

# Epidemic Networks

Kelsey Andersen & Karen Garrett

[garrettlab.com](http://garrettlab.com)



Institute for  
Sustainable Food  
Systems



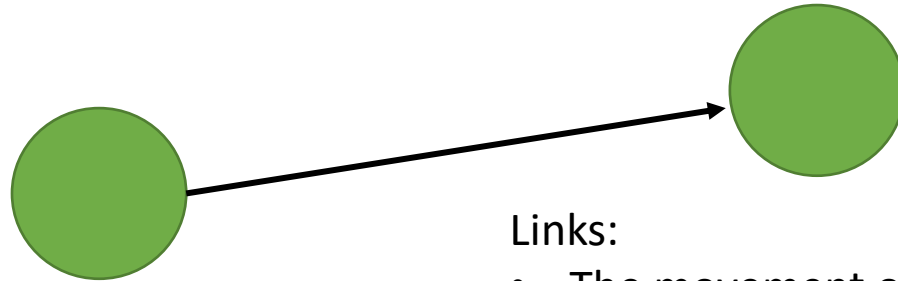
# Learning Goals

- 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



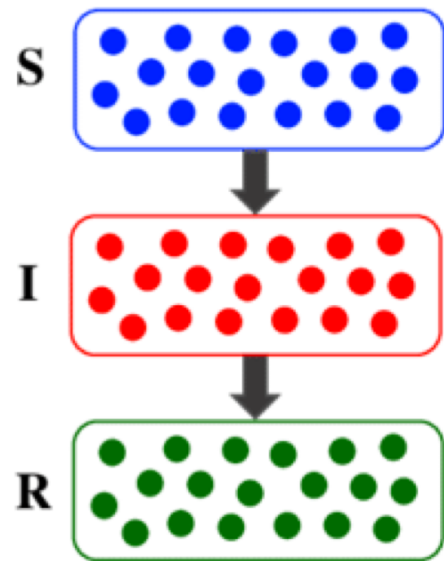
- Nodes can be:
  - Genes
  - Plants
  - Farms
  - Farmers
  - Communities
  - Landscapes
  - Nations

## Links:

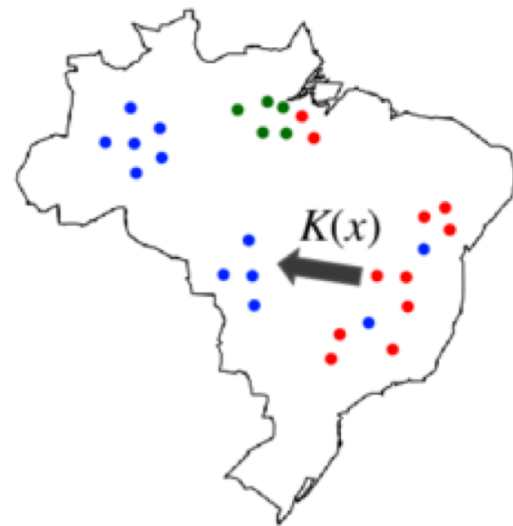
- The movement of pathogens between nodes
- Probability of infection
- Links can be weighted
- Can be dynamic or fixed
- Can vary over time (formation and dissolution)
- Can be generated based on empirical trade data
- Or modeled

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

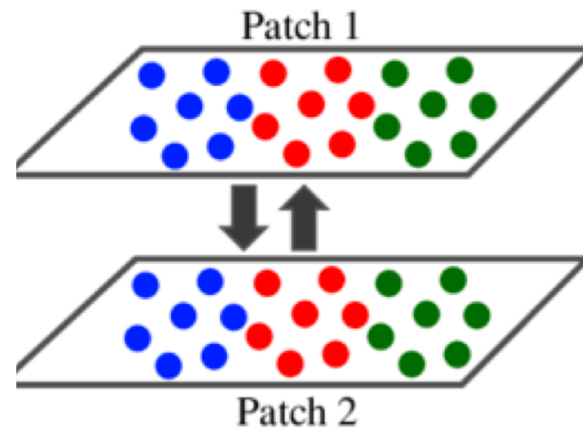
A. Compartmental model



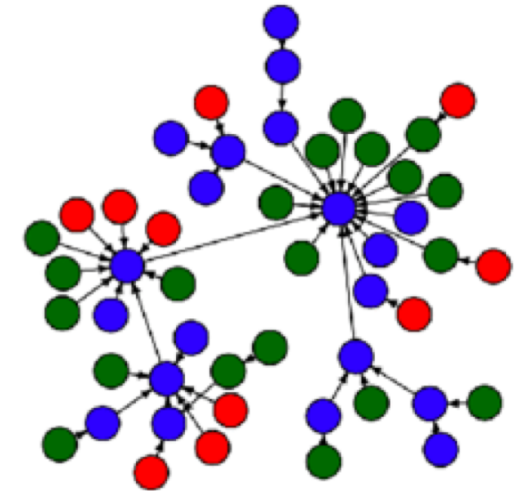
C. Spatial model



D. Metapopulation model



E. Network model



# Advantages and disadvantages of network models

## Advantages

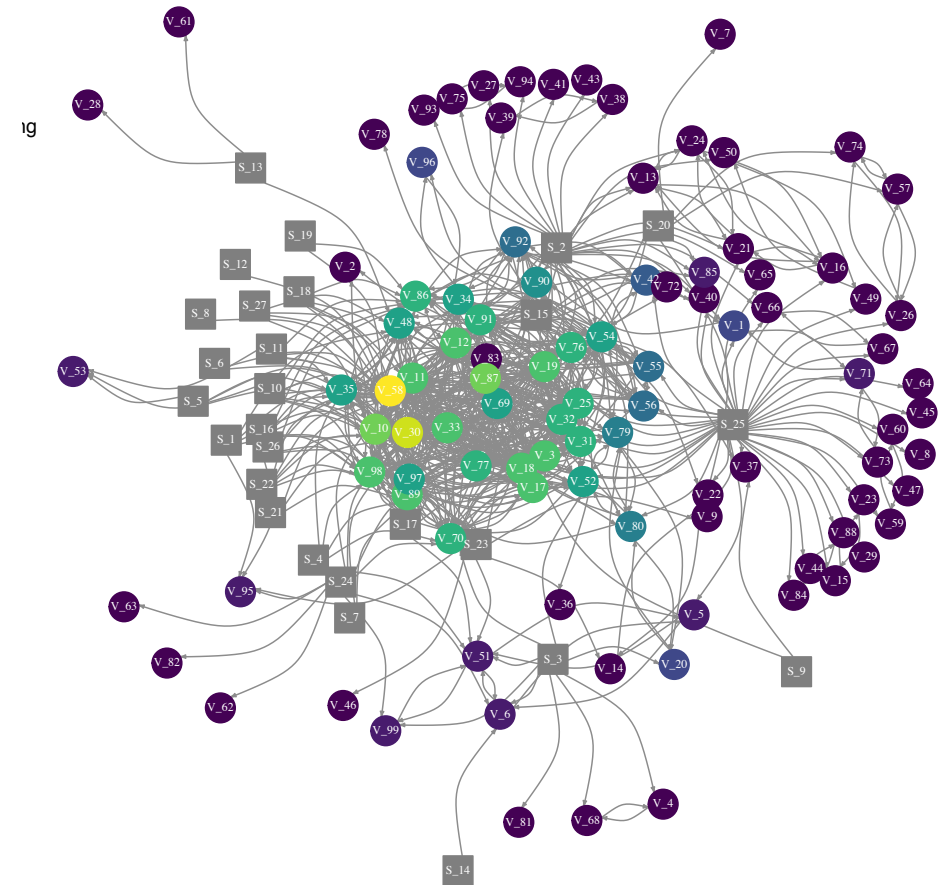
- Realistic contact structures between individuals
- Can accurately capture trade patterns
- Geo-spatial connectivity
- Social influence

## 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?



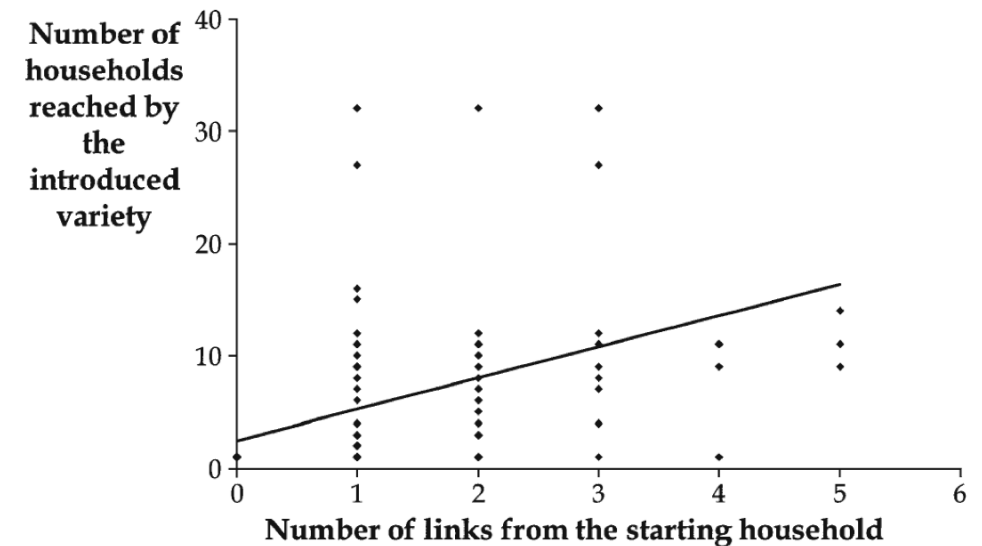
# Network descriptives may tell us some about epidemic risk in networks

Measure	Definition	Relevance for Epidemics
<b>Degree Centrality</b>	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

# 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



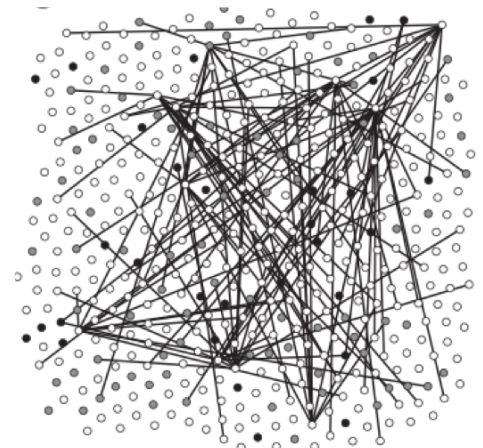
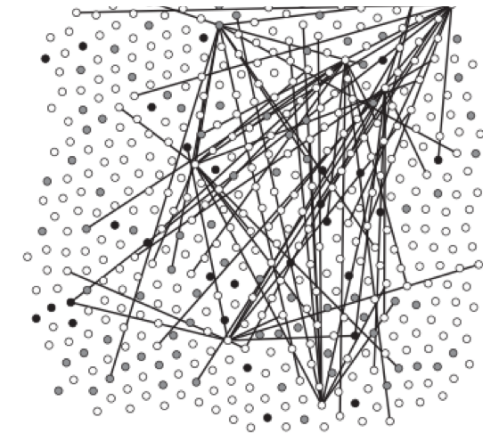
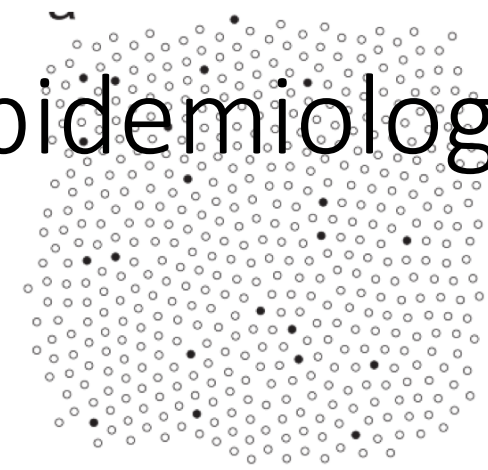
**Fig. 4** Number of households reached by an introduced variety (using the simulation model described in the article) as a function of the number of outgoing connections of the household to whom the variety was first given or that first created it ( $n=197$ ,  $r^2=0.18$ ,  $y=2.48+2.77x$ ,  $p<0.001$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



# Used for many years in human epidemiology

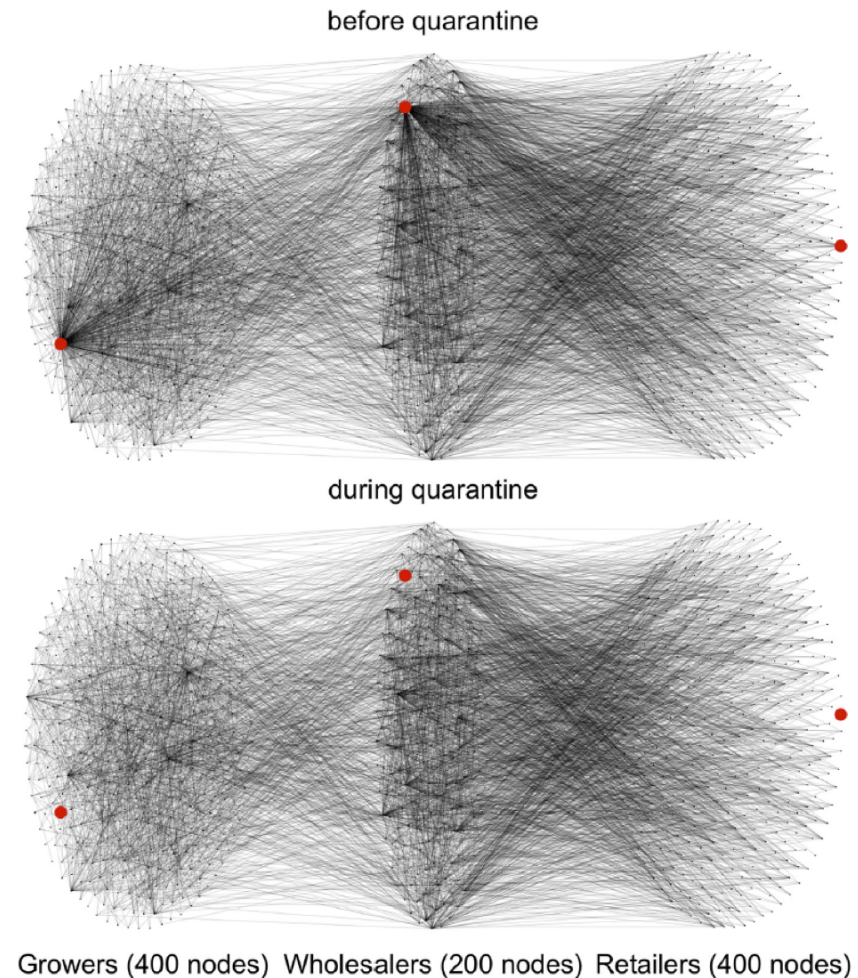
## Case: 2014 Ebola Outbreak

- 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

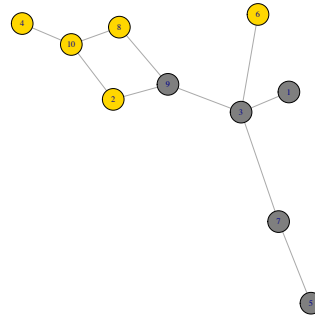


# Epidemics in horticultural trade networks

- 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



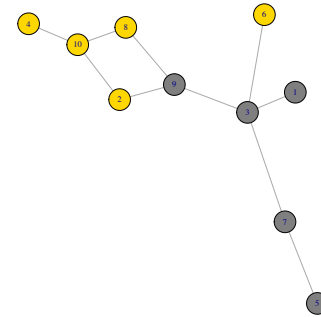
# Exponential Random Graph Models (ERGMs)



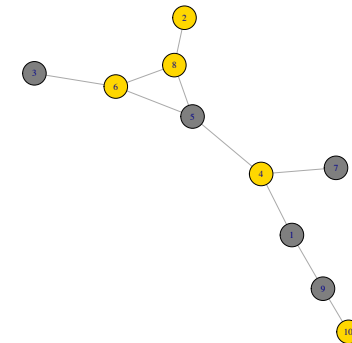
# Concept: Exponential Random Graph Models (ERGMs)

- 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  $\sim$  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**



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

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

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



UF | IFAS  
UNIVERSITY of FLORIDA





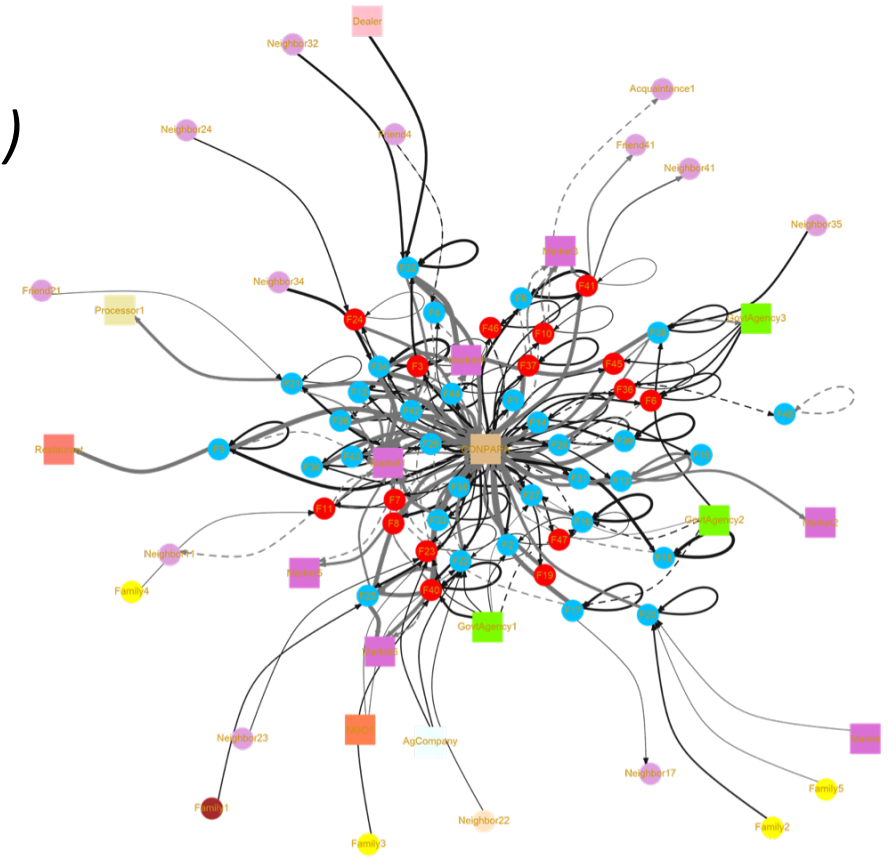


# In the potato seed system in Ecuador, what is the effect of node type on the probability that a link exists?

Model:

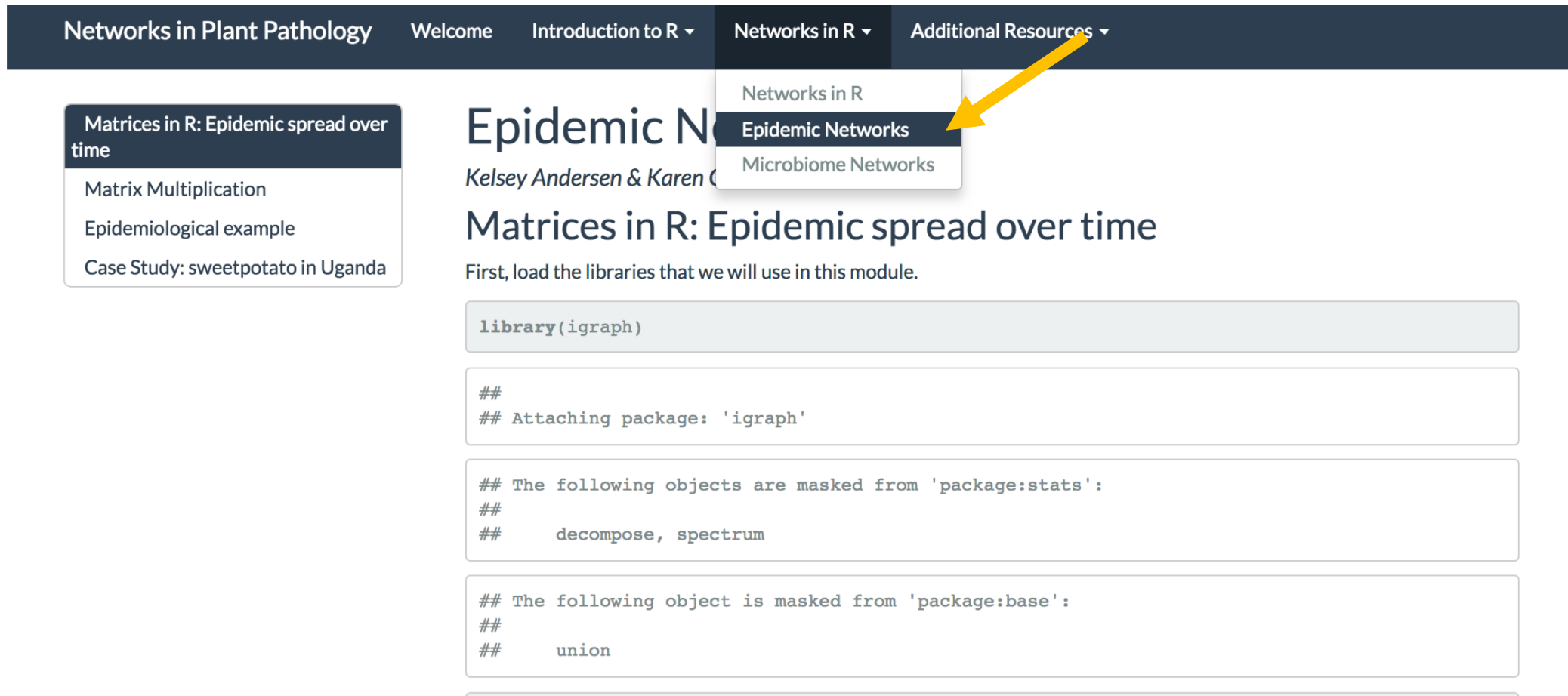
$$Y \sim \text{edges} + \text{nodematch}(\text{"gender"}) + \text{nodematch}(\text{"type"}) + \text{mutual} + \text{idegrange}(1) + \text{gwidegree}(1)$$

- **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/



The screenshot shows a website navigation bar with the following items: "Networks in Plant Pathology", "Welcome", "Introduction to R", "Networks in R", and "Additional Resources". The "Networks in R" menu is open, showing three options: "Networks in R", "Epidemic Networks", and "Microbiome Networks". A yellow arrow points to the "Epidemic Networks" option.

Matrices in R: Epidemic spread over time

- Matrix Multiplication
- Epidemiological example
- Case Study: sweetpotato in Uganda

## Epidemic Networks

Kelsey Andersen & Karen Coombes

### 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
```