CONABIO

SPECIES:

A platform for modelling spatial data and identifying ecological interactions

Chris Stephens, Constantino Gonzalez and Raul Sierra
C3-Centro de Ciencias de la Complejidad y Instituto de Ciencias Nucleares, UNAM
CONABIO


Just how many interactions can we directly observe?
Can we infer interactions without direct observation?
SCIENCEphotoLIBRARY
gettyimages
Frank Lukasseck
ARKive
Importancia médica
T. Infestans
T. barberi
T. pallidipennis
T. longipenis
T. recurva
T. neotomac
Culex quinquefasciatus

Ecology is the scientific analysis and study of

<u>interactions</u>

---

among organisms and their environment

Physics is the scientific analysis and study of

<u>interactions</u>

between matter and energy

How have we understood <u>interactions</u> in physics?

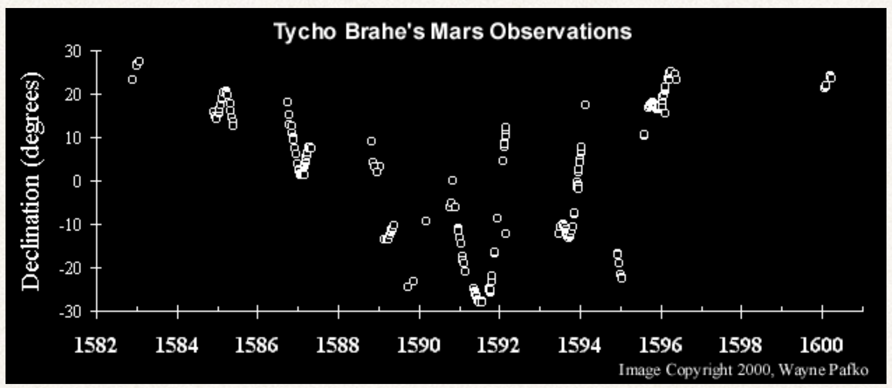
Through Spatial Modeling!

Studying where things are, and when,

relative to each other.

Spatial Modeling in the past…

Data —> Phenomenology —> Taxonomy —> Theory



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*).

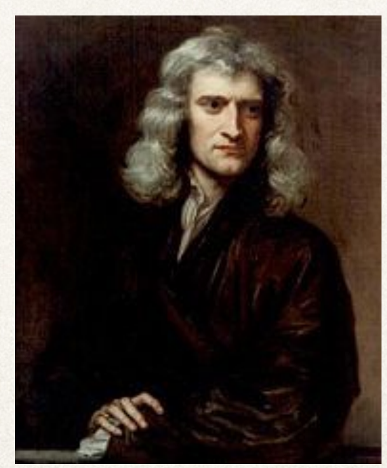
This implies that the Sun may be the physical cause of the acceleration of planets.
Newton defined the force acting on a planet to be the product of its mass and the acceleration. So:

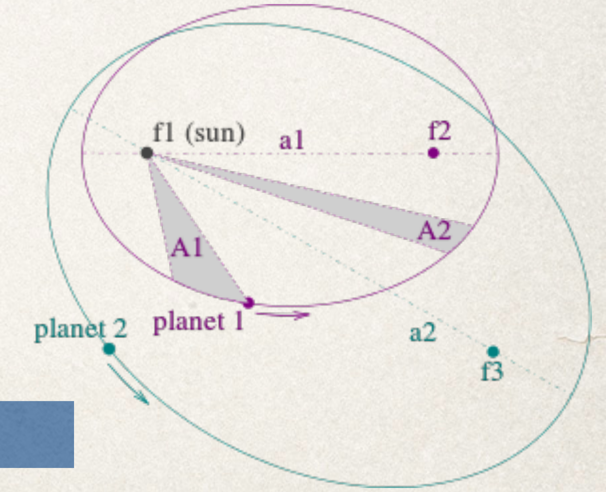
1. Every planet is attracted towards the Sun.
2. The force acting on a planet is in direct proportion to the mass of the planet and in inverse proportion to the square of its distance from the Sun.

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

CONABIO

"Keplerian" Ecological models

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

$$P(C \mid X)$$

S(C | X)
Risk score

Characterizes niche and "anti-niche"

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!

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,…

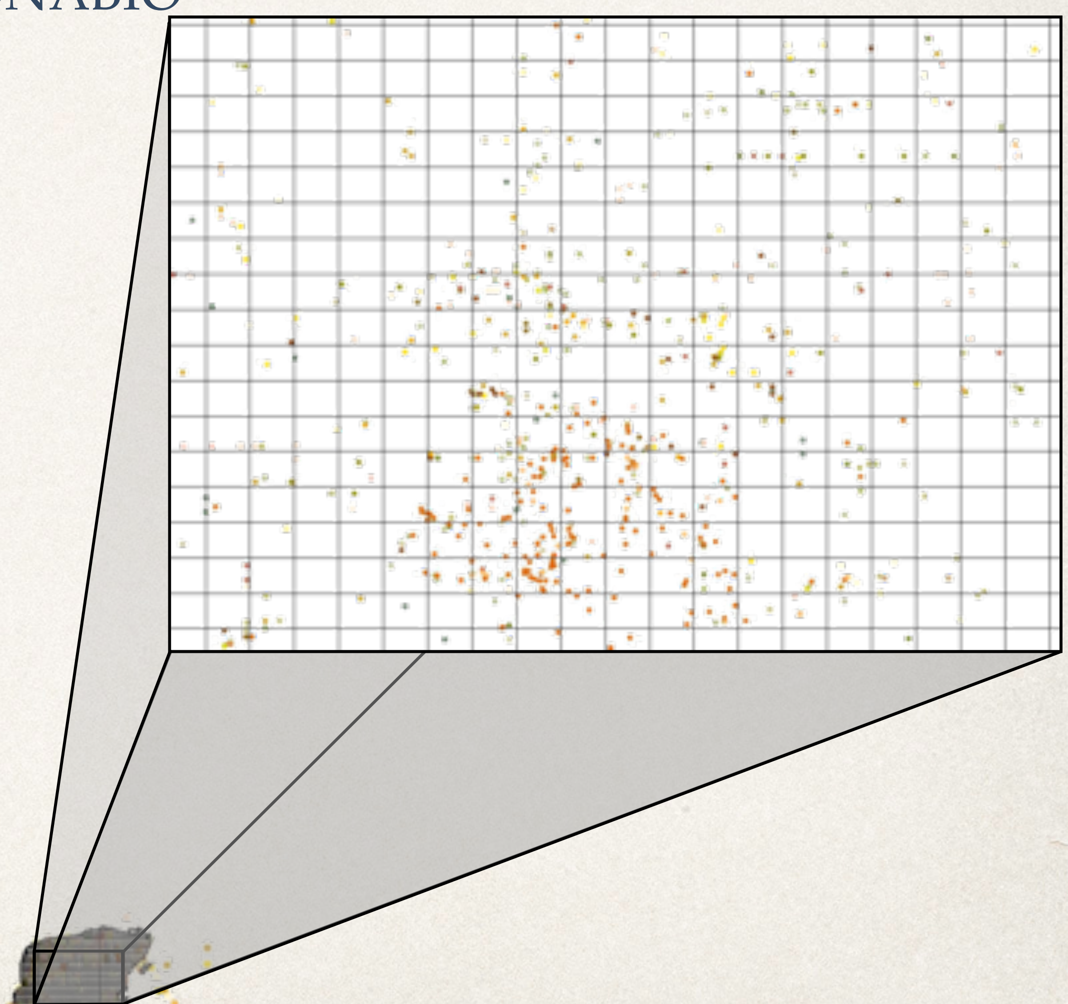
Different spatial resolution

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

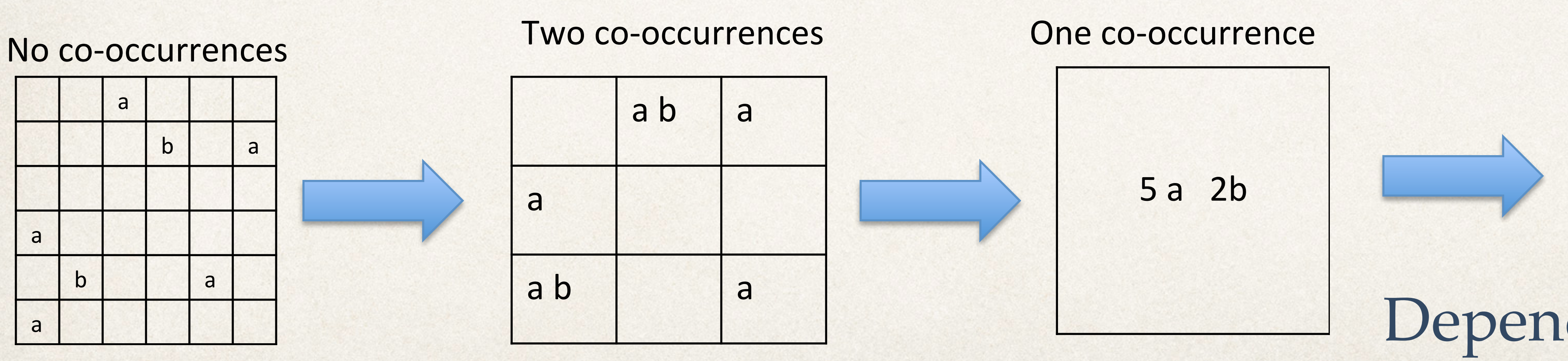


The data are represented in space and time – spatial data mining

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.

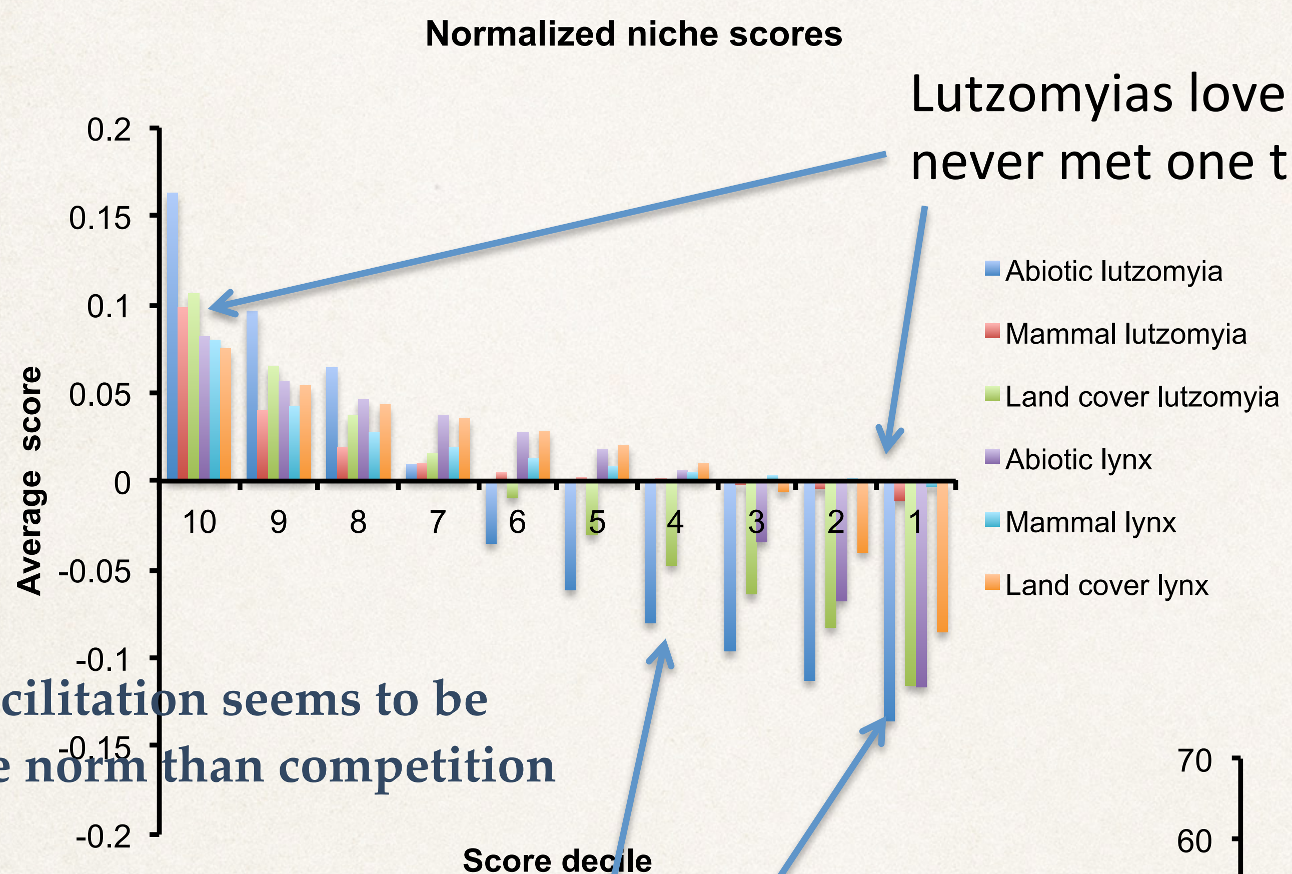


N(ab)/N(b)
= P(a | b)

Dependence of species a on niche variable b

Here we're in geographic space

Two Example Niches

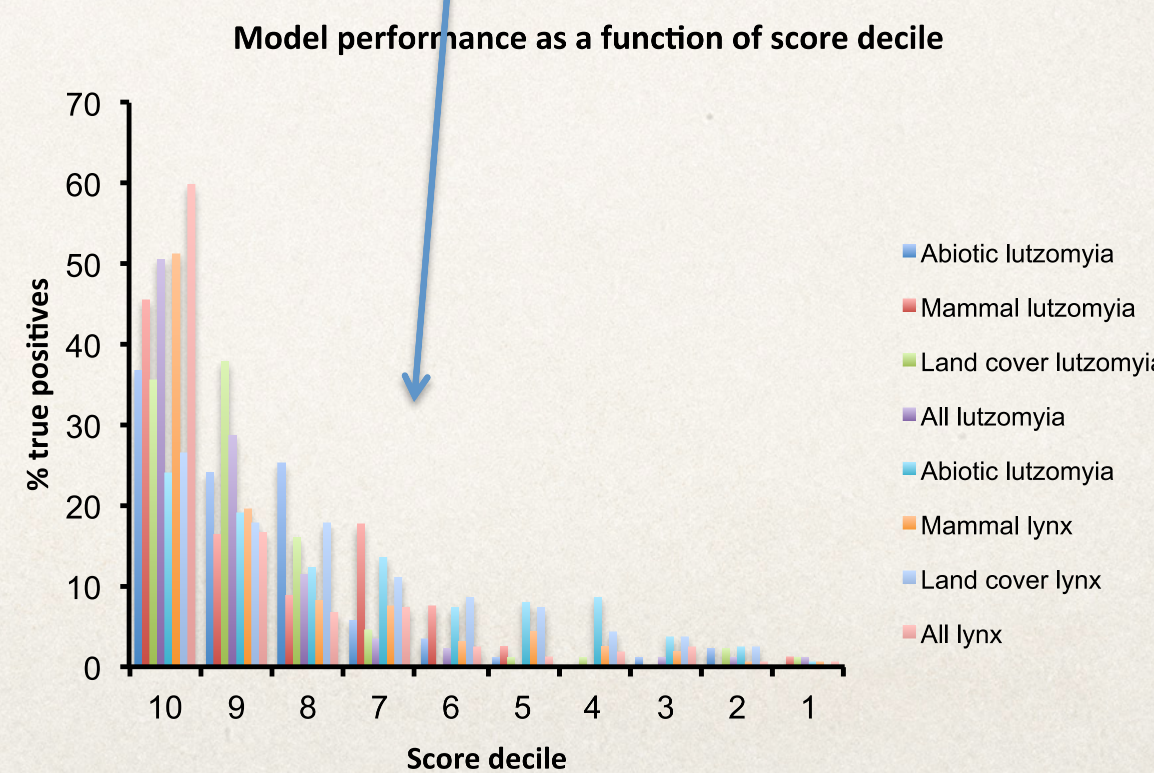


Lutzomyias love mammals, never met one they didn't like

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

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

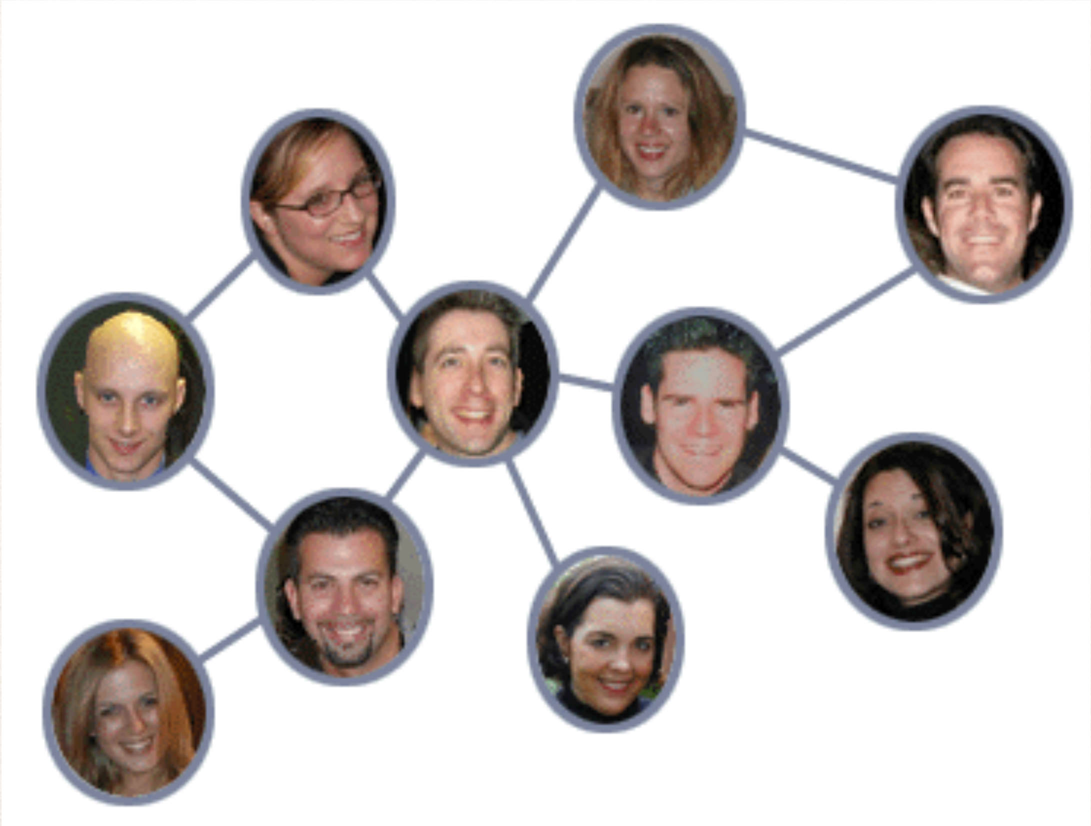


Chains of causality

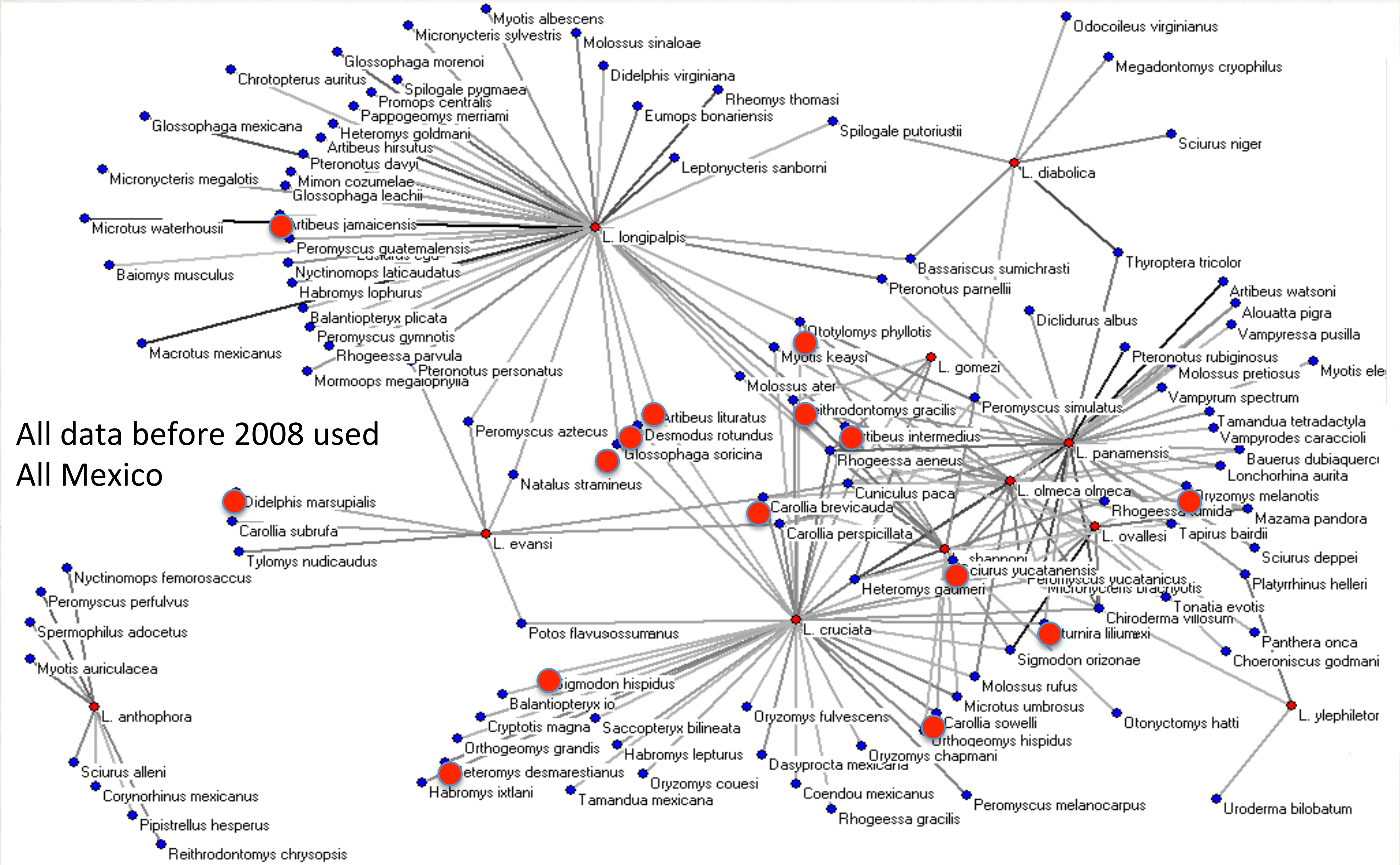
Now for Communities…

You can judge a man by his “friends”

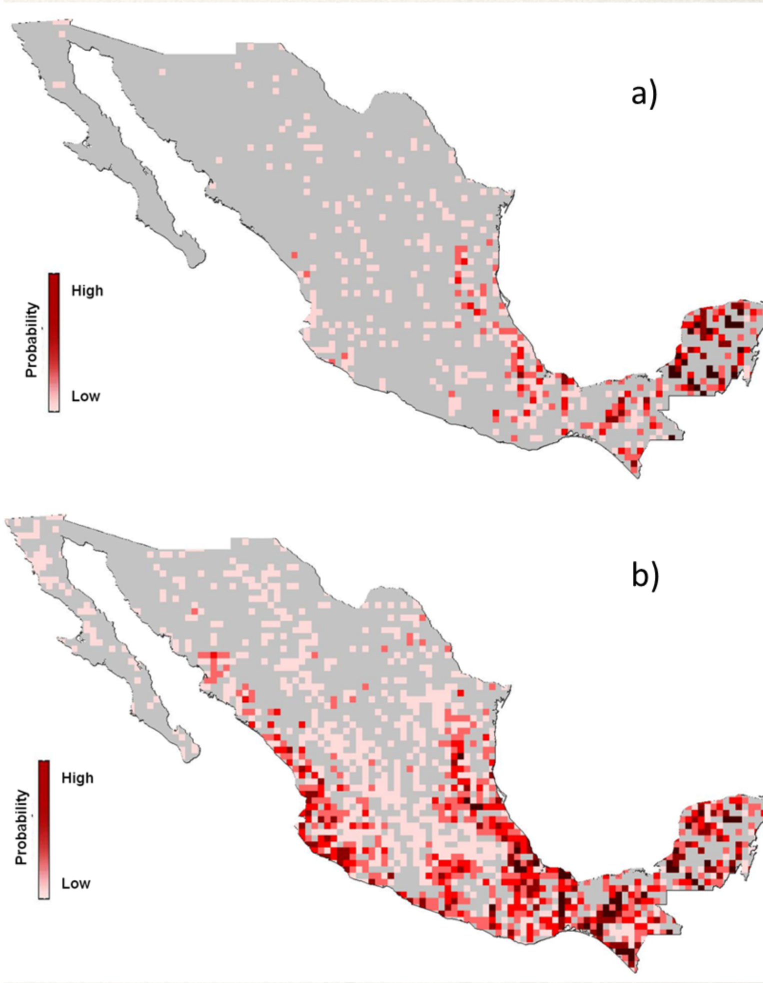
or his “enemies”, or “parasites”, or “prey” or “predators” or…



The Ecology of Leishmaniasis



All data before 2008 used
All Mexico



What does this tell us about vector control?

Who are we?

Grupo de Trabajo

C3 - Centro de Ciencias de la Complejidad, UNAM; CONABIO;

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. Everardo Robredo
7.- Lic. Juan Carlos Salazar Carrillo
8.- Ma. Juan Barrios
9.- Ing. Raúl Jiménez *

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