

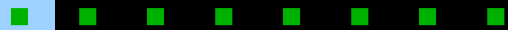


les gènes du démon de maxwell





antoine danchin
唐善・安東

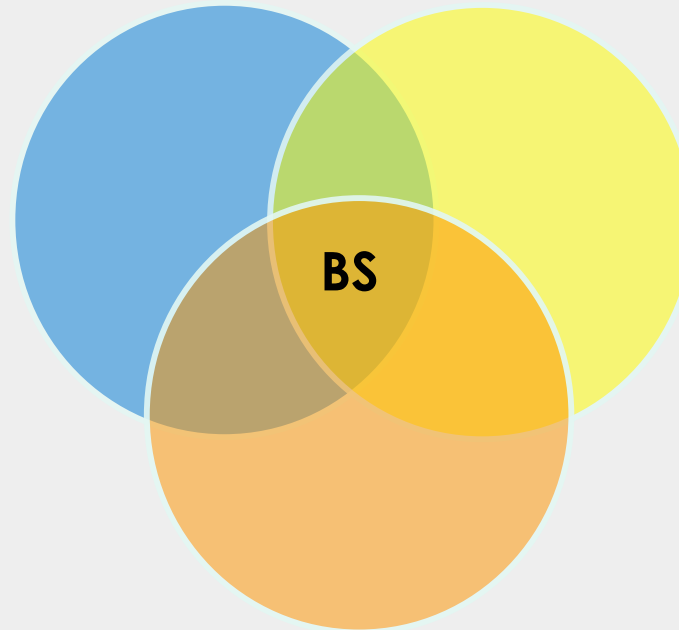


hôpital necker enfants-malades
22 mars 2010

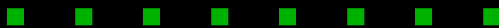
situation de la biologie synthétique

ingénierie
calcul
modélisation

biologie moléculaire
phylogénomique
biotechnologie

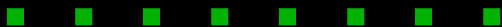


origine de la vie
vie artificielle
vie orthogonale



les objectifs de la biologie synthétique

- reconstruire la vie, et mesurer notre ignorance
- avec le point de vue de l'ingénieur, classer et normaliser des « biobriques » pour construire une « usine cellulaire »
- appliquer ce que nous avons compris à des objets de nature physico-chimique différente



le pendant pour les ordinateurs

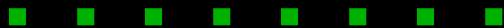
RepRap (Replicating Rapid-prototyper, 2004) vise à créer une imprimante laser 3D auto - reproductrice :

la machine produit la plupart de ses composants (= "biobriques")

ce qui manque :

- o le programme
- o la ligne d'assemblage (gestion du temps et de l'espace, et fonctions spécifiques comme la lubrification)

<http://reprap.org/>



le paradigme électronique

1. abstraction

- retenir l'essentiel

2. conception

- circuits génétiques homéostatiques ou de comportement compliqué : oscillations, multistabilité, résistance au bruit, etc.

3. standardisation

- projet "biobricks" au MIT.

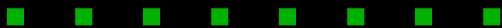
4. modélisation

- simuler la construction et analyser son comportement dynamique et sa sensibilité au bruit, aux conditions initiales etc.

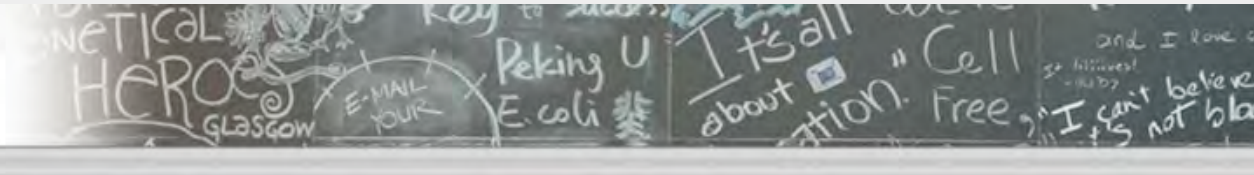
5. construction et test

- combinaison des diverses biobriques et introduction dans un "chassis". analyse du comportement de la construction et retour d'expérience. --

retour à 1

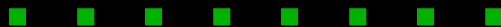
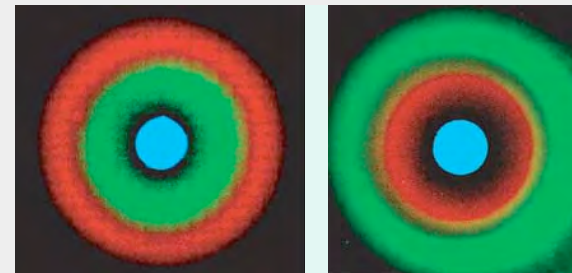


MIT : iGEM et ses « biobriques »



iGEM (International genetically engineered machines) **pose la question :**

peut-on construire des systèmes biologiques à partir d'éléments standardisés, interchangeables et mis en œuvre dans des cellules vivantes ? ou la biologie est-elle trop compliquée pour être reconstruite ainsi ?



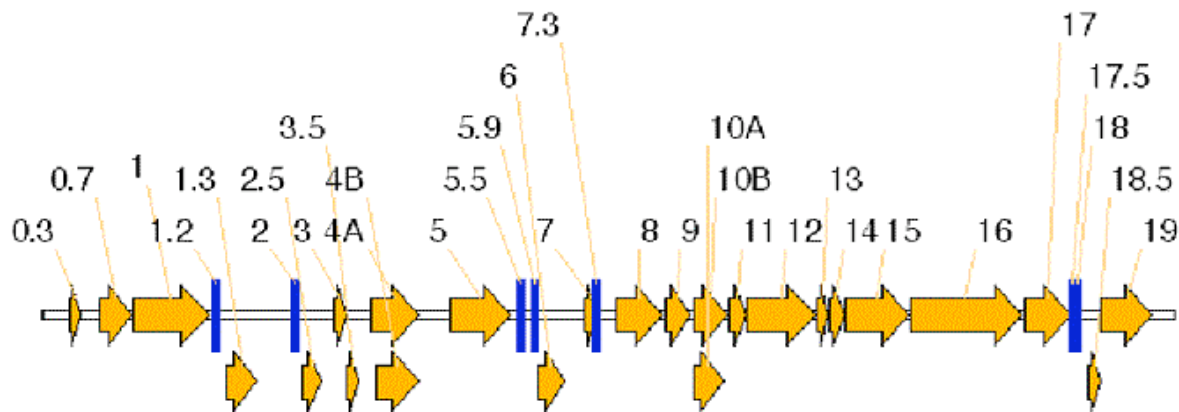
Basu S, Gerchman Y, Collins CH, Arnold FH, Weiss R.
A synthetic multicellular system for programmed pattern formation.

Nature. 2005 434:1130-1134

exemple princeps : le bactériophage T7



Les gènes connus du bactériophage T7



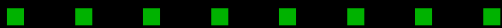
•Prise de contrôle•

- Destruction •
- Réplication •
- Genèse de la capside •
- Encapsidation •
- Préparation de la sortie de la cellule •
- Lyse •

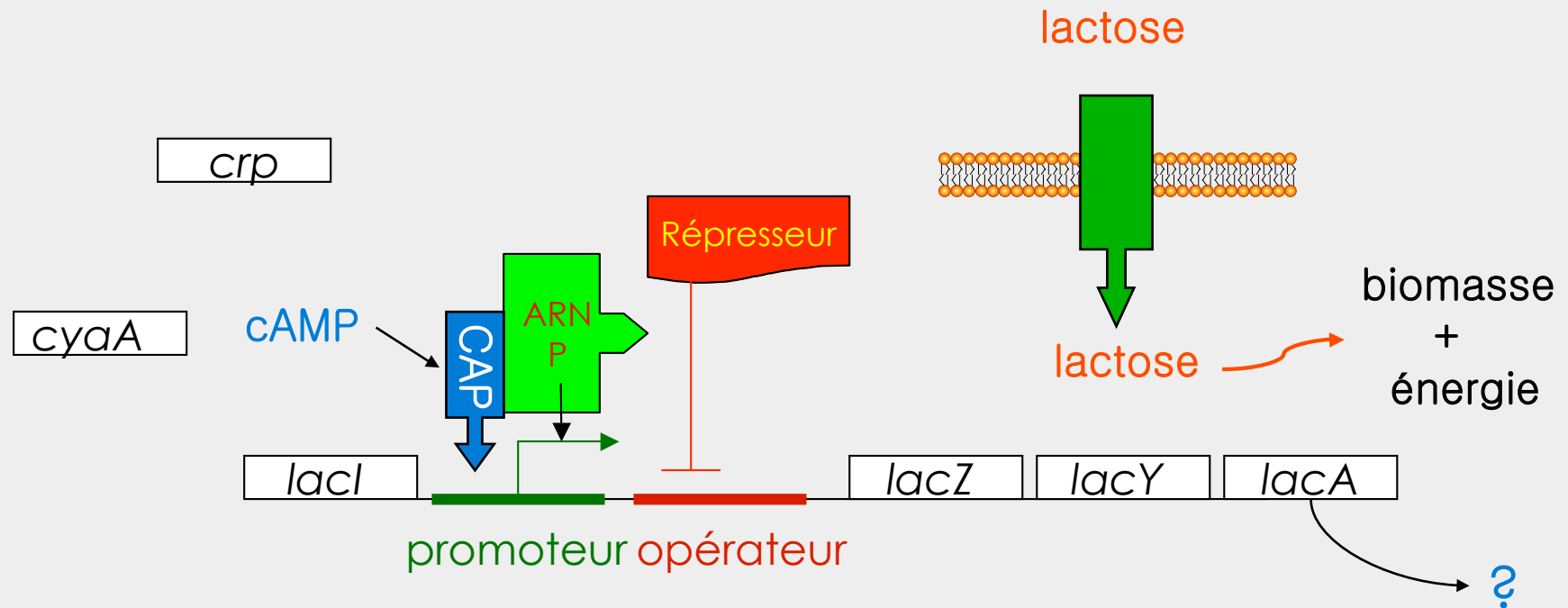
reprogrammer le bactériophage T7



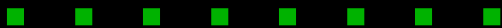
- les zones de contrôle des différentes phases sont superposées
- elles ne sont pas standardisées
- on sait les redessiner selon des principes d'ingénierie, et tester leurs propriétés par des modèles mathématiques
- les modèles permettent de prédire un comportement du phage synthétique semblable à celui du phage naturel
- le phage synthétique forme des plages de lyse, mais un peu plus petites que celles du phage naturel
- l'évolution de ce phage vers des formes spontanément virulentes efface la construction humaine



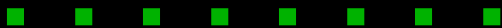
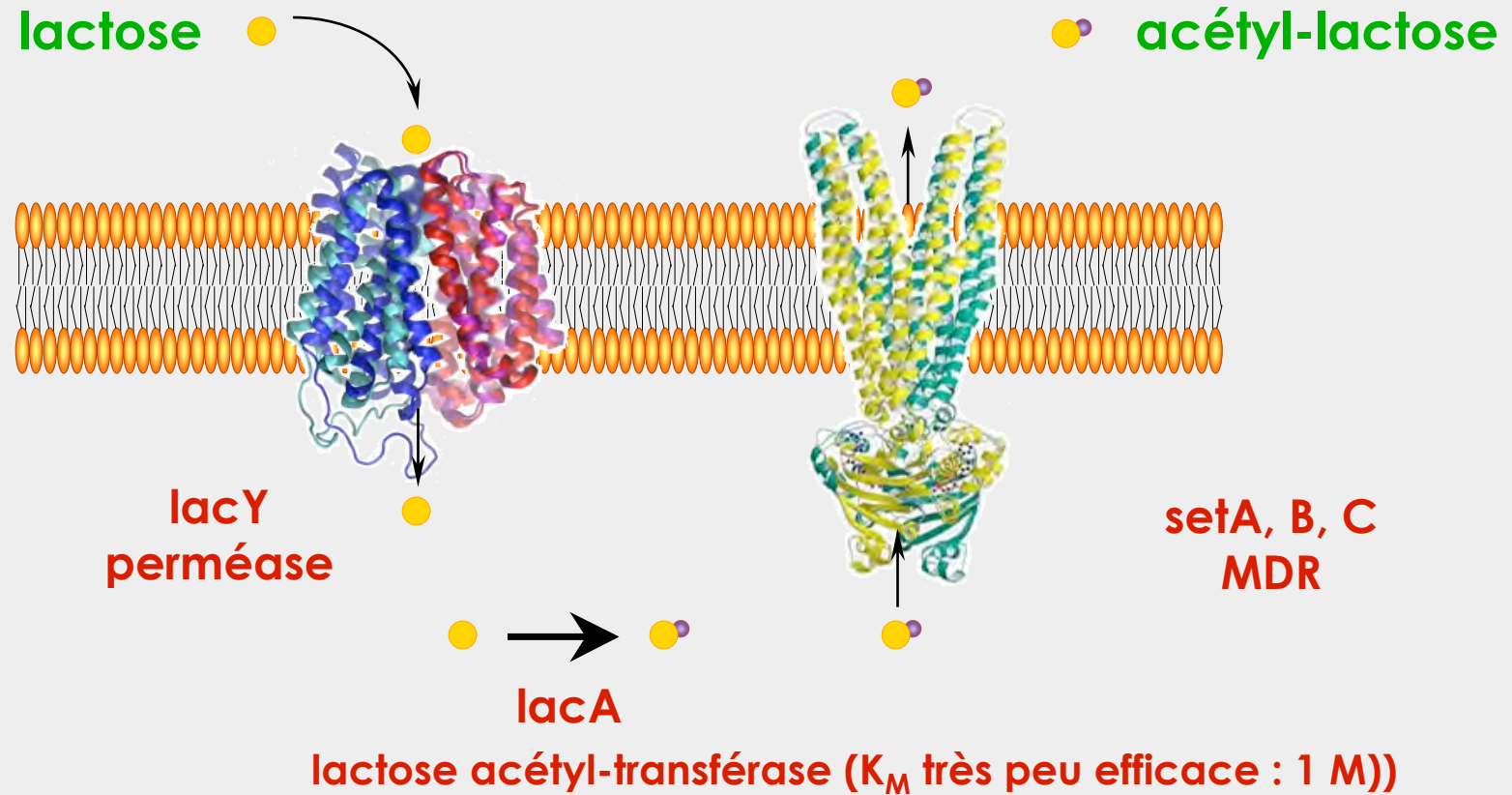
l'opéron lactose revisité



**quelle est la fonction de la lactose acétyl-transférèse ?
pourquoi a-t-il fallu 60 ans pour poser la question ?**

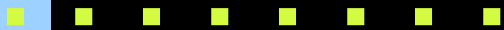


s o u p a p e d e s é c u r i t é





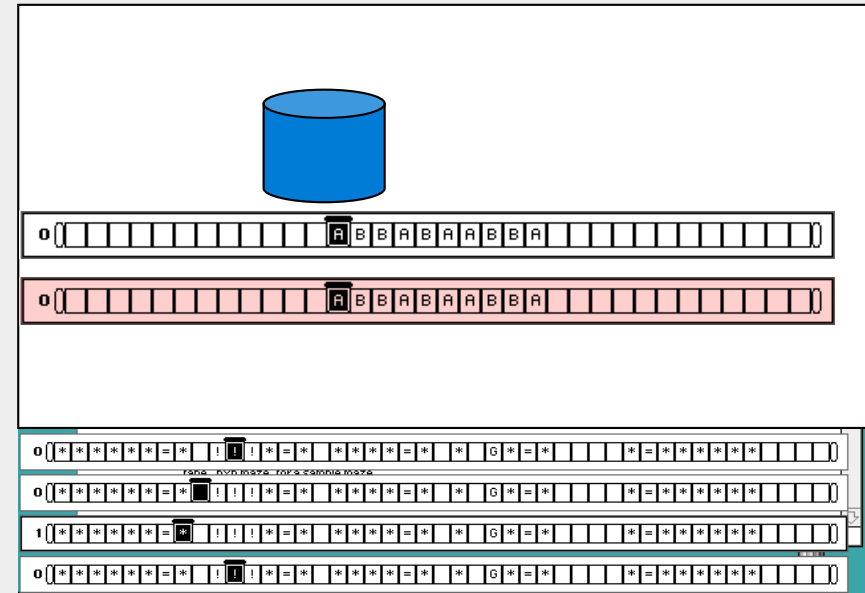
du programme à la cellule



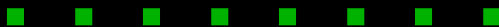
ce qu'est le calcul

deux entités :

- une machine pour lire et écrire
- un programme (sur un support physique), séparé (par l'esprit humain) en deux entités :
 - programme (définissant l'"objectif")
 - données (définissant le "contexte")



la machine est distincte de l'entité données/programme



cellules et ordinateurs

la génétique repose sur la description des génomes comme textes écrits au moyen d'un alphabet : mais **les cellules se comportent - elles comme des ordinateurs ?**

→ transfert génétique horizontal

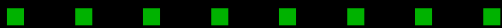
→ virus

→ génie génétique

→ **transplantation d'un génome nu dans une cellule réceptrice** suivie du changement de la machine réceptrice en une nouvelle (2007)

tout concourt à séparer

"machine" (usine cellulaire) et "données/programme" (le génome)

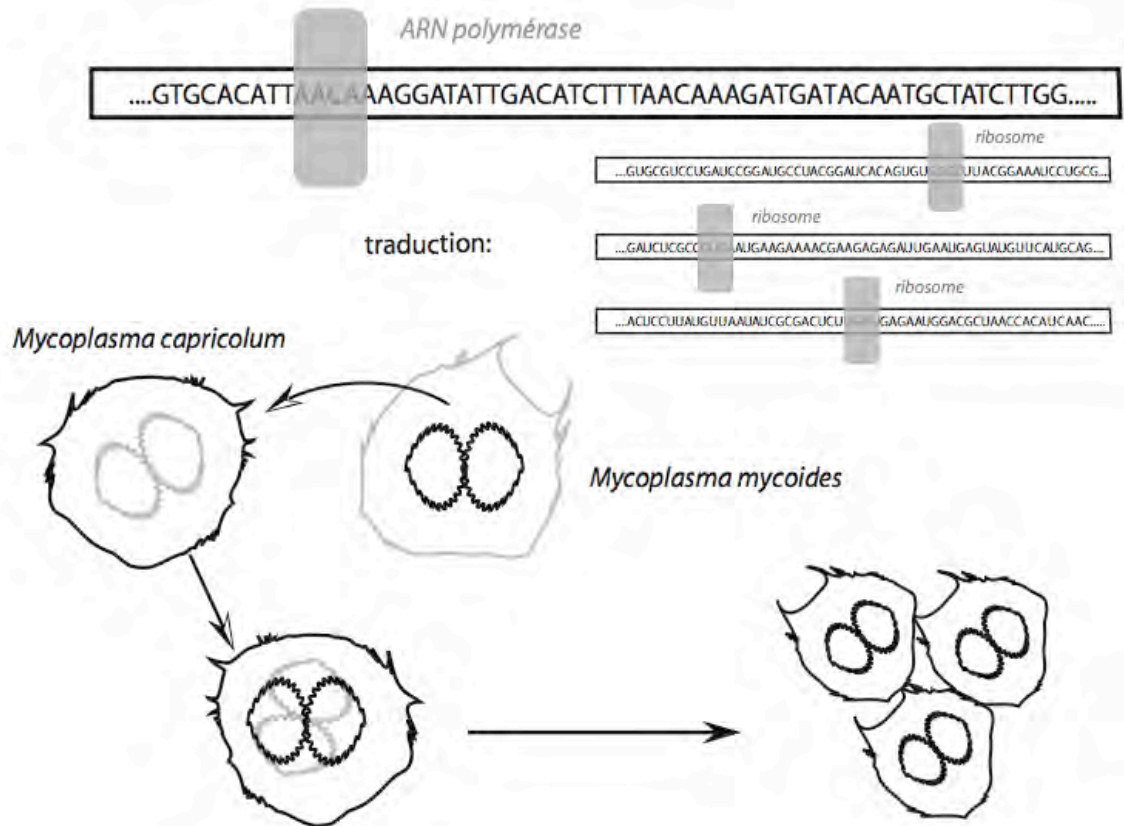


la démonstration de l'artigue-venter

la machine de turing

peut être parallélisée

transplantation d'un génome



■ ■ ■ ■ ■ ■ ■ ■

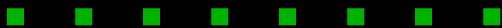
genome transplantation in bacteria: changing one species to another
Lartigue C, Glass JI, Alperovich N, Pieper R, Parmar PP, Hutchison CA 3rd, Smith HO, Venter JC
Science (2007) 317: 632-638

objection au modèle de la cellule-ordinateur

“En plus du programme génétique, la cellule comporte une quantité considérable d'information...”

Mais c'est exactement la même chose pour un ordinateur et son système d'exploitation !

Cette observation — valide — signifie que notre réflexion sur le concept d'information est incomplète



dans un ordinateur, n'oublions pas le support physique

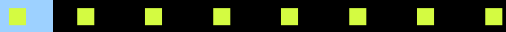


il ne suffit pas d'avoir une molécule d'ADN avec la bonne séquence, il doit encore être correctement replié !

■ ■ ■ ■ ■ ■ ■ ■ ■ ■
Gibson DG, Benders GA, Axelrod KC, Zaveri J, Algire MA, Moodie M, Montague MG, Venter JC, Smith HO, Hutchison CA 3rd. One-step assembly in yeast of 25 overlapping ADN fragments to form a complete synthetic *Mycoplasma genitalium* genome. *Proc Natl Acad Sci U S A.* (2008) 105:20404-20409

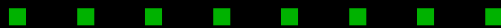
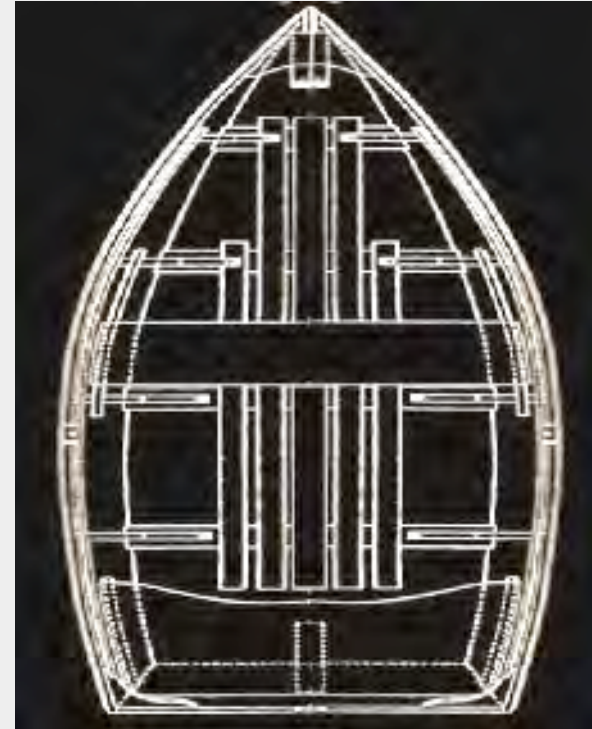


l'information



l a b a r q u e d e d e l p h e s

- la biologie est une science des relations entre objets
- elle est **symplectique** (de **συν** ensemble, **πλεκτειν**, tisser), le même mot que « complexe » en Latin ; utilisé pour éviter les connotations malheureuses associées au mot « complexité » ; une connotation géométrique n'aura pas d'inconvénient...
- c'est une information qui exprime ce qui est conservé dans la barque et non la matière de ses planches !



A. Danchin The Delphic Boat, Harvard University Press, 2003
La barque de Delphes, Odile Jacob, 1998

V. de Lorenzo, A. Danchin Synthetic Biology: discovering new worlds and new words 9: 822-827. EMBO Reports, 2008

symplectic biology

rapid research notes in systems and synthetic biology

[Home](#) [News](#) [About Us](#) [Copyright](#)

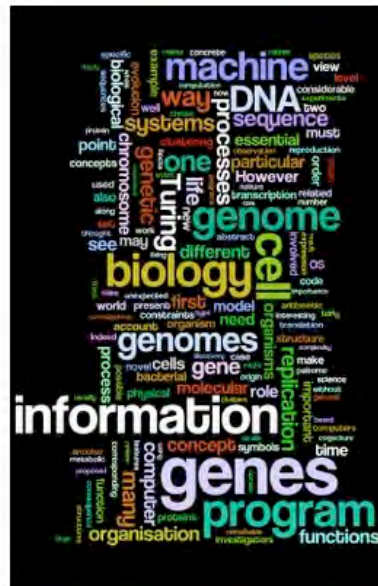
[Log in](#)

News

-  Summer Course on Synthetic Biology in Valencia, Spain
Mar 06, 2010
 -  Orthogonal translation via evolution of a quadruplet-decoding ribosome
Feb 23, 2010
 -  Biomimetic studies inspire japanese urbanists.
Jan 29, 2010
 -  Non-retroviral virus sequences in mammalian genomes.
Jan 28, 2010
- [More news...](#)

Welcome to Symplectic Biology

A peer-reviewed journal for innovative publications in Systems and Synthetic Biology



Aims and Scope

Symplectic Biology is a peer reviewed journal fostering the integration of Synthetic Biology with the more traditional Systems Biology. It aims at rapid publication of novel experiments and concepts. As its name indicates (« symplectic » is the greek equivalent of the latin « complex » without its fuzzy connotations) it endeavours to integrate physics, chemistry, information sciences and other mathematics-based disciplines into a rapidly accessible network of experiments and models permitting to combine in a challenging way the various aspects of what is traditionally named the complexity of living phenomena. Its primary aim is to promote construction of synthetic life via the quantitative characterization and understanding of biological systems at different levels of integration, ranging from the role of structure and dynamics of a single molecule to the organization and evolution of molecular and cellular networks.

Featured Articles

[Under Construction](#)

Challenges

[Under Construction](#)

Search

[Advanced Search...](#)

Editorial Board

Editors in chief

[Antoine Danchin](#)
[Victor de Lorenzo](#)

Editors

[Sigal Ben-Yehuda](#)
[Philippe Binder](#)
[Jesús Blázquez](#)
[Jason W. Chin](#)
[Gang Fang](#)
[Jiandong Huang](#)
[Jose Ignacio Jimenez](#)
[Frédérique Lisacek](#)
[Philippe Marière](#)
[Vitor Martins dos Santos](#)
[Kenta Nakai](#)
[Christos Ouzounis](#)
[Sven Panke](#)
[Manuel Porcar](#)
[Alfonso Valencia](#)
[Anil Wipat](#)
[Jeffrey Tze-Fei Wong](#)
[Huanming Yang](#)

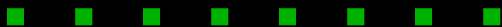
au delà des catégories usuelles

fait remarquable, la durée de l'exploration mentale varie de manière linéaire avec les distances réelles des points marqués par le sujet sur la carte, de la plage au cocotier, à la hutte, au trésor ; la carte mentale contient donc la même information sur les distances que la carte réelle

jean-pierre changeux (1983) collège de france

des systèmes génétiques reconstruits ont été utilisés pour manipuler l'information, construire des matériaux, retraiter des produits chimiques, produire de l'énergie, produire de la nourriture, et aider à conserver et améliorer la santé humaine et notre environnement

drew endy (2005) mit



explorer l'information : l'infotaxie



© P. Dubois - 11/12/2004
Saturnia pyri
Remuzat (20) - Village

comment un papillon trouve un partenaire 1 000 mètres plus loin ?

remonter un gradient chimique est impossible à une telle distance (turbulence de l'air, obstacles...)

vergassola et collègues ont montré que maximiser la collecte d'information permet d'atteindre cet objectif...

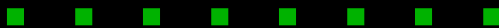
'Infotaxis' as a strategy for searching without gradients

Vergassola M, Villermaux E, S
hraman BI

Nature (2007) 445: 406-409

grand paon de nuit saturnia pyri

<http://pdubois.free.fr/>

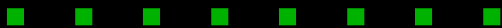


h e u r i s t i q u e n o u v e l l e

matière / énergie / espace / temps

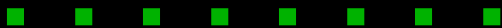
- physique classique
- physique quantique
- chimie
- biologie
 - développement
 - neurobiologie
 - linguistique
- mathématique (informatique)

information



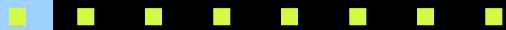
l'information comme nouvelle catégorie du réel

- 1929 **leo szilard** analyse le démon de maxwell, a l'intuition (fausse) de la relation entre énergie et information : pour lui la création d'1 bit consomme $kT \log 2$
- 1949 **claudes shannon** crée une théorie de la communication
- **1961 rolf landauer** prouve que le calcul est réversible ; aucune énergie n'est requise pour la création d'information ; elle est requise pour effacer la mémoire du processus de création
- ~1974 **andrey kolmogorov**, **gregory chaitin**, **ray solomonoff** définissent la complexité algorithmique
- 1988 **charles bennett** définit la profondeur logique (lien entre temps et complexité algorithmique) pour prendre en compte la valeur d'une information et illustre concrètement le théorème de landauer
- 1989 **wojciech zurek** lie la complexité algorithmique et l'énergie, en réfléchissant à ce que serait le démon de maxwell
- 2007 **scott muller** définit l'information comme attribut quelconque permettant de déterminer l'état d'un système, via une asymétrie
- 2009 **takahiro sagawa** et **masahito ueda** reprennent le théorème de landauer liant énergie et effacement de la mémoire





les fonctions minimales



à la recherche des fonctions ubiquistes

variation / sélection / amplification

↳ stabilisation ↻

évolution



crée (l'information apparaît ici)

fonction



capture ("recrute")

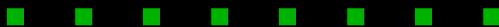
structure



code

séquence

l'ubiquité fonctionnelle n'implique pas l'ubiquité structurale

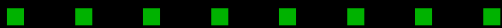


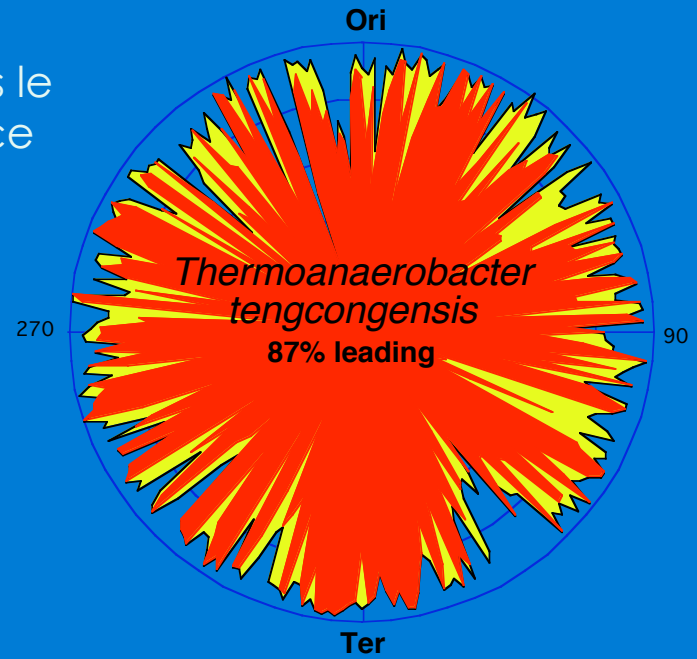
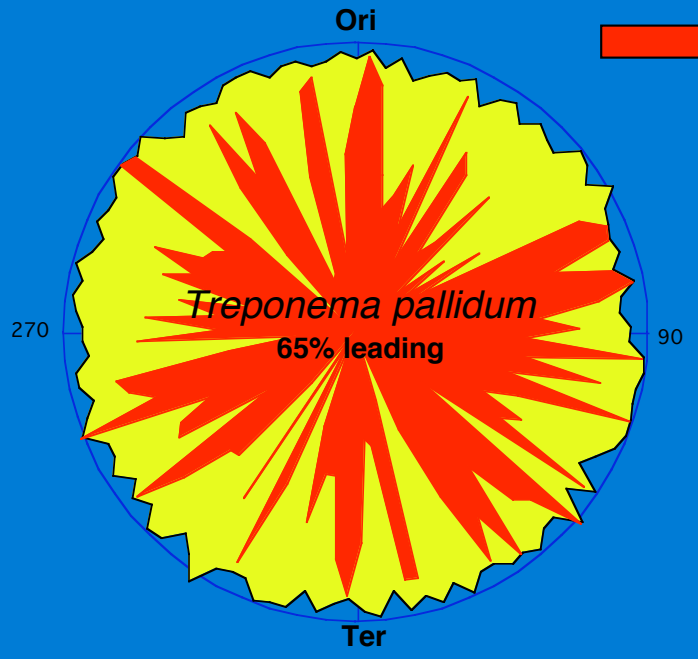
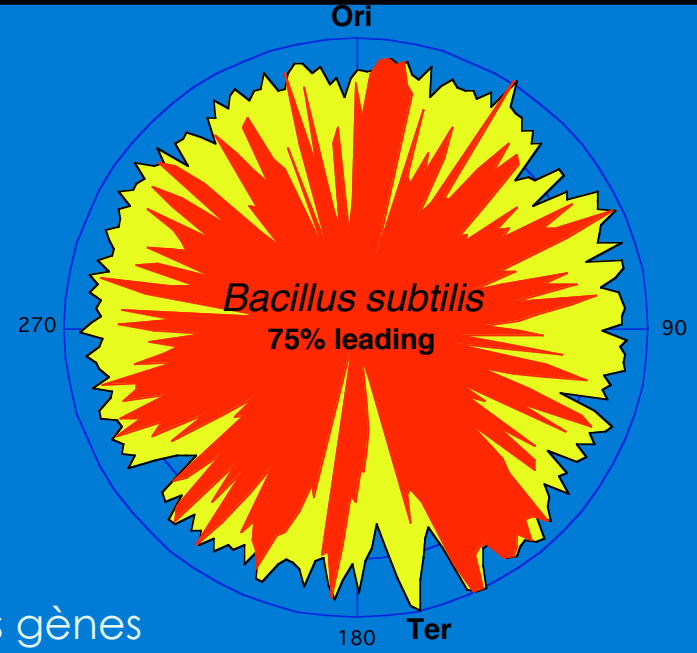
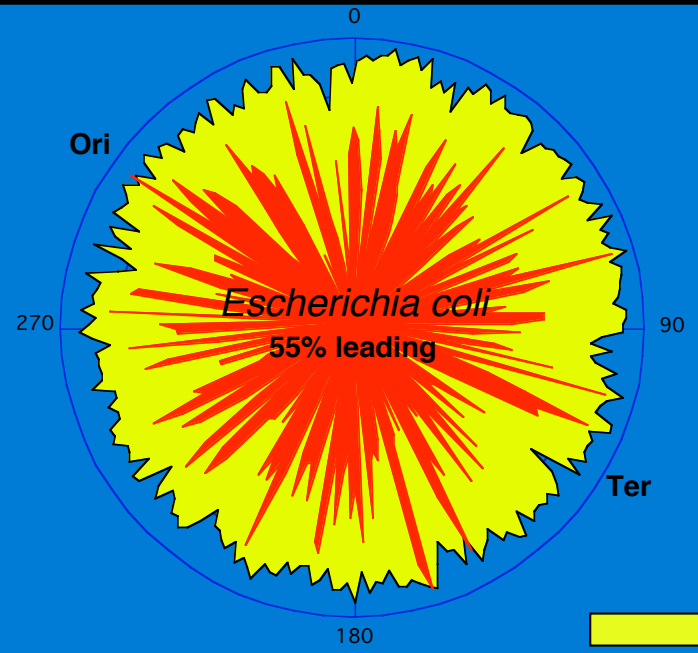
de l'ubiquité fonctionnelle à la persistance des gènes



l'ubiquité fonctionnelle n'implique pas l'ubiquité des gènes
les objets efficaces tendent à persister au cours des générations :

- rechercher la « persistance » permet d'identifier la plupart des fonctions ubiquistes
- « ubiquiste » est-il un synonyme d'« essentiel » ? Les gènes « essentiels en laboratoire » sont situés dans le brin direct de l'ADN

~ 500 gènes persistent dans les génomes bactériens ; ils sont impliqués non seulement dans les trois processus nécessaires à la vie, mais aussi dans la **maintenance** et l'**adaptation** aux **phénomènes transitoires** ; une fraction gère l'**évolution** de l'organisme

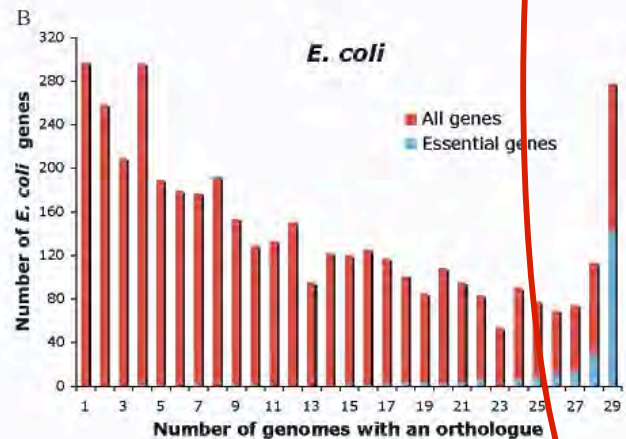
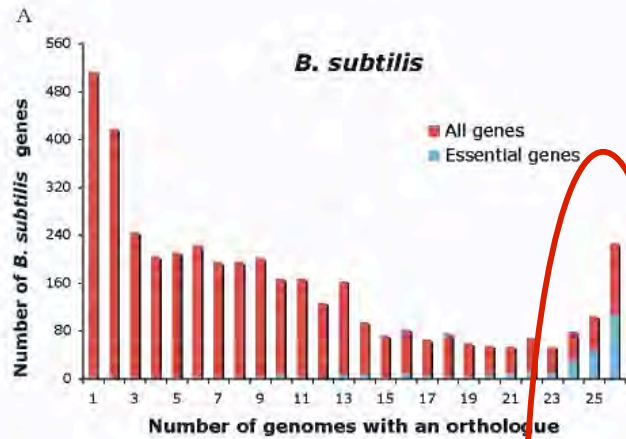




 densité des gènes
 densité des gènes dans le brin précoce



persistance : des gènes en trop !



gènes persistants

gènes essentiels et

dégradation utilisant de l'énergie

rustines métaboliques

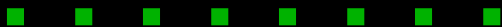
origines de la vie

freeman dyson a montré qu'il ne peut y avoir **une seule** origine à la vie

- o la **réplication** conduit presque inmanquablement à une avalanche d'erreurs
- o la **reproduction**, qui concerne des relations entre objets, peut s'améliorer au cours du temps

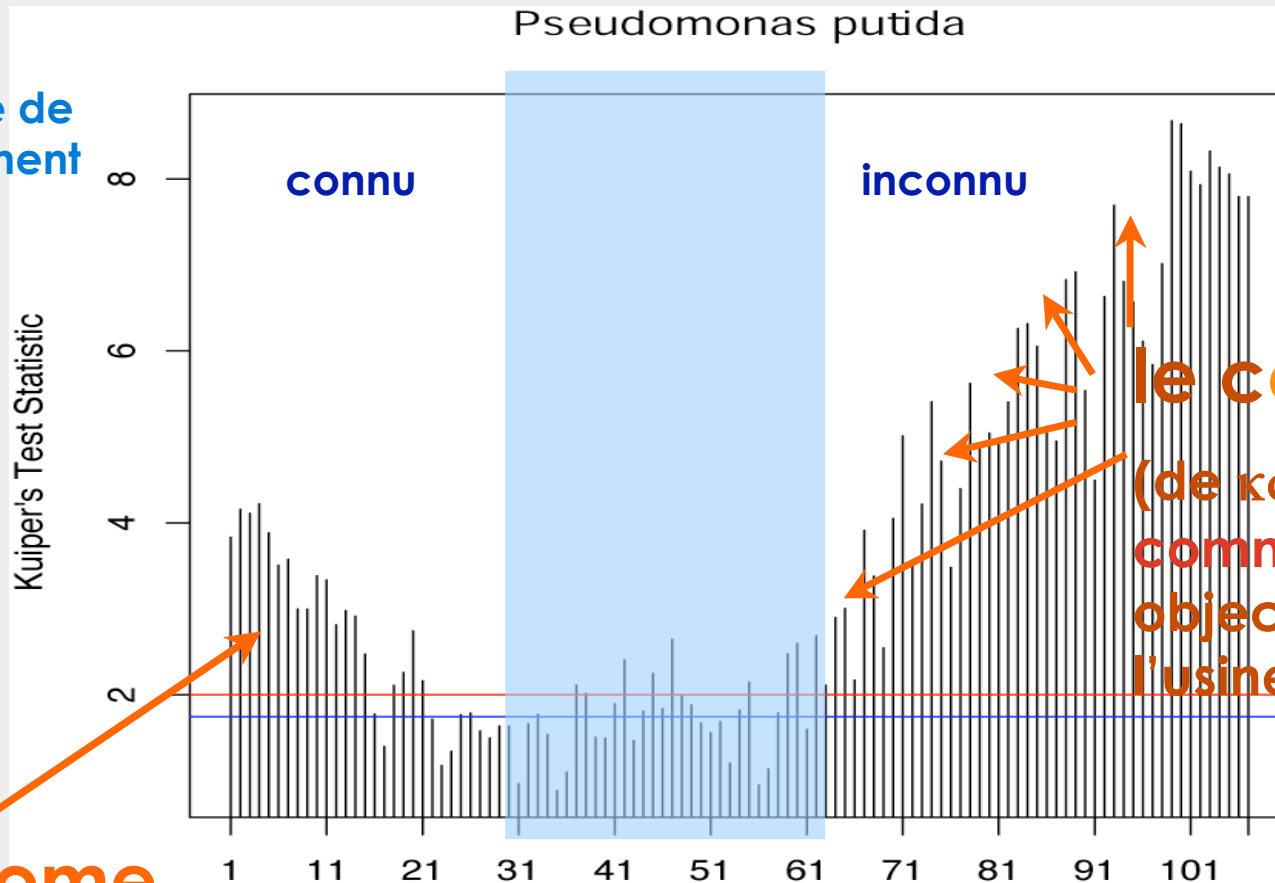
la reproduction a donc dû précéder la réplication

cela implique que le métabolisme est premier ; les gènes persistants reconstituent un scénario métabolique de l'origine de la vie



organisation des génomes bactériens

fréquence de regroupement



le cénome

(de κοινος, commun): l'objectif de l'usine cellulaire

fréquence dans les génomes

le paléome

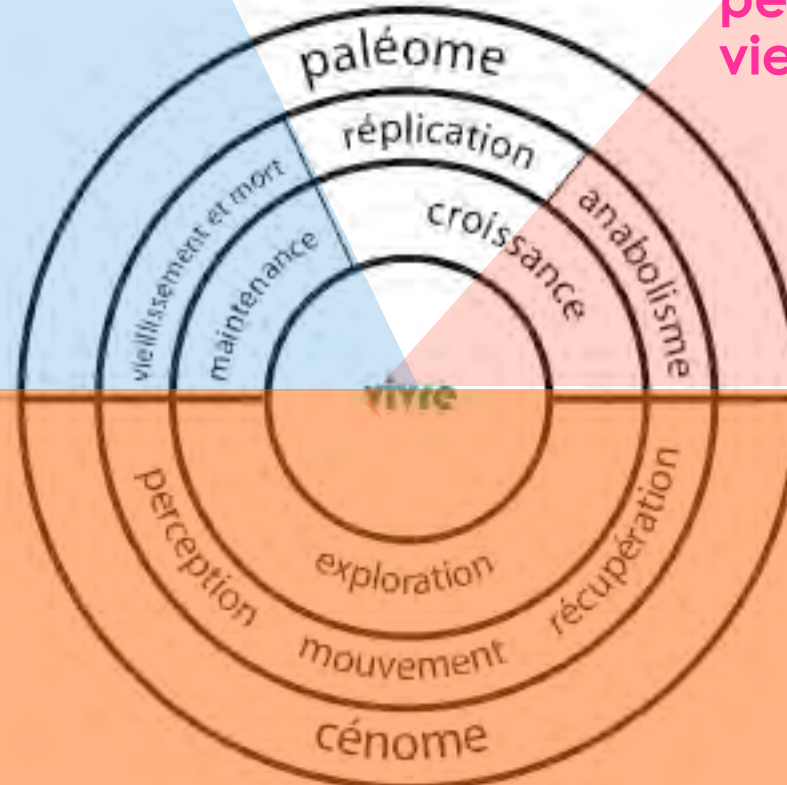
(de παλαιος, ancien): l'usine cellulaire

histoire de deux génomes

métagénomique

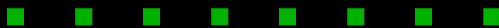
survivre

perpétuer la vie



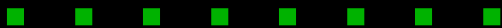
vivre en contexte

$\Sigma = \{\text{pan-génome}\}$



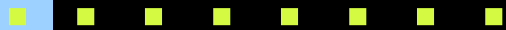
le paléome et le cénome

- la structure du paléome
 - fonctions essentielles ; évolution de la machinerie de l'expression des gènes
 - dégradation utilisant de l'énergie
 - métabolisme du soufre (anabolisme, récupération, catabolisme)
 - « frustration » chimique (« rustines » métaboliques)
- le cénome : occupation d'une niche écologique





les gènes du démon de maxwell



les bébés naissent très jeunes !

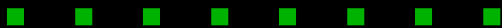
- **la machine se reproduit**

- la reproduction peut s'améliorer au cours du temps : c'est toujours un organisme âgé qui donne naissance à un jeune (cela implique **création d'information**)

- **le programme se réplique**

- la réplication accumule progressivement des erreurs

quels gènes permettent l'accumulation d'information ?

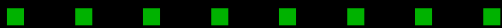


revisiter l'information

l'intuition dit que la création d'information demande de l'énergie

faux : la **création d'information est réversible** (landauer, 1961 ; bennett, 1982, 1988, zurek, 1989); accumuler l'information requiert un processus **énergivore pour "faire de la place"**

question ouverte : "faire de la place" est nécessaire pour accumuler l'information ; comment est-ce réalisé ? pouvons-nous identifier dans les génomes les gènes codant les fonctions qui mettent en œuvre ce processus ? pouvons-nous trouver une source ubiquiste et stable d'énergie ?

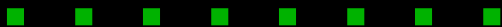


valeur de l'information

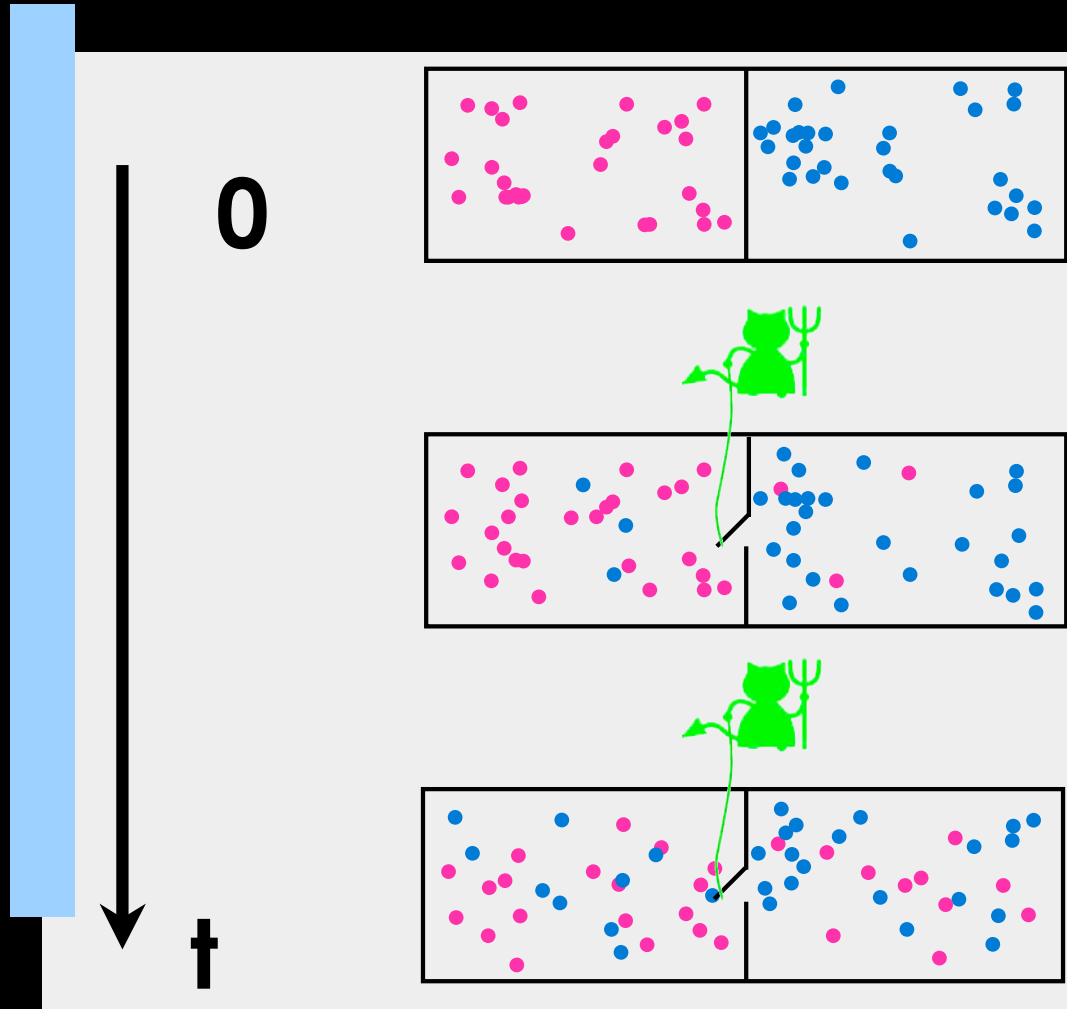
dans les modèles classiques de l'information on ne se soucie pas de la signification, ou de la **valeur** de l'information

l'information du programme est transmise telle quelle par la réplication, sans valeur, mais d'où vient donc l'information de la machine ? L'idée est qu'il faut reconsidérer le concept dès qu'on a affaire à des objets individualisés :

peut-on **imaginer les gènes d'un démon de maxwell** qui ferait le tri entre ce qui est fonctionnel (localement) et ce qui ne l'est pas ?

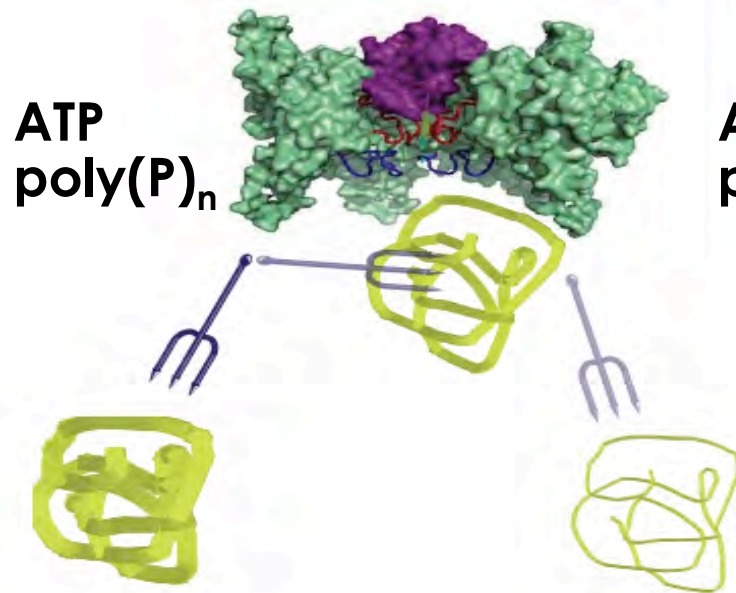


le démon de maxwell



le démon peut accumuler de l'information ou renverser le temps s'il peut mesurer la vitesse et la position des atomes de gaz ; récoltant une information pour calculer quand il doit fermer la trappe

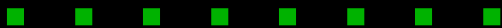
les gènes du démon de maxwell



la machinerie de dégradation
utilise de l'énergie pour rejeter
intacte une entité fonctionnelle

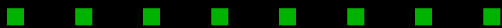


les entités non fonctionnelles
sont reconnues et dégradées



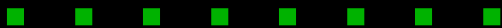
i n f o r m a t i o n

- les processus dégradatifs dépendant de l'énergie font de la place pour des entités nouvellement synthétisées ; de l'énergie est consommée pour **prévenir** la dégradation des entités fonctionnelles
- ce processus accumule de l'information, quelle que soit son origine, par effet de cliquet
- comme ce processus est ubiquiste, on attend que les fonctions correspondantes soient codées dans le paléome, y compris la gestion des sources d'énergie postulées ici



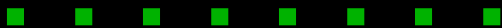
une ressource ubiquiste : les polyphosphates

- la synthèse et la dégradation des poly-P est codée dans les gènes non-essentiels persistants ; ce processus peu connu est associé à la dégradation de l'ARN
- le poly-P est un **minéral**, donc extrêmement stable ; il est présent de façon ubiquiste dans les cellules
- les NTPs sont régénérés à partir des NMP et de poly-P ; la protéase Lon peut utiliser le poly-P au lieu d'ATP ; le NADP (anabolisme) peut être engendré à partir de NAD et de poly-P...



innovation : les mutations adaptatives

- l'accumulation d'information dépendant de l'énergie est aveugle ; elle ignore d'où l'information proviendra
- l'information peut provenir d'une mémoire, celle du génome pré-existant ; elle peut aussi être créée de novo
- les **mutations adaptatives** sont des créations de novo d'information ; elles dépendent donc des gènes impliqués dans l'accumulation d'information



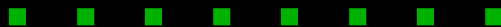
mutations adaptatives

construction de bactéries "intelligentes"

mises en croissance sur un milieu limitant en éléments nutritifs, elles forment des colonies de 10^7 bactéries environ ; le milieu contient aussi des nutriments **qu'elles ne savent pas utiliser**.

au bout de quelques semaines apparaissent des papilles qui se mettent à croître et qui utilisent les éléments ajoutés ; ce sont des **mutations adaptatives**.

elles ne préexistaient pas et supposent donc la **création d'information**.



des mutations apparaissent au cours du vieillissement



séquençage de 14 génomes + 30 PCRs

le nombre total des mutations croît avec l'âge des colonies

les mutations sont réparties dans tout le chromosome, et concentrées dans un gène => PCR de plusieurs colonies

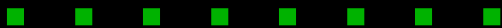
dans ce gène on trouve différentes mutations dans les différentes papilles, 2 mutations in 30% des cas

dans certains cas une des deux mutations est silencieuse

avec une source de carbone particulière, c'est au moins un autre gène qui est impliqué

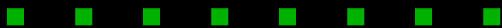
vivre et perpétuer la vie

- le paléome comprend à la fois les gènes permettant de construire l'usine cellulaire, et les gènes permettant de contrecarrer le vieillissement par l'accumulation d'une information nouvelle dans la descendance
- les processus dégradatifs dépendant de l'énergie font de la place pour des entités nouvellement synthétisées ; de l'énergie est consommée pour **prévenir** la dégradation des entités fonctionnelles
- ce processus accumule de l'information, quelle que soit son origine, par effet de cliquet
- ce processus est **myope**: il ne peut avoir un projet, d'où l'aspect "bricolé" de la vie et de son évolution



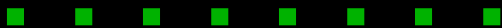
prédictions diverses

- la persistance bactérienne dans un hôte dépend de gènes persistants non-essentiels
- l'initiation du cancer provient de cellules (souches) qui ont découvert des mutations adaptives leur permettant de créer une descendance immortelle
- l'accumulation d'information dans le cerveau (mémoire et apprentissage) implique des processus pour faire de la place tout en préservant les connexions fonctionnelles, d'une façon qui doit dépendre de l'énergie



une cellule synthétique ?

- la vue d'ingénieur de la BS exclut que les cellules artificielles soient innovantes
- il est possible d'**exclure les gènes permettant l'accumulation d'information**
- la conséquence est que, comme les usines, l'usine cellulaire vieillira et devra être systématiquement reconstruite
- cela a l'avantage social que les risques sont minimisés
- en revanche cela pose problème pour l'utilisation à grande échelle de cellules synthétiques



c o n t r i b u t i o n s

in silico

gang fang, eduardo rocha, tingzhang wang

in vivo

agnieszka sekowska, evelyne turlin, andrew martens

collaborations

genoscope, beijing genome institute, fudan university, the university of hong kong, hong kong university of science and technology

