutral
3	28.35028897	0.057748823	2.565336481	0.02472052	0.543737366	Neutral
4	54.4519145	0.058161893	4.462506196	0.006577184	-0.297716652	Neutral
5	29.52415799	0.0449925	1.332280042	-0.027307941	-0.424597011	Neutral
6	13.8002455	0.043015337	3.827085961	0.052868925	0.00038215	Filtering
7	68.51513963	0.05221236	3.773466742	0.006122832	-0.373019012	Filtering
8	6.409247373	0.067539483	1.27370363	0.056550081	-0.419361409	Filtering
9	7.869379132	0.044893904	1.621625795	0.005944114	-0.225941885	Filtering
10	56.74215972	0.042376631	3.870139564	-0.010551762	1.612683176	Filtering
11	32.21489305	0.039088122	4.436192932	-0.005017495	-0.608370405	Competition
12	35.31406132	0.055897014	1.802319164	0.00665565	0.575928644	Competition
…	15.17747744	0.052302964	2.833102254	-0.007315425	-0.343133625	Competition
10,000	176.8646799	0.046759559	3.404803136	0.003965272	-0.492954288	Competition

Ruffley *et al.* 2019 *Ecol. & Evol. Accepted*

Results

1. *Power Analysis*

2. *Parameter Estimation*

Results

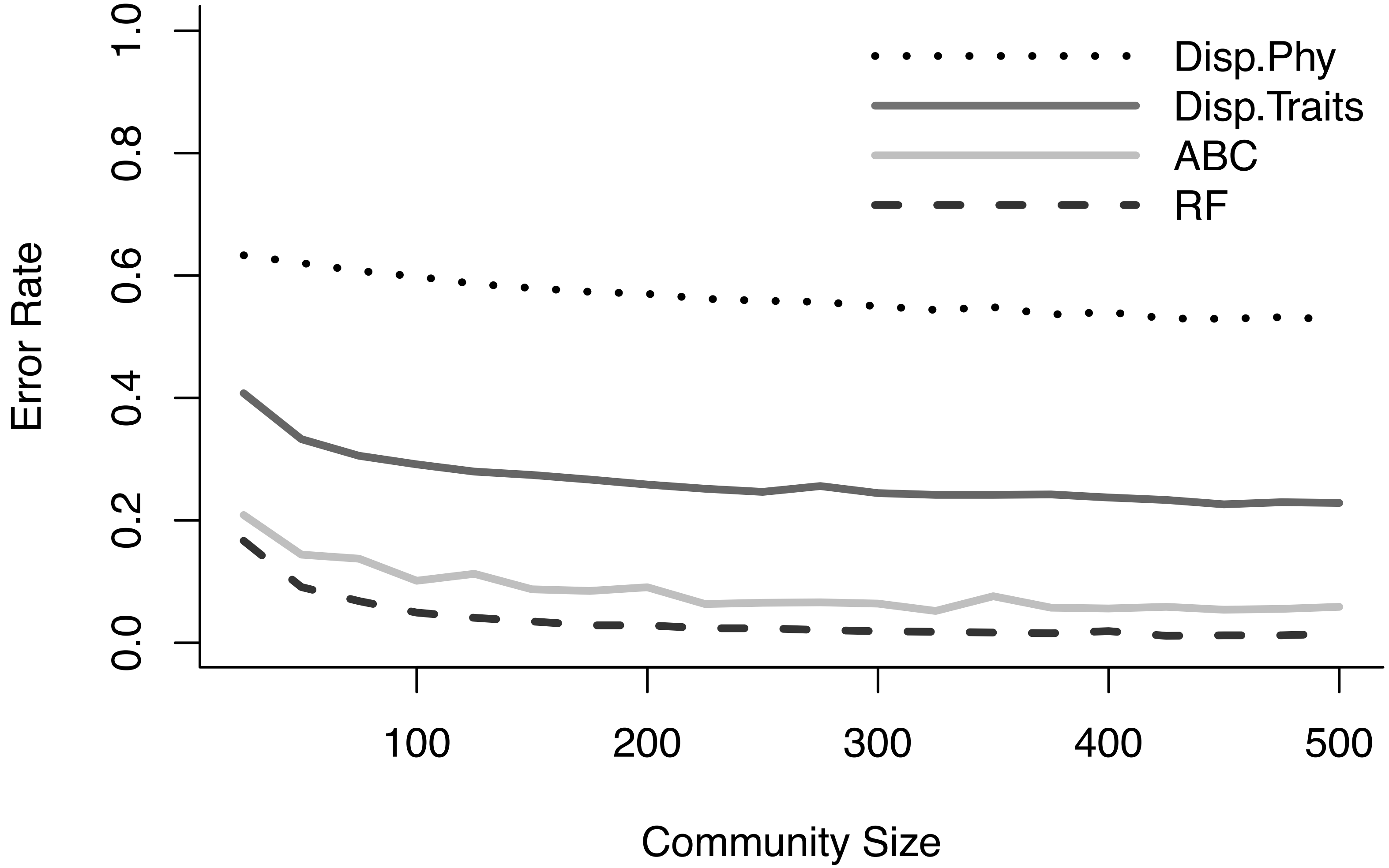
1. *Power Analysis*
   - ***Dispersion Metrics***

     Phylogenetic data

     Functional trait data

   - ***randomForest (RF)***
   - ***ABC***

Results: Power Analysis



Ruffley *et al.* 2019 *Ecol. & Evol. Accepted*

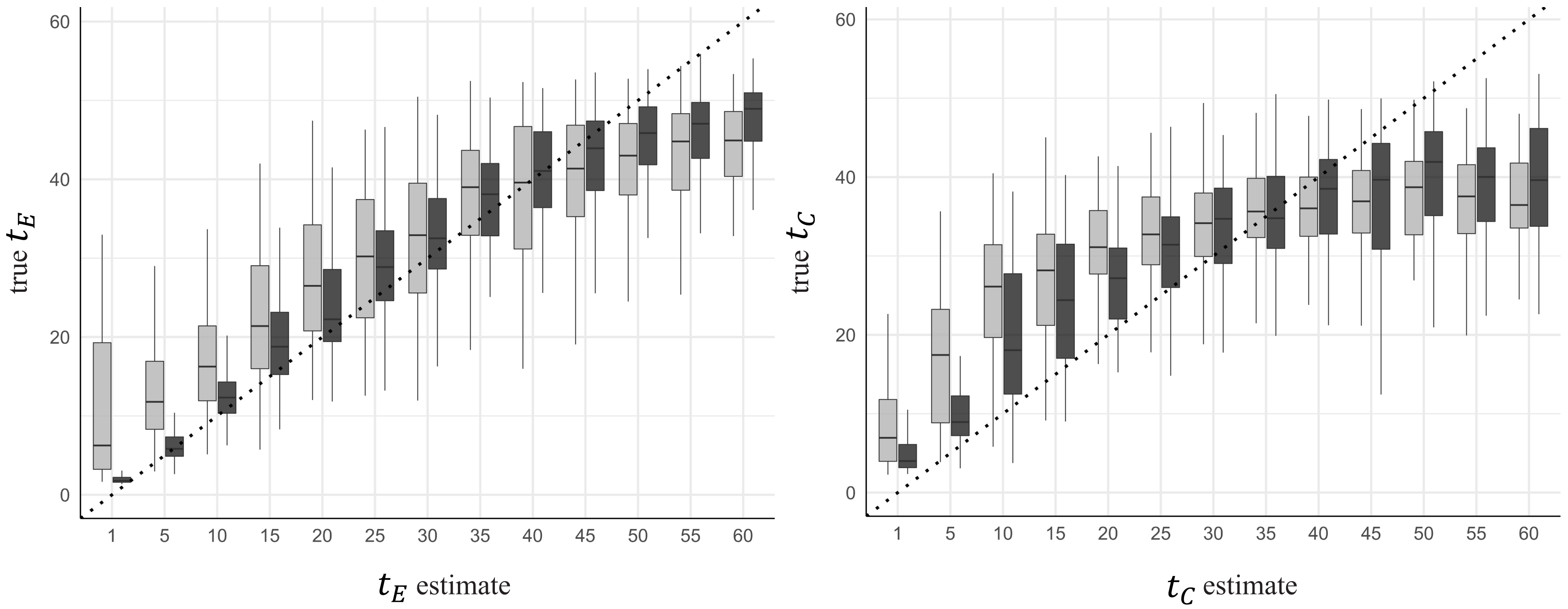
Results

1. *Power Analysis*

2. ***Parameter Estimation***

   -Estimate τ_E and t_C

Results: Parameter Estimation



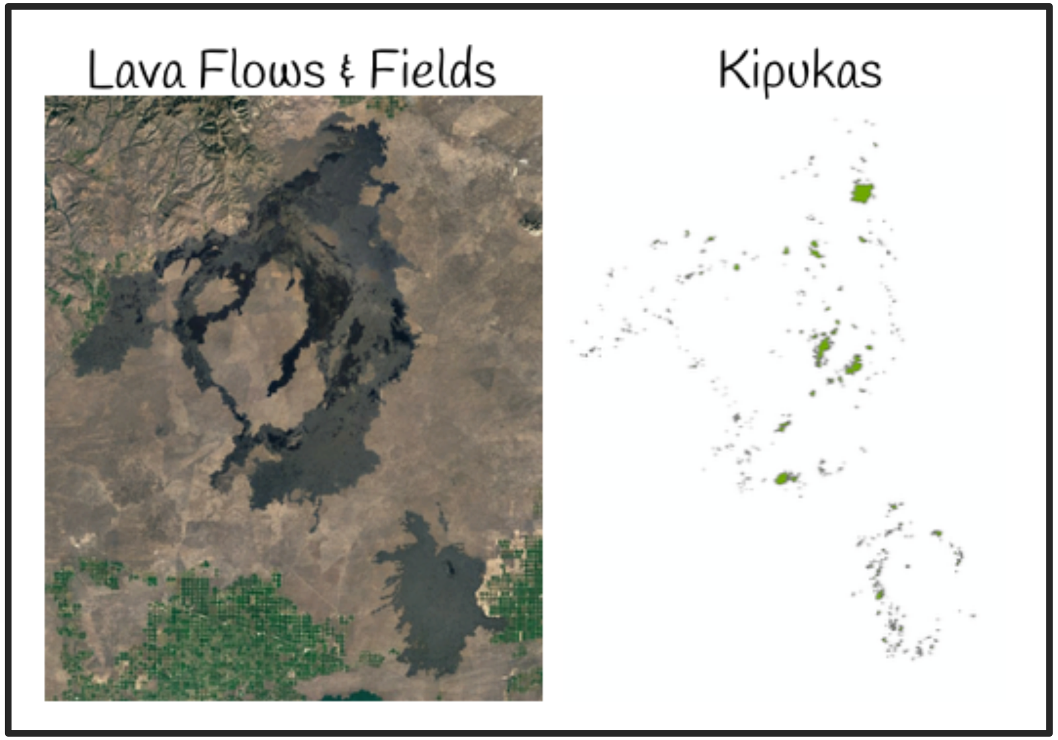
Ruffley *et al.* 2019 *Ecol. & Evol. Accepted*

CAMI introductory tutorial

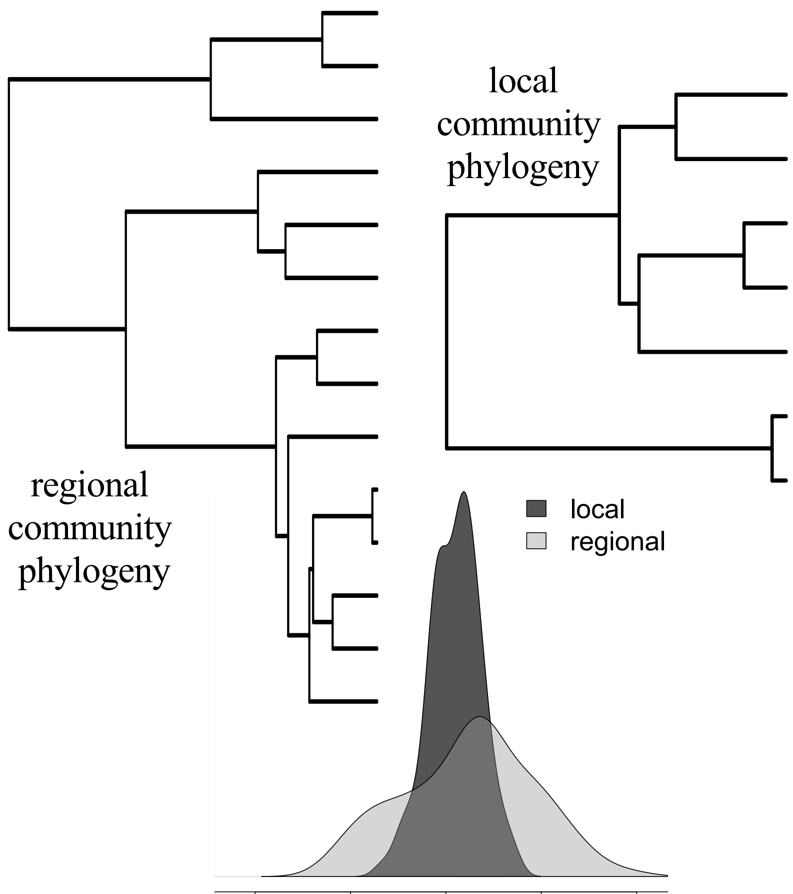
- compphylo.github.io/Oslo2019/CAMI_files/CAMI_1.html

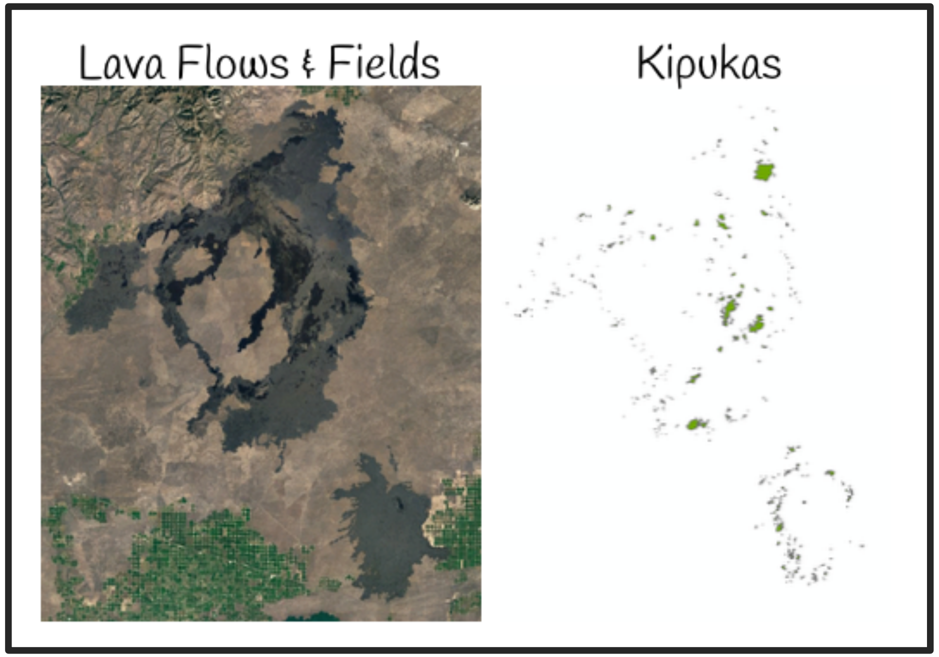
Results: Lava-flow Islands in Craters of the Moon National Monument and Preserve

Katie Peterson
Dr. Christine Parent lab
University of Idaho



Results: Lava-flow Islands in Craters of the Moon National Monument and Preserve





Results: Lava-flow Islands in Craters of the Moon National Monument and Preserve

Smith, S. A., and J. W. Brown. 2018. Constructing a broadly inclusive seed plant phylogeny. American Journal of Botany 105(3): 302-314.

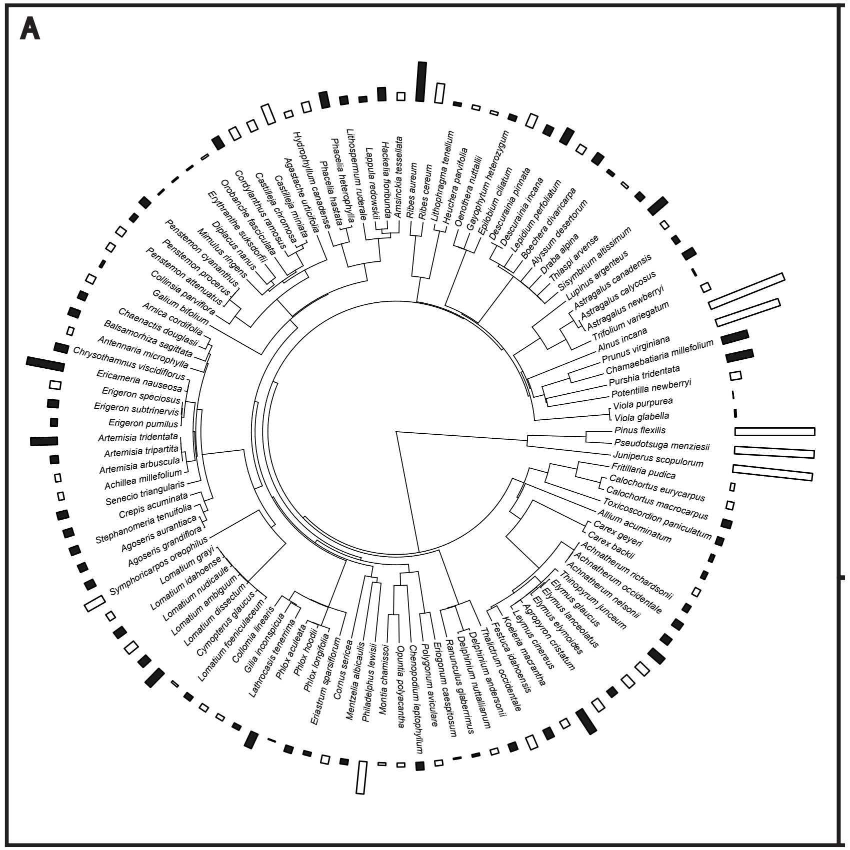
- Spermatophyta phylogeny 79,881 tips
- Discarded 79,768 tips (99.9%)

- 8 kipuka communities
- 113 species in regional community
- 18-60 species in local community

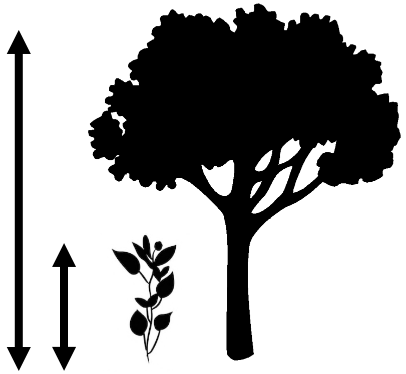
Try it with your community!

https://github.com/ruffleymr/CommunityPhylogeneticExample

Results: Lava-flow Islands in Craters of the Moon National Monument and Preserve

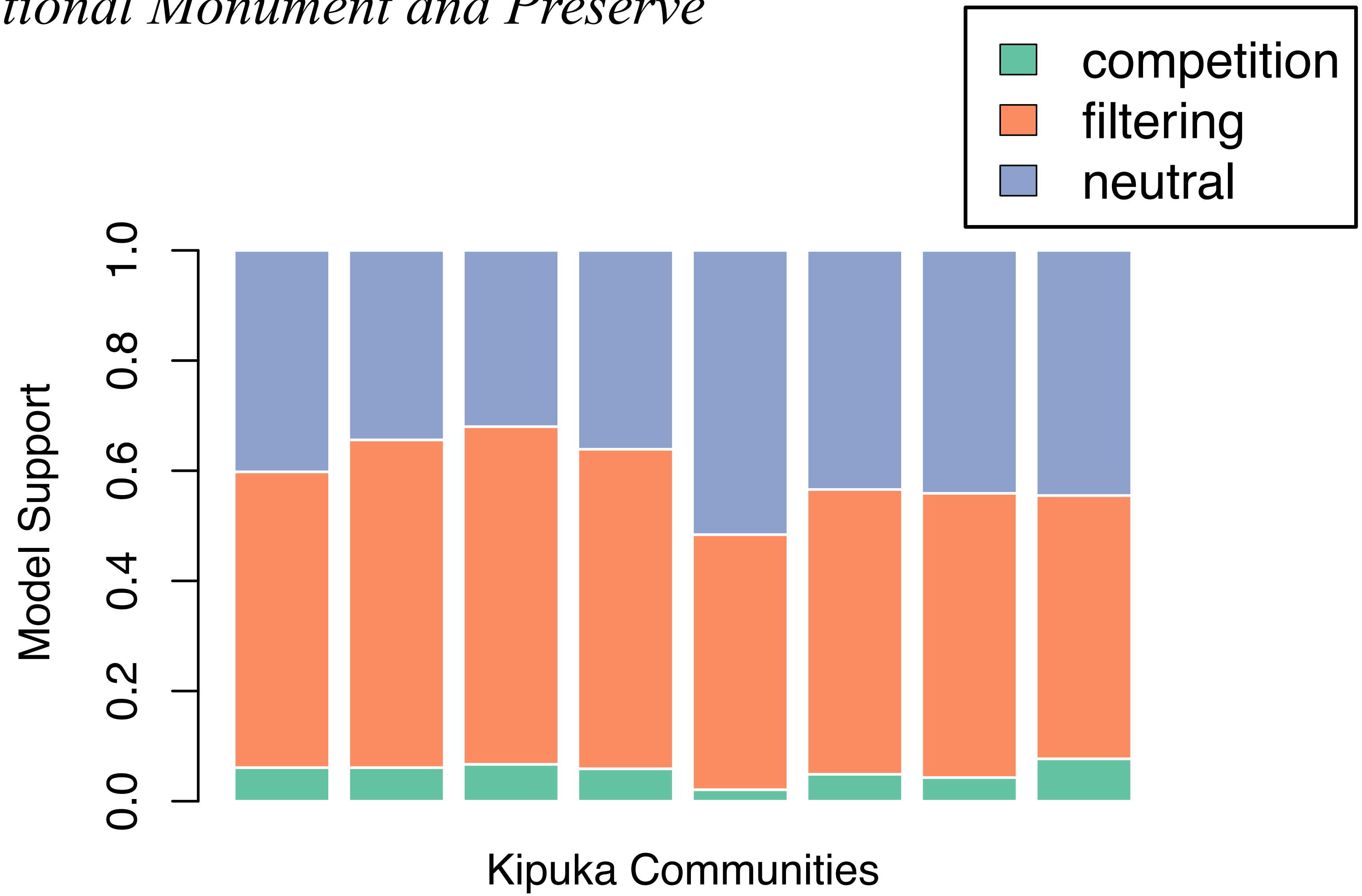


Maximum vegetative height

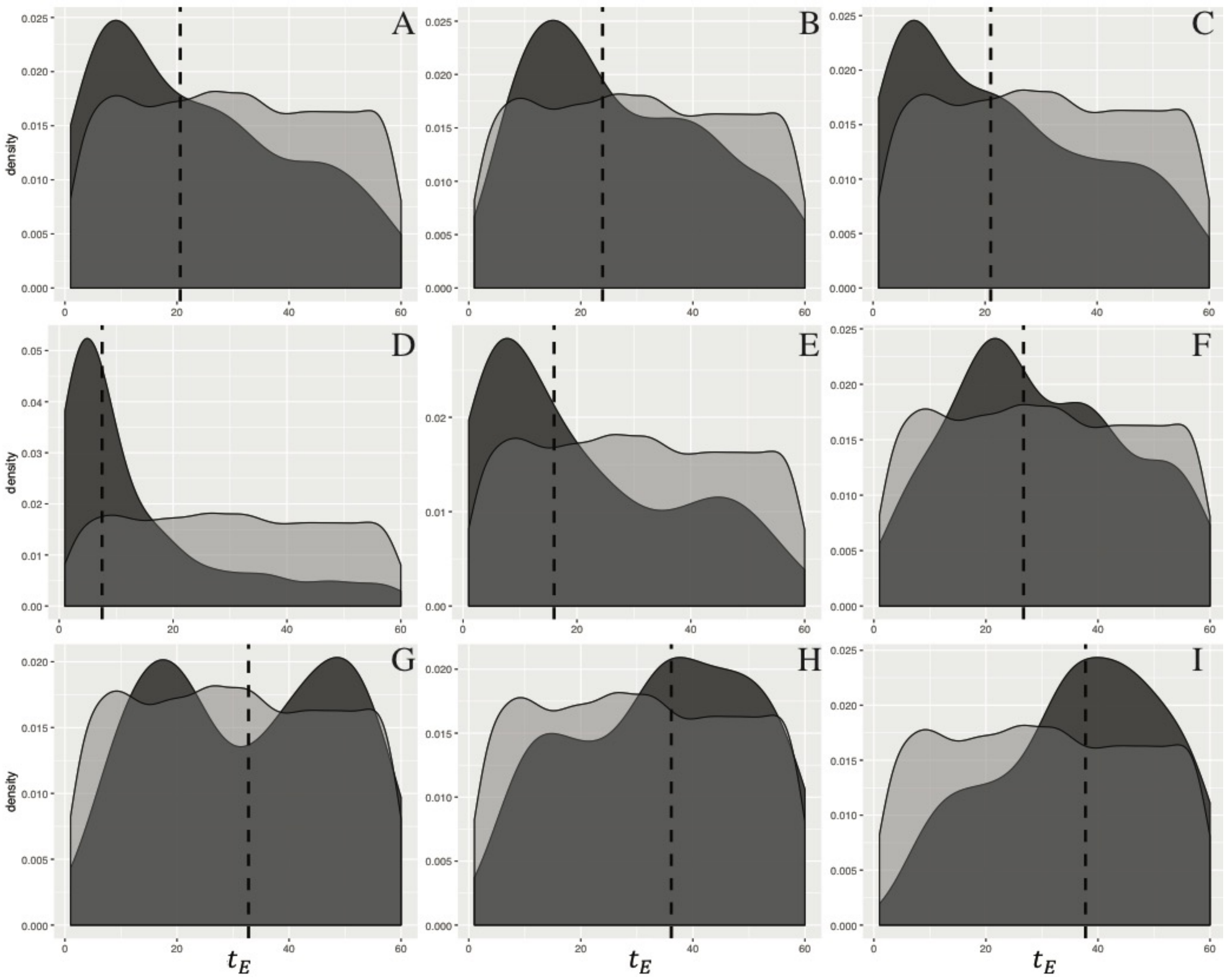


Ruffley *et al.* 2019 *Ecol. & Evol. Accepted*

Results: Lava-flow Islands in Craters of the Moon National Monument and Preserve



Results: Lava-flow Islands in Craters of the Moon National Monument and Preserve



Ruffley *et al.* 2019 *Ecol. & Evol. Accepted*

CAMI tutorial 2: San Juan Islands

- compphylo.github.io/Oslo2019/CAMI_files/CAMI_2.html

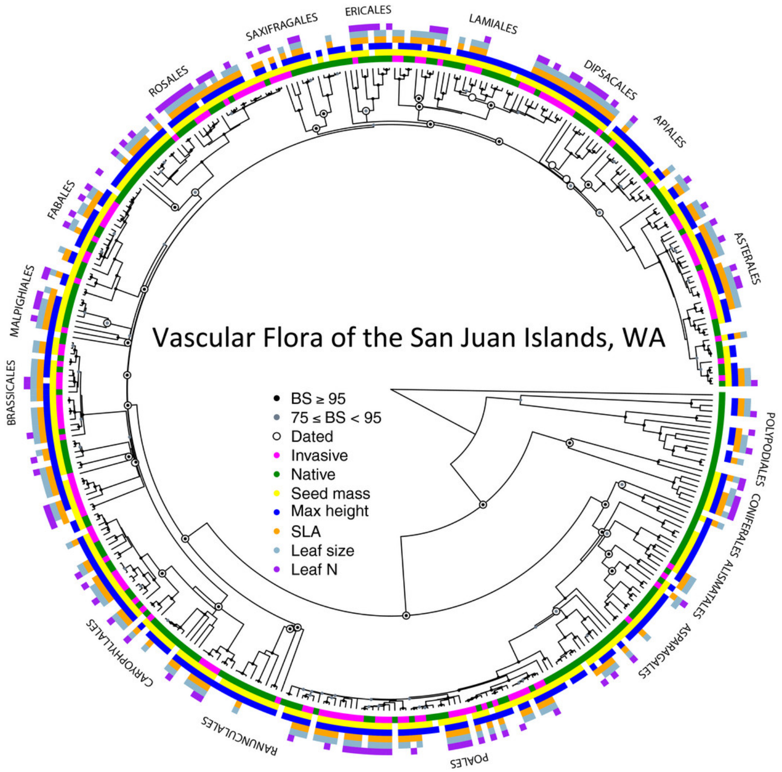
Results: San Juan Islands in the Pacific Northwest

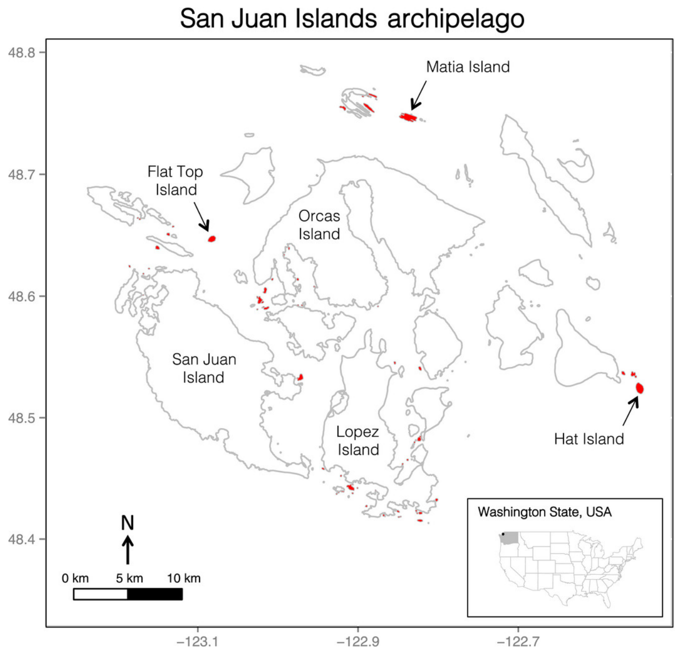
Diversity and Distributions, (Diversity Distrib.) (2016) **22**, 318–331

BIODIVERSITY RESEARCH

Deconstructing Darwin's Naturalization Conundrum in the San Juan Islands using community phylogenetics and functional traits

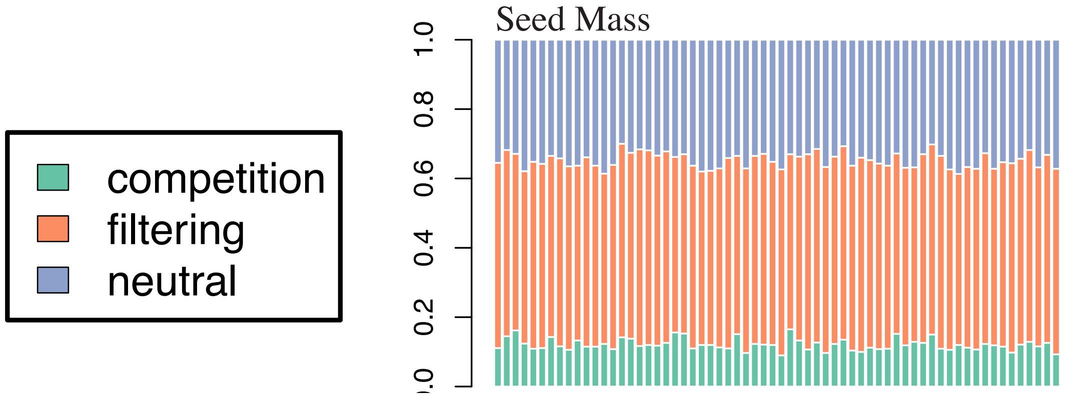
Hannah E. Marx[1]*, David E. Giblin[2], Peter W. Dunwiddie[2] and David C. Tank[1]



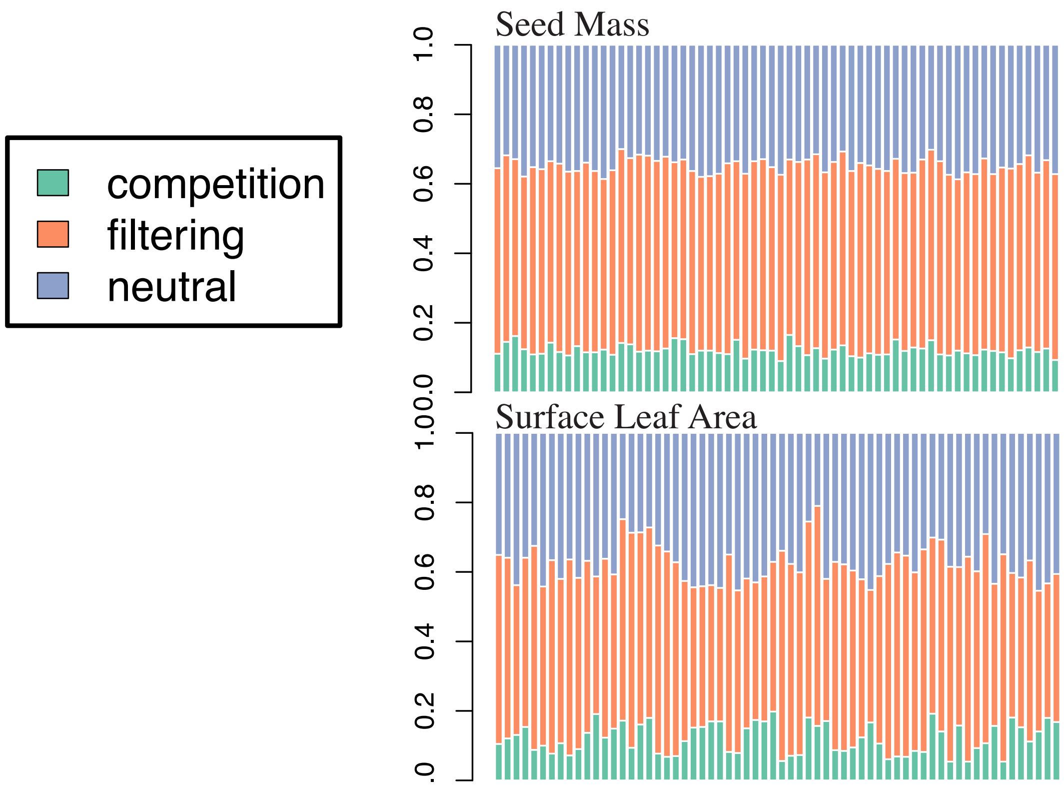


- 322 regional species
- 50 island communities
- 25-150 species in each
- 3 traits

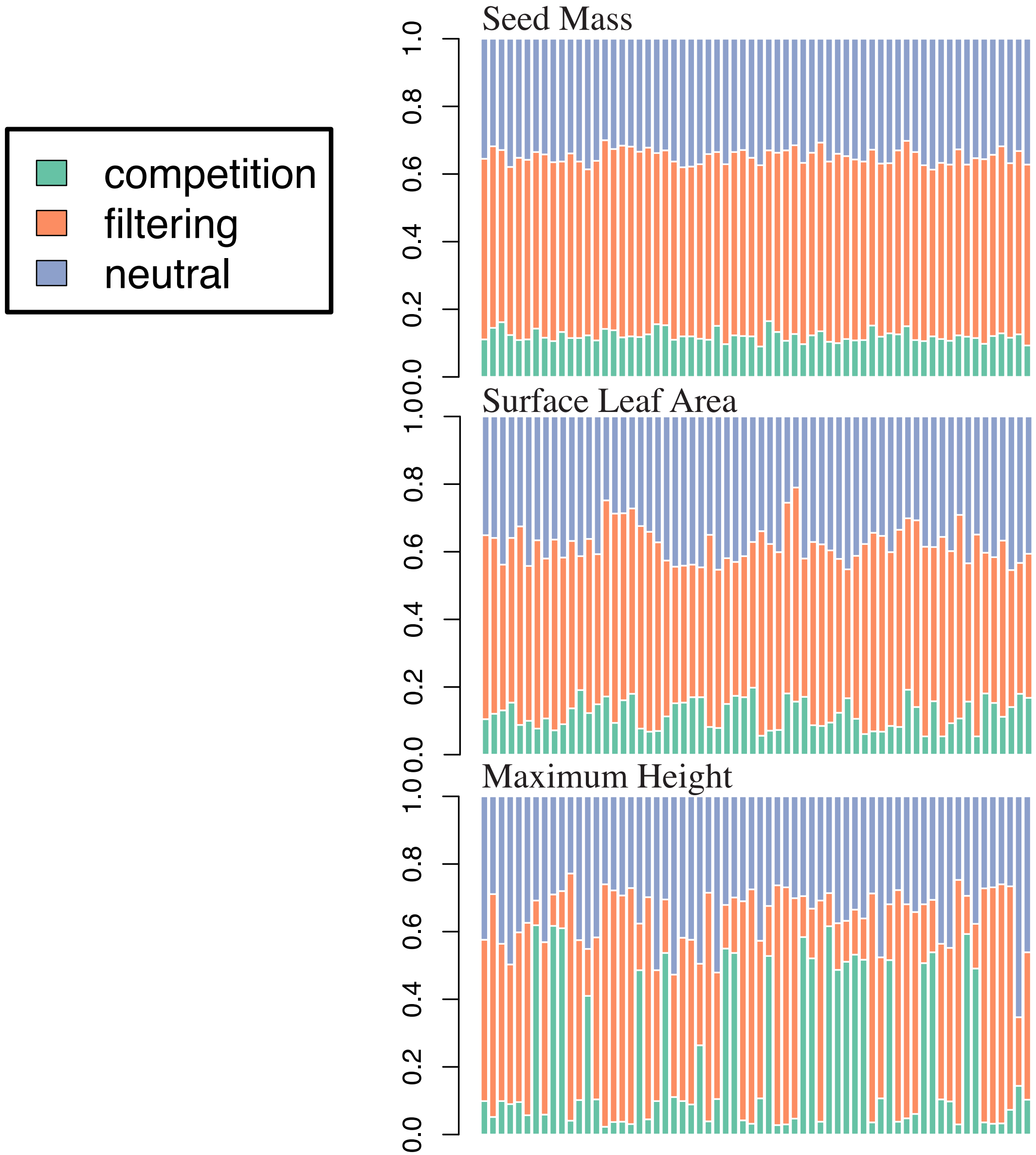
Results: San Juan Islands in the Pacific Northwest



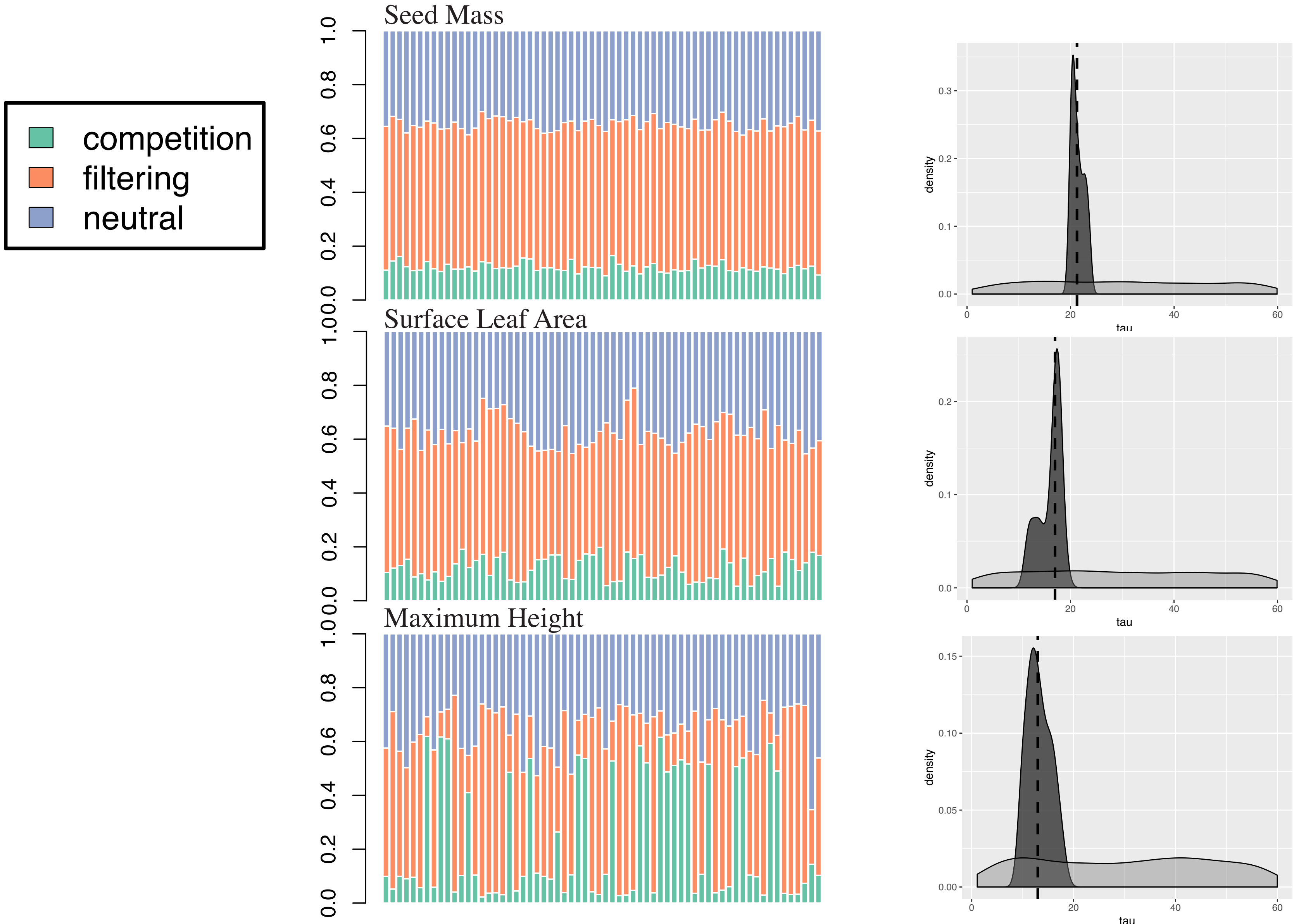
Results: San Juan Islands in the Pacific Northwest



Results: San Juan Islands in the Pacific Northwest



Results: San Juan Islands in the Pacific Northwest



CAMI Conclusions

- Simulate phylogenetic and phenotypic community assembly data under neutral and non-neutral models of assembly

- Implements a model-based inference procedure with parameter estimation

CAMI: Community Assembly Model Inference

R package and docs available now @ https://github.com/ruffleymr/CAMI

CAMI: Community Assembly Model Inference

R package and docs available now @

www.github.com/ruffleymr/CAMI

Received: 2 January 2017 | Accepted: 21 July 2017
DOI: 10.1111/2041-210X.12918

RESEARCH ARTICLE

Methods in Ecology and Evolution BRITISH ECOLOGICAL SOCIETY

ecolottery: Simulating and assessing community assembly with environmental filtering and neutral dynamics in R

François Munoz[1] | Matthias Grenié[2] | Pierre Denelle[2] | Adrien Taudière[2] | Fabien Laroche[2,3] | Caroline Tucker[2,4] | Cyrille Violle[2]

APPLICATION

VirtualCom: a simulation model for eco-evolutionary community assembly and invasion

Tamara Münkemüller[1]* and Laure Gallien[1,2,3]

ECOLOGY
ECOLOGICAL SOCIETY OF AMERICA

Article | Full Access

A new modeling approach estimates the relative importance of different community assembly processes

Fons van der Plas, Thijs Janzen, Alejandro Ordonez, Wimke Fokkema ... See all authors

REVIEW

Methods in Ec

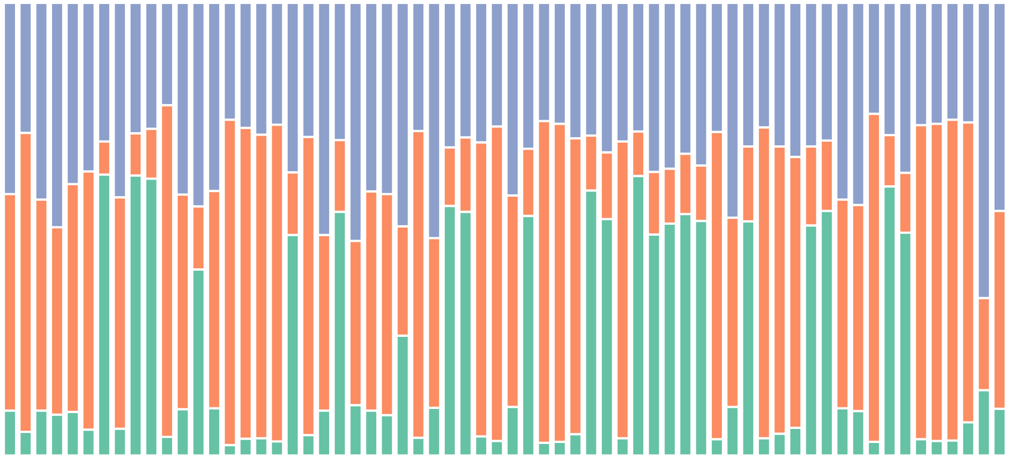
Inferring community assembly processes from macroscopic patterns using dynamic eco-evolutionary models and Approximate Bayesian Computation (ABC)

Mikael Pontarp[1,2,3] | Åke Brännström[4,5] | Owen L. Petchey[2]

Identifying Models of Trait-Mediated Community Assembly

using random forests and approximate Bayesian computation

Megan Ruffley, Katie Peterson,
Bob Week, David C. Tank,
Luke J. Harmon



Society of Systematic Biologists

University of Idaho

SSE
SOCIETY for the STUDY of EVOLUTION

iDiv
German Centre for Integrative Biodiversity Research

NSF

SANTA FE INSTITUTE

THE OHIO STATE UNIVERSITY

BOTANICAL SOCIETY OF AMERICA

FOR BIO

the City College of New York

UiO : Life Science
University of Oslo

ibest
Institure for Bioinformatics and Evolutionary Studies

ID
University of Idaho Stillinger Herbarium