SMT Solving and modeling for biology

Denis Kuperberg
CNRS, LIP, ENS Lyon

GdT Plume
11/12/2023

Part 1

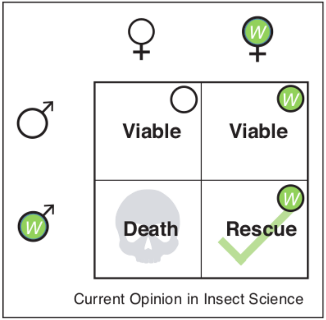
Wolbachia-induced infertility in mosquitoes

With
Sylvain Charlat, Alice Namias , Mathieu Sicard, Mylène Weill

LBBE LABORATOIRE DE BIOMÉTRIE ET BIOLOGIE ÉVOLUTIVE

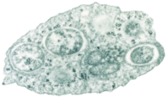
isem Institute of Evolutionary Science of Montpellier

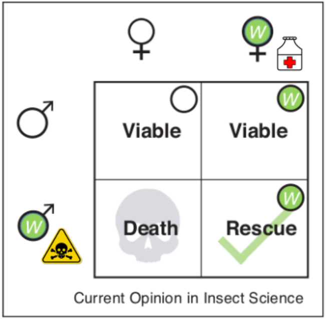
Cytoplasmic Incompatibility



Credit: [Namias et al 2022]

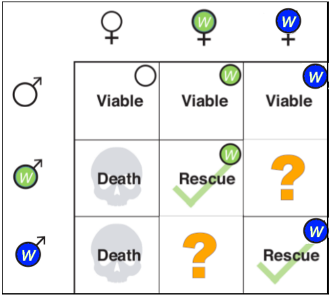
The toxin/antidote hypothesis



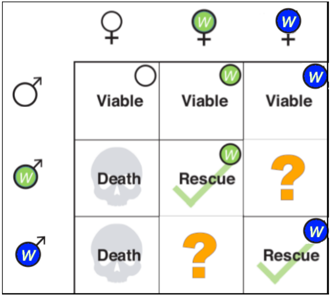


Credit: [Namias et al 2022]

Different Wolbachia strains



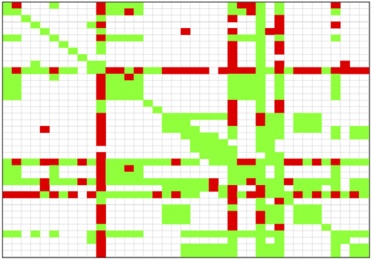
Different Wolbachia strains



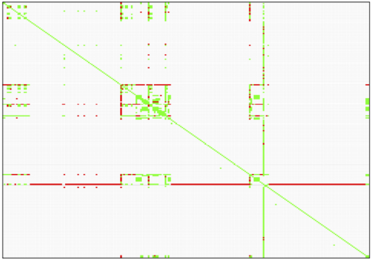
Hypothesis: each strain has its own cocktail.

The data

39 lines with phenotypic and molecular data



239 lines with phenotypic data



Optimizing the parameters

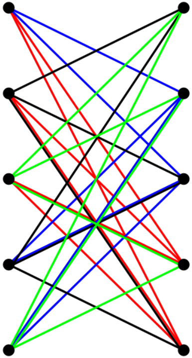
How many kind of toxins needed ?

Optimizing the parameters

How many kind of toxins needed ?

NP-complete [Nor et al 2012]:

biclique cover

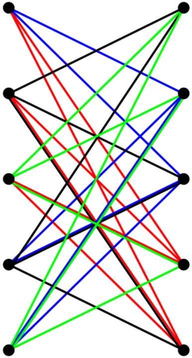


Optimizing the parameters

How many kind of toxins needed ?

NP-complete [Nor et al 2012]:

biclique cover



Variant: quantitative model

Using SAT/SMT Solver

Before SAT solving [Nor et al 2012]:

- ad-hoc heuristics for 19*19 matrix
- fill missing data manually
- approximate results

Using SAT/SMT Solver

Before SAT solving [Nor et al 2012]:

- ad-hoc heuristics for 19*19 matrix
- fill missing data manually
- approximate results

With minisat/Z3 (this work):

- deal with 239*239 matrix
- robust to missing data
- explore various models and questions

Results

Boolean

Toxs per strain	Types
1	14
2	10
3	9

Quantitative

Levels	Types
1	14
2	7
3	5

Results

Boolean

Toxs per strain	Types
1	14
2	10
3	9

Quantitative

Levels	Types
1	14
2	7
3	5

And more:

- test robustness
- predict missing data
- find cells discriminating models

Perspectives

- Confront predictions with new data
- Prescribe tests to discriminate models
- Find correlations with genetic data
- Evolutionary explanations

Part 2

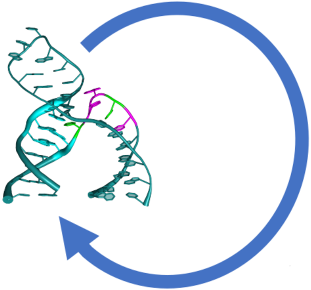
Detection of autocatalytic cycles

With
Sylvain Charlat, Etienne Rajon, Nicolas Lartillot

LBBE LABORATOIRE DE BIOMÉTRIE ET BIOLOGIE ÉVOLUTIVE

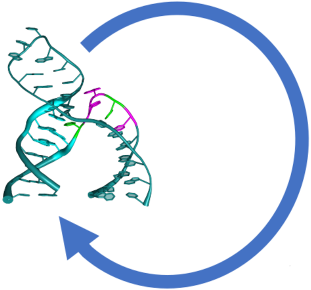
Origin of Life dichotomy

Replicator first

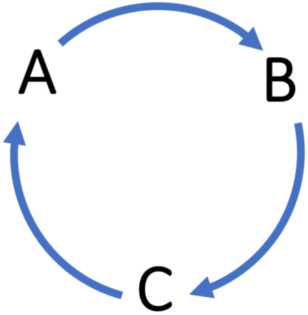


Origin of Life dichotomy

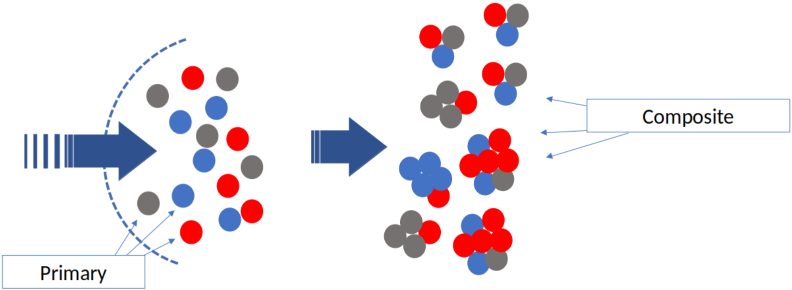
Replicator first



Metabolism first

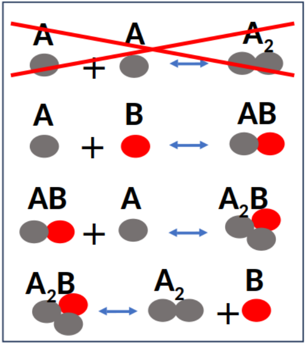


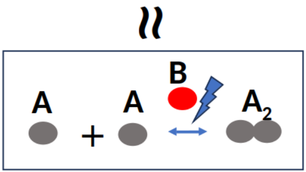
Soup of chemical reactions



Spontaneous (auto)catalysis

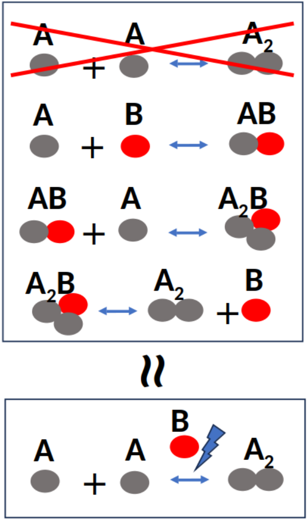
Catalysis



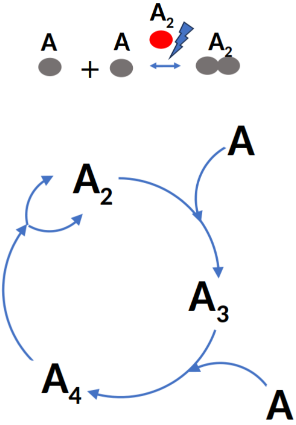


Spontaneous (auto)catalysis

Catalysis



Autocatalysis



Reaction matrix

$$
\begin{array}{ll}
(R_1) & A + B \rightleftharpoons AB \\
(R_2) & A + A \rightleftharpoons AA
\end{array}
$$

Reaction matrix

$$
\begin{array}{ll}
(R_1) & A + B \rightleftharpoons AB \\
(R_2) & A + A \rightleftharpoons AA
\end{array}
$$

	R_1	R_2
A	-1	-2
B	-1	0
AA	0	1
AB	1	0

Reaction matrix

$$(R_1) \quad A + B \rightleftharpoons AB$$
$$(R_2) \quad A + A \rightleftharpoons AA$$

	R_1	R_2
A	-1	-2
B	-1	0
AA	0	1
AB	1	0

$\vec{v} \in \mathbb{R}^2$ flow vector $\Rightarrow$ $M \cdot \vec{v}$ balance for each entity.

Autocatalysis via matrices

Autocatalytic core [Blockhuis et al 2022]:
Submatrix N of M such that

- each column and line contains coef< 0 and coef> 0
- $\exists \vec{v} \in \mathbb{R}^k$ such that $N \cdot \vec{v} \in (\mathbb{R}^{*+})^k$
- N minimal

Autocatalysis via matrices

Autocatalytic core [Blockhuis et al 2022]:
Submatrix N of M such that

- each column and line contains coef< 0 and coef> 0
- $\exists \vec{v} \in \mathbb{R}^k$ such that $N \cdot \vec{v} \in (\mathbb{R}^{*+})^k$
- N minimal

Example: $A_2 \to A_3 \to A_4 \to A_2 + A_2$

A_2	-1	0	2
A_3	1	-1	0
A_4	0	1	-1

Autocatalysis via matrices

Autocatalytic core [Blockhuis et al 2022]:
Submatrix N of M such that

- each column and line contains coef< 0 and coef> 0
- $\exists \vec{v} \in \mathbb{R}^k$ such that $N \cdot \vec{v} \in (\mathbb{R}^{*+})^k$
- N minimal

Example: $A_2 \to A_3 \to A_4 \to A_2 + A_2$

A_2	-1	0	2
A_3	1	-1	0
A_4	0	1	-1

$$N \cdot \begin{pmatrix} 6 \\ 5 \\ 4 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}$$

Autocatalysis via matrices

Autocatalytic core [Blockhuis et al 2022]:
Submatrix N of M such that

- each column and line contains coef< 0 and coef> 0
- $\exists \vec{v} \in \mathbb{R}^k$ such that $N \cdot \vec{v} \in (\mathbb{R}^{*+})^k$
- N minimal

Example: $A_2 \to A_3 \to A_4 \to A_2 + A_2$

A_2	-1	0	2
A_3	1	-1	0
A_4	0	1	-1

$$N \cdot \begin{pmatrix} 6 \\ 5 \\ 4 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}$$

Goal: Find autocatalytic cores in a chemical soup.

Autocatalysis via matrices

Autocatalytic core [Blockhuis et al 2022]:
Submatrix N of M such that

- each column and line contains coef< 0 and coef> 0
- $\exists \vec{v} \in \mathbb{R}^k$ such that $N \cdot \vec{v} \in (\mathbb{R}^{*+})^k$
- N minimal

Example: $A_2 \to A_3 \to A_4 \to A_2 + A_2$

A_2	-1	0	2
A_3	1	-1	0
A_4	0	1	-1

$$N \cdot \begin{pmatrix} 6 \\ 5 \\ 4 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}$$

Goal: Find autocatalytic cores in a chemical soup.

NP-complete ?

From concentrations to flows

$$A + B \rightleftharpoons AB$$

How to compute flow ?

From concentrations to flows

$$A + B \rightleftharpoons AB$$

How to compute flow ?

- $k_{\text{assoc}} = [A][B]$
- $k_{\text{dissoc}} = [AB]$
- flow $= k_{\text{assoc}} - k_{\text{dissoc}}$

From concentrations to flows

$$A + B \rightleftharpoons AB$$

How to compute flow ?

- $k_{\text{assoc}} = [A][B]$
- $k_{\text{dissoc}} = [AB]$
- flow $= k_{\text{assoc}} - k_{\text{dissoc}}$

Consistent core: Flow witness must be realistic

Concentration vector $\vec{u} \rightsquigarrow$ flow vector $\vec{v}$.

From concentrations to flows

$$A + B \rightleftharpoons AB$$

How to compute flow ?

- $k_{\mathrm{assoc}} = [A][B]$
- $k_{\mathrm{dissoc}} = [AB]$
- flow $= k_{\mathrm{assoc}} - k_{\mathrm{dissoc}}$

Consistent core: Flow witness must be realistic

Concentration vector $\vec{u} \rightsquigarrow$ flow vector $\vec{v}$.

Compatible cores: share the same witness

What we want

- design systems
- sample systems
- find autocatalytic cores
- mark consistent cores
- find sets of compatible cores

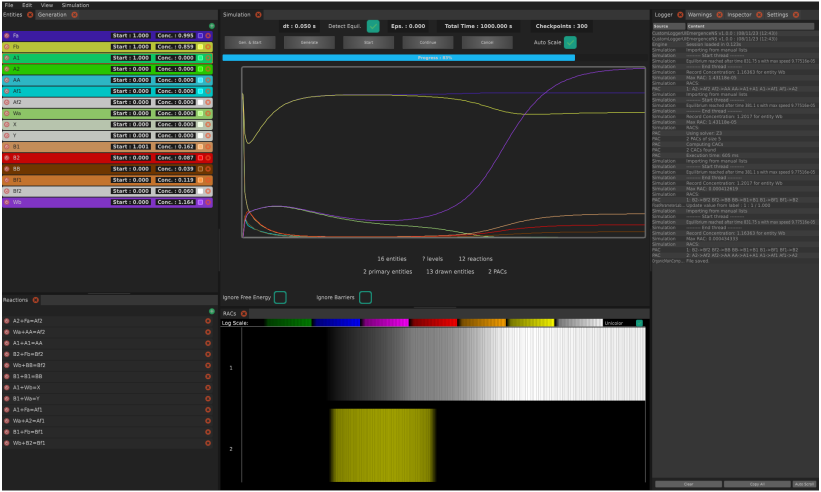
What we want

- design systems
- sample systems
- find autocatalytic cores
- mark consistent cores
- find sets of compatible cores

SMT Solver Z3

JUCE OrganicUI

The program


File
Edit
View
Simulation
Entities
Generation
Fa
Fb
A1
A2
AA
Af1
Af2
Wa
X
Y
B1
B2
BB
Bf1
Bf2
Wb
Start : 1.000
Conc. : 0.995
Conc. : 0.859
Start : 1.001
Conc. : 0.162
Conc. : 0.087
Conc. : 0.039
Conc. : 0.119
Conc. : 0.060
Conc. : 1.164
Reactions
A2+Fa=Af2
Wa+AA=Af2
A1+A1=AA
B2+Fb=Bf2
Wb+BB=Bf2
B1+B1=BB
A1+Wb=X
B1+Wa=Y
A1+Fa=Af1
Wa+A2=Af1
B1+Fb=Bf1
Wb+B2=Bf1
Simulation
dt : 0.050 s
Detect Equil.
Eps. : 0.000
Total Time : 1000.000 s
Checkpoints : 300
Auto Scale
16 entities
7 levels
12 reactions
2 primary entities
13 drawn entities
2 PACs
Ignore Free Energy
Ignore Barriers
RACs
Log Scale:
Unicolor
Logger
Warnings
Inspector
Settings

Perspectives

- Spectrum of interactions between cores
- Quantitative analysis (expectancy of autocatalysis,...)
- Links with multiplicity of equilibria
- Identification of scales of individuality
- Spatialisation

Perspectives

- Spectrum of interactions between cores
- Quantitative analysis (expectancy of autocatalysis,. . .)
- Links with multiplicity of equilibria
- Identification of scales of individuality
- Spatialisation

Very long term goals:

- Quantify natural selection
- Criteria for life
- Possible build-up scenario

Thank you !