



# Vivaldi Project

## Data management plan

**Genetic markers**

**Key words:** *SNP markers, genotyping tables, assignment rates, Crassostrea gigas*

## DATA MANAGEMENT PLAN

### Template sheet for each dataset

<b>Partner name</b>	LABOGENA
<b>Data category</b>	
<b>Concerned WP</b>	WP3 Choisissez un élément. Choisissez un élément.
<b>Genetic markers</b>	
<b>Name of the VIVALDI referent(s)</b>	Genetic markers: Mathias Wegner & JB Lamy
<b>Reference of the dataset</b> Please refer to the DMP table to find the appropriate reference. Ex: Genome-Patho/SubTaskN*/Pathogen/PartnerN*	Genetic-Marker/SubTaskN*3.1.2/Crassostrea/Labogéna
<b>Description of the data</b>	Rows are the individuals and the columns are the genotyped loci. The genotype is encoded using the Nucleotide of each alleles at a given genomic position (ex A/C). The assignment will be also recorded for each individuals regarding the expected pedigree.
<b>Type</b>	Tabular files
<b>Period and frequency of data collection</b>	Unknown
<b>Geographical site of data collection (if applicable)</b>	N/A
<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.	All the description of the dataset is given in the D3.1. Including the methodology to process the data.
<b>Protocols</b> Example: 16S ribosomal RNA gene sequencing by NGS Please refer to the DMP table* for more examples	DNA extraction: QIASymphony DSP DNA Kit Genotyping: Illumina platform or Mass spectrometry MALDI-TOF (AGENA)
<b>Nature of the collected/generated data</b>	Genotyping data in .txt format and assignment files in xls format

<p><i>Example: Raw dataset in .blc/.fastqc/.fasta formats for genomic information, and processed data set will be .vcf/.bed formats.</i></p> <p><b>Please refer to the DMP table* for more examples</b></p>	
<p><b>Coverage (if applicable)</b></p> <p><i>Example: random genomic regions covered at 50 X</i></p> <p><b>Please refer to the DMP table* for more examples</b></p>	N/A
<p><b>What are the prerequisites allowing to use the data as such?</b></p> <p><i>Example: Any person able to use .fastqc file and .fasta file</i></p> <p><b>Please refer to the DMP table* for more examples</b></p>	Partners as specified in the agreement (IFREMER, SYSAAF, CNRS...)
<p><b>Sharing of main data</b></p>	<p><i>Saved and shared after publication</i></p> <p><i>Please specify</i></p>
<p><b>Archiving and preservation</b></p> <p><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></p> <p><b>Please refer to the DMP table* for more examples</b></p>	<p><i>Data will be archived on an internal server that is backed up (Labogéna server).</i></p>
<p><b>List, description and storage of associated data (metadata)</b></p> <p><i>Examples: environmental data, mortality monitoring, genotyping...</i></p>	Not relevant
<p><b>Sharing of metadata (if relevant)</b></p>	<p><i>Saved but will remain confidential even after publication</i></p> <p><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](#)

