

Vivaldi Project

Data management plan

Marine amoebae and their associated bacteria diversity

02/07/2018



DATA MANAGEMENT PLAN

Template sheet for each dataset

Data category Microbiome (environment) WP4 Choisissez un élément. Choisissez un élément. Choisissez un élément. Name of the VIVALDI referent(s) Microbiome: Alberto Pallavicini Reference of the dataset Please refer to the DMP toble to find the appropriate reference. Ex-Genome Pathon (Jubi Taskit / Pathogeny Partnent)* 18s and 16s barcoding of marine amoebae diversity and diversity of their associated bacteria Description of the data Type Sequences Period and frequency of data collection Geographical site of data collection (if applicable) 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. Description of the material from which the dataset is generated information will be obtained from individuals, which can come from i	Partner name	CNRS
Concerned WP Choisissez un élément. Choisissez un élément. Choisissez un élément. Name of the VIVALDI referent(s) Microbiome: Alberto Pallavicini Reference of the dataset Please refer to the DMP toble to find the appropriate reference. En: Genome Pathon/SubToskNr/Prothogen/Protracent* 18s and 16s barcoding of marine amoebae diversity and diversity of their associated bacteria 18s and 16s barcoding of marine amoebae diversity and diversity of their associated bacteria 18s and 16s barcoding of marine amoebae diversity and diversity of their associated bacteria Type Sequences Samples were collected monthly over a year of time from 2017 to 2018 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. Description of the material from which the dataset is generated Information will be obtained from individuols, 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. Protocols Example: 16s ribosomal RNA gene sequencing by NGS Flease refer to the DMP table* for more examples fastqc and fasta fastqc and fasta fastqc and fasta fastqc and fasta		
Choisissez un élément. Reference of the dataset Please refer to the DMP table to find the appropriate reference. Ex-Genome-bathol/Jubiosakhi/Pathogan/Partinent* 18s and 16s barcoding of marine amoebae diversity and diversity of their associated bacteria 18s and 16s barcoding of marine amoebae diversity and diversity of their associated bacteria 18s and 16s barcoding of marine amoebae diversity and diversity of their associated bacteria 18s and 16s barcoding of marine amoebae diversity and diversity of their associated bacteria 18s and 16s barcoding of marine amoebae diversity and diversity of their associated bacteria 18s and 16s barcoding of marine amoebae diversity and diversity of their associated bacteria 18s and 16s barcoding of marine amoebae diversity and diversity of their associated bacteria 18s and 16s barcoding of marine amoebae diversity and diversity of their associated bacteria 18s and 16s barcoding of marine amoebae diversity and diversity of their associated bacteria 18s and 16s barcoding of marine amoebae diversity and diversity of their associated bacteria 18s and 16s barcoding of marine amoebae diversity and diversity of their associated bacteria 18s and 16s barcoding of the 18s were collected area. 18s and 16s barcoding of set en amoebae diversity and diversity of their associated bacteria 18s and 16s barcoding of the 18s was sequenced for amoebae, and volta ferming table in the Thau lagoon, outside of Sète harbor, and in Banyuls sur mer near the protected area. 18s and 16s barcoding of the 18s was sequenced for amoebae, and volta ferming table in the Thau lagoon, outside of Sète harbor, and in Banyuls sur mer near the protected area. 18s and 16s area ferming table in the Thau lagoon, outside of Sète harbor, and in Banyuls sur mer near the protected area. 18s and 16s area ferming table in the Thau lagoon, outside of Sète harbor, and in Ba	Data category	Microbiome (environment)
Reference of the dataset Please refer to the DMP table to find the appropriate reference. Ext. Genome. Pathof/Sub/TaskW/Pathogen/Partnern* 18s and 16s barcoding of marine amoebae diversity and diversity of their associated bacteria 18s and 16s barcoding of marine amoebae diversity and diversity of their associated bacteria Type Sequences Period and frequency of data collection Geographical site of data collection (if applicable) 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. Description of the material from which the dataset is generated 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. Protocols Example: 16s ribosomal RNA gene sequencing by NGS Microbiome: Alberto Pallavicini Microbiome: Alberto Pallavic/A-4/Crassostrea_gigas/2 Microbiome-bivalve/4.4/Crassostrea_gigas/2 18s and 16s barcoding of marine amoebae diversity and diversity of their associated bacteria Type Sequences Samples were collected monthly over a year of time from 2017 to 2018 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. 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 Standard protocol for 18s and 16s ribosomal RNA gene sequencing by NGS (MiSeq illumina) Standard protocol for 18s and 16s ribosomal RNA gene sequencing by NGS (MiSeq illumina) From the 16s was sequenced for bacteria Standard protocol for 18s and 16s ribosomal RNA gene sequencing by NGS (MiSeq illumina) From the 16s was sequenced for bacteria	Concerned WP	
Reference of the dataset Please refer to the DMP table to find the appropriate reference. Br. Genome Pathol-Substank / Plathologen/Partnersh* 18s and 16s barcoding of marine amoebae diversity and diversity of their associated bacteria 18s and 16s barcoding of marine amoebae diversity and diversity of their associated bacteria Type Sequences Period and frequency of data collection Geographical site of data collection (if applicable) Three Mediterranean geographical sites were chosen: One oyster farming table in the Thau lagoon, outside of Sete harbor, and in Banyuls sur mer near the protected area. Description of the material from which the dataset is generated 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. Protocols Example: 16s ribosomal RNA gene sequencing by NGS (MiSeq illumina) Standard protocol for 18s and 16s ribosomal RNA gene sequencing by NGS (MiSeq illumina) Standard protocol for 18s and 16s ribosomal RNA gene sequencing by NGS (MiSeq illumina) Standard protocol for 18s and 16s ribosomal RNA gene sequencing by NGS (MiSeq illumina) Formation with the collected/generated data Example: Row dataset in		Choisissez un élément.
Please refer to the DMP table to find the appropriate reference. Ex. Genome Patho/SubTisskii/Puthagen/Partmeth* 18s and 16s barcoding of marine amoebae diversity and diversity of their associated bacteria Type Sequences Period and frequency of data collection Geographical site of data collection (if applicable) Description of the material from which the dataset is generated 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. Protocols Example: 16s ribosomal RNA gene sequencing by NGS Please refer to the DMP table* for more examples fastqc and fasta	Name of the VIVALDI referent(s)	Microbiome: Alberto Pallavicini
18s and 16s barcoding of marine amoebae diversity and diversity of their associated bacteria Type Sequences Period and frequency of data collection Geographical site of data collection (if applicable) 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. Description of the material from which the dataset is generated Information will be obtained from individuals, which can come from individuals, which can come from individuals, which can come from family produced in hatchery. Animals can be infected (naturally or experimentally). DNA extraction can be done from the whole animal, tissue. Protocols Example: 16S ribosomal RNA gene sequencing by NGS Please refer to the DMP table* for more examples Nature of the collected/generated data Example: Row dataset in	Please refer to the DMP table to find the appropriate reference.	Microbiome-bivalve/4.4/Crassostrea_gigas/2
their associated bacteria Sequences Samples were collected monthly over a year of time from 2017 to 2018 the mother possible from eyster farming table in the Thau lagoon, outside of Sète harbor, and in Banyuls sur mer near the protected area. Describe the material from which the data is generated V4 region of the 18s was sequenced for amoebae, and V3-V4 region from the undersological from the 16s was sequenced for bacteria their associated bacteria Samples were collected monthly over a year of time from 2017 to 2018 experimently on eyster farming table in the Thau lagoon, outside of Sète harbor, and in Banyuls sur mer near the protected area. 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 Standard protocol for 18s and 16s ribosomal RNA gene sequencing by NGS (MiSeq illumina) Standard protocol for 18s and 16s ribosomal RNA gene sequencing by NGS (MiSeq illumina) Standard protocol for 18s and 16s ribosomal RNA gene sequencing by NGS (MiSeq illumina) Standard protocol for 18s and 16s ribosomal RNA gene sequencing by NGS (MiSeq illumina) From the 16s ribosomal RNA gene sequencing by NGS (MiSeq illumina)	Ex. Genome ratio/subrasity /ratiogen/ration	
Period and frequency of data collection Samples were collected monthly over a year of time from 2017 to 2018 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. Description of the material from which the dataset is generated 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. Protocols Example: 16S ribosomal RNA gene sequencing by NGS Please refer to the DMP table* for more examples Nature of the collected/generated data Example: Raw dataset in	Description of the data	
Period and frequency of data collection Samples were collected monthly over a year of time from 2017 to 2018 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. Description of the material from which the dataset is generated 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. Protocols Example: 16S ribosomal RNA gene sequencing by NGS Please refer to the DMP table* for more examples Standard protocol for 18s and 16S ribosomal RNA gene sequencing by NGS Please refer to the DMP table* for more examples fastqc and fasta fastqc and fasta	Description of the data	
Geographical site of data collection (if applicable) 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. Description of the material from which the data is generated 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. Protocols Example: 16S ribosomal RNA gene sequencing by NGS Please refer to the DMP table* for more examples Nature of the collected/generated data Example: Raw dataset in	Туре	Sequences
Geographical site of data collection (if applicable) 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. Description of the material from which the data is generated 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. Protocols Example: 16S ribosomal RNA gene sequencing by NGS Please refer to the DMP table* for more examples Nature of the collected/generated data Example: Raw dataset in	Period and frequency of data	Samples were collected monthly over a year of time from 2017 to
farming table in the Thau lagoon, outside of Sète harbor, and in Banyuls sur mer near the protected area. Description of the material from which the dataset is generated 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. Protocols Example: 165 ribosomal RNA gene sequencing by NGS Please refer to the DMP table* for more examples Nature of the collected/generated data Example: Raw dataset in		
which the dataset is generated 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. Protocols Example: 16S ribosomal RNA gene sequencing by NGS Please refer to the DMP table* for more examples Nature of the collected/generated data Example: Raw dataset in		farming table in the Thau lagoon, outside of Sète harbor, and in
which the dataset is generated 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. Protocols Example: 16S ribosomal RNA gene sequencing by NGS Please refer to the DMP table* for more examples Nature of the collected/generated data Example: Raw dataset in	5 (1)	
Example: 16S ribosomal RNA gene sequencing by NGS Please refer to the DMP table* for more examples Nature of the collected/generated data Example: Raw dataset in	which the dataset is generated 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	V4 region of the 18s was sequenced for amoebae, and V3-V4 region
data Example: Raw dataset in	Example: 16S ribosomal RNA gene sequencing by NGS Please refer to the DMP table* for	, , ,
data Example: Raw dataset in	Noture of the collected / conserved	factor and facto
.blc/.fastqc/.fasta formats for genomic	data Example: Raw dataset in	jastąc ana jasta



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

^{*}To access the <u>DMP table</u>, 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