HostSeq COVID-19 Databank:
Consent elements for prospective and retrospective datasets

The HostSeq Databank will accept datasets from cohorts specifically consented for HostSeq (**prospective datasets**) as well as samples for sequencing and datasets from cohorts not specifically consented for HostSeq (**retrospective datasets**), provided they meet the core consent elements identified below.

I) CORE RESEARCH CONSENT ELEMENTS FOR PROSPECTIVE COLLECTION OF DATASETS

Consent to the collection of tissue samples by local sites, followed by the genetic analysis of tissue samples and sharing of research data through the HostSeq Databank, should be obtained by researchers at the local site. This consent can be obtained from the tissue donor (first person consent) or, where applicable, by obtaining consent from their legally authorized representative (substituted consent). Due to the nature of COVID-19, in many cases, consent may need to be obtained verbally, and necessary adjustment to the consent form and consent guidelines should be implemented, as recommended by the local site.

Prospective cohort consent forms should minimally include the following **core consent elements** to foster broad data sharing through the HostSeq Databank[1]:

	If a researcher at a local site wishes to deposit datasets in **HostSeq COVID-19 controlled-access Databank,** research consent should be obtained for:
Research data	Whole genome sequencing of the sample and the ongoing collection of clinical data from participant's medical records/chart, administrative databases, etc.
International sharing	International sharing of genetic and clinical data
Future research use	Future health research on COVID-19 and other health outcomes
Commercial use	Use of genetic and clinical data for commercial purposes
Controlled access	Sharing of genetic and clinical data through a controlled-access mechanism
Storage on cloud servers in Canada	Storage of genetic and clinical data in the HostSeq Databank, on centralized Canadian cloud servers
Duration of storage	Indefinite storage of genetic and clinical data
Data withdrawal	Not possible to withdraw data that has already been distributed and used
Re-identification	Low risk that the participant could be re-identified in the future
Option for recontact of participants, (not mandatory, strongly recommended)[2]	Optional recontact of participants, (yes/no) It is strongly recommended that, where applicable, mature minors be included in the consent process and the option for recontact.
Assent	Assent for children, where applicable

[1] Examples of appropriate consent language for each of these core elements are provided in Appendix A.

[2] HostSeq strongly recommends that contributing collections include an option to re-contact participants in their consent forms. HostSeq will accept datasets from participants who have agreed or refused re-contact, however, the databank will track whether participants have agreed to be re-contact or not in order to determine whether future data/sample collection is possible.

If any of the items listed above are not included in the consent documents, datasets should not be deposited as-is in the HostSeq Databank without obtaining appropriate approvals from your local research ethics committee.

II) ASSESSMENT OF CONSENT ELEMENTS FOR INCLUSION OF DATASETS FROM RETROSPECTIVE COHORTS

The steps below aim to provide guidance on determining whether samples or datasets from retrospective cohorts are suitable for inclusion in the HostSeq COVID-19 Databank. Data submitted to the HostSeq COVID-19 Databank will be maintained in a controlled access database. Retrospective datasets include for example:

- Datasets that could be generated from already collected tissue samples (e.g. archival samples, samples collected in the context of another research project/biobank, samples collected, etc.);
- Datasets that were generated before the creation of the HostSeq COVID-19 Databank.

Retrospective consent materials may have used different language and in some cases may be ambiguous, or be silent, as to sample and data sharing and their potential uses. To help researchers from local sites determine whether samples and/or data can be used for the HostSeq COVID-19 Database, the following assessment tool was developed, based on the principles found in the Global Alliance for Genomics and Health's Consent Policy[3] and other best practice guidelines[4].

***Question 1*:** Does the consent form permit genetic sequencing of the tissue sample collected?

- If the question is **Not Applicable** because the tissue sample has already been analyzed and genetic data has already been derived with appropriate authorization, please proceed to ***Question 2***.
- If the answer is **Yes**, please proceed to ***Question 2***.
- If the answer is **No**, please proceed to ***Question 3* for re-consent or consent waivers**.

Question 2:

Does the consent form pertaining to datasets from retrospective cohorts indicate that:	***Yes***	***No***
a) Genetic analyses will be undertaken on the tissue sample and relevant clinical data will be collected?		
b) Data will be shared internationally?		
c) Data will be used for future health research on COVID-19 and related disorders, or for broader health/biomedical research uses?		
d) Data may be transferred and stored outside of the institution where it was collected and be made available to researchers through that external platform?		
e) Data will be stored for an indefinite period of time?		

[3] Global Alliance for Genomics and Health: Consent Policy (September 2019), available online at: https://www.ga4gh.org/wp-content/uploads/GA4GH-Final-Revised-Consent-Policy_16Sept2019.pdf

[4] For example, Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans – TCPS2 (2018), Section 5.5A (Consent and secondary use of identifiable information for research purposes) and 12.3A (Consent and secondary use of human biological materials for research purposes).

f) It may not be possible to withdraw data that has already been distributed and used by researchers?	
g) There is a minimal risk of re-identification in the future?	
h) Data can be used for commercial purposes?	
i) Genetic and clinical data will be shared through a controlled-access mechanism, meaning that research applicants are required to apply to an access committee to request access to the data, or similar process/consent language?	

- If **you have answered <u>yes to all of the above</u>**, core consent elements are deemed to be met to submit datasets to the HostSeq COVID-19 Databank. (In case of doubt or ambiguity in the consent language, please consult your local research ethics committee for guidance).
- If **you have answered <u>no to any of the above</u>**, please proceed to ***Question 3*** to determine if re-consent of participants is possible, or whether a consent waiver should be sought from your institutional ethics committee to deposit data in the HostSeq COVID-19 Databank.

Question 3:

Re-contact / re-consent	***Yes***	***No***
a) Does your consent form allow for re-contact of participants?		
b) Is it feasible for you to re-contact and re-consent your participants for inclusion of their data in the HostSeq COVID-19 Databank?		

- If **you have answered <u>yes to both questions</u>**, please re-contact and re-consent the participants and include the HostSeq COVID-19 Databank Consent Elements for prospective cohorts in your consent material.
- If **you have answered <u>no to either question</u>**, please proceed to ***Question 4***.

Question 4:

Requesting a waiver of consent	***Yes***	***No***
a) Is it possible for you to apply to your local ethics committee (or equivalent) to obtain a waiver of consent requirement in order to deposit your dataset in the HostSeq COVID-19 Databank? (See TCPS, article 5.5A and 12.3A)		

- If **you have answered <u>yes</u>**, please request and obtain a consent waiver you're your institutional research ethics committee, according to your local procedures.
- If **you have answered <u>no</u>**, your data cannot by deposited in the HostSeq COVID-19 Databank.

****

Appendix A: Example of consent clauses for the prospective Core Consent Elements

	Core Consent Elements	Example of consent clause language:
Research data	Whole genome sequencing of the sample and the ongoing collection of clinical data from participant's medical records/chart, administrative databases, etc.	You are being asked to consent for the whole genome sequencing of the DNA from your sample and to the storage of your **genetic data** in the HostSeq COVID-19 Databank, to be used for research purposes. You will also be asked to provide **clinical data** which includes some personal information about yourself and your health, as well information such as your age, ethnicity, and family's health history. We will also request health information about you from your family doctor and from other institutions or registries that may have your health information, for example, [*where applicable, include any relevant governmental/administrative health data repository in your province*]. We may get research data from other studies that you were involved in or future studies, if you agree.
International sharing	International sharing of genetic and clinical data	The HostSeq Databank will share your genetic and clinical data with approved researchers in Canada and around the world (which may include national and international researchers from academia, charitable organisations, hospitals, and 'for-profit' private companies, such as drug companies).
Future research use	Future health research on COVID-19 and other health outcomes	Your stored genetic and clinical data will be shared by the CGen Databank for future research on COVID-19 and other health outcomes.
Commercial use	Use of genetic and clinical data for commercial purposes	It is possible that future research conducted using your data combined with data from others will eventually lead to the development of new diagnostic tests, new drugs or other commercial products. If this happens, you will not receive any part of the profits from such products. The rights to the commercial products will belong to the sponsor, collaborators or future unknown researchers.
Controlled access	Sharing of genetic and clinical data through a controlled-access mechanism	Your data will only be shared with researchers approved by CGen (the sponsor) after review by a Data Access Committee (DAC). This DAC will verify, among other criteria, that the proposed research use is in conformity with the objectives of the HostSeq Databank, and that the research team applying for access has obtained the proper research ethics approval. Approved researchers will sign agreements. These agreements will control how the data will be used.
Storage on cloud servers in Canada	Storage of genetic and clinical data in the HostSeq Databank, on centralized Canadian cloud servers	Data in the HostSeq Databank is under the responsibility of CGen. CGen is a Canadian, federally funded national platform for genome sequencing and analysis. Data in the HostSeq Databank is stored on cloud servers (i.e. online), based in Canada.
Duration of storage	Indefinite storage of genetic and clinical data	The data stored in the HostSeq Databank will be kept indefinitely, or until it is no longer useful for research.
Data withdrawal	Not possible to withdraw data that	If you decide to withdraw from the HostSeq Databank, your data will no longer be shared, and no new data will be collected. If you decide

	has already been distributed and used	to withdraw from the study, data stored in the HostSeq Databank will be destroyed as of the time of your notification. However, it may be impossible to withdraw the results once they have been compiled with the results of others participating in the study or if they have been published. Furthermore, if some of the data have been shared with other researchers, it may not be possible to remove this part of the data. In these cases of total withdrawal being impossible, your identity will still be protected.
Re-identification	Low risk that the participant could be re-identified in the future	There is always a small risk that your data may lead to you being re-identified one day. Genetic information is unique to every person, just like a fingerprint. This means that theoretically you can be identified by your genetic code. However, this is not easy to do. As technology advances, there may be new ways of linking data back to you that we cannot foresee today, in spite of the strict security measures. The potential re-identification or unintentional release of your genetic and clinical research data could lead to loss of privacy and to possible future discrimination against you or your biological relatives.
Recontact (includes mature minors)	Option for recontact of participants (not mandatory, strongly recommended) It is strongly recommended that, where applicable, mature minors be included in the consent process and given the option for recontact as well.	I agree to be re-contacted to update my personal information, to obtain additional health information, or, to be invited to participate in new research projects: _____**Yes** _____ **initials** _____**No** _____**initials**
Assent	Option for obtaining assent of children, where applicable	ASSENT If you decide you want to be in this study, please print/write your name. If you decide that you don't want to be in the study, then all you have to do is tell me [insert name]. Assent of minor, capable of understanding the nature of the research (signature or print name) or obtain verbal assent.

No	Version	Date	Description
1	v.1.0	08 May 2020	-First draft circulated to the HostSeq Scientific Committee
2	v.2.0	15 May 2020	-Changed CGenSeq to Host Seq -Changes to aligned with consent template language
3	v.3.0	23 June 2020	-Addition of recontact option as a core element for prospective cohorts
4	v.4.0	27 July 2020	-Clarification of recontact option
5	v.5.0	26 August 2020	-Correction of mistakes and typos (future use of data – changed to “COVID-19 and other health outcomes” throughout)
6	v.6.0	17 November 2020	- Include options for obtaining assent for children and including mature minors in the consent process and recontact option