Supporting information for Vaccine-escape and fast-growing mutations in the United Kingdom, the United States, Singapore, Spain, India, and other COVID-19-devastated countries

Rui Wang[1], Jiahui Chen[1], Kaifu Gao[1], and Guo-Wei Wei[1,2,3*]
[1] Department of Mathematics,
Michigan State University, MI 48824, USA.
[2] Department of Electrical and Computer Engineering,
Michigan State University, MI 48824, USA.
[3] Department of Biochemistry and Molecular Biology,
Michigan State University, MI 48824, USA.

April 29, 2021

Contents

*Corresponding author. Email: weig@msu.edu

S1 BFE changes following 551 non-degenerate mutations on the S protein RBD

Figure S1-Figure S5 plot the BFE changes for the complexes of S protein and ACE2 for all 551 non-degenerate mutations on the S protein RBD. Frequency information has been indicated. It is worthy to note the disparity between the BFE magnitudes for binding-strengthening mutations and binding-weakening mutations. Such a disparity indicates that the SARS-CoV-2 has been relatively adapted for human infection.

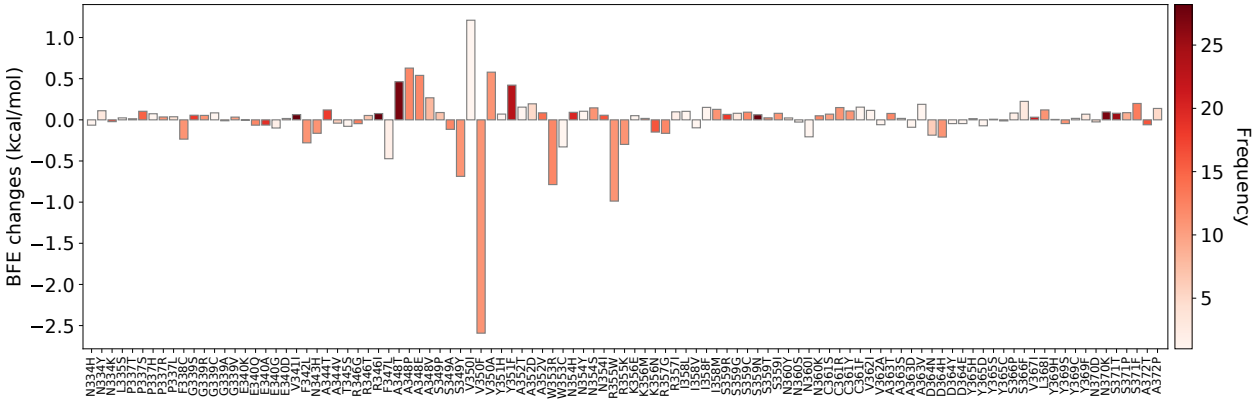

Figure S1: Illustration of the BFE changes for the complexes of S protein and ACE2 following all 551 non-degenerate mutations on the S protein RBD. The frequency shows the number of detections in the database.

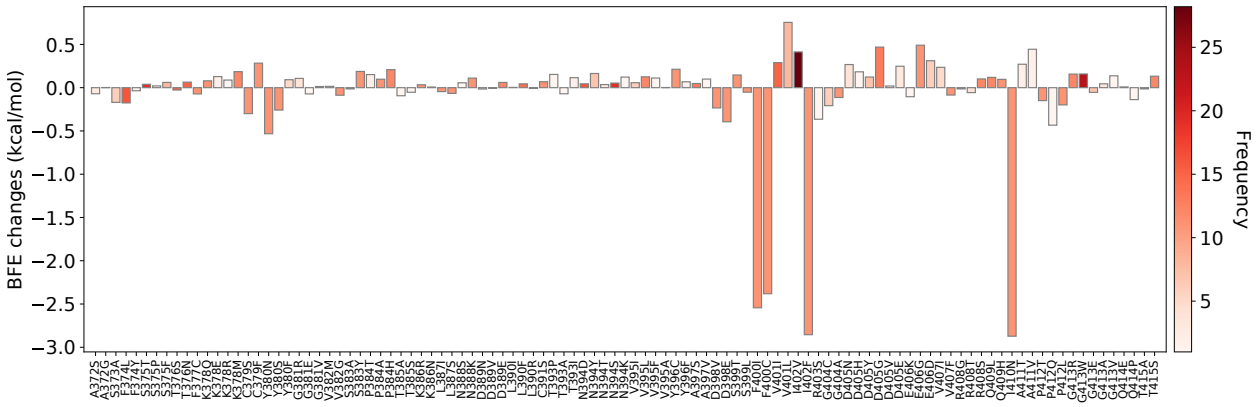

Figure S2: (Continue) Illustration of the BFE changes for the complexes of S protein and ACE2 following all 551 non-degenerate mutations on the S protein RBD. The frequency shows the number of detections in the database.

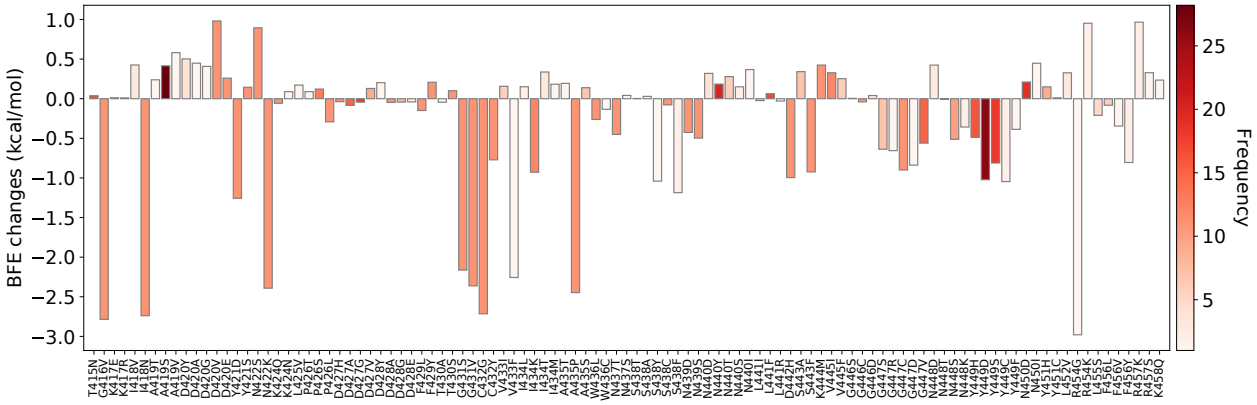

Figure S3: (Continue) Illustration of the BFE changes for the complexes of S protein and ACE2 following all 551 non-degenerate mutations on the S protein RBD. The frequency shows the number of detections in the database.

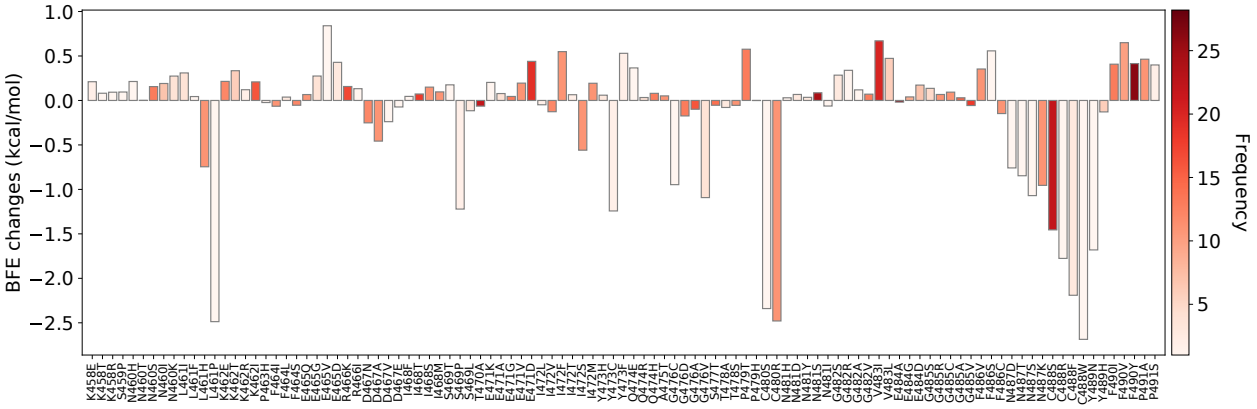

Figure S4: (Continue) Illustration of the BFE changes for the complexes of S protein and ACE2 following all 551 non-degenerate mutations on the S protein RBD. The frequency shows the number of detections in the database.

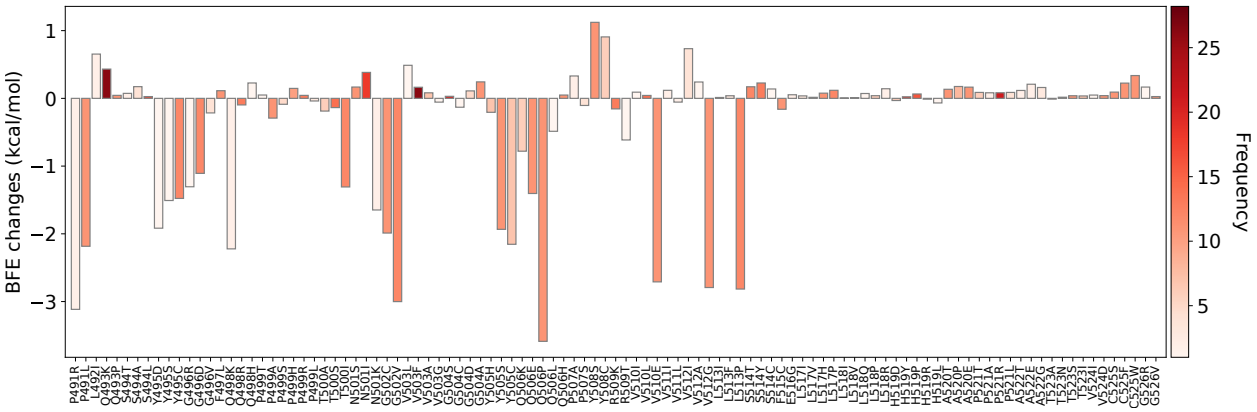

Figure S5: (Continue) Illustration of the BFE changes for the complexes of S protein and ACE2 following all 551 non-degenerate mutations on the S protein RBD. The frequency shows the number of detections in the database.

S2 106 antibodies with their corresponding PDB IDs

PDB ID	Antibodies	PDB ID	Antibodies	PDB ID	Antibodies
6W41	CR3022	6WPS	S309	6XC2	CC12.1
6XC3	CC12.1/CR3022	6XC4	CC12.3	6XC7	CC12.3/CR3022
6XCM	C105	6XDG	REGN10933/10987	6XE1	CV30
6XEY	Fab 2-4	6XKP	CV07-270	6XKQ	CV07-250
6YZ5	H11-D4	6Z2M	CR3022/H11-D4	6ZBP	H11-H4
6ZCZ	EY6Z/Nb	6ZER	EY6Z	7A29	Sb23
7B3O	STE90-C11	7BWJ	P2B-2F6	7BYR	BD23
7BZ5	B38	7C01	CB6	7C8V	SR4
7C8W	MR17	7CAH	H014	7CAN	MR17-K99Y
7CDI	P2C-1F11	7CDJ	P2C-1A3	7CH4	BD-604
7CH5	BD-629	7CHB	BD-236	7CHE	BD-236/BD-368-2
7CHF	BD-604/BD-368-2	7CHH	BD-368-2	7CJF	A fab
7CM4	CT-P59	7CWM	P17	7CWN	P17/H014
7JMO	COVA2-04	7JMP	COVA2-39	7JMW	COVA1-16
7JV6	S2H13	7JVA	S2A4	7JW0	S304
7JWB	VH binder	7JX3	S309/S2H14/S304	7K43	S2M11
7K45	S2E12	7K8M	C102	7K8T	C002
7K8U	C104	7K8V	C110	7K8W	C119
7K8X	C121	7K8Z	C135	7K90	C144
7K9Z	Fabs 298/52	7KFV	C1A-B12	7KFW	C1A-B3
7KFX	C1A-C2	7KFY	C1A-F10	7KMG	LY-CoV555
7KMH	LY-CoV488	7KMI	LY-CoV481	7KN5	VHH E/U
7KN6	VHH V/Fab CC12.3	7KN7	VHH W/Fab CC12.3	7KZB	CR3014-C8/CR3022
7LD1	DH1047	7EAM	7D6	7LSS	Fab 2-7
7LAA	DH1041	7LS9	1-57	7D0B	P5A-3C12_1B
7CZW	P5A-2G7	7D00	P5A-1B8	7D0C	P5A-3A1
7D0D	P5A-3C12_2B	7CZX	P5A-1B9	7CZY	P5A-2F11_2B
7CZV	P5A-1B6_3B	7CZZ	P5A-2F11_3B	7CZP	P2B-1A1
7CZQ	P2B-1A10	7CZT	P5A-2G9	7CZU	P5A-1B6_2B
7CZR	P5A-1B8_2B	7CZS	P5A-1B8_3B	7BEL	COVOX-88/-45
7BEM	COVOX-269 scFV	7BEJ	COVOX-158	7BEP	COVOX-384/S309
7BEN	COVOX-253/-75	7BEO	COVOX-253H55L/-75	7BEH	COVOX-316
7BEI	COVOX-150	7NDA	COVOX-253H55L	7NDB	COVOX-253H165L
7ND8	COVOX-384	7ND6	COVOX-40	7ND4	COVOX-88
7LOP	CV05-163/CR3022	7DPM	MW06	7KS9	910-30
7L5B	2-51	6M0J	ACE2		

S3 Supplementary Data

The Supplementary_Data.zip contains two folders:

1. SNP Data: A total of 31 CSV files for the SARS-CoV-2 S protein RBD SNP data from 31 different countries.
2. Fast Grow: A total of 31 HTML files for the log growth rates and log frequencies of specific SARS-CoV-2 S protein RBD mutations in 31 different countries.

S4 Supplementary Figures

Figure S6 - Figure S28 illustrate the log growth rate and log frequency of mutations on S protein RBD in the Germany, Canada, Sweden, Switzerland, Australia, France, Belgium, Italy, Spain, Ireland, Iceland, Luxembourg, Norway, Poland, Portugal, Latvia, Lithuania, Slovenia, Finland, Turkey, Czech Republic, United Arab Emirates, and Austria, respectively.

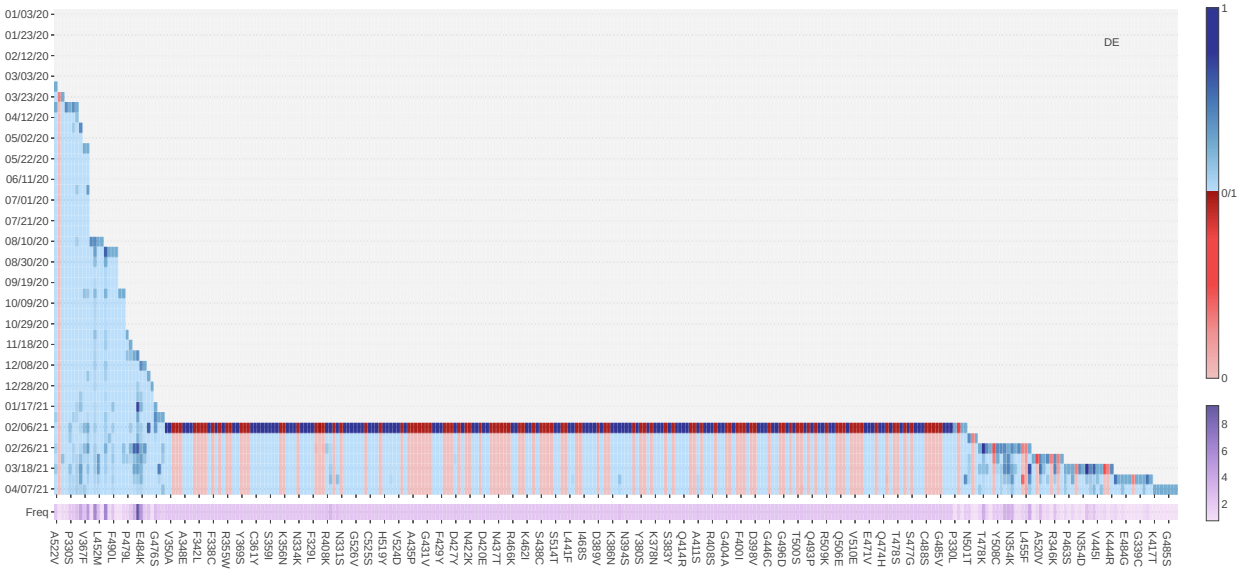

Figure S6: The log growth rate and log frequency of mutations on S protein RBD in the Germany. The blue and red colors respectively represent the binding-strengthening and binding-weakening mutations on RBD. The darker blue/red means the binding-strengthening/binding-weakening mutations with a higher growth rate in a specific 10-day period. The darker purple represents the mutation with a higher log frequency.

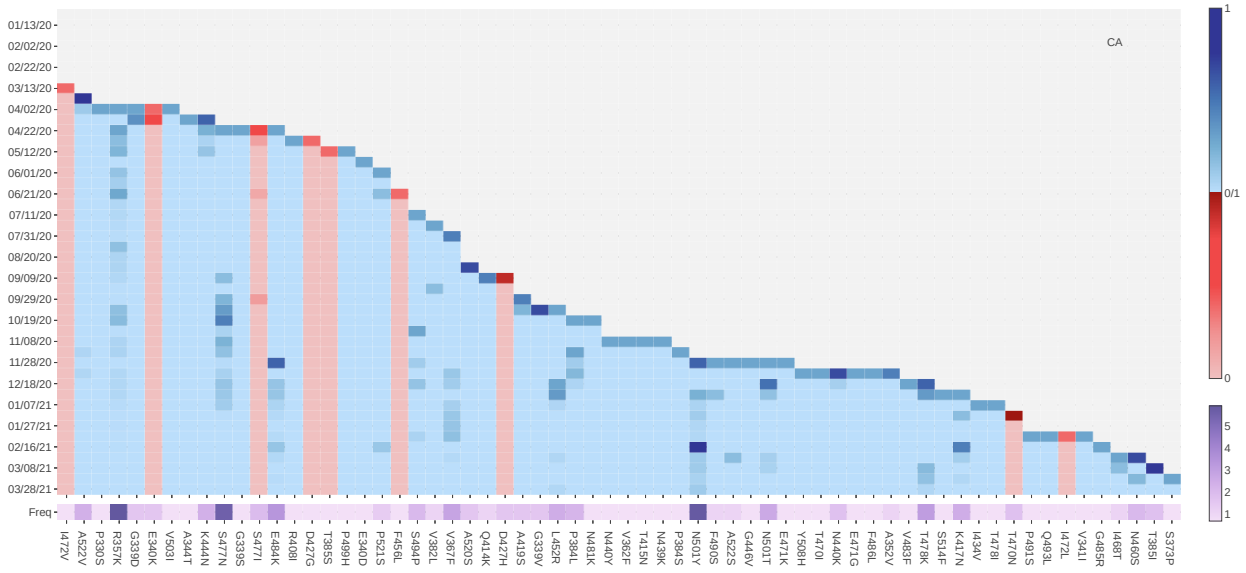

Figure S7: The log growth rate and log frequency of mutations on S protein RBD in the Canada. The blue and red colors respectively represent the binding-strengthening and binding-weakening mutations on RBD. The darker blue/red means the binding-strengthening/binding-weakening mutations with a higher growth rate in a specific 10-day period. The darker purple represents the mutation with a higher log frequency.

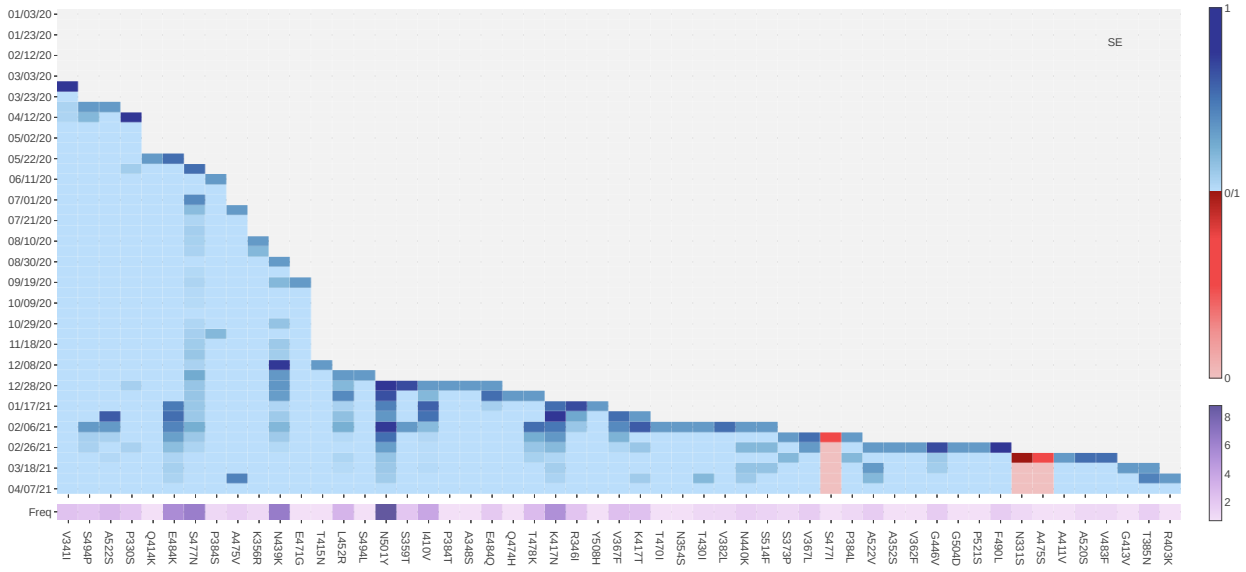

Figure S8: The log growth rate and log frequency of mutations on S protein RBD in the Sweden. The blue and red colors respectively represent the binding-strengthening and binding-weakening mutations on RBD. The darker blue/red means the binding-strengthening/binding-weakening mutations with a higher growth rate in a specific 10-day period. The darker purple represents the mutation with a higher log frequency.

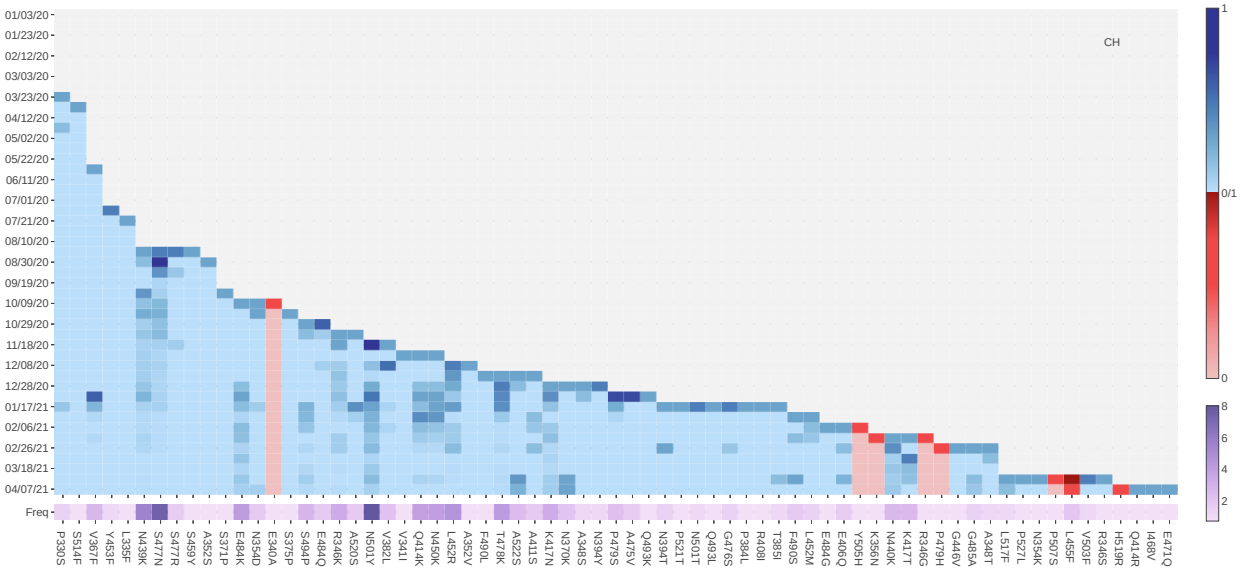

Figure S9: The log growth rate and log frequency of mutations on S protein RBD in the Switzerland. The blue and red colors respectively represent the binding-strengthening and binding-weakening mutations on RBD. The darker blue/red means the binding-strengthening/binding-weakening mutations with a higher growth rate in a specific 10-day period. The darker purple represents the mutation with a higher log frequency.

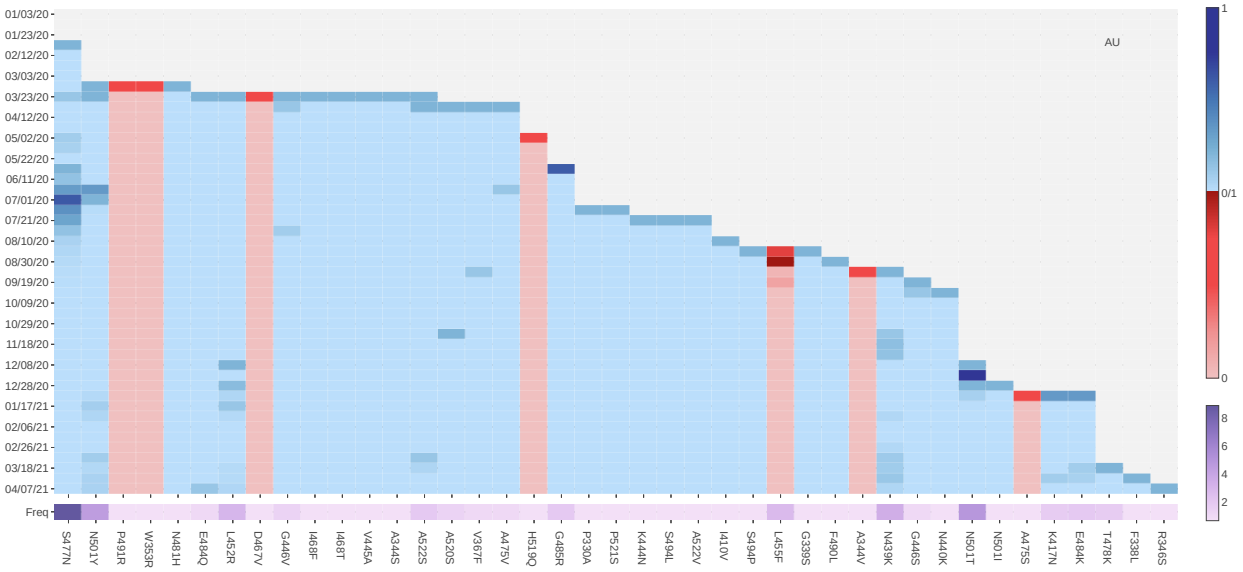

Figure S10: The log growth rate and log frequency of mutations on S protein RBD in the Australia. The blue and red colors respectively represent the binding-strengthening and binding-weakening mutations on RBD. The darker blue/red means the binding-strengthening/binding-weakening mutations with a higher growth rate in a specific 10-day period. The darker purple represents the mutation with a higher log frequency.

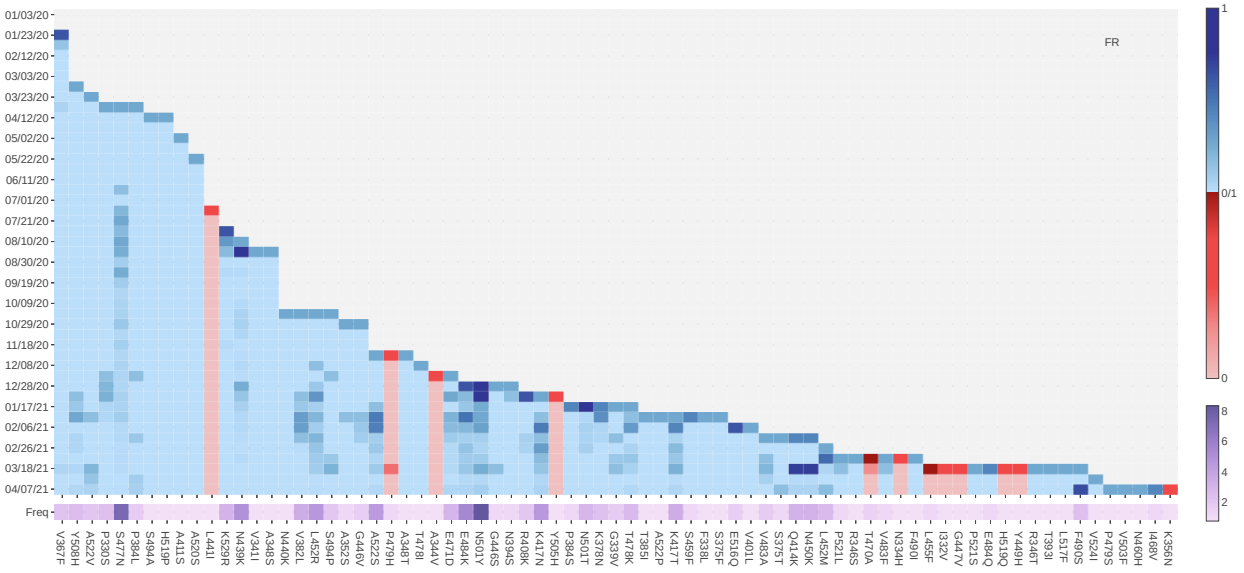

Figure S11: The log growth rate and log frequency of mutations on S protein RBD in the France. The blue and red colors respectively represent the binding-strengthening and binding-weakening mutations on RBD. The darker blue/red means the binding-strengthening/binding-weakening mutations with a higher growth rate in a specific 10-day period. The darker purple represents the mutation with a higher log frequency.

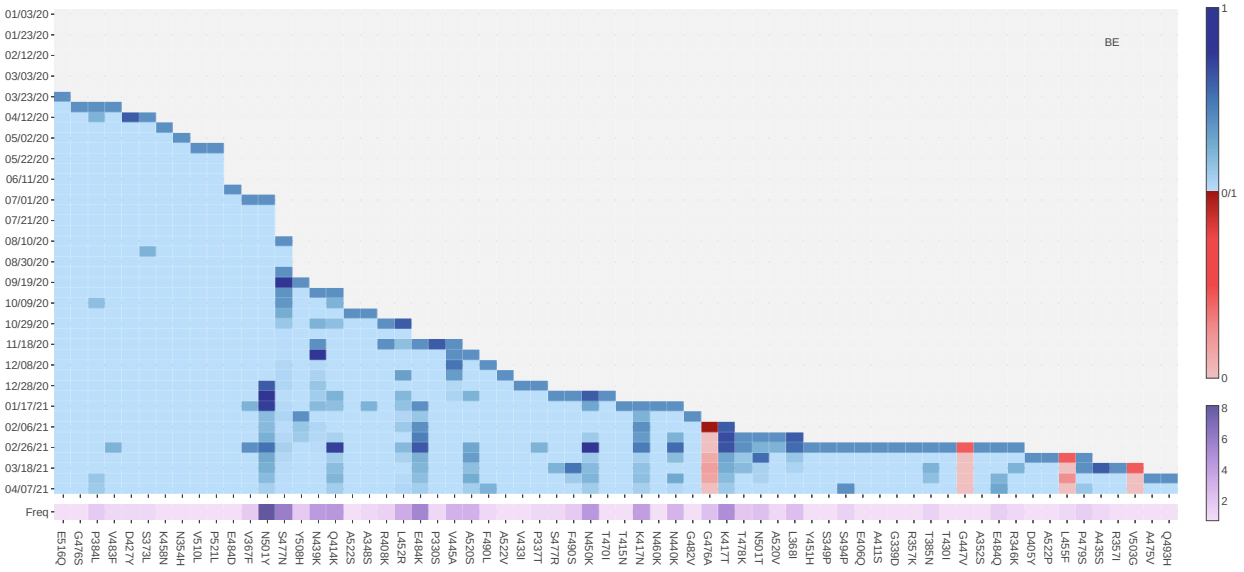

Figure S12: The log growth rate and log frequency of mutations on S protein RBD in the Belgium. The blue and red colors respectively represent the binding-strengthening and binding-weakening mutations on RBD. The darker blue/red means the binding-strengthening/binding-weakening mutations with a higher growth rate in a specific 10-day period. The darker purple represents the mutation with a higher log frequency.

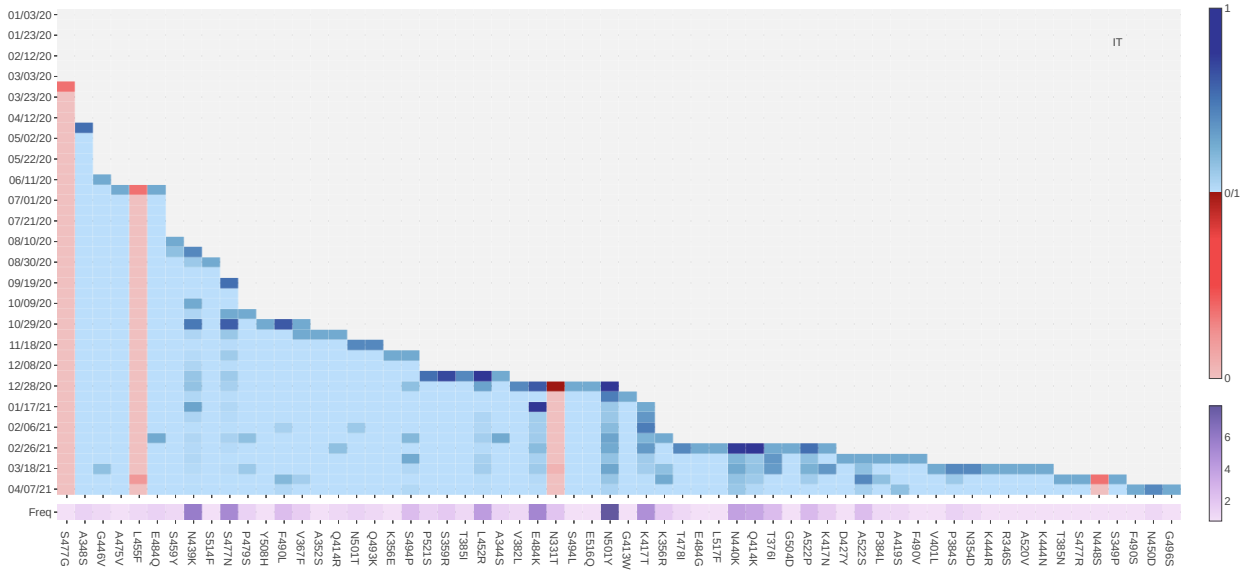

Figure S13: The log growth rate and log frequency of mutations on S protein RBD in the Italy. The blue and red colors respectively represent the binding-strengthening and binding-weakening mutations on RBD. The darker blue/red means the binding-strengthening/binding-weakening mutations with a higher growth rate in a specific 10-day period. The darker purple represents the mutation with a higher log frequency.

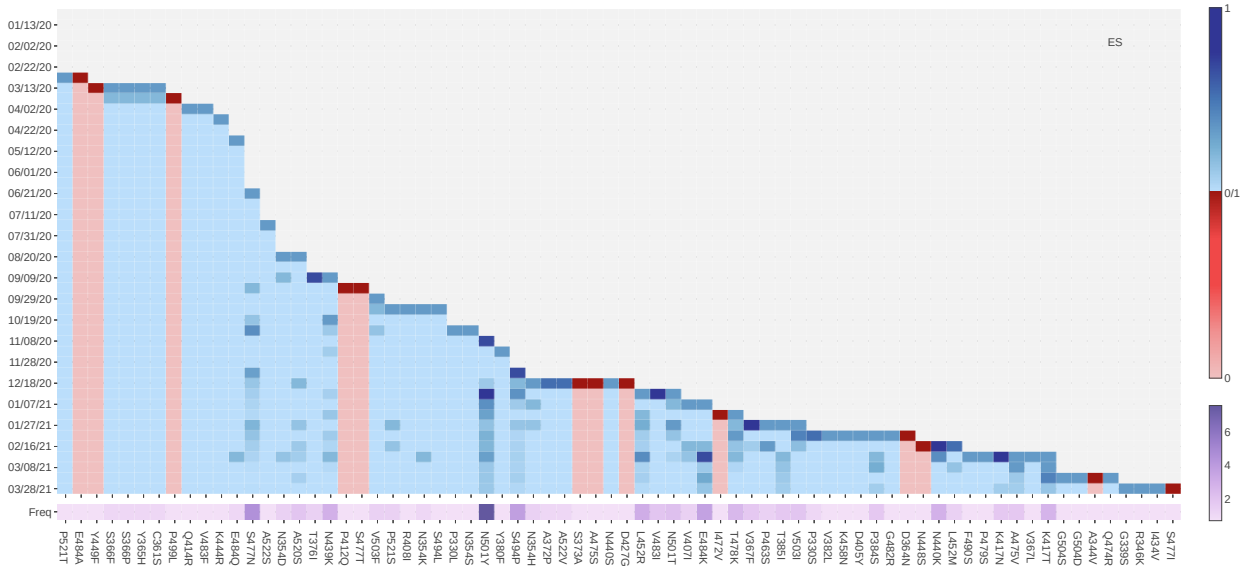

Figure S14: The log growth rate and log frequency of mutations on S protein RBD in the Spain. The blue and red colors respectively represent the binding-strengthening and binding-weakening mutations on RBD. The darker blue/red means the binding-strengthening/binding-weakening mutations with a higher growth rate in a specific 10-day period. The darker purple represents the mutation with a higher log frequency.

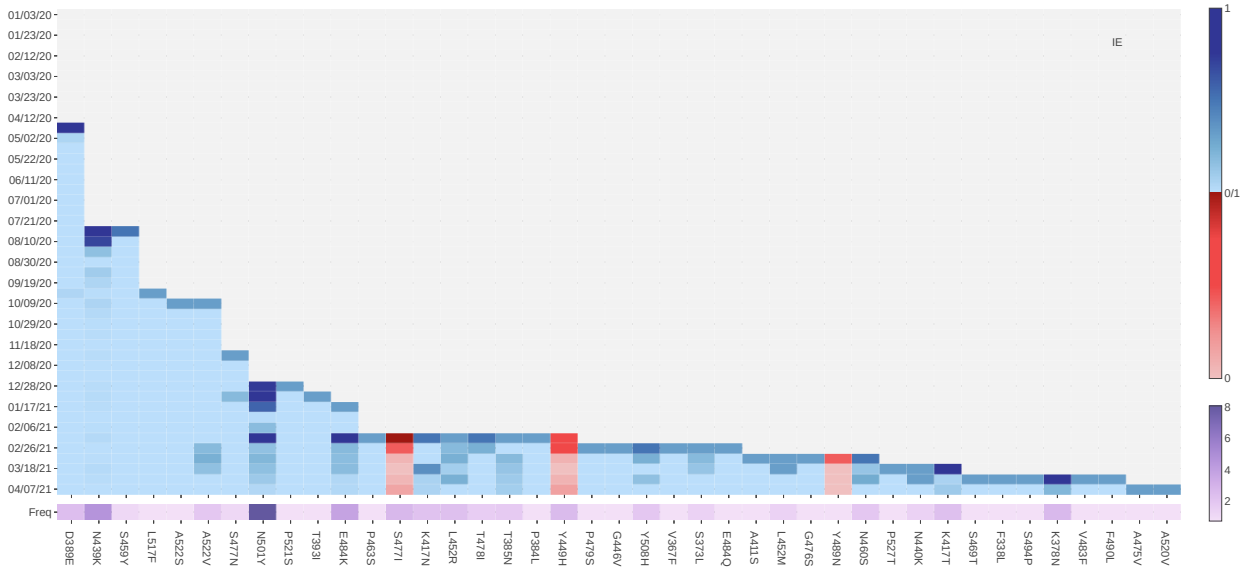

Figure S15: The log growth rate and log frequency of mutations on S protein RBD in the Ireland. The blue and red colors respectively represent the binding-strengthening and binding-weakening mutations on RBD. The darker blue/red means the binding-strengthening/binding-weakening mutations with a higher growth rate in a specific 10-day period. The darker purple represents the mutation with a higher log frequency.

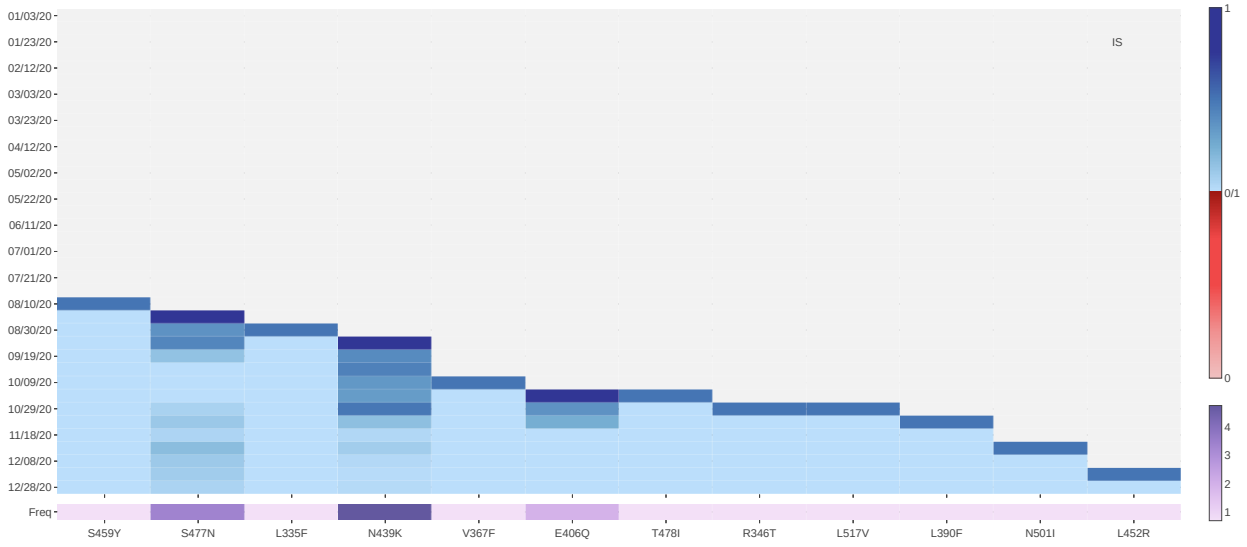

Figure S16: The log growth rate and log frequency of mutations on S protein RBD in the Iceland. The blue and red colors respectively represent the binding-strengthening and binding-weakening mutations on RBD. The darker blue/red means the binding-strengthening/binding-weakening mutations with a higher growth rate in a specific 10-day period. The darker purple represents the mutation with a higher log frequency.

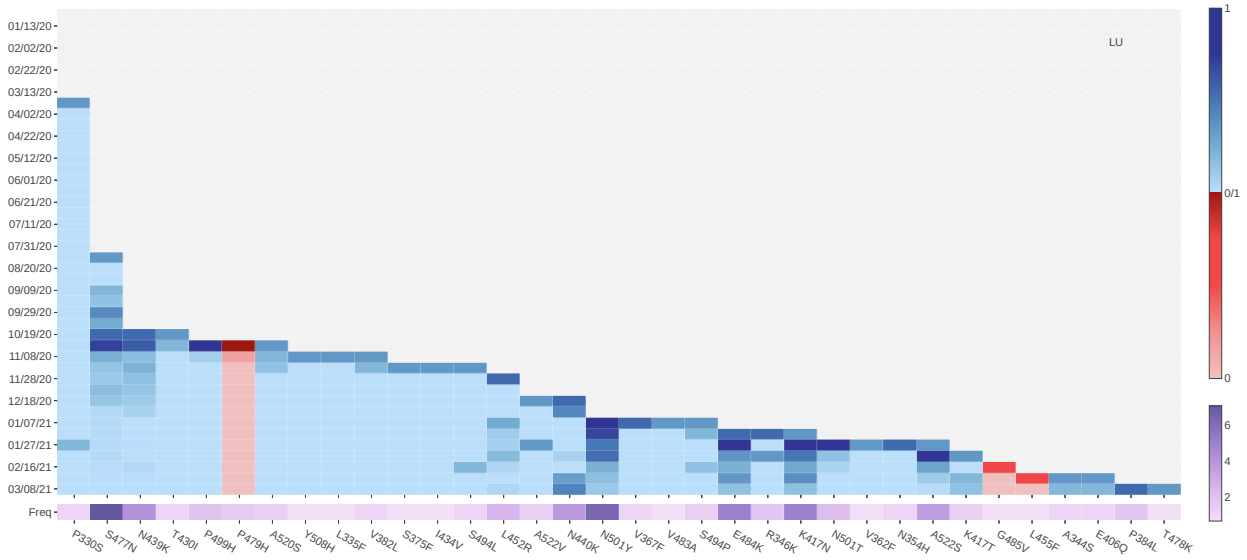

Figure S17: The log growth rate and log frequency of mutations on S protein RBD in the Luxembourg. The blue and red colors respectively represent the binding-strengthening and binding-weakening mutations on RBD. The darker blue/red means the binding-strengthening/binding-weakening mutations with a higher growth rate in a specific 10-day period. The darker purple represents the mutation with a higher log frequency.

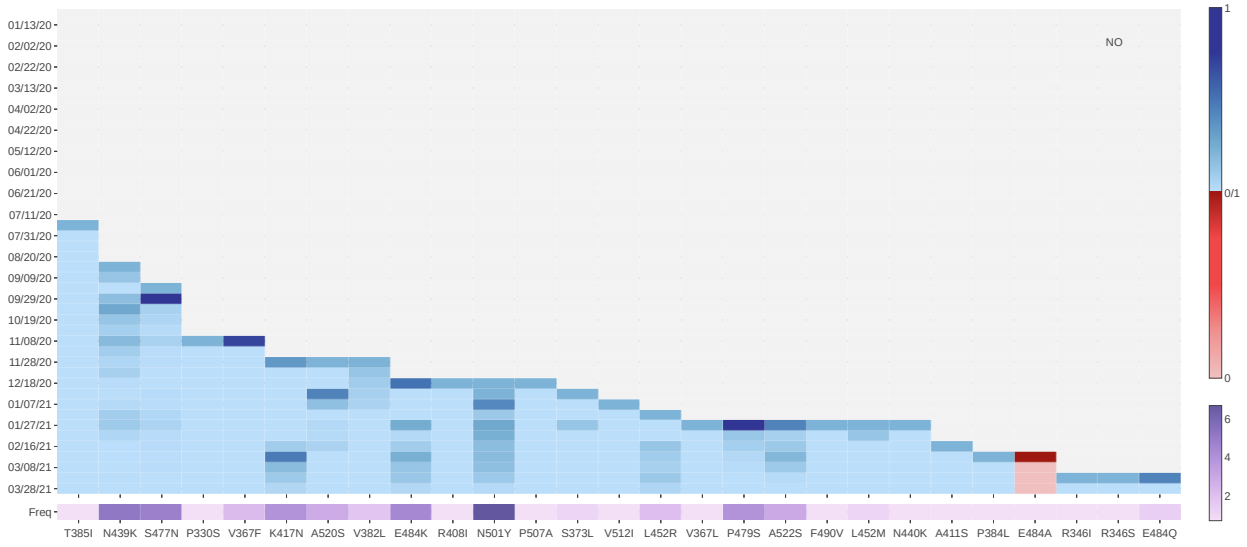

Figure S18: The log growth rate and log frequency of mutations on S protein RBD in the Norway. The blue and red colors respectively represent the binding-strengthening and binding-weakening mutations on RBD. The darker blue/red means the binding-strengthening/binding-weakening mutations with a higher growth rate in a specific 10-day period. The darker purple represents the mutation with a higher log frequency.

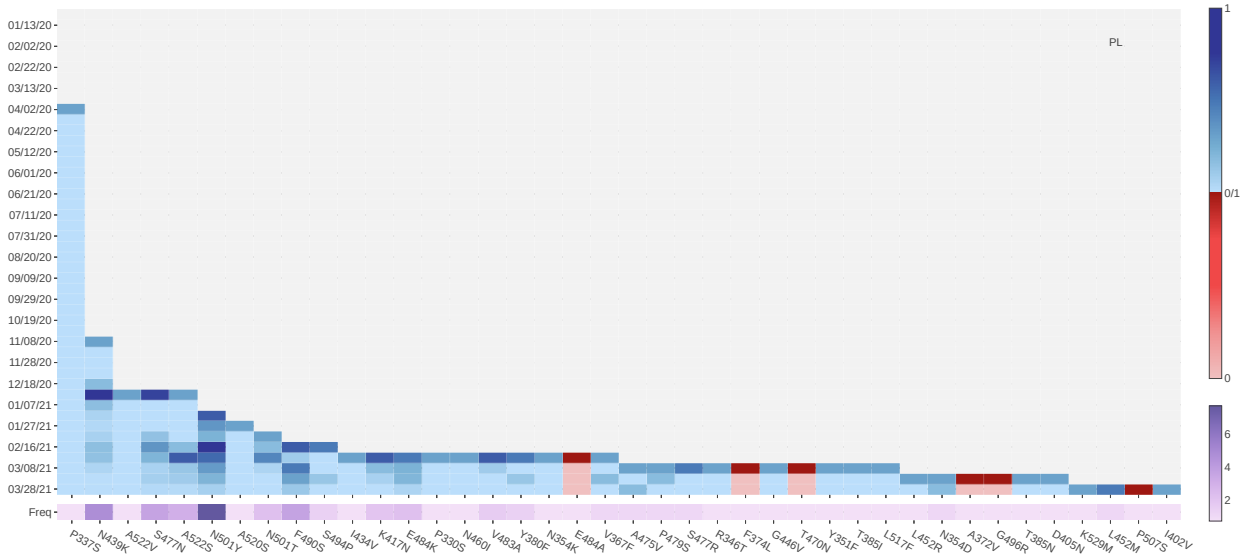

Figure S19: The log growth rate and log frequency of mutations on S protein RBD in the Poland. The blue and red colors respectively represent the binding-strengthening and binding-weakening mutations on RBD. The darker blue/red means the binding-strengthening/binding-weakening mutations with a higher growth rate in a specific 10-day period. The darker purple represents the mutation with a higher log frequency.

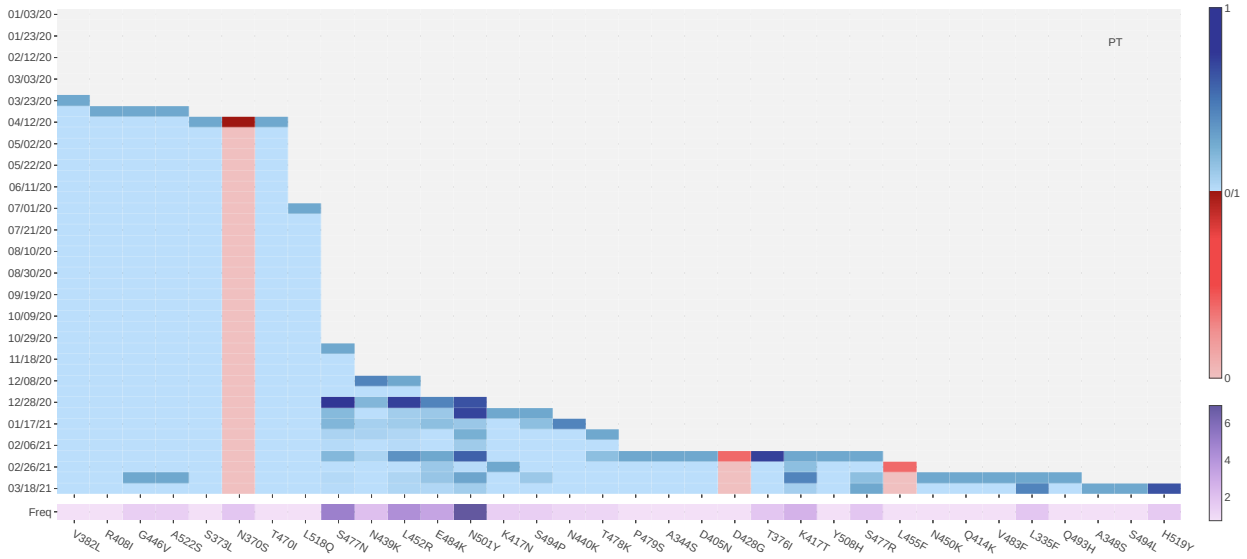

Figure S20: The log growth rate and log frequency of mutations on S protein RBD in the Portugal. The blue and red colors respectively represent the binding-strengthening and binding-weakening mutations on RBD. The darker blue/red means the binding-strengthening/binding-weakening mutations with a higher growth rate in a specific 10-day period. The darker purple represents the mutation with a higher log frequency.

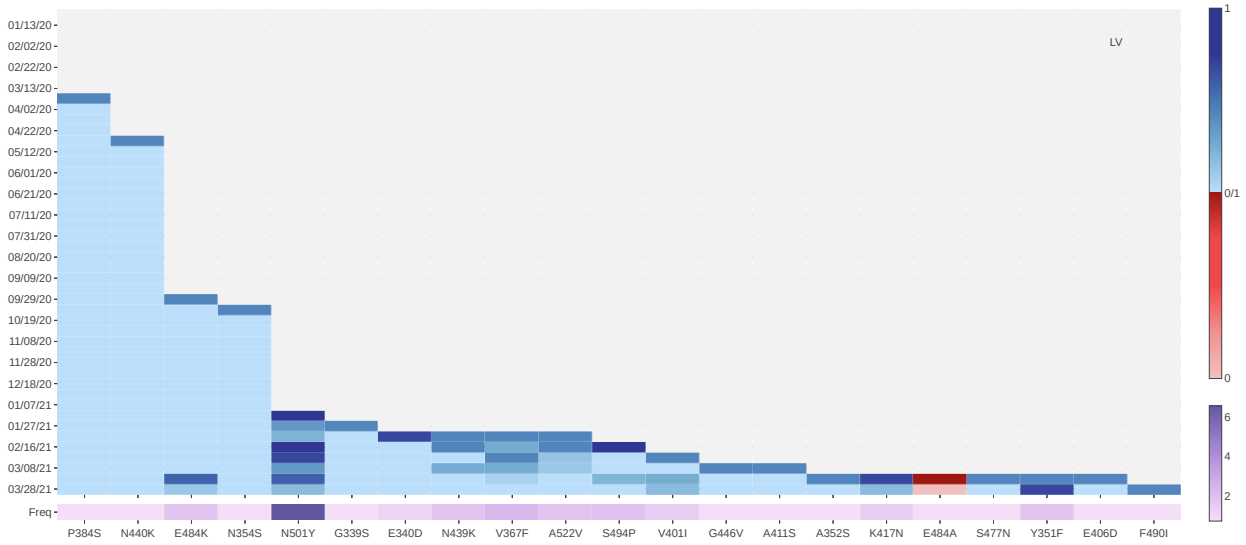

Figure S21: The log growth rate and log frequency of mutations on S protein RBD in the Latvia. The blue and red colors respectively represent the binding-strengthening and binding-weakening mutations on RBD. The darker blue/red means the binding-strengthening/binding-weakening mutations with a higher growth rate in a specific 10-day period. The darker purple represents the mutation with a higher log frequency.

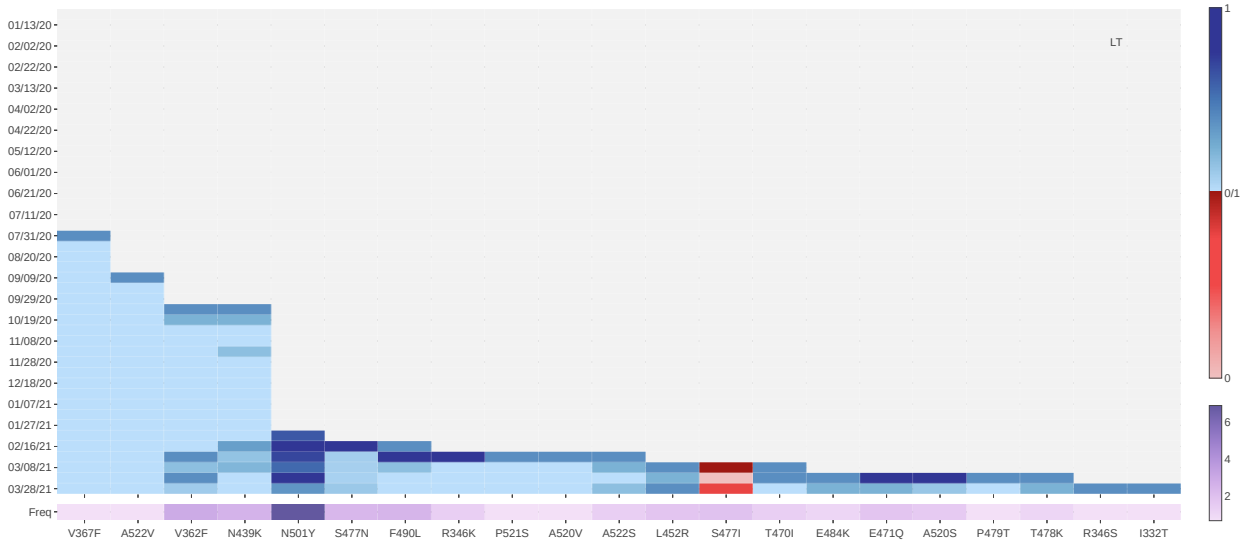

Figure S22: The log growth rate and log frequency of mutations on S protein RBD in the Lithuania. The blue and red colors respectively represent the binding-strengthening and binding-weakening mutations on RBD. The darker blue/red means the binding-strengthening/binding-weakening mutations with a higher growth rate in a specific 10-day period. The darker purple represents the mutation with a higher log frequency.

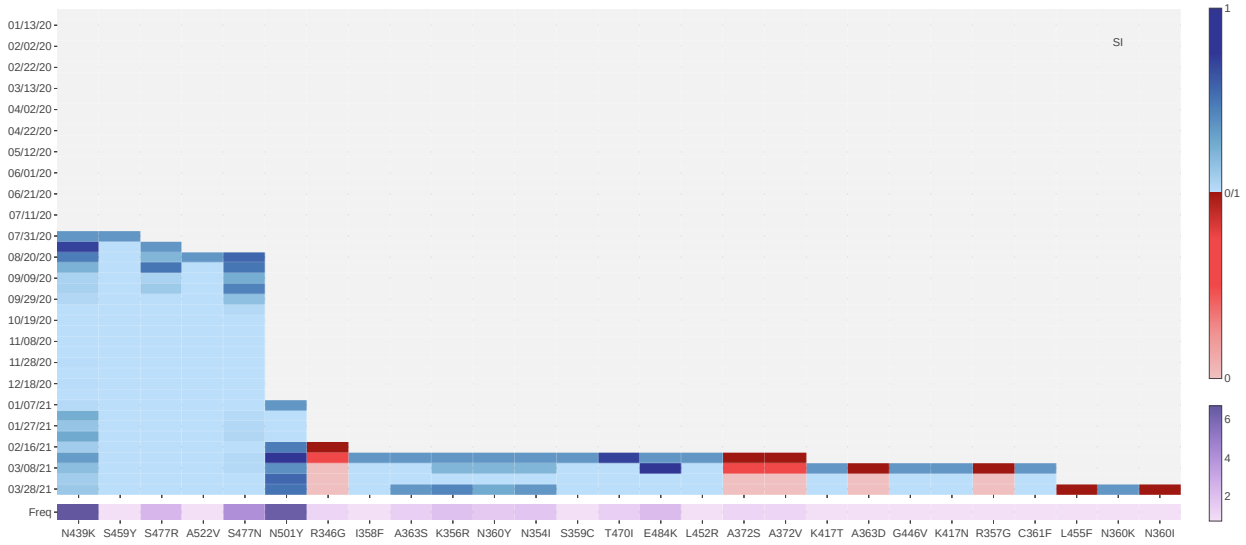

Figure S23: The log growth rate and log frequency of mutations on S protein RBD in the Slovenia. The blue and red colors respectively represent the binding-strengthening and binding-weakening mutations on RBD. The darker blue/red means the binding-strengthening/binding-weakening mutations with a higher growth rate in a specific 10-day period. The darker purple represents the mutation with a higher log frequency.

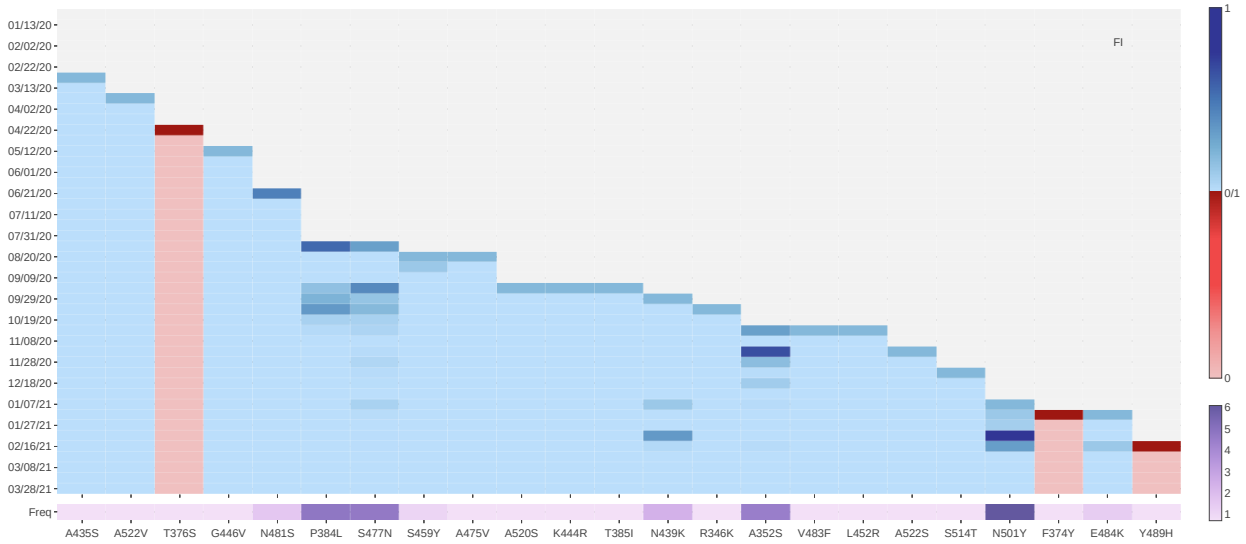

Figure S24: The log growth rate and log frequency of mutations on S protein RBD in the Finland. The blue and red colors respectively represent the binding-strengthening and binding-weakening mutations on RBD. The darker blue/red means the binding-strengthening/binding-weakening mutations with a higher growth rate in a specific 10-day period. The darker purple represents the mutation with a higher log frequency.

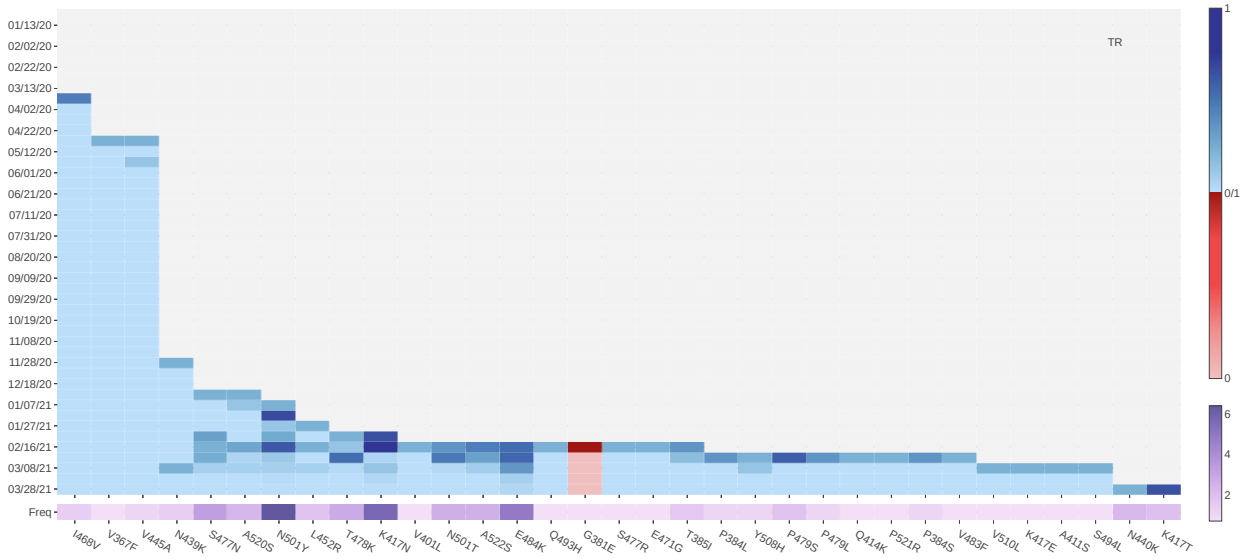

Figure S25: The log growth rate and log frequency of mutations on S protein RBD in the Turkey. The blue and red colors respectively represent the binding-strengthening and binding-weakening mutations on RBD. The darker blue/red means the binding-strengthening/binding-weakening mutations with a higher growth rate in a specific 10-day period. The darker purple represents the mutation with a higher log frequency.

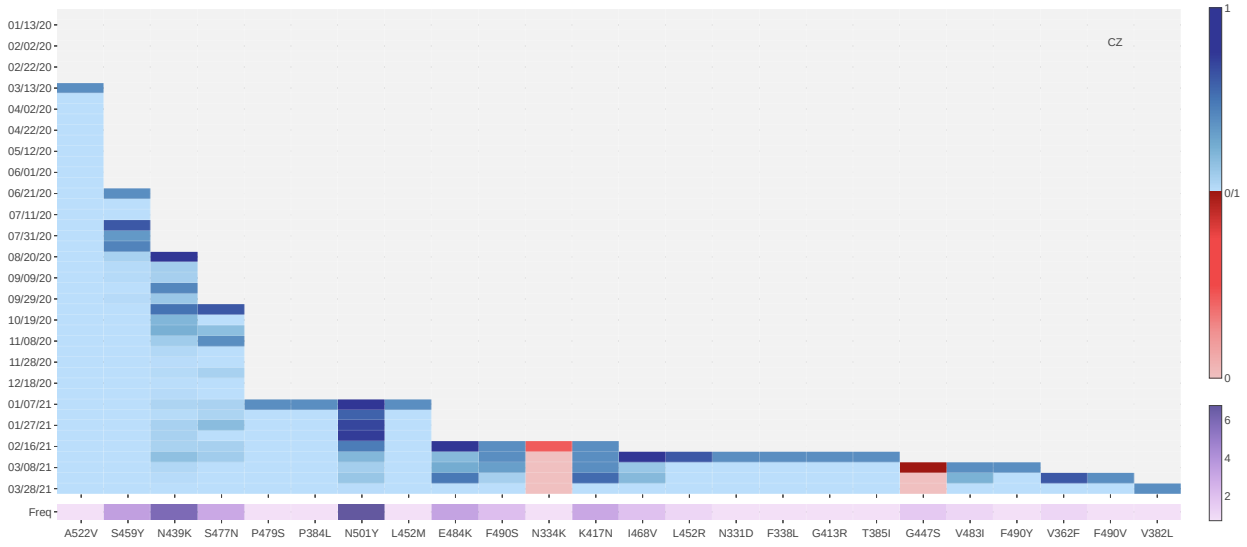

Figure S26: The log growth rate and log frequency of mutations on S protein RBD in the Czech Republic. The blue and red colors respectively represent the binding-strengthening and binding-weakening mutations on RBD. The darker blue/red means the binding-strengthening/binding-weakening mutations with a higher growth rate in a specific 10-day period. The darker purple represents the mutation with a higher log frequency.

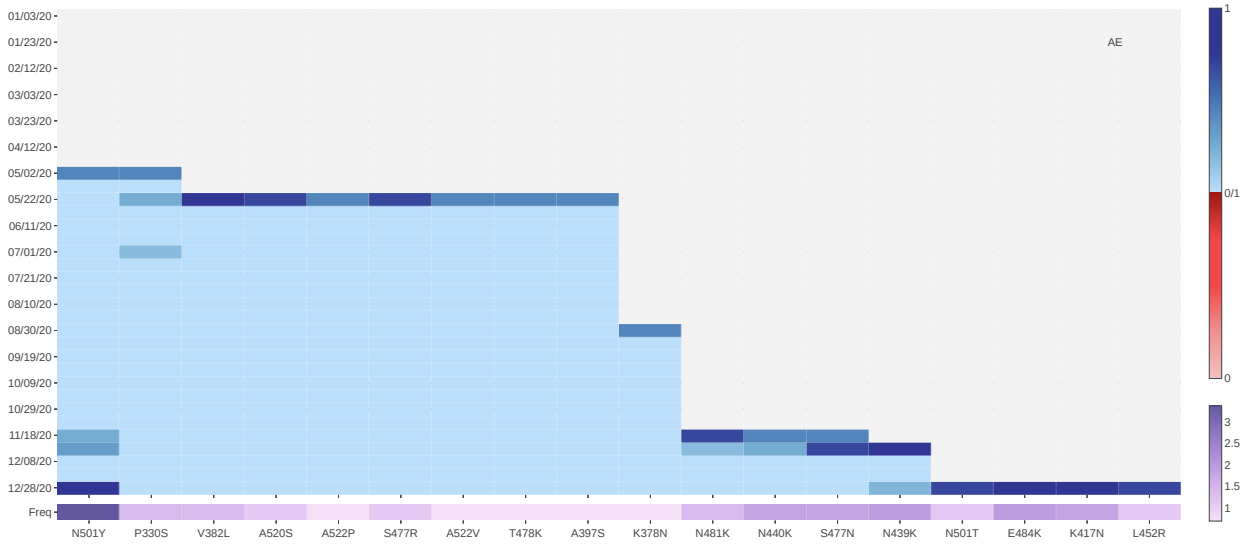

Figure S27: The log growth rate and log frequency of mutations on S protein RBD in the United Arab Emirates. The blue and red colors respectively represent the binding-strengthening and binding-weakening mutations on RBD. The darker blue/red means the binding-strengthening/binding-weakening mutations with a higher growth rate in a specific 10-day period. The darker purple represents the mutation with a higher log frequency.

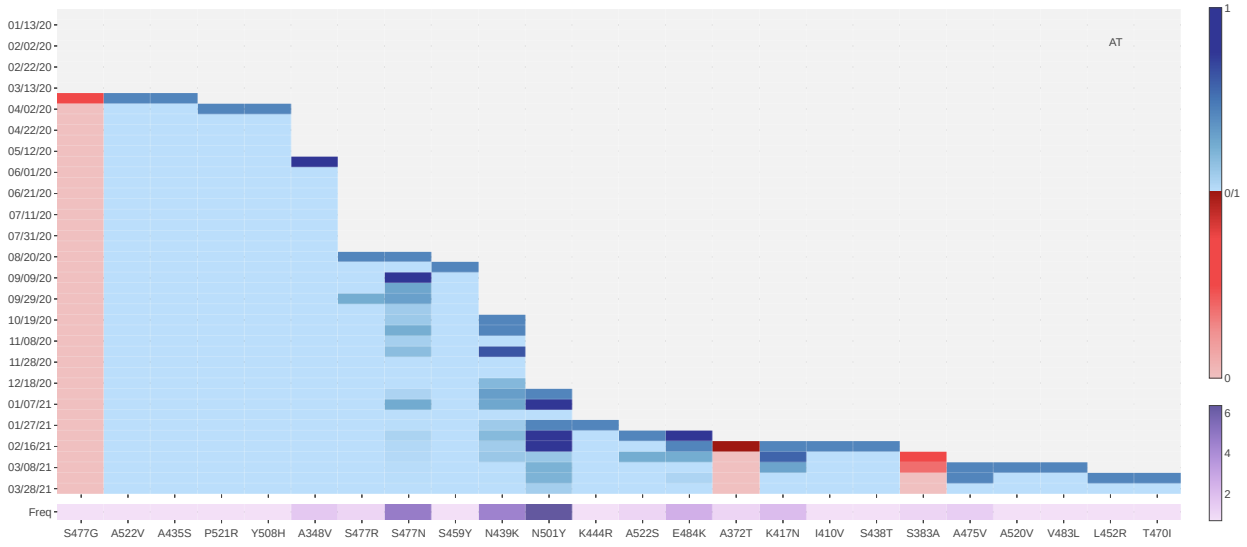

Figure S28: The log growth rate and log frequency of mutations on S protein RBD in the Austria. The blue and red colors respectively represent the binding-strengthening and binding-weakening mutations on RBD. The darker blue/red means the binding-strengthening/binding-weakening mutations with a higher growth rate in a specific 10-day period. The darker purple represents the mutation with a higher log frequency.