hCoV-2019/nCoV-2019 Version 3 Amplicon Set

The ARTIC project released a protocol for sequencing hCoV-2019 on January 22nd (https://www.protocols.io/view/ncov-2019-sequencing-protocol-bbmuik6w). This protocol has been rapidly adopted by labs worldwide and resulted in a large number of nearly complete genomes deposited into GISAID.

In testing the protocol, several groups noticed a systematic dropping out of amplicons 18 and 76 with this protocol, resulting in no more than 98% genome coverage being achievable on the V1 protocol. Whilst investigating the cause of this, Itokawa *et al.* (https://www.biorxiv.org/content/10.1101/2020.03.10.985150v1.full) noted that nCoV-2019_18_LEFT and nCoV-2019_76_RIGHT might form a dimer, and suggested an alternative called 2019_76_RIGHTv2. We incorporated this suggestion in our V2 protocol substituting nCoV-2019_18_LEFT for nCoV-2019_18_LEFT_alt2. The reason we chose to substitute nCoV-2019_18_LEFT instead was because of the availability of a high scoring alternative from the original primer3 output, however users quickly reported that this change caused other amplicons to drop out instead resulting in a sort of *amplicon whack-a-mole!* This illustrates the unpredictable interactions between primers within a multiplex PCR reaction.

To address this issue, we have developed a V3 protocol which tries to improve coverage in problematic amplicons by addition of alternative primers (*alts*). To achieve this we extracted all high scoring primers for regions 7, 9, 14, 15, 18, 21, 44, 45, 46, 76 and 89 from the primer3 output. Of these we selected the alternative with the furthest coordinates from the original primer (to avoid very similar primers). Rather than substituting these both are added in an equimolar fashion to the V1 scheme which generates four possible primer combinations for each region with alternates. This dramatically improves the performance of these weaker regions and requires only 100,000 reads per sample to achieve 50x minimum coverage across all regions. This change does require an analysis pipeline change to correctly trim the data at the outermost coordinates of the primer set for each region. We are hugely grateful to John Tyson at University of British Columbia for his efforts testing these alternate primers, working with Richard Harrigan and his team at the British Columbia Centers for Disease Control. We are also grateful to the groups around the world for providing feedback and data to help us refine the amplicon scheme.

While we are satisfied with the current performance of the scheme, accumulation of genetic diversity in hCoV-2019 over time may necessitate further changes. We are also investigating a proposal to recover the 3’ and 5’ untranslated regions.

Generating the primer set

If you do not have any primers, the full V3 primer list is accessible via the Github repository: https://github.com/artic-network/artic-ncov2019/blob/master/primer_schemes/nCoV-2019/V3/nCoV-2019.tsv

If you already have V1 or V2 primer sets, you can upgrade these to V3 by generating 100μM pools of the alternates for both pools, diluting them 10μM then adding 0.45μL (pool A) and 0.38μL (pool B) to the respective PCR reactions in place of nuclease-free water.

Name	Pool	Sequence
nCoV-2019_7_LEFT_alt0	nCoV-2019_1	CATTTGCATCAGAGGCTGCTCG
nCoV-2019_7_RIGHT_alt5	nCoV-2019_1	AGGTGACAATTTGTCCACCGAC
nCoV-2019_9_LEFT_alt4	nCoV-2019_1	TTCCCACAGAAGTGTTAACAGAGG
nCoV-2019_9_RIGHT_alt2	nCoV-2019_1	GACAGCATCTGCCACAACACAG
nCoV-2019_15_LEFT_alt1	nCoV-2019_1	AGTGCTTAAAAAGTGTAAAAGTGCCT
nCoV-2019_15_RIGHT_alt3	nCoV-2019_1	ACTGTAGCTGGCACTTTGAGAGA
nCoV-2019_21_LEFT_alt2	nCoV-2019_1	GGCTATTGATTATAAACACTACACACCCT
nCoV-2019_21_RIGHT_alt0	nCoV-2019_1	GATCTGTGTGGCCAACCTCTTC
nCoV-2019_45_LEFT_alt2	nCoV-2019_1	AGTATGTACAAATACCTACAACTTGTGCT
nCoV-2019_45_RIGHT_alt7	nCoV-2019_1	TTCATGTTGGTAGTTAGAGAAAGTGTGTC
nCoV-2019_89_LEFT_alt2	nCoV-2019_1	CGCGTTCCATGTGGTCATTCAA
nCoV-2019_89_RIGHT_alt4	nCoV-2019_1	ACGAGATGAAACATCTGTTGTCACT
nCoV-2019_14_LEFT_alt4	nCoV-2019_2	TGGCAATCTTCATCCAGATTCTGC
nCoV-2019_14_RIGHT_alt2	nCoV-2019_2	TGCGTGTTTCTTCTGCATGTGC
nCoV-2019_18_LEFT_alt2	nCoV-2019_2	ACTTCTATTAAATGGGCAGATAACAACTGT
nCoV-2019_18_RIGHT_alt1	nCoV-2019_2	GCTTGTTTACCACACGTACAAGG
nCoV-2019_44_LEFT_alt3	nCoV-2019_2	CCACAGTACGTCTACAAGCTGG
nCoV-2019_44_RIGHT_alt0	nCoV-2019_2	CGCAGACGGTACAGACTGTGTT
nCoV-2019_46_LEFT_alt1	nCoV-2019_2	CGCTTCCAAGAAAAGGACGAAGA
nCoV-2019_46_RIGHT_alt2	nCoV-2019_2	CACGTTCACCTAAGTTGGCGTAT
nCoV-2019_76_LEFT_alt3	nCoV-2019_2	GGGCAAACTGGAAAGATTGCTGA
nCoV-2019_76_RIGHT_alt0	nCoV-2019_2	ACCTGTGCCTGTTAAACCATTGA

Table 1: Primer sequences to upgrade a V1 pool to V3.

Name	Pool	Sequence
nCoV-2019_7_LEFT_alt0	nCoV-2019_1	CATTTGCATCAGAGGCTGCTCG
nCoV-2019_7_RIGHT_alt5	nCoV-2019_1	AGGTGACAATTTGTCCACCGAC
nCoV-2019_9_LEFT_alt4	nCoV-2019_1	TTCCCACAGAAGTGTTAACAGAGG
nCoV-2019_9_RIGHT_alt2	nCoV-2019_1	GACAGCATCTGCCACAACACAG
nCoV-2019_15_LEFT_alt1	nCoV-2019_1	AGTGCTTAAAAAGTGTAAAAGTGCCT
nCoV-2019_15_RIGHT_alt3	nCoV-2019_1	ACTGTAGCTGGCACTTTGAGAGA
nCoV-2019_21_LEFT_alt2	nCoV-2019_1	GGCTATTGATTATAAACACTACACACCCT
nCoV-2019_21_RIGHT_alt0	nCoV-2019_1	GATCTGTGTGGCCAACCTCTTC
nCoV-2019_45_LEFT_alt2	nCoV-2019_1	AGTATGTACAAATACCTACAACTTGTGCT
nCoV-2019_45_RIGHT_alt7	nCoV-2019_1	TTCATGTTGGTAGTTAGAGAAAGTGTGTC
nCoV-2019_89_LEFT_alt2	nCoV-2019_1	CGCGTTCCATGTGGTCATTCAA
nCoV-2019_89_RIGHT_alt4	nCoV-2019_1	ACGAGATGAAACATCTGTTGTCACT
nCoV-2019_14_LEFT_alt4	nCoV-2019_2	TGGCAATCTTCATCCAGATTCTGC
nCoV-2019_14_RIGHT_alt2	nCoV-2019_2	TGCGTGTTTCTTCTGCATGTGC
nCoV-2019_18_LEFT	nCoV-2019_2	TGGAAATACCCACAAGTTAATGGTTTAAC
nCoV-2019_18_RIGHT_alt1	nCoV-2019_2	GCTTGTTTACCACACGTACAAGG
nCoV-2019_44_LEFT_alt3	nCoV-2019_2	CCACAGTACGTCTACAAGCTGG
nCoV-2019_44_RIGHT_alt0	nCoV-2019_2	CGCAGACGGTACAGACTGTGTT
nCoV-2019_46_LEFT_alt1	nCoV-2019_2	CGCTTCCAAGAAAAGGACGAAGA
nCoV-2019_46_RIGHT_alt2	nCoV-2019_2	CACGTTCACCTAAGTTGGCGTAT
nCoV-2019_76_LEFT_alt3	nCoV-2019_2	GGGCAAACTGGAAAGATTGCTGA
nCoV-2019_76_RIGHT_alt0	nCoV-2019_2	ACCTGTGCCTGTTAAACCATTGA

Table 2: Primer sequences to upgrade a V2 pool to a V3

Expected results

On a test sample, with 100,000 reads randomly subsamples per barcode, we obtained the following results on a typical clinical sample:

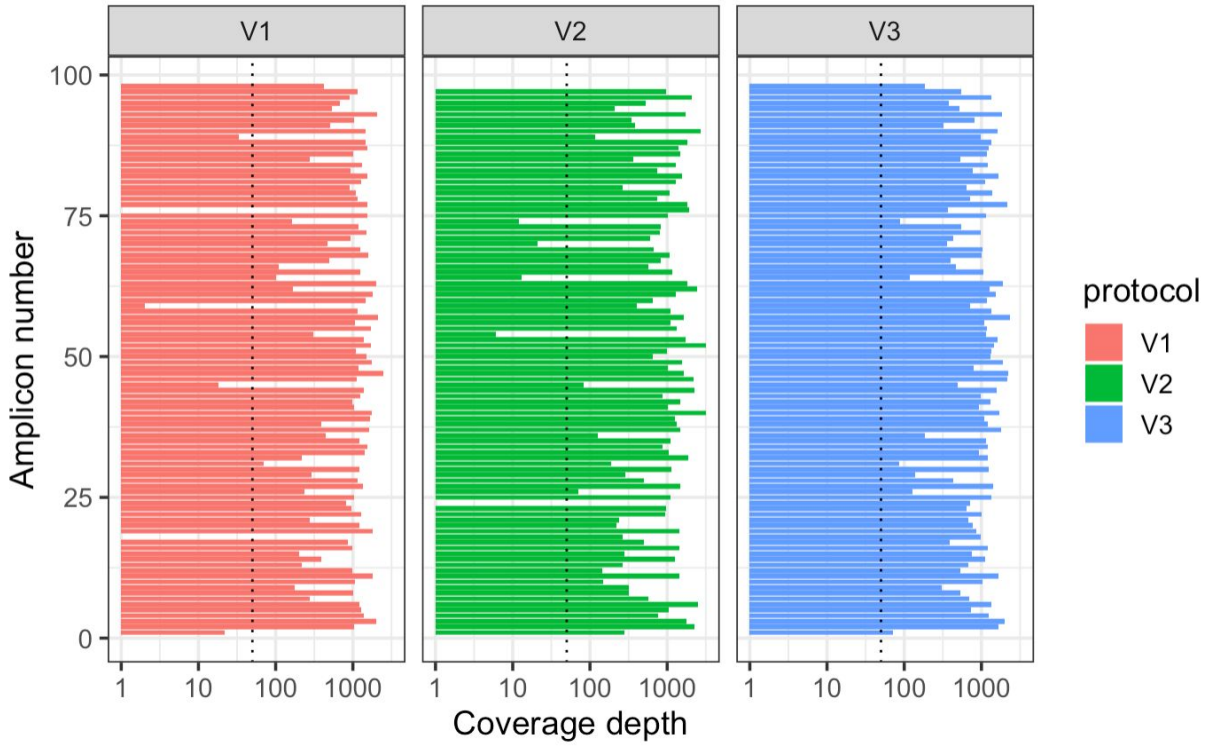

Figure 1: Comparison of V1, V2 and V3 protocols run on the same clinical sample with a barcoded library. For comparison, 100,000 randomly subsampled reads were used. As a guide to sufficient coverage, the vertical dotted line indicates 50x coverage.

Future supply

Currently oligonucleotide synthesis facilities such as IDT are working at high capacity due to high demand for qPCR reagents and orders may take longer than usual. We are organising a bulk production of these V3 oligos via a manufacturer which will allow the supply of 100µM pools ready to use. This has the benefits of reducing duplicate small-scale synthesis runs and mean you won't have to pool the primers yourself. It is also becoming more difficult for us to orchestrate DHL shipments from Birmingham as collection services have now stopped. We expect these to be available over the coming weeks. If you need to order these oligos now please consider joining up with other people also doing ARTIC protocol sequencing to reduce the number of orders being placed and the burden on overstretched manufacturing capacity during the COVID-19 pandemic.

Josh Quick and Nick Loman on behalf of the ARTIC Consortium

24th March 2020