Maxwell's demon's genes:

Is it possible to construct a synthetic living cell?

Antoine Danchin 唐善・安東

Collège de France, May 5th, 2009

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 (including processes) 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

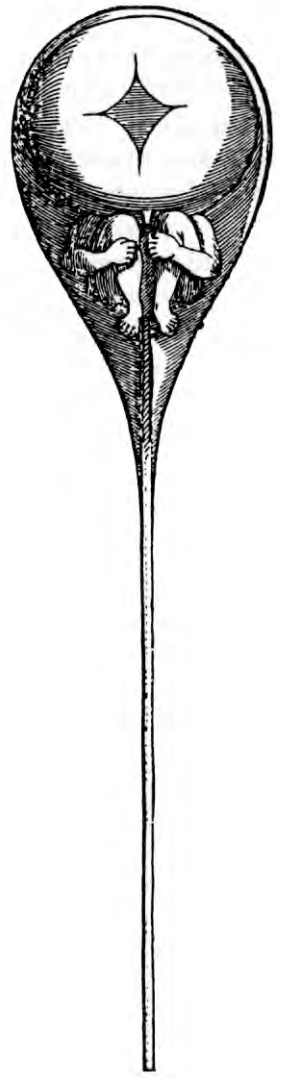
• **Evolving.** SB combines design and evolution to use (poorly understood) principles that drive adaptation

"Living organisms as information traps"

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

Synopsis

- A construction algorithm, not an organism, is the entity transmitted through generations
- The machine running the algorithm is physically separated from the algorithm's support
- This process is conceptually similar to that unfolding in computers
- Yet, computer do not (at present) make computers: how is the ageing process tackled?
- Comparative analysis of genetic programs permits discovery of the genes necessary to make young organisms from old ones. It also present living organisms as information traps



Computers making computers?

“Living organisms as information traps”

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

A 3D self-reproducing printer

Project RepRap (Replicating Rapid-prototyper, 2004) **aims at creating a laser 3D self-reproducing printer :**

- **The machine produces most of its components (= “biobricks”)**
- **What is missing :**
  - **The program**
  - **The assembly (managing space and time - sequence of events, and specific functions such as lubrication)**

http://reprap.org/

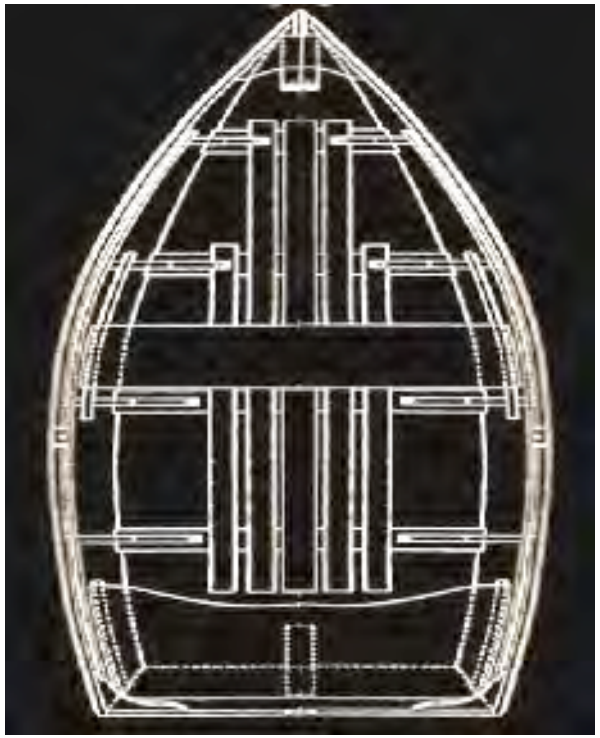
“Living organisms as information traps”

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

Information in the machine

- Biology is a science of relationships between objects

- 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

A. Danchin **Bacteria as computers making computers**. *FEMS Microbiology Reviews* 2009, **33**: 3-26

Making information concrete: infotaxis

Saturnia pyri

http://pdubois.free.fr/

How does a moth find a partner one kilometer away?

Climbing up a chemical gradient is impossible at such a distance (air turbulence, obstacles…)

Vergassola and co-workers have shown that maximising information collection permits reaching that goal…

'Infotaxis' as a strategy for searching without gradients

Vergassola M, Villermaux E, Shraiman BI

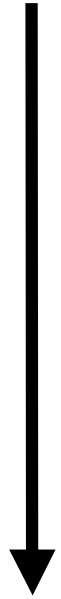
Nature (2007) **445:** 406-409

A fifth category of Reality

Matter / Energy / Space / Time

- Classical physics
- Quantum physics
- Chemistry
- Biology
  - Development
  - Neurobiology
  - Linguistics
- Mathematics

Information



"Living organisms as information traps"

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

Reproduction vs Replication

In his analysis of the physical constraints required to permit life to begin, Freeman Dyson has shown that we need to separate reproduction — which can improve over time — from replication — which is doomed to accumulate errors in the absence of external specific resources

Hence, reproduction must have predated replication

Reproduction supposes networks of interactions, and it is most easily accounted for as a type of compartmentalised metabolism

Replication works in a template-like fashion

We recognise here easily the general properties of life

Machine and program

Life requires:

- **A machine ("chassis" in SB jargon) allowing the program to be enacted (reproduces)**
  - **1. Metabolism (a dynamic process)**
  - **2. Compartmentalisation (casings, defining inside and outside)**

- **A program (a "book of recipes", which replicates)**
  - **3. Recursive information transfer and trapping => coding from one level to a second one as an essential element (fundamentally different from feedback and the like)**

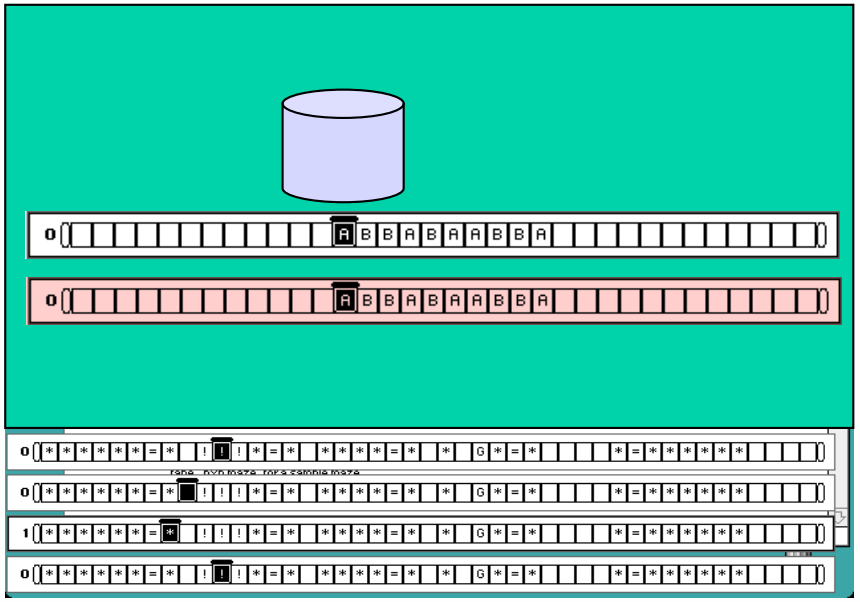
The cell is the atom of life

"Living organisms as information traps"

What computing is

Two entities permit computing:

- A machine able to read and write
- A program on a physical support, split by the human mind (not conceptually!) into two entities:
  - Program (providing the “goal”)
  - Data (providing some of the context)



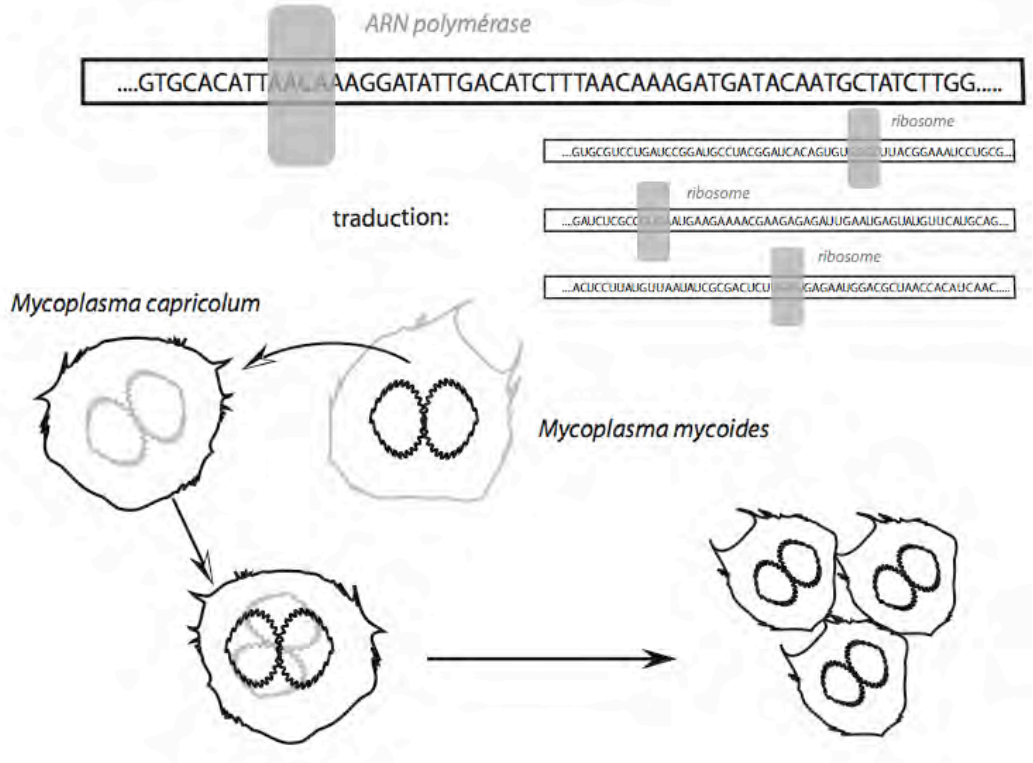
The machine is distinct from the data/program

“Living organisms as information traps”

Lartigue-Venter's demonstration

The Turing machine

May exist in a parallel set up



Genome transplantation

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

"Living organisms as information traps"

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

Objection to the computer model of the cell

“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”

“Living organisms as information traps”

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

Even in authentic computers, mind the physical support!

It is not enough to have a DNA molecule with the right sequence, it needs to be correctly folded!

"Living organisms as information traps"

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

Babies are born very young!

- **The machine reproduces**
  - Reproduction can improve over time: it is always an aged organism that gives birth to a young one (this implies creation of information)

- **The program replicates**
  - Replication keeps accumulating errors

Which genes permit accumulation of information?

"Living organisms as information traps"

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

Looking for ubiquitous functions

Variation / Selection / Amplification

↳ Stabilisation ↱

Evolution

↓ *creates (information comes in)*

Function

↓ *captures (recruits)*

Structure

↕ *codes*

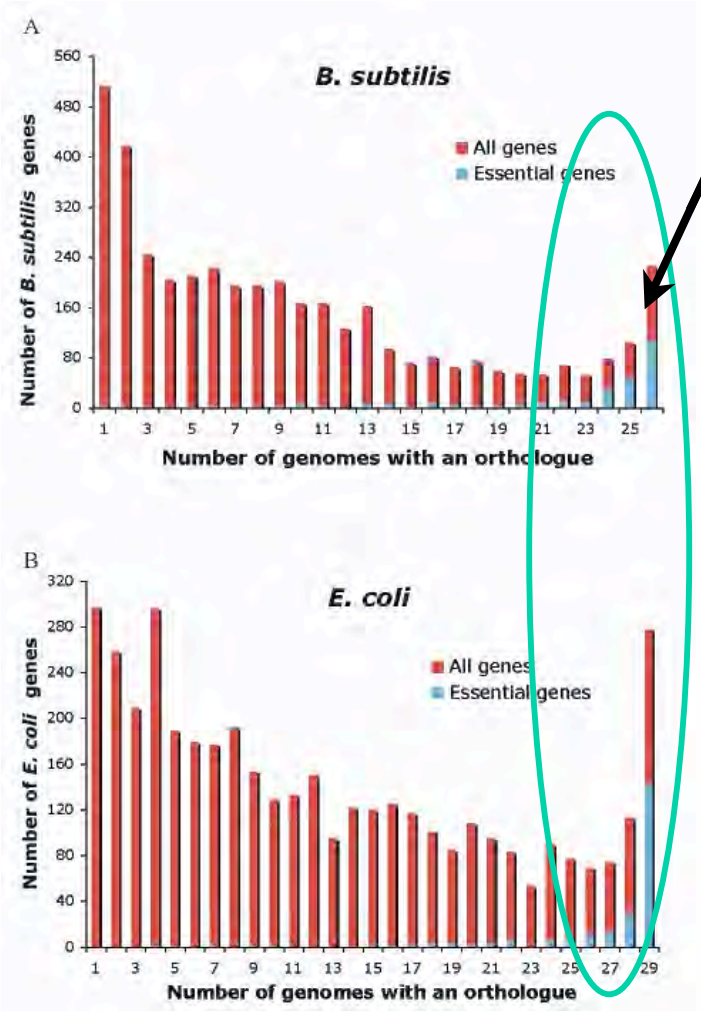
Sequence

Functional ubiquity does not imply structural ubiquity

"Living organisms as information traps"

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

Gene persistence: too many genes



Persistent genes

Essential genes and ….

Stress, maintenance and repair

Energy-dependent degradation

Metabolic patches (serine effect)

"Living organisms as information traps"

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

Persistent genes are clustered together

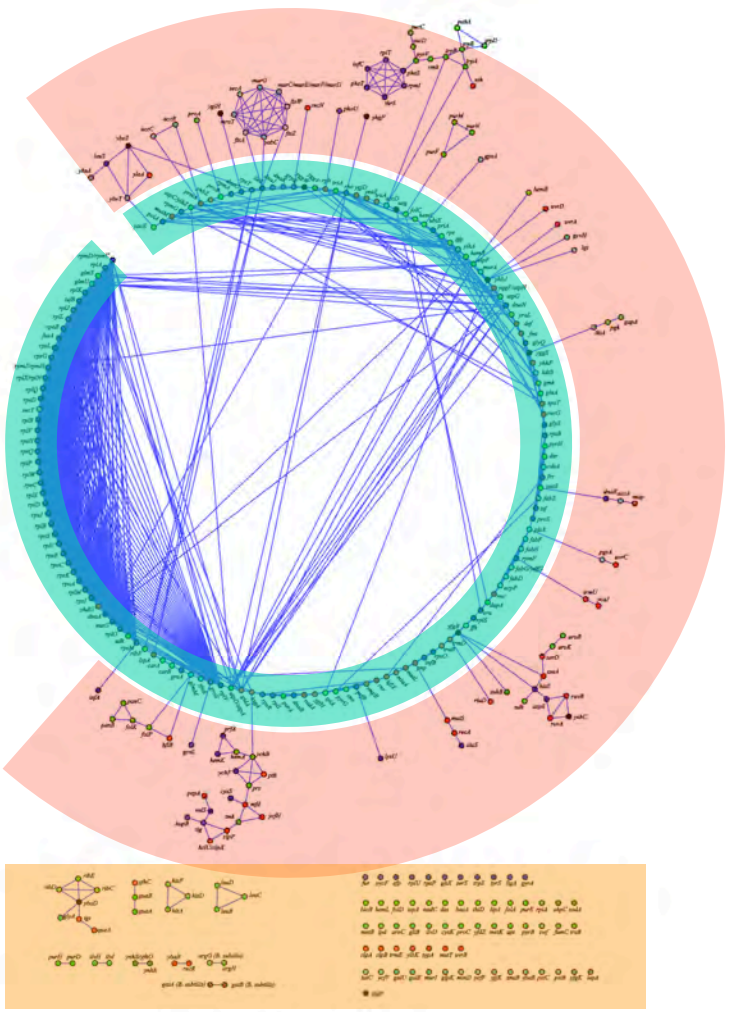
Persistent genes are functionally defined. They are located in the leading strand of DNA replication

Depending of their tendency to remain clustered in genomes (in > 250 bacteria with genome length > 1,500) they form three families that suggest a scenario of the origin of life

This group of core genes form the paleome (from παλαιος, ancient)

Persistent genes recapitulate the origin of life

The external network, made of genes of intermediary metabolism (nucleotides and coenzymes, lipids), is highly fragmented; the middle network is built around class I tRNA synthetases, and the inner network, almost continuous, organized around the ribosome, transcription and replication manages information transfers

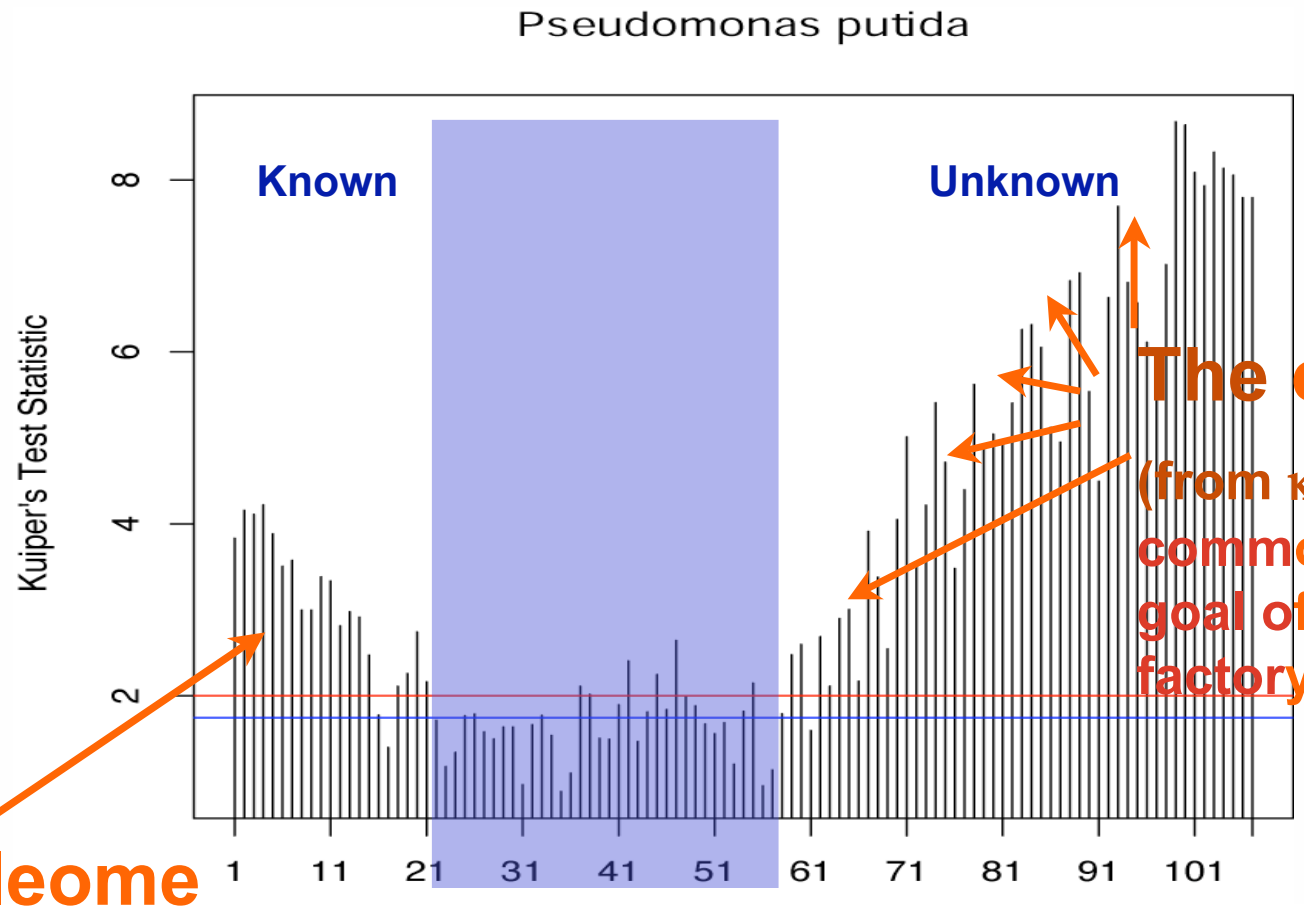


A Danchin, G Fang, S Noria
The extant core bacterial proteome is an archive of the origin of life
Proteomics. (2007) **7**:875-889

"Living organisms as information traps"

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

Organisation of bacterial genomes



Twice too many persistent genes

Functional ubiquity does not imply structural ubiquity

Yet, efficient objects tend to persist through generations:

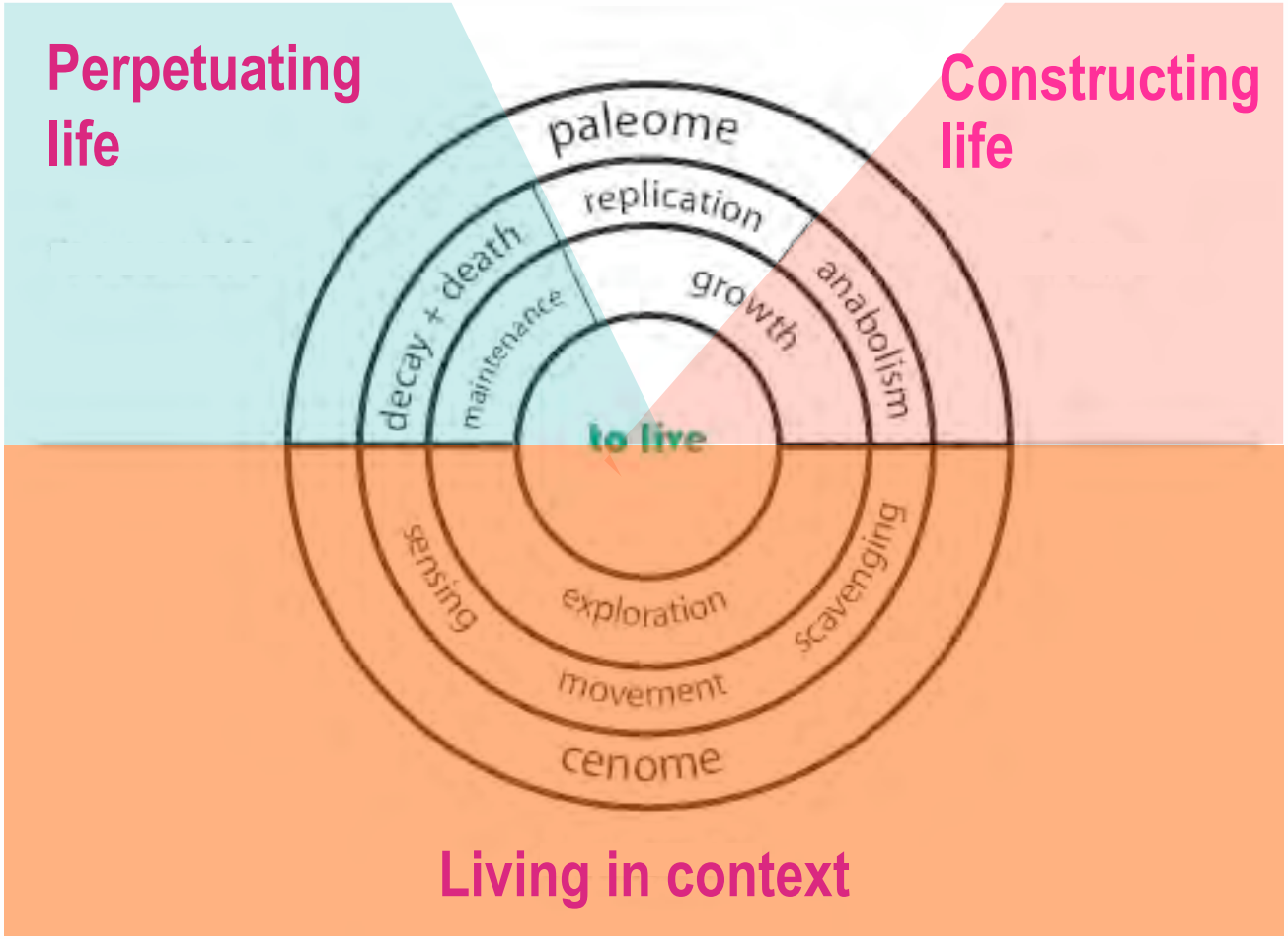
- Looking for « persistence » permits identification of (most) ubiquitous functions
- Is « ubiquitous » a synonym of « essential »?

~ 500 genes persist in bacterial genomes, forming the paleome; only ~250 are essential

A variable number permits to occupy a niche (cenome)

A tale of two genomes

metagenomics



"Living organisms as information traps"

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

A split paleome

- **Paleome 1 (essential genes)**
  - **Constructor**: DNA specifies proteins which form the machine that constructs the cell (reproduction)
  - **Replicator**: DNA specifies proteins that replicate DNA (replication)

- **Paleome 2 (persistent non-essential genes)**
  - Perennisation of life, energy-dependent degradation
  - Metabolic patches (chemical frustration)

"Living organisms as information traps"

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

Revisiting information

Intuition tells that creation of information requires energy. Yet, in an endeavour to calculate the limits of practical computation, Landauer demonstrated that creation of information is reversible (*i.e.* does **not** require energy: Landauer, 1961; Bennett, 1982, 1988); however, accumulating information requires an energy-dependent process to make room for this ratchet-like accumulation.

Open question: if « making room » is needed to accumulate information, how is it obtained? Can we identify in genomes the genes coding for the functions required to put this process in action? Can we find a ubiquitous and stable energy source?

Maxwell's demon's genes

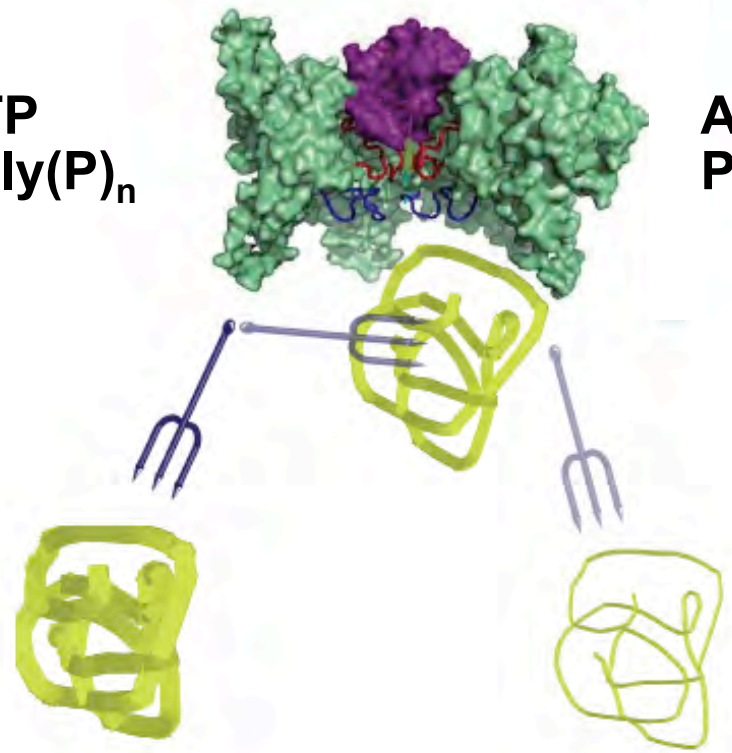
- Energy-dependent degradative processes make room for newly synthesised entities; energy is used to **prevent** degradation of functional entities
- This process accumulates information, whatever its origin, in a ratchet-like manner
- As this process is ubiquitous, we expect that the corresponding functions are encoded in the paleome, including management of the major energy sources postulated here

"Living organisms as information traps"

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

Maxwell's demon's genes

ATP
$Poly(P)_n$



ADP + Pi
$Poly(P)_{n-1}$ + Pi

The degradation machinery uses energy to reject a functional entity

Non functional entities are recognised and degraded

A prediction: Adaptive mutations

- Energy-dependent accumulation of information is blind; it cannot know where information will come from
- Information can just come from memory; it can also be created de novo
- **Prediction:** adaptive mutations are a de novo creation of information; they should depend on the genes involved in accumulation of information

Putting the conjecture to test: Adaptive mutations

Construction of "intelligent" bacteria

Placed to grow on a medium with limited nutrient supply. Form colonies of approximately 10^8 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

Natural selection traps information

- Energy-dependent degradative processes make room for newly synthesised entities; energy is used to **prevent** degradation of functional entities

- This process accumulates information, whatever its origin, in a ratchet-like manner

- Because the process is ubiquitous, the corresponding functions are expected to be coded in the paleome, including the possible energy source

- This process is **myopic**: it cannot have any grand design, hence the “tinkering” feature of the evolution of life

“Living organisms as information traps”

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

A synthetic cell?

- The engineering view of SB precludes innovation in synthetic cells
- It is possible to exclude genes permitting accumulation of information
- The consequence is that, as factories, cell factories will age and have to be systematically reconstructed
- This has the considerable societal advantage that the associated risks are minimised

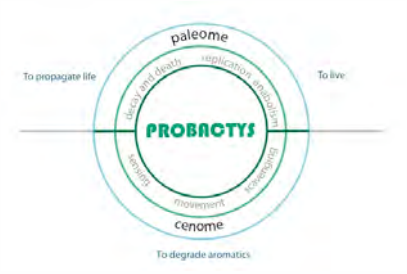
"Living organisms as information traps"

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

Authors and Funding

Gang Fang, Eduardo Rocha — in silico

Andrew Martens, Agnieszka Sekowska — in vivo



FONDATION
SCIENTIFIQUE
FOURMENTIN-GUILBERT

POUR LE RAYONNEMENT DE LA BIOLOGIE

INSTITUT PASTEUR

cnrs