Vivaldi Project

Data management plan

Maldi-tof spectra
Key words: bacteria, vibrio, ribosomal proteins, MSP

13/09/2018

DATA MANAGEMENT PLAN

Template sheet for each dataset

Partner name	*IFREMER*

Data category	*Maldi-Tof spectra*

Concerned WP	*WP1* *Choisissez un élément.* *Choisissez un élément.*

Name of the VIVALDI referent(s)	*Maldi-tof: Céline Garcia (Ifremer)*

Reference of the dataset *Please refer to the DMP table to find the appropriate reference.* Ex: Genome-Patho/SubTaskN°/Pathogen/PartnerN°	Maldi-TOF/SubTaskN°1.21.1/Vibrio/PartnerN°1

Description of the data	Maldi-TOF profiles from different *Vibrio* species and strains

Type	*Spectra*

Period and frequency of data collection	The bacteria collection of the French National Reference Laboratory (NRL) has been used. Some bacteria were purchased from official collections, others were kindly provided by C. Paillard (CNRS) and F. Leroux (Ifremer Roscoff) but most of them came from different bivalve molluscs and were mainly isolated during French mollusc mortality events.

Geographical site of data collection (if applicable)	*N/A*

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.*	Bacteria collected in molluscs during mortality events
Protocols *Example: 16S ribosomal RNA gene sequencing by NGS* **Please refer to the DMP table* for more examples**	*Bruker protocole for MSP creation*

Nature of the collected/generated data	*Spectra*

Example: Raw dataset in .blc/.fastqc/.fasta formats for genomic 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**	*mainly ribosomal proteins (between 2000 and 20000 dalton)*
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**	*All laboratories with a Bruker MALDI-TOF*
Sharing of main data	*Saved and shared after publication* All main spectra of the Vibrio database will be available as btmsp-file.
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**	*Please explain how the data will be archived and preserved, including the type of storage (hard drive, cloud, server…) and the foreseen back-up.* The data are stored on a hard drive and on a local computer server. If possible, this database will be public via SINOE.
List, description and storage of associated data (metadata) *Examples: environmental data, mortality monitoring, genotyping…*	16S, gyrB, ldh and pyrH sequences of the different bacteria used for the creation of MSP and isolated by the French LNR during French mortality event. These data are stored on a hard drive and on a local computer server. They will put on GenBank after publication.
Sharing of metadata (if relevant)	*Choisissez un élément.* *Please specify*

*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**