

High Throughput sequencing and  
biological research:

Paradigm switches from genetics to  
ecophysiology

Sébastien Massart

Gembloux Agro Bio Tech  
University of Liège



gembloux  
agro bio tech

Université  
de Liège





90'ies

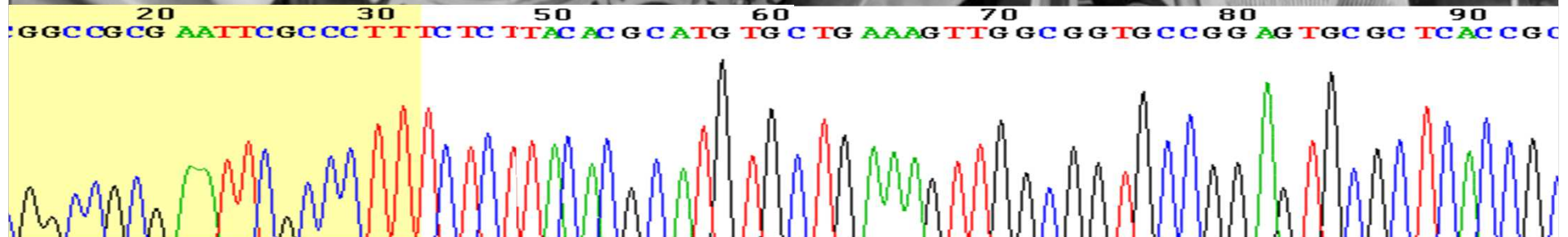
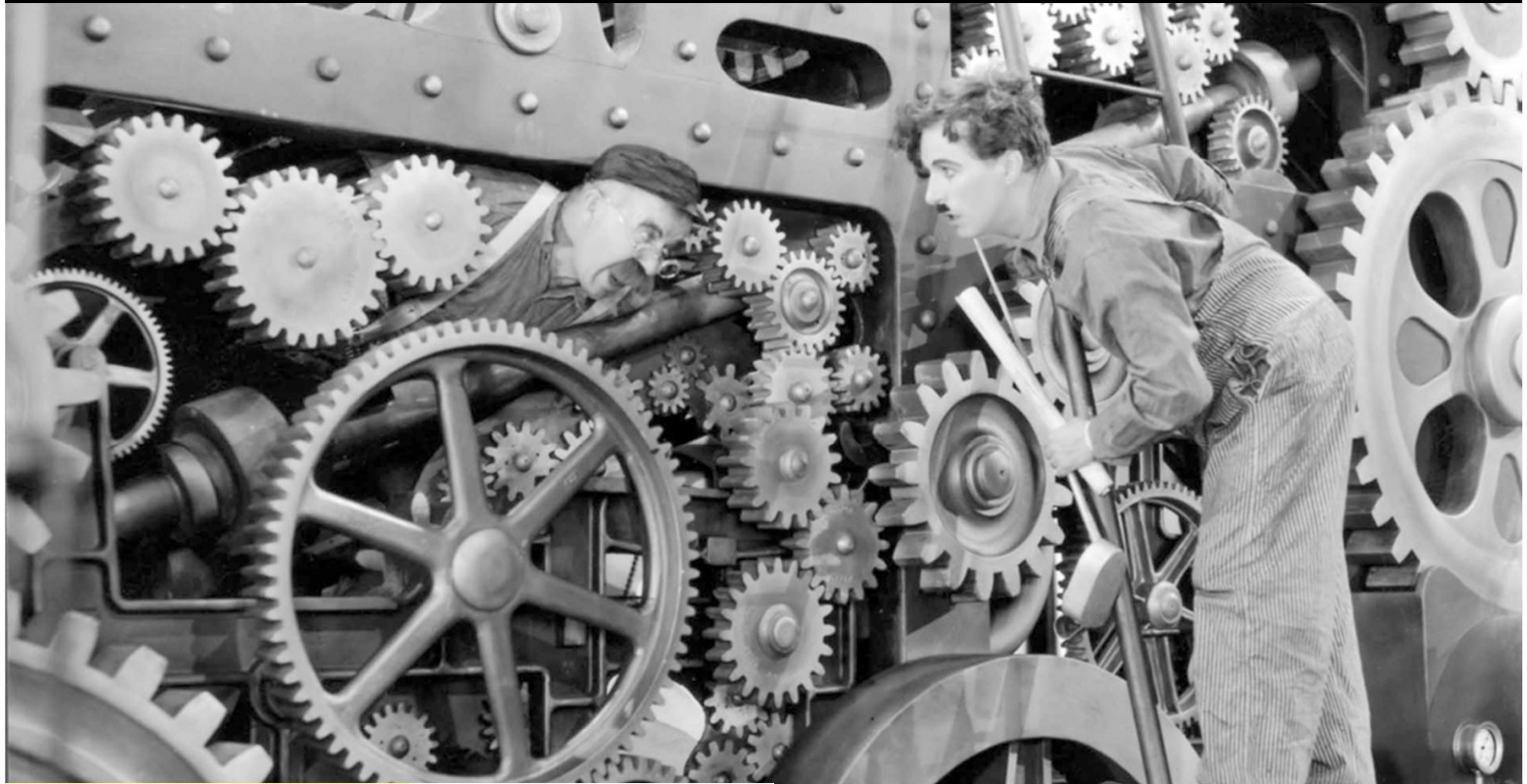
1.000

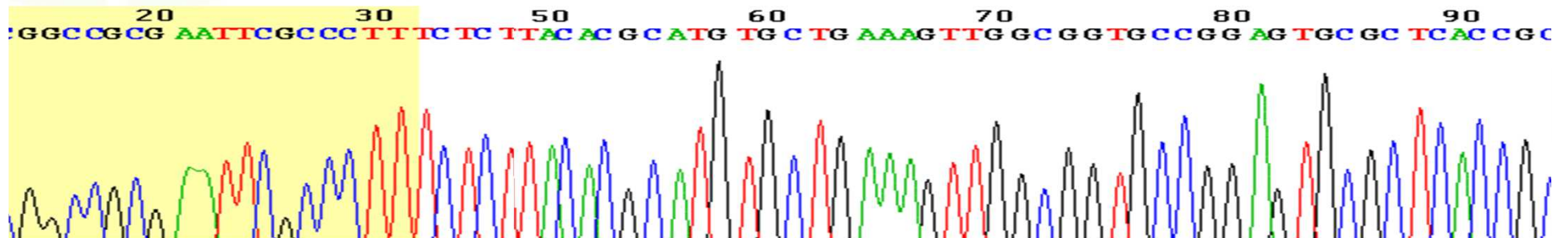




2000'ies

1.000.000







2014

4.500.000.000.000

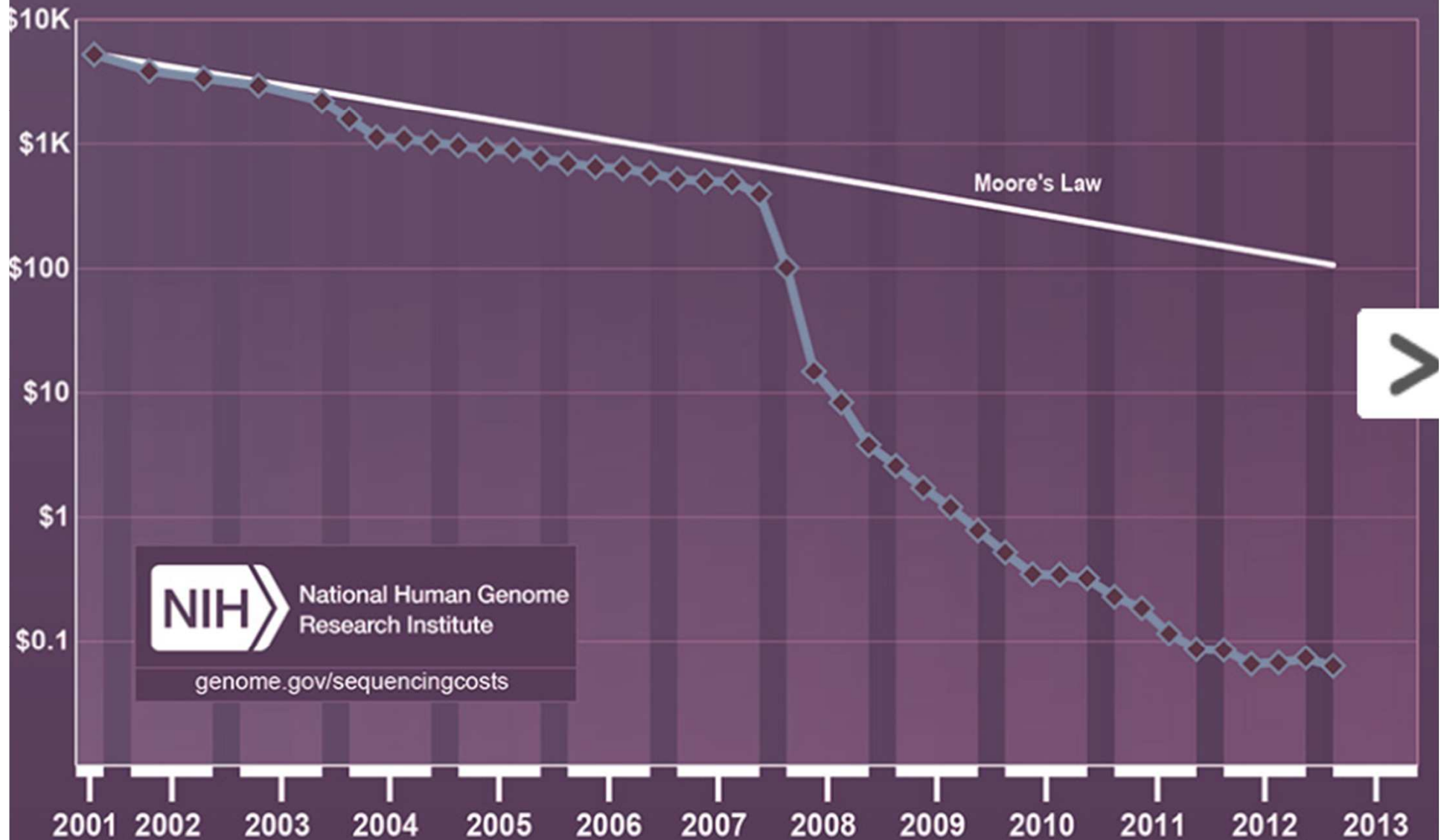


2015

Next Big Thing



# Cost per megabase



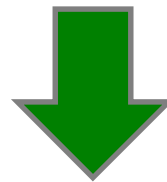
# Sequencing as « Big Science »



Encyclopedia of life is available



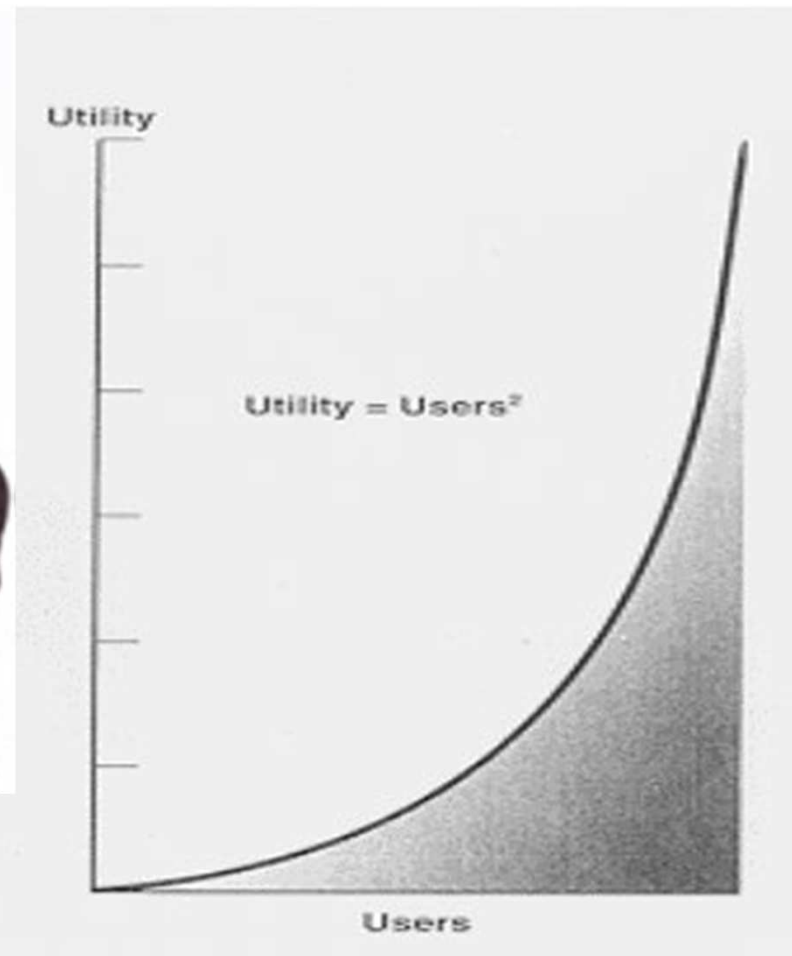
Smart harvest is revolutionizing R&D



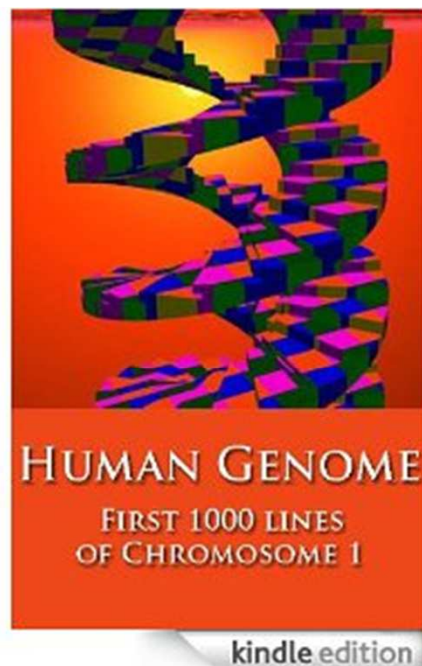
Genome-based R&D ?



# Genome knowledge & phone





[illegible]

## Human Genome: First 1000 lines of Chromosome 1 [Kindle Edition]

Eliana Brown (Author)

Be the first to review this item |  Like (0)

Kindle Price: **\$33.34** includes VAT & free international wireless delivery via **Amazon Whispernet**

- Text-to-Speech: Enabled 
- Don't have a Kindle? [Get your Kindle here.](#)

TGGAGGAGGAGACG  
 CAGAGGAGAGGG  
 CAGTACCGCGCG  
 TGTGTGGGTCCG  
 GCGTCAGGATTC  
 CTCAGGTTTGA  
 GCGAGGAGGGG  
 CACGCGGAGCTG  
 TCGCAGCAGGTC  
 CACTTAAAGCTT  
 CACAGCCAGAG  
 AGATTTTAGGG  
 CAGTGTCTCTG  
 AATCGGCAAA  
 CAGAGCGCGTC  
 CCGAGGAGCCG  
 CCCACAGCAAC  
 CGCGCCCGCGC  
 GCGAGAGGAAC  
 GCGGACCAGGA  
 TACTGGGAGCCT  
 TATTAAGTAC  
 TAAGAAGTACA  
 AACGTCTCGCT  
 TCATCGAGCGA  
 CAGCAAGGCCG

[illegible]

osome

ireless



# Bio-IT in 2010



**GEEK OF THE WEEK**



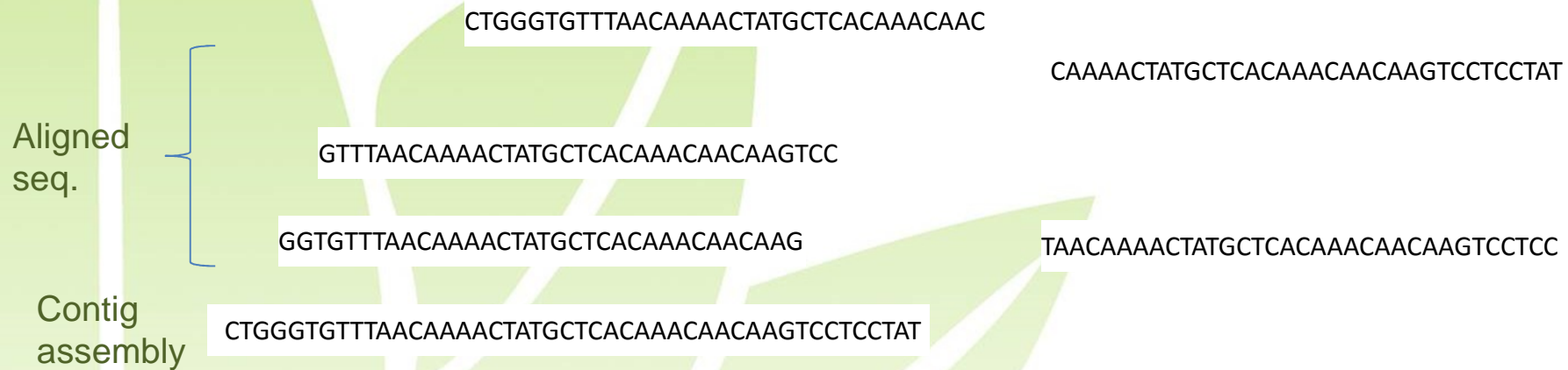
# Bio-IT in 2014



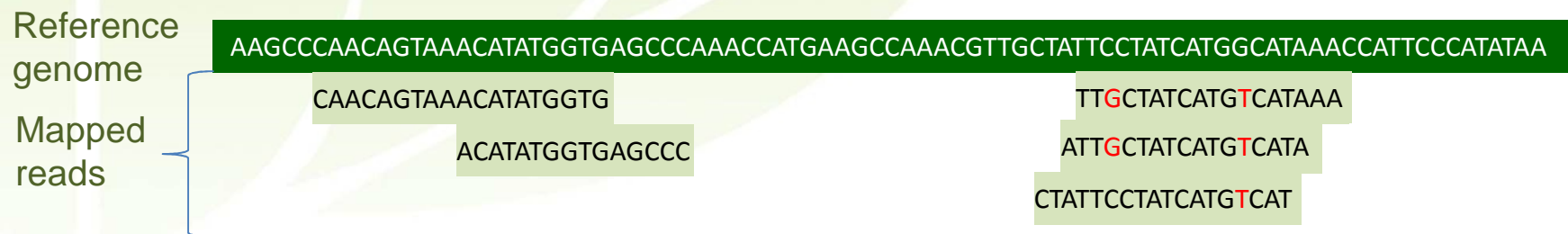


# BioIT: how does it works ?

## de novo assembly :



## Mapping on reference genome :



# Balance between wet lab and computer

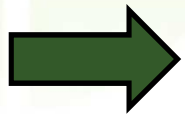




# Examples: medical practice

## Prenatal detection of Trisomy 21 :

- ✓ Clinical test in Belgium from 2014
- ✓ Easy, quick and cheap pre-screening test from maternal blood
- ✓ More than trisomy



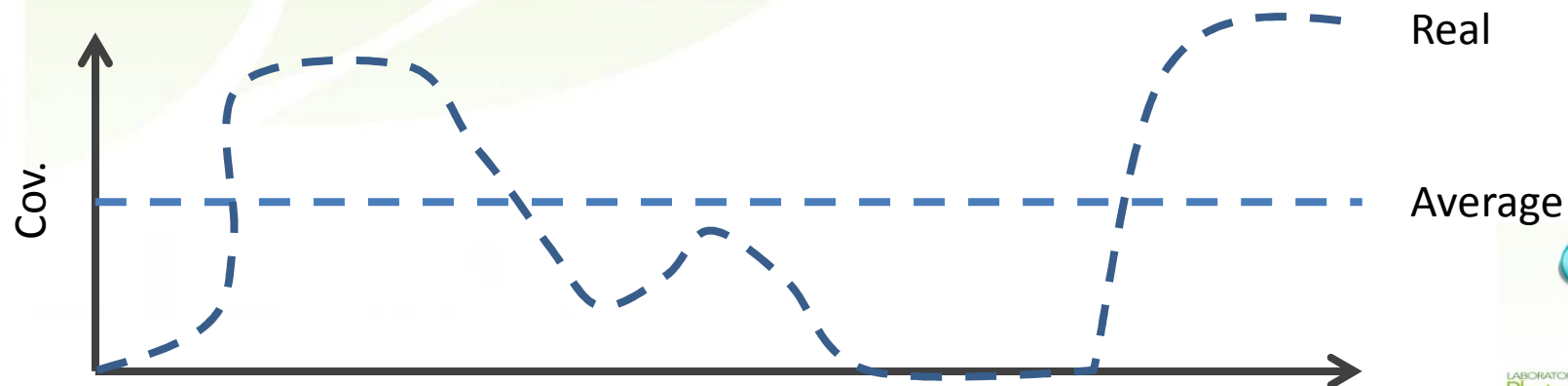
ethical & societal implications ....

.... which will be by-passed

# Examples: medical practice

## Oncology:

- ✓ Millions of tumor genomes scheduled: map of mutation
- ✓ Boom of gene panels for personalized medicine
- ✓ False positives / false negatives / misinterpretations
- ✓ Genome coverage: behind the 50X average





# Examples: medical practice

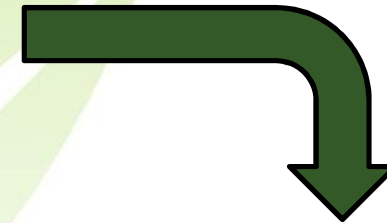
## Oncology:

- ✓ Millions of tumor genomes scheduled: map of mutation
- ✓ Boom of gene panels for personalized medicine
- ✓ False positives / false negatives / misinterpretations
  - ✓ Genome coverage: behind the 50X average
  - ✓ Probability - heterozygosity - low coverage

# Examples: medical practice

## Heterozygosity – low coverage - probability

| Real | Seq 1 | Seq 2 | Seq 3 |
|------|-------|-------|-------|
| A/T  | A     | AA    | AAA   |
|      |       |       | AAT   |
|      |       | AT    | ATA   |
|      |       |       | ATT   |
|      | T     | TA    | TAA   |
|      |       |       | TAT   |
|      |       | TT    | TTA   |
|      |       |       | TTT   |



25% probability of missing heterozygosity !!

# Examples: medical practice

## Oncology:

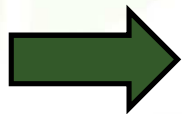
- ✓ Millions of tumor genomes scheduled: map of mutation
- ✓ Boom of gene panels for personalized medicine
- ✓ False positives / false negatives / misinterpretations
  - ✓ Genome coverage: behind the 50X average
  - ✓ Probability - heterozygosity - low coverage
  - ✓ Sequencing errors
- ✓ Activable mutations for other cancer



# Examples: medical practice

## Oncology:

- ✓ Millions of tumor genomes scheduled: map of mutation
- ✓ Boom of gene panels for personalized medicine
- ✓ False positives / false negatives / misinterpretations
- ✓ No yet a medical utility but opens the way to « charlatanism » and « cobaye patients »

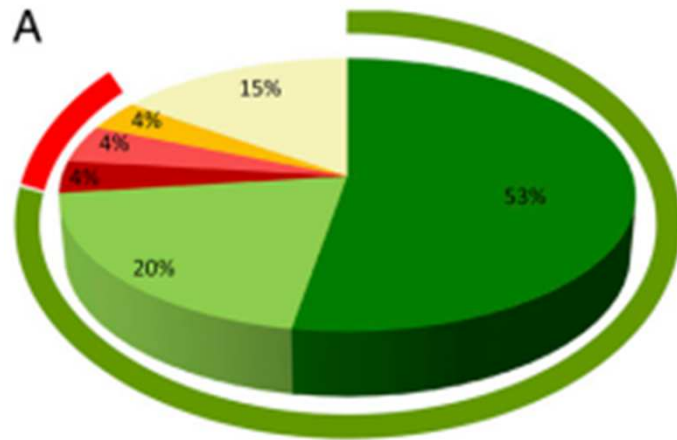


Still a long way for tumors but interest for early detection based on blood signature

# Examples: microbiome

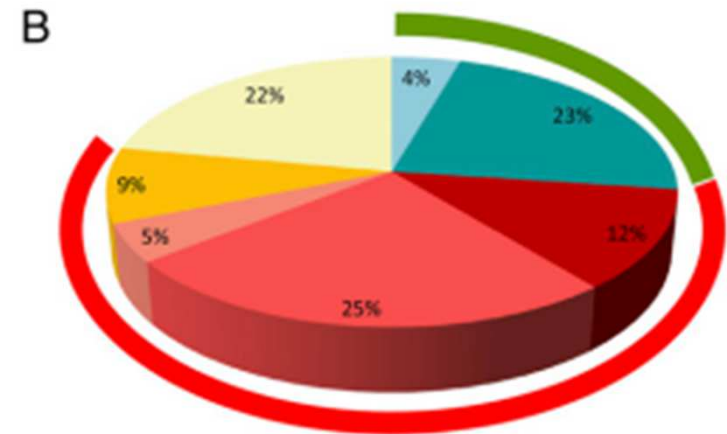
## Stone age microbiome:

- ✓ Healthy children: urban Italy vs. rural Burkina Faso
- ✓ Analysis of 16S rDNA gene
- ✓ 15,000 sequences per sample



BF

|                    |                 |
|--------------------|-----------------|
| ■ Prevotella       | ] Bacteroidetes |
| ■ Xylanibacter     |                 |
| ■ Acetitomaculum   | ] Firmicutes    |
| ■ Faecalibacterium |                 |
| ■ Subdoligranulum  |                 |
| ■ Others           |                 |



EU

|                    |                 |
|--------------------|-----------------|
| ■ Alistipes        | ] Bacteroidetes |
| ■ Bacteroides      |                 |
| ■ Acetitomaculum   | ] Firmicutes    |
| ■ Faecalibacterium |                 |
| ■ Roseburia        |                 |
| ■ Subdoligranulum  |                 |
| ■ Others           |                 |

- ✓ Very interesting but .... frustrating as only Genders...

# ***Homo sapiens***





# Examples: microbiome

## Future:

- ✓ Simple tests by 16s sequencing :
- ✓ Deep analysis of the genes and genomes of the microflora
- ✓ Same process as genome sequencing but incomplete genomes (prevalence, size and complexity)
- ✓ Reconstruction of metabolic pathway, understanding of the functionality

# Examples: microbiome

## Future: functional analysis

- ✓ Aphids symbionts: *Buchnera* sp. and *Wolbachia* sp.
- ✓ Sampling and sequencing (30 M sequences per sample)
- ✓ Genome reconstruction and comparison with other strains or close species

# Examples: microbiome

## Ecophysiology:

- ✓ Project from Laura Llorens and Dolors Verdaguer
- ✓ Climate change and natural ecosystems
- ✓ Plant physiology, botanic composition, soil analysis

+ microbial population (with Lluís Banyeras)





# Examples: e-DNA

- Environmental DNA
- Application of genetic to study ecosystem from a decade but...



& now :



- Scientific knowledge :



- 78.620 arthropods
- 11.407 insects
- 3.883 aves
- 1.850 mammals



# Examples: e-DNA

## ➤ Biodiversity:

✓ Ecosystems

✓ Species



✓ Genes and genomes

} NGS opportunities !!!

# Examples: e-DNA

## ➤ Answered questions:

- ✓ Monitor invasion of alien species through water sampling (bullfrog in France: better sensitivity)
- ✓ Study the species diversity in water (filtering, 50 ml ...) for nearly any branch of the tree of life
- ✓ What is the plant taxonomic diversity of any environment?
  - ✓ Soil (from tundra to rainforest)
  - ✓ Pollen sampling (*Graminaceae*)



# Examples: e-DNA

## ➤ Answered questions:

- ✓ Determine the diet of any animal from fecal samples (lemur, herbivore,...)
- ✓ History of ecosystem by scanning lake sediments
- ✓ Study animal diversity through coprophagous insects
- ✓ Confirming presence of endangered species by coprophagous insects

# Examples: e-DNA



VS.



New promising research area with complementarities which cannot be ignored

# Key concepts

- Garbage in / garbage out (design & quality)
- Minimal a priori... but hypotheses
- Complexity reduction & prioritization
- Iterative, trials & errors
- « Loteria de Navidad » for heterozygous
- No clearcut: Black & White -> grey
- Confirmation of results



# Conclusion

- Technological revolution with decreasing costs
- Open new research areas in traditional and new disciplines
- Many research needed to exploit full potential of the technologies