Looking locally to see globally:

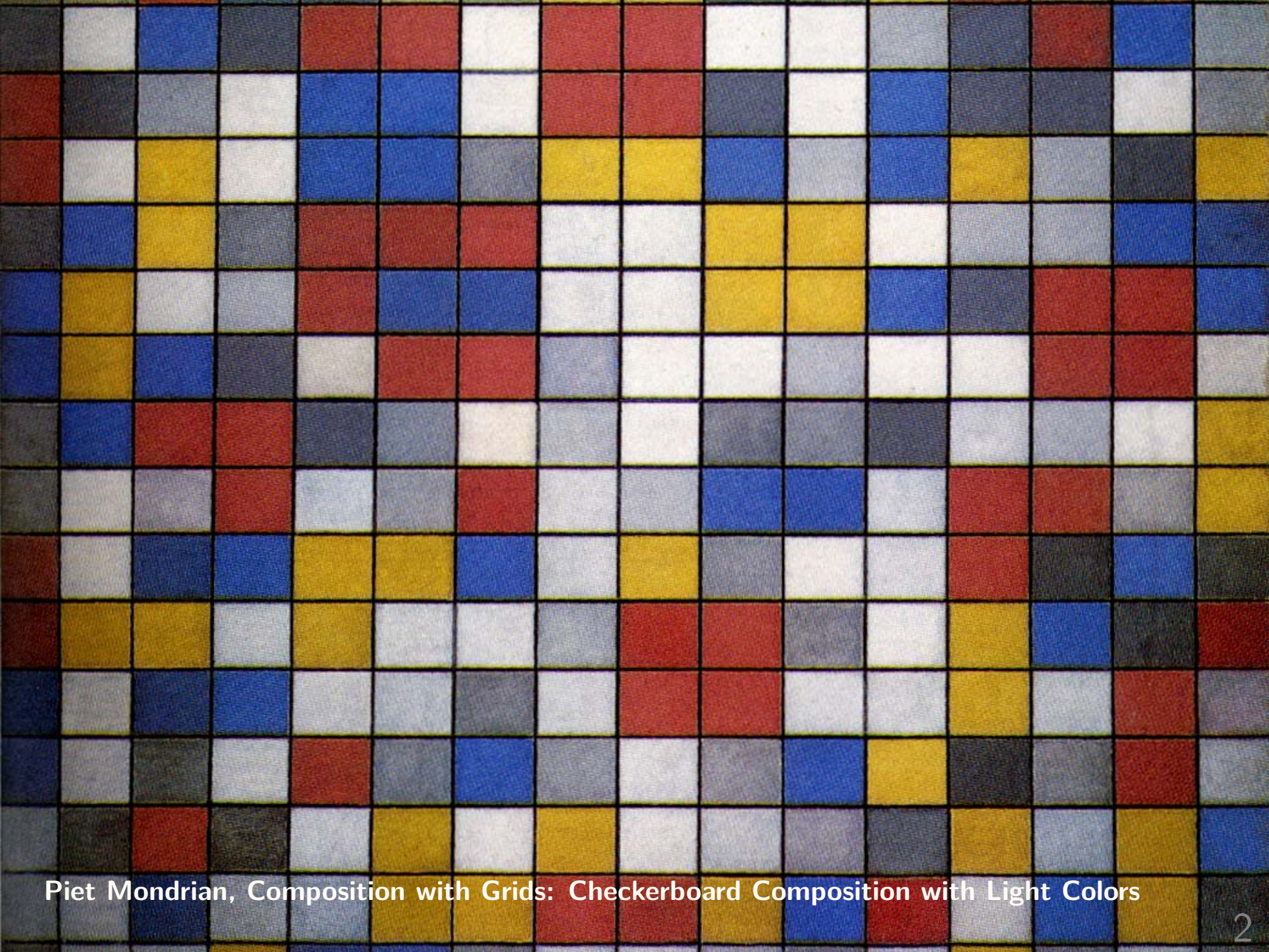
motif sampling to distinguish interaction type and predict dynamical properties of whole networks

Matthew J. Michalska-Smith

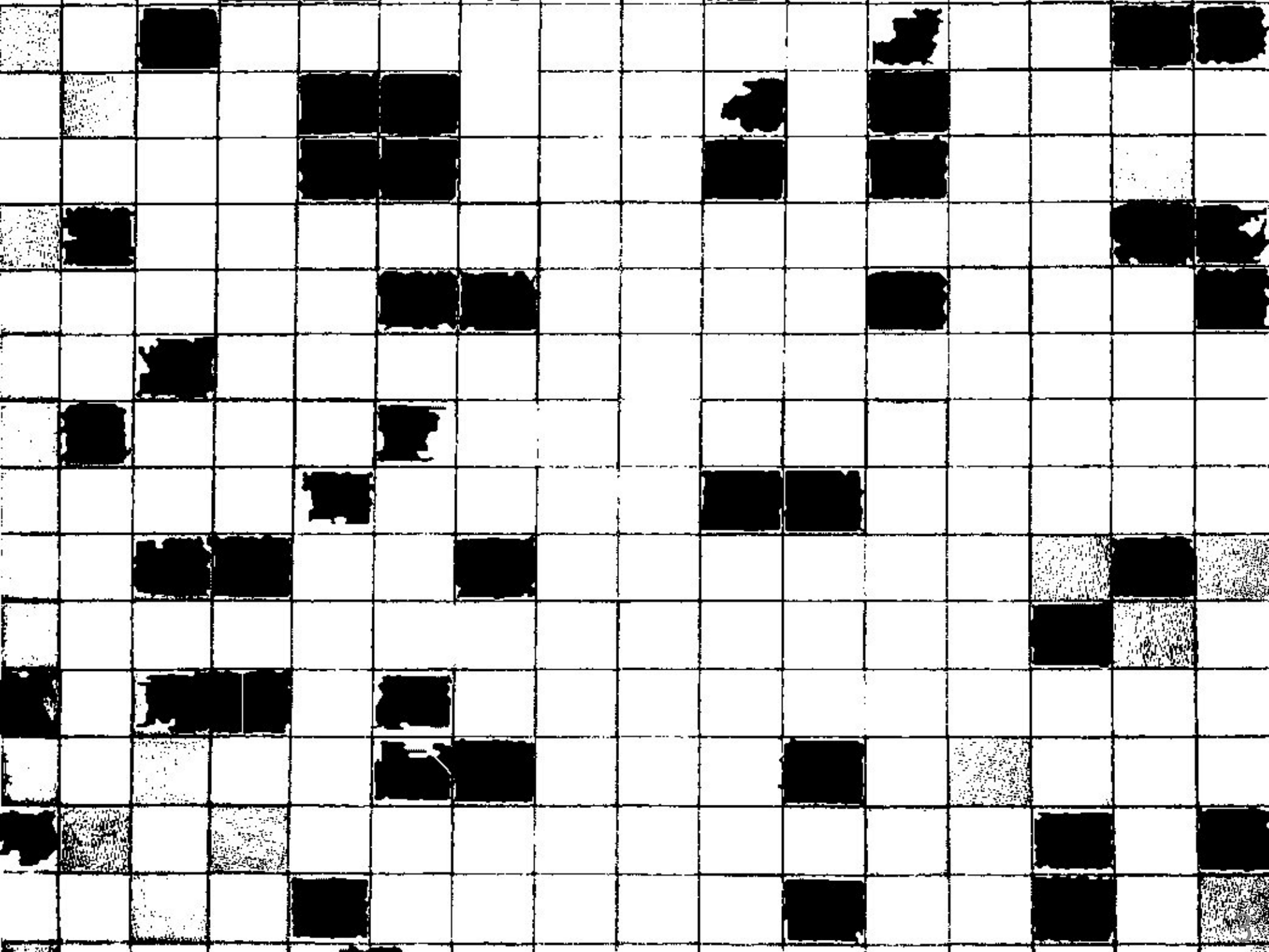
University of Chicago

distinguish interaction type

Gustav Klimt, Mohnfeld



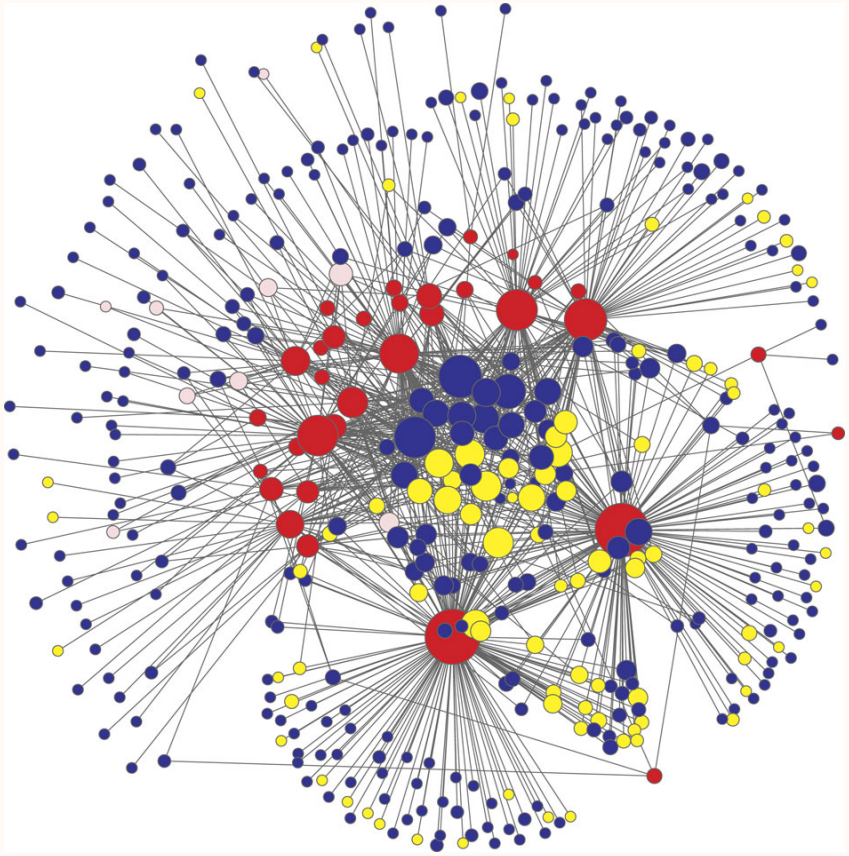
Piet Mondrian, Composition with Grids: Checkerboard Composition with Light Colors



How can we tell where this matrix has come from?

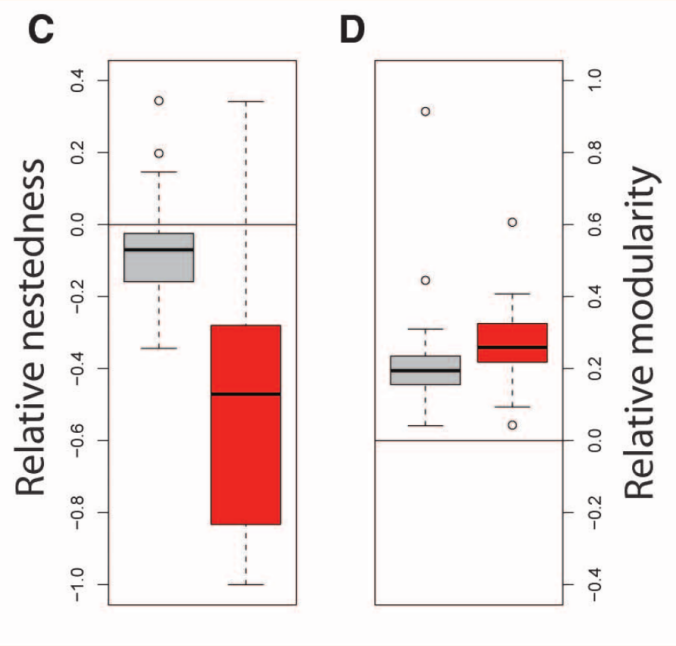
Why might this be desirable/interesting?

- Insights into mechanisms of network formation
- Unknown/ambiguous interactions



Toju, *et al.*, Nature Communications, 2014

Why might this be possible?



- Observed diffrences (*sensu* Thébault & Fontaine, Science, 2010)
- Different reactions/ coevolutions to mutualistic *vs* antagonistic interactions

Thebáult & Fontaine, Science, 2010

Data

- 323 Bipartite Networks
  - 3-21977 rows
  - 2-888 columns
  - 0.008-0.93 connectance
- Subdivided into 4 types, 14 subtypes

Antagonism	Mutualism	Biogeography	Non-biological
Bacteria-Phage	Ant-Plant	Species-Islands	Awards-Recipients
Host-Parasitoid	Plant-Seed Disperser	Species-Mountains	Participant-Answer
Host-Parasite	Plant-Pollinator	OTU-Body Site	Cultural Evolution
Plant-Herbivore	Anemone-Fish		

Methods: Measuring network properties

36 nonindependent metrics for each network

- Number of rows, columns, edges, connectance
- Leading eigenvalues
- Gaps between leading eigenvalues
- Traces of matrix powers
- Subgraph (motif) counts
- Nestedness
- Modularity

Methods: Controlling for unreliable variation

Randomized Networks

- Erdős-Rényi
  - Preserves **average** number of interactions/species
- Configuration
  - Preserves **actual** number of connections/species

Normalization

- p-value

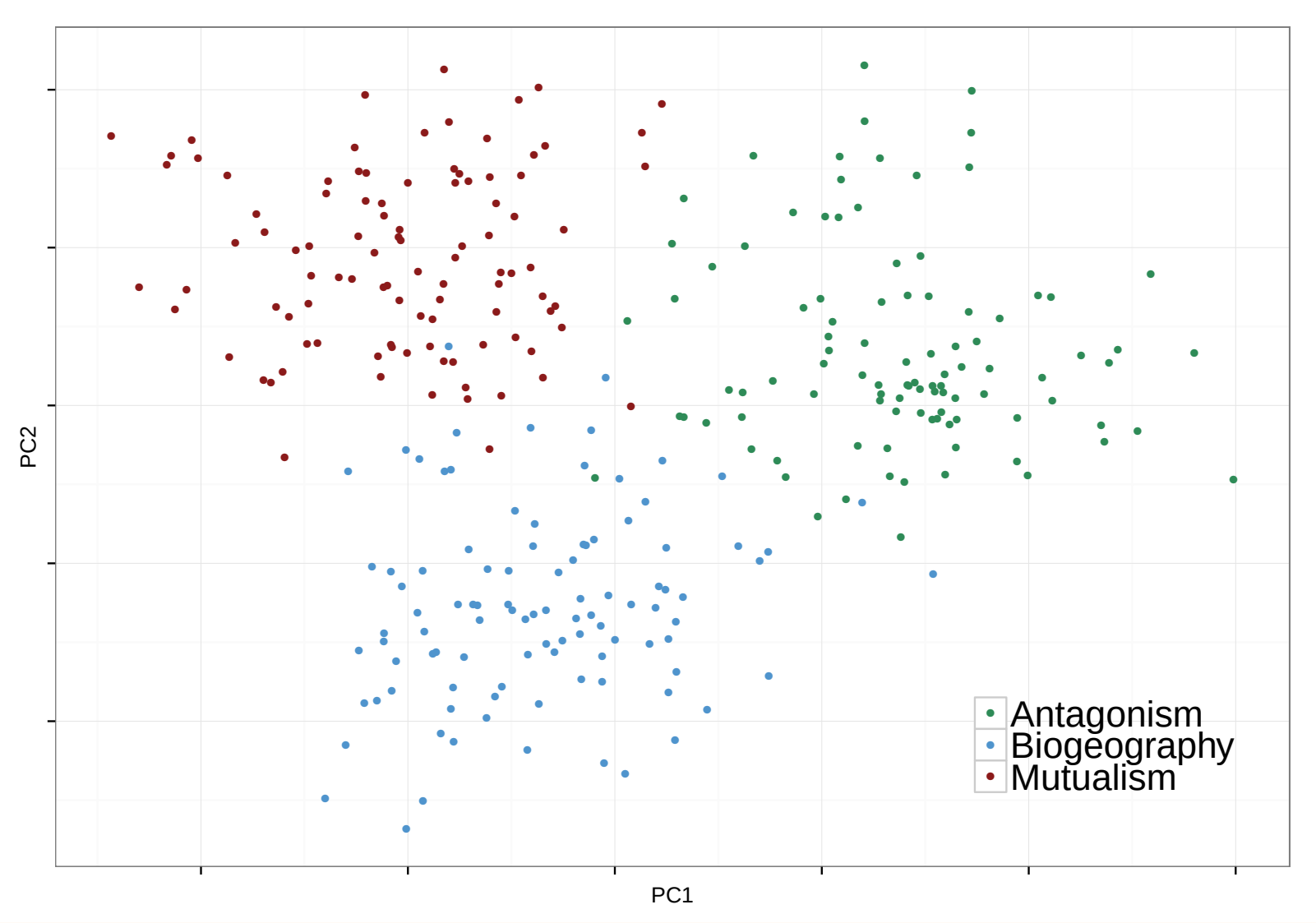
$$\frac{\text{Number of random networks with Value} < \text{Observed Value}}{\text{Number of randomizations}}$$

- z-score

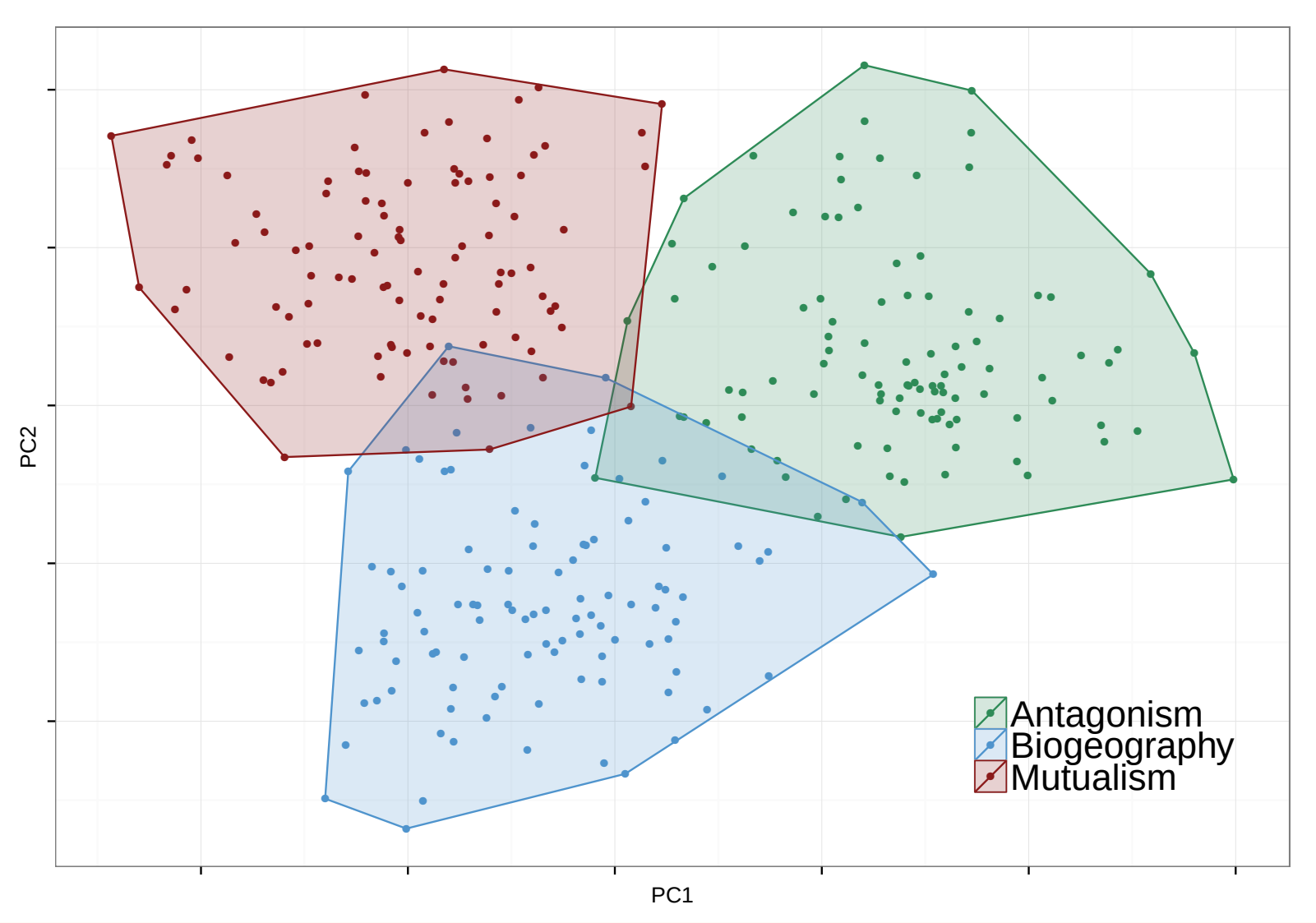
$$\frac{\text{Observed value} - \text{Mean of random network values}}{\text{Standard deviation of random values}}$$

Principal Component Analysis

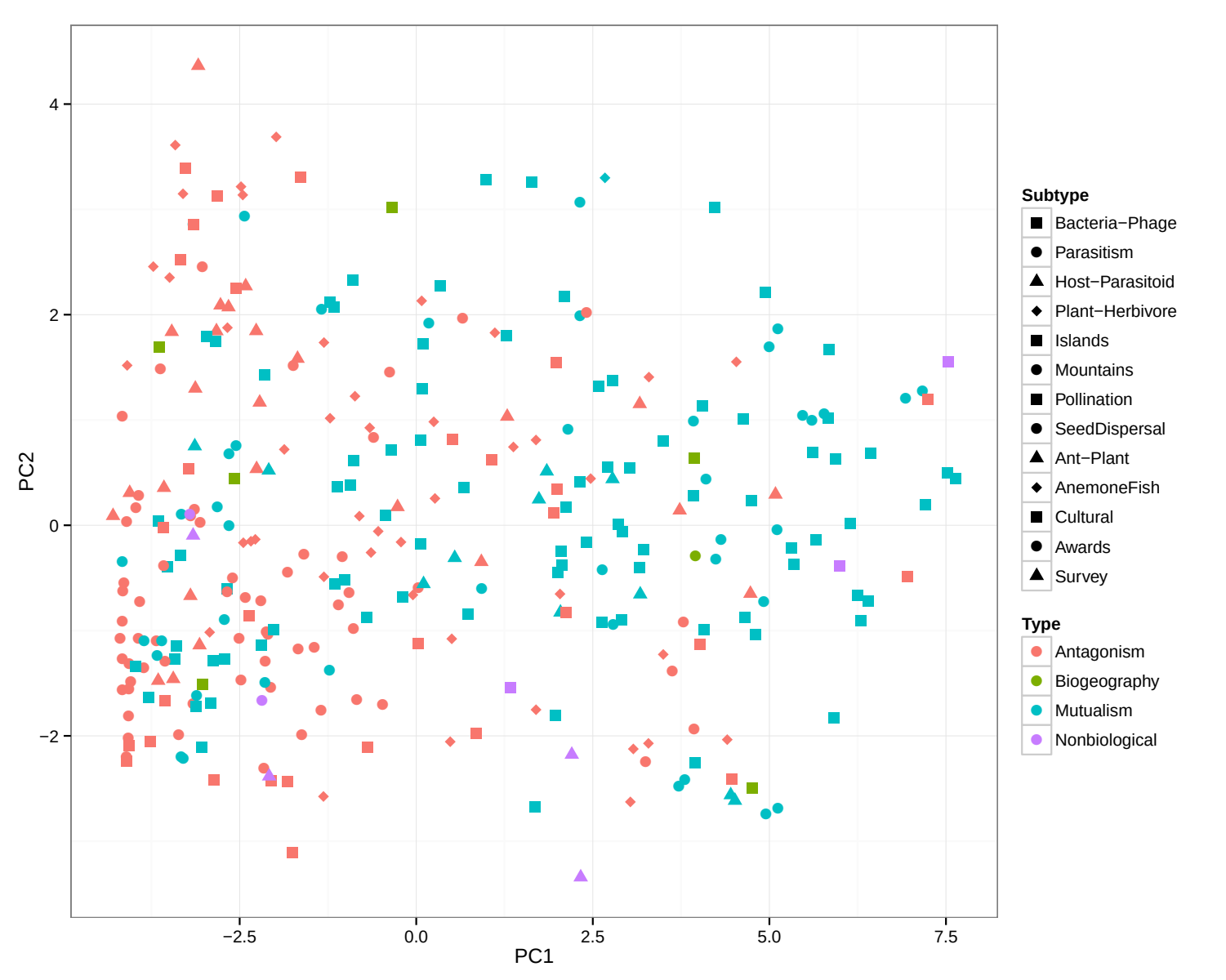
Desired results



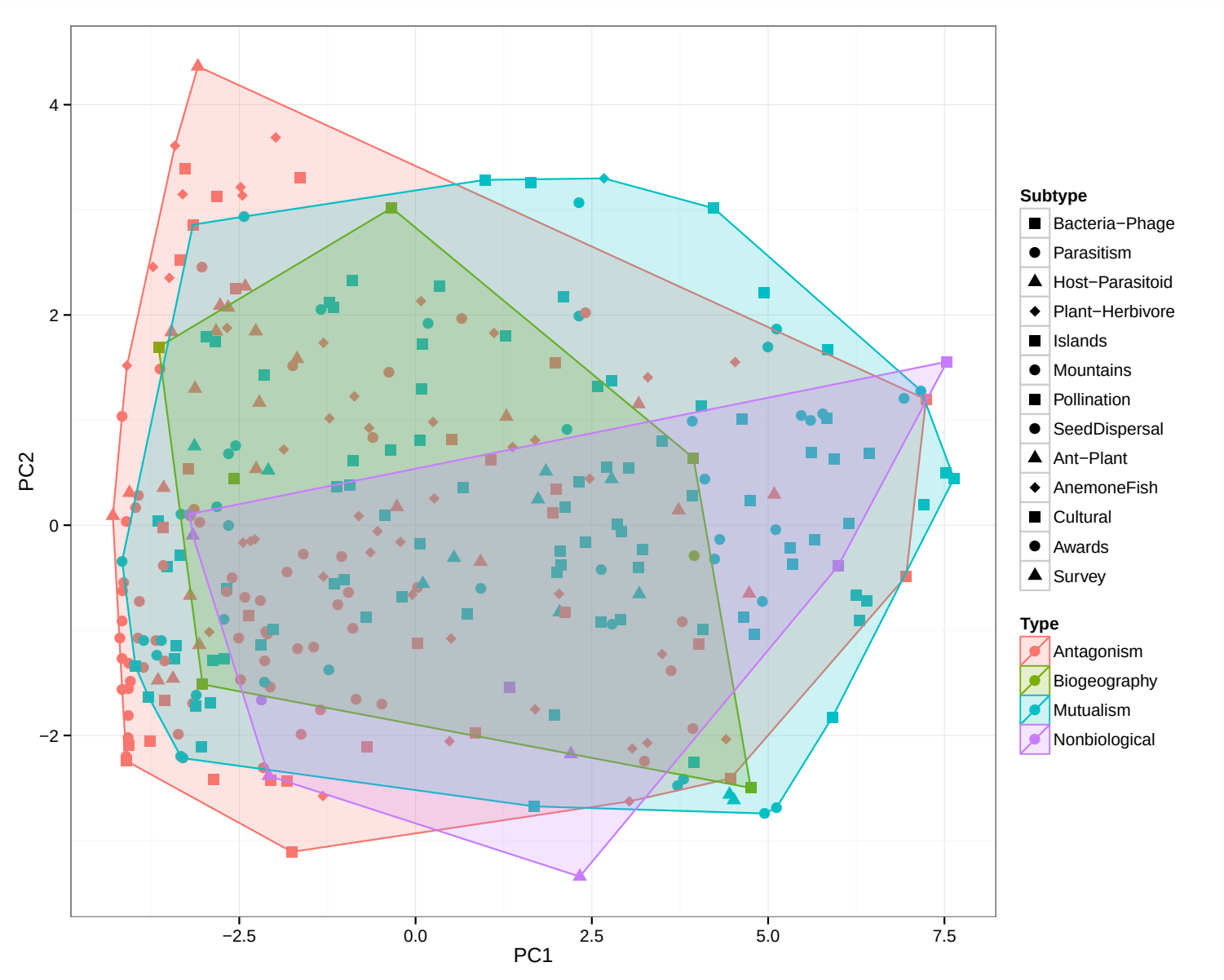
Desired results



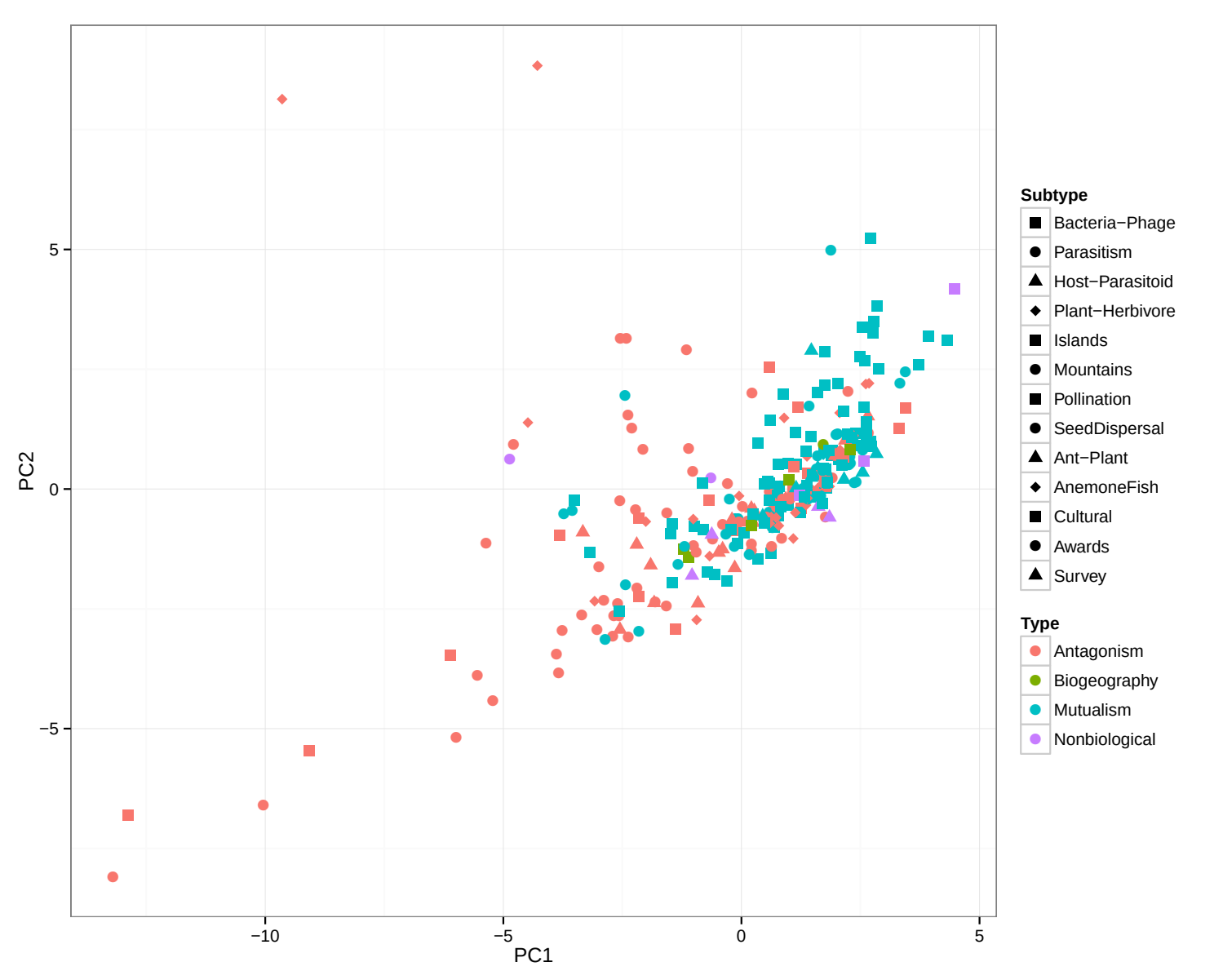
P-Values


PC2
PC1
4
2
0
−2
−2.5
0.0
2.5
5.0
7.5
Subtype
Bacteria−Phage
Parasitism
Host−Parasitoid
Plant−Herbivore
Islands
Mountains
Pollination
SeedDispersal
Ant−Plant
AnemoneFish
Cultural
Awards
Survey
Type
Antagonism
Biogeography
Mutualism
Nonbiological

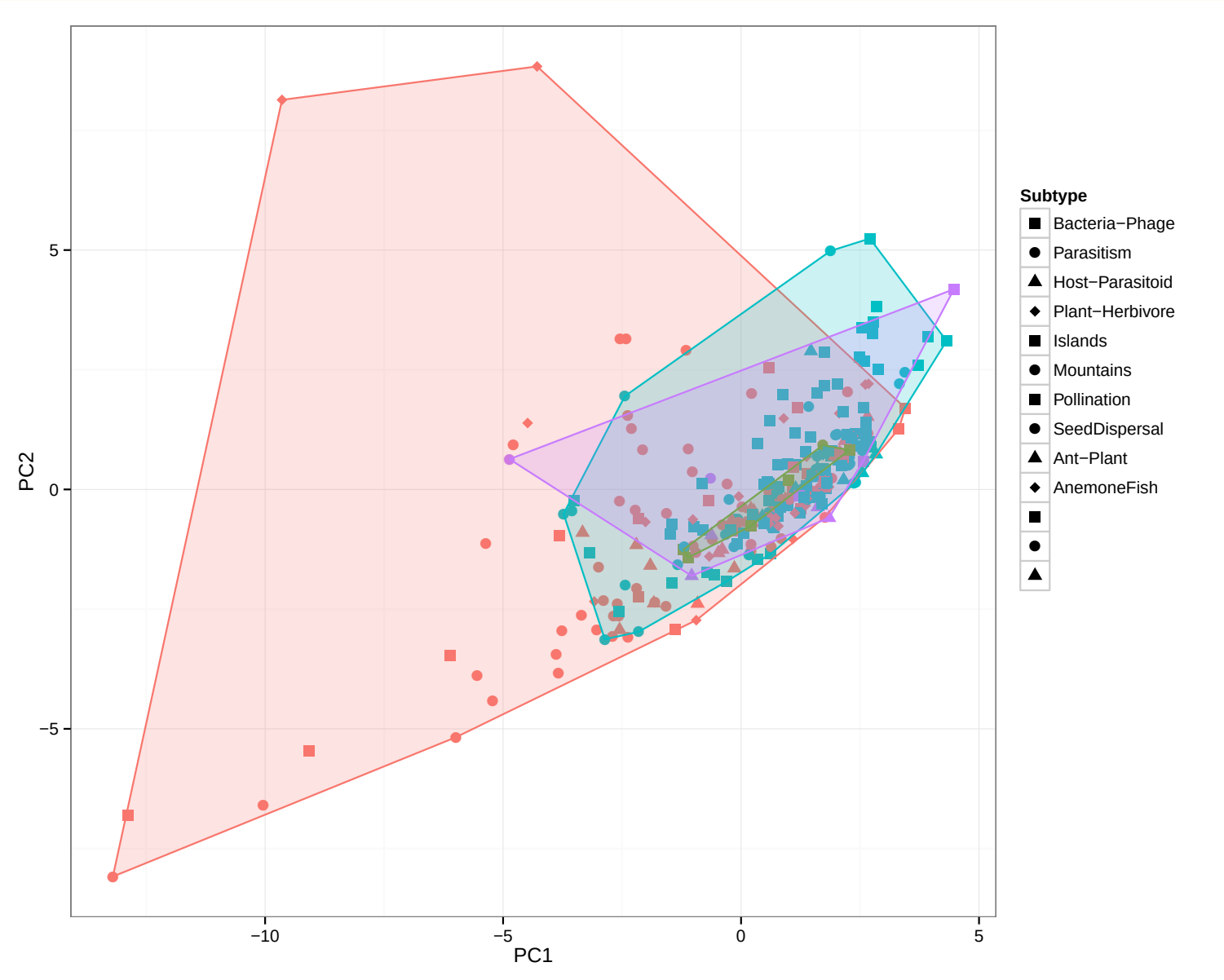
P-Values


PC2
PC1
4
2
0
−2
−2.5
0.0
2.5
5.0
7.5
Subtype
Bacteria−Phage
Parasitism
Host−Parasitoid
Plant−Herbivore
Islands
Mountains
Pollination
SeedDispersal
Ant−Plant
AnemoneFish
Cultural
Awards
Survey
Type
Antagonism
Biogeography
Mutualism
Nonbiological

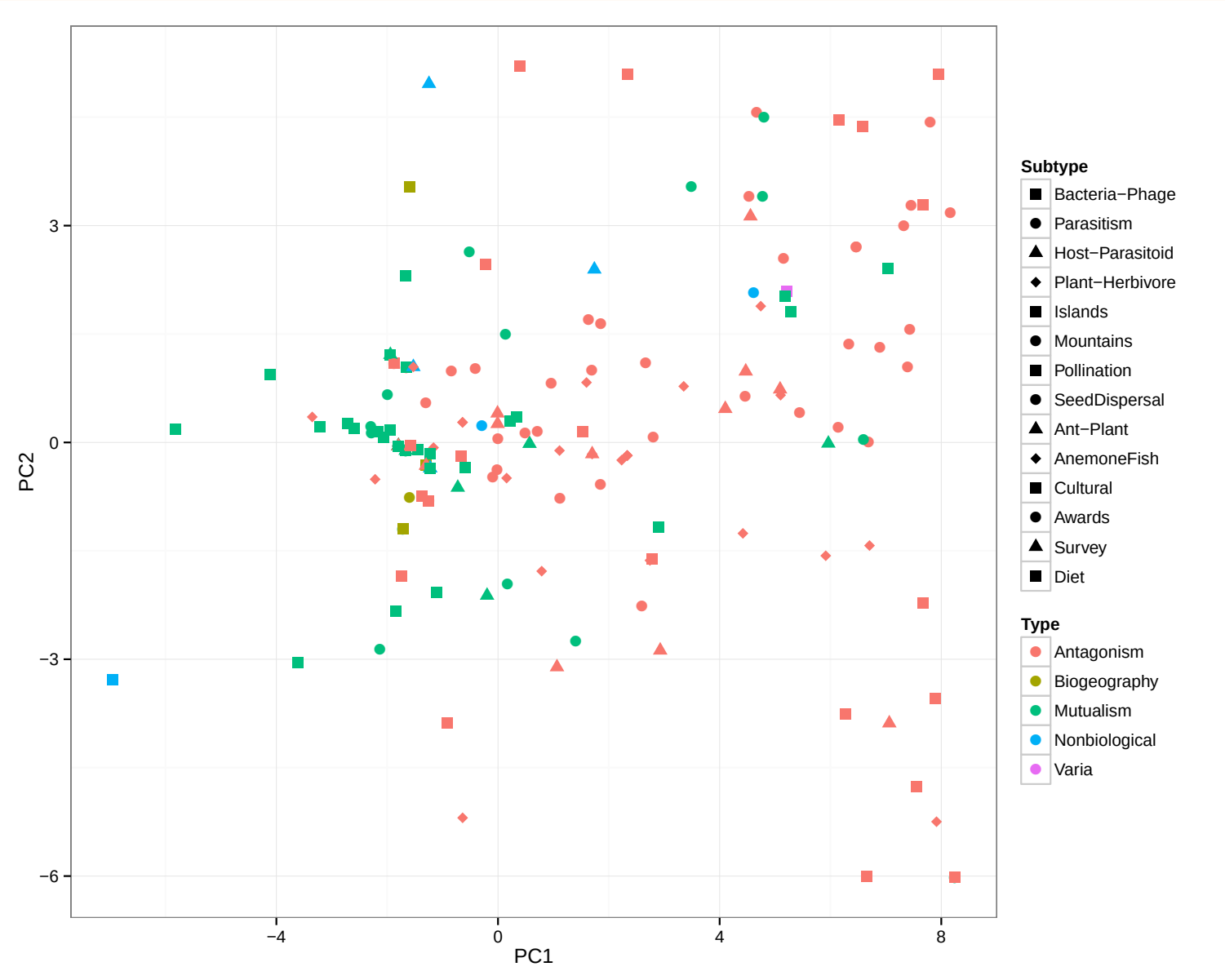
Z-Scores


PC2
5
0
−5
−10
−5
0
5
PC1
Subtype
Bacteria−Phage
Parasitism
Host−Parasitoid
Plant−Herbivore
Islands
Mountains
Pollination
SeedDispersal
Ant−Plant
AnemoneFish
Cultural
Awards
Survey
Type
Antagonism
Biogeography
Mutualism
Nonbiological

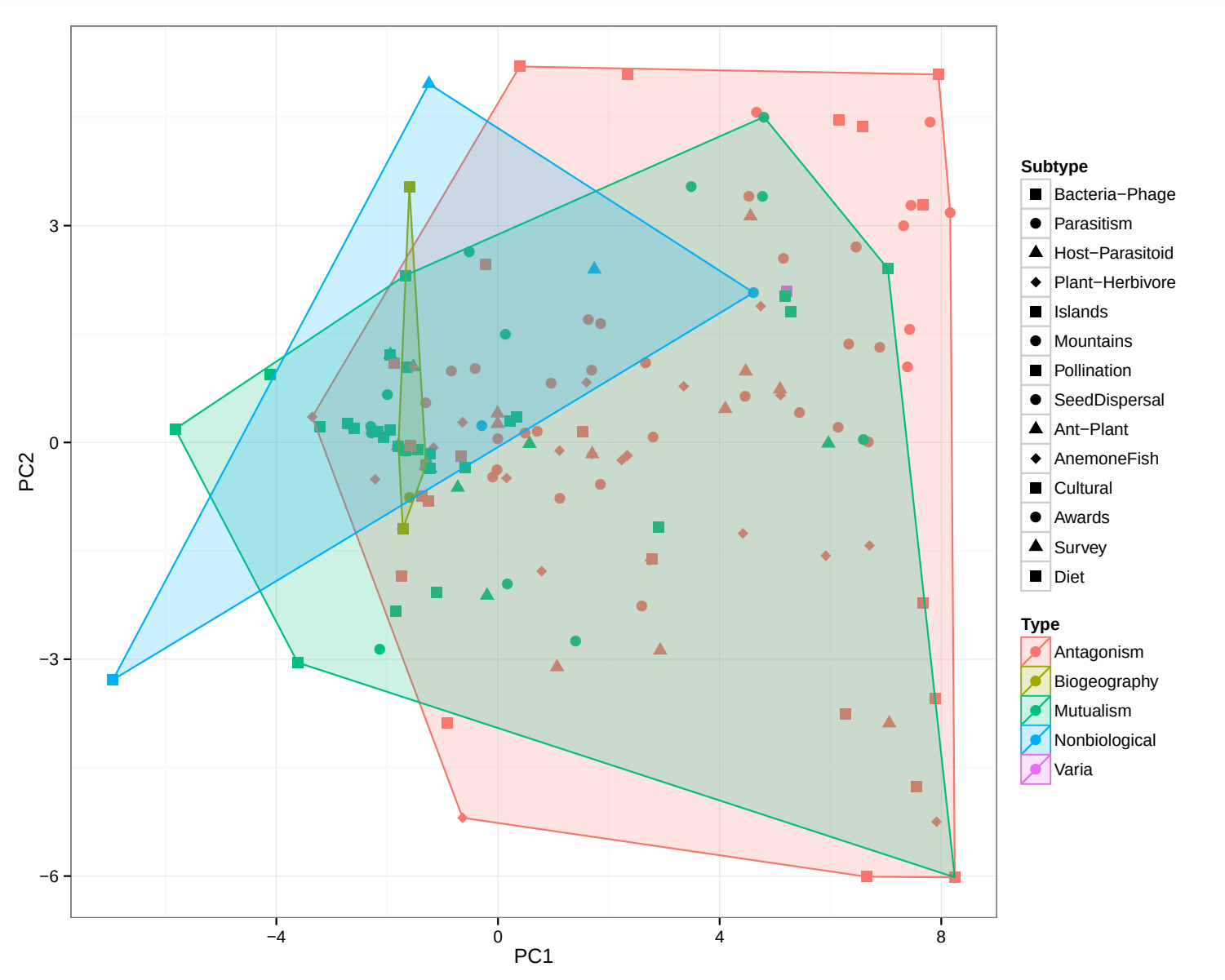
Z-Scores


PC2
5
0
−5
−10
−5
0
5
PC1
Subtype
Bacteria−Phage
Parasitism
Host−Parasitoid
Plant−Herbivore
Islands
Mountains
Pollination
SeedDispersal
Ant−Plant
AnemoneFish

Significantly Over/Under



Significantly Over/Under



What went wrong?

Insufficient Power?

- More data
- More metrics

Improper normalization?

- Conserve subgraph distributions

Low resolution?

What if nothing went wrong?

Variation within network types/subtypes is greater than variance between them

(Biological) networks are more similar than we thought

The processes which we invoked to form our expectation are not (as) important to network structure

Take-Home Message

There is no silver bullet

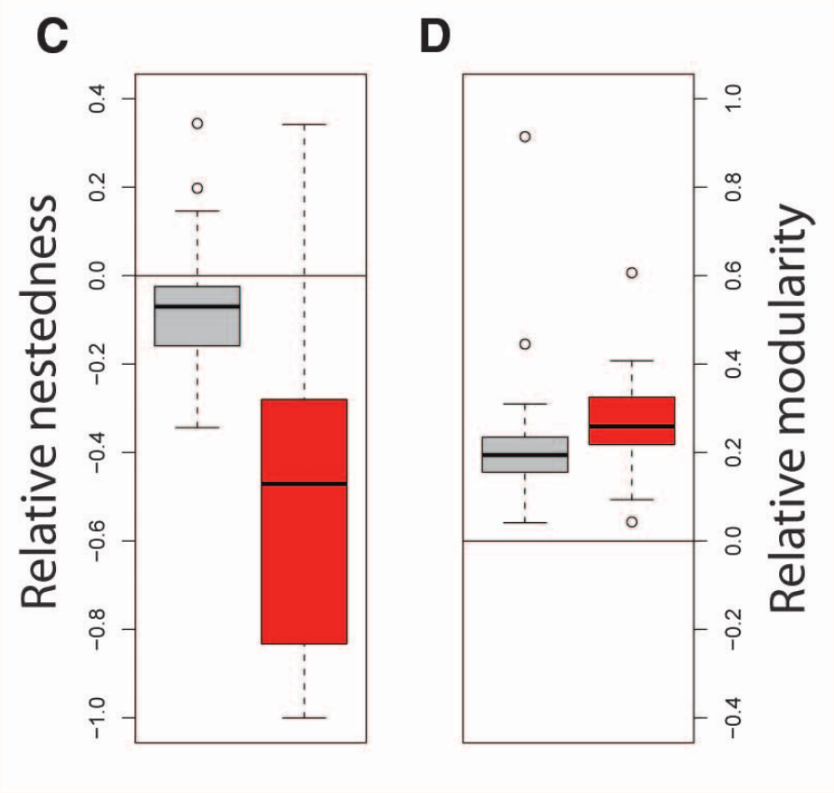
- Biological networks are hard to disentangle using network structure
- Further exploration is needed to identify which (if any) structural properties are important in delineating different types of biological networks

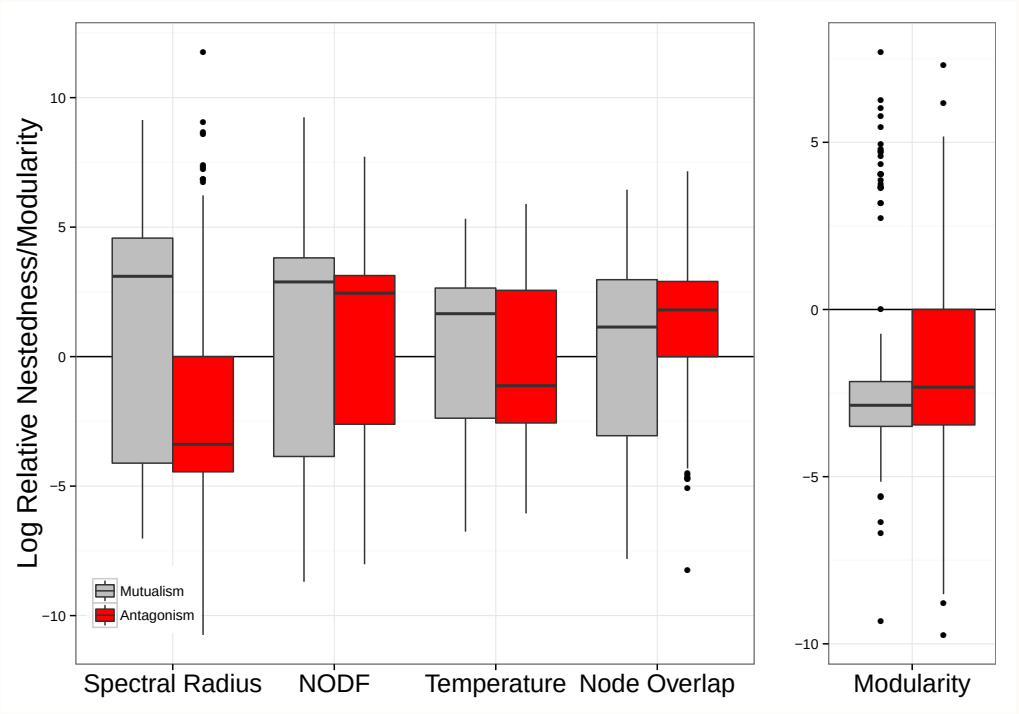
thank you

Stefano Allesina
Elisa Thébault
Jacopo Grilli
Jakez Rolland

THE UNIVERSITY OF CHICAGO

What about “known” differences?





Thebáult & Fontaine, Science, 2010