

Analyzing determinants of gut microbiota using interaction network methods

François Massol, Emilie Macke, Martijn Callens,
Luc De Meester, Ellen Decaestecker

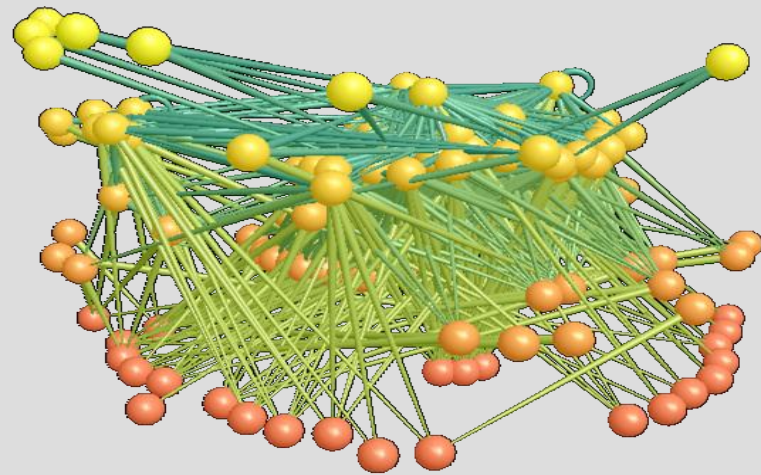


The network era of ecology

Now mapping more ecological objects as networks

Increase in the amount of network data:

- field data more open
- NGS/metabarcoding data



- How to analyze this data?
- Can we transpose methods to other types of networks?

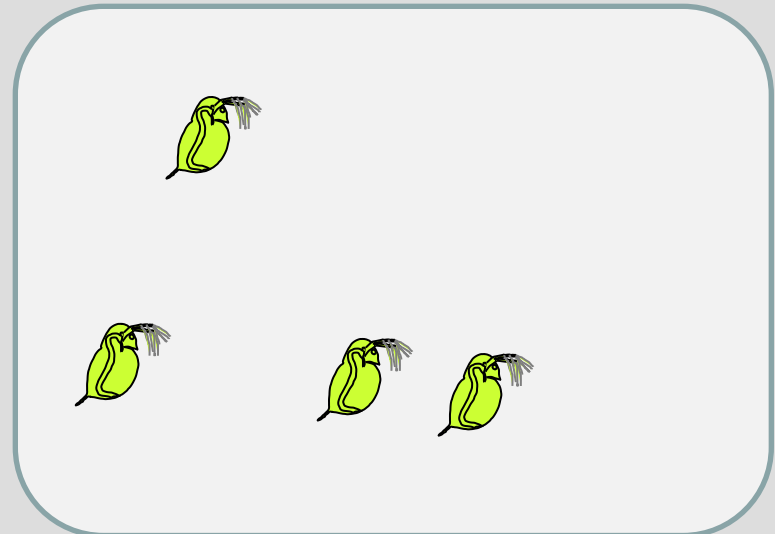
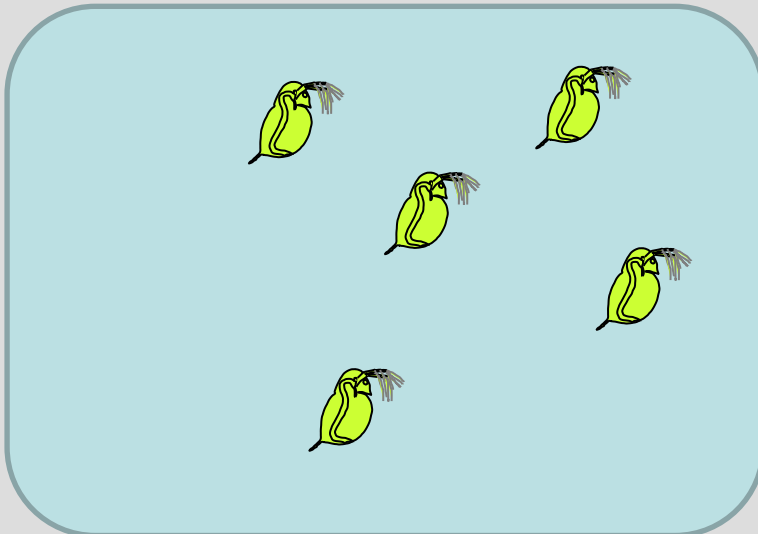
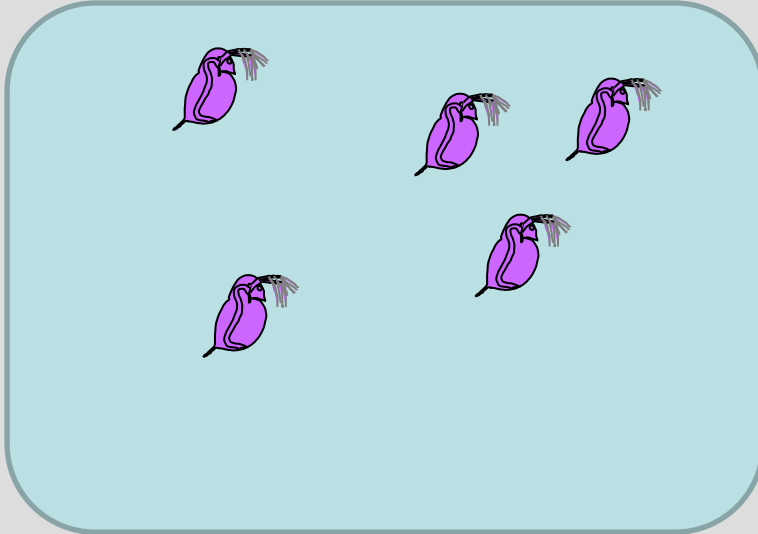
An example: *Daphnia magna*'s microbiotas



Diet



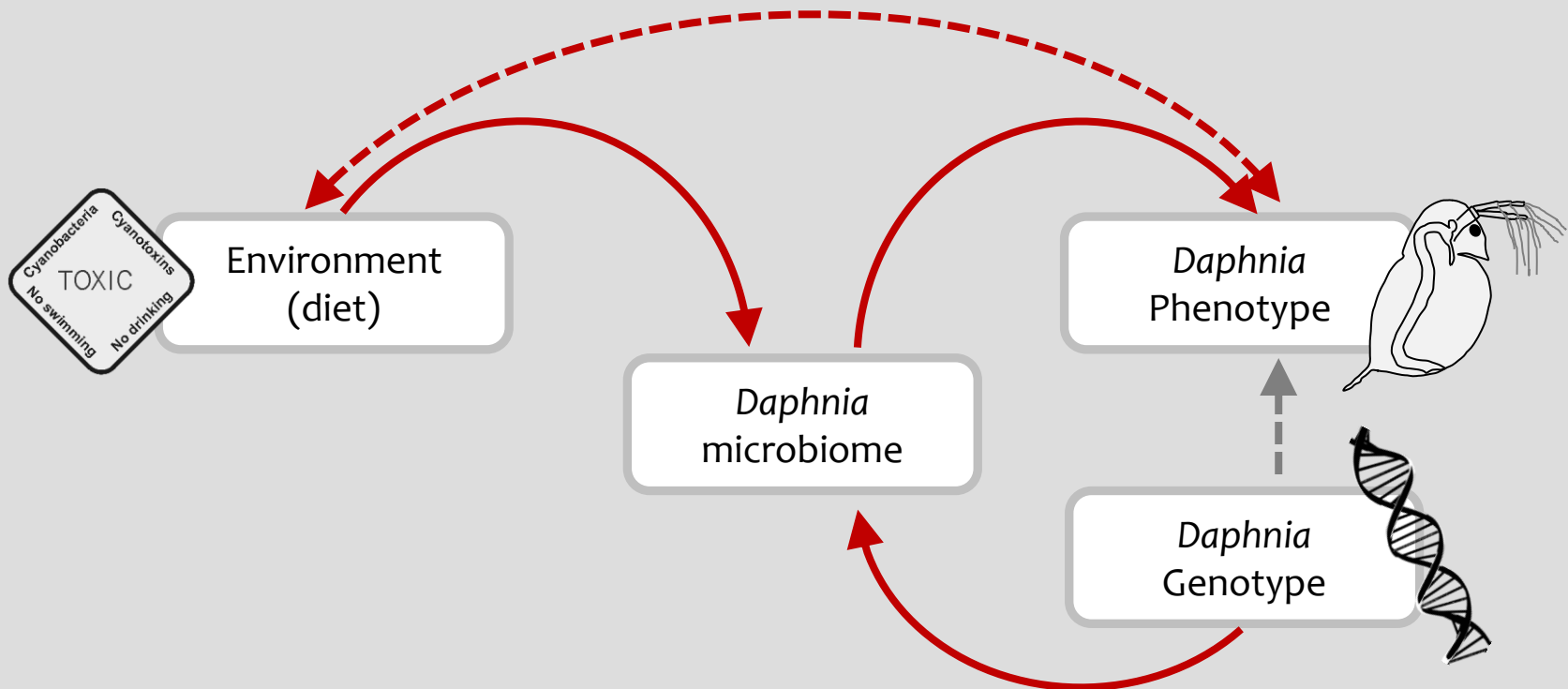
Daphnia genotype



Surviving a toxic diet

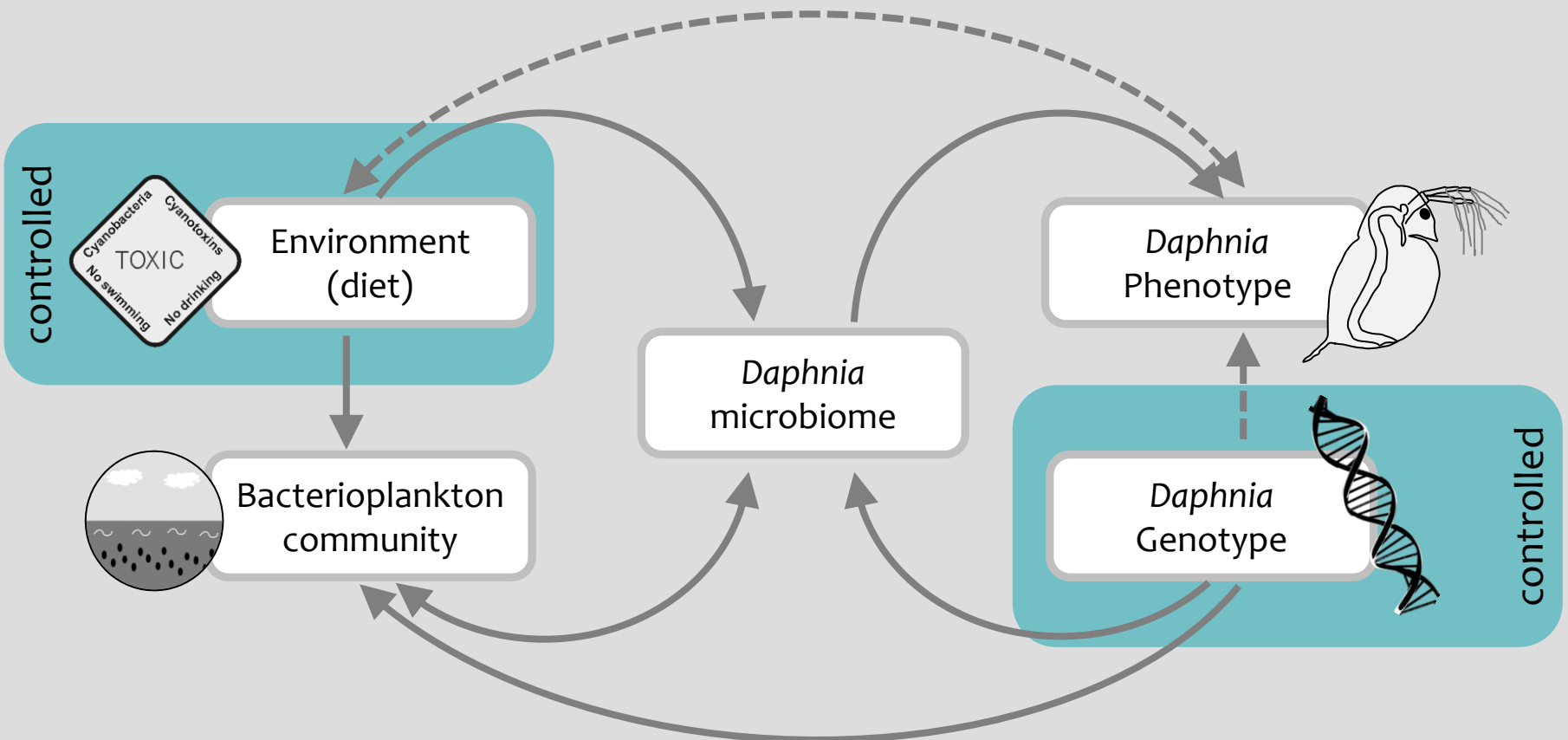
Earlier results (cross-inoculation):

- gut microbiota controls adaptation to toxic diet
- genotype affects recruitment of adapted microbiota



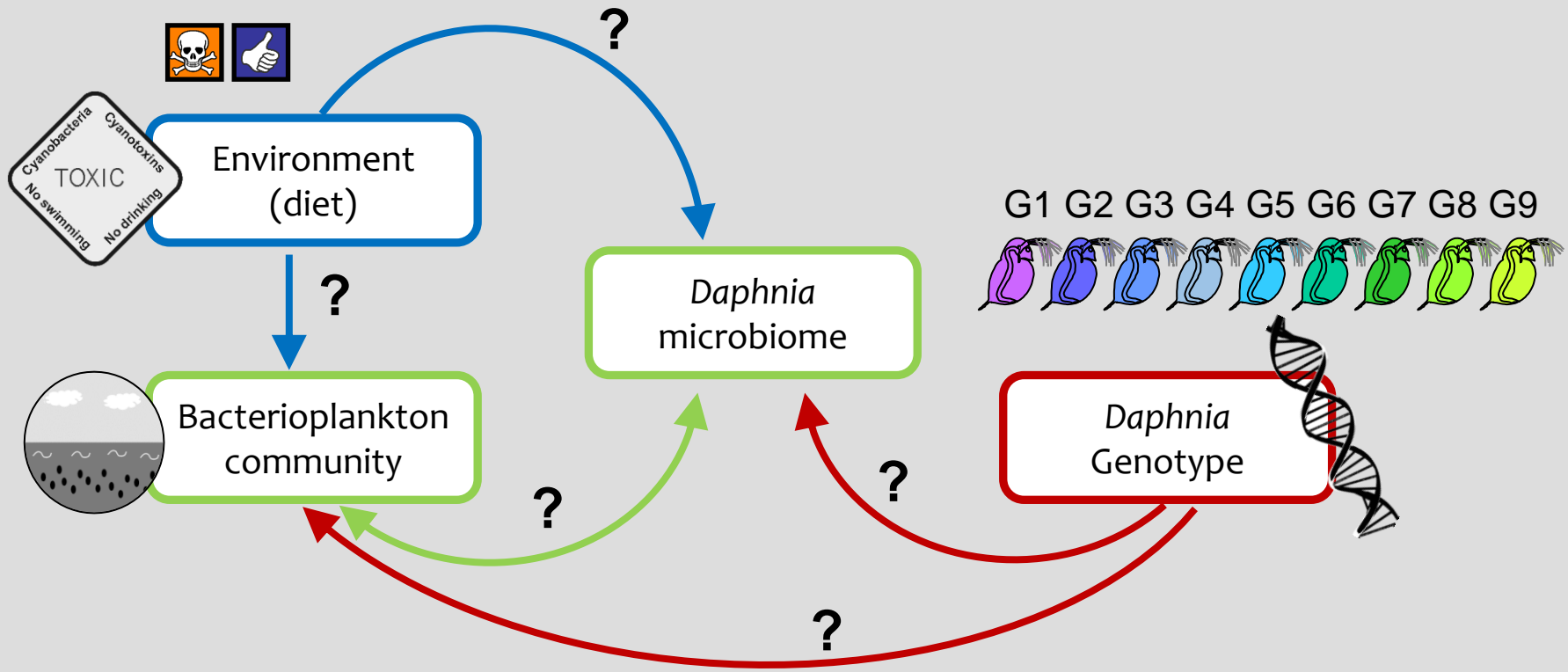
General question

Disentangling the links between host (*Daphnia magna*) genotype, diet, their gut microbiome and the bacterioplankton community



Specific biological question

What determines interactions between *Daphnia* hosts and their two microbiotas (gut microbiota and bacterioplankton)?





E. Macke



Genotypes exposed to a cyanobacteria diet



Genotypes exposed to a green algae diet

Microcystis diet



Scenedesmus diet



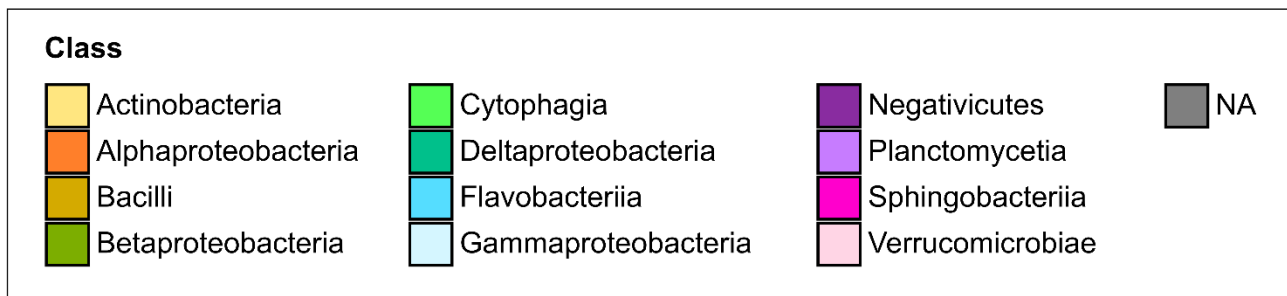
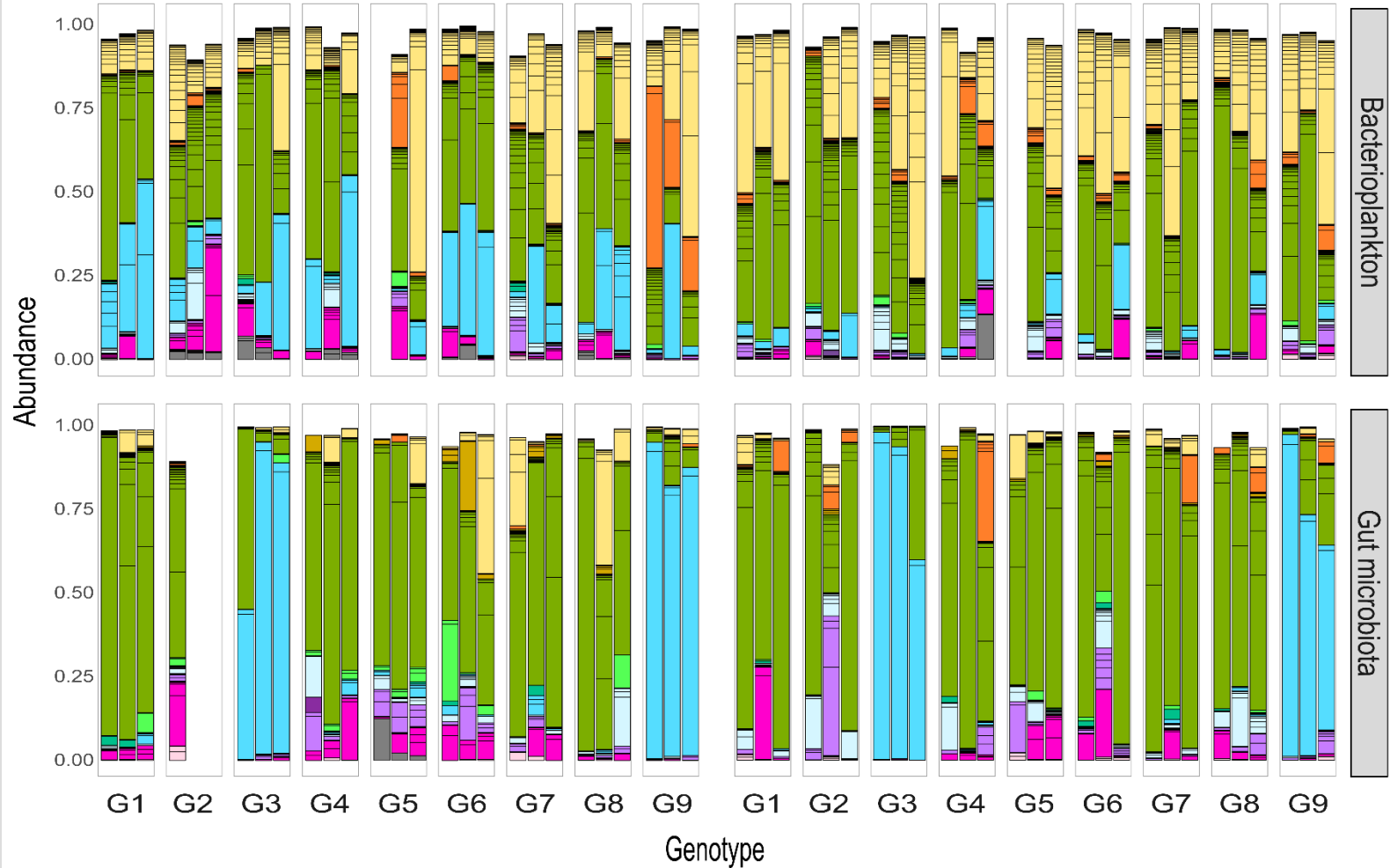
After 1 year of exposure:

Sequencing on
gut microbiota +
bacterioplankton

9 genotypes x 2 diets x 3 replicates

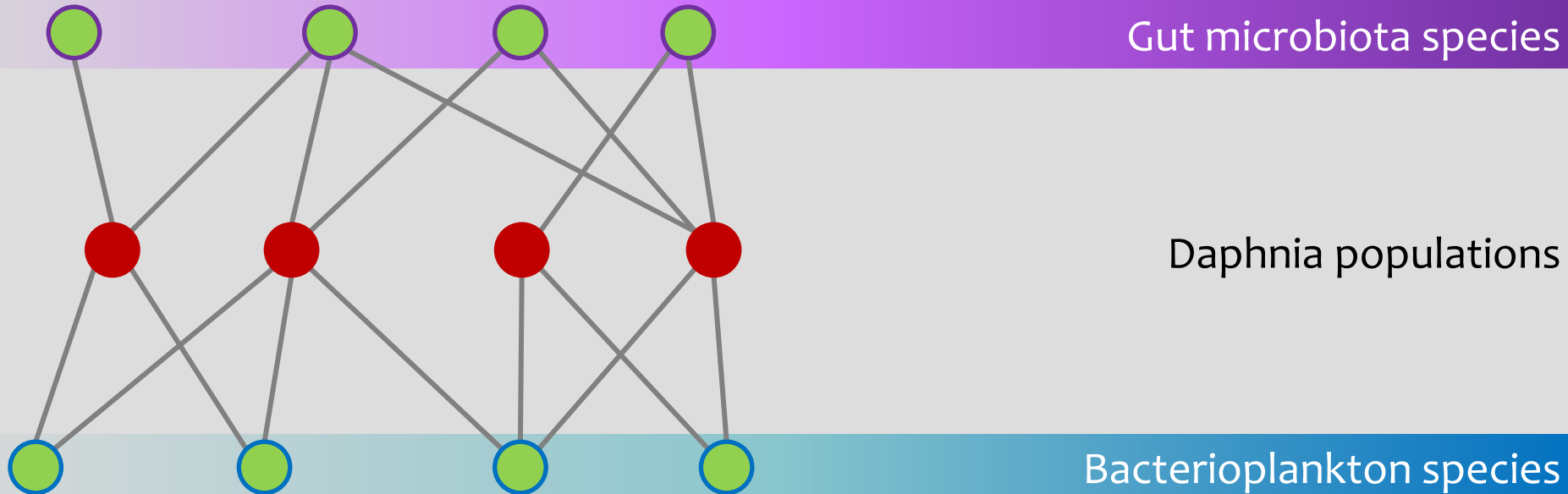
Genotypes exposed to a cyanobacteria diet

Genotypes exposed to a green algae diet



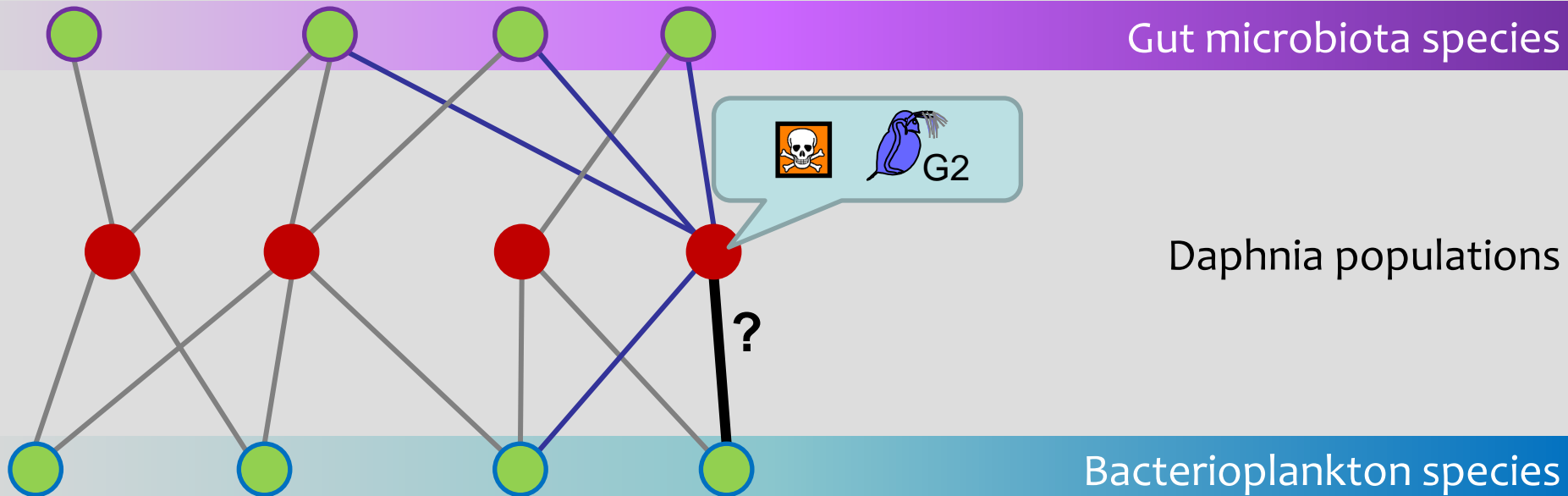
The data

Daphnia-Gut microbiota-Bacterioplankton
= tripartite (or doubly bipartite) network



The question

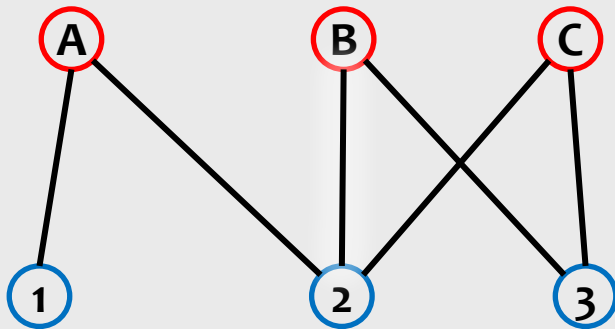
What determines interactions?



Main issue: interactions are not independent

An inferential framework

1. A network is equivalent to a matrix



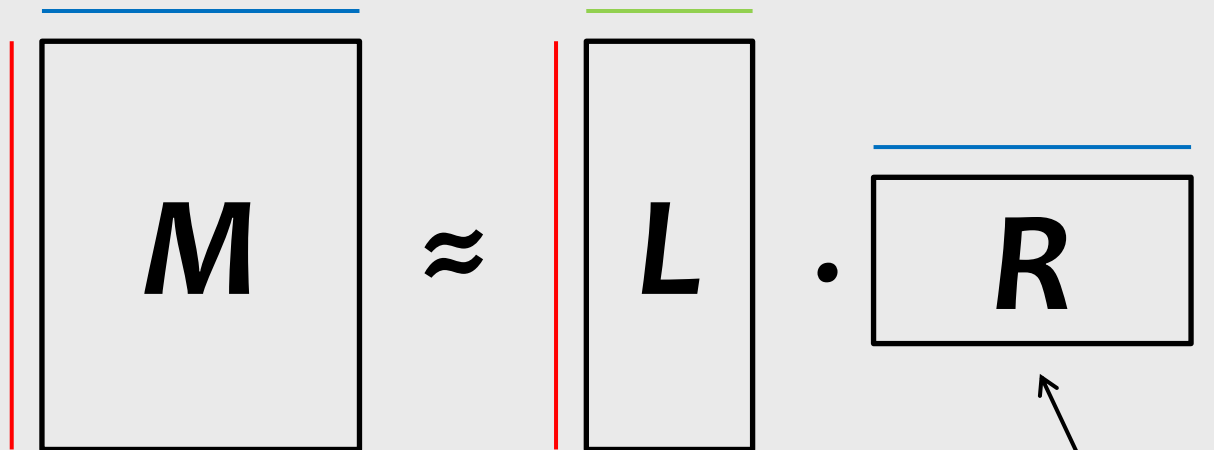
	1	2	3
A	1	1	0
B	0	1	1
C	0	1	1

The incidence matrix

An inferential framework

2. The matrix can be approximated

Idea: approximate the incidence matrix as the product of independent matrices



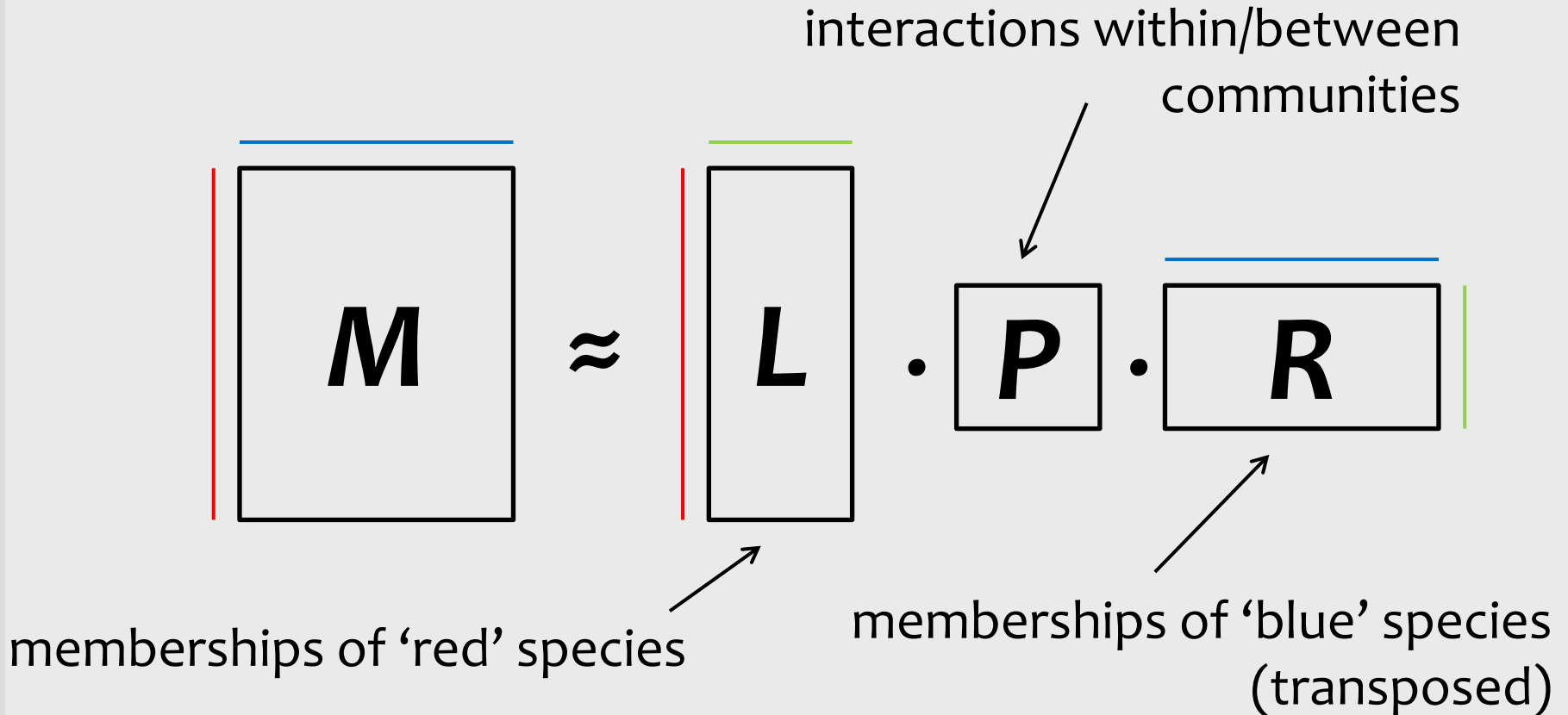
latent traits of 'red' species

latent traits of 'blue' species
(transposed)

Approximation #1

Using community memberships

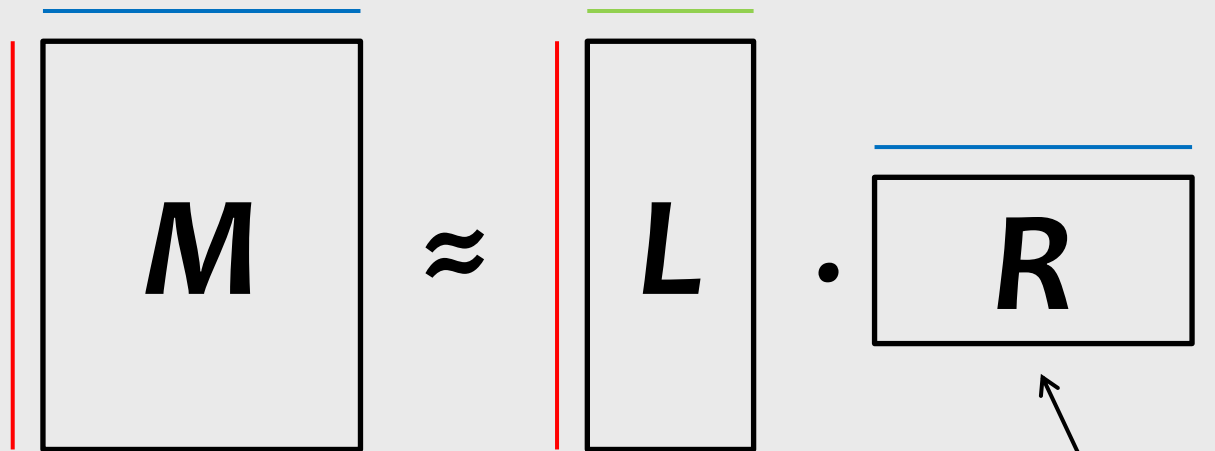
can be performed with 'modules' or 'blocks'



Approximation #2

Singular value decomposition

very similar to a PCA



latent traits of 'red' species

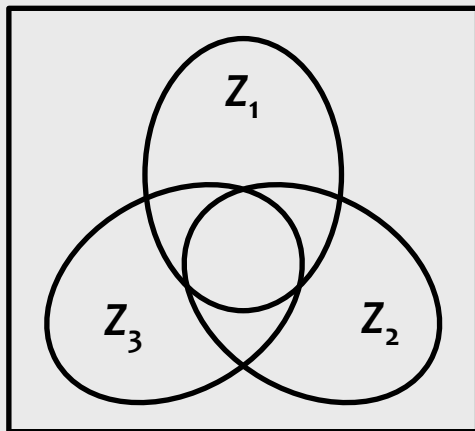
latent traits of 'blue' species
(transposed)

Explanations = Dalla Riva & Stouffer 2016 *Oikos*

An inferential framework

3. Each approximation matrix (L and R) can be 'regressed' on external variables (redundancy or canonical correspondence analyses)

$$\boxed{X} \sim \boxed{Z_1} + \boxed{Z_2} + \boxed{Z_3}$$



R^2 / χ^2 explained by the different fractions (e.g. $Z_1 \mid Z_2 + Z_3$)

Testable in two ways:

- row permutation (classic RDA test)
- graph permutation (configuration model)

How do we choose the vectors?

- In orthogonal explanatory matrices:
 - method of Blanchet et al. (2008)

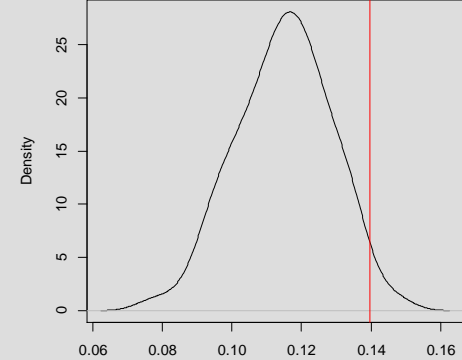
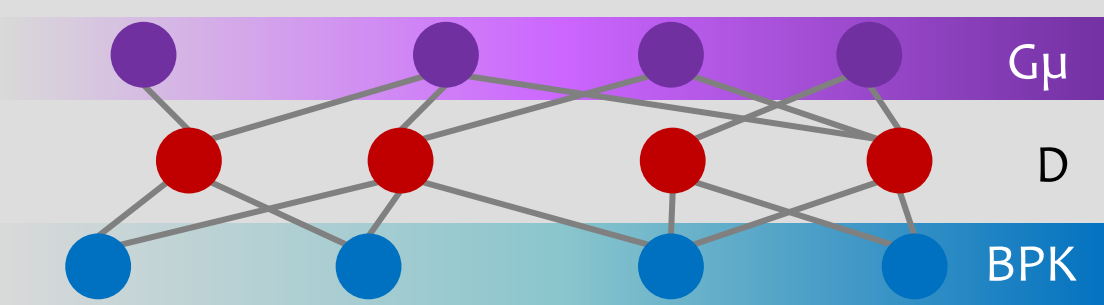
Principle: forward search with double stopping criterion (after checking that the whole matrix explains something significant...)

- adjusted R^2 cannot be higher than that found with all vectors
- the last added vector must contribute significantly to the adjusted R^2

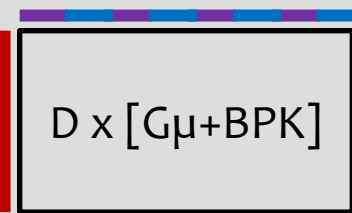
How do we choose the vectors?

- In orthogonal explained matrices
... depends on the question / what we want to use the approximation for

In our case (Daphnia): keeping enough SVectors to 'recover' an incidence matrix that has similar communities

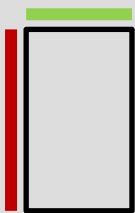


1



2

SVD approx

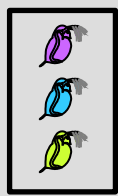


3

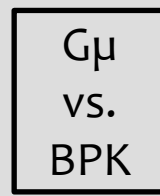
RDA



diet

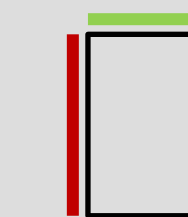


genotype



sample type

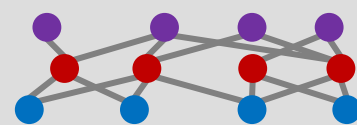
$\times 10^4$



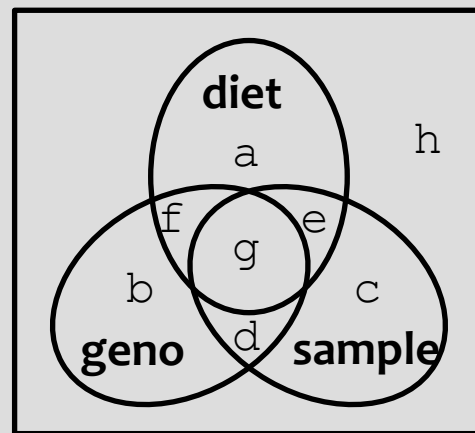
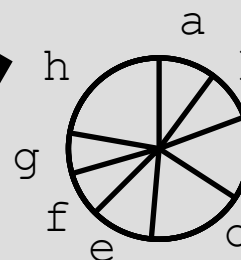
residual permutation

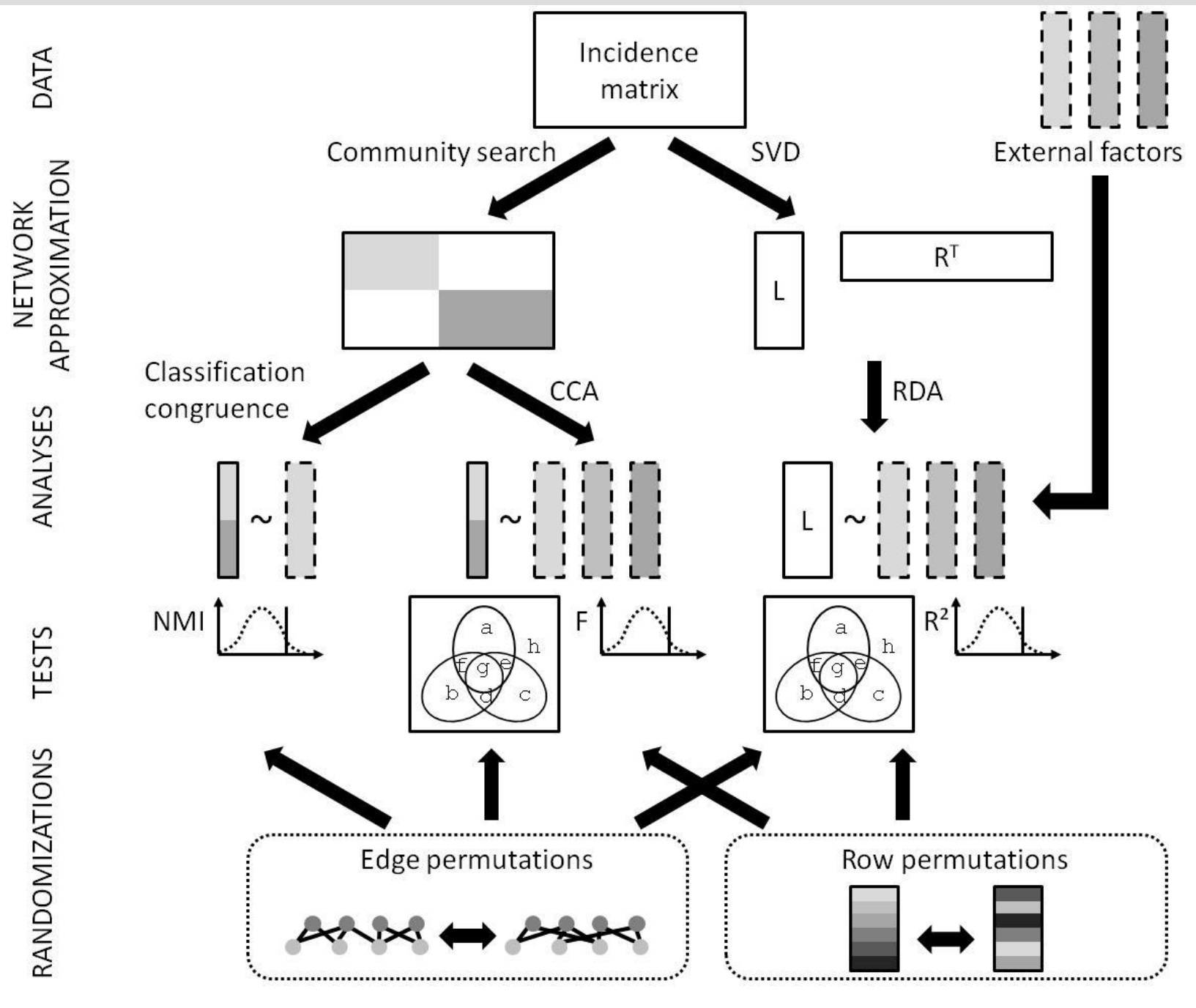
test of fractions

$\times 10^4$

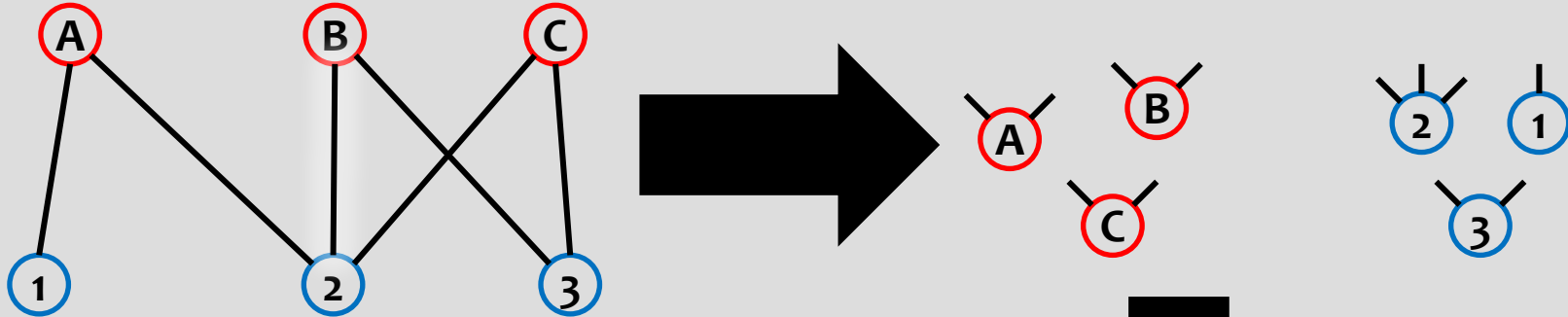


edge permutation



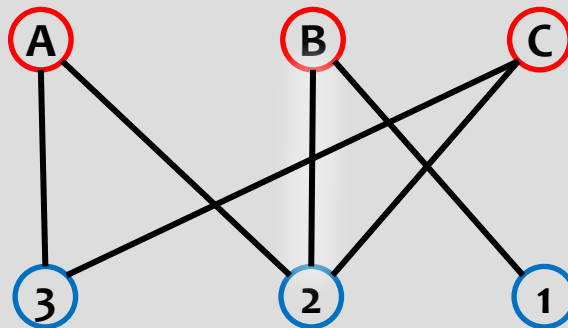


Graph permutations

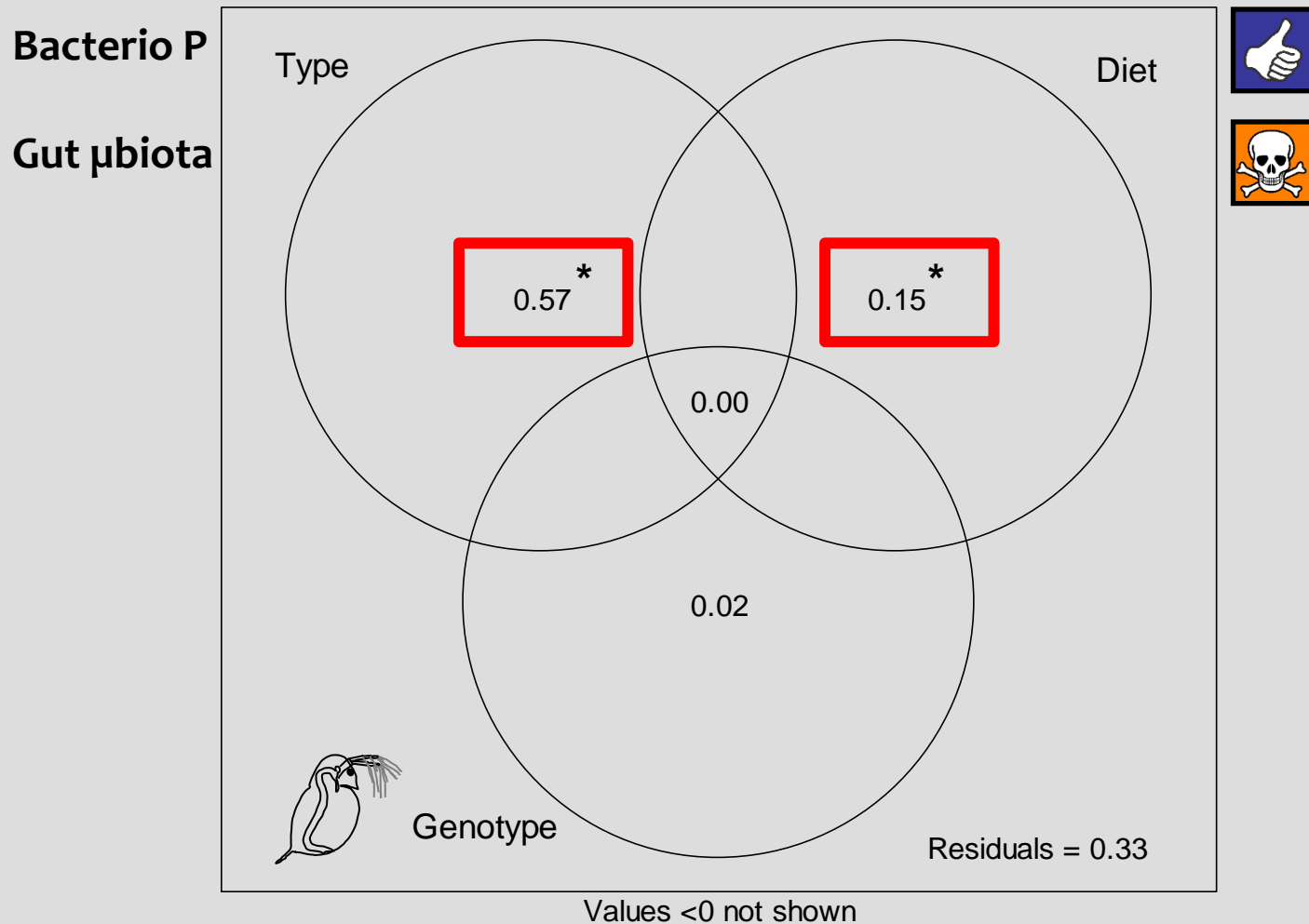


Also called ‘configuration model’

Can be performed efficiently using the ‘curveball’ algorithm



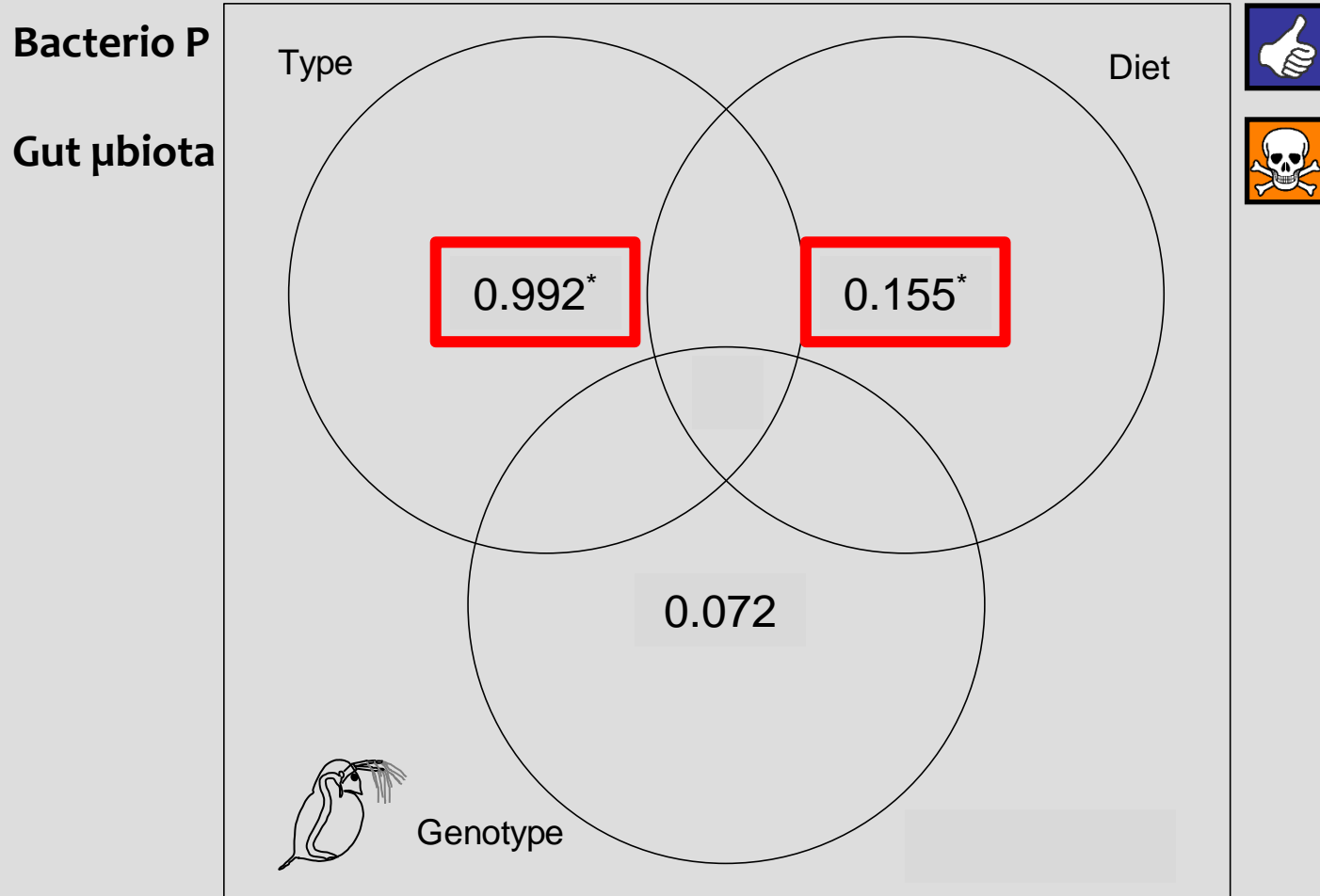
SVD-RDA on whole network



3 SVectors kept

Type and Diet fractions significant (row permutations and edge permutations)

Modules-CCA on whole network

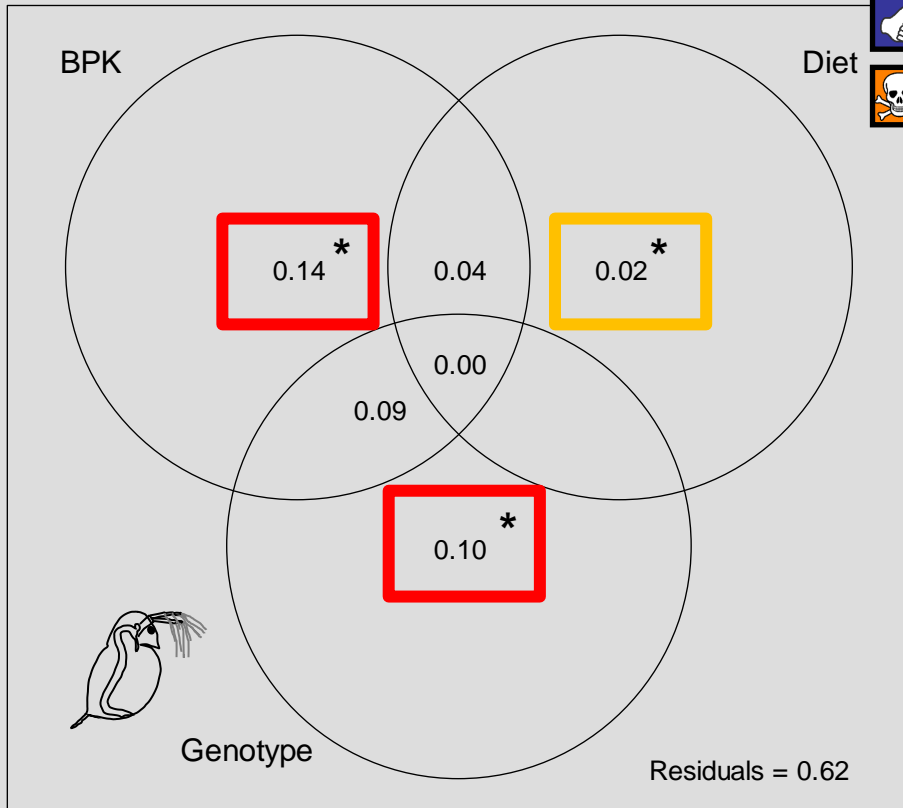


χ^2 values

Type and Diet fractions significant (row permutations and edge permutations)

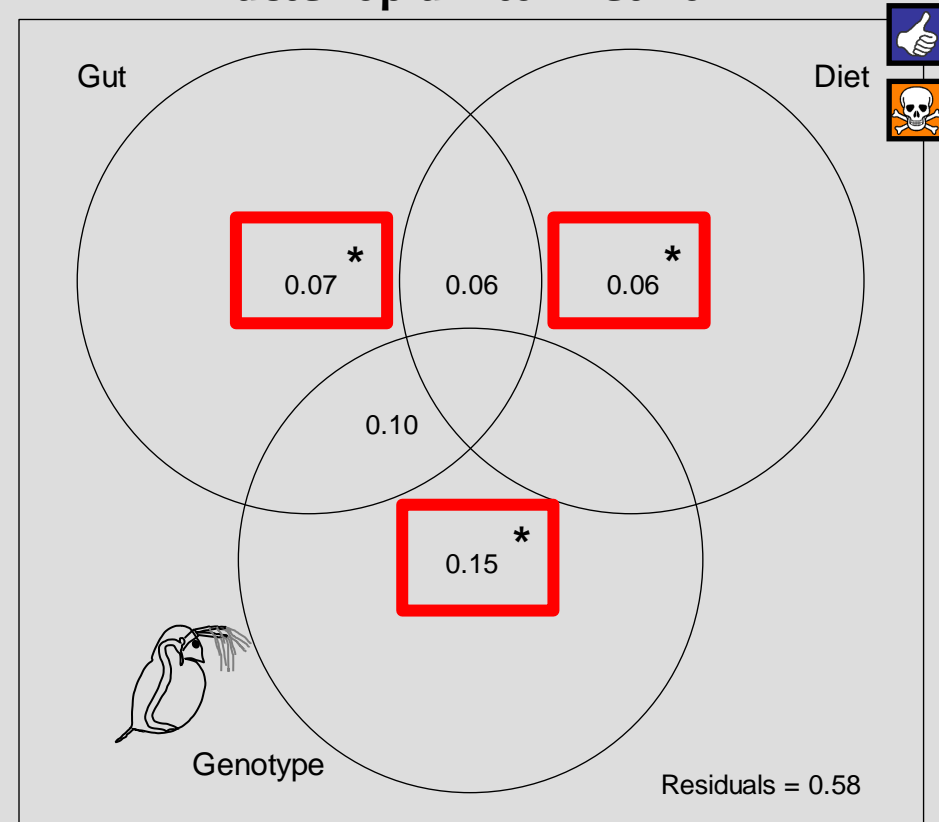
SVD-RDA on sub-networks

Gut microbiota network



Values <0 not shown

Bacterioplankton network



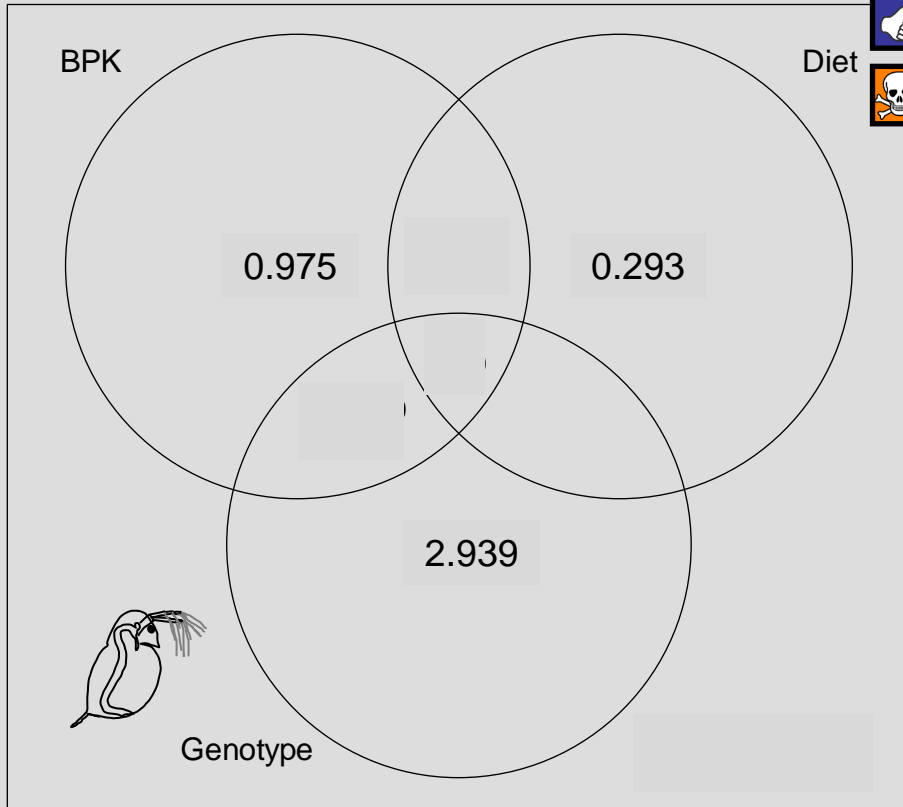
Values <0 not shown

10 Sectors kept
Genotype & BPK doubly significant
Diet only with row permutations

11 Sectors kept
All effects doubly significant

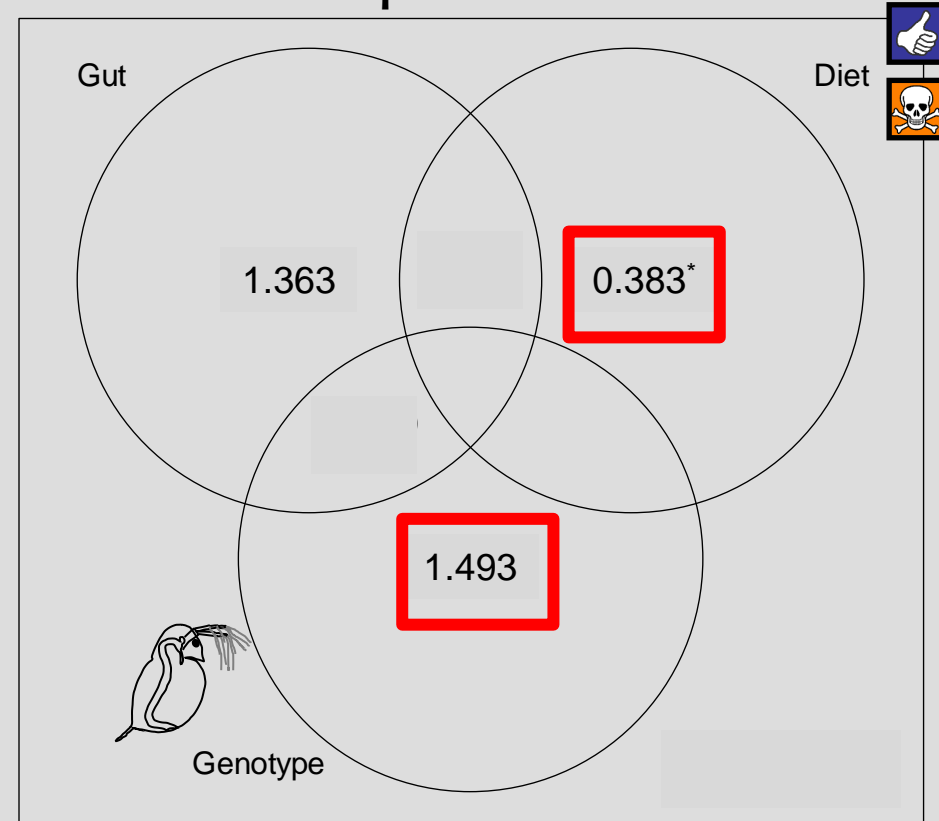
Modules-CCA on sub-networks

Gut microbiota network



Not significant

Bacterioplankton network



Diet and genotype effects doubly significant

Take-home message (method)

Network data can be decomposed to remove dependencies and then effects can be tested through RDA/CCA:

- allows disentangling the effects of external drivers
- two tests => effects due to variation in degrees or more generally variation in network structure

Take-home messages (biology)

- Bacterioplankton and Gut microbiota are different communities
- Diet selects BPK but not Gut microbiota
- *Daphnia* genotype selects Gut microbiota (but at a scale finer than modules) and BPK
- BPK and gut microbiota have some reciprocal effects, but only perceived below the scale of modules

Perspectives / Discussion

- Understanding genomic differences between *Daphnia* genotypes, in particular regarding immune genes
- Effect of diet on BPK: introducing different bacteria with different foods?
- Disentangling causes and consequences: merging present approach with structural equation models?

Acknowledgements & Advertisement

Co-authors

Emilie Macke, Martijn Callens, Ellen Decaestecker

Funding sources



DOI: 10.1111/1365-2656.13297

ANIMAL SOCIAL NETWORKS
Research Article

Journal of Animal Ecology

A methodological framework to analyse determinants of host-microbiota networks, with an application to the relationships between *Daphnia magna*'s gut microbiota and bacterioplankton

François Massol^{1,2} | Emilie Macke³ | Martijn Callens^{3,4} | Ellen Decaestecker³

Submit your preprints to PCI! (<https://peercommunityin.org>)



- Independent of publishing companies – made by and for researchers
- Totally free for authors and readers
- Transparent: reviews, data, script, methods of recommended preprints are public
- Not exclusive: can be submitted to journal afterwards
- Efficient: more than 30 journals may use PCI evaluations
- Sponsored by > 50 research institutions

