

Ecole de Porquerolles

Modélisation individu-centrée de systèmes biologiques complexes

Application à la simulation de l'évolution de réseaux génétiques bactériens

Guillaume Beslon

INSA – INRIA – LIRIS – IXXI



Introduction

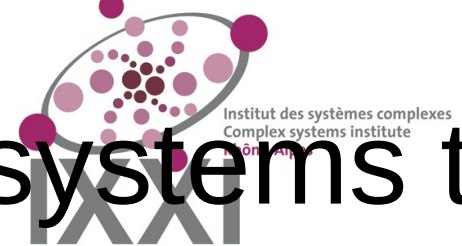
- Aim of the course (first part):
 - Modélisation individu-centrée de systèmes biologiques complexes
 1. Complexes ?
 2. Systèmes Biologiques complexes ?
 3. Modélisation individu-centrée ?
 4. Modélisation ?
- Who am I?
 - Guillaume BESLON (guillaume.beslon@liris.cnrs.fr)
 - Professor at the INSA-Lyon, LIRIS Lab. (Laboratoire d'Informatique en Image et Systèmes d'Information),
 - Head of the INRIA COMBINING Team
 - Director of the IXXI (Rhône-Alpes Complex Systems Institute)
 - Research topics: Individual-based modeling of complex biological systems (mainly evolution)

What is a complex system?

- Numerous definitions but general agreement on:
 - The structure of the system (“many elements”)
 - Some subjective judgment (not always clearly accepted)
 - Something “emerges” (but no general agreement on “what is emergence”)
 - Something is dynamic and “self-organized”



“A system is a complex system if it is made of a large number of interacting elements and if the dynamics of these interactions govern the behavior of the system, giving to it an appearance of unity from the point of view of an external observer.”



From complex systems to the science of complex systems

- We can (more or less) define a “complex system” but what is the “science of complex systems”
 - ~any system is a complex system at some levels of description
 - Does it implies that “science of complex systems = science”?
 - Hope you’ll agree that it is absurd! (at least)
- How can you define a science?
 - E.g., Biology, Chemistry and Physics are all working on DNA
 - A science is not defined by its objects but rather by its questions
- The science of complex systems is **NOT** the science of systems that are complex!
 - It is the science of questions that are specific to complex systems

Back to the definition

“A system is a complex system if it is made of a large number of interacting elements and if the dynamics of these interactions govern the behavior of the system, giving to it an appearance of unity from the point of view of an external observer.”

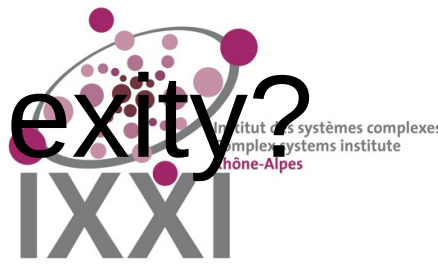


Back to the definition

“A system is a complex system if it is made of a large number of interacting elements and if the dynamics of these interactions govern the behavior of the system, giving to it an appearance of unity from the point of view of an external observer.”

- So the question is:
 - Given the elements and their interactions, how can we quantify/understand/reproduce the appearance of unity?
- Two main questions can be derived from the central one:
 1. Description: What is the “unity” of the system? What are the elements? How can we describe both levels accurately?
 2. Understanding: What is the link between the dynamic of local interactions and the unity of the global system?

Biocomplexity?



second
minute
year
Miller

Scales

nanometre
micrometre
metre
kilometre

“Nothing in biology makes sense except in the light of evolution”

(Dobzhansky, 1973)

... And evolution is pragmatic

Number of elements

nucleotides 5.10^9
genes 3.10^5
proteins 10^{10}
cells 10^{14}
neurons 10^{12}
humans 5.10^9

Heterogeneity

kind of proteins 10^6
kind of cells 10^3
species 10^7

Complex Systems Institute
5 rue du Vercors 69007 Lyon France
Tél. +33 4 72 72 85 81 | Email: ixxi@ens-lyon.fr | Site: www.ixxi.fr

(me, now)

Biologie des systèmes complexes



Institut des systèmes complexes
Complex systems institute
Rhône-Alpes

- D'où vient l'apparence d'unité des systèmes biologiques ?
 - Biologie des systèmes, biologie intégrative, vie artificielle, ...
- Comment étudier cette « apparence » d'unité ?
 - Manque d'outils adaptés à la complexité des systèmes biologiques
- Démarche de modélisation (biologie « in silico »)
 - Quels modèles ?
 - Modèles individu-centrés (« description locale, observation globale »)
 - Par définition ces modèles manipulent un grand nombre d'éléments
 - Par définition ces modèles permettent d'explorer les interactions multi-échelles (mais rarement plus de deux niveaux d'organisation)
 - Malheureusement rarement (très) hétérogènes ...
 - Ils permettent un dialogue efficace (parlant) entre le modèle et la biologie (les biologistes)

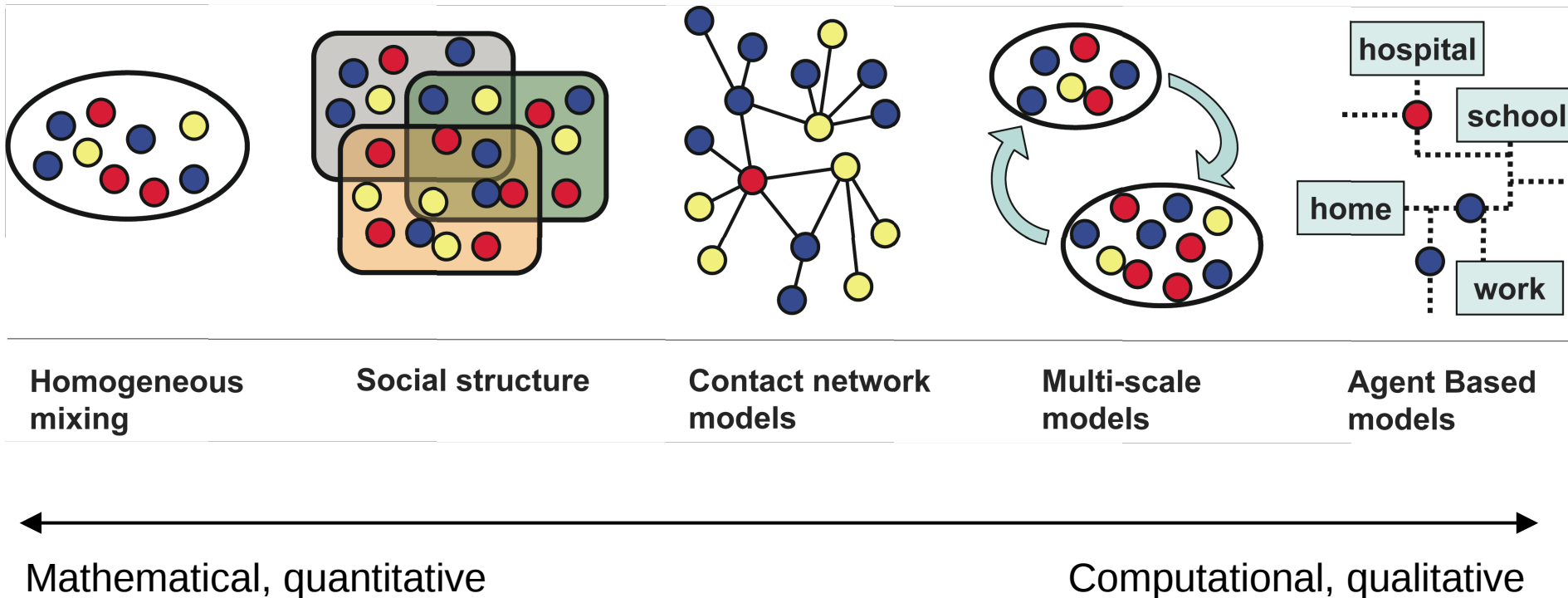
What is a model?

- No clear definition (again!), strong polysemy
 - Models in science (formal sciences $\neq$ empirical sciences), models in art
- But always a clear link with idealization or imitation...

“To an observer B , an object A^* is a model of an object A to the extent that B can use A^* to answer questions that interest him about A .” (Marvin Minsky)

- So scientific models are instruments for scientific discovery
 - Used to explore properties of systems through virtual experiments
 - What is the epistemological status of a virtual experiment?
- Computational models are those which uses computation to perform the experiments
 - The model typically uses an algorithm to compute the state at type t from the state at time $t-1$

Archipelago of models in complex systems science



[From Barthelemy, 2008]

What is IBM/ABM?

- Agent-Based Modeling is a kind of computational models based on an explicit description of the agents.
- “Bottom-Up” modeling:
 - Describe the system at the local level with some formalism
 - Simulate it (computational model)
 - Observe and analyze the results (at both levels!)

“In agent-based modeling (ABM), a system is modeled as a collection of autonomous decision-making entities called agents. Each agent individually assesses its situation and makes decisions on the basis of a set of rules. Agents may execute various behaviors appropriate for the system they represent -- for example, producing, consuming, or selling. Repetitive competitive interactions between agents are a feature of agent-based modeling, which relies on the power of computers to explore dynamics out of the reach of pure mathematical methods.”

[Bonabeau, 2002]

References

- Bonabeau, E. (2002). Agent-based modeling : Methods and techniques for simulating human systems. *Proceedings of the National Academy of Sciences of the USA (PNAS)*, 99(suppl. 3):7280–7287.
- Grimm, V. (1999). Ten years of individual-based modelling in ecology : what have we learned and what could we learn in the future ? *Ecological Modelling*, 115:129–148.
- Bankes, S. (2002). Agent-based modeling : A revolution ? *Proceedings of the National Academy of Sciences of the USA (PNAS)*, 99:7199–7200.
- Grimm, V., Revilla, E., Berger, U., Jeltsch, F., Mooij, W.M., Railsback, S.F., Thulke, H.-H., Weiner, J., Wiegand, T., DeAngelis, D.L. (2005) Pattern-Oriented Modeling of Agent-Based Complex Systems: Lessons from Ecology. *Science*, 310:987-991.
- Macal, C. M. et North, M. J. (2006). Tutorial on agent-based modeling and simulation part 2 : how to model with agents. In WSC06 : *Proceedings of the 38th Winter simulation conference*, Monterey (USA), pages 73–83.

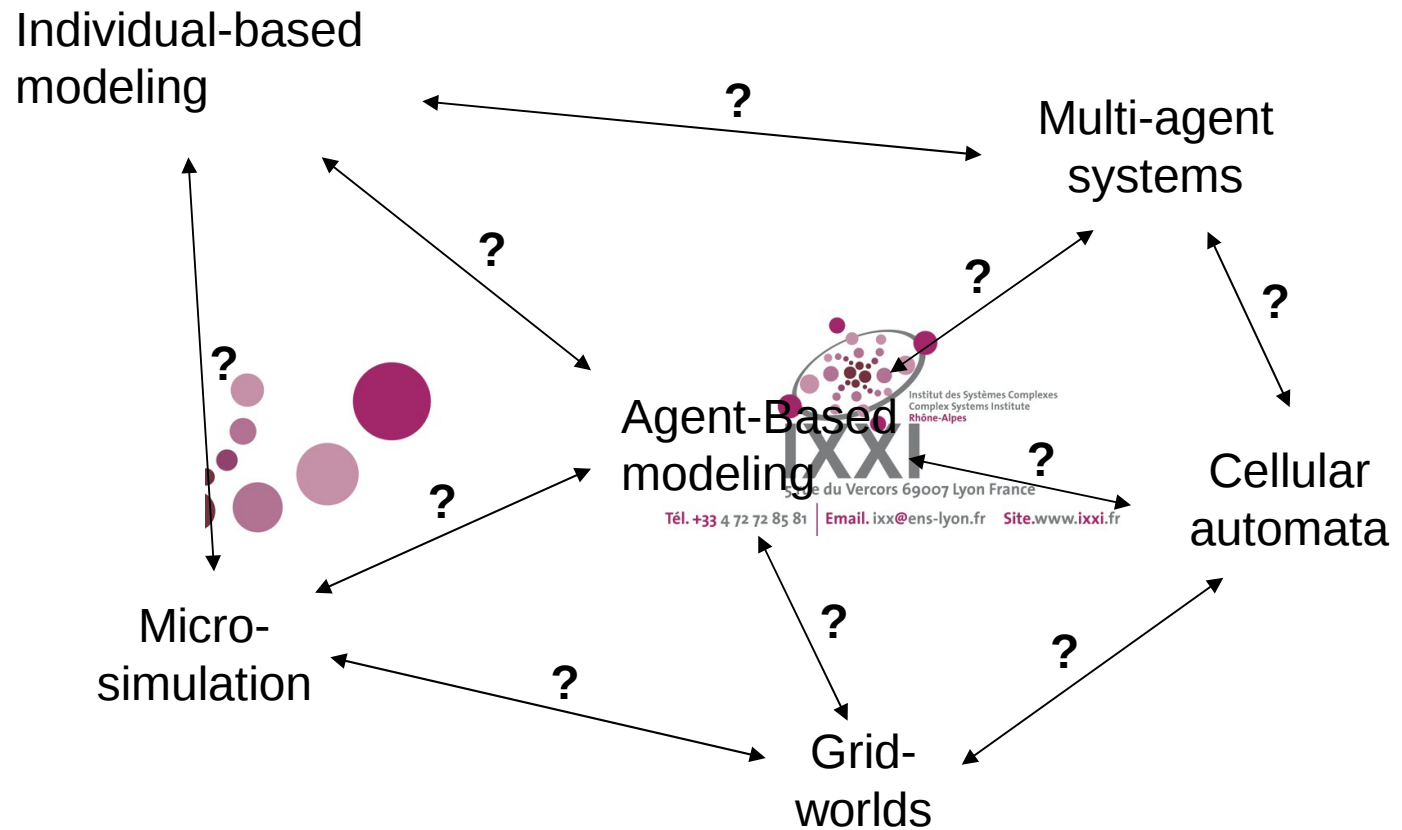
What is ABM?

- Agent-Based Modeling is more a methodology than a precise technique
 - You can choose the formalism you “want” at the agent level (dynamical models, set of rules, discrete/continuous coordinates, punctual particles or not, ...)
 - The only thing you need is a way to compute the interactions and, thus, the resulting behavior
 - But this may not be a trivial question!
 - You’ll have to use computational tools that can be very diverse...

“Agent-Based Model is a mindset more than a technology.”

[Bonabeau, 2002]

What is ABM?



What is ABM?

- Consensus for the principles
- Diversity of the appellations!
 - Micro-simulation (physics)
 - Agent-Based Modeling (computer science, social science)
 - Individual-Based Modeling (biology, ecology)
 - Bottom-Up simulation
- The only real difference is with MAS
 - Multi-Agent Systems are NOT Agent-Based Models
 - MAS are IT technologies trying to use CS approaches to improve the behavior of programs and computers
 - MAS are NOT models
 - MAS can be used to implement ABM but... why?

ABM, Cellular Automata and Grid Worlds

- 2D cellular automata are often presented as ABM
 - In CA rules are associated with the places, not with the agents
 - CA are not ABM, except when dealing with fixed agents (one place-one agent)
- Grid world are 2D worlds (sometimes 3D) where objects move on a grid-based space according to rules
 - The rules are local to the objects, not to the places
 - Probably the simplest ABM
 - E.g., DLA ...

What is an agent?

- Classical definition (North & Macal)
 - A discrete entity/program with its own goals and behaviors
 - Autonomous, with a capability to adapt and modify its behaviors
 - Some key aspect of behaviors can be described.
 - Mechanisms by which agents interact can be described.
- Examples
 - People, groups, organizations, insects, swarms, robots...
- But this definition is strongly rooted in MAS and social systems

What is an agent?

- You will often find figures like:



QuickTime^a and a
decompressor
are needed to see this picture.

OR



Institut des Systèmes Complexes
Complex Systems Institute
Rhône-Alpes

5 rue du Vercors 69007 Lyon France

Tél. +33 4 72 72 85 81 | Email. ixx@ens-lyon.fr | Site. www.ixxi.fr

QuickTime^a and a
decompressor
are needed to see this picture.

What is an agent?

- An agent is (only) the unit of description of the micro-level
 - Again, “Agent” is more a methodological concept than a technological concept!
- What is agent (or not) depends on your point of view!
 - What is really important is what is local and what is not!
- It is often very difficult to decide what is an attribute, what is a memory, what is a resource ...
 - E.g. xcor, ycor, speed, energy, ...

What is an agent?

- Care the difference between:
 - “Anthropomorphic” definition: An entity that senses its environment and acts upon it in order to achieve a goal
 - Technical definition: A persistent autonomous software entity dedicated to a specific purpose (e.g. a program, a thread or a robot)
 - Methodological definition: The conceptual unit of interest, defines a boundary between what is modelled and what is observed (hum ... often the observed system is the agent...)

Life-cycle of an ABM

- Developing an ABM seems straightforward!
 - Describe the system at the agent level ; describe the interactions between the agents
 - Create a population of agents
 - Use some simulation method/software to let the agents and the population run
 - Observe the result(s)
 - Draw conclusion
- Actually it is (quite) as simple as this...
 - But some steps may be difficult ;)

Designing the agent level

- NOT YET: As in every model, define very carefully your system and your aim FIRST!
 - Generally a scientific question but...
 - ABM can also be used to help to define a scientific question!
- Choose the agent level, the agents behavior and the agent interactions
 - Take care: the devil is in the details!
 - You need a good knowledge and skill in order to be able to select the appropriate description at the appropriate level!
 - Care habits, transfer of models from a domain to another one, code reuse, ...
 - Care implicit choices

How to design the agents?

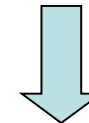
- Actually no real methodology...
 - ABM skill helps,
 - A precise question helps a lot,
 - Domain knowledge helps enormously!
- The only methodology is trial and errors!

- Examples of agents

- Molecules
- Planets/stars
- Humans
- Insects
- Companies
- Cars
- Drops of water
- Birds...



Both have similar properties:
inanimate objects following
physical (Newtonian) laws



Can we use the same
agents models?

How to choose the “level of complexity” ?

QuickTime[®] and a
decompressor
are needed to see this picture.



Tél. +33 4 72 72 85 81 | Email. ixx@ens-lyon.fr | Site. www.ixxi.fr

[Grimm et al., 2005]

From agents to multi-agents

- Once you have designed the agents, you still have important choices to make
 - These choices are often forgotten (often implicit!)
- Agent will “live” in a spatio-temporal world
 - Real world is continuous
 - Agents’ world is not!
 - It creates risks and difficulties
- How to model time?
- How to model space?

The time model

- Time is often (always?) neglected in MAS approaches
 - Generally considered as a non-problem
- Discrete Time
 - Synchronous, asynchronous, discrete-events
 - What is the correct time step?
 - The higher the time step, the higher the error
 - The lower the time step, the slower the simulation
 - Practitioners are generally NOT able to estimate the correct time step of their systems!
 - The correct time step depends on the movement and on the interaction models
- The time model may strongly influences the global behavior

Tél. +33 4 72 72 85 81 | Email. ixxi@ens-lyon.fr | Site. www.ixxi.fr



The space model

- Space is often at the core of ABM
 - Space is mainly a constraint on agents' neighborhood
 - Very often, you will use ABMs to test the behavior of analytical models in a given spatial framework
- Lots of different space models are possible
 - From “Soup model” to GIS models
 - You often have to mix different space models (e.g. continuous space for agents + diffusion on a grid)



5 rue du Vercors 69007 Lyon France
Tél. +33 4 72 72 85 81 | Email. ixx@ens-lyon.fr | Site. www.ixxi.fr

QuickTime[®] and a
decompressor
are needed to see this picture.

The space model

- Care: like for time, there are often implicit assumptions for the space model
 - Is 2D sufficient?
 - How to model the borders of the space?
 - Absorbing, reflecting, static, periodic...
 - How to model infinite spaces?

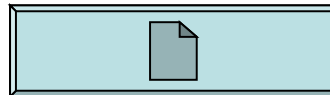


“Diffusion is not a perfectly mixing process in low dimension because the diffusing molecule will return to its initial position with probability 1, whereas, for $d > 2$, there is a significant probability that the diffusing molecule will never return to its origin.”

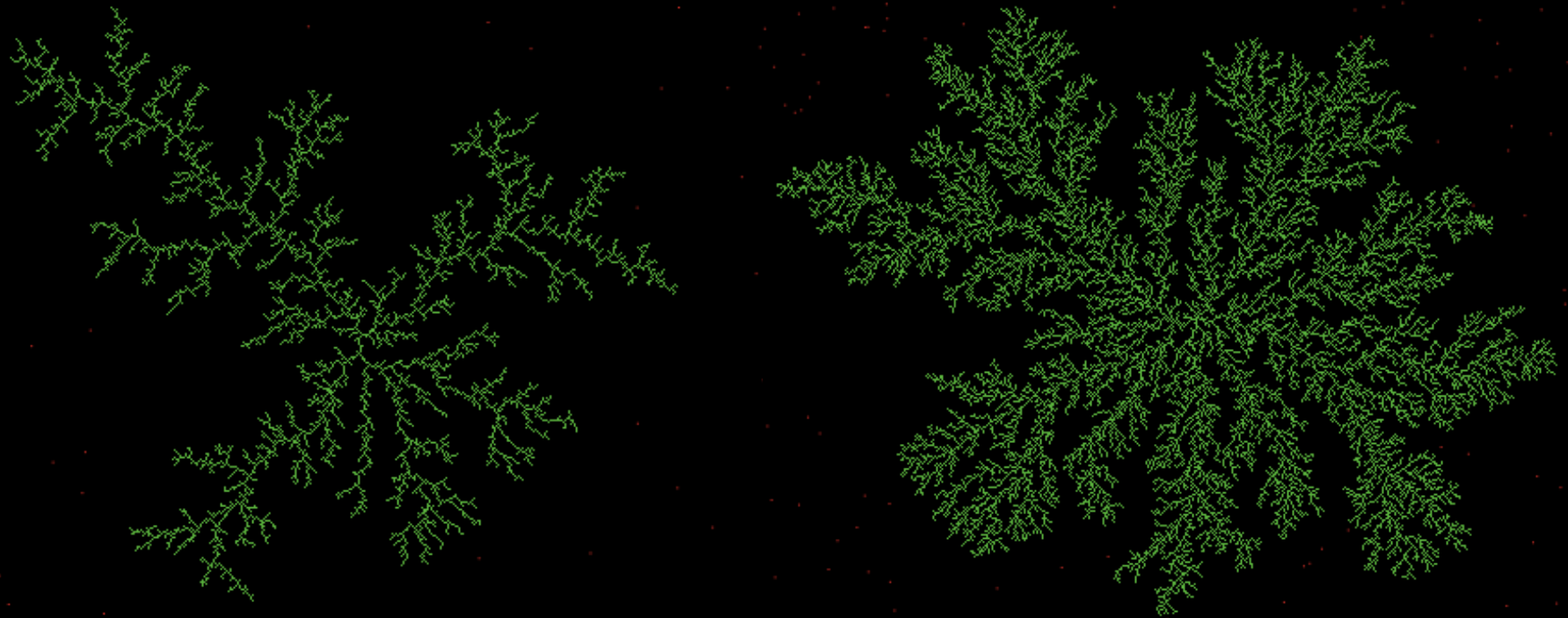
[berry, 2002]

Movement

- Agents will often move in “a” space
 - The laws of movement are generally supposed simple
 - Very often they are not!
 - Care not to reuse implicitly macroscopic laws of motion into a microscopic world (e.g., planets and molecules)
 - Sometimes the laws of motion explains the “emergent” results by themselves!
- E.G., DLA
 - Agents explore differently their vicinity depending on the laws of motion!



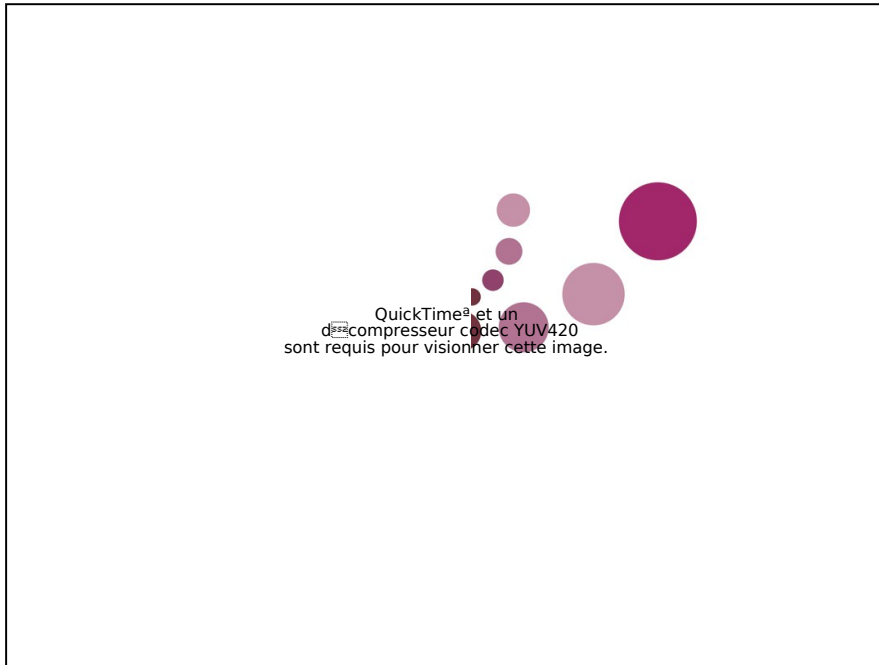
Fractals structures created by DLA



Two different laws of movement,
which one is correct?

Law of motion matters!

- Coral morphogenesis [Merks et al., 2003]
 - Same agents
 - Different diffusion parameters leads to different shapes



Slow diffusion



Fast diffusion

Implementation step

- Once you have designed your agents and their relations, how can you implement them and run the simulation?
 - Plate-forms, frameworks,
 - Programming from scratch (which language),
 - Reuse a previous model
- Take care: the implementation phase is NOT the most difficult nor the most time consuming!
- Choose the methods/tools such that they
 - Respect the modeling phase
 - Will be efficient during the experimental phase
 - Enable to follow “strictly” a scientific experimental methodology
- Then, you’ll probably have to program “a little”...

Implementation

- You will often find figures like:



The pro&cons of visualization

- Complex systems are based on subjective judgment
 - We need a visual feedback!
- We often have no mean to decide what is correct and what is not
 - We need a visual feedback!
- We have to care natural interpretations
 - Care visual feedback! (“I like it!”)
- We have to repeat the experiments
 - Visual feedback are often slow!
- We have to repeat experiments
 - Visual feedback cannot be aggregated
- Conclusion
 - Care to visualize easily and to emphasize what is important
 - Care not to focus only on visualization: data output are important



Institut des systèmes complexes
Complex systems institute
Rhône-Alpes

Experiments

- Agent-Based Models often have MANY parameters
 - Most of them are often implicit ...
 - E.g., in my own model (Aevol) : 53 parameters!
- Agent-Based Models are generally slow
 - Need lots of computational resources
- It is NOT possible to test all parameters
 - Again, no hint! (except your own knowledge and experiments)
- Don't explore randomly the parameter space
 - Use “good practices” of experimental science
 - Actually ABM is an experimental approach (digital experiments)
 - Having a laboratory notebook is a VERY good practice!
 - Log all your experiments ; finish all your experiments
- Making the model is often less “difficult” than running the model...
 - Plan resources and time from the beginning of your project

The meta-life-cycle of ABM

- Actually, ABM are not so difficult to build!
- The difficulty is (again) to produce knowledge with them!
- Meta-life cycle of ABM
 - Identify a good question
 - Build different simple models and play with them to identify what matters or not
 - Build YOUR model and make it stable
 - Make experiments with the model (experimental method helps!)
 - Analyze the results (statistical skill helps!)
 - Hopefully, acquire new knowledge (model the model)
 - Communicate, confront, publish
 - **FORGET YOUR MODEL**

Forget your model?

- Two reasons:
- The model is not the knowledge

“It could be argued that a criterion to determine good models is that they are no longer needed afterwards; The decisive thing with modeling is not the model per se, but what the model and working with the model does to our mind.”



[V. Grimm, 1999]

- Remember that a model depends on a question...
 - If you change the question you MUST change the model
 - Of course, you can reuse some pieces of software but be careful on implicit choice
 - The software is not the model
 - Take care not to jump steps in the meta-life-cycle!

So, is there a methodology?

- Definitely not
 - Modeling is an art
 - A counterfeiter is NOT an artist (though a skilled person!)
- But we can give hints
 - Be a VERY skilled with your modeling tools
 - Start from a good true question (i.e. that interests someone)
 - Be rigorous in your “experiments”
 - *“Avoid the temptation to run tomorrow’s computer simulations before yesterday’s has been fully understood”* (miller, 1995)
 - Use multiple complementary models rather than a big one
 - Confront your results with the specialists ; (try to) publish in the journal they read

ABM validation

- Verification: The program is doing what you want it to do
 - Very difficult problem! (+/- software engineering)
- Validation: The model produces the “correct” behavior
 - Impossible problem: A model is never “valid”

“Essentially, all models are wrong, but some are useful.”

[G. Box]

- Actually it depends on what you want to do with the model!
 - Predictive models can be tested (but never proved!)
 - Scientific models generally cannot
 - A good model is a model that enables *me* to construct a scientific discourse

Applications



Table 1: Agent-based Modeling Applications

Business and Organizations • Manufacturing • Consumer markets • Supply chains • Insurance Economics • Artificial financial markets • Trade networks Infrastructure • Electric power markets • Hydrogen economy • Transportation Crowds • Human movement • Evacuation modeling	Society and Culture • Ancient civilizations • Civil disobedience Terrorism • Social determinants • Organizational networks Military • Command & control • Force-on-force Biology • Ecology • Animal group behavior • Cell behavior • Sub cellular molecular behavior
--	--

+ evolution
 + hydrology
 + membrane models
 + soil models
 + agriculture
 + diffusion of innovation
 + ...

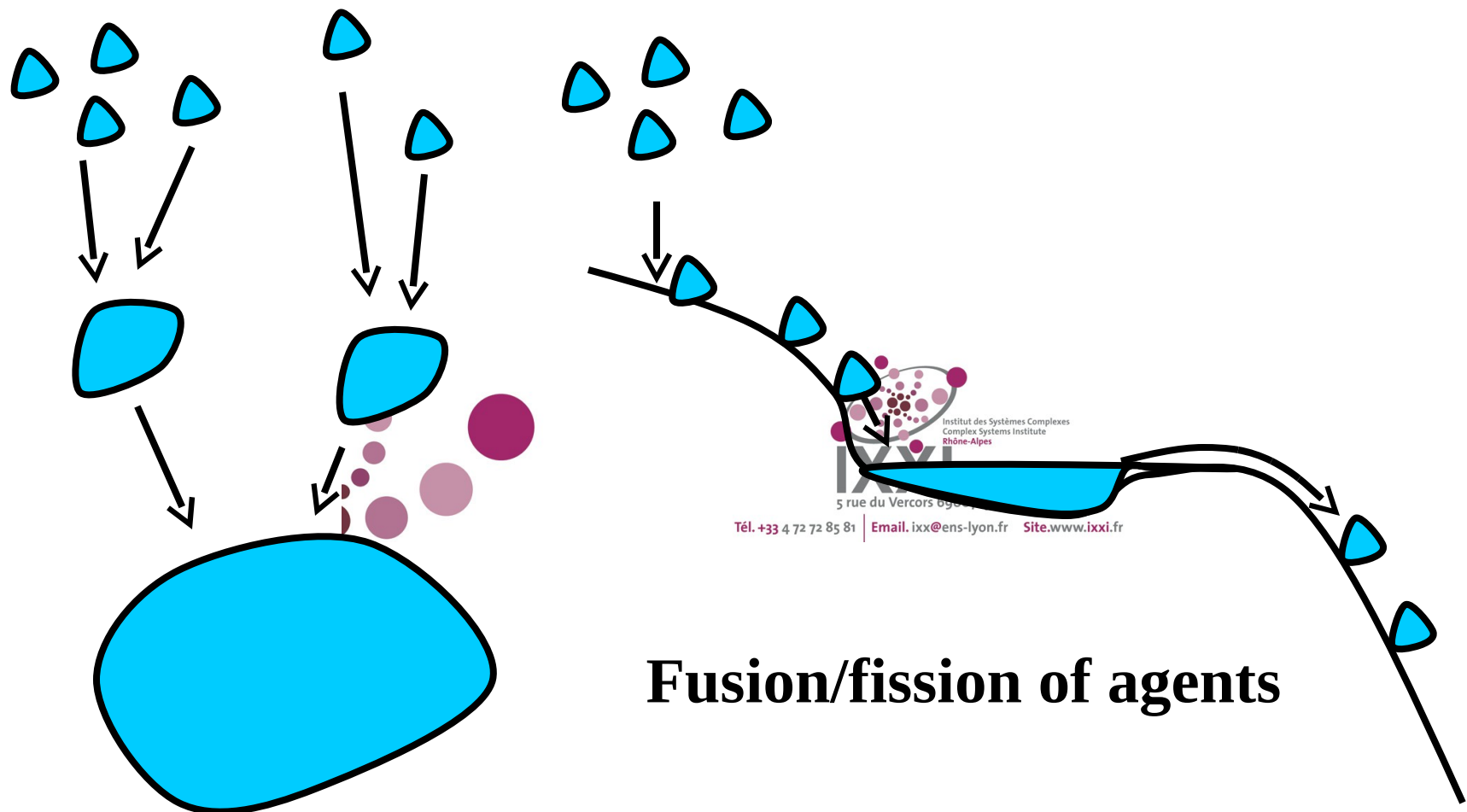
ixxi.fr

Note that businessmen are not as “narrow-minded” as scientists ;)

No need of “proofs”, just need to sell!

[North & Macal, 2006]

Grand challenge of ABM



When/why using ABM?

- [Grimm, 1999]
 - Pragmatic motivation: ABM can model phenomenon impossible to model with other approaches (“another tool in the modelers toolbox”)
 - Paradigmatic motivation: State variables modeling gives a false vision of reality since individuality, discreteness, locality or space matter
- Hum, not clear ... real motivations are more basic
 - Easy to construct, manipulate and extent (easy to change/add/remove parameters, rules,...) ... too easy?
 - Can model unknown phenomenon (if you have knowledge at the lower level)
 - ABM use a domain-based ontology (they are good interfaces between disciplines) easy to describe and to explain ... too easy?
 - “Looks like” (pleasant models) ... too pleasant?

Why/when using ABM?

- Very often, it is claimed that ABM must be used when analytical models fails but
 - Analytical models have a long history in ~every scientific domain (are you sure they fail?)
 - Can we (computer scientists) really know when analytical models can or cannot be used
- In practice, always try to use ABM in parallel with analytical models...
 - ABM can be use before analytical model (to propose hypothesis)
 - ABM can be used after analytical model (to validate hypothesis)

ABM vs. Analytical models

QuickTime^a et un
decompresseur
sont requis pour visionner cette image.

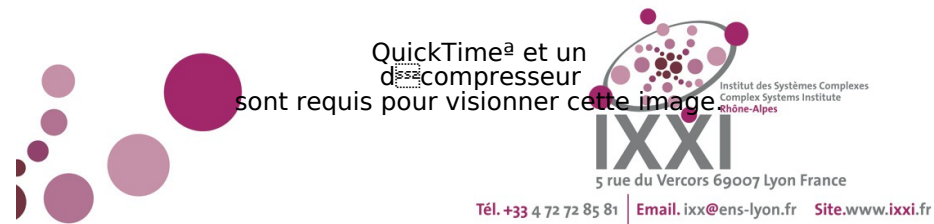


QuickTime^a et un
decompresseur
sont requis pour visionner cette image.



Tél. +33 4 72 72 85 81 | Email. ixx@ens-lyon.fr | Site. www.ixxi.fr

The BIG risk!



Another BIG risk!



QuickTime® and a
decompressor
are needed to see this picture.



Tél. +33 4 72 72 85 81 | Email. ixx@ens-lyon.fr | Site. www.ixxi.fr

The BIGGEST risk!



The future of ABM?

QuickTime^a et un
décompresseur
sont requis pour visionner cette image.



Institut des Systèmes Complexes
Complex Systems Institute
Rhône-Alpes

5 rue du Vercors 69007 Lyon France

Tél. +33 4 72 72 85 81 | Email. ixx@ens-lyon.fr | Site. www.ixxi.fr

Retour à la vraie question !

- L'usage des modèles en science est tout sauf clair ...
- Le modèle est intimement lié à l'imitation, à l'analogie, à la ressemblance
 - Mais il peut représenter aussi bien l'objet à imiter que l'imitation de l'objet ou un intermédiaire entre l'objet et l'imitation ...
 - Modèles comme médiateurs ...
- La modélisation est souvent considéré comme une démarche interdisciplinaire ...
 - Pourtant, chaque discipline a sa propre conception des modèles ...
 - Les modèles sont souvent à l'interface entre sciences appliquées et sciences expérimentales ...
 - Dialogues de sourds autour des modèles (e.g. modèle de données, modèles d'objets)
 - Modèle normatif/modèle descriptif ...



Tél. +33 4 72 72 85 81 | Email: ixxi@ens-lyon.fr | Site: www.ixxi.fr

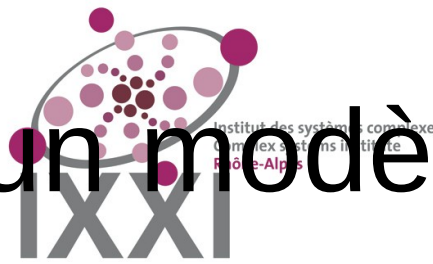


Qu'est-ce qu'un modèle ?

- Définitions « courantes » :
 - Ce qui sert ou doit servir d'objets d'imitation pour faire ou reproduire quelque chose,
 - Personne ou objet dont l'artiste reproduit l'image,
 - Objet, fait, personne possédant au plus haut point certaines qualités et caractéristiques et à laquelle peuvent se rapporter des faits ou des objets réels,
 - Objet, type déterminé selon lequel des objets semblables peuvent être reproduits en de multiples exemplaires,
 - Objet de même forme qu'un autre objet mais exécuté en réduction
 - Représentation simplifiée d'un processus, d'un système



Qu'est-ce qu'un modèle ?



« To an observer B , an object A^* is a model of an object A to the extent that B can use A^* to answer questions that interest him about A . »

Marvin Minsky

- Définition très permissive : est-ce que tout est modèle ?
 - Non : le modèle doit *servir* à produire de la connaissance ...
 - Le modèle est donc un instrument scientifique
 - Il doit être utilisé *comme* un instrument
 - Est-il un instrument comme un autre ?
 - Non : selon la définition c'est un instrument personnel
- Paradoxe : si le modèle est un instrument, il doit être accepté par une communauté scientifique ...
 - Le modèle doit être considéré comme un instrument valide ...
 - Il doit se conformer aux pratiques scientifiques correspondant au champs d'étude de A (et de B ? Et de A^* ?)
 - Mais chaque modèle est un instrument différent ...





Une pièce à deux faces

- Le modèle est un instrument personnel
 - En ce sens son usage est TRES permissif ...
- Le modèle est un instrument collectif
 - En ce sens son usage est TRES restrictif ...
- Dans les deux cas son usage est très dangereux
 - Car en tant qu'instrument systématiquement nouveau, il doit être faire systématiquement ses preuves (et non faire preuve) ...
 - Risque personnel (preuve insuffisante ou fausse)
 - Risque collectif (preuve non reconnue par la communauté)
- Or, la modélisation a toujours un caractère interdisciplinaire
 - L'usage individuel et l'usage collectif peuvent être conduits au sein de disciplines différentes ...
 - En particulier dans les systèmes complexes ...



Un instrument personnel

- Comment le modèle peut-il « faire preuve »
 - « Ce qui est simple est toujours faux. Ce qui ne l'est pas est inutilisable » (P. Valery)
 - « The decisive thing with modelling is not the model per se, but what the model and working with the model does to our mind » (V. Grimm, 1999)
 - « It could be argued that a criterion to determine good models is that they are no longer needed afterward » (V. Grimm, 1999)
 - Le seul critère de qualité d'un modèle est son « utilité » (J.-M. Legay, 1973) ou sa « pertinence » (J.-L. Le Moigne, 1977)
- Le modèle ne fait donc *jamais* preuve
 - Mais ça n'interdit pas son utilité
- C'est le modélisateur qui incarne le lien entre le modèle et l'objet modélisé
 - Mais cela ne suffit pas ...

Un instrument personnel

- Le modèle est indissociable de sa conception et de son utilisation (i.e., de son interprétation)
 - « La connaissance-projet se produit – et se représente – par conception de modèles (...) et non plus par analyse. Le modèle alors, qu'il soit iconique ou symbolique, devient source de connaissance et non plus résultat. Il ne décrit plus, ex-post, une connaissance-objet tenue pour ex-ante ; il représente a priori une connaissance-projet qui n'existe que par lui. » (J.-L. Le Moigne, 1987).
- Le modèle n'est donc pas un résultat, un objectif scientifique *en soi*
 - Le modèle n'est pas une (simple) copie
- Il n'est modèle que par rapport à une question sur un objet et par rapport à un *interprète* ...
 - On ne peut pas dissocier le modèle du modélisateur ...
 - Pourtant la pratique scientifique nous impose de communiquer le modèle à une communauté





Un instrument collectif

- Le modèle est un instrument personnel mais qui doit autoriser les échanges avec le collectif ...
 - Sinon, risque de dérive intuitionniste ...
 - La science qui se fait est la science qui se communique ...
 - A qui ?
 - Que doit-on communiquer ? Le modèle, l'intuition ou la « conclusion » ?
 - La communication change-t-elle le statut du modèle ?
- « Il y a peu de controverses entre simulateurs car il y a peu de travail collectif. Les simulateurs sont rassemblés par l'équipement informatique qui leur est nécessaire, mais ils fonctionnent plutôt à la manière de petits artisans : chacun son problème, son modèle, son programme » (I. Stengers et B. Bousaude-Vincent, 2003)

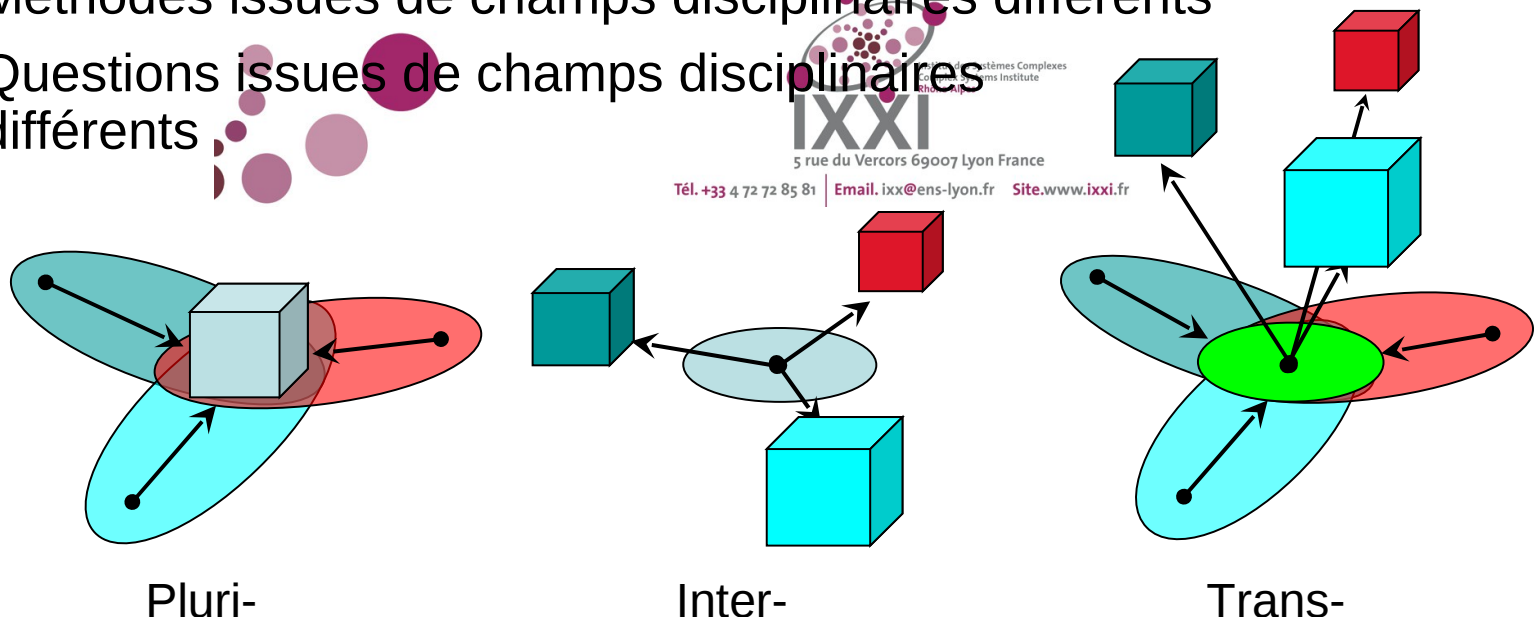
Un instrument collectif

- Chaque champ d'application, chaque domaine scientifique, va exiger du modèle (et du modélisateur) qu'il se plie aux règles (implicites) du domaine
 - Sous peine de ne pas être considéré comme un instrument valide
 - Qu'est-ce qui fait la validité d'un instrument ?
 - Un modèle peut-il être un instrument valide puisqu'il est toujours un instrument ad-hoc ?
 - Attendez-vous à devoir convaincre ...
- Le modèle doit être intégré à la connaissance du domaine et non à la connaissance « des modèles »
 - Imagine-t-on Galilée communiquer ses résultats uniquement à des opticiens ?
 - Galilée a dû convaincre que les lois de l'optique sont valides pour l'astronomie
 - Le modèle doit définitivement s'insérer dans la pluridisciplinarité ...



Inter- pluri- trans-disciplinarité

- Modéliser implique de dépasser les frontières traditionnelles entre les disciplines scientifiques
- Des collaborations sont indispensables
 - Expérimentateurs/modélisateurs, spécialistes du local/du global
 - Méthodes issues de champs disciplinaires différents
 - Questions issues de champs disciplinaires différents



Inter- pluri- trans-disciplinarité

- L'inter- pluri- trans-disciplinarité est souvent défendue ... dans les discours
 - Beaucoup plus rarement en pratique
 - E.g. : « Je ne prends que les meilleurs » ...
- Traverser les frontières entre disciplines scientifiques est difficile ! Cela demande du temps, du tact et cela implique des risques !
 - Soyez modestes : toutes les disciplines sont TRES avancées
 - Soyez tolérants : toutes les disciplines ont des habitudes (bizarres ;)
 - Soyez clairs : quel est votre objectif ? Qui voulez-vous convaincre ? (où voulez-vous publier ?)
 - **Ne croyez jamais pouvoir apporter une connaissance de l'extérieur d'une discipline!**

“The burden of proof is on us to explain our results to biologists in their own language and in their own journals”

[Miller, 1995]

Ecole de Porquerolles

Modélisation individu-centrée de systèmes biologiques complexes

Application à la simulation de l'évolution de réseaux génétiques bactériens

Guillaume Beslon

INSA – INRIA – LIRIS – IXXI



Introduction

- Aim of the course (first part):
 - Application à la simulation de l'évolution de réseaux génétiques bactériens
 1. Evolution ?
 2. Simulation de l'évolution ? (digital genetics)
 3. Simulation de l'évolution de réseaux génétiques
- Who am I?
 - Guillaume BESLON (guillaume.beslon@liris.cnrs.fr)
 - Professor at the INSA-Lyon, LIRIS Lab. (Laboratoire d'Informatique en Image et Systèmes d'Information),
 - Head of the INRIA COMBINING Team
 - Director of the IXXI (Rhône-Alpes Complex Systems Institute)
 - Research topics: Individual-based modeling of complex biological systems (mainly evolution)

Evolutionary systems biology ?

- Every biological system is the result of an evolutionary story:
 - Understanding the story may help to understand the system
- Systems biology aims at explaining the global structure and organization of biological systems
 - +/- reverse engineering applied to biological systems
 - BUT: in reverse engineering, we have clues on the aims/wills/wishes/methods of the engineers
 - We don't have such clues in the case of biological systems
 - Our “natural interpretations” are likely to be false (care anthropomorphisms...)
 - “Evolutionary systems biology” can guide us, help us avoiding natural interpretations, give the organization clues ...

"But with regard to the material world, we can at least go so far as this—we can perceive that events are brought about not by insulated interpositions of Divine power, exerted in each particular case, but by the establishment of general laws."

W. WHEWELL: *Bridgewater Treatise*.

"To conclude, therefore, let no man out of a weak conceit of sobriety, or an ill-applied moderation, think or maintain, that a man can search too far or be too well studied in the book of God's word, or in the book of God's works; divinity or philosophy; but rather let men endeavour an endless progress or proficience in both."

BACON: *Advancement of Learning*.

Down, Bromley, Kent,
October 1st, 1859.

ON
THE ORIGIN OF SPECIES

BY MEANS OF NATURAL SELECTION,

OR THE
PRESERVATION OF FAVOURED RACES IN THE STRUGGLE
FOR LIFE.

By CHARLES DARWIN, M.A.,

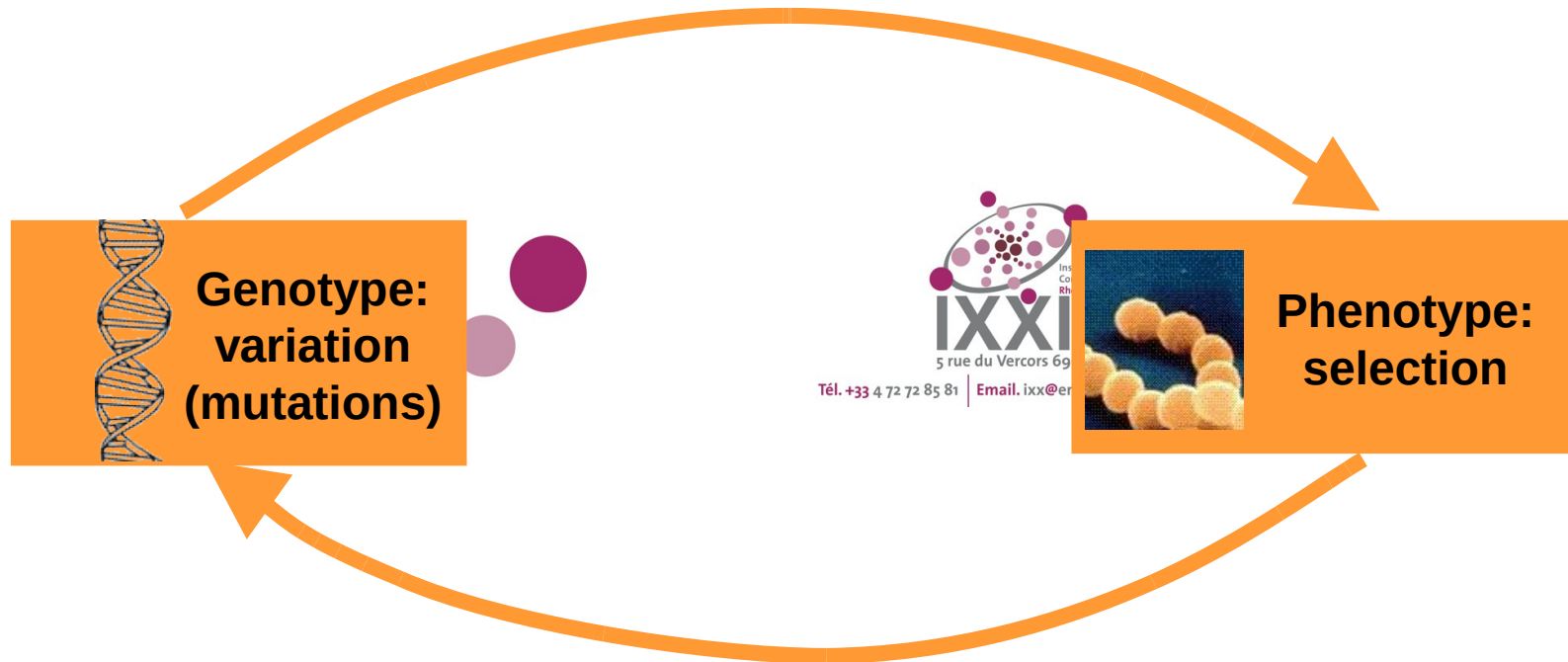
FELLOW OF THE ROYAL, GEOLOGICAL, LINNEAN, ETC., SOCIETIES;
AUTHOR OF 'JOURNAL OF RESEARCHES DURING H. M. S. BEAGLE'S VOYAGE
ROUND THE WORLD.'

LONDON:
JOHN MURRAY, ALBEMARLE STREET.
1859.

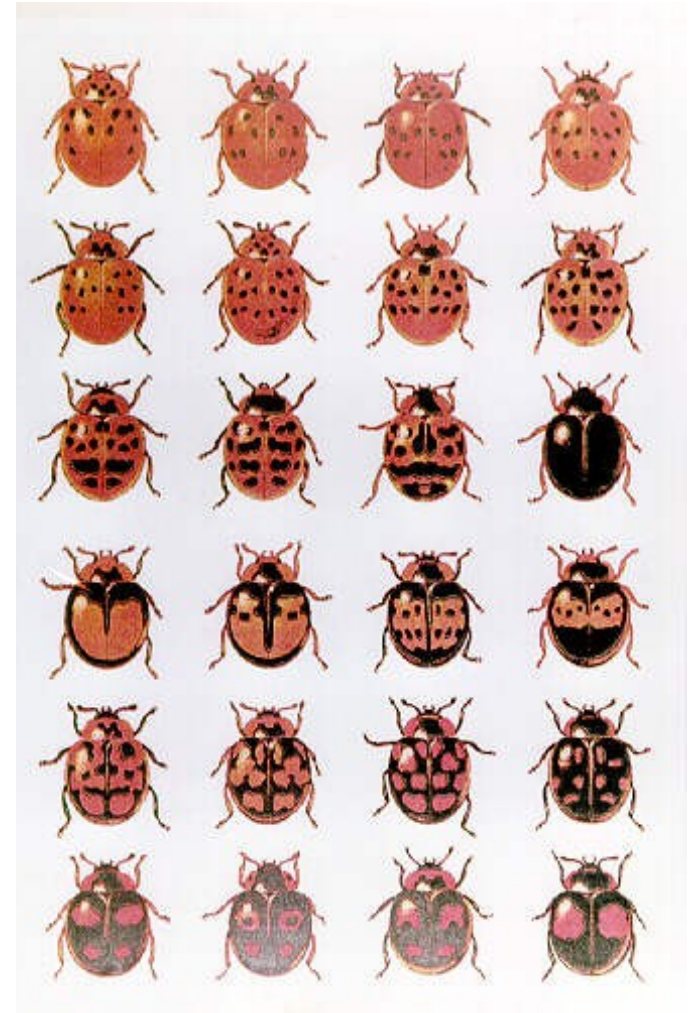
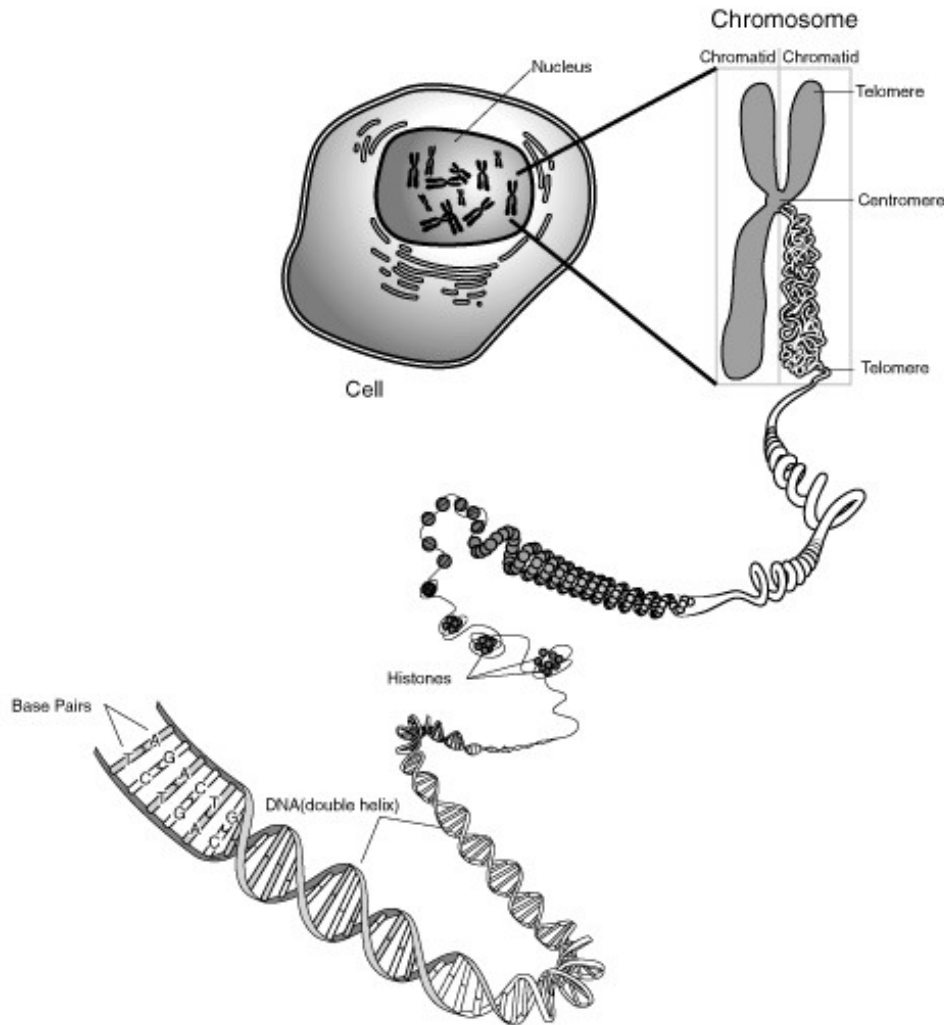
Evolution in two words

*“Evolution will occur **whenever and wherever** three conditions are met: replication, variation (mutation), and differential fitness (competition).”*

[Daniel Dennett]

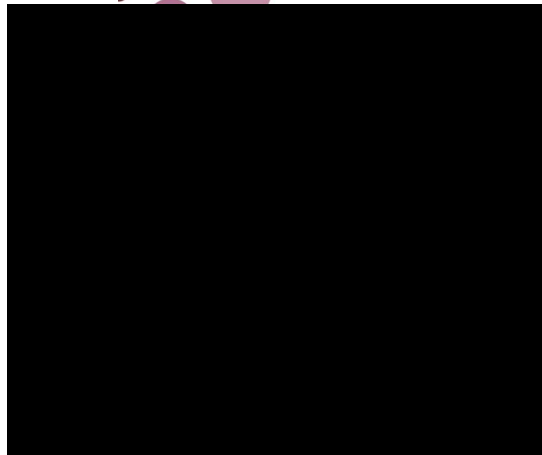


Genetic variability





Natural selection



Tél. +33 4 72 72



The **fitness** measures
the probability of
survival and
reproduction

Example of “natural” evolution

- *Biston betularia* (Peppered moth)
 - 1848: first (known) occurrence of the black morph (*carbonaria*)
 - 1898: *carbonaria* represents 98% of the population (industrial melanism)



Example of “natural” evolution

- *Biston betularia* (Peppered moth)
 - 1848: first (known) occurrence of the black morph (*carbonaria*)
 - 1898: *carbonaria* represents 98% of the population (industrial melanism)



Introduction

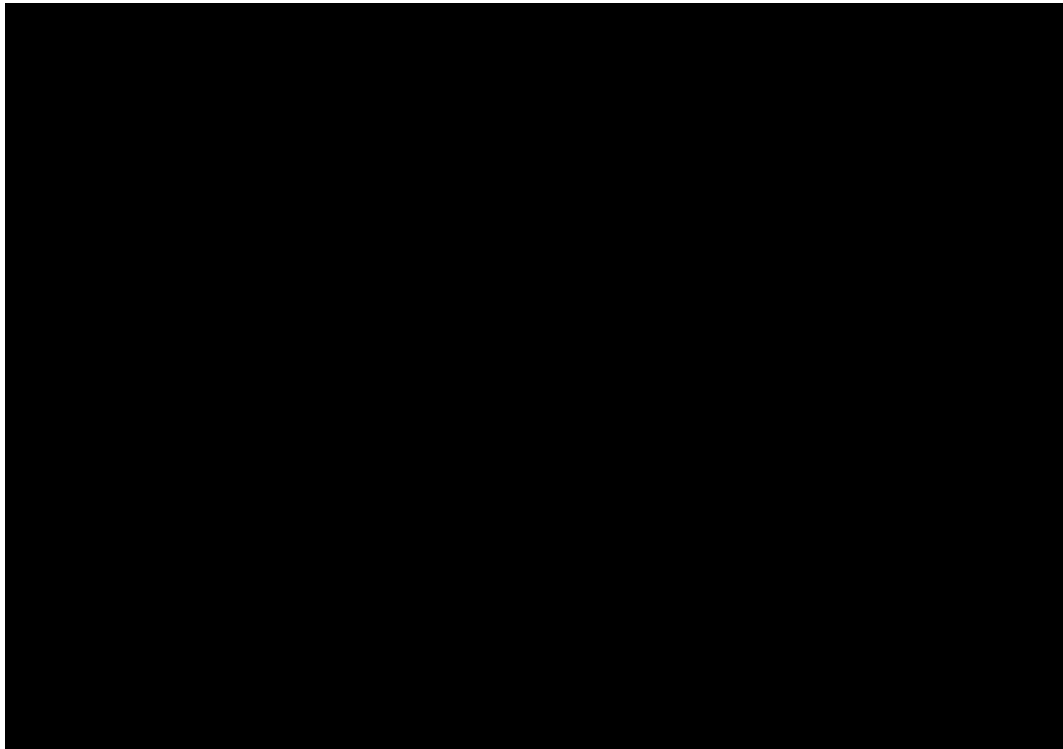
- Although it can be described in a few words, evolution give rise to many complex phenomenon that can be very difficult to understand
 - Evolution of cooperation, evolution of sex, evolution of complexity...
- Evolution is difficult to study
 - Well known snapshot (today)
 - Few fossil records
 - Difficult experiments
- Some evolutionary pressures are well-known but their relative contribution is almost impossible to assess
 - Modeling needed!



Institut des systèmes complexes
Complex systems institute
Rhône-Alpes

The fitness landscape metaphor

(Sewall Wright, 1932)

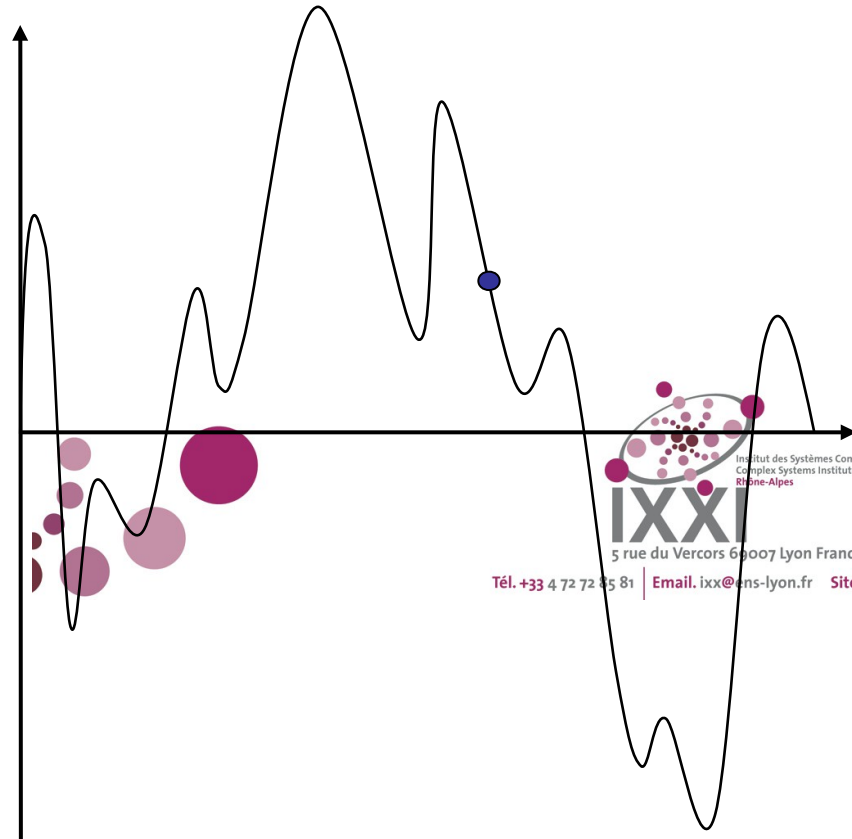


Institut des Systèmes Co
mplex Systems Instit
Rhône-Alpes

007 Lyon Fran
s-lyon.fr Sit

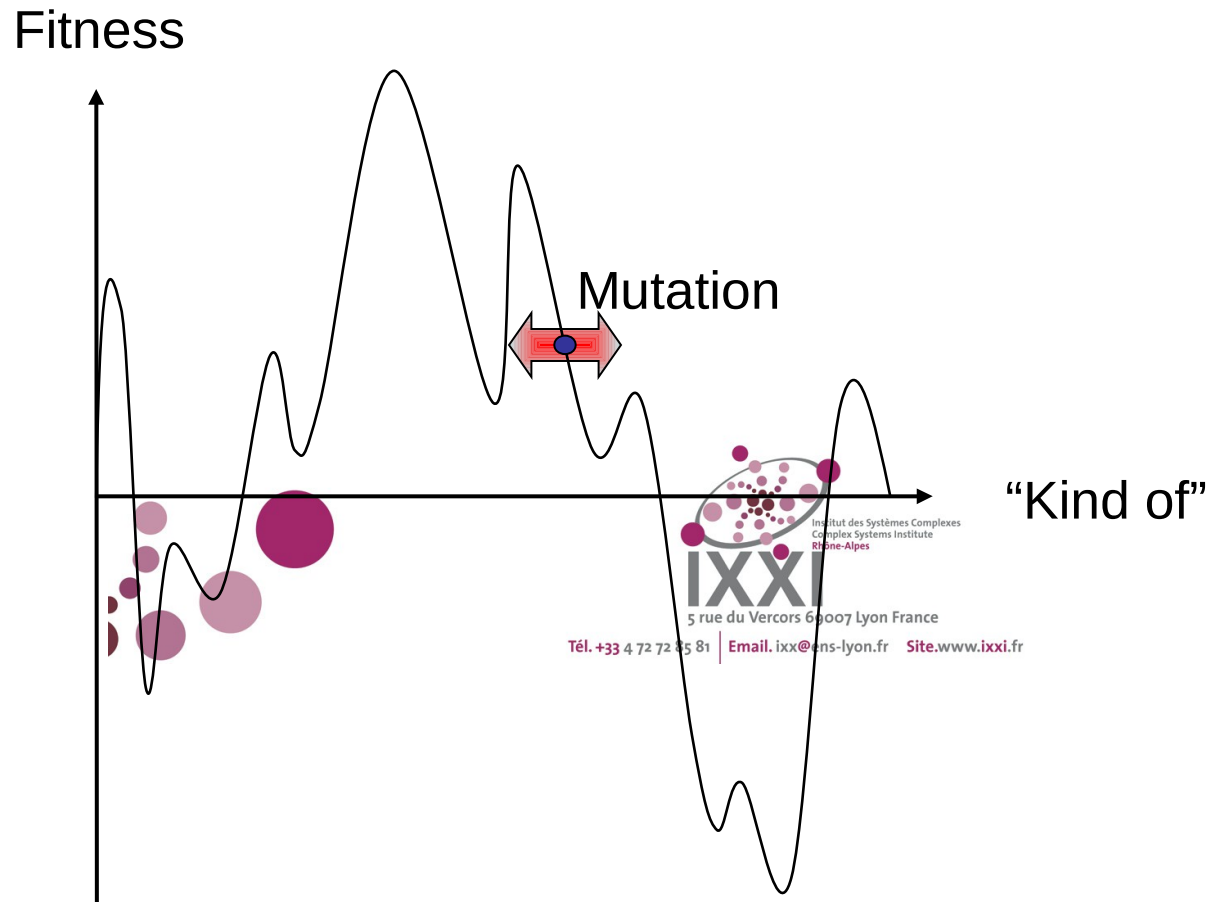
The fitness landscape metaphor

Fitness

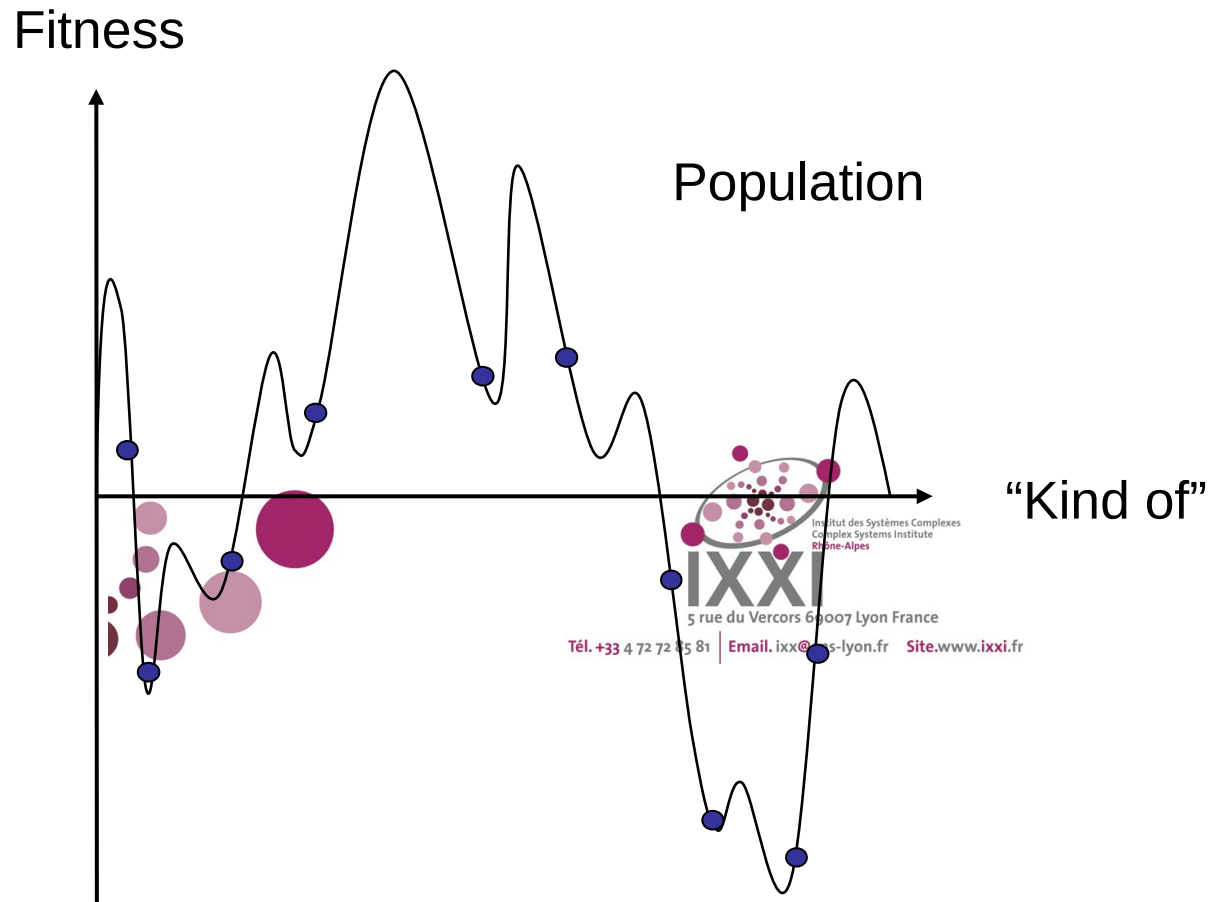


“Kind of”

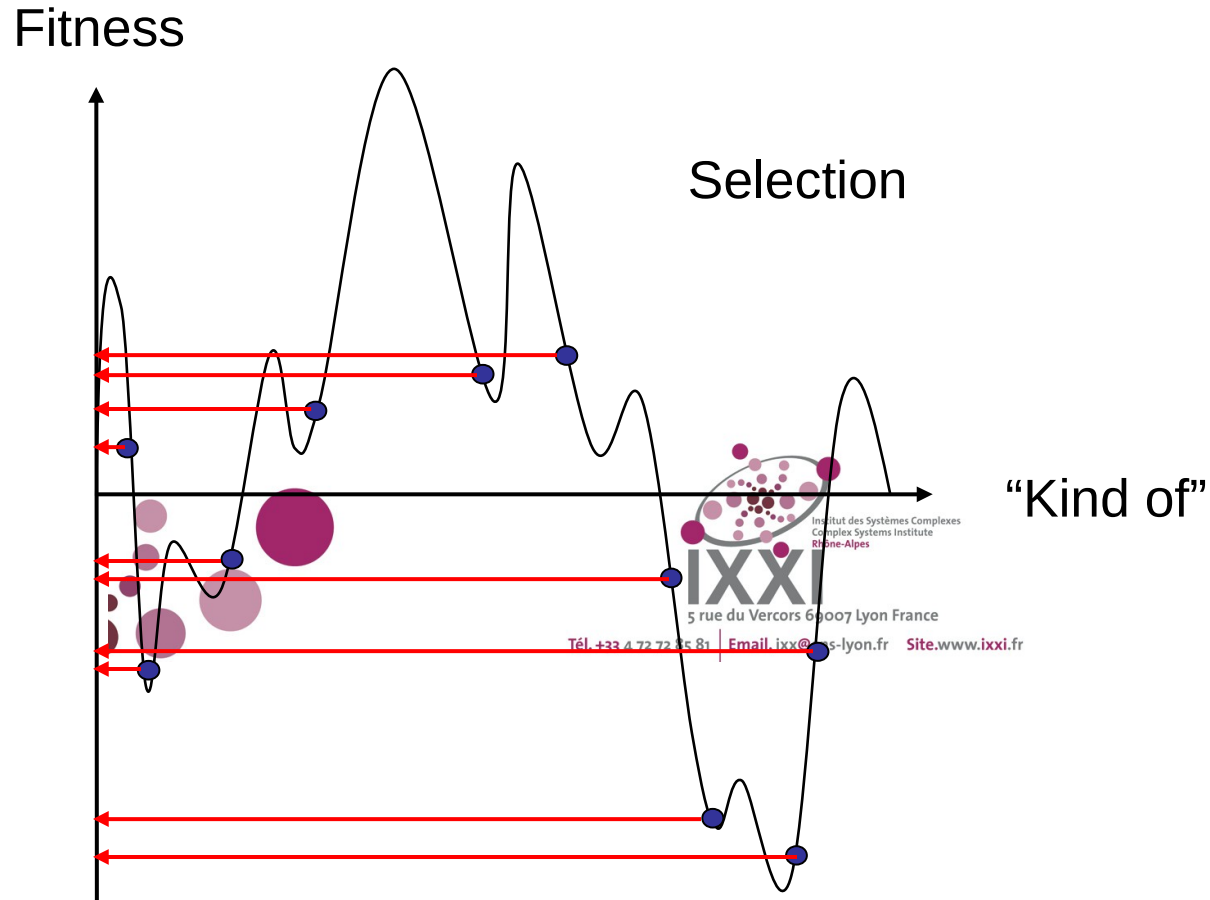
The fitness landscape metaphor



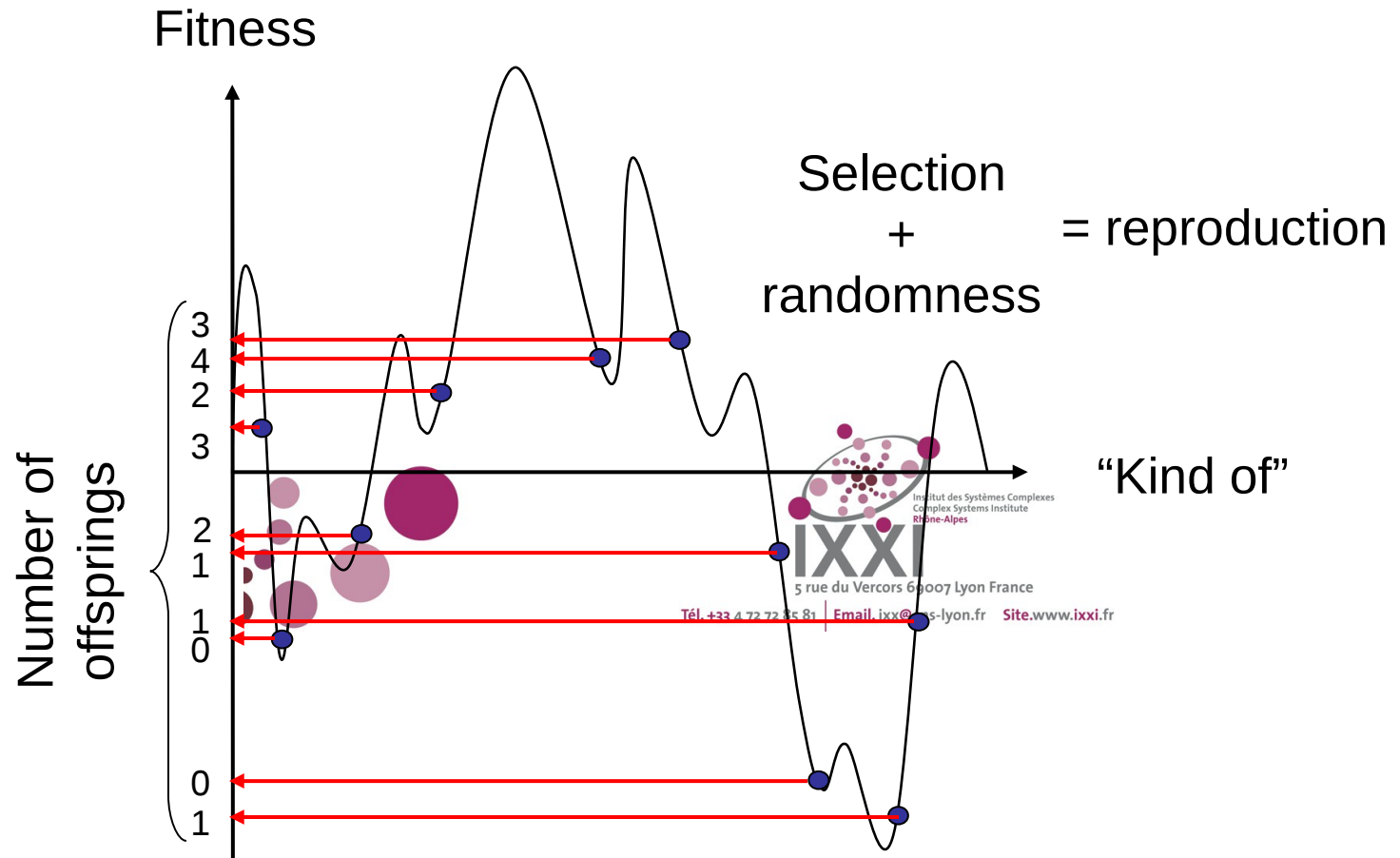
The fitness landscape metaphor



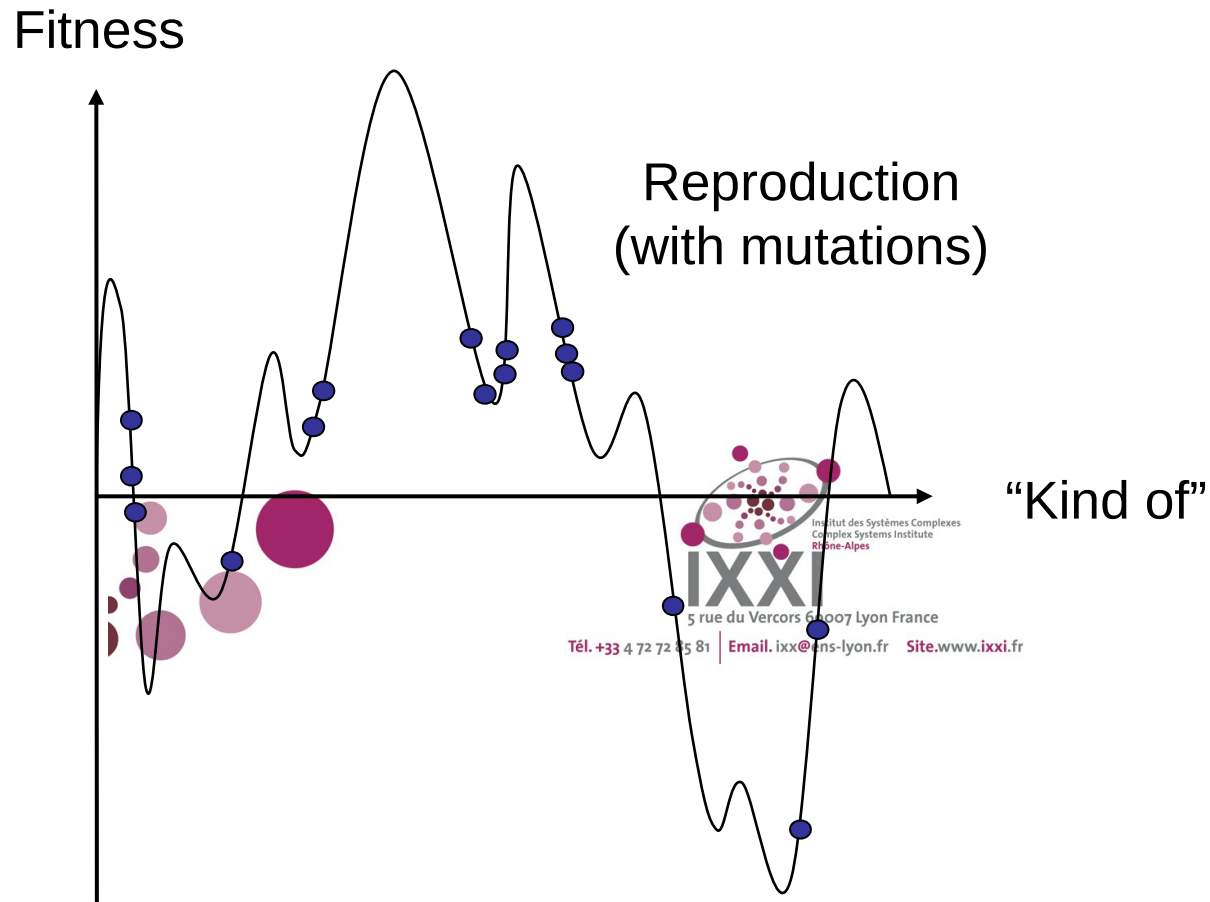
The fitness landscape metaphor



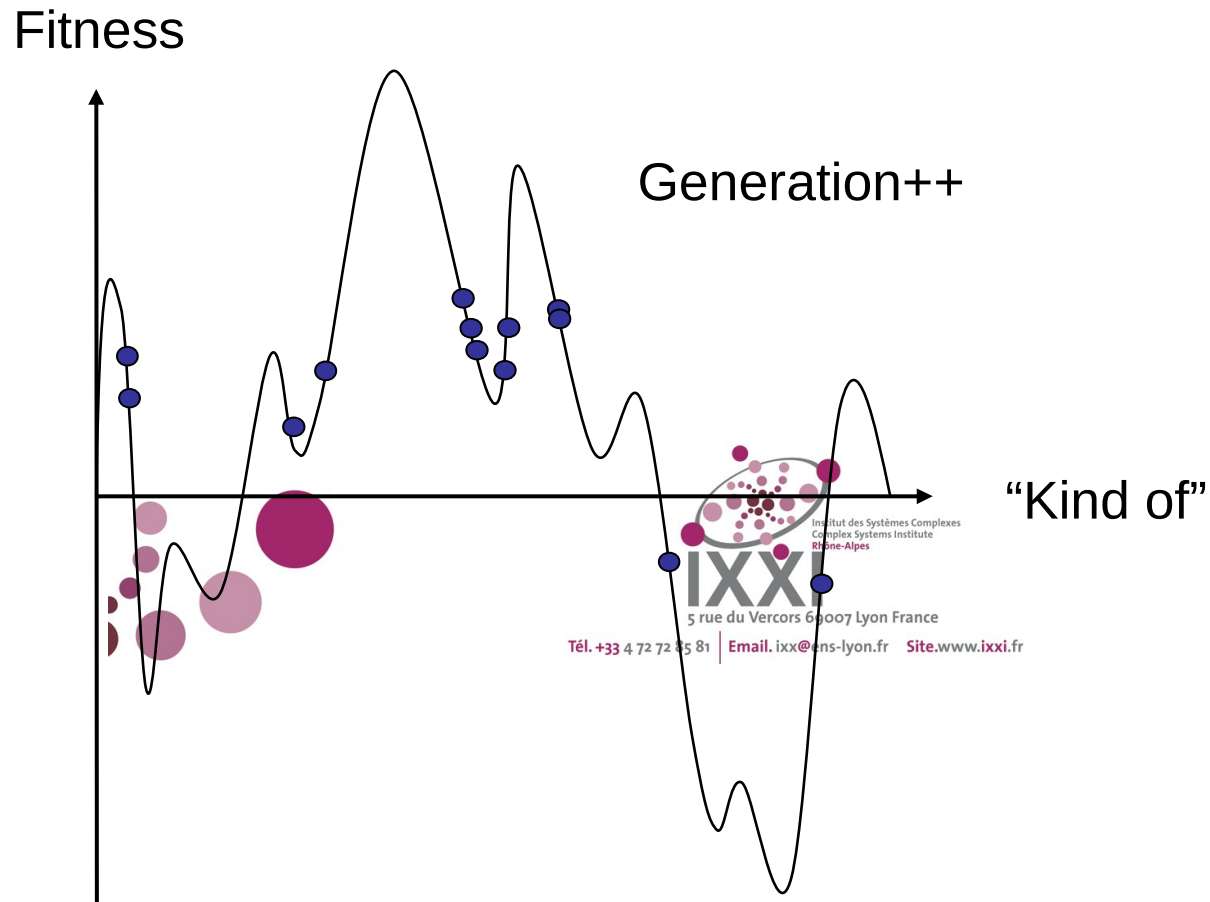
The fitness landscape metaphor



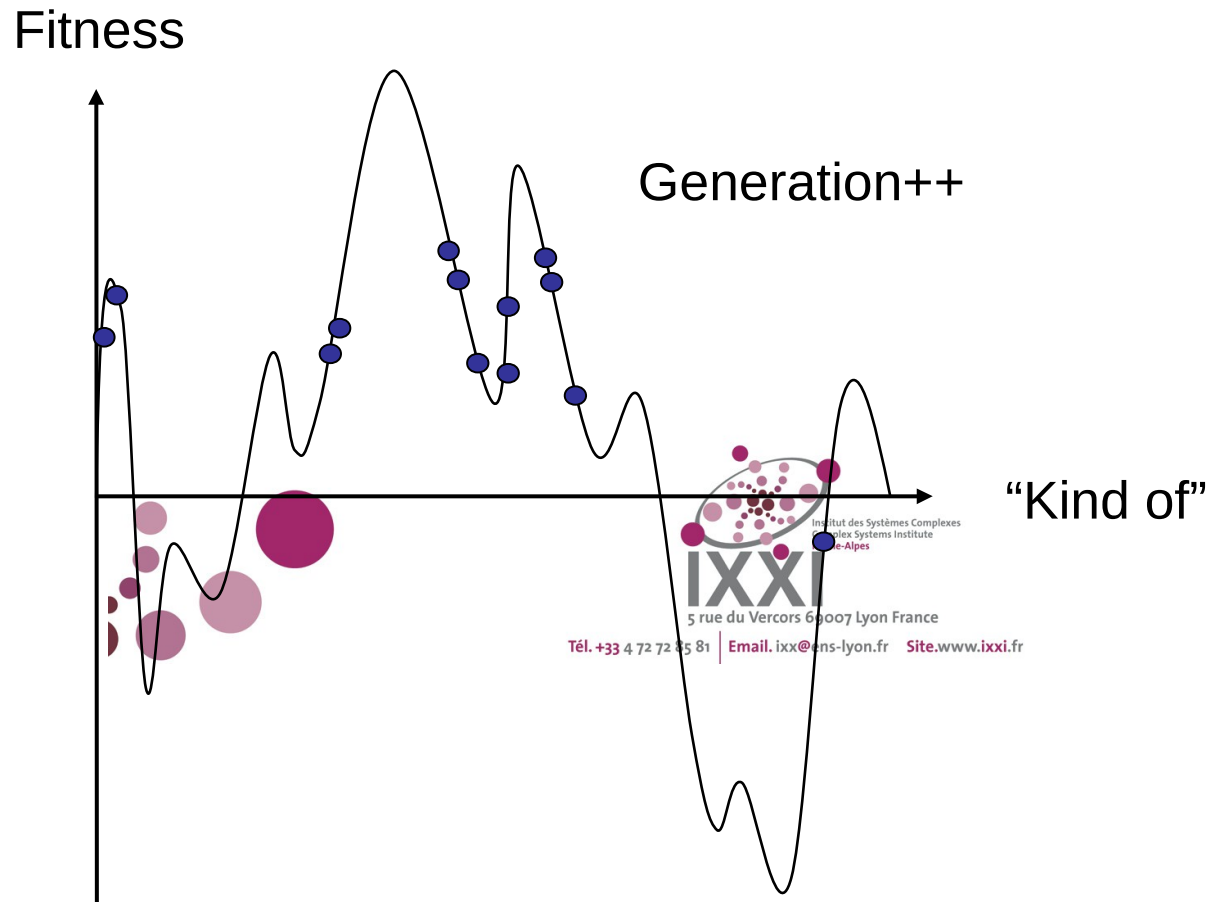
The fitness landscape metaphor



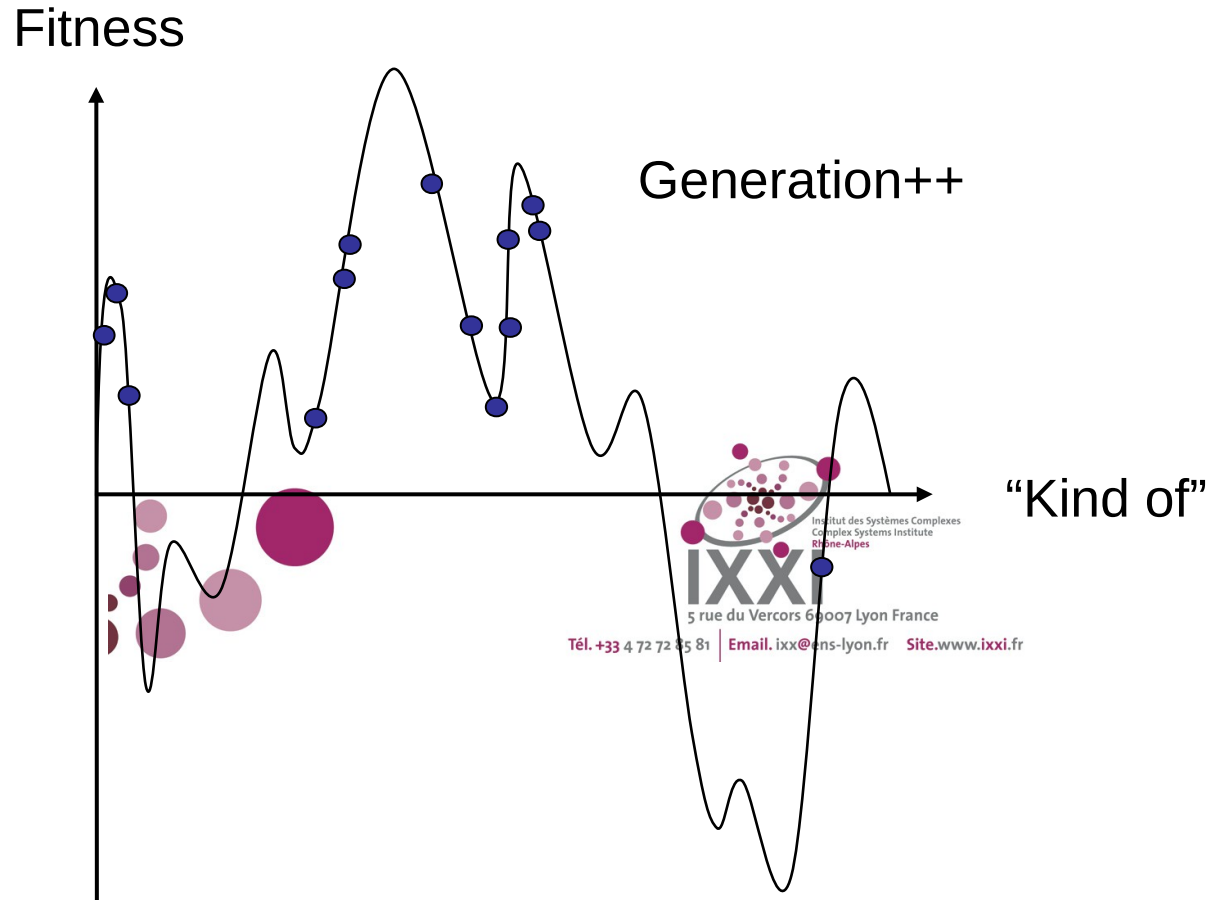
The fitness landscape metaphor



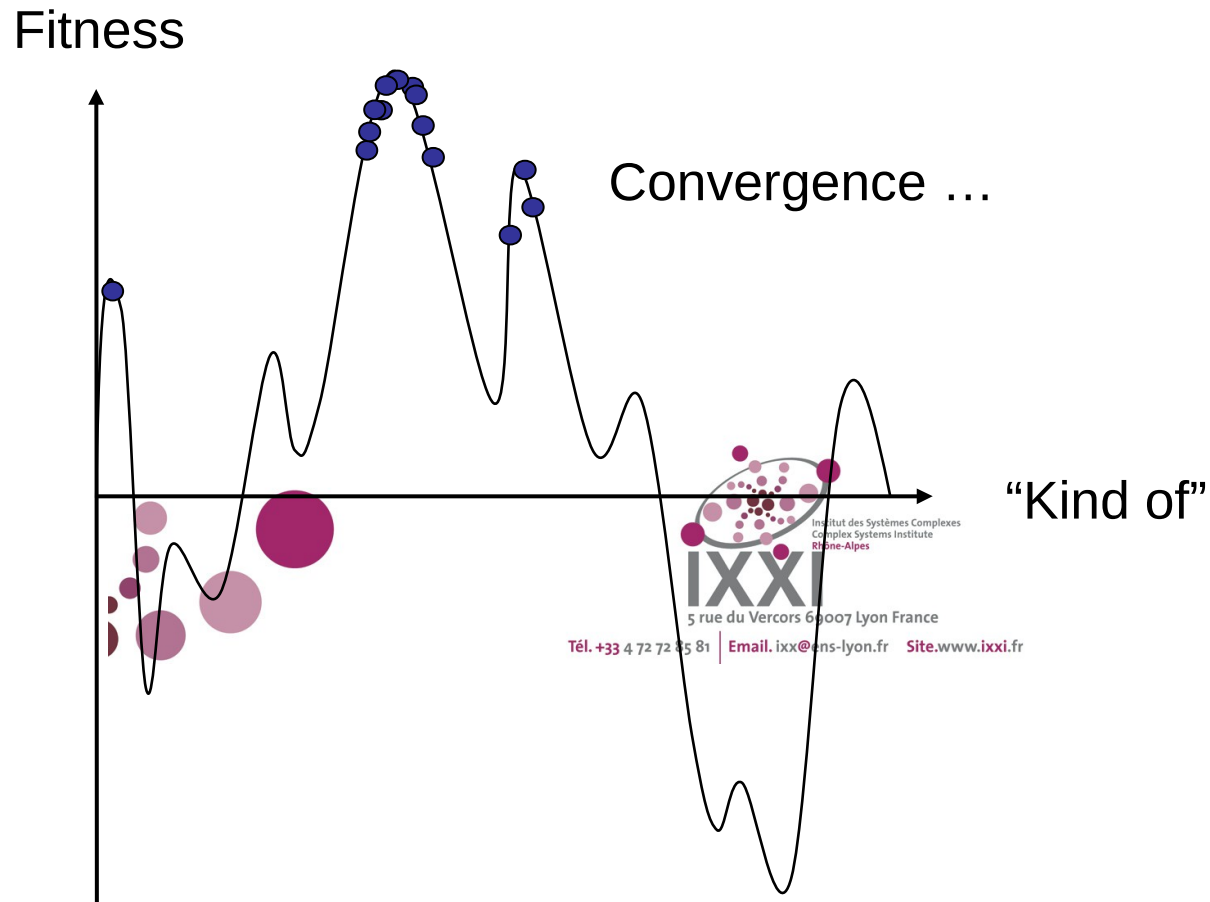
The fitness landscape metaphor



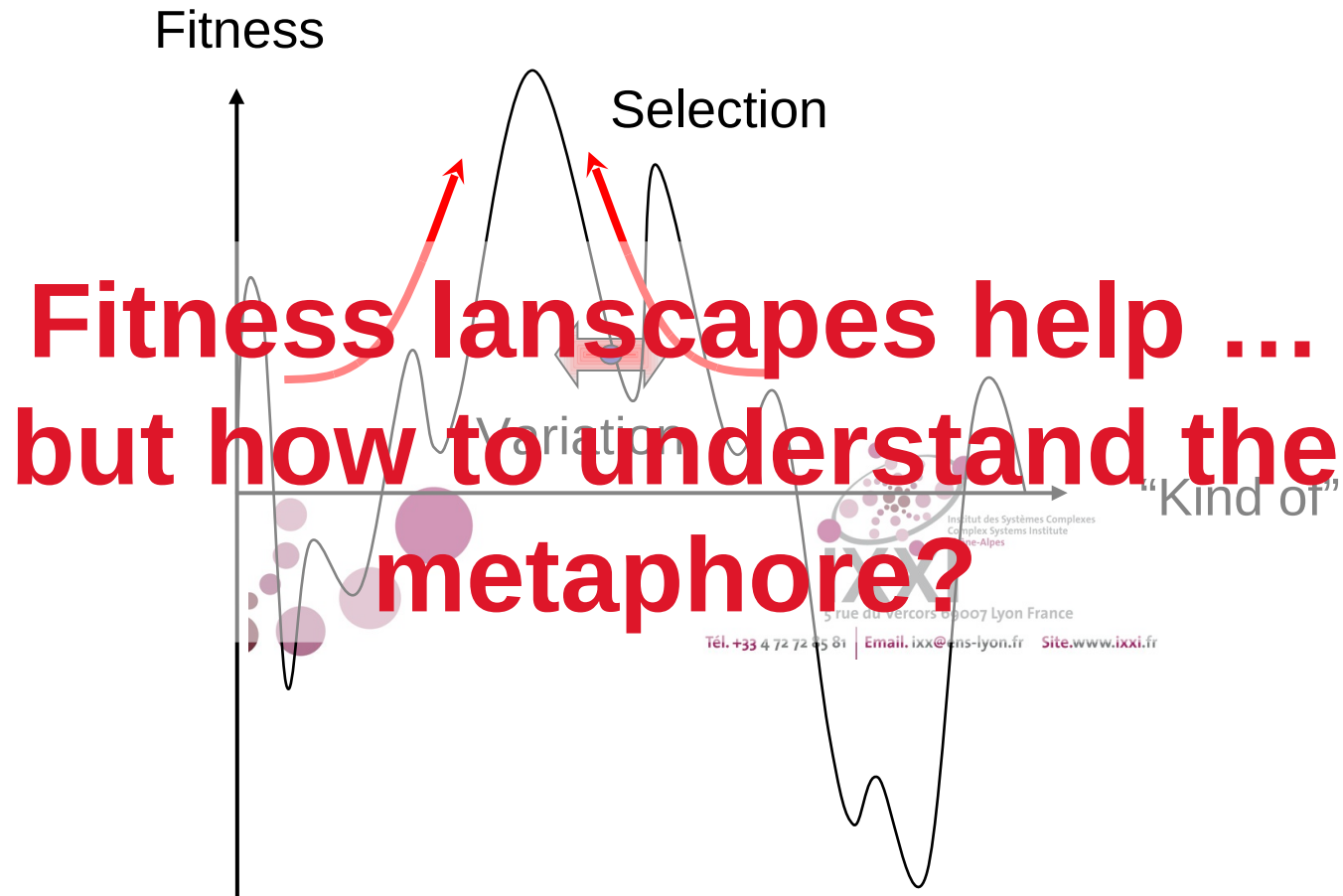
The fitness landscape metaphor



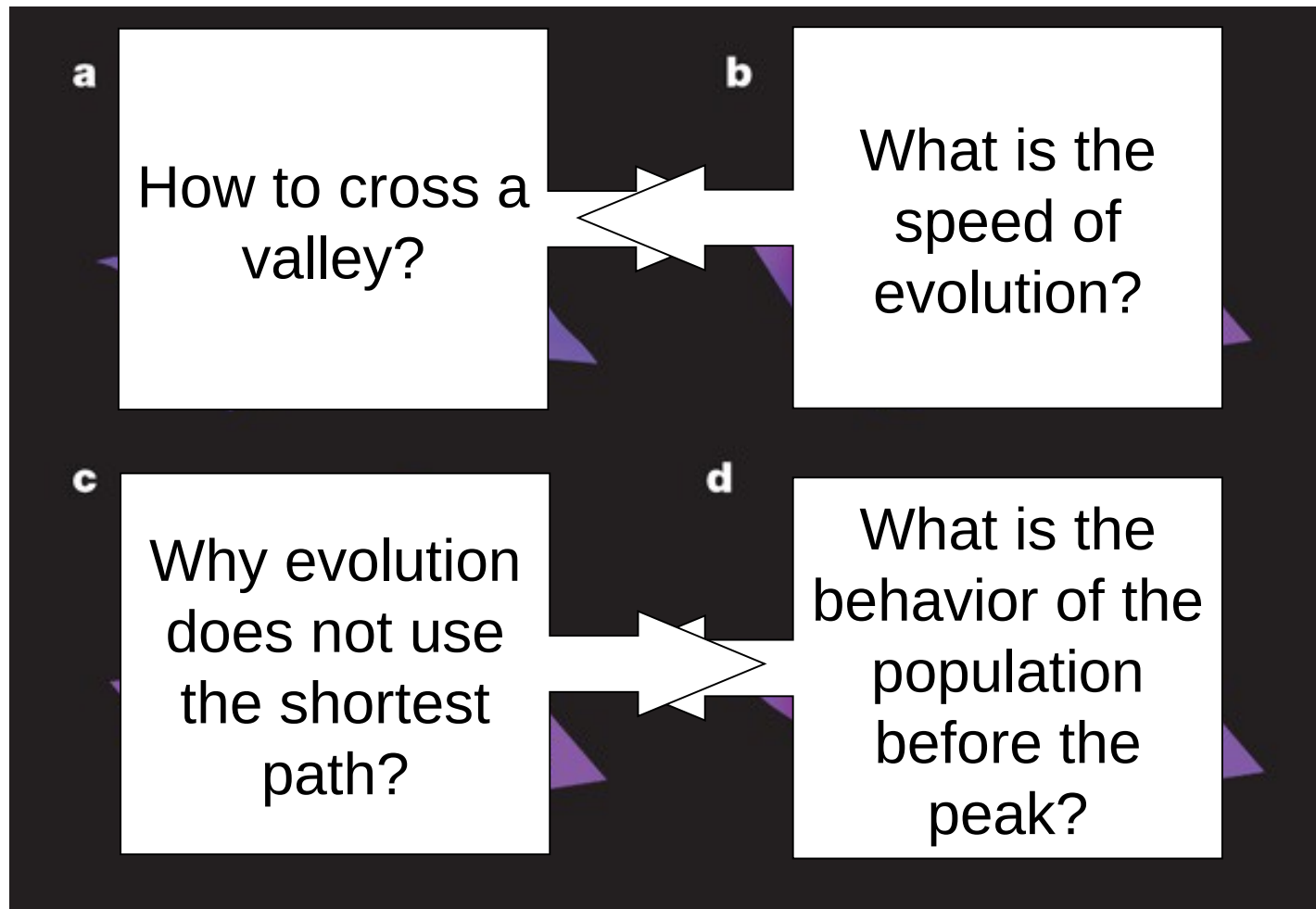
The fitness landscape metaphor



Two antagonist forces



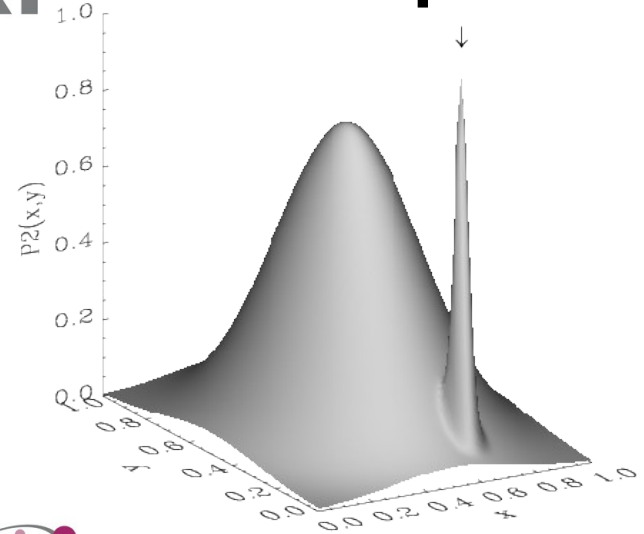
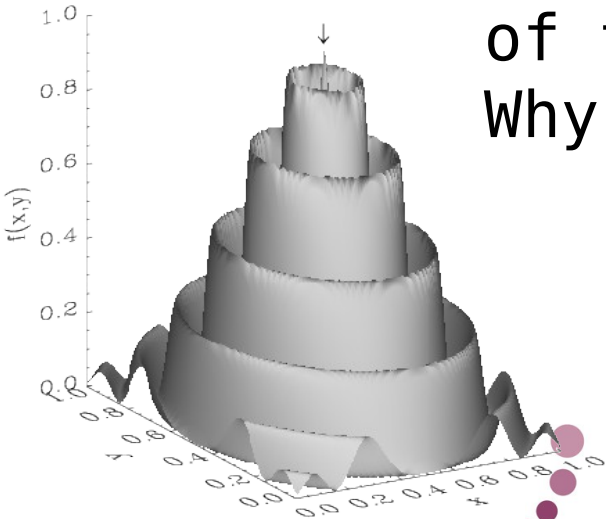
Fitness landscapes help thinking



[Poelwijk et al. 2007]

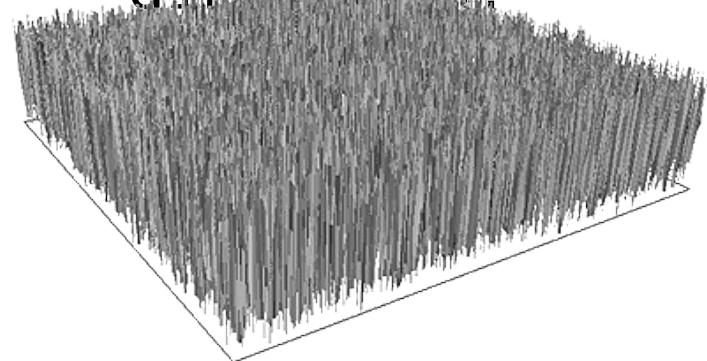
Questions of fitness landscape

What is the shape of the landscape?
Why?



Is the landscape static? If not, what triggers changes of the landscape shape?

What is the correct number of dimension?





Need for experimental evolutionary studies

- Evolution is a general mechanism that relies on many random events
 - How can we distinguish between the effect of the mechanism and the effect of the random events?
 - We only have a single “experiment” at our disposal!
- Many questions cannot be addressed without experiments (or only hardly addressed!)
 - Is there a trend in the evolution of biological complexity?
 - What if we start again?
 - Is evolution predictable?
 - Is evolution really universal? (Cf. Dennett)
 - What is true for *E. coli* is true for the elephant...

Experimental evolution

- Controlled experiments ARE possible for organisms which are
 - Cheap, small, abundant, controllable (organism and environment), fast (short generational time), measurable (sequence, fitness, ...), freezable ...
 - E.g., bacteria (*E. coli*, *salmonella*, ...), viruses and phages, yeast, *C. elegans*, *Drosophila*, ...
- Longest experiment in evolution
 - 12 strains of *E. coli* evolved during 40.000 generations in R. Lenski lab. at Michigan State University

<http://myxo.css.msu.edu/index.html>



QuickTime® et un
décodeur
sont requis pour visionner cette image.



Experimental evolution is not enough

- All known organisms share parts of their evolutionary history
 - We all come from LUCA (~3.5 billion years ago)
- Conditions are always changed by the experimental setup
 - What are the consequences on the evolutionary process?
- How can we analyze the results?
 - Real organisms are too complex for us!

“So far, we have been able to study only one evolving system and we cannot wait for interstellar flight to provide us with a second. If we want to discover generalizations about evolving systems, we have to look at artificial ones.”

[John Maynard Smith, 1992]

Artificial life



- Life inside a computer?
 - Free forms ...
- Digital experiment on controlled organisms (artificial life)

QuickTime^a et un
decompresseur
sont requis pour visionner cette image.

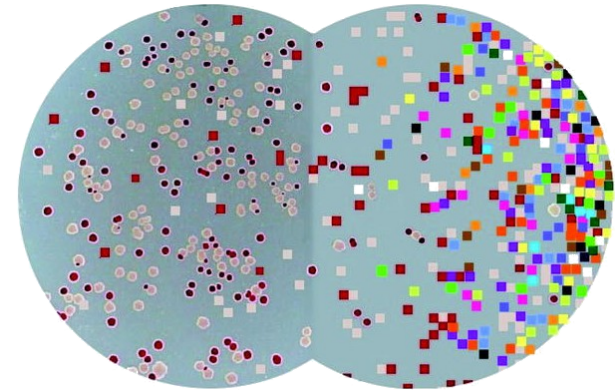


QuickTime^a et un
decompresseur
sont requis pour visionner



Artificial Life in few steps

- 1978 First attempts (C. Langton, LANL)
 - “*Life as it could be*”
- 1990 *Venus* simulator (S. Rasmussen, LANL)
- 1991 *Tierra* (T. Ray, U. of Delaware)
- 1992 *Creatures* (K. Sims, digital corp.)
- 1993 *Avida* (C. Adami., C.T. Brown, C. Ofria, Caltech)
 - Probably the most classical digital genetic software today
- 1996 *Amoeba* (A. Pargellis, Lucent)
- 2000 *Golem* project (H. Lipson, J. B. Pollack, Brandeis Univ.)
- 2005 *Aevol* (G. Beslon, C. Knibbe, INSA-Lyon)
- 2006 *Evolving robots* (D. Floreano, L. Keller, EPFL/UNIL)



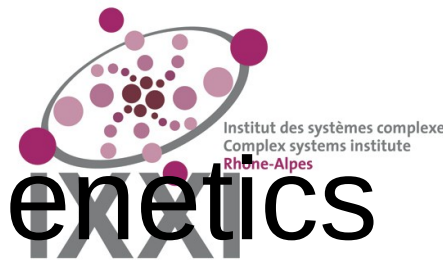
- Note 1: Lots of researchers don't use the term but construct models close to these ones (e.g., Paulien Hogeweg, Uri Alon, ...)
- Note 2: Artificial life not only focuses on evolution but evolution is the heart of artificial life



QuickTime^a and a
decompressor
are needed to see this picture.



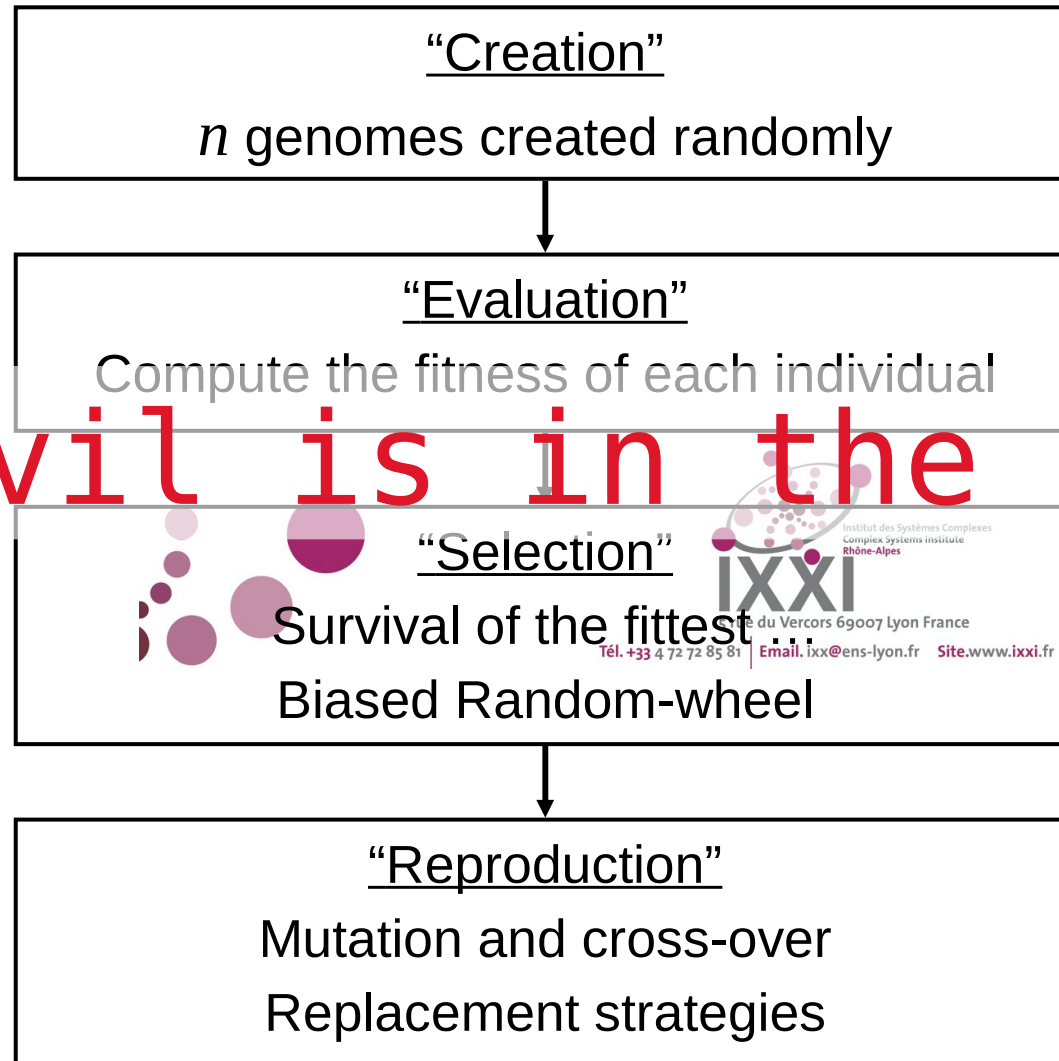
QuickTime^a and a
decompressor
are needed to see this picture.



Digital genetics

- Software that creates environment inside of a computer for populations of *self-replicating* elements, subject to *mutation* and *survival of the fittest*
 - “Real evolution of false organisms” (real Darwinism)
- This software can be used as an experimental setup
 - Modify some parameters of the simulation, look at the consequences on the organisms and/or on the ecosystem
 - Look for regularities...
- Experiments can be *repeated* many times for statistical accuracy.
 - All mutational events are known
- Digital Genetics = Agents-Based Modeling applied to evolution

Pseudo-code



the devil is in the detail

Evolved Virtual Creatures

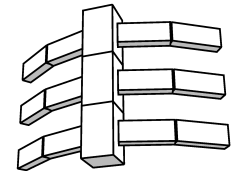
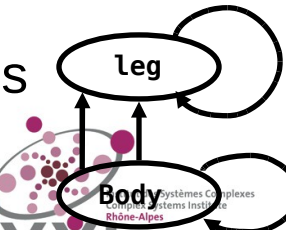
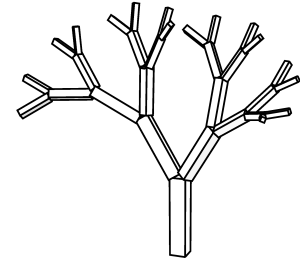
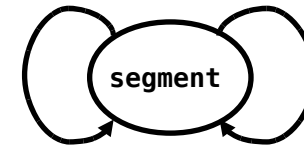
(Karl Sims 1994)

- Each creature is defined by a graph

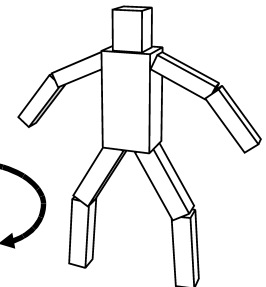
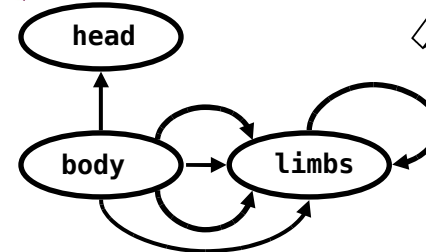
- One node = one body element
- One link = one joint
- Dual-links = multiple bodies
- Recursive links = repeated structures

- Nodes and links are valued

- Dimensions
- Joint limits
- Relative position
- Recursion control
- Joint control
- ...



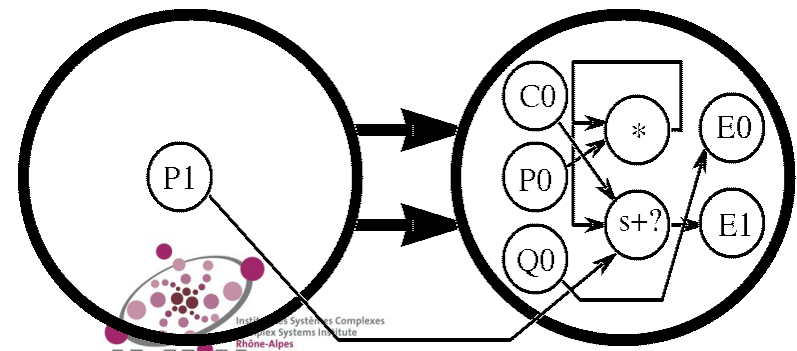
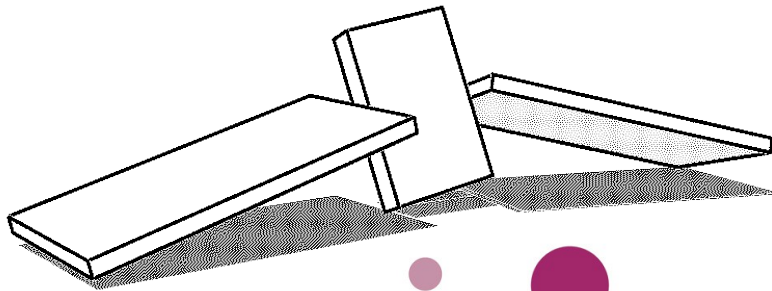
5 rue du Vercors 69007 Lyon France
Tél. +33 4 72 72 85 81 | Email. ixx@ens-lyon.fr | Site. www.ixxi.fr



Evolved Virtual Creatures

(Karl Sims 1994)

- Each creature owns a distributed brain that receives stimuli and produces motor output at the joints ...



- Example:
 - P1: body light-sensor
 - C0, P0, Q0 : “wings” light-sensors
 - *, s+? : computation elements
 - E0, E1 : joint motor control

Evolved Virtual Creatures

(Karl Sims 1994)

- Each creature “lives” in a precisely controlled world (viscosity, gravity, obstacles, light, ...)
 - The emergent morphology and behavior is strongly dependent on the environment condition (although highly variable)
- The main difficulty is the computation of the fitness values (i.e. the simulation part!)
 - Each simulation error is rapidly detected and used by the creatures!
- Nice! What can we conclude?
 - Hmm ... good question
 - It is almost impossible to disentangle the effect of evolution and environmental conditions from the effect of the (very complicated) genotype to phenotype mapping!
 - But Sims paved the way for many models (Framsticks, Golem...)

Too complex to comprehend?

- Creatures and similar models aim at simulating real “high level” organisms like mammals, birds, worms or snakes
 - The genotype-phenotype mapping is too complex
 - Interesting for engineering and computer graphics
 - Actually very few “real results” in evolutionary biology
- We need a more simple genotype-phenotype mapping
 - Models based on artificial chemistries
- Artificial chemistries
 - Computer instructions or sequences interpreted by a virtual CPU to produce the behavior of the organism
 - Historically artificial chemistries come from “core-war” games
 - Various formalisms [Dittrich et al., 2001] ...

Tierra: the ancestor

(Tom Ray, 1992)

“In Tierra, the self-replicating entities are executable machine code programs, which do nothing more than make copies of themselves in the RAM memory of the computer. Thus the machine code becomes an analogue of the nucleic acid based genetic code of organic life”

[T. Ray]

- Tierra enables to study the evolutionary behavior of evolving entities engaged in an “open-ended evolution”
 - No goal but (implicitly) survive and reproduce
 - Need to be sowed by some predefined code able to self-reproduce
- Tierra is an evolving ecological system

[<http://life.ou.edu/pubs/fatm/fatm.html>]

Tierra: the ancestor

(Tom Ray, 1992)

- Evolution of host-parasite systems (time 1)

QuickTime^a et un
décompresseur
sont requis pour visionner cette image.



Tél. +33 4 72 72 85 81 | Email. ixx@ens-lyon.fr | Site. www.ixxi.fr

Tierra: the ancestor

(Tom Ray, 1992)

- Evolution of host-parasite systems (time 1)

QuickTime^a et un
décompresseur
sont requis pour visionner cette image.



Tél. +33 4 72 72 85 81 | Email. ixx@ens-lyon.fr | Site. www.ixxi.fr

Tierra: the ancestor

(Tom Ray, 1992)

- Evolution of host-parasite systems (time 1)

QuickTime^a et un
décompresseur
sont requis pour visionner cette image.



Tél. +33 4 72 72 85 81 | Email. ixx@ens-lyon.fr | Site. www.ixxi.fr

Tierra: the ancestor

(Tom Ray, 1992)

- Evolution of host-parasite systems (time 1)

Nice but ...
Still we cannot conclude



Avida: the maturity

(Chris Adami, 1997)

- Avida is not only a “better” model; it also starts with better questions
 - Chris Adami interacts with biologists on an almost daily basis ...
 - Important collaboration with Richard Lenski
- Avida uses a simpler artificial chemistry than Tierra
 - Each “avidian” contains its own CPU (no interaction during code execution)
 - Avidians are immersed in a 2D space
 - The evolution is no more open-ended (the “fitness” don’t have the same meaning!) but the results are easier to analyze!
 - Better trade-of between simplicity and complexity of the model
- Many results in biology
 - See e.g., C. Adami, T. Collier, S. F. Elena, C. Ofria, C. Wilke, R. Lenski, D. Misevic...

“Avidians”



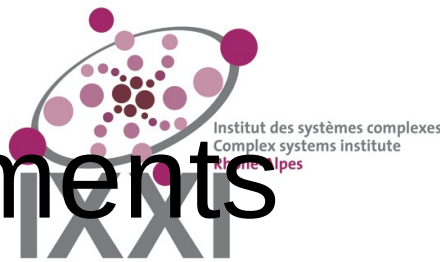
QuickTime^a and a
decompressor
are needed to see this picture.



Tél. +33 4 72 72 85 81 | Email. ixx@ens-lyon.fr | Site. www.ixxi.fr

[Adami, 2006]

Experiments



- Two different organisms, same conditions
 - Yellow organism: good but not robust
 - Blue organism = not so good but robust

Mutation rate: 0.5



QuickTime^a et un
dssacompresseur codec YUV420
sont requis pour visionner cette image.

Mutation rate: 1.5



Tél. +33 4 72 72 85 81 | Email. ixx@ens-lyon.fr | Site. www.ixxi.fr

QuickTime^a et un
dssacompresseur codec YUV420
sont requis pour visionner cette image.

“Survival of the flattest”

QuickTime^a and a
decompressor
are needed to see this picture.



QuickTime^a and a
decompressor
are needed to see this picture.

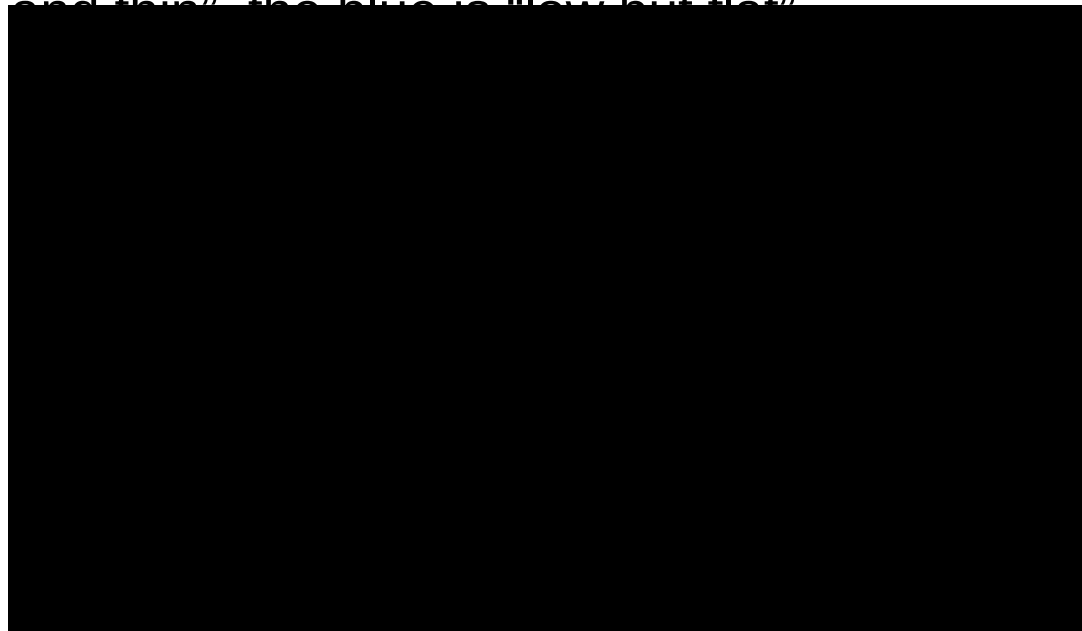


Tél. +33 4 72 72 85 81 | Email. ixx@ens-lyon.fr | Site. www.ixxi.fr

QuickTime^a and a
decompressor
are needed to see this picture.

Modeling the model

- “Survival of the flattest” [Wilke et al., Nature, 2001]
 - Under strong mutational pressure, sharp peaks are disadvantaged
 - When understood, the mechanism can be explained without the computational model (“*the model is no longer needed afterwards*”)
 - E.g., interpretation in terms of fitness landscape ... the yellow is “high and thin”, the blue is “low but flat”

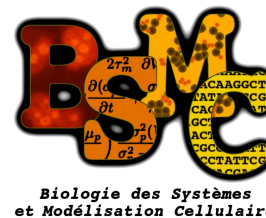


The aevol/R-aevol model

An individual-based model of genome and regulation networks evolution and structuring

Knibbe, C. (2006) Structuration de génomes par sélection indirecte de la variabilité mutationnelle, une approche par modélisation et simulation, PhD Thesis, INSA-Lyon, décembre 2006, 174 p.

Sanchez-Dehesa, Y. (2009) R-aevol, un modèle de génétique digitale pour étudier l'évolution des réseaux de régulation génétiques, PhD Thesis, INSA-Lyon, décembre 2009, 175 p.



Origin of genomic structures?

Institut des systèmes complexes
Complex systems Institute
Rhône-Alpes



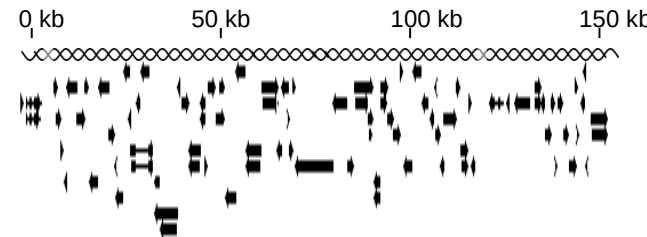
Homo sapiens

~3 billions bp
~25 000 genes



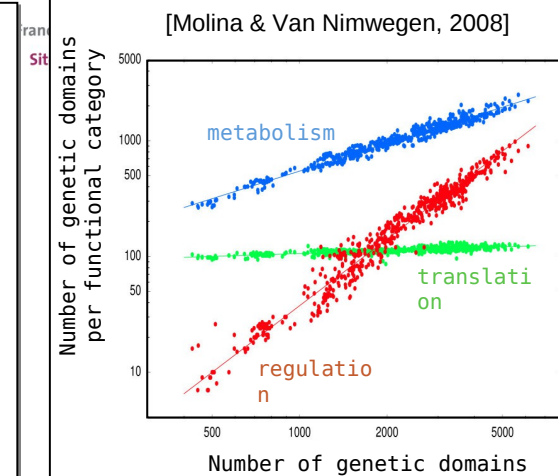
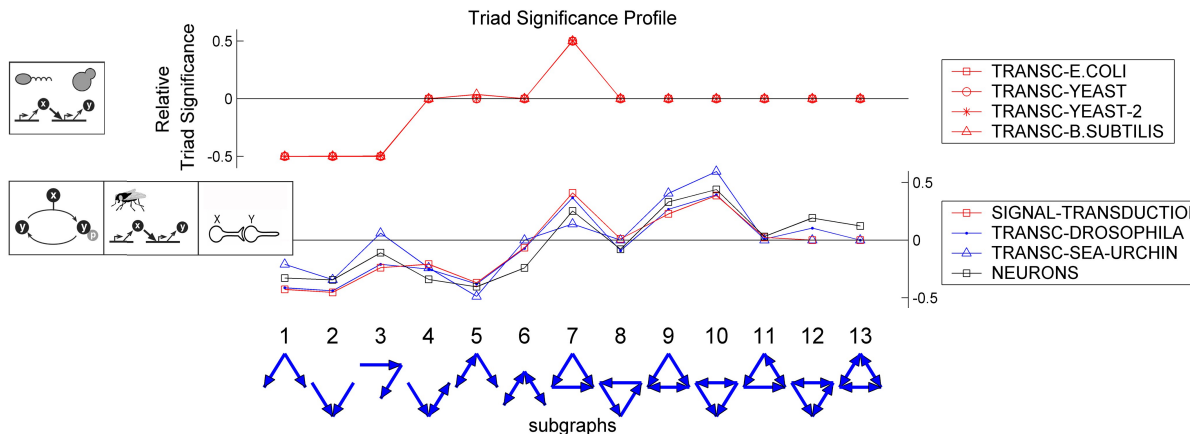
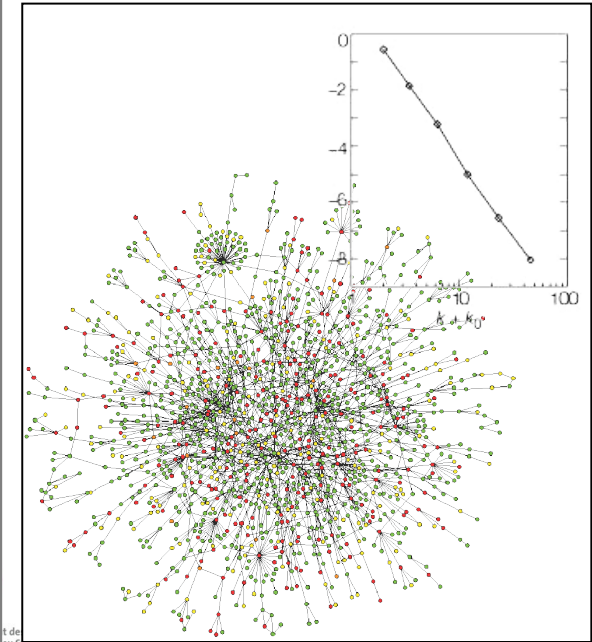
Neisseria meningitidis

~2 millions bp
~2 000 genes

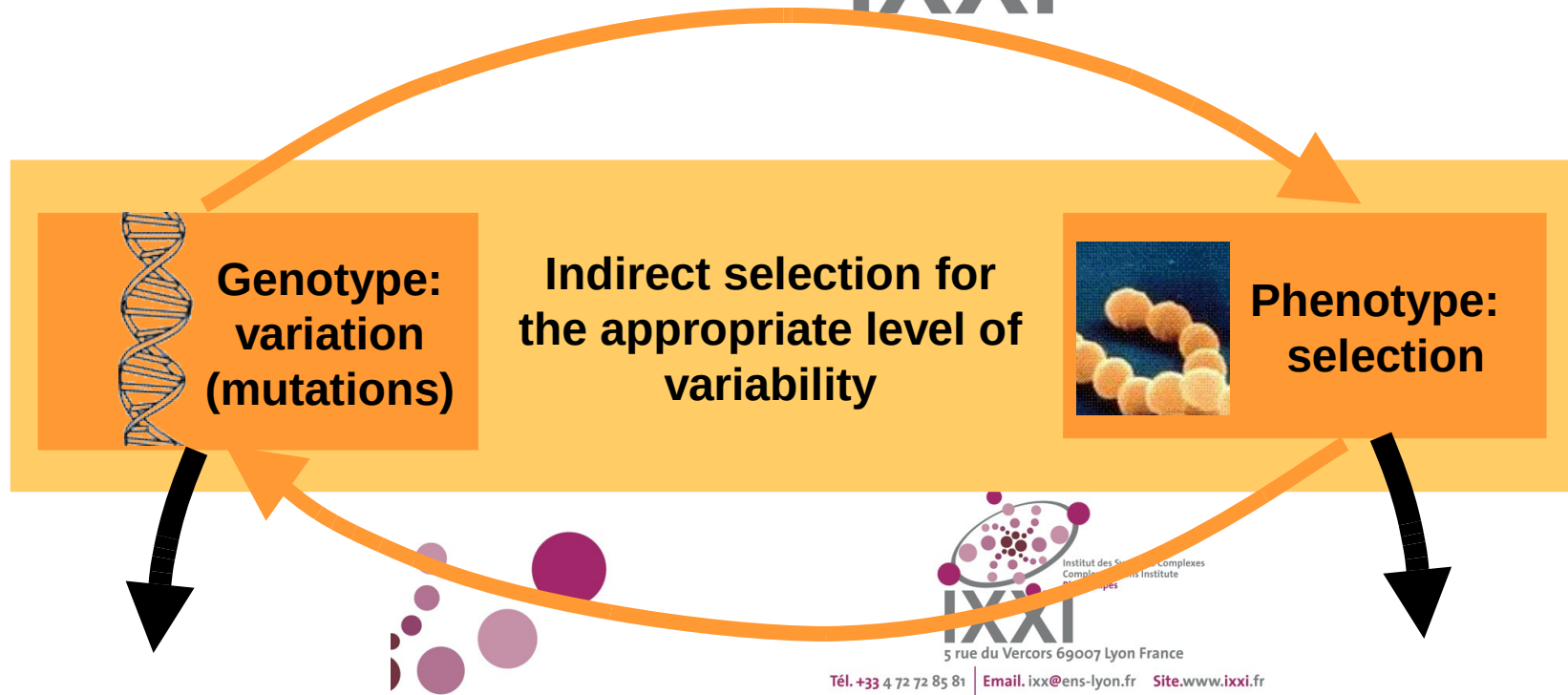


Herpes HSV-1

~150 000 bp
~100 genes



Origin of genomic structures?



Mutational biases:

“Homo Sapiens genome spontaneously undergoes more insertions than deletions”

Selective costs:

“A long genome can be disadvantageous for a bacteria or a virus”

Indirect selection

(a thought experiment)

institut des systèmes complexes
Complex systems institute
Rhône-Alpes

High variability level

(Low probability to reproduce)
 $F_v \approx 0$

Mid variability level

Low variability level

(High probability to reproduce neutrally):
 $F_v \gg 1$

Too frequent mutations:

Lineage extinction

→ indirect selection for robustness

Organisms can be (indirectly) selected depending on their robustness and evolvability (i.e. depending on their ability to evolve; second-order selection)

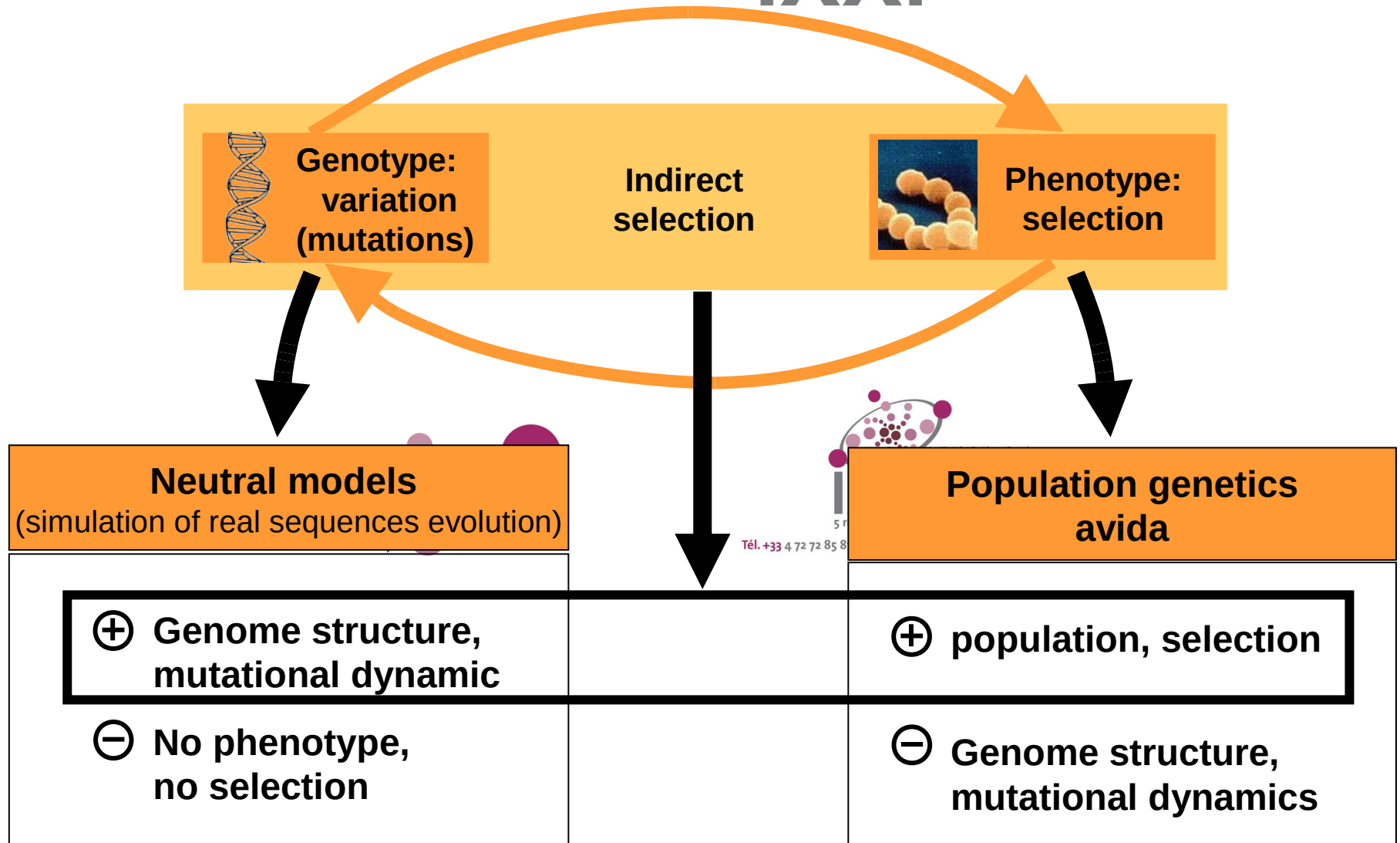
... But what are (i) the relative influence of direct and indirect selection? (ii) the effect of indirect selection on genome architecture? (iii) the range of parameters in which indirect selection occurs? (and many others)

Evolutionary dead end

→ indirect selection for variability

generations

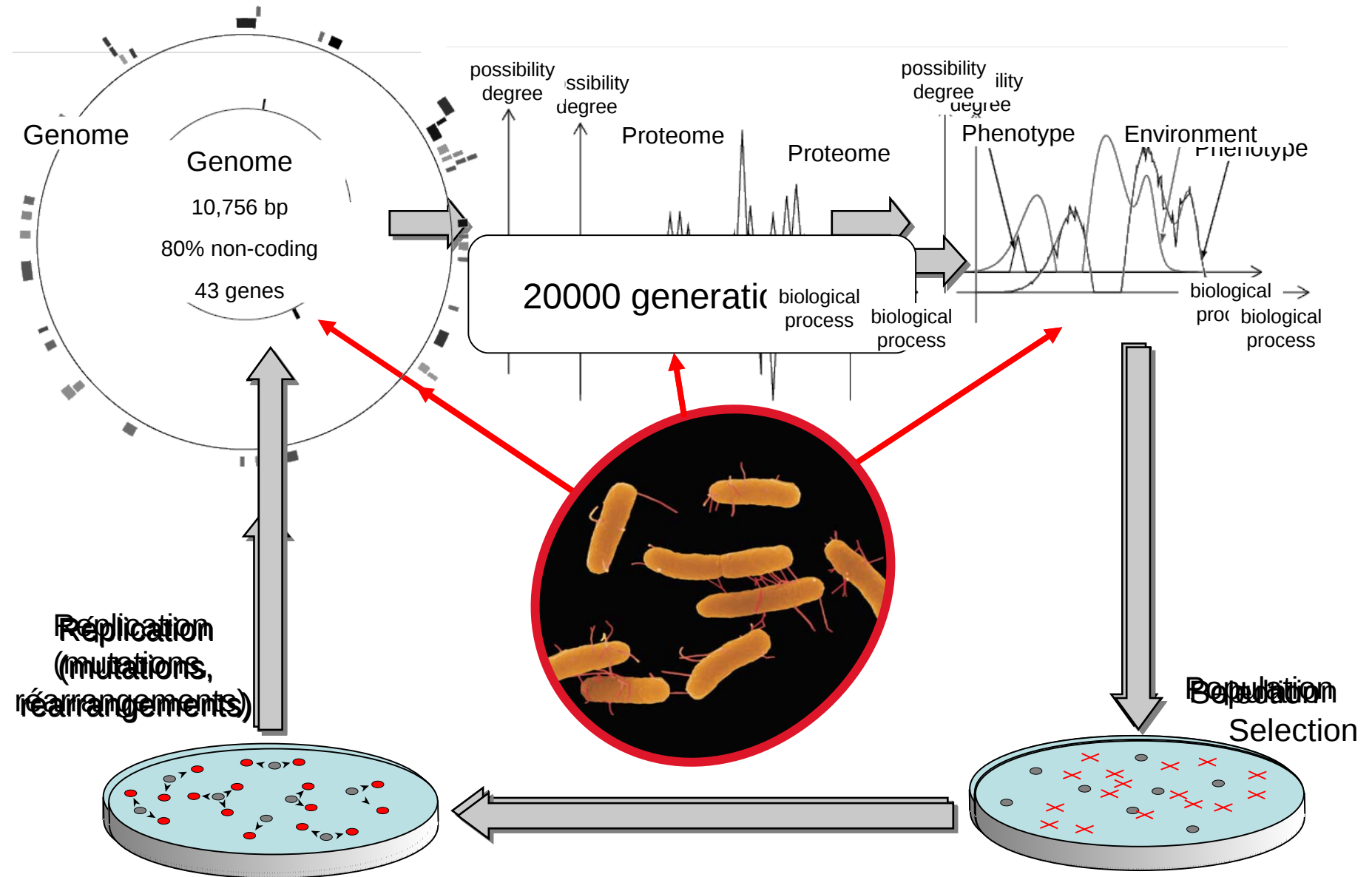
How to model indirect selection?



Contraintes sur le modèle

- Les structures moléculaires doivent être interprétables en termes biologiques
 - Génome, gènes, protéines, promoteurs, ...
- Elles doivent être “évoluables”
 - Nombre de gènes variable, taille du génome variable, ...
- Les structures moléculaires doivent être soumises à un processus mutationnel “réaliste”
 - Mutations ponctuelles
 - Remaniements chromosomiques
 - Transfert horizontal
- Les organismes ne doivent pas être sélectionnés en fonction de leur structure moléculaire
 - La sélection doit opérer sur le phénotype

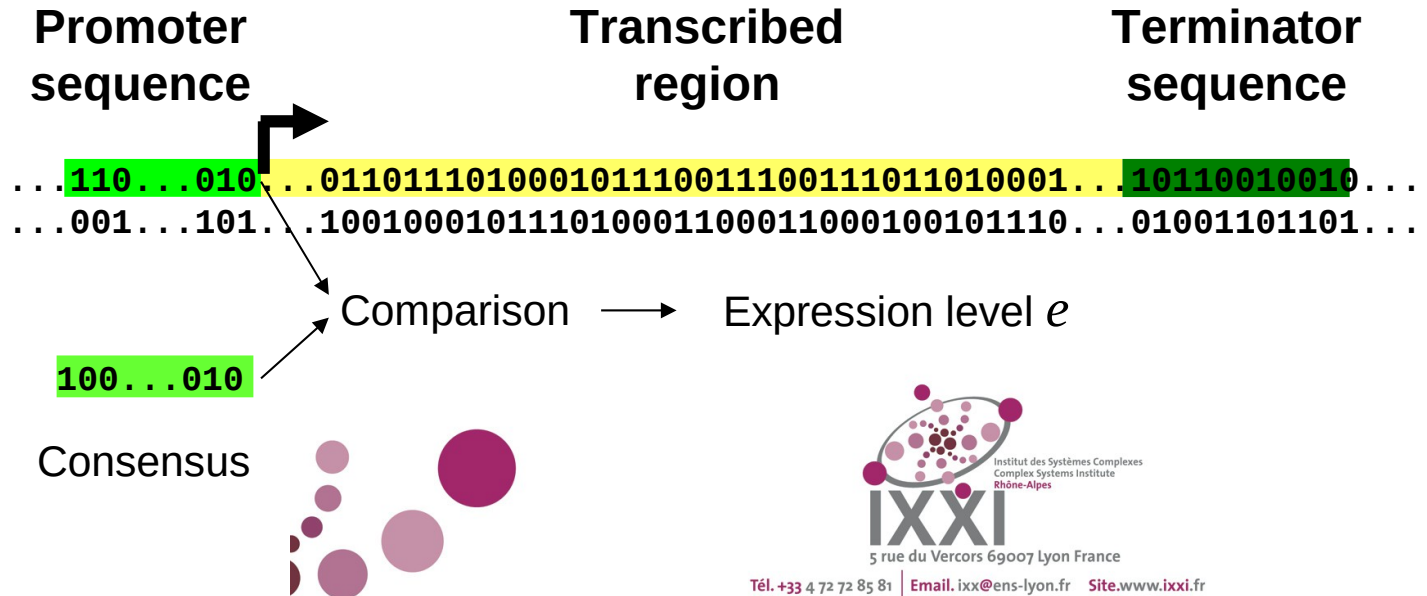
The aevo model



Transcription in aevol



Institut des Systèmes Complexes
Complex Systems Institute
Rhône-Alpes



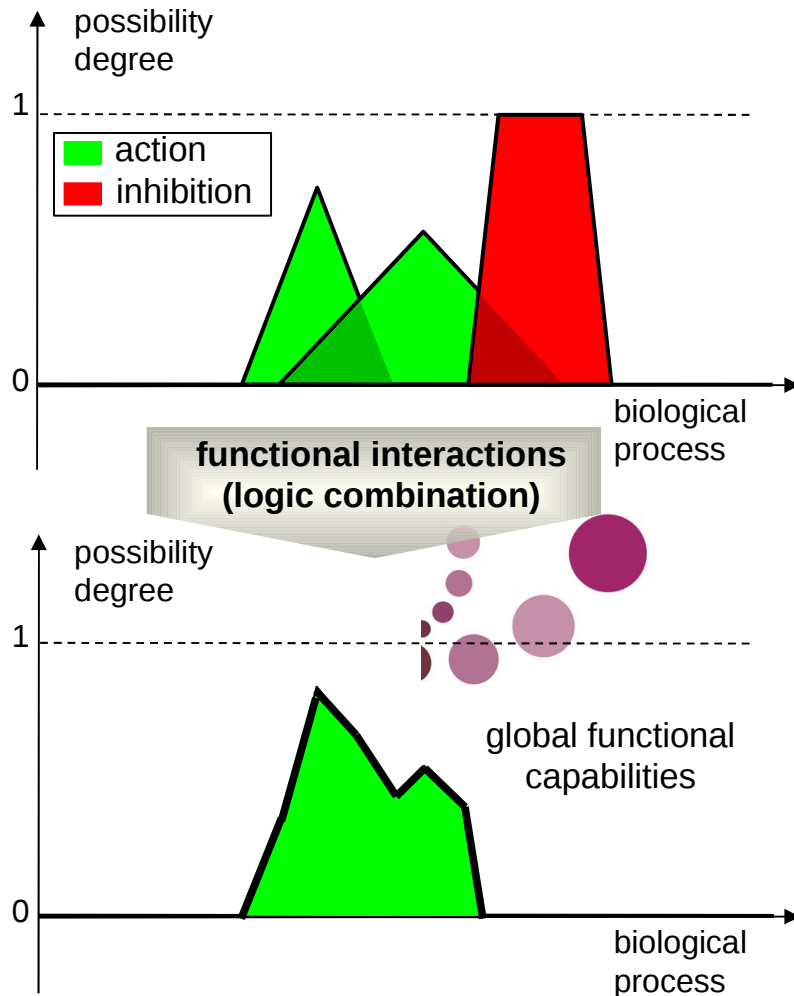
Institut des Systèmes Complexes
Complex Systems Institute
Rhône-Alpes

5 rue du Vercors 69007 Lyon France

Tél. +33 4 72 72 85 81 | Email. ixx@ens-lyon.fr | Site. www.ixxi.fr



Protein-protein interactions



- Interactions by triangles overlap
 - Pleiotropy
 - Polygeny
- Fuzzy sets combination
 - Phenotype = set of activated functions minus set of inhibited functions
 - Lukasiewicz operators

$$P = (\cup A_i) \cap \overline{(\cup I_j)}$$

Ævol: Reproduction cycle

Random or
clonal
initialization

N individuals

Reproduction
mutational process

Phenotype
computation

Comparison with
environmental
reference

Computation of W
(number of offspring)

$$W \approx N \cdot \text{prob}(\text{reproduction})$$

- Punctual mutations
- Small insertions
- Small deletions

- Translocations
- Inversions
- Duplications
- Large deletions

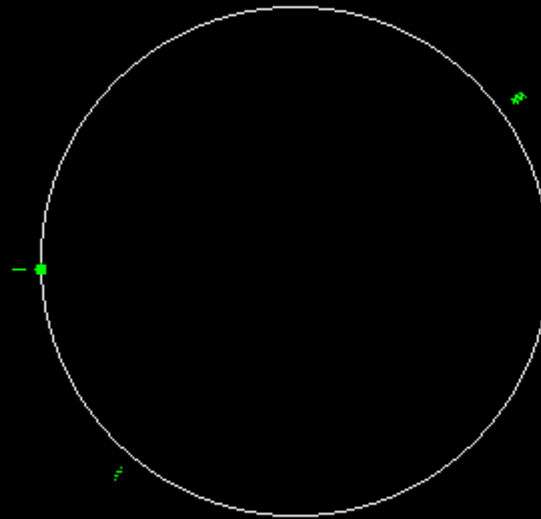
In mean, uL
per
reproduction

Ævol: The movie (« winning » lineage)

Genome length = 25993 bp

Generation = 129

Small insertion at 19406 of sequence 01



In-silico experimental evolution



- Mutation rate u :
 - Six mutation rates from $u = 5.10^{-6}$ to $u = 2.10^{-4}$ per bp
 - Same mutation rates for point mutations and rearrangements
- Selection:
 - Two selection modes (fitness proportional or rank-based)
 - Different selection strength (here $k = 250$ or $k = 1000$)
- Experimental evolution during 20000 generations
 - Populations: 1000 individuals
 - Steady environment
- Three repetitions per couple (u, k)
 - More than 100 simulations
 - It's *really* an experimental approach ...



Ævol: The movie (II) ...



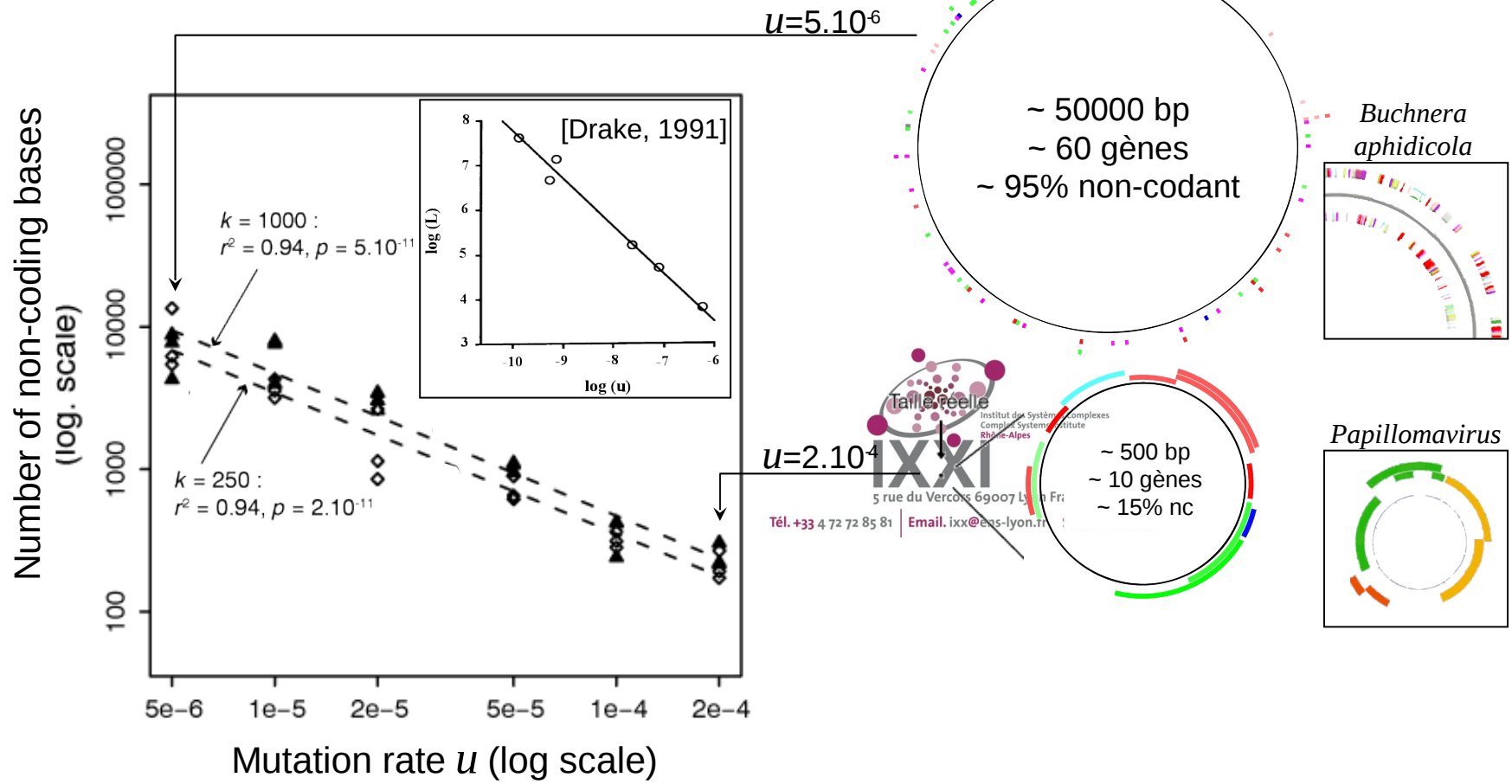
High mutation rates : $2 \cdot 10^{-4}$ / pb



Low mutation rates : $5 \cdot 10^{-6}$ / pb

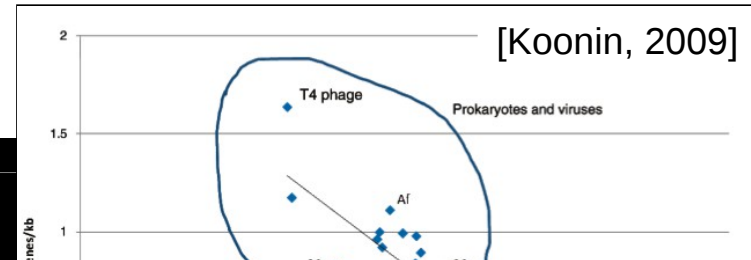
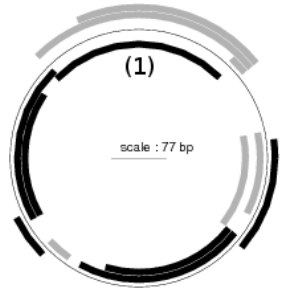
Scaling laws emerge *in silico*

IXXI



Yet another model explaining everything ;)

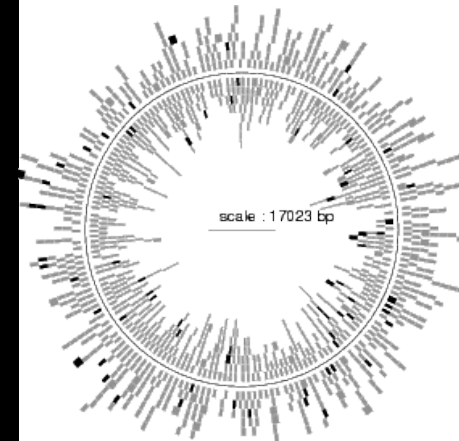
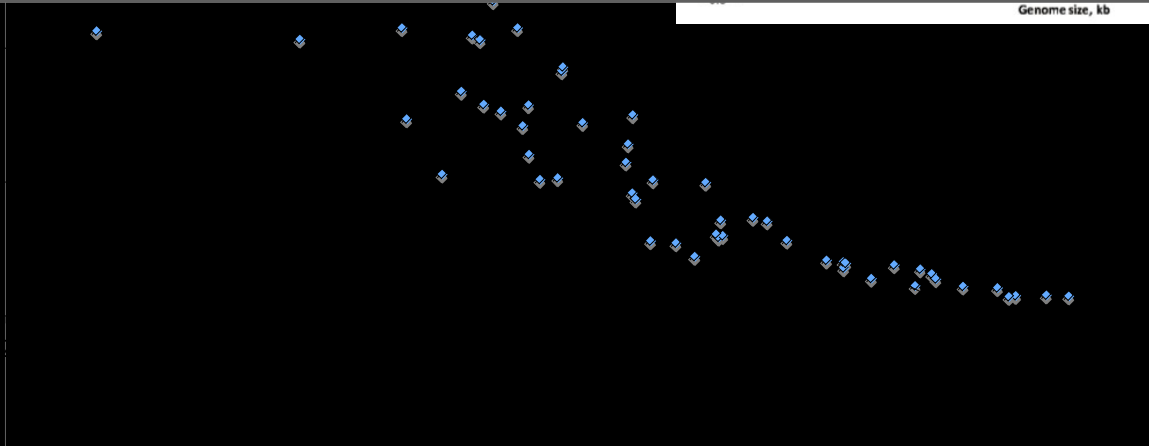
Institut des systèmes complexes
Complex systems institute
Rhône-Alpes



The model is able to reproduce known (but unexplained) data ...

But “*Prédire n’est pas expliquer*” (R. Thom) ...

Gene density

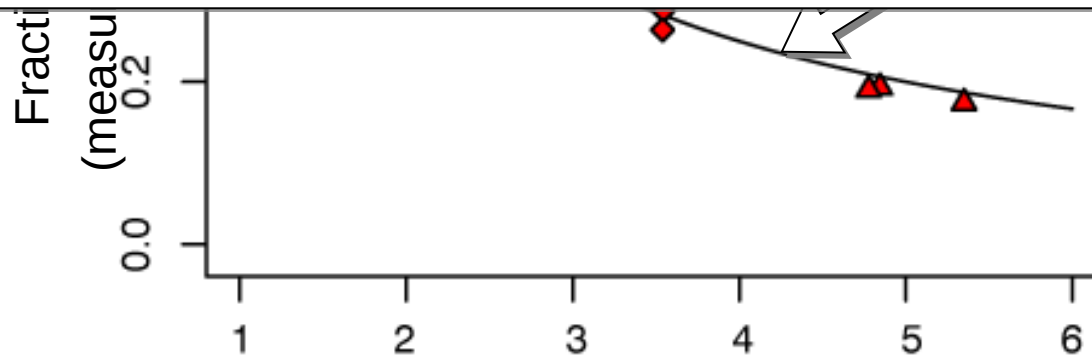


Experiments in the model



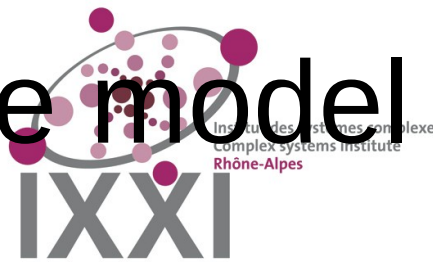
The Regulation of the number of neutral offspring is the hallmark of an indirect selection process; the link between the mutation rate u and the size of the non-coding sequences show that the indirect selection depends (at least partly) on these sequences...

... But what is the link? Where does the burden come from?



Number of reproductive trials : W
(depends on the competitors)

Modeling the model



- Mathematical model of reproduction
 - The math model represents aevol *AND* the “real world”...
- F_v : Probability of neutral reproduction as a function of genome size, (L), mutation rate (u) and neutrality of each kind i of mutation (v_i):

• v_i : Probability for a mutation of type i to be neutral depending on the genome structure:

If: (i) genomes undergo large duplications and deletions, (ii) the number and the average size of these events increase with genome size, Then: the mutational variability of a lineage depends on the amount of non-coding DNA (it is mutagenic for the genes it surrounds).

Thus the indirect selection for an appropriate level of variability actually selects for a specific amount of non-coding DNA

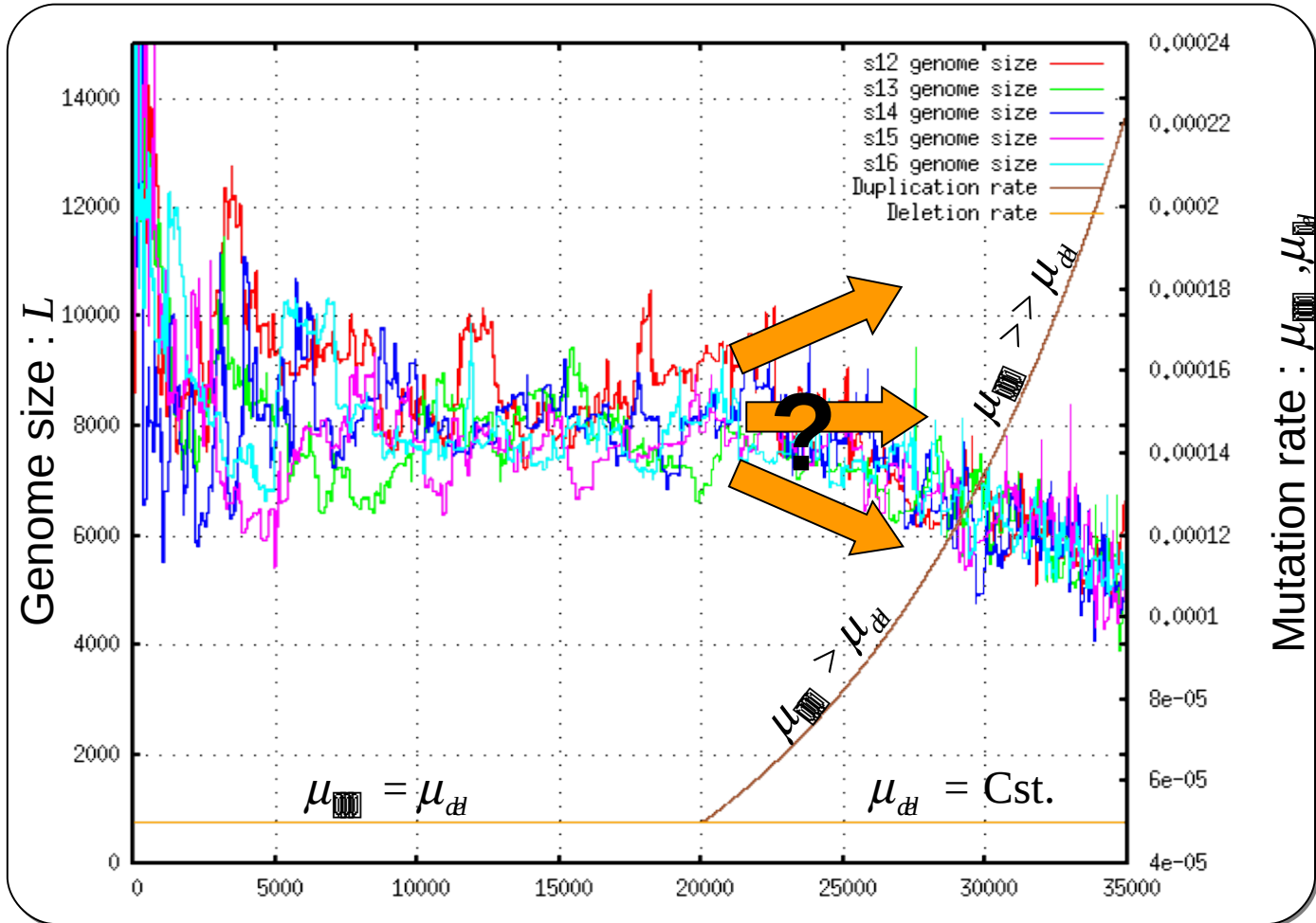
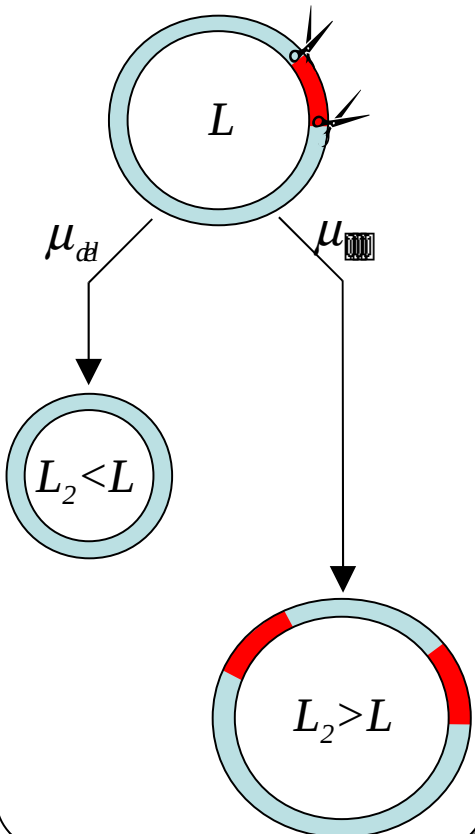
$$\left\{ \begin{array}{l} \tilde{\nu}_{\text{punct}} = \tilde{\nu}_{\text{ins}} = \tilde{\nu}_{\text{del}} = 1 - \frac{l}{L} \\ \tilde{\nu}_{\text{inv}} = \left(1 - \frac{l}{L}\right)^2 \\ \tilde{\nu}_{\text{transloc}} = \left(1 - \frac{l}{L}\right)^3 \end{array} \right. \quad \left\{ \begin{array}{l} \tilde{\nu}_{\text{gdel}} = \frac{1}{2L^2} \sum_{j=1}^{N_G} \lambda_j (\lambda_j + 1) \\ \tilde{\nu}_{\text{dup}} = \frac{1}{2L^2} \left(1 - \frac{l}{L}\right) \sum_{j=1}^{N_G} \lambda_j (\lambda_j + 1) \end{array} \right.$$

Care natural interpretations ;)



« It is simply a truism that the observed genome size is the result of a balance between the rate of DNA gain and loss » (Gregory, 2004)

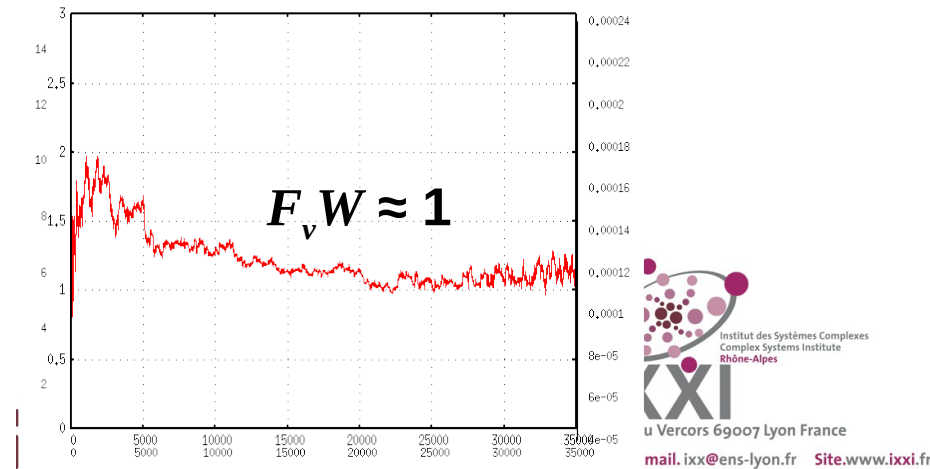
DNA gain: duplications
DNA loss: deletions



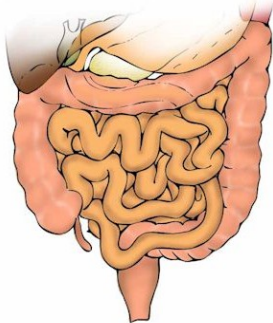
Care natural interpretations ;)



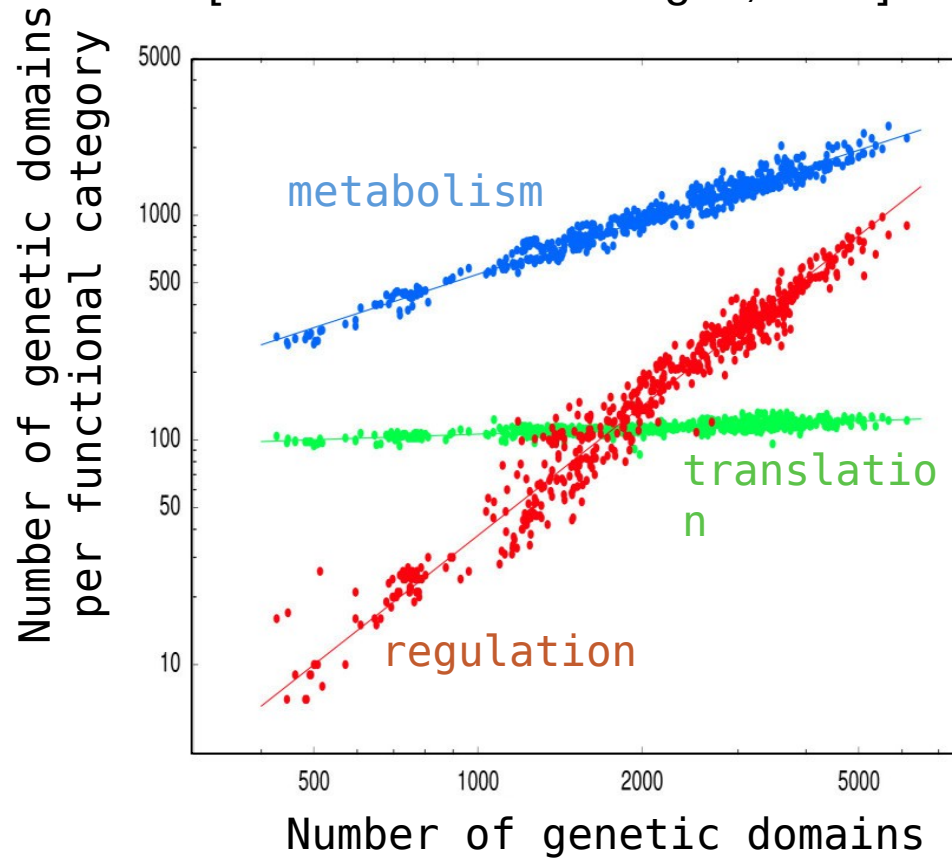
- Metabolic error ~ inverse of the fitness value
 - After generation 20 000, the metabolic error increases!
 - Evolution DECREASES the fitness ! (how is it possible?)



What about gene networks ?



[Molina & Van Nimwegen, 2008]



Acyrthosiphon pisum

bp

transcription factors

Escherichia coli

bp

s

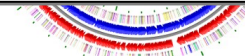
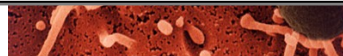
transcription factors

Escherichia coli

bp

es

275 transcription factors



BAOys: Friday April 15 09:42:20 2005

Length: 4,630,675 bp; Genes: 4,254



Institut des systèmes complexes
Complex systems institute
Rhône-Alpes

Contraintes sur le modèle

- L'évolution des réseaux de gènes est un “hot-topic”
 - Voir les travaux de W. Banzhaf, D. Floreano, P. Hogeweg, U. Alon, J. Knabe ...
- Pourquoi un modèle de plus sur la base de aevol ?
 - Un modèle de l'évolution des réseaux de régulation est d'abord un modèle de l'évolution !
- Les contraintes posées pour aevol sont toujours valides !
 - C'est le génome qui mute
 - C'est le phénotype qui évolue
 - Le réseau est “entre les deux” : il évolue indirectement !
- Contraintes supplémentaires :
 - Le réseau évolue en *CIS* et en *TRANS* ...
 - Pas de mutation directe des liens !
 - Attention : ici la différence procaryotes/eucaryotes est fondamentale !



R-aevol: regulation in aeol



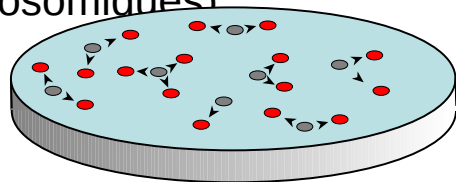
In R-aevol, the organisms own a genome and a regulation network. The network is made of metabolic genes and transcription factors

...

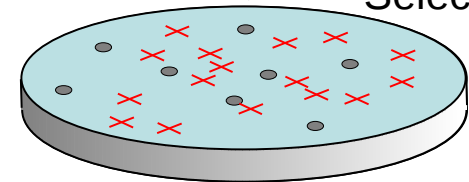
Experiments in R-aevol: How does the network structure depend on the evolutionary conditions? (e.g., environment complexity)

Processus
logique

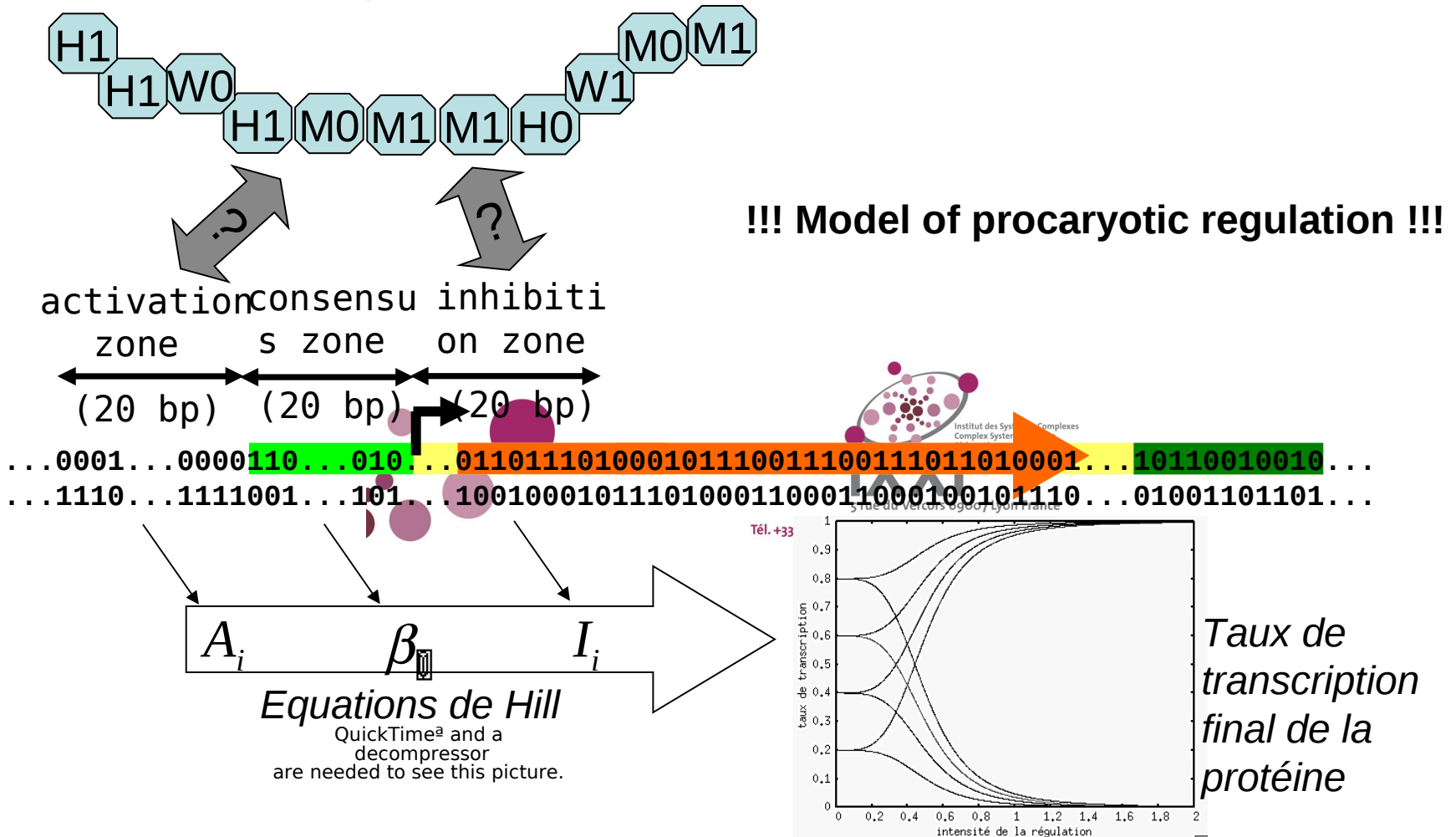
Réplication
(mutations, indels,
réarrangements
chromosomiques)



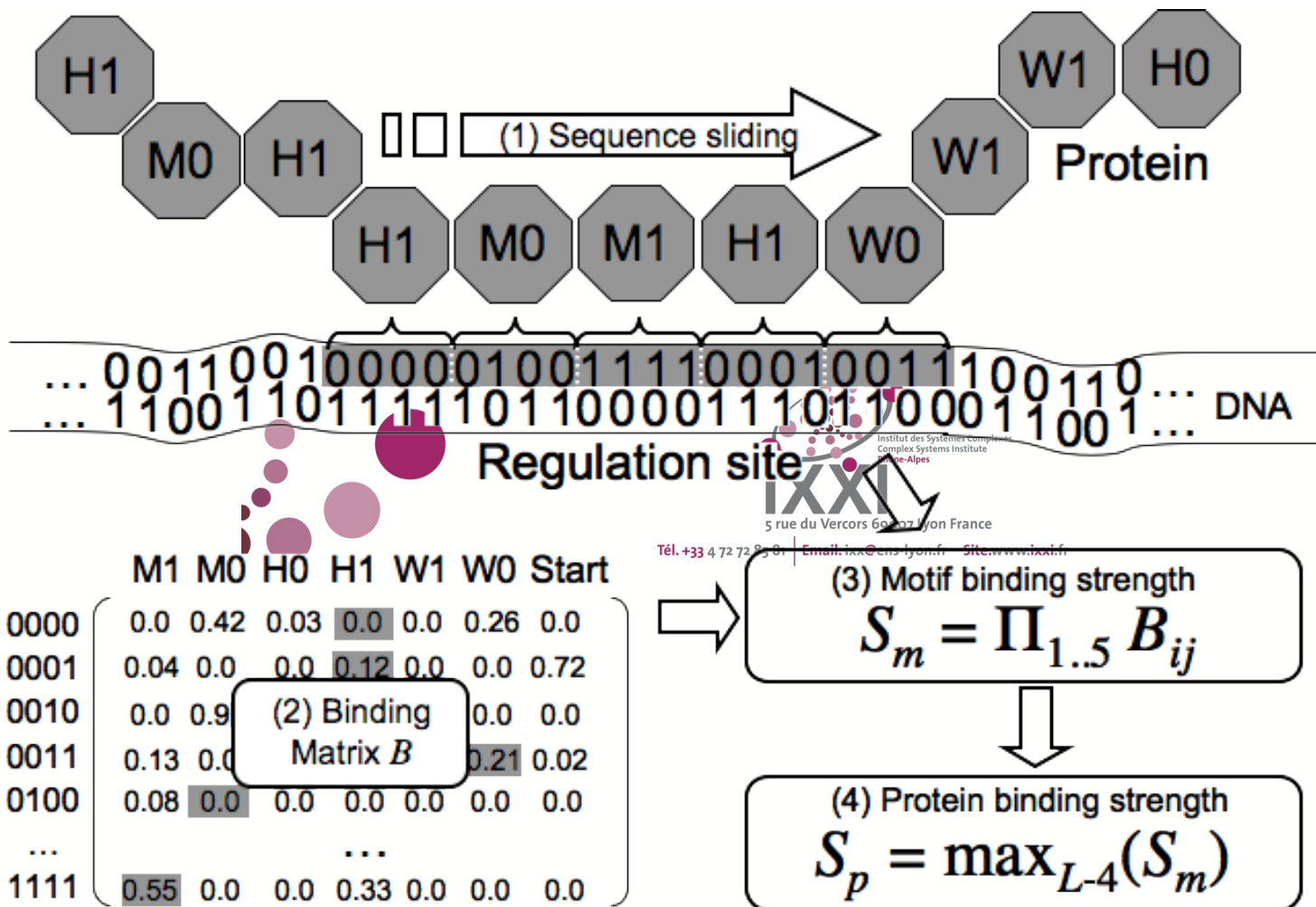
Population
Sélection



R-aevol: introducing regulation into aevo



R-aevol: introducing regulation into aevo

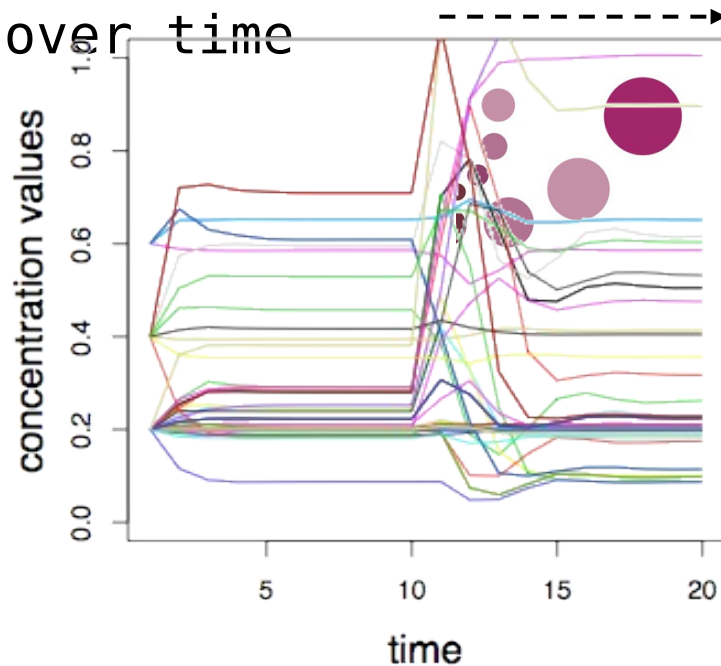


R-aevol: regulation in aevo

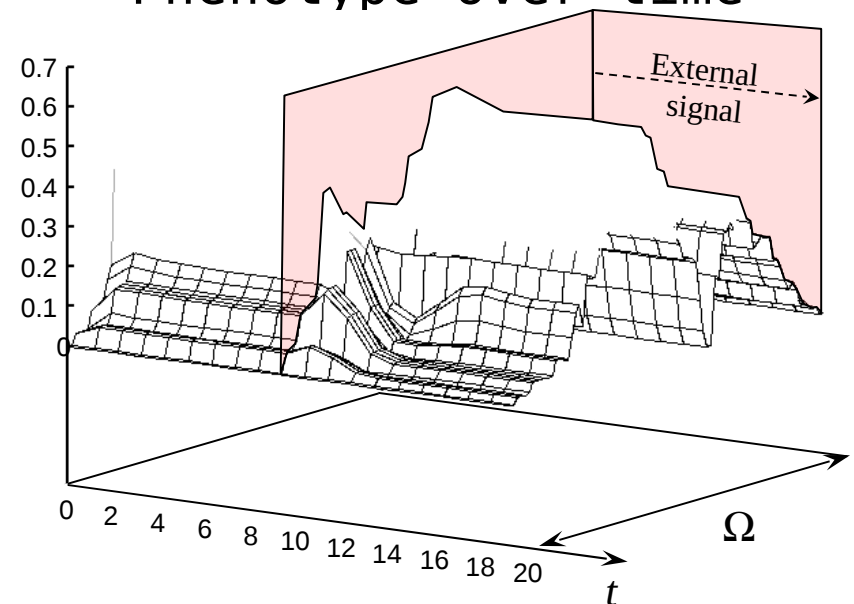


- In R-aevo, the organism's phenotype becomes a function of time
 - Organisms have a “life”; they can interact with their environment
 - Experiments in a two-states environment; the metabolic error is computed at $t = 10$ and $t = 20$

Protein concentrations over time

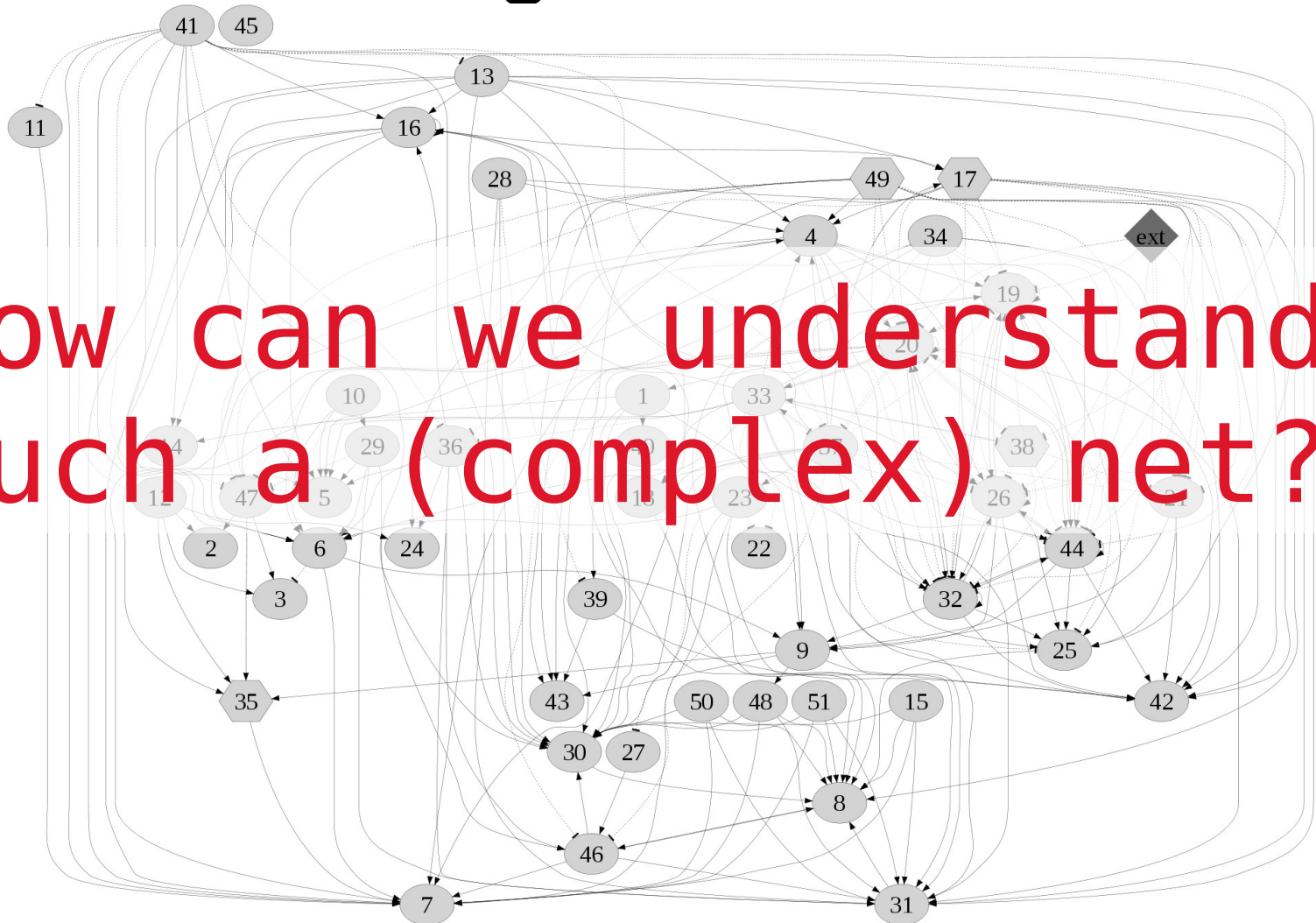


Phenotype over time



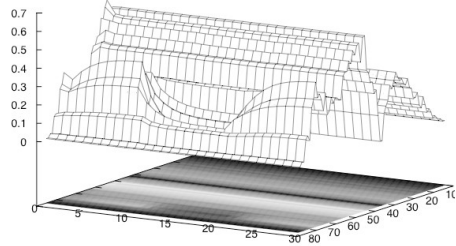
Evolved network after 15000 generations

How can we understand
such a (complex) net?

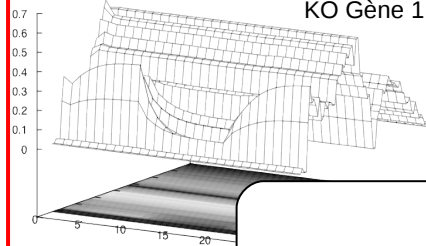


Systematic Knock-Out experiments

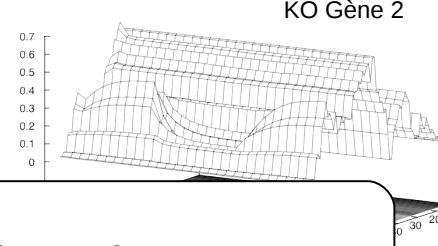
Wild Type



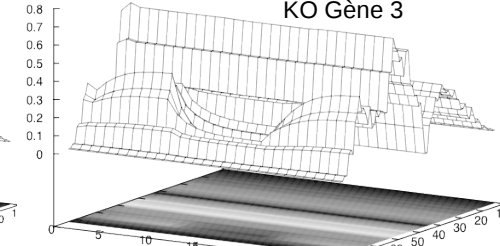
KO Gène 1



KO Gène 2

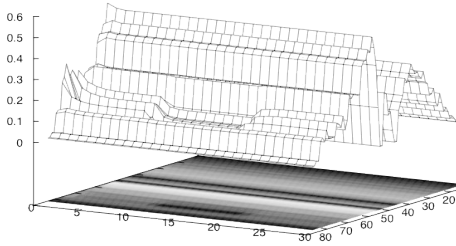


KO Gène 3

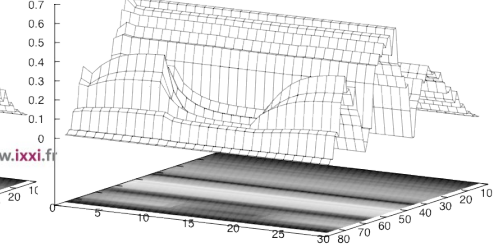


Clustering...

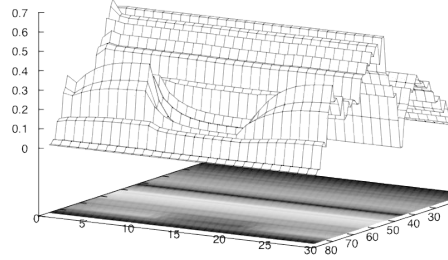
KO Gène 4



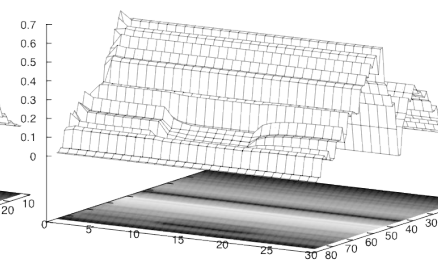
KO Gène 7



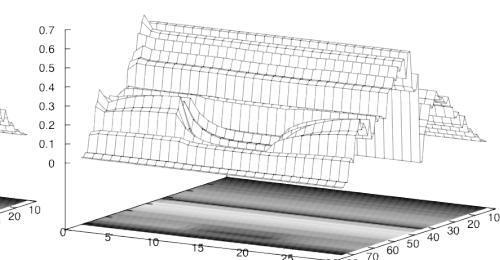
KO Gène 31



KO Gène 32

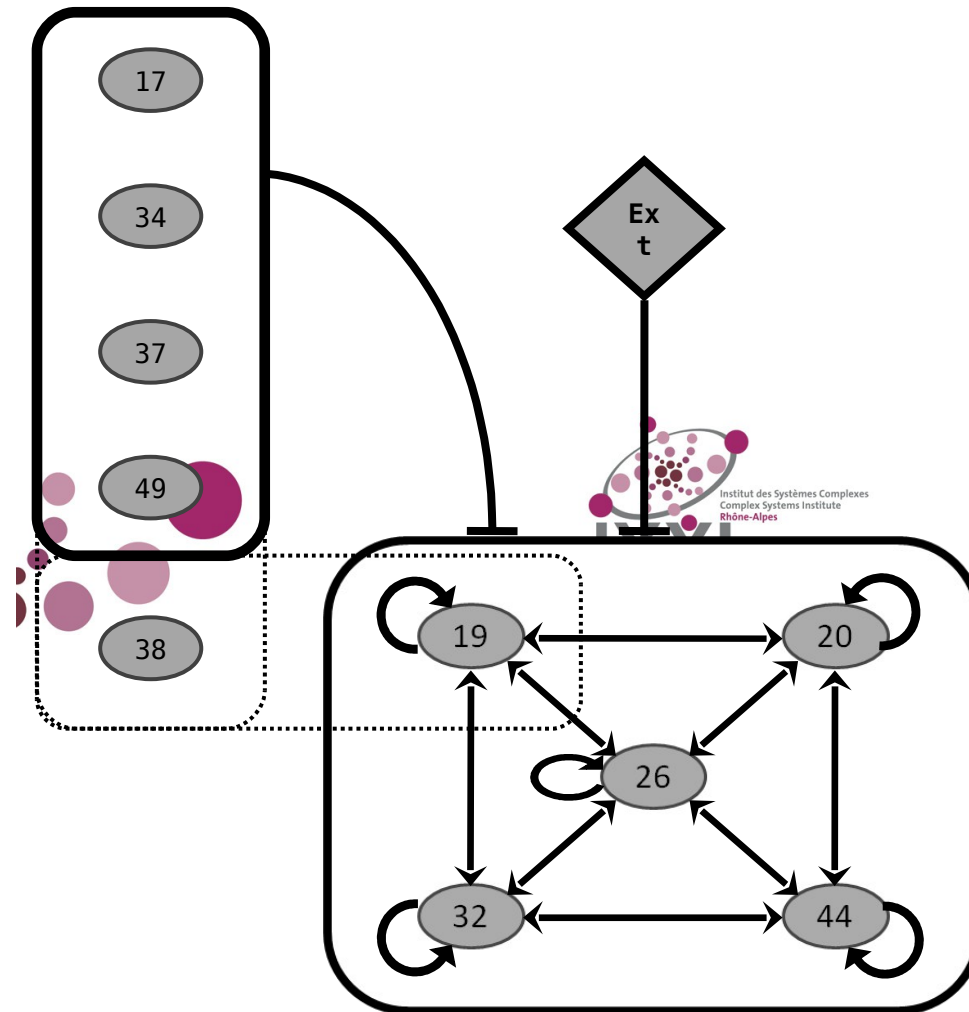


KO Gène 33

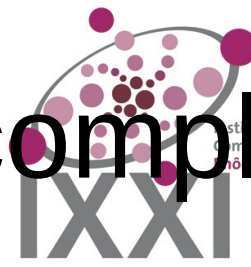


...

Reduced “schematic” network with two modules



Origin of complexity



- Where does the network complexity come from?
 - [In less stable, more changing environments, transcription factors are over-represented] ... *This suggests that in ever-changing, highly competitive environments, there is a strong selective pressure towards regulated and coordinated gene expression, compared with very stable environments. (Cases et al., 2003)*
- According to this view, the origin of (transcriptomic) complexity is another (environmental) complexity!
 - But in our experiments, the complex network emerged in a simple environment (two states)
- What about the effect of the mutation rate?
 - Similar experimental protocol as in aevol ...
 - Six mutation rates, three repetitions, 40 000 generations



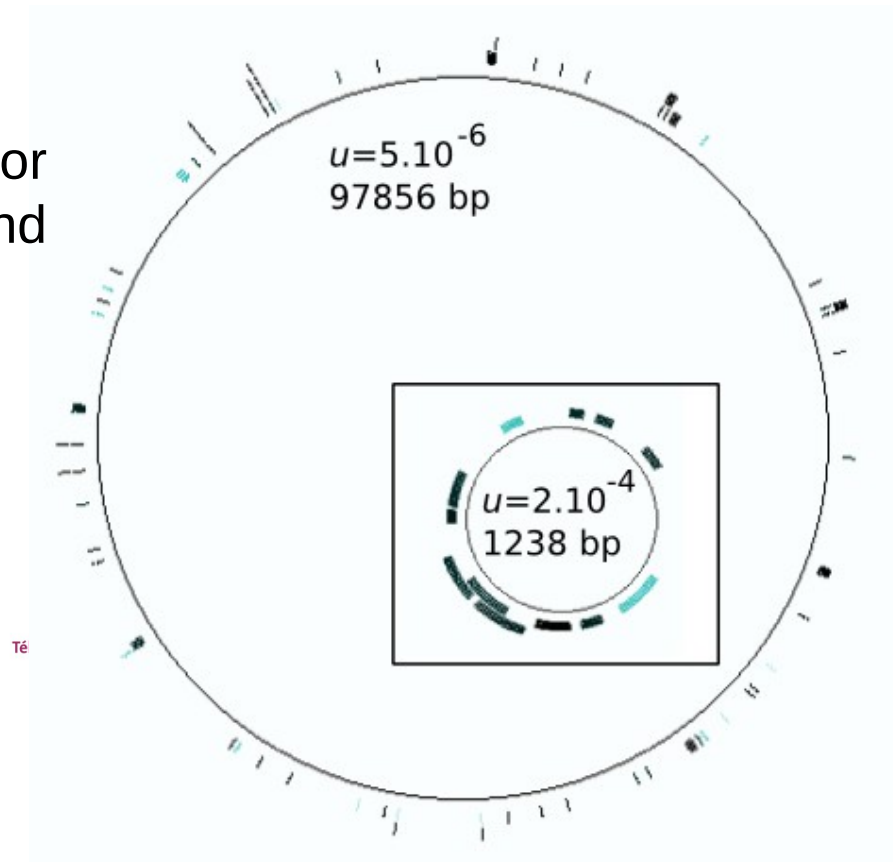
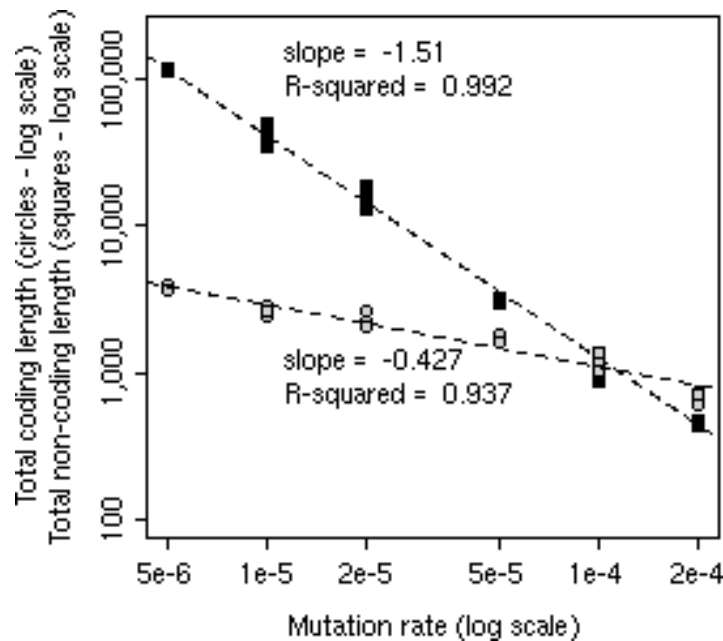
Institut des Systèmes Complexes
Complex Systems Institute
Rhône-Alpes

5 rue du Vercors 69007 Lyon France

Tél. +33 4 72 72 85 81 | Email. ixx@ens-lyon.fr | Site. www.ixxi.fr

Impact of mutation rates on genomic structures

- The mutation rate is a major determinant of genome size and gene number.

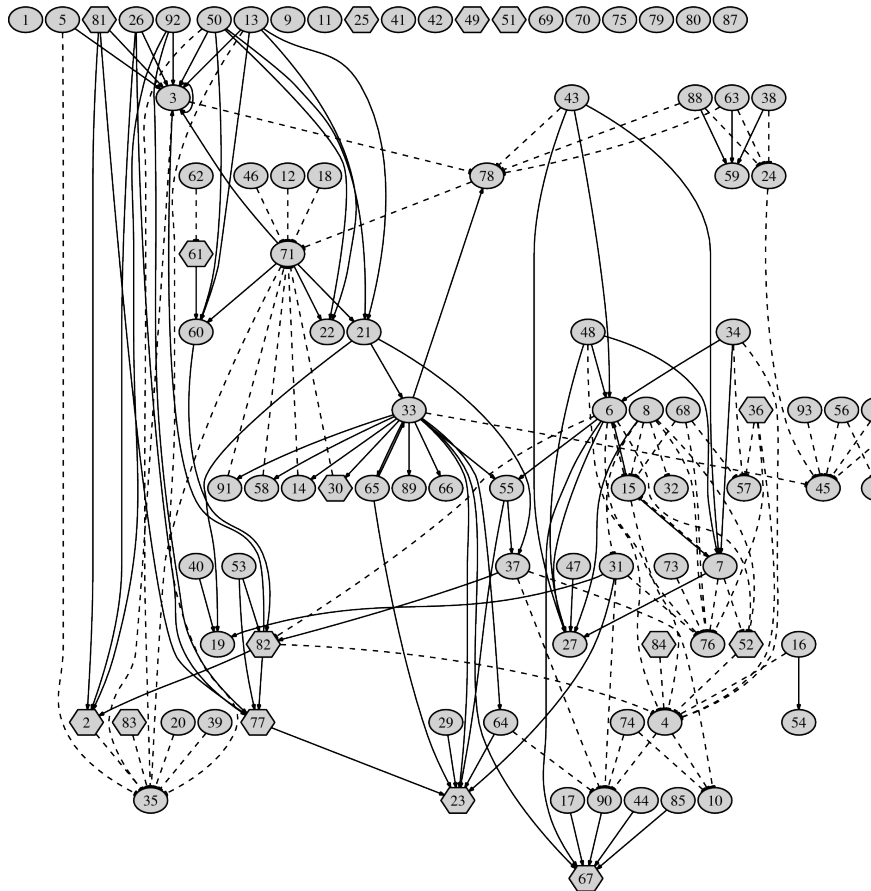


[Beslon *et al.*, IPCAT'09]

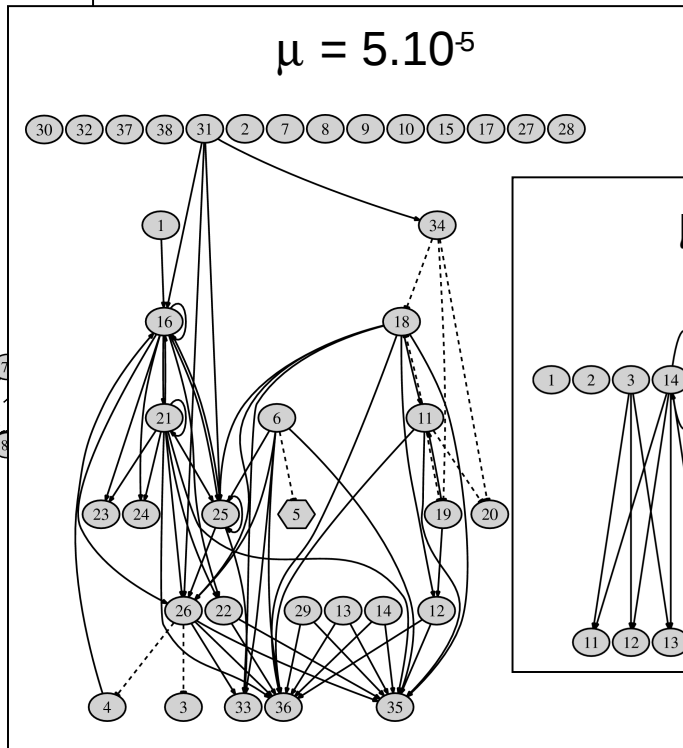
Impact of mutation rates on transcriptomic structures



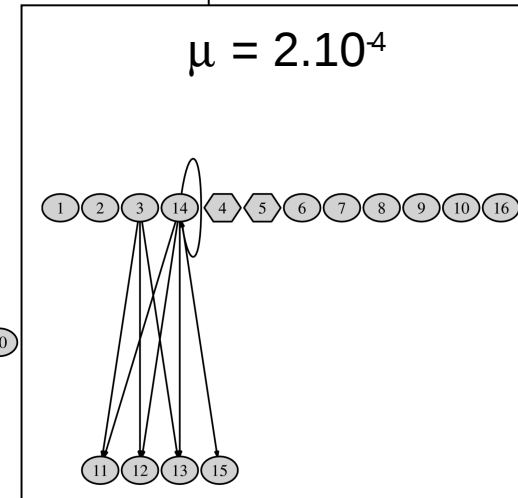
$\mu = 5.10^6$



$\mu = 5.10^5$



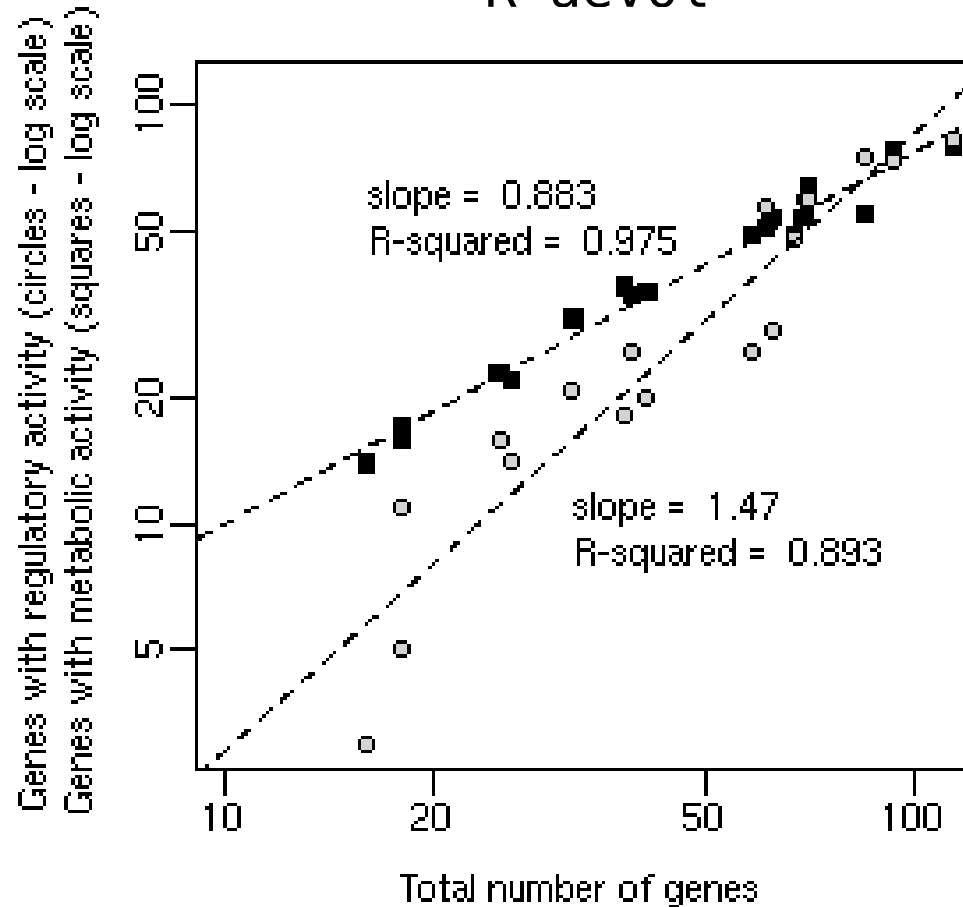
$\mu = 2.10^4$



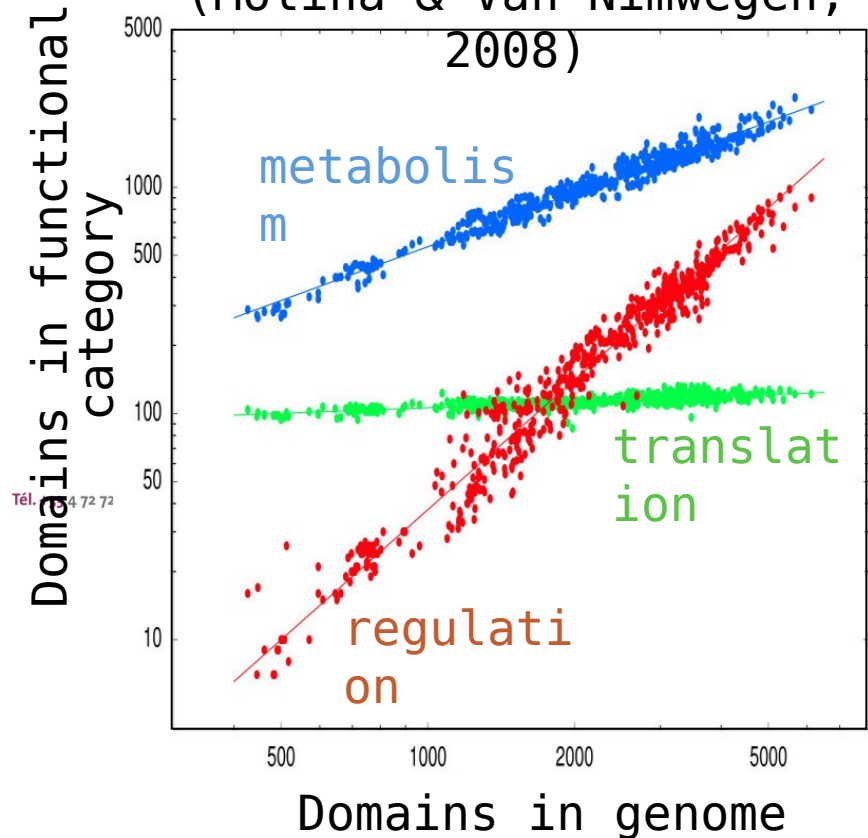
[Beslon *et al.*, IPCAT'09]
 [Beslon *et al.*, BioSystems 2010]

R-aevol: emergence of scaling laws (1)

R-aevol



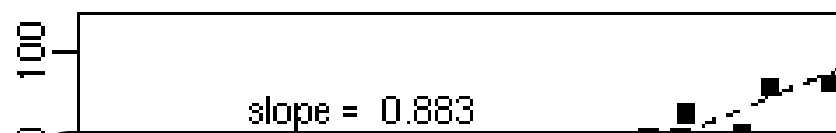
Biological data
(Molina & Van Nimwegen, 2008)



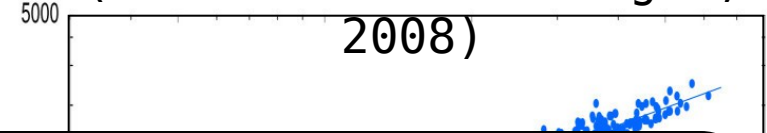
R-aevol: emergence of scaling laws (2)

R-aevol

Biological data
(Molina & Van Nimwegen, 2008)

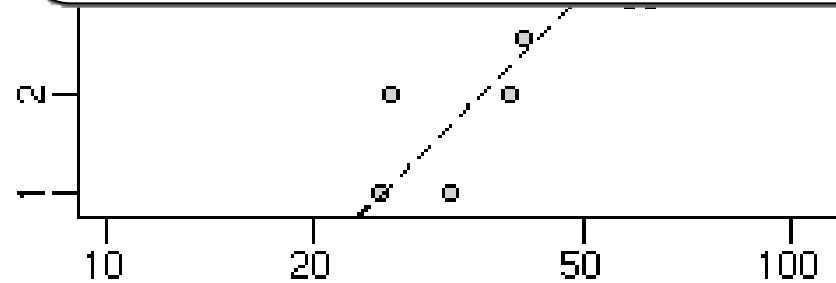


ional

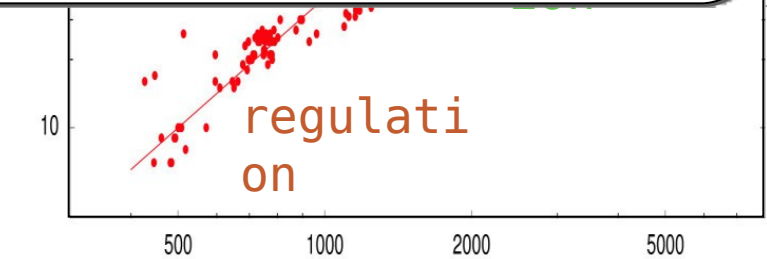


Complexity emerges “for free” (environmental complexity is NOT a necessary condition)

Can indirect selection explain this result?



Dom

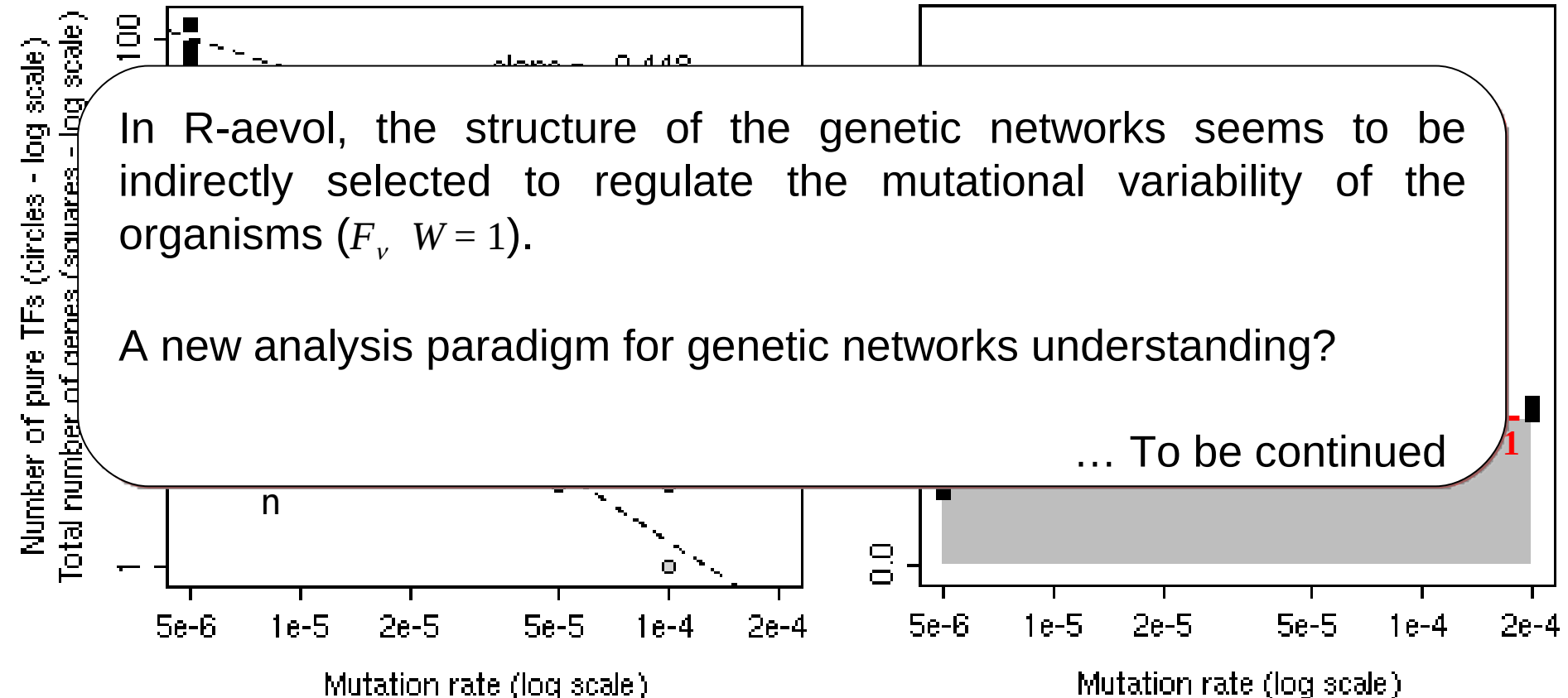


Domains in genome

R-aevol: emergence of scaling laws (3)

R-aevol

$F_v \sim \text{constant}$



Take-Home Message



Evolutionary systems biology...

- Provides essential clues to understand biological systems
 - But models are necessary

Digital genetics...

- Opens a new window on evolution
 - Enables experimental studies in evolution

Evolutionary history can explain the genomic diversity

- "Survival of the flattest"
 - Indirect selection of genetic structures (robustness/evolvability)

Complexity can emerge "for free"

- Emergence of a new paradigm in systems biology? (complexity first)
 - What is the environment of an organism?
 -



Tél. +33 4 72 72 85 81 | Email. ixx@ens-lyon.fr | Site. www.ixxi.fr



QuickTime[®] et un
décompresseur
sont requis pour visionner cette image.



Conclusion



C'est l'interdisciplinarité qui nous permet de conclure, pas le modèle

« *Dans les champs de l'observation, le hasard ne favorise que les esprits préparés* » (Louis Pasteur)

