

Assessing pathogen invasion based on community evenness and metabolic similarity.

Wai Kit Tsang

Promotors: Prof. dr. ir. Nico Boon
Prof. dr. Willem Waegeman
Tutor: Ir. Michiel Stock
Dr. Ramiro Vílchez Vargas

Ghent University

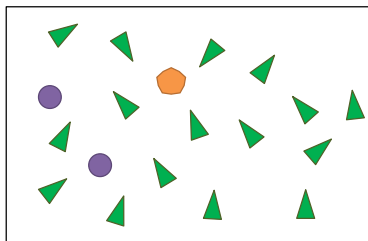
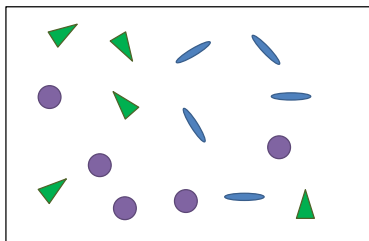
waikit.tsang@ugent.be

25 juni 2015

Problem Statement

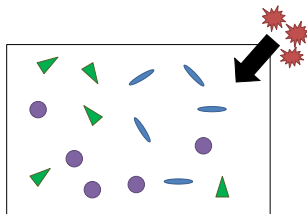
Problem Statement

Relate pathogen invasion to similarity-based diversity



Methodology

Labwork

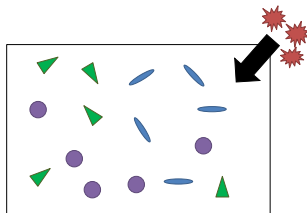


Modelling

- assembling synthetic communities
- 10 species
- pathogen invasion experiments

Methodology

Labwork

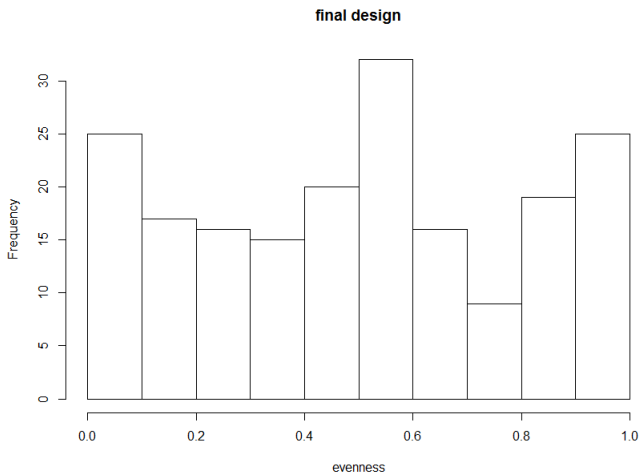


- assembling synthetic communities
- 10 species
- pathogen invasion experiments

Modelling

- Data from invasion experiments
- Evaluate similarity-based diversity

Labwork: Experimental Design



- Stratified and random design
- Test for evenness effect

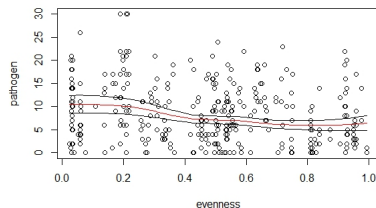
Lab Results

Negative Binomial Regression

```
glm.nb(formula = pathogen ~ ., data = df,
        init.theta = 1.325429331,
        link = log)
```

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	5.474	0.7227	7.574	3.61e-14 ***
evenness	-0.952	0.1725	-5.522	3.36e-08 ***
cellcount	-0.499	0.1159	-4.310	1.63e-05 ***



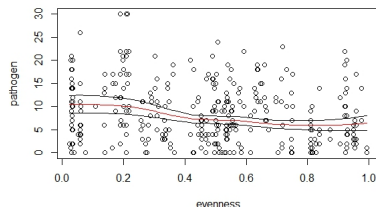
Lab Results

Negative Binomial Regression

```
glm.nb(formula = pathogen ~ ., data = df,
       init.theta = 1.325429331,
       link = log)
```

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	5.474	0.7227	7.574	3.61e-14 ***
evenness	-0.952	0.1725	-5.522	3.36e-08 ***
cellcount	-0.499	0.1159	-4.310	1.63e-05 ***



- Higher evenness decreases potential for pathogen invasion
- Worked with four pathogen, only one was significant

Similarity-based Diversity

Hill Diversity

$${}^qD(\mathbf{p}) = \left(\sum_{i=1}^S p_i^q \right)^{1/(1-q)}$$

Similarity-based Diversity

Hill Diversity

$${}^qD(\mathbf{p}) = \left(\sum_{i=1}^S p_i^q \right)^{1/(1-q)}$$

$$q = 0$$

$${}^0D(\mathbf{p}) = \text{Species Richness}$$

Similarity-based Diversity

Hill Diversity

$${}^qD(\mathbf{p}) = \left(\sum_{i=1}^S p_i^q \right)^{1/(1-q)}$$

$$q = 0$$

$${}^0D(\mathbf{p}) = \text{Species Richness}$$

$$q \rightarrow 1$$

$${}^1D = \exp(\text{Shannon})$$

Similarity-based Diversity

Hill Diversity

$${}^qD(\mathbf{p}) = \left(\sum_{i=1}^S p_i^q \right)^{1/(1-q)}$$

$$q = 0$$

$${}^0D(\mathbf{p}) = \text{Species Richness}$$

$$q \rightarrow 1$$

$${}^1D = \exp(\text{Shannon})$$

$$q = 2$$

$${}^2D = \text{inverse Simpson}$$

Similarity-based Diversity (Leinster and Cobbold, 2012)

Similarity-based Diversity

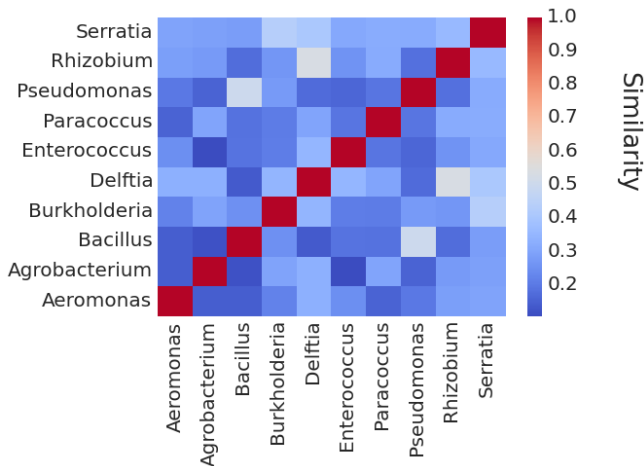
$${}^qD^{\mathbf{Z}}(\mathbf{p}) = \left(\sum p_i(\mathbf{Zp})_i^{q-1} \right)^{\frac{1}{1-q}},$$

where q is the community's order and $q \neq 1$ (but its limit $q \rightarrow 1$ exists) so that

$$(\mathbf{Zp})_i = \sum_{j=1}^S Z_{ij} p_j.$$

Example Similarity Matrix

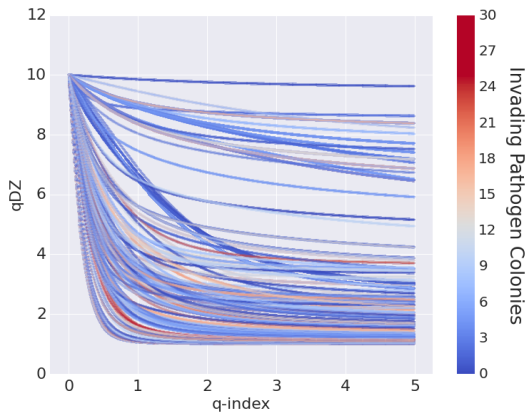
**Example Similarity matrix: co-occurrence in ecosystems
(Chaffron et al., 2010)**



Similarity-based Diversity

Similarity-based diversity for invasion experiments

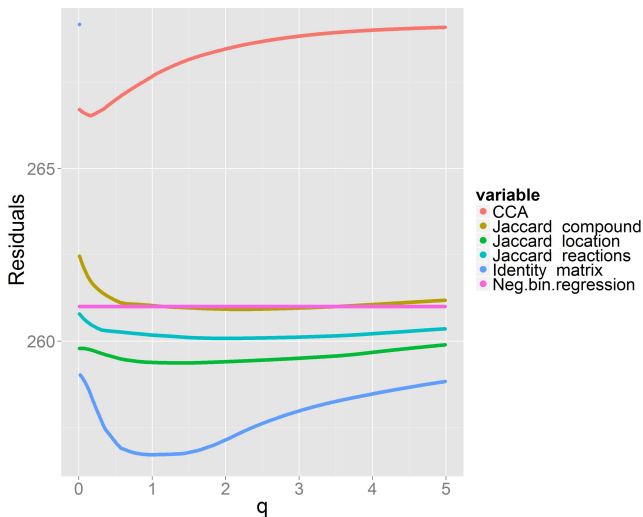
Similarity-based diversity with identity matrix



- less diverse communities have more invasion
- $q \rightarrow 0$: sensitive to rare species
- $q \rightarrow \infty$: stresses relative abundance

Similarity-based Diversity

Finding optimal q for predicting invasion potential

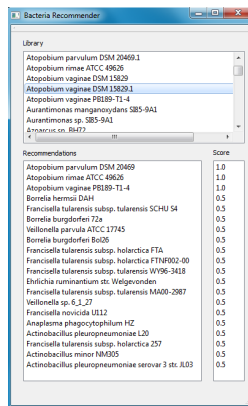


- Shannon diversity is best
- No similarity adequately describes invasion

Recommender Engine

"Bacteria Who Like These bacteria Also Like ... "

- Finding suitable matches between bacteria
- Based on co-occurrence in ecosystems (Chaffron et al., 2010)
- Assembling communities very intensive
- A recommender engine can decrease workload



Conclusion

Similarity-based index

- Similarity-based diversity helps the statistical analysis of pathogen invasion
- Shannon diversity is most optimal to describe pathogen invasion
- Similarity information can help in design of experiments

Recommender Engine

- Recommendations can decrease workload in the lab

Thank you for your attention!

References I

- Samuel Chaffron, Hubert Rehrauer, Jakob Pernthaler, and Christian von Mering. A global network of coexisting microbes from environmental and whole-genome sequence data. *Genome research*, 20(7):947–959, 2010.
- Tom Leinster and Christina A Cobbold. Measuring diversity: the importance of species similarity. *Ecology*, 93(3):477–489, 2012.