**DATA ACCESS POLICY**

**Indonesian Genome Diversity Project Phase 2**

**EGAS00001003654 and EGAS00001003670**

These terms and conditions govern access to the managed access datasets (details of which are set out in Appendix I) to which the User Institution has requested access. The User Institution agrees to be bound by these terms and conditions.

**DEFINITIONS**

*Authorised Personnel*: The individuals at the User Institution to whom the Indonesian Genome Diversity Project Phase 2 (IGDP2) committee grants access to the Data. This includes the User, the individuals listed in Appendix II and any other individuals for whom the User Institution subsequently requests access to the Data. Details of the initial Authorised Personnel are set out in Appendix II.

*Data*: The managed access datasets to which the User Institution has requested access.

*Data* *Producers*: The Indonesian Genome Diversity Project Phase 2 (IGDP2) committee and the collaborators listed in Appendix I responsible for the development, organisation, and oversight of these Data.

*External* *Collaborator*: A collaborator of the User, working for an institution other than the User Institution.

*Project*: The project for which the User Institution has requested access to these Data. A description of the Project is set out in Appendix II.

*Publications*: Includes, without limitation, articles published in print journals, electronic journals, reviews, books, posters and other written and verbal presentations of research.

*Research* *Participant*: An individual whose data form part of these Data.

*Research* *Purposes*: Shall mean research that is seeking to advance the understanding of genetics and genomics, including the treatment of disorders, and work on statistical methods that may be applied to such research.

*User*: The principal investigator for the Project.

*User Institution(s)*: The Institution that has requested access to the Data.

*Indonesian Genome Diversity Project Phase 2 (IGDP2) committee*: The Indonesian Genome Diversity Project Phase 2 committee comprises three members: Prof Herawati Sudoyo from the Mochtar Riady Institute for Nanotechnology, Jakarta, Indonesia; Prof J. Stephen Lansing from the Santa Fe Institute, USA; and Prof Murray P. Cox from Massey University, New Zealand. These members will retain their role, even if they relocate to other institutions. Membership of this committee may change over time. For legal purposes, the Indonesian committee member has precedence on all matters relating to this document.

**DATA ACCESS AGREEMENT**

1. The User Institution agrees to only use these Data for the purpose of the Project (described in Appendix II) and only for Research Purposes. The User Institution further agrees that it will only use these Data for Research Purposes which are within the limitations (if any) set out in Appendix I.

2. The User Institution agrees to preserve, at all times, the confidentiality of these Data. In particular, it undertakes not to use, or attempt to use these Data to compromise or otherwise infringe the confidentiality of information on Research Participants. Without prejudice to the generality of the foregoing, the User Institution agrees to use at least the measures set out in Appendix I to protect these Data.

3. The User Institution agrees to protect the confidentiality of Research Participants in any research papers or publications that they prepare by taking all reasonable care to limit the possibility of identification.

4. The User Institution agrees not to link or combine these Data to other information or archived data available in a way that could re-identify the Research Participants, even if access to that data has been formally granted to the User Institution or is freely available without restriction.

5. The User Institution agrees only to transfer or disclose these Data, in whole or part, or any material derived from these Data, to the Authorised Personnel. Should the User Institution wish to share these Data with an External Collaborator, the External Collaborator must complete a separate application for access to these Data.

6. The User Institution agrees that the Data Producers, and all other parties involved in the creation, funding or protection of these Data: a) make no warranty or representation, express or implied as to the accuracy, quality or comprehensiveness of these Data; b) exclude to the fullest extent permitted by law all liability for actions, claims, proceedings, demands, losses (including but not limited to loss of profit), costs, awards damages and payments made by the Recipient that may arise (whether directly or indirectly) in any way whatsoever from the Recipient’s use of these Data or from the unavailability of, or break in access to, these Data for whatever reason and; c) bear no responsibility for the further analysis or interpretation of these Data.

7. The User Institution agrees to follow the Fort Lauderdale Guidelines (<https://www.wtccc.org.uk/wtccc/assets/wtd003207.pdf>) and the Toronto Statement (<http://www.nature.com/nature/journal/v461/n7261/full/461168a.html>). This includes but is not limited to recognising the contribution of the Data Producers and including a proper acknowledgement in all reports or publications resulting from the use of these Data.

8. The User Institution agrees to follow the Publication Policy in Appendix III. This includes respecting the moratorium period for the Data Producers to publish the first peer-reviewed report describing and analysing these Data.

9. The User Institution agrees not to make intellectual property claims on these Data and not to use intellectual property protection in ways that would prevent or block access to, or use of, any element of these Data, or conclusion drawn directly from these Data.

10. The User Institution can elect to perform further research that would add intellectual and resource capital to these data and decide to obtain intellectual property rights on these downstream discoveries. In this case, the User Institution agrees to implement licensing policies that will not obstruct further research and to follow the U.S. National Institutes of Health Best Practices for the Licensing of Genomic Inventions (2005) (<https://www.icgc.org/files/daco/NIH_BestPracticesLicensingGenomicInventions_2005_en.pdf>) in conformity with the Organisation for Economic Co-operation and Development Guidelines for the Licensing of the Genetic Inventions (2006) (<http://www.oecd.org/science/biotech/36198812.pdf>).

11. The User Institution agrees to destroy/discard the Data held, once it is no longer used for the Project, unless obliged to retain the data for archival purposes in conformity with audit or legal requirements.

12. The User Institution will notify the Indonesian Genome Diversity Project Phase 2 (IGDP2) committee within 30 days of any changes or departures of Authorised Personnel.

13. The User Institution will notify the Indonesian Genome Diversity Project Phase 2 (IGDP2) committee prior to any significant changes to the protocol for the Project.

14. The User Institution will notify the Indonesian Genome Diversity Project Phase 2 (IGDP2) committee as soon as it becomes aware of a breach of the terms or conditions of this agreement.

15. The Indonesian Genome Diversity Project Phase 2 (IGDP2) committee may terminate this agreement by written notice to the User Institution. If this agreement terminates for any reason, the User Institution will be required to destroy any Data held, including copies and backup copies. This clause does not prevent the User Institution from retaining these data for archival purpose in conformity with audit or legal requirements.

16. The User Institution accepts that it may be necessary for the Data Producers to alter the terms of this agreement from time to time. As an example, this may include specific provisions relating to the Data required by Data Producers other than the Indonesian Genome Diversity Project Phase 2 (IGDP2) committee. In the event that changes are required, the Data Producers or their appointed agent will contact the User Institution to inform it of the changes and the User Institution may elect to accept the changes or terminate the agreement.

17. If requested, the User Institution will allow data security and management documentation to be inspected to verify that it is complying with the terms of this agreement.

18. The User Institution agrees to distribute a copy of these terms to the Authorised Personnel. The User Institution will procure that the Authorised Personnel comply with the terms of this agreement.

19. This agreement (and any dispute, controversy, proceedings or claim of whatever nature arising out of this agreement or its formation) shall be construed, interpreted and governed by the laws of the Republic of Indonesia and shall be subject to the exclusive jurisdiction of Indonesian courts.

*AGREED FOR USER INSTITUTION*

Name:

Title:

Date:

Signature:

*PRINCIPAL INVESTIGATOR*

I confirm that I have read and understood this Agreement

Name:

Title:

Date:

Signature:

*AGREED FOR THE INDONESIAN GENOME DIVERSITY PROJECT PHASE 2 (IGDP2) COMMITTEE*

Name:

Title:

Date:

Signature:

**APPENDIX I – IGDP2 DATASET DETAILS**

High-coverage whole genome sequences were collected to study patterns of genomic variation across the broad geography of Indonesia and New Guinea. This region has experienced an extremely complex demographic history, including repeated bouts of admixture with archaic and modern human groups. This dataset reports whole genome sequences for 161 individuals from 14 populations spanning this geographical region, from communities close to mainland Asia through to New Guinea. Particular attention has been paid in the original study to genomic signals that are informative for population history, including admixture with archaic hominins and the role of modern human admixture during the late Pleistocene and Holocene.

*Contact person:*

Prof Herawati Sudoyo, Mochtar Riady Institute for Nanotechnology, Jakarta, Indonesia

herawati.sudoyo@mrinstitute.org

*The Indonesian Genome Diversity Project Phase 2 (IGDP2) project, which is a consortium comprising:*

Prof Herawati Sudoyo, Mochtar Riady Institute for Nanotechnology, Jakarta, Indonesia (Lead Organization)

Prof J. Stephen Lansing, Santa Fe Institute, NM, USA

Prof Murray P. Cox, Massey University, New Zealand

*Names of other data producers/collaborators:*

Massey University, New Zealand:

Georgi Hudjashov

Complexity Institute, Nanyang Technological University, Singapore:

Guy S. Jacobs

Estonian Biocentre, University of Tartu, Estonia:

Mait Metspalu, Georgi Hudjashov

Eijkman Institute, Jakarta, Indonesia:

Pradiptajati Kusuma

*Specific limitations on areas of research:*

Users must be formally affiliated with an officially recognized Institution.

The User can replicate existing studies published by the Indonesian Genome Diversity Project Phase 2 (IGDP2) research program, using similar techniques, approaches and methods, to ensure that the published science is reproducible. Approval will be automatically granted for such use.

The User cannot publicly release IGDP2 data. All rights regarding data release remain with the IGDP2 committee.

The User cannot use IGDP2 data for for-profit purposes.

The User cannot undertake studies for personal use, such as family history research, or perform this research for others.

The User cannot undertake studies of a medical or clinical nature without first seeking the approval of the IGDP2 committee. Evidence of specific ethical approvals, including documentation from an Indonesian ethics board, will likely be necessary for approval to be granted.

The User can undertake new demographic studies, including studies focusing on the history of archaic hominins and modern humans, as long as this does not directly compete with ongoing studies by the Indonesian Genome Diversity Project Phase 2 (IGDP2) program. All research projects must be approved by the IGDP committee.

The User can undertake studies of selection, including on alleles with archaic and modern ancestry, as long as this does not directly compete with ongoing studies by the Indonesian Genome Diversity Project Phase 2 (IGDP2) program. All research projects must be approved by the IGDP2 committee.

Note that all uses of the data must have specific prior approval from the IGDP2 committee. Evidence of ethical approvals, including documentation from an Indonesian ethics board, may be necessary for approval to be granted in some cases.

A moratorium on publication until a given date may be a condition of data access and use.

*Minimum protection measures required:*

Data can be held in unencrypted files on an institutional compute system, with Unix user group read/write access for one or more appropriate groups but not Unix world read/write access behind a secure firewall. Laptops holding these data should have password protected logins and screen locks (set to lock after 5 min of inactivity). If held on USB keys or other portable hard drives, the data must be encrypted.

**APPENDIX II – PROJECT DETAILS (to be completed by the Requestor)**

*Brief, but detailed, abstract of the Project in which the Data will be used (500 words max)*

*All Individuals who the User Institution to be named as registered users (repeat as needed)*

Name of Registered User:

Job Title:

Email:

Supervisor:

*All Individuals that should have an account created at the EGA (repeat as needed)*

Name of Registered User:

Job Title:

Email:

**APPENDIX III – PUBLICATION POLICY**

The Indonesian Genome Diversity Project Phase 2 (IGDP2) project intends to publish the results of their analysis of this dataset and do not consider its deposition into public databases to be the equivalent of such publications. The Indonesian Genome Diversity Project Phase 2 (IGDP2) project anticipates that the dataset could be useful to other qualified researchers for a variety of purposes. However, some areas of work are subject to a publication moratorium.

The publication moratorium covers any publications (including oral communications) that describe the use of the dataset. For research papers, submission for publication should not occur until the IGDP2 committee has provided written consent for publication on or after a given date, either in a separate written document, or more commonly, as part of this agreement.

In any publications based on these data, please describe how the data can be accessed, including the name of the hosting database (e.g., The European Genome-phenome Archive at the European Bioinformatics Institute) and its accession numbers (e.g., EGAS00001003654 and EGAS00001003670), and acknowledge its use in a form agreed by the User Institution with the Indonesian Genome Diversity Project Phase 2 (IGDP2) project committee.

**Project Title:**

Indonesian Genome Diversity Project Phase 2

**Repository and Accession Number:**

European Genome-phenome Archive

Whole genome sequencing:  EGAD00001005059 (linked to study EGAS00001003654)

DNA genotyping array: EGAD00010001713 (linked to study EGAS00001003670)

**Citation:**

Natri, H.M., G. Hudjashov, G. Jacobs, P. Kusuma, L. Saag, C.C. Darusallam, M. Metspalu, H. Sudoyo, M.P. Cox, I. Gallego Romero and N.E. Banovich. 2022. Genetic architecture of gene regulation in Indonesian populations identifies QTLs associated with global and local ancestries. *American Journal of Human Genetics* 109:50-65.

Natri, H.M., K.S. Bobowik, P. Kusuma, C.C. Darusallam, G.S. Jacobs, G. Hudjashov, J.S. Lansing, H. Sudoyo, N.E. Banovich, M.P. Cox and I. Gallego Romero. 2020. Genome-wide DNA methylation and gene expression patterns reflect genetic ancestry and environmental differences across the Indonesian archipelago. *PLoS Genetics* 16:e1008749.