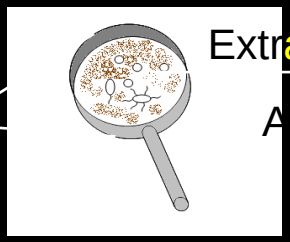
Mining Metagenomes for Novel Enzymes

Timothy M. Vogel

www.GenomEnviron.org

Ampère
CENTRALE LYON
CNRS
CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE
INSA LYON
UCBL

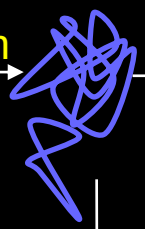
Metagenomic approach:



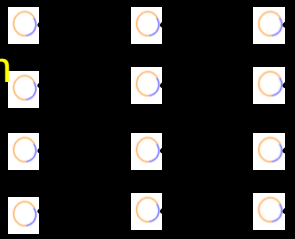
Extraction

ADN

Cloning

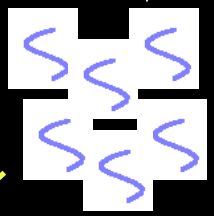


Transformation

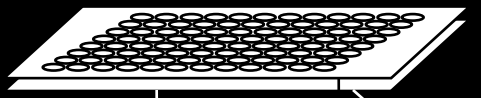


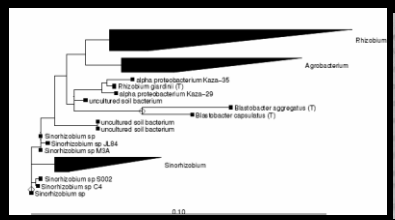
vector

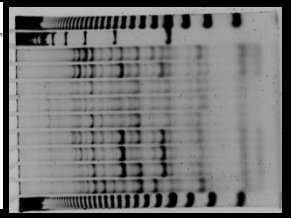
PCR

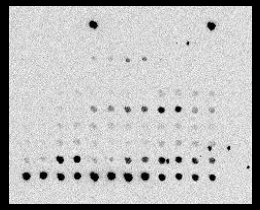


Clone Library









Culture *in vitro*

Cloning and/or sequencing

RISA, T-RFLP DGGE, Phylochip

Molecular screening

Chemical screening

Biological activity

Lombard *et al.*, 2006

METAGENOME PYROSEQUENCING

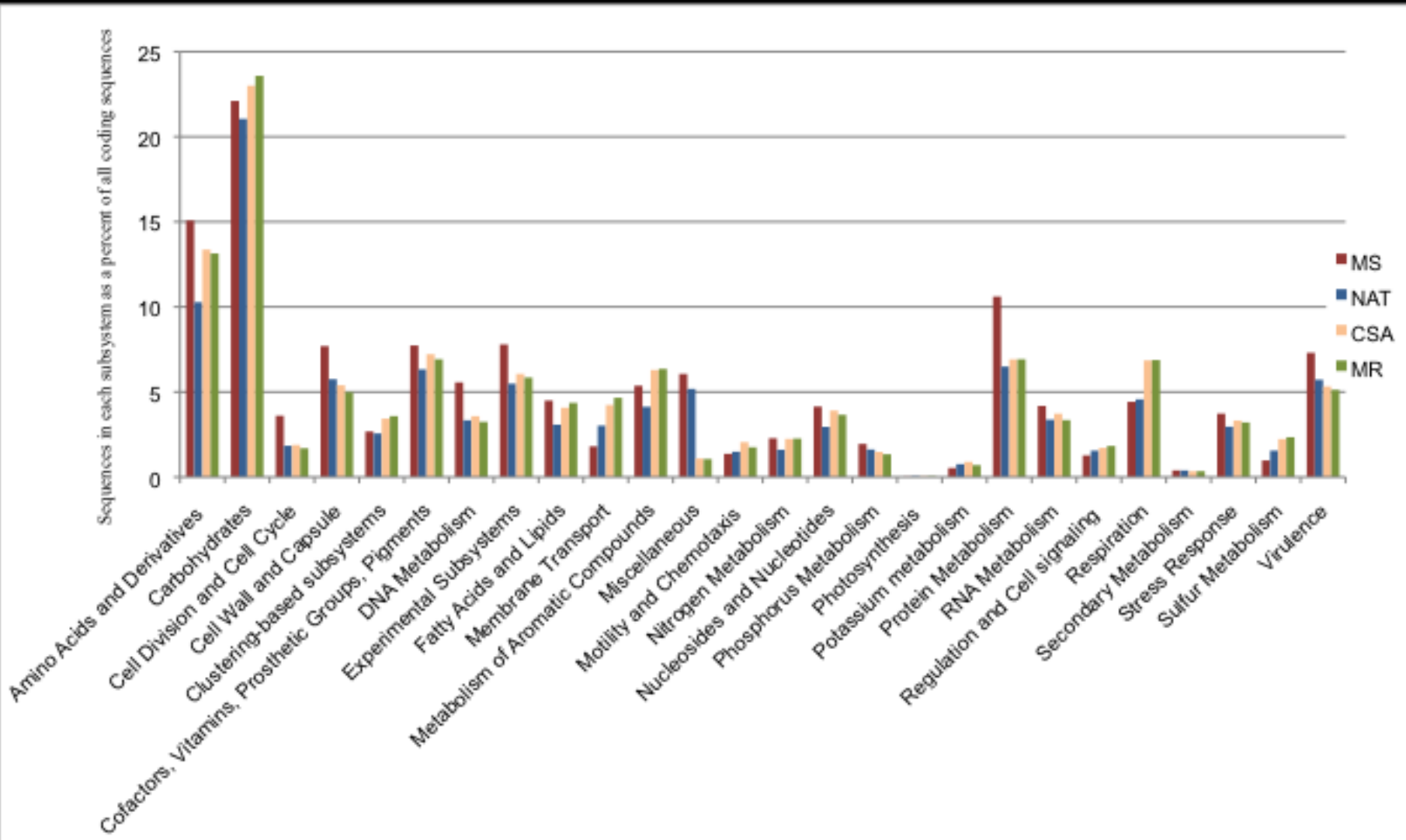
454 Life science Technology

	MS	NAT	CSA	MR
Number of sequences	182 710	163 581	181 022	162 290
Total Length of Sequences (Mb)	46	41	43	38
Average Length of Sequences (bp)	251	247	235	234

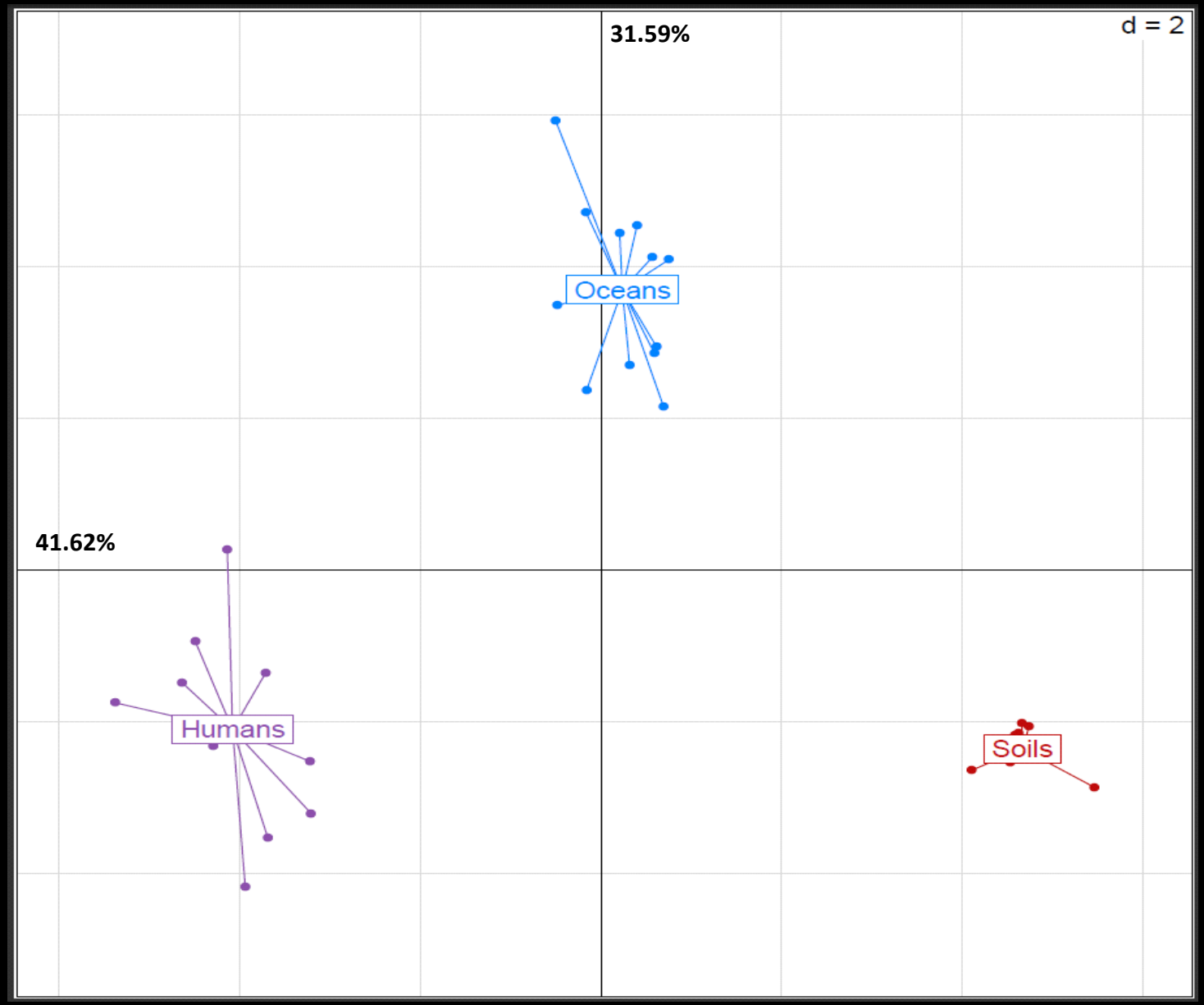
Sequence annotation : MG-RAST server (Meyer et al, 2008)

	MS	NAT	CSA	MR
Number of coding sequences	82 504	80 828	121 381	100 524
Number of subsystems	566	583	562	555
Feature in subsystems (%)	55	46	52	51
Hypothetical proteins (%)	20	27	20	20

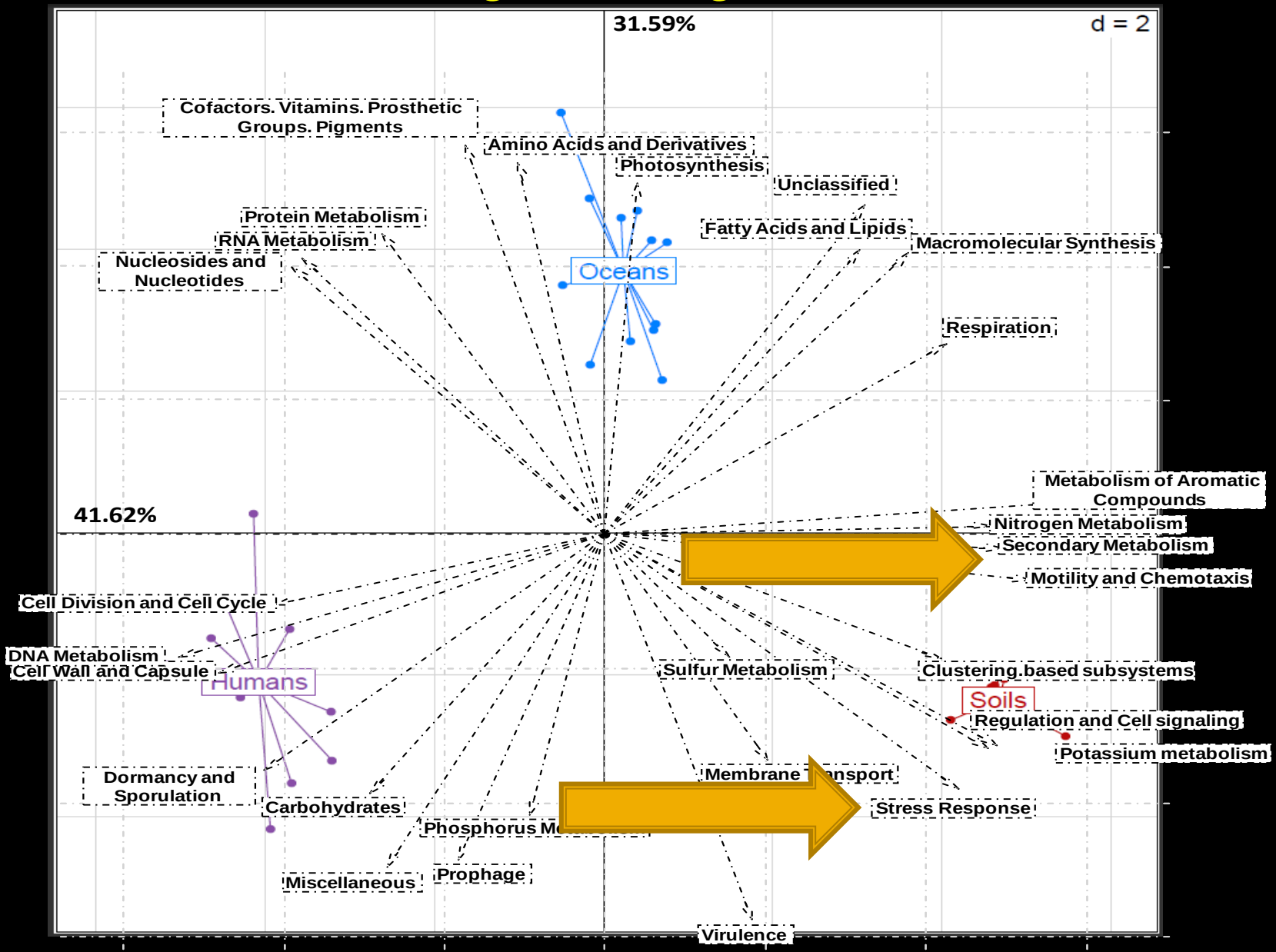
SUBSYSTEM COMPARISON



PCA comparing general functional subsystems distributions among 32 metagenomes



PCA comparing general functional subsystems distributions among 32 metagenomes



“Species”/g soil		Number of bp		Number of clones
10^4	→	4×10^{10}	→	10^6
10^7	→	4×10^{13}	→	10^9

Total b&a/g soil		Number of bp		Number of clones
10^9	→	4×10^{15}	→	10^{11}

40 kb inserts; average genome size 4×10^6

Scientific Committee on Problems of the Environment

SCOPE

http://www.icsu-scope.org/

↓

SCOPE program on Microbial Environmental Genomics

MicroEnGen III

↓

Soil Metagenome International Consortium

METASTED

↓

TERRAGENOME

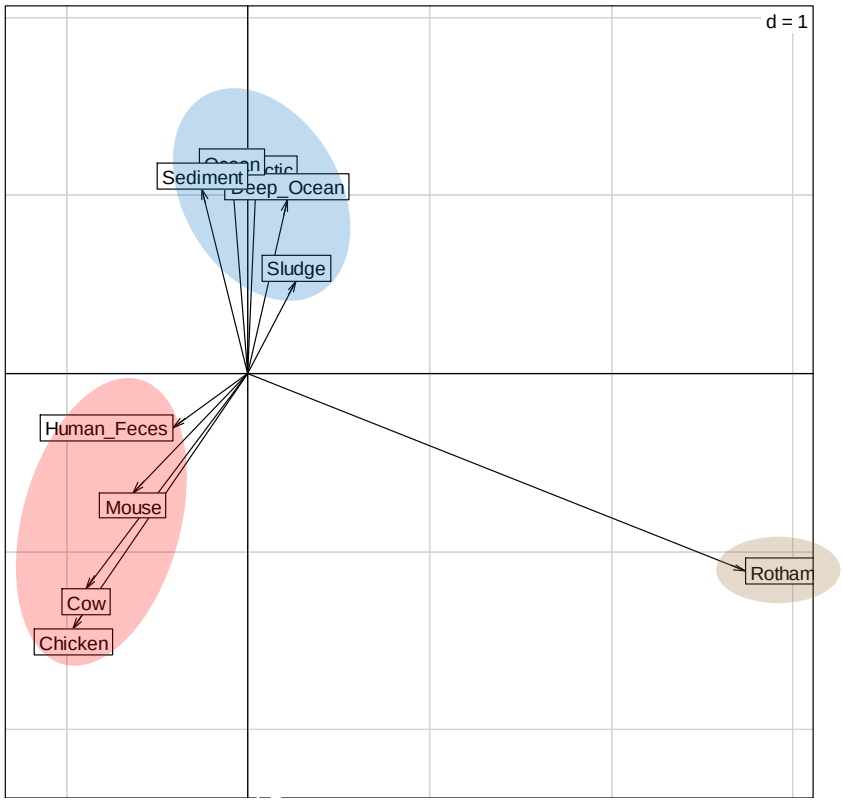
http://www.terragenome.org/

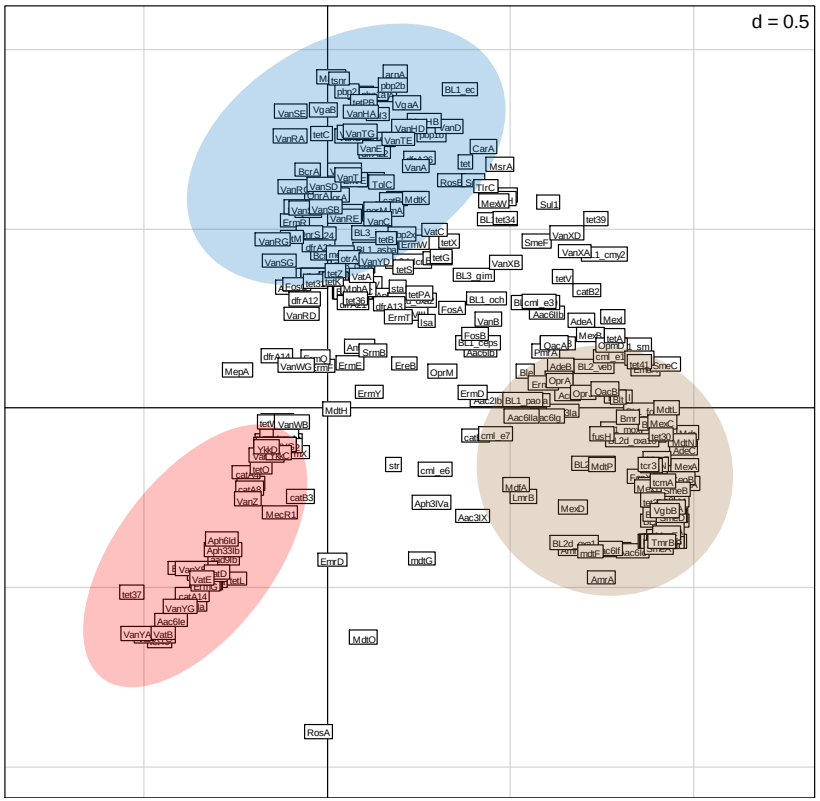
The long-term experimental site in the UK: Rothamsted http://www.rothamsted.ac.uk/

- Extensive metadata
- From 50 to 140 years of controlled experiments

Comparison of the relative distribution of ARGD in the different metagenomes by principal component analysis

Metagenomes *ARGD*





Critical issues

Composite or single samples, Sample size and suitability, Enrichments, SIP

DNA molecular weight, yield, purity and integrity

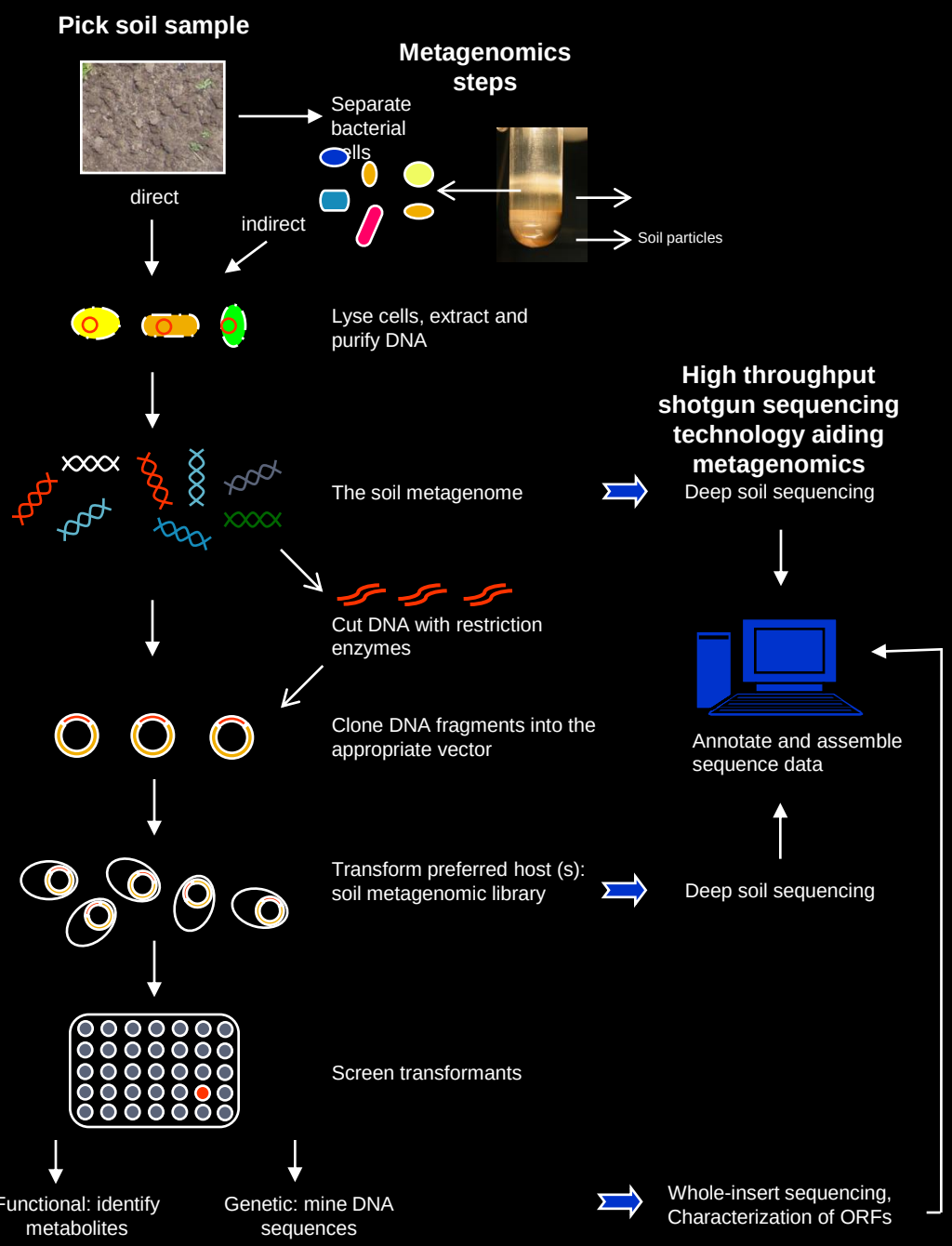
Calibration of DNA fragment size

Vector choice: insert and library size, host range, selection markers, GFP labeling

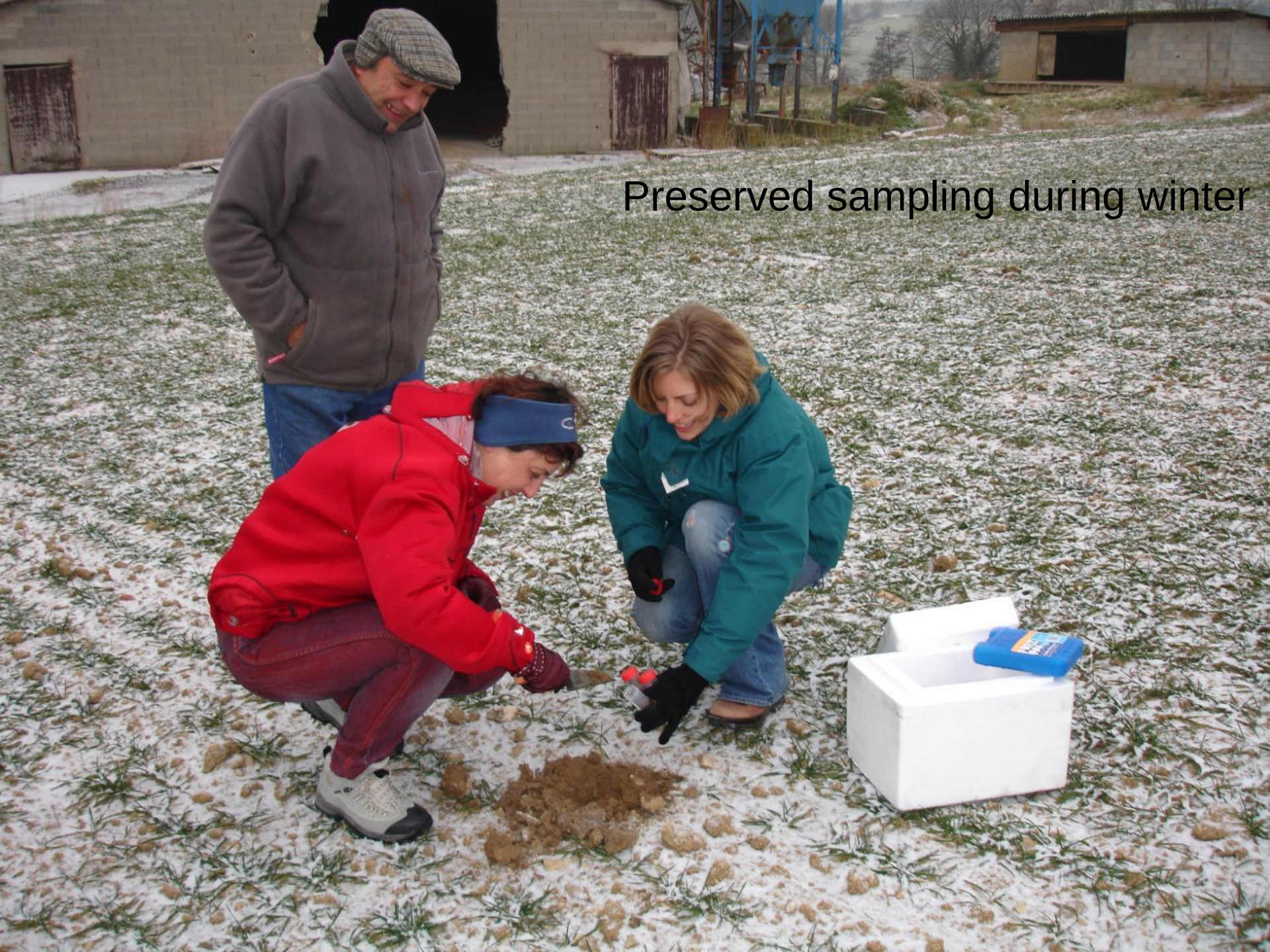
Multiple expression hosts in functional screenings

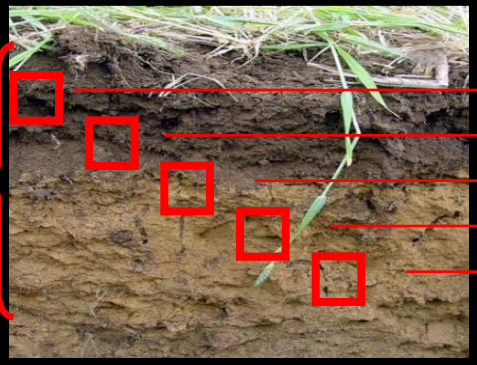
Screening efficiency for bacterial genes / metabolites

Clone pooling, Microarrays, SIGEX-FACS and other high throughput methods



van Elsas et al


Preserved sampling during winter



MIX

Plot 1

Plot 2

Plot 3

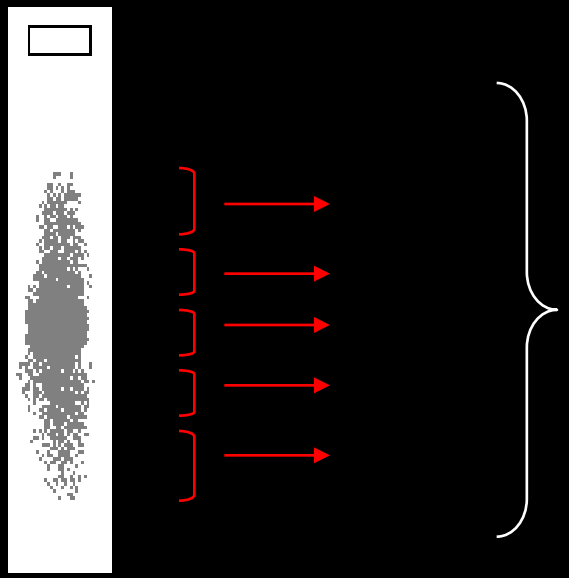
Plot 4

Plot 5

1 EXTRACTION

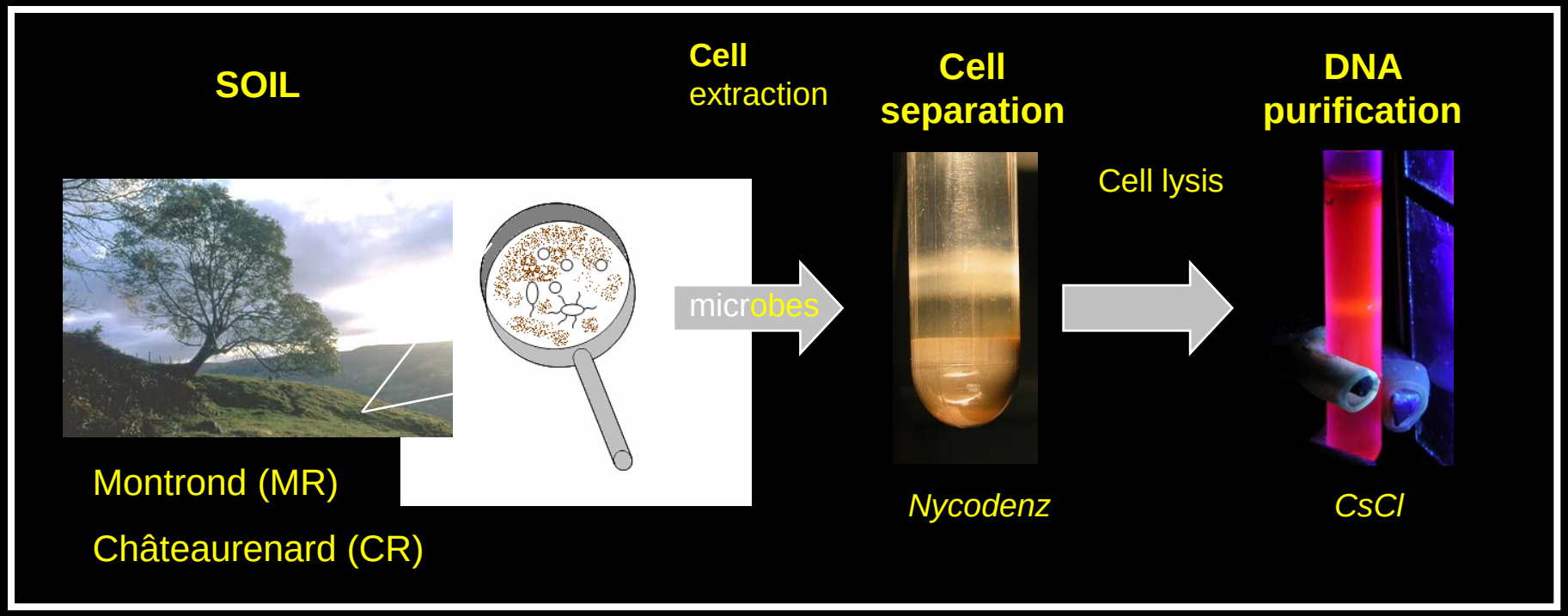
5 EXTRACTIONS

PFGE :

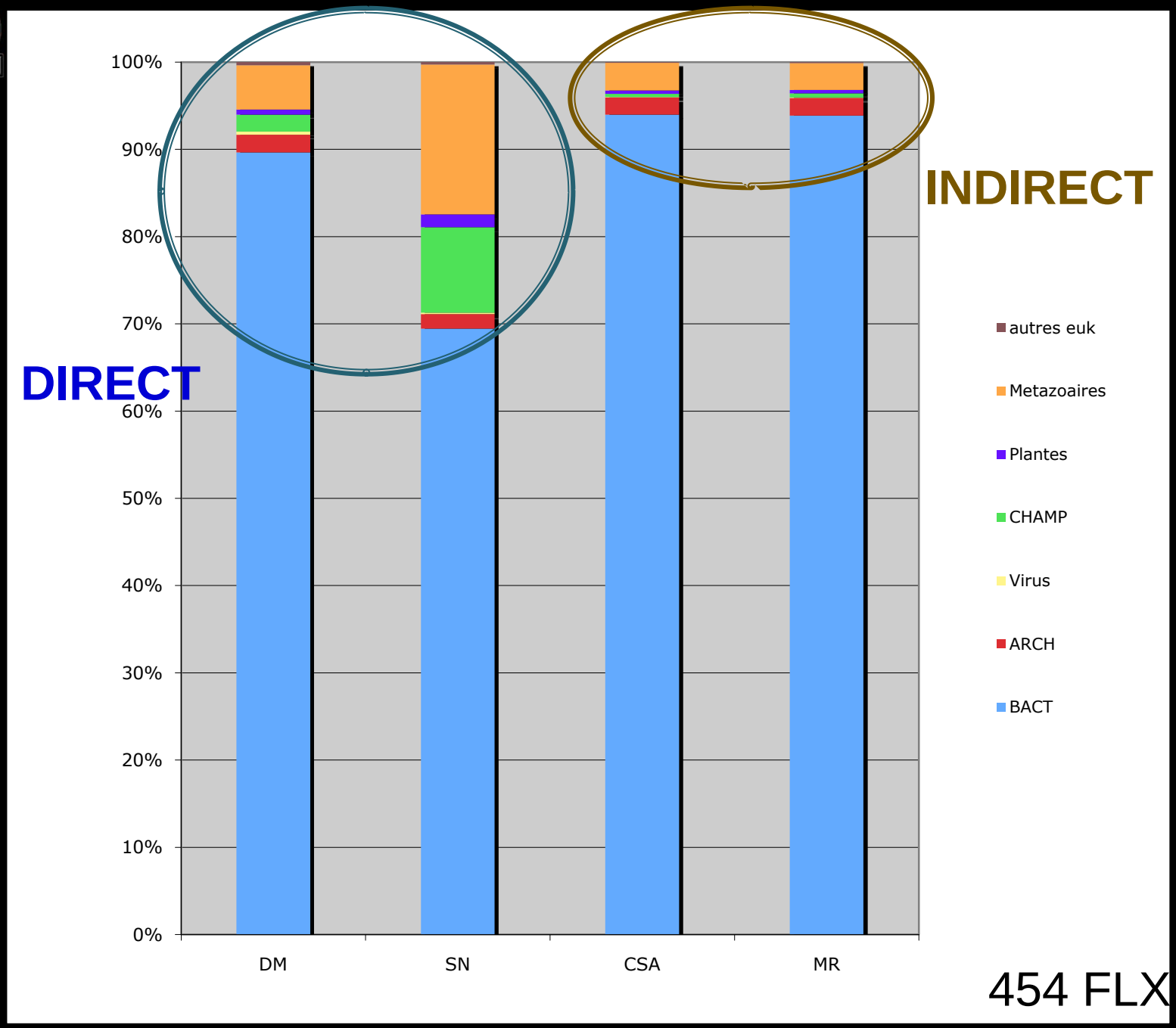


Diversity differences and distribution

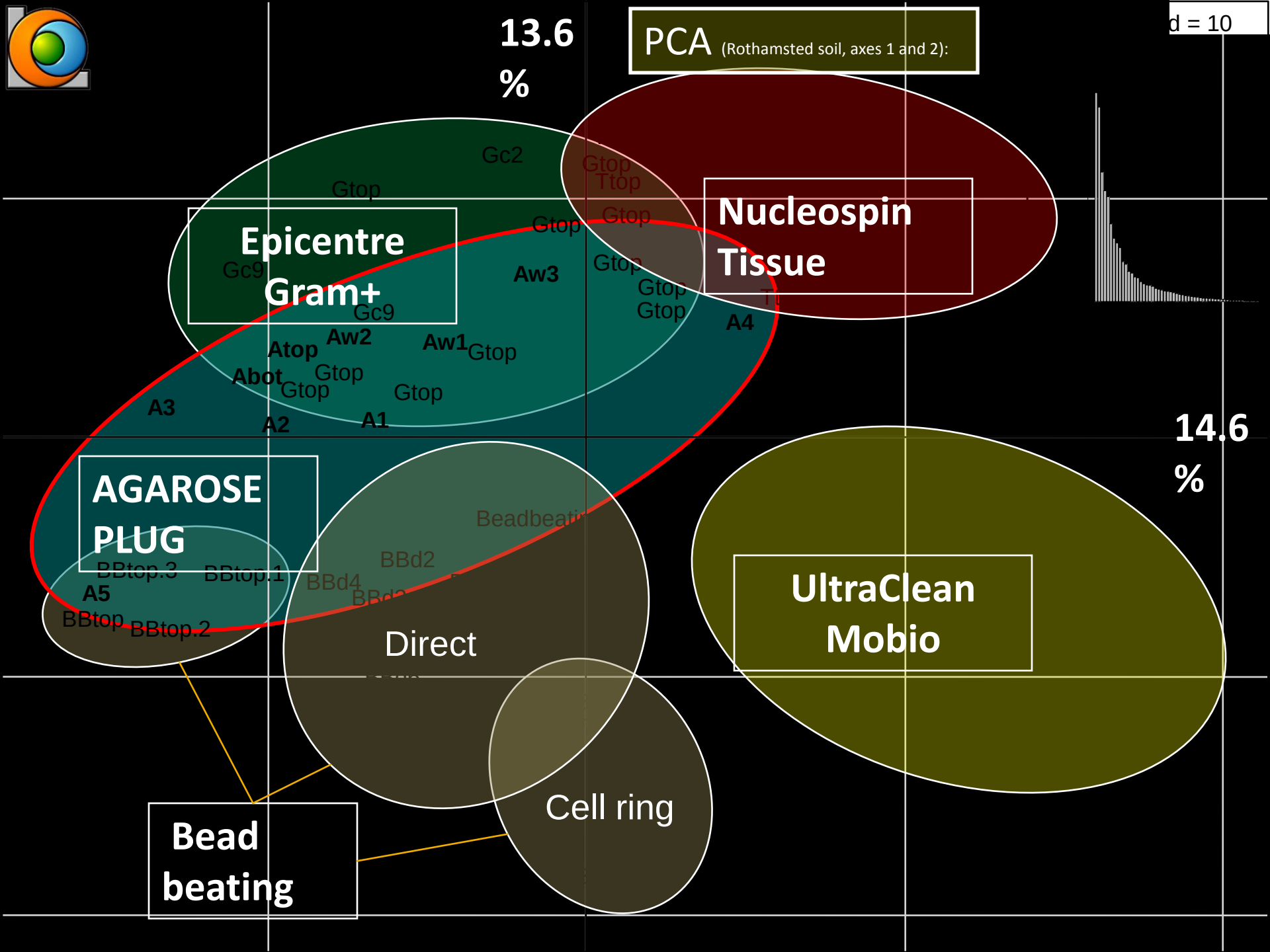
Cell separation by Nycodenz density gradient



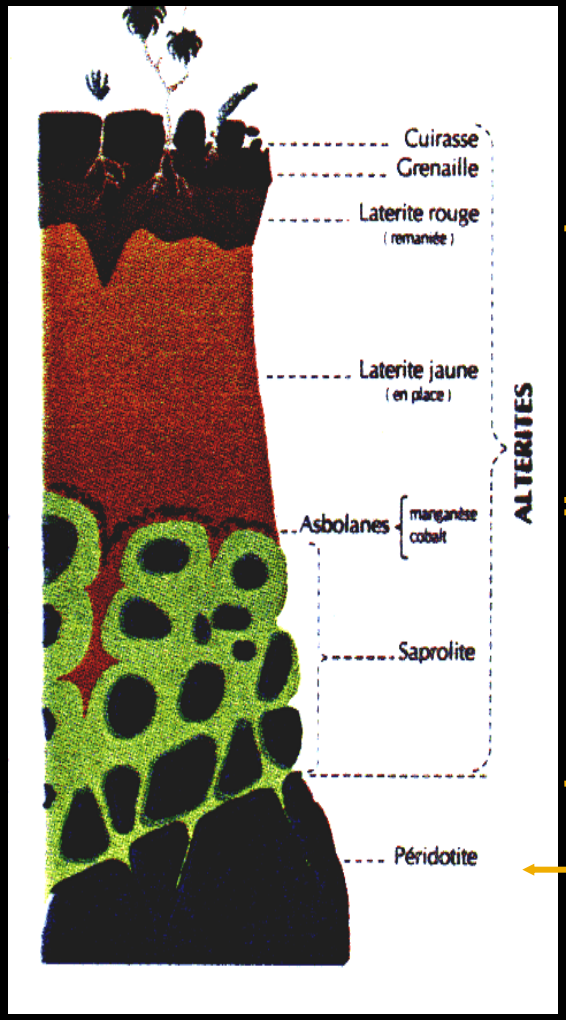
Bertrand et al


DIRECT
INDIRECT
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
DM
SN
CSA
MR
autres euk
Metazoaires
Plantes
CHAMP
Virus
ARCH
BACT

454 FLX Data


13.6 %
PCA (Rothamsted soil, axes 1 and 2):
d = 10
Nucleospin Tissue
Epicentre Gram+
AGAROSE PLUG
Direct
Bead beating
Cell ring
UltraClean Mobio
14.6 %
Gc2
Gtop
Ttop
Gc9
Aw3
Aw2
Aw1
Atop
Abot
A1
A2
A3
A4
A5
BBd2
BBd4
BBtop
BBtop.1
BBtop.2
BBtop.3

Hunting for Nickel resistance



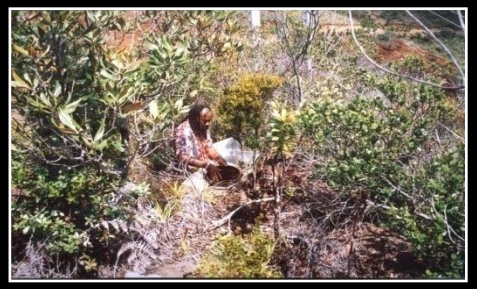
1-1,5% of Ni

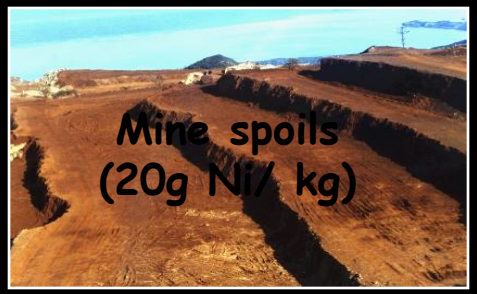
100 m

2-3% of Ni

Ultramafic rock

Natural ecosystem (4g Ni/kg)





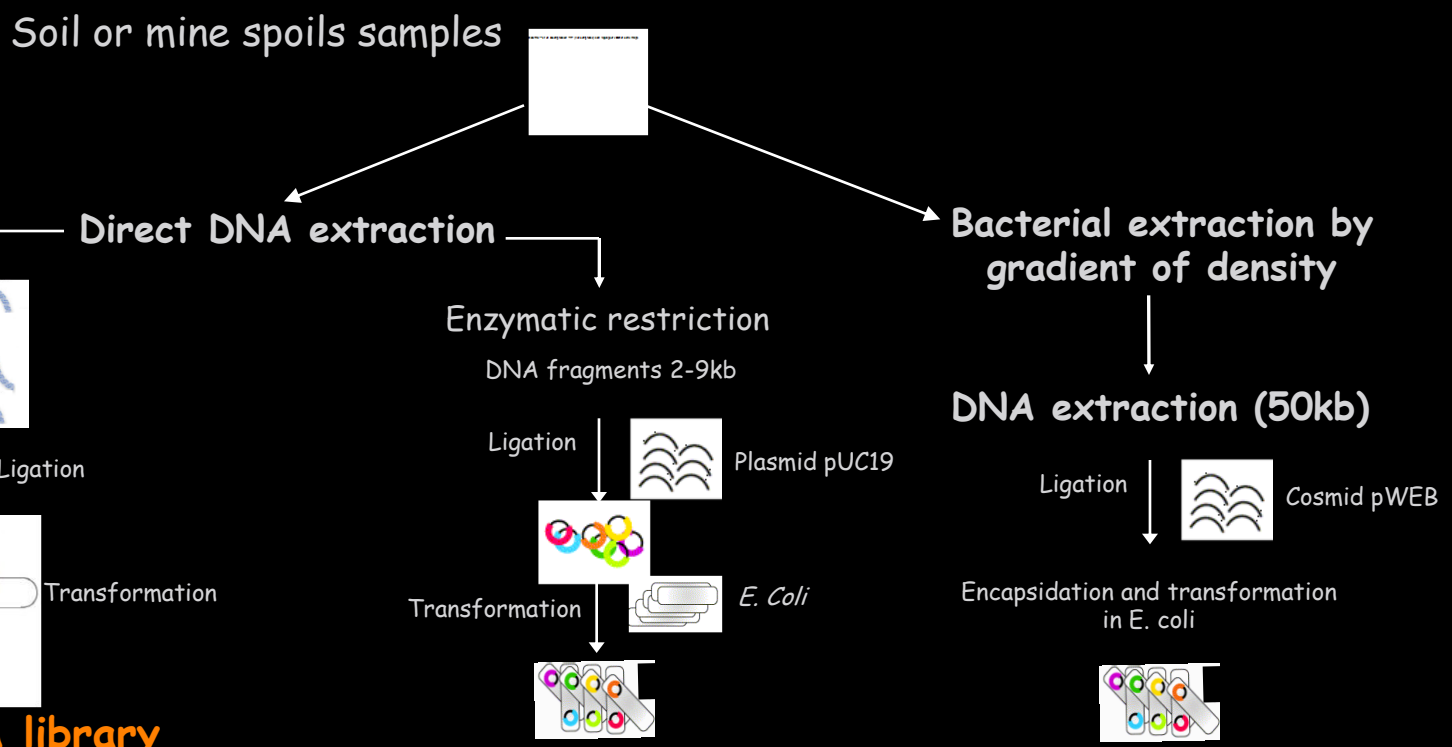
High levels of heavy metals

Ni
Co Fe

C, N...

Essential element deficiencies

Metagenomic approach



1. 16S rDNA library

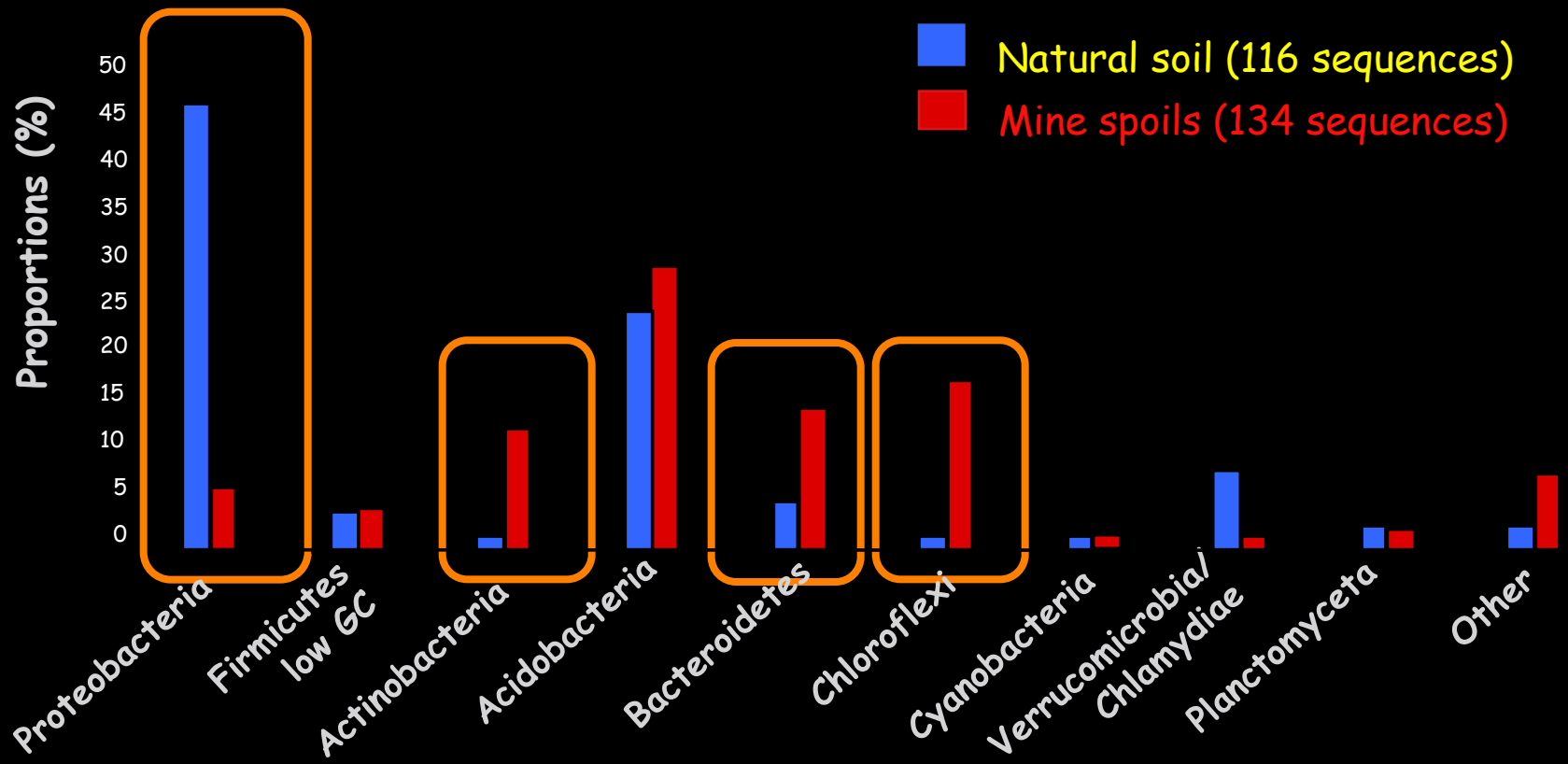
2. Total plasmidic DNA library

3. Total cosmidic DNA library

Analysis by sequencing : to determine the bacterial community structure

Analysis by screening based on activity: to select nickel resistant clones

Bacterial community structure analysis Based on 16S rRNA library sequences



Hery et al

Metagenomic insert identification

DME3

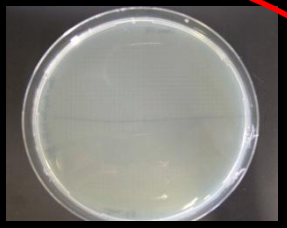
Sub-cloning

Sequencing

Transposon insertion
(EZ::Tn5<Kan 2>)

Gene inactivation

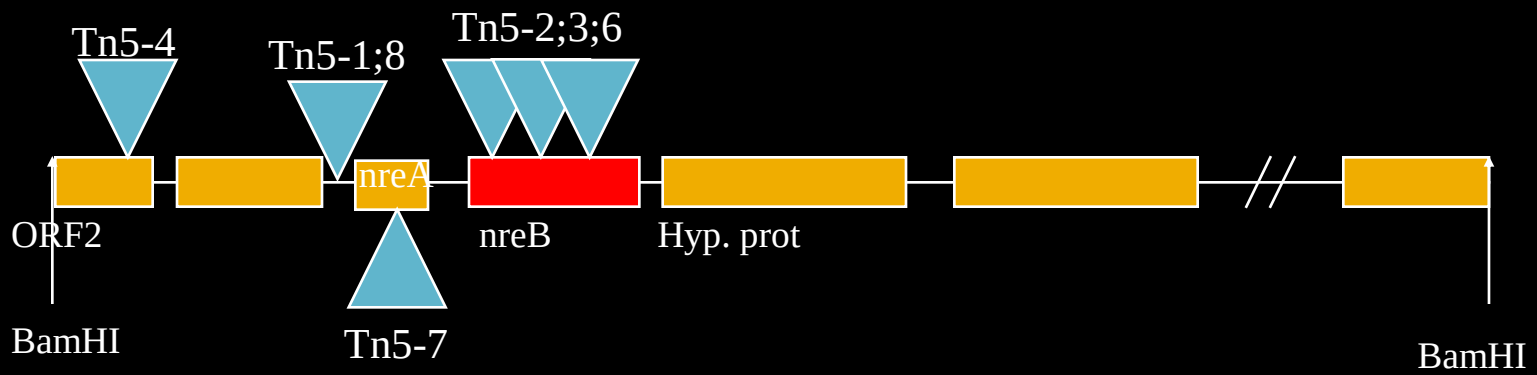
2 sensible clones on nickel



Without nickel

Nickel (4 mM)

Transposons to access novel genes in metagenomic libraries

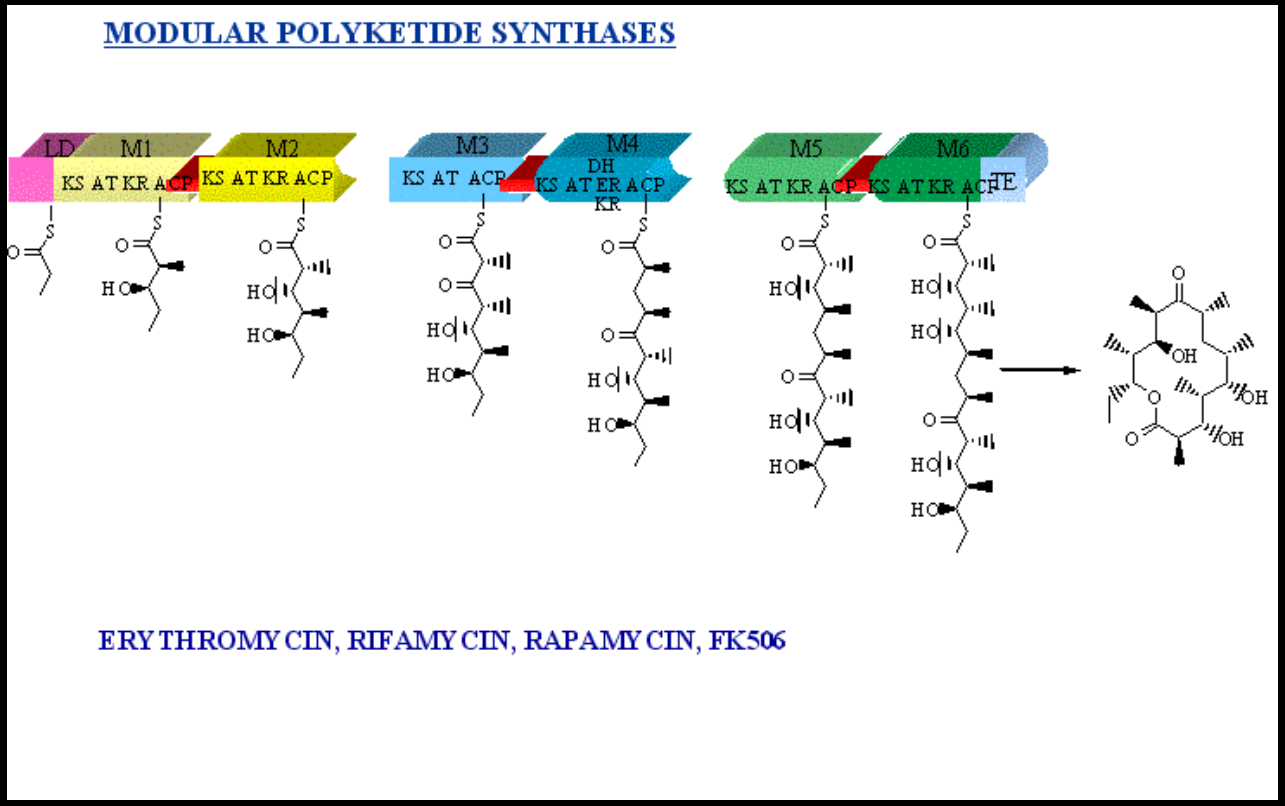


Fall et al

Antibiotic Production and Antibiotic Resistance

- Numbers of producers vs numbers of resistors
- Production diversity vs resistance diversity
- Gene transfer and gene diversity
- Hunt for new molecules

Antibiotic Production Diversity

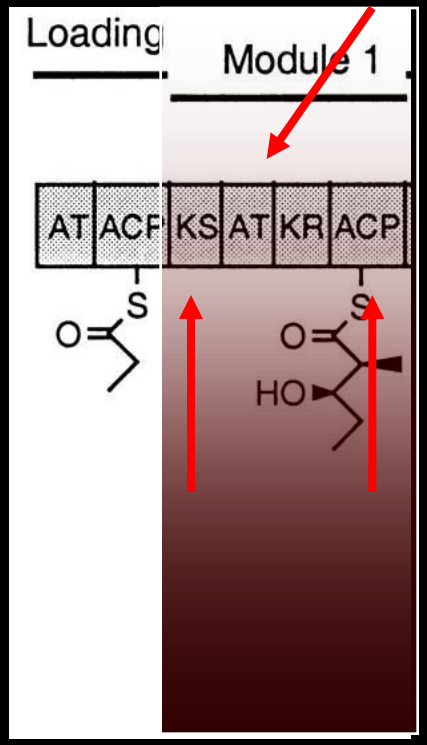


Type I PKS (PKSI)

The structure of polyketides produced by PKSI is related to the specific linear organization in domains and modules

Ginolhac et al

A minimal extender module is composed with at least three domains:



*** A ketosynthase domain (KS) for decarboxylative condensation of the extender unit onto the growing chain**

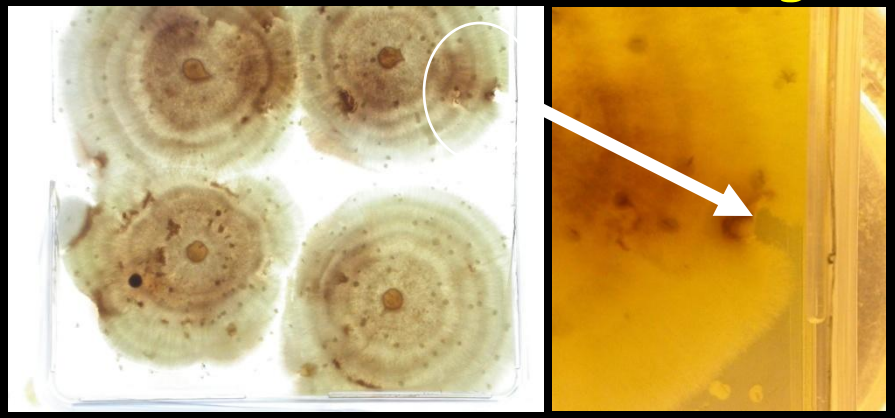
*** An acyltransferase domain (AT) for substrate selection and transfer**

*** An acyl carrier protein (ACP) which loads the growing chain**

Ginolhac et al

Functional and molecular screening

Functional:
antagonism against
R. solani AG3



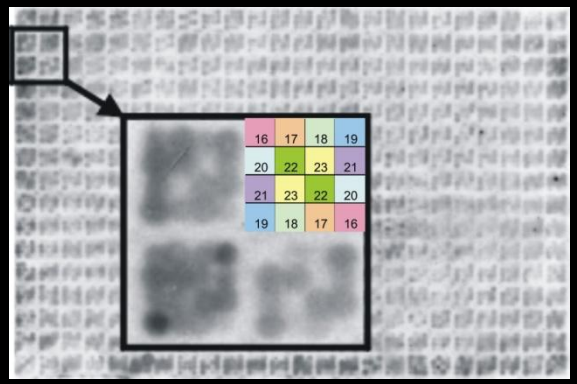
Slipstream inhibition of mycelial growth

van Elsas et al

Metagenomic clone library: 16,000 clones
77 000 clones

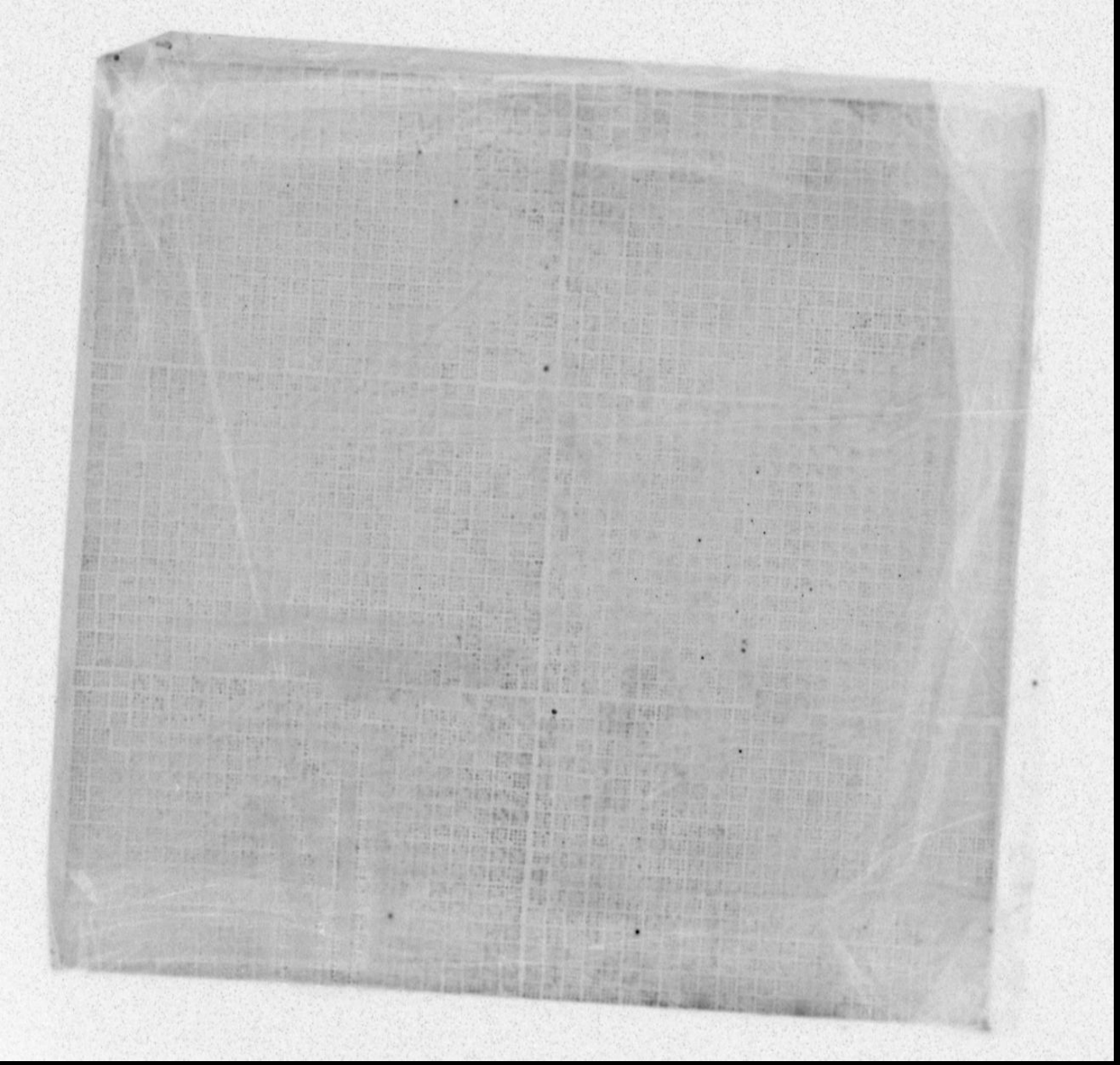
Clone selection

Hybridization with PKS1 probe



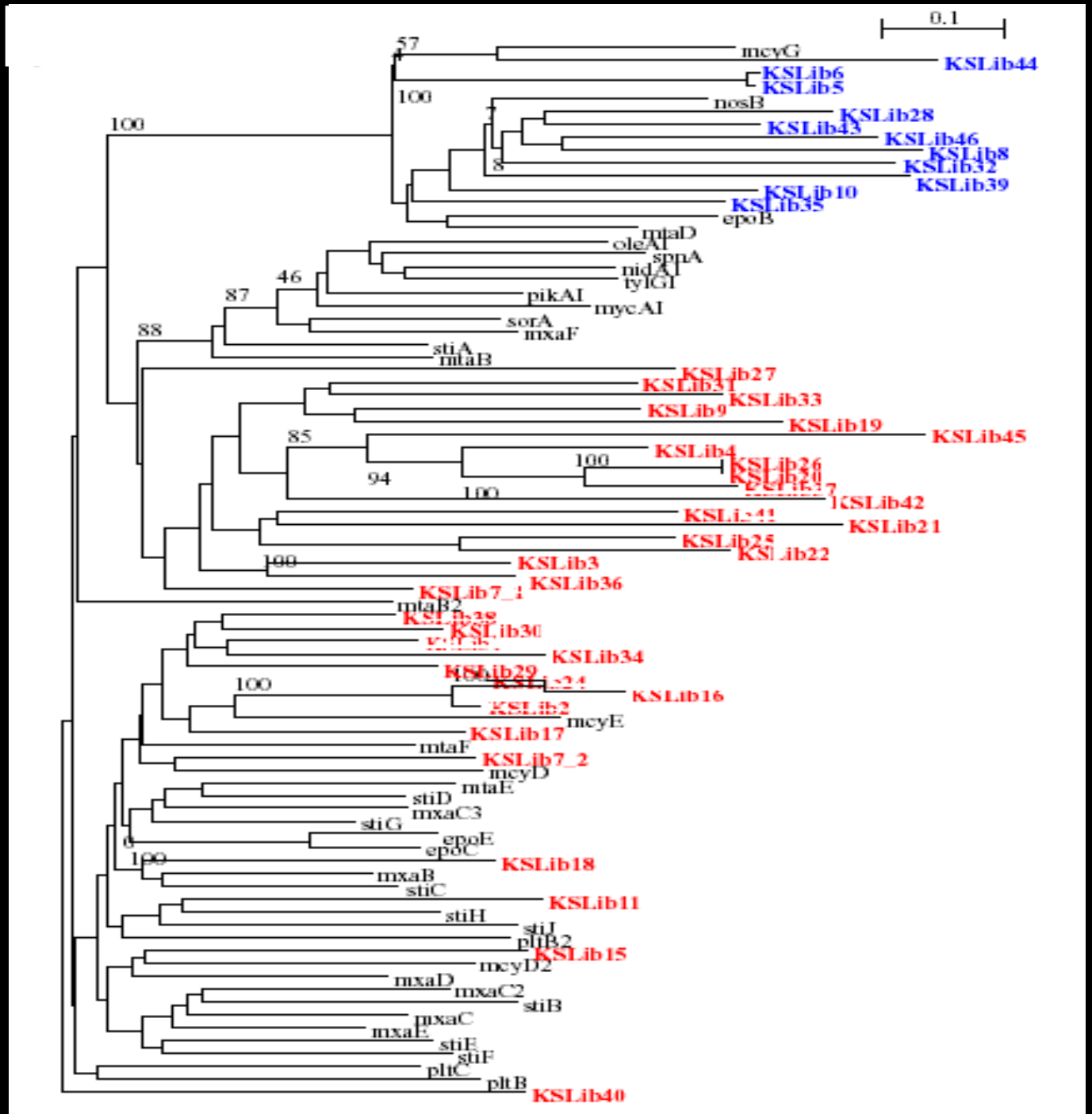
Molecular:
screening for PKS1
biosynthetic genes

Ginolhac et al



Sequenced clones and original DNA by 454 FLX

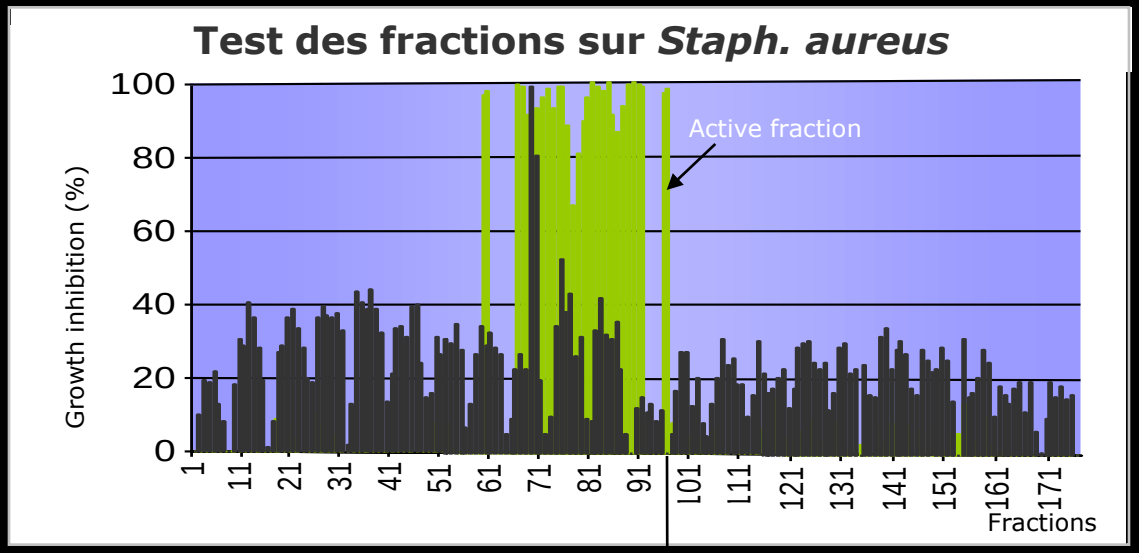
Demanèche et al

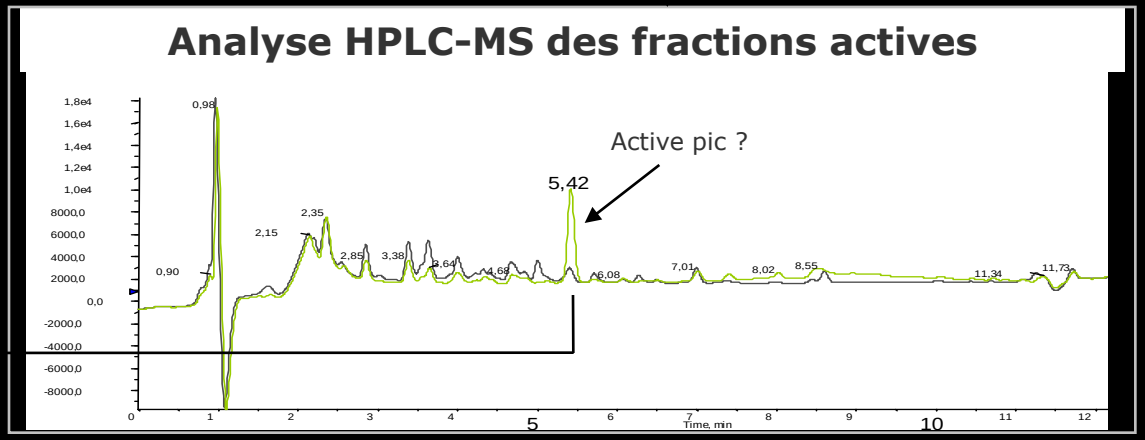


Ginolhac et al

Identification of active fractions

Bioguided Purification





Libragen S.A.

CONCLUSIONS

Metagenomics provides access to « novel » genes

1. Sequencing helps target appropriate environments and possible gene families
    - Sequence interpretation is heavily dependent on genome sequences in database – need more genomes sequenced

2. Clone librairies provide phenotypic and genotypic screening
    - Phenotypic screening is dependant on host cell, etc, but can be used to produce targeted enzyme – compound purification

Environmental Microbial Genomics Group

www.genomenviron.org