

Carlos J. Melián,
National Center for Ecological Analysis and Synthesis, University of California, Santa Barbara.

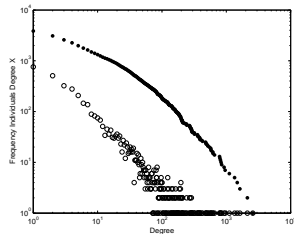
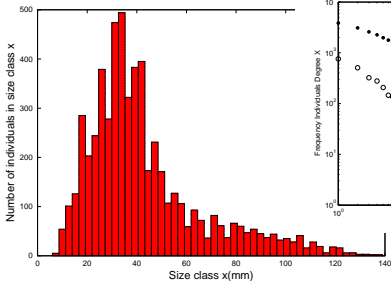
Towards a General Framework in Food Webs Driven by Data at Multiple Biological Levels

The Sven Lovén Centre for Marine Sciences, University of Gothenburg, Strömstad, Sweden,
June 5, 2009

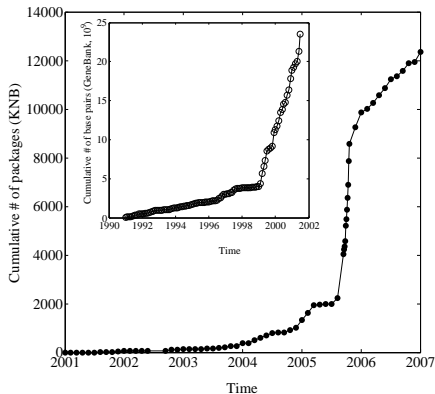
Thanks!

- Computing-scientist staff at NCEAS
- Microsoft Research Ltd., Cambridge, UK.
- Drew Allen, Jennifer Dunne, Jonathan Davies, Stanley Harpole, Stephen Hubbell, Pablo Marquet, Brad McRae, Mark Urban, and Tommaso Zillio,

Food webs as a focus for unifying ecological theory



Data accumulation in Networks of data at multiple levels (GeneBank, KNB)



Food webs as a focus for unifying ecological-evolutionary theory

- 1 Models with explicit DNA evolution and speciation (Kimura:1968,Higgs & Derrida:1992,Gavrilets:2004)

Food webs as a focus for unifying ecological-evolutionary theory

- ① Models with explicit DNA evolution and speciation (Kimura:1968,Higgs & Derrida:1992,Gavrilets:2004)
- ② Aggregation: Link raw data with dynamical models of interacting individuals and populations

1. Models with explicit DNA evolution and speciation

- Why are there so many species?

1. Models with explicit DNA evolution and speciation

- Why are there so many species?
- Ecological views have focused on the mechanisms that enable or constraint species coexistence (Hutchinson:1959)

1. Models with explicit DNA evolution and speciation

- Why are there so many species?
- Ecological views have focused on the mechanisms that enable or constraint species coexistence (Hutchinson:1959)
- Why are there so few kinds of animals?

1. Models with explicit DNA evolution and speciation

- Why are there so many species?
- Ecological views have focused on the mechanisms that enable or constraint species coexistence (Hutchinson:1959)
- Why are there so few kinds of animals?
- Additional constraints on the process of speciation, constraints set by the genetics rather than ecology (Felsenstein:1981)

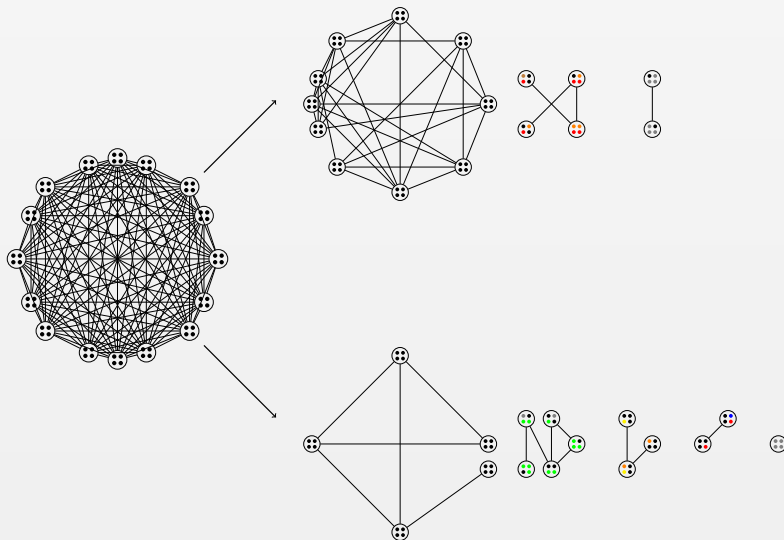
Two broad categories of speciation:

- Ecological speciation: Reproductive isolation (RI) evolves between populations by divergent (positive, neutral, negative) selection arising from differences between ecological environments and ecological processes (Schluter:2009).

Two broad categories of speciation:

- Ecological speciation: Reproductive isolation (RI) evolves between populations by divergent (positive, neutral, negative) selection arising from differences between ecological environments and ecological processes (Schluter:2009).
- Genetic speciation: Evolution of RI by the fixation of different types of mutations experiencing similar environments (Lynch:2007).

1. Models with explicit DNA evolution and speciation



2. Aggregation



2. Aggregation



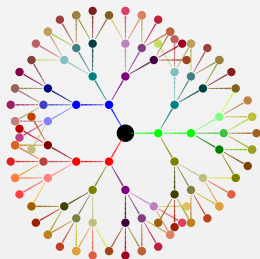
2. Aggregation



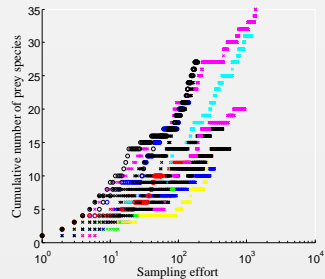
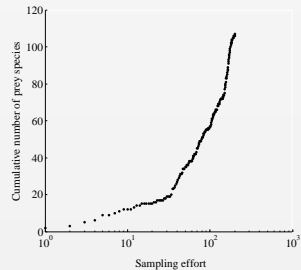
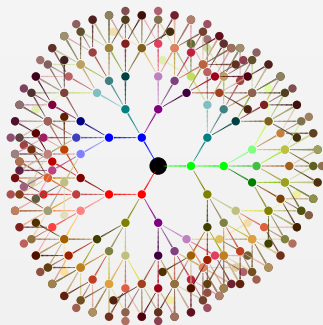
2. Aggregation

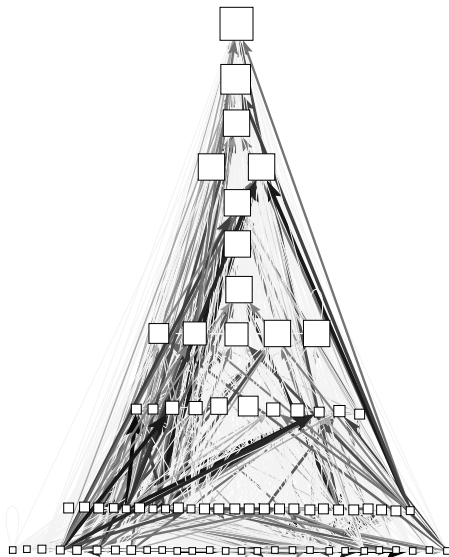


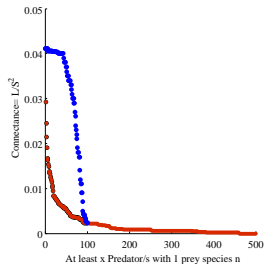
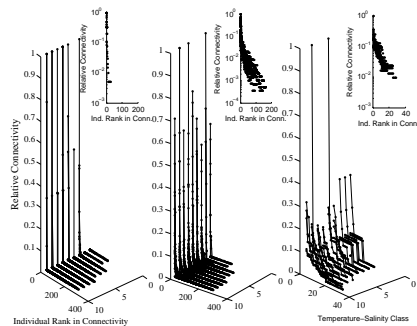
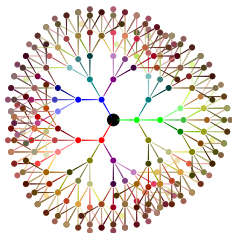
2. Aggregation



2. Aggregation







Summary

- Networks of Data (at individual level) from ONE (type, location, level)...

Summary

- Networks of Data (at individual level) from **ONE** (type, location, level)...
- to **MULTIPLE** (types, locations, levels).

Summary

- Networks of Data (at individual level) from **ONE** (type, location, level)...
- to **MULTIPLE** (types, locations, levels).
- From **ONE** specific question...

Summary

- Networks of Data (at individual level) from **ONE** (type, location, level)...
- to **MULTIPLE** (types, locations, levels).
- From **ONE** specific question...
- to **SEVERAL** questions

Summary

- Networks of Data (at individual level) from **ONE** (type, location, level)...
- to **MULTIPLE** (types, locations, levels).
- From **ONE** specific question...
- to **SEVERAL** questions
- From models with **ONE** Output

Summary

- Networks of Data (at individual level) from **ONE** (type, location, level)...
- to **MULTIPLE** (types, locations, levels).
- From **ONE** specific question...
- to **SEVERAL** questions
- From models with **ONE** Output
- to models with **SEVERAL** Outputs

Summary

- Networks of Data (at individual level) from **ONE** (type, location, level)...
- to **MULTIPLE** (types, locations, levels).
- From **ONE** specific question...
- to **SEVERAL** questions
- From models with **ONE** Output
- to models with **SEVERAL** Outputs
- From methods to select **one model** according to **ONE** specific output

Summary

- Networks of Data (at individual level) from **ONE** (type, location, level)...
- to **MULTIPLE** (types, locations, levels).
- From **ONE** specific question...
- to **SEVERAL** questions
- From models with **ONE** Output
- to models with **SEVERAL** Outputs
- From methods to select **one model** according to **ONE** specific output
- to methods to select **models** according to **SEVERAL** outputs

Questions

- What is the minimum information we need to model diversification and coexistence at molecular and ecological levels?
It is possible to estimate a maximum likelihood for the general structure of the individual interactions in such a framework?

Questions

- What is the minimum information we need to model diversification and coexistence at molecular and ecological levels?
It is possible to estimate a maximum likelihood for the general structure of the individual interactions in such a framework?
- Does (neutral, positive, negative) selection and evolution at molecular and ecological levels predict individual rank in connectance and species rank in abundance?

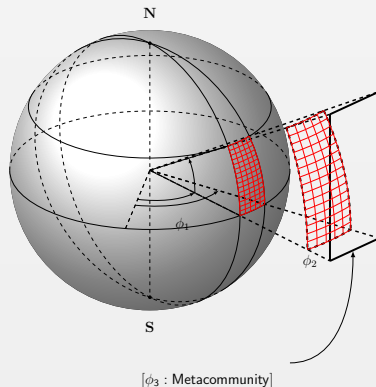
Questions

- What is the minimum information we need to model diversification and coexistence at molecular and ecological levels?
It is possible to estimate a maximum likelihood for the general structure of the individual interactions in such a framework?
- Does (neutral, positive, negative) selection and evolution at molecular and ecological levels predict individual rank in connectance and species rank in abundance?
- Dream question: Does (neutral, positive, negative) selection predict size-spectrum and species level food web properties?

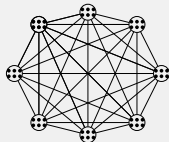
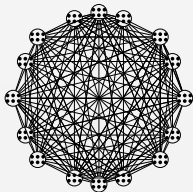
Sampling individual diets, size and abundance at several s-t conditions

30520 inds after 144 samplings/year, Fish (5725/54); Shrimps (365/258); Mysids (18393)

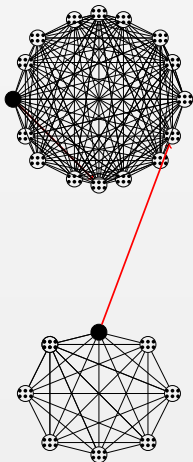
688×10^6 inds/ $10^5 m^3$, Fish 9×10^6 ; Shrimps 4556; Mysids 679×10^6



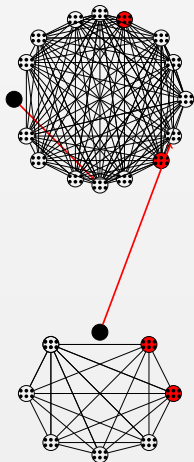
DNA evolution model with explicit speciation: steps



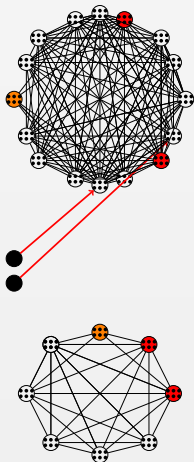
DNA evolution model with explicit speciation: steps



DNA evolution model with explicit speciation: steps



DNA evolution model with explicit speciation: steps



DNA evolution model with explicit speciation: formal steps

- Genetic similarity; $q^{ij} = \frac{1}{L} \sum_{u=1}^L S_u^i S_u^j$, thus $Q = [q^{ij}]$, where $q^{ij} \in [-1, 1]$ and $S_1^i, S_2^i, \dots, S_L^i$ is the genome of individual i :

DNA evolution model with explicit speciation: formal steps

- Genetic similarity; $q^{ij} = \frac{1}{L} \sum_{u=1}^L S_u^i S_u^j$, thus $Q = [q^{ij}]$, where $q^{ij} \in [-1, 1]$ and $S_1^i, S_2^i, \dots, S_L^i$ is the genome of individual i :
- q^{ij} as the fraction of identical sites (f^{ij}), thus $q^{ij} = \frac{1}{L} [L f^{ij} - L(1 - f^{ij})] = 2 f^{ij} - 1$, and $f^{ij} = \frac{1+q^{ij}}{2}$

DNA evolution model with explicit speciation: formal steps

- Genetic similarity; $q^{ij} = \frac{1}{L} \sum_{u=1}^L S_u^i S_u^j$, thus $Q = [q^{ij}]$, where $q^{ij} \in [-1, 1]$ and $S_1^i, S_2^i, \dots, S_L^i$ is the genome of individual i :
- q^{ij} as the fraction of identical sites (f^{ij}), thus $q^{ij} = \frac{1}{L} [L f^{ij} - L(1 - f^{ij})] = 2 f^{ij} - 1$, and $f^{ij} = \frac{1+q^{ij}}{2}$
- Which is the expected frequency of nucleotides that a new offspring k will share with each individual j in the population ($E[f^{kj}]$)?

DNA evolution model with explicit speciation: formal steps

- Genetic similarity; $q^{ij} = \frac{1}{L} \sum_{u=1}^L S_u^i S_u^j$, thus $Q = [q^{ij}]$, where $q^{ij} \in [-1, 1]$ and $S_1^i, S_2^i, \dots, S_L^i$ is the genome of individual i :
- q^{ij} as the fraction of identical sites (f^{ij}), thus $q^{ij} = \frac{1}{L} [L f^{ij} - L(1 - f^{ij})] = 2 f^{ij} - 1$, and $f^{ij} = \frac{1+q^{ij}}{2}$
- Which is the expected frequency of nucleotides that a new offspring k will share with each individual j in the population ($E[f^{kj}]$)?
- $q^{ki} = \frac{e^{-2\mu}}{2} (q^{G_1(k)i} + q^{G_2(k)i})$ and $q^{kk} = 1$.

DNA evolution model with explicit speciation: formal steps

- Genetic similarity; $q^{ij} = \frac{1}{L} \sum_{u=1}^L S_u^i S_u^j$, thus $Q = [q^{ij}]$, where $q^{ij} \in [-1, 1]$ and $S_1^i, S_2^i, \dots, S_L^i$ is the genome of individual i :
- q^{ij} as the fraction of identical sites (f^{ij}), thus $q^{ij} = \frac{1}{L} [L f^{ij} - L(1 - f^{ij})] = 2 f^{ij} - 1$, and $f^{ij} = \frac{1+q^{ij}}{2}$
- Which is the expected frequency of nucleotides that a new offspring k will share with each individual j in the population ($E[f^{kj}]$)?
- $q^{ki} = \frac{e^{-2\mu}}{2} (q^{G_1(k)i} + q^{G_2(k)i})$ and $q^{kk} = 1$.
- Matrix Q approaches a mean value $Q^* = \frac{1}{4J\mu}$

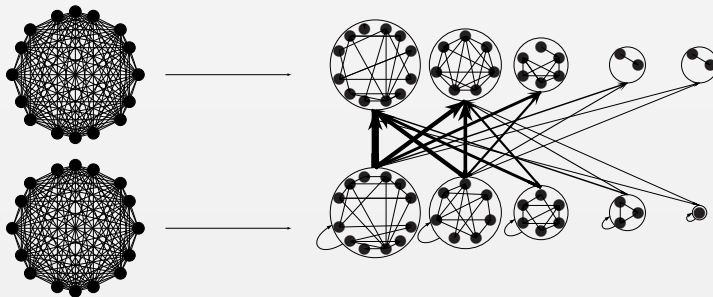
DNA evolution model with explicit speciation: formal steps

- Genetic similarity; $q^{ij} = \frac{1}{L} \sum_{u=1}^L S_u^i S_u^j$, thus $Q = [q^{ij}]$, where $q^{ij} \in [-1, 1]$ and $S_1^i, S_2^i, \dots, S_L^i$ is the genome of individual i :
- q^{ij} as the fraction of identical sites (f^{ij}), thus $q^{ij} = \frac{1}{L} [L f^{ij} - L(1 - f^{ij})] = 2 f^{ij} - 1$, and $f^{ij} = \frac{1+q^{ij}}{2}$
- Which is the expected frequency of nucleotides that a new offspring k will share with each individual j in the population ($E[f^{kj}]$)?
- $q^{ki} = \frac{e^{-2\mu}}{2} (q^{G_1(k)i} + q^{G_2(k)i})$ and $q^{kk} = 1$.
- Matrix Q approaches a mean value $Q^* = \frac{1}{4J\mu}$
- If $q^{ij} < q^{\min}$, then reproductive isolation occurs. Note that q^{\min} implicitly captures the ecological (i.e., prezygotic) and genetic (i.e., postzygotic) mechanisms of speciation.

DNA evolution model with explicit speciation: formal steps

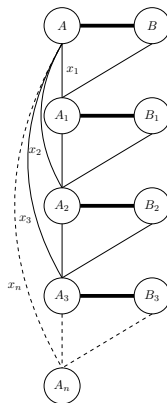
- Genetic similarity; $q^{ij} = \frac{1}{L} \sum_{u=1}^L S_u^i S_u^j$, thus $Q = [q^{ij}]$, where $q^{ij} \in [-1, 1]$ and $S_1^i, S_2^i, \dots, S_L^i$ is the genome of individual i :
- q^{ij} as the fraction of identical sites (f^{ij}), thus $q^{ij} = \frac{1}{L} [L f^{ij} - L(1 - f^{ij})] = 2 f^{ij} - 1$, and $f^{ij} = \frac{1+q^{ij}}{2}$
- Which is the expected frequency of nucleotides that a new offspring k will share with each individual j in the population ($E[f^{kj}]$)?
- $q^{ki} = \frac{e^{-2\mu}}{2} (q^{G_1(k)i} + q^{G_2(k)i})$ and $q^{kk} = 1$.
- Matrix Q approaches a mean value $Q^* = \frac{1}{4J\mu}$
- If $q^{ij} < q^{\min}$, then reproductive isolation occurs. Note that q^{\min} implicitly captures the ecological (i.e., prezygotic) and genetic (i.e., postzygotic) mechanisms of speciation.
- If $q^{\min} > Q^*$ then speciation happens (Higgs & Derrida 1992)

Nucleotide heterozygosity, Speciation rate and Relative species abundance



Can we predict speciation rate in the neutral case?

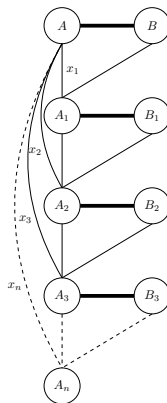
What is the number of steps (x_n) at which $q^{min} > q^{AA_n}$?



$$\bullet \quad E[q^{kj}] = \frac{e^{-2\mu}}{2} (q^{G_1(k)j} + q^{G_2(k)j}),$$

Can we predict speciation rate in the neutral case?

What is the number of steps (x_n) at which $q^{min} > q^{AA_n}$?

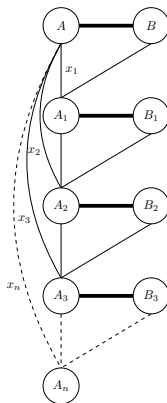


$$\bullet E[q^{kj}] = \frac{e^{-2\mu}}{2} (q^{G_1(k)j} + q^{G_2(k)j}),$$

$$\bullet q^{AA_1} = x_1 = e^{-2\mu} \frac{(q^{AA} + q^{AB})}{2},$$

Can we predict speciation rate in the neutral case?

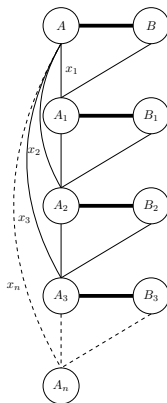
What is the number of steps (x_n) at which $q^{min} > q^{AA_n}$?



- $E[q^{kj}] = \frac{e^{-2\mu}}{2} (q^{G_1(k)j} + q^{G_2(k)j}),$
- $q^{AA_1} = x_1 = e^{-2\mu} \frac{(q^{AA} + q^{AB})}{2},$
- $q^{AA_2} = x_2 = e^{-2\mu} \left(\frac{(1+E[AB])}{2} \right)^2,$
where $E[AB] = X = (1 + q^{min})/2$

Can we predict speciation rate in the neutral case?

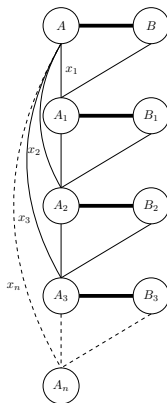
What is the number of steps (x_n) at which $q^{\min} > q^{AA_n}$?



- $E[q^{kj}] = \frac{e^{-2\mu}}{2} (q^{G_1(k)j} + q^{G_2(k)j}),$
- $q^{AA_1} = x_1 = e^{-2\mu} \frac{(q^{AA} + q^{AB})}{2},$
- $q^{AA_2} = x_2 = e^{-2\mu} \left(\frac{(1+E[AB])}{2} \right)^2,$
where $E[AB] = X = (1 + q^{\min})/2$
- $q^{\min} > e^{-2\mu n} \left(\frac{(1+X)}{2} \right)^n,$

Can we predict speciation rate in the neutral case?

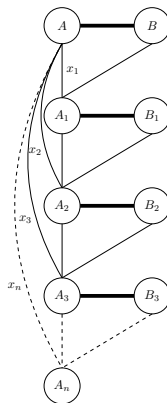
What is the number of steps (x_n) at which $q^{\min} > q^{AA_n}$?



- $E[q^{kj}] = \frac{e^{-2\mu}}{2} (q^{G_1(k)j} + q^{G_2(k)j}),$
- $q^{AA_1} = x_1 = e^{-2\mu} \frac{(q^{AA} + q^{AB})}{2},$
- $q^{AA_2} = x_2 = e^{-2\mu 2} \left(\frac{(1+E[AB])}{2} \right)^2,$
where $E[AB] = X = (1 + q^{\min})/2$
- $q^{\min} > e^{-2\mu n} \left(\frac{(1+X)}{2} \right)^n,$
- and applying logarithms result in:
$$n = \frac{\log(q^{\min})}{-2\mu + \log(\frac{1+X}{2})},$$

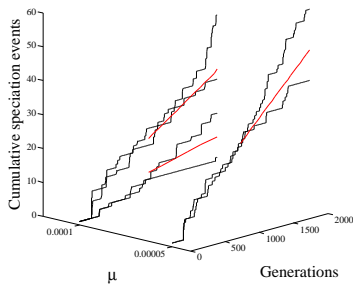
Can we predict speciation rate in the neutral case?

What is the number of steps (x_n) at which $q^{\min} > q^{AA_n}$?



- $E[q^{kj}] = \frac{e^{-2\mu}}{2} (q^{G_1(k)j} + q^{G_2(k)j}),$
- $q^{AA_1} = x_1 = e^{-2\mu} \frac{(q^{AA} + q^{AB})}{2},$
- $q^{AA_2} = x_2 = e^{-2\mu 2} \left(\frac{(1+E[AB])}{2} \right)^2,$
where $E[AB] = X = (1 + q^{\min})/2$
- $q^{\min} > e^{-2\mu n} \left(\frac{(1+X)}{2} \right)^n,$
- and applying logarithms result in:
$$n = \frac{\log(q^{\min})}{-2\mu + \log(\frac{1+X}{2})},$$
- and the expected speciation rate is:
$$\nu = \frac{1}{n} = \frac{-2\mu + \log(\frac{1+X}{2})}{\log(q^{\min})}.$$

This expression gives an accurate prediction of the speciation rate...



Can we predict the expected relative species abundance at steady state?

- $\nu = \phi(1)N(0, 1)$

Can we predict the expected relative species abundance at steady state?

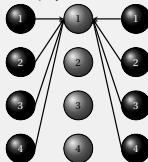
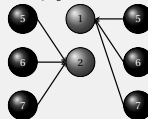
- $\nu = \phi(1)N(0, 1)$
- $\phi(1) N(2, 1) + \phi(1) N(0, 1) = \sum_{i=2}^J \phi(i) N(1, i)$

Can we predict the expected relative species abundance at steady state?

- $\nu = \phi(1)N(0,1)$
- $\phi(1) N(2,1) + \phi(1) N(0,1) = \sum_{i=2}^J \phi(i) N(1,i)$
- .
- .
- .
- .
- .
- $\phi(j) N(j+1,j) + \phi(j) N(j-1,j) = \sum_{i=j+1}^J \phi(i) N(j,i)$

Maximum likelihood

	a) $A_1, \text{Predator}$		b) $A_2, \text{Predator}$
A_1, Prey	$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 1 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 3 & 1 & 0 & 0 \\ 4 & 1 & 0 & 0 \end{pmatrix}$	A_1, Prey	$\begin{pmatrix} 1 & 2 \\ 5 & 1 \\ 6 & 1 \\ 7 & 1 \end{pmatrix}$
A_2, Prey	$\begin{pmatrix} 1 & 1 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 3 & 1 & 0 & 0 \\ 4 & 1 & 0 & 0 \end{pmatrix}$	A_2, Prey	$\begin{pmatrix} 5 & 0 & 1 \\ 6 & 0 & 1 \\ 7 & 0 & 1 \end{pmatrix}$

c) $A_1 \text{ Prey } A_1 \text{ Predator } A_2 \text{ Prey}$ d) $A_1 \text{ Prey } A_2 \text{ Predator } A_2 \text{ Prey}$ 

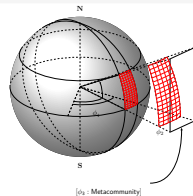
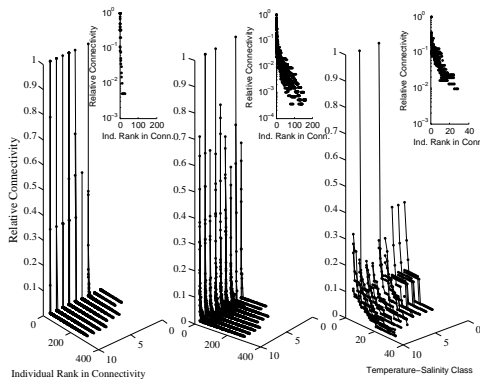
$$P(N_{ind}(L)|p_{12}, p_{21}, p_{11}, p_{22}) =$$

$$\begin{aligned} & p_{12}^{L_{12}} (1 - p_{12})^{N_1^{prey} - L_{12}} \\ & p_{21}^{L_{21}} (1 - p_{21})^{N_2^{prey} - L_{21}} \\ & p_{11}^{L_{11}} (1 - p_{11})^{N_1^{prey} - L_{11}} \\ & p_{22}^{L_{22}} (1 - p_{22})^{N_2^{prey} - L_{22}}, \quad (1) \end{aligned}$$

$$\ell(\vec{p} | N_{ind}(L)) = \prod_{i=1}^S \prod_{j=1}^S p_{ij}^{L_{ij}} (1 - p_{ij})^{N_i^{prey} - L_{ij}},$$

$$\ell(\mu_R, \mu_P, q_R^{min}, q_P^{min} | N_{ind}(L)) = \prod_{i=1}^S \prod_{j=1}^S p_{ij}^{L_{ij}} (1 - p_{ij})^{N_i^{prey} - L_{ij}}. \quad (2)$$

Individual rank in connectance



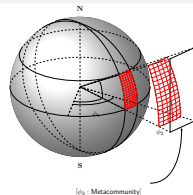
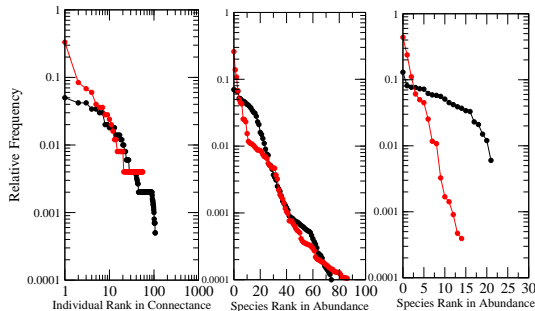
Neutral predictions at individual and species levels

(Left) $J_P = \text{Fish (5725)}$; $J_R = \text{Shrimps(258)}$;

(Center) (Fish) $J_{PM} = 2.4 \times 10^6 \text{ inds}/10^5 m^3$; $\mu_P = 10^{-7}$; $q_P^{\min} = 0.57$;

(Right) (Shrimps) $J_{RM} = 4556 \text{ inds}/10^5 m^3$; $\mu_R = 2.5 \times 10^{-4}$; $q_R^{\min} = 0.45$;

$\log \ell(\vec{p}) = -1273.47$



Conclusions

- 1 A sampling theory of food webs in space and time to test genetic and ecological drift with the observed individual rank in connectance and species rank in abundance.

Conclusions

- 1 A sampling theory of food webs in space and time to test genetic and ecological drift with the observed individual rank in connectance and species rank in abundance.
- 2 Those two distributions depart from neutral expectations in the Guadalquivir estuary food web in all the s-t conditions (sampling of 30520 inds. and an estimated abundance of 688×10^6 inds/ $10^5 m^3$).

General Conclusions

- 1 Theoretical framework to understand diversification and coexistence at multiple biological levels.

General Conclusions

- 1 Theoretical framework to understand diversification and coexistence at multiple biological levels.
- 2 Birth-death process with neutral genome evolution (genetic and ecological drift) and a molecular filter that regulates the production of fertile offspring allow approximations at steady state that can be tested using thousands of individuals simultaneously.

General Conclusions

- 1 Theoretical framework to understand diversification and coexistence at multiple biological levels.
- 2 Birth-death process with neutral genome evolution (genetic and ecological drift) and a molecular filter that regulates the production of fertile offspring allow approximations at steady state that can be tested using thousands of individuals simultaneously.
- 3 Future extensions include MLE to test the expected individual interactions according to their size and the abundance of each species.