

Supplementary Materials

Rapid viral metagenomics using SMART-9N amplification and nanopore sequencing

Ingra M. Claro, Mariana S. Ramundo, Thais M. Coletti, Camila A. M. da Silva, Ian N. Valenca, Darlan S. Candido Flavia C. S. Sales, Erika R. Manuli, Jaqueline G. de Jesus, Anderson de Paula, Alvina Clara Felix, Pamela dos Santos Andrade, Mariana C. Pinho, William M. de Souza, Mariene R. Amorim, José Luiz Proenca-Modena, Esper G. Kallas, Jose Eduardo Levi, Nuno Rodrigues Faria, Ester C. Sabino, Nicholas J. Loman, Joshua Quick.

Correspondence to: j.quick@bham.ac.uk and n.j.loman@bham.ac.uk

This file includes:

Tables S1 to S9
Figure S1 and S2

Table S1 - Description of samples collected and protocol realized to each sample.

Sample	Multiplex PCR	SMART-9N	Rapid SMART-9N
ZIKV reference	X	X	X
YFV clinical sample	X	X	X
SARS-CoV-2 clinical sample	X		X

Legend: ZIKV = Zika virus;
YFV = Yellow fever virus;
SARS-CoV-2 = Severe acute respiratory syndrome coronavirus 2.

Table S2 - Description of samples positive for Zika virus reference sample strain BeH815744 (ZIKV) (n=1), yellow fever virus (YFV) (n=41), and Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) (n=10) by real-time quantitative reverse transcription PCR with the corresponding sample types, Ct-values, estimated focus forming units (FFU) per milliliter or estimated genome copies per mL, and the virus reference size (nts).

Sample	Virus	Sample type	Ct-value	Estimate focus forming units (FFU) per milliliter or genome copies per mL	Reference size (nts)
ZIKV Ref	ZIKV	Isolate	15.1	6E+07	10,808
1002	YFV	Plasma	33.4	2.50E+00	10,859
1003	YFV	Plasma	37.0	1.00E+00	10,859
1004	YFV	Plasma	36.0	1.20E+00	10,859
1005	YFV	Plasma	20.1	1.80E+04	10,859
1007	YFV	Plasma	36.8	1.20E+00	10,859
1009	YFV	Plasma	19.1	3.90E+04	10,859
1010	YFV	Plasma	33.7	2.50E+00	10,859
1019	YFV	Plasma	17.2	2.10E+05	10,859
1020	YFV	Plasma	29.5	5.80E+01	10,859
1023	YFV	Plasma	28.0	2.00E+02	10,859
1024	YFV	Plasma	33.0	2.50E+00	10,859
1026	YFV	Plasma	28.4	1.10E+02	10,859
1028	YFV	Plasma	28.2	1.10E+02	10,859
1029	YFV	Plasma	17.4	2.10E+05	10,859
1030	YFV	Plasma	4.6	1.50E+10	10,859
1032	YFV	Plasma	22.9	5.50E+03	10,859
1036	YFV	Plasma	23.3	2.75E+03	10,859
1037	YFV	Plasma	31.9	1.00E+01	10,859
1042	YFV	Plasma	23.3	2.75E+03	10,859
1043	YFV	Plasma	25.8	8.00E+02	10,859
1045	YFV	Plasma	23.7	2.75E+03	10,859
1053	YFV	Plasma	26.6	2.80E+02	10,859
1055	YFV	Plasma	26.9	2.80E+02	10,859
1059	YFV	Plasma	31.0	2.80E+01	10,859
1063	YFV	Plasma	32.0	7.00E+00	10,859
1065	YFV	Plasma	33.0	2.80E+00	10,859
1066	YFV	Plasma	29.3	3.00E+01	10,859
1068	YFV	Plasma	27.3	1.50E+02	10,859
1069	YFV	Plasma	30.1	2.80E+01	10,859
1070	YFV	Plasma	28.0	1.10E+02	10,859

1072	YFV	Plasma	27.6	1.50E+02	10,859
1073	YFV	Plasma	32.5	7.00E+00	10,859
1074	YFV	Plasma	27.9	1.50E+02	10,859
1075	YFV	Plasma	24.1	1.70E+03	10,859
1076	YFV	Plasma	34.8	2.30E+00	10,859
1078	YFV	Plasma	28.6	1.10E+02	10,859
1079	YFV	Plasma	25.5	6.50E+02	10,859
1080	YFV	Plasma	31.6	1.10E+02	10,859
1081	YFV	Plasma	25.2	6.50E+02	10,859
1082	YFV	Plasma	33.8	2.80E+00	10,859
1086	YFV	Plasma	28.2	1.10E+02	10,859
2159	SARS-CoV-2	NPS	32.2	2.80E+02	29,903
2163	SARS-CoV-2	NPS	27.6	5.40E+03	29,903
2168	SARS-CoV-2	NPS	33.3	1.30E+02	29,903
2173	SARS-CoV-2	NPS	25.5	2.10E+04	29,903
2182	SARS-CoV-2	NPS	28.7	2.80E+03	29,903
2187	SARS-CoV-2	NPS	31.1	5.80E+02	29,903
2192	SARS-CoV-2	NPS	27.5	5.80E+03	29,903
2207	SARS-CoV-2	NPS	28.9	2.40E+03	29,903
5148	SARS-CoV-2	NPS	22.6	1.50E+05	29,903
5157	SARS-CoV-2	NPS	21.8	2.40E+05	29,903

Legend: ZIKV = Zika virus;

YFV = Yellow fever virus;

SARS-CoV-2 = Severe acute respiratory syndrome coronavirus 2;

Ct = cycle threshold

FFU = focus forming units

nts = nucleotides;

NPS = nasopharyngeal swab

Table S3 - Summary of virus nanopore sequencing data using the tiling multiplex PCR approach of Zika virus reference strain BeH815744 (ZIKV) (n=1), yellow fever virus (YFV) (n=21), and severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) (n=10) samples with the corresponding Ct-values.

Sample	Virus	Ct-value	Number of reads	% Mapped reads	Average depth	% Genome coverage	% Identity	N50	%20x coverage
ZIKV Ref	ZIKV	15.1	7,432	55.99	326.98	98.74	83.3	445	97.10
1003	YFV	37.0	31,07	51.85	76.65	99.84	88.9	659	86.36
1005	YFV	20.1	68,63	24.72	782.25	99.89	88.5	599	91.64
1009	YFV	19.1	20,15	91.08	852.99	99.89	88.9	593	94.09

1019	YFV	17.2	54,51	83.20	2,11	99.95	89.1	595	91.88
1020	YFV	29.5	124,52	1.71	100.93	99.82	88.7	656	89.11
1023	YFV	28.0	64,27	69.62	1,95	92.77	91.5	563	61.13
1024	YFV	33.0	179,93	15.14	1,22	87.63	91.3	560	55.70
1029	YFV	17.4	35,83	91.41	1,51	99.89	89.1	585	91.68
1030	YFV	4.6	65,57	97.47	3,10	99.89	88.9	595	91.85
1032	YFV	22.9	61,57	42.48	1,20	99.82	89.0	583	91.51
1042	YFV	23.3	27,53	6.25	82.43	99.82	88.8	643	87.95
1045	YFV	23.7	51,04	90.14	2,11	99.83	88.9	585	91.73
1053	YFV	26.6	78,44	91.50	3,37	99.85	89.0	594	93.92
1055	YFV	26.9	103,99	4.83	238.17	78.43	91.9	587	31.66
1059	YFV	31.0	25,86	5.87	72.50	99.82	88.8	660	88.77
1063	YFV	32.0	38,97	4.17	77.05	99.82	88.8	637	91.02
1065	YFV	33.0	39,73	5.51	104.05	99.82	88.9	625	91.35
1068	YFV	27.3	245,11	2.40	221.91	86.40	91.6	525	44.53
1073	YFV	32.5	49,49	11.24	247.57	78.11	90.7	567	35.01
1074	YFV	27.9	80,31	3.09	117.94	99.82	88.7	634	91.55
1075	YFV	24.1	171,54	21.26	1,57	87.22	91.5	551	65.64
2159	SARS-CoV-2	32.2	77,634	96.03	864.03	99.03	92.7	501	87.32
2163	SARS-CoV-2	27.6	86,279	94.48	928.92	96.66	92.7	500	87.17
2168	SARS-CoV-2	33.3	68,839	96.36	821.77	99.04	92.7	500	89.23
2173	SARS-CoV-2	25.5	139,582	94.51	1,34	95.90	93.4	505	84.49
2182	SARS-CoV-2	28.7	129,854	97.04	1,52	99.90	93.5	505	94.70
2187	SARS-CoV-2	31.1	79,728	95.47	852.58	98.16	93.5	506	84.86
2192	SARS-CoV-2	27.5	125,292	96.34	1,45	99.92	93.4	504	95.62
2207	SARS-CoV-2	28.9	128,745	96.57	1,53	99.89	93.5	506	97.62
5148	SARS-CoV-2	22.6	129,957	97.27	1,57	99.89	92.2	501	98.17
5157	SARS-CoV-2	21.8	133,087	97.04	1,57	99.90	92.1	498	96.81

Legend: ZIKV = Zika virus;
YFV = Yellow fever virus;
SARS-CoV-2 = Severe acute respiratory syndrome coronavirus 2;
Ct = cycle threshold

N50: Half of the genome sequence is covered by contigs larger than or equal to the N50 contig size.

Identity: Identity compared to the genome reference.

Table S4 - Sequencing summary and alignment statistics results for Zika virus reference sample strain BeH815744 (ZIKV) using the SMART-9N method during development (n = 1 sample) according to the material input (FFU/mL).

Sample	Input (FFU/mL)	n°Reads	% Mapped reads	Average depth	% Genome coverage	% Identity	N50	Longest read	%20x coverage
ZIKV_Ref	6x10 ⁷	465,48	48.27	10,01	99.97	84.9	853	10,07	99.73
ZIKV_Ref	6x10 ⁶	239,58	46.14	4,73	99.80	84.7	741	8,05	99.67
ZIKV_Ref	6x10 ⁵	166,47	39.74	4,56	99.96	84.4	867	9,91	99.57
ZIKV_Ref	6x10 ⁴	163,28	56.29	7,48	99.68	83.9	1,15	7,82	99.18
ZIKV_Ref	6x10 ³	228,49	25.55	7,02	99.57	83.5	1,66	10,83	99.00
ZIKV_Ref	6x10 ²	79,66	6.51	719.08	99.67	83.4	2,14	6,16	97.54
ZIKV_Ref	6x10 ¹	298,98	0.52	154.25	99.02	83.1	1,74	9,64	90.70

Legend: ZIKV = Zika virus;
FFU = focus forming units.

N50: Half of the genome sequence is covered by contigs larger than or equal to the N50 contig size.

Longest read: Size of the aligned longest read in the data.

Identity: Identity of the aligned reads compared to the genome reference.

Table S5 - Sequencing summary and alignment statistics results for Zika virus reference sample strain BeH815744 (ZIKV) using the Rapid SMART-9N method during development (n = 1 sample) according to the material input (FFU/mL).

Sample	Input (FFU/mL)	n°Reads	% Mapped reads	Average depth	% Genome coverage	% Identity	N50	Longest read	%20x coverage
ZIKV_Ref 1µl	6x10 ⁷	51,88	84.51	5,55	99.62	83	1,84	7,68	99.00
ZIKV_Ref 0,5µl	6x10 ⁷	12,96	86.15	1,58	99.69	82.6	2,11	8,43	98.82
ZIKV_Ref	6x10 ⁶	23,39	85.11	5,55	99.96	83	2,08	7,20	99.02
ZIKV_Ref	6x10 ⁵	17,27	72.10	2,88	99.47	83.1	2,14	7,09	99.02
ZIKV_Ref	6x10 ⁴	79,87	49.67	1,87	99.78	83.2	2,50	7,62	99.06
ZIKV_Ref	6x10 ³	104,04	9.48	7,42	99.30	83.2	2,34	8,44	99.02

ZIKV_Ref	6x10 ²	189,46	0.16	1,76	99.56	83.3	2,10	5,85	94.73
ZIKV_Ref	6x10 ¹	134,34	0.06	57	87.58	83.1	2,91	6,01	44.60

Legend: ZIKV = Zika virus;
FFU = focus forming units.

N50: Half of the genome sequence is covered by contigs larger than or equal to the N50 contig size.

Longest read: Size of the aligned longest read in the data.

Identity: Identity of the aligned reads compared to the genome reference.

Table S6 - Sequencing summary and alignment statistics results for yellow fever virus (YFV) plasma samples (n=7) using the SMART-9N protocol during method validation according to the Ct-values.

Sample	Ct value	Number of reads	% Mapped	Average depth	% Genome coverage	% Identity	N50	Longest read	%20x coverage
1019	17.2	30,05	32.2	630.1	99.91	91.6	998	3,92	98.84
1020	29.5	53,8	2.0	40.6	99.91	92.9	659	4,46	64.28
1029	17.4	16,83	85.7	1,87	99.99	92.3	1,58	7,95	99.91
1030	4.6	31,2	98.6	3,48	99.93	92.4	1,36	7,72	99.91
1045	23.7	981	21.4	240.0	96.96	92.3	1,39	6,94	63.81
1053	26.6	1,74	20.4	40.4	99.91	92.2	1,4	10,08	80.50
1065	33.0	15,82	0.3	3.2	95.11	90.1	1,01	2,57	0.00

Legend: Ct = cycle threshold;

N50: Half of the genome sequence is covered by contigs larger than or equal to the N50 contig size.

Longest read: Size of the aligned longest read in the data.

Identity: Identity of the aligned reads compared to the genome reference.

Table S7 - Sequencing summary and alignment statistics results for yellow fever virus (YFV) plasma samples (n=7) using the Rapid SMART-9N protocol during method validation according to the Ct-values.

Sample	Ct value	Number of reads	% Mapped	Average depth	% Genome coverage	% Identity	N50	Longest read	%20x coverage
1019	17.2	368,59	0.44	66.50	99.91	89.9	705	6,7	95.58
1020	29.5	441,13	0.05	26.35	99.10	89.4	2,05	6,15	81.68
1029	17.4	9,21	38.18	175.87	99.91	90.2	792	9,12	98.95

1030	4.6	212,07	98.26	25,3	99.95	89.8	1,8	3,96	99.91
1045	23.7	29,97	5.13	21.44	99.91	91.0	2,1	6,16	99.04
1053	26.6	1,005,93	0.31	323.37	99.92	91.3	1,61	6,86	99.74
1065	33.0	109,774	0.16	22.81	94.28	91.8	2,16	7,21	50.50

Legend: Ct = cycle threshold;

N50: Half of the genome sequence is covered by contigs larger than or equal to the N50 contig size.

Longest read: Size of the aligned longest read in the data.

Identity: Identity of the aligned reads compared to the genome reference.

Table S8- Sequencing summary and alignment statistics results for Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) clinical samples (n=10) using the Rapid SMART-9N protocol during method validation according to the Ct-values.

Sample	Ct Value	Number of reads	% Mapped reads	Average depth	% Genome coverage	% Identity	N50	Longest read	%20x coverage
2159	32.2	187,35	68.30	556.96	100	91.9	2,42	6,50	99.80
2163	27.6	127,93	97.87	368.68	100	91.9	2,38	8,50	99.84
2168	33.3	82,53	66.01	281.04	100	91.9	2,54	7,21	99.79
2173	25.5	117,53	52.10	97.51	100	91.5	1,13	5,05	99.57
2182	28.7	142,95	64.40	333.64	100	92.0	2,38	8,77	99.80
2187	31.1	64,77	44.54	109.53	100	91.7	2,56	8,09	99.80
2192	27.5	130,99	48.16	98.49	100	91.3	1,09	5,92	99.81
2207	28.9	191,79	98.22	628.12	100	91.4	2,04	8,65	99.87
5148	22.6	842,57	12.15	291.97	100	91.7	1,10	6,30	99.81
5157	21.8	935,75	41.75	1,28	100	91.0	1,24	18,48	99.89

Legend: Ct = cycle threshold;

N50: Half of the genome sequence is covered by contigs larger than or equal to the N50 contig size.

Longest read: Size of the aligned longest read in the data.

Identity: Identity of the aligned reads compared to the genome reference.

Table S9 - Proportion in the percentage of unclassified, Eukaryota, bacteria, archaea, and viruses reads, for each sample according to the Kraken classification distribution and metagenomics methodologies.

Sample	Virus type	Methodology	Number of reads	Unclassified (%)	Eukaryota (%)	Bacteria (%)	Archaea (%)	Virus (%)
1019	YFV	SMART-9N	30,05	58.12	37.63	1.37	0.01	2.18
1020	YFV	SMART-9N	53,8	46.17	45.23	5.08	0.01	0.14
1029	YFV	SMART-9N	16,83	72.75	11.03	1.97	0.02	13.87
1030	YFV	SMART-9N	31,2	84.46	2.40	1.91	0.02	11.17
1045	YFV	SMART-9N	981	30.48	35.27	4.79	0	29.26
1053	YFV	SMART-9N	1,74	28.90	42.31	1.15	0.06	27.00
1065	YFV	SMART-9N	15,82	46.91	52.23	0.56	0.01	0.01
1019	YFV	Rapid SMART-9N	368,59	83.10	1.70	16.85	0.01	0.04
1020	YFV	Rapid SMART-9N	441,13	78.60	18.85	2.26	0.01	0.04
1029	YFV	Rapid SMART-9N	9,21	84.76	12.74	1.55	0	0.73
1030	YFV	Rapid SMART-9N	212,07	92.86	0.54	2.33	0.02	4.79
1045	YFV	Rapid SMART-9N	29,97	80.47	18.86	0.84	0.01	0.65
1053	YFV	Rapid SMART-9N	1,005,9	77.88	21.15	0.73	0.01	0.03
1065	YFV	Rapid SMART-9N	109,77	75.42	23.77	0.41	0.01	0.01
2159	SARS-CoV-2	Rapid SMART-9N	187,35	58.72	2.84	20.35	0.04	17.91
2163	SARS-CoV-2	Rapid SMART-9N	127,93	62.76	3.69	3.22	0.04	30.13
2168	SARS-CoV-2	Rapid SMART-9N	82,53	63.46	6.59	14.42	0.02	15.34
2173	SARS-CoV-2	Rapid SMART-9N	117,53	74.84	17.37	1.68	0	5.90
2182	SARS-CoV-2	Rapid SMART-9N	142,95	56.45	6.00	17.96	0.04	19.37
2187	SARS-CoV-2	Rapid SMART-9N	64,77	61.90	12.43	12.01	0	13.39
2192	SARS-CoV-2	Rapid SMART-9N	130,99	74.93	18.21	1.51	0.03	5.24
2207	SARS-CoV-2	Rapid SMART-9N	191,79	74.11	2.25	2.10	0.04	21.43
5148	SARS-CoV-2	Rapid SMART-9N	842,57	75.46	1.43	22.16	0.02	0.88
5157	SARS-CoV-	Rapid	935,75	75.31	1.70	19.37	0.01	3.48

	2	SMART-9N						
--	---	----------	--	--	--	--	--	--

Legend: YFV = yellow fever virus

SARS-CoV-2 = Severe acute respiratory syndrome coronavirus 2

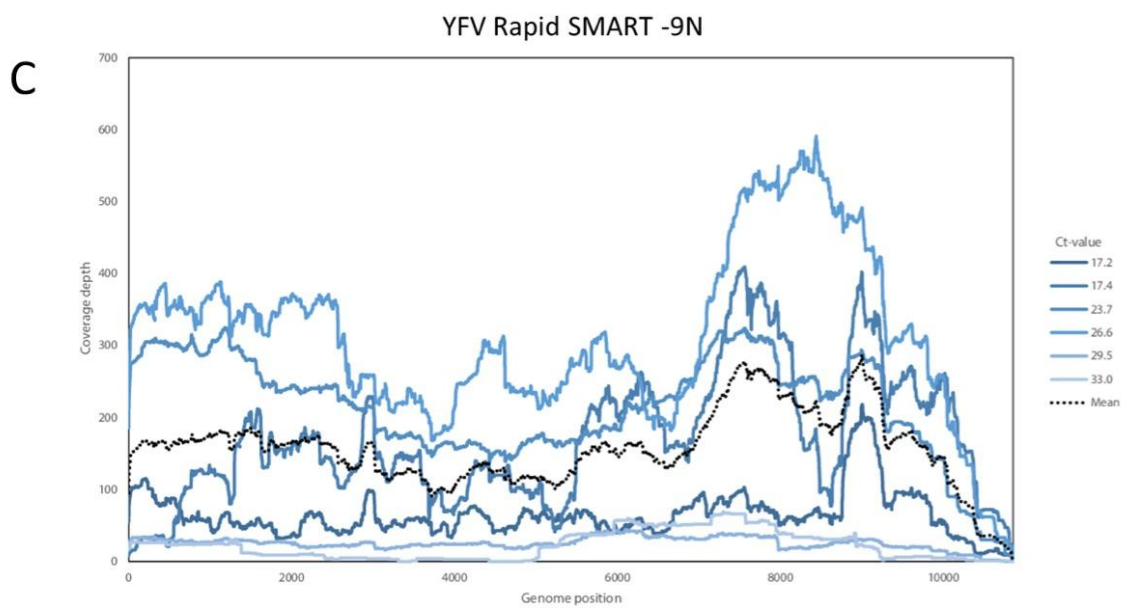
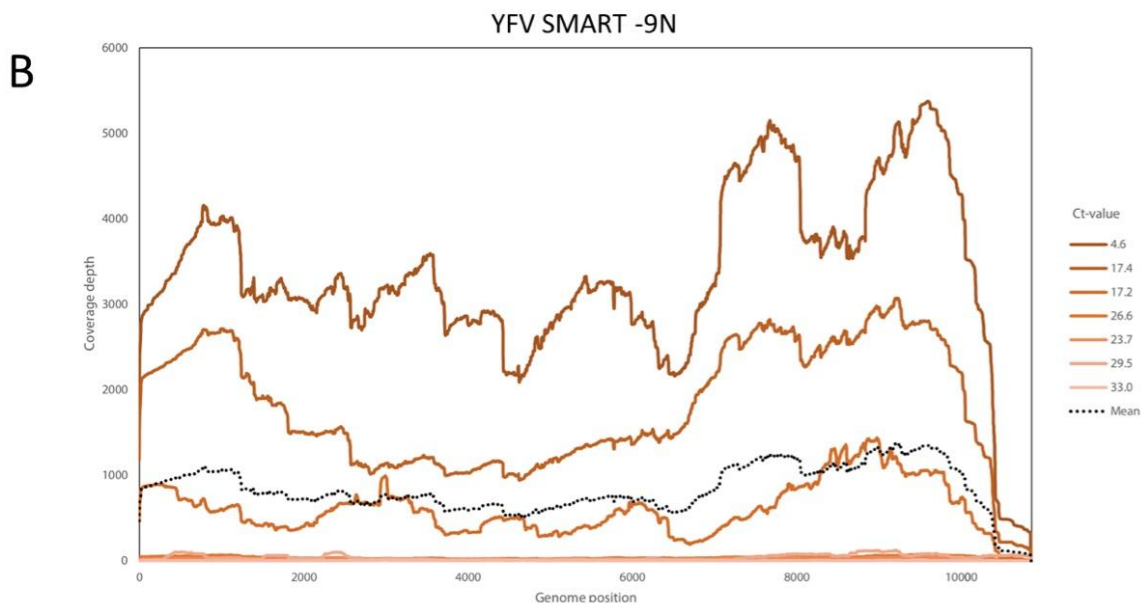
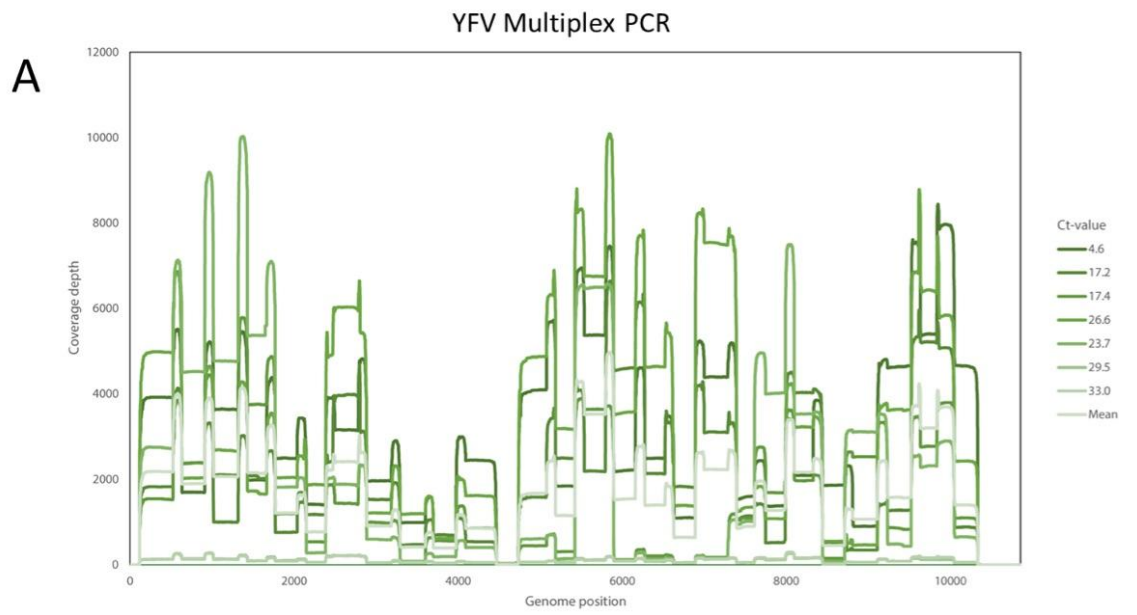


Figure S1 - Comparison of genome coverage depth across the yellow fever virus (YFV) genome for different methods (i.e., multiplex PCR, SMART-9N, and Rapid SMART-9N) in all clinical samples tested with different Ct-values. YFV (n=7)

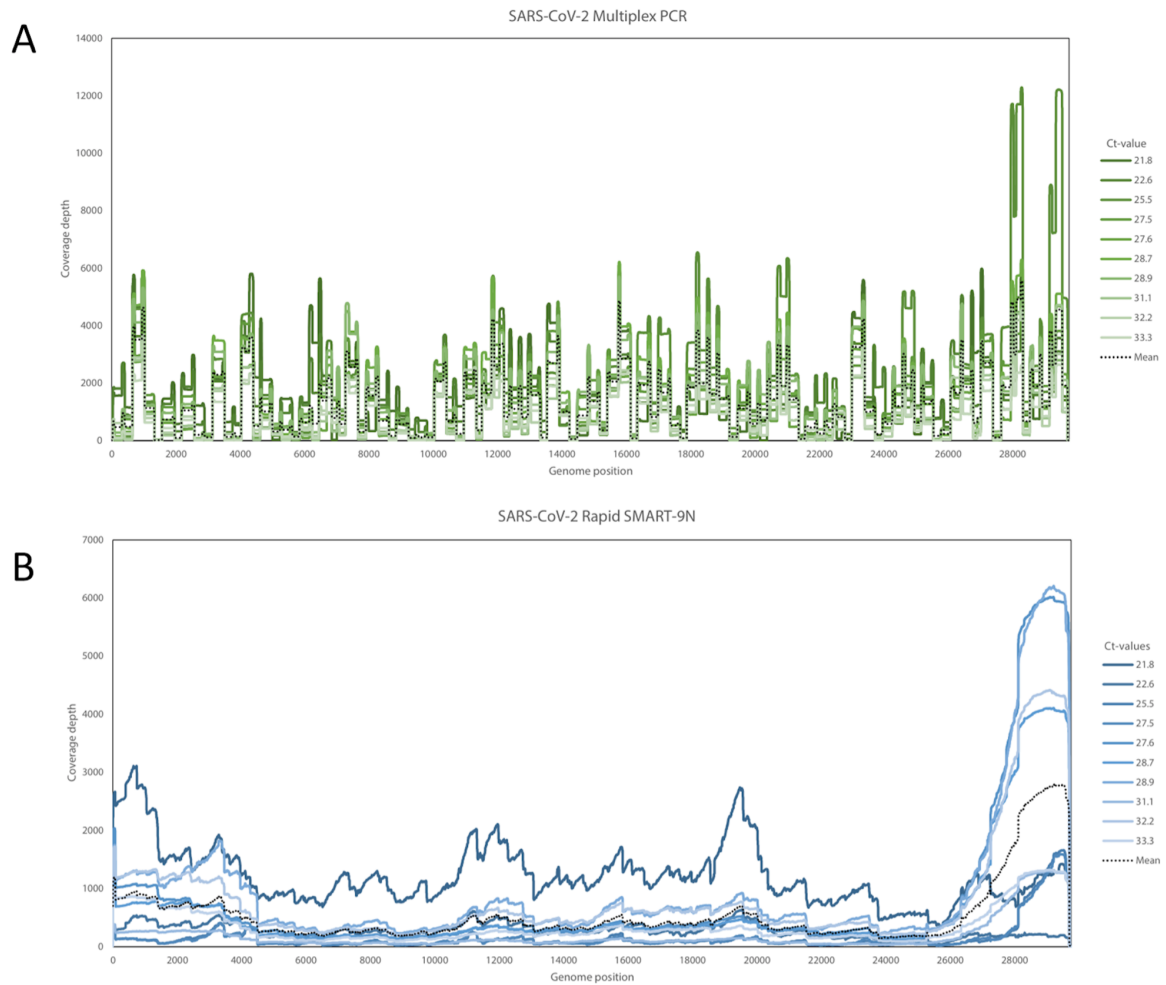


Figure S2 - Comparison of genome coverage depth across the Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) virus genome for different methods (i.e., multiplex PCR, and Rapid SMART-9N) in all clinical samples tested with different Ct-values. SARS-CoV-2 (n=10).