

# Efficient screening of long oligonucleotides against hundred thousands of SARS-CoV-2 genome sequences

## SUPPLEMENTARY MATERIAL

### Supplementary Figure S1

Analysis of mutant sequences. **(A)** Structure of RPA amplicon and oligonucleotides (RdRP FP/RP (32-mers), RdRP P (54-mer) and complete amplicon (128-mer) [1]). **(B)** Total number of SARS-CoV-2 sequences analysed and number of sequences containing nucleotide changes in the RPA oligonucleotides. **(C)** Percentage of mutant sequences per month (January to April 2021). **(D)** Sequence of the Alpha variant common mutation in the RPA reverse oligonucleotide.

### Supplementary Figure S2

Monthly analysis of SARS-CoV-2 variant genomes available from GISAID. The total number of sequences per month is depicted by the blue bars (scale on the right-hand side of each panel). Mutated target sequences are red dots, while discarded sequences of poor quality are shown by orange dots. The red line in the Alpha variant depicts the emerging and declining mutant sequence for the forward primer (.....c.....).

### Supplementary Figure S3

Observed sequence discard and mutation frequencies of all sequences analysed. **(A)** Distribution across months. **(B)** Boxplot of all values. Discarded sequences (orange) showed a mean of 1.26%  $\pm$  0.59% and (range 0.27%-3.46%); mutated sequences (red) showed a mean of 0.51%  $\pm$  0.48% and (range 0.08%-2.64%). Epsilon variant discard rates in blue.

### Supplementary Table S1

RdRP RT-RPA assay oligonucleotides.

### Supplementary Table S2

Data for Figure S1. Monthly analysis of variant sequences downloaded through GISAID search window by the end of May.

### Supplementary Table S3

Monthly analysis of Delta variant sequences downloaded on 2021-11-29.

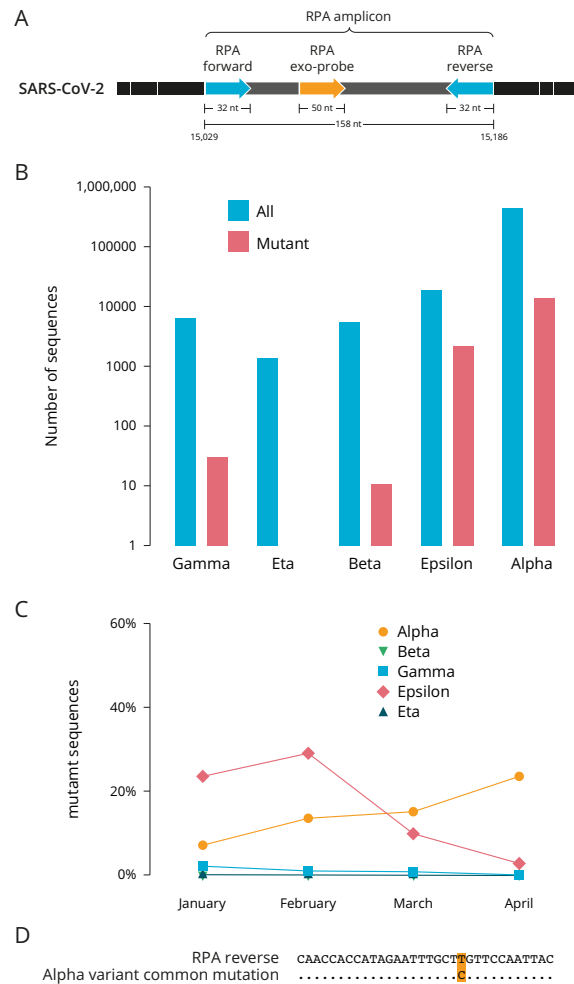
### Supplementary Table S4

Samples and reads details. Dominant changed reverse primer sequence only.

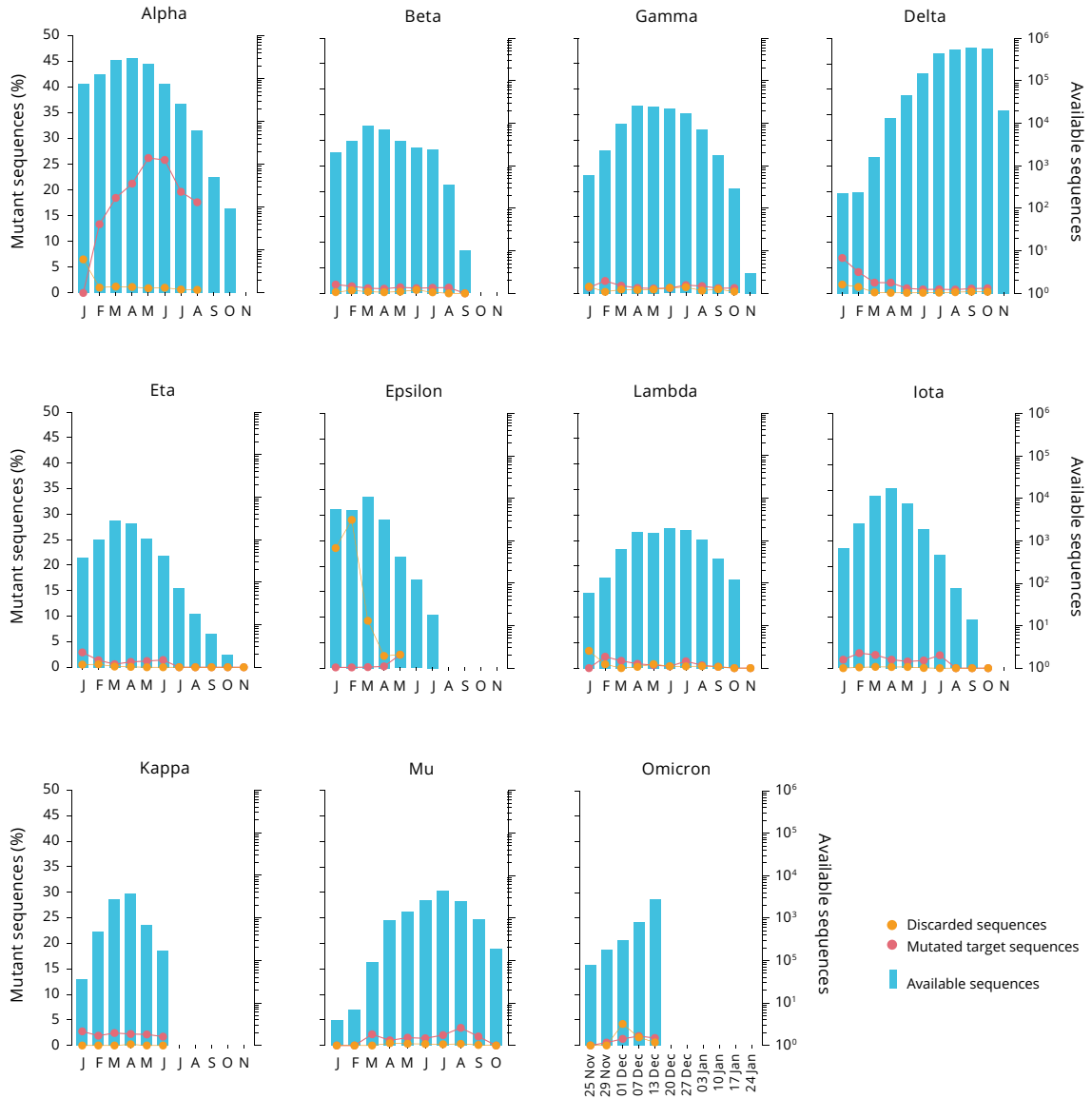
### Supplementary Table S5

Variations found in the Delta variant in June 2021. A total of 177 (0.12%) variant sequences were identified of the 149,576 screened sequences.

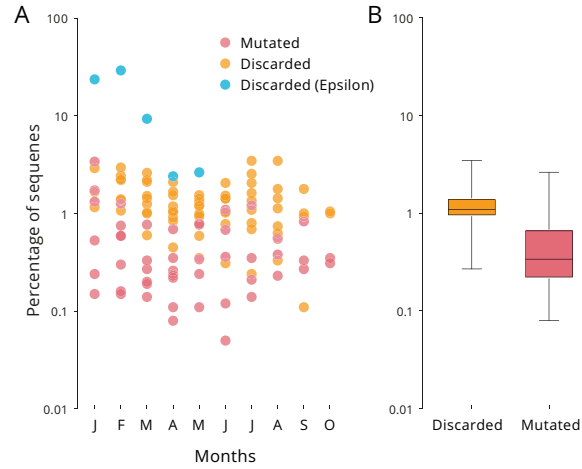
## SUPPLEMENTARY FIGURES



**Supplementary Figure S1.** Analysis of mutant sequences. **(A)** Structure of RPA amplicon and oligonucleotides (RdRP FP/RP (32-mers), RdRP P (54-mer) and complete amplicon (128-mer) [1]). **(B)** Total number of SARS-CoV-2 sequences analysed and number of sequences containing nucleotide changes in the RPA oligonucleotides. **(C)** Percentage of mutant sequences per month (January to April 2021). **(D)** Sequence of the Alpha variant common mutation in the RPA reverse oligonucleotide.



**Supplementary Figure S2.** Monthly analysis of SARS-CoV-2 variant genomes available from GISAID. The total number of sequences per month is depicted by the blue bars (scale on the right-hand side of each panel). Mutated target sequences are red dots, while discarded sequences of poor quality are shown by orange dots. The red line in the Alpha variant depicts the emerging and declining mutant sequence for the forward primer (.....c.....).



**Supplementary Figure S3.** Observed sequence discard and mutation frequencies of all sequences analysed. **(A).** Distribution across months. **(B)** Boxplot of all values. Discarded sequences (orange) showed a mean of  $1.26\% \pm 0.59\%$  and (range 0.27%-3.46%); mutated sequences (red) showed a mean of  $0.51\% \pm 0.48\%$  and (range 0.08%-2.64%). Epsilon variant discard rates in blue.

## SUPPLEMENTARY TABLES

**Supplementary Table S1.** RdRP RT-RPA assay oligonucleotides. --- represents three non-detected nucleotides spanned by the tetrahydrofuran bridge.

Name	Sequence (5'-3')	Length
Forward	TATGCCATTAGTGCAAAGAATAGAGCTCGCAC	32 nt
Reverse	CAACCACCATAGAATTTGCTTGTTCCAATTAC	32 nt
Exo-Probe	ATTCTCTATTGAGCATAAGAATATTGCTGC---TACATTCTTGTTGCCG	47 (50)

**Supplementary Table S2.** Data for Figure S1. Monthly analysis of variant sequences downloaded through GISAID search window by the end of May. \* deletion in reverse primer.

	Month	Total no. sequences	Changed target sequences	Discarded sequences
Alpha	J	68,192	4,774 (7.00%)	309 (0.45%)
	F	99,797	13,314 (3.34%)	1,402 (1.40%)
	M	167,069	24,978 (15.13%)	2,900 (1.38%)
	A	131,558	31,130 (23.36%)	1,551 (1.18%)
Beta	J	202	0 (0%)	3 (1.49%)
	F	322	1 (0.31%)	3 (0.93%)
	M	3,528	2 (0.06%)	42 (1.19%)
	A	1,558	5 (0.32%)	18 (1.16%)
Gamma	J	274	5 (1.82%)	3 (1.09%)
	F	630	5 (0.79%)	3 (0.48%)
	M	2,270	18 (0.79%)	47 (2.07%)
	A	3,529	3 (0.09%)	49 (1.39%)
Epsilon	J	5,477	9 (19.50%*)	214 (3.91%)
	F	5,265	29 (22.24%*)	332 (6.31%)
	M	10,853	13 (0.16%)	1,027 (9.46%)
	A	3,168	5 (0.16%)	81 (2.56%)
Eta	J	199	0 (0%)	3 (1.49%)
	F	318	1 (0.31%)	3 (0.93%)
	M	870	0 (0%)	5 (0.57%)
	A	580	0 (0%)	10 (1.72%)

**Supplementary Table S3.** Monthly analysis of Delta variant sequences downloaded on 2021-11-29. \* analysis postponed until numbers have been updated.

Month	Total no. sequences	Changed target sequences	Discarded sequences
J	31	4 (1.73%)	14 (6.93%)
F	239	3 (1.26%)	10 (4.18%)
M	1614	3 (0.19%)	34 (2.11%)
A	13,066	15 (0.11%)	274 (2.10%)
M	45,613	51 (0.11%)	452 (0.99%)
J	149,576	177 (0.12%)	1,162 (0.78%)
J	443,308	633 (0.14%)	3,534 (0.80%)
A	539,653	1,231 (0.23%)	4,014 (0.74%)
S	602,118	2,009 (0.33%)	5,615 (0.93%)
O	574,656	1,784 (0.31%)	5,773 (1.00%)
N	20,484*		

**Supplementary Table S4.** Samples and reads details. Dominant changed reverse primer sequence only. \* not plotted in Figure 1.

Month	Total no. sequences	Changed target sequences	Discarded sequences
J	74,983	19 (0.03%)	4915 (6.55%)
F	121,112	16,195 (13.37%)	1301 (1.07%)
M	268,703	49,660 (18.48%)	3,372 (1.25%)
A	297,289	63,172 (21.25%)	3,500 (1.18%)
M	216,525	6,790 (26.23%)	2,003 (0.93%)
J	72,493	18,744 (25.86%)	752 (1.04%)
J	24,749	4,968 (19.67%)	170 (0.69%)
A	6,156	1,086 (17.64%)	39 (0.63%)
S	500	212 (42.40%*)	219 (43.80%*)
O	170	75 (44.12%*)	0 (0%*)
N	1	0 (0%*)	1 (100.00%*)

**Supplementary Table S5.** Variations found in the Delta variant in June 2021. A total of 177 (0.12%) variant sequences were identified of the 149,576 screened sequences.

Nb. changes	RPA forward variation	RPA reverse variation
1	...T.....	
4	....T.....	
9	.....T.....	
1	.....A.....	
3	.....G.....	
5	.....G.....	
1	.....C.....	
4	.....C.....	
5	.....G.....	
5	.....T.....	
24	.....T.....	
8	.....G.....	
5	.....C.....	
8	.....T.....	
6	.....A.....	
1		..G.....
4		...G.....
4		....C.....
1		.....a.....
4		.....G.....
12		.....T.....
2		.....G.....
2		.....a.....
9		.....T.....
2		.....G.....
32		.....T.....
9		.....C.....
6		.....C.....

## REFERENCES

1. El Wahed AA, Patel P, Maier M, Pietsch C, Rüster D, Böhlken-Fascher S, et al. Suitcase lab for rapid detection of SARS-CoV-2 based on recombinase polymerase amplification assay. *Analytical chemistry* **93** (2021) 2627–2634. doi:10.1021/acs.analchem.0c04779.