List of Data types received to date

Clinical / Phenotypic
Whole Genome Sequencing (WGS)
Transcriptomics Analysis (RNA-Seq)
Genotyping (GWAS)
Proteomics
Serology
PRNT50 and PRNT90 titers (NML)
Metabolomics

General acknowledgments

Whole Genome Sequencing (WGS)

Data Producer
Which lab/institution generated the data

McGill Genome Center

Experiment type
Small paragraph describing the experiment methods, the libraries preparation methods, etc.

Blood samples were shipped to the McGill Genome Centre in ACD tubes via Globex Courrier. The DNA extraction protocol is semi-automated. First, samples are treated with a lysis buffer in the Biosafety Level 2 (BSL2) room and then extracted using the CMG-1091 DNA extraction kit from Perkin Elmer on the Chemagic MSM-I instrument. DNA concentration is measured using picogreen. An aliquot of 25 ul of each sample at 16 ng/ul is prepared for sequencing using the PerkinElmet Janus liquid handler. A library is prepared for each sample with a unique index (code bar) with the Illumina DNA PCR-Free Prep, Tagmentation kit. The quality of each library is validated by pPCR on the LabChip and 27 libraries are mixed together at an equimolar amount. The pool of libraries is loaded into an Illumina S4 flowcell and sequenced on the Illumina NovaSeq 6000.

Experiment technology
Description of the instruments being used, and link to the manufacturer.

The [Chemagic MSM-I](#) is used for DNA extraction. Quantification of DNA is done with the [Quant-iT PicoGreen dsDNA Assay Kit](#) from ThermoFisher. Preparation of samples is done on th[e PerkinElmer Janus Workstation model AJM8001](#)(link provided to more recent models). The [Illumina DNA PCR-Free Prep, Tagmentation](#) is used for library preparations and quality control of libraries is done by qPCR on the [PerkinElmer LabChip GX](#) using the [DNA High Sensitivity Reagent Kit](#) from PerkinElmer. Sequencing is performed on the [NovaSeq 6000](#) using the [NovaSeq 6000 S4 Reagent Kit v1.5](#) (300 cycles).

Analysis
Small paragraph describing the way the raw data was analyzed, to produce the files being downloaded. This should also include all tools with their versions and parameters used.

Whole genome sequencing data was analyzed for variants using the GenPipes DnaSeq pipeline which implements an adaptation of the GATK Best Practices. Reads were aligned to the human reference genome (build GRCh38) using the BWA-mem aligner.62 Mapped reads were further refined using GATK and [Picard program](#) suites to improve mapping near insertions and deletions (indels; GATK IndelRealigner), mark duplicate reads (Picard MarkDuplicates) and to improve quality scores (GATK BaseRecalibrator). Single Nucleotide Variants were called using GATK HaplotypeCaller in gvcf mode to allow efficient downstream merging of multiple samples into one variant file. All the samples within a cohort were combined using GATK CombineGVCFs, and finally genotyped using GenotypeGVCFs.

Quality control steps accomplished

What was done to ensure the quality of the datasets provided.

- All samples should attain a mean coverage of 30x; top ups were applied if coverage was below
- Contamination within a sample was estimated using verifyBAMid2 and those samples above a FREEMIX value of 0.04 were flagged.
- Sample mix up was assessed through concordance between NGS and snp array as well as sex concordance between NGS vs snp array, NGS versus metadata contained both locally and in redcap database. Additional methods like NGScheckmate and GATK CrosscheckFingerprints were used if mix ups were suspected.
- The absence of each sample within the merged vcf was used to ensure all samples in the cohort had similar number of variant.

Accompanying metadata files

List here all files available in the Globus endpoint that provide additional information about the experiment.

NA

Acknowledgements

Describe how use of the data should be acknowledged in future publications. (Paper, research group, etc.)

McGill Genome Center MaGiC platform for the sequencing
C3G for the data analysis

Description of the data

Please include information on each of the file types that is provided in the dataset. You can repeat the table below multiple times if there is more than one file type. (e.g. instrument result files, and a metrics or metadata file)

Data type 1

File type (e.g. CSV, Excel, VCF, PLINK)

CRAM

What does the file type represent?

Aligned reads from the sequencing experiments

Is the data raw or processed(the result of an analysis/transformation of the data)? If processed, please describe the transformation.

Raw data. Raw FASTQs were aligned on the hg38 reference genome assembly to generate CRAM files.

How to open the data file to consult/analyze it?

The file can be opened using samtools.

Does the data need to be transformed in any way before release?

No.

Data type 2

File type (e.g. CSV, Excel, VCF, PLINK)

VCF

What does the file type represent?

Variant calls over the whole genome

Is the data raw or processed(the result of an analysis/transformation of the data)? If processed, please describe the transformation.

Processed data. Sequencing reads were aligned over the GRCH38 reference assembly and variants were called out of these results.

How to open the data file to consult/analyze it?

The file can be opened using bcftools.

Does the data need to be transformed in any way before release?

No.

Transcriptomics Analysis (RNA-Seq)

Data Producer
Which lab/institution generated the data

Centre d'expertise et de services Génome Québec

Experiment type
Small paragraph describing the experiment methods, the libraries preparation methods, etc.

A PAXgene RNA tube (BD) was collected by blood draw, inverted ten times, kept at room temperature for 24h, then transferred at -20C for another 24h before storage at -80C until RNA extraction. RNA was extracted manually. Manual RNA extraction was performed with PAXgene Blood RNA Kit - Qiagen according to manufacturer's instructions. Total RNA was quantified and its integrity assessed on a LabChip GXII (PerkinElmer) instrument. Libraries were generated from 250 ng of total RNA as follows: mRNA enrichment was performed using the NEBNext Poly(A) Magnetic Isolation Module (New England BioLabs). cDNA synthesis was achieved with the NEBNext RNA First Strand Synthesis and NEBNext Ultra Directional RNA Second Strand Synthesis Modules (New England BioLabs). The remaining steps of library preparation were done using the NEBNext Ultra II DNA Library Prep Kit for Illumina (New England BioLabs). Adapters and PCR primers were purchased from New England BioLabs. Libraries were quantified using the Kapa Illumina GA with Revised Primers-SYBR Fast Universal kit (Kapa Biosystems). Average size fragment was determined using a LabChip GXII (PerkinElmer) instrument. The libraries were normalized and pooled and then denatured in 0.05N NaOH and neutralized using the HT1 buffer. The pool was loaded at 225pM on a Illumina NovaSeq S4 lane as per the manufacturer's recommendations. The run was performed for 2x100 cycles (paired-end mode). A phiX library was used as a control and mixed with libraries at 1% level. Base calling was performed with RTA v3.4.4 . Program bcl2fastq2 v2.20 was then used to demultiplex samples and generate fastq reads.

Experiment technology
Description of the instruments being used, and link to the manufacturer.

Illumina NovaSeq, with S4 lanes.

Total RNA was quantified and its integrity assessed on a LabChip GXII (PerkinElmer) instrument. Libraries were generated from 250 ng of total RNA as follows: mRNA enrichment was performed using the NEBNext Poly(A) Magnetic Isolation Module (New England BioLabs). cDNA synthesis was achieved with the NEBNext RNA First Strand Synthesis and NEBNext Ultra Directional RNA Second Strand Synthesis Modules (New England BioLabs). The remaining steps of library preparation were done using the NEBNext Ultra II DNA Library Prep Kit for Illumina (New England BioLabs). Adapters and PCR primers were purchased from New England BioLabs. Libraries were quantified using the Kapa Illumina GA with Revised Primers-SYBR Fast Universal kit (Kapa Biosystems). Average size fragment was determined using a LabChip GXII (PerkinElmer) instrument. The libraries were normalized and pooled and then denatured in 0.05N NaOH and neutralized using the HT1 buffer. The pool was loaded at 225pM on a Illumina NovaSeq S4 lane as per the manufacturer's recommendations. The run was

performed for 2x100 cycles (paired-end mode). A phiX library was used as a control and mixed with libraries at 1% level.

Analysis

Small paragraph describing the way the raw data was analyzed, to produce the files being downloaded. This should also include all tools with their versions and parameters used.

Base calling was performed with RTA v3.4.4 . Program bcl2fastq2 v2.20 was then used to demultiplex samples and generate fastq reads.

Quality control steps accomplished

What was done to ensure the quality of the datasets provided.

The average base quality score for each sample dataset was verified to be Q33 or above and the % of aligned reads on reference sequence homo sapiens:hg19 was verified to be 90% or above.

Accompanying metadata files

List here all files available in the Globus endpoint that provide additional information about the experiment.

Metrics file in csv format

Acknowledgements

Describe how use of the data should be acknowledged in future publications. (Paper, research group, etc.)

The transcriptomics sequencing was performed at the Centre d'expertise et de services Génome Québec (Montreal, Quebec, Canada)

Description of the data

Please include information on each of the file types that is provided in the dataset. You can repeat the table below multiple times if there is more than one file type. (e.g. instrument result files, and a metrics or metadata file)

Data type 1

File type (e.g. CSV, Excel, VCF, PLINK)

Fastq. There are 2 files, one for Read 1 or Forward and Read 2 or Reverse

What does the file type represent?

FASTQ format is a text-based [format](#) for storing both a biological sequence (usually [nucleotide sequence](#)) and its corresponding quality scores. Both the sequence letter and quality score are each encoded with a single [ASCII](#) character for brevity. It is labeled: InstrumentType.Run#.Index#.SampleID_read#.fastq.gz

Is the data raw or processed(the result of an analysis/transformation of the data)? If processed, please describe the transformation.

You could consider it raw or processed depending on how you see it (it is not really possible to work with anything upstream of fastq, but of course there are many transformation of the raw signal, the description is in the analysis text

How to open the data file to consult/analyze it?

You need to decompress the gzip files. There are many programs that will then use these fastq files as input.

Does the data need to be transformed in any way before release?

No. Any transformation of data should be determined by the user.

Genotyping (GWAS)

Data Producer

Which lab/institution generated the data

Brent Richards lab, Jewish General Hospital

Experiment type

Small paragraph describing the experiment methods, the libraries preparation methods, etc.

Blood samples are shipped to the McGill Genome Centre in ACD tubes via Globex Courrier. The DNA extraction protocol is semi-automated. First, samples are treated with a lysis buffer in the Biosafety Level 2 (BSL2) room and then extracted using the CMG-1091 DNA extraction kit from Perkin Elmer on the Chemagic MSM-I instrument. DNA concentration is measured using picogreen. An aliquot of 20 ul of each sample at 20 ng/ul is prepared for genotyping using the PerkinElmet Janus liquid handler. Samples are genotyped using the Axiom Precision Medicine Research Array (PMRA) Kit, which includes genomic DNA preparation, target preparation of the samples on the Hamilton Nimbus and array processing on the GeneTitan.

Experiment technology

Description of the instruments being used, and link to the manufacturer.

The [Chemagic MSM-I](#) is used for DNA extraction. Quantification of DNA is done with the [Quant-iT PicoGreen dsDNA Assay Kit](#) from ThermoFisher. Preparation of samples is done on the [PerkinElmer Janus Workstation model AJM8001](#)(link provided to more recent models). Preparation of samples for genotyping is done on [Hamilton Nimbus](#) and the genotyping assays are performed on the [Affymetrix GeneTitan](#) using the [PMRA panel](#).

Analysis

Small paragraph describing the way the raw data was analyzed, to produce the files being downloaded. This should also include all tools with their versions and parameters used.

Initial data processing is done with the [Axiom Analysis software](#).

Quality control steps accomplished

What was done to ensure the quality of the datasets provided.

Marker quality control tests were performed on a subset of ancestrally homogeneous participants genotyped at 861,229 markers, who were determined via comparison to 2,504 individuals across 5 super populations from 1000 Genomes phase 3. We performed batch effect quality control, replicate discordance check, and finally variants with low allele frequency (MAF<0.001), low genotyping call rate (<98%), and with departure from Hardy-Weinberg equilibrium (HWE) (p value<$1x10^{-6}$) were removed.

For sample quality control, we used the genetic markers selected above and further filtered them using the following additional criteria: MAF>0.01, marker-wise missingness <0.01, single nucleotide substitutions with single character allele-codes (A, C, G, or T) (PLINK --snps-only 'just-acgt' option), and excluded variants within high linkage disequilibrium (LD) region. The above filtering resulted in a total of 407,158 genetic markers, which were used to determine low quality samples. We removed individuals with extreme values in heterozygosity and missingness. These outliers are defined as high genotype missing rate (overall missing genotype rate >0.03) or high/low heterozygosity rate on autosomal chromosomes. We then determined sex chromosome composition by estimating heterozygosity of X chromosome markers using PLINK --check-sex. We removed individuals with discordant self-reported gender and genetic sex. The samples passed sample QC were used for

genotype imputation. We performed genotype phasing and imputation on the BQC19 participants using the TOPMed reference panel at the University of Michigan Imputation Service.

Accompanying metadata files

List here all files available in the Globus endpoint that provide additional information about the experiment.
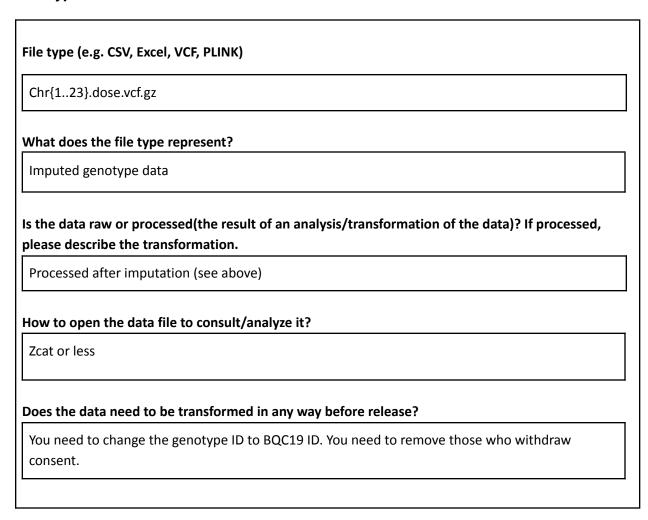
NA

Acknowledgements

Describe how use of the data should be acknowledged in future publications. (Paper, research group, etc.)

BQC19 imputation data was generated by Tomoko Nakanishi at Brent Richards lab, McGill.
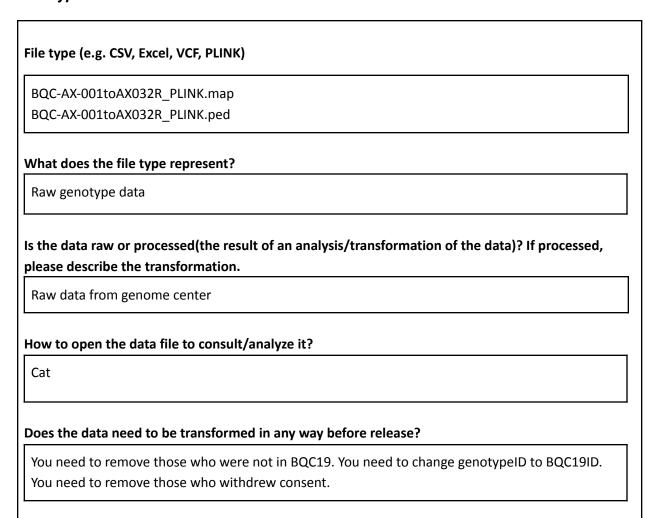See here for detailed codes used for generating the data.
https://github.com/richardslab/BQC19_genotype_pipeline

Description of the data

Please include information on each of the file types that is provided in the dataset. You can repeat the table below multiple times if there is more than one file type. (e.g. instrument result files, and a metrics or metadata file)

Data type 1

File type (e.g. CSV, Excel, VCF, PLINK)

Chr{1..23}.dose.vcf.gz

What does the file type represent?

Imputed genotype data

Is the data raw or processed(the result of an analysis/transformation of the data)? If processed, please describe the transformation.

Processed after imputation (see above)

How to open the data file to consult/analyze it?

Zcat or less

Does the data need to be transformed in any way before release?

You need to change the genotype ID to BQC19 ID. You need to remove those who withdraw consent.

Data type 2

File type (e.g. CSV, Excel, VCF, PLINK)

Chr{1..23}.info.gz

What does the file type represent?

Imputation quality file

Is the data raw or processed(the result of an analysis/transformation of the data)? If processed, please describe the transformation.

Processed after imputation (see above)

How to open the data file to consult/analyze it?

Zcat or less

Does the data need to be transformed in any way before release?

No this does not contain any individual-level data.

Data type 3

File type (e.g. CSV, Excel, VCF, PLINK)

BQC-AX-001toAX032R_PLINK.map
BQC-AX-001toAX032R_PLINK.ped

What does the file type represent?

Raw genotype data

Is the data raw or processed(the result of an analysis/transformation of the data)? If processed, please describe the transformation.

Raw data from genome center

How to open the data file to consult/analyze it?

Cat

Does the data need to be transformed in any way before release?

You need to remove those who were not in BQC19. You need to change genotypeID to BQC19ID. You need to remove those who withdrew consent.

Metabolomics

Data Producer

Metabolon for Brent Richards Lab

`

Experiment type

The goal of the study is to assess metabolomic biomarker signatures and profiles in a COVID Cohort.

Following receipt, samples were inventoried, immediately stored at -80°C(until processed) , each sample accessioned into the Metabolon LIMS system and assigned a unique identifier. Once protein removed, to recover chemically diverse metabolites, proteins were precipitated with methanol under vigorous shaking for 2 min (Glen Mills GenoGrinder 2000) followed by centrifugation. The resulting extract was divided into five fractions: two for analysis by two separate reverse phase (RP)/UPLC-MS/MS methods with positive ion mode electrospray ionization (ESI), one for analysis by RP/UPLC-MS/MS with negative ion mode ESI, one for analysis by HILIC/UPLC-MS/MS with negative ion mode ESI, and one sample was reserved for backup.

Experiment technology

Samples are prepared using the automated MicroLab STAR® system from Hamilton Company. All methods utilized a Waters ACQUITY ultra-performance liquid chromatography (UPLC) and a Thermo Scientific Q-Exactive high resolution/accurate mass spectrometer interfaced with a heated electrospray ionization (HESI-II) source and Orbitrap mass analyzer operated at 35,000 mass resolution.

The hardware and software foundations for the informatics system were the LAN backbone, and a database server running Oracle 10.2.0.1 Enterprise Edition. Raw data was extracted, peak-identified and QC processed using Metabolon's hardware and software built on a web-service platform utilizing Microsoft's .NET technologies, which run on high-performance application servers and fiber-channel storage arrays in clusters to provide active failover and load-balancing.

Analysis

Compounds are identified by comparison to library entries of purified standards or recurrent unknown entities. This library is maintained based on authenticated standards that contains the retention time/index (RI), mass to charge ratio (m/z), and chromatographic data (including MS/MS spectral data) on all molecules present in the library. Furthermore, biochemical identifications are based on three criteria: retention index within a narrow RI window of the proposed identification, accurate mass match to the library +/- 10 ppm, and the MS/MS forward and reverse scores between the experimental data and authentic standards. The MS/MS scores are based on a comparison of the ions present in the experimental spectrum to the ions present in the library spectrum. While there may be similarities between these molecules based on one of these factors, the use of all three data points can be utilized to distinguish and differentiate biochemicals.

Metabolon data analysts use proprietary visualization and interpretation software to confirm the consistency of peak identification among the various samples. For studies spanning multiple days, a data normalization step was performed to correct variation resulting from instrument inter-day tuning differences.

Quality control steps accomplished

A variety of curation procedures were carried out to ensure that a high-quality data set was made available for statistical analysis and data interpretation. Several types of controls were analyzed in concert with the experimental samples: a pooled matrix sample generated by taking a small volume of each experimental sample served as a technical replicate throughout the data set; extracted water samples served as process blanks; and a cocktail of QC standards that were carefully chosen not to interfere with the measurement of endogenous compounds were spiked into every analyzed sample, allowed instrument performance monitoring and aided chromatographic alignment

Accompanying metadata files

QC_report.docx (on Globus endpoint)

References:
https://doi.org/10.1093/jalm/jfz026

Acknowledgements

Describe how use of the data should be acknowledged in future publications. (Paper, research group, etc.)

Description of the data

Please include information on each of the file types that is provided in the dataset. You can repeat the table below multiple times if there is more than one file type. (e.g. instrument result files, and a metrics or metadata file)

Data type 1

File type (e.g. CSV, Excel, VCF, PLINK)

Excel

What does the file type represent?

The file type represents relative quantification data of metabolites detected in "plasma" for COVID cohort.

Is the data raw or processed(the result of an analysis/transformation of the data)? If processed, please describe the transformation.

The data provided includes raw data (area under the curve), batch normalized, and batch-normalized imputed data. They are organized in separate tabs in the document depending on the type of data.

How to open the data file to consult/analyze it?

These are normal excel files and can be assessed using Microsoft Excel software.

Does the data need to be transformed in any way before release?

No. Any transformation of data should be determined by the user.

Proteomics(SomaLogic data)

Data Producer

Which lab/institution generated the data

SomaLogic

Experiment type

Small paragraph describing the experiment methods, the libraries preparation methods, etc.

SomaScan Assay v4 was used for quantifying serum protein levels of 4,701 human proteins, a description of which can be found on the SomaLogic website and in our pre-print: https://www.medrxiv.org/content/10.1101/2021.10.04.21264015v1

Experiment technology

Description of the instruments being used, and link to the manufacturer.

Aptamer based proteomics technology (SomaScan): https://somalogic.com/technology/

Analysis

Small paragraph describing the way the raw data was analyzed, to produce the files being downloaded. This should also include all tools with their versions and parameters used. When available, provide the citation describing the method

Protein levels were quantified in relative fluorescence units (RFUs).
Normalization and calibration steps were performed by SomaLogic. The normalization procedure involved three steps: hybridization control normalization, intraplate median signal normalization, and plate scaling and calibration.
A second file for each batch contains data with one more normalization step: median signal normalization to an internal reference in each batch.

Further details can be found in:
https://www.mcgill.ca/genepi/files/genepi/bqc19_jgh_prt_tech_note_0.pdf

Quality control steps accomplished

What was done to ensure the quality of the datasets provided.

No further quality controls were applied to the data

Accompanying metadata files

List here all files available in the Globus endpoint that provide additional information about the experiment.

NA

Acknowledgements

Describe how use of the data should be acknowledged in future publications. (Paper, research group, etc.)

References:
https://www.nature.com/articles/s41591-021-01281-1
https://www.medrxiv.org/content/10.1101/2021.10.04.21264015v1

Description of the data

Please include information on each of the file types that is provided in the dataset. You can repeat the table below multiple times if there is more than one file type. (e.g. instrument result files, and a metrics or metadata file)

Data type 1

File type (e.g. CSV, Excel, VCF, PLINK)

.adat

What does the file type represent?

It is a SomaLogic proprietary data file, which can be converted to .CSV by the package listed below

Is the data raw or processed(the result of an analysis/transformation of the data)? If processed, please describe the transformation.

Raw, directly received from SomaLogic

How to open the data file to consult/analyse it?

This file needs to be opened by SomaDataIO R package:
https://github.com/SomaLogic/SomaDataIO

Does the data need to be transformed in any way before release?

No. Any transformation of data should be determined by the user.

Serology

Data Producer

Which lab/institution generated the data

Finzi Lab

Experiment type

Small paragraph describing the experiment methods, the libraries preparation methods, etc.

ELISA assays were performed as described: https://pubmed.ncbi.nlm.nih.gov/33067385/
Cell-based ELISA assays were performed as described: https://pubmed.ncbi.nlm.nih.gov/33969322/.

Experiment technology

Description of the instruments being used, and link to the manufacturer.

ELISA. Recombinant SARS-CoV-2 S RBD proteins (2.5 µg/ml), or bovine serum albumin (BSA) (2.5 µg/ml) as a negative control, were prepared in PBS and were adsorbed to plates (MaxiSorp Nunc) overnight at 4°C. Coated wells were subsequently blocked with blocking buffer (Tris-buffered saline [TBS] containing 0.1% Tween20 and 2% BSA) for 1h at room temperature. Wells were then washed four times with washing buffer (Tris-buffered saline [TBS] containing 0.1% Tween20). CR3022 mAb (50ng/ml) or plasma from SARS-CoV-2-infected or uninfected donors (1/250 dilutions) were prepared in a diluted solution of blocking buffer (0.1 % BSA) and incubated with the RBD-coated wells for 90 minutes at room temperature. Plates were washed four times with washing buffer followed by incubation with secondary Abs (diluted in a diluted solution of blocking buffer (0.4% BSA)) for 1h at room temperature, followed by four washes. HRP enzyme activity was determined after the addition of a 1:1 mix of Western Lightning oxidizing and luminol reagents (Perkin Elmer Life Sciences). Light emission was measured with a LB941 TriStar luminometer (Berthold Technologies). Signal obtained with BSA was subtracted for each plasma and was then normalized to the signal obtained with CR3022 mAb present in each plate. The seropositivity threshold was established using the following formula: mean of all COVID-19 negative plasma + (3 standard deviation of the mean of all COVID-19 negative plasma).

Cell-based ELISA. Detection of the trimeric SARS-CoV-2 Spike at the surface of HOS cells was performed by cell-based enzyme-linked immunosorbent assay (ELISA). Briefly, parental HOS cells or HOS-Spike cells were seeded in 384-well plates (2.8×104 cells per well) overnight. Cells were blocked with blocking buffer (washing buffer [1.8 mM CaCl2, 1 mM MgCl2, 25 mM Tris (pH 7.5), and 140 mM NaCl] supplemented with 10mg/mL non-fat dry milk and 5mM Tris [pH 8.0]) for 30min. CR3022 mAb (1 µg/ml) or plasma from SARS-CoV-2-infected or uninfected donors (at a dilution of 1/250) were prepared in blocking buffer and incubated with the cells for 1h at room temperature. Respective HRP-conjugated secondary antibodies were then incubated with the samples for 45 min at room

temperature. The binding of CR3022 IgG was quantified with HRP-conjugated antibodies specific for the Fc region of human IgG. For all conditions, cells were washed 6 times with blocking buffer and 6 times with washing buffer. HRP enzyme activity was determined after the addition of a 1:1 mix of Western Lightning oxidizing and luminol reagents (PerkinElmer Life Sciences). Light emission was measured with an LB 942 TriStar luminometer (Berthold Technologies). Signal obtained with parental HOS was subtracted for each plasma and was then normalized to the signal obtained with CR3022 mAb present in each plate. The seropositivity threshold was established using the following formula: mean of all COVID-19 negative plasma + (3 standard deviation of the mean of all COVID-19 negative plasma).

Analysis

Small paragraph describing the way the raw data was analyzed, to produce the files being downloaded. This should also include all tools with their versions and parameters used. When available, provide the citation describing the method

NA. The values recorded were results of the experiment done on the BQC19 samples based on the method described above. No analysis was done on the data before or after to reach the result. In addition, flow cytometry or ADCC are out of scope for BQC19 samples.

Quality control steps accomplished

What was done to ensure the quality of the datasets provided.

Quality check was done internally at the Finzi lab and equivocal results (such as positive for RBD but negative for full Spike, removed)

Accompanying metadata files

List here all files available in the Globus endpoint that provide additional information about the experiment.

NA

Acknowledgements

Describe how use of the data should be acknowledged in future publications. (Paper, research group, etc.)

Description of the data

Please include information on each of the file types that is provided in the dataset. You can repeat the table below multiple times if there is more than one file type. (e.g. instrument result files, and a metrics or metadata file)

Data type 1

File type (e.g. CSV, Excel, VCF, PLINK)

Excel

What does the file type represent?

Antibody measures

Is the data raw or processed(the result of an analysis/transformation of the data)? If processed, please describe the transformation.

Results were normalized to the internal control present in each plate (CR3022) as described https://pubmed.ncbi.nlm.nih.gov/33969322/

How to open the data file to consult/analyze it?

The data can be opened using Microsoft Excel software

Does the data need to be transformed in any way before release?

No. Any transformation of data should be determined by the user.

PRNT50 and PRNT90 titers (NML)

Data Producer

Which lab/institution generated the data

National Microbiology Laboratory, Public Health Agency of Canada

Experiment type

Small paragraph describing the experiment methods, the libraries preparation methods, etc.

The Plaque Reduction Neutralization Test (PRNT) is used to quantify the titer of neutralizing antibody against SARS-CoV-2 in vitro. The SARS-CoV-2 PRNT is conducted in six key stages: (1) cell culture, (2) serum dilution, (3) serum-virus incubation (neutralization), (4) inoculation, (5) visualization, and (6) plaque quantification and interpretation. Briefly, serial dilutions of serum samples are incubated with infectious SARS-CoV-2 to allow serum antibodies to interact with the virus prior to inoculation of a susceptible VeroE6 cell monolayer. SARS-CoV-2 neutralizing antibodies present in the sample bind to the virus, preventing SARS-CoV-2 from infecting the monolayer. In the absence of neutralizing antibodies, SARS-CoV-2 is able to infect the monolayer and cause cytopathic effect (CPE), which can be visualized as plaques using crystal violet dye. The neutralizing antibody titer is measured as the highest dilution of each serum sample resulting in a ≥50% (PRNT50) and ≥90% (PRNT90) reduction in the number of plaques compared to the back-titration control.

Experiment technology

Description of the instruments being used, and link to the manufacturer.

This is an in-house assay that is performed manually in containment level 3.

Analysis

Small paragraph describing the way the raw data was analyzed, to produce the files being downloaded. This should also include all tools with their versions and parameters used. When available, provide the citation describing the method

Citation describing the SARS-CoV-2 PRNT can be accessed here
-https://journals.asm.org/doi/10.1128/Spectrum.00886-21

Quality control steps accomplished

What was done to ensure the quality of the datasets provided.

All SARS-CoV-2 PRNTs include two control antibodies (two monoclonal antibodies that neutralize the ancestral SARS-CoV-2 virus) as well as the virus back-titration which is performed in duplicate. Validation of the PRNT was completed and published in Microbiology Spectrum (see attachment).

Accompanying metadata files

List here all files available in the Globus endpoint that provide additional information about the experiment.

NA

Acknowledgements

Describe how use of the data should be acknowledged in future publications. (Paper, research group, etc.)

The National Microbiology Laboratory should be acknowledged in any future publications.

Description of the data

Please include information on each of the file types that is provided in the dataset. You can repeat the table below multiple times if there is more than one file type. (e.g. instrument result files, and a metrics or metadata file)

Data type 1

File type (e.g. CSV, Excel, VCF, PLINK)

Excel

What does the file type represent?

PRNT50 and PRNT90 titer results

Is the data raw or processed(the result of an analysis/transformation of the data)? If processed, please describe the transformation.

Plaque counts are used to determine the PRNT50 and PRNT90 titers. The neutralizing antibody titer is highest dilution of each serum sample resulting in a ≥50% (PRNT50) and ≥90% (PRNT90) reduction in number of plaques compared to the back-titration control.

How to open the data file to consult/analyze it?

Not applicable

Does the data need to be transformed in any way before release?

No