Pollination community structure revisited: a critical examination of networking statistics and a proposed modeling technique to permit cross-site networking comparisons of field data

Andrew Moldenke
Dept of Botany & Plant Pathology
Oregon State University
Corvallis, OR USA 97331
moldenka@science.oregonstate.edu

Overview 1

Foodwebs

Pollination interactions

Recent popularity (R-bipartite)

Overview 2

Serious conceptual limitations

- empty matrix cells & rare species & ‘specialists’
- density dependence & density independence
- competition and extinction analysis

Potential better approach

Primacy of ‘niche-overlap’ statistic

Foodweb analysis

The past 25 years the emphasis has shifted from
a descriptive who-eats-whom….
to an analysis of *ecosystem stability*

in terms of:

1) species extinction,
2) conservation mitigation, and
3) climate change.

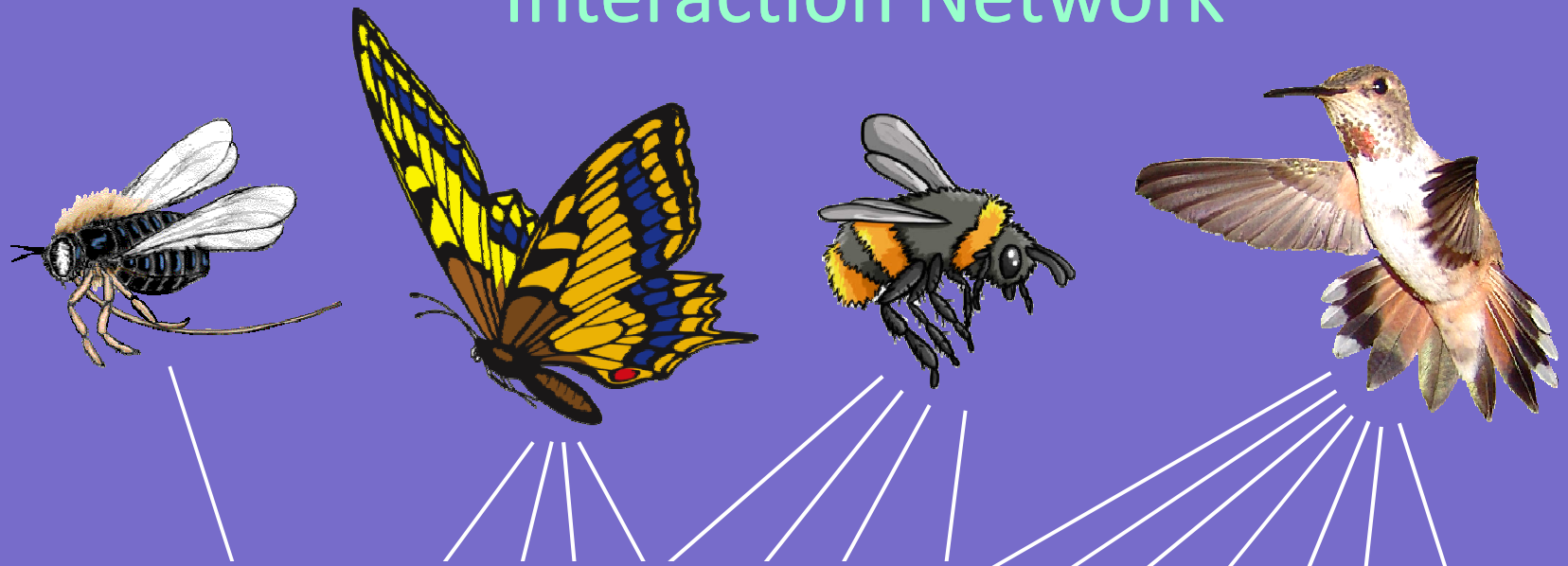
Most recently

Foodweb analysis has been expressed in terms of *networking* statistics, all collated for researchers as R-bipartite on the web.

Most recently

Several dozen fascinating papers using R-bipartite have been published on pollinator networking, but they all suffer from certain fundamental methodological problems and the unsubstantiated tendency to *interpret* competitive or extinction consequences **directly** from the superficial networking statistics.

Interaction Network



Number of plant species = 95

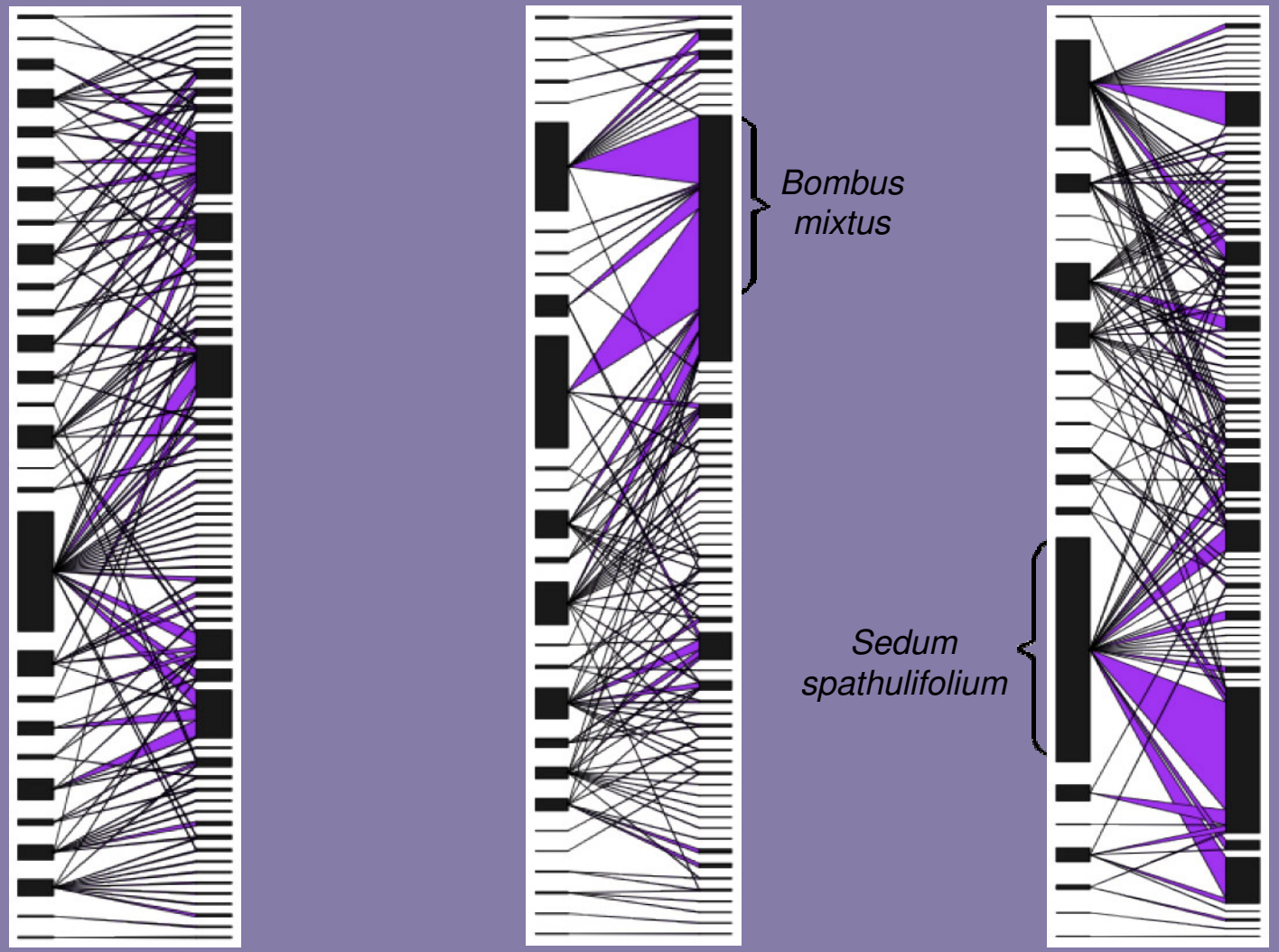
Number of pollinator species = 248

Number of potential different interactions = 23,560

(# flowers x # pollinators)



R-bipartite networking representation



Lots of bugs and plants CnE

One dominant bug CB

One dominant plant CRS

Necessity to transform data

In our studies of 25 subalpine meadows in Oregon, I found that all 19 of the most widely used networking statistics were auto-correlated.

All samples sizes must be made equivalent to those of other comparable studies ***BEFORE*** analyses can be run.

Basic lessons

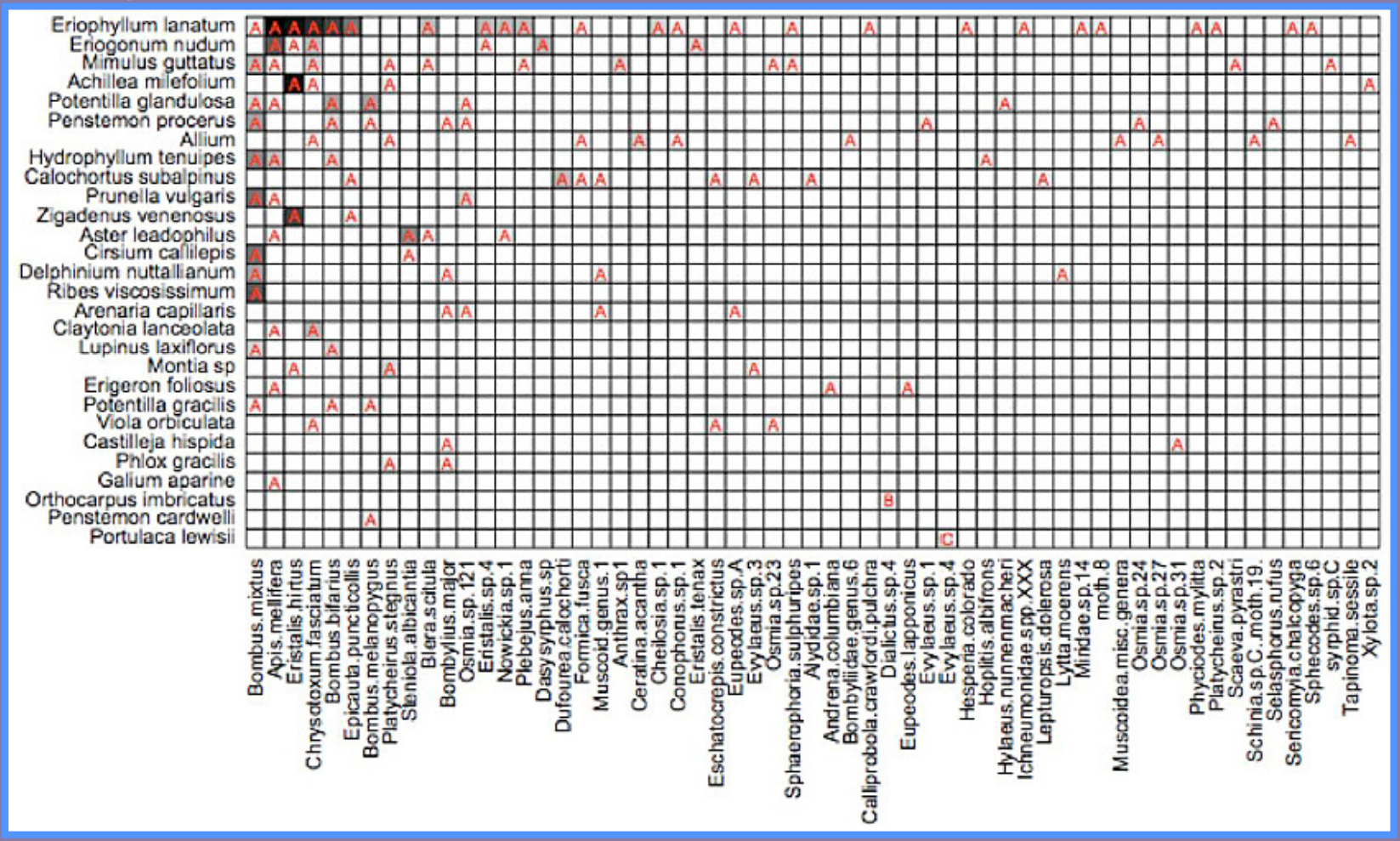
Lesson #1) pollination interactions are very diverse (i.e., species rich)

The next slide is the representation of a tiny subalpine meadow (ca. 100m x 30m) and it contains 2 dozen plants and 5 dozen potential pollen vectors (1700 interaction cells).

Matrix representation of interactions

Rows and columns are ordered by decreasing abundance

Plant species



Cone Edge Network 2011

Pollinator species

Basic lessons

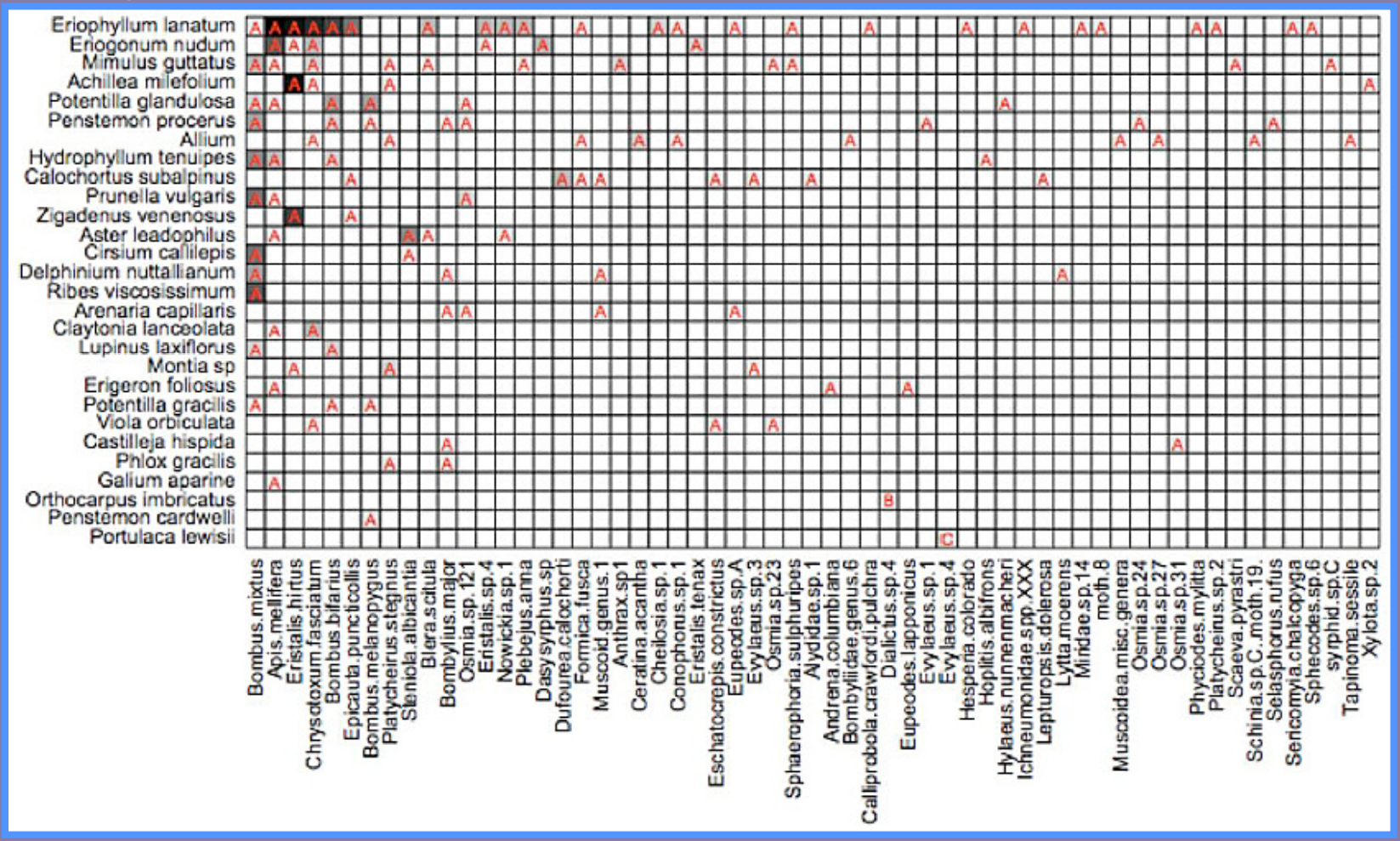
Lesson #2) you can see that most of the potential interaction combinations are blank (= not observed)

Most of the interactions that have been observed are between the most abundant species of plants and the most abundant species of insects

Matrix representation of interactions

Rows and columns are ordered by decreasing abundance

Plant species



Cone Edge Network 2011

Pollinator species

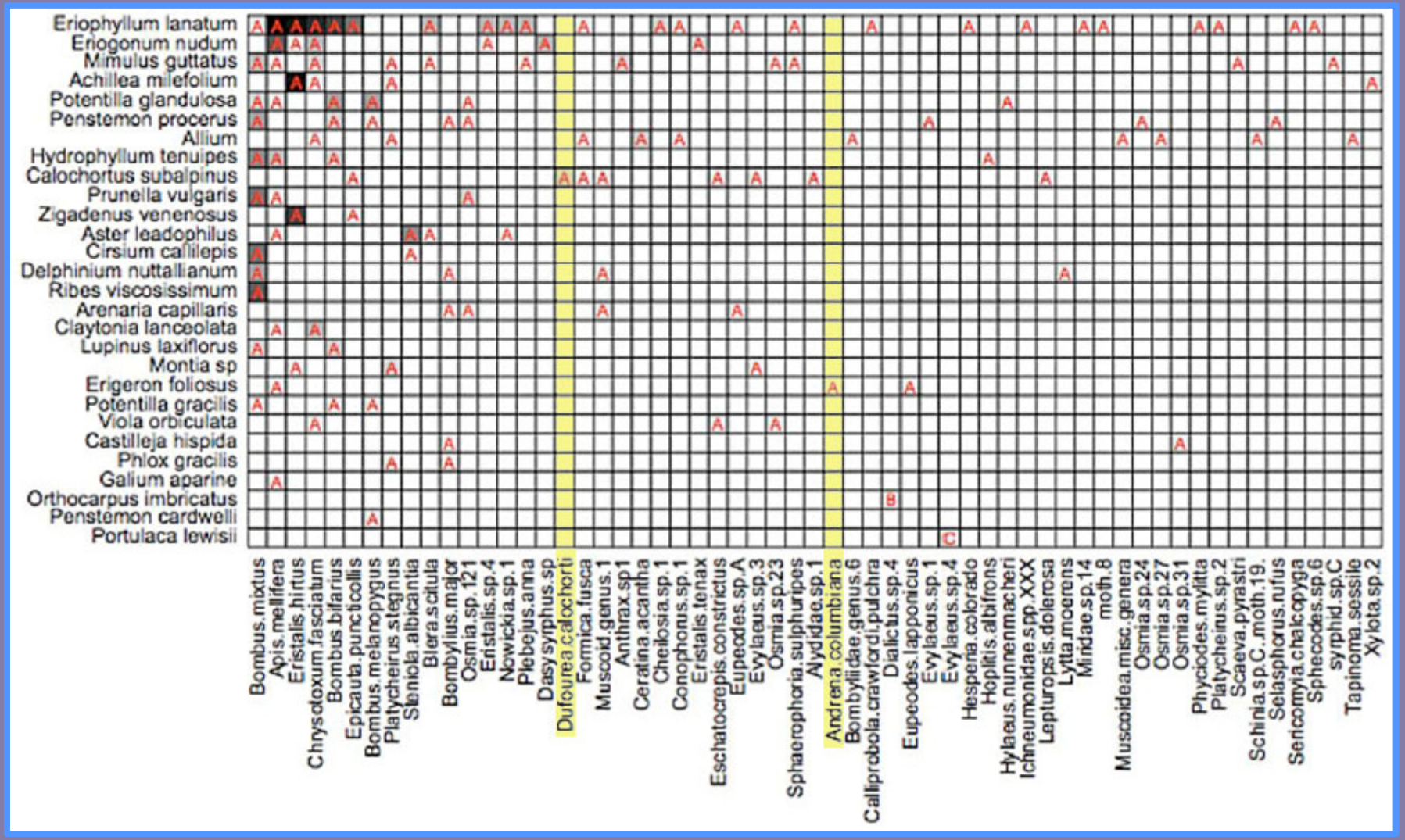
Basic lessons

Lesson #3) The majority of foodweb interactions between uncommon plants and uncommon insects *imply* specificity, i.e., that the insect is a *specialist* visitor of only 1 species of plant.

(35/58 = 60% vector species have abundances <4)

But in fact, there are only 2 actual specialists.

Matrix representation of interactions





Obvious solution



Make more observations! Then it would be obvious if the rare vector was a specialist or not!

1 tiny web:

Insects	Plants 1	2	3	4	5	6	7	8	9
1	97	75	54	20	8	3	1	1	1
2	45	33	25	15	8	5	3	2	0
3	0	0	0	0	0	94	0	0	0
4	0	0	0	0	0	0	0	1	0

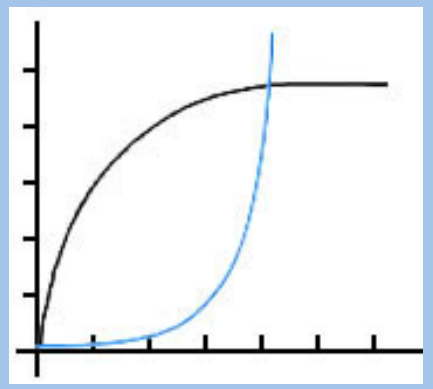
Basic lessons

Lesson #4) The more you observe/collect, the greater the species richness of pollinators in your samples.

However, the blank matrix cells still do not fill in because richness increases even *more rapidly* than the observations of generalist pollinators -- so the *apparent* numbers of specialists increases

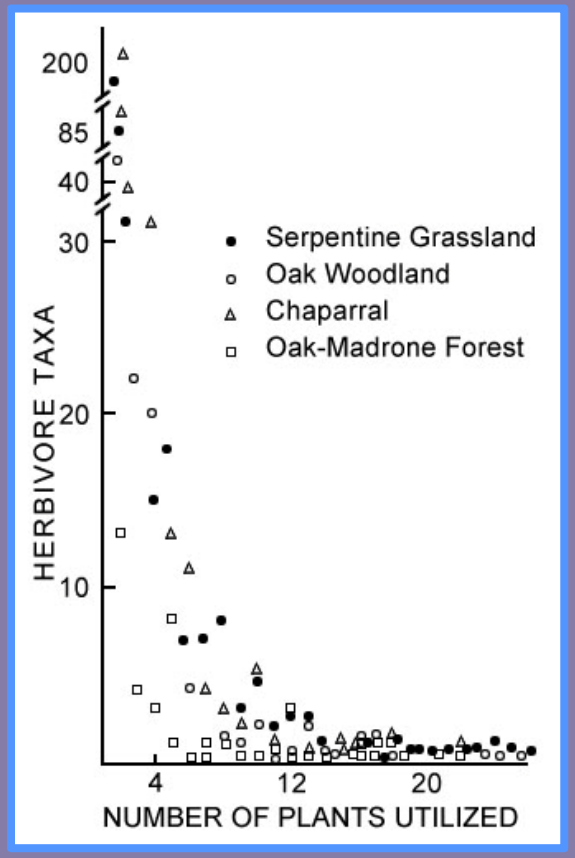
(true of everywhere I have studied, and most of the study sites reported in the literature)

Species area curve does not approach a horizontal asymptote. There are always a few very common species, fewer moderately common ones, and *very many* uncommon species.



negative logarithmic curve

When measuring abundance, 68% of vectors are found only 1-4 times per site, even when the total count is well in excess of 10,000-25,000.



So the more you make observations, the more rare species you encounter!

Specialized vector feeders

Determination of vector specificity must be made *independently from field data* because your small sample subset does not encompass the variability of the possible habitats.

For North America we have a good general idea of the feeding preferences of native bees

(bees = most likely group for genetically-determined specialists; but also there are herbivore/host-plant specialization to consider)

Specialized vector feeders

In fact for western N America where I have worked, specialist-feeding bees vary from 5-55% of the local bee fauna) (Moldenke, 1978)

BUT:

The highest richness of specialist-feeding bees is in the southwestern deserts; it is usually <<20% in most other plant communities.

Basic lessons

Lesson #5) Really: in most habitats everybody is a generalist

(or potentially a generalist)

When looking at your data, with the notable exception of a few specialist bees, *all other vectors* are potentially *generalists*.

So, the basic model should be a *null model* of generalists

Interactions between generalist vectors feeding upon generalist plants is simply the product of:

vector density x plant density

(with a few exceptions due to specialized morphology and behavior)

Then compare the predictions of the null generalist model with the field observations to determine how well you have in fact described the system.

(Note: most manuscripts convert their matrices to presence/absence rather than numerical density data – converting emphasizes the biases we just discussed, it does NOT diminish them – keep the matrices quantitative!!)

Programming exceptions into the null generalist model





VECTORS:

- Flight times not overlapping blooming times
- Too big to alight on/enter into the flower
- Inappropriate tongue length
- Known specialist-feeders
  on only a single genus



Programming exceptions into the null generalist model



PLANTS:

- Tubular flowers (tongue-length of vectors)
- Upside-down flowers
- Closed flowers
- Darkly colored flowers
- Nocturnal flowers







(1, or more than 1, exclusionary floral device)

Different guilds of flowers

GENERALIST FLOWER (any pollinator morphology)

-- open dish 34 %

SPECIALIST FLOWER (exclusion types)

-- closed flower 9 %

-- long tubular flower 25 %

-- hanging flower 7 %

-- dark red/purple color 11 %

-- tiny flower 17 %

(weak = practically no pollinator support)

-- etc. (nocturnal; really weird, etc) trace %



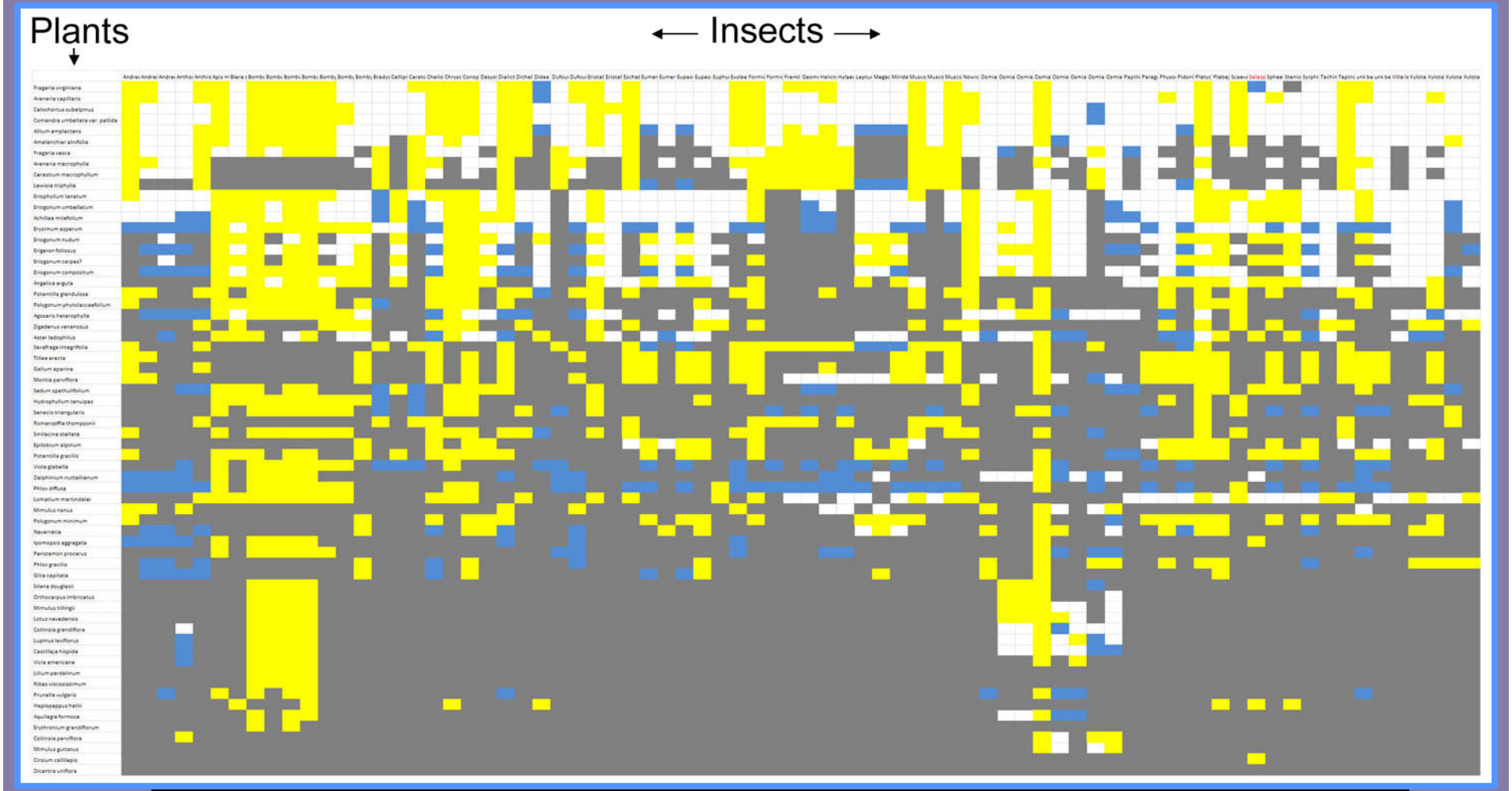








So, we can then combine the density-independent exclusionary covariates of the bugs and the plants with the density-dependent null matrix of ‘everyone is a generalist’ – this is what we get as a prediction of matrix



Yellow = highly probable	White = likely	Blue = unlikely	Gray = highly unlikely

Therefore, to correct pitfalls in the literature

Field Data Matrix

(observational data)

Normalized Probability Matrix

(transformed observational data)

Null Generalist Matrix

(transformed field data * density dependence)

Exclusionary behavior/morphology Matrix

(density-independent p & v covariates)

Null * Exclusionary = Covariate Matrix

(dens-dependent * dens-independent)

Then, simulation model with inter-annual variability

(at each step you compare the predictions with the actual observations and appropriately adjust the model)

So, as the American hamburger advertisements say –
"Where is the beef?"

With this approach to describing and analyzing pollination interactions, it is easy to see that the *density-independent factors* with which individual species of plants and vectors compete through time are defined and quantified by the *'Exclusionary Matrix'.*

Simulation modeling

The *exclusionary matrix* is the paramount driver through time -- densities fluctuate, immigration occurs, local extinction of vector populations occur -- but the most important interactions between species are those which determine the most potentially intensely competing species regardless of densities

(the smaller portions of the matrix with both exclusive and overlapping plant/vector combinations = ‘modules’)

Competition and local extinction



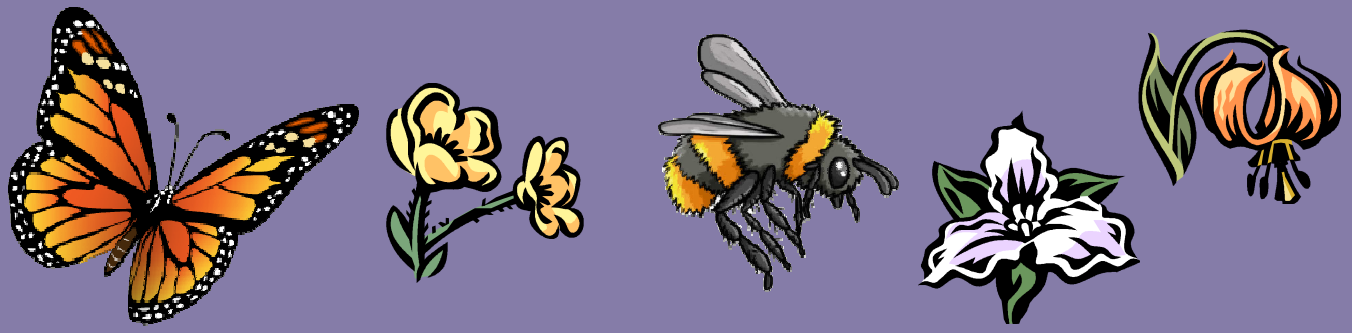
The most heuristic statistic is the matrix-adaptation of the Shannon Diversity Index called the *'niche-overlap index',* when I first proposed it in 1975.

It quantifies how any 2 species differ/overlap in their use of resources:

plants using ratio of unique/generalist pollen vectors

insects using unique/generalist plants

thereby, quantitatively defining the modules of the entire community with the most likely direct competitive interactions.



Niche overlap

For vectors:

$$(d_{ij}) = \frac{\text{\# food resources utilized by vector}_i \text{ (but not by vector}_j\text{)}}{\text{total \# food sources of vector}_i}$$

Produces a unique value for each cell in the entire non-symmetric matrix

Values range from = 0 (nothing uniquely visited, or vector species are indistinguishable, i.e., 2 different species of bumblebees) to = 1

(vector species are completely unique, do not share any food resources, i.e., a tumbling flower-beetle and a mason bee).

Biological consequences of networking

Competition and extinction are *not* direct consequences of networking; networking analysis only helps to better delineate the potential for direct competition.

Many other factors are likely to *overwhelm* simple foodweb interactions relative to species persistence (examples: nest-site availability, disease, annual weather variation).

Take-home messages about pollination networking

1) The biological consequences are best approached with *quantitative* (NOT binary +/-) *matrices*.

2) Field samples must always be *scaled to samples with equal* size, richness or matrix volume in order to be compared (R-bipartite stats are mostly auto-correlated with N, or SR).

3) *Density-independent matrix covariates* must be separated from density-dependent ones and are most easily quantified and tested for predictive value in the sequential method just presented.

4) *Niche-overlap values* for individual plants and vectors have the most importance for productive *adaptation to competitive studies*.

REFERENCE:

Once the modeling is complete all the data, both original and transformed, together with all the covariates and the history of tweeking the model, will be available at the homepage of the Andrews Forest Long Term Ecological Research (LTER) site

andrewsforest@oregonstate.edu