



Vivaldi Project

Data management plan

Genome (pathogen)

Key words: *Crassostrea gigas*, *OsHV-1*, *Illumina TruSeq*,

07/08/2018

DATA MANAGEMENT PLAN

Template sheet for each dataset

Partner name	IFREMER
Data category	
Concerned WP	WP1 Choisissez un élément. Choisissez un élément. <i>Genome (pathogen)</i>
Name of the VIVALDI referent(s)	Genome (OsHV-1): Benjamin
Reference of the dataset <i>Please refer to the DMP table to find the appropriate reference.</i> <small>Ex: Genome-Patho/SubTaskN*/Pathogen/PartnerN*</small>	Genome-pathogene/WP1.2.2/OsHV1/Ifremer
Description of the data	Whole genome sequencing of infected oyster. Grands-parents, parents, and 5 offsprings per F2 families produced at Ifremer La Tremblade (France) (80 individuals).
Type	Sequences
Period and frequency of data collection	N/A
Geographical site of data collection (if applicable)	N/A
Description of the material from which the dataset is generated <i>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.</i>	Grands-parents were from highly resistant versus highly sensible families. We add a pedigree files to describe individuals relatedness
Protocols <i>Example: 16S ribosomal RNA gene sequencing by NGS</i> Please refer to the DMP table* for more examples	Nucleic acid extraction was performed using the QIAamp DNA Mini Kit (Qiagen) according to the manufacturer's handbook. All the samples have been prepared using the Illumina TruSeq PCR free protocol on a HiSeq 4000 paired-end 150x2 with 350 insert size.
Nature of the collected/generated data	.fastq.gz We will add finale files as soon they available.



<p><i>Example: Raw dataset in .blc/.fastqc/.fasta formats for genomic information, and processed data set will be .vcf/.bed formats.</i> Please refer to the DMP table* for more examples</p>	
<p>Coverage (if applicable) <i>Example: random genomic regions covered at 50 X</i> Please refer to the DMP table* for more examples</p>	<p><i>Grands-parents and parents have been sequenced at 15 X and offsprings have sequenced at 15 X.</i></p>
<p>What are the prerequisites allowing to use the data as such? <i>Example: Any person able to use .fastqc file and .fasta file</i> Please refer to the DMP table* for more examples</p>	<p><i>After publication anybody that could process fastq.gz files.</i></p>
<p>Sharing of main data</p>	<p><i>Saved and shared after publication</i> <i>Please specify</i></p>
<p>Archiving and preservation <i>Example: data will be stored on a hard drive + online back up and then will be released on public database (Sinoe, Dryad) after publication.</i> Please refer to the DMP table* for more examples</p>	<p><i>The dataset will be stored on datarmor in the dataref disk: /dataref/ref2/intranet/ifremer/sg2m/vivaldi/data/dna-sequence-raw/</i></p> <p><i>The dataset will be public (ftp server) after publication. It will be also referenced in Sinoe/SISMER/SEXTANT catalogue. https://sextant.ifremer.fr/fr/geoservices/catalogue</i></p>
<p>List, description and storage of associated data (metadata) <i>Examples: environmental data, mortality monitoring, genotyping...</i></p>	<p><i>Individuals relatedness is described in the associated pedigree file and phenotype file</i> <i>DISKII_G</i> <i>G:\IFREMER\02-PROJETS\H2020\Tackling_disease\08_QTL_detection_VIVALDI\03_sampling_design\01_input</i></p>
<p>Sharing of metadata (if relevant)</p>	<p><i>Saved and shared after publication</i> <i>Please specify</i></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](#)**

