Exploration de la biodiversité génétique des sols : programme Terragenome

Pascal SIMONET

« Environmental Microbial Genomics » group
Laboratoire Ampère, UMR CNRS 5005
Ecole Centrale de Lyon
Ecully (69)

Pascal.simonet@ec-lyon.fr

Microorganisms: « quantitative level »

- Microscopical Eukaryotes
- **Prokaryotes**

Whitman *et al.* 1998 PNAS, 95, 6578-6583. *Prokaryotes: The unseen majority.*

4-6 x 10^{30} bacterial cells

1,2 x	**10^{29}**	**Ocean**
2,6 x	**10^{29}**	**Soil**
3,5 x	**10^{30}**	**Deep soil**
0,25 - 2,5 x	**10^{30}**	**Ocean deep soil**

Microorganisms: « qualitative level »

1 g of soil

Number of bacterial cells : 1 000 000 000

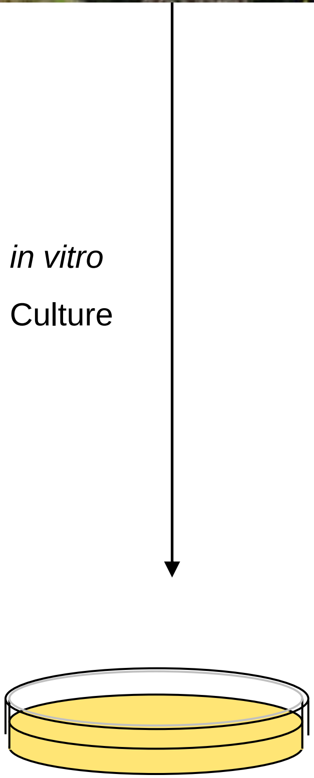
Number of bacterial species:

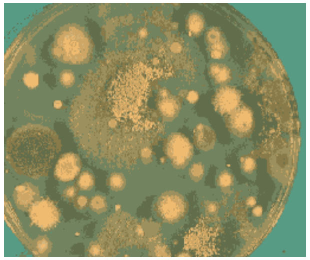
10 000 Torsvik 2002 (Science)

Roesch 2007 (ISME J.)

>10 000 000 Gans 2005 (Science)

Soil?





in vitro Isolation of cultivable bacteria

Percentage of cultivable bacteria?

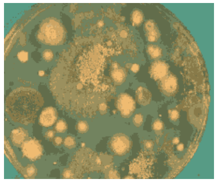
From 0,1 to 1%

(meta-) genomics approach

in vitro

Culture

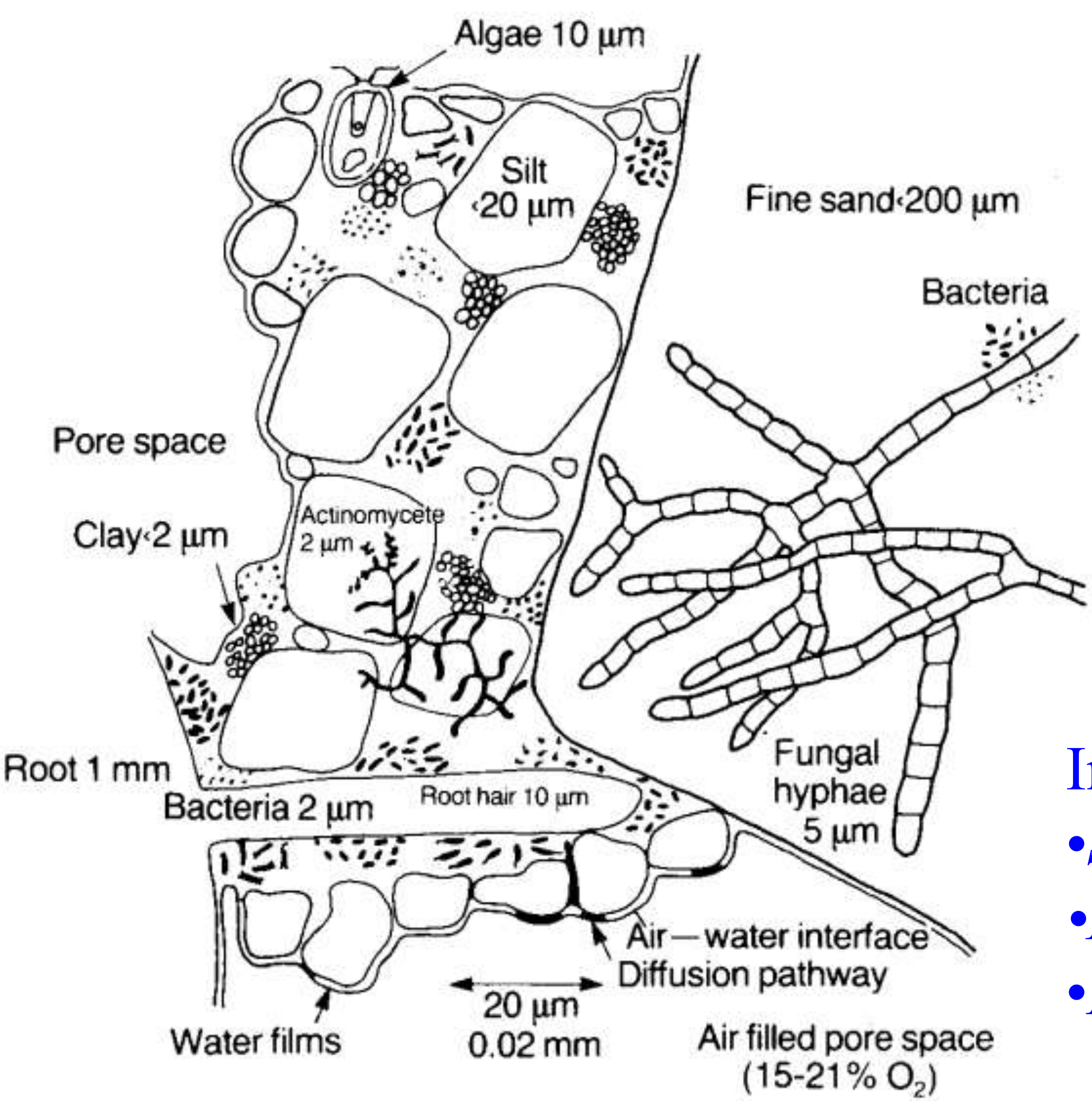
«Environmental DNA»



***in vitro* culture approach**

Cultivable bacteria? From 0,1 to 1%

Soil, a Composite of Communities



Its very diverse

Complex gradients

Unmixed

- microaggregates
- rhizosphere
- fungalsphere
- fauna
- pore surfaces
- OM coatings

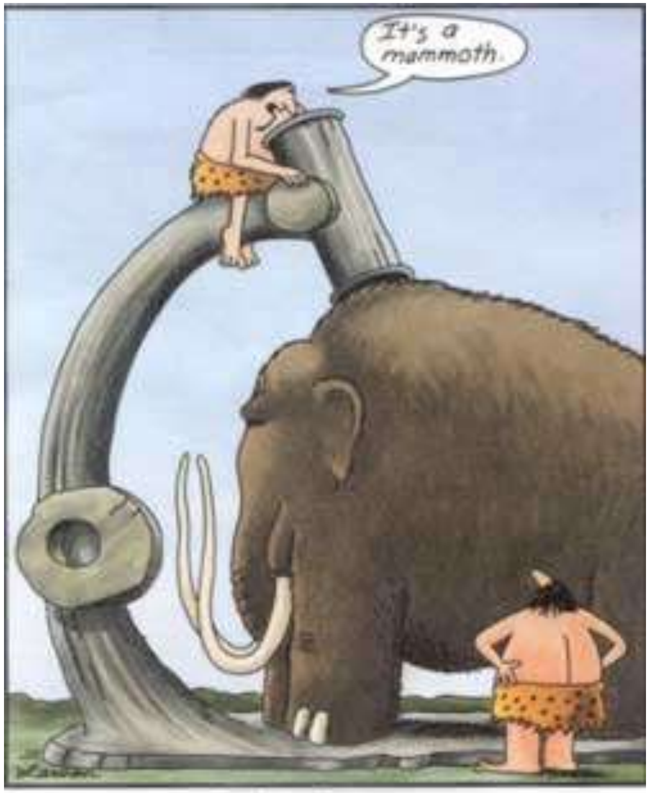
Implications:

•Spatial isolation

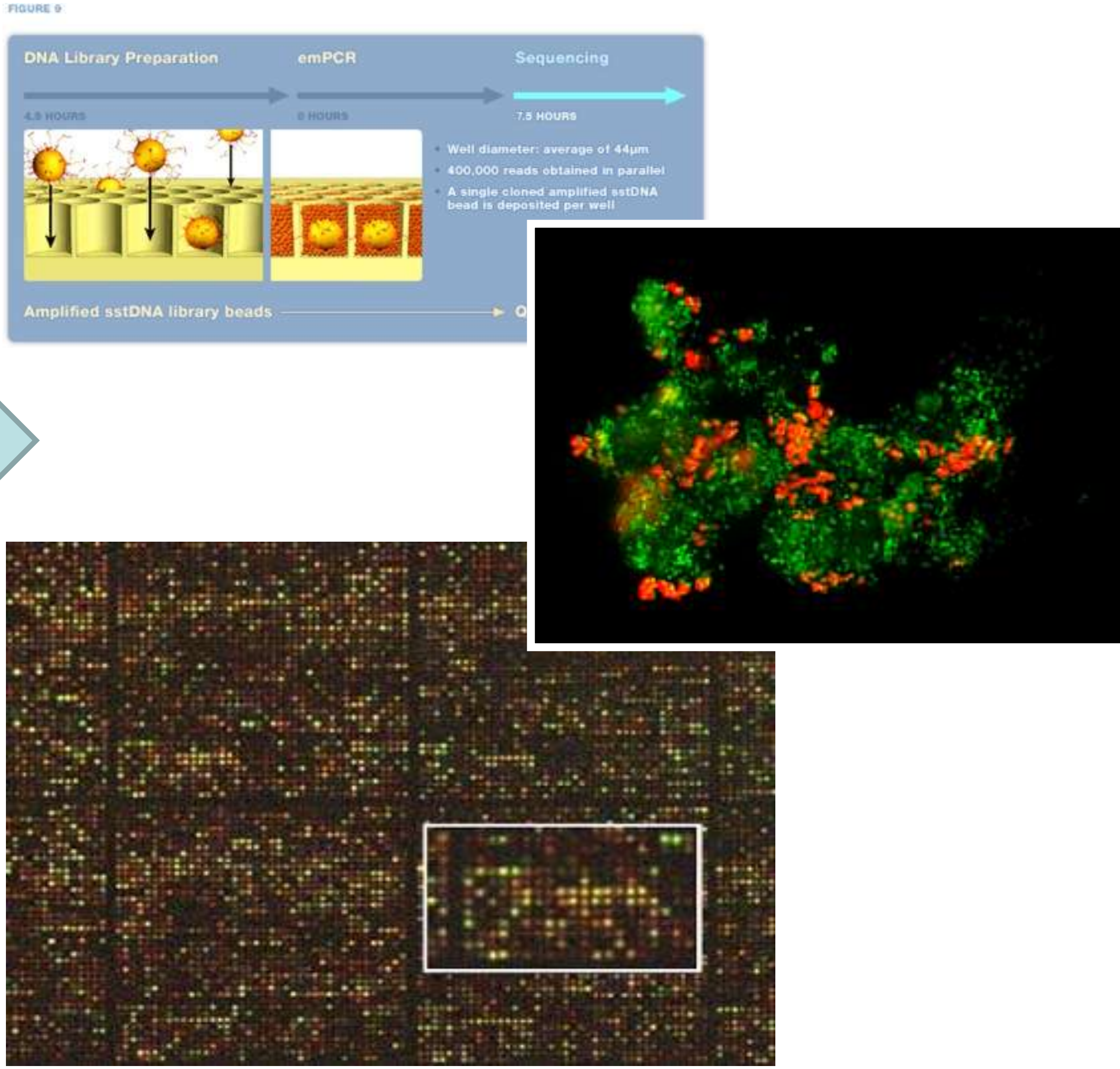
•Minimizes competition

•How to sample since multiple communities?

How 30 years has changed our technology to perceive life on Earth



Early microscopes



Courtesy of George Kowalchuk

De nombreuses raisons de s'intéresser aux communautés bactériennes

- impact sur les équilibres biogéochimiques
    - quels sont les acteurs ?
    - nouvelles étapes des cycles biologiques des éléments

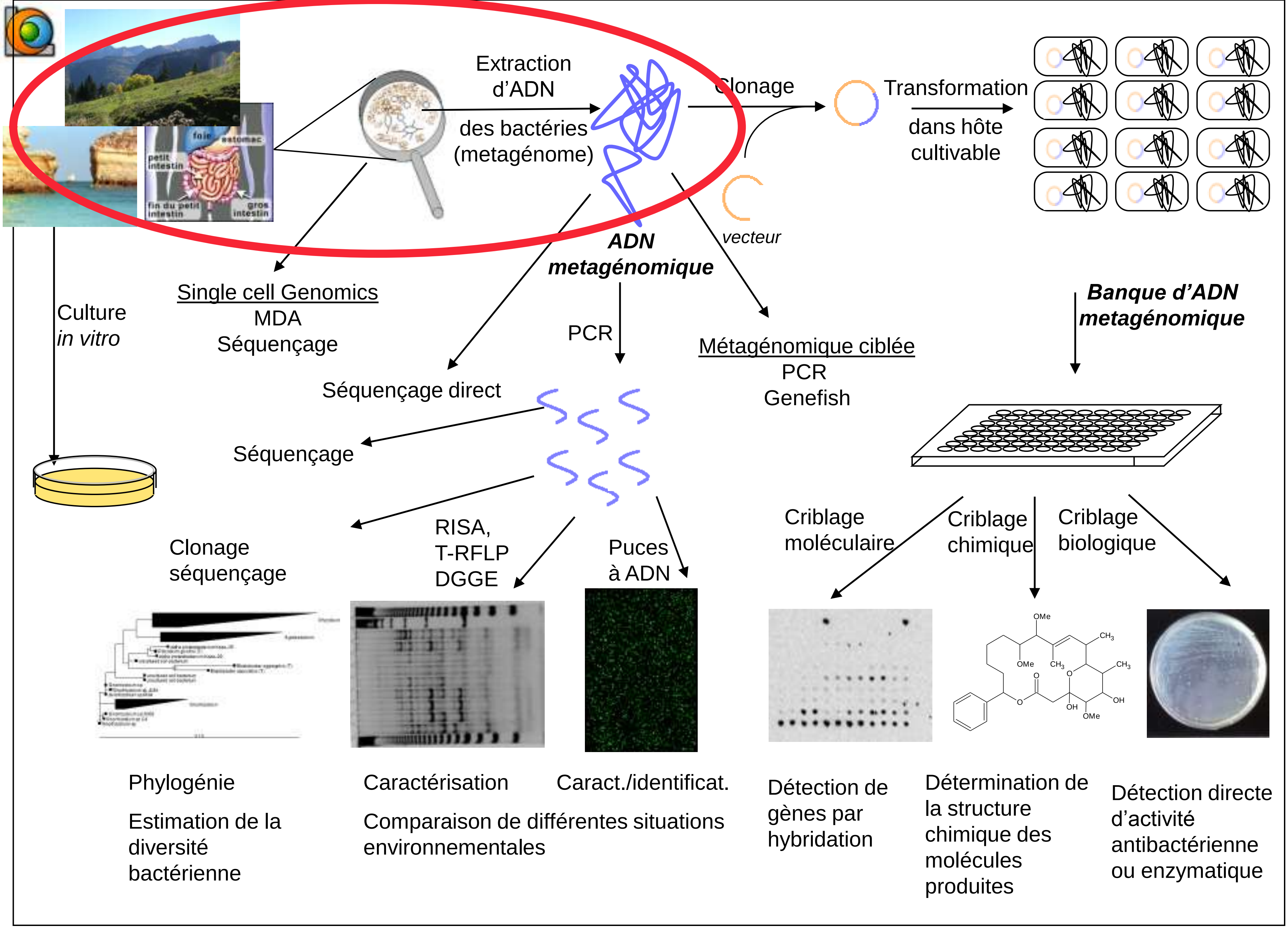
- impact sur la santé (flores microbiennes humaines)

- modèles d'écosystèmes (structure des communautés bactériennes)

- utilisation de la biodiversité à des fins d'applications
    - substances thérapeutiques
    - substances d'intérêt industriel
    - enzymes utiles pour la chimie de synthèse
    - bioremédiation

- nouveaux éclairages sur l'évolution

Courtesy of Denis LePaslier



Extraction d'ADN des bactéries (metagénome)

ADN metagénomique

Clonage

Transformation dans hôte cultivable

Banque d'ADN metagénomique

Culture *in vitro*

Single cell Genomics
MDA
Séquençage

Séquençage direct

PCR

Métagénomique ciblée
PCR
Genefish

Séquençage

Clonage séquençage

RISA,
T-RFLP
DGGE

Puces à ADN

Criblage moléculaire

Criblage chimique

Criblage biologique

Phylogénie

Estimation de la diversité bactérienne

Caractérisation

Caract./identificat.

Comparaison de différentes situations environnementales

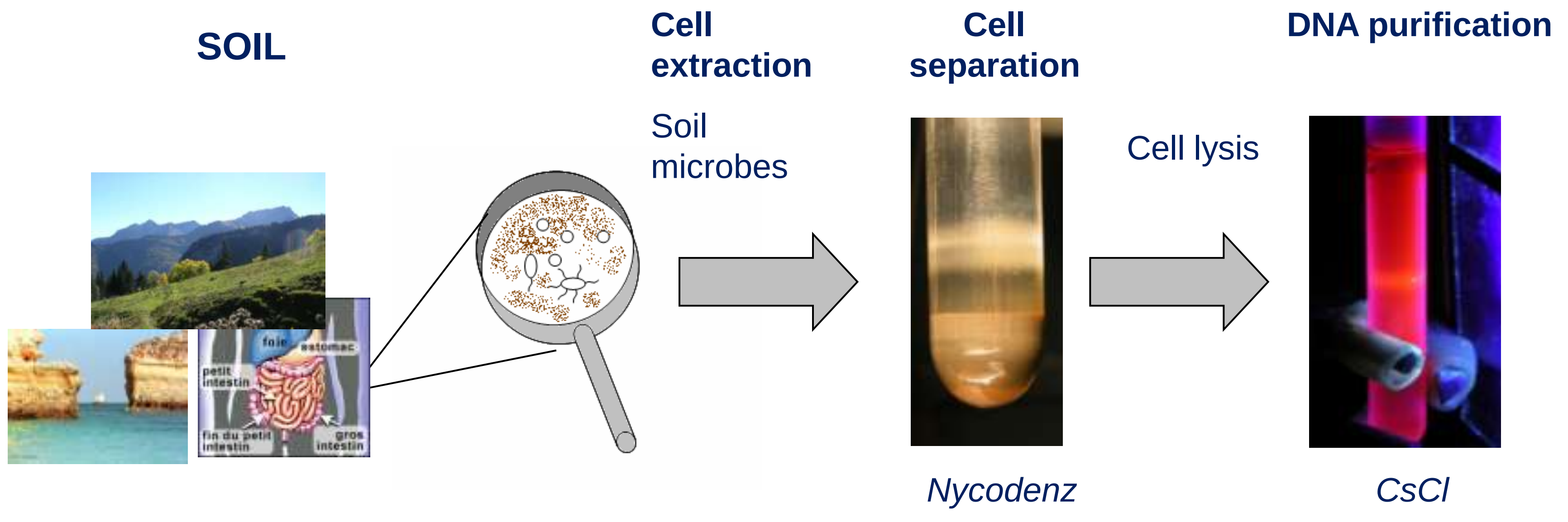
Détection de gènes par hybridation

Détermination de la structure chimique des molécules produites

Détection directe d'activité antibactérienne ou enzymatique

Metagenome DNA extraction

- *In situ* lysis and total DNA extraction
- Cell extraction and lysis



SOIL

Cell extraction

Soil microbes

Cell separation

Nycodenz

Cell lysis

DNA purification

CsCl

Optimization of bacterial DNA recovery

Sampling strategies

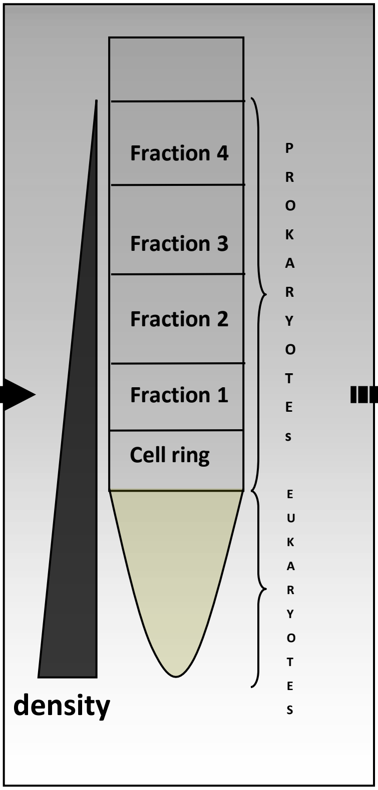
- **Time of the year**
- **Depth**

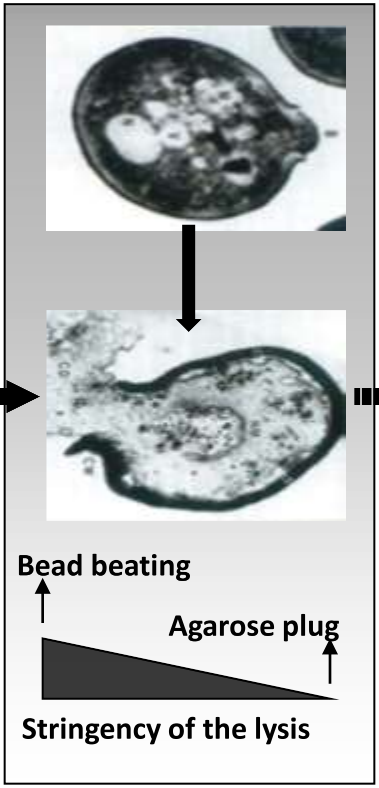
Improvement of **cell recovery (Nycodenz)**

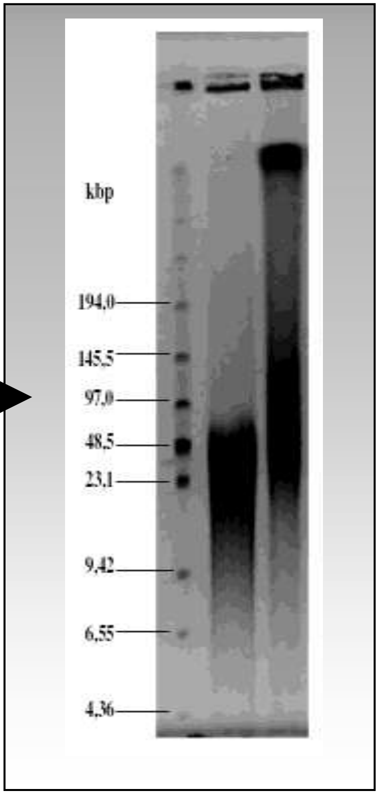
Improvement of DNA recovery (sensitivity to lysis treatments)

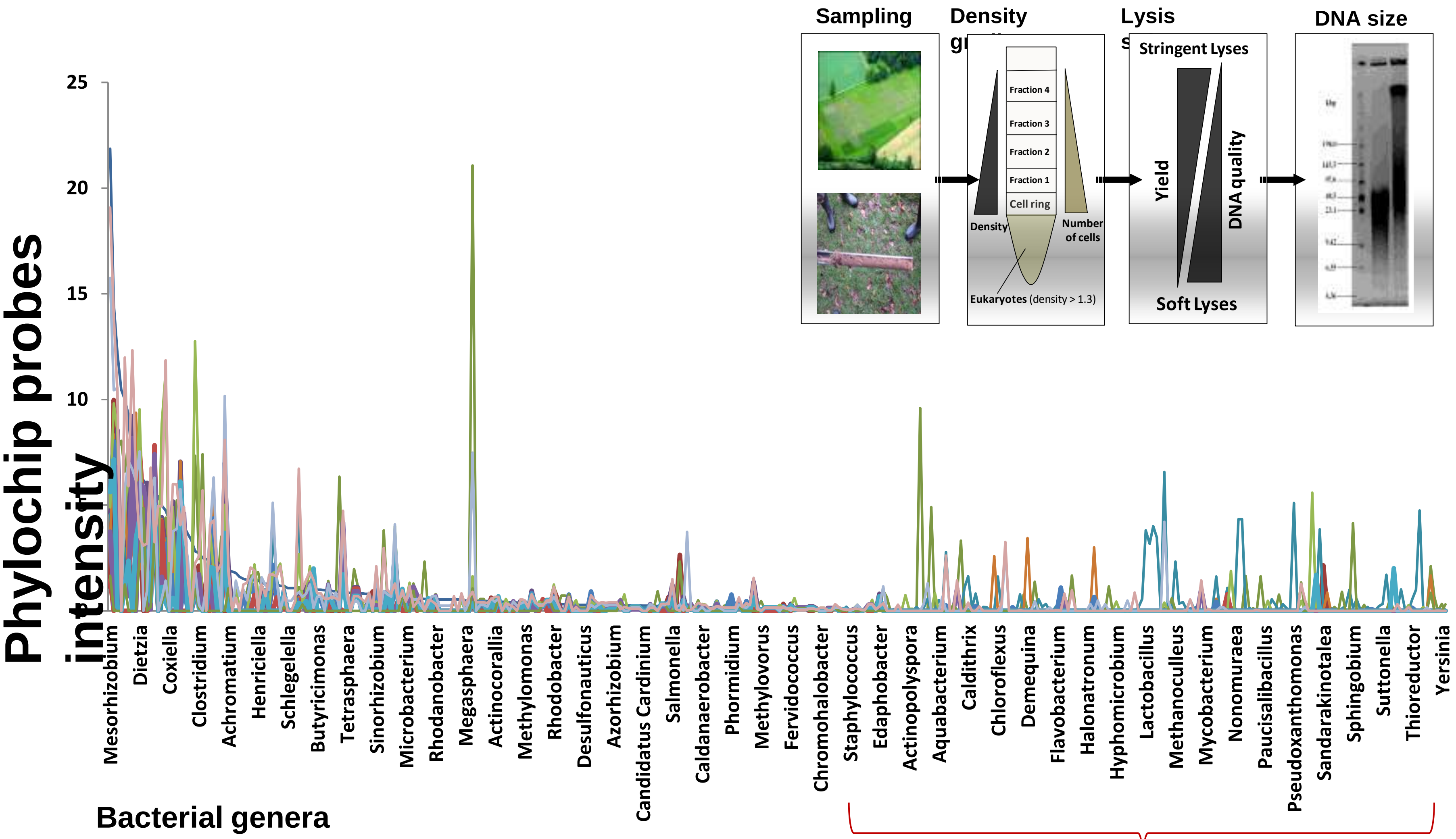
Improvement

of DNA recovery (DNA degradation)





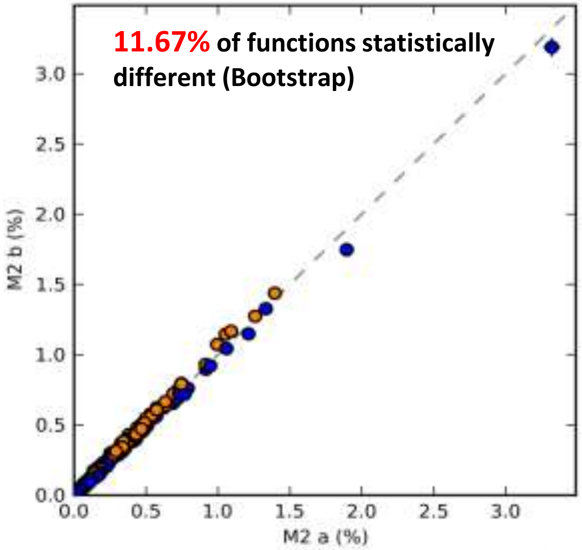




Undetected with one DNA extraction method

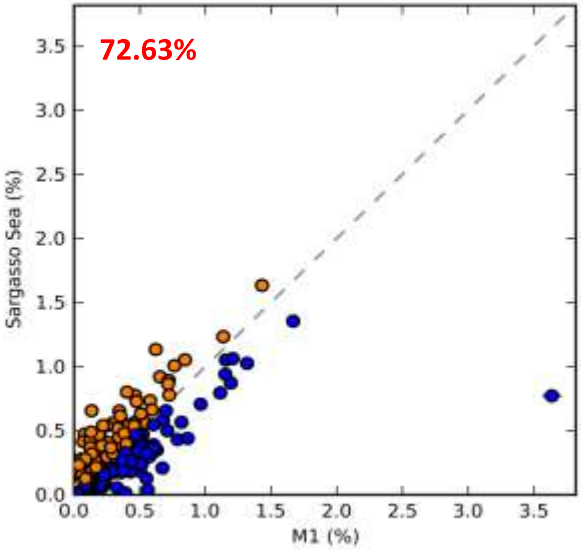
Functional comparison using MG RAST annotation and STAMP statistical analyses

1. technological reproducibility



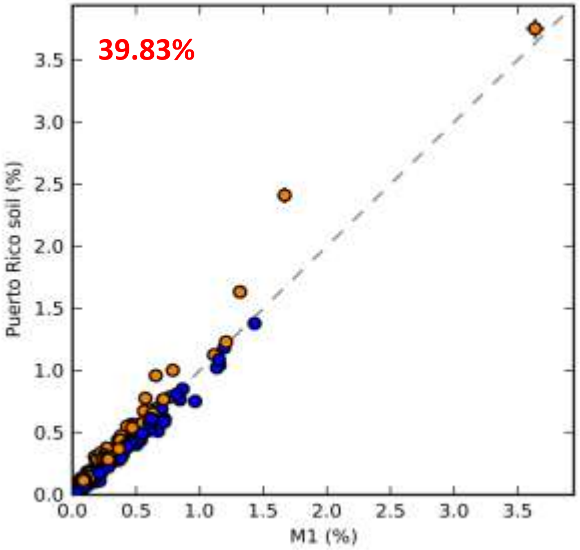
11.67% of functions statistically different (Bootstrap)

2. comparison with an ocean



72.63%

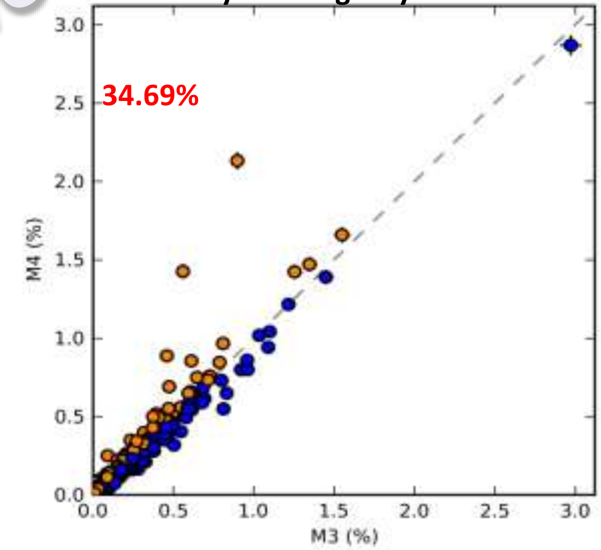
3. comparison with another soil



39.83%

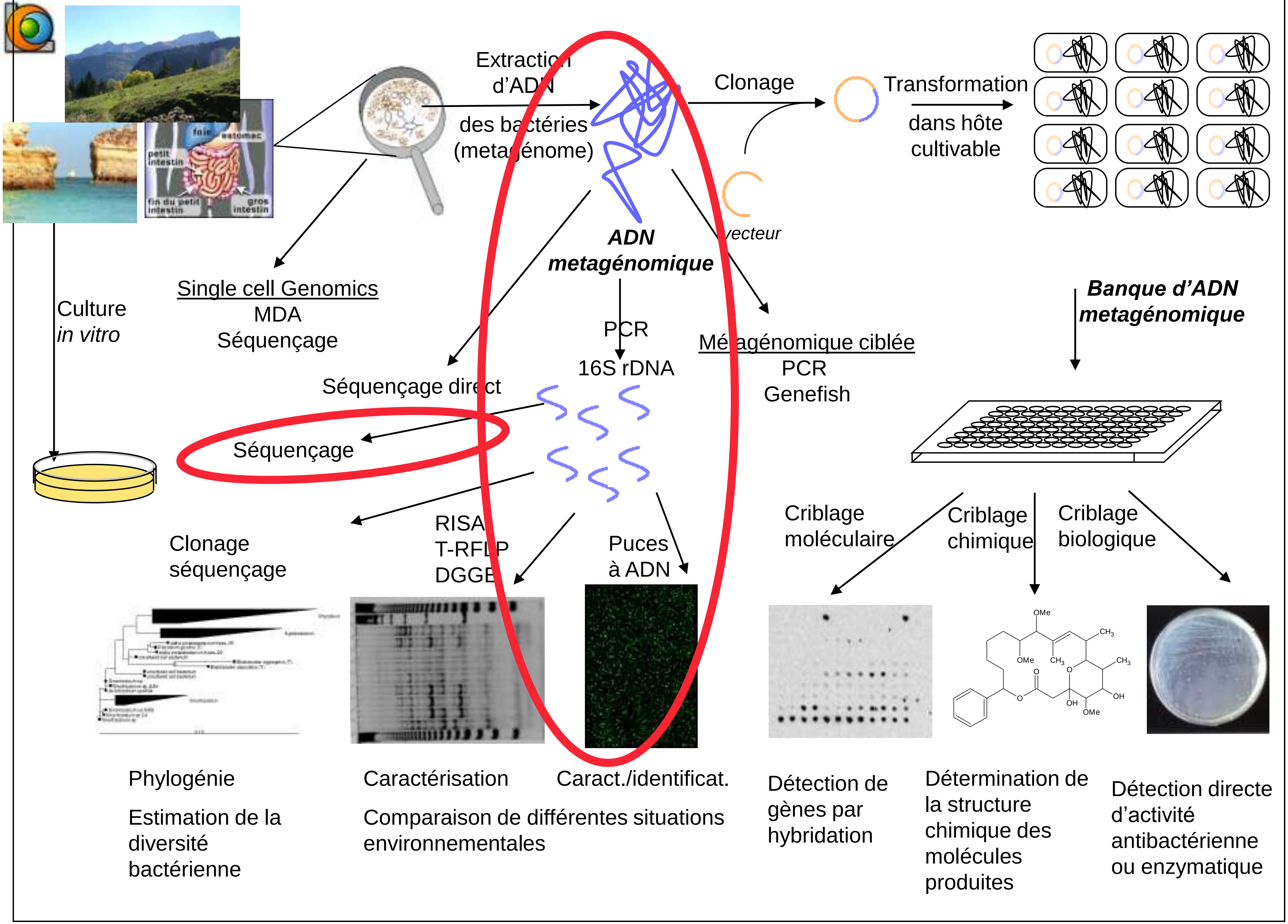
Functional level

4. Cell lysis stringency effect



34.69%

DNA Extraction bias



Extraction d'ADN des bactéries (metagénome)

Clonage

vecteur

Transformation

dans hôte cultivable

ADN metagénomique

Culture *in vitro*

Single cell Genomics
MDA
Séquençage

PCR
16S rDNA

Métagénomique ciblée
PCR
Genefish

Banque d'ADN metagénomique

Séquençage direct

Séquençage

Clonage séquençage

RISA
T-RFLP
DGGE

Puces à ADN

Criblage moléculaire

Criblage chimique

Criblage biologique

Phylogénie

Estimation de la diversité bactérienne

Caractérisation

Caract./identificat.

Comparaison de différentes situations environnementales

Détection de gènes par hybridation

Détermination de la structure chimique des molécules produites

Détection directe d'activité antibactérienne ou enzymatique

Progression of the Sequencing Era

First generation (1995- 2007)

Instrument	Length	Nos	Output	Time/run
Sanger	700bp	96	60Kb	

Second generation (2006 - now)

Roche/titanium	400bp	I mil	400Mb	1-2da
Illumina GAII	75bp x 2	30mil	2.2Gb	8da
Illumina HiSeq	100bp	1-2bil	200 Gb	8da
ABI SOLiD	75bp	3-4 bil	300Gb	14 da

Third generation (single molecule read) (late 2010)

PacBio	>1000bp	0.1-0.3bil	1Gb	15 min

Metasoil Project

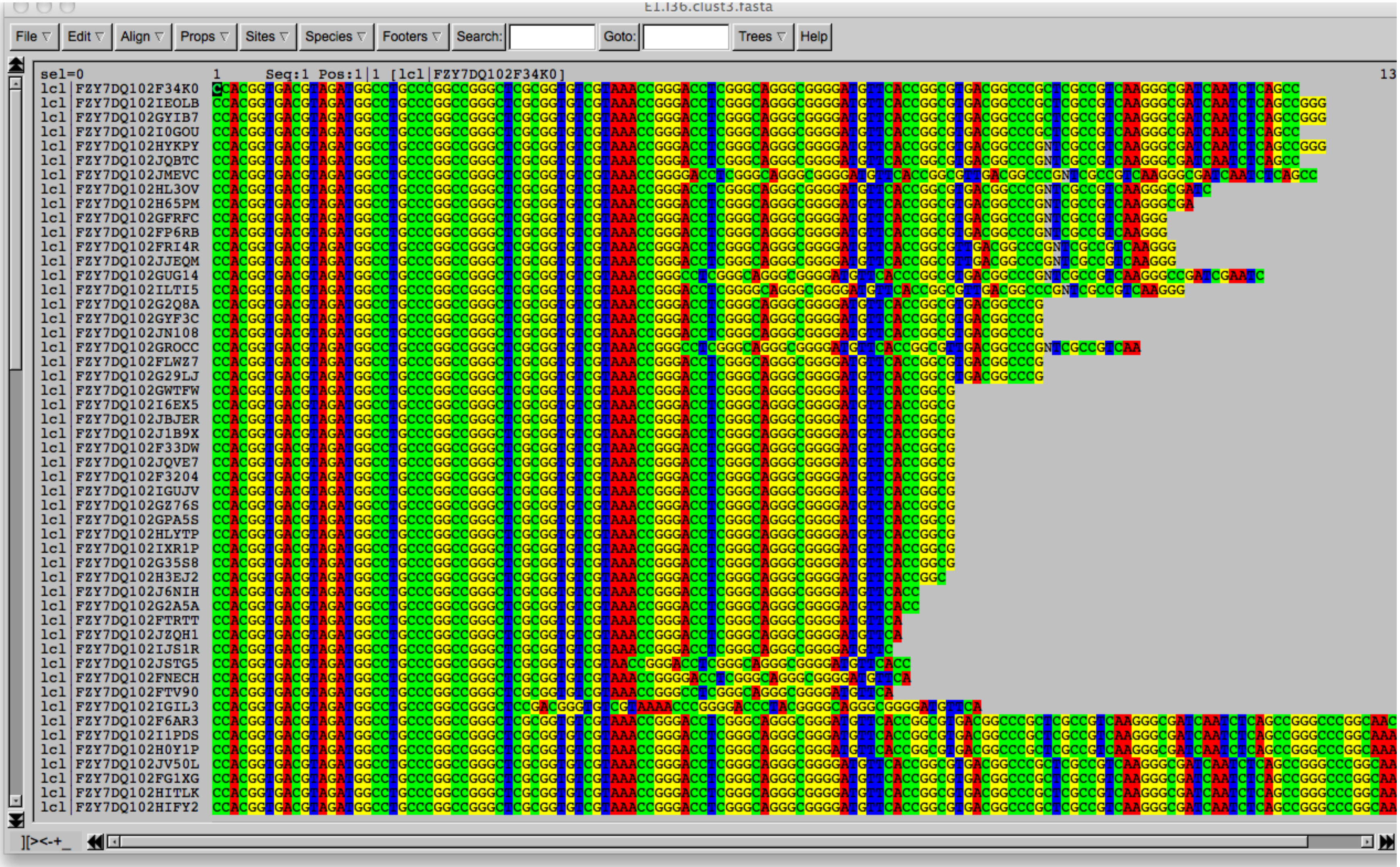
The initial support for Terragenome (complete sequencing of a reference soil metagenome) :

Objective:

- Optimization of bacterial DNA recovery.
- Metagenomic DNA library construction
- Pyrosequencing of directly extracted DNA

Park Grass, Rothamsted: an internationally recognized agroecology field experiment for 150 years

Too many copies



```
sel=0                1     Seq:1 Pos:1|1 [lcl|FZY7DQ102F34K0]
lcl|FZY7DQ102F34K0
lcl|FZY7DQ102IEOLB
lcl|FZY7DQ102GYIB7
lcl|FZY7DQ102I0GOU
lcl|FZY7DQ102HYKPY
lcl|FZY7DQ102JQBTC
lcl|FZY7DQ102JMEVC
lcl|FZY7DQ102HL3OV
lcl|FZY7DQ102H65PM
lcl|FZY7DQ102GFRFC
lcl|FZY7DQ102FP6RB
lcl|FZY7DQ102FRI4R
lcl|FZY7DQ102JJEQM
lcl|FZY7DQ102GUG14
lcl|FZY7DQ102ILTI5
lcl|FZY7DQ102G2Q8A
lcl|FZY7DQ102GYF3C
lcl|FZY7DQ102JN108
lcl|FZY7DQ102GROCC
lcl|FZY7DQ102FLWZ7
lcl|FZY7DQ102G29LJ
lcl|FZY7DQ102GWTFW
lcl|FZY7DQ102I6EX5
lcl|FZY7DQ102JBJER
lcl|FZY7DQ102J1B9X
lcl|FZY7DQ102F33DW
lcl|FZY7DQ102JQVE7
lcl|FZY7DQ102F3204
lcl|FZY7DQ102IGUJV
lcl|FZY7DQ102GZ76S
lcl|FZY7DQ102GPA5S
lcl|FZY7DQ102HLYTP
lcl|FZY7DQ102IXR1P
lcl|FZY7DQ102G35S8
lcl|FZY7DQ102H3EJ2
lcl|FZY7DQ102J6NIH
lcl|FZY7DQ102G2A5A
lcl|FZY7DQ102FTRTT
lcl|FZY7DQ102JZQH1
lcl|FZY7DQ102IJS1R
lcl|FZY7DQ102JSTG5
lcl|FZY7DQ102FNECH
lcl|FZY7DQ102FTV90
lcl|FZY7DQ102IGIL3
lcl|FZY7DQ102F6AR3
lcl|FZY7DQ102I1PDS
lcl|FZY7DQ102H0Y1P
lcl|FZY7DQ102JV50L
lcl|FZY7DQ102FG1XG
lcl|FZY7DQ102HITLK
lcl|FZY7DQ102HIFY2
```

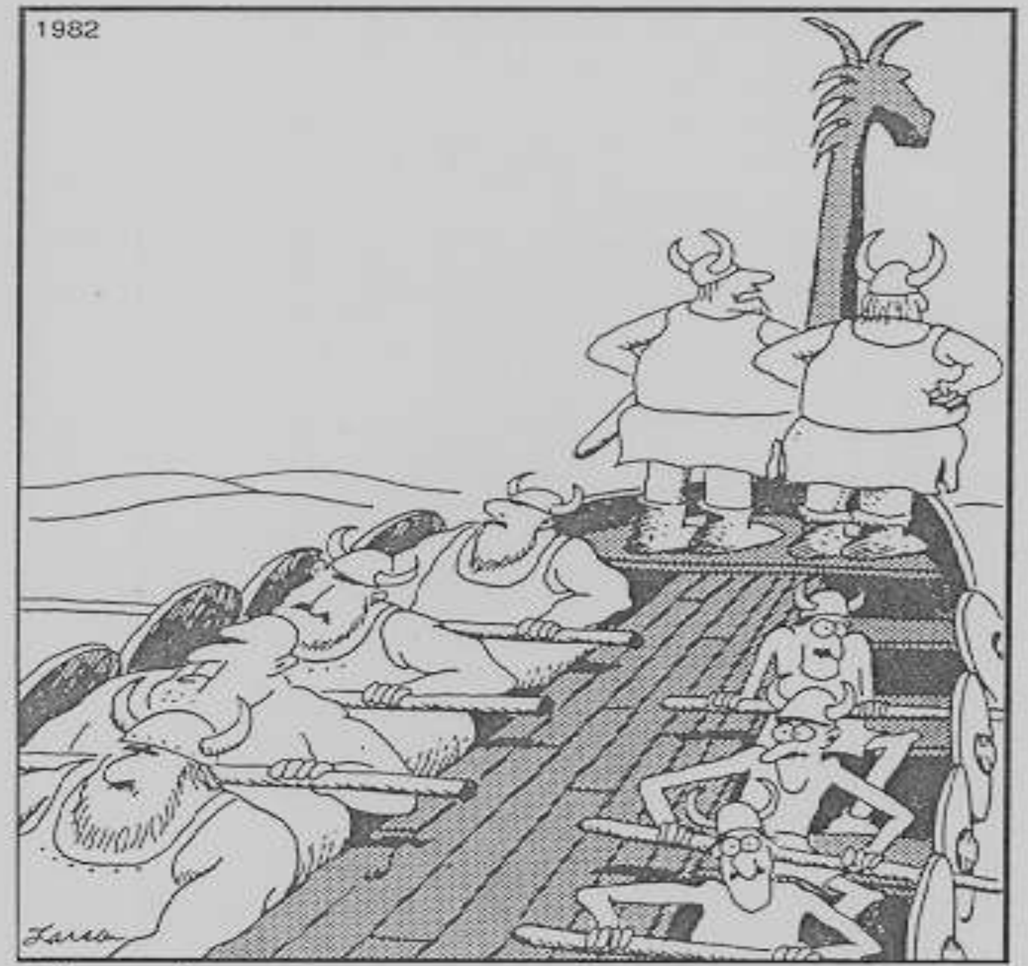
Use cd-hit-454

Niu et al. Artificial and natural duplicates in pyrosequencing reads of metagenomic data. BMC Bioinformatics (2010) vol. 11 pp. 187


TOO MUCH
DATA - NEED
HELP !

We are now in a position to generate huge quantities of sequence (and other) information … but are we traveling efficiently?

Bulk data
generation
capacity

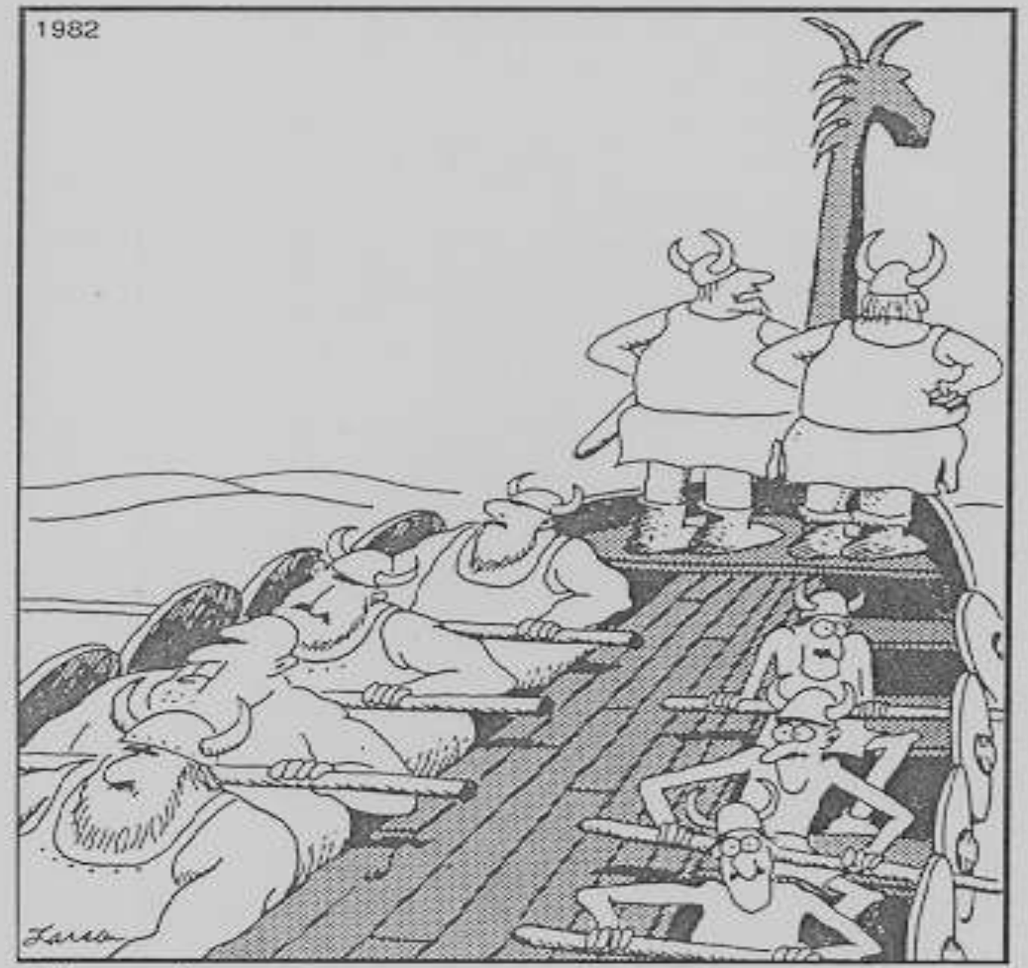


“I've got it, too, Omar. . . a strange feeling like we've just been going in circles.”

Screening &
bioinformatic
strategies

We are now in a position to generate huge quantities of sequence (and other) information … but are we traveling efficiently?

Technology

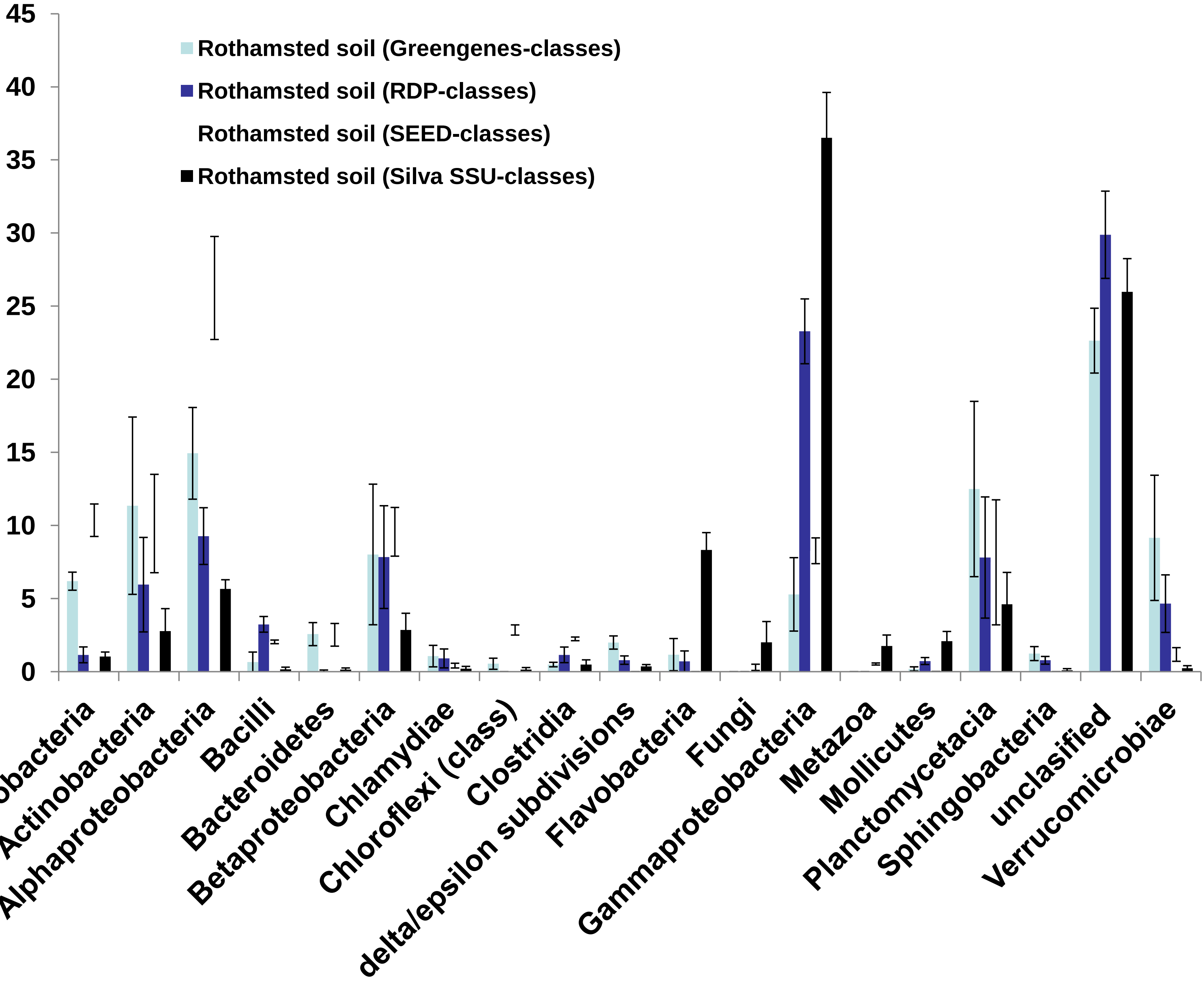


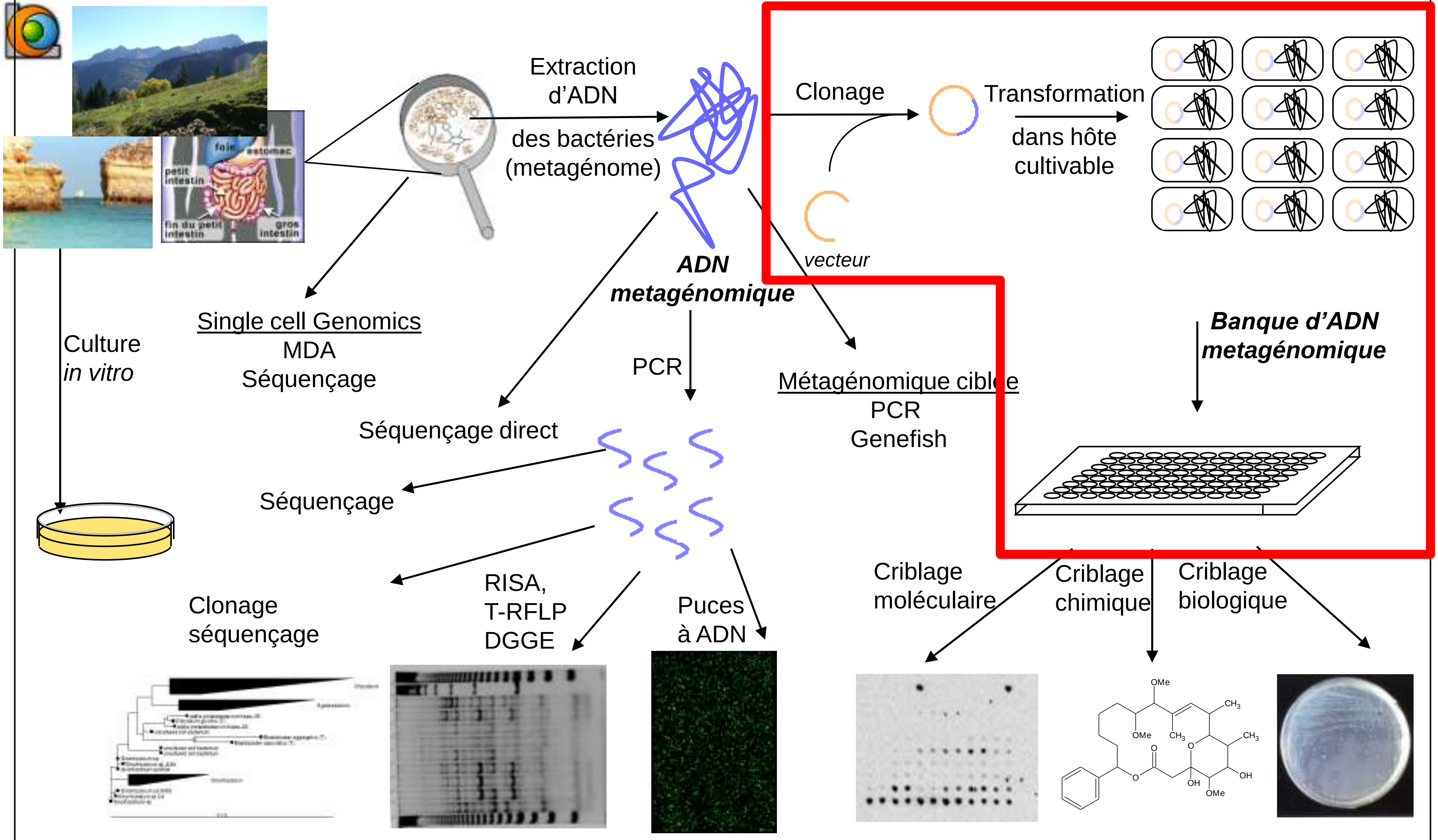
“I've got it, too, Omar. . . a strange feeling like we've just been going in circles.”

Science

Courtesy of Jim Prosser and George Kowalchuk


HAVE SOFTWARE
NEED DATA


45
40
35
30
25
20
15
10
5
0
Rothamsted soil (Greengenes-classes)
Rothamsted soil (RDP-classes)
Rothamsted soil (SEED-classes)
Rothamsted soil (Silva SSU-classes)
Acidobacteria
Actinobacteria
Alphaproteobacteria
Bacilli
Bacteroidetes
Betaproteobacteria
Chlamydiae
Chloroflexi (class)
Clostridia
delta/epsilon subdivisions
Flavobacteria
Fungi
Gammaproteobacteria
Metazoa
Mollicutes
Planctomycetacia
Sphingobacteria
unclasified
Verrucomicrobiae



Extraction d'ADN des bactéries (metagénome)

ADN metagénomique

Clonage

vecteur

Transformation dans hôte cultivable

Banque d'ADN metagénomique

Culture *in vitro*

Single cell Genomics
MDA
Séquençage

Séquençage direct

PCR

Métagénomique ciblée
PCR
Genefish

Séquençage

Clonage séquençage

RISA,
T-RFLP
DGGE

Puces à ADN

Criblage moléculaire

Criblage chimique

Criblage biologique

Phylogénie

Estimation de la diversité bactérienne

Caractérisation

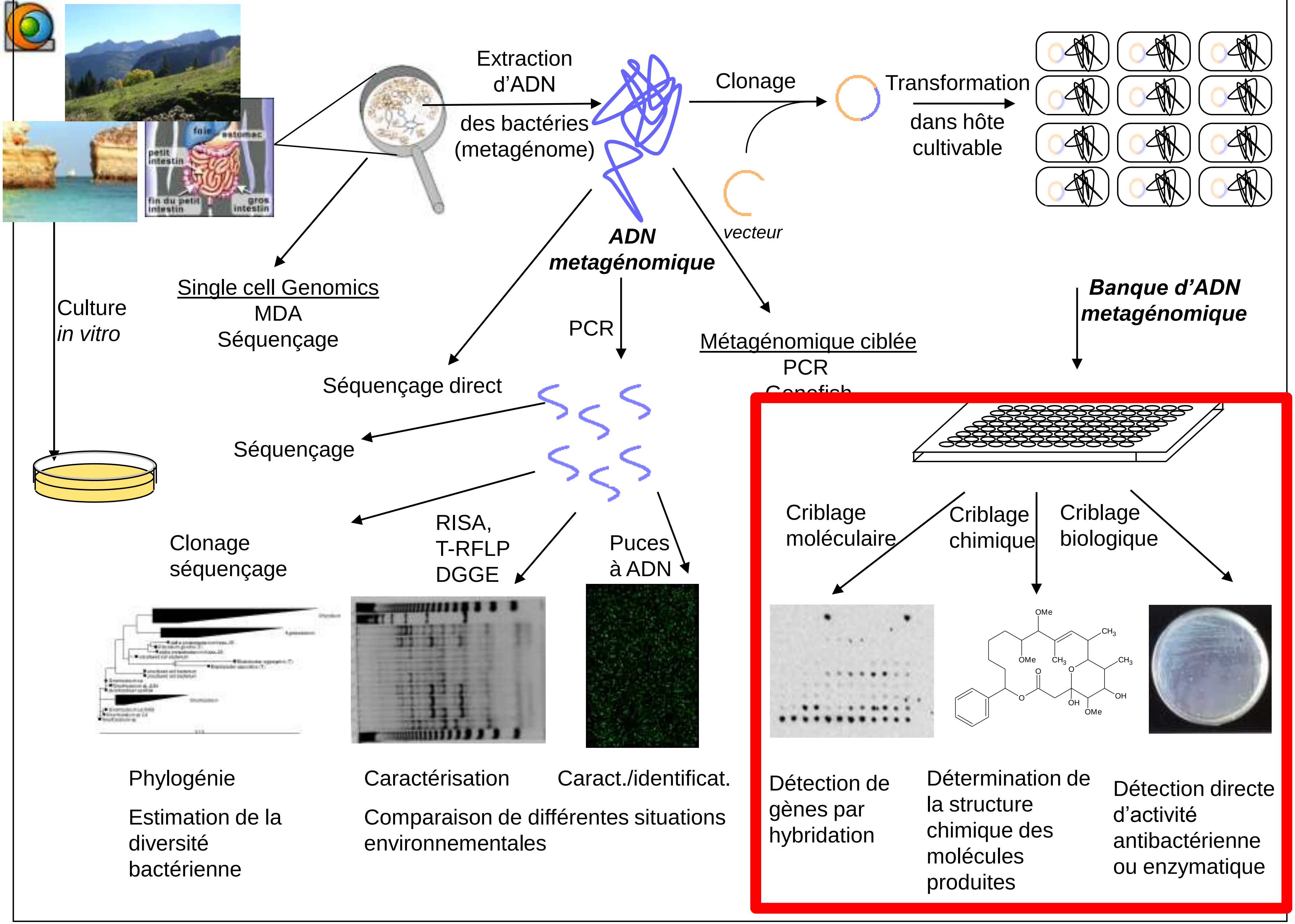
Caract./identificat.

Comparaison de différentes situations environnementales

Détection de gènes par hybridation

Détermination de la structure chimique des molécules produites

Détection directe d'activité antibactérienne ou enzymatique



Extraction d'ADN des bactéries (metagénome)

Clonage

Transformation dans hôte cultivable

vecteur

ADN metagénomique

Culture *in vitro*

Single cell Genomics
MDA
Séquençage

Séquençage direct

PCR

Métagénomique ciblée
PCR
Genefish

Banque d'ADN metagénomique

Séquençage

Clonage séquençage

RISA,
T-RFLP
DGGE

Puces à ADN

Criblage moléculaire

Criblage chimique

Criblage biologique

Phylogénie

Estimation de la diversité bactérienne

Caractérisation

Caract./identificat.

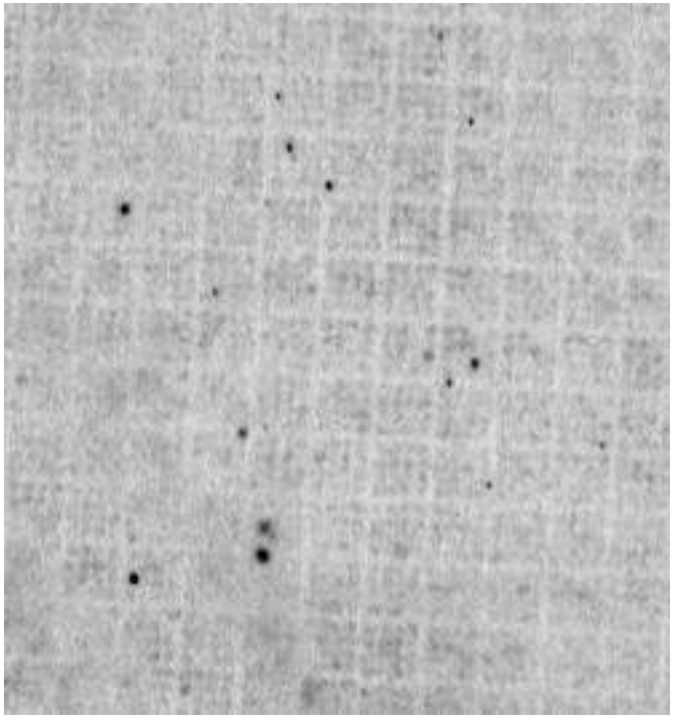
Comparaison de différentes situations environnementales

Détection de gènes par hybridation

Détermination de la structure chimique des molécules produites

Détection directe d'activité antibactérienne ou enzymatique

Molecular screening (60 000 clones)



LibraGen

Molecular screening

• 60 000 clones	139 « KS » +++
• 40 « KS »: sequencing	No redundancy Novelty (similarity < 67%)

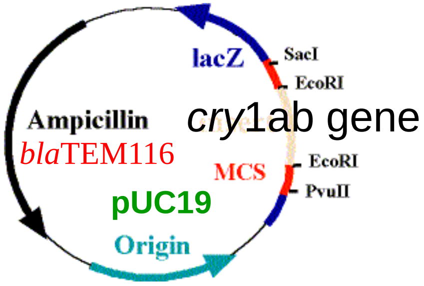
Ginolhac A., Jarrin C., Gillet B., Robe P., Pujic P., Tuphile K., Bertrand H., Vogel T. M., Perriere G., Simonet P., and Nalin R 2004.
Phylogenetic Analysis of Polyketide Synthase I Domains from Soil Metagenomic Libraries Allows Selection of Promising Clones.
Appl. Environ. Microbiol. 2004;70 5522-5527

La métagénomique appliquée à la problématique des plantes OGM !

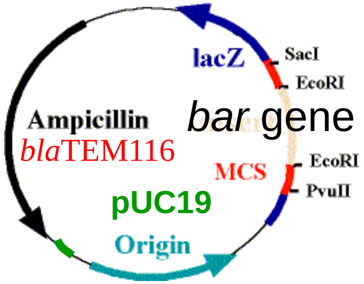
Experimental station: Bazièges, France

Zea mays L.

pCIB4431



pCIB3064



Event Bt176

Field with transgenic Bt 176 corn (*bla*TEM-116) 10-years culture

Ampicillin-resistant bacteria isolation and bacterial metagenomic DNA extraction

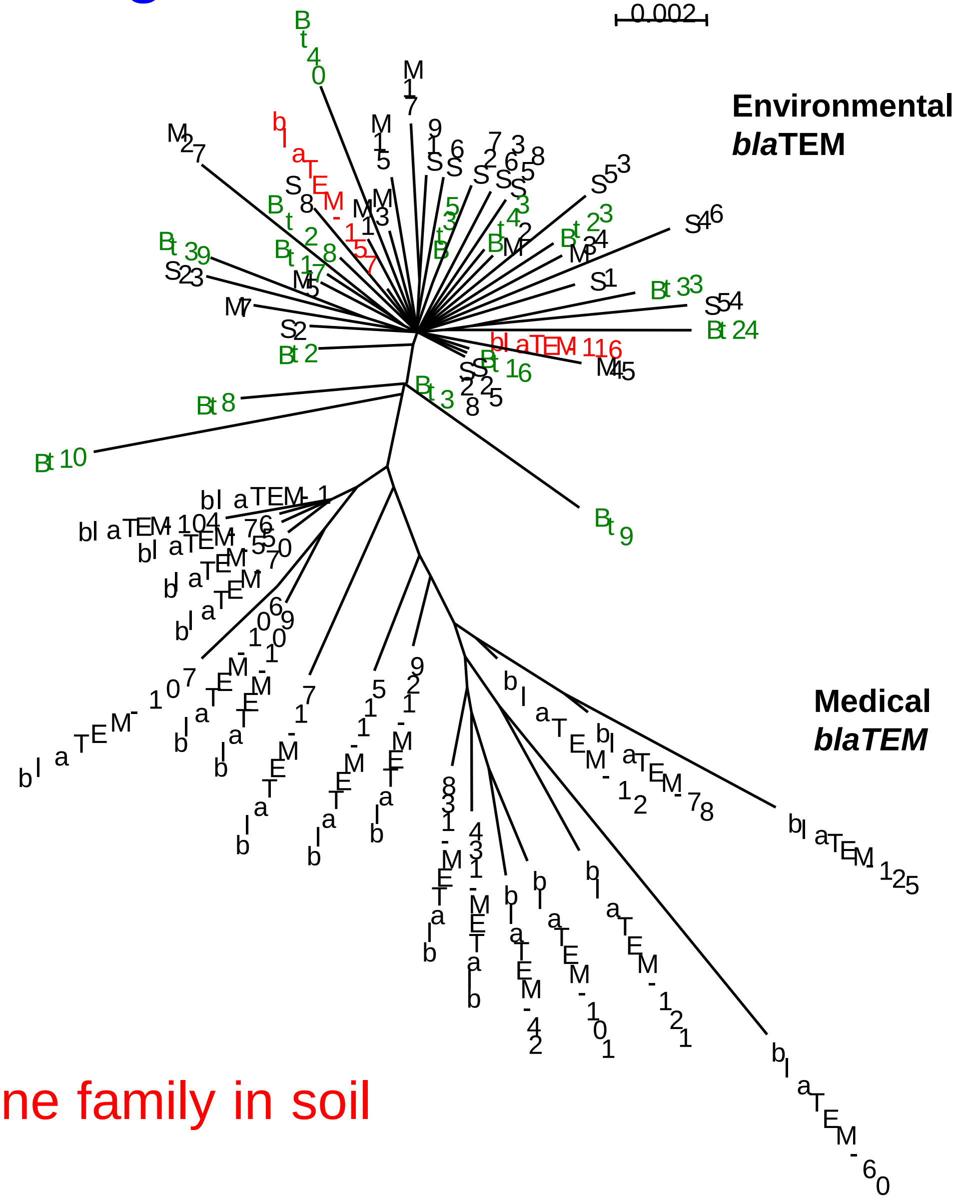
Beta-lactam resistance tests

Diversity of *bla*TEM genes
(PCR + cloning)
Comparisons of bacterial communities
(microarray data)

Controls with a traditional corn field and a prairie in the same location

Diversity of *bla*TEM genes in soils

Metagenomic approach
=
DNA extracted from soil



10 DNA inserts that originated from transgenic and prairie soil DNA exhibited the *bla*TEM116 gene sequence, confirming its natural occurrence in environmental bacteria.

153 different *bla*TEM sequences were found in the PCR-amplified bacterial DNA.

High polymorphism level of this gene family in soil bacterial populations

Transfert de gènes des plantes (OGM) aux bactéries?

- Est-ce que l'acquisition d'ADN étranger par des bactéries est possible?

- Si oui, est-ce qu'il y a une régulation de l'acquisition de l'ADN?
  Par quels mécanismes?

- Est-ce que l'ADN de la plante peut être transféré aux bactéries?

- Est-ce qu'il y a une spécificité pour l'ADN des plantes OGM ?

- Est-ce que l'on peut détecter un transfert de gènes des plantes vers les bactéries?
  Dans quels environnements? A quelle fréquence?

- **Est-ce qu'il y a un impact sur la communauté des bactéries?**

Conclusions

√ Indigenous bacteria possess a wide diversity of beta-lactam resistance genes

↳ Involved in evolutionary processes

√ Soil is a reservoir of genetic determinants providing bacteria with the ability to adapt rapidly to present and future antibiotics!

√Acquisition of *bla* gene from transgenic plants would not modify existing gene diversity in soil nor confer specific advantage

↳ The risk that antibiotic resistances in GMP can pose to clinical strains could be considered as almost null

Aurélie Faugier, Sébastien Cécillon,
Davide Francioli, Tom Delmont,
Emmanuel Prestat, Jean-Michel
Monier, Timothy M Vogel,

Environmental Microbial Genomics

www.GenomEnviron.org