



## HostSeq Data Sharing Policy and Procedures CGEn HostSeq

### 1. Introduction

CGEn's HostSeq is a Databank ("Databank") of The Hospital for Sick Children, McGill University and BC Cancer (collectively, "CGEn"), funded by the Government of Canada through Innovation, Science, and Economic Development Canada (ISED) and administered by Genome Canada, pursuant to which CGEn conducted the whole genome sequencing ("WGS") of more than 10,000 human genomes using samples donated by persons confirmed or at risk to have been infected with the novel coronavirus SARS-CoV2.

The goal of HostSeq was to provide the WGS results and their associated meta-data to the researcher community, in a cloud-based and/or local server data environment for research and analysis. Through the Databank, authorized researchers have access to a large database of WGS and meta data for thousands of SARS-CoV2 positive cases, which should speed discovery of COVID-19 disease severity biomarkers, treatments, and ultimately vaccines. HostSeq data is also consented for other health research and it is the policy and objective of HostSeq to make the Databank available to the broadest possible research community.

The Databank contains Research Data, WGS Data, Contributor Provided Data and other data or information that CGEn elects to make available. Data is obtained from (or derived from biomaterials provided by) retrospectively and prospectively recruited cohorts of Participants, where appropriate Review Board waivers are obtained. CGEn shall have access to any Data for non-research, operational and quality assurance purposes only. Any access to the CGEn HostSeq Databank and/or Data for research purposes, will be established in accordance with this HostSeq Data Sharing Policy and Procedures.

This HostSeq Data Sharing Policy and Procedures (this "Policy") states the policy of CGEn HostSeq with respect to (1) the process for approving proposals from Researchers to access the Program Databank, and (2) subject to such approval, the procedures for a Researcher to download copies of Program Data from Program Databank. This controlled access is to be offered to Researchers from around the world who are conducting research on COVID-19 and other health conditions.

This Policy may be amended by CGEn from time to time. The governing version of this Policy will be the version posted on the CGEn HostSeq Website at any given time.

### 2. Definitions

*"Access Authorization Letter"* means the formal letter sent by the DACO to Researcher to advise of the decision of the DACO in connection with Researcher's Researcher Application.

*"Access Renewal Letter"* means the formal letter sent by the DACO to Researcher to advise of the decision of the DACO in connection with Researcher's Renewal/Completion Form.

*"CGEn HostSeq Website"* means the CGEn HostSeq website at <https://www.cgen.ca/project-overview>

*"Consultant"* means Policy Partnerships Project for Genomic Governance (P3G2) at McGill University or such other independent consultant as CGEn may appoint from time to time.



*“Contributor”* means the contributing organization/institution and/or individual that made the Participant’s biological materials available for WGS and the associated Research Data available for deposit into the Databank.

*“Contributor Provided Data”* means any and all data provided by any other researcher other than Researcher to CGEn for inclusion into the Databank.

*“Coordinating Center”* means the Coordinating Center at SickKids.

*“DACO”* means the Program’s independent Data Access Compliance Office.

*“Databank Access Agreement”* or *“DAA”* means the databank access agreement between Researcher, the Institution (if any) and CGEn that sets forth the terms and conditions regarding access by a Researcher to the Databank.

*“Institution”* means the organization at which Researcher is employed, affiliated or enrolled.

*“Participant”* means a provider of Biological Samples and/or Research Data that is used as part of HostSeq.

*“Data”* means Research Data, WGS Data and Contributor Provided Data and other data or information that CGEn elects to make available through the Databank in accordance with the Policy.

*“Databank”* means the cloud based and/or local-server data environment maintained by CGEn as part of HostSeq that stores Data.

*“Research Data”* means de-identified information concerning Participants, which may include family configuration, age at time of testing, sex, diagnosis, family and medical history, and any other clinically relevant information collected by Contributor and deposited into the Databank. For clarity, Research Data shall not include any personally identifying information about the Participant, his/her family or its members.

*“Renewal/Completion Form”* means the form submitted by Researcher to renew access to the Databank, which will be reviewed and approved by DACO in accordance with this Policy.

*“Research Project”* means the research project that a Researcher intends to conduct using Data accessed through the Program Databank as described in the Researcher Application.

*“Research Team”* means the research staff employed by Institution and assisting Investigator in the conduct of the Research Project.

*“Researcher”* means a person and/or institution seeking access to the Databank in accordance with this Policy.

*“Researcher Application”* means the application submitted by Researcher to obtain access to the Databank, which will be reviewed and approved by DACO in accordance with this Policy.

“*Review Board*” means the Institutional Review Board, Research Ethics Committee, Research Ethics Board or another equivalent body under the laws or regulations governing a Researcher and a Research Project.

“*WGS Data*” means whole genome sequencing data derived from Biological Samples of Participants.

### **3. Governance**

The Coordinating Centre and DACO will have the responsibility for implementing this Policy as set forth herein.

The independent Data Access Committee (DAC) includes members listed on the CGEn HostSeq website.

### **4. Application Procedure**

In order to access the Databank, Researcher must apply for and receive approval for access to the Databank in accordance with the following procedures.

#### **Step 1: Completion of Researcher Application**

Researcher must prepare a completed Researcher Application via the HostSeq application portal (<https://www.cgen.ca/hostseq-databank-access-request>). The Researcher Application requires Researcher to provide the following information:

1. Name and credentials of Researcher;
2. Institutional affiliation;
3. List of up to 3 peer reviewed publications by Researcher;
4. Outline of Research Project, including: Title of Research Project; Research question and aims; Background and rationale; Methods and how the HostSeq data will be used; and Expected outcomes and impact of the project
5. Summary of Research Project suitable for a lay audience;
6. Description of Program Data proposed to be downloaded from the Program Databank;
7. A copy of the IRB or Ethics Committee approval letter and the associated Data Collection Form to confirm that a Review Board has approved Researcher’s use and receipt of Program Data \*;
8. List of all individuals in Research Team with their respective job titles, institutional affiliations and institutional e-mail addresses.

\*Should Researcher seek an exemption from a Research Board review, Researcher shall request a Review Board to consider such exemption.

#### **Step 2: Review by the Data Access Compliance Office**

The DAC will have the discretion to approve or decline a Researcher Application based on ethical, programmatic or other relevant considerations. Among other things, the DAC may consider the following criteria:

- The Researcher Application is complete;

- Researcher is qualified to conduct the Research Project and undertake the proposed analysis;
- The Research Project is feasible given the resources in the Program Databank;
- The scope of the Research Project is relevant to COVID-19 or other health conditions;
- Researcher has confirmed that Review Board approval has been obtained, or, if no confirmation of Review Board approval is provided, whether the explanation as to why Review Board approval is not required is adequate; and
- The lay summary of the Research Project is comprehensible to the general public.

For Researcher Applications from Researcher who are affiliated with known institutions, the DACO will endeavor to render its decision within thirty (30) days after receipt by the DACO of the Researcher Application. For the avoidance of doubt, all Researcher Applications will be considered from Researchers who do not have an institutional affiliation, but review of such Researcher Applications is likely to take more time.

### Step 3: Notification of Decision

The DACO will send an Access Authorization Letter to Researcher and Coordinating Centre advising whether a Researcher Application is approved, declined or conditionally approved. If the DACO has approved a Researcher Application, the Access Authorization Letter issued to Researcher will indicate that the Coordinating Centre will be in contact with Researcher to begin the process of Data Access Agreement (DAA) execution.

If the DACO has conditionally approved a Researcher Application, the Access Authorization Letter will set forth the additional information required to be submitted to the DACO. Researcher will provide additional information, as requested, to the DACO. Upon receipt of such additional information, the DACO will review the Researcher Application, together with the additional information in accordance with Step 2 above, and a notification of the decision of the DACO will be provided in accordance with this Step 3.

### Step 4: Execution of Data Access Agreement (DAA)

Upon approval of Research Project by DACO, Coordinating Centre will contact Researcher and Institution to review DAA. DAA has been drafted in accordance with CGEn policies and international guidelines. No changes will be made to DAA unless demonstrably required by law. Any comments on the draft DAA made by Institution will be directed to SickKids legal department. Once there is agreement on the terms of DAA, DAA will be fully executed.

DAA will govern the Researcher's access to the Databank. The terms of DAA will prevail over any inconsistent terms of this Policy, the CGEn HostSeq Website or elsewhere, and over any oral or written statement made by the staff of the Coordinating Center or of the DACO or any other representative of CGEn.

### Step 5: Data Access

Upon full execution of DAA, Researcher and/or Research Team members will be given instructions to access the Data, which may include assigning login credentials to the Researcher and/or Research Team members using the email address provided by Researcher in the Researcher Application, for an account enabling the Researcher to download the Data from an online platform. This account will be owned by

Researcher and all costs associated with work undertaken in that account will be the responsibility of Researcher.

## **5. Duration of Access and Renewal Procedure**

Each DAA will have a term of one (1) year from the date of execution of the DAA by CGEn. Any changes to the Research Project or Research Team over the course of the project must be reported to the DACO as described in the DAA. Automatic renewal of DAA for another one (1) year term is possible upon DACO approval of Renewal/Completion Form as set forth below.

To renew a DAA and continue access to the Databank after expiry date of the existing DAA, Researcher will be required to submit a Renewal/Completion Form. The Renewal/Completion Form will require Researcher to provide:

- Confirmation of current status of Research Project (whether Research Project is ongoing or complete, status of Data analysis);
- Confirmation of ongoing or renewed Ethics Approval, including renewed Ethics Approval Letter if required;
- Brief update on Research Project progress, including early outcomes and publications;
- Report on all publications prepared using the results of the Research Project;
- Confirmation that the content of the Researcher Application originally submitted remains correct and complete, or any required modifications to Research Project;
- An updated list of and contact information for the members of the Research Team, or confirmation that the list of and contact information for the members of the Research Team remains correct and complete;
- Incidents within the course of Research Project;
- Report on any complaints or allegations of mismanagement.

The Renewal/Completion Form will be provided to Researcher by DACO approximately six (6) weeks prior to expiration of DAA. Researcher must submit the completed Renewal/Completion Form, signed by Institution, to the DACO by e-mail ([daco@cgen.ca](mailto:daco@cgen.ca)). It is recommended that the completed Renewal/Completion Form is submitted to DACO no later than four (4) weeks prior to the expiration of the DAA Term.

Upon approval of the Renewal/Completion Form by DACO, DACO will send an Access Renewal Letter to Researcher and Coordinating Centre to advise them of the renewal. The Access Renewal Letter will renew the DAA and Researcher and the Research Team can continue to access the Databank using the existing log-on credentials. The renewal term will be for one (1) year.

## **6. Project Termination Procedure**

If Researcher wishes to terminate their access and use of Data, Researcher indicates in the Renewal/Completion Form (or at any point in writing to Coordinating Centre and DACO) that the project is complete, and access to the Databank is no longer required. The final, completed Renewal/Completion Form signed by Institution, including certification in writing the destruction of Program Data archived, copied, downloaded or otherwise extracted by Researcher and the Research Team from the Program Databank, must be submitted to DACO and Coordinating Centre prior to expiry of the DAA. The DAA will then expire without being renewed (or be terminated in accordance with the terms of the DAA), and the log-on credentials of Researcher and the Research Team will expire simultaneously. The obligations of

Researcher that survive termination of the DAA will remain in effect, including, without limitation, the obligation to follow the HostSeq Publication Policy.

## 6. CGEn HostSeq Researcher Application Registry

The Coordinating Centre will create a registry of the names, affiliated institutions and Research Project lay summaries for all Researchers who have been approved for access to the Databank. This registry will be available for public access. Other aspects of the Researcher Application will not be disclosed to the public.

## 7. Data Security Assessment

In order to avoid, among other things, inadvertent disclosure of the data of Participants to unauthorized individuals, the DAA requires that Researcher implement data security safeguards, including appropriate technical and organizational measures to protect Data against accidental or unlawful destruction or accidental loss, alteration, unauthorized disclosure or access, and which provide a level of security appropriate to the risk represented by the processing and the nature of the data to be protected.

Minimum data security safeguards are as follows:

- *Physical Security* – Downloaded copies of Data must be maintained on physically secure computer systems, such as a locked office. If downloaded copies of Data are stored on a portable device (e.g., laptop, tablet, USB thumb drive, external drives), the portable device must be encrypted to avoid its disclosure in case of loss or theft.
- *Access Security* – Only Researcher and Research Team should have access to Data. If downloaded copies of Data are stored locally on a shared computer system or a file server, then the downloaded data must be password or encryption protected so that only Researcher and the Research Team have access to the downloaded copies. Any backup media must either be encrypted or stored in a physically secure location. Any remote access to Data must be done using multi factor authentication.
- *Network Security* – If downloaded copies of Data are stored on a network-accessible computer, some combination of network firewalls, network intrusion monitoring and virus scanning software must be implemented. If Data is used on multiple systems, Researcher must ensure that the data access security policies are retained throughout the processing of Data on all other systems.
- *Destruction of Data* – Upon completion of the Research Project (or the earlier expiration or termination of the DAA), Researcher must delete and destroy all copies of Data in accordance with the DAA.
- *Data Security Policies* – Researcher must have an information technology (IT) policy in place that addresses at a minimum the following items:
  - Logging and auditing of access to data and the computer network;
  - Password protection of the computer network;
  - Virus and malware protection of the computer network;
  - Auditable data destruction procedure;



- Secure data backup procedure; and
- Strong encryption on any portable device which may store or provide access to Data.

Upon reasonable request of CGEn, Researcher will submit its data processing facilities, data files and documentation needed for processing to reviewing, auditing and/or certifying by Researcher (or any independent or impartial inspection agents or auditors, selected by CGEn and not reasonably objected to by the Researcher) to ascertain compliance with the warranties and undertakings in these clauses, with reasonable notice.

#### **8. Amendment of this Policy**

CGEn may at any time and without notice amend this Policy in any manner in which CGEn determines to be in its best interest.