CRyPTIC DATA SHARING POLICY

CRyPTIC, Wellcome Trust and Bill and Melinda Gates Foundation are committed to ensuring the outputs of the project, including research data, are managed and used in ways that maximise public benefit.

Making data available to all CRyPTIC Consortium members and the research community in a timely manner ensures these data can be verified, built upon and used to advance knowledge. The success of this is dependent upon the data users adopting good research practice, with integrity and transparency in managing, using and sharing research data. Our aim is to come to a shared agreement on how best to do this.

1. Types of data and management

Raw sequencing files and variant calls generated within CRyPTIC

1. Raw sequencing files will be sent to European Bioinformatics Institute (EBI) via FTP (or another agreed method) for standardised analysis and variant calling. After variant-calling QC, for sequences newly generated for the project, the variant calls and raw sequences will be uploaded to the EBI open access data repository under embargo and will only be available to the Analysis Group until a date set by the CRyPTIC Collaborators providing the samples allowing general access (which maybe immediately i.e. on submission)..

2. The variant calling files will be available immediately via FTP to the CRyPTIC Collaborator generating the related sequences. In the event CRyPTIC Collaborator would like access to their sequences a request should be made to EBI to transfer them to the FTP site. EBI is only able to transfer sequences to the CRyPTIC Collaborator generating them. .

3. CRyPTIC Consortium Members are encouraged to release their raw sequencing files and variant calls on an open access basis as soon as possible, but in any event at the time of the first publication using the sequences or within 1 year of the project ending, whichever is sooner.

Metadata and Vizion Images generated within CRyPTIC

1. Metadata (primarily DST but also limited anonymised patient/sample information) collected under CRyPTIC funding will be input into the CRyPTIC database. Access to this data will be restricted to the CRyPTIC Collaborator owning the data, unless otherwise specified by the CRyPTIC Collaborator, with exception of the Analysis Group who will have unrestricted access to the data to perform analyses, and the Data Oversight Group who will have limited access to the data for the sole purpose of quality control. No patient identifiable information will be collected.

2. CRyPTIC Consortium Members are encouraged to release metadata linked to the sequences held in EBI (or similar open access data repository) as soon as possible. In any event, the metadata must be released for general access within 6 months of publication (in accordance

with the funder's terms and conditions) or within 1 year of the project ending, whichever is sooner. If appropriate, the metadata may be released under licenced terms.

3. Vizion Images of phenotype results using UKMYC5 or MYCOTBI plates generated under CRyPTIC will be stored locally and on the CRyPTIC database. These images may be released into the public domain at any point for the sole purpose of collecting plate interpretation data.

Sequences and metadata contributing to but generated outside of CRyPTIC

1. Unpublished raw sequences, variant calls and metadata will be uploaded to EBI (or similar open access data repository) embargoed from general access until a date set by the CRyPTIC Consortium Member, unless such data have already been uploaded by the CRyPTIC Consortium Member themselves. For WGS this should not be later than the date of first publication using the data. For WGS and metadata, CRyPTIC Consortium Members would be strongly encouraged to make data generally available on the same timelines as for data generated within CRyPTIC.

Analysed data

1. Publications shall be made in accordance with the CRyPTIC publication policy. Full datasets shall be released on publically accessible databases within 6 months of their publication in an open access journal, if the data are not already available through the repositories above.

2. Use of data

Data Oversight Group

1. The Data Oversight Group shall have access to data on a limited basis to ensure data quality is maintained throughout CRyPTIC.

Analysis Group

1. The Analysis Group will have access to all data generated within and contributing to the CRyPTIC project.

2. Primary quality and 'first pass' analysis of genomic and metadata will be carried out by the Analysis Group under the terms of the CRyPTIC project and data sharing agreement. At least every second month, initial results will be presented to CRyPTIC Consortium Members. CRyPTIC Consortium Members will discuss and agree interpretation of the results, further analysis and potential CRyPTIC Secondary Studies.

final

<u>Other CRyPTIC Consortium Members</u>

1. CRyPTIC Collaborators may use the data they generated under the Project in studies falling outside the CRyPTIC Primary Objective and CRyPTIC Secondary Studies, providing the objectives of such study does not jeopardise the ability of the Project to meet CRyPTIC Primary Objective and hence the integrity of the Project.

   CRyPTIC Consortium Members may not use the data generated by any other CRyPTIC Participant without permission from the CRyPTIC Consortium Member generating such data, unless the data is already open access.

2. In the event a CRyPTIC Consortium Member wishes to use unpublished data generated by another CRyPTIC Consortium Member, this should be discussed during an analysis meeting (held every second month) and between the CRyPTIC Consortium Member(s) owning the data. If all CRyPTIC Consortium Members are in agreement, a very short CRyPTIC Secondary Study summary will be jointly submitted to the Executive Committee for confirmation that data embargoes should be lifted for the relevant data and CRyPTIC Consortium Members. The Executive Committee will also confirm that the secondary study proposal does not harm the integrity of the Project or undermine any pre-existing country-specific study. This confirmation should not be unreasonably withheld, and any disagreements between CRyPTIC Consortium Member and the Executive Committee would be escalated to the General Assembly. Any unpublished data will only be released to the CRyPTIC Collaborator generating such data even when permission has been granted by the data owner to a CRyPTIC Collaborator to use the data.

3. In addition, if the proposed secondary use of data includes requests for support from the Analysis Group, the Executive Committee will determine whether time can be allocated to support the CRyPTIC Secondary Study.