Epidemic Networks

Garrettlab.com

garrettlab.com

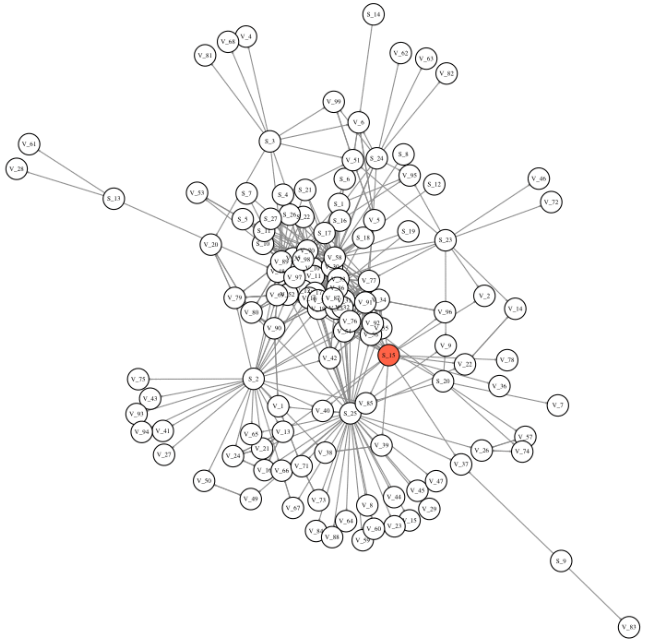
UF | IFAS
UNIVERSITY of FLORIDA

Institute for
Sustainable Food
Systems

PLANT PATHOLOGY
UF/IFAS

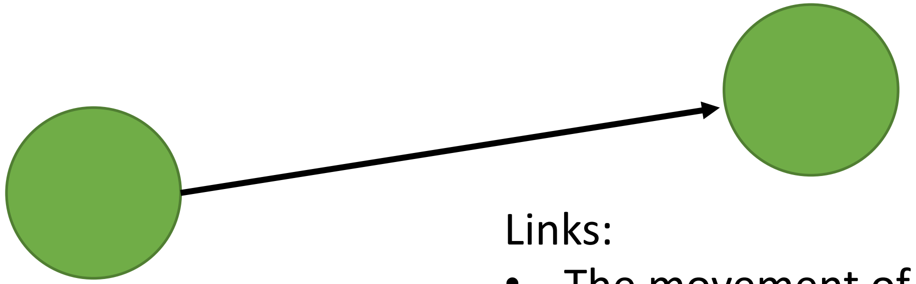
Learning Goals

- Why networks can be useful for modeling epidemics
- Advantages and disadvantages of network analysis
- Network property examples
- R code!



Network Models for Epidemics

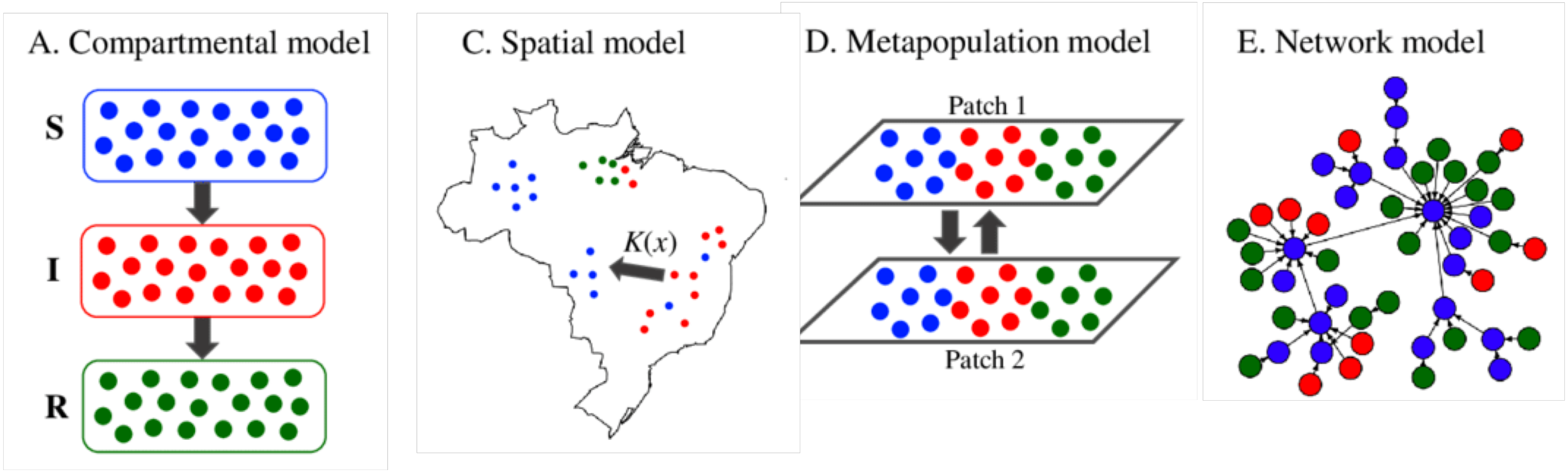
- Network models are effective tools for modeling human, animal, and plant diseases



Links:
- The movement of pathogens between nodes
- Probability of infection
- Links can be weighted
- Can be dynamic or fixed
- Can vary over time (formation and dissolution)
- Can be generated based on empirical trade data
- Or modeled

- Nodes can be:
  - Genes
  - Plants
  - Farms
  - Farmers
  - Communities
  - Landscapes
  - Nations

Network models vs. Standard Differential Equation Models (SIR, for example)



Anuwat Wiratsudakul, et al 2018

Advantages and disadvantages of network models

Advantages

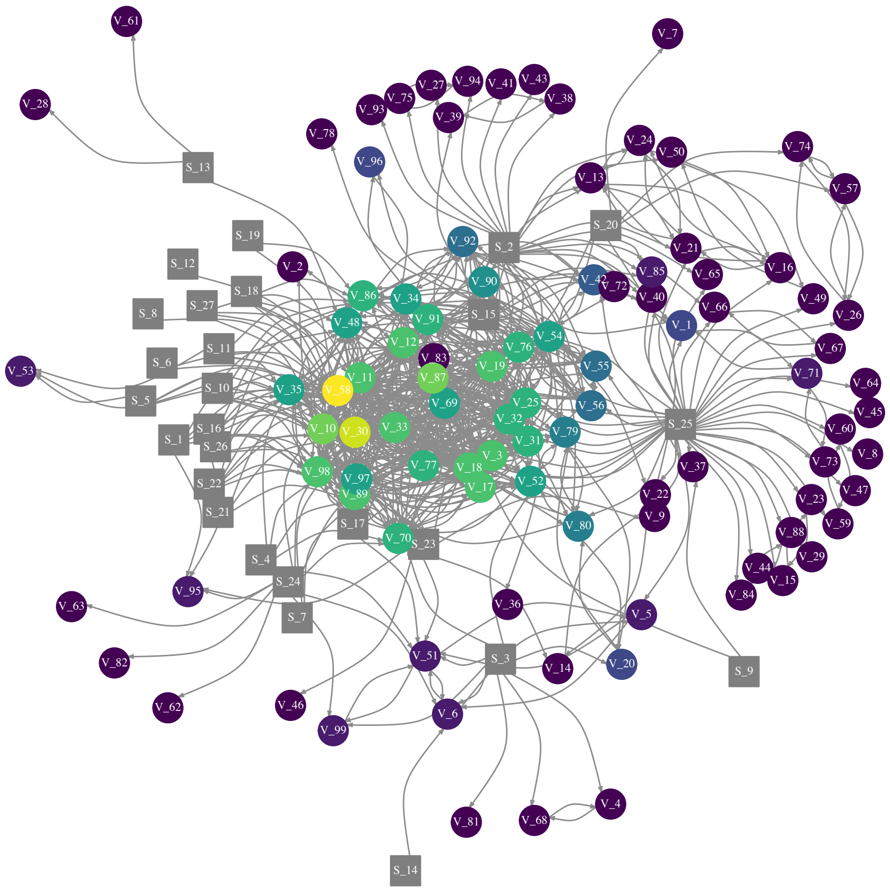
- Realistic contact structures between individuals
- Can accurately capture trade patterns
- Geo-spatial connectivity
- Social influence

Disadvantages

- Data availability - difficult to sample a complete large network
- Sampling
- Effect of aggregation

Examples of Questions One Might Ask with Network Models of Epidemics:

- Given what we know about a geographic network, what are the best locations for risk based pathogen surveillance?
- Are there locations that are super-spreaders?
- How do social processes affect plant disease dispersal?



Network descriptives may tell us some about epidemic risk in networks

Measure	Definition	Relevance for Epidemics
Degree Centrality	The number of links a node has to other nodes in the network (both incoming and outgoing)	Node degree of epidemic starting point (number of links) has been shown to influence epidemic outcomes. Those with high degree may be "super spreaders" once infected
Betweenness centrality	The number of shortest paths through the network of which a node is a part	A measure of how much a node serves as a "bridge" to new nodes. Removal of nodes with high betweenness may contain an epidemic within a region
Eigenvector centrality	A weighted sum reflecting both direct links to a node (degree) and the node degree of neighbors	If a node does not itself have a high node degree, but is connected to nodes with high degree, it may be at increased risk of infection and spreading infection
Closeness Centrality	The inverse of the average length of the shortest path to/from all the other nodes in the network	Efficiency of the node to spread disease to any other node in the network

Node statistics and epidemic outcomes
Case: variety spread in seed networks

- Study looking at barley seed flows in Ethiopia
- Findings indicate that the node degree of the starting households drives the final number of households reached
- Important implications for epidemic spread in seed systems

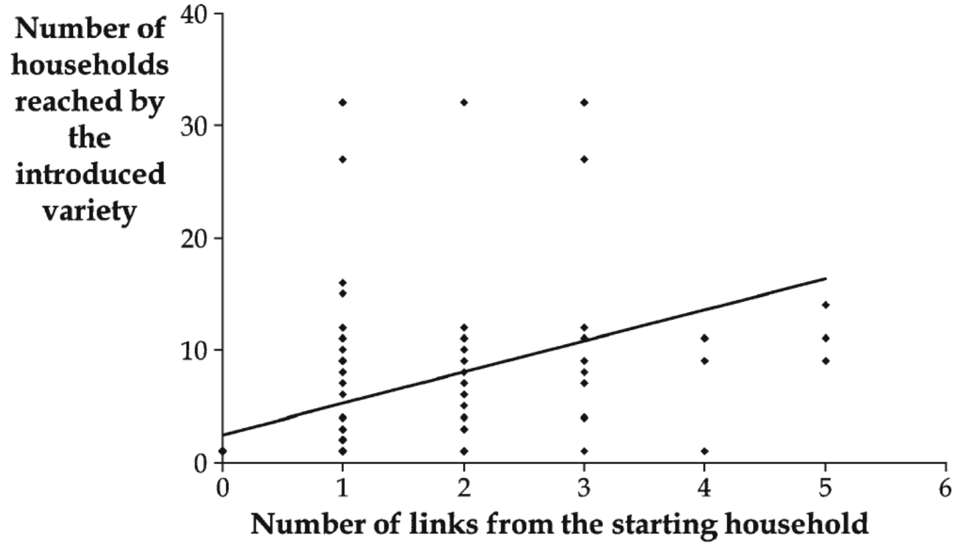
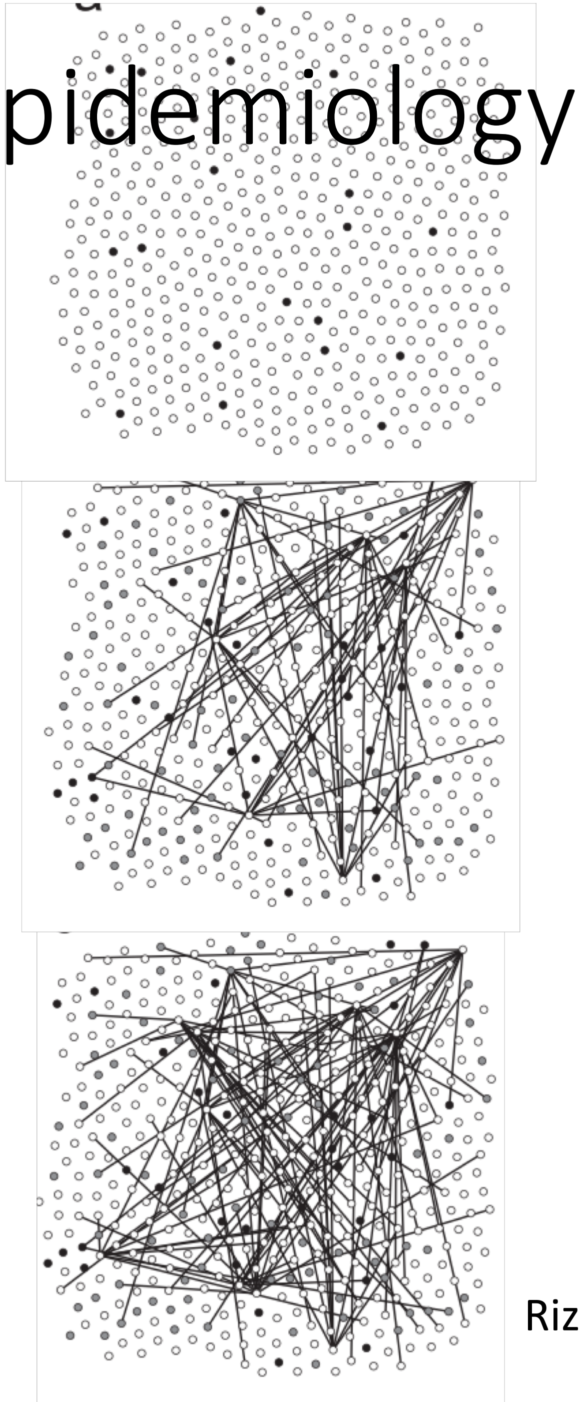


Fig. 4 Number of households reached by an introduced variety (using the simulation model described in the article) as a function of the number of outgoing connections of the household to whom the variety was first given or that first created it (n=197, r^2=0.18, y=2.48+2.77x, p<0.001)

Marco Pautasso 2015

Used for many years in human epidemiology
Case: 2014 Ebola Outbreak

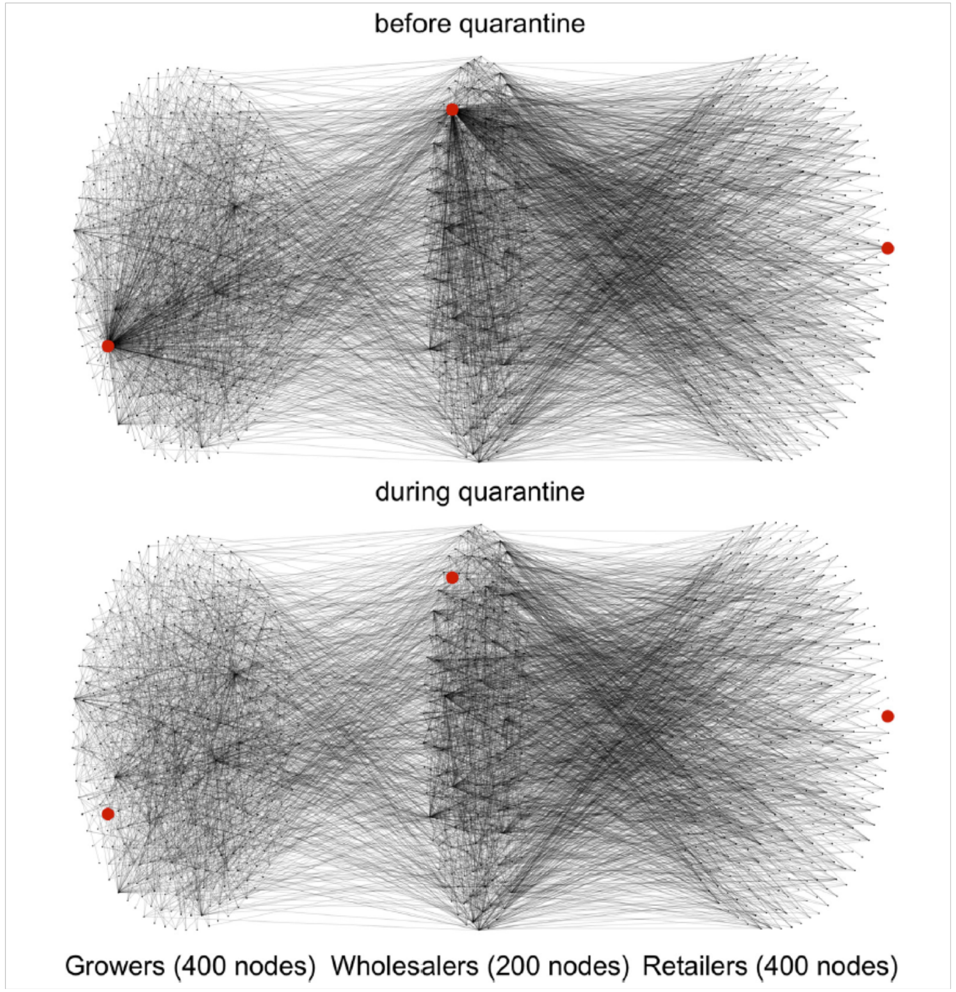
- Modeling Ebola spread combining SEIR model with network-based model
- Overcomes the assumption of homogeneous mixing of other Ebola models
- Calibrated with field data from 2014 Liberia
- Modeled a one year forecast of disease spread along with what-if scenario analysis to assess intervention strategies



Rizzo et al 2016

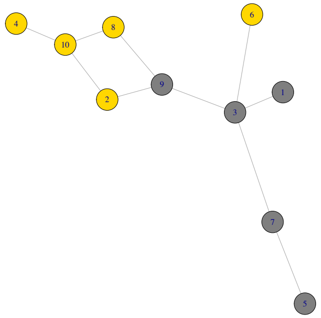
Epidemics in horticultural trade networks

- Multi-layer network that looked at growers, wholesalers, and retailers
- Once infection was detected, infected nurseries were isolated until infection dropped below a detection threshold
- Focusing epidemics on hub growers is most useful for controlling spread



Nelson & Boone 2015

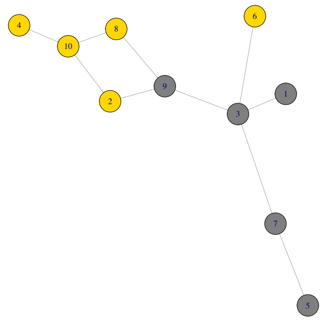
Exponential Random Graph Models (ERGMs)



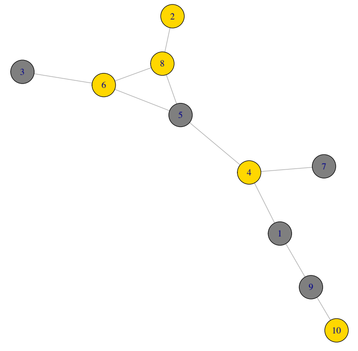
Concept: Exponential Random Graph Models (ERGMs)

- Statistical modeling framework to understand the generative process driving link formation in a network that deviates from random
- Predicts probability of a link, conditional on the structure of the rest of the network

Consider sampled network data on edges between yellow and gray nodes



Does this differ from simple random graphs with the same number of nodes and links (10 million possible)?



Exponential random graph models (ERGMs)

- Fit using Monte Carlo Markov Chain maximum likelihood estimation

- Predictor types (and examples of hypotheses)
  - Node - H: node of type X is more likely to have a link than other types of nodes
  - Dyad- H: nodes of type X and Y are more likely to have a link between them
  - Relation- H: a node with link type X is more likely to have link type Y

ERGM Model Specification

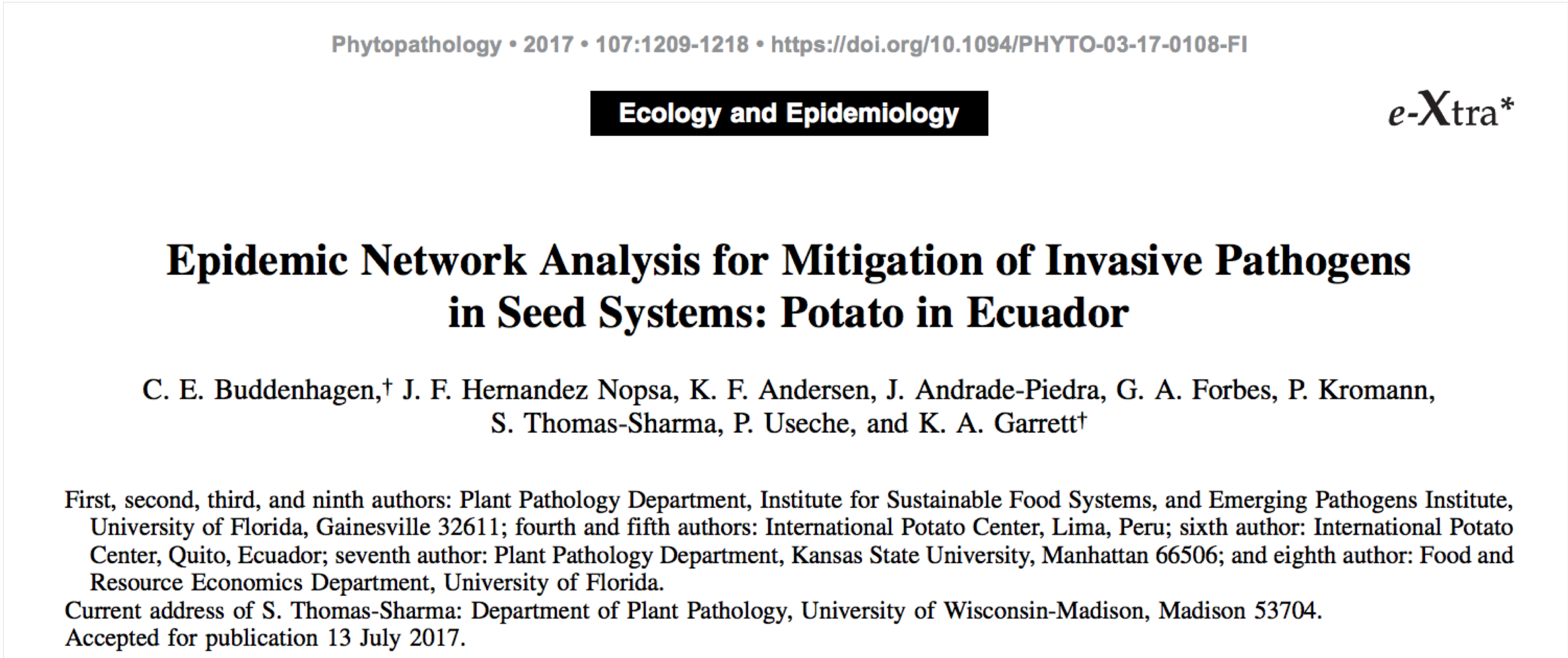
Useful R Packages:
- statnet()
- ergm()
- stergm()

- Null model: a random graph model
  - Network ~ edges
  - Single model term analogous to the intercept in a regression model

- The null model fit can be compared to the fit of more complicated models

- Compare models using statistics like AIC and BIC, to evaluate the trade-off between adding more parameters and improving the model fit

- **Simulate networks based on specified network models- important for using sampled seed system data to estimate the larger network**



Phytopathology • 2017 • 107:1209-1218 • https://doi.org/10.1094/PHYTO-03-17-0108-FI

Ecology and Epidemiology

e-Xtra*

Epidemic Network Analysis for Mitigation of Invasive Pathogens in Seed Systems: Potato in Ecuador

C. E. Buddenhagen,† J. F. Hernandez Nopsa, K. F. Andersen, J. Andrade-Piedra, G. A. Forbes, P. Kromann,
S. Thomas-Sharma, P. Useche, and K. A. Garrett†

First, second, third, and ninth authors: Plant Pathology Department, Institute for Sustainable Food Systems, and Emerging Pathogens Institute, University of Florida, Gainesville 32611; fourth and fifth authors: International Potato Center, Lima, Peru; sixth author: International Potato Center, Quito, Ecuador; seventh author: Plant Pathology Department, Kansas State University, Manhattan 66506; and eighth author: Food and Resource Economics Department, University of Florida.

Current address of S. Thomas-Sharma: Department of Plant Pathology, University of Wisconsin-Madison, Madison 53704.

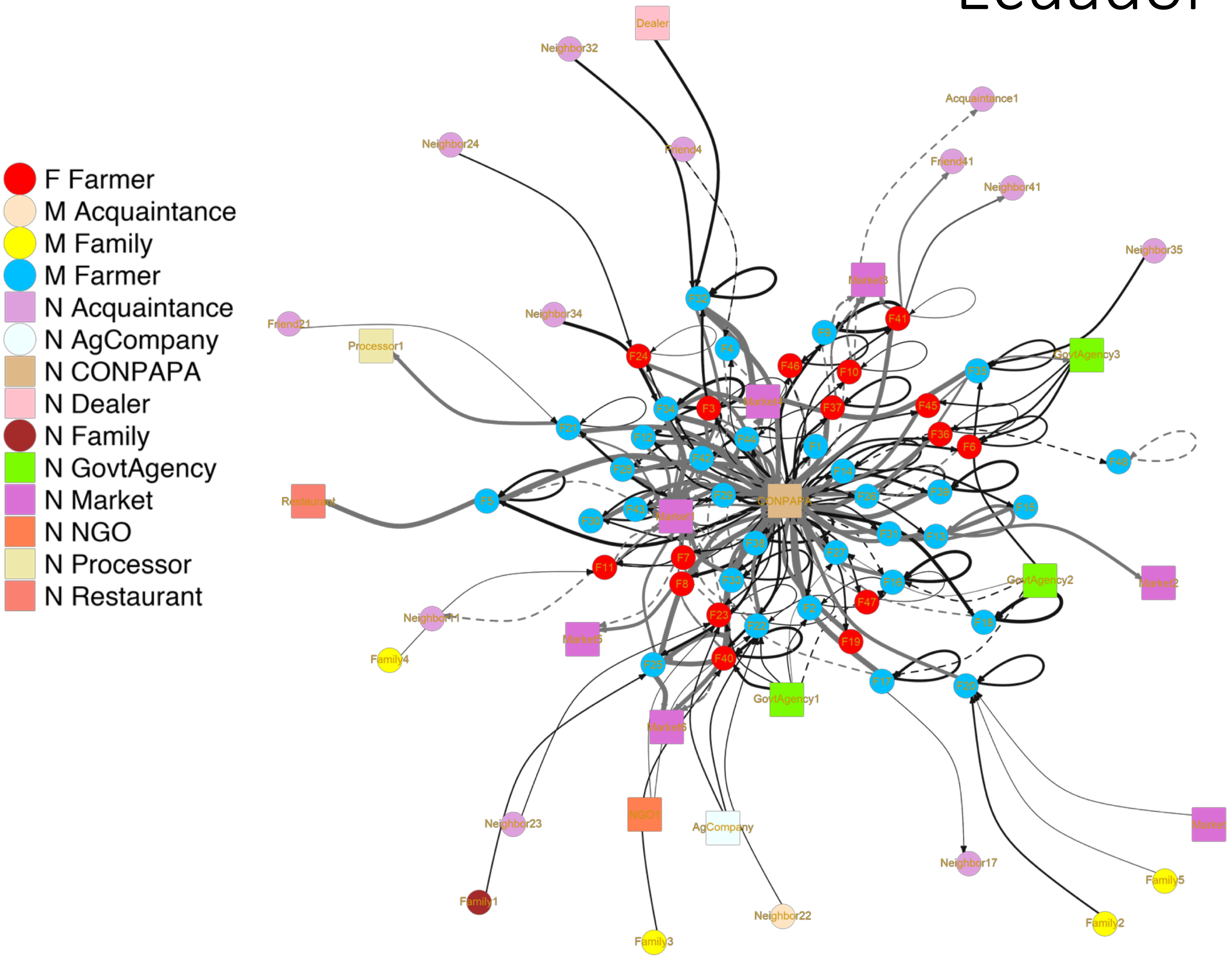
Accepted for publication 13 July 2017.

RESEARCH PROGRAM ON Roots, Tubers and Bananas
CGIAR

• INTERNACIONAL • CENTRO DE LA PAPA • CIP

UF IFAS UNIVERSITY of FLORIDA

Ecuador Seed Network



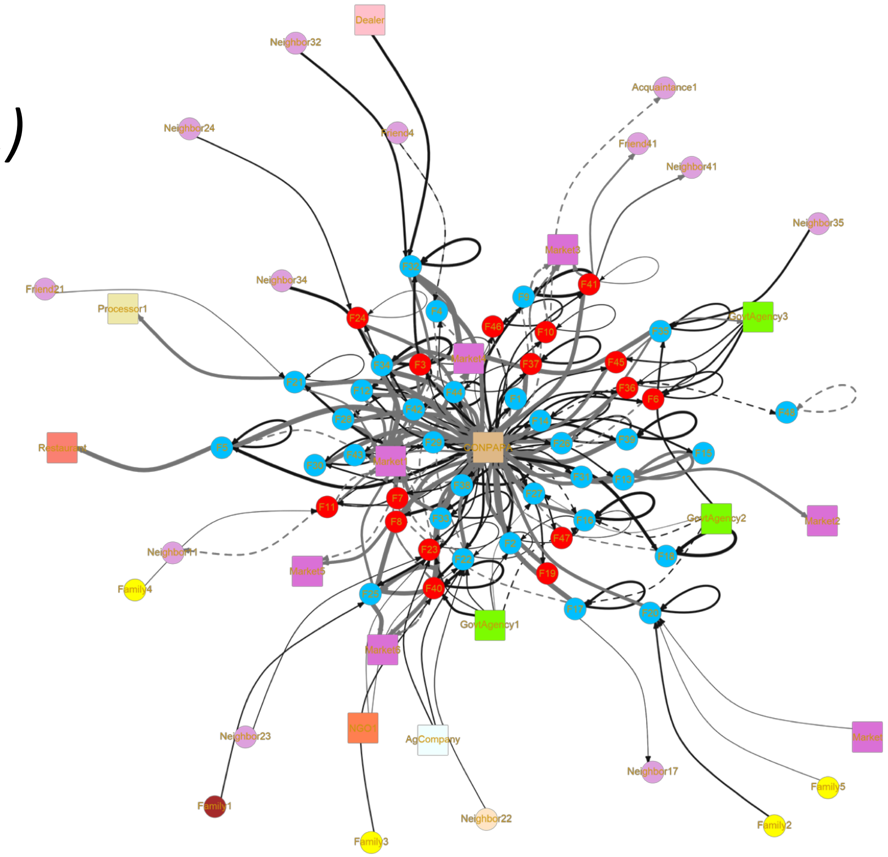
In the potato seed system in Ecuador, what is the effect of node type on the probability that a link exists?

Model:

$$Y \sim edges + nodematch(\text{“gender”}) + nodematch(\text{“type”}) + mutual + idegrange(1) + gwidegree(1)$$

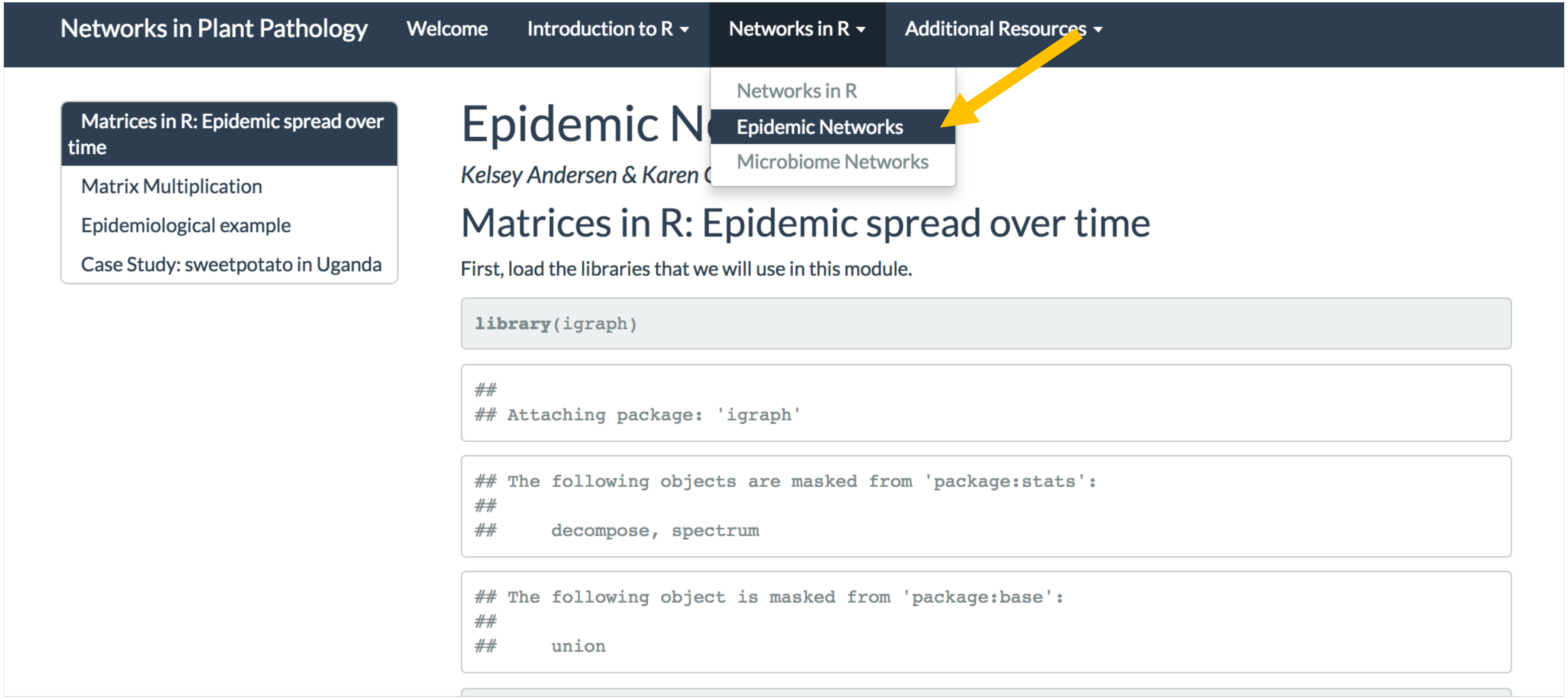
- **Node type** had an effect on the likelihood that reported links exist in the potato transaction network ($P < 0.0001$).

- Farmers were much more likely to report transaction links with institutions (like CONPAPA) than with other farmers.



Let’s move to the markdown!

kelseyandersen.github.io/NetworksPlantPathology/



Epidemic N

Kelsey Andersen & Karen G

Matrices in R: Epidemic spread over time

First, load the libraries that we will use in this module.

```
library(igraph)
```

```
##
## Attaching package: 'igraph'
```

```
## The following objects are masked from 'package:stats':
##
##     decompose, spectrum
```

```
## The following object is masked from 'package:base':
##
##     union
```