



# Vivaldi Project

## Data management plan

Marine amoebae and their associated bacteria  
diversity

02/07/2018

## DATA MANAGEMENT PLAN

### Template sheet for each dataset

<b>Partner name</b>	CNRS
<b>Data category</b>	Microbiome (environment)
<b>Concerned WP</b>	WP4 Choisissez un élément. Choisissez un élément.
<b>Name of the VIVALDI referent(s)</b>	Microbiome: Alberto Pallavicini
<b>Reference of the dataset</b> Please refer to the DMP table to find the appropriate reference. Ex: Genome-Patho/SubTaskN*/Pathogen/PartnerN*	Microbiome-bivalve/4.4/Crassostrea_gigas/2
<b>Description of the data</b>	18s and 16s barcoding of marine amoebae diversity and diversity of their associated bacteria
<b>Type</b>	Sequences
<b>Period and frequency of data collection</b>	Samples were collected monthly over a year of time from 2017 to 2018
<b>Geographical site of data collection (if applicable)</b>	Three Mediterranean geographical sites were chosen: One oyster farming table in the Thau lagoon, outside of Sète harbor, and in Banyuls sur mer near the protected area.
<b>Description of the material from which the dataset is generated</b> Information will be obtained from individuals, which can come from natural/hatchery population and/or from family produced in hatchery. Animals can be infected (naturally or experimentally). DNA extraction can be done from the whole animal, tissue.	Describe the material from which the data is generated V4 region of the 18s was sequenced for amoebae, and V3-V4 region from the 16s was sequenced for bacteria
<b>Protocols</b> Example: 16S ribosomal RNA gene sequencing by NGS <b>Please refer to the DMP table* for more examples</b>	Standard protocol for 18s and 16S ribosomal RNA gene sequencing by NGS (MiSeq illumina)
<b>Nature of the collected/generated data</b> Example: Raw dataset in .blc/.fastqc/.fasta formats for genomic	fastqc and fasta



<p>information, and processed data set will be .vcf/.bed formats. <b>Please refer to the DMP table* for more examples</b></p>	
<p><b>Coverage (if applicable)</b> Example: random genomic regions covered at 50 X <b>Please refer to the DMP table* for more examples</b></p>	<p>Specify the coverage</p>
<p><b>What are the prerequisites allowing to use the data as such?</b> Example: Any person able to use .fastqc file and .fasta file <b>Please refer to the DMP table* for more examples</b></p>	<p>Any person able to use .fastqc file and .fasta file</p>
<p><b>Sharing of main data</b></p>	<p>Saved and shared after publication Please specify</p>
<p><b>Archiving and preservation</b> Example: data will be stored on a hard drive + online back up and then will be released on public database (Sinoe, Dryad) after publication. <b>Please refer to the DMP table* for more examples</b></p>	<p>data will be stored by each partner and then will be released on public database (NCBI) after publication</p>
<p><b>List, description and storage of associated data (metadata)</b> Examples: environmental data, mortality monitoring, genotyping...</p>	<p>The metadata file contains all available information related to the environmental data for this study: sample type, season, GPS data. All these metadata will be available in both the SRA bioproject (to be released after publication) and the associated publication.</p>
<p><b>Sharing of metadata (if relevant)</b></p>	<p>Saved and shared after publication All data will be available with the publication</p>

\*To access the [DMP table](#), please login on the VIVALDI online platform

Once completed, this sheet has to:

1. Be sent to the referent(s) identified above for a final check
2. Be uploaded on the [VIVALDI online platform](#)

