INSTITUT PASTEUR

Nature and Artifice: A reflection on domestication and GMOs

Ambassade de France, Copenhagen
23 november 2006

Genetics of Bacterial Genomes
http://www.pasteur.fr/recherche/unites/REG/
adanchin@pasteur.fr

INSTITUT PASTEUR

Take home message

- Culture, not Nature, is the hallmark of humanity
  - Culture differentiates Man from Nature
  - Initially, humans protect Nature against humans (« *Tristes tropiques* » C. Lévy-Strauss); Artifice is still perceived as dangerous
  - Nature has no purpose, it is not friendly, but needs to be tamed
  - Domestication of plants and animals marks humanisation, starting the Neolithic Age
  - GMOs are the latest avatar of domestication
  - The fear of Artifice leads to the dangerous thought that plant GMOs are dangerous while animal GMOs are safe
  - Often, going to Artifice is safer
- The next revolution will be **synthetic biology**

adanchin@pasteur.fr

INSTITUT PASTEUR

A few dates

- 200,000-100,000 BP Birth of *Homo sapiens* ssp. Sapiens
- 17,000-12,000 BP Domestication of dog (*Canis canis*)
- 12,000 BP Domestication of fermentation microbes (lactobacilli and fungi)
- 10,000-9,000 BP Domestication of rice (*Oryza sativa*) and wheat (*Triticum aestivum*)
- 10,000 BP Domestication of cattle (*Bos taurus*)
- 9,500 BP Domestigation of pig (*Sus scrofa*)
- 9,000 BP Domestication of maize (*Zea mays* ssp. Parviglumis)
- 6,000 BP Domestication of horse (*Equus caballus*)
- 5,000 BP Domestication of silkworm (*Bombyx mori*)

- 1764 Kolreuter fertilizes plants artificially
- 1866 Mendel establishes the first laws of genetics
- 1940-1960 The Green Revolution (maize and wheat breeding, er, and seed development based on irradiation and chemical mutag
- 1973 The first bacterial GMO
- 1984 The first transgenic mice
- 1987 Quantitative Trait Loci are used to speed up selection of
- 1994 The first GMO food reaches the food market

1800 1 billion
1928 2 billion
1961 3 billion
1974 4 billion
1987 5 billion
1999 6 billion

- **What life is**
- Living and computing
- Biological information
- Genetically Modified Organisms

Genetics of Bacterial Genomes
http://www.pasteur.fr/recherche/unites/REG/
adanchin@pasteur.fr

INSTITUT PASTEUR

What life is

Three processes make life:

- **Information transfer => genomics decyphers the programme associated to the cell**

Forces coupling the genome structure to the structure of the cell:

- **Metabolism**
- **Compartmentalisation**

The cell is the atom of life, with two compartmentalisation strategies: a simple envelope (bacteria...), or the multiplication of membranes and skins (plants and animals...)

adanchin@pasteur.fr

INSTITUT PASTEUR

Information transfer manipulates strings of symbols

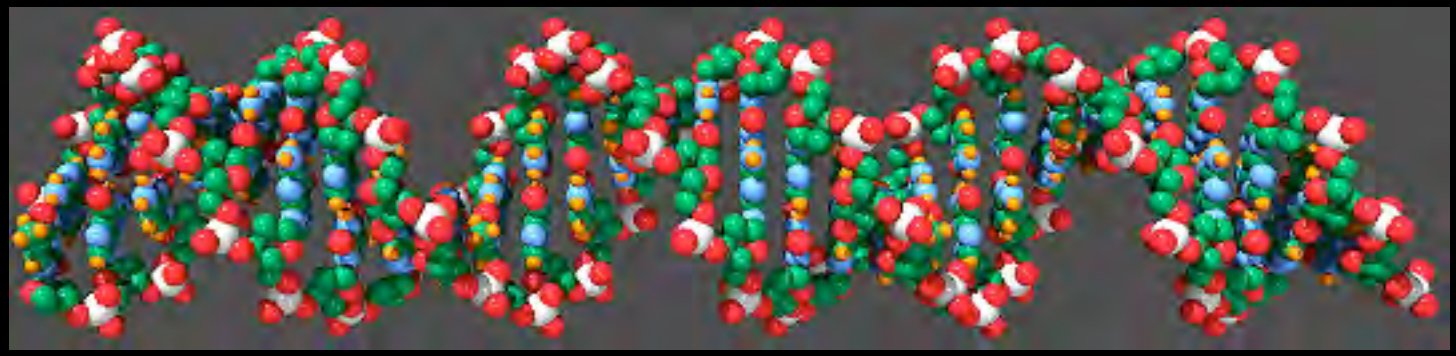
- **Replication** (law: “complementarity”)
- **Transcription** (law: “complementarity”)
- **Translation** (law: a cypher, the “genetic code”)

adanchin@pasteur.fr

INSTITUT PASTEUR

Information transfer

Deoxyribonucleic acid (DNA) stores the memory of the genetic programme. It chains four types of basic building blocks. Two strands are intertwined, making a double helix.



Remarkably, the genome programme reads like a text:

....GCGGTATTTTGATGGAGTTATACGGAAGGGATGTTC....

....CGCCATAAAACTACCTCAATATGCCTTCCCTACAAG....

Genetics of Bacterial Genomes
http://www.pasteur.fr/recherche/unites/REG/
adanchin@pasteur.fr

INSTITUT PASTEUR

Complementarity: DNA replication

DNA

The text of one half of the helix entirely specifies the text of the other half, as the positive and negative views of a photograph specify each other

A => T

C => G

T => A

G => C

HKU - PASTEUR RESEARCH CENTRE
香港大學—巴斯德研究中心

=>

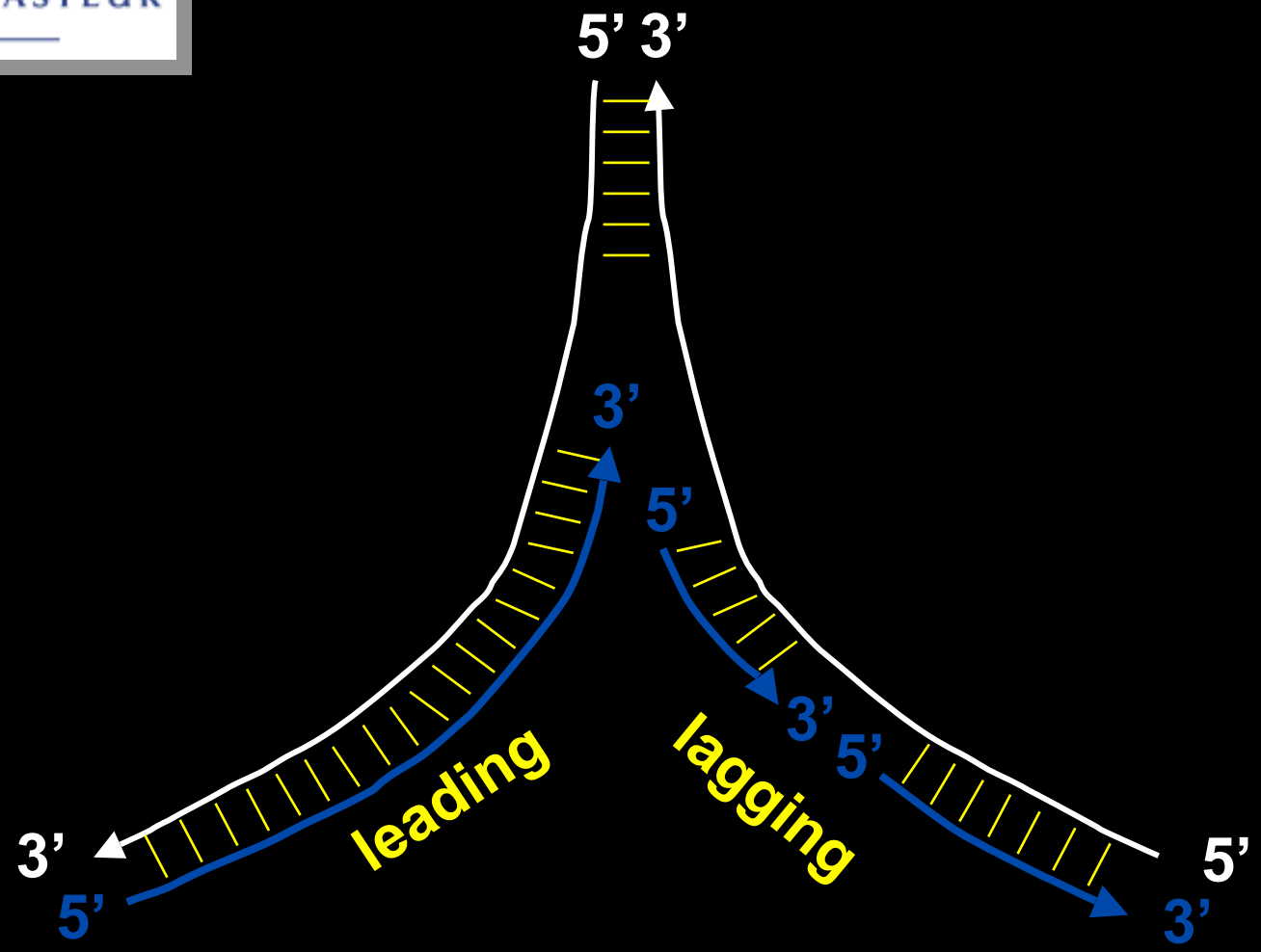
<=

HKU - PASTEUR RESEARCH CENTRE
香港大學—巴斯德研究中心

Errors result in evolution

adanchin@pasteur.fr

INSTITUT PASTEUR



adanchin@pasteur.fr

INSTITUT PASTEUR

The objects of the cell factory

The factory reads a « magnetic tape » DNA, which chains four types of chemicals, noted A, T, C and G, the bases.

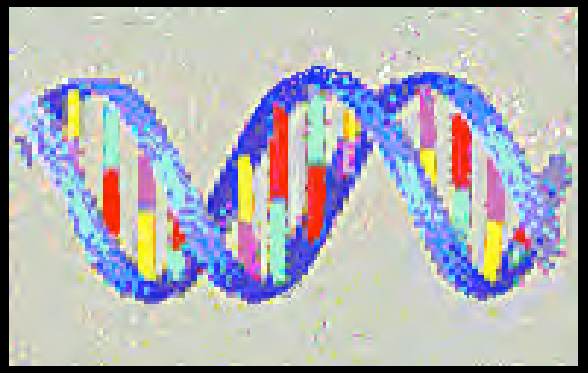
The machine is enclosed in a solid casing, the membrane (connecting the inside with the outside) and small « heads » read the message and transform it into active component, the proteins. The proteins are also chains of standard chemicals, of twenty types, the amino acids.

adanchin@pasteur.fr

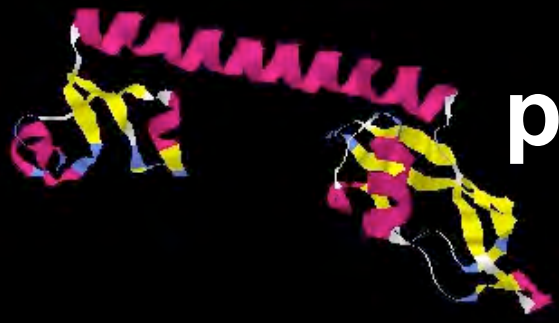
INSTITUT PASTEUR

From the genetic programme to the cell

When the machine reads the programme, it performs actions. A special machinery reads the DNA and copies it into active objets, the proteins (enzymes are proteins).



DNA



protein

adanchin@pasteur.fr

INSTITUT PASTEUR

The genetic code

Only meaningful parts of the four-letter DNA text (genes) are translated into a twenty letters text, the text of a protein. The cypher driving the correspondence is named the genetic code.

For example “ TCA ” in DNA means “ S ” (a chemical residue, “Serine”, among the twenty amino acid types) in the protein.

adanchin@pasteur.fr

INSTITUT PASTEUR

The cell factory

A cell looks like a computer which would programme the construction of similar computers

It has a magnetic tape, or hard disk (the « genetic programme ») and reading devices which allow it to read the programme and put it into action

The « cloning » of the ewe Dolly performed that action: moving the programme from a machine (a cell) to another one (an egg without a nucleus)

adanchin@pasteur.fr

INSTITUT PASTEUR

- What life is
- Living and computing
- Biological information
- Genetically Modified Organisms

Genetics of Bacterial Genomes
http://www.pasteur.fr/recherche/unites/REG/
adanchin@pasteur.fr

INSTITUT PASTEUR

Information Transfer

As for building up a machine, one needs a book of recipe to build up a cell.

This construction changes the text of the recipe into something concrete: this transfers « information ».

In a cell, information transfer is managed by the genetic programme

adanchin@pasteur.fr

INSTITUT PASTEUR

Context: the "genetic programme"

- **Physics:** ***matter, energy, time***
- **Statistical physics:** **Physics +** ***information***
- **Biology:** **Physics +** ***information, coding, control...***
- **Arithmetics:** ***sequences of integers, recursivity, coding…***
- **Computation:** **Arithmetics +** ***programmes*** **+** ***machine...***

The « genetic programme » metaphor has practical consequences: we know how to manipulate genes and gene products, can we push the metaphor to its ultimate consequences?

adanchin@pasteur.fr

INSTITUT PASTEUR

A genetic computer

- In a computer the machine is separated from the data and the programme
- Data and programme play the same role (*i.e.* they can be thought of as ‘ declarations ’)

adanchin@pasteur.fr

INSTITUT PASTEUR

What is computing?

Two processes are needed for computing:

- A machine able to read and write
- A programme on a physical support (a perforated or magnetic tape illustrates the sequence of symbols that make the programme), split (in practice, but not conceptually) into two entities :
  - Programme (providing the goal)
  - Data (providing the context)

The machine is distinct from the programme

adanchin@pasteur.fr

INSTITUT PASTEUR

Cells and computers

Genetics describes genomes as texts written with a four letter alphabet: do cells behave as computers?

Horizontal Gene Transfer
Viruses
Genetic engineering
Animal cloning

all points to separation between

« Machine » (the cell factory)
and
Data + programme

adanchin@pasteur.fr

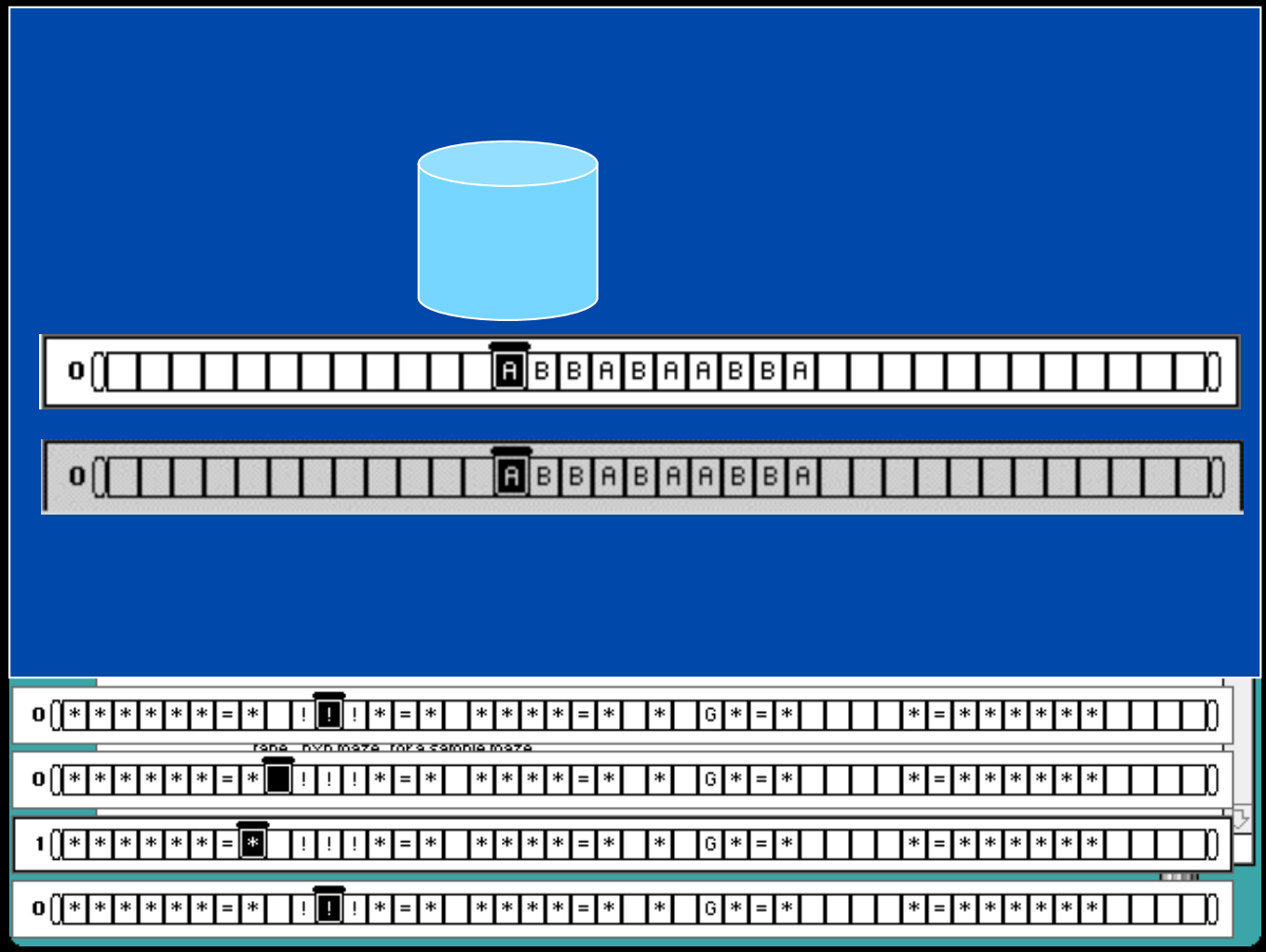
INSTITUT PASTEUR

The Turing machine

machine (read/write)

programme (data)

in the form of a linear sequence of symbols



Genetics of Bacterial Genomes
http://www.pasteur.fr/recherche/unites/REG/
adanchin@pasteur.fr

INSTITUT PASTEUR

An algorithmic view of biological actions

Replication, transcription, translation: high parallelism

"Begin, Repeat, Check Points, End"

The action is always oriented, with a beginning and an end

The processes of time control (check points) are essential to allow coordination of multiple actions in parallel

Genetics of Bacterial Genomes
http://www.pasteur.fr/recherche/unites/REG/
adanchin@pasteur.fr

INSTITUT PASTEUR

- What life is
- Living and computing
- Biological information
- Genetically Modified Organisms

Genetics of Bacterial Genomes
http://www.pasteur.fr/recherche/unites/REG/
adanchin@pasteur.fr

INSTITUT PASTEUR

Three types of information

Biology is an abstract science (!) and you need to prepare to make a considerable effort of abstraction to understand it

Information is at its core, but what is information?

Three types of information (not exhaustive!):

1. Shannon's information (Shannon's entropy)
2. Algorithmic complexity
3. Logical depth
4. Critical depth

adanchin@pasteur.fr

INSTITUT PASTEUR

Shannon's entropy (1)

The information in a string of symbols (a sequence) can be appreciated by some measure of the probability of occurrence of particular motifs

A commonplace description of genome sequences calls « entropy » the information of the sequence, assuming that life is a fight against entropy: both are wrong assignments

Caveat: **Myron Tribus relates that von Neumann, to whom Shannon had turned to help him find a name for his information, proposed prophetically: "*You should call it entropy for two reasons. In the first place your uncertainty function has been used in statistical mechanics under that name, so it already has a name. In the second place, and more important, no one knows what entropy really is, so in a debate you will always have the advantage*", thus opening a Pandora's box of intellectual confusion.**

adanchin@pasteur.fr

INSTITUT PASTEUR

Shannon's entropy (2)

Shannon's probability law is similar to that written on the grave of Ludwig Bolzmann. This is the source of enormous confusion

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

The validity of this formula rests on very strong hypotheses about the nature of the signals (for example, the signals should fit standard Laplace-Gauss probability laws)

adanchin@pasteur.fr

INSTITUT PASTEUR

Algorithmic complexity

- Shannon's « entropy » does not care about the meaning of a sequence (it can represent replication, however)
- Algorithmic complexity is the length of the shortest programme needed to generate the sequence. A sequence is random if it cannot be described by a programme with a length shorter than the sequence (Kolmogorov and others); e.g. 0101010101... has a low complexity
- Bacterial genomes look « random »; plant and animal genomes look « repeated » => both look « uninteresting »!

adanchin@pasteur.fr

INSTITUT PASTEUR

Unexpected properties of strings of symbols

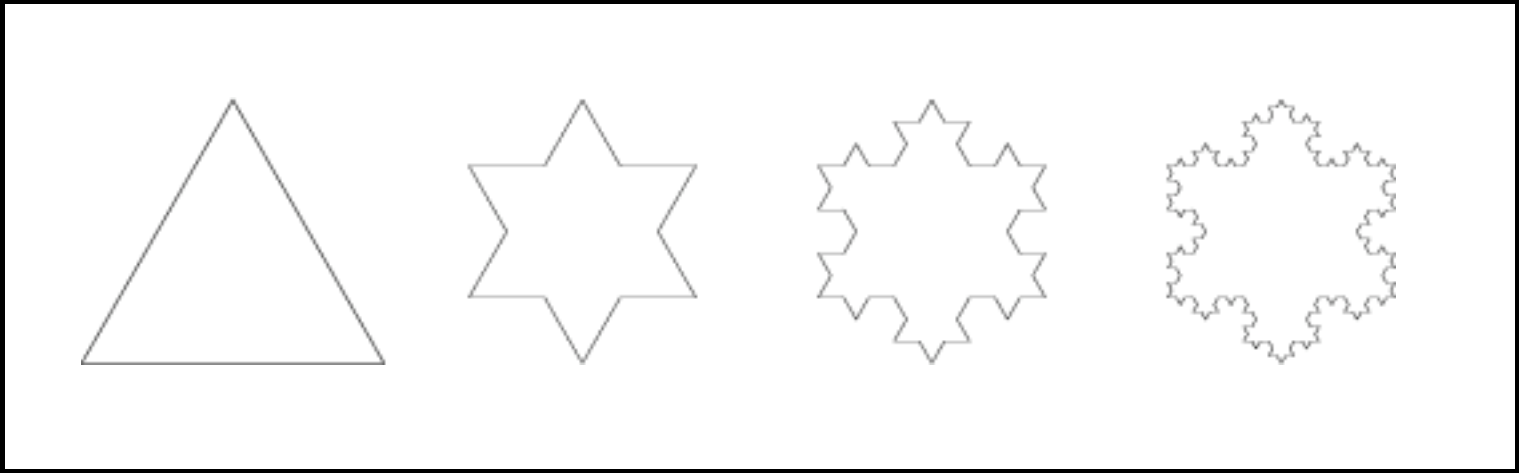
- The DNA « text » can be considered as a string of symbols
- Strings can play the role of « programmes » with remarkable properties (e.g. when « recursive », i.e. calling themselves as routines). A machine run by a recursive programme usually does not have standard mechanistic (i.e. predictable) properties

adanchin@pasteur.fr

INSTITUT PASTEUR

Logical Depth (1)

A very short programme (low algorithmic complexity) can describe a repeated sequence (trivial), but also a fractal figure such as Koch's snowflake: as shown in the figure, this is a stepwise process, asking for time to be introduced into the concept of information



adanchin@pasteur.fr

INSTITUT PASTEUR

Logical Depth (2)

The consequence of logical depth is that, when a programme is complicated enough (branching and recursive), it becomes impossible to predict its outcome. The only way to know it is to run the programme…

Evolution has evolved DNA from DNA, from DNA... in such a way that every single base has a certain « depth », that makes that living organisms are, in principle and by construction, poised to be ultimately unpredictable…

This allows them to create some progeny that can survive in an unpredictable future.

adanchin@pasteur.fr

INSTITUT PASTEUR

- What life is
- Living and computing
- Biological information
- Genetically Modified Organisms

Genetics of Bacterial Genomes
http://www.pasteur.fr/recherche/unites/REG/
adanchin@pasteur.fr

A general constraint hindering engineering: evolution

Variation / Selection / Amplification

Stabilisation

Evolution

creates

Function

capture (recruits)

Structure

codes

Sequence

adanchin@pasteur.fr

INSTITUT PASTEUR

Plant GMOs: improving crop yields and product quality

- After Kolreuter fertilized plants, Mendel found how characters are inherited, starting the first green revolution. **Crosses between variants permitted screening for new breeds**
- **Plants are mutagenized** (often irradiation) to produce more variants
- **Variants are unpredictable**. Together with « useful » mutations they carry **many** other mutations, often deleterious, requiring a long selection process to obtain the required plants
- **Genetic engineering modifies only one, or a small number** of targeted genes, directly going to the desired properties. **GMOs are more predictable than mutants obtained by the more traditional (also artificial!) approaches**
- Genetic engineering aims at stabilizing the phenotype: this goes against the ability of plants to evolve. Hence, **GMOs are less fit to the non-humanized (wild) environment**

adanchin@pasteur.fr

INSTITUT PASTEUR

Animal GMOs: humanisation of organs

- The initial goal of animal GMOs was similar to that of plant GMOs, improving quality of animal products, or using the animal as a factory
- Interest for health care created a new **Unmet Need**, associated to phylogenetic proximity between animals and humans: using GMOs as providing tissue substitutes.
- Humanizing organs is widely accepted by the public, despite its obvious danger:
  - **Animal tissues contain retroviruses**
  - **Emerging diseases are often caused by transfer from animals to humans (eg HIV, SARS, Hantah etc)**

Nature is preadapted through evolution, it can be more dangerous in principle than Artifice

Genetics of Bacterial Genomes
http://www.pasteur.fr/recherche/unites/REG/
adanchin@pasteur.fr

INSTITUT PASTEUR

Nature and Artifice

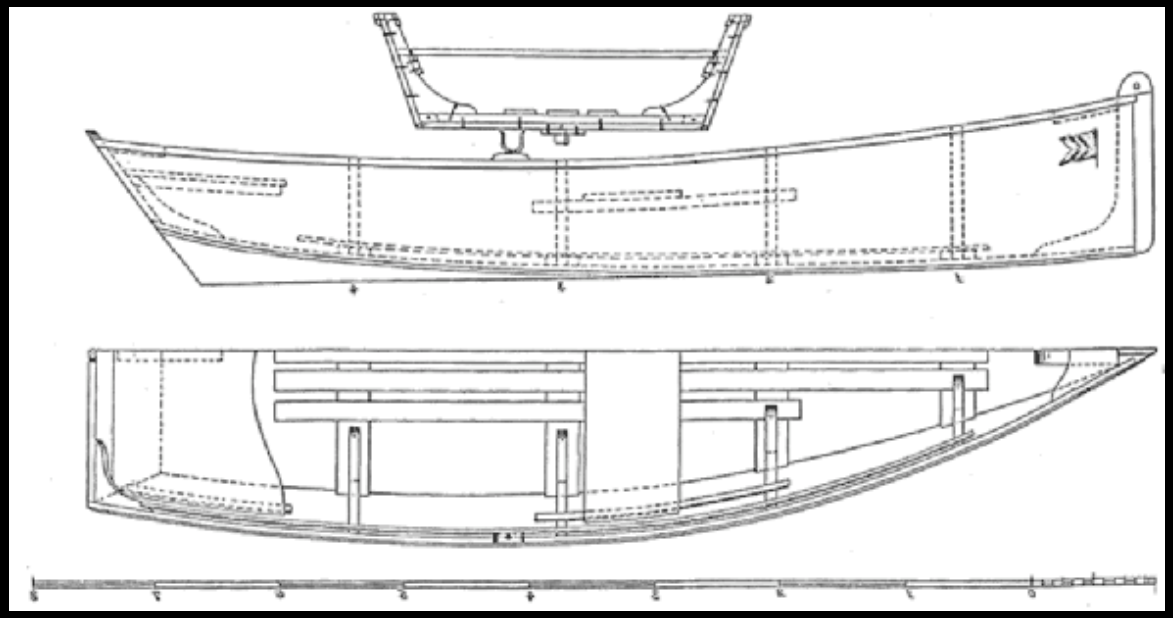
- Biological adaptation is the easier, the closer the organism from humans
- Coming from Humans, human blood is dangerous
- Natural processes are more difficult to control than artificial processes
- Progress in applications of biological knowledge is linked to increasing our control over the processes, i.e. getting ever more artificial, e.g. inventing efficient artificial blood would be an immense progress (cf HIV, BSE, Hepatitis, etc)
- Future progresses will go to « synthetic biology », where the evolution potential of artificial organisms will be lower

adanchin@pasteur.fr

INSTITUT PASTEUR

Caveat: The Delphic Boat

- We spoke of isolated genes
- But genes do not operate in isolation!
- Proteins are part of complexes, as are parts in an engine
- It is important to understand their relationships, as those in the planks which make a boat



adanchin@pasteur.fr

Thank you

Genetics of Bacterial Genomes
http://www.pasteur.fr/recherche/unites/REG/

adanchin@pasteur.fr