
maxwell demon's genes

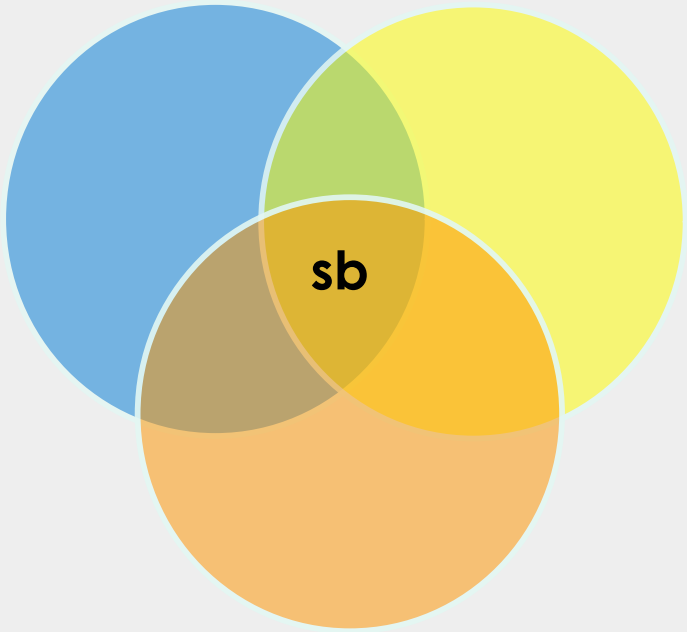
antoine danchin

唐善・安東

interfacing systems biology and synthetic biology
belgian society for microbiology
bruxelles, march 25, 2010

situation of synthetic biology

engineering
computation
modelisation



molecular biology
phylogenomics
biotechnology

origin of life
artificial life
orthogonal life

goals of synthetic biology

- to reconstruct life, finding out what is missing via our failures
- taking the engineer viewpoint, to classify and normalise « biobricks » to construct a « cell factory »
- to apply our knowledge to build life with objects of a different physico-chemical nature

the parallel for computers

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

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

missing:

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

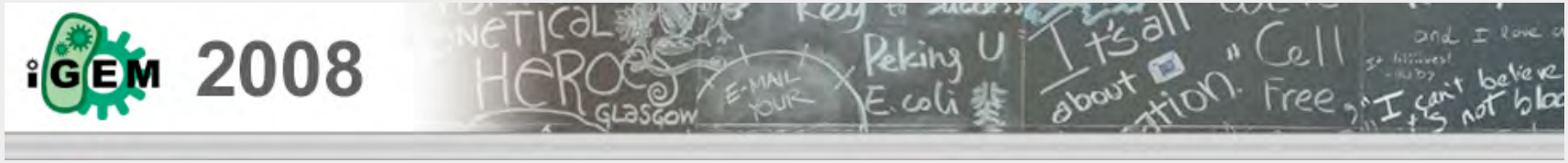
http://reprap.org/

the electronic paradigm

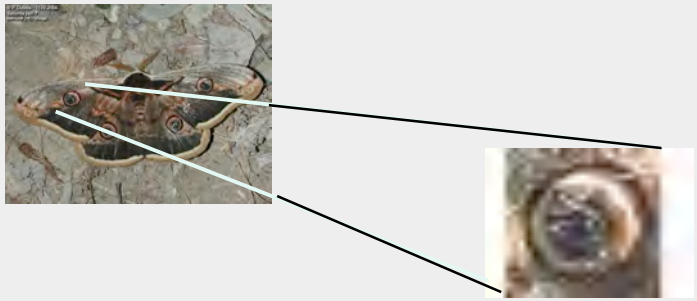
1. abstraction
   - isolation of the essential features of what is (re)constructed.
2. conception
   - conception of circuits with predictable properties.
3. standardisation
   - use and reuse of "biobricks".
4. modelisation
   - simulation of the construction and prediction of its dynamics, sensitivity to noise etc.
5. construction and test
   - combination of biobricks into a "chassis"; analysis of the behaviour of the construct and lessons from experience. --

back to 1

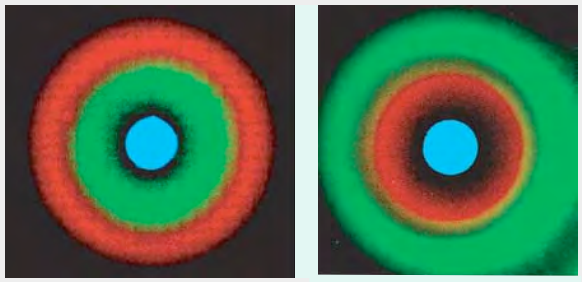
MIT : iGEM and its « biobricks »



iGEM (International genetically engineered machines) **asks the question:**



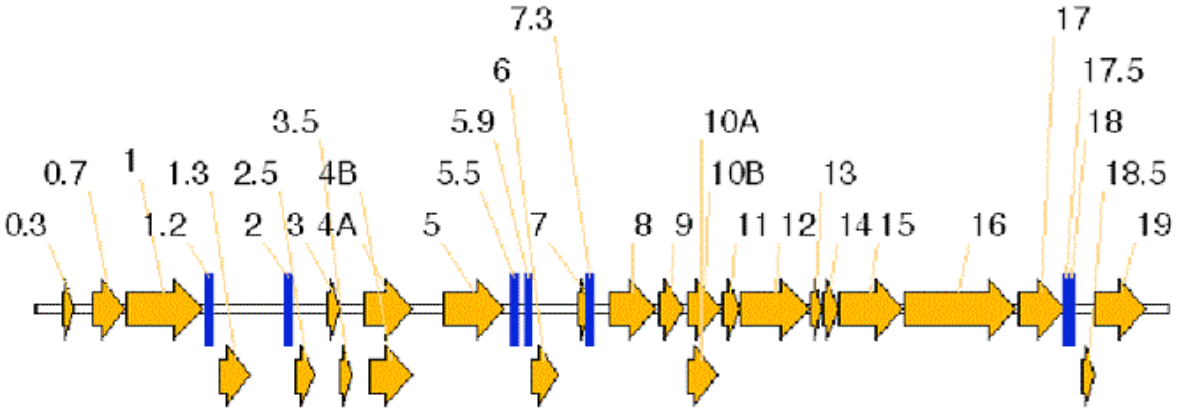
can we construct construct biological systems from standardised elements, placed within living cells? or, is biology too complicated to be reconstructed that way?



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

reprogramming bacteriophage T7

known genes of bacteriophage T7



•taking control•

• destruction •

• replication •

• synthesis of the capsid •

•encapsidation•

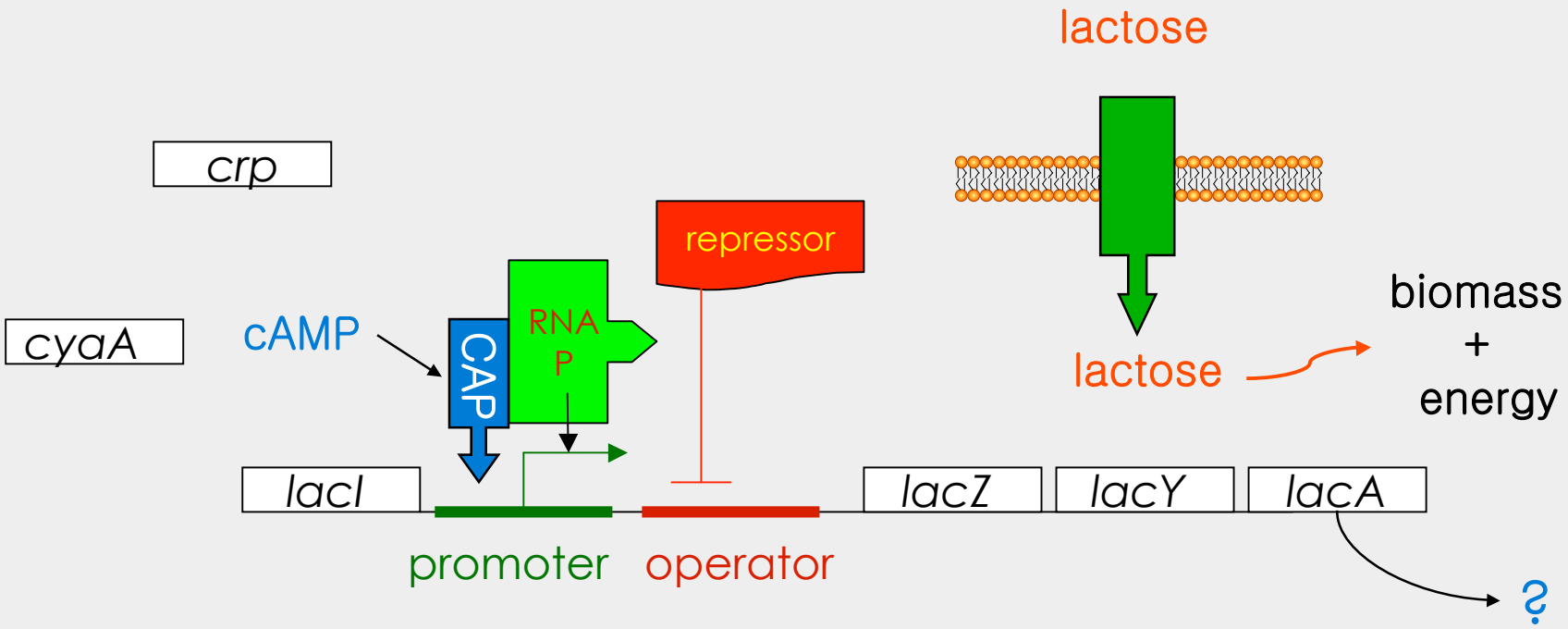
• getting out of the cell lysis •

drew endy: http://online.kitp.ucsb.edu/online/infobio01/endy/

reprogramming bacteriophage T7

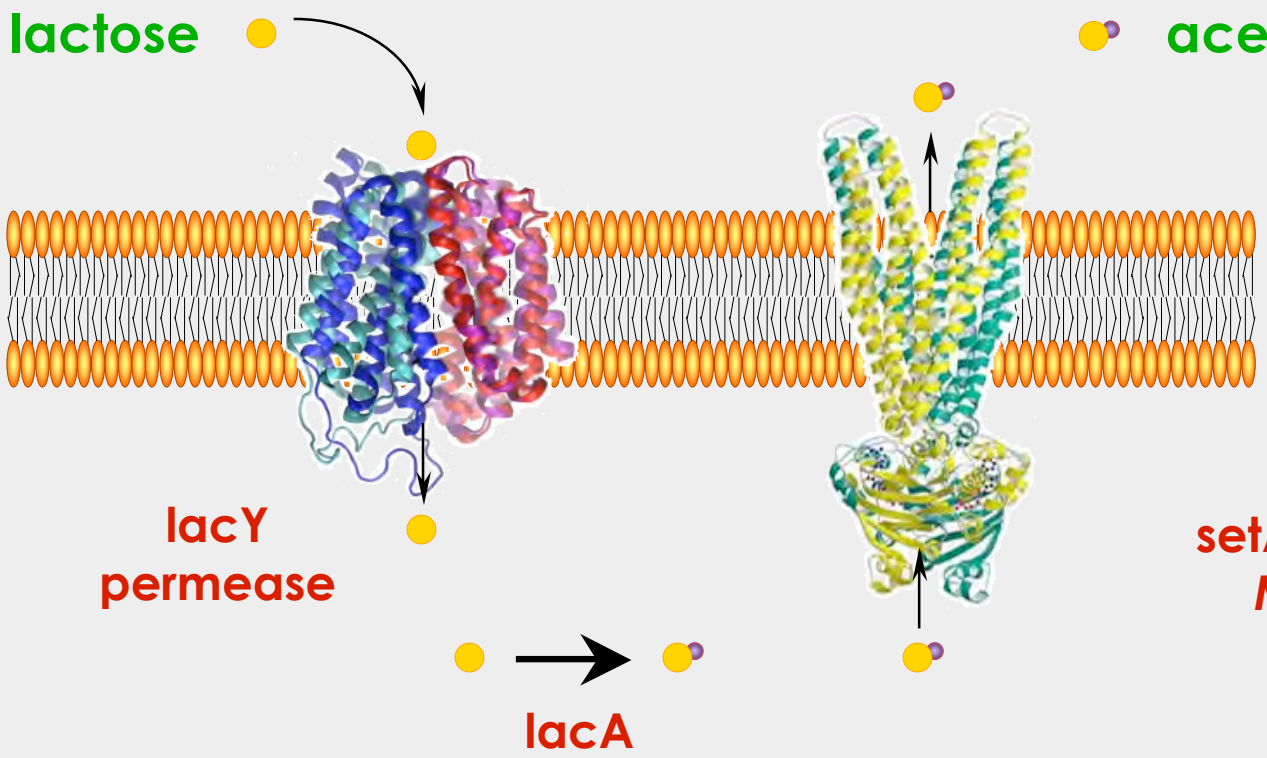
- control regions overlap
- they are not standardised
- they may be redesigned according to engineering rules, and tested using mathematical models
- models predict the synthetic phage behaviour and compare it with the natural phage
- the synthetic phage forms smaller lysis plaques than its natural counterpart
- the evolution the synthetic phage to more virulent forms erases the human construct…

the lactose operon revisited

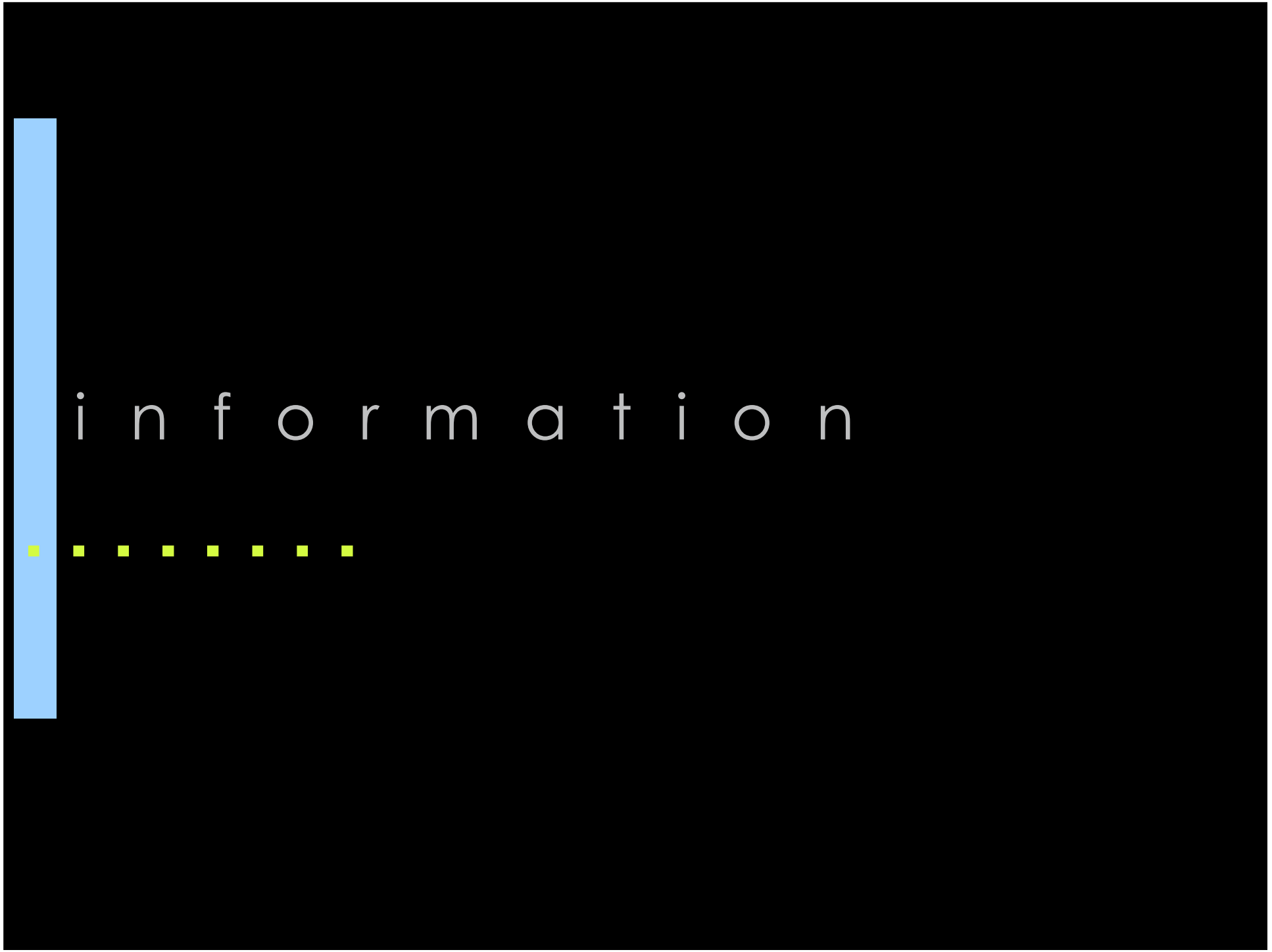


what is the function of lactose acetyl-transferase?
why did we need 60 years to ask the question?

s a f e t y v a l v e



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


information

w h a t l i f e i s

life requires:

- a machine ("chassis") allowing the program to be expressed (reproduces)
  - 1. metabolism (a dynamic process)
  - 2. compartmentalisation (casings, defining inside and outside)

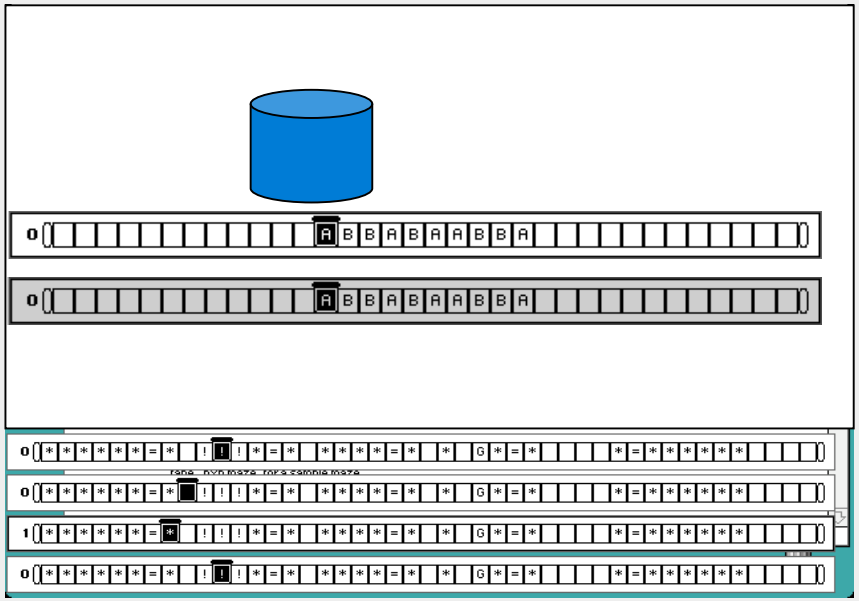
- a program (a “book of recipes”, which is replicated)
  - 3. recursive information transfer and trapping => coding from one level to a second one introduces an essential asymmetry (fundamentally different from feedback)

the cell is the atom of life

what computing is

two entities :

- a machine to read and write
- a program (sequence of symbols on a physical support), split into two entities:
  - program (defining the goal)
  - data (defining the context)



the machine is distinct from the entity data/program

cells and computers

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

- horizontal gene transfer
- virus
- genetic engineering
- transplantation of a naked genome in a recipient cell changing the host recipient into a new one (2007)

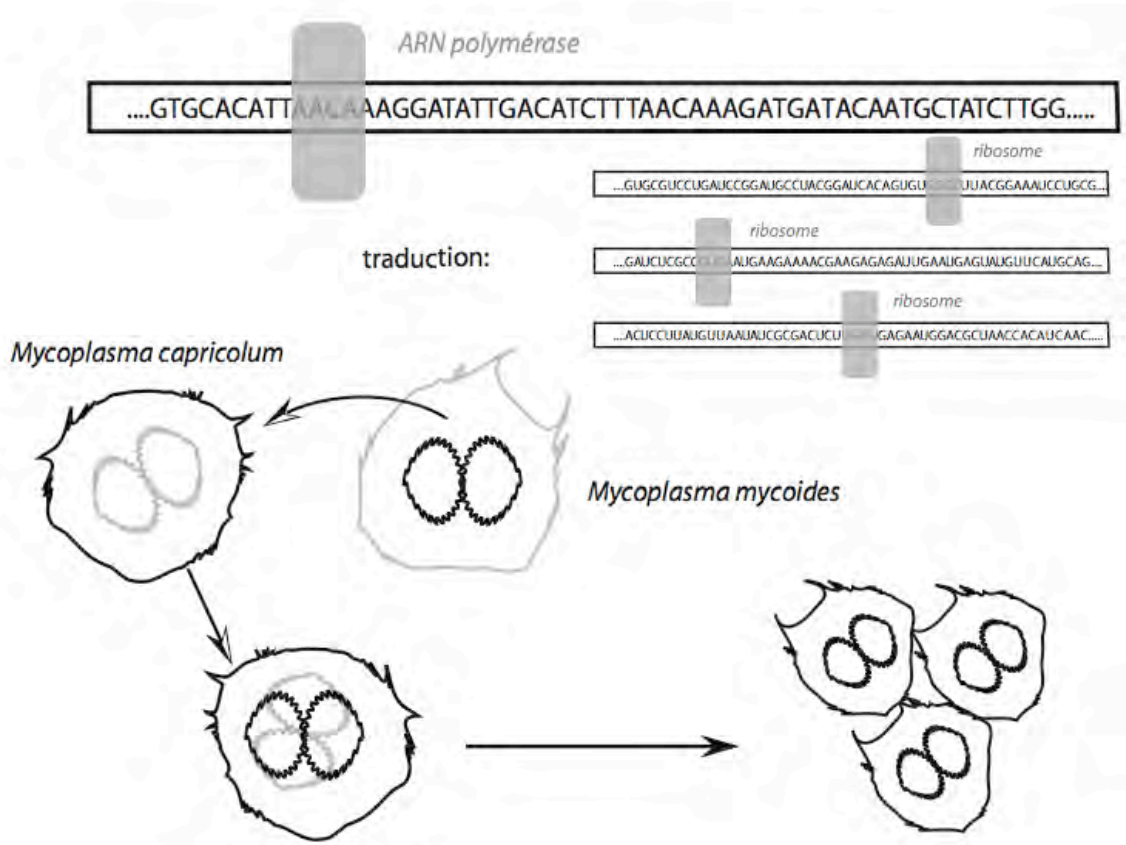
everything separates

"machine" (cell factory) et "data/program" (the genome)

lartigue-venter's demonstration

the turing machine may be parallelised

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

two types of information

standard information: carries its own forces along with it (e.g. information in DNA replication)

contextual information: the presence of a flame elicits a response, such as the flight of a moth, in the part of an external, autonomous agent which provides all necessary forces and energy

it is in this second case that we can best see information as a category distinct from mass or energy; the theory does not exist yet…

information of the machine

"beside the genetic program, the cell carries a considerable amount of information…"

true: but in a computer as well

this requires construction of an entirely novel theory of "machine-information"

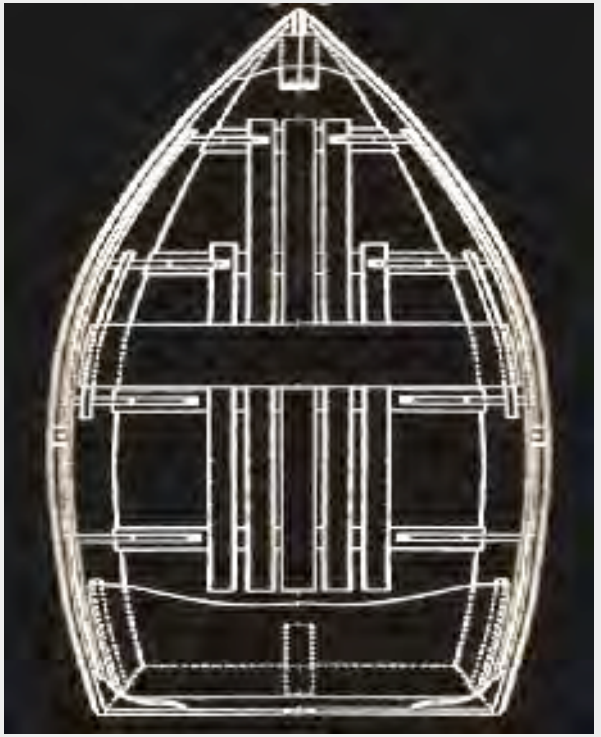
in a computer, do not forget the physical support

it is not enough to have a DNA molecule with the right sequence, it needs to be correctly folded!
also, this implies that animal cloning is a further support of the cell as a computer hypothesis

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

t h e d e l p h i c b o a t

- biology is a science of relationships between objects
- it is symplectic (συν together, πλεκτειν, to weave), same word as « complex » in latin;
- it is an information that expresses what is conserved in the boat, not the matter of its planks !



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

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

symplectic biology

rapid research notes in systems and synthetic biology

News

Summer Course on Synthetic Biology in Valencia, Spain

Mar 06, 2010

Orthogonal translation via evolution of a quadruplet-decoding ribosome

Feb 23, 2010

Biomimetic studies inspire japanese urbanists.

Jan 29, 2010

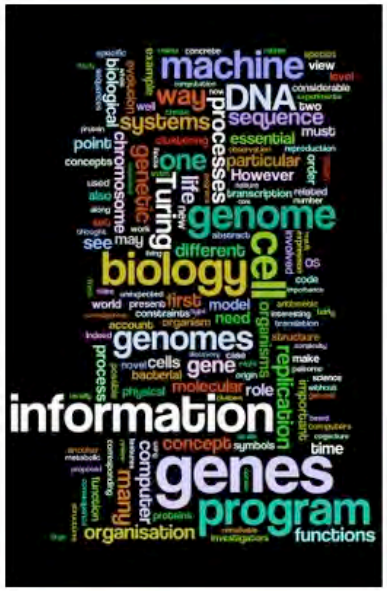
Non-retroviral virus sequences in mammalian genomes.

Jan 28, 2010

More news...

Welcome to Symplectic Biology

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



Featured Articles

Under Construction

Challenges

Under Construction

Aims and Scope

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

Search

Search Site | Search

Advanced Search...

Editorial Board

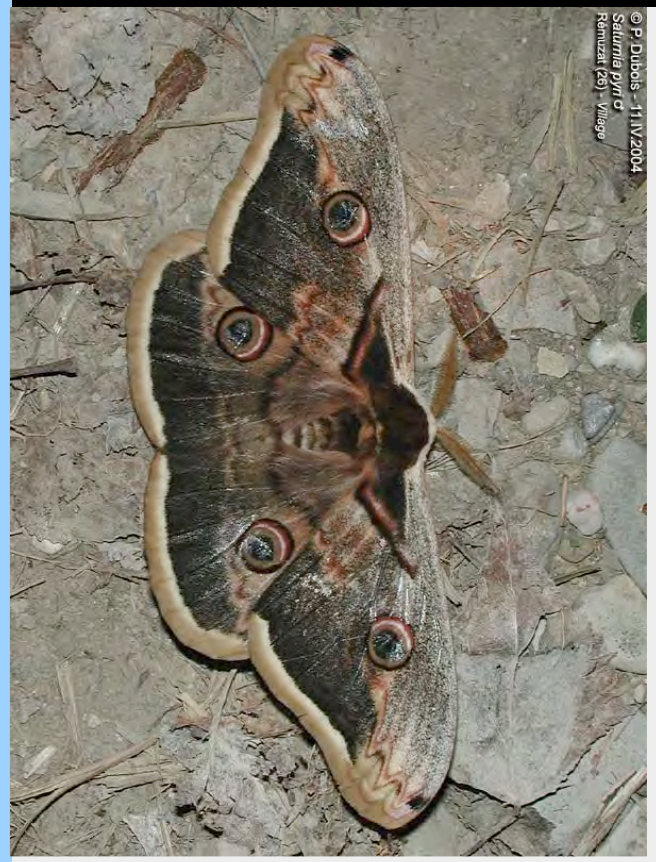
Editors in chief

Antoine Danchin
Victor de Lorenzo

Editors

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

exploring information: infotaxis



large peacock *saturnia pyri*
http://pdubois.free.fr/

how can a moth find a partner 1,000 meters away?

to climb a chemical gradient is impossible (air turbulence, obstacles…)

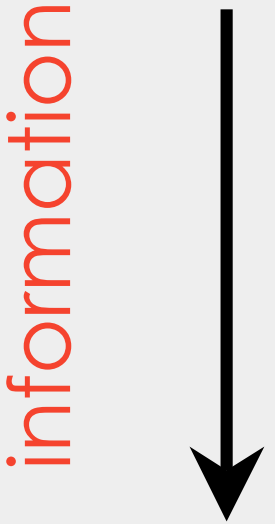
vergassola and co-workers have shown that maximising the collection of information allows the animal to achieve this goal…

'Infotaxis' as a strategy for searching without gradients
Vergassola M, Villermaux E, Shraiman BI
Nature (2007) 445: 406-409

a new heuristics

matter / energy / space / time

- classic physics
- quantum physics
- chemistry
- biology
  - development
  - neurobiology
  - linguistics
- mathematics (informatics)



information as a new currency of reality

- 1929 leo szilard (wrong) intuition of the link between energy and information: creation of 1 bit requires kTlog2, while analysing maxwell's demon
- 1949 claude shannon theory of communication
- **1**961 rolf landauer proof of computation reversibility (no energy is required for creation of information); energy is required for erasing memory
- ~1974 andrey kolmogorov, gregory chaitin, ray solomonoff define algorithmic complexity
- 1988 charles bennett defines logical depth (links time and algorithmic complexity) to define value of information
- 1989 wojciech zurek links algorithmic complexity and energy, reflecting on maxwell's demon
- 2007 scott muller defines information as any attribute that helps determine the state of a system, via asymmetry
- 2009 takahiro sagawa and masahito ueda reassess landauer's theorem

a bridge with matter and energy

- information does not derive from matter
  - orthogonal synthetic biology
- information does not derive from energy
  - creation of information is reversible
  - accumulation of information requires energy
- information bridges digital and analogic engineering

a minimal set of functions

looking for ubiquitous functions

variation / selection / amplification

↳ stabilisation ↱

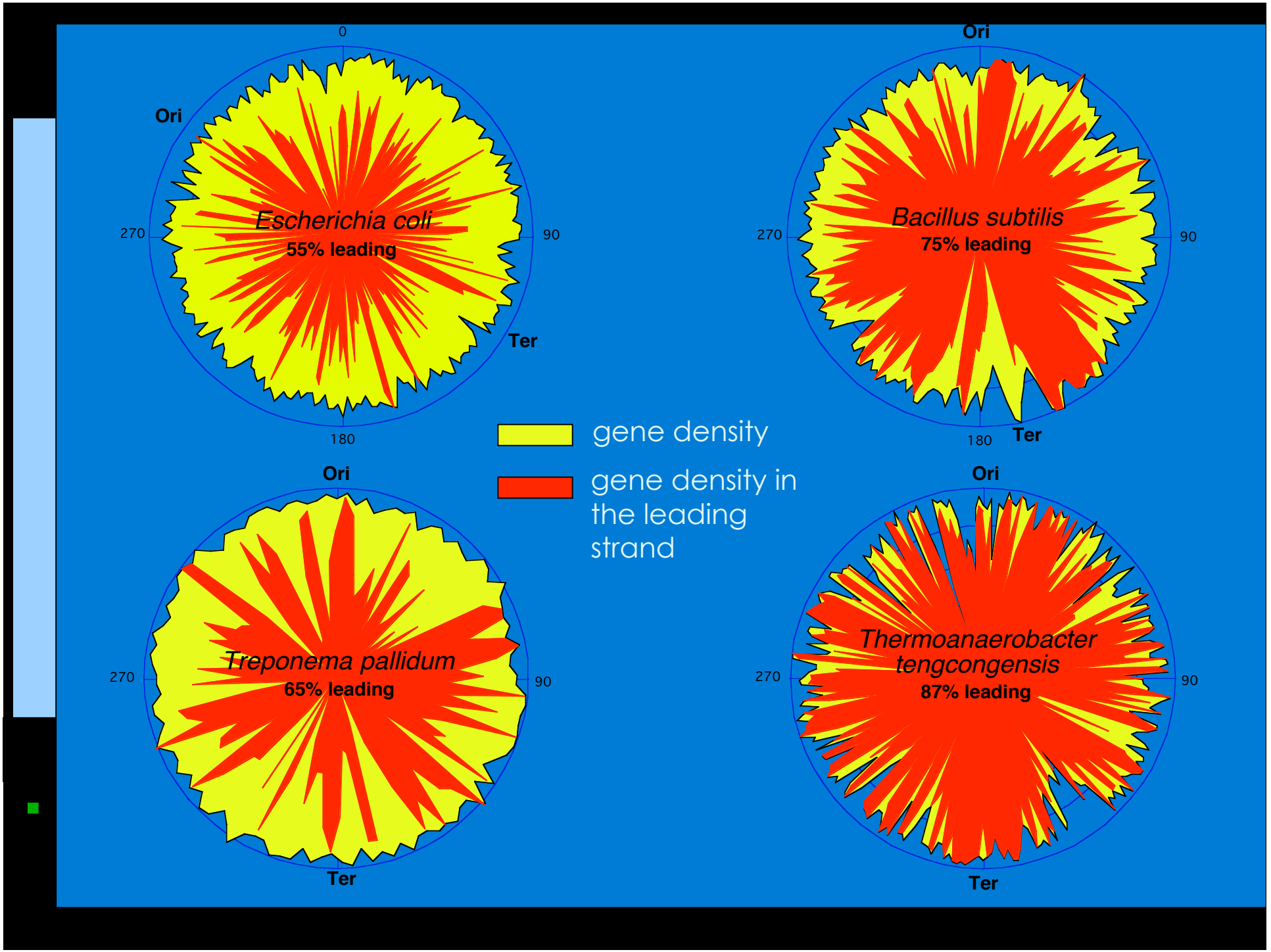
functional ubiquity does not imply structural ubiquity

from functional ubiquity to gene persistence

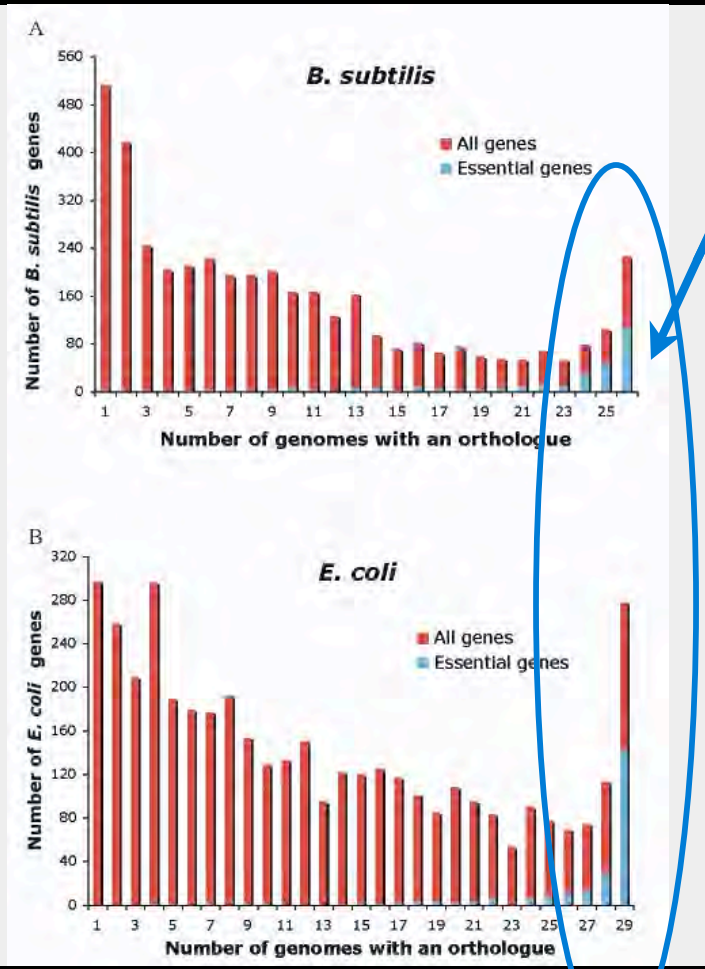
functional gene ubiquity does not imply gene structural ubiquity; efficient entities tend to persist through generations

- looking for « persistence » permits identification of most ubiquitous functions
- is « ubiquitous » synonymous to « essential »?
- genes that are « essential in the laboratory » are located in the leading DNA strand

~ 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


0
Ori
270
90
180
Ter
Escherichia coli
55% leading
Ori
270
90
180
Ter
Bacillus subtilis
75% leading
gene density
gene density in the leading strand
Ori
270
90
Ter
Treponema pallidum
65% leading
Ori
270
90
Ter
Thermoanaerobacter tengcongensis
87% leading

persistence: too many genes!



persistent genes

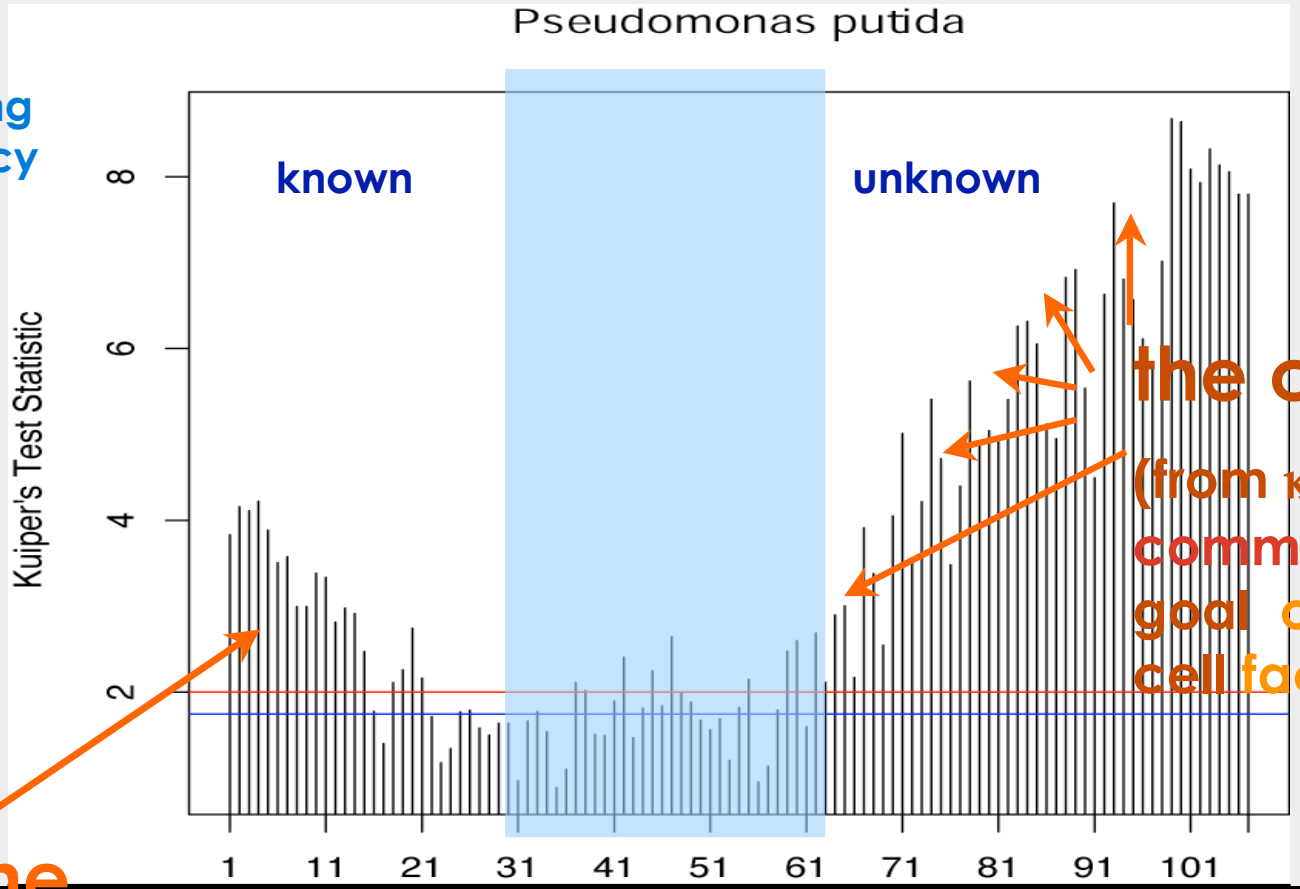
genes essentiels and ….

energy-dependent degradation

metabolic patches

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

organisation of bacterial genomes



the cenome

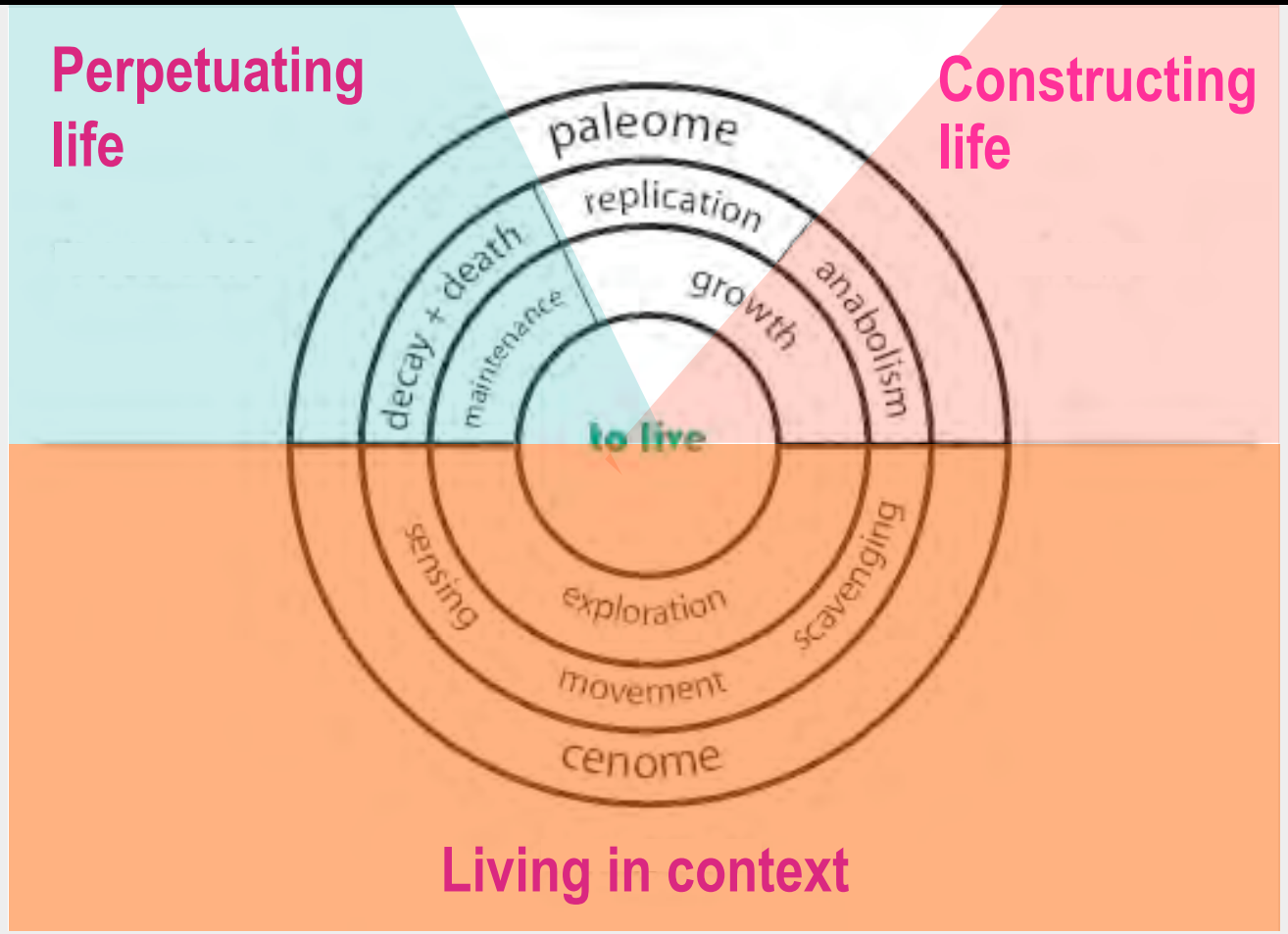
(from κοινος, common): the goal of the cell factory

paleome

(from παλαιος, ancient): the cell factory

a tale of two genomes

metagenomics



Σ = {pan-genome}

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

the paleome and the cenome

- the structure of the paleome
  - essential function; evolution of the gene expression machinery
  - energy-dependent degradation
  - sulfur metabolism anabolism, salvage, catabolism
  - chemical « frustration » (metabolic "patches")

- the cenome : occupation of a niche

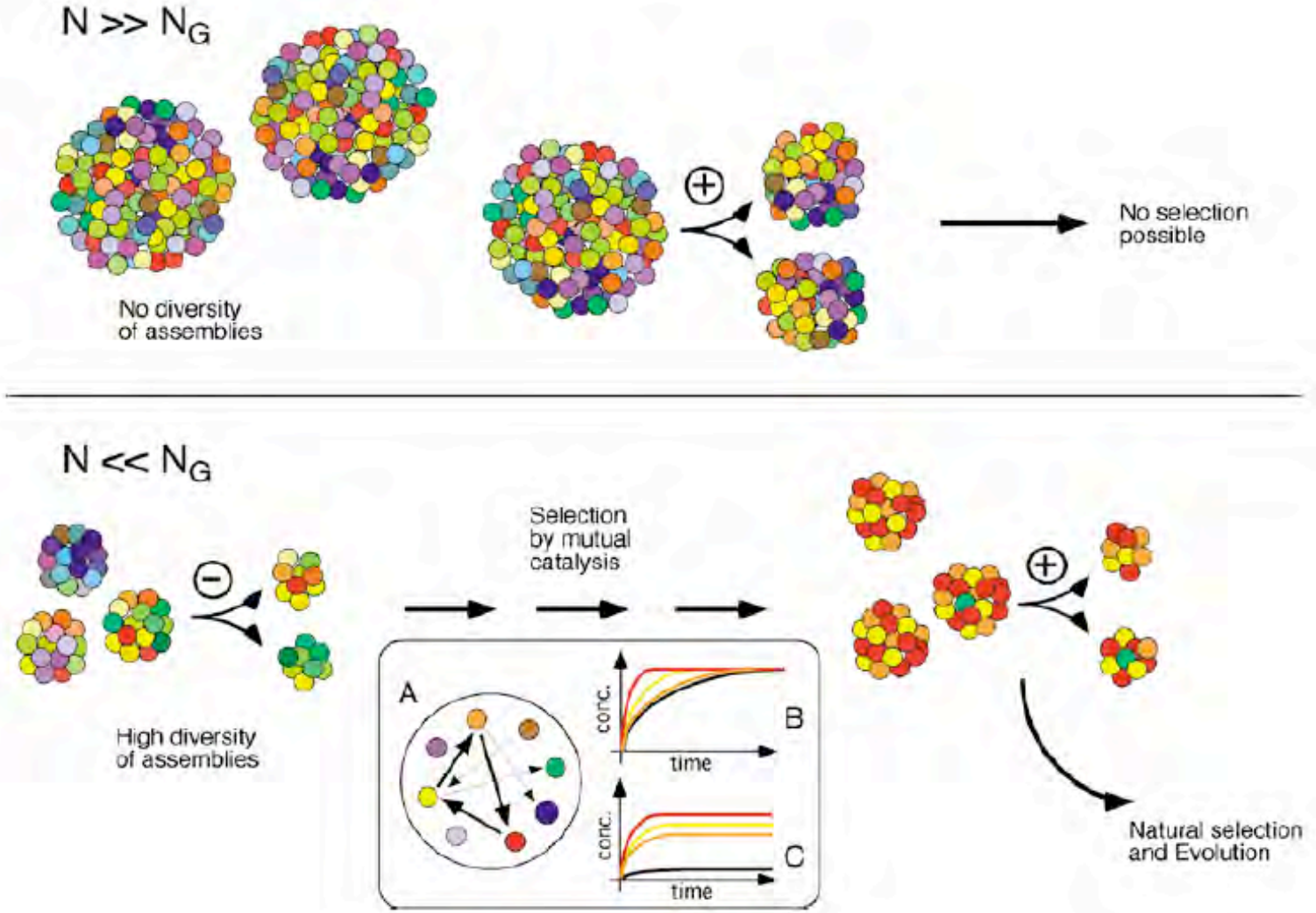
maxwell's demon's genes

babies are born very young!

- **the machine reproduces**
  - reproduction can improve over time: it is always an old organism that gives birth to a young one (this implies creation and accumulation of information)
- **the program replicates**
  - replication progressively accumulates errors

which genes permit accumulation of information?

contextual information and reproduction



r e v i s i t i n g i n f o r m a t i o n

intuition tells us that creation of information asks for energy

wrong: creation of information is reversible (landauer, 1961; bennett, 1982, 1988, zurek, 1989); to accumulate information requires an energy-dependent processs to "make room"

open question: "to make room" is necessary to accumulate information; how is this performed? can we identify in genomes the genes coding for the functions that permit this process? can we find a ubiquitous and stable energy source?

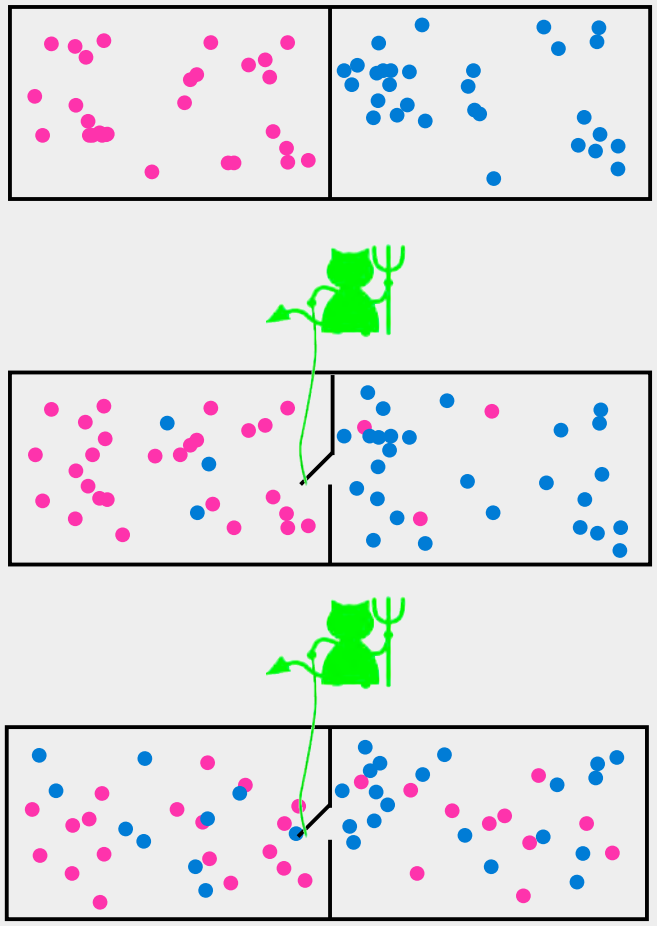
value of information

in the classical models of information one does not take meaning into account, nor the value of an information

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

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

m a x w e l l ' s d e m o n



the demon can accumulate information or reverse time if it can measure the speed and the position of the atoms of gas, collecting an information to calculate when it must close the trap

0

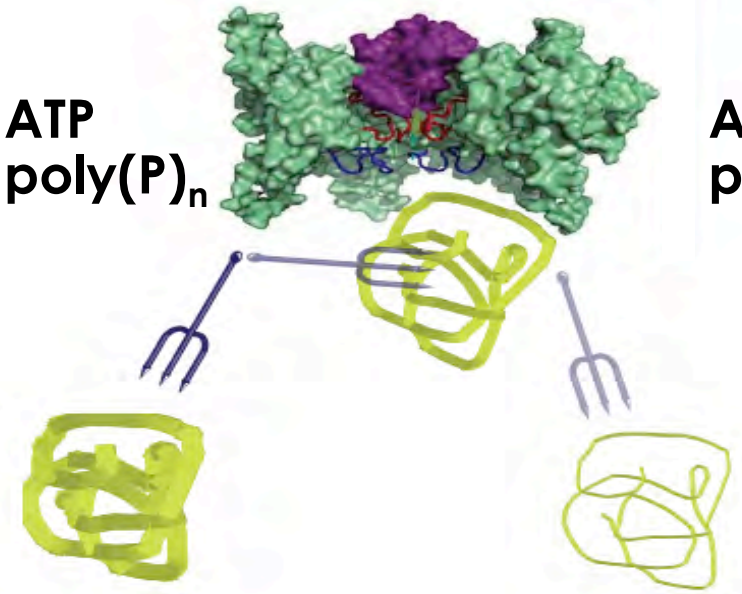
t

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

a ubiquitous ressource: polyphosphates

- synthesis and degradation of poly-P is encoded in the non essential persistent genes; this little known process is associated to the degradation of RNA
- poly-P is a **mineral**, therefore extremely stable; it is a ubiquitous component of cells
- NTPs can be regenerated from NMP and poly-P; protease Lon can use poly-P instead of ATP; NADP (anabolism) can be generated from NAD and poly-P…

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 weeks time, papillae appears that begin to grow and invade the medium, using supplied "unusable" nutrients. They derive from adaptive mutations

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

a d a p t i v e m u t a t i o n s

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

to live and to perpetuate life

- the paleome comprises both genes permetting the construction of the cell factory, and genes permetting counteracting ageing via the accumulation of a novel information in the progeny
- the 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

predictions

- bacterial persistence in a host depends on persistent non-essential genes
- initiation of cancer comes from cells (stem cells) that discovered adaptive mutations that permit them to generate an immortal progeny
- accumulation of information in the brain (memory and learning) implies the existence of processes to make room while preserving functional connections, in a way which must be energy-dependent

a synthetic cell?

- the engineering view of SB precludes that artificial cells be innovative
- it is possible to exclude the genes permetting accumulation of information
- the consequence is that, as all factories, the cell factory will age and will need to be systematically rebuilt
- this has a in-built societal benefit, as, by construction, 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

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