Generalism drives abundance:
a computational causal discovery approach

Chulian Song, Marie-Josée Fortin, Andrew Gonzalez

Department of Biology, McGill University
Department of Ecology and Evolutionary Biology, University of Toronto

CSEE-SCEE Virtual Conference, 2021
More abundant species are also more generalized

**Generalism**

- **Generalist**

- **Specialist**
A chicken-and-egg dilemma: Generalism drives abundance, or abundance drives generalism

**Generalism**

- **Generalist**
  - Selection process: Generalist species have competitive advantage
- **Specialist**
  - Drift process: Abundant species are generalist by chance

**Abundance**

- **Abundant**
- **Rare**
Computational causal discovery approach

Controlled experiments

- Initial condition
- Fixed generalism
- 5 plants
- 2 plants

Causal discovery

- A causes B
- No Causal Relation
- B causes A
Causal discovery with formal logic
Example: Detecting causal direction between dodo and extinct species
Causal discovery with formal logic
Example: Detecting causal direction between dodo and extinct species

Dodo $\rightarrow$ Extinct species
Dodo $\leftarrow$ Extinct species

<table>
<thead>
<tr>
<th>it is a dodo</th>
<th>it is extinct</th>
<th>Implication</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>T</td>
<td>It is a dodo and is extinct</td>
</tr>
<tr>
<td>F</td>
<td>F</td>
<td>It is not a dodo and is not extinct</td>
</tr>
<tr>
<td>T</td>
<td>F</td>
<td>It is a dodo and is not extinct</td>
</tr>
<tr>
<td>F</td>
<td>T</td>
<td>It is not a dodo and is extinct</td>
</tr>
</tbody>
</table>

Simmons et. al., *Oikos* (2019)
Causal discovery with formal logic
Detection of causal direction in abundance and generalism

<table>
<thead>
<tr>
<th>dodo/abundant</th>
<th>extinct/generalist</th>
<th>Abundant-Generalist</th>
<th>Rare-Specialist</th>
<th>Abundant-Specialist</th>
<th>Rare-Generalist</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>T</td>
<td>✓</td>
<td></td>
<td>If generalism drives abundance</td>
<td></td>
</tr>
<tr>
<td>F</td>
<td>F</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T</td>
<td>F</td>
<td></td>
<td></td>
<td>If abundance drives generalism</td>
<td></td>
</tr>
<tr>
<td>F</td>
<td>T</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Simmons et. al., *Oikos* (2019)
Formal logic on binary variables does not automatically apply to continuous variables

---

Abundance

Generalization

Normalization

0.00 0.25 0.50 0.75 1.00

0.25 0.50 0.75 1.00

Abundance

Generalization

Song et. al., submitted
Generalism drives abundance
(selection process is generally stronger than drift process)

Data bias was **corrected**

![Graph 1](Image 1)

- Rare-Generalist (Abundance drives generalism)
- Abundant-Specialist (Generalism drives abundance)
- Song et al., *submitted*

Data bias was **uncorrected**

![Graph 2](Image 2)

- Rare-Generalist (Abundance drives generalism)
- Abundant-Specialist (Generalism drives abundance)
Two other causal discovery methods confirm that generalism drives abundance.

Nonlinear additive noise model based on nonparametric regression:

\[ Y = f(X) + \epsilon(Y) \]

- Nonlinear function of cause
- Noise is independent of \( X \)

Geometric-information inference based on information theory:

\[ H(X) \geq H(Y) \]

- Entropy of \( X \)
- Entropy of \( Y \)

<table>
<thead>
<tr>
<th>Hummingbird</th>
<th>Plant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abundance</td>
<td>Generalism</td>
</tr>
<tr>
<td>0.001</td>
<td>0.740</td>
</tr>
<tr>
<td>Dependent with noise</td>
<td>Independent with noise</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Hummingbird</th>
<th>Plant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abundance</td>
<td>Generalism</td>
</tr>
<tr>
<td>0.001</td>
<td>0.608</td>
</tr>
<tr>
<td>Dependent with noise</td>
<td>Independent with noise</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Song et. al., submitted</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.77</td>
</tr>
<tr>
<td>2.36</td>
</tr>
<tr>
<td>0.41</td>
</tr>
<tr>
<td>4.58</td>
</tr>
</tbody>
</table>
The strength of selection processes increases when local temperatures are more variable

Song et al., J. Anim. Ecol. (2017); Song et. al., submitted
Take-home message

• Our computational approach allows us to use the relative strength of the causal directions as a proxy of the relative roles of either selection or drift process.

• In contrast to previous findings, all three causal discovery methods consistently found strong evidence that generalism drives abundance in pollinator-hummingbird communities and reef fish datasets.

• Selection processes act more strongly than drift processes when local temperatures are more variable. This generalizes previous known results in two-species communities to multispecies communities.
Thanks!

Marie-Josée Fortin  
University of Toronto  

Andrew Gonzalez  
McGill University  

@Chuliang_Song  
clsong.com  
clsong.ecology@gmail.com