



**HAL**  
open science

# **rpoB, a promising marker for analyzing the diversity of bacterial communities by amplicon sequencing**

Jean-Claude Ogier, Sylvie Pages, Matthieu Barret, Sophie Gaudriault

## ► To cite this version:

Jean-Claude Ogier, Sylvie Pages, Matthieu Barret, Sophie Gaudriault. rpoB, a promising marker for analyzing the diversity of bacterial communities by amplicon sequencing. 9. Colloque de l'Association Francophone d'Ecologie Microbienne, Nov 2019, Bussang, France. hal-02463373

**HAL Id: hal-02463373**

**<https://hal.science/hal-02463373>**

Submitted on 5 Jun 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# RPOB, A PROMISING MARKER FOR ANALYZING THE DIVERSITY OF BACTERIAL COMMUNITIES BY AMPLICON SEQUENCING (1)

J.C. Ogier<sup>1</sup>, S. Pagès<sup>1</sup>, M. Barret<sup>2</sup> and S. Gaudriault<sup>1</sup>

<sup>1</sup>INRA, Université Montpellier, UMR1333 DGIMI, CC054, 34095 Montpellier Cedex 05, France  
<sup>2</sup>INRA, Agrocampus Ouest, Université d'Angers, UMR-IRHS, 49071 Beaucouzé, France

**OBJECTIVES:** Microbiome composition is frequently studied by the amplification and high-throughput sequencing of specific molecular markers (metabarcoding). The 16S rRNA gene is classically used to estimate bacterial diversity, but its low discriminating power for certain bacterial genera and its variable copy number in prokaryotic genomes constitute important limitations (2). In this study, we assessed the potential benefit of a portion of the *rpoB* gene as an alternative genetic marker. We first analyzed the sequence data generated by metabarcoding with *rpoB* and 16S (V3V4 region) markers on an artificial bacterial DNA complex corresponding to 19 different phylogenetic taxa. We then compared the performance of the *rpoB* and V3V4 markers in an animal ecosystem model, the infective juveniles (IJs) of the entomopathogenic nematode *Steinernema glaseri* carrying the symbiotic bacteria *Xenorhabdus poinarii*.

## 1. Experimental design for metabarcoding analysis

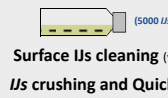
### Preparation of 5 mock communities

Mix of genomic DNAs extracted from 19 bacterial species  
5 mocks differing in the proportions of the taxa



### DNA extraction from Nematodes

Surface IJs cleaning (water)  
IJs crushing and Quick extract Kit



### Amplification of two genetic markers

16S v3v4 450 pb  
*rpoB* 430 pb



ILLUMINA MISEQ SEQUENCING

### Data analysis

- Raw read processing
- Quality check
- Chimera removal
- Read abundance filter (>0.005% of the whole data set)

• Taxonomic assignment (RDP Blast, 97% similarity)

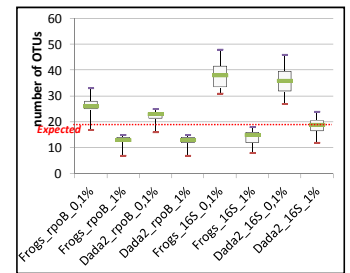
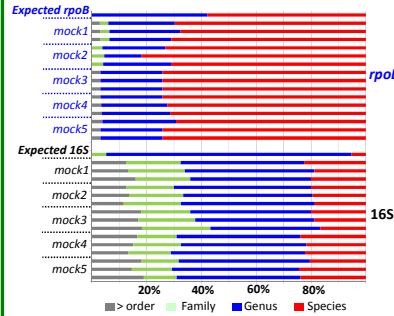
SILVA database

*rpoB* database\* (45 000 seq)

• Read abundance filter for individual sample (>0.1% and >1%)

## 2. Analysis of Mock communities

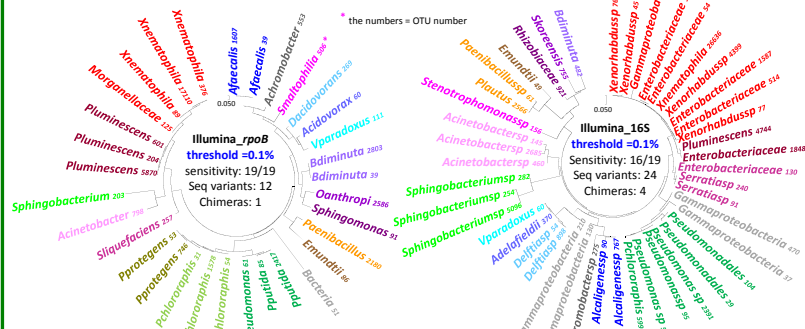
### Impact on OTU taxonomic affiliations and observed OTU richness



• Taxonomic assignment level is better with *rpoB* marker

- OTU richness overestimated, especially for 16S marker
- The optimal read abundance threshold to individual sample is 0.1%

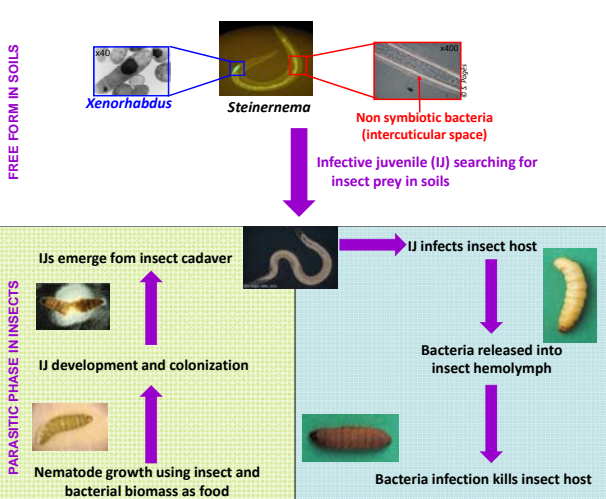
### Phylogenetic tree of the overall OTUs



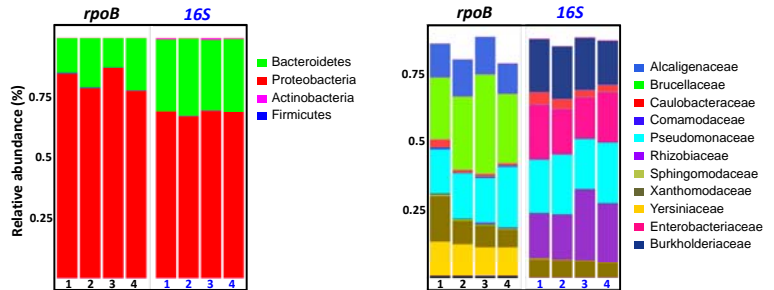
- Sensitivity is better with *rpoB* marker, i.e. the 19 species which compose the mock are detected
- Many additional OTUs are observed with V3V4 marker, corresponding to chimeras and variant sequences

## 3. Description of the nematodes IJ microbiota (*Steinernema glaseri*, 4 replicates)

### Biological cycle of entomopathogenic nematodes

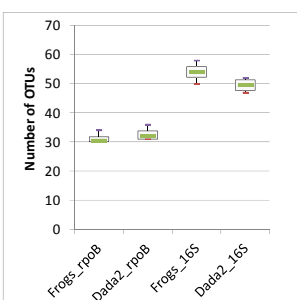


### OTU composition at the Phylum and Family level



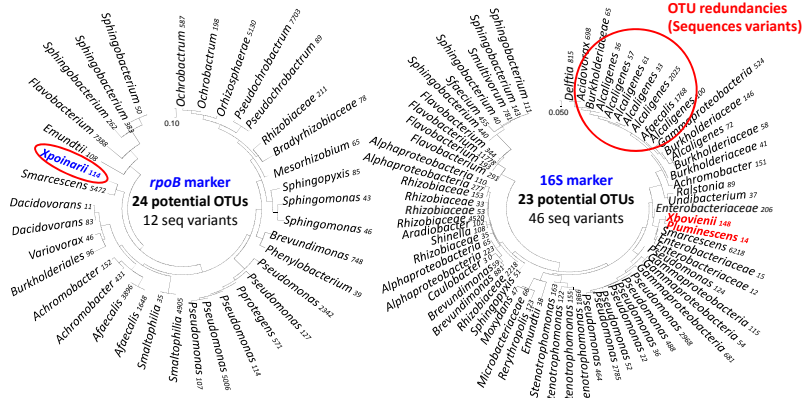
- Similar bacterial compositions were obtained with both markers at the Phylum level (mainly Proteobacteria), but differences appear at the Family level

### OTU richness



- The pipeline tools (FROGS vs DADA2) does not influence the OTU richness
- The OTU richness is dependent on the used marker (vary from 30 to 55 OTUs)
- Probable over estimation of OTU richness with 16S marker

### Phylogenetic tree of the overall OTUs



- Only *rpoB* marker detects the bacterial symbiont *X. poinarii*
- Numerous sequence variants with the 16S marker (overestimation of the OTU diversity)
- these sequence variants generate OTU identification errors (e.g. identification of *P. luminescens* and *X. bovienii* instead of *X. poinarii*)

**CONCLUSIONS:** The use of *rpoB* gene for metabarcoding analysis is a promising approach to accurately explore the diversity of bacterial communities because of its best discriminating power and the reduction of bias compared to the 16S marker. We created a database which includes 45,000 *rpoB* sequences covering the large diversity of available prokaryotic genomes. This database is available from the FROGS website (<http://frogs.toulouse.inra.fr/>).

**References:** (1) Ogier et al., 2019. *rpoB*, a promising marker for analyzing the diversity of bacterial communities by amplicon sequencing. BMC Microbiol. 2019 Jul 29;19(1):171. doi: 10.1186/s12866-019-1546-z.  
(2) Roux et al., 2011. Comparison of 16S rRNA and protein-coding genes as molecular markers for assessing microbial diversity (Bacteria and Archaea) in ecosystems. FEMS Microbiol Ecology. 78. 617-628