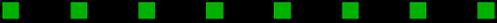




les gènes du démon de maxwell





antoine danchin
唐善・安東



école des mines de paris
18 janvier 2010

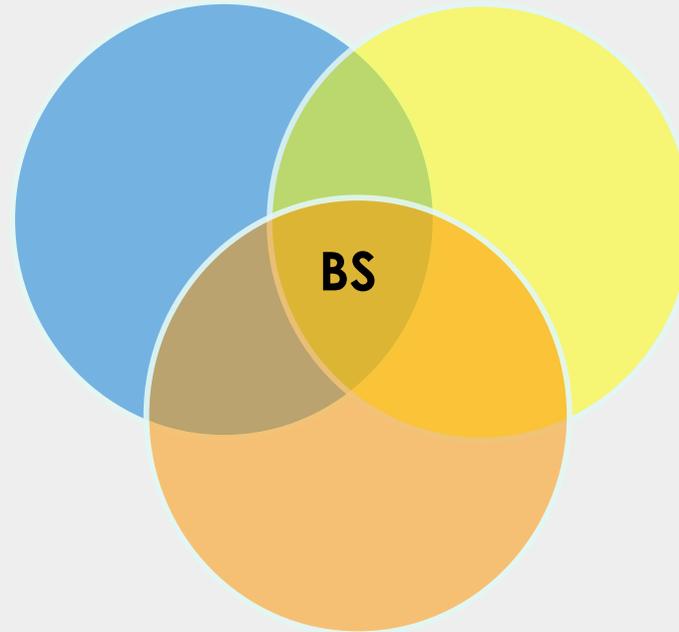


le projet de la biologie synthétique



position 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

- un premier objectif de la BS est de **reconstruire la vie**, pour explorer si nous comprenons ce qu'elle est, en découvrant, à partir de nos échecs, ce qui manque
- 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 »**



I n f o r m a t i q u e a u s s i !

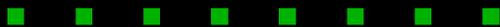
le projet RepRap (Replicating Rapid - prototyper, 2004) vise à créer une imprimante laser 3D auto - reproductrice :

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 de ce qu'on veut (re)construire.

2. conception

- concevoir des circuits génétiques capables d'homéostasie ou de comportements plus compliqués comme des oscillations, de la multistabilité, de la résistance au bruit, etc.

3. standardisation

- étape importante pour le futur de la BS, la standardisation est à la base du 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

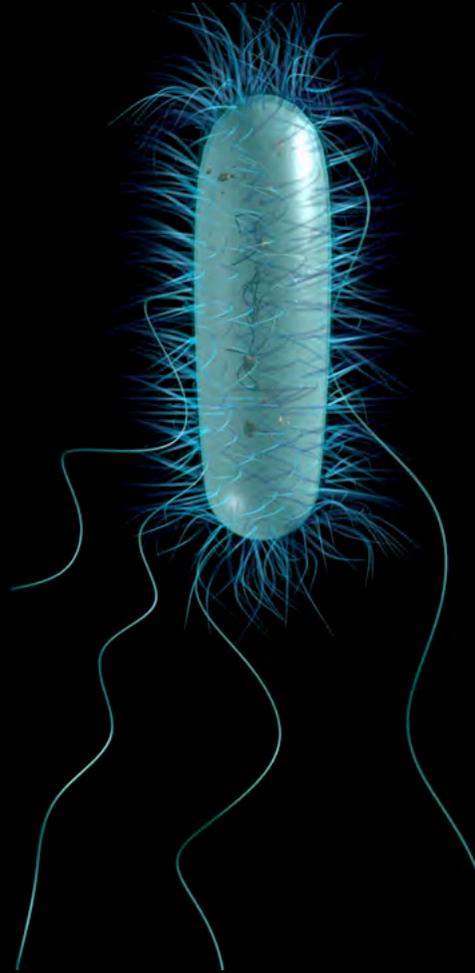




la logique du vivant



le chassis universel : *escherichia coli*



principes du génie génétique

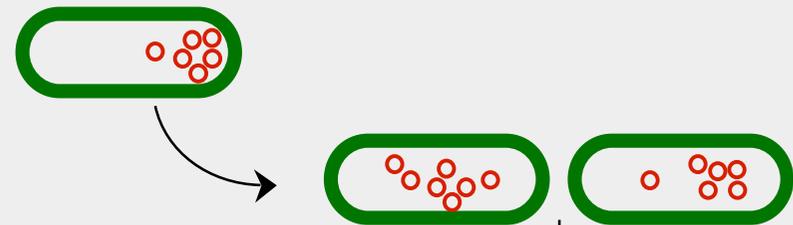
construction d'un vecteur chimérique



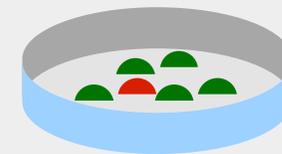
transformation de l'hôte



multiplication et division de l'hôte

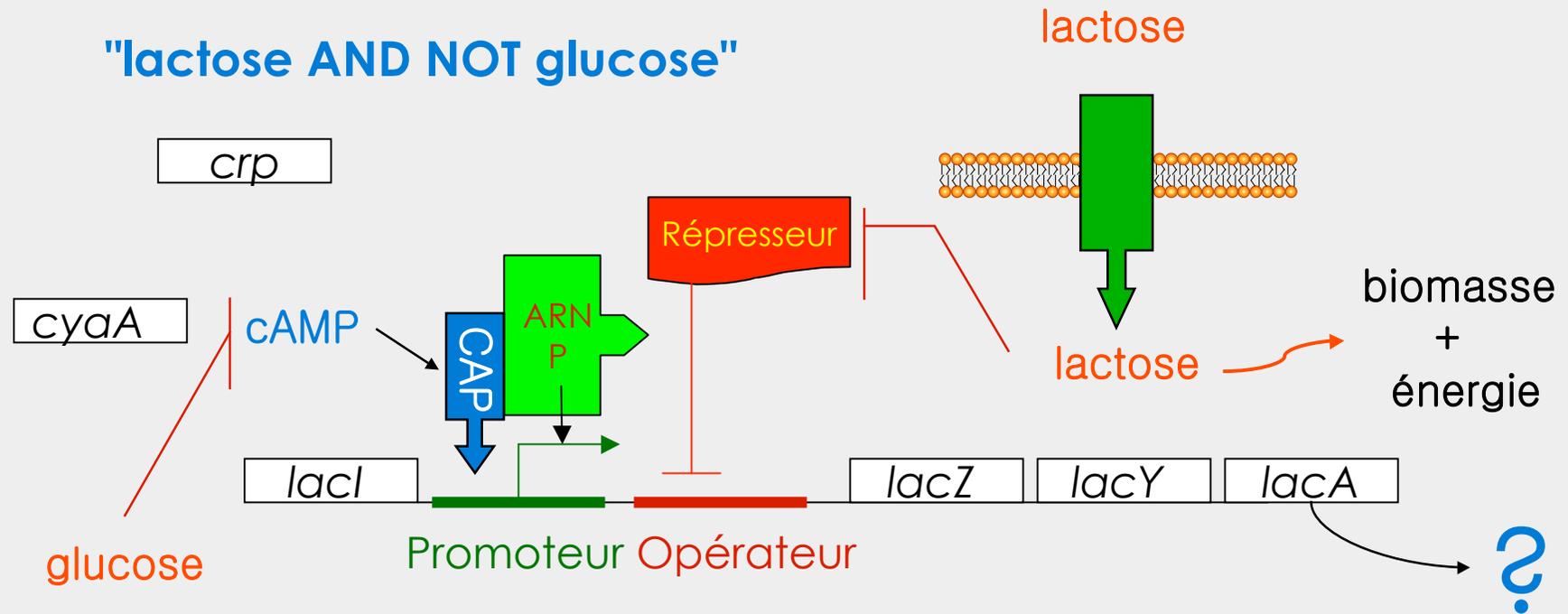


criblage ou sélection identifiant un clone positif



l'opéron lactose, circuit génétique modèle

"lactose AND NOT glucose"



François Jacob, David Perrin, Carmen Sanchez, Jacques Monod

L'opéron : groupe de gènes à expression coordonnée par un opérateur [C. R. Acad. Sci. Paris 250 (1960) 1727-1729]



régulation de l'expression génétique

contrôle transcriptionnel

démarrage et fin de la transcription

vitesse de transcription

durée de vie du messenger (dégradosome)

contrôle traductionnel

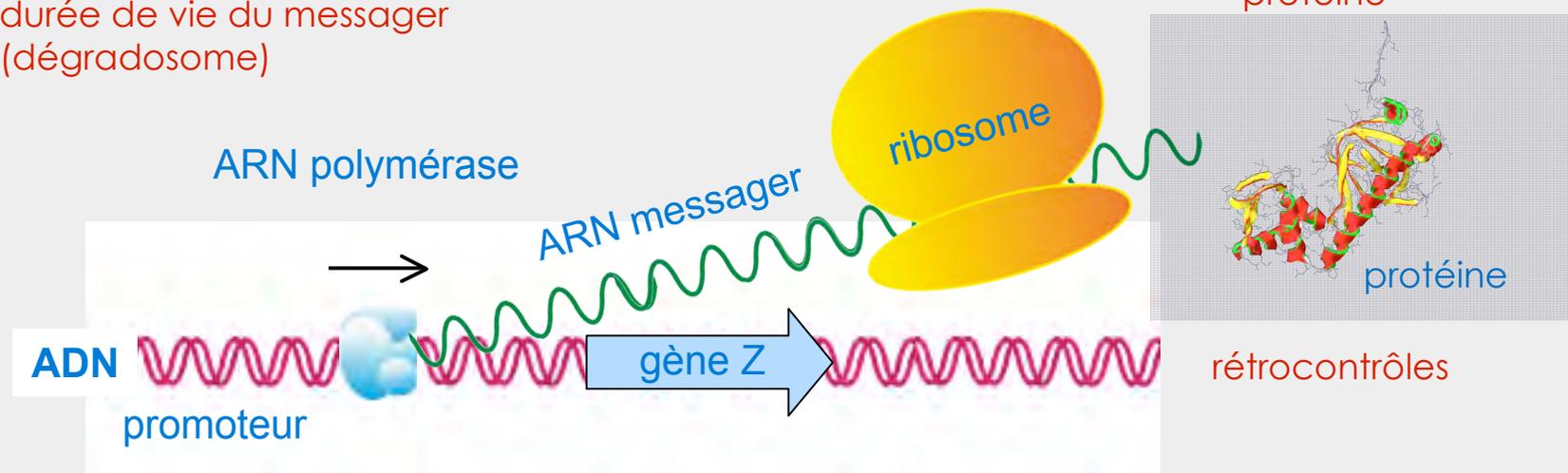
démarrage et fin de la traduction

vitesse de traduction

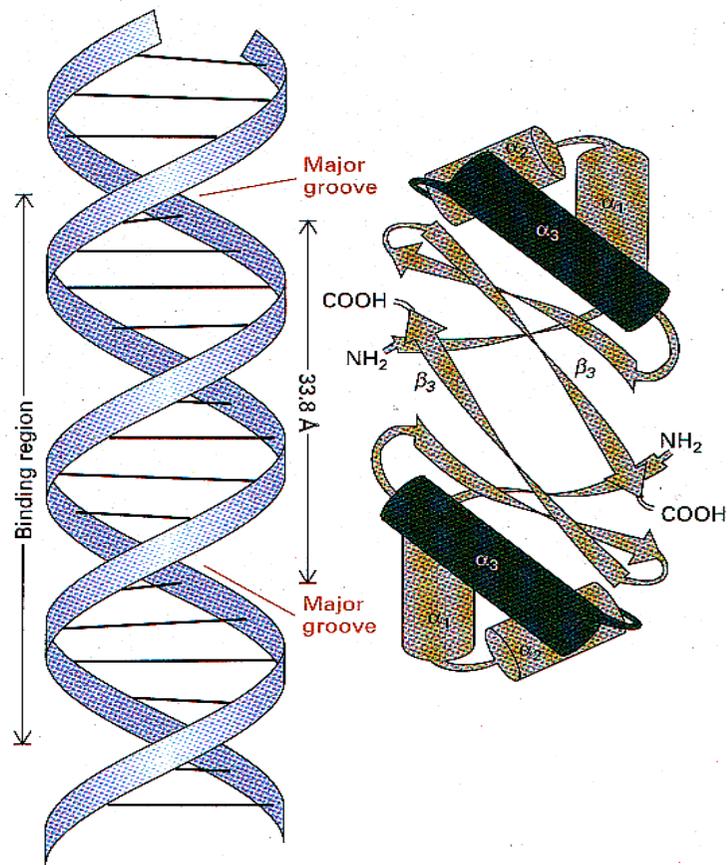
contrôle post-traductionnel

modification chimique de la protéine

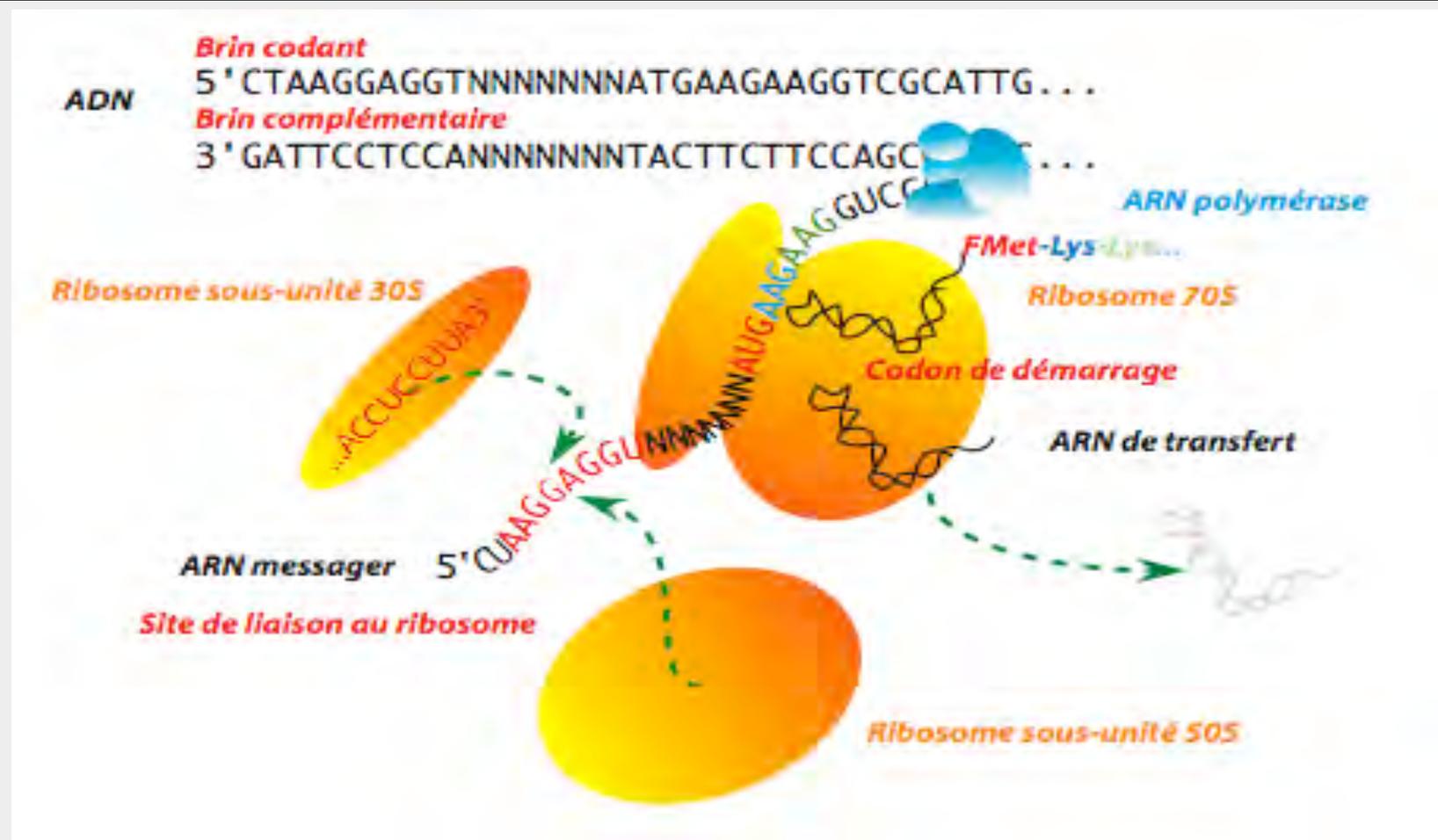
durée de vie de la protéine



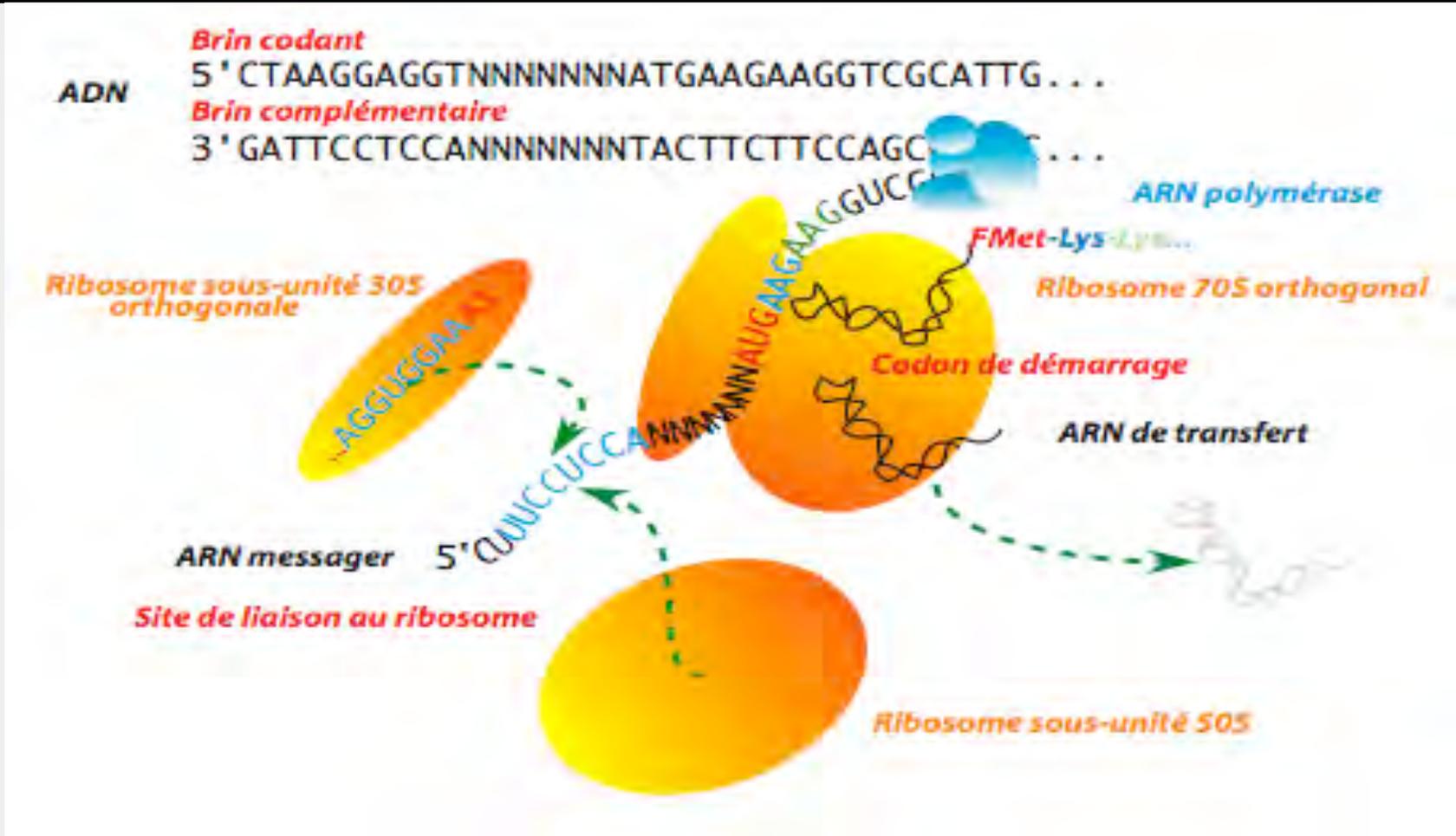
liaison physique présence / absence



synthèse des protéines



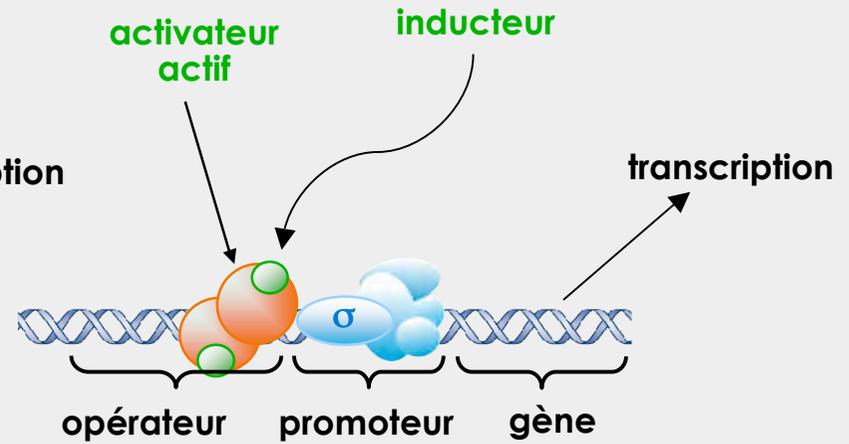
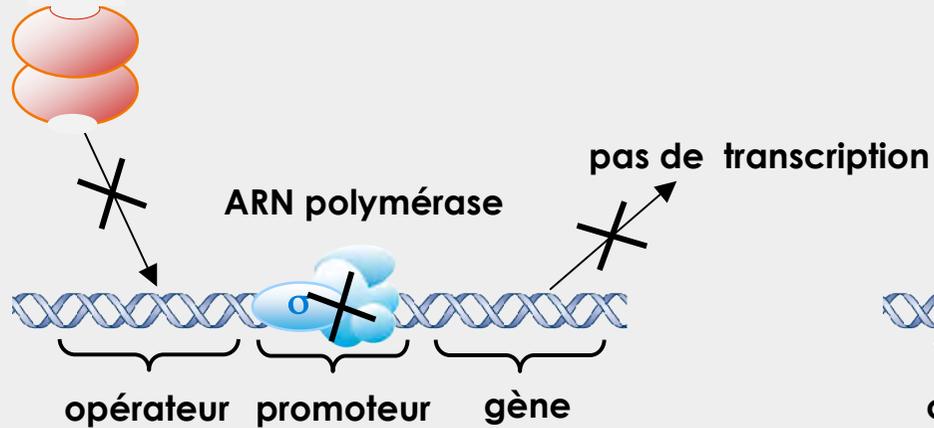
synthèse orthogonale



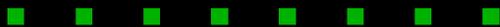
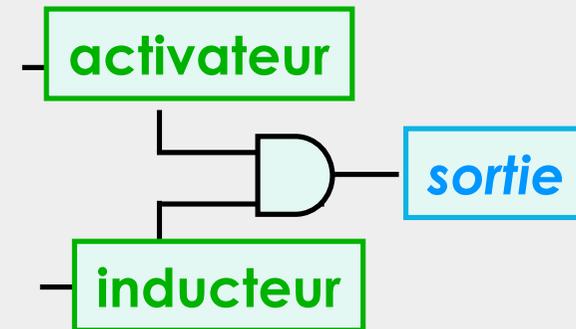
A network of orthogonal ribosome x mRNA pairs.
 Rackham O, Chin JW. Nat Chem Biol 2005 1: 159-166.

régulation par activation

activateur
inactif

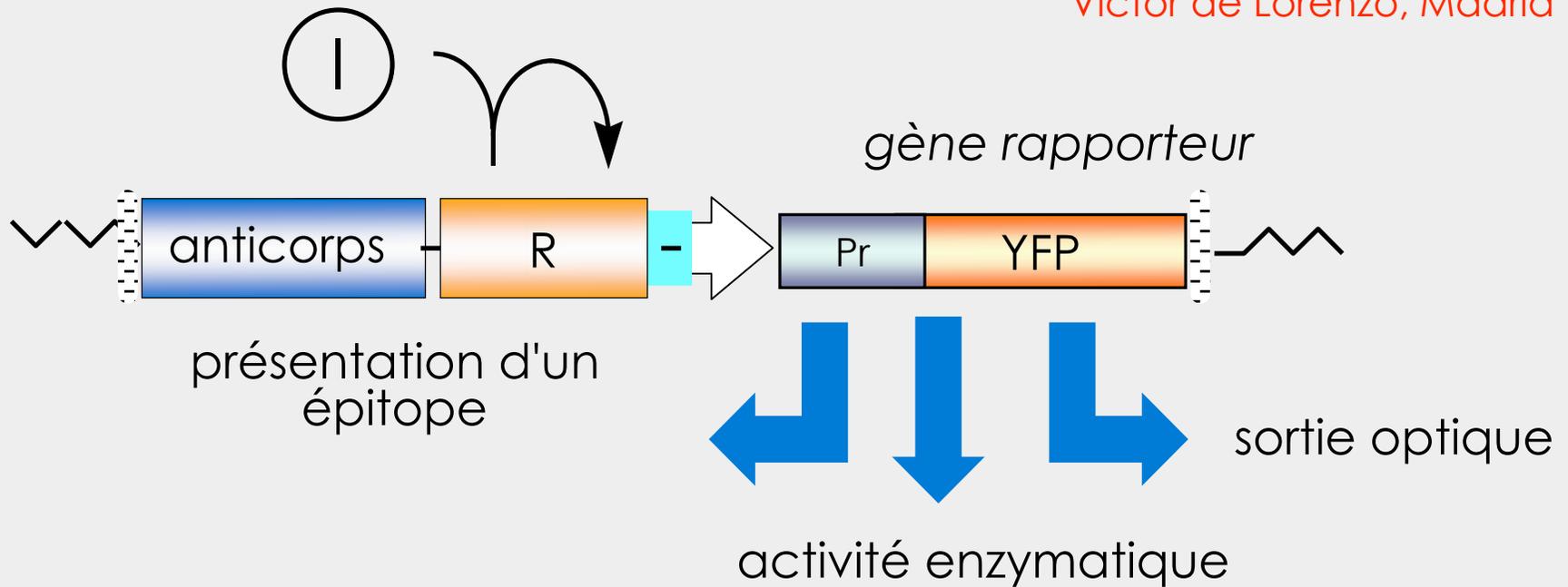


activateur	inducteur	sortie
0	0	0
0	1	0
1	0	0
1	1	1



détecteurs associés à des promoteurs métaboliques

Victor de Lorenzo, Madrid



coloration



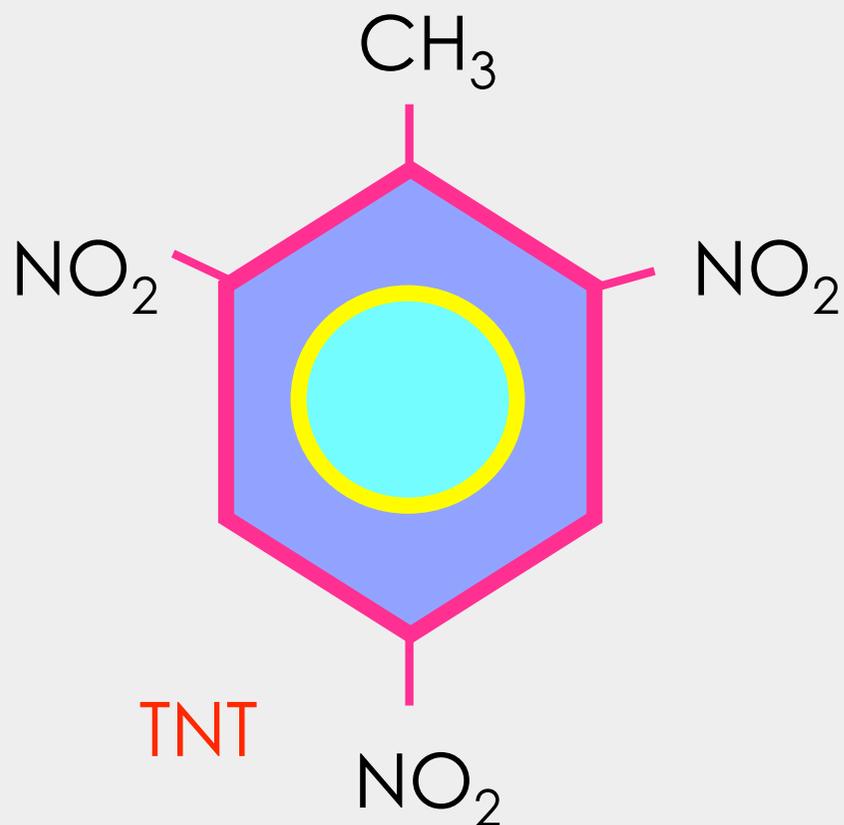
fluorescence



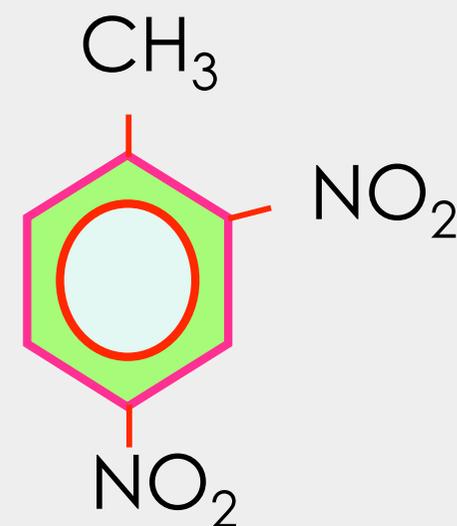
bioluminescence



vers un détecteur d'explosifs

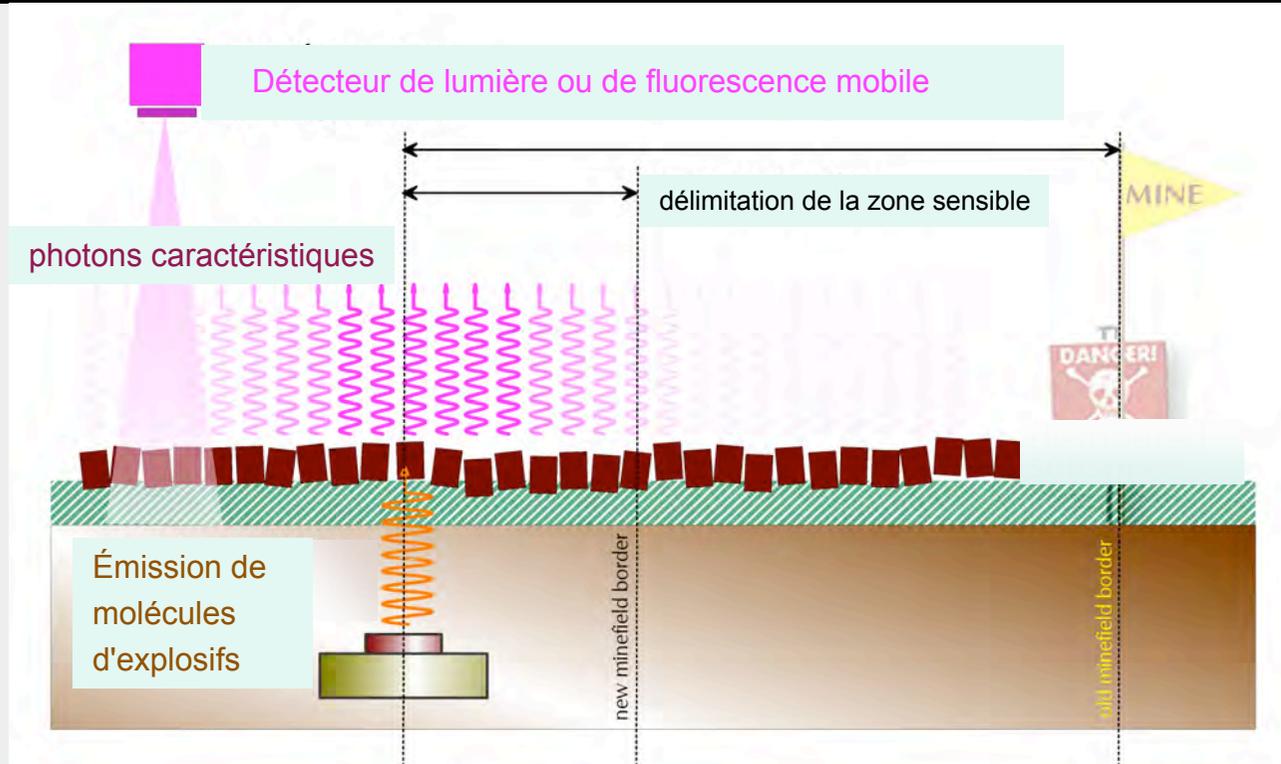


la molécule test
2,4-DNT



travaux du groupe de victor de lorenzo à madrid

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)

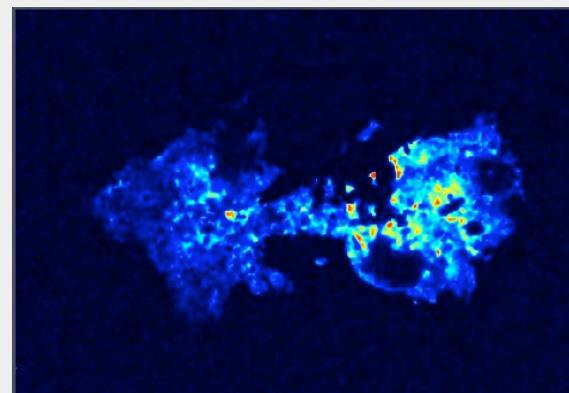
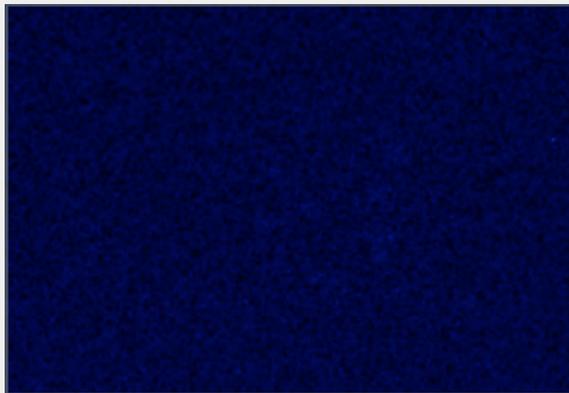


dispersion de bactéries sentinelles dans un sol modèle

wt XylR

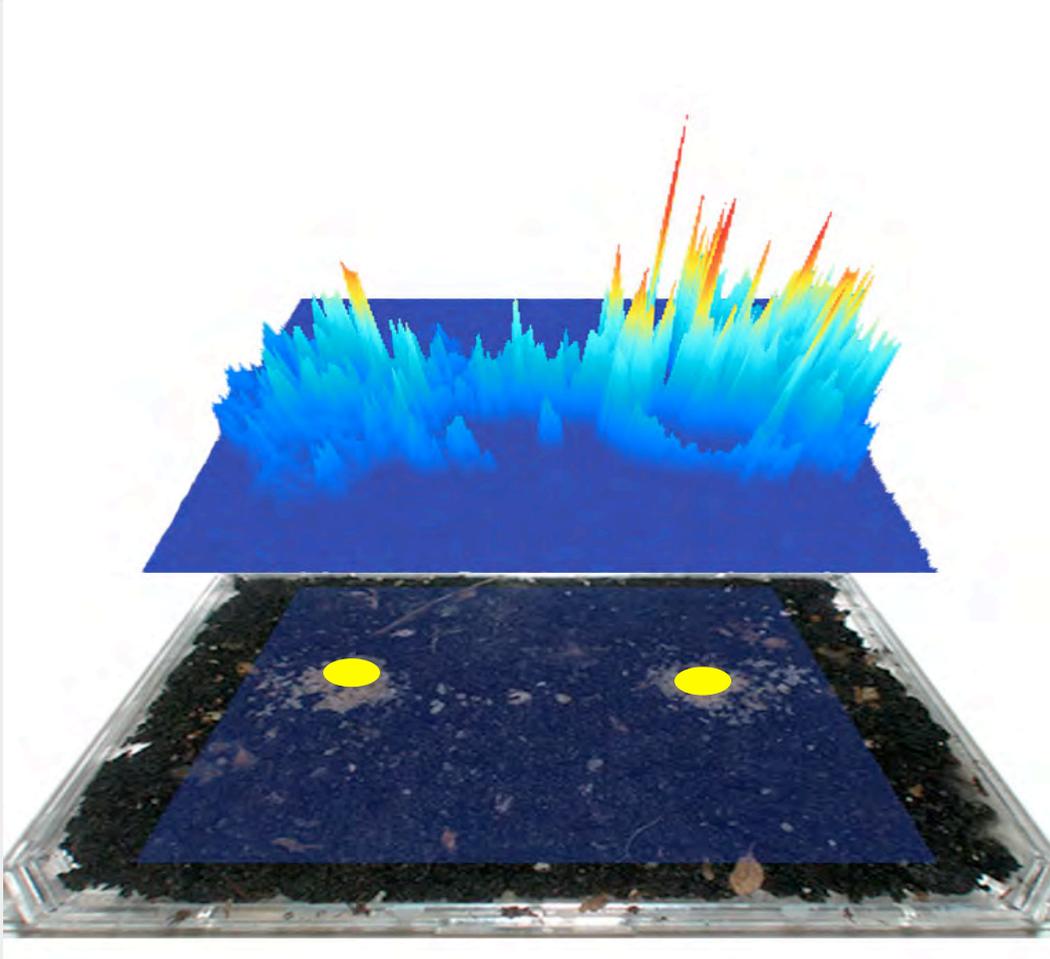


XylR5



■ ■ ■ ■ ■ ■ ■ ■
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

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

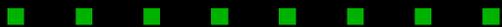
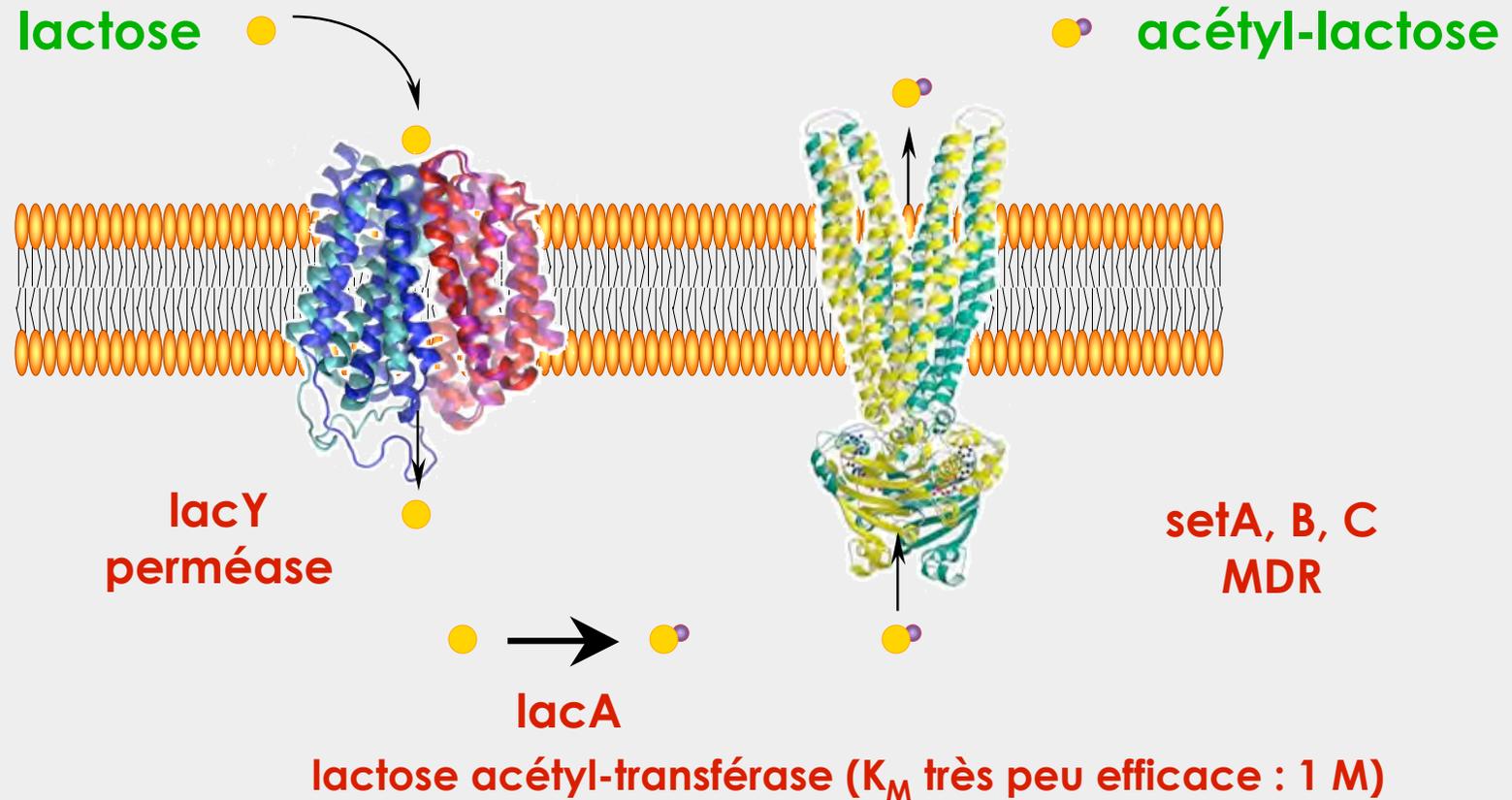




d i f f i c u l t é s



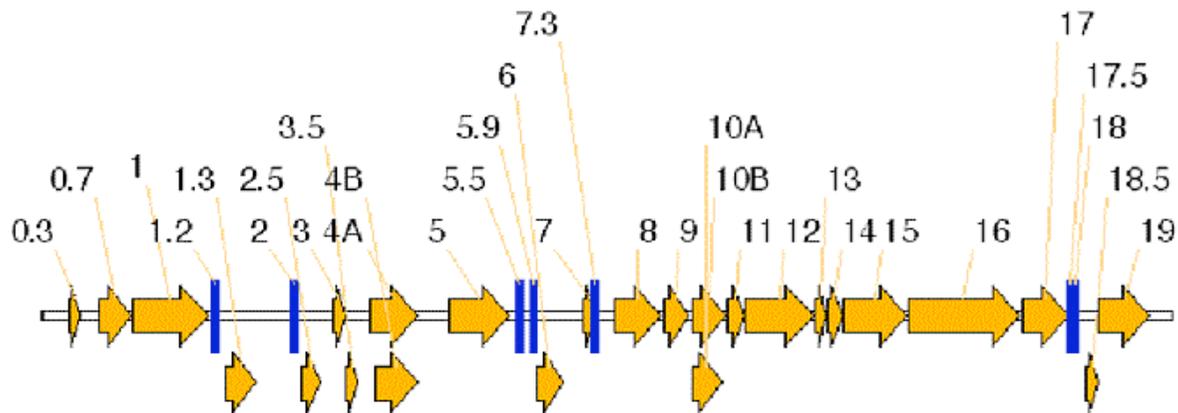
s o u p a p e d e s é c u r i t é



reprogrammer 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

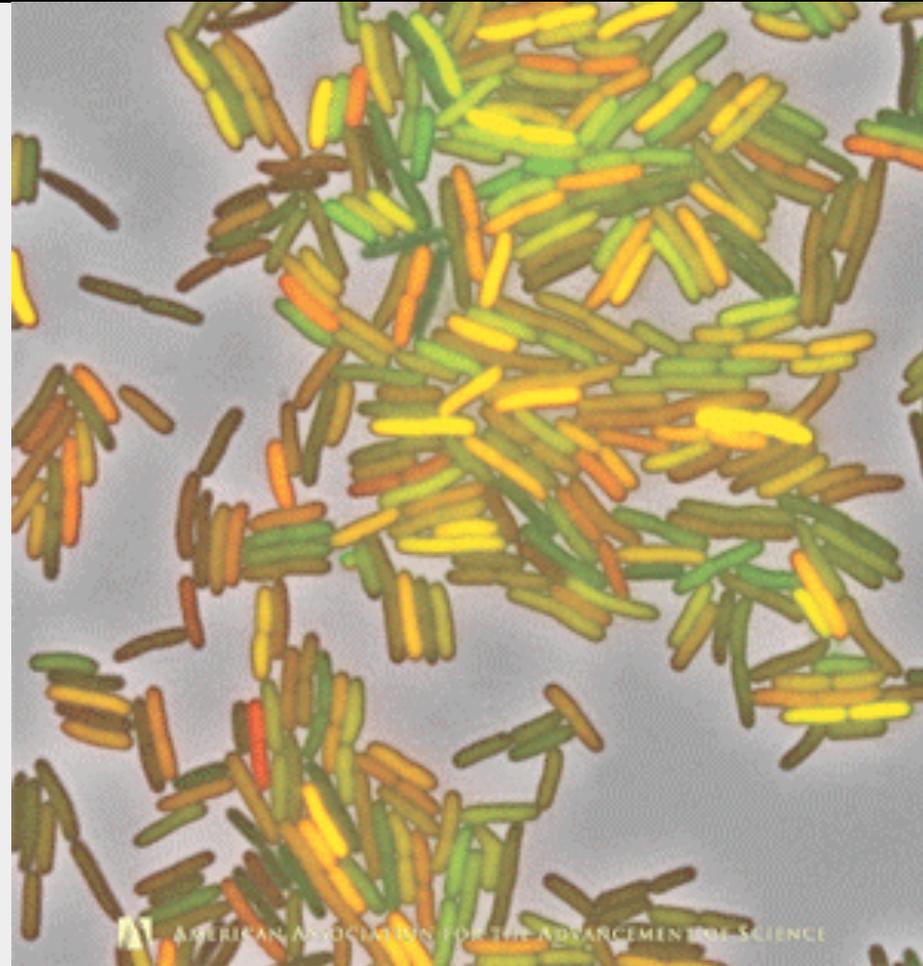


- les zones de contrôle des différentes phases sont partiellement superposées
- elles ne sont pas standardisées
- on sait les redessiner selon des principes d'ingénierie, et tester par des modèles mathématiques leurs propriétés
- les modèles permettent de prédire un comportement du phage synthétique 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



effet du bruit

l'expérience montre que la synthèse varie énormément d'une cellule à l'autre ; la plupart des modèles construits en biologie des systèmes est profondément irréaliste





du programme à la cellule



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



u n o r d i n a t e u r g é n é t i q u e

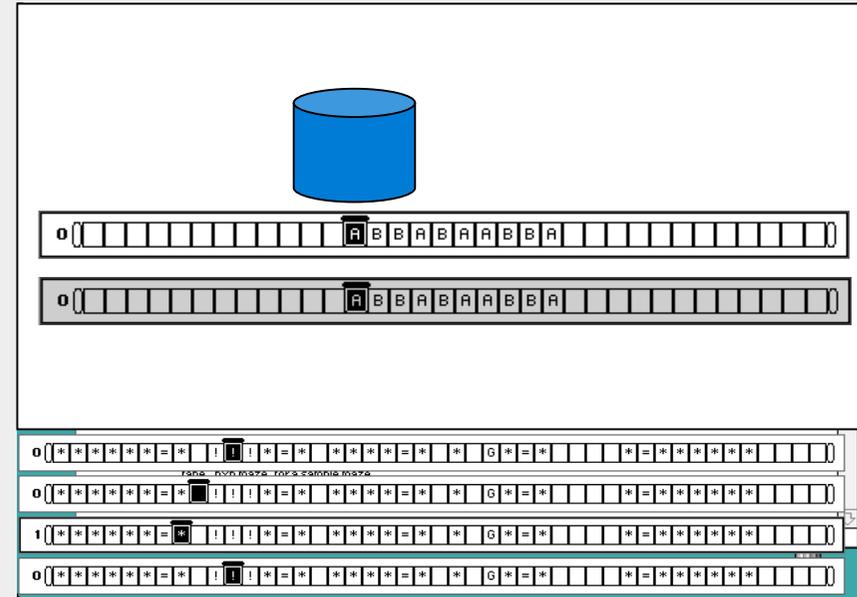
- dans un ordinateur, la machine est séparée des données et du programme
- données et programme jouent le même rôle (càd on peut les considérer comme des ' déclarations ') ; ils peuvent être modifiés par la machine
- une réflexion générale due à alan turing et de nombreux autres penseurs considère les actions de la machine, mais pas la façon dont elle est construite



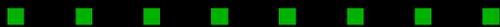
ce qu'est le calcul

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



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



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



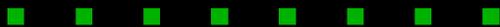
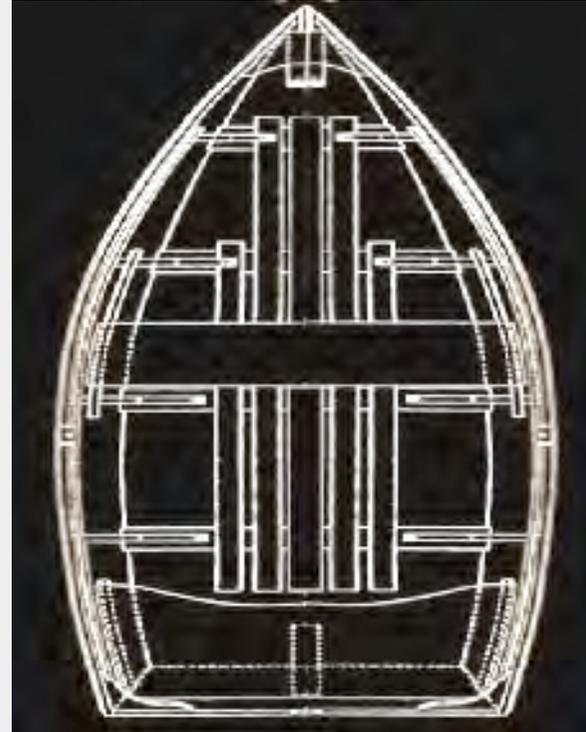
évolution et biologie synthétique

- la bs repose sur le modèle du programme génétique, où un **algorithme de construction** est **répliqué** au cours des générations
- une **machine (« chassis »)**, qui se **reproduit** au cours des générations, exécute l'algorithme, dont elle est physiquement séparée du support, comme dans les ordinateurs
- mais, si les ordinateurs font des ordinateurs, comment font-ils pour gérer le vieillissement ?
- l'analyse comparative des programmes génétiques met au jour des gènes (**gènes du démon de maxwell**) nécessaires pour **faire des organismes jeunes à partir d'organismes âgés** ; elle montre que les organismes vivants sont des **pièges à 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
- c'est une information qui exprime ce qui est conservé dans la barque et non la matière de ses planches !
- mais il faut une machine pour remplacer les planches, et la machine vieillit...



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

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



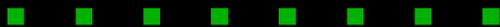
grand paon de nuit saturnia pyri
<http://pdubois.free.fr/>

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



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

matière / énergie / espace / temps

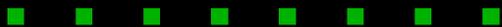
- physique classique
- physique quantique
- chimie
- biologie
 - développement
 - neurobiologie
 - linguistique
- mathématique (informatique)

information



l'information comme nouvelle catégorie du réel

- 1929 **leo szilard**, analysant le démon de maxwell, a l'intuition (fausse) de la relation entre énergie et information: la création d'1 bit consomme $1/2$ kT
- 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) ; l'énergie 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





comprendre grâce à l'évolution



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



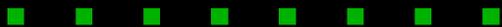
de l'ubiquité fonctionnelle à la persistance des gènes

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

→ rechercher la « persistance » permet d'identifier la plupart des fonctions ubiquistes

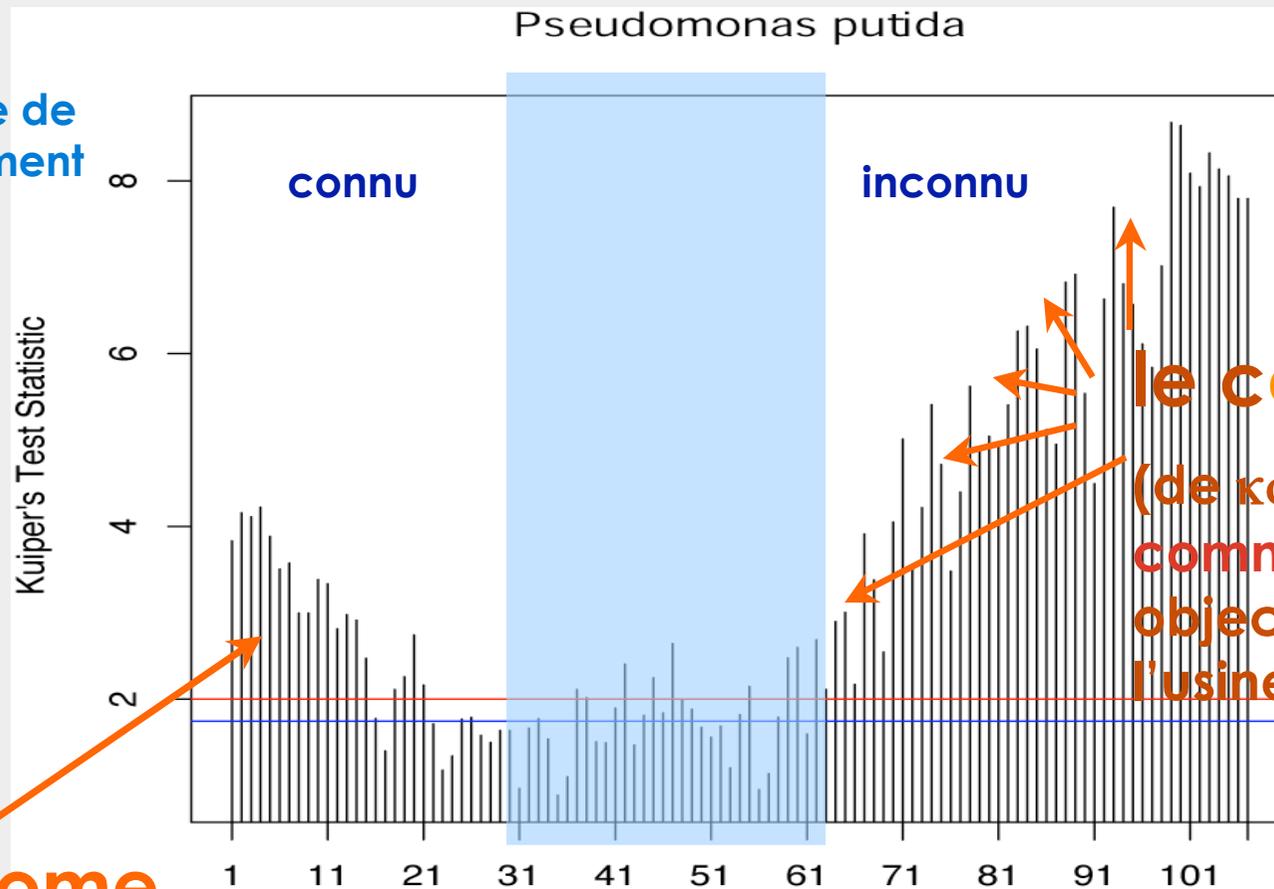
→ « ubiquiste » est-il un synonyme d'« essentiel » ? Les gènes « essentiels en laboratoire » sont situés dans le brin direct de l'ADN

~ 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



organisation des génomes bactériens

fréquence de regroupement



le cénome

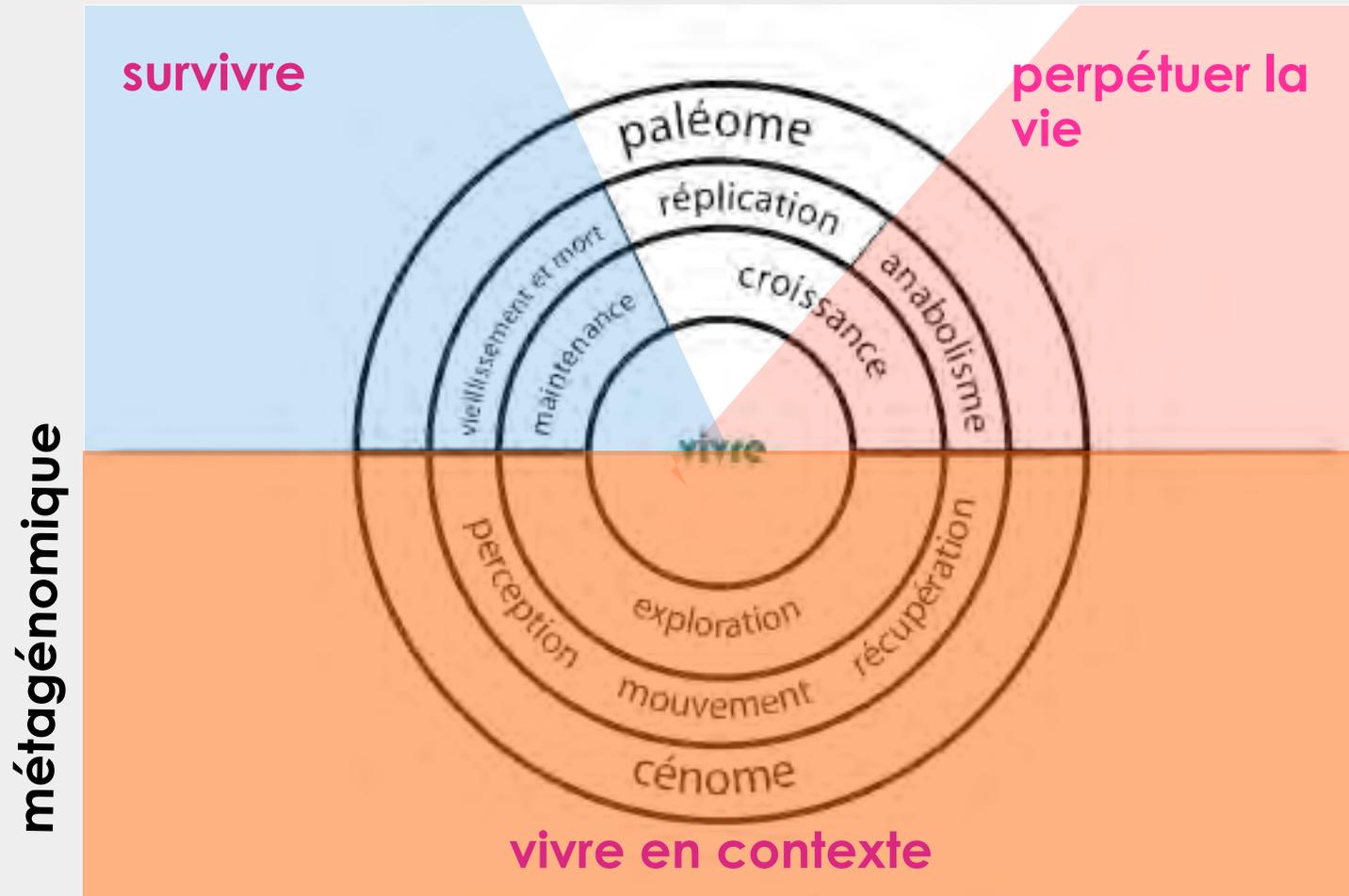
(de κοινος, commun): l'objectif de l'usine cellulaire

fréquence dans les génomes

le paléome

(de παλαιος, ancien): l'usine cellulaire

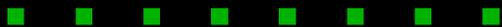
histoire de deux génomes



origines de la vie

- le physicien freeman dyson a montré qu'il ne peut y avoir **une seule** origine à 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 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 ?



un paléome diversifié

paléome 1 (gènes essentiels)

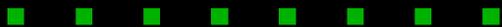
constructeur : l'ADN spécifie les protéines qui forment la machine qui construit la cellule (reproduction)

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)



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 ?



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

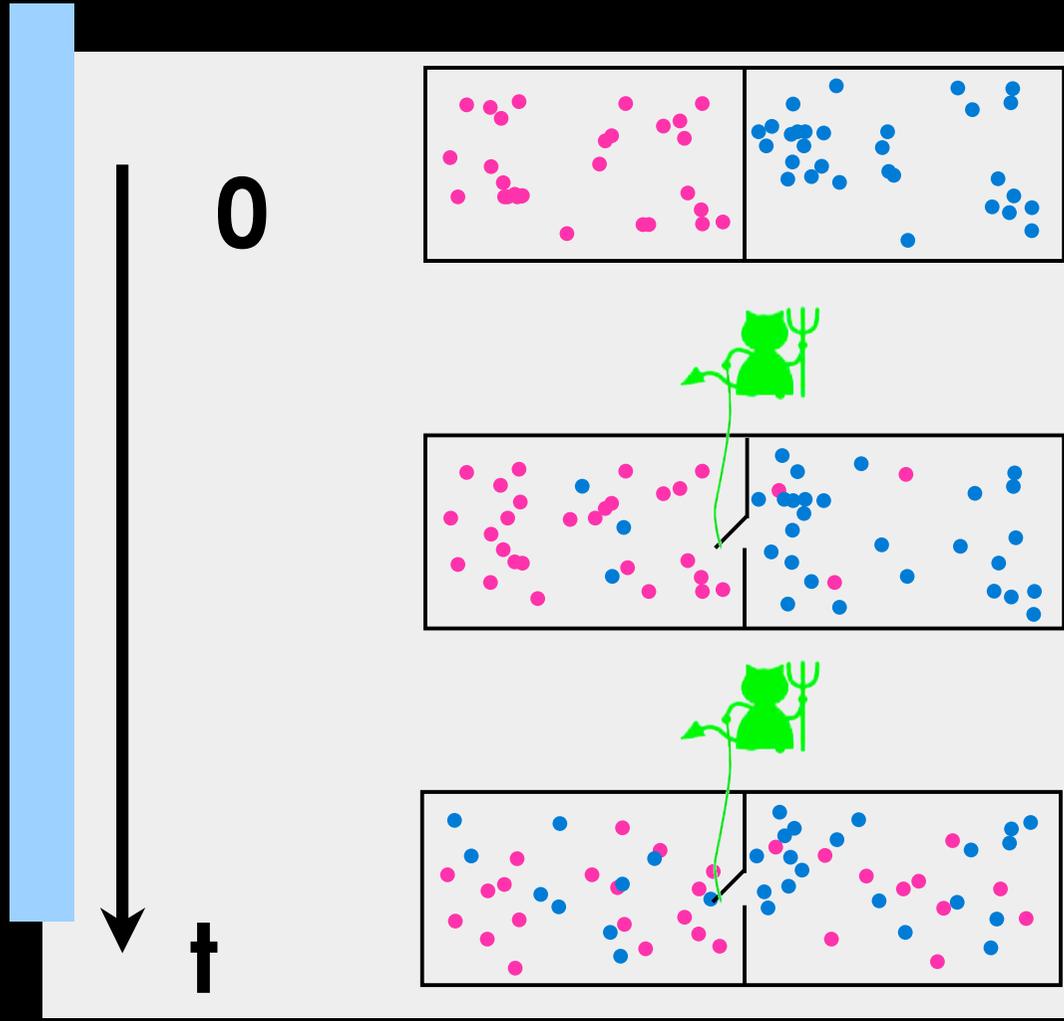


valeur de l'information

- dans les modèles classiques de l'information on ne se soucie pas de la signification, ou de la **valeur** de l'information
- 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 ?
- l'analyse comparative des programmes génétiques met au jour des gènes (**gènes du démon de maxwell**) nécessaires pour **faire des organismes jeunes à partir d'organismes âgés** ; elle montre que les organismes vivants sont des **pièges à information**



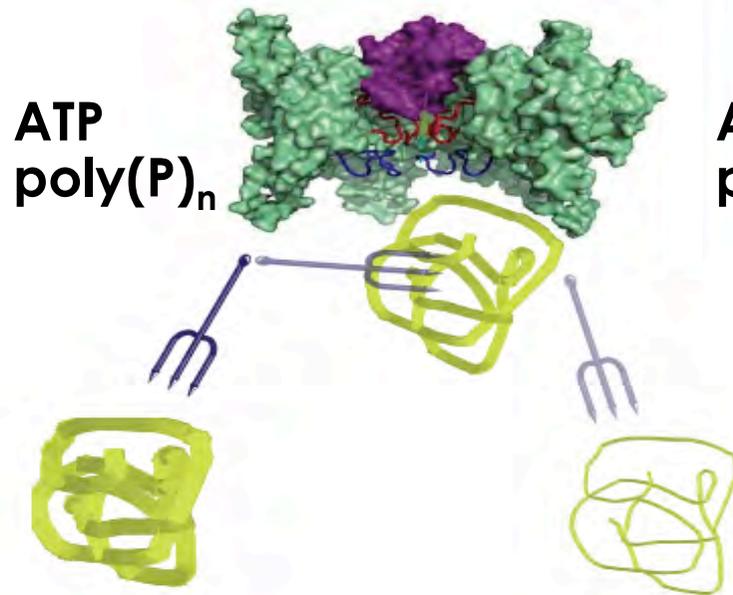
le démon de maxwell



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

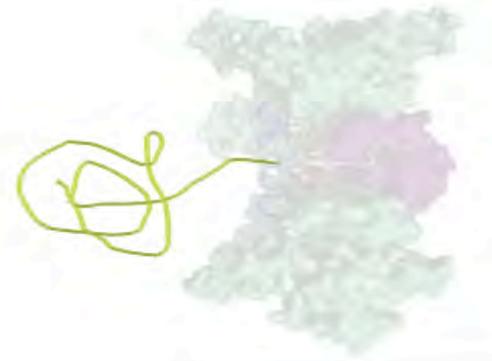


les gènes du démon de maxwell

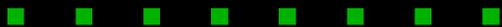


la machinerie de dégradation
utilise de l'énergie pour rejeter
intacte une entité fonctionnelle

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



les entités non fonctionnelles
sont reconnues et dégradées



innovation : les mutations adaptatives

- l'accumulation d'information énergivore 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



un modèle du cancer

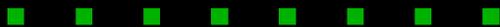
construction de bactéries "intelligentes"

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.

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

ces mutations ne préexistaient pas ; elles supposent donc la création d'information

les cellules mutantes envahissent la boîte comme des cellules cancéreuses



vivre et perpétuer la vie

- le génome 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 biologie synthétique exclut que les cellules artificielles soient innovantes
- il est possible d'omettre 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

