CONFERENCE CAHR 2021

CONGRÈS DE L'ACRV 2021

An Inter-Laboratory Genomic Cross-Validation of a COVID-19 Outbreak in a Long-Term Care Facility

Hope R Lapointe[1], Gordon Ritchie[2,3], Winnie Dong[1], Kimia Kamelian[3], Shannon Russell[3], Jihyun Choi[4], Nancy Matic[2,3], Don Kirkby[1], Samuel D Chorlton[3], Mel Krajden[3,5], Julio SG Montaner[1,6], Victor Leung[3,6], Marc Romney[2,3], Zabrina L. Brumme[1,7], Natalie Prystajecky[3,5*], Christopher F Lowe[2,3*], Chanson J. Brumme[1,6*]

[1] BC Centre for Excellence in HIV/AIDS, Vancouver, British Columbia, Canada
[2] Division of Medical Microbiology and Virology, St. Paul's Hospital, Vancouver, British Columbia, Canada
[3] Department of Pathology and Laboratory Medicine, University of British Columbia, Vancouver, British Columbia, Canada
[4] Department of Medicine, Division of Social Medicine, University of British Columbia, Vancouver, British Columbia, Canada
[5] Public Health Laboratory, British Columbia Centre for Disease Control, Vancouver, British Columbia, Canada
[6] Department of Medicine, Division of Infectious Diseases, University of British Columbia, Vancouver, British Columbia, Canada
[7] Faculty of Health Sciences, Simon Fraser University, Burnaby, British Columbia, Canada
*equal contribution

Poster No.: 249 (Basic Science)

Correspondence: hlapointe@bccfe.ca

BRITISH COLUMBIA CENTRE *for* EXCELLENCE *in* HIV/AIDS

BC Centre for Disease Control

SFU FACULTY OF HEALTH SCIENCES

UBC

Providence HEALTH CARE How you want to be treated.

BRITISH COLUMBIA Ministry of Health

Background

- Between June 9 and August 31 2020, a long-term care facility in British Columbia experienced a prolonged COVID-19 outbreak.
- We describe our inter-laboratory effort to cross-validate whole-genome SARS-CoV-2 sequencing and bioinformatic methods and characterize outbreak viral dynamics.

Methods

- Nucleic acid extracts from confirmed SARS-CoV-2 positive diagnostic specimens linked to the outbreak were sent to at least one of three independent laboratories for whole-genome sequencing:
  - BC Centre for Excellence in HIV/AIDS (BCCFE)
  - St. Paul's Hospital – Virology (SPHV)
  - BC Centre for Disease Control/UBC Harrigan Lab (BCCDC)
- For each specimen with sufficient material, at least two labs attempted sequencing, with discrepancies resolved by a third.
- Amplicons were generated using the ARTIC protocol version 3 (BCCDC), or an in-house, adapted version of the ARTIC protocol (SPHV, BCCFE).
- Amplicons were sequenced using **Illumina MiSeq** (BCCFE) or **Oxford Nanopore MinION** platforms (SPHV/BCCDC).
- Raw sequence data were processed using the MiCall (BCCFE - MiSeq), BugSeq (SPHV - MinION), or ARTIC (BCCDC - MinION) bioinformatic pipelines
- Consensus sequences representing the majority nucleotide at each position were compared for concordance, ignoring any positions with missing data. Single nucleotide polymorphisms were identified relative to the presumed outbreak founder virus.
- Phylogenetic trees were generated and visualized in Nextstrain using default settings, with sub-sampling disabled. SARS-CoV-2 sequences originating from BC laboratories up to August 31st 2020 were extracted from GISAID. Lineages were assigned using Pangolin (v.2.3.9).
- Sequences with <80% breadth of coverage (i.e. genome completeness) were excluded from phylogenetic analysis. However, all sequences were retained for nucleotide analysis regardless of coverage.

BRITISH COLUMBIA CENTRE for EXCELLENCE in HIV/AIDS

Results

Providence HEALTH CARE How you want to be treated. | BRITISH COLUMBIA | Ministry of Health

- Eighty-nine individuals had presumed epidemiological links to the outbreak (Table 1).
  - Long-term care residents: N=53
  - Long-term care staff: N=31
  - Other putatively-linked infections: N=5
- Sequencing success was highly dependent on SARS-CoV-2 virus loads
  - Sequencing success was low (25%) in samples with C_t ≥30
- In total, 65 (73.0%), 54 (60.7%) and 25 (28.1%) samples were sequenced in ≥1, ≥2 or all three labs respectively (Figure 1).
- Non-identical sequences differed by a median of 1 [Q1-Q3: 1-2] nucleotides (Figure 2).

Table 1: Sequencing success stratified by cycle threshold

	C_t value		
	C_t < 30	C_t ≥ 30	All
Count	65	24	89
Sequenced	59 (90.7%)	6 (25%)	65 (73.0%)
BCCFE	53 (81.5%)	0	53 (59.6%)
SPHV	46 (70.8%)	2 (8.3%)	48 (53.9%)
BCCDC	39 (60.0%)	4 (16.7%)	43 (48.3%)

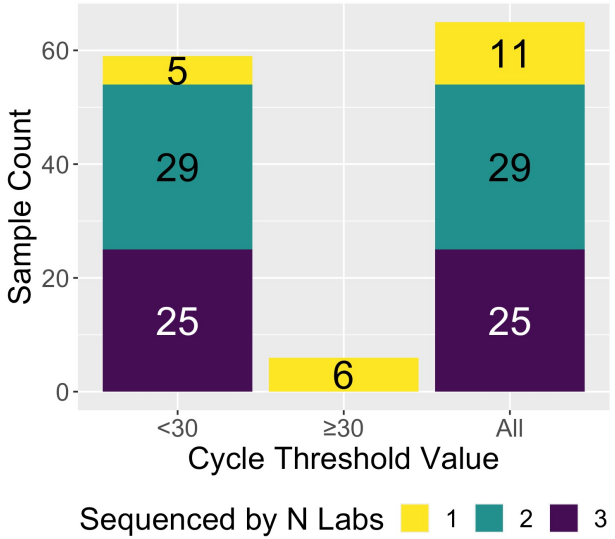


Figure 1: Available sequences generated by independent laboratories per sample, according to cycle threshold value

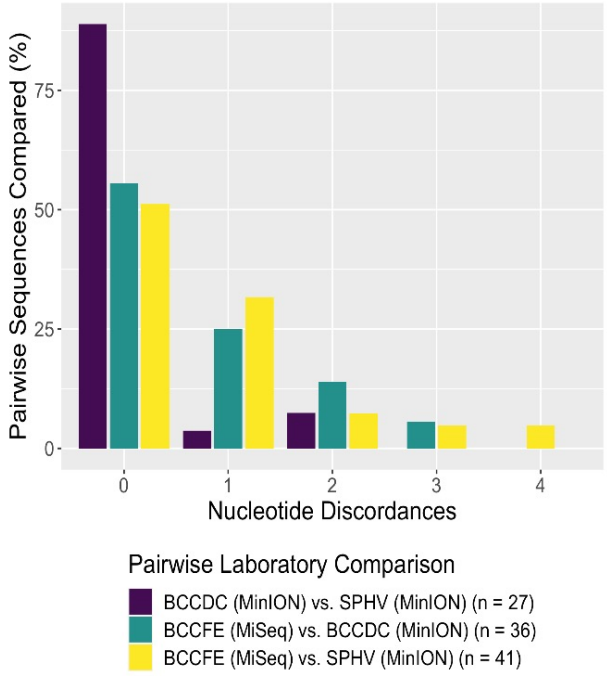


Figure 2: Number of discordant nucleotides observed in pairwise comparisons of consensus sequences

Results

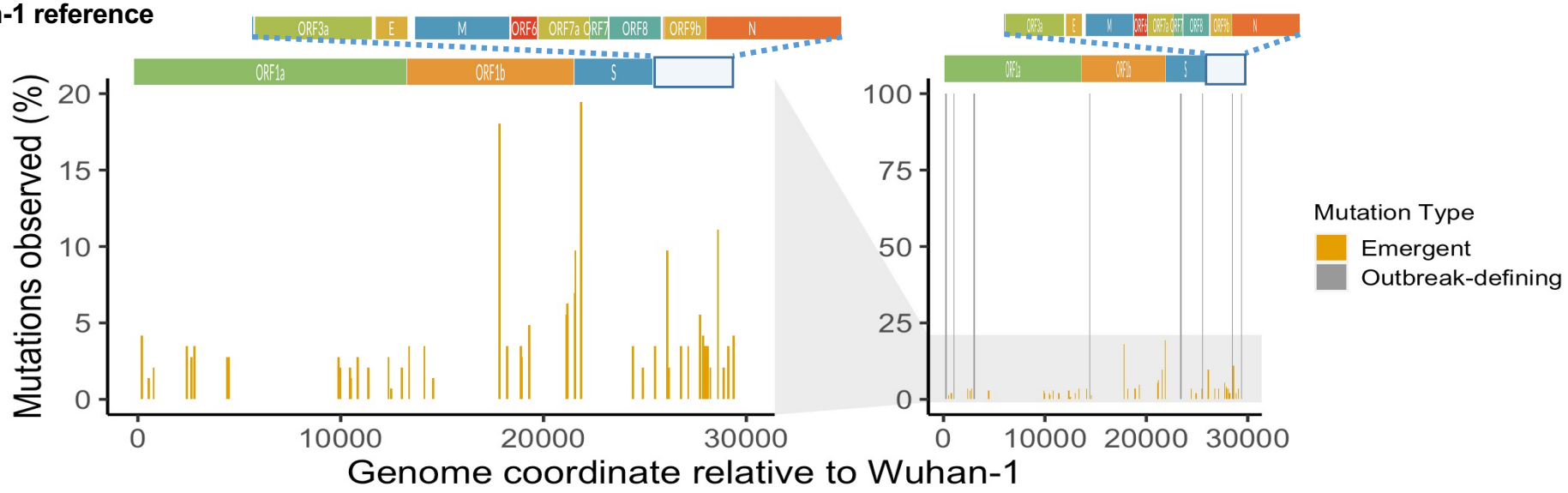
- Pairwise nucleotide concordance between labs was ≥99.99% (Table 2).
  - Overall, 62.5% (n=65/104) sequences compared were identical.
- Greater inter-lab concordance was observed for sequences collected on the same sequencing platform ($p<0.01$).
- Outbreak sequences shared eight nucleotide substitutions relative to the Wuhan-1 reference.
  - C241T, C1059T, C3037T, C14408T, A23403G, G25563T, C28500T, C29364T
- Additional nucleotide substitutions throughout the genome were selected - and in some cases transmitted - during the outbreak (Figure 3).

Table 2: Sequence concordance observed in pairwise inter-laboratory comparisons

	Sequencing laboratory		
	BCCFE (MiSeq)	BCCDC (MinION)	SPHV (MinION)
BCCFE (MiSeq)	-	20/36 (55.6%)	21/41 (51.2%)
BCCDC (MinION)	1,076,475/ 1,076,508 (99.9%)	-	24/27 (88.9%)
SPHV (MinION)	1,225,982/ 1,226,023 (99.9%)	807,376/ 807,381 (99.9%)	-

*Values above the diagonal (in green) represent the number of pairwise comparisons between laboratories that yielded identical sequences. Values below the diagonal (in blue) represent the total number of identical nucleotides across all samples and genome positions compared.

Figure 3: Nucleotide positions with variation relative to the Wuhan-1 reference



BRITISH COLUMBIA CENTRE for EXCELLENCE in HIV/AIDS

Outbreak Investigation

- Three samples collected early in the outbreak yielded identical sequences; this was presumed to be the outbreak founder virus.
- Outbreak sequences belonged to the B.1 lineage
- Subsequent samples collected over a two month period acquired up to four mutations (median 1; Q1-Q3: 0-1.5) relative to the presumed founder.
- Phylogenetic analysis confirmed the outbreak as a single cluster within the larger provincial epidemic, consistent with a single source (Figure 4).
- A total of 50 unique nucleotide substitutions (relative to the founder) were observed across all outbreak sequences.

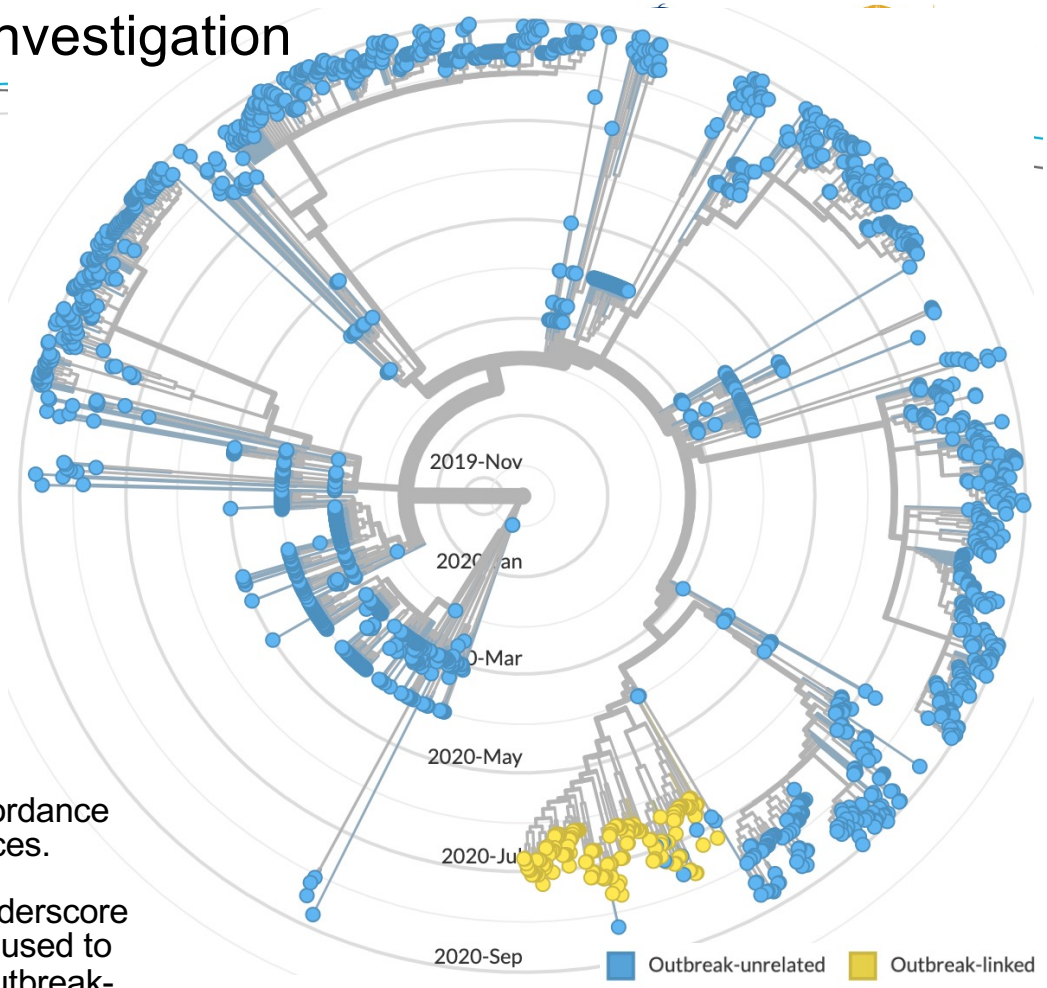


Figure 4: Time-scaled phylogeny of SARS-CoV-2 sequences from British Columbia until August 31st, 2020.

Conclusion & Future Directions

- Inter-laboratory whole-genome SARS-CoV-2 sequence concordance was high despite sequencing/bioinformatics platform differences.
- Minor sequence discrepancies between labs nevertheless underscore the importance of laboratory cross-validation if sequencing is used to characterize emerging variants or to classify sequences as outbreak-related, as determination of genetic relatedness for SARS-CoV-2 can be influenced by as few as one nucleotide polymorphism.
- Sequence discordances identified between laboratories require further investigation to assess intra-host sequence variability, potential sources of amplification/sequencing error, and implications of using different bioinformatic approaches.
- Phylogenetic analysis suggests a prolonged single-source outbreak with subsequent diversification. Further phylogenetic investigation will assess evolutionary rates within this outbreak in the context of BC's provincial epidemic.

We acknowledge and thank the BCCDC Public Health Laboratory and associated authors as the GISAID Originating and Submitting Laboratory for the outbreak-unrelated SARS-CoV-2 genome sequences from BC.

BRITISH COLUMBIA CENTRE for EXCELLENCE in HIV/AIDS

BC Centre for Disease Control

SFU FACULTY OF HEALTH SCIENCES