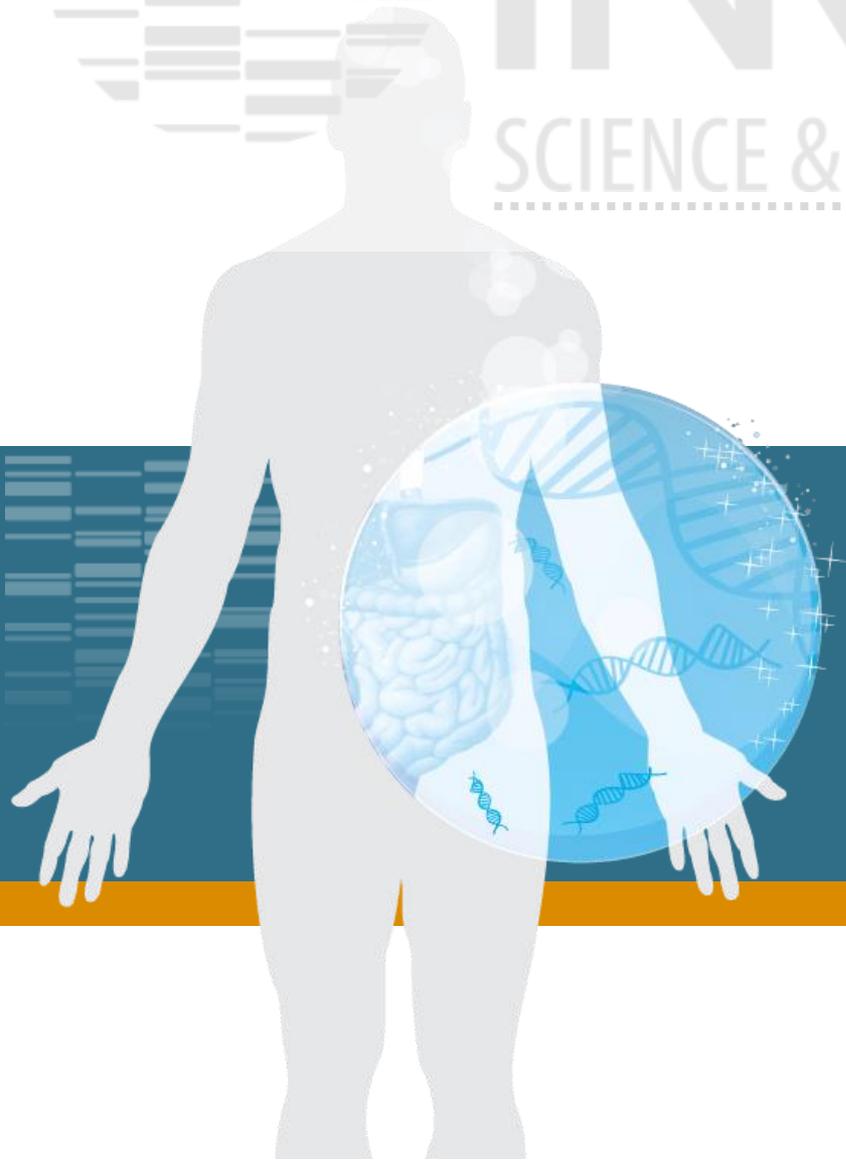




INRA
SCIENCE & IMPACT



metagenopolis
mgps.eu



RARe

Projet Ring Test 16S

10 décembre 2020

Christian Morabito,

Mathieu Almeida, Hugo Roume, Joël Doré



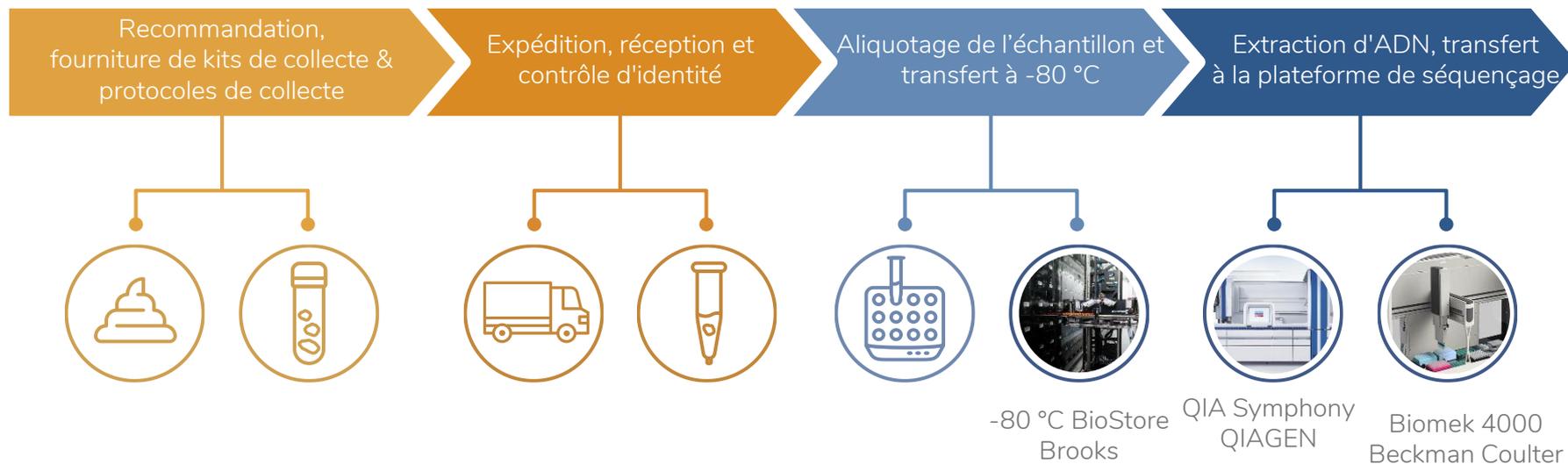


Responsable scientifique : Joël Doré

Responsable opérationnel : Christian Morabito

Equipe composée de 3 assistants-ingénieurs

Espace confiné (niveau 2) composé de deux laboratoires, d'une salle de biobanking et d'une pièce des robots



Plateforme SAMBO

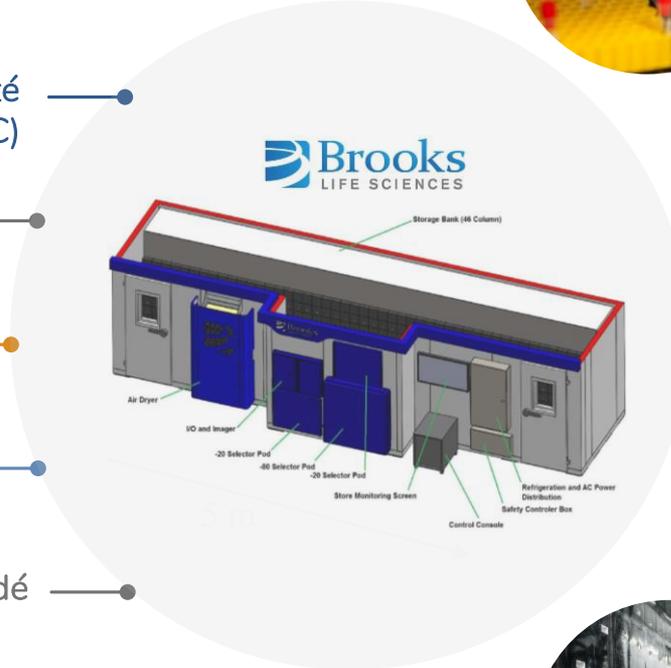
Stockage sécurisé et de qualité
de vos échantillons (-80° C)

Pipeline standardisé
et automatisé

Capacité de stockage :
600 000 échantillons

Base de données intégrée
pour la traçabilité
des échantillons

Tube 2D barcodé



Qu'est ce qu'un ring test?

A **ringtest** is part of an external **quality assurance programme** for a measuring method.

Usually a **reference institute (MGP)** sends **identical samples (DNA)** which have to be analysed for **special parameters (Bacterial profile)** to **different laboratories (INRAE & private)**.

The industrial, medical or research laboratory gets a **limited time (3 months)** with a deadline to send in the **analysis results (Bacterial count tables)**.

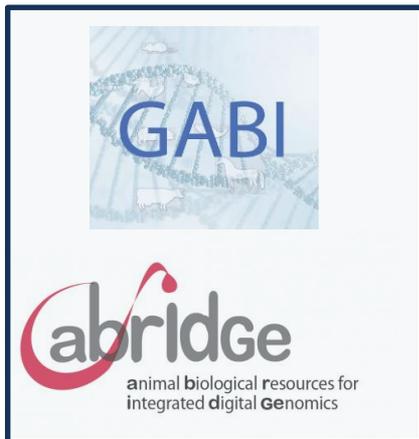
The statistical evaluation and interpretation of the laboratories' results is a great help for all participating labs as it **opens the possibility to assess the quality of their analysis compared to other laboratories**.

For accredited laboratories (e.g. ISO 17025, etc.) the participation in ringtests is obligatory. Nevertheless, also non accredited labs take part in ringtests

Reference:

<https://en.wikipedia.org/wiki/Ringtest>

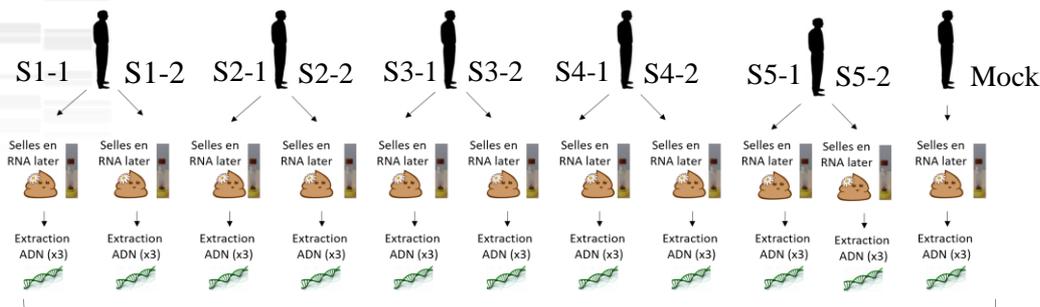
Remerciements



Joel Doré initiateur du projet

Design experimental: 5 individus, 2 temps, 3 réplicats

- 5 individus (S) prélevés à 2 temps (1 & 2)
- ADN séquencés en triplicats



Travail réalisé sur la plateforme extraction ADN de Métagénopolis

Travail réalisé sur la plateforme séquençage et analyse bio informatique de Métagénopolis (3 mois)

Séquençage shotgun

Abondance des espèces métagénomiques (MGS)

Agglomération des MGS par genre

Travail réalisé sur les plateformes de séquençage et analyse bioinformatique de chaque partenaire du projet (3 mois)

Séquençage 16S

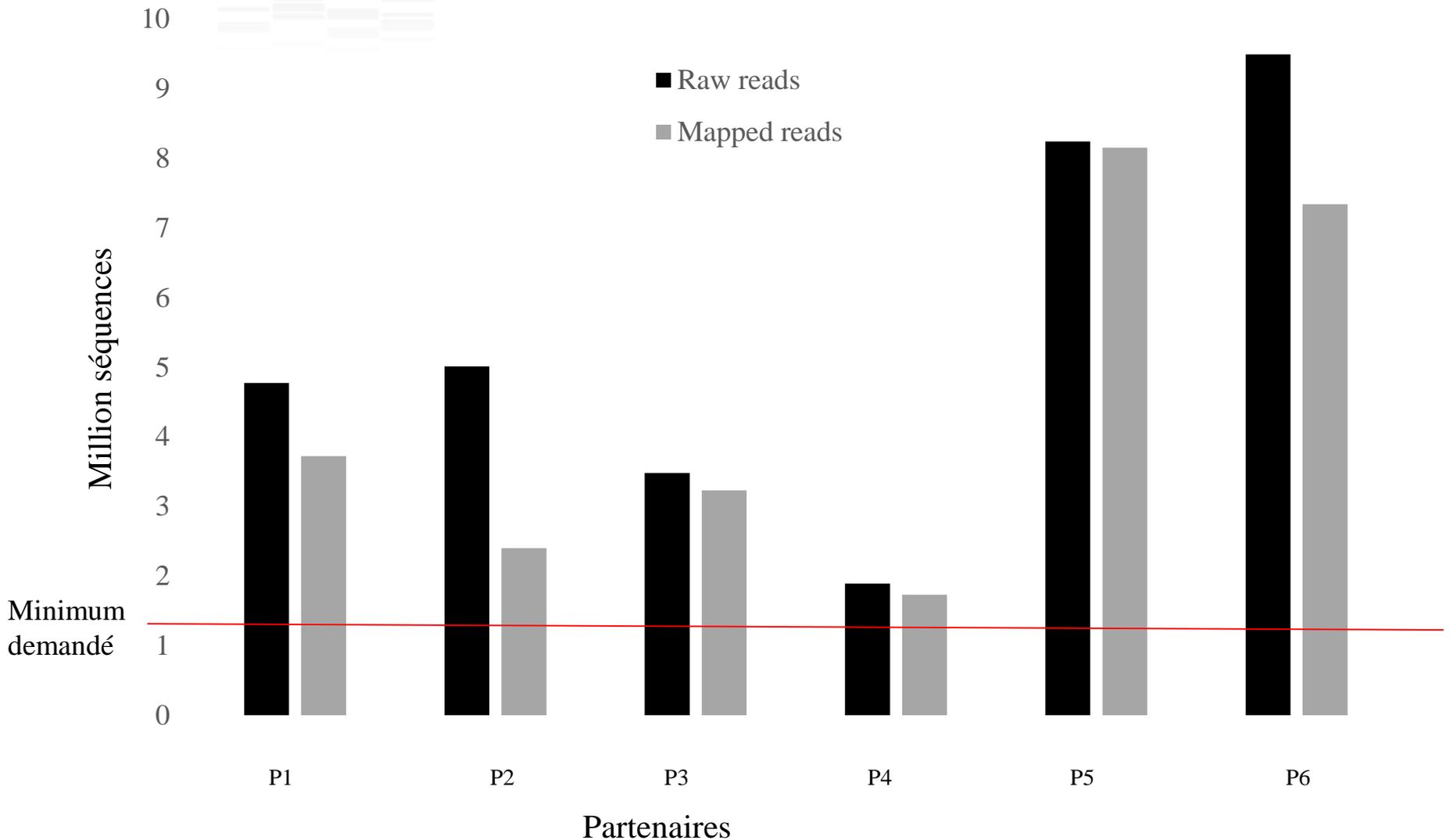
Abondance des OTU

Tableau récapitulatif des pipelines de séquençage et d'analyse en bioinformatique utilisés

Séquençage région hypervariable V3-V4 amplifié par PCR et séquençé sur le séquenceur MiSeq (Illumina) en 2x250 bp (ou de façon très minoritaire 2x300, P3). **P2 a analysé les mêmes données de séquençage que P1**

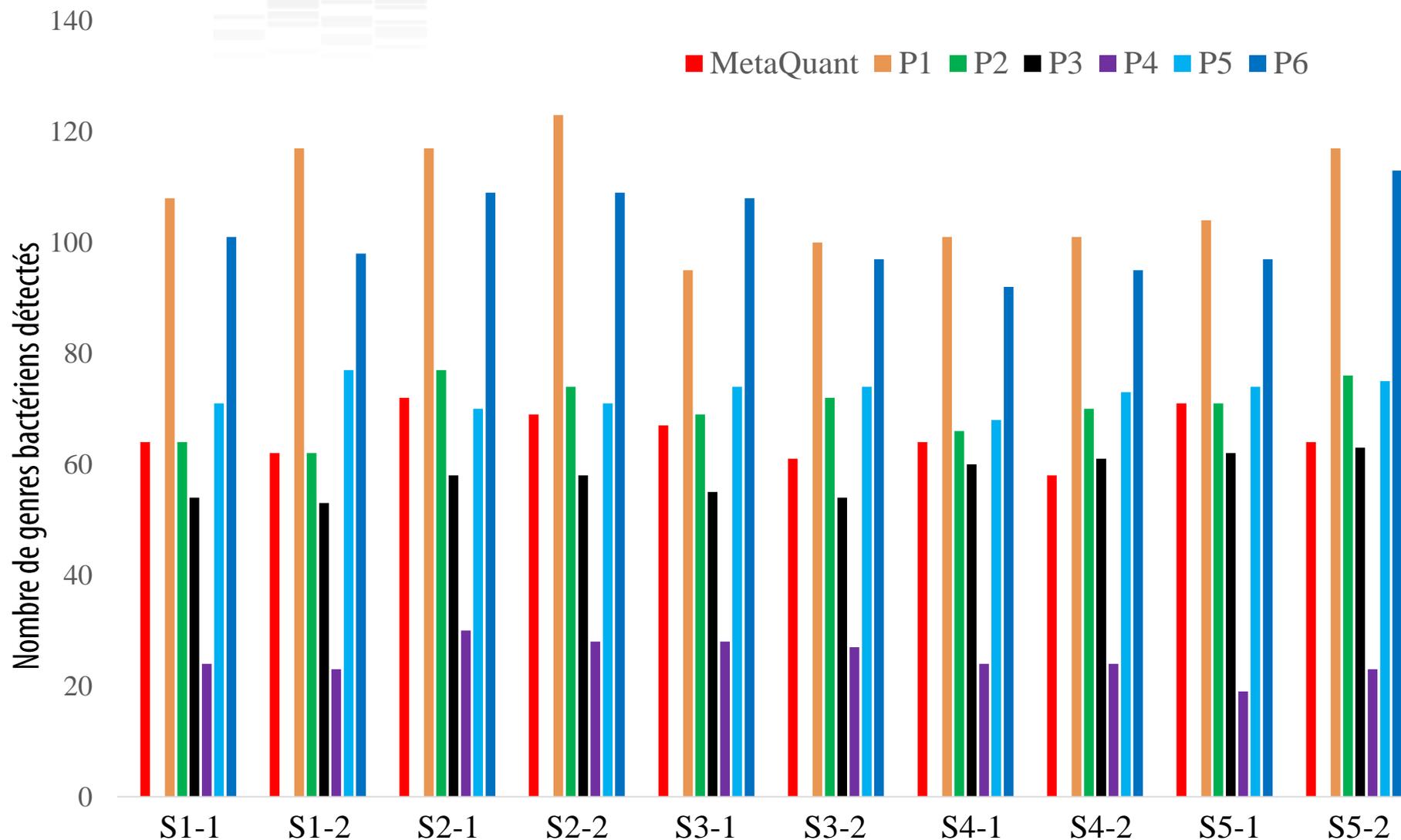
Partenaires	Pipeline	Clustering	Base de donnée taxo	Elimination des séquences
P1	mothur 1.36.1	homologie Vsearch	Greengenes	chimères + homopolymères
P2	« maison »	homologie Vsearch	RDPTools	Chimère
P3	« maison »	différences Swarm	SILVA 132	Chimères + peu abondante
P4	QIIME v1.9.1	homologie Usearch	Greengenes v13.8	Chimères
P5	QIIME v1.9.1	homologie Usearch	SILVA v.123	peu abondante
P6	mothur 1.40.0	homologie Vsearch	Greengenes v13.5.99	Chimères + peu abondant

Profondeur de séquençage minimum atteinte pour tous les partenaires

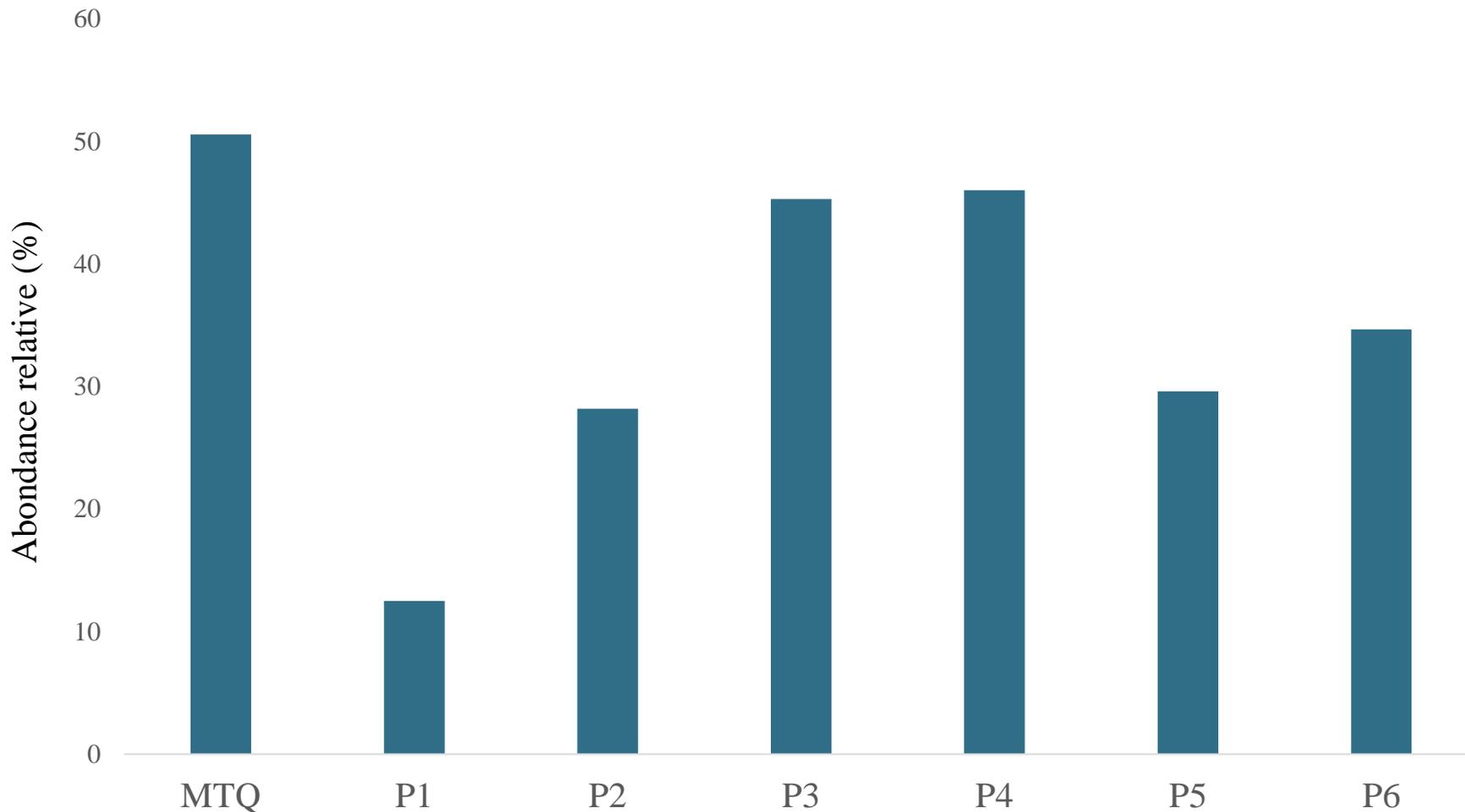


Comparaison du nombre de genre bactérien

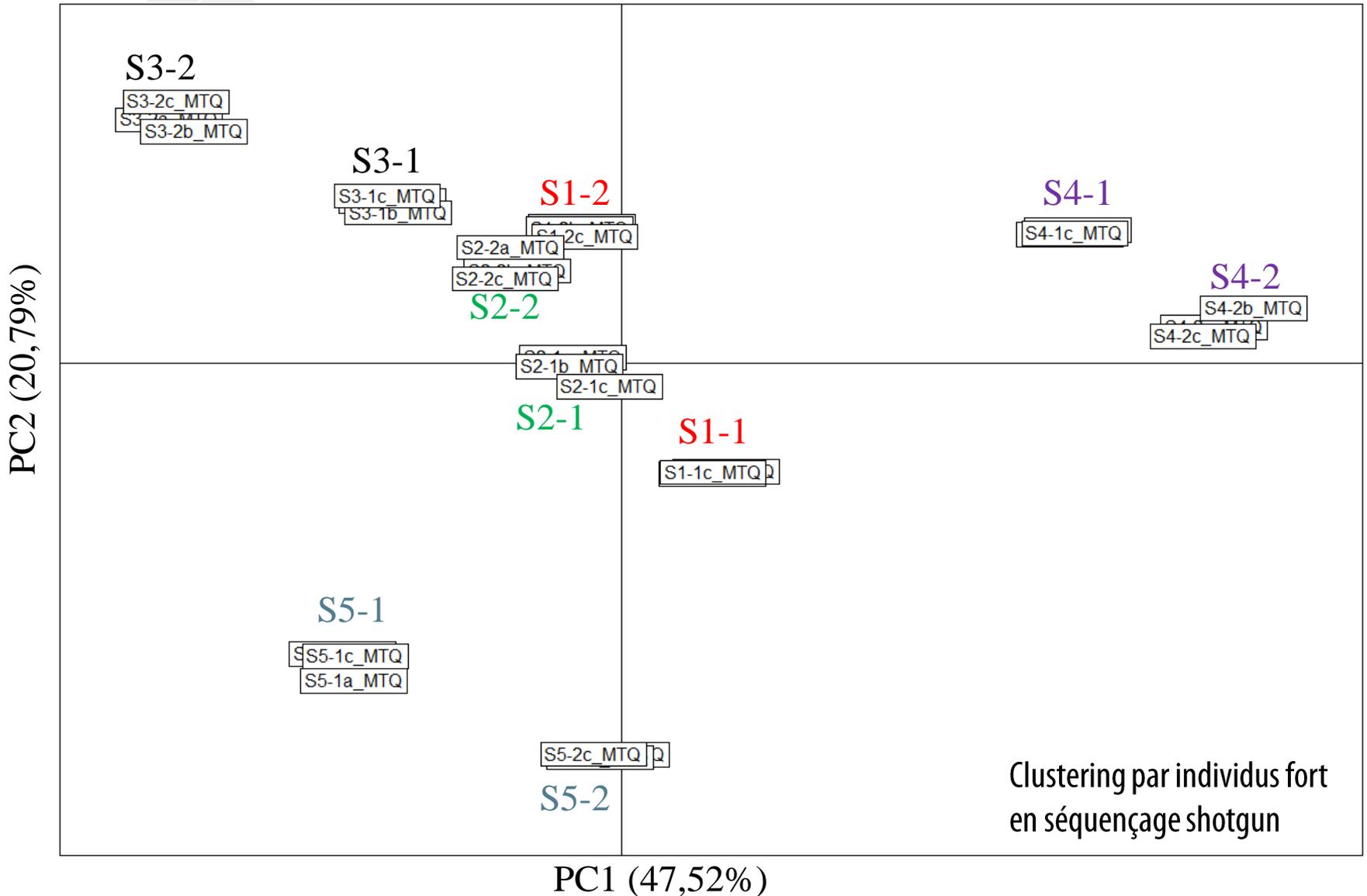
Séquençage complet vs séquençage 16S



Moyenne de l'abondance relative des genres bactérien non classifiés au genre par partenaires en comparaison au données shotgun



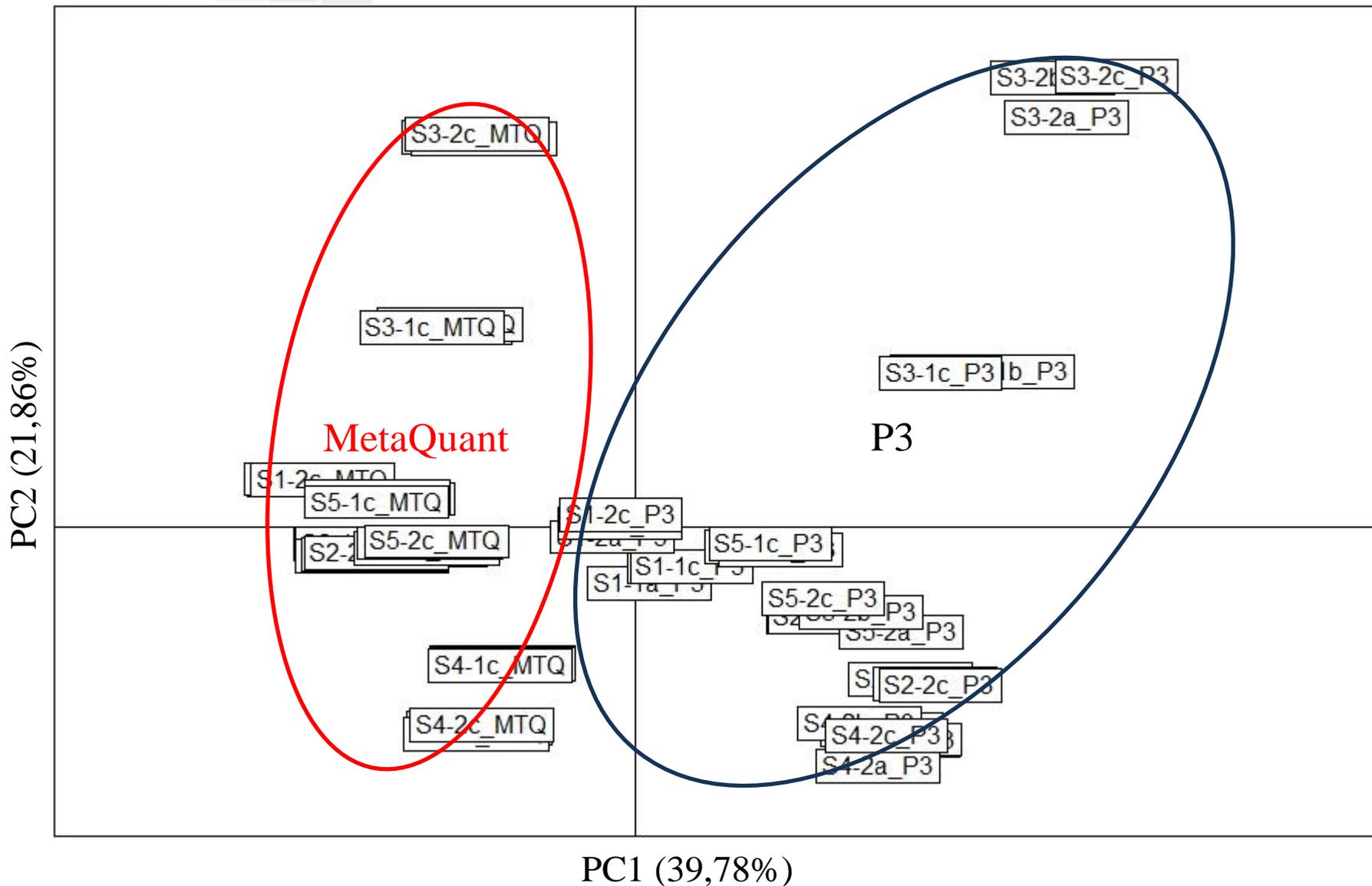
Analyse en beta-diversité des genres bactérien à partir des données shotgun (PCoA – Bray-Curtis)





Shotgun *vs.* P3

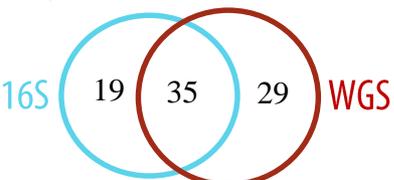
Mesure de similarité du profile bactérien au genre 16S vs. Shotgun du partenaire 3 (P3) - ACoP



Profile bactérien au genre 16S vs. Shotgun: sujet S1

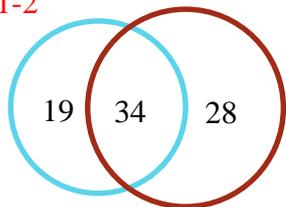
Gemmiger, Akkermansia, Methanobrevibacter et Candidatus Cibiobacter sont absent

S1-1



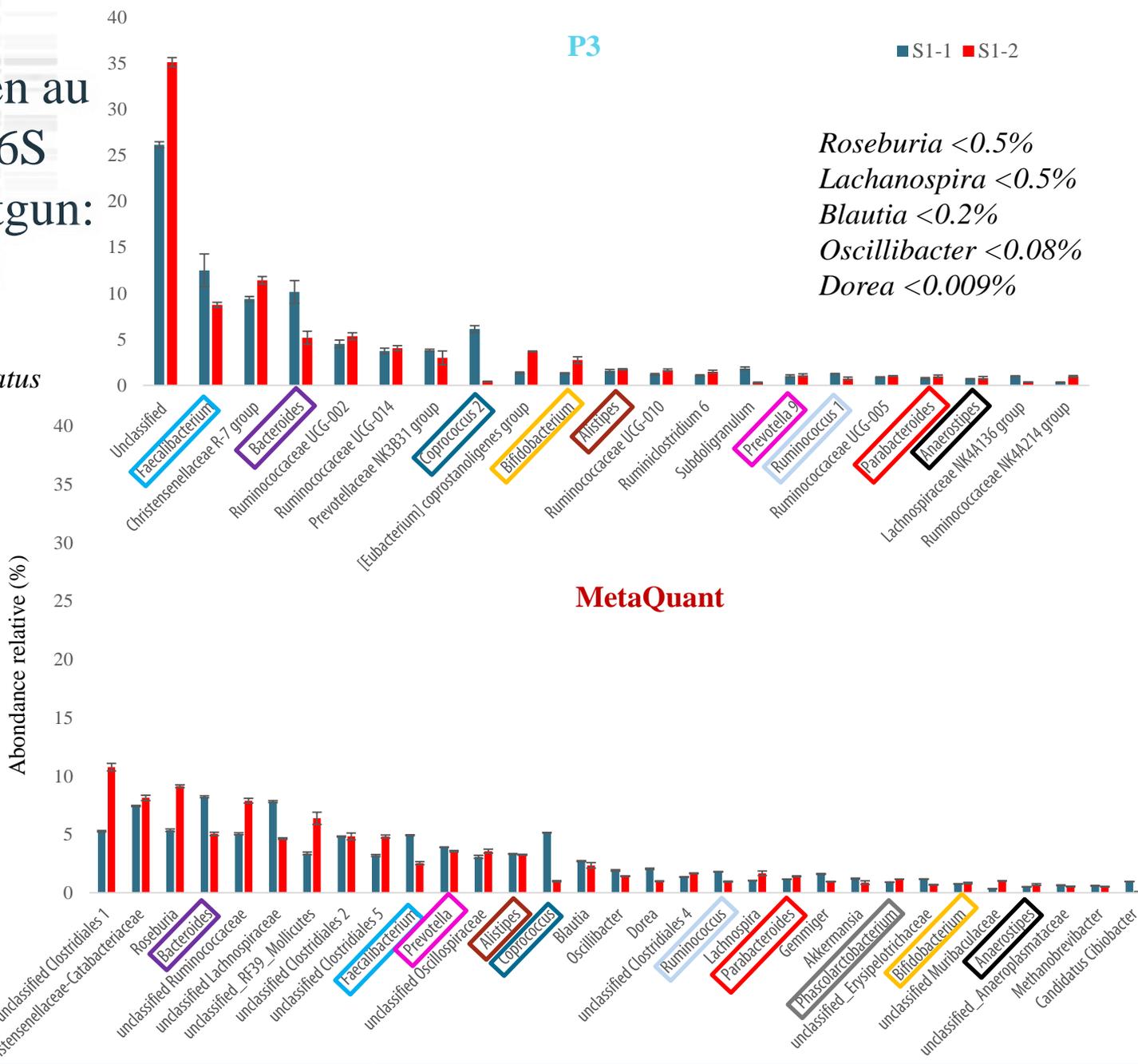
N=84

S1-2



N=82

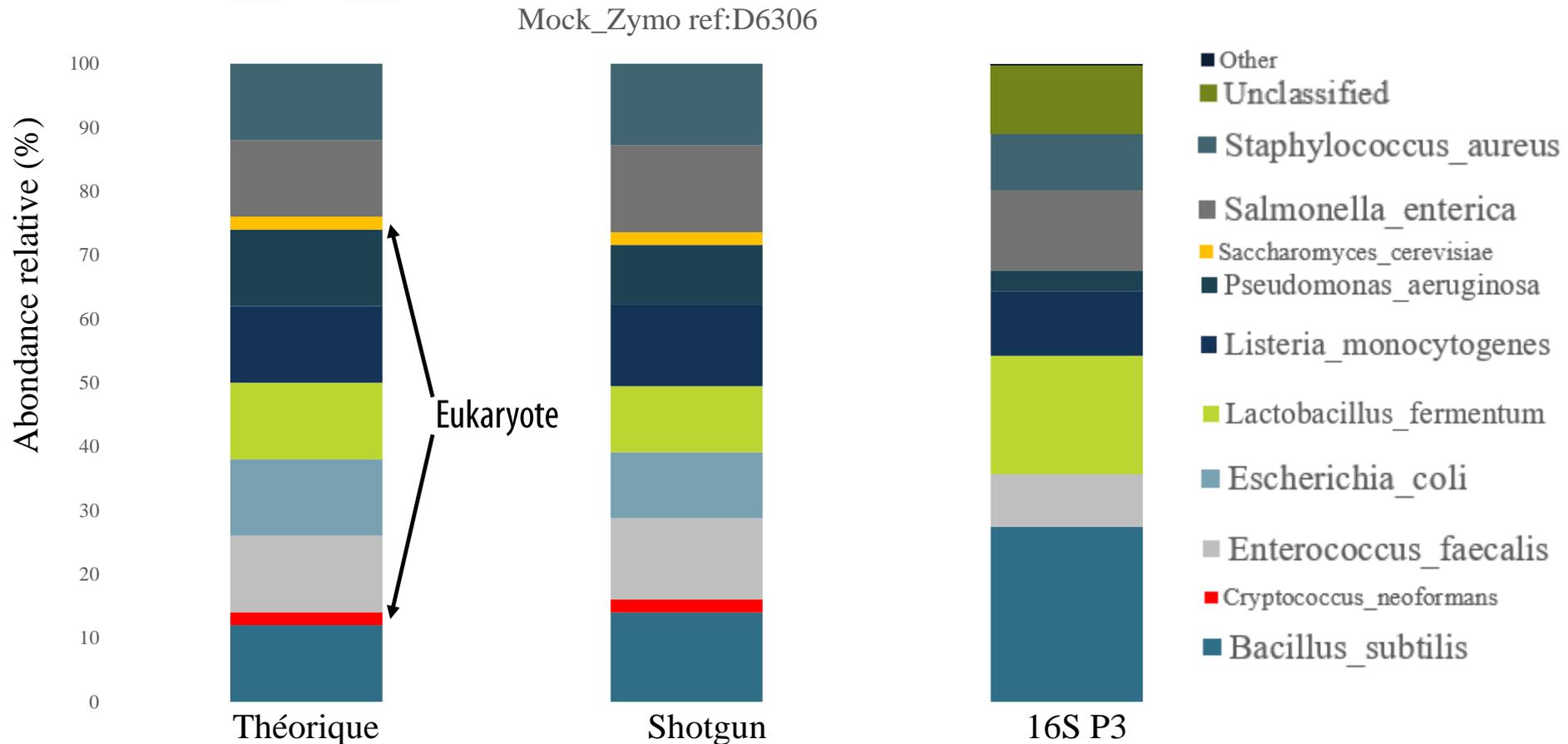
Ruminoclostridium, et Subdoligranulum sont absent



Profils des communautés Mock 16S vs. Shotgun



Other= 44 genres

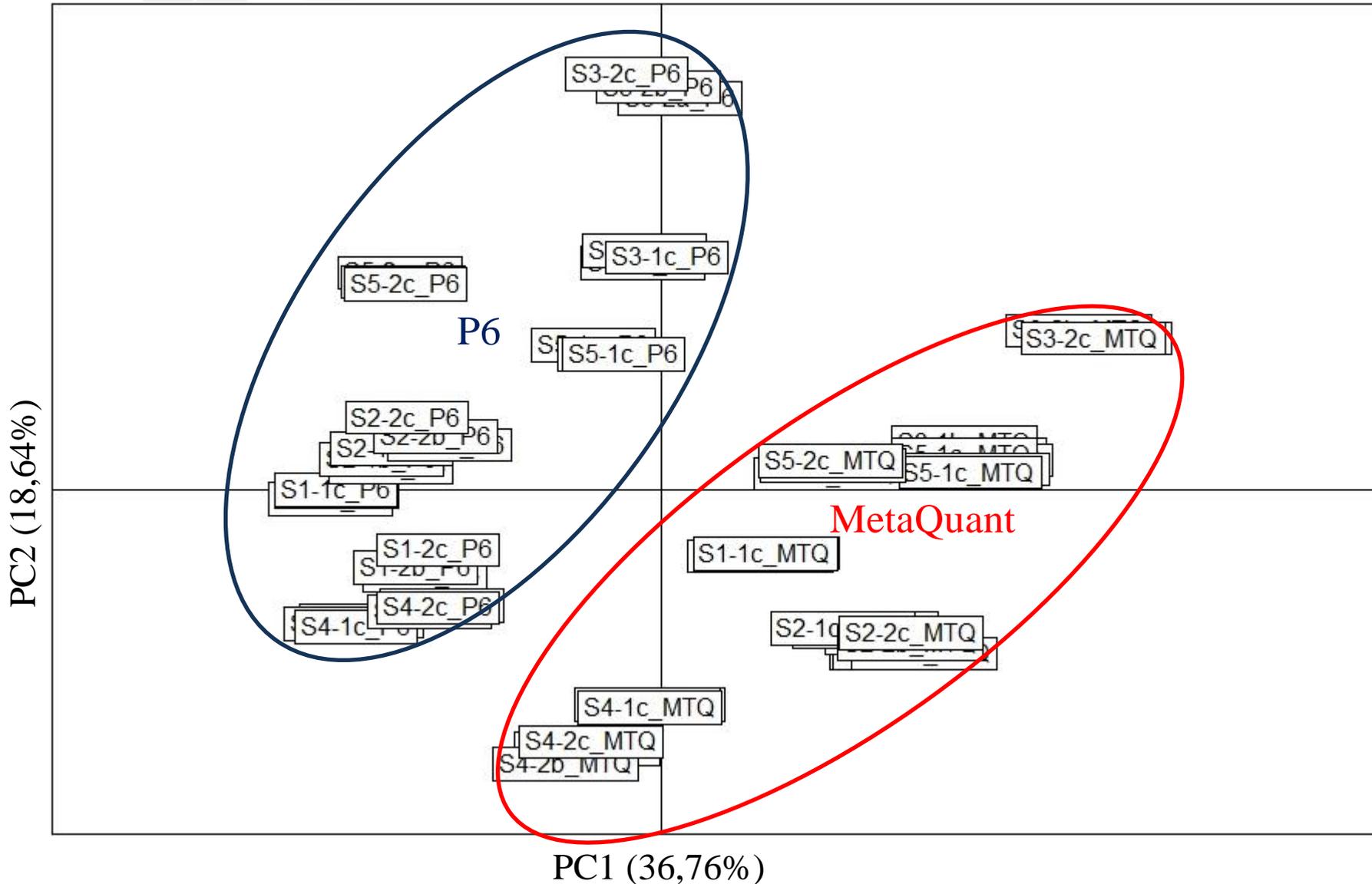


- Pour P6 en 16S: *Escherichia coli* est absent, remplacé par la catégorie unclassified.
- *Saccharomyces* et *Cryptococcus* pas amplifié en 16S car *Eukaryote* (ne possède pas d'ADNr 16S)
- *Bacillus subtilis* surestimé



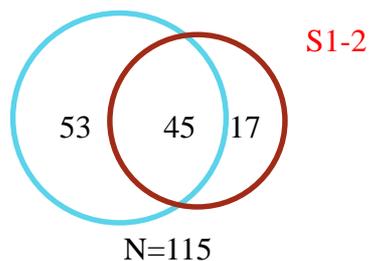
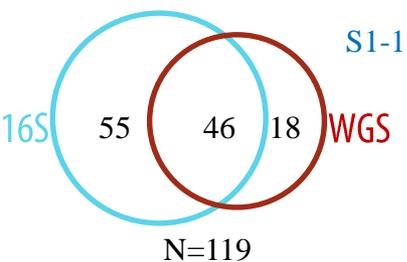
Shotgun vs. P6

Mesure de similarité du profile bactérien au genre 16S vs. Shotgun du partenaire 6 (P6) - ACoP



Profile bactérien au genre 16S vs. Shotgun: sujet S1

Candidatus Cibiobacter absent
Akkermansia surévalué en 16S

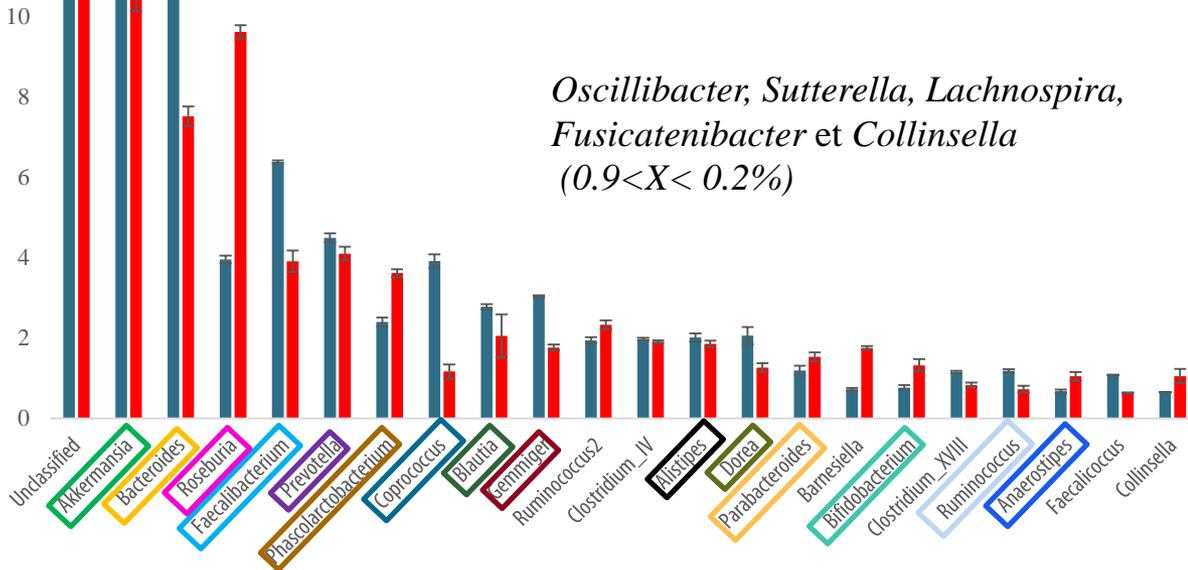


Barnesiella et *Faecalicoccus* absent en shotgun

P6

■ S1-1 ■ S1-2

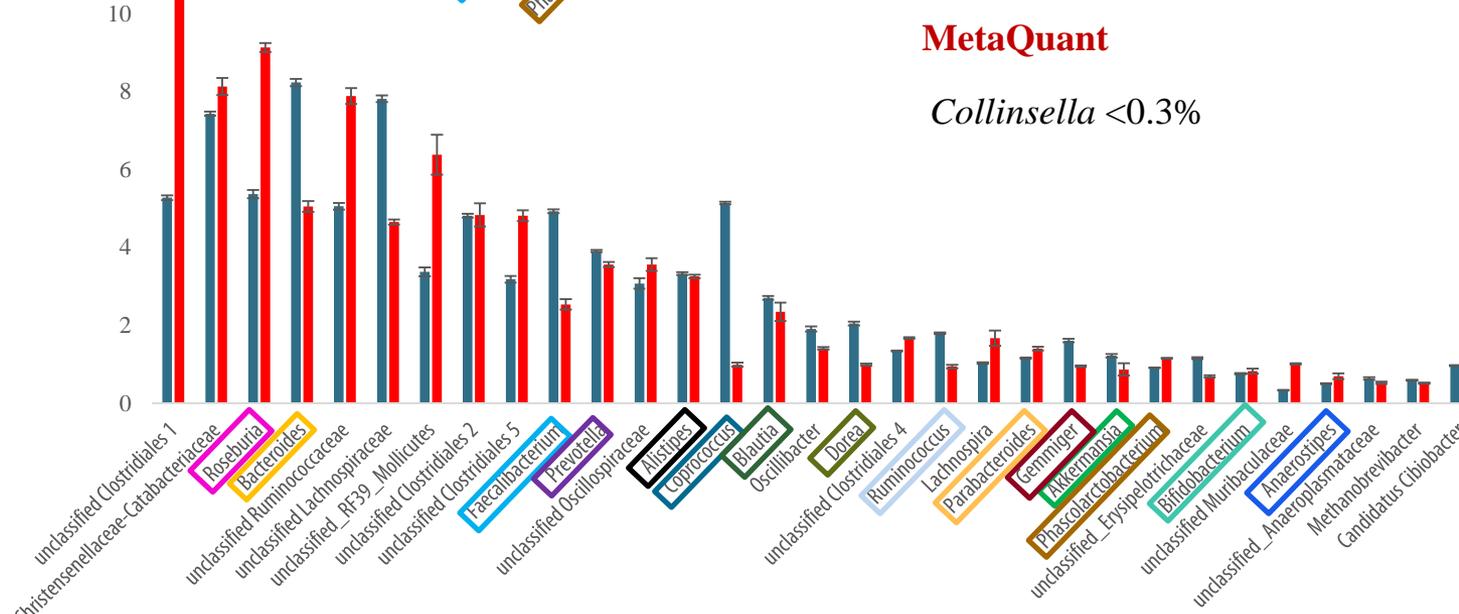
Abundance relative (%)



Oscillibacter, *Sutterella*, *Lachnospira*,
Fusicatenibacter et *Collinsella*
($0.9 < X < 0.2\%$)

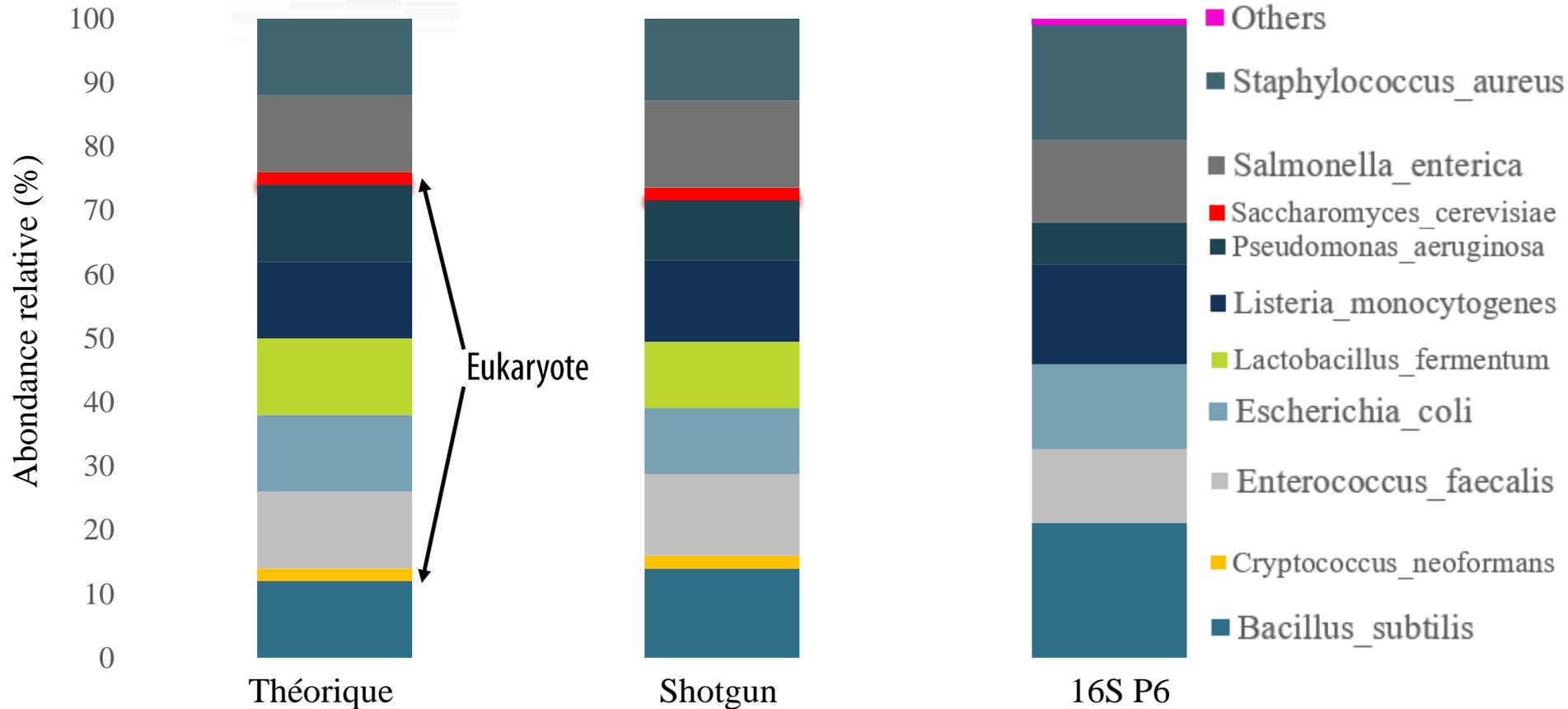
MetaQuant

Collinsella < 0.3%



Profils des communautés Mock ADN 16S vs. Shotgun

Mock_Zymo ref: D6306

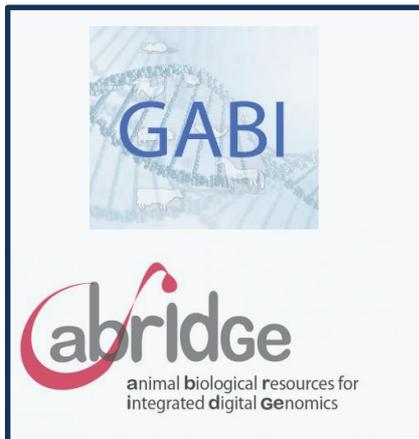


- *Saccharomyces* et *Cryptococcus* pas amplifié en 16S car Eukaryote (ne possède pas d'ADNr 16S)
- *Lactobacillus* quasiment pas détecté par P6 (<0.01%): amplification ? pipeline ?



- Effet méthode > effet sujet → histoire biologique différente
- P1 et P2: même séquençage, pipeline bioinformatique différent: P1 outlier et P2 assez proche → importance du pipeline bioinformatique choisi!!!
- Même pour partenaires « proches » (P3, P6), effet méthode > effet sujet.
→ Akkermansia chez S1: perdu avec P3, surévalué avec P6 (> 10% vs < 2% en shotgun)
- Le shotgun reste la méthode à privilégier (voir résultats Mock)
- Si pas possible, travailler sur le pipeline bioinformatique pour se rapprocher du shotgun
- Shallow sequencing: un bon compromis ?

Remerciements



Joel Doré initiateur du projet



ANNEXES



Description des pipelines d'analyse

Séquençage et analyse du gène 16S rRNA par le partenaire 16S 1 (P1)

Analyse en métabarcoding conduite sur la régions V3-V4 amplifié par PCR et séquencé sur le séquenceur MiSeq (Illumina)

Pipeline basé sur le logiciel mothur (Schloss *et al.*, 2009)

- logiciel : mothur **1.36.1**
- Base de donnée alignement: Greengenes (bactéries/archées)
- Elimination des chimères + **homopolymères (>8pb)** + Singleton (1 seq/OTU)
- Clustering des OTUs: les séquences présentant 100% d'homologie, entre elles, ont été regroupées en séquences uniques, puis en OTUs regroupant les séquences ayant au moins 97% d'homologie entre elles.

Analyse des gènes 16S rRNA par le partenaire 16S 2 (P2)

Analyse en métabarcoding conduite sur la régions V3-V4 amplifiée par PCR et séquencée sur le séquenceur MiSeq (Illumina) par P1

Pipeline développé basé sur un logiciel « **maison** »

- Recherche et élimination primers: cutadapt
- Correction erreurs des séquences: SPAdes
- Merging des reads: PEAR
- Clustering des OTUs: Vsearch (unoise)
- Elimination Chimère: uchime 3
- Base de donnée taxonomie : RDPTools suit

Séquençage et analyse du gène 16S rRNA par le partenaire 16S 3 (P3)

Analyse en métabarcoding conduite sur la régions V3-V4 amplifiée par PCR et séquencée sur le séquenceur MiSeq (Illumina) 2 x 300 pb (29 ADN) et 2 x 250 (4 ADN; <40k séquences). Nombre de séquences: 86k +/- 20k

Pipeline développé basé sur un logiciel « **maison** »

- Qualité des séquences: FastQC 11.7 et MultiQC 1.5
- Contigage des séquences: Vsearch v2.6.2
- Recherche et élimination primers: cutadapt 1.14
- Clustering des OTUs: logiciel Swarm v2.2.2 (autorisation d'une différence pour agréger les séquences, puis les graines des clusters précédent sont agrégées en OTUs jusqu'à 3 différences)
- Elimination des chimères: Vsearch v2.6.2
- Délétion des cluster trop peu abondant: outil Filter (élimine les OTUs dont l'abondance cumulé est inférieur à $5 \cdot 10^{-5}$ de l'abondance totale de tous les OTUs) + OTUs non partagé par au moins 3 des tiplicats
- Base de donnée taxonomie : **SILVA 132**, filtré sur le score de pintail de 50

Séquençage et analyse du gène 16S rRNA par le partenaire 16S 4 (P4)

Analyse en métabarcoding conduite sur la régions V3-V4 amplifiée par PCR et séquencée sur le séquenceur MiSeq (Illumina) 2 x 250.

Pipeline développé basé sur le logiciel **QIIME v1.9.1** (Caporaso *et al.*, 2010)

- Qualité des séquences: FastQC
- Démultiplexage: CASAVA v1.0
- Merging: Flash (>30 b de chevauchement, à 97% identité)
- Elimination des chimères: Usearch v6.1
- Clustering des OTUs: Uclust v1.2.22q (97% similarité)
- Base de donnée alignement et taxonomie: Greengenes v13.8

Séquençage et analyse du gène 16S rRNA par le partenaire 16S 5 (P5)

Analyse en métabarcoding conduite sur la régions V3-V4 amplifiée par PCR et séquencée sur le séquenceur MiSeq (Illumina, 2x300 pb)

Pipeline basé sur le logiciel QIIME (Caporaso *et al.*, 2010)

- logiciel : **QIIME 1.9.1**
- Base de donnée alignement: SILVA v.123
- Clustering des OTUs: open-reference avec la méthode usearch61 et seuil de clustering de 0,97
- Singleton: élimination des OTUs d'abondance <0.005%

Séquençage et analyse du gène 16S rRNA par partenaire 16S 6 (P6)

Analyse en métabarcoding conduite sur la régions V3-V4 amplifiée par PCR et séquencée sur le séquenceur MiSeq (Illumina)

Pipeline développé basé sur le logiciel mothur (Schloss *et al.*, 2009)

- logiciel : mothur 1.40.0
- Base de donnée alignement: Greengenes v13.5.99
- Base de donnée taxonomie : RDP v16
- Elimination des chimères: Vsearch
- Clustering des OTUs: Opticlust
- Seuil de clustering OTUs: 0,03
- Délétion des sigletons (1 séquence/OTU)