molecular traffic jams and the reproduction vs replication dilemma

antoine danchin 唐善・安東

amabiotics sas

international workshop on synthetic biology
copenhagen, august 27, 2010

goals of synthetic biology

- reconstructing and understanding: forgetting the “black box” sb reconstructs life to explore whether we understand what life is and learn missing entities from our failures

- abstracting: sb keeps the laws defining life, and applies them using objects of a different physico-chemical nature

- engineering: sb designs and standardises « biobricks » to construct a « cell factory » with man's interests' drive

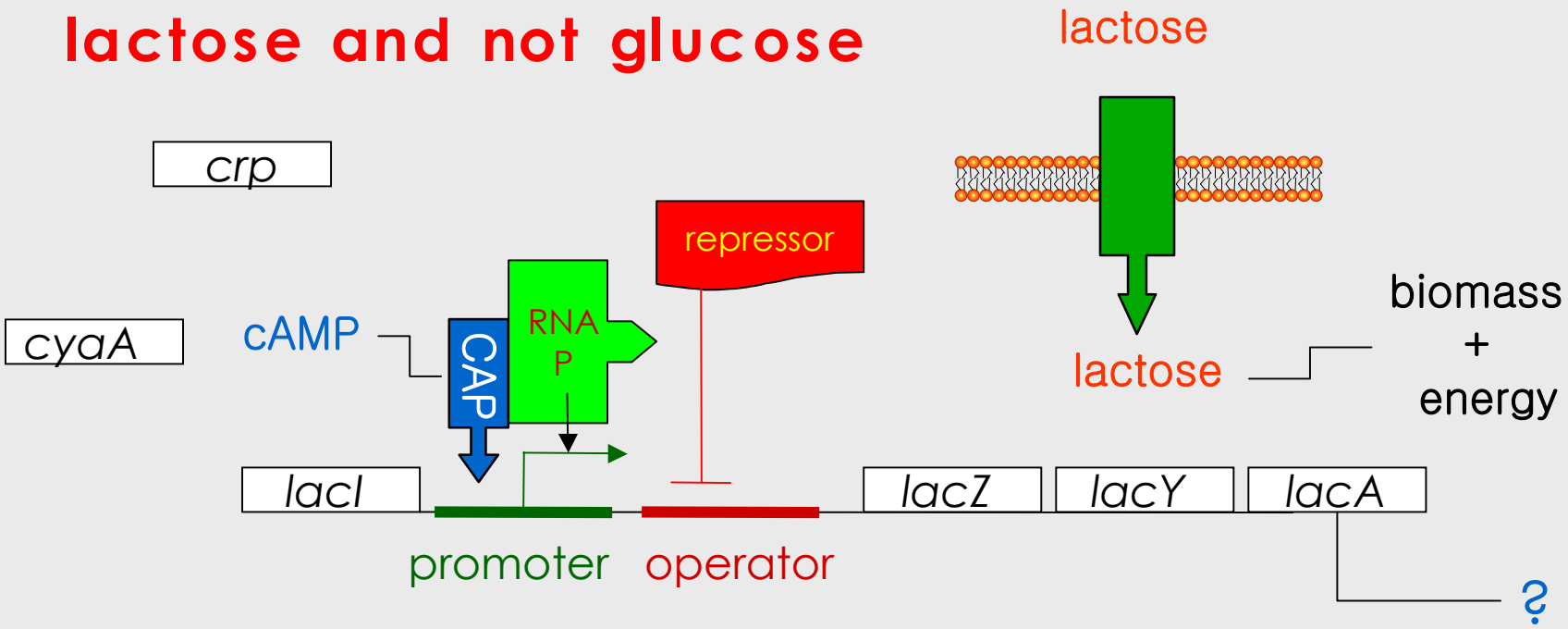
- evolving: sb combines design and evolution to use (poorly understood) principles that drive adaptation

reasoning as engineers do

. . . . . .

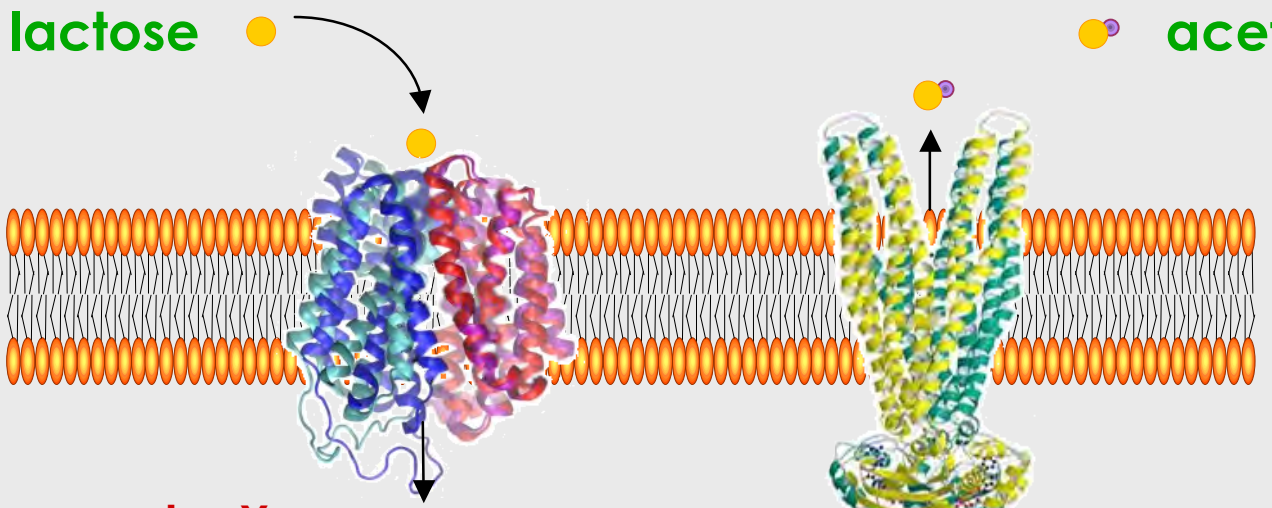
"the missing staircase"

lactose and not glucose



what is the function of lactose acetyl-transferase?
why did we need 60 years to ask the question?

cells need safety valves, not a leaks



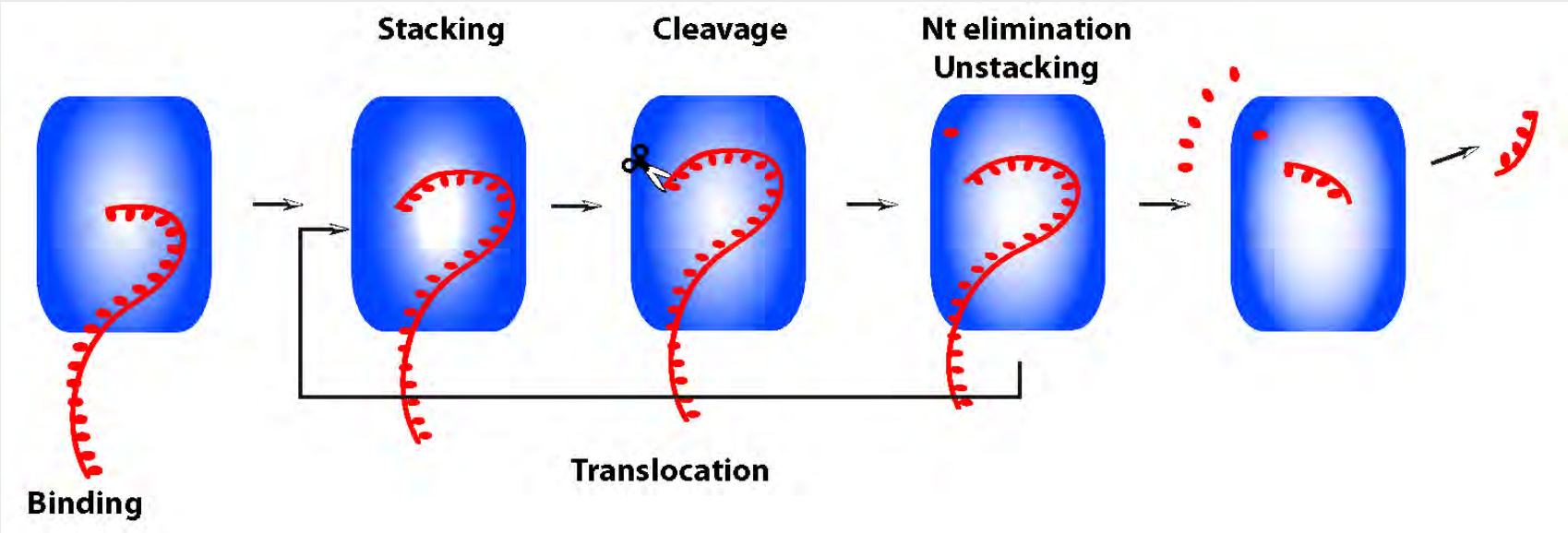
lactose acetyl-transferase (very inefficient K_M: 1 M)

coping with leftovers

nanornase is an essential function

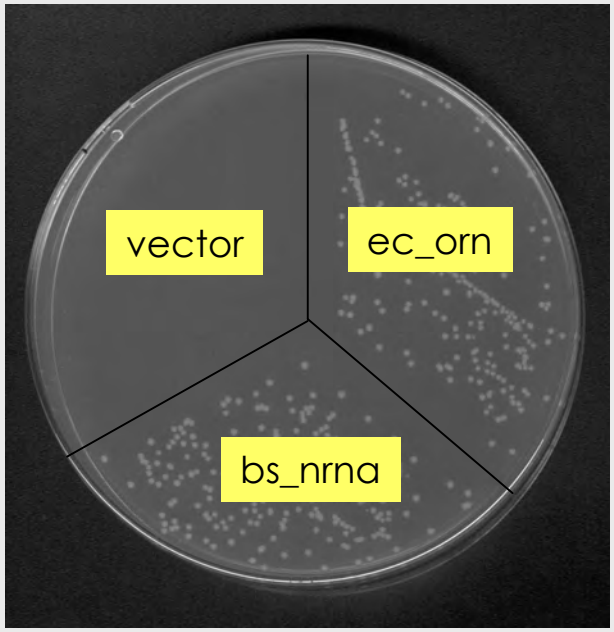
exoribonucleases are processive

this allows them to stick to their substrate and chew it up until they reach an end point where they can no longer proceed, yielding a leftover 2-5 nt long (usually 3nt)



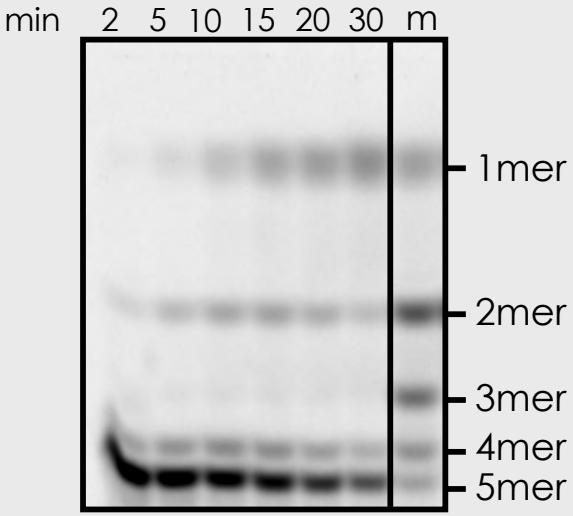
nano-rnases: a case of convergent evolution

in vivo



b. subtilis nrna complements
e. coli orn^{-}, no common origin

in vitro



degradation of nanorna 5mers
(cy5-ccccc-3')

an unlimited list of functions

- storage (location, address)
- coping with errors
  - metabolic interference (alpha-dicarbonyl)
  - misfolding
  - modifications: programmed or accidental
- robustness and promiscuity (functional leaks…)
- aggregation (what about crystals?)
- lubrication

computers making computers

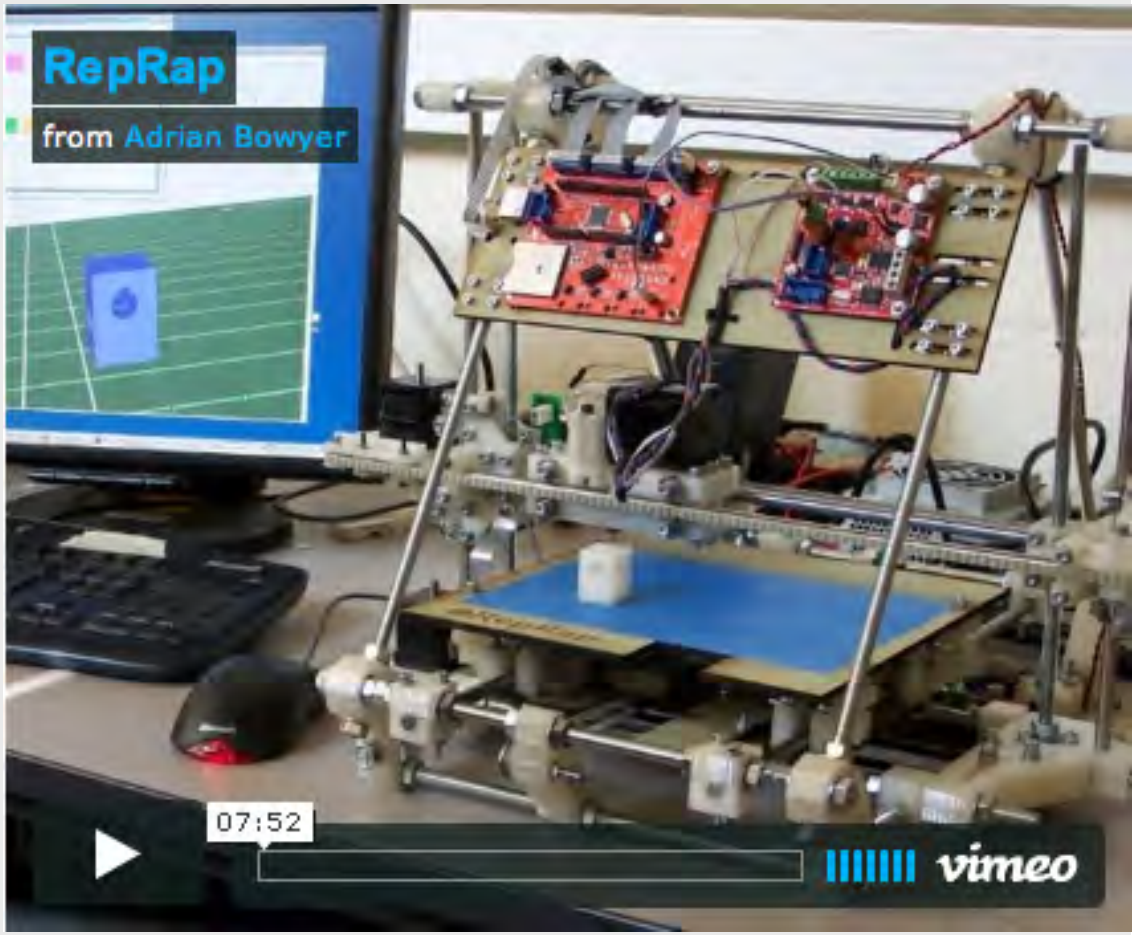
reprap (replicating rapid prototyper, 2004) aims at creating an auto - reproducing laser 3d printer:

the machine produces most of its components (= "biobricks")

missing:

- the program
- the assembly line (management of time and space, and specific functions such as lubrication)

http://reprap.org/



spatial constraints

cells and computers

genetics rests on the description of genomes as texts written with an alphabet: but do cells behave as computers?

horizontal gene transfer

viruses

genetic engineering

transplantation of a naked genome in a recipient cell changing the host recipient into a new one (2007)

everything separates

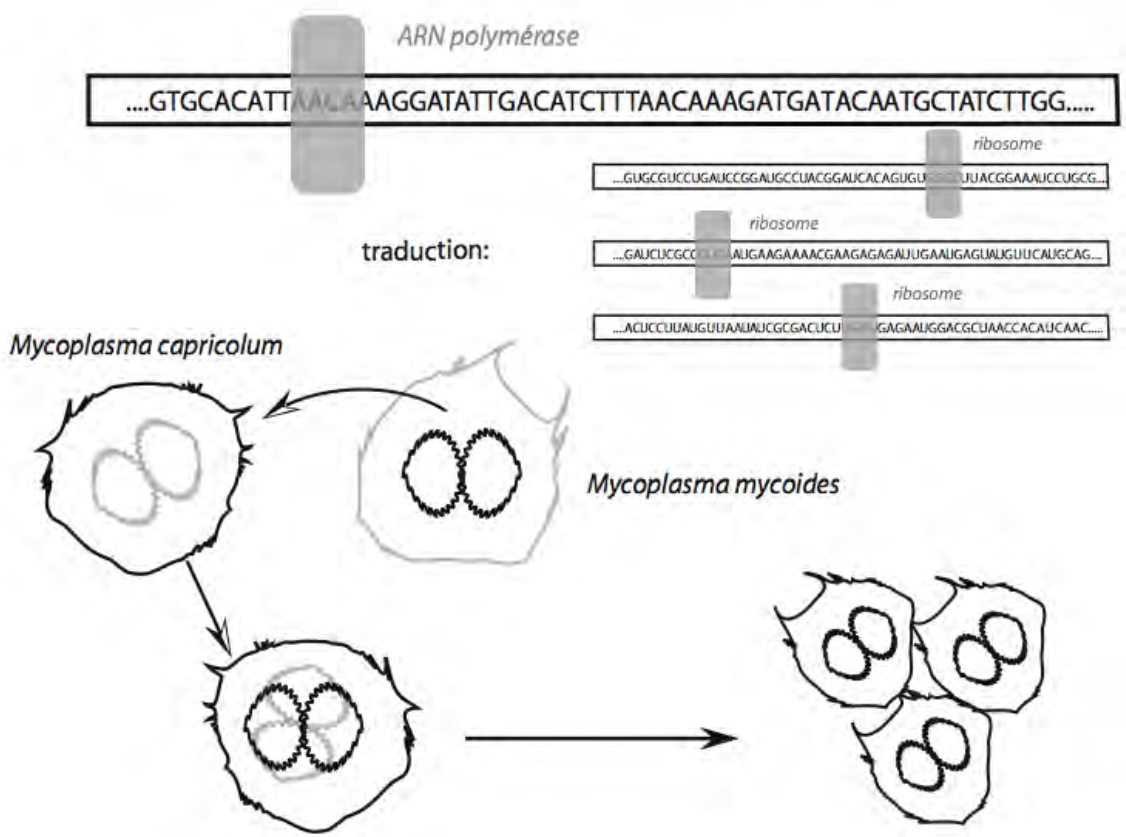
"machine" (chassis) and "data/program" (genome)

need for an operating system, and for constraints in the chassis

cells as computers

computers implement the turing machine concept, with a finite program

genome transplantation



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

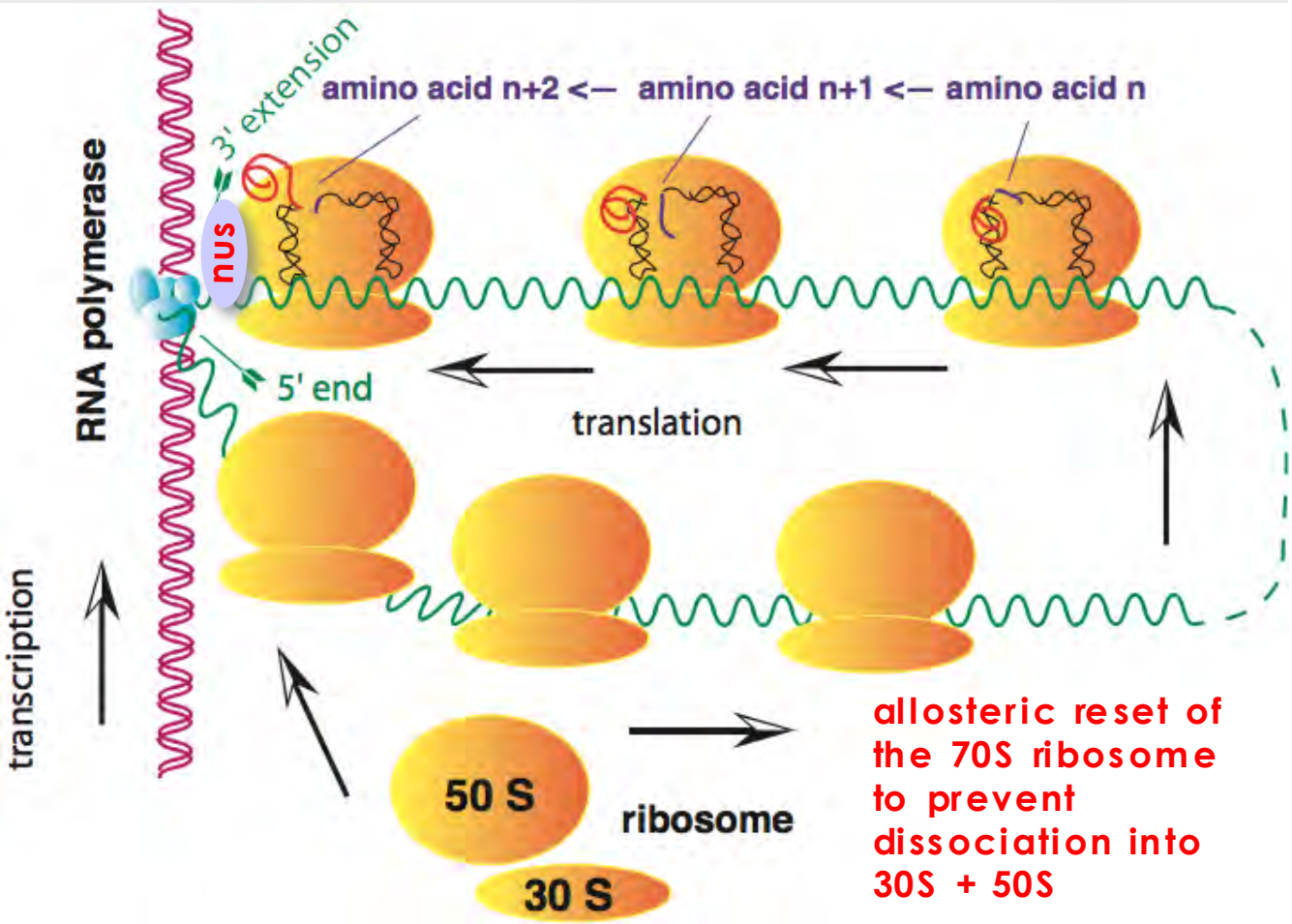
where are the ribosomes in the cell?

DNA persistence length > 10 times the diameter of the cell; + "handles" for replication

mRNA: free => tangles

1 fixed point => still a great many degrees of freedom

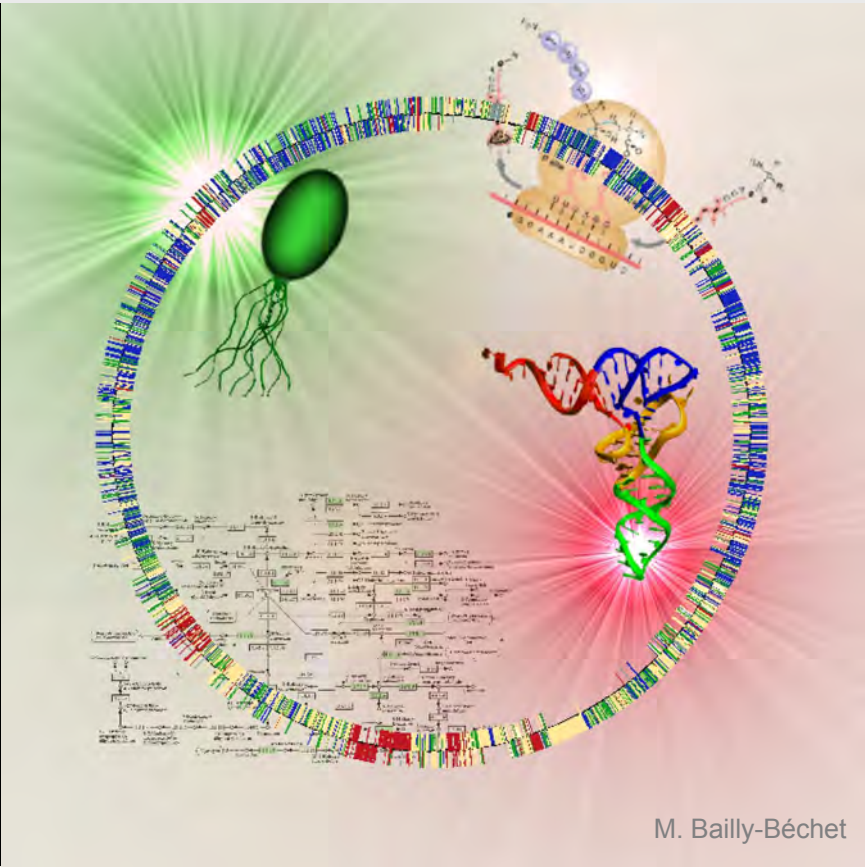
2 fixed points => dramatic collapse of degrees of freedom



evidence: translation islands

one groups is associated to high expression (blue)

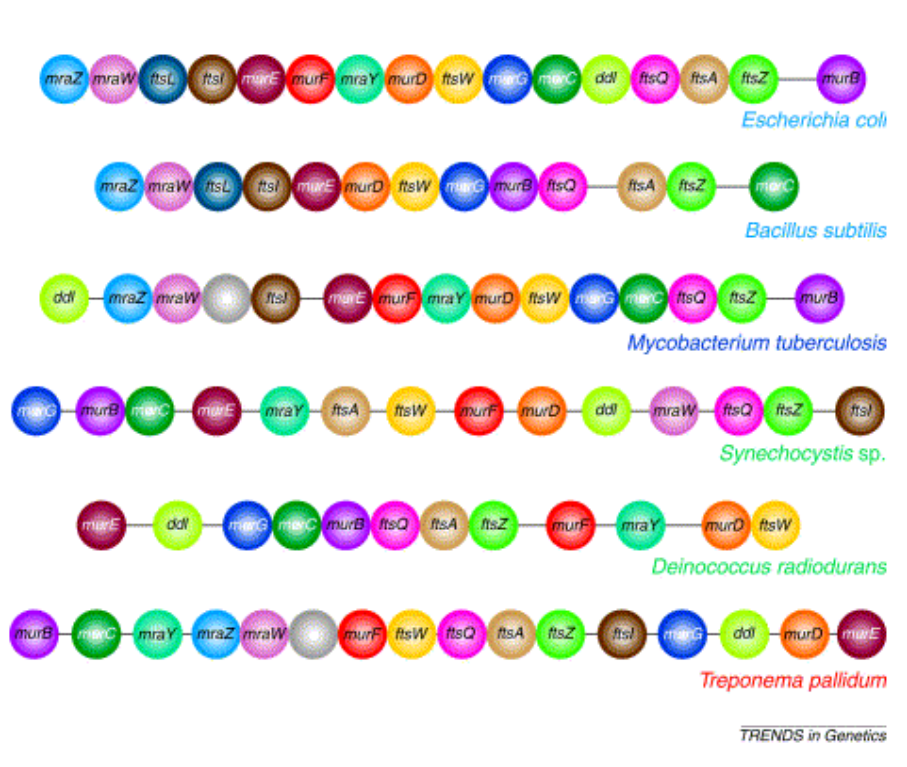
the other groups are also functionnally consistent: horizontally transferred genes (red) motility (yellow) and intermediary metabolism (green).

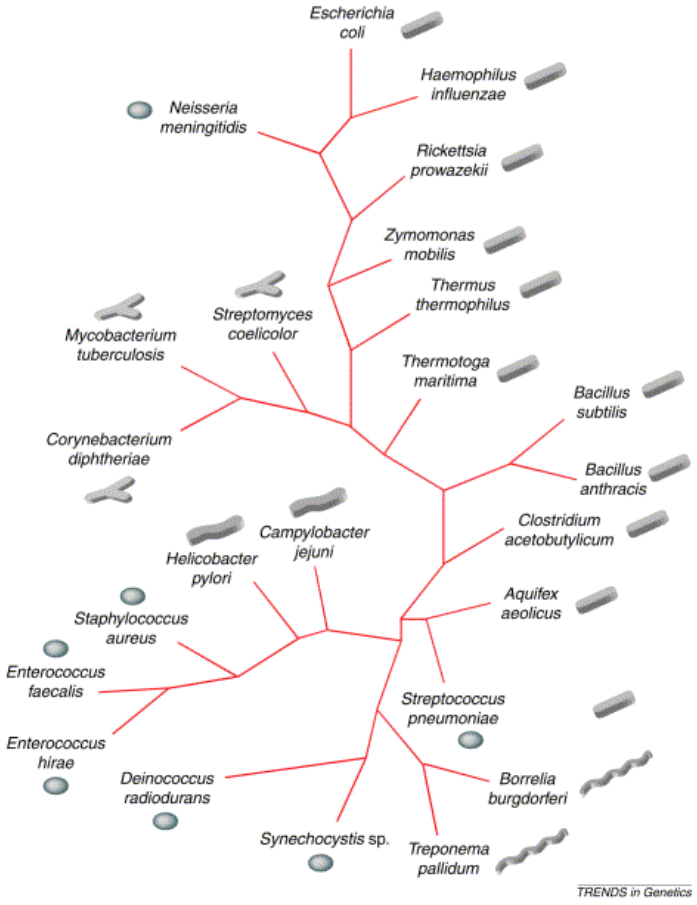


m bailly-bechet, a danchin, m iqbal, m marsili, m vergassola
codon usage domains over bacterial chromosomes
plos computational biology (2006) **2**: e37

gene order and cell shape

mur-fts islands



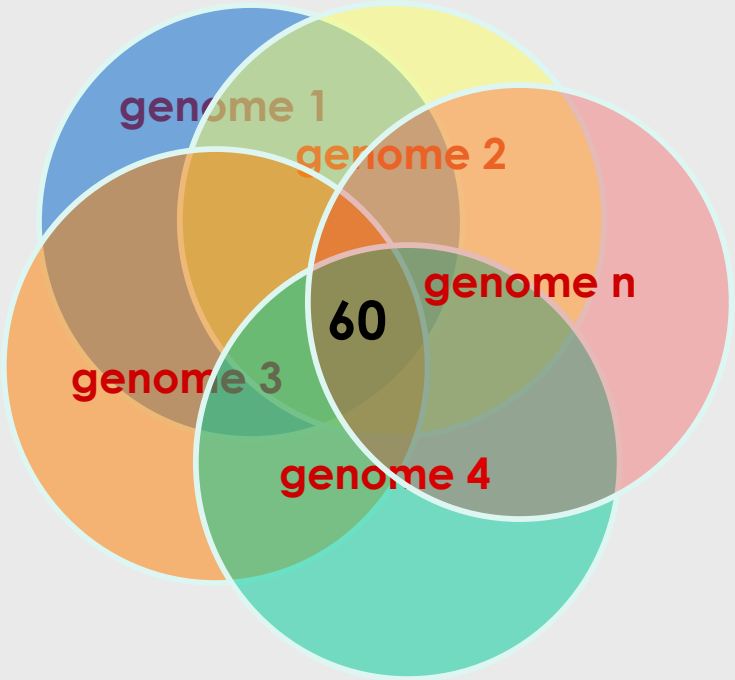


tamames j, gonzalez-moreno m, mingorance j, valencia a, vicente m
bringing gene order into bacterial shape
trends in genetics (2001) **17**: 124-126

a minimal set of functions

. . . . . .

2003: 60 conserved proteins



the number of conserved genes tends to zero!

from functional ubiquity to gene persistence

functional gene ubiquity does not imply gene ubiquity

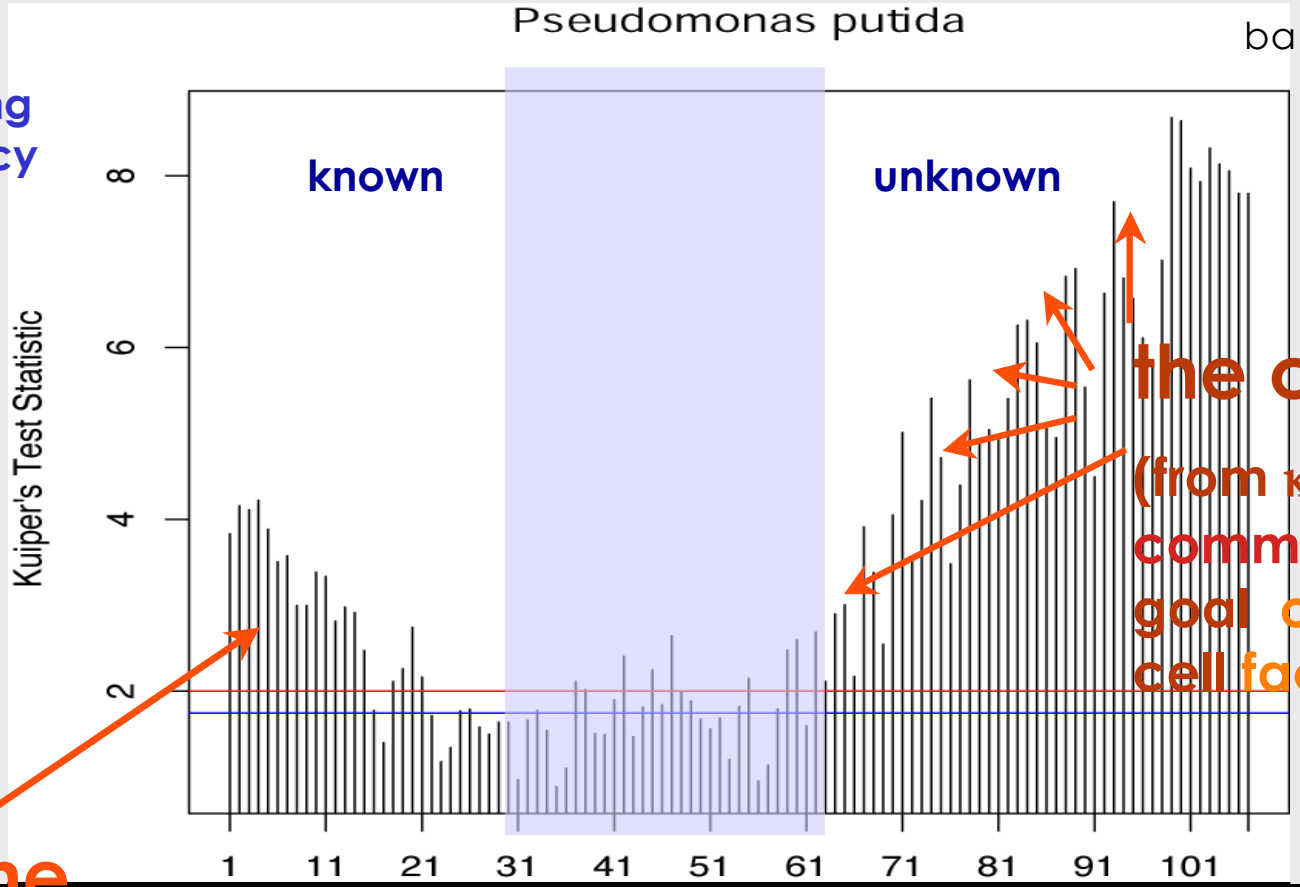
yet, efficient entities tend to persist through generations

→ looking for « persistence » identifies most ubiquitous functions

~ 500 genes persist in bacterial genomes; they are involved not only in the three processes required for life but also in maintenance and adaptation to transient phenomena ; a fraction manages the evolution of the organism, via energy-dependent degradative processes

an unexpected common structural feature: persistent genes are located in the leading DNA strand

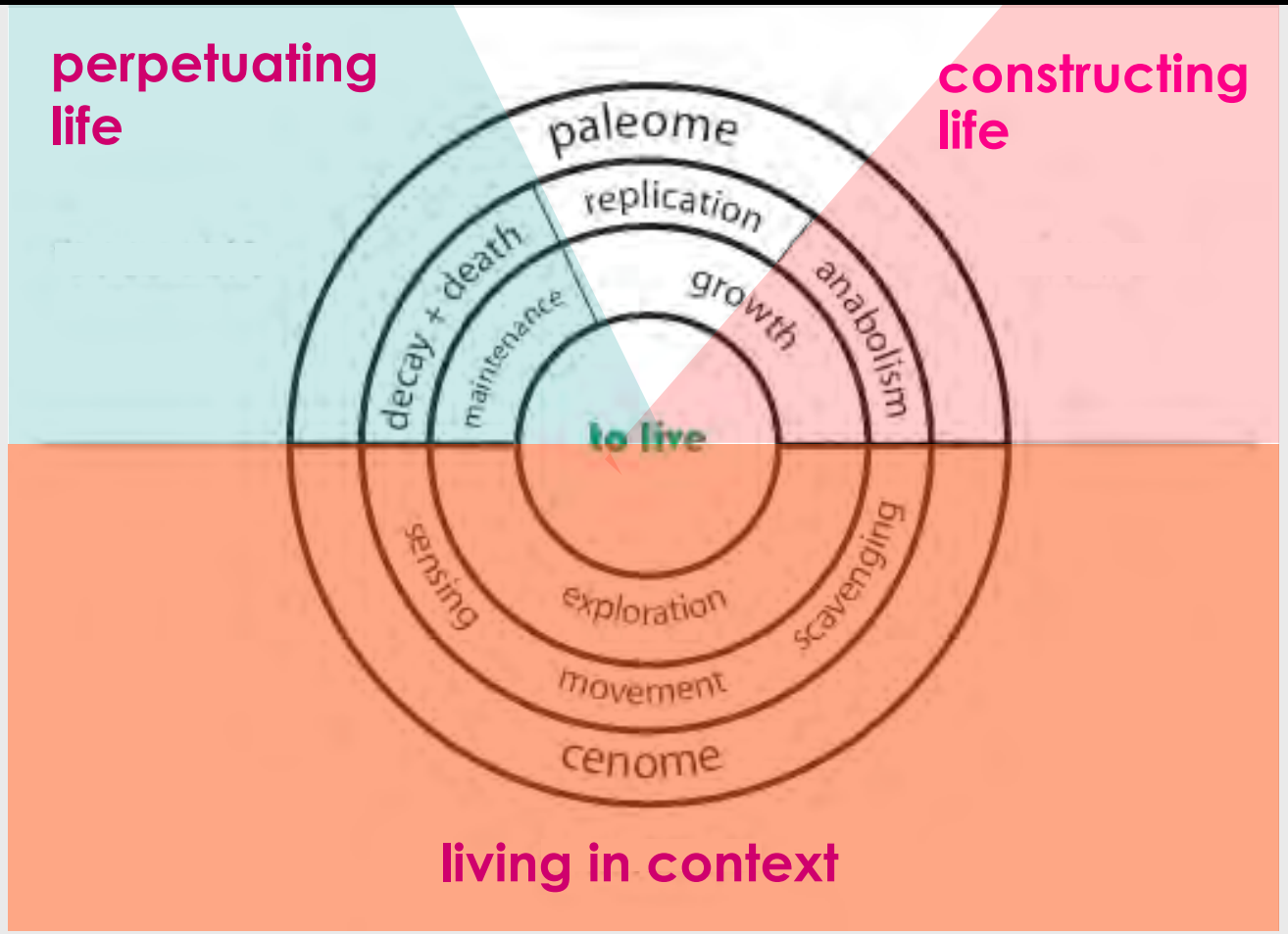
syntenies of orthologs



paleome

(from παλαιος, ancient): the cell factory

a tale of two genomes



organised genome dynamics in the *escherichia coli* species results in highly diverse adaptive paths
touchon m, hoede c, tenaillon o, barbe v, ..., medigue c, rocha ep, denamur e.
plos genet. 2009 jan;5:e1000344

maxwell's demon's genes

revisiting information

living organisms are information gathering and using systems (igus) that aim at maintaining their activity by building a *record* of the relevant measurements they have performed (zurek, 1989) => genetic and epigenetic heredity

to say that a system occupies a certain state implies that one has the information necessary to generate a complete description of that state: the information gathering process is reversible (i.e. does not use energy) provided it is allowed to save a copy of the input (landauer, 1961)

landauer's theorem: to erase a bit of information in an environment at temperature T requires dissipation of energy $\geq k_B T \ln 2$; this demands that information be granted a physical status as a negative contribution to free energy

$$\mathcal{F} = F - K_B T \ln 2 \, I = E - k_B T \ln 2 \, (H+I)$$

value of information

the information of the program is transmitted "as is" during replication, with no value associated to particular sequences: where does the information of the machine (and of the environment) come from?

to accumulate information requires an energy-dependent processs to "make room", without erasing valuable information

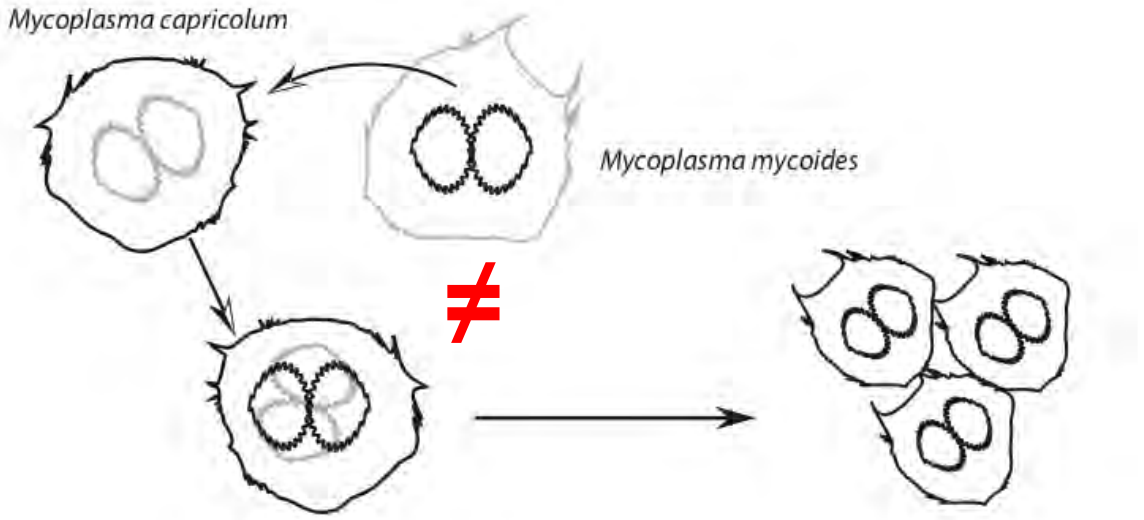
can we imagine the genes of a maxwell's demon which would select among what is functional or young (locally) and what does not work?

replication and reproduction are not the same

the program replicates (makes an identical copy)

the cell reproduces (makes a similar copy)

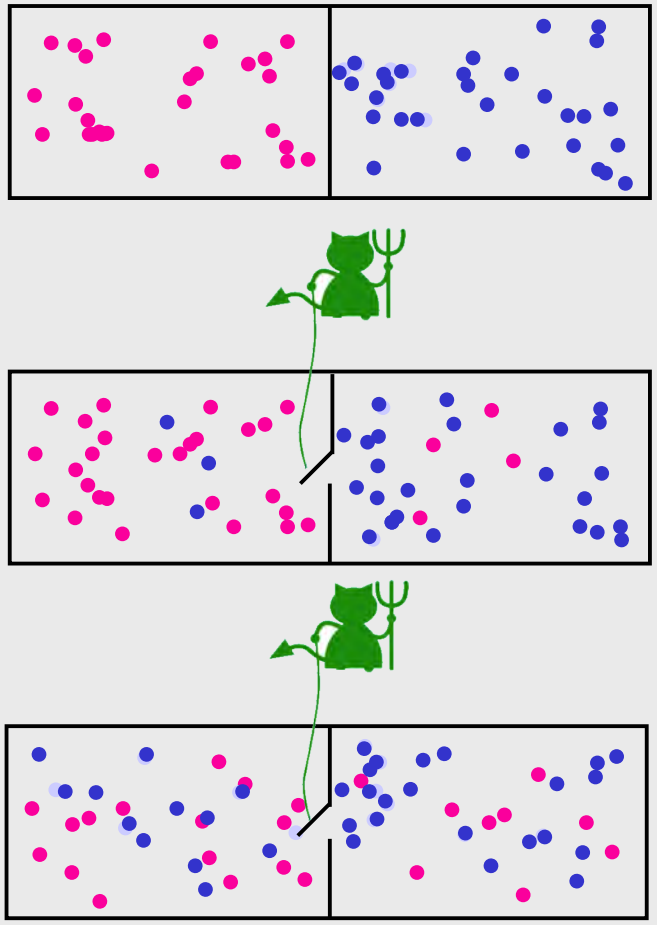
this split is the basis of evolution



maxwell's demon

hot

cold



0

t

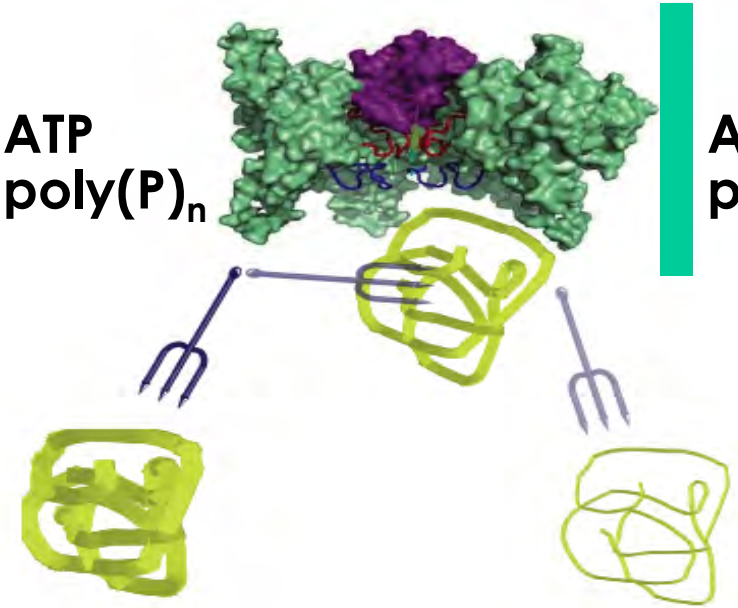
the demon reverses time while measuring the speed of the atoms of gas, recording an information to calculate when it must close the trap, thus permitting temperature-dependent generation of energy

0

t

maxwell's demon's genes

ATP
poly(P)$_n$



ADP + Pi
poly(P)$_{n-1}$ + Pi **<= in the paleome**

the degradation machinery uses energy to reject unaltered a functional entity

non functional entities are recognised and degraded

innovation: adaptive mutations

- energy-dependent accumulation of information is blind; it ignores the source of information
- information can come from a memory, that of the pre-existing genome; it can also be created de novo

- adaptive mutations are de novo creations of information; therefore they dependent on genes involved in accumulation of information

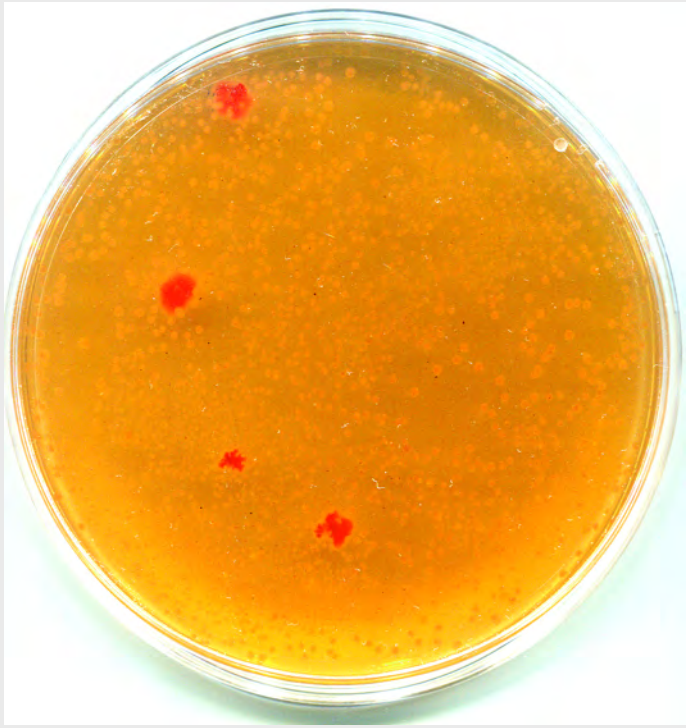
adaptive mutations

construction of "intelligent" bacteria

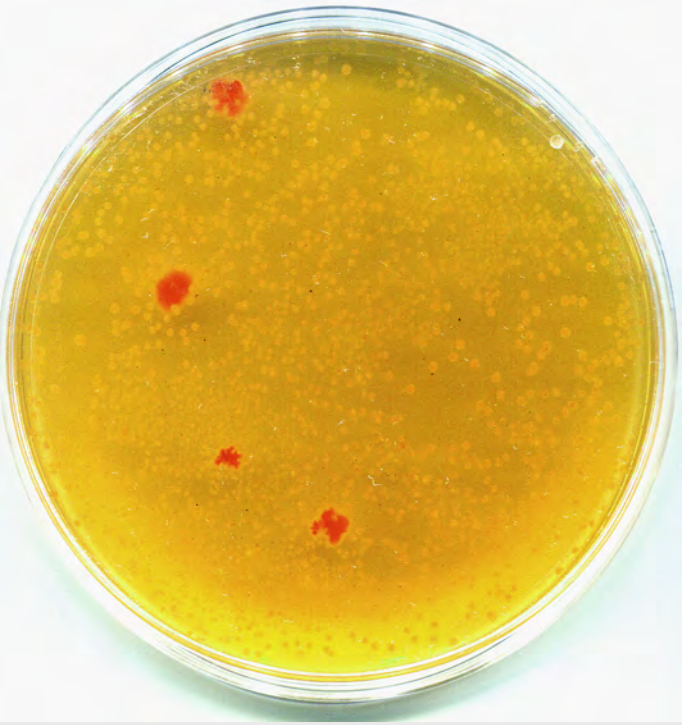
placed to grow on a medium with limited nutrient supply; form colonies of approximately 10^7 bacteria; the medium also contains nutrients that they cannot use

after a few days/weeks time, papillae appear that begin to grow and invade the plate, using supplied "unusable" nutrients; they derive from adaptive mutations

they did not pre-exist, and this supposes creation and recording of information



adaptive mutations



sequencing seven genomes + 30 pcrs

the total number of mutations is higher in older colonies

mutations are spread throughout the chromosome, and concentrated in one gene => pcr of many colonies

in this particular gene one finds different mutations in different papillae, 2 mutations in 30% of the cases

in some cases one of the two mutations is silent

on a particular carbon source, there is a least one other gene involved

natural selection is a principle of physics

- **natural selection**: making room using energy to avoid erasing context-dependent functional information
- energy-dependent degradative processes make room for newly synthesised entities; energy is consumed to **prevent** degradation of functional entities
- this process accumulates information, whatever its origin, in a ratchet-like process
- this process is **myopic**: it cannot have a design, hence the "tinkering" feature of life and its evolution

a synthetic cell?

- the engineering view of sb precludes that artificial cells be innovative
- we can exclude the genes permetting accumulation of information
- the consequence is that the cell factory will age and will need to be systematically rebuilt
- this has a in-built societal benefit, as risks are minimised
- but this poses problems when applications require that industrial processes are scaled-up: this may not be possible, unless we can harness the function of the maxwell's demon's genes to the human goals

a novel publication concept

symplectic biology

rapid research notes in systems and synthetic biology

Home | About Us | Copyright | 1st Max Planck Society - Chinese Academy of Science Exploratory Round Table Conference (ERTC) on Synth

Log in

News

International Symposium on Synthetic Biology, october 18-19, 2010, Orchard Hotel, Singapore

Jul 10, 2010

Synthetic Biology Copenhagen 2010 International Workshop, August 25-27, 2010, Nano-Science Center, University of Copenhagen

Jul 03, 2010

EC-US task Force on Biotechnology Research, Workshop on Standards in Synthetic Biology, 4-6 june 2010 - Spain

May 10, 2010

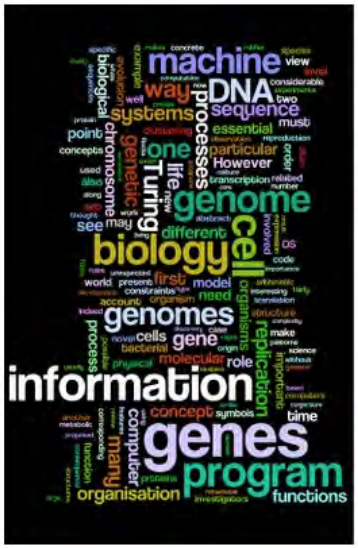
Workshop on Synthetic Biology, Denmark, August 25-27, 2010

May 10, 2010

More news...

Welcome to Symplectic Biology

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



Featured Articles

The "missing" experiment showing separation of the program from the chassis has been performed by Craig Venter and his co-workers

Challenges

Challenge n°1: Is the synteny around the pyrH gene in bacterial genomes significant?

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

Search

Search Site | Search

Advanced Search...

Board of Editors

Editors in chief

Antoine Danchin
Victor de Lorenzo

Editors

Sigal Ben-Yehuda
Philippe M. Binder
Jesús Blázquez
Nediljko Budisa
Jason W. Chin
Gang Fang
Jian-Dong Huang
Jose Ignacio Jimenez
Frédérique Lisacek
Philippe Marlière
Kenta Nakai
Christos Ouzounis
Sven Panke
Manuel Porcar
Alfonso Valencia
Anil Wipat
J. Tze-Fei Wong
Huanming Yang

http://knol.google.com/k/symplectic-biology#

contributions

in silico

gang fang, eduardo rocha

in vivo

undine mechold, agnieszka sekowska

collaborations

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