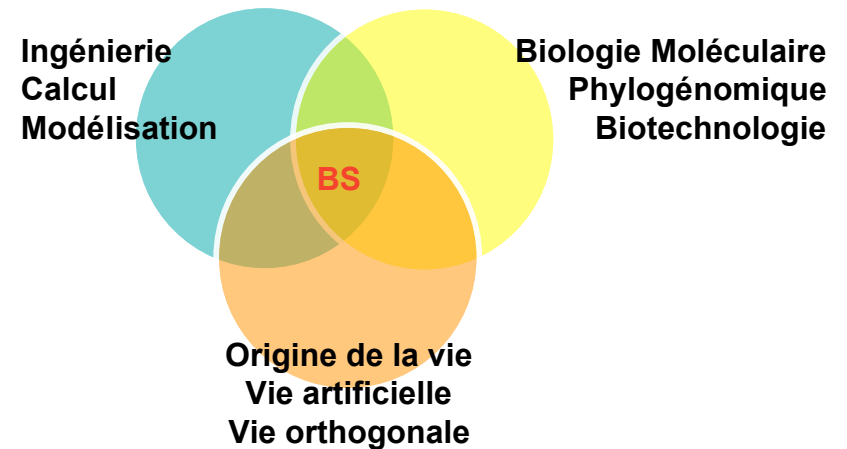


L'usine cellulaire synthétique : réalité ou fiction ?

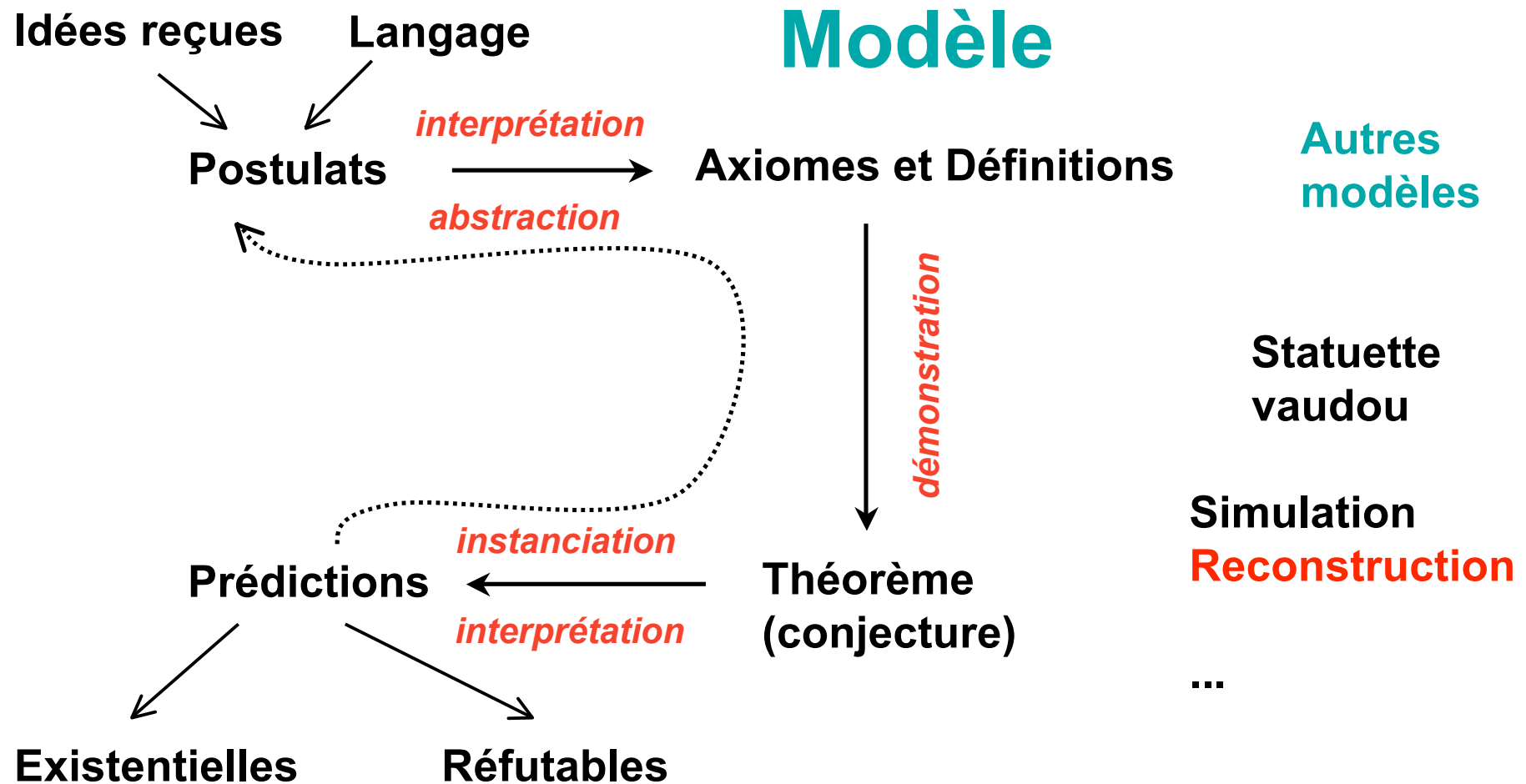


Antoine Danchin

唐善 安東

Ecole des Mines de Paris 15 juin 2009

Ce qu'est la Science: la Méthode Critique Générative



- - - - **BIOLOGIE SYNTHÉTIQUE** - - - -

Objectifs de la Biologie Synthétique

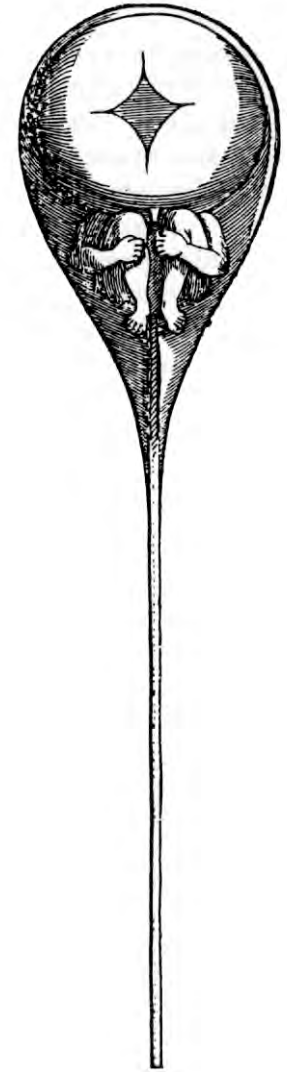
- ➔ Un premier objectif de la SB est de reconstruire la vie, dans un effort pour explorer si nous comprenons ce qu'elle est, en découvrant ce qui manque à partir de nos échecs
- ➔ Dans un second effort, nous conservons ce que nous avons compris pour l'appliquer à des objets d'une nature physico-chimique différente
- ➔ Un troisième objectif voit la vie du point de vue de l'ingénieur, et cherche à classer et normaliser des « biobriques » pour construire une « usine cellulaire »

“Les organismes vivants comme pièges à information”

<http://www.normalesup.org/~adanchin>

Petit bagage

- Ce n'est pas un organisme qui est transmis au cours des générations, mais une recette pour le construire et lui permettre de vivre
- La machine qui met en œuvre la recette est physiquement séparée du livre de recettes
- Ce processus est formellement semblable à celui qui se déroule dans les ordinateurs
- Mais les ordinateurs ne font pas (encore) des ordinateurs : que faut-il de plus ?
- L'étude comparative du programme génétique nous permet de découvrir les gènes nécessaires à cette prouesse : vivre se distingue de perpétuer la vie





Des ordinateurs faisant des ordinateurs?

“Les organismes vivants comme pièges à information”

<http://www.normalesup.org/~adanchin>

Une imprimante 3D auto-reproductrice

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

- La machine produit la plupart de ses composants (= “biobriques”)
- Ce qui manque :
 - Le program
 - La ligne d’assemblage (gestion du temps et de l’espace, et fonctions spécifiques comme la lubrification)

<http://reprap.org/>



“Les organismes vivants comme pièges à information”

<http://www.normalesup.org/~adanchin>

In vivo, in vitro, in silico

- Biologie

- Biochimie

- Génétique

➔ Biologie Moléculaire

The singular fate of genetique in the history of French biology, 1900-1940.

Burian RM, Gayon J, Zallen D J Hist Biol (1988) **21**: 357-402

- Théorie des Nombres et Informatique

- Machines de Turing

- Algorithmique

➔ Théorie de l'Information

“Les organismes vivants comme pièges à information”

<http://www.normalesup.org/~adanchin>

Ce qu'est la vie

La vie demande :

- ➔ Une machine permettant à un programme de s'exprimer (qui se reproduit)
 - ➔ 1. Métabolisme (processus dynamique)
 - ➔ 2. Compartimentation (définissant un intérieur et un extérieur)
- ➔ Un programme (“livre de recettes”, qui se réplique)
 - ➔ 3. Transfert récursif d'information => codage d'un premier niveau vers un second comme caractéristique essentielle

La cellule est l'atome de vie

“Les organismes vivants comme pièges à information”

<http://www.normalesup.org/~adanchin>

Le métabolisme

La vie est en perpétuel devenir « chimique »

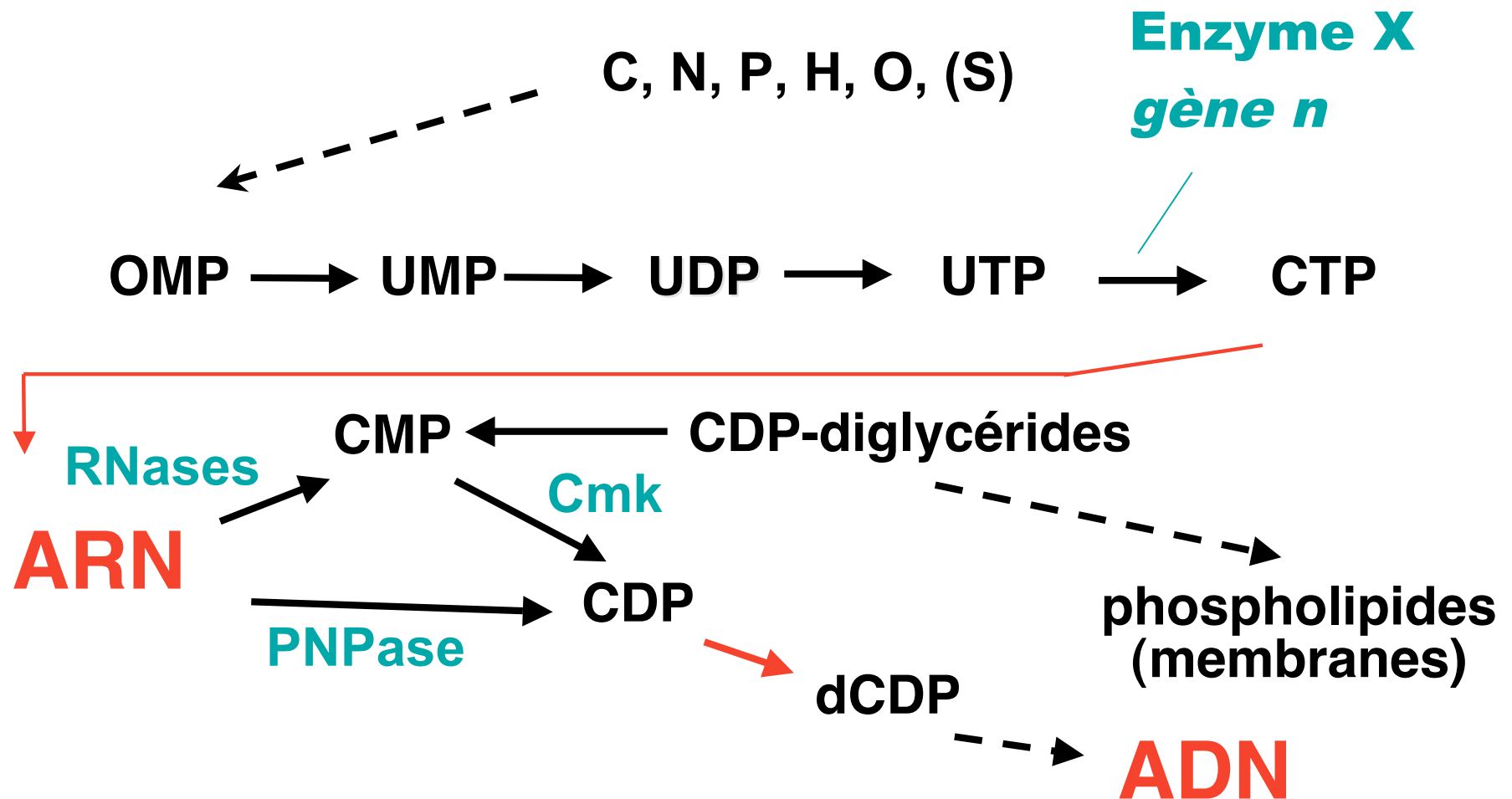
- Environ 800 molécules interconvertibles sont nécessaires à toute forme de vie (“anabolisme et catabolisme”)
- Assimilation de C, N, O, S, P et d'ions
- Gestion de l'énergie (ATP et transferts d'électrons)

Les coenzymes (souvent connus sous le nom de ‘vitamines’) sont des composés essentiels du métabolisme, mais ils sont généralement oubliés dans les scénarios de l'origine de la vie !

“Les organismes vivants comme pièges à information”

<http://www.normalesup.org/~adanchin>

Une voie métabolique



“Les organismes vivants comme pièges à information”

<http://www.normalesup.org/~adanchin>

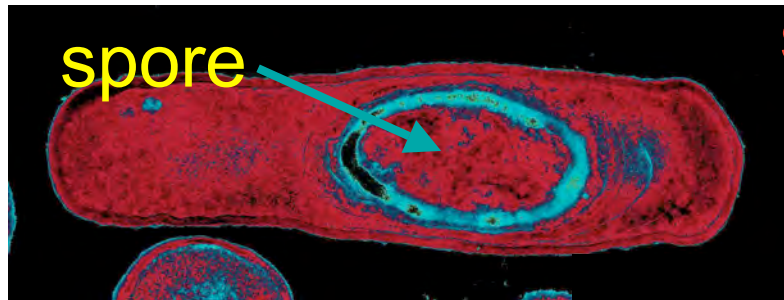
Compartimentation

Deux stratégies principales distinguent les organismes

→ Une seule enveloppe (plus ou moins compliquée) : les **procaryotes (Archaea et Bacteria)**

→ Multiplication des membranes et des peaux : les **eucaryotes (Eukarya)**

Même chez les bactéries....



Bacillus subtilis



Stigmatella aurantiaca

Transfert d' Information

Comme pour la construction d'une machine, il faut un « manuel », une recette, pour construire une cellule

Il faut pour cela changer le texte de la recette en quelque chose de concret : ce processus transfère de l'« information »

Dans la cellule, le transfert d'information est organisé par le **programme génétique**

“Les organismes vivants comme pièges à information”

<http://www.normalesup.org/~adanchin>

Règles de réécriture

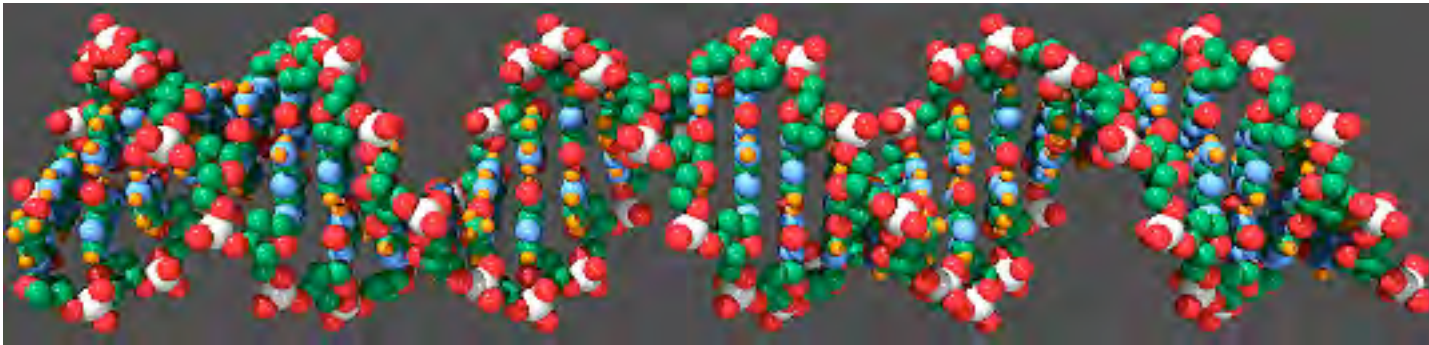
- **Réplication (ADN $\rightarrow$ ADN) loi:**
“complémentarité”
- **Transcription (ADN $\rightarrow$ ARN) loi:**
“complémentarité”
- **Traduction (ARN $\rightarrow$ protéine) loi:**
“code génétique”

“Les organismes vivants comme pièges à information”

<http://www.normalesup.org/~adanchin>

Mémoire

L'Acide Désoxyribonucléique (ADN) conserve la mémoire du programme génétique. Il est fait du chaînage de quatre types de maillons élémentaires. Deux chaînes **complémentaires orientées** se font face, formant une double hélice.



A => T

C => G

T => A

G => C



HKU - PASTEUR RESEARCH CENTRE

香港大學—巴斯德研究中心

=>

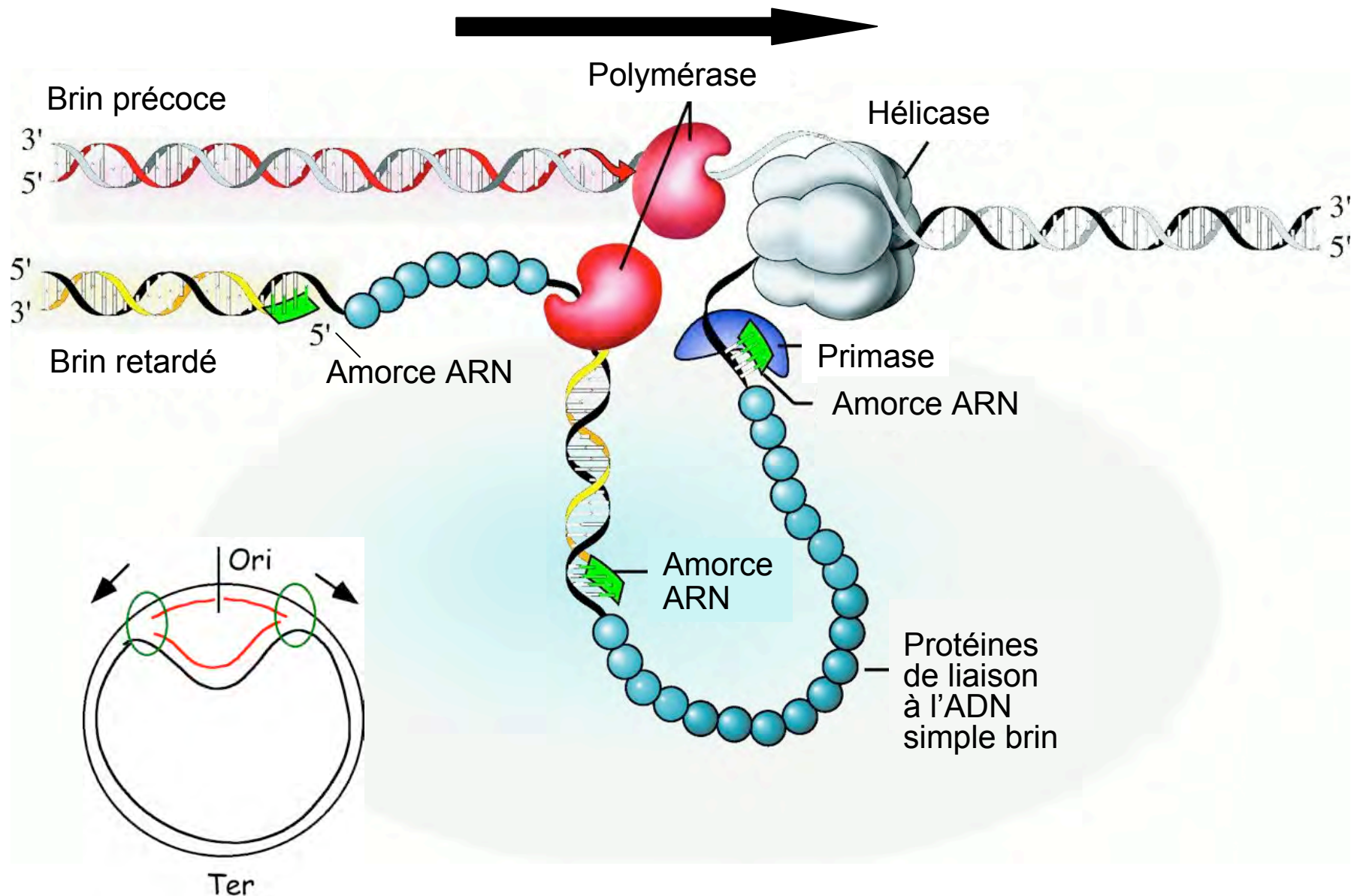
<=



HKU - PASTEUR RESEARCH CENTRE

香港大學—巴斯德研究中心

Réplication de l'ADN



Deuxième loi de complémentarité : la transcription

« L'opéron lactose »

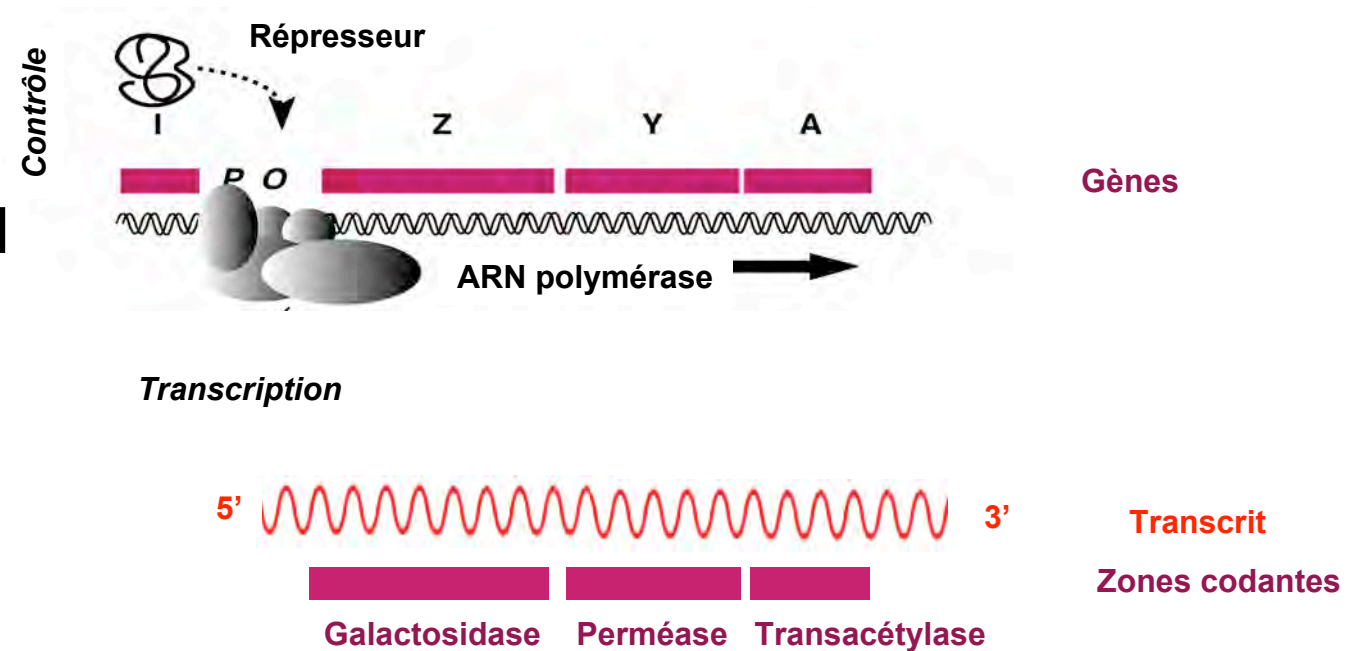
ADN => ARN

dA => rU

dC => rG

dT => rA

dG => rC



JACOB F, PERRIN D, SANCHEZ C, MONOD J.C R L'opéron: groupe de gènes à expression coordonnée par un opérateur

C R Hebd Seances Acad Sci. (1960) **250**:1727-9

JACOB F, MONOD J. Gènes de structure et gènes de régulation dans la biosynthèse des protéines

C R Hebd Seances Acad Sci. (1959) **249**:1282-1284.

Le code génétique et la traduction

La séquence génomique est vue comme un texte écrit avec un alphabet de quatre lettres. Ce texte est associé à un code (au sens de « code secret »), le code génétique, qui conduit d'un alphabet à quatre lettres à un alphabet de 20 lettres, les acides aminés.

Ce codage permet le retour sur soi, la **récurtivité**, puisque ce sont des protéines, codées par des gènes écrits sous forme de séquences de nucléotides dans l'ADN qui le manipulent. Ce processus crée des propriétés originales des liens entre objets.

Douglas Hofstadter (original : Basic Books, 1979 ; traduction, 2000)
Gödel, Escher, Bach : Les Brins d'une Guirlande Eternelle, Dunod.

“Les organismes vivants comme pièges à information”

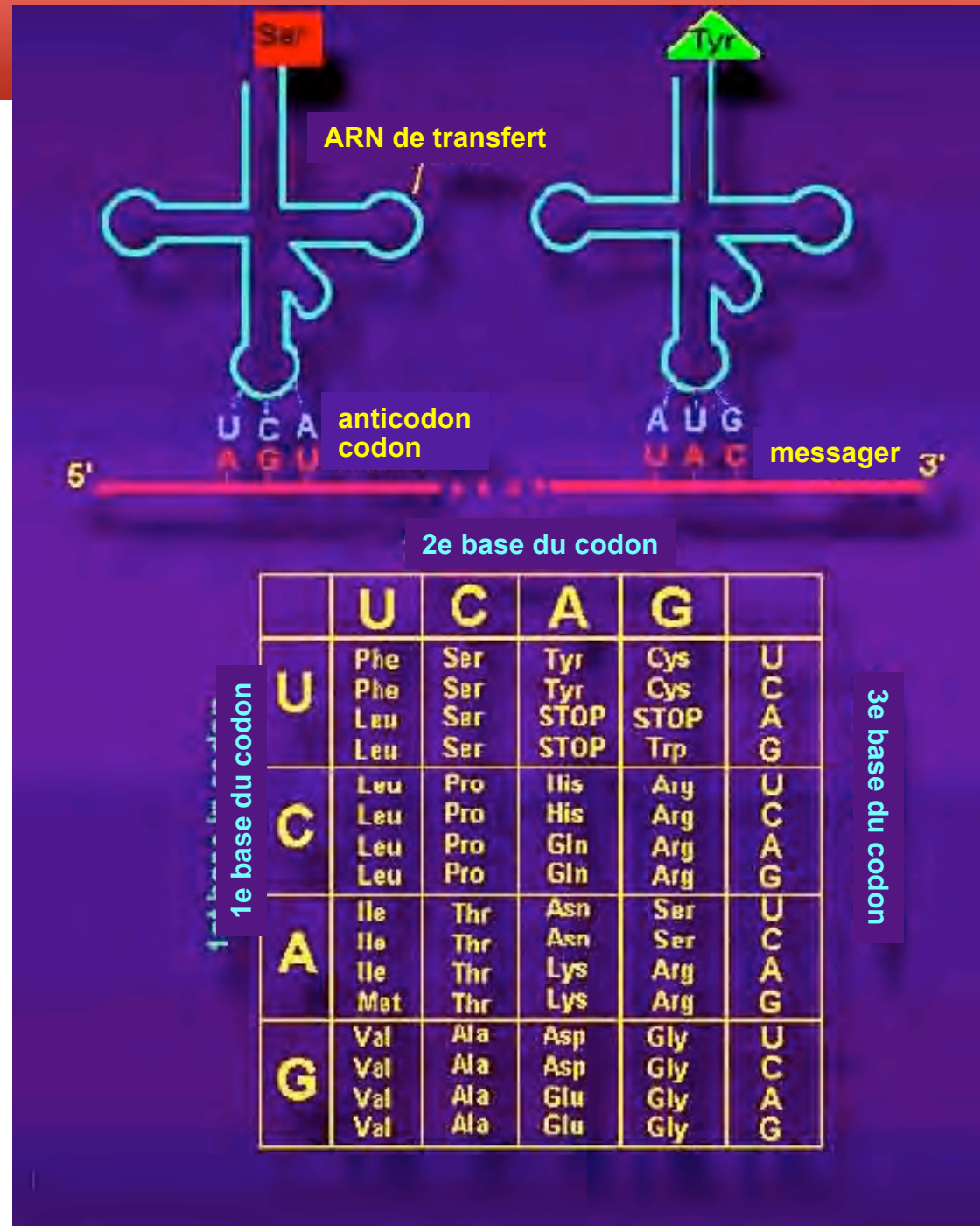
<http://www.normalesup.org/~adanchin>

Protéine

Adaptateur
ARN de transfert

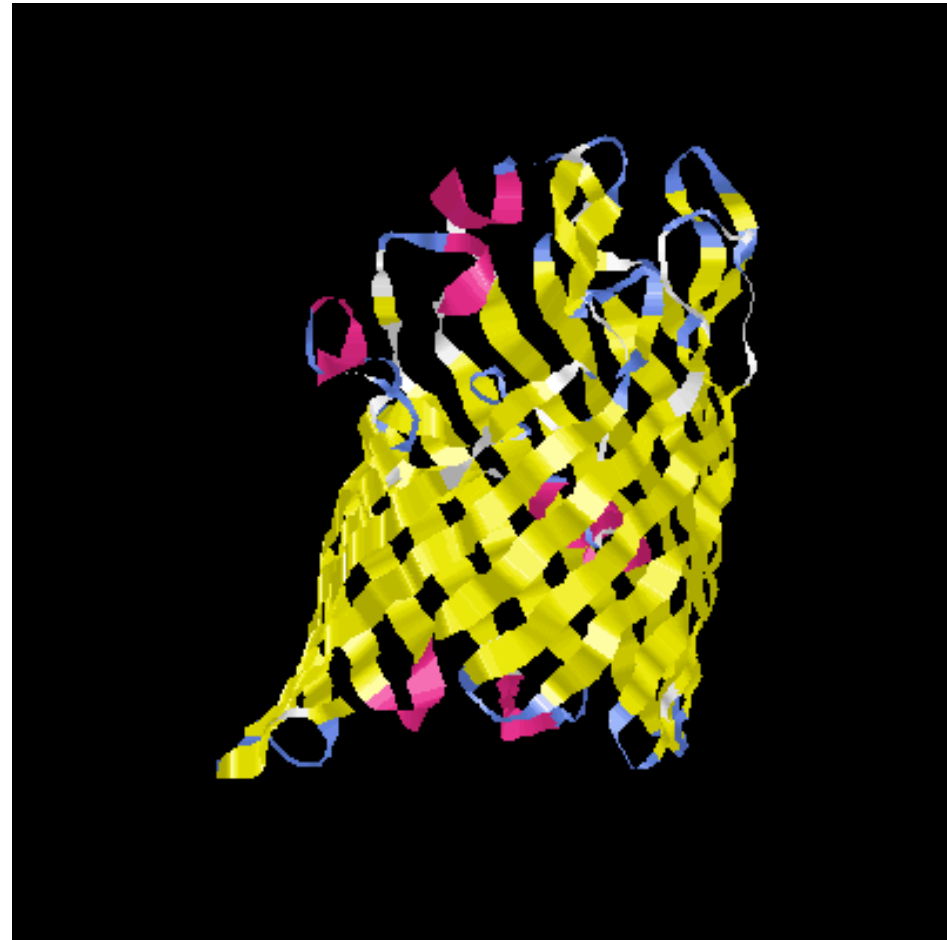
ARN messenger

Le code
génétique



Manipulation

Les protéines **se replient en trois dimensions** et participent à toutes sortes de processus : métabolisme, compartimentation, contrôle...

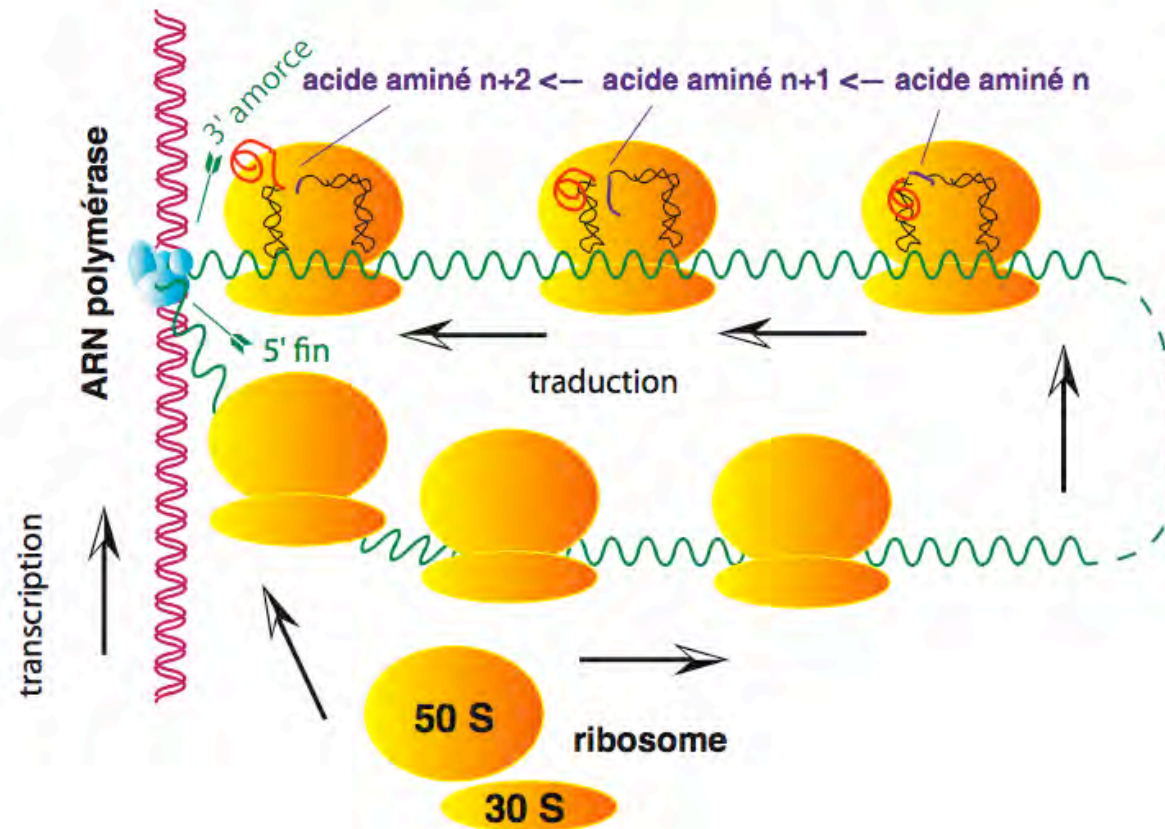


“Les organismes vivants comme pièges à information”

<http://www.normalesup.org/~adanchin>

L'usine cellulaire : "l'expression génétique"

Les ribosomes sont les « têtes de lecture » qui lisent les ARN transcrits en ARN messagers en utilisant un à un des ARN de transfert pour lire les codons



Cellules et ordinateurs

La génétique repose sur la description des génomes comme des 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» (l'usine cellulaire) et «Données/Programme» (le génome)

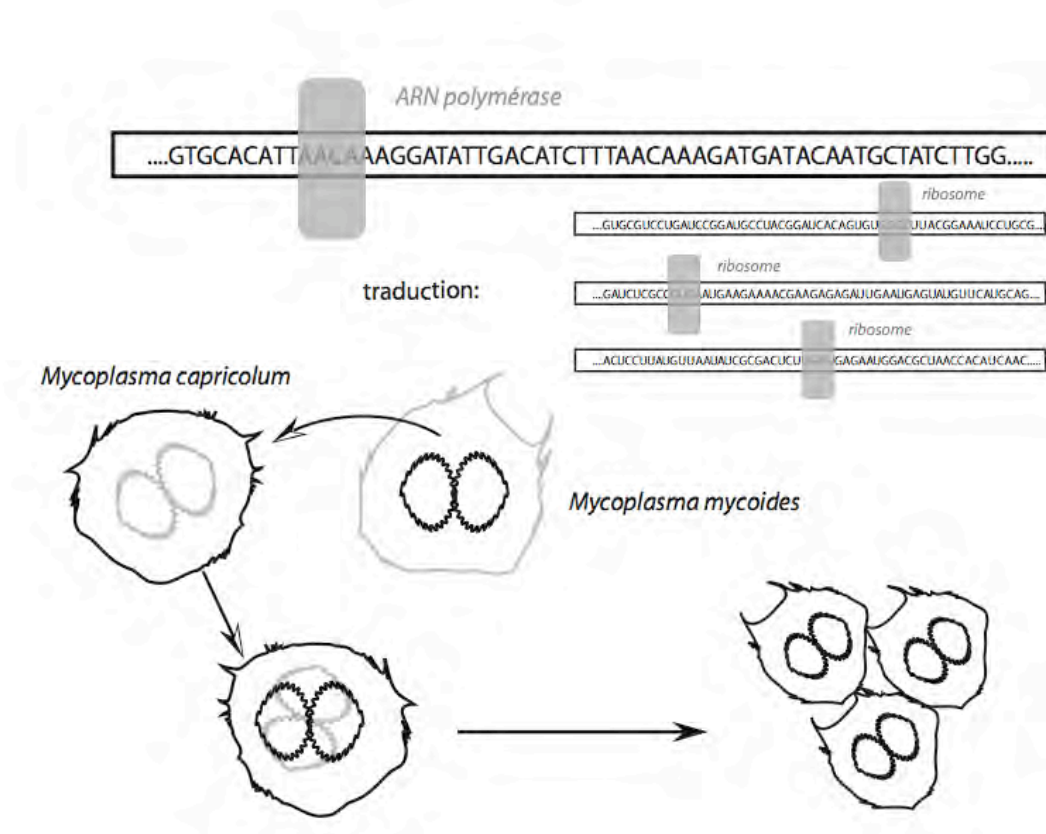
CONJECTURE: les organismes vivants sont des pièges à information, et nous pouvons identifier les processus concrets qui permettent l'accumulation de l'information

La démonstration de Lartigue-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

“Les organismes vivants comme pièges à information”

<http://www.normalesup.org/~adanchin>

Un ordinateur génétique

- Dans un ordinateur, la machine est séparée des données et du programme
- Données et programme jouent le même rôle (*càd* on peut les considérer comme des ‘déclarations’); ils peuvent être modifiés par la machine
- Une réflexion générale (due à Alan Turing et de nombreux autres penseurs) considère les actions de la machine, mais pas la façon dont elle est construite

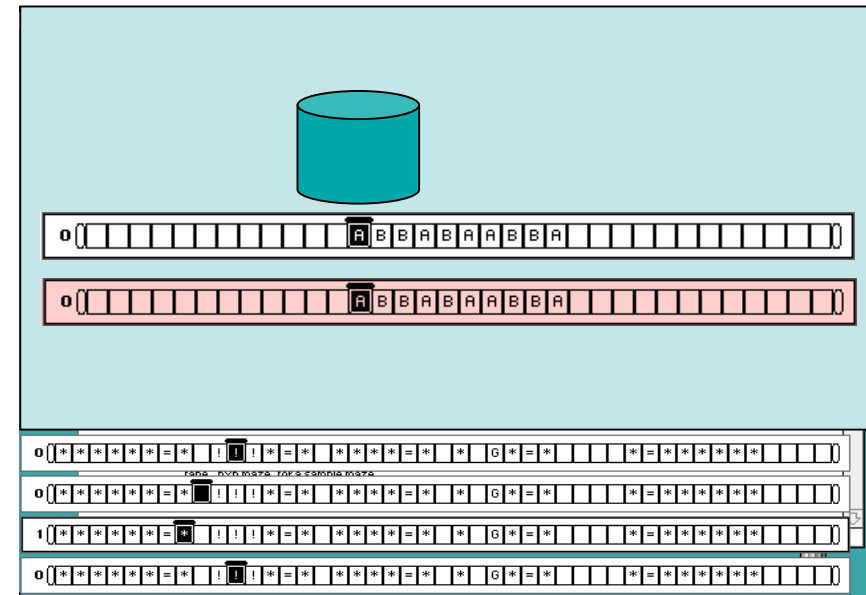
“Les organismes vivants comme pièges à information”

<http://www.normalesup.org/~adanchin>

Ce qu'est le calcul

Deux entités permettent le calcul :

- Une machine capable de lire et d'écrire
- Un programme (sur un support physique), séparé par l'esprit humain (pas conceptuellement !) en deux entités :
 - Programme (définissant l'"objectif")
 - Données (définissant le contexte)



La machine est distincte de l'entité données/programme

“Les organismes vivants comme pièges à information”

<http://www.normalesup.org/~adanchin>

Virus et programme

Comme un virus informatique, un virus biologique est fait d'un ou plusieurs acides nucléiques (son génome), codant une ou plusieurs protéines permettant sa réplication, sa protection dans la cellule et l'environnement, et sa propagation.

Souvent il code aussi des activités qui interfèrent avec le métabolisme de la cellule hôte, le détournant au profit de sa propre construction. Comme dans les ordinateurs il peut piéger des fonctions de la machine et les transmettre à sa descendance.

Cependant il ne code pas la totalité des objets permettant l'expression de son génome. **Il dépend entièrement de l'usine cellulaire pour se perpétuer.**

“Les organismes vivants comme pièges à information”

<http://www.normalesup.org/~adanchin>

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 simplement que notre réflexion sur le concept d’information est incomplète

Même pour 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 DNA fragments to form a complete synthetic *Mycoplasma genitalium* genome. *Proc Natl Acad Sci U S A.* (2008) 105:20404-20409

“Les organismes vivants comme pièges à information”

<http://www.normalesup.org/~adanchin>

Organisation des génomes

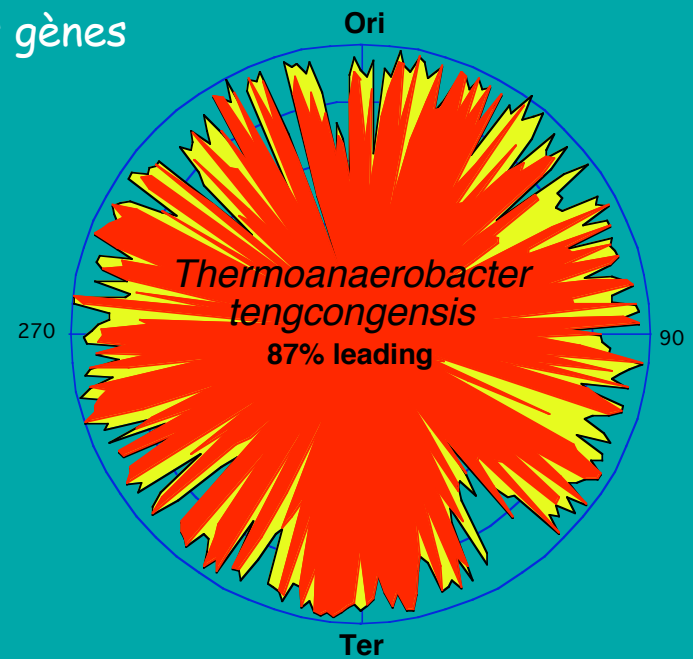
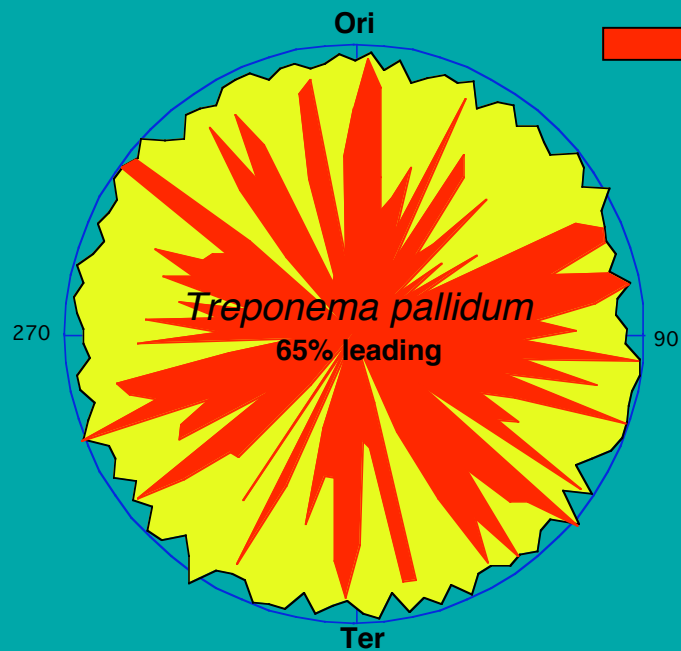
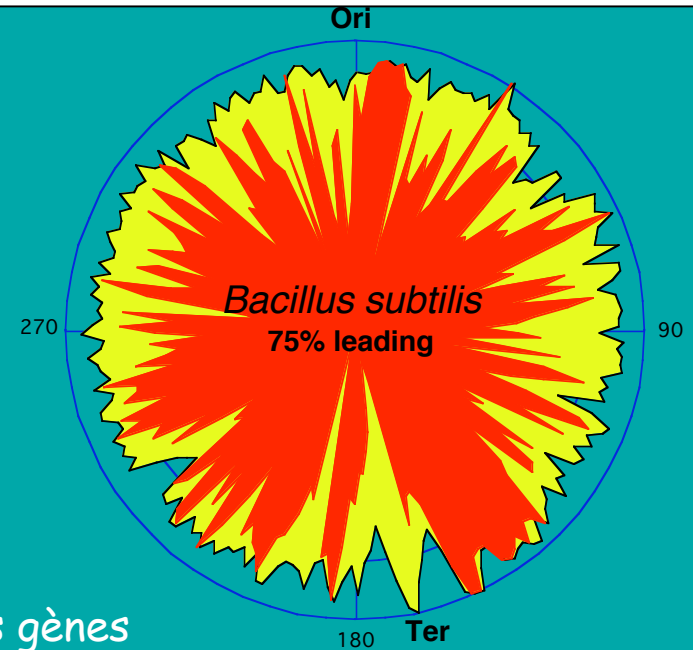
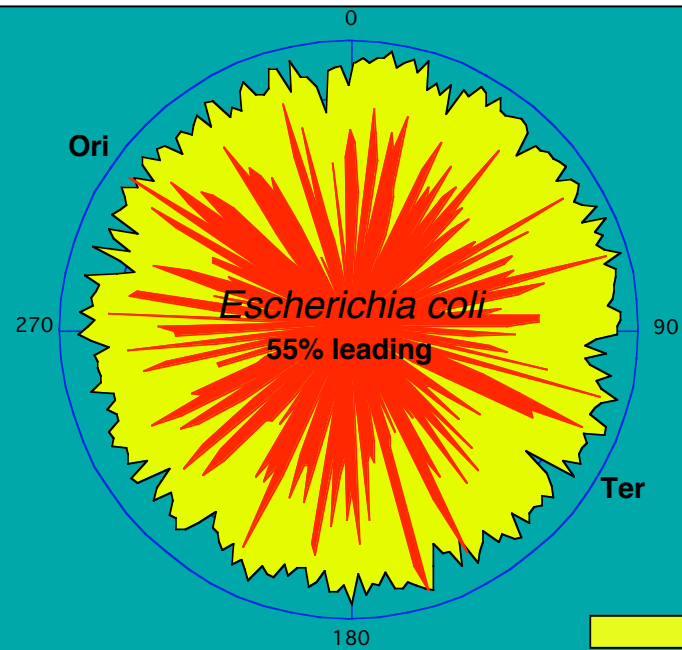
L'ordre des gènes est-il aléatoire dans les chromosomes ?

En plus d'une gestion différente de l'ADN dans les divers organismes les gènes semblent répartis plus au moins au hasard dans le génomes

Pourtant, des groupes de gènes comme les **opérons** ou les **îlots de pathogénicité** se regroupent, et codent des protéines dont les fonctions sont reliées

“Les organismes vivants comme pièges à information”

<http://www.normalesup.org/~adanchin>



■ Densité des gènes
■ Densité des gènes dans le brin précoce

Organisation du génome

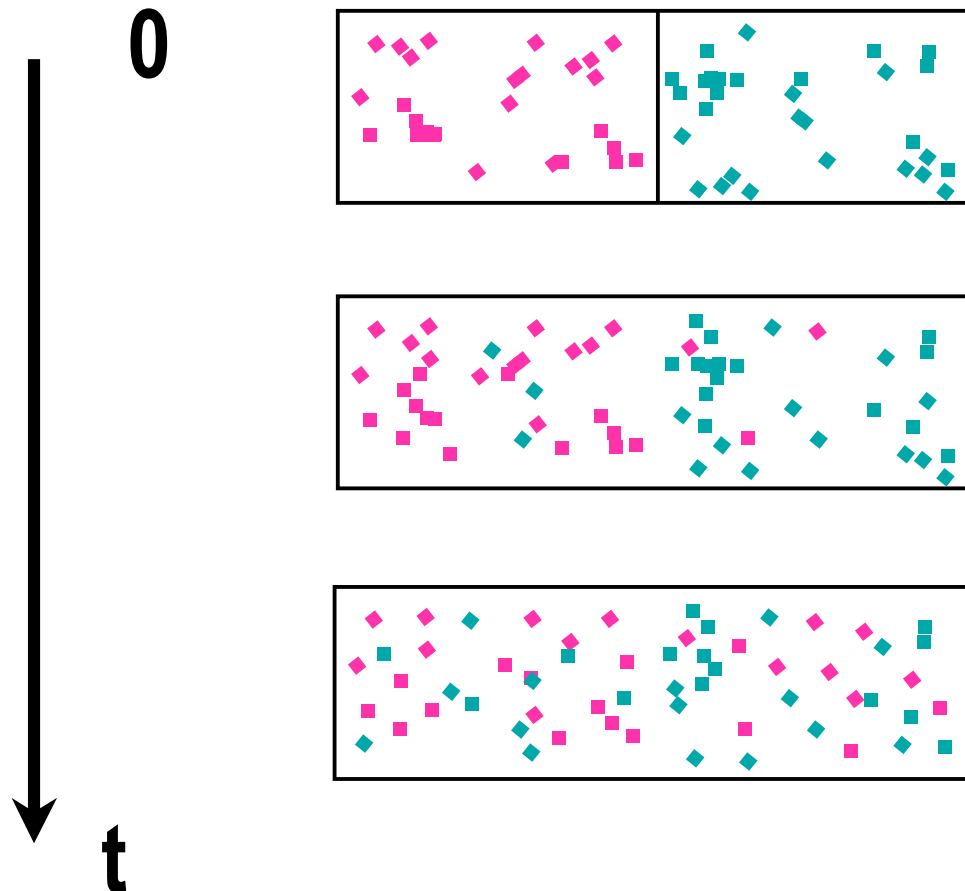
Quelles forces physiques permettent la séparation des chromosomes, l'organisation des gènes, leur regroupement ?

“Les organismes vivants comme pièges à information”

<http://www.normalesup.org/~adanchin>

Une vue familière de l'entropie

Pour des particules isolées, l'évolution spontanée va de l'hétérogène vers l'homogène

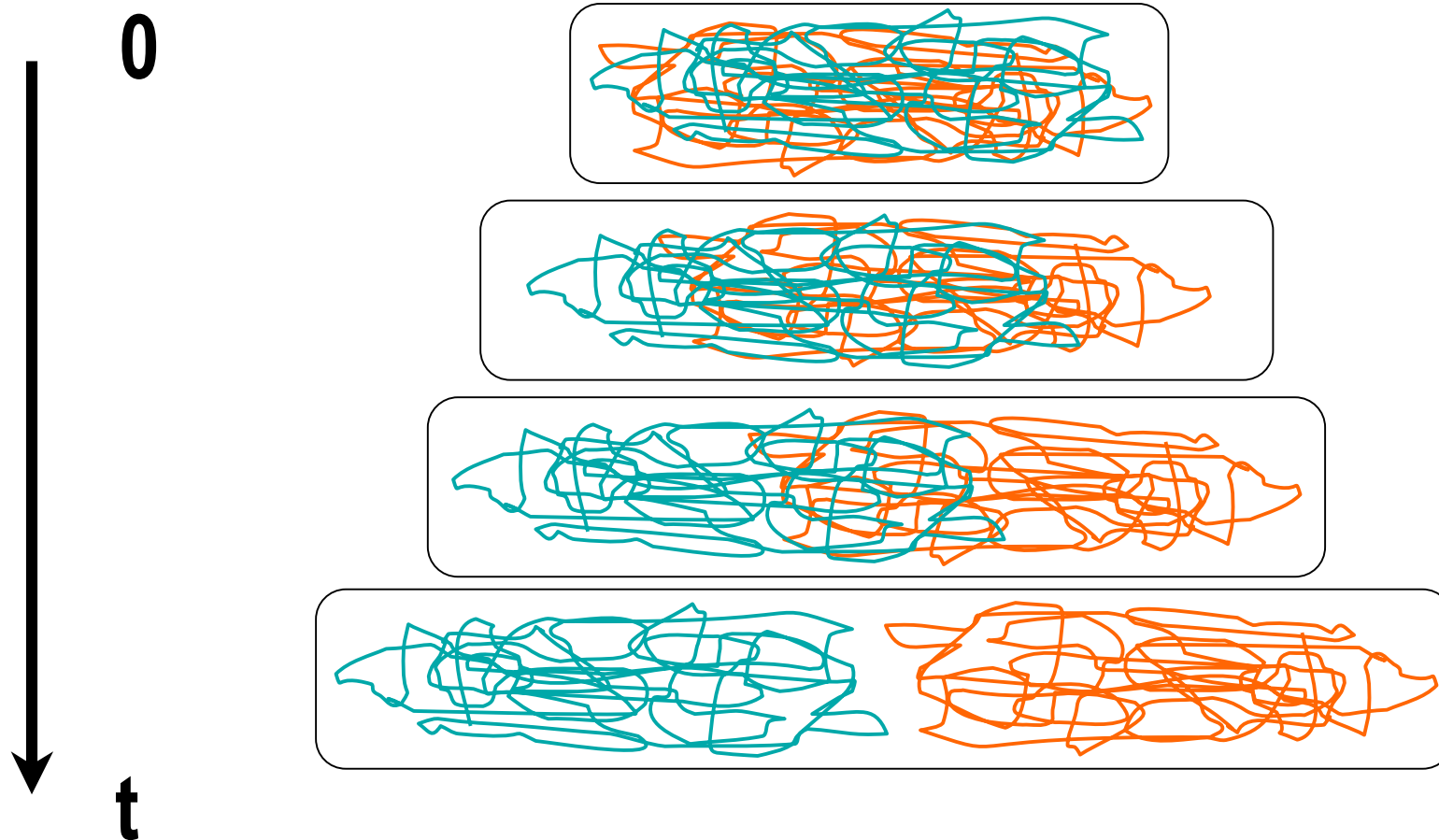


$$S = k \log \Omega$$



Benjamin Crowell, licensed under the Creative Commons Attribution-ShareAlike license

L'accroissement d'entropie sépare les chromosomes !



Danchin A, Guerdoux-Jamet P, Moszer I, Nitschke P. Mapping the bacterial cell architecture into the chromosome. *Philos Trans R Soc Lond B Biol Sci* (2000), **355**:179-190

Jun S, Mulder B. Entropy-driven spatial organization of highly confined polymers: lessons for the bacterial chromosome. *Proc Natl Acad Sci U S A.* (2006) 103:12388-12393

Une image de la cellule dans le chromosome ?

Si la machine doit non seulement se comporter comme un ordinateur mais doit aussi construire la machine, il faut trouver quelque part une image de la machine (John von Neumann)

Y a-t-il une image de la cellule (de l'organisme) dans le génome ? L'ordre des gènes est-il aléatoire ?

“Les organismes vivants comme pièges à information”

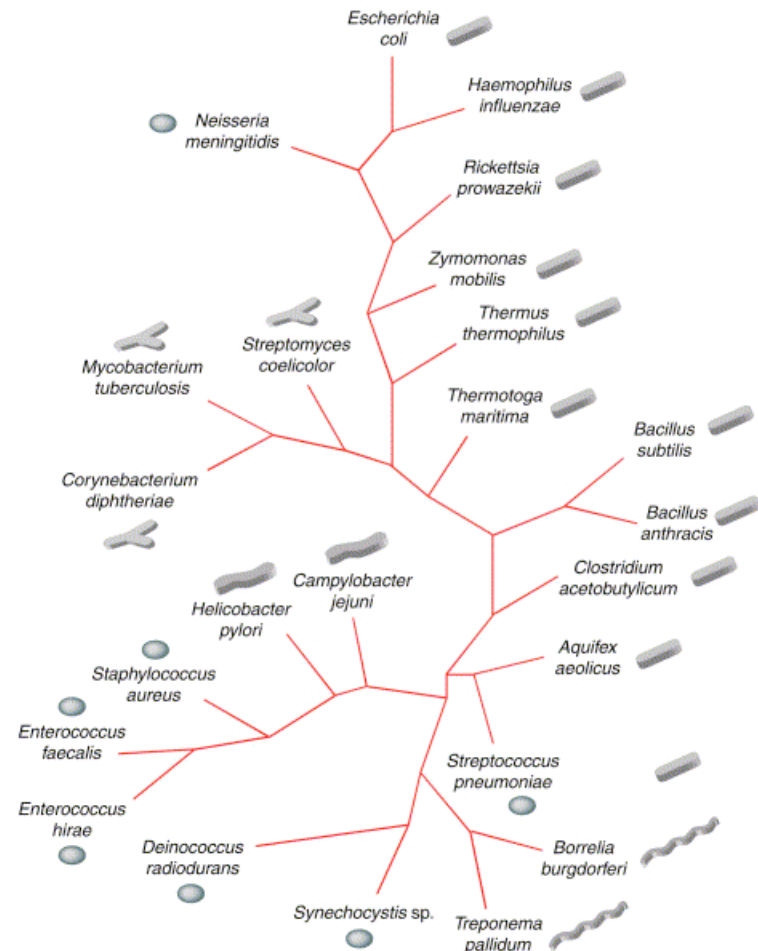
<http://www.normalesup.org/~adanchin>

Ordre des gènes et forme de la cellule

Les îlots *mur-fts*



TRENDS in Genetics



TRENDS in Genetics

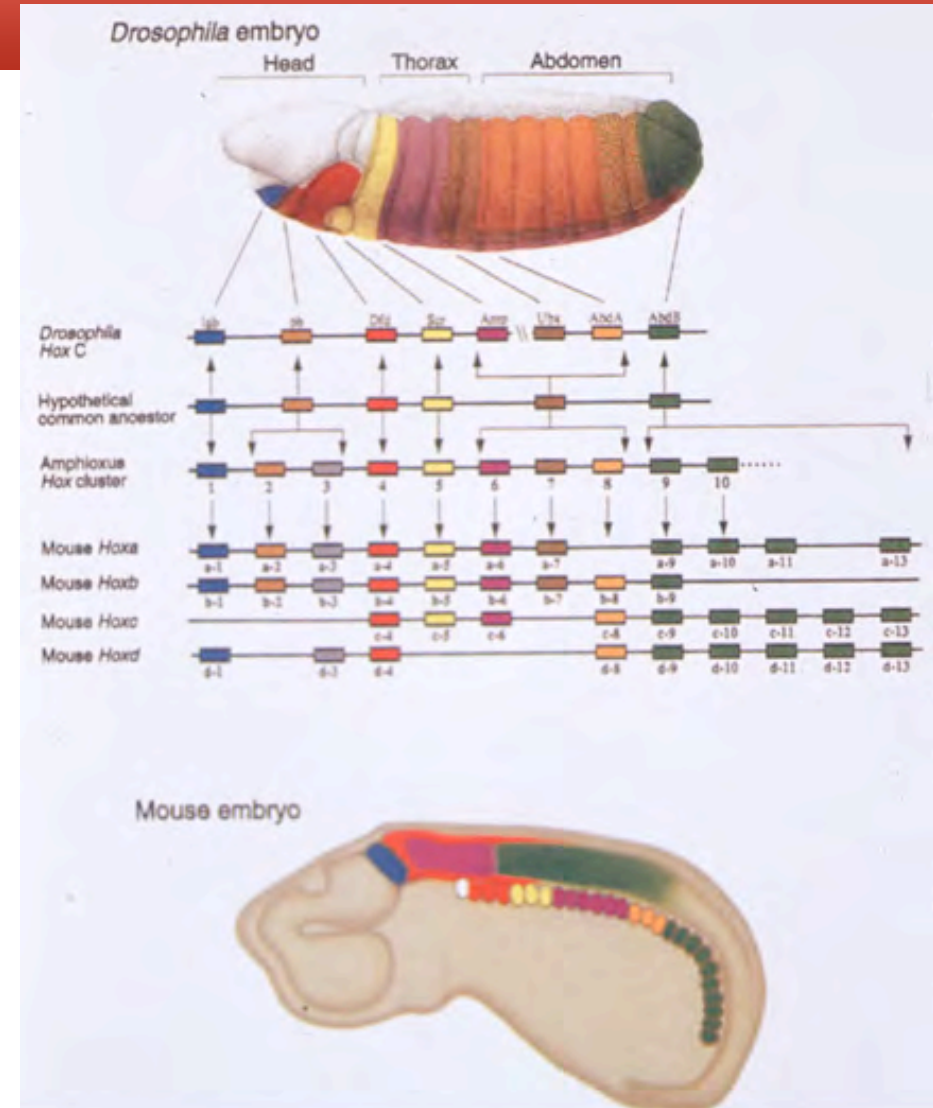
Tamames J, Gonzalez-Moreno M, Mingorance J, Valencia A, Vicente M

Bringing gene order into bacterial shape

Trends in Genetics (2001) 17: 124-126

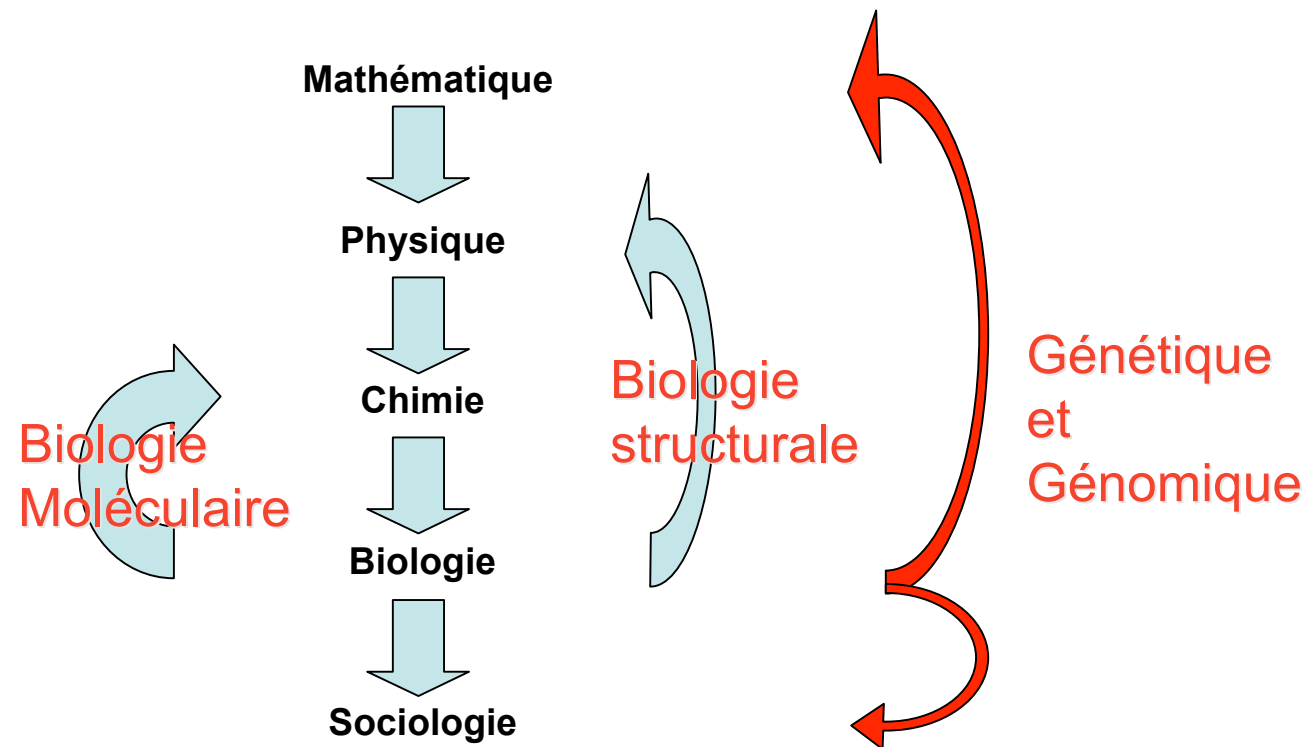
Les homéogènes

Drosophiloculus,
Homunculus?
Celluloculus?



From Carroll, S. B. (1995) *Nature* 376, 479

Une idéologie sous-jacente



“Les organismes vivants comme pièges à information”

<http://www.normalesup.org/~adanchin>

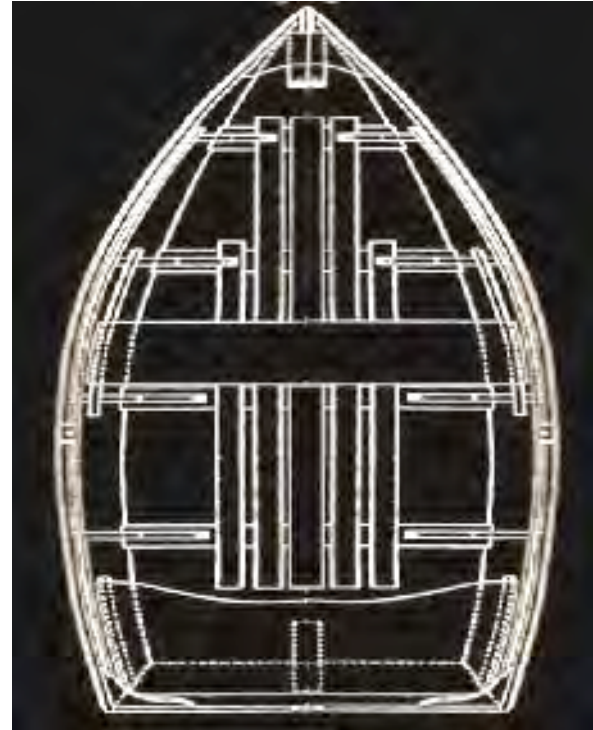
« Complexité » : un terme à éviter

Ce mot signifie en effet des choses diamétralement opposées

Il signifie ou bien qu'un phénomène est d'un ordre très riche et très compliqué parce qu'il est *hautement organisé* mais comporte de multiples éléments (d'un nombre généralement supposé dépasser l'entendement), ou tout au contraire, parce qu'il est très embrouillé, très *désordonné*

La barque de Delphes

- La biologie est une science des relations entre objets
- Elle est **symplectique** (de $\sigma\upsilon\nu$ ensemble, $\pi\lambda\epsilon\kappa\tau\epsilon\iota\nu$, tisser), le même mot que « complexe » en Latin
- 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

“Les organismes vivants comme pièges à information”

<http://www.normalesup.org/~adanchin>

Au delà des catégories usuelles

La loi du code génétique ne peut absolument pas être conçue *a priori* si l'on connaît seulement l'équation qui règle la vie des atomes, l'équation de Schrödinger. Elle est de l'ordre de la *symbolique*. Il n'y a rien que d'arbitraire dans la correspondance qui existe entre un acide aminé d'une protéine, et les codons qui le spécifient.

En 1933, dans un article intitulé *Light and Life* Niels Bohr remarquait: *"Reconnaître l'importance essentielle de caractéristiques fondamentalement atomiques dans le fonctionnement des organismes vivants ne suffit nullement, cependant, à fournir l'explication globale des phénomènes biologiques, nécessaire pour que nous puissions atteindre la compréhension de la vie à partir de l'expérience physique seule"*.

Information

Dans l'histoire, la physique fondamentale s'est occupée de découvrir les particules fondamentales de la nature et les équations qui décrivent leurs mouvements et leurs interactions. Il apparaît désormais qu'un programme différent peut avoir une importance équivalente : découvrir les modes que la nature prend pour moduler et prévenir l'expression et la manipulation de l'information plutôt que s'occuper du mouvement des particules

Andrew Steane (1998) Université d'Oxford

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

“Les organismes vivants comme pièges à information”

<http://www.normalesup.org/~adanchin>

Explorer l'information : l'infotaxie



Grand Paon de Nuit
Saturnia pyri
<http://pdubois.free.fr/>

Comment un papillon peut-il trouver 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...

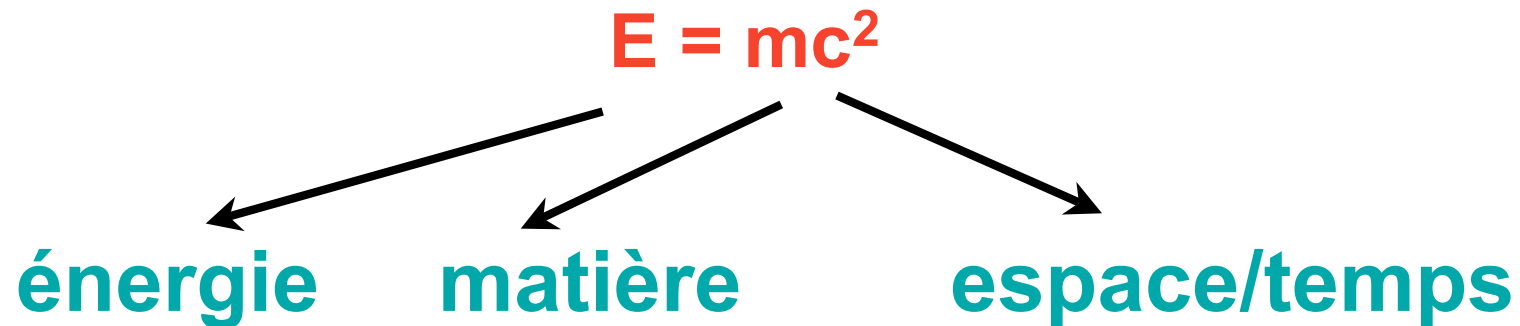
'Infotaxis' as a strategy for searching without gradients

Vergassola M, Villermaux E, Shraiman BI
Nature (2007) **445**: 406-409



L'information, une cinquième catégorie de la nature ?

Physique classique



Physique Quantique

$$\Delta x \Delta p \geq h/4\pi$$

indétermination = manque d'**information**

"Les organismes vivants comme pièges à information"

<http://www.normalesup.org/~adanchin>

Revisiter la hiérarchie des sciences

Matière / Énergie / Espace / Temps

- Physique Classique
- Physique Quantique
- Chimie
- Biologie
 - Développement
 - Neurobiologie
 - Linguistique
- Mathématique

Information



“Les organismes vivants comme pièges à information”

<http://www.normalesup.org/~adanchin>

Différents niveaux d'information

- Ce qui est vu par la réplication : aucune « signification », information de Shannon
- Ce qui est vu par la machinerie de l'expression des gènes et par l'évolution
 - **Complexité algorithmique (espace)**
 - **Profondeur logique (temps)**
 - **(Profondeur critique) (compaction)**

Danchin, A. 1996. On genomes and cosmologies. In *Integrative Approaches to Molecular Biology* (eds. J. Collado-Vides B. Magasanik, and T. Smith), pp. 91-111. The MIT Press, Cambridge (USA)

“Les organismes vivants comme pièges à information”

<http://www.normalesup.org/~adanchin>

Ce que voit la réplication

Au cours de la réplication la séquence de l'ADN est recopiée exactement. La copie ne se soucie nullement de la signification de la séquence

L'enveloppe génétique (la collection des séquences recopiées) peut être étudiée à un premier niveau de l'information qui lui correspond. Typiquement on utilise à ce stade l'« information » ou « entropie » de Shannon

L' entropie de Shannon (1)

Une collection de messages (suite de symboles dans un alphabet fini) communiqués au travers d'un canal bruité a la forme probabiliste suivante :

$$H(p_i) = - \sum \{p_i \log_2 p_i \mid i \in I\}$$

Cette formule est si générale que de nombreux modèles d'autres phénomènes que la communication de suite de symboles peuvent être ainsi représentés. **Danger des extrapolations** : la sphère est caractéristique des étoiles, des oranges, et des bulles de savon...

La **validité** de cette formule repose sur des hypothèses très fortes sur la nature des signaux en cause (en particulier ils doivent satisfaire aux lois des probabilité de Laplace-Gauss)

L' entropie de Shannon (2)

Mise en garde: Myron Tribus raconte que von Neumann, à qui Shannon avait demandé de l'aide pour trouver un nom pour sa fonction définissant l'information, lui proposa de façon prophétique :

*“Vous pouvez l'appeler **entropie** pour deux raisons. D'abord votre fonction d'incertitude a bien été utilisée en mécanique statistique sous ce nom, ce qui fait qu'elle a déjà un nom. Ensuite, et c'est plus important, personne ne sait réellement ce qu'est l'entropie, ce qui fait que dans le moindre débat vous aurez l'avantage”*

créant ainsi la boîte de Pandore de la plus grande confusion intellectuelle.

“Les organismes vivants comme pièges à information”

<http://www.normalesup.org/~adanchin>

Complexité algorithmique

- ➔ L'« entropie » de Shannon ne se soucie pas de la signification (réplication).
- ➔ Kolmogorov et d'autres ont proposé de définir un autre niveau d'information par la **longueur du plus petit programme définissant exactement une suite de symboles** : la complexité algorithmique. On a pensé que cela permet de définir le hasard : **ce qui est aléatoire est non-compressible**
- ➔ Connaissant une séquence de symbole, on ne peut en connaître la complexité... Mais cela nous fournit un **programme de recherche** : afin d'approcher la complexité algorithmique d'une séquence, on s'intéresse à la façon dont elle a été engendrée (au cours de l'évolution, au cours des synthèses, etc)

Procaryotes et eucaryotes

→ **Procaryotes** : une seule enveloppe, plus ou moins compliquée, génome superficiellement « aléatoire » ; superposition des signaux (combinatoire pour l'interaction avec des protéines impossible)

→ **Eucaryotes** : multiplication des membranes et des peaux ; génome riche en répétitions, juxtaposition des signaux (riche combinatoire des protéines qui contrôlent la transcription)

Dans les deux cas cela semble inintéressant ; n'est-ce pas la marque d'une faille dans notre description ?

Le contexte : “valeur” d’une information

La complexité algorithmique est insuffisante. Elle ne suffit nullement à définir le hasard. Tout dépend du **contexte. Une séquence quelconque doit se trouver quelque part dans le développement des décimales de π : si l’on sait en donner la position, alors la complexité algorithmique de la séquence diminue !**

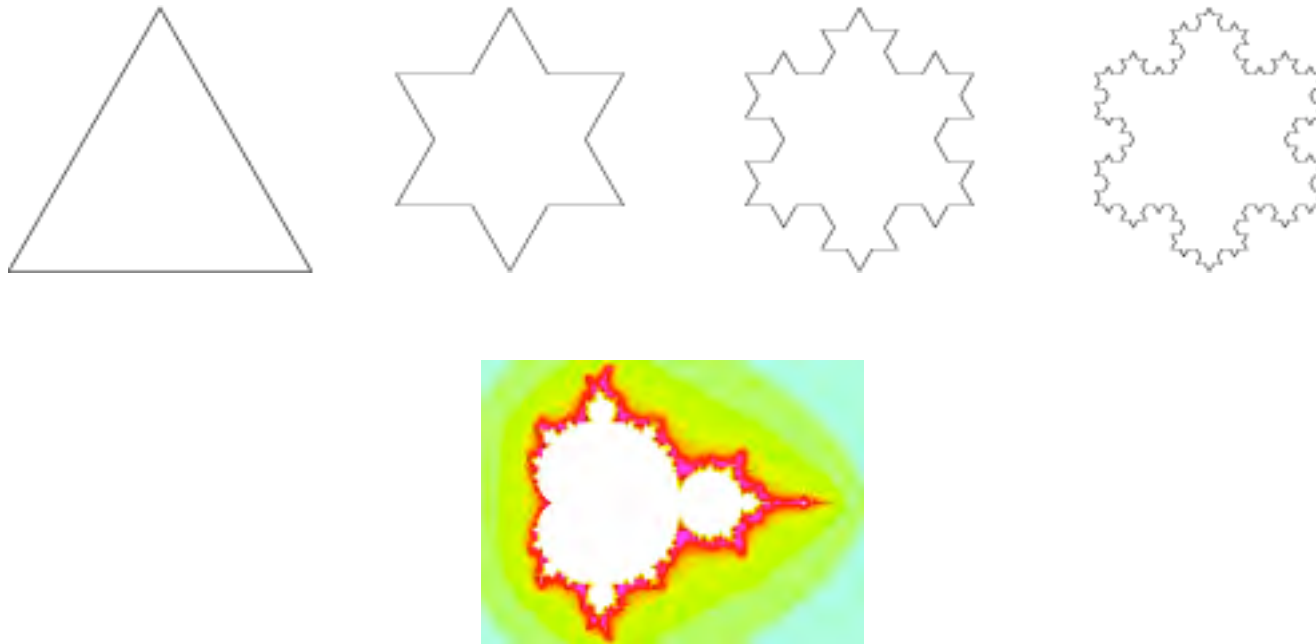
De même, les répétitions ont un sens, une valeur :

Les poules du couvent couvent

Pour un ordinateur, un locuteur français, un locuteur non-francophone ?

La profondeur logique (1)

Un programme très court (**faible complexité algorithmique**) peut décrire une simple séquence répétée, mais aussi une figure fractale comme le flocon de Koch ou l'ensemble de Mandelbrot



“Les organismes vivants comme pièges à information”
<http://www.normalesup.org/~adanchin>

La profondeur logique (2)

- La différence entre ces deux cas de même complexité algorithmique concerne la **valeur** de l'information associée à la $n^{\text{ième}}$ étape
- Les programmes récursifs impliquent qu'il est nécessaire d'exécuter le programme pour obtenir l'information
- Bennett a nommé « profondeur logique » le temps nécessaire pour avoir accès à cette information

“Les organismes vivants comme pièges à information”

<http://www.normalesup.org/~adanchin>

Vision phylogénétique du monde

- ➔ Les programmes génétiques évoluent par mutations (parfois de simples réarrangements, sans changement du « texte » génomique)
- ➔ Ceux qui ne sont pas totalement incapables de survivre et de se multiplier ont une descendance : survivre implique que les conditions locales de l'environnement **créent une information qui sculpte une image du monde** dans les génomes qui perdurent
- ➔ Ce processus crée, via la mise à l'épreuve du fonctionnement concret de chaque organisme sa stabilité dans l'environnement où il est né

Trois niveaux reliés sculptent une image du monde

Trois processus se superposent pour constituer l'individu final :

→ **Phylogenèse (Evolution)**

→ **Ontogenèse (Différentiation: les homéogènes)**

→ **Epigenèse (Individuation)**

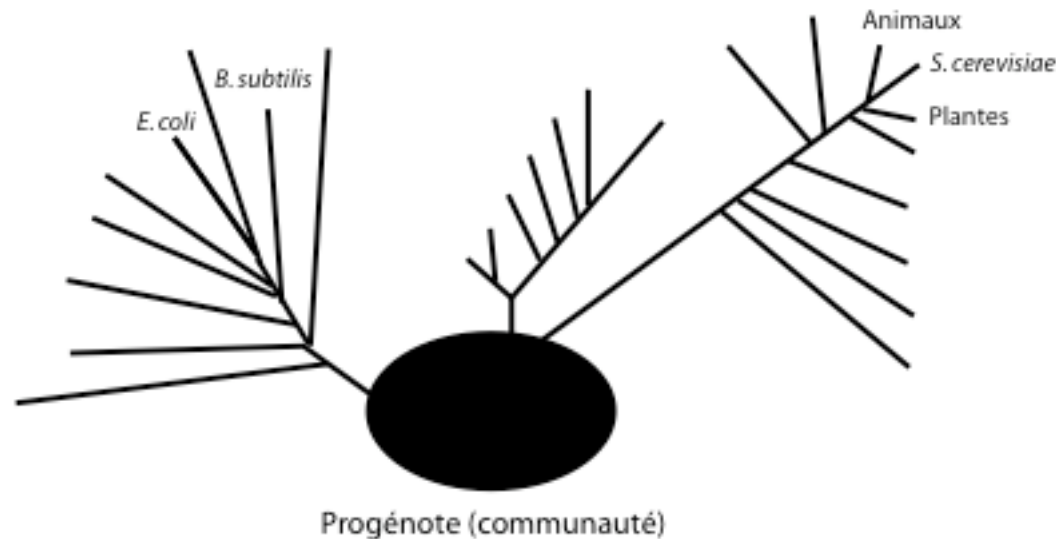
Chacun est stabilisé par la pression de sélection des événements qui agissent sur les fonctions en cause

Phylogenèse

Bactéries

Archées

Eucaryotes



(Woese, 1990)

Tout organisme est le fruit d'une longue évolution, la profondeur logique implique qu'il n'y a pas d'ADN « poubelle » (« junk » DNA)

L'origine des fonctions est floue, **elle distingue la machine du programme** : ne soyons pas « adamistes » !

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

"Les organismes vivants comme pièges à information"

<http://www.normalesup.org/~adanchin>

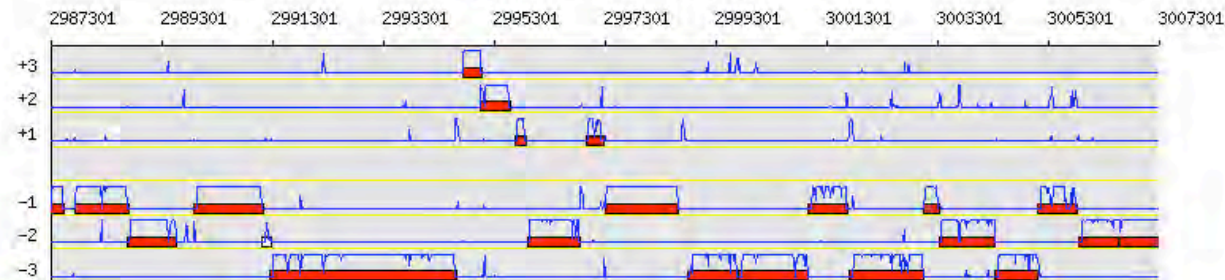
La pierre de Rosette de la génomique comparative



Bacillus subtilis 168 chromosome BSU 843

2987301 -- 3007301

(sequence length : 4215606 bases)



▼ Synteny results



C. Médigue

De l'ubiquité fonctionnelle à la persistance des gènes

L'entropie de Shannon (2)

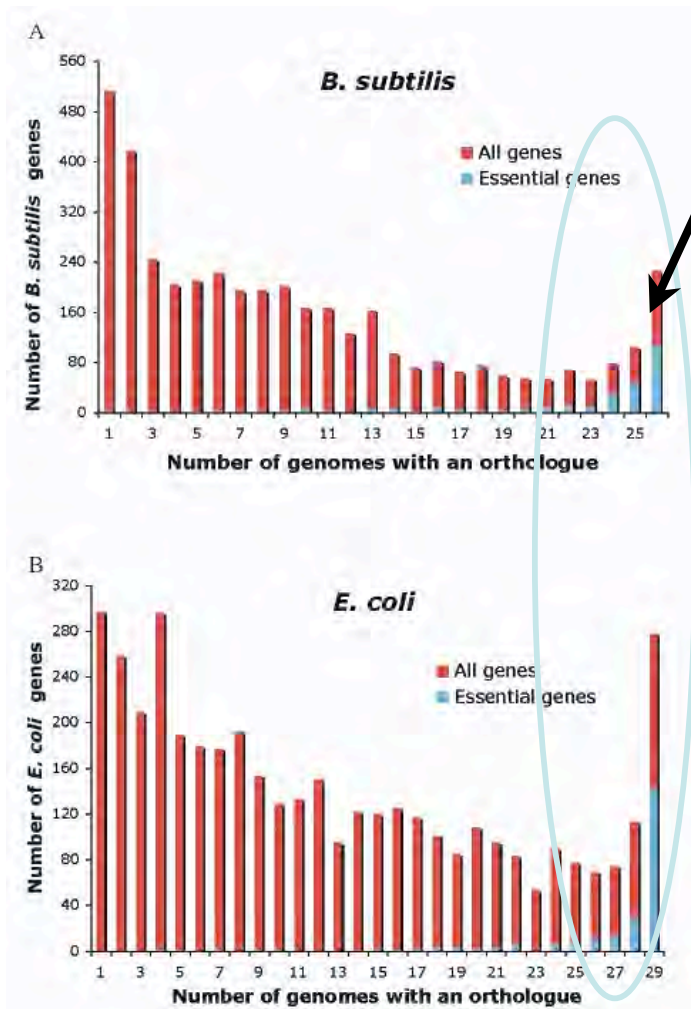
L'ubiquité fonctionnelle n'implique pas l'ubiquité des gènes
Les objets efficaces tendent à persister au cours des générations :

→ 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

Persistance : des gènes en trop !



Gènes persistants

Gènes essentiels et

Stress, maintenance et réparation

Dégradation utilisant de l'énergie

Rustines métaboliques (effet sérine)

G Fang, EPC Rocha, A Danchin
How essential are non-essential genes?
Mol Biol Evol (2005) 22: 2147-2156

“Les organismes vivants comme pièges à information”

<http://www.normalesup.org/~adanchin>

Une origine minérale de la vie

La nécessité d'une sélection de molécule dans un mélange complexe, et de la polymérisation de ces molécules conduit à penser à une origine du métabolisme **à la surface de solides**. Ce scénario met l'accent sur la séparation du métabolisme de la réplication, cette dernière étant une invention **tardive** de l'évolution prébiotique :

**Maillons élémentaires => nucléotides => ARNt
=> ribosome => ADN**

“Les organismes vivants comme pièges à information”

<http://www.normalesup.org/~adanchin>

Le monde des génomes à ARN

Le "monde ARN", si populaire, est impossible sans pré-existence de **molécules remarquablement instables**, les nucléotides

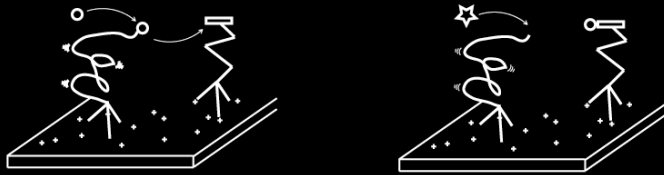
Les ARN ont deux rôles : ce sont des **substrats et catalyseurs** de réactions chimiques, et ce sont des **matrices**, qui permettent leur propre réplication

Le vrai "monde ARN" est le "**monde des génomes à ARN**", qui apparaît avec leur fonction de matrice répliquatives

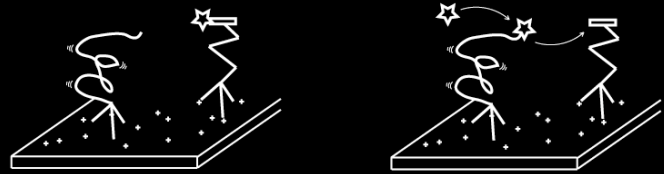
"Les organismes vivants comme pièges à information"

<http://www.normalesup.org/~adanchin>

Variations sur un thème



Transformation homéotopique



Origin



Origine



ARNt hors de la traduction

Formation d'un lien peptidique

Gly~ARNt^{Gly} murein
Phe~ARNt^{Phe} N-terminal
Leu~ARNt^{Leu} N-terminal
Ara~ARNt^{Ara} Ubiquitine

Homéotopie

Met~ARNt^{Met} → FMet~ARNt^{Met}
Glu~ARNt^{Glu} → Gln~ARNt^{Gln}
Ser~ARNt^{Ser} → SeCys~ARNt^{Ser}

Lipides et coenzymes

Lys~ARNt^{Lys} Lipides
Glu~ARNt^{Glu} Aminolevulinate

Autres modifications

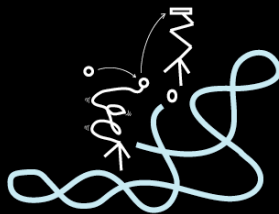
Base modifications

Présence d'un groupement phosphate sans raison chimique

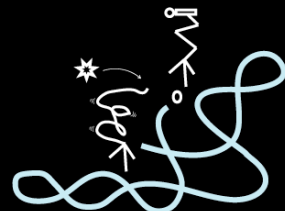
SERINE (SERINE~PHOSPHATE)
PYRIDOXAL PHOSPHATE
ISOPRENYL ~ PYROPHOSPHATE
4-PHOSPHOPANTETHEINE
THIAMINE ~ PYROPHOSPHATE

Réactions utilisant un nucléotide

CYTIDYLATE (LIPIDES)
ADENYLATE, GUANYLATE, URIDYLATE (SUGARS)
COENZYMES (NAD, CoA...)



Les ARN de transfert
se substituent aux surfaces
minérales



Origin



Origine



Un scénario minéral de l'origine de la vie

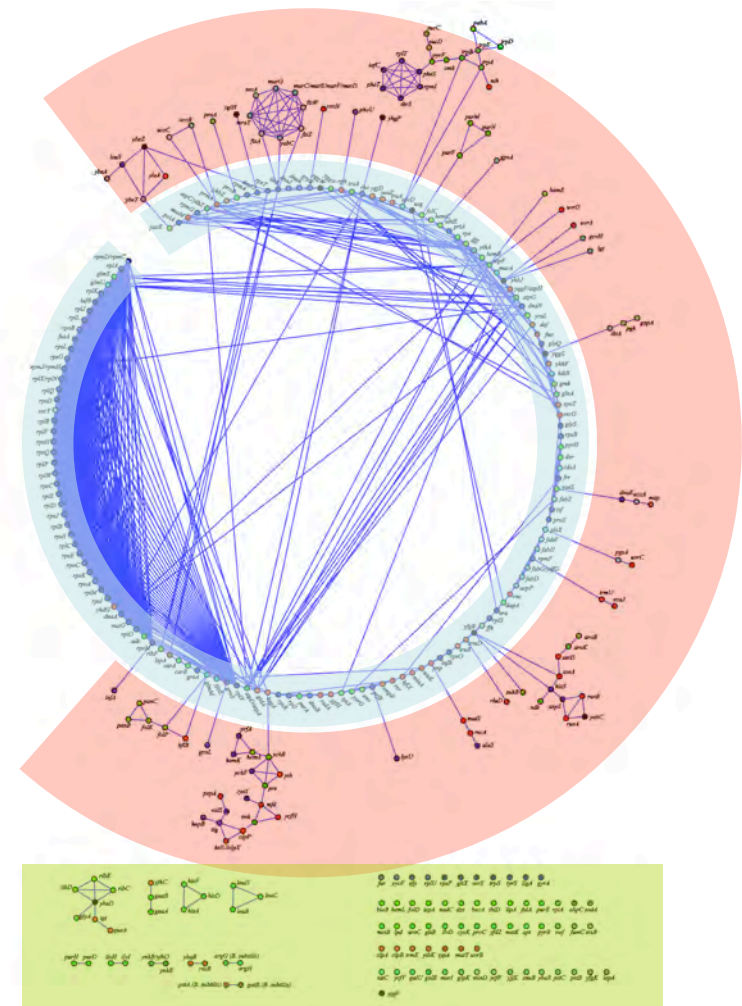
- La surface de solides chargés (comme la pyrite (Fe-S)) sélectionne et compartimente des molécules chargées ; cette étape forme quelques acides aminés, les principaux coenzymes, les acides gras et les ribonucléotides ; la polymérisation avec élimination de molécules d'eau est favorisée par l'augmentation de l'entropie
- Un métabolisme compartimenté crée des substituts aux surfaces via la polymerisation de ribonucléotides en présence de peptides (le monde ARN, avec les ancêtres des ARNt)
- Les ARN découvrent la loi de complémentarité, et le code génétique est inventé (des substrats aux matrices). Les acides nucléiques sont stabilisés par l'invention des désoxyribonucléotides, à un moment où les règles du contrôle des transferts d'information sont découvertes, d'abord au sein du monde des génomes ARN où des vésicules comprenant les ancêtres des gènes se scindent et fusionnent au hasard, avant la formation des premiers génomes

Les gènes persistants récapitulent l'origine de la vie

Le **réseau externe**, composé de gènes du métabolisme intermédiaire (nucléotides et coenzymes, lipides), est très fragmenté; le **réseau intermédiaire** est construit autour des tRNA synthétases de classe I, et le **réseau central**, presque continu, organisé autour du ribosome, de la transcription et de la réplication gère les transferts d'information

A Danchin, G Fang, S Noria

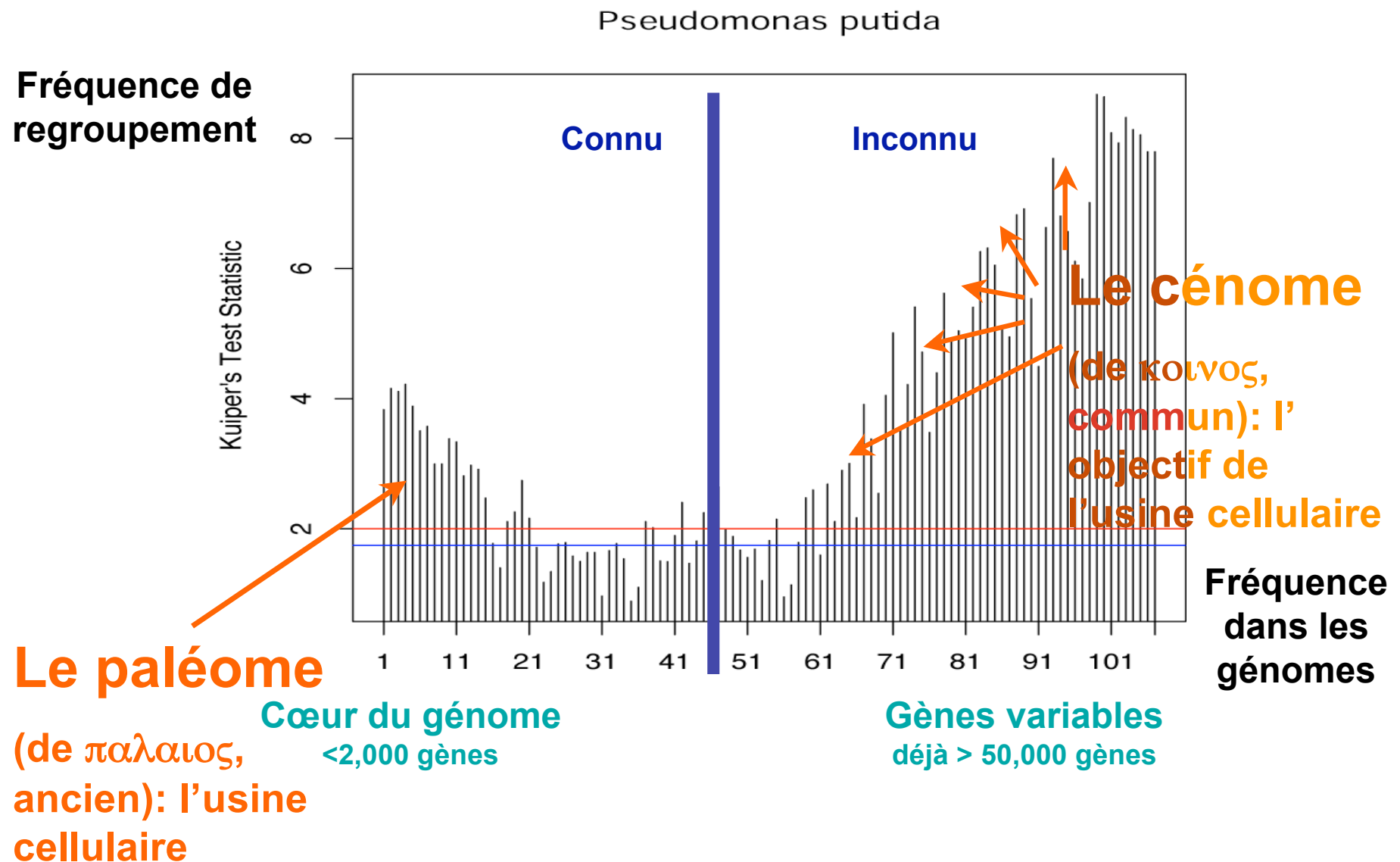
The extant core bacterial proteome is an archive of the origin of life
Proteomics. (2007) 7:875-889



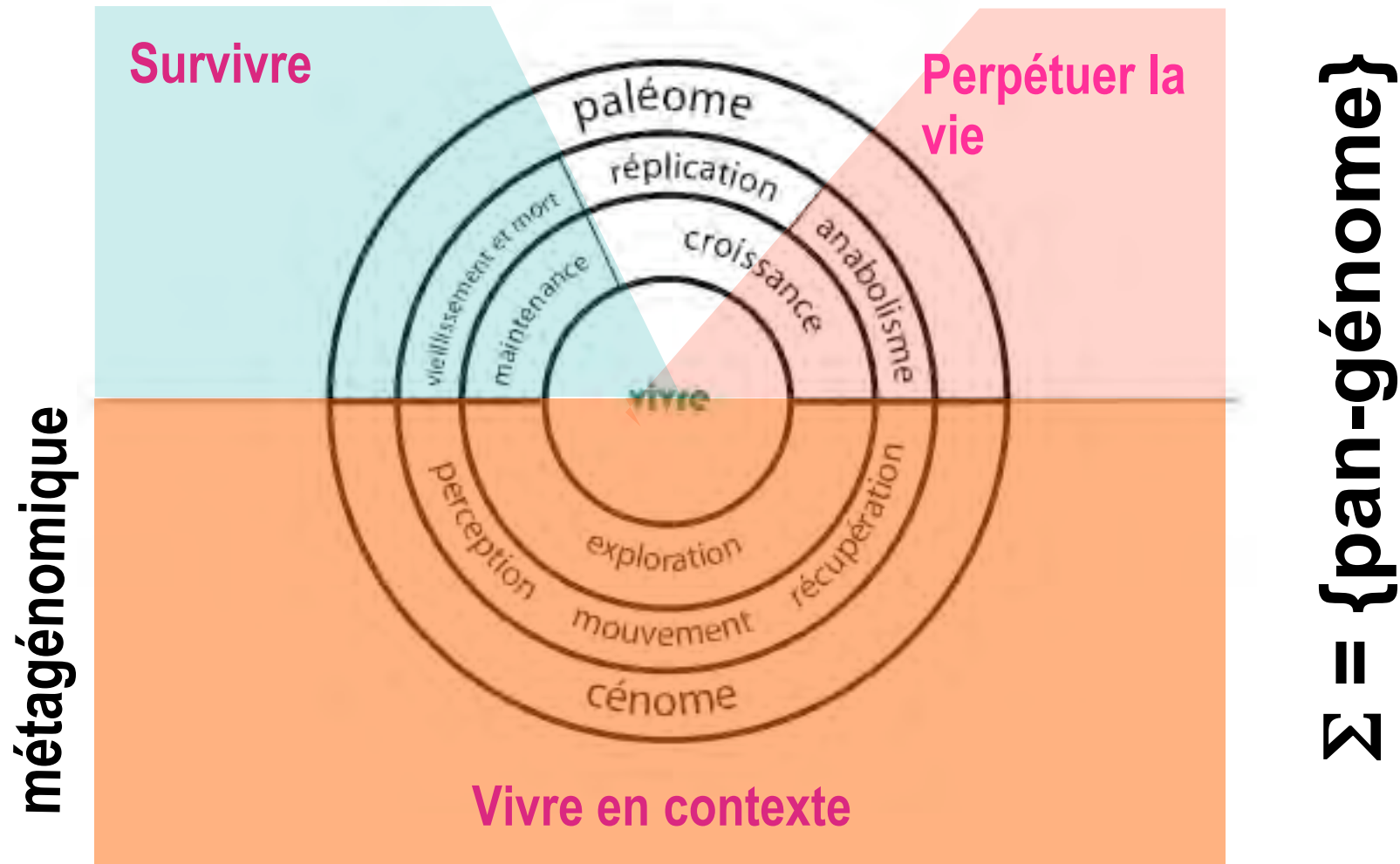
“Les organismes vivants comme pièges à information”

<http://www.normalesup.org/~adanchin>

Organisation des génomes bactériens



Histoire de deux génomes



“Les organismes vivants comme pièges à information”

<http://www.normalesup.org/~adanchin>

« Etre » implique le regroupement

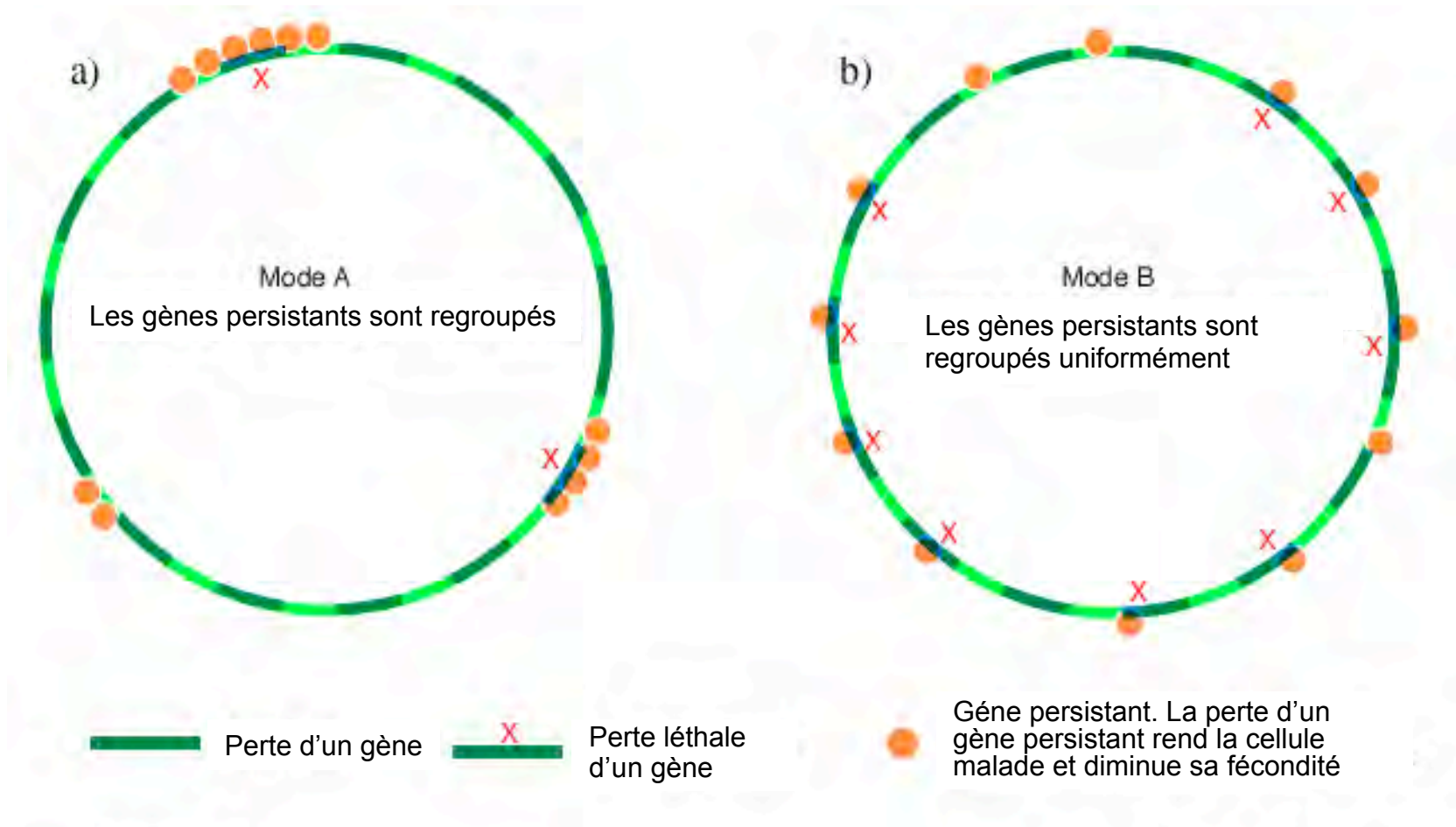
Pourquoi les gènes persistants sont-ils regroupés ? Un modèle très simple montre que si, pour compenser leur transfert « horizontal », les gènes disparaissent en groupes, alors tout gène contribuant à l'adéquation de l'organisme à son environnement assez fréquemment se regroupera avec les gènes ayant la même propriété

Cela explique le regroupement des gènes essentiels, mais aussi celui des gènes de résistance aux antibiotiques....

En conséquence le regroupement **précède et ne résulte pas de la** co-transcription ou des interactions protéine-protéine !

Il n'y a aucune prescience dans l'organisation du génome

L'existence implique le regroupement



Fang G, Rocha EP, Danchin A.
Persistence drives gene clustering in bacterial genomes.
BMC Genomics. (2008) **9**:4.

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 organismes vivants comme pièges à information”

<http://www.normalesup.org/~adanchin>

Origines de la vie

- Le physicien Freeman Dyson a montré qu'il ne peut y avoir une seule origine de la vie
- La **réplication** conduit presque inmanquablement à une avalanche d'erreurs (sauf en cas d'échanges avec l'extérieur qui permet de revenir en arrière)
- La **reproduction**, qui concerne des relations entre objets, peut s'améliorer au cours du temps.
- La reproduction a donc dû précéder la réplication

“Les organismes vivants comme pièges à information”

<http://www.normalesup.org/~adanchin>

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 ?

“Les organismes vivants comme pièges à information”

<http://www.normalesup.org/~adanchin>

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); accumuler l'information requiert un processus dépendant de l'énergie 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 pour mettre en œuvre ce processus ? Pouvons-nous trouver une source ubiquiste et stable d'énergie ?

"Les organismes vivants comme pièges à information"

<http://www.normalesup.org/~adanchin>

Un paléome diversifié

- **Paléome 1 (gènes essentiels)**
 - **Constructeur**: l'ADN spécifie les protéines qui forment la machine qui construit la cellule (reproduction)
 - **Répliqueur**: l'ADN spécifie les protéines qui répliquent l'ADN (réplication)
- **Paléome 2 (gènes persistants non-essentiels)**
 - Pérennisation de la vie, **dégradation dépendant de l'énergie**
 - Rustines métaboliques (« frustration » chimique)

Information

- 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

“Les organismes vivants comme pièges à information”

<http://www.normalesup.org/~adanchin>

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 ?

“Les organismes vivants comme pièges à information”

<http://www.normalesup.org/~adanchin>

Conjecture : les polyphosphates

- La synthèse et la dégradation des poly-P se retrouve codée dans les gènes non-essentiels persistants ; ce processus est encore peu connu et 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 peuvent être 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...

“Les organismes vivants comme pièges à information”

<http://www.normalesup.org/~adanchin>

Les gènes du démon de Maxwell

ATP
Poly(P)_n



ADP + Pi
Poly(P)_{n-1} + Pi <= dans le paléome



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



Les entités non fonctionnelles sont reconnues et dégradées

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

“Les organismes vivants comme pièges à information”

<http://www.normalesup.org/~adanchin>

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

“Les organismes vivants comme pièges à information”

<http://www.normalesup.org/~adanchin>

Mutations adaptatives

Construction de bactéries "intelligentes"

Mises en croissance sur un milieu limitant en éléments nutritifs. Forment des colonies de 10^8 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**



A. Sekowska

Mutations adaptatives et CVNC

- ⇒ **Prédiction** : les mutations adaptatives sont des créations *de novo* d'information ; elles dépendent donc des gènes impliqués dans l'accumulation d'information
- ⇒ Le statut des « cellules viables mais non cultivables » est prédit par le modèle (manque de potentiel régénératif)
- ⇒ L'accumulation d'information dépendant de l'énergie est aveugle ; elle ignore d'où l'information proviendra
- ⇒ L'information peut provenir d'une mémoire ; elle peut aussi être créée *de novo*

“Les organismes vivants comme pièges à information”

<http://www.normalesup.org/~adanchin>

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

“Les organismes vivants comme pièges à information”

<http://www.normalesup.org/~adanchin>

Le métabolisme précède la réplication

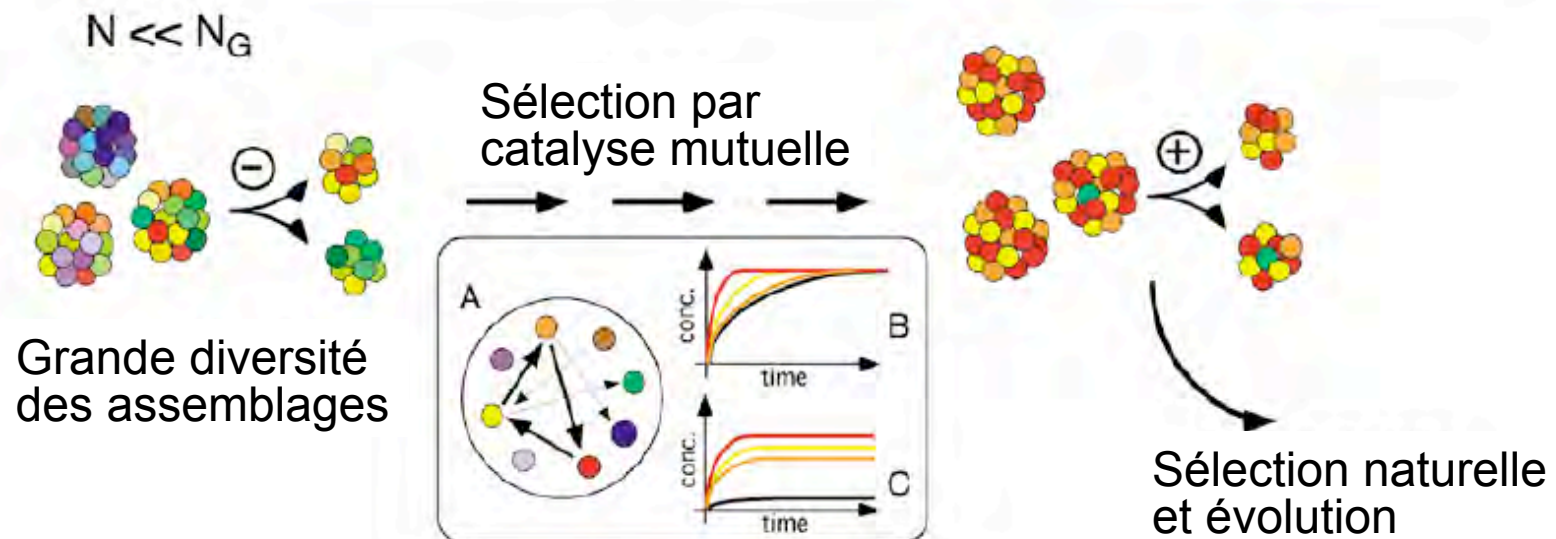
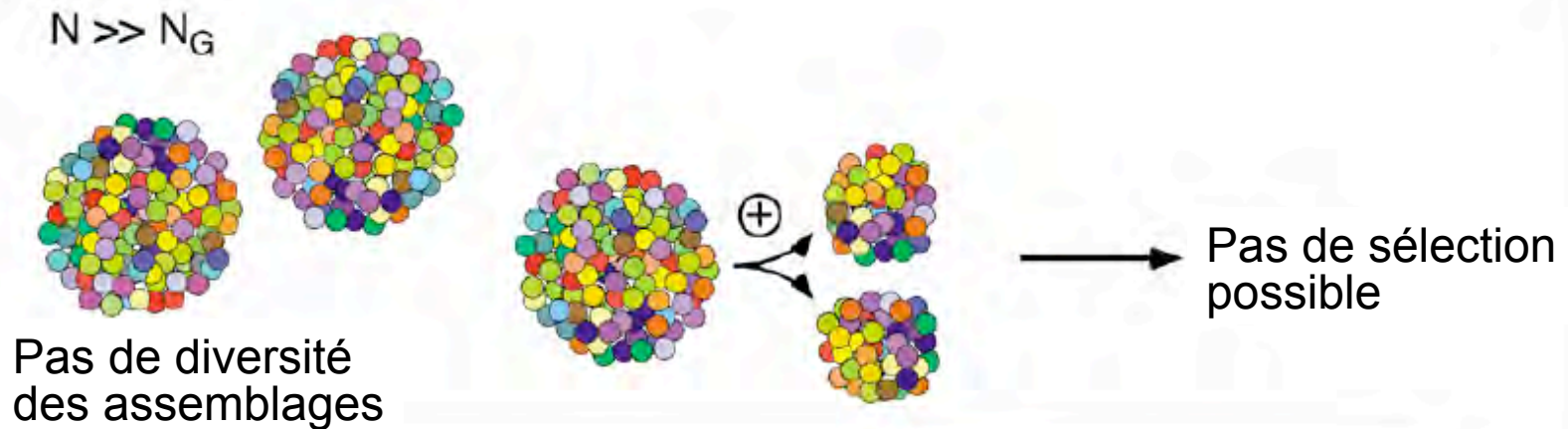
La **reproduction**, qui peut s'améliorer au cours du temps, doit précéder la réplication. Ce qui peut se reproduire est un ensemble de relations entre objets (chimiques). C'est exactement ce qui se produit au cours du métabolisme

Les scénarios qui permettent de reconstruire la machine cellulaire, fondés sur la reproduction, doivent en premier lieu explorer l'origine du métabolisme

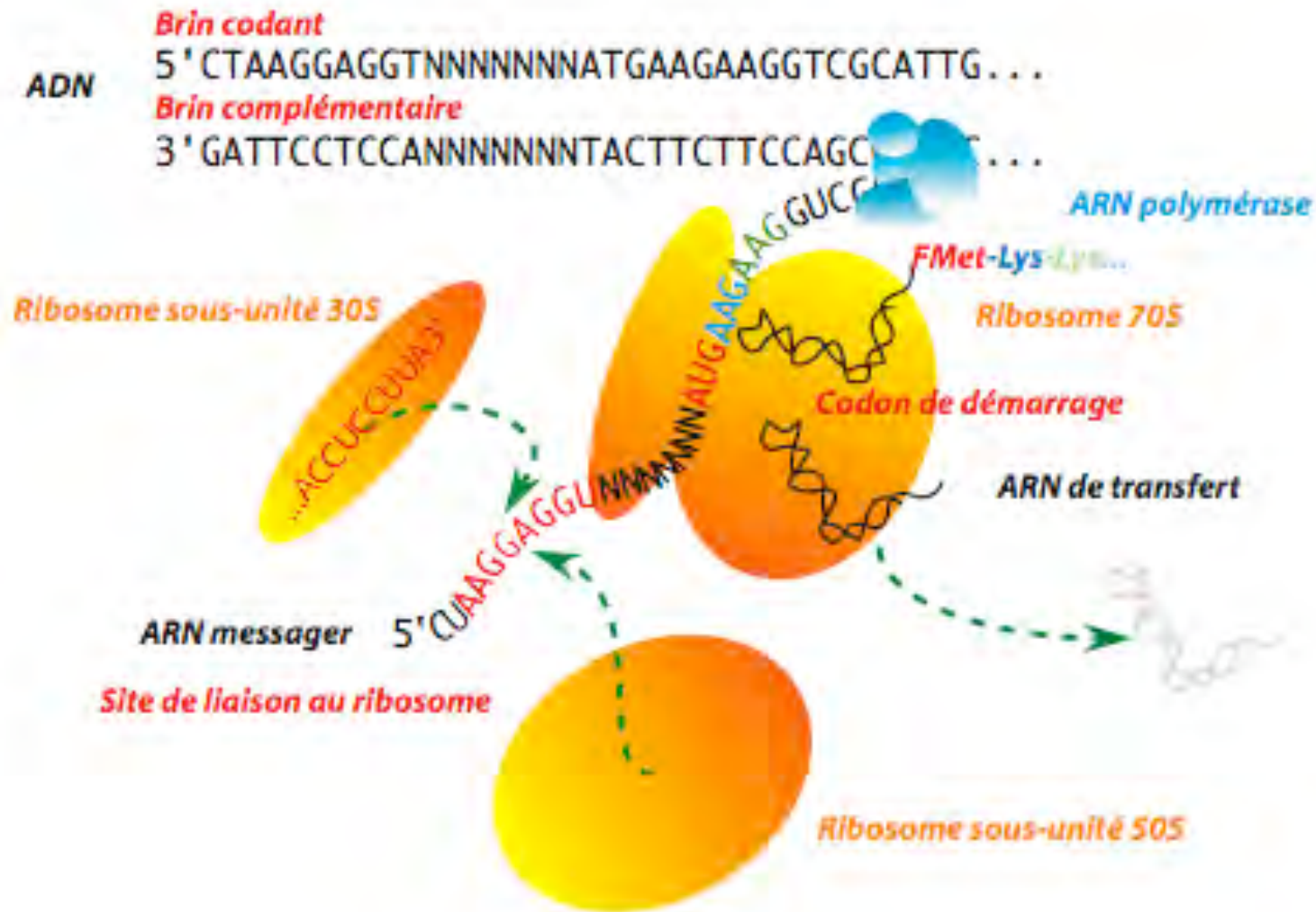
“Les organismes vivants comme pièges à information”

<http://www.normalesup.org/~adanchin>

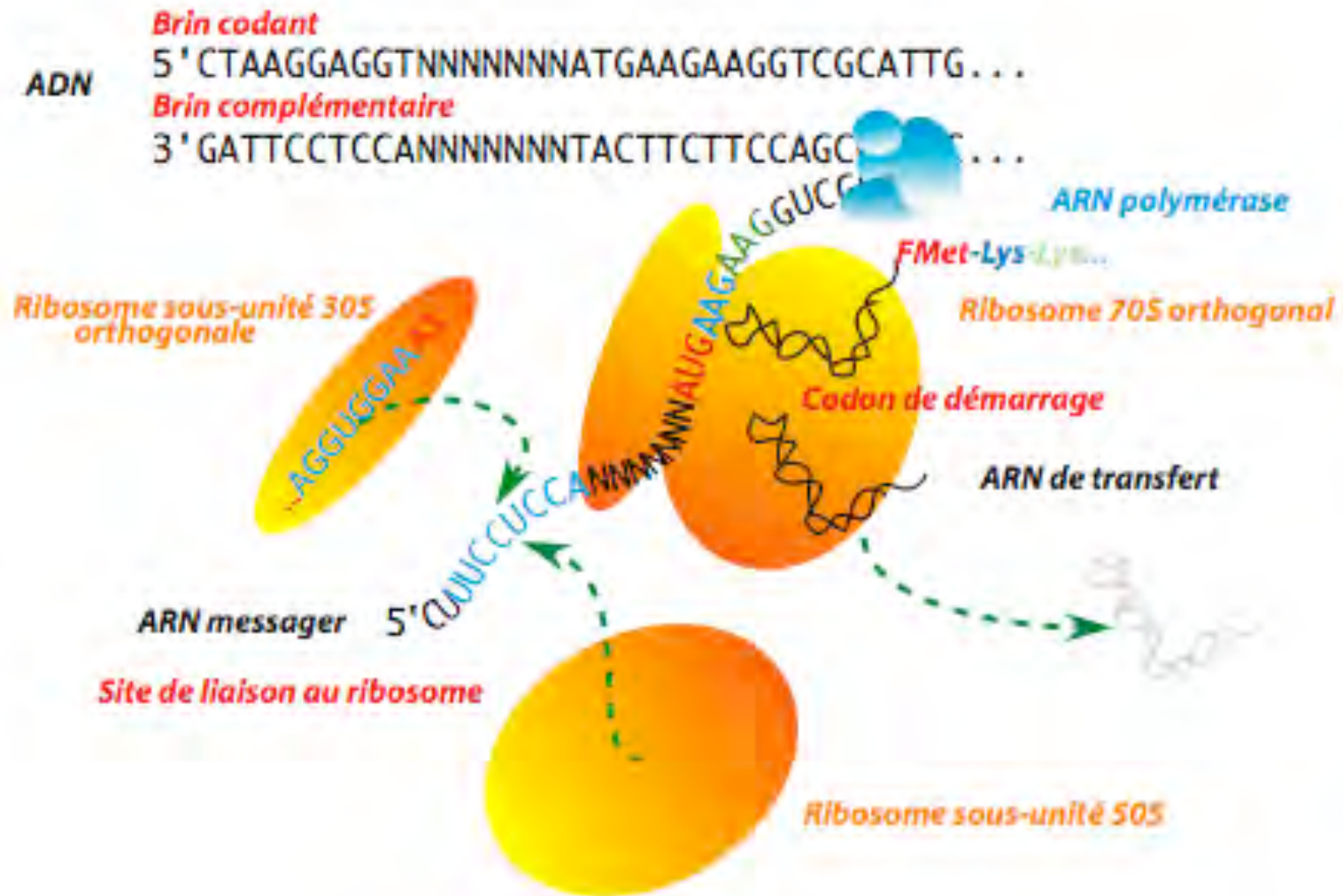
Développer la vie en tube à essais



Synthèse des protéines in vitro



Synthèse orthogonale



A network of orthogonal ribosome x mRNA pairs.

Rackham O, Chin JW. *Nat Chem Biol* 2005 1: 159-166.

La biologie synthétique

→ Comprendre et conserver les principes et les composants de base

- Encore réduire le plus petit génome : vers le génome minimal (*Mycoplasma genitalium*, Smith, Venter)
- Reconstruire ce qu'on connaît (repressilator, Leibler)
- Réécrire en simplifiant (génie logiciel) (T7, Endy ; promoteurs, Hwa)
- Reprogrammer : l'usine cellulaire (HCV, Liang ; carbone, Marlière)
- Reconstruire une cellule ancestrale « Jurassic Park moléculaire » (protéines ancestrales, Benner)
- ...

→ Conserver les principes et introduire une nouvelle chimie

- Changer les acides aminés (Cohen, Wong...)
- Changer le code génétique (Schultz, Marlière)
- Changer la chimie du support de l'hérédité (S-2L, PNA, TNA, GNA, etc)
- ...

La compétition iGEM au MIT

International Genetically Engineered Machine (iGEM)

→ Standardisation

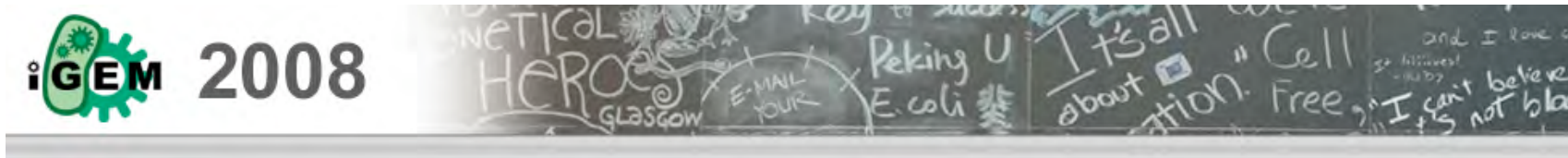
<http://2008.igem.org/>

→ Découplage

→ Abstraction

→ Design: Combiner des composants de base pour construire à volonté des circuits de régulation, à la manière de l'ingénierie électronique

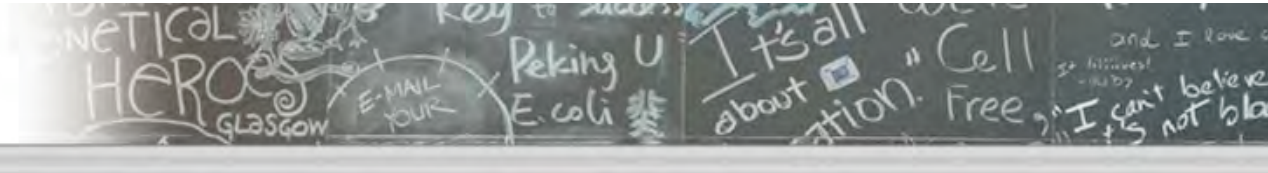
→ Évolution



“Les organismes vivants comme pièges à information”

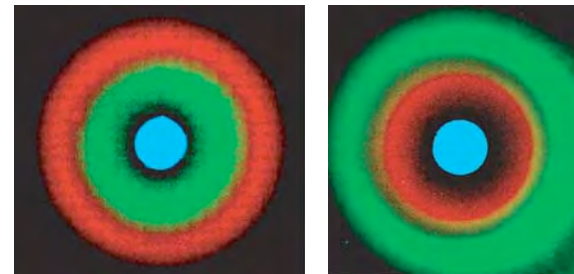
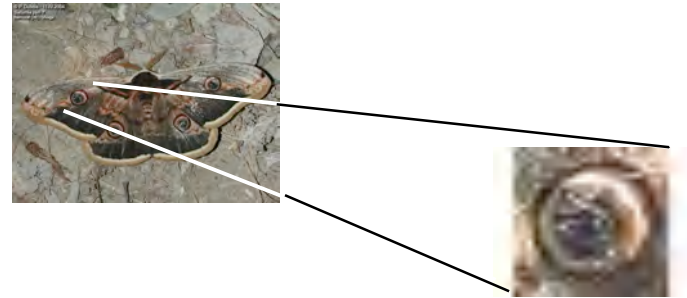
<http://www.normalesup.org/~adanchin>

MIT : iGEM et ses « biobriques »



iGEM (International Genetically Engineered Machines) **pose la question:**

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

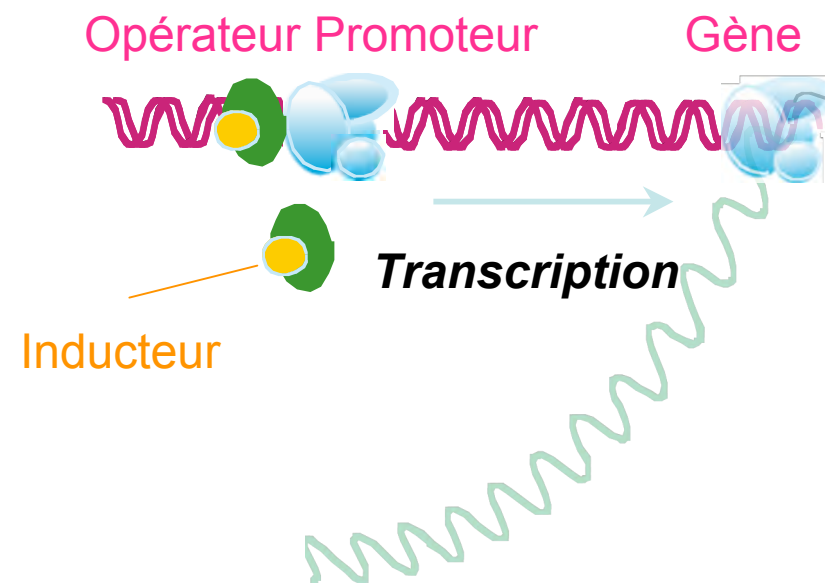
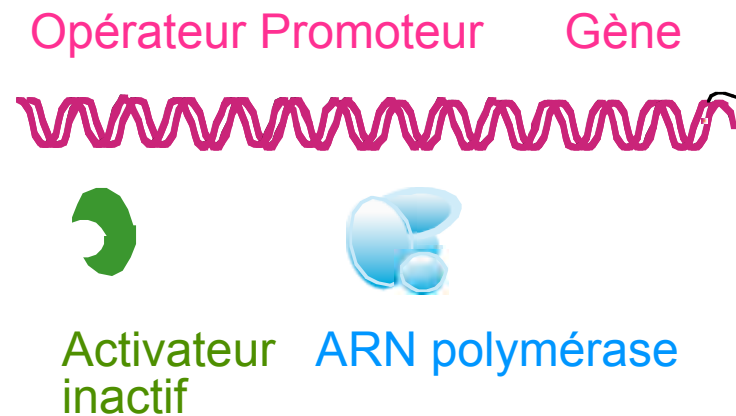
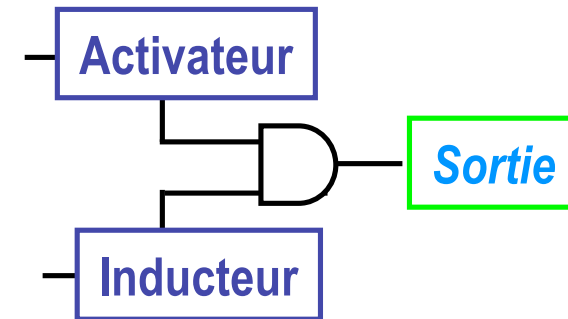


<http://2008.igem.org/>

Basu S, Gerchman Y, Collins CH, Arnold FH, Weiss R.
A synthetic multicellular system for programmed pattern
formation.
Nature. 2005 434:1130-1134

Activation de la transcription

Activateur	Inducteur	Sortie
0	0	0
0	1	0
1	0	0
1	1	1

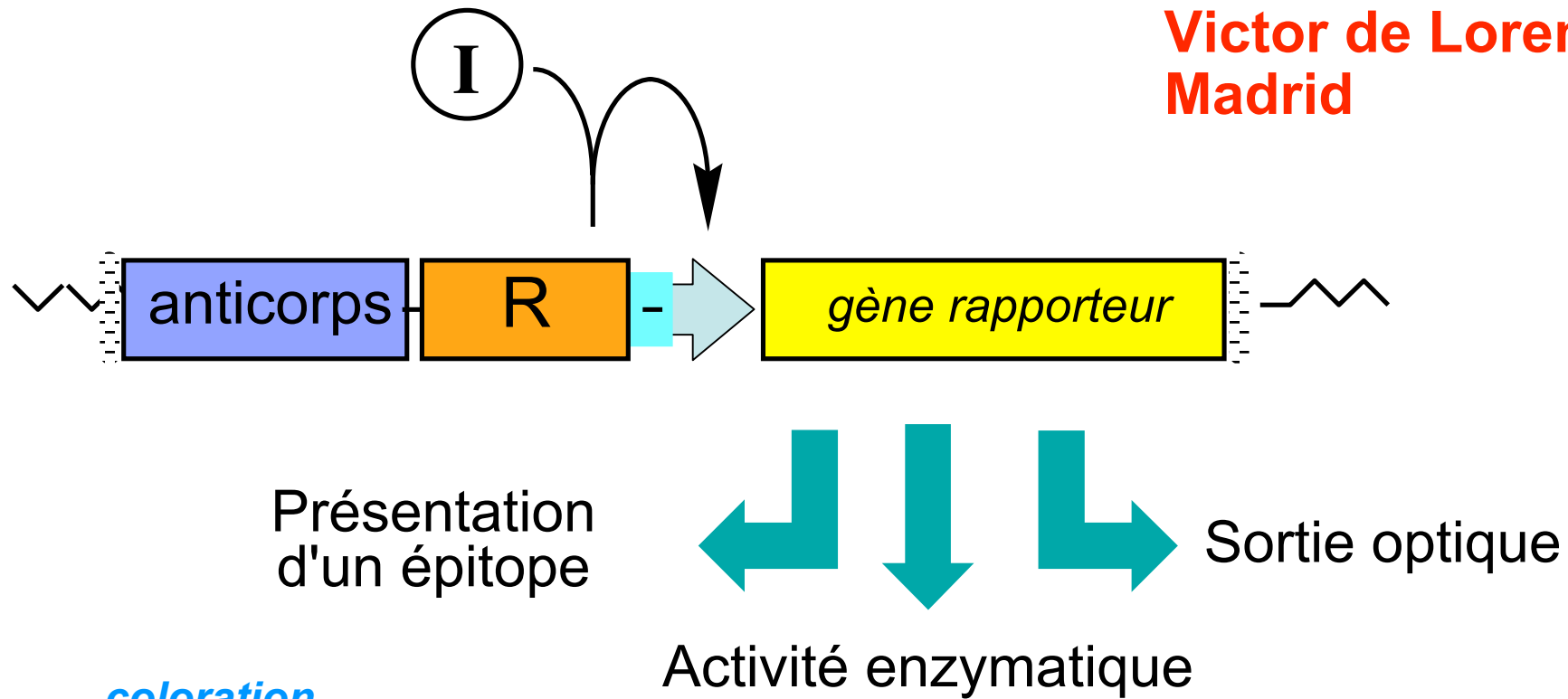


“Les organismes vivants comme pièges à information”

<http://www.normalesup.org/~adanchin>

Détecteurs associés à des promoteurs métaboliques

Victor de Lorenzo
Madrid



coloration



fluorescence



bioluminescence

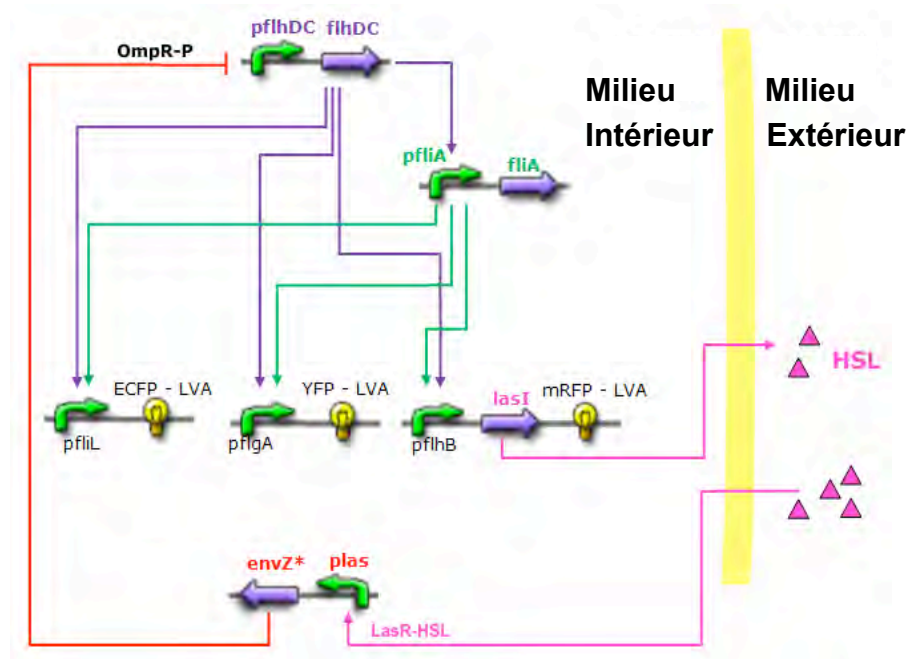


Une horloge bactérienne

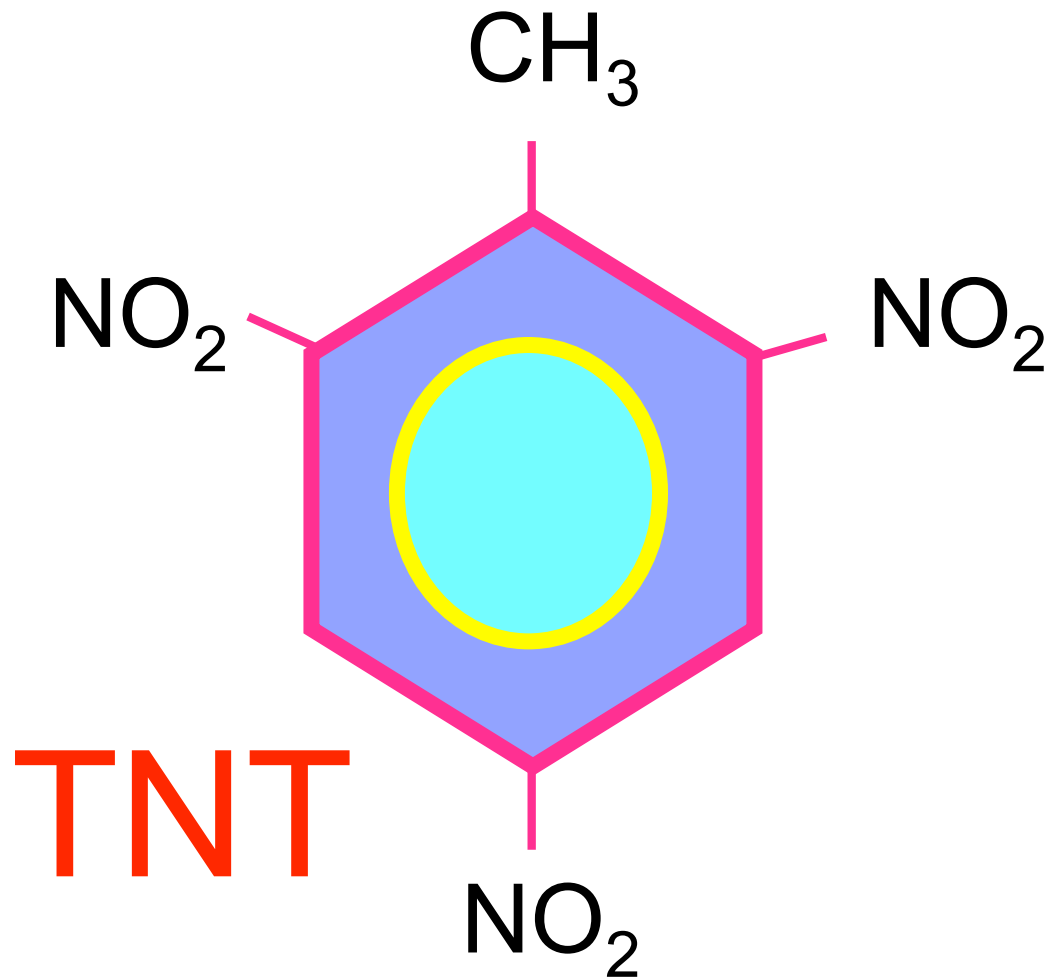
<http://2008.igem.org/Team:Paris>

BacteriO'Clock

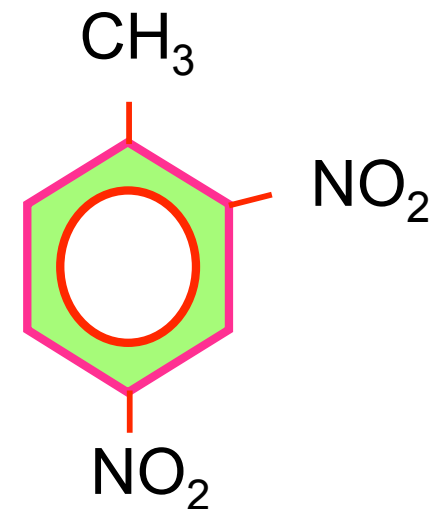
Un simple tube à essais contenant les bactéries reprogrammées pourrait donner l'heure, en changeant de couleur au cours du temps grâce à des oscillations créées par un circuit génétique.



Vers un détecteur d'explosifs

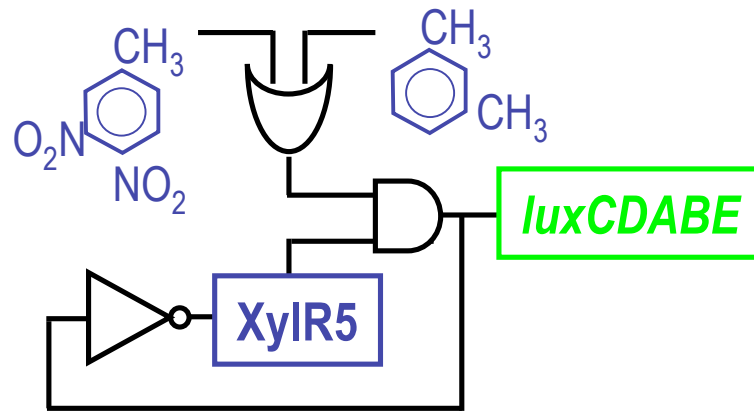


La molécule test
2,4-DNT

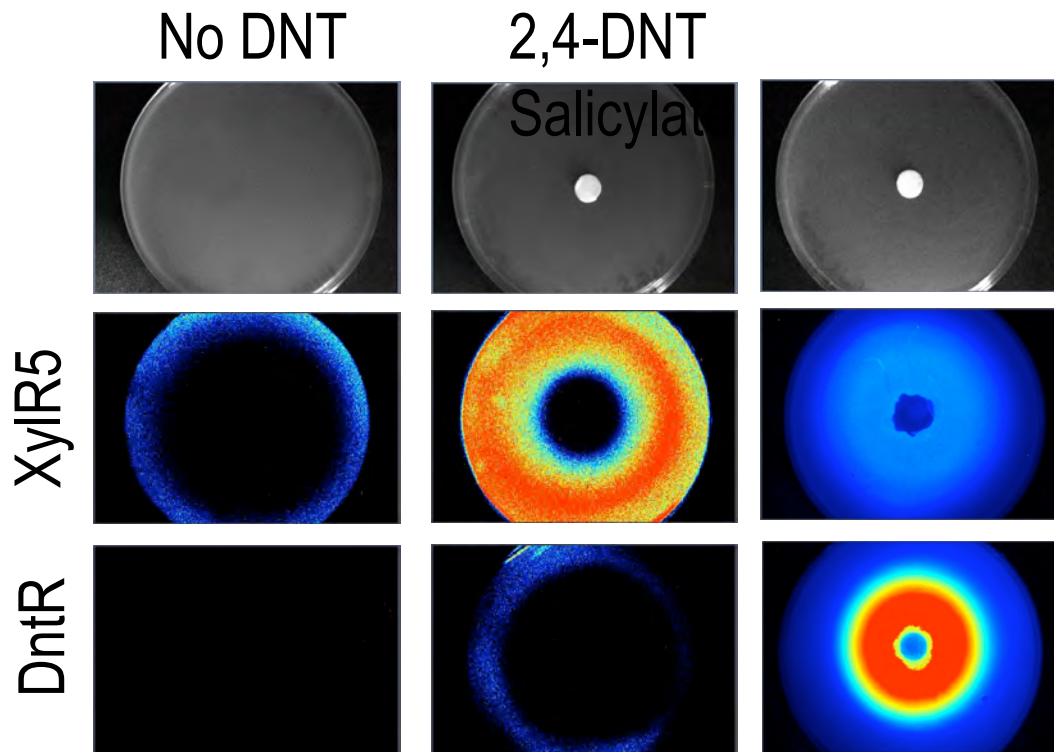


Travaux du groupe de Victor de Lorenzo à Madrid

Circuits formels et réalisations

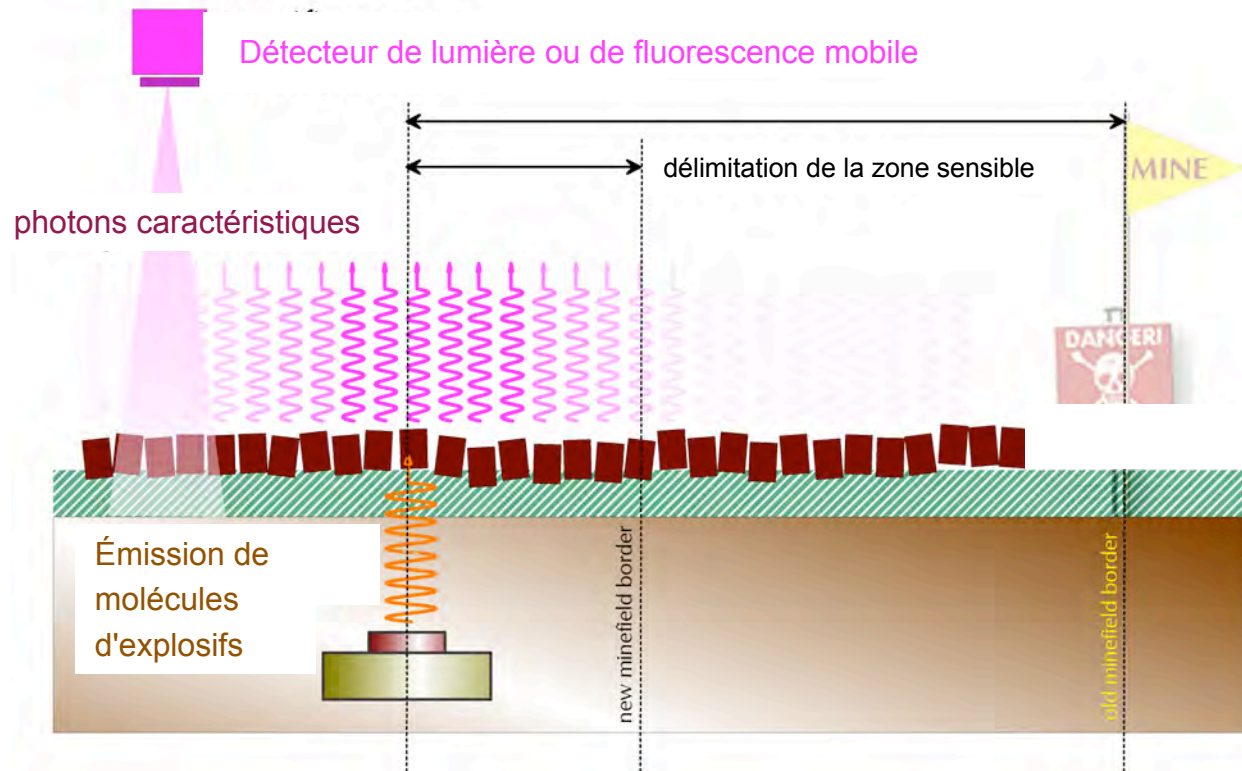


Le régulateur XylR5 reconnaît aussi bien le xylène que le dinitrotoluène. Il faut un inducteur **ET** le régulateur pour activer le promoteur *lux*. Le régulateur réprime sa propre synthèse, ce qui est représenté par une porte logique **NON**



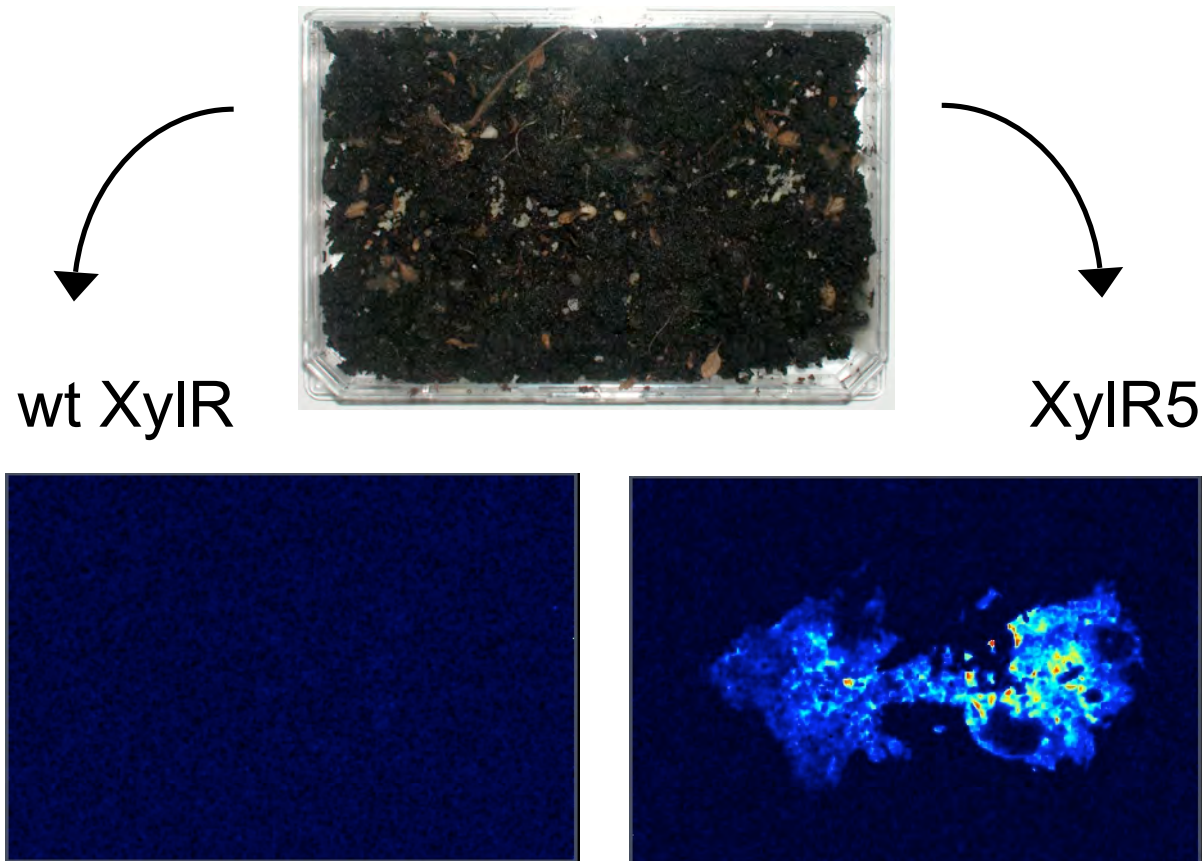
Mining logic gates in prokaryotic transcriptional regulation networks.
Silva-Rocha R, de Lorenzo V.
FEBS Lett. 2008 582: 1237-44

Construction d'un détecteur de mines



L'épandage de microbes sensibles aux émissions de molécules d'explosifs permet de délimiter la zone spécifique où se trouvent les mines par l'émission de lumière ou de fluorescence qui peut être détectée par un détecteur de lumière mobile (la détection des photons est très sensible)

Dispersion de bactéries sentinelles dans un sol modèle

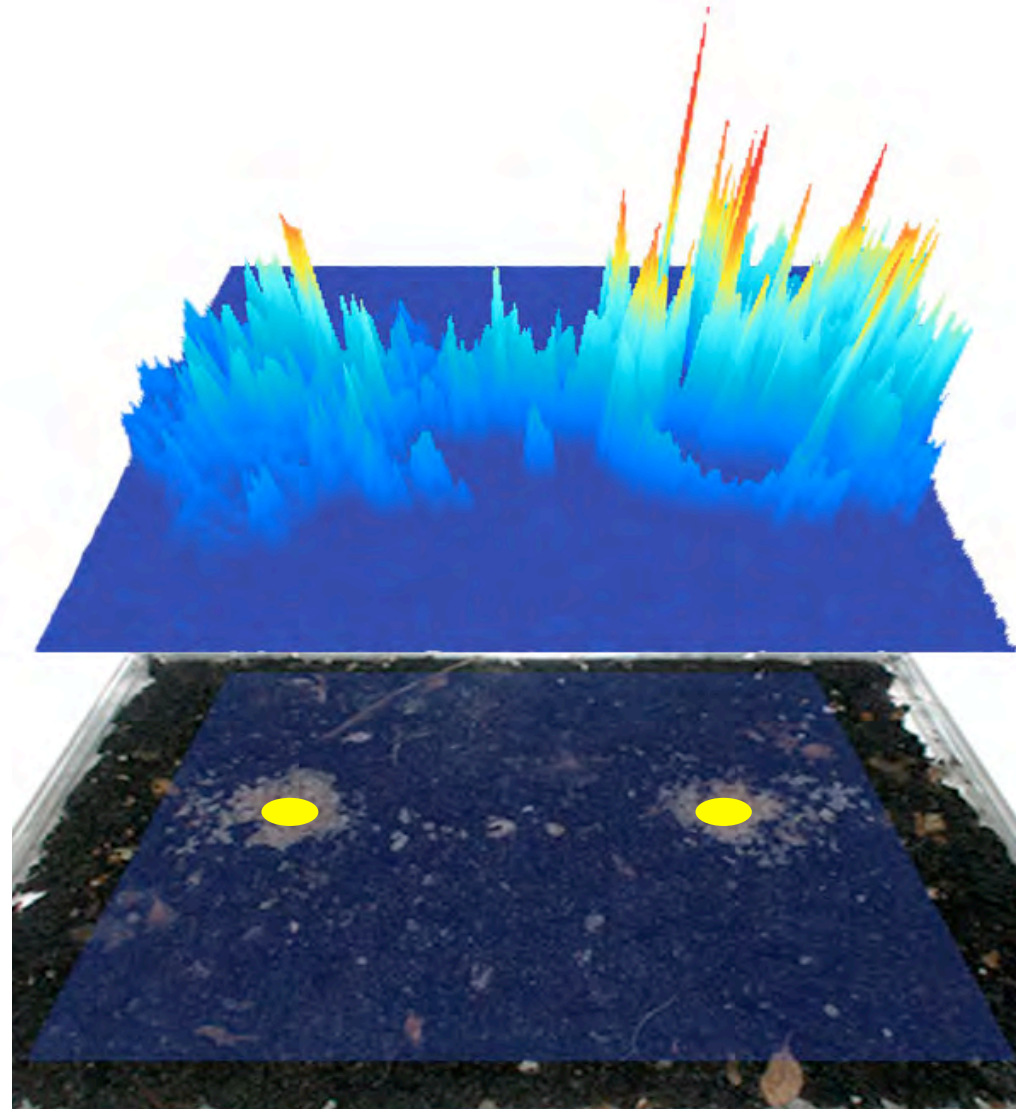


Stable implantation of orthogonal sensor circuits in Gram-negative bacteria for environmental release.
de las Heras A, Carreño CA, de Lorenzo V
Environ Microbiol (2008) **10**: 3305-3316

“Les organismes vivants comme pièges à information”

<http://www.normalesup.org/~adanchin>

Détection de dinitrotoluène dans un sol modèle



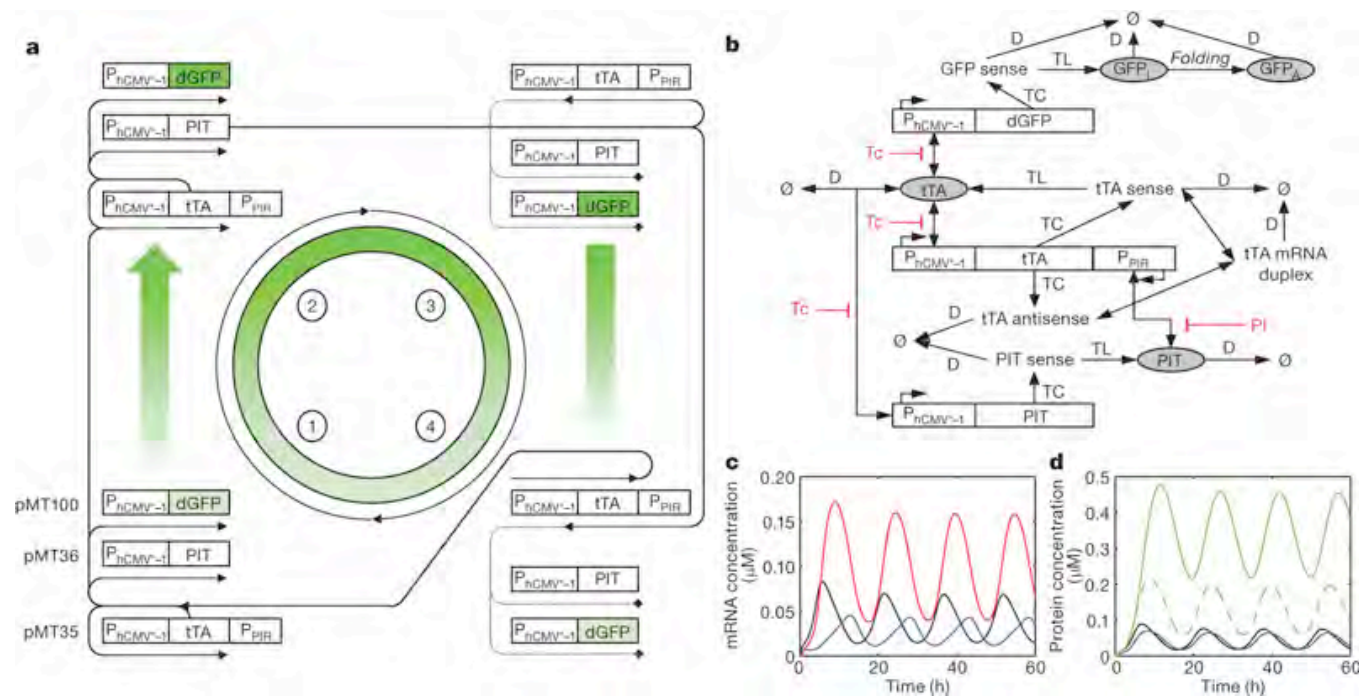
Stable implantation of orthogonal sensor circuits in Gram-negative bacteria for environmental release.

de las Heras A, Carreño CA, de Lorenzo V

Environ Microbiol (2008) **10**: 3305-3316

synthétique dans les cellules de ma

Horloge moléculaire dans des cellules de mammifères et prédiction des oscillations produites à partir d'ARN régulateurs



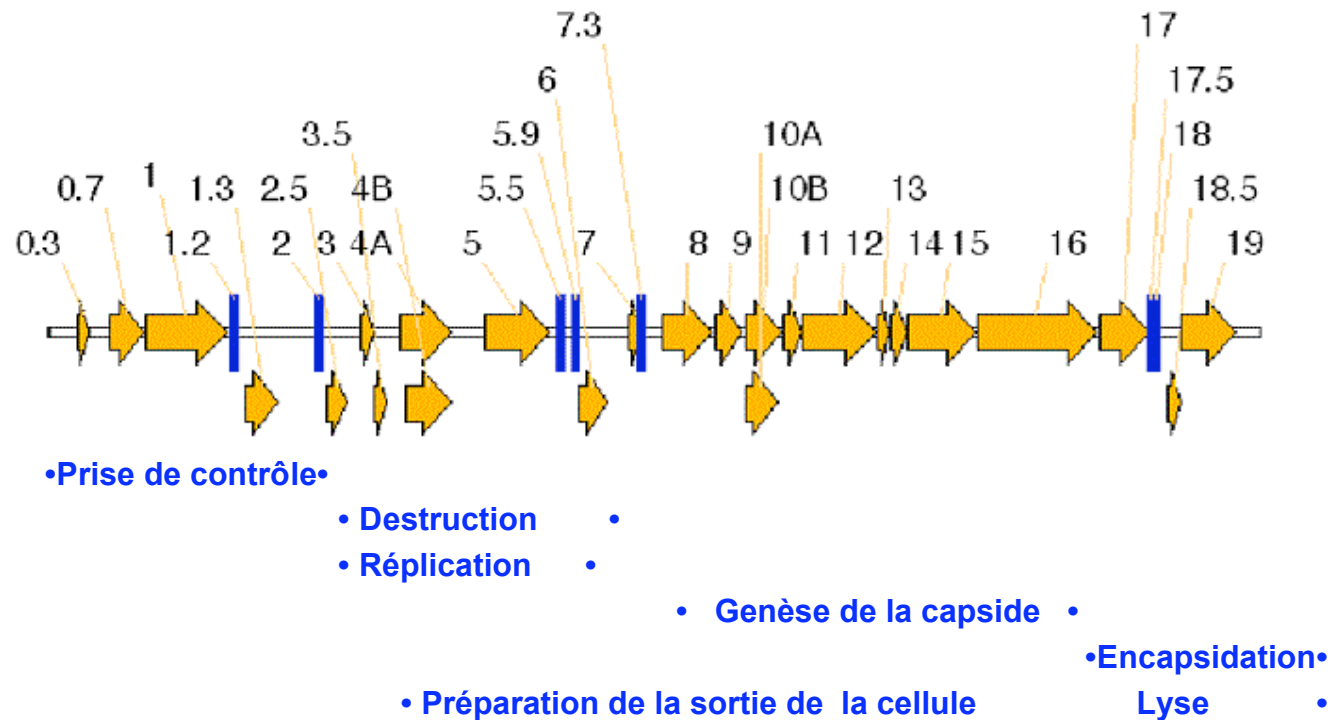
M Tigges *et al.* *Nature* **457**, 309-312 (2009) doi:10.1038/nature07616

nature

Reprogrammer le bactériophage T7



Les gènes connus du bactériophage T7



Drew Endy: <http://online.kitp.ucsb.edu/online/infobio01/endy/>

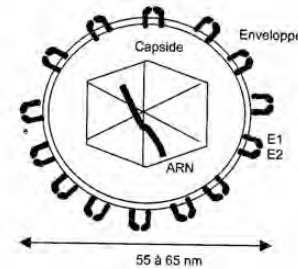
Reprogrammer le bactériophage T7



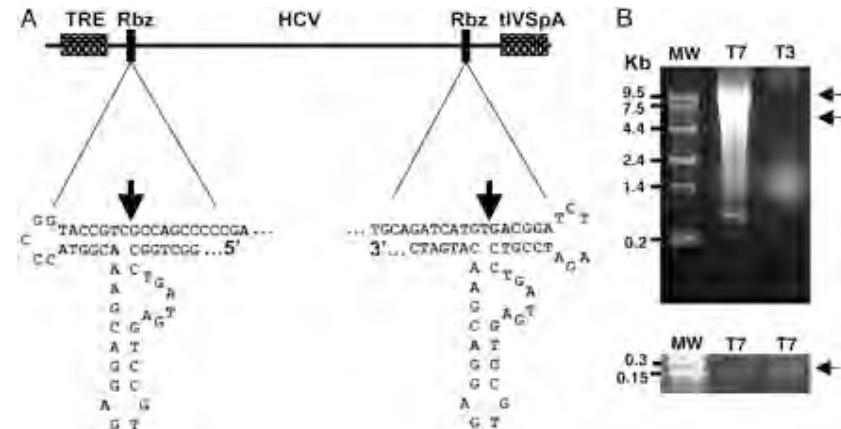
- Les zones de contrôle des différentes phases sont partiellement superposées
- Elles ne sont pas standardisées
- On sait les redessiner selon des principes d'ingénierie, et tester par des modèles mathématiques leurs propriétés
- Les modèles permettent de prédire un comportement du phage synthétique semblable à celui du phage naturel
- L'expérience montre que 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

Reprogrammer un virus, pas encore une cellule...

- Le virus de l'hépatite C comporte un **ARN** à brin unique de 9600 nucléotides, codant directement une protéine multifonctionnelle de 3010 acides aminés qui est secondairement scindée en au moins 10 protéines tardives de maturation virale. **Fragile, il se réplique mal**
- Par génie génétique, Liang et collègues ont construit un **ADN** autorépliatif codant l'ARN du virus, entouré de deux régions capables sous forme d'ARN de se cliver spontanément en libérant l'ARN viral.
- Cette protection/déprotection, entièrement artificielle, donne lieu à la production continue de virus normalement infectieux**



Pr. Amine SLIM Laboratoire de Microbiologie - CHU Charles Nicolle Tunis



Heller T et al An in vitro model of hepatitis C virion production. Proc Natl Acad Sci U S A. 2005 102:2579-2583.

“Les organismes vivants comme pièges à information”

<http://www.normalesup.org/~adanchin>

Biosécurité et biosûreté

- Il faut distinguer les problèmes du risque associé à toute construction humaine destinée par principe à améliorer nos conditions de vie : c'est la **biosécurité**. Typiquement on utilise des explosifs pour extraire des minéraux dans les carrières.
- Et les problèmes liés à l'usage malfaisant de nos constructions : c'est la **biosûreté**. Typiquement les entreprises terroristes utilisent des explosifs pour terroriser en tuant.

“Les organismes vivants comme pièges à information”

<http://www.normalesup.org/~adanchin>

Il ne faut pas crier "au loup" sans raison

La vie évolue comme le ferait un avion dont les moteurs et les ailes sont modifiés **en vol**

La Biologie Synthétique (les OGM sont une première approche dans ce domaine) nous permet de prendre une attitude différente, celle de quelqu'un qui ne veut pas mettre sa vie en danger lorsqu'il fait des expériences

"Les organismes vivants comme pièges à information"

<http://www.normalesup.org/~adanchin>

Petit bagage : Nature et Artifice

→ Jadis l'Homme cherchait à protéger la Nature contre lui-même ; l'Artifice était perçu comme dangereux

→ Mais la Nature est par essence imprévisible et inamicale

→ Une façon de contrôler les organismes vivants est de **réduire leur capacité à évoluer**

→ La domestication des plantes et des animaux marque un nouvel état de l'humanisation, qui commence au Néolithique avec le contrôle de la Nature grâce à l'Artifice

→ Les OGM sont le dernier avatar de la domestication

→ Ce qui est Naturel est souvent bien plus dangereux (parce qu'incontrôlable) que ce qui est Artificiel

→ La peur de l'Artifice conduit à la pensée très dangereuse que les Plantes Génétiquement Modifiées sont dangereuses alors que les animaux Génétiquement Modifiés ne le sont pas

Une cellule synthétique ?

- La vision d'ingénieur de la BS **exclut par définition que les cellules synthétiques innovent**
- Il est possible d'écarter dans nos constructions les gènes qui permettent l'accumulation d'information
- Il s'en suit que les cellules ainsi construites, comme les usines, vieilliront et devront être renouvelées
- **L'avantage en termes de biosécurité est que ces cellules seront par construction incapables d'innover**

L'information est contagieuse

- Toute vérité n'est pas bonne à dire
- L'OMS tente d'éradiquer les maladies contagieuses spécifiques à l'homme, essentiellement virales
- Cette éradication repose sur la vaccination et suppose que la majorité de la population mondiale se vaccine
- Elle suppose aussi qu'on ne connaisse pas assez bien l'organisme pour ne pouvoir le reconstruire
- C'est ainsi que la connaissance du génome du virus de la variole a rendu son éradication (obtenue en 1977) inopérante... La séquence du virus peut en effet se propager par l'internet, par exemple...

Not every truth is good. The dangers of publishing knowledge about potential bioweapons. Danchin A. EMBO Reports (2002) 3: 102-104.

Remerciements

Le paléome

TingZhang Wang analyse du protéome et vieillissement

Gang Fang, Eduardo Rocha persistance des gènes

Isabelle Martin-Verstraete *et al* anabolisme des composés soufrés

Anne-Marie Gilles métabolisme des nucléotides

Undine Mechold, Sandra Cescau *and al* métabolisme de l'ARN

Le paléome et le cénome

Agnieszka Sekowska, Conghui You, Andrew Martens récupération du soufre et rustines métaboliques

Le cénome

Evelyne Krin, Evelyne Turlin, Sabina Chalabaev, Jean-Baptiste Masson

Autres collaborations

Francis Biville, Massimo Vergassola et collaborateurs

GenoScope

Beijing Genome Institute, Fudan University, The University of Hong Kong

Hong Kong University of Science and Technology

