Vivaldi Project

Data management plan

Sequences of the C2-C6 and ORF49 / 50 Regions for UK isolates of OsHV-1

Key words: OsHV-1, UK, sequences

31/05/2019

DATA MANAGEMENT PLAN

Template sheet for each dataset

Partner name	*CEFAS*
Data category	*Genome (pathogen)*
Concerned WP	*WP1* *Choisissez un élément.* *Choisissez un élément.*
Name of the VIVALDI referent(s)	*Genome (OsHV-1): Benjamin*
Reference of the dataset *Please refer to the DMP table to find the appropriate reference.* *Ex: Genome-Patho/SubTaskN°/Pathogen/PartnerN°*	Genome-Patho/subtask 1.2.2/Pathogen/CEFAS
Description of the data	*Describe the data* Sequences of the C2-C6 and ORF49 / 50 Regions for UK isolates of OsHV-1
Type	*Sequences*
Period and frequency of data collection	*Samples collected during outbreaks between 2008 and 2016*
Geographical site of data collection (if applicable)	Samples collected from 4 sites in the UK experiencing OSHV-1 disease outbreaks between 2010-2015; Whitstable bay, Poole Harbour, the River Teign, and the Blackwater Essex (River Blackwater, River Roach, River Colne and River Crouch)
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.*	DNA extracted from gills / mantle of naturally infected oysters during outbreaks of OsHV-1.
Protocols *Example: 16S ribosomal RNA gene sequencing by NGS* ***Please refer to the DMP table* for more examples***	DNA extractions performed using the Qiagen DNA tissue kit from 5mg of tissue and suspended in a 50µl volume. The initial confirmation of OsHV-1 infection was achieved using conventional PCR using the C2/C6 primers (Arzul et al 2001) and the internal primers OsHV-1 fint (TTCTAGGATATGGAGCTGCG) and OsHV-1 rint

This project has received funding from the European Union's Horizon 2020 Research and innovation programme under grant agreement N° 678589

	(GACTTCTATGGGTATGTCAG). The ORF49/50 primers used for the subsequent virus characterisation, were designed and optimised for use in the Weymouth laboratory; ORF49/50 IGR F - TATGGGTATATTGAGTGAACAC, ORF49/50 IGR Fint CAAATTAATTCCTTGCCGCTC, ORF49/50 IGR R TCAACCGGAAGTTCCATAGG and ORF49/50 IGR Rint TCCATGTTAAAGTTRACCTC DNA was amplified in 50µl reactions containing; 1x GoTaq flexi buffer (#M3171, Promega, UK) 2.5mM MgCl2, 1mM dNTP mix, 50 pmol each of the primers, 1.25 units of GoTaq® DNA Polymerase and 2.5 µl of the purified DNA. The reaction mix was subjected to 40 temperature cycles (1 min at 95°C, 1 min at 55°C and 1 min at 72°C) after an initial denaturing step (5 min at 95°C) in a Peltier PTC-225 thermal cycler, followed by a final extension step of 10 min at 72°C. The two round nested PCR was performed as above using the internal primers. Products were electrophoresced on a 2% agarose gel and the bands excised. The DNA was extracted and purified by ethanol precipitation, and both strands were sequenced using the ABI PRISM Big Dye Terminator v3.1 cycle sequencing kit using the same primers used in the amplification and analysed on a 3500xl genetic analyser. A consensus sequence for each product was generated using CLC workbench software (Qiagen) Multiple alignments were performed using Clustal W (Thompson et al. 1997) with the following Clustal parameters: a gap opening penalty of 15 and gap extension penalty of 6.66. Phylogenetic analyses were conducted using MEGA version 4 (Tamura et al. 2007). The neighbour-joining tree was constructed using a maximum composite likelihood model, and the robustness of the tree was tested using 1000 bootstrap replicates.
Nature of the collected/generated data *Example: Raw dataset in .blc/.fastqc/.fasta formats for genomic information, and processed datas set will be .vcf/.bed formats.*	Raw data as an ABI .Seq file and the analysed data as a .fasta file.

This project has received funding from the European Union's Horizon 2020 Research and innovation programme under grant agreement N° 678589

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***	Genome regions: C-region (500nt) & Orf 49 / ORF 50 (350nt)
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***	Once published and submitted to GenBank any person able to use the fasta file
Sharing of main data	Saved and shared after publication
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***	Stored on Cefas server until published, then released to GenBank
List, description and storage of associated data (metadata) *Examples: environmental data, mortality monitoring, genotyping…*	Associated mortality data held by Cefas
Sharing of metadata (if relevant)	n/a

*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**

This project has received funding from the European Union's Horizon 2020 Research and innovation programme under grant agreement N° 678589