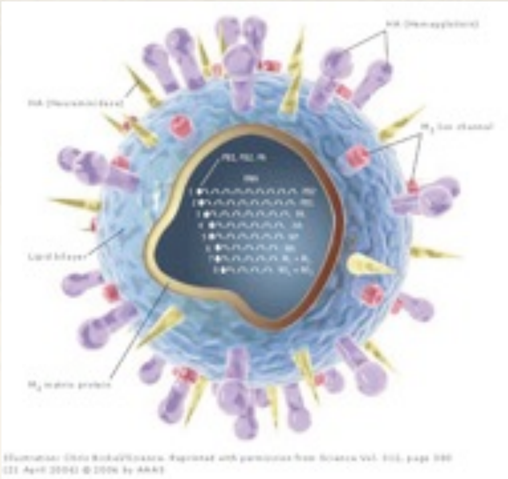
Ecological Modelling Using Big, Deep Spatial Data

Chris Stephens
C3-Centro de Ciencias de la Complejidad y Instituto de Ciencias Nucleares, UNAM

Seminar, Stanford University 28/04/2017

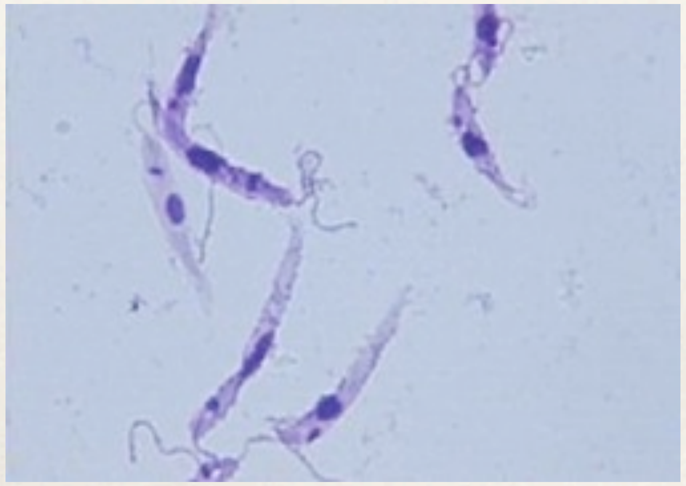


Ecology is the scientific analysis and study of **interactions** among organisms and their environment

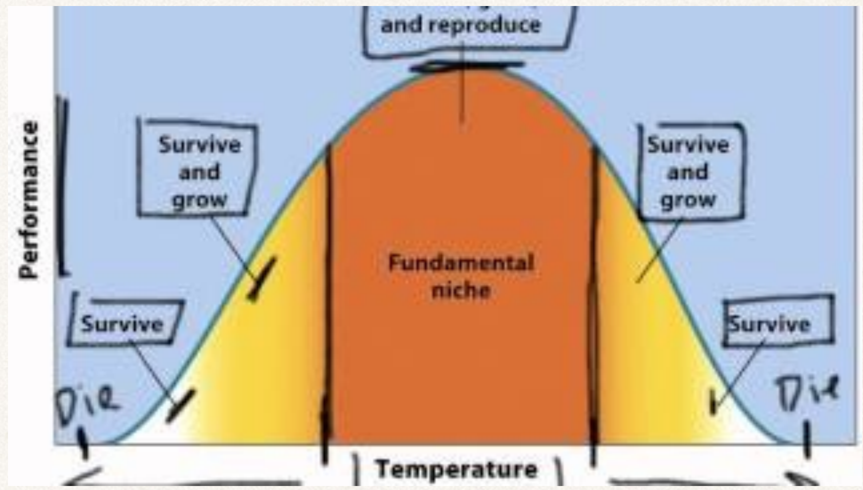
Multifactorial with changing interactions

An Ecology is a Complex Adaptive System

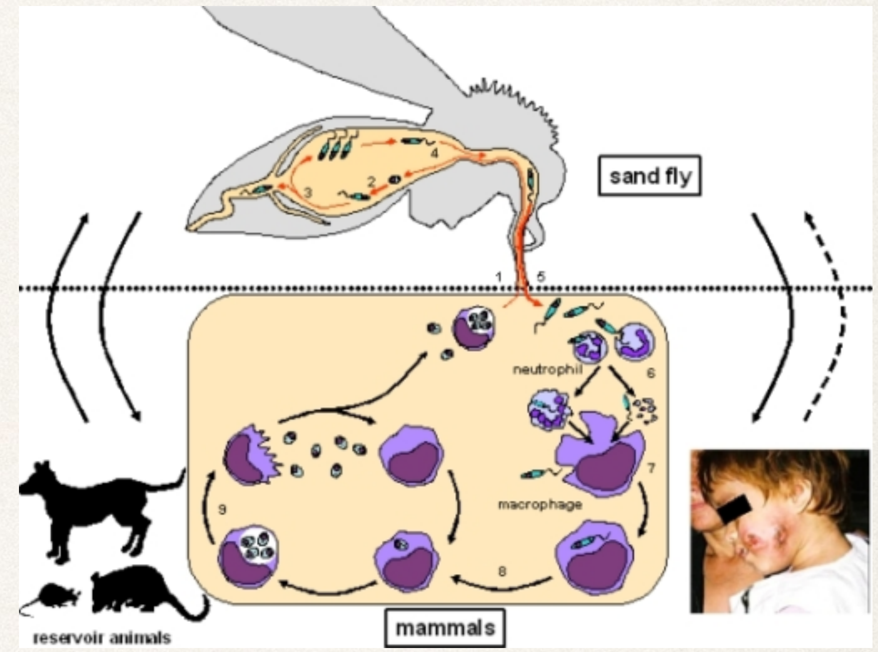
Type of interaction	Sign	Effects
mutualism	+/+	both species benefit from interaction
commensalism	+/0	one species benefits, one unaffected
competition	-/-	each species affected negatively
predation, parasitism, herbivory	+/-	one species benefits, one is disadvantaged



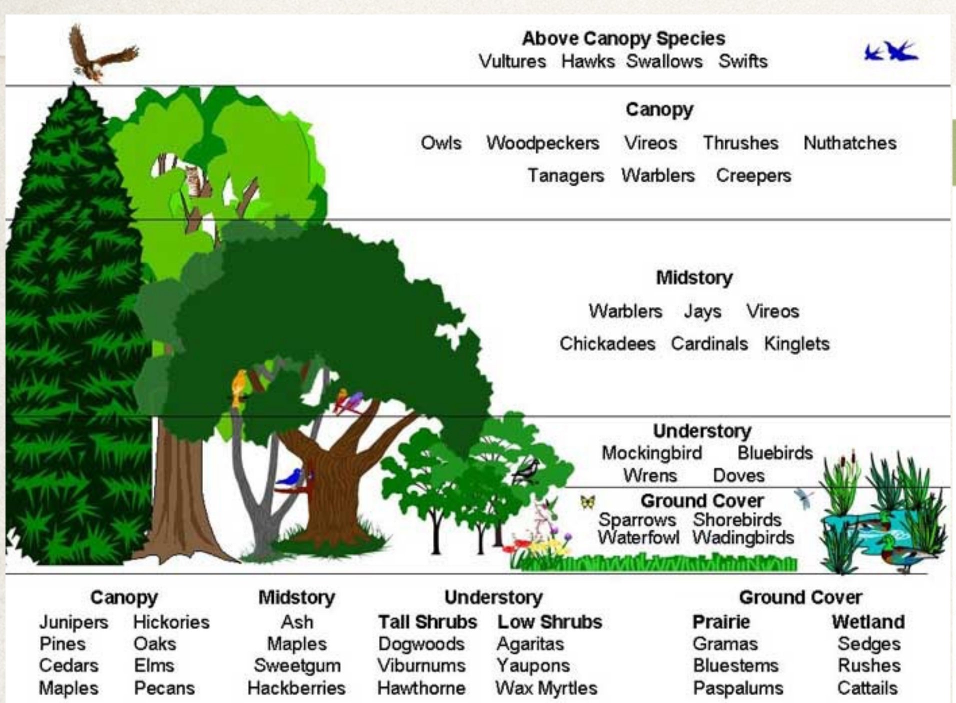
Niche versus Community

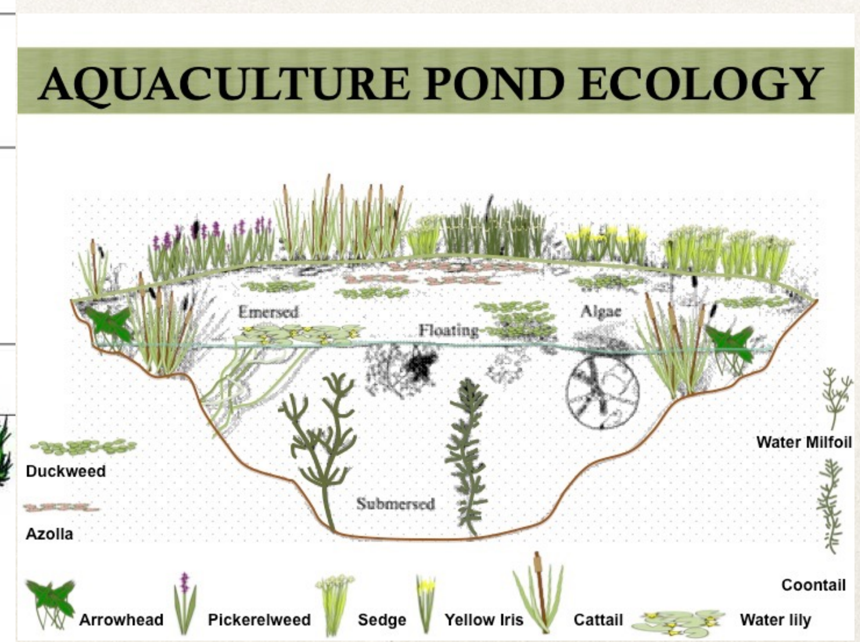


Hutchinson: "the set of biotic and abiotic conditions in which a species is able to persist and maintain stable population sizes."



Community ecology examines how interactions among species and their environment affect the abundance, distribution and diversity of species within communities.



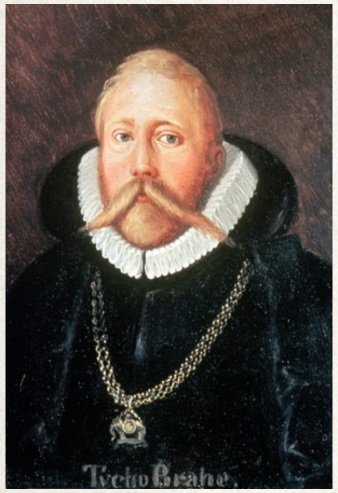


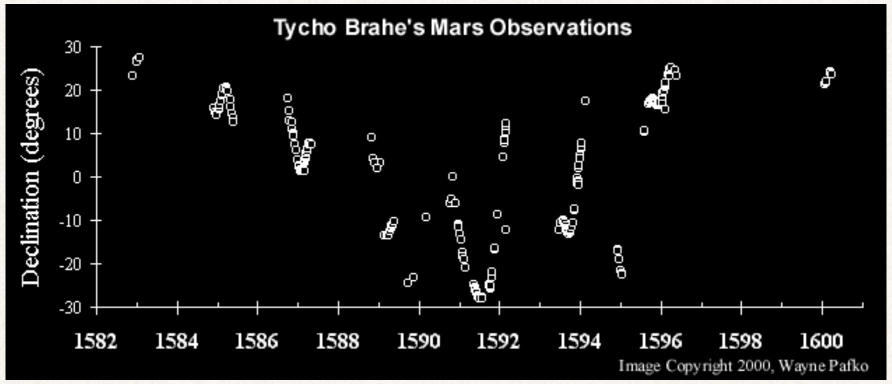

Can we infer ecological interactions without direct observation?
Just how many interactions can we directly observe?
Importancia médica
T. Infestans
T. barberi
T. pallidipennis
T. longipenis
T. recurva
T. neotomae
Culex quinquefasciatus
gettyimages
Frank Lukasseck
Roberto L.M. Novaes
ARKive
www.arkive.org
SCIENCEphotoLIBRARY

Inferring Interactions from Spatial Data…

A famous historical antecedent

Data —> Phenomenology —> Taxonomy —> Theory —> Isn't all science data science?





Data —> **Phenomenology**

Kepler's Laws

1. The orbit of a planet is an ellipse with the Sun at one of the two foci.
2. A line segment joining a planet and the Sun sweeps out equal areas during equal intervals of time.
3. The square of the orbital period of a planet is proportional to the cube of the semi-major axis of its orbit.

Isaac Newton computed the acceleration of a planet moving according to Kepler's first and second law.

1. The *direction* of the acceleration is towards the Sun.
2. The *magnitude* of the acceleration is inversely proportional to the square of the planet's distance from the Sun (the *inverse square law*).

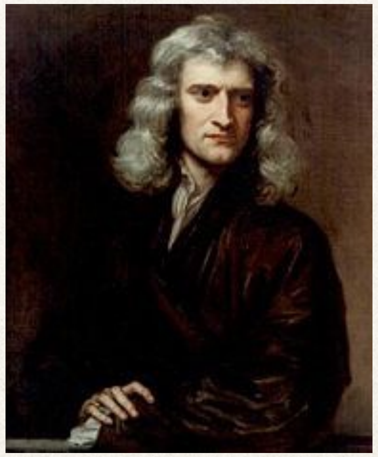
This implies that the Sun may be the physical cause of the acceleration of planets.
Newton defined the force acting on a planet to be the product of its mass and the acceleration. So:

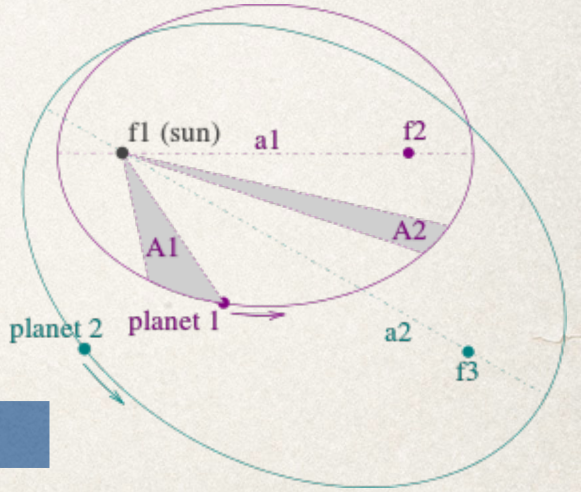
1. Every planet is attracted towards the Sun.
2. The force acting on a planet is in direct proportion to the mass of the planet and in inverse proportion to the square of its distance from the Sun.

The Sun plays an unsymmetrical part, which is unjustified. So he assumed, in Newton's law of universal gravitation:

1. All bodies in the solar system attract one another.
2. The force between two bodies is in direct proportion to the product of their masses and in inverse proportion to the square of the distance between them.

As the planets have small masses compared to the Sun, the orbits conform approximately to Kepler's laws. Newton's model fits actual observations more accurately.





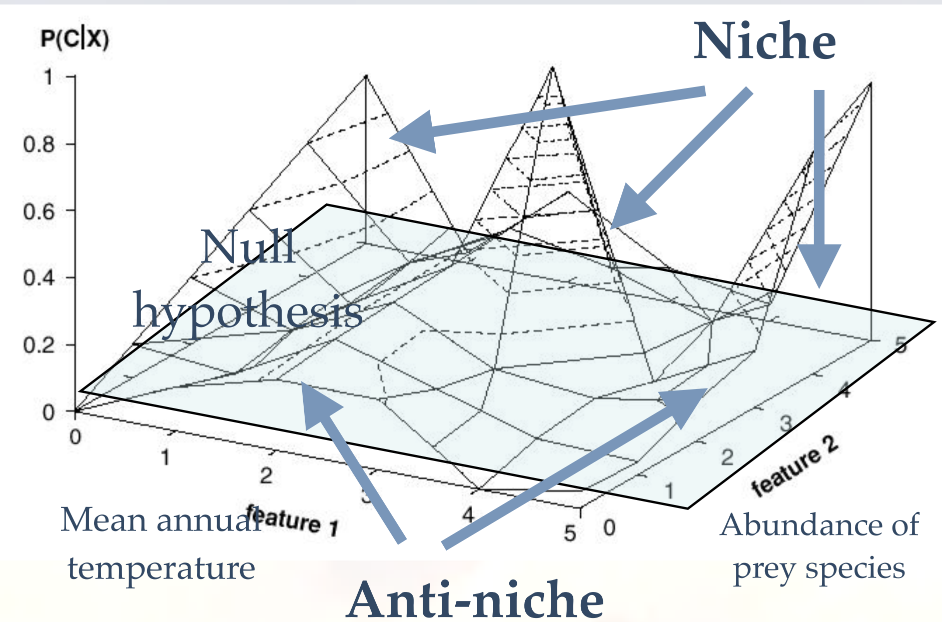
$$F = ma$$

$$F = GMm / r^2$$

Theory

“Keplerian” Ecological models

CONABIO



What do we want to predict?
C = (C1, C2, C3, …, CN)
the presence, or abundance, or,… of one or more populations or taxa

P(C | X)

S(C | X)
Risk score

What affects it?
The “niche”
X = (X1, X2, X3, …, XM)

A large part of the complexity is in the multi-factoriality of both C and X. Adaptation is inherent in the fact that P(C | X) can change in time.

$$\mathbf{X} = X(sd)+X(se)+X(n)+X(ev)+X(g)+X(af)+X(hm)+X(i)+X(sp)+\ldots$$

Macro-Climactic factors

Micro-Climatic factors

Hydrography

Prey species

Human activity

Behavioural characteristics

Phenotypic characteristics

Competitor species

Predator species

Problems of co-dependence and causality

And the data? Where are the "Brahes"? There's lots of them!

Thanks to the Data Revolution

Normally data mining takes place in a "categorical" space (the equivalent in ecology is a niche space). However, most ecological data is spatio-temporal at multiple scales. Spatial data mining is much less developed than standard data mining.

- Collection data ← SNIB, CONABIO
- Ecological niche data
- Ecological niche model data
- Socio-economic data
- Socio-demographic data
- Phenotypic data
- Vegetable and crop cover
- Geographical data
- Medical and public health data…

Problems with spatial data:

Different sources

Different location, data base, access,…

Different data types

categorical, metric, continuous, discrete,…

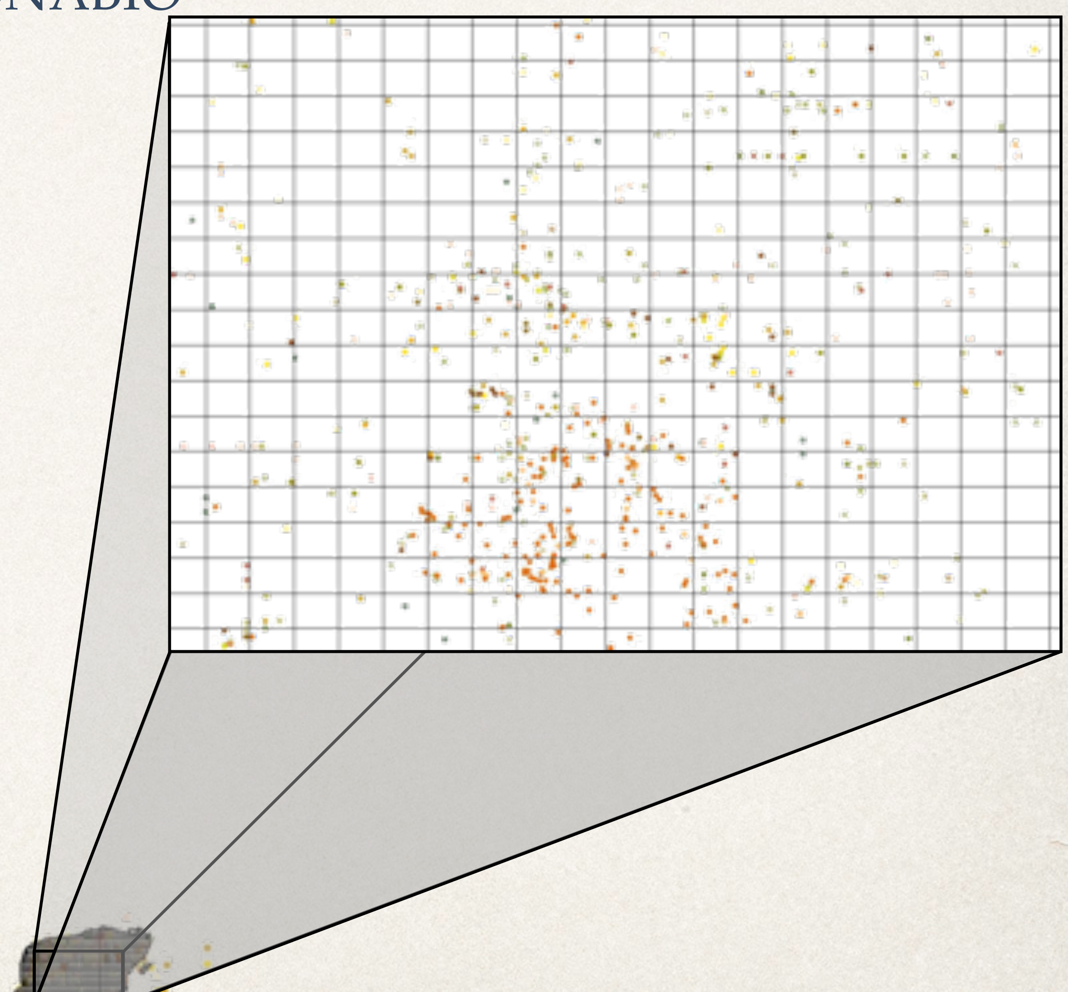
Different spatial resolution

Explicit – e.g., pixel by pixel in environmental layers

Implicit – 30,000,000 data points versus 30

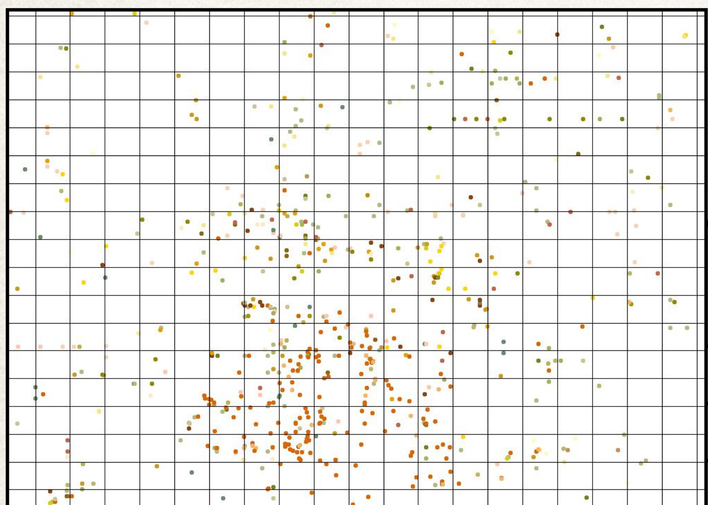
"Quality" (e.g. Phenotypic characteristic) versus "quantity"

Abiotic versus biotic



The data are represented in space and time – spatial data mining

A Democracy of the Data:
To infer interactions from where "things" are



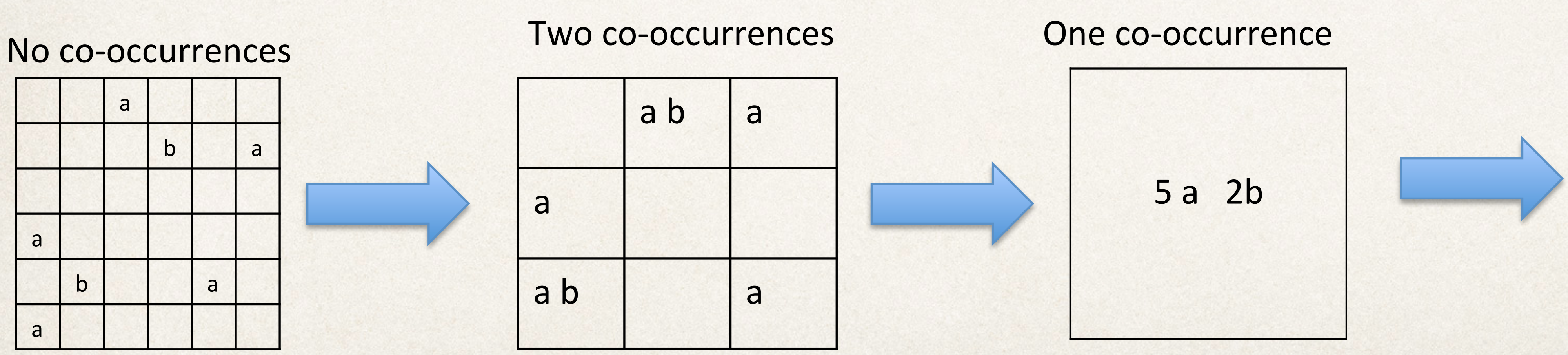
Choose a spatial resolution: give everyone one vote there.
The "Senate" versus the "Congress" approach!

Cuadrante	Sigmodon hispidus	Dipetalogaser maxima	Casos Chagas	Precipitación anual	Temperatura promedio	GARP Triatoma maximus	GARP Diptaloster maxima	Perfil agricola
A1	1	3	1	23	18.6	1	1	4
A2	0	1	0	23	18.6	1	1	4
A3	0	2	0	23.7	18.7	1	1	1
A4	0	4	0	23.7	18.7	1	1	3
A5	0	2	1	23.7	18.7	1	1	3
A6	2	5	2	23.7	18.7	1	1	2
A7	0	1	0	23.3	18.4	1	1	5
A8	0	2	0	22.8	18.8	1	1	3
A9	1	3	1	22.8	18.8	1	1	1
A10	0	1	0	22.8	18.8	0	1	1
A11	0	0	0	22.8	18.8	0	1	1
A12	0	0	0	22.8	18.8	0	1	2
A13	0	0	0	22.8	18.8	0	0	4
A14	0	0	0	22.8	18.8	0	0	3
A15	0	2	0	22.8	18.8	0	1	4
A16	0	1	0	22.8	18.8	0	1	2
A17	0	0	0	22.8	18.8	0	1	1
A18	0	0	0	22.8	18.8	0	0	1

Now we can make statistical inferences

In standard data mining, for example: P(death | age) = N(death,age)/N(age); P(death | diabetes); P(death | age,diabetes); to **infer** that age is a risk factor for death, as is diabetes. Here, we count individuals who have different traits. There is a preferred statistical unit - the individual within which we can look for coincidences/co-occurrences. In spatial data mining this is not the case.

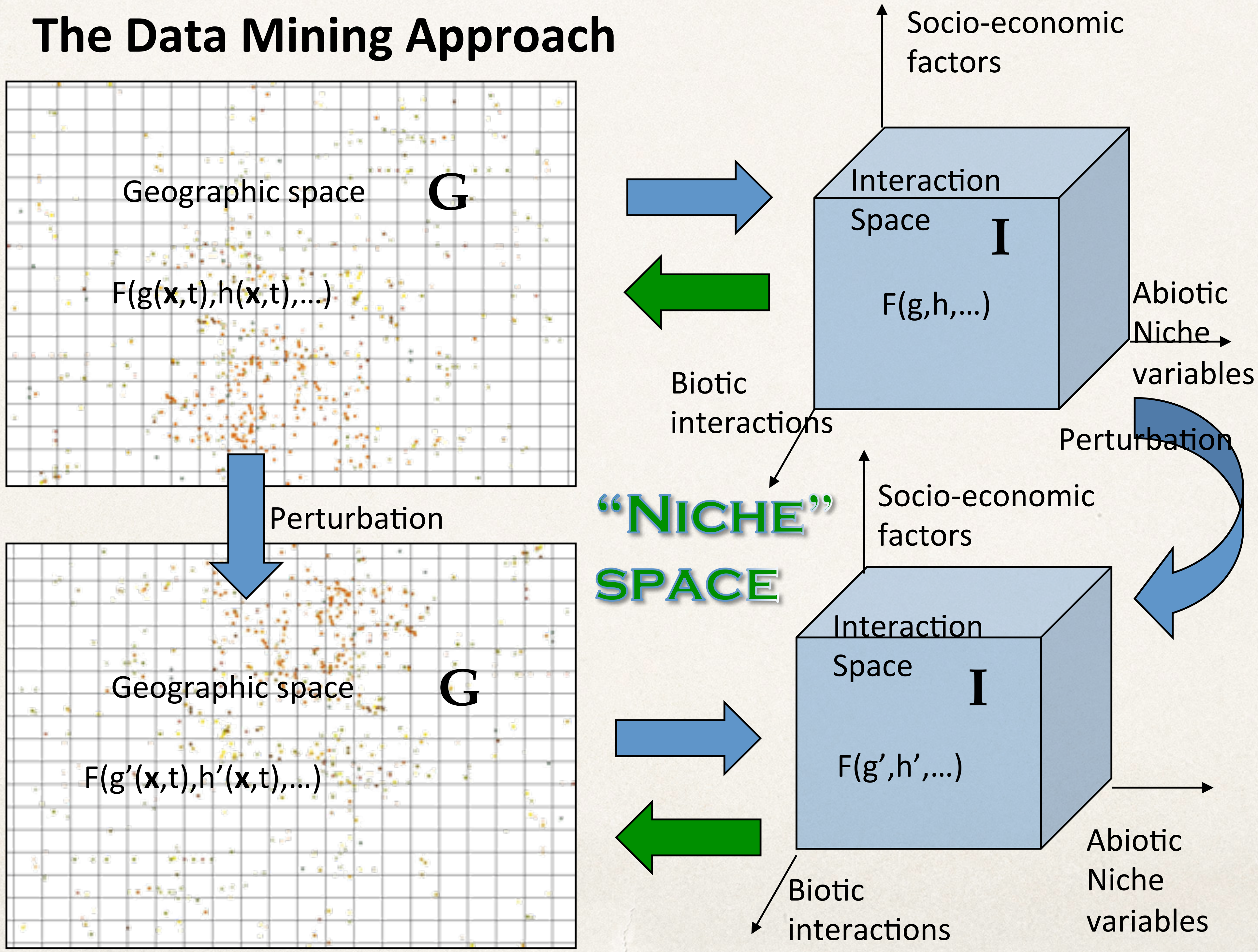
We must define coincidences/co-occurrences using an appropriate **uniform** spatio-temporal scale.



N(ab)/N(b) = P(a | b)

Dependence of species a on niche variable b

And we can pass to Niche Space: Or can we?



The Technical Part

For niche construction

P(**C** | **X**) = P(C | X1,X2,X3,…,XN)
= N(CX1,X2,X3,…,XN) / N(X1,X2,X3,…,XN)

But… N(CX1,X2,X3,…,XN) = 0, 1
the "curse of dimensionality"

Use Bayes' theorem
P(**C** | **X**) = P(**X** | C)P(C) / P(**X**)
and assume

$P_{NB}(\mathbf{X}|C) = \prod_{i=1}^{N} P(X_i|C)$

Naive Bayes Approximation
Total factorisation

$$P_{GB}(\mathbf{X}|C) = P(\xi^{(i)}|C) = \prod_{\alpha=1}^{N^{C}_{\xi^{(i)}}} P(\xi^{\alpha}|C)$$

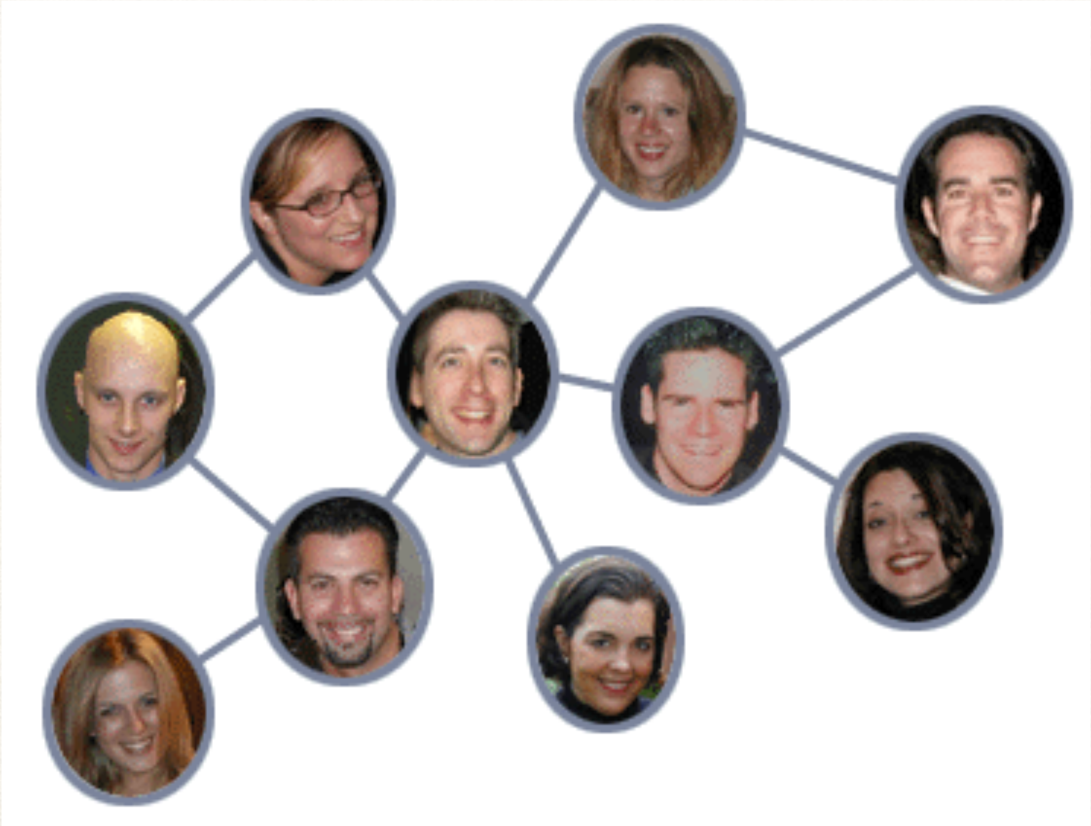
Generalised Bayes Approximation
Takes into account correlations

$$P_{GB}(\mathbf{X}|\bar{C}) = P(\xi^{(j)}|\bar{C}) = \prod_{\alpha=1}^{N^{\bar{C}}_{\xi^{(j)}}} P(\xi^{\alpha}|\bar{C})$$

Now for Communities…

You can judge a man by his "friends"

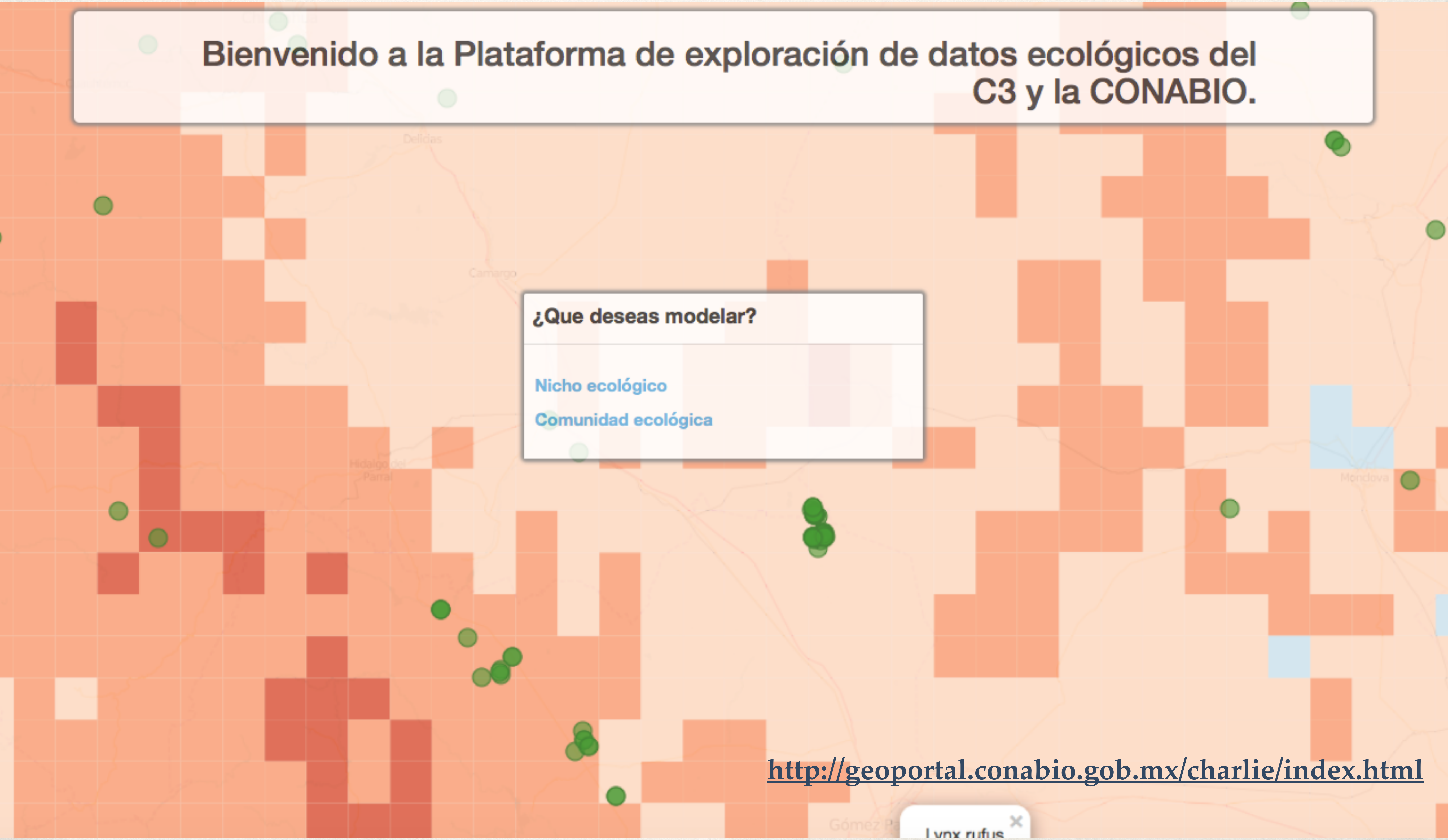
or his "enemies", or "parasites", or "prey" or "predators" or…



Use Complex Inferential Networks to Represent Community Interactions

- Take nodes to be…
  - Species, other taxonomic or phylogenetic groupings, groupings by phenotypic characteristics,
- Take links to be a statistical measure of spatial (temporal) co-occurrence
  - P(Y|X), epsilon(Y|X), P(A,B|C,D), epsilon(Z|X,Y)
  - What is a high/low degree of co-occurrence? (Choosing a null hypothesis)
  - What spatial (temporal) resolution? (When do things co-occur?)

and some results…



http://geoportal.conabio.gob.mx/charlie/index.html

Two Example Niches: Lutzomyia

TOP DECILE Optimal niche conditions for *Lutzomyia*				BOTTOM DECILE Suboptimal niche conditions for *Lutzomyia*			
ABIOTIC VARIABLES	RANGE	Epsilon	Score contribution	ABIOTIC VARIABLES	RANGE	Epsilon	Score contribut
BIO17	88-219	8.960	5.013	BIO12	42-507	-5.604	-2.279
BIO1	23.3-26.4	8.938	1.006	BIO16	18-218	-5.001	-2.328
BIO11	22.2-25.3	8.873	2.322	BIO18	1-249	-3.839	-3.799
BIO14	26-63	8.782	4.916	BIO6	3.1-3.4	-3.761	-2.931
BIO4	25.35-33.09	7.543	2.152	BIO7	26.3-28.4	-3.544	-8.853
BIO6	13.4-16.6	7.524	3.293	BIO2	16.5-18.4	-3.535	-2.997
BIO13	392-774	7.107	12.913	BIO11	2.9-12.5	-3.271	-4.482
BIO7	28.5-30.6	7.012	3.803	BIO4	3310-7184	-2.971	-9.551
BIO16	1019-2019	6.925	12.175	BIO19	192-383	-2.940	-0.448
BIO19	192-383	6.618	4.157	BIO10	28.9-32.3	-2.669	-0.916
BIO12	1906-3302	6.314	8.701	BIO1	10.3-19.9	-2.189	-1.033
BIO2	9.8-10.8	6.130	4.458	BIO3	3.7-5.5	-2.130	-3.576
BIO18	623-746	5.748	1.260	BIO8	28.4-31.7	-1.964	-0.731
RESERVOIRS				**RESERVOIRS**			
Reithrodontomys gracilis		8.892	2.640	*Sigmodon hispidus*		6.946	1.244
Heteromys gaumeri		8.800	2.234				
Heteromys desmarestianus		8.716	2.381				
Ototylomys phyllotis		7.559	2.028				
Peromyscus yucatanicus		7.249	2.116				
Sigmodon hispidus		6.946	1.244				
Didelphis marsupialis		5.774	1.662				
Oryzomys melanotis		3.494	1.387				
Marmosa mexicana		2.773	1.541				
LAND COVER				**LAND COVER**			
Cloud forest		6.642	1.408	Subtropical scrub		-1.675	-1.527
Tropical evergreen forest		6.603	4.476	Subtropical scrub with secondary vegetation		-1.849	-1.658
Cloud forest with secondary vegetation		6.028	1.459	Xeric scrub with secondary vegetation		-2.092	-3.640
Tropical evergreen forest with secondary vegetation		6.007	4.344	Xeric scrub		-2.924	-4.044
Agriculture areas		5.966	1.736	Mesquite		-3.337	-1.714
Human settlement		4.947	0.577	Grassland		-3.734	-1.874
Deciduous tropical forest with secondary vegetation		4.081	1.013	Mangroves		-4.063	-2.000

Two Example Niches: Lynx Rufus

TOP DECILE — Optimal niche conditions for L. rufus

ABIOTIC VARIABLES	RANGE	Epsilon	Score contribution
BIO1	-2.7 - 16.7	5.488	6.109
BIO6	-9.4 - 3.4	5.327	3.005
BIO8	2.2 - 14.7	4.797	1.096
BIO4	25.35- 48.95	4.704	1.393
BIO9	-3.5 - 16.4	4.687	5.758
BIO11	-3.6 - 16.5	4.632	7.050
BIO16	219 – 418	4.602	0.524
BIO5	7.7 - 30.5	4.330	1.777
BIO10	-2.7 – 22	4.266	2.33

LAND COVER	Epsilon	Score contribution
Grassland	4.883	0.629
Plantation forest	4.738	1.934
Xeric scrub with secondary vegetation	4.283	1.094
Oyamel forest	4.274	1.256
High mountain meadow	4.042	1.812
Agriculture areas	3.903	0.745
Xeric scrub	3.955	0.678
Coniferous forest	3.878	0.565
Quercus forest	3.858	0.475
Human establishment	3.661	0.356
Coniferous forest with secondary vegetation	3.631	0.591
Quercus forest with secondary vegetation	3.457	0.468

BOTTOM DECILE — Suboptimal niche conditions for L. rufus

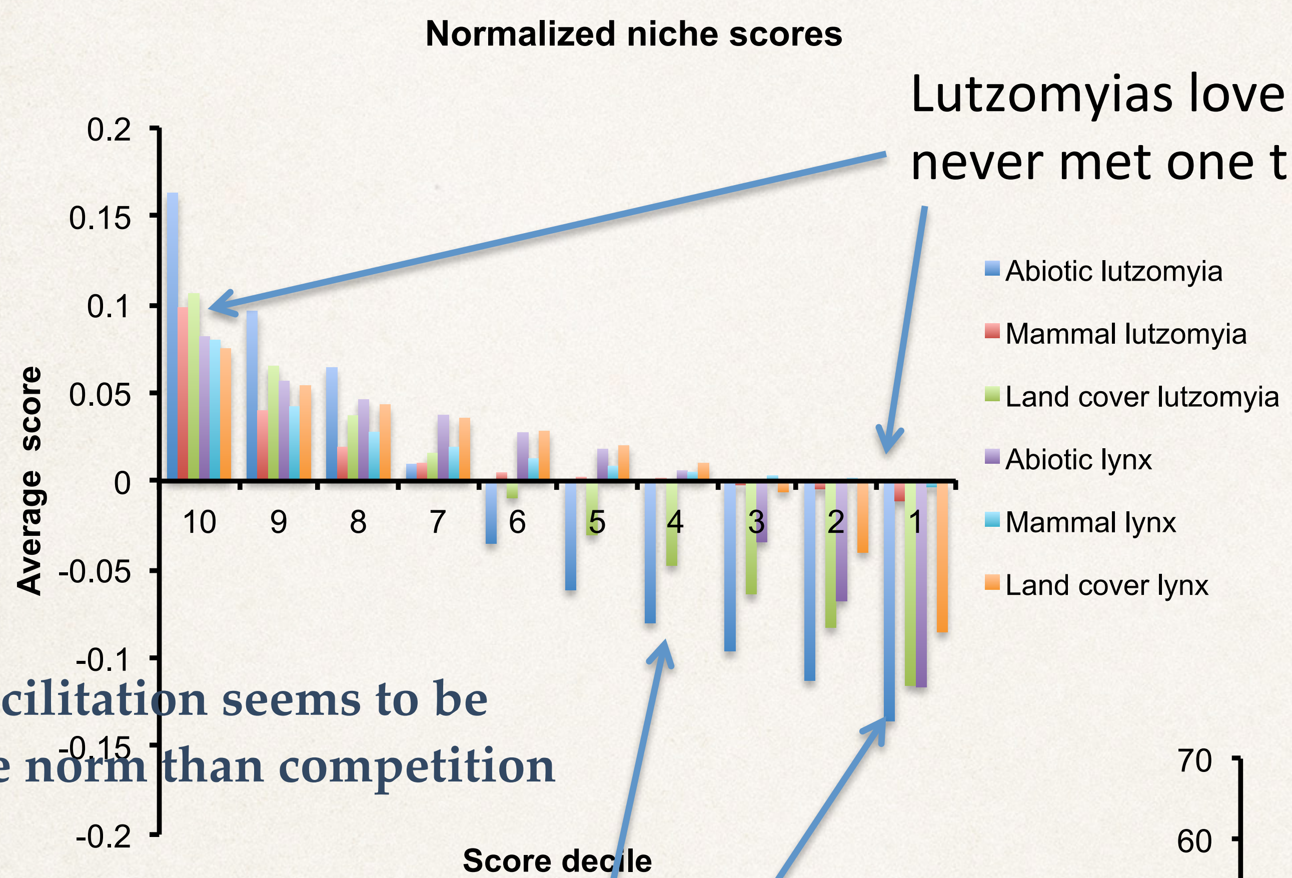
ABIOTIC VARIABLES	RANGE	Epsilon	Score contribution
BIO9	19.8 - 29.7	-4.177	-0.821
BIO11	19 - 28.6	-3.930	-5.379
BIO6	6.8 - 19.9	-3.578	-1.902
BIO1	23.3 - 29.7	-3.452	-3.128
BIO16	619 - 1618	-3.060	-3.268
BIO7	11.5 - 21.4	-2.853	-1.656
BIO17	88- 219	-2.782	-1.091
BIO2	7.3 - 11.9	-2.594	-0.954
BIO13	238 - 620	-2.59	-3.996
BIO12	974- 3302	-2.512	-1.413
BIO14	26-63	-2.253	-4.666
BIO18	374 - 870	-2.219	-1.068

LAND COVER	Epsilon	Score contribution
Low forest evergreen with secondary vegetation	-2.088	-0.430
Savannah	-2.202	-1.907
Agriculture areas	-2.245	-0.395
Cloud forest with secondary vegetation	-2.439	-2.061
Mangrove	-2.506	-1.191
Tropical evergreen forest with secondary vegetation	-2.540	-3.532
Tropical evergreen forest	-2.566	-3.575
Deciduous tropical forest	-2.924	-1.816
Deciduous tropical forest with secondary vegetation	-3.143	-2.471

PREYS		
Spermophilus variegatus	13.824	1.883
Sylvilagus floridanus	11.004	1.439
Neotoma albigula	9.143	1.604
Microtus mexicanus	8.846	1.776
Dipodomys ordii	8.636	1.565
Dipodomys merriami	8.618	1.306
Neotoma mexicana	8.034	1.378
Sigmodon leucotis	6.275	1.982
Sylvilagus audubonii	5.972	1.556
Didelphis virginiana	5.553	1.054
Cratogeomys merriami	5.385	2.031
Nasua narica	5.270	1.147
Dipodomys deserti	5.057	2.059
Dipodomys nelsoni	4.972	1.453
Odocoileus virginianus	4.457	1.589
Romerolagus diazi	4.427	4.362
Dipodomys gravipes	4.296	2.465
Dipodomys spectabilis	4.039	1.366
Neotomodon alstoni	3.860	1.589
Ammospermophilus harrisii	3.700	2.128
Dipodomys agilis	3.469	1.248
Spermophilus tereticaudus	2.332	1.366
Dipodomys simulans	1.875	1.877
Mustela frenata	1.810	0.928
Sylvilagus cunicularius	1.743	1.030
POTENTIAL COMPETITORS		
Leopardus pardalis	3.373	1.147
Panthera onca	2.559	0.928
Leopardus wiedii	1.597	0.735
Herpailurus yagouaroundi	1.138	0.524

PREYS		
Sylvilagus floridanus	11.004	1.439
Neotoma mexicana	8.034	1.378
Didelphis virginiana	5.553	1.054
Nasua narica	5.270	1.147
Odocoileus virginianus	4.457	1.589
POTENTIAL COMPETITORS		
Leopardus pardalis	3.373	1.147
Panthera onca	2.559	0.928
Leopardus wiedii	1.597	0.735
Herpailurus yagouaroundi	1.138	0.524

Two Example Niches

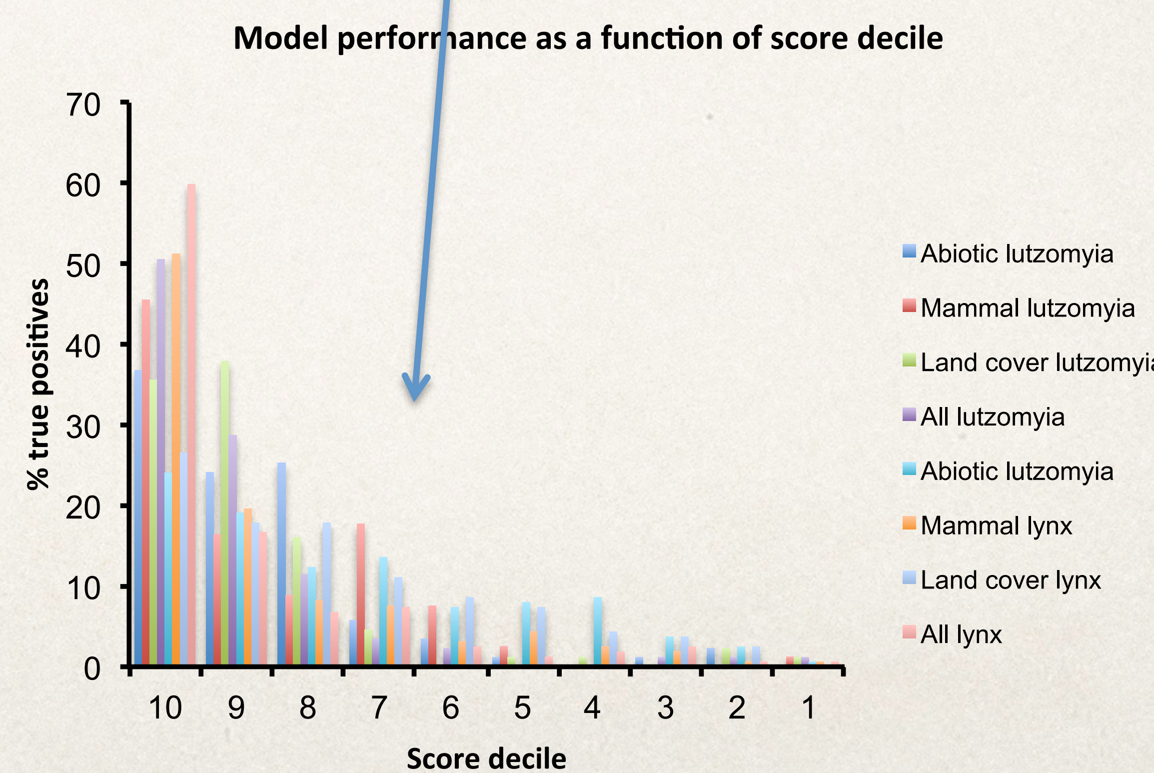


Lutzomyias love mammals, never met one they didn't like

Biotic facilitation seems to be more the norm than competition

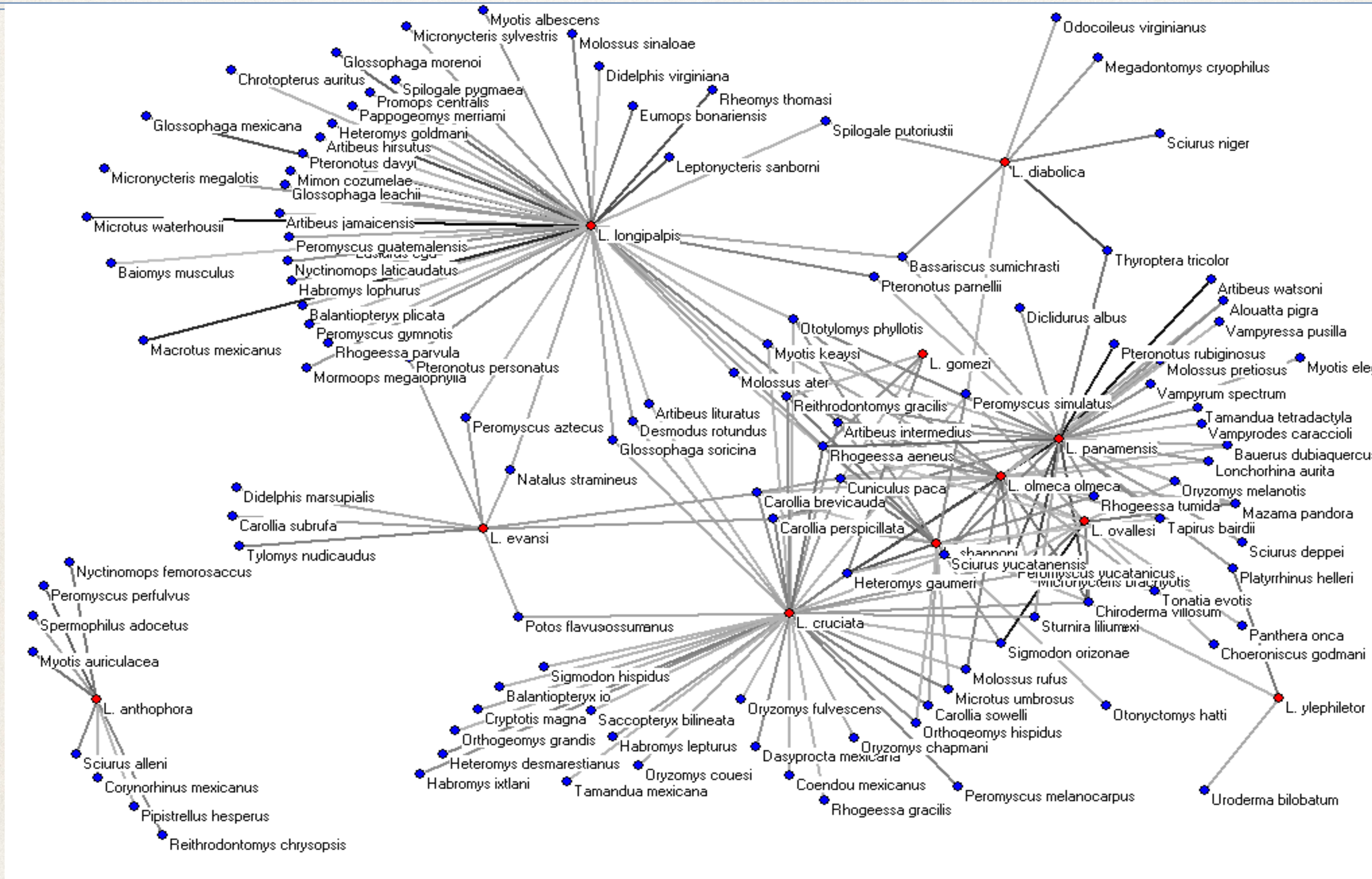
Climatic factors are more important for determining where Lutzomyias aren't rather than where they are

Including in a fuller, richer Niche Space leads to more predictive models (less false positives/negatives)



Chains of causality

The Ecology of Leishmaniasis



Species	ε	Negative	Positive	Total	% positive	Confidence (95%)
Carollia sowelli	8.83	43	2	45	4.4	-1 - 14
*Heteromys gaumeri**	8.8	5	0	5	0	-15 – 29
Peromyscus mexicanus	8.79	115	6	121	5	2 - 11
*Heteromys desmarestianus**	8.72	30	0	30	0	-2 - 16
Molossus rufus	8.63	1	0	1	0	-42 - 56
Glossophaga soricina	8.57	19	7	26	26.9	-3 - 16
Carollia perspicillata	8.5	8	0	8	0	-11 - 24
Pteronotus parnellii	8.16	4	0	4	0	-18 - 31
Desmodus rotundus	8.15	13	1	14	7.1	-6 - 20
Sturnira lilium	8.03	56	7	63	11.1	1 - 13
Artibeus phaeotis	8.01	35	1	36	2.8	-1 - 15
Oryzomys couesi	7.73	2	0	2	0	-28 - 41
*Ototylomys phyllotis**	7.56	9	1	10	10	-9 - 22
*Sigmodon hispidus**	7.28	36	4	40	10	-1 - 14
*Peromyscus yucatanicus**	7.25	3	0	3	0	-22 - 35
Didelphis virginiana	7.12	3	0	3	0	-22 - 30
Didelphis marsupialis	6.44	11	0	11	0	-8 - 21
Philander opossum	6.25	6	1	7	14.3	-12 - 25
Centurio senex	6.01	1	0	1	0	-42 - 56
Artibeus jamaicensis	5.98	81	5	86	5.8	1 - 12
Artibeus lituratus	5.84	38	3	41	7.3	-1 - 14
Myotis keaysi	5.61	2	0	2	0	-28 - 41
Chiroderma villosum	5.56	5	0	5	0	-15 - 29
Saccopteryx bilineata	5.3	1	0	1	0	-42 - 56
Sciurus aureogaster	5.23	71	8	79	7.3	1 - 12
Baiomys musculus	5.21	2	0	2	0	-28 - 41
Artibeus watsoni	5.13	2	0	2	0	-28 - 41
Choeroniscus godmani	5.05	10	3	13	23.1	-7 - 20
Pteronotus personatus	5.03	3	1	4	25	-18 - 31
Reithrodontomys mexicanus	4.91	1	0	1	0	-42 - 56
Oryzomys rostratus	4.87	22	1	23	4.3	-4 - 17
Micronycteris microtis	4.23	1	0	1	0	-42 - 56
Oligoryzomys fulvescens	4.2	6	0	6	0	-13 - 27
Peromyscus leucopus	3.8	22	4	26	15.4	-3 - 16
Sturnira ludovici	3.79	24	1	25	4	-3 - 17
Vampyrodes caraccioli	3.69	1	0	1	0	-42 - 56
Liomys pictus	3.61	47	1	48	2.1	0 - 14
Glossophaga commissarisi	3.49	2	6	8	75	-11 - 24
Lonchorhina aurita	3.48	1	0	1	0	-42 - 56
Phyllostomus discolor	3.48	0	1	1	100	-42 - 56
Platyrrhinus helleri	3.36	5	0	5	0	-22 - 35
Uroderma bilobatum	3.34	4	0	4	0	-18 - 31
Urocyon cinereoargenteus	2.97	1	0	1	0	-42 - 56
Procyon lotor	2.95	1	0	1	0	-42 - 56
Myotis velifer	2.58	3	0	3	0	-18 - 31
Microtus mexicanus	2.53	16	0	16	0	-6 - 19
Myotis nigricans	2.47	2	0	2	0	-28 - 41
Leptonycteris yerbabuenae	2.43	1	1	2	50	-28 - 41
Reithrodontomys fulvescens	2.08	20	0	20	0	-4 - 18
Neotoma mexicana	1.99	5	0	5	0	-15 - 29
Eptesicus fuscus	1.82	1	0	1	0	-42 - 56
Peromyscus levipes	1.34	1	0	1	0	-42 - 56
Sorex saussurei	1.29	3	0	3	0	-22 - 35
Osgoodomys banderanus	1.21	9	0	9	0	-10 - 23
Liomys irroratus	1.16	8	0	8	0	-11 - 24
Myotis auriculus	0.22	2	0	2	0	-28 - 41
Tadarida brasiliensis	-0.09	1	0	1	0	-42 - 56
Peromyscus hylocetes	-0.28	2	0	2	0	-28 - 41
Antrozous pallidus	-0.34	1	0	1	0	-42 - 56
Peromyscus zarhynchus	-0.46	2	0	2	0	-28 - 41
Chaetodipus hispidus	-0.71	4	0	4	0	-18 - 31
Peromyscus pectoralis	-0.73	2	0	2	0	-28 - 41
Neotomodon alstoni	-0.9	17	0	17	0	-5 - 19
Baiomys taylori	-1.16	10	3	13	23.1	-7 - 20
Chaetodipus nelsoni	-1.24	3	0	3	0	-22 - 35
Neotoma micropus	-1.27	16	0	16	0	-6 - 19
Peromyscus maniculatus	-1.37	58	2	60	3.3	0 - 13
Peromyscus eremicus	-1.41	0	1	1	100	-42 - 56
Perognathus flavus	-1.52	1	0	1	0	-42 - 56
Dipodomys merriami	-2.01	1	0	1	0	-42 - 56

Formatte

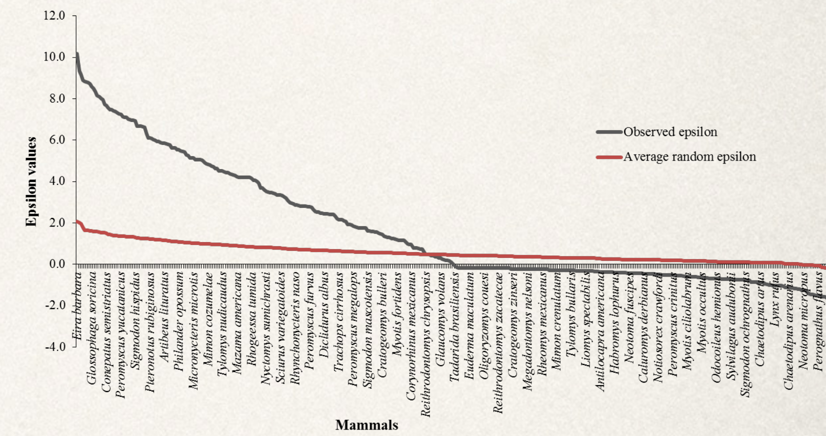
- Only about 50 (2.5%) of mammals on the American continent have been identified as hosts of Leishmania
- In Mexico only 8 out of 419 (2.1%) had been identified as hosts
- We collected 922 individuals from 70 species
- Predicted and confirmed 21 new species of mammal as carriers of Leishmania in Mexico
- 13 of them are bats, identified for the first time in Mexico
- Squirrels identified as carriers
- 33% of collected species were confirmed as hosts
- Overall infection rate was 6.7%
- No species could be rejected as a host at this infection rate at the 95% confidence level
- Changes the picture for control of Leishmania totally;
- Leishmania and Lutzomyias are eclectic in their host source.
- Linnean classification is NOT ecologically relevant

Prediction at the Ecosystemic Level: Disease reservoirs

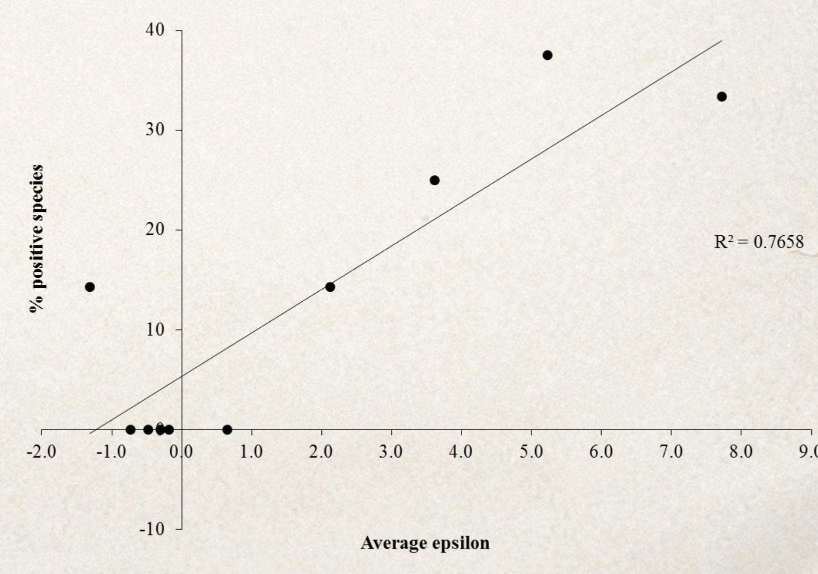
	Mammals	Epsilon	Conf.
1	Eira barbara	10.1683	
2	Rhogeessa aeneus	9.3649	
3	Artibeus intermedius	9.1628	
4	Reithrodontomys gracilis	8.8921	Yes
5	Carollia sowelli	8.8303	
6	Heteromys gaumeri	8.8000	Yes
7	Peromyscus mexicanus	8.7859	
8	Heteromys desmarestianu	8.7164	Yes
9	Molossus rufus	8.6277	
10	Glossophaga soricina	8.5713	
11	Carollia perspicillata	8.5030	
12	Orthogeomys hispidus	8.3468	
13	Pteronotus parnellii	8.1632	
14	Desmodus rotundus	8.1519	
15	Dasyprocta mexicana	8.1128	
16	Sturnira lilium	8.0290	
17	Dermanura phaeotis	8.0055	
18	Dasyprocta punctata	7.9678	
19	Oryzomys couesi	7.7253	
20	Potos flavus	7.7246	
21	Conepatus semistriatus	7.6879	
22	Ototylomys phyllotis	7.5587	Yes
23	Ateles geoffroyi	7.4787	
24	Cryptotis magna	7.4207	
25	Cuniculus paca	7.3220	
26	Lampronycteris brachyotis	7.2852	
27	Sigmodon hispidus	7.2805	Yes
28	Peromyscus yucatanicus	7.2486	Yes
29	Oryzomys chapmani	7.1242	
30	Didelphis virginiana	7.1150	
31	Peromyscus melanocarpus	7.0260	
32	Microtus umbrosus	6.9630	
33	Thyroptera tricolor	6.9630	
34	Nasua narica	6.8953	
35	Megadontomys cryophilus	6.6830	
36	Oryzomys alfaroi	6.6816	
37	Sorex veraepacis	6.6797	
38	Carollia subrufa	6.6316	
39	Peromyscus aztecus	6.6173	
40	Didelphis marsupialis	6.4390	Yes
41	Sciurus yucatanensis	6.3865	
42	Philander opossum	6.2546	
43	Habromys ixtlani	6.1120	
44	Microtus waterhousii	6.1120	
45	Pteronotus rubiginosus	6.1120	
46	Reithrodontomys microdor	6.0967	
47	Coendou mexicanus	6.0268	
48	Centurio senex	6.0076	
49	Artibeus jamaicensis	5.9786	
50	Glossophaga morenoi	5.8847	

	Mammals	Epsilon	Conf.
51	Molossus sinaloae	5.8518	
52	Artibeus lituratus	5.8422	
53	Mormoops megalophylla	5.8374	
54	Habromys lepturus	5.7848	
55	Myotis keaysi	5.6148	
56	Chiroderma villosum	5.5562	
57	Tamandua mexicana	5.4845	
58	Tylomys nudicaudus	5.4510	
59	Saccopteryx bilineata	5.2984	
60	Macrotus mexicanus	5.2472	
61	Sciurus aureogaster	5.2267	
62	Baiomys musculus	5.2092	
63	Rhogeessa tumida	5.1950	
64	Sciurus deppei	5.1414	
65	Dermanura watsoni	5.1338	
66	Otonyctomys hatti	5.1338	
67	Orthogeomys grandis	5.0556	
68	Alouatta palliata	5.0457	
69	Choeroniscus godmani	5.0457	
70	Peropteryx macrotis	5.0457	
71	Pteronotus personatus	5.0266	
72	Lontra longicaudis	4.9330	
73	Reithrodontomys mexicanu	4.9120	
74	Oryzomys rostratus	4.8681	
75	Mimon cozumelae	4.8327	
76	Pteronotus davyi	4.7943	
77	Herpailurus yagouaroundi	4.7100	
78	Glossophaga leachii	4.6849	
79	Rhogeessa gracilis	4.6317	
80	Sylvilagus brasiliensis	4.6317	
81	Hodomys alleni	4.5155	
82	Leopardus wiedii	4.4420	
83	Peromyscus simulatus	4.4195	
84	Sigmodon alleni	4.3707	
85	Bassariscus sumichrasti	4.3110	
86	Oryzomys fulvescens	4.3110	
87	Diphylla ecaudata	4.3013	
88	Oryzomys melanotis	4.2907	Yes
89	Micronycteris microtis	4.2338	
90	Mazama americana	4.2274	
91	Microtus oaxacensis	4.2061	
92	Rheomys thomasi	4.2061	
93	Oryzomys saturatior	4.2061	
94	Myotis elegans	4.2024	
95	Oligoryzomys fulvescens	4.1984	
96	Natalus stramineus	4.0626	
97	Balantiopteryx io	4.0522	
98	Nyctinomops laticaudatus	4.0522	
99	Tlacuatzin canescens	4.0119	
100	Odocoileus virginianus	3.9265	

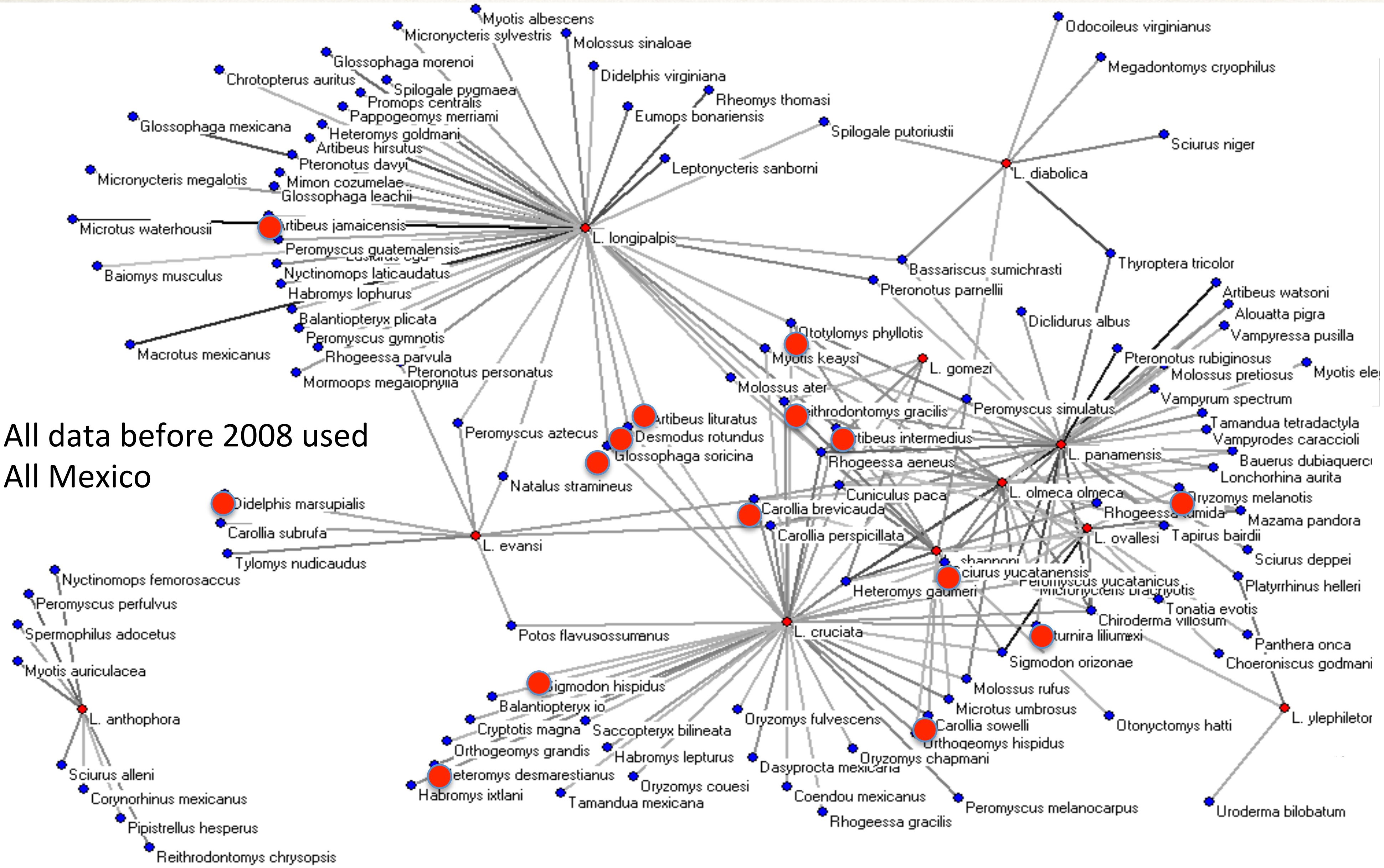
	Mammals	Epsilon	Conf.
101	Balantiopteryx plicata	3.8590	
102	Peromyscus leucopus	3.7994	
103	Sturnina ludovici	3.7888	
104	Enchisthenes hartii	3.6929	
105	Vampyrodes caraccioli	3.6929	
106	Eptesicus furinalis	3.6453	
107	Liomys pictus	3.6107	
108	Glossophaga commissaris	3.4861	
109	Lonchorhina aurita	3.4781	
110	Phyllostomus discolor	3.4781	
111	Peromyscus gymnotis	3.4516	
112	Anoura geoffroyi	3.4201	
113	Platyrrhinus helleri	3.3586	
114	Eumops bonariensis	3.3398	
115	Sciurus variegatoides	3.3398	
116	Uroderma bilobatum	3.3373	
117	Lasiurus intermedius	3.2197	
118	Lasiurus ega	3.1739	
119	Peromyscus megalops	3.1410	
120	Eumops glaucinus	3.0564	
121	Urocyon cinereoargenteus	2.9697	
122	Procyon lotor	2.9502	
123	Hylonycteris underwoodi	2.9343	
124	Rhynchonycteris naso	2.8580	
125	Eptesicus brasiliensis	2.8106	
126	Myotis albescens	2.8106	
127	Lophostoma evotis	2.8106	
128	Tapirus bairdii	2.8106	
129	Vampyrum spectrum	2.8106	
130	Marmosa mexicana	2.7731	Yes
131	Peromyscus furvus	2.7731	
132	Myotis velifera	2.5757	
133	Spilogale putorius	2.5411	
134	Microtus mexicanus	2.5268	
135	Dasypus novemcinctus	2.4725	
136	Myotis nigricans	2.4704	
137	Lophostoma brasiliense	2.4407	
138	Diclidurus albus	2.4407	
139	Sciurus niger	2.4407	
140	Leptonycteris curasoae	2.4268	
141	Nyctomys sumichrasti	2.4026	
142	Sigmodon mascotensis	2.3815	
143	Alouatta pigra	2.3374	
144	Peromyscus melanophrys	2.2204	
145	Dermanura tolteca	2.1920	
146	Trachops cirrhosus	2.1663	
147	Bauerus dubiaquercus	2.1612	
148	Spilogale pygmaea	2.1612	
149	Leptonycteris nivalis	2.1402	
150	Sylvilagus floridanus	2.1002	



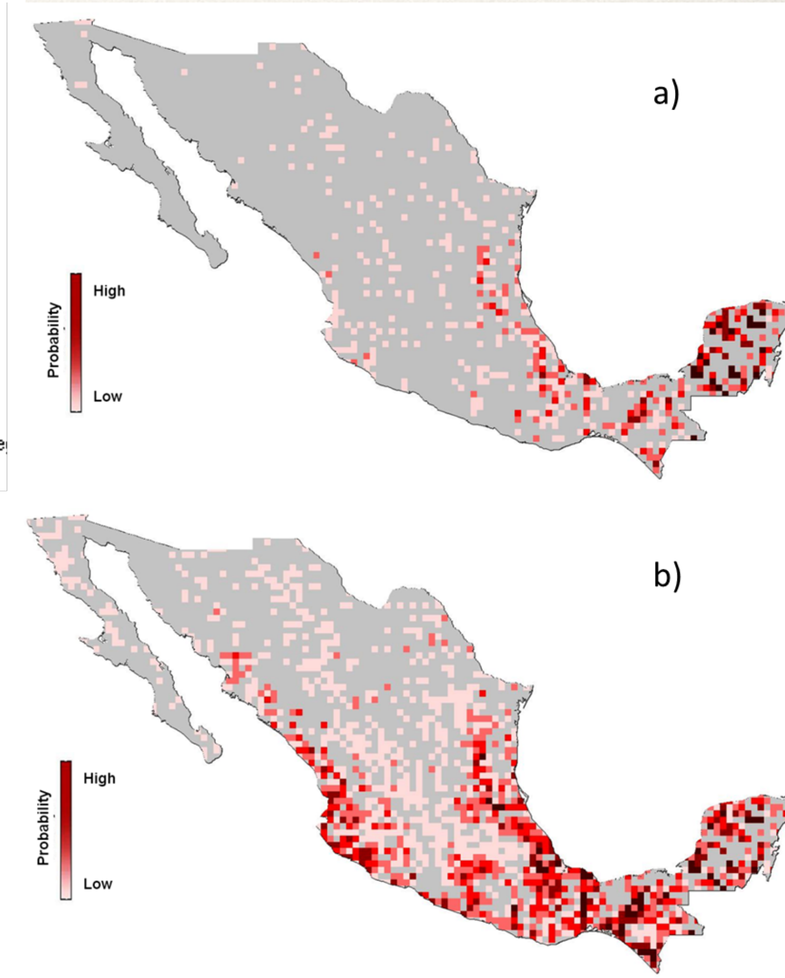
Biotic facilitation seems to be the norm. Species are not distributed randomly



The Ecology of Leishmaniasis

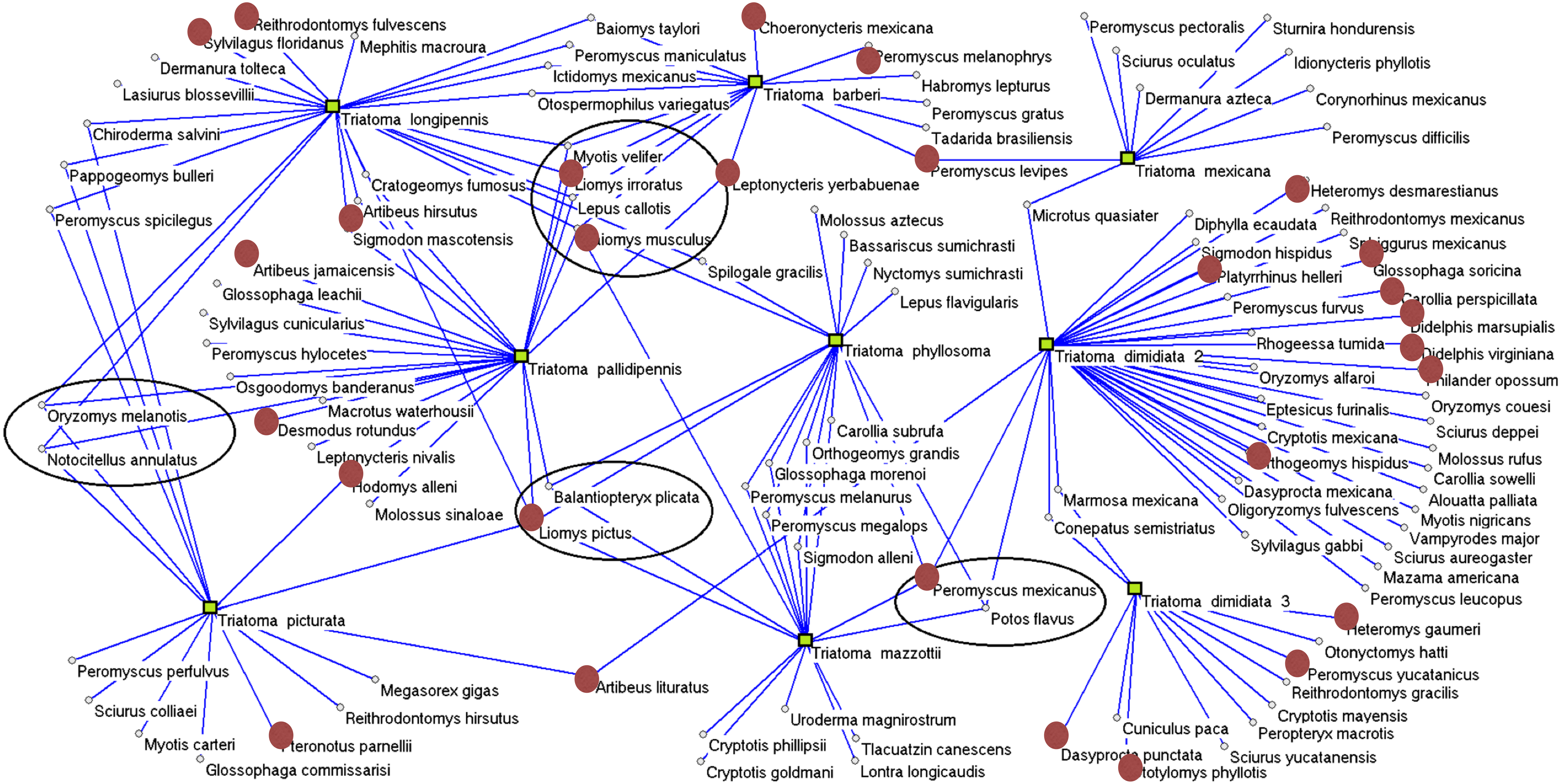


All data before 2008 used
All Mexico



What does this tell us about vector control?

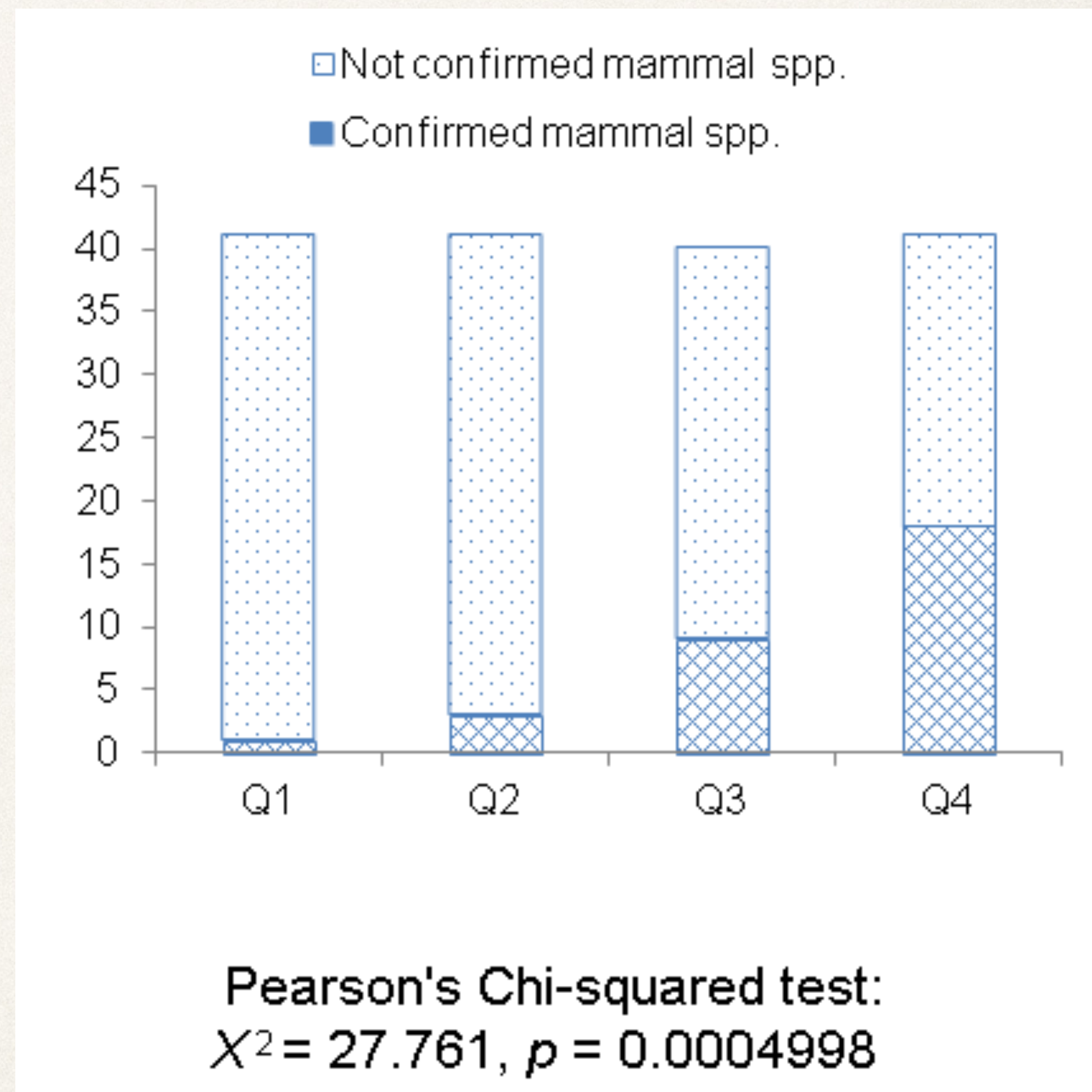
The Ecology of Chagas



The Ecology of Chagas

CONFIRMED MAMMAL	Q	ε2
Baiomys musculus[a,b,c,d]	4	12.63
Liomys irroratus[a,b,c,d,e]	4	11.20
Artibeus jamaicensis[a,b]	4	10.57
Glossophaga soricina[a]	4	10.02
Desmodus rotundus[b]	4	9.91
Peromyscus mexicanus[f]	4	9.76
Didelphis virginiana[b,e,f,g]	4	9.76
Leptonycteris yerbabuenae (curasoae)[b]	4	8.91
Sturnira lilium[a,b]	4	8.64
Orthogeomys hispidus[h]	4	7.75
Pteronotus parnellii[a,b]	4	7.60
Reithrodontomys fulvescens[i]	4	7.52
Sigmodon hispidus[c,d,j]	4	7.01
Didelphis marsupialis[e,h,j]	4	6.60
Carollia perspicillata[j]	4	6.59
Nasua narica[k]	4	6.45
Peromyscus leucopus[h]	4	6.36
Sigmodon mascotensis[e]	4	6.33
Tylomys nudicaudus[j]	3	6.07
Choeronycteris mexicana[a]	3	6.06
Peromyscus melanophrys[b]	3	5.75
Philander opossum[e,j]	3	5.74
Mephitis macroura[e]	3	5.59
Peromyscus levipes[c,d]	3	5.26
Dasypus novemcinctus[i,j]	3	4.82
Procyon lotor[j,k]	3	4.26
Hodomys alleni[l]	3	3.74
Sylvilagus floridanus[h]	2	3.50
Urocyon cinereoargenteus[h]	2	3.42
Heteromys desmarestianus[f]	2	3.21
Neotoma mexicana[a,c]	1	2.64
Dasyprocta punctata[h]	-	NS
Heteromys gaumeri[h]	-	NS
Lynx rufus[i]	-	NS
Neotoma micropus[j]	-	NS
Otospermophilus (Spermophilus) variegatus[b]	-	NS
Ototylomys phyllotis[h,j]	-	NS
Peromyscus yucatanicus[h]	-	NS
Spilogale angustifrons (putorius)[h]	-	NS

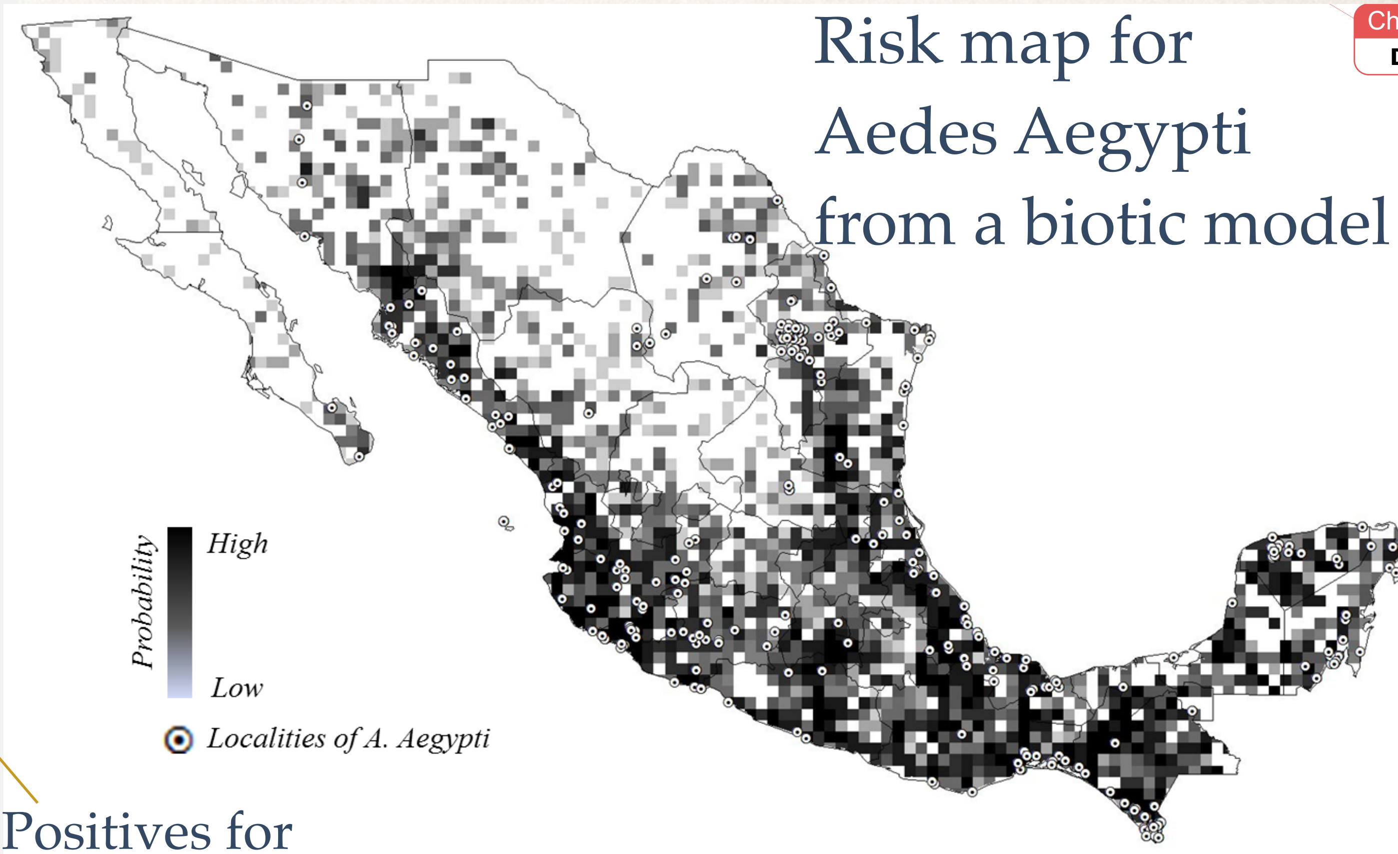
signifies also a confirmed host for Leishmania



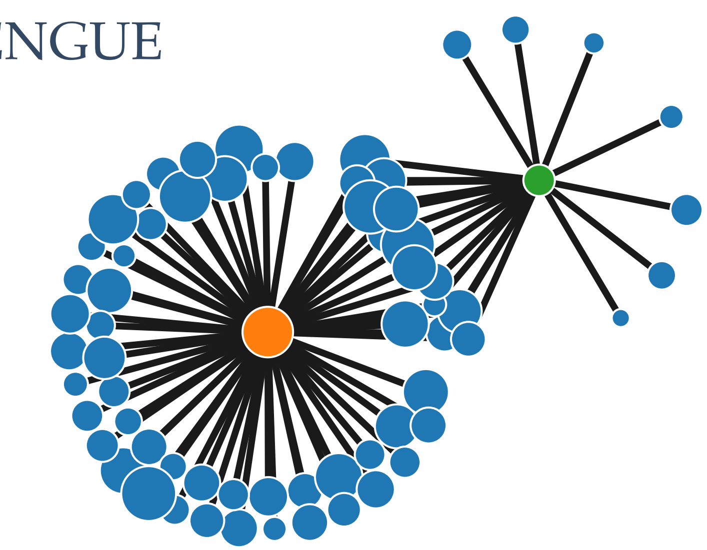
La Ecología de Dengue/CHIKV/ZIKV

Rank	Mammal	epsilon	Rank	Mammal	epsilon
1	*Glossophaga soricina*	12.78	38	*Dasypus novemcinctus*	7.11
2	*Molossus rufus*	11.99	39	*Sigmodon hispidus*	7.02
3	*Artibeus jamaicensis**	11.68	40	*Uroderma bilobatum*	6.82
4	*Liomys pictus*	11.06	41	*Leptonycteris curasoae*	6.75
5	*Oryzomys couesi*	11.04	42	*Carollia perspicillata*	6.71
6	*Carollia subrufa*	10.49	43	*Centurio senex*	6.61
7	*Sturnira lilium*	10.28	44	*Sciurus colliaei*	6.59
8	*Artibeus lituratus**	9.91	45	*Lontra longicaudis*	6.49
9	*Choeroniscus godmani*	9.42	46	*Didelphis marsupialis*	6.49
10	*Liomys salvini*	9.33	47	*Cratogeomys bulleri*	6.35
11	*Oligoryzomys fulvescens*	9.15	48	*Carollia sowelli**	6.27
12	*Dermanura phaeotis*	9.12	49	*Myotis elegans*	6.12
13	*Rhogeessa tumida*	9.06	50	*Myotis nigricans**	6.06
14	*Pteronotus personatus*	9.05	51	*Sigmodon arizonae*	6.00
15	*Baiomys musculus*	8.97	52	*Rhynchonycteris naso*	5.95
16	*Glossophaga commissarisi*	8.80	53	*Tlacuatzin canescens*	5.87
17	*Didelphis virginiana*	8.58	54	*Leopardus pardalis*	5.84
18	*Pteronotus parnellii**	8.58	55	*Caluromys derbianus*	5.78
19	*Orthogeomys hispidus*	8.53	56	*Molossus molossus*	5.76
20	*Sciurus aureogaster*	8.52	57	*Oryzomys rostratus*	5.76
21	*Molossus sinaloae*	8.51	58	*Osgoodomys banderanus*	5.76
22	*Desmodus rotundus*	8.23	59	*Myotis carteri*	5.66
23	*Saccopteryx bilineata*	8.22	60	*Micronycteris microtis*	5.52
24	*Lasiurus intermedius*	8.15	61	*Sylvilagus brasiliensis*	5.47
25	*Phyllostomus discolor*	8.12	62	*Sylvilagus floridanus*	5.37
26	*Philander opossum*	8.10	63	*Spermophilus annulatus*	5.36
27	*Peromyscus gymnotis*	7.90	64	*Peromyscus leucopus*	5.30
28	*Balantiopteryx plicata*	7.81	65	*Conepatus leuconotus*	5.30
29	*Eptesicus furinalis*	7.69	66	*Chaetodipus pernix*	5.27
30	*Pteronotus davyi*	7.55	67	*Sciurus yucatanensis*	5.23
31	*Dermanura tolteca*	7.48	68	*Sigmodon mascotensis*	5.13
32	*Sciurus variegatoides*	7.48	69	*Eira barbara*	5.12
33	*Mormoops megalophylla*	7.45	70	*Ateles geoffroyi*	5.11
34	*Oryzomys melanotis*	7.42	71	*Neotoma phenax*	5.07
35	*Artibeus intermedius*	7.40	72	*Noctilio leporinus*	5.06
36	*Chaetodipus artus*	7.20	73	*Reithrodontomys fulvescens*	4.95
37	*Nasua narica*	7.18			

Positives for DENGUE



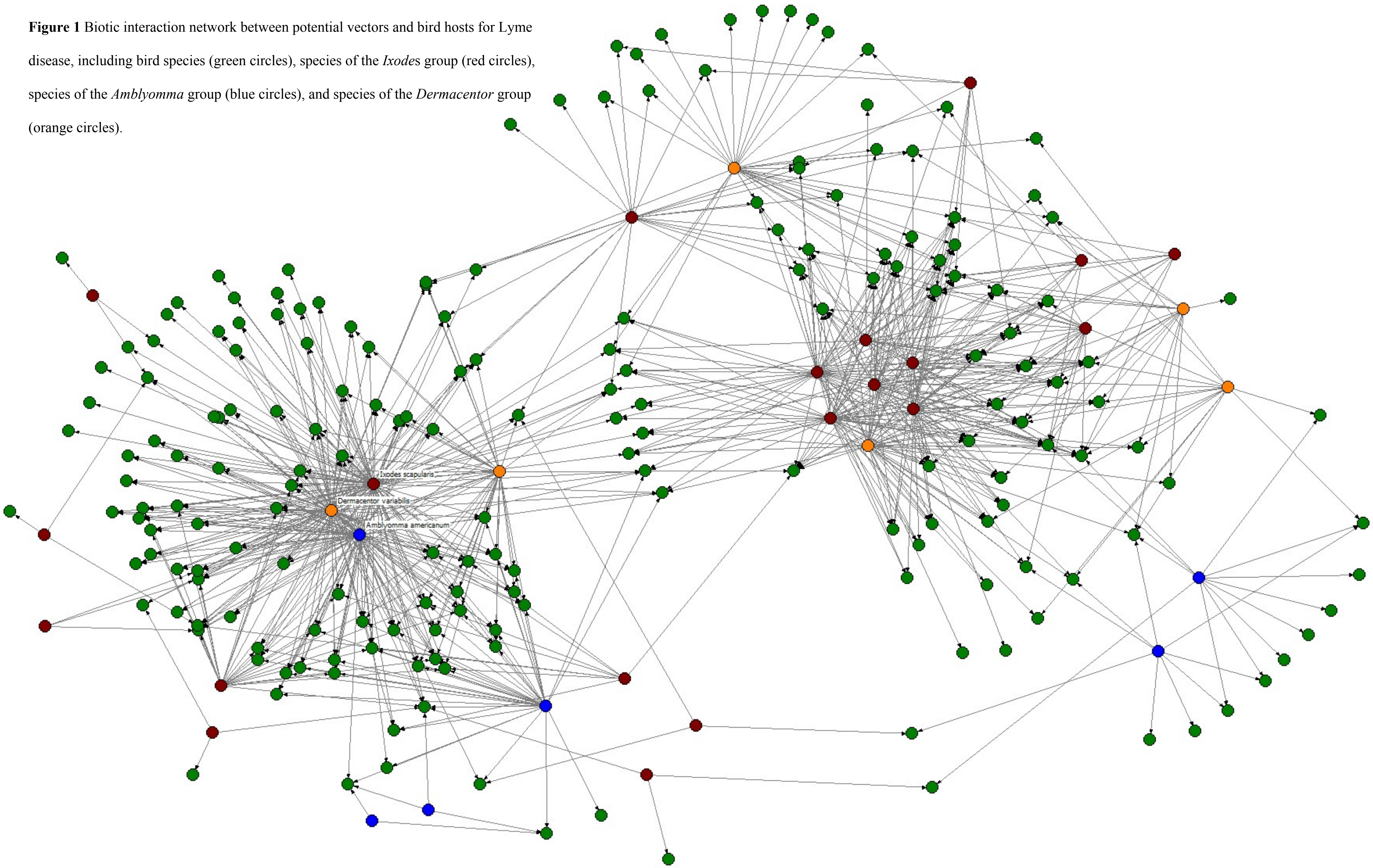
Risk map for Aedes Aegypti from a biotic model



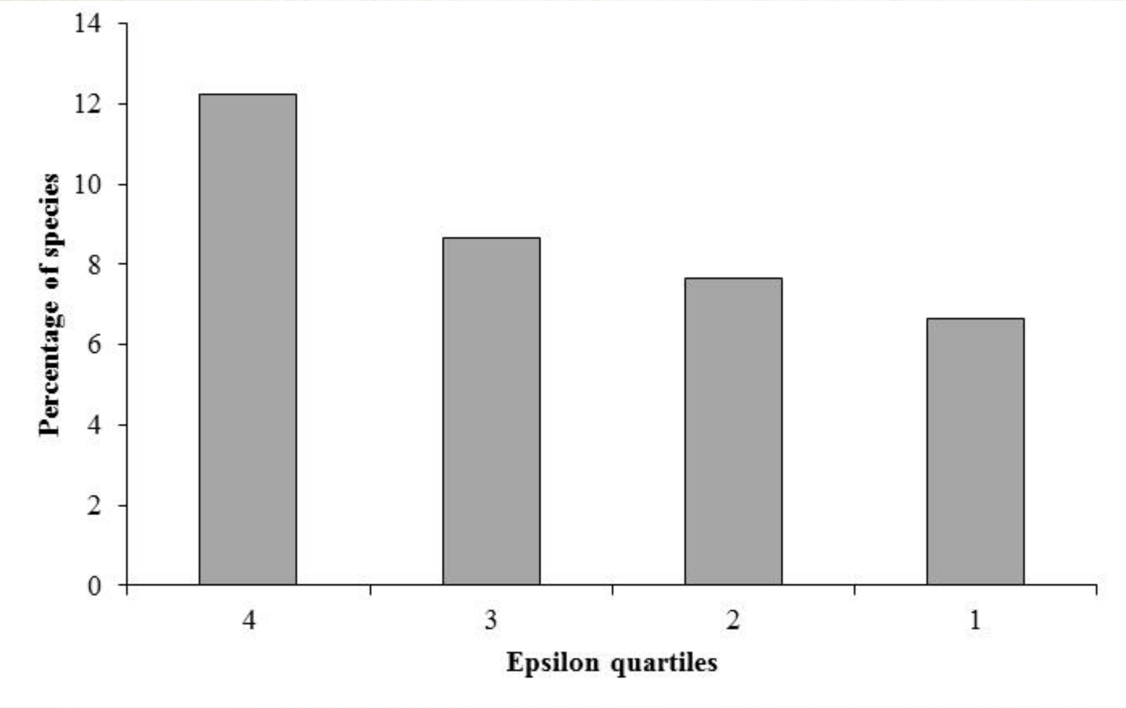
Complex Inference Network for Aedes aegypti and Aedes albopictus

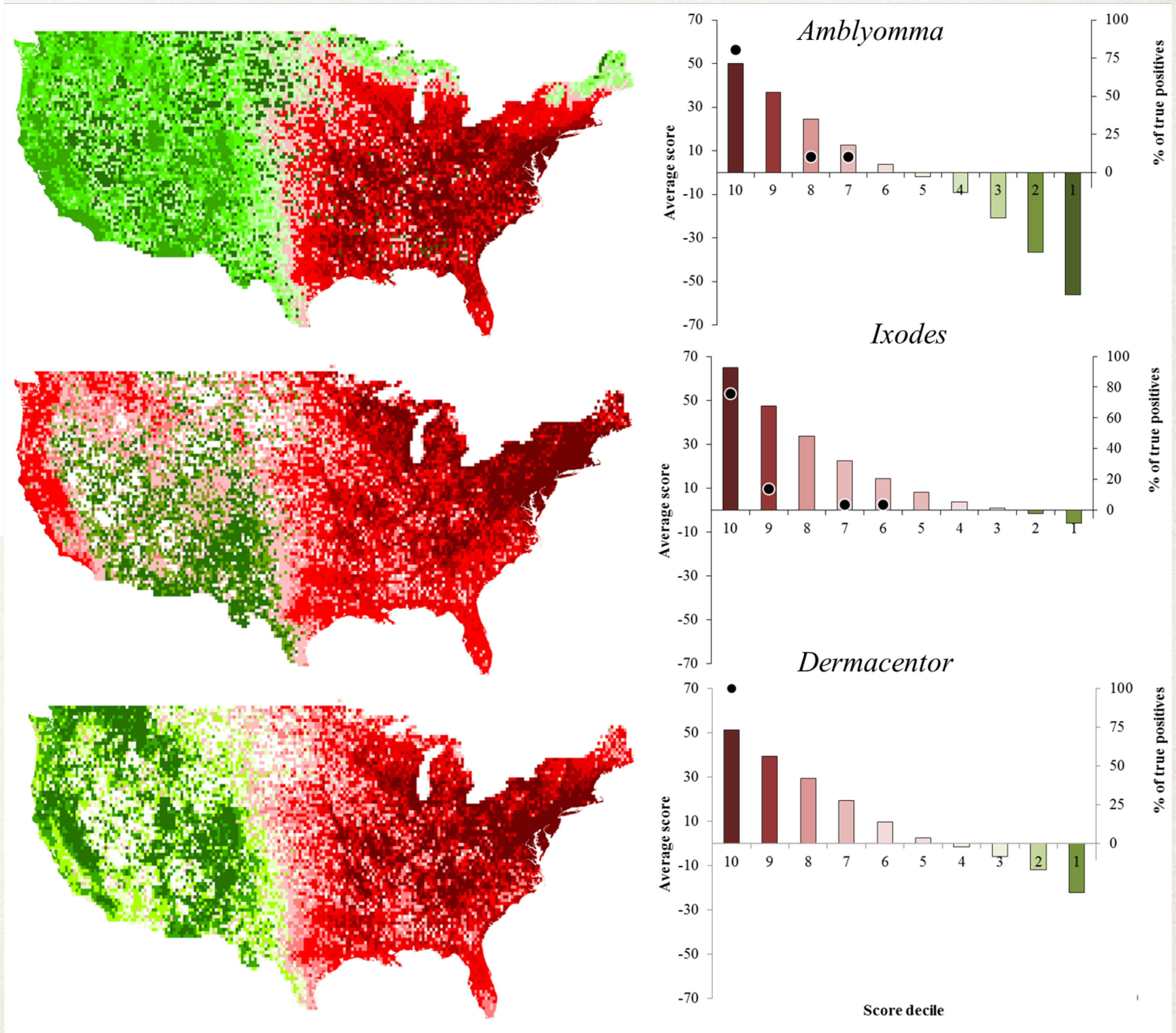
The Ecology of Lyme

Figure 1 Biotic interaction network between potential vectors and bird hosts for Lyme disease, including bird species (green circles), species of the *Ixodes* group (red circles), species of the *Amblyomma* group (blue circles), and species of the *Dermacentor* group (orange circles).



The Ecology of Lyme





Conclusions

- Ecologies are Complex Adaptive Systems - highly multifactorial
  - There are far too many relevant ecological interactions associated with a zoonosis to be experimentally identified and quantified
- Spatio-temporal data about organisms, relative to each other (biotic) and relative to the environment (abiotic), can be used to deduce the nature of the interactions between them and with the environment
  - This can be done at the niche level (one to many) and at the community level (many to many)
  - Our formalism allows for the incorporation of any data type, data format and data resolution
- The labels for "things" - family, genus, species, DTU, population, sylvatic, peri-domestic, competent, guild, cases, biomarkers etc. - allow us to detect heterogeneity in the interactions
  - The more labels we have the more we can compare different hypotheses
  - More labels means more data: clinical cases, phylogenetic,…
  - Data has to be incorporated in the SPECIES platform

Conclusions

- The Niche "fitness" landscape of a taxon C can be characterised quantitatively by $P(C | \mathbf{X})$ using spatio-temporal data mining
  - What are their general topological and geometrical characterisations?
  - How rugged/smooth are they?
  - What is the distribution of epistasis
    - Are distributions random?
    - Facilitation versus competition
  - What are the right coordinates?
  - What is the dynamics of Niche landscapes? How do they evolve?
  - How do we determine and characterise causal chains in ecology?
- At the community level, spatio-temporal data can be used to construct Complex Inference Networks (CIN) as representations of ecosystems
  - How to distinguish causality from correlation?
  - How to determine co-dependencies?
- The niches and community relations of diseases can be determined via CIN
  - Identification of transmission cycles and host range
    - Leishmania, Chagas, Lyme, Dengue, Zika, West Nile,…
  - Many zoonoses are multi-host, multi-vector, multi-pathogen systems.

Grupo de Trabajo

C3 - Centro de Ciencias de la Complejidad, UNAM; Instituto de Biología, UNAM; CONABIO; Universidad Catolica de Chile; Facultad de Medicina, UNAM

1.- Dr. Christopher R. Stephens
2.- Dr. Raúl Sierra Alcocer
4.- Dr. Constantino González Salazar
5.- M. en C. Enrique del Callejo
6.- M. en C. Ana Ruiz
7.- Dr. Víctor Sánchez-Cordero
8.- Dr. Ángel Rodríguez Moreno
9.- Dra. Inge Becker
10.- Dr. Gabriel Granados Gutiérrez
11.- Dr. Carlos Napoleón Ibarra Cerdeña
12.- M. en C. Laura Rengifo
13.- Dr. Pablo Marquet

Publications

Competitive interactions between felid species may limit the southern distribution of bobcats Lynx rufus
V Sánchez-Cordero, D Stockwell, S Sarkar, H Liu, CR Stephens, ...
Ecography 31 (6), 757-764, 2008

Using biotic interaction networks for prediction in biodiversity and emerging diseases
CR Stephens, JG Heau, C González, CN Ibarra-Cerdeña, ...
PLoS One 4 (5), e5725, 2009

Exploratory analysis of the interrelations between co-located boolean spatial features using network graphs
R Sierra, CR Stephens
International Journal of Geographical Information Science 26 (3), 441-468, 2012

Constructing ecological networks: a tool to infer risk of transmission and dispersal of Leishmaniasis
C González-Salazar, CR Stephens
Zoonoses and public health 59 (s2), 179-193, 2012

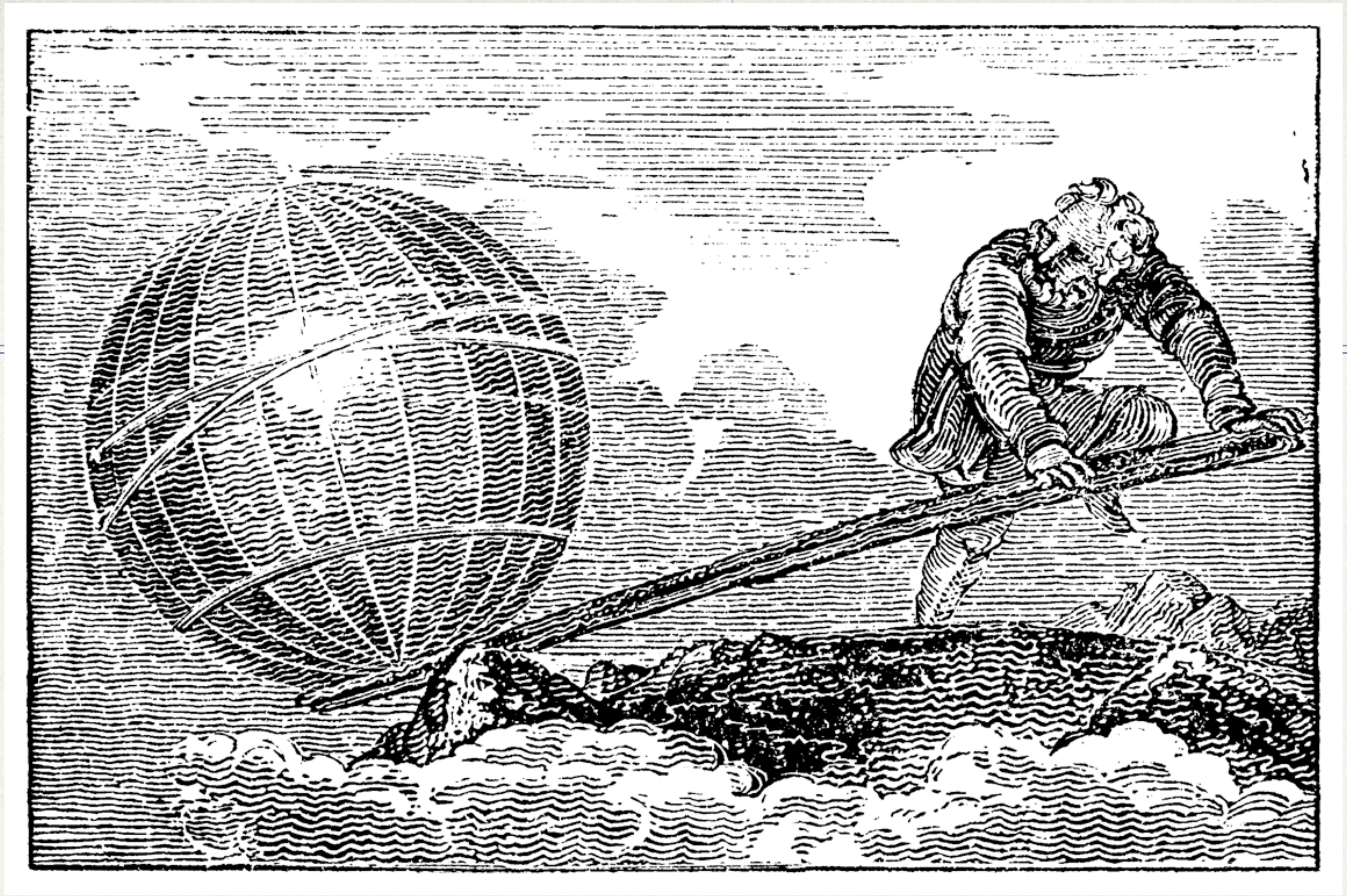
Comparing the relative contributions of biotic and abiotic factors as mediators of species' distributions
C González-Salazar, CR Stephens, PA Marquet
Ecological Modelling 248, 57-70, 2013

Leishmania (L.) mexicana Infected Bats in Mexico: Novel Potential Reservoirs
M Berzunza-Cruz, Á Rodríguez-Moreno, G Gutiérrez-Granados, ...
PLoS neglected tropical diseases 9 (1), e0003438-e0003438, 2015

Predicting the potential role of non-human hosts in Zika virus maintenance
C González-Salazar, CR Stephens and V. Sanchez-Cordero
submitted to Eco-health

UNDERSTANDING TRANSMISSIBILITY PATTERNS OF CHAGAS DISEASE THROUGH COMPLEX VECTOR-HOST NETWORKS
Laura Rengifo-Correa, Constantino González-Salazar, Juan J. Morrone, Juan Luis Téllez-Rendón, Christopher Stephens, submitted to PLoS Neglected Tropical diseases

Can you judge a disease host by the company it keeps? Predicting disease hosts and their relative importance using complex networks
CR Stephens et al, submitted to PLoS Neglected Tropical diseases



δῶς μοι πᾶ στῶ καὶ τὰν γᾶν κινάσω

Give me a place to stand on and I´ll move the earth

Give me enough data and I´ll predict anything

The Data Revolution will revolutionise our ability to model and understand ecology

Table 1.Bioclimatic variables from WorldClim: BIO1= Annual Mean Temperature; BIO2= Mean Diurnal Range (Mean of monthly (max temp - min temp)); BIO3= Isothermality [((BIO2/BIO7) * 100)]; BIO4= Temperature Seasonality (standard deviation *100); BIO5= Max Temperature of Warmest Month; BIO6= Min Temperature of Coldest Month; BIO7= Temperature Annual Range (BIO5-BIO6); BIO8= Mean Temperature of Wettest Quarter ; BIO9= Mean Temperature of Driest Quarter; BIO10= Mean Temperature of Warmest Quarter ; BIO11= Mean Temperature of Coldest Quarter; BIO12= Annual Precipitation; BIO13= Precipitation of Wettest Month; BIO14= Precipitation of Driest Month; BIO15= Precipitation Seasonality (Coefficient of Variation); BIO16= Precipitation of Wettest Quarter; BIO17= Precipitation of Driest Quarter; BIO18= Precipitation of Warmest Quarter; BIO19= Precipitation of Coldest Quarter. These bioclimatic variables were derived from the average monthly mean temperature (°C * 10), average monthly minimum temperature (°C * 10), average monthly maximum temperature (°C * 10) and average monthly precipitation (mm) (Hijmans et al., 2005).

Range	BIO1	BIO2	BIO3	BIO4	BIO5	BIO6	BIO7
R1	-27-5	73-97	37-44	210-984	38-76	-98--65	115-166
R2	6-37	98-108	45-48	985-1759	77-114	-64--32	167-189
R3	38-70	109-119	49-51	1760-2534	115-152	-31-1	190-214
R4	71-102	120-130	52-55	2535-3309	153-190	2-34	215-238
R5	103-135	131-141	56-60	3310-4084	191-229	35-67	239-262
R6	136-167	142-153	61-64	4085-4859	230-267	68-100	263-284
R7	168-199	154-164	65-67	4860-5634	268-305	101-133	285-306
R8	200-232	165-174	68-71	5635-6409	306-343	134-166	307-329
R9	233-264	175-184	72-76	6410-7184	344-381	167-199	330-355
R10	265-297	185-207	77-84	7185-7959	382-420	200-232	356-392

	BIO8	BIO9	BIO10	BIO11	BIO12	BIO13	BIO14
R1	-22-11	-35--2	-20-14	-36--4	42-507	8-84	0-12
R2	12-45	-1-31	15-48	-3-28	508-973	85-161	13-25
R3	46-79	32-64	49-82	29-60	974-1439	162-237	26-37
R4	80-113	65-97	83-117	61-92	1440-1905	238-314	38-50
R5	114-147	98-131	118-151	93-125	1906-2371	315-391	51-63
R6	148-181	132-164	152-185	126-157	2372-2836	392-467	64-75
R7	182-215	165-197	186-220	158-189	2837-3302	468-544	76-88
R8	216-249	198-230	221-254	190-221	3303-3768	545-620	89-100
R9	250-283	231-263	255-288	222-253	3769-4234	621-697	101-113
R10	284-317	264-297	289-323	254-286	4235-4700	698-774	114-126

	BIO15	BIO16	BIO17	BIO18	BIO19
R1	37-45	18-218	0-43	1-125	0-95
R2	46-54	219-418	44-87	126-249	96-191
R3	55-63	419-618	88-131	250-373	192-287
R4	64-72	619-818	132-175	374-497	288-383
R5	73-81	819-1018	176-219	498-622	384-479
R6	82-89	1019-1218	220-262	623-746	480-575
R7	90-98	1219-1418	263-306	747-870	576-671
R8	99-107	1419-1618	307-350	871-994	672-767
R9	108-116	1619-1818	351-394	995-1118	768-1016
R10	117-125	1819-2019	395-438	1119-1243	1017-1927