Figure S1. Percentage of nucleotides in 250-nt segments of the SARS-CoV genome

nucleotide G nucleotide T nucleotide C nucleotide A

Non-UTR regions of all the isolates are considered; positions (x-axis) are on the ClustalX output scale; percentage of nucleotides in the corresponding blocks are represented on y-axis
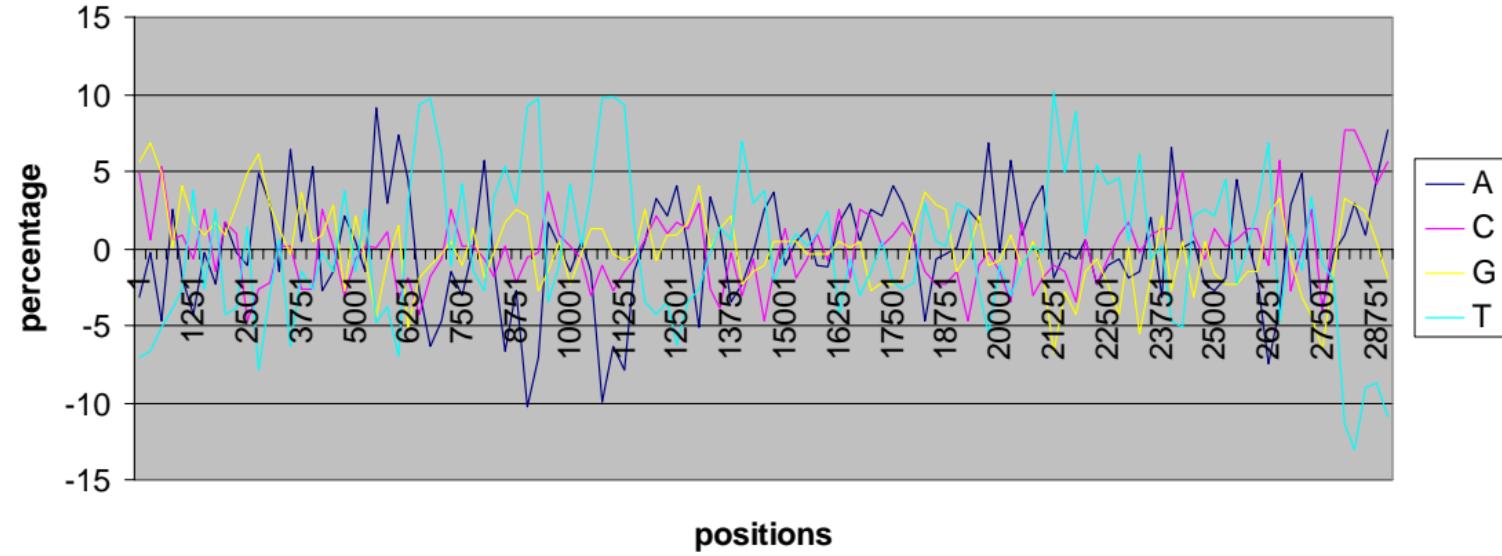
Figure S2. Deviation of percentages of nucleotides over 250-nt blocks from the corresponding percentages in the whole dataset
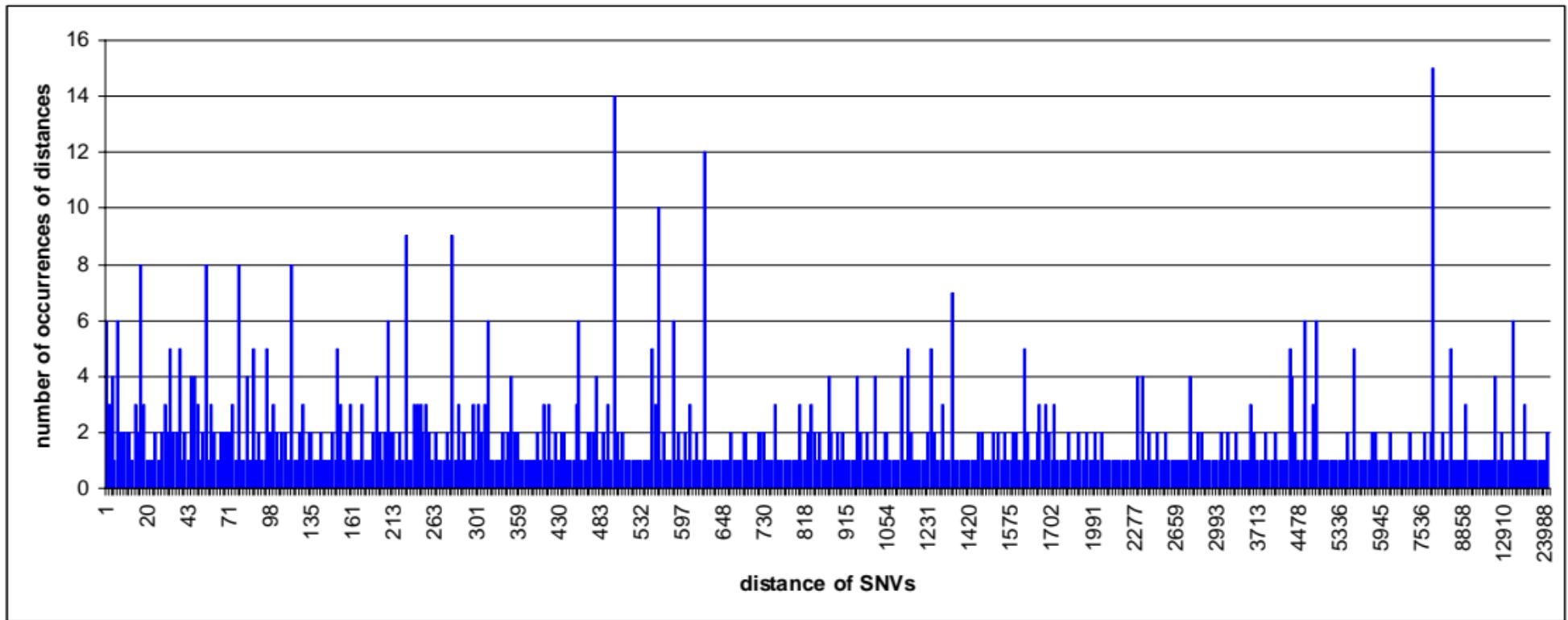
Figure S3. Distribution of distances between the neighboring SNVs, and numbers of their occurrences
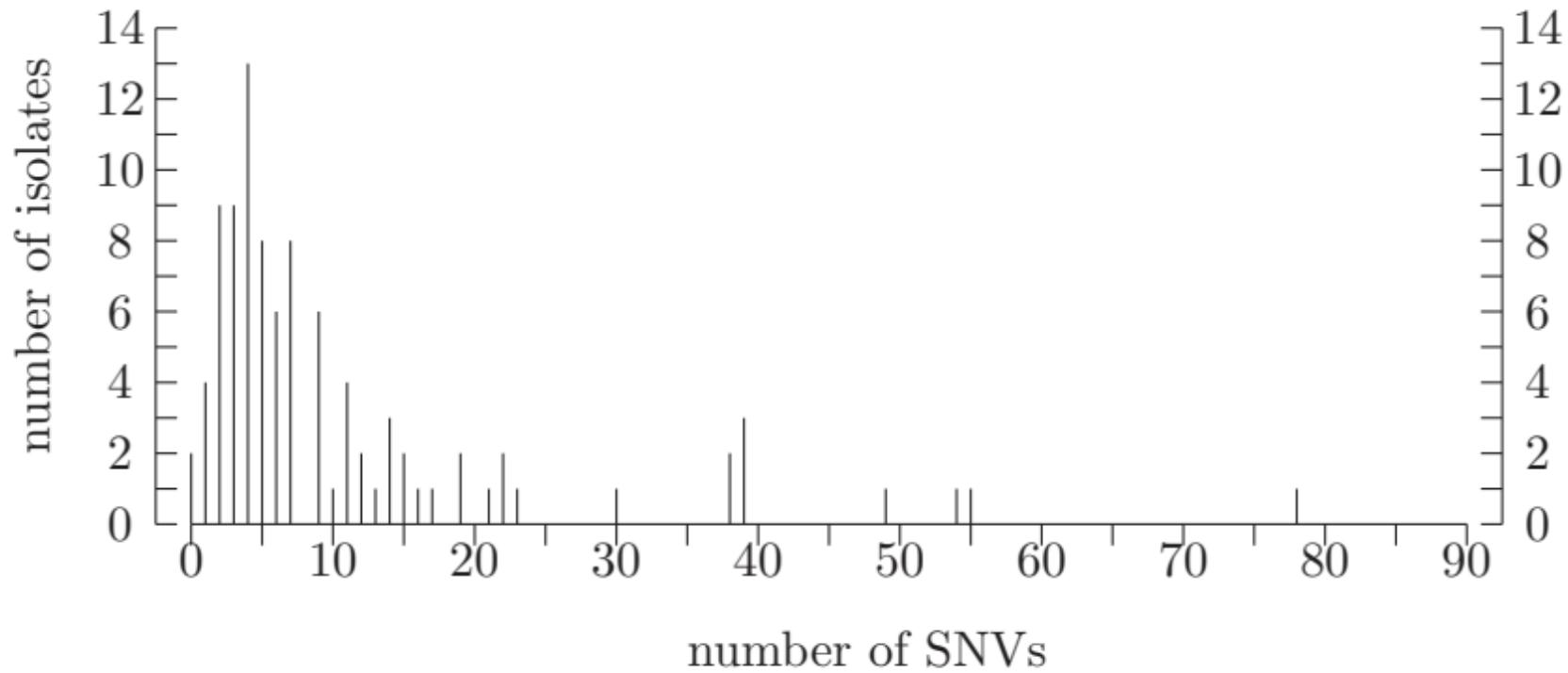
**Figure S4.** Distrubution of isolates per number of SNVs
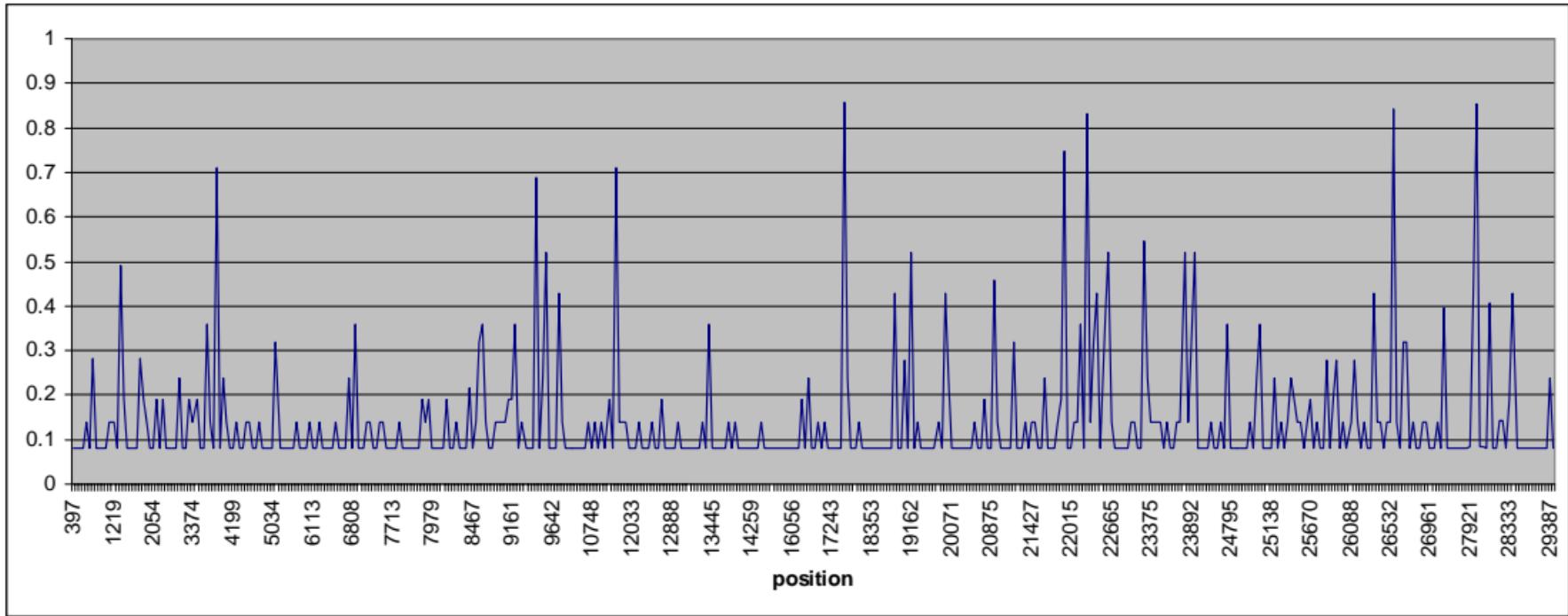
Figure S5. Entropy of genome nucleotide positions
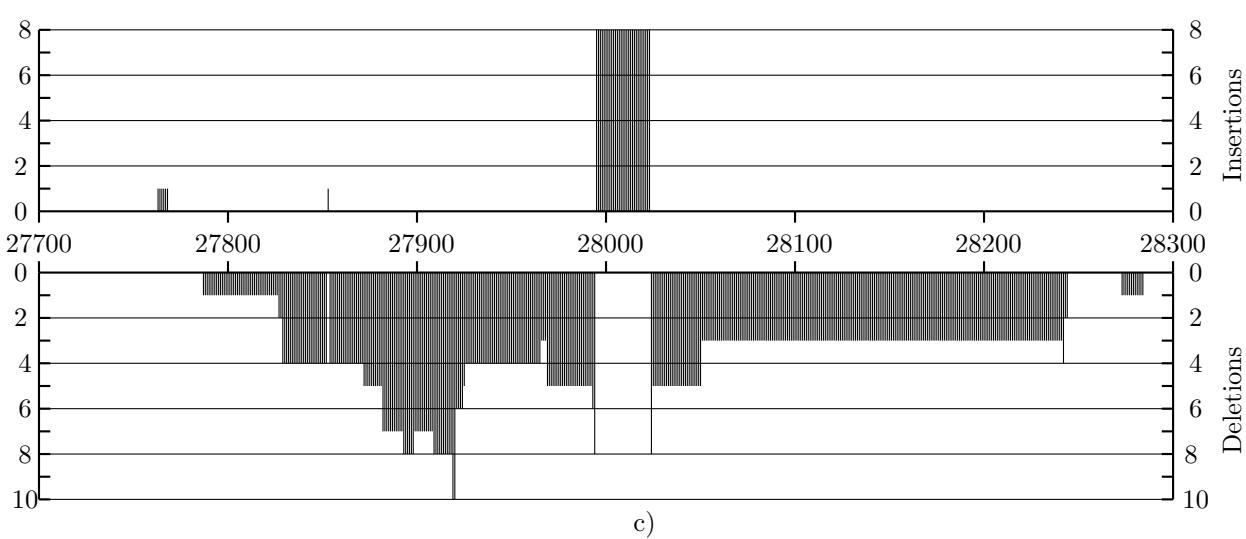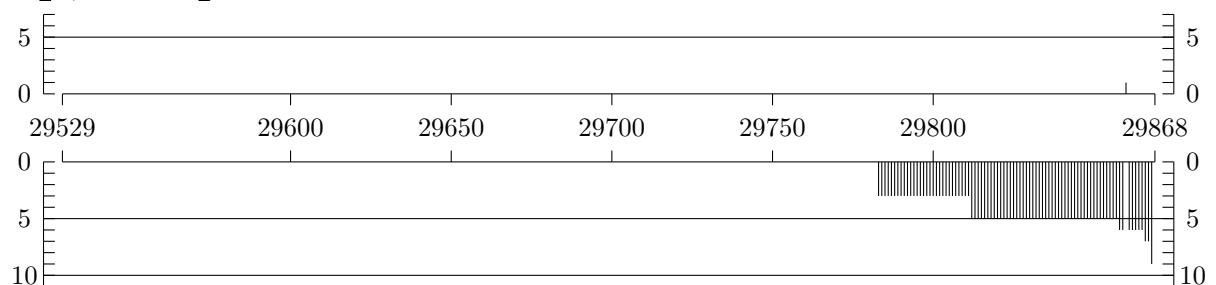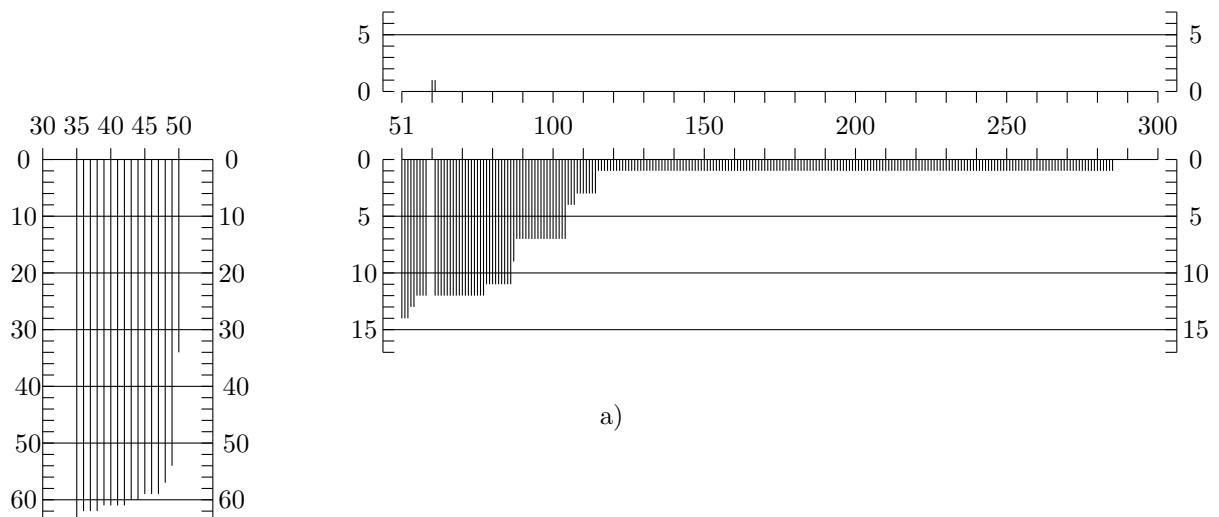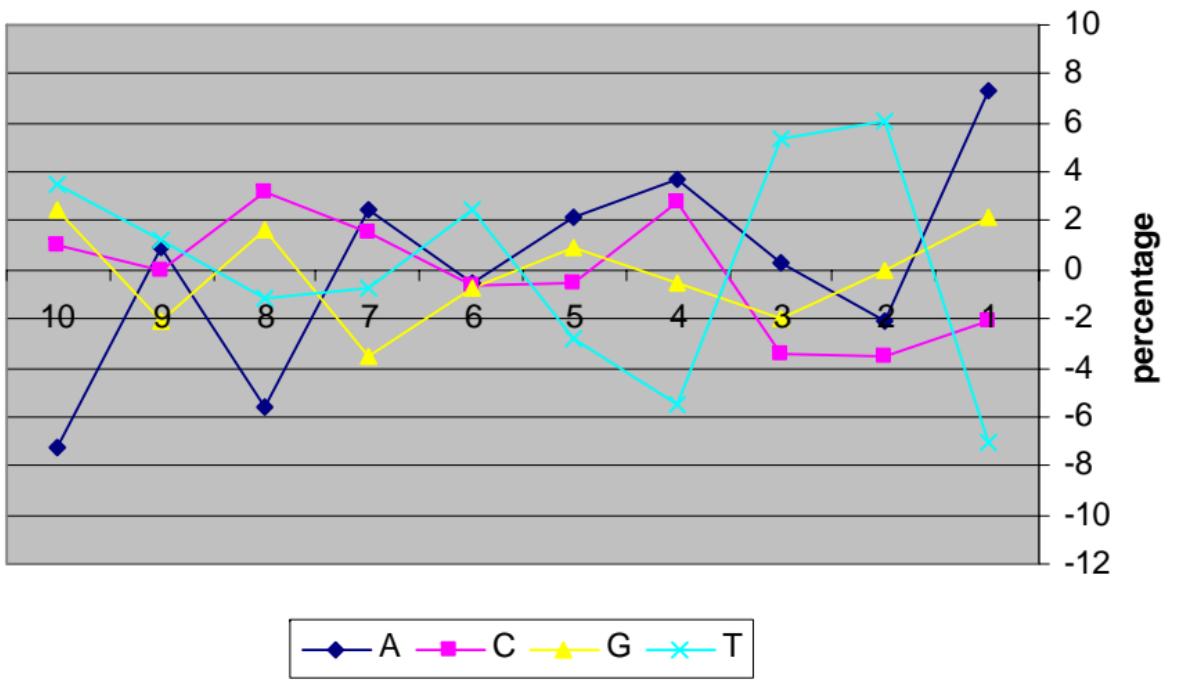
**Figure S6.** Distribution of INDELs in 5'UTR (a), 3'UTR (b) and in interval 27700-28300 (c)
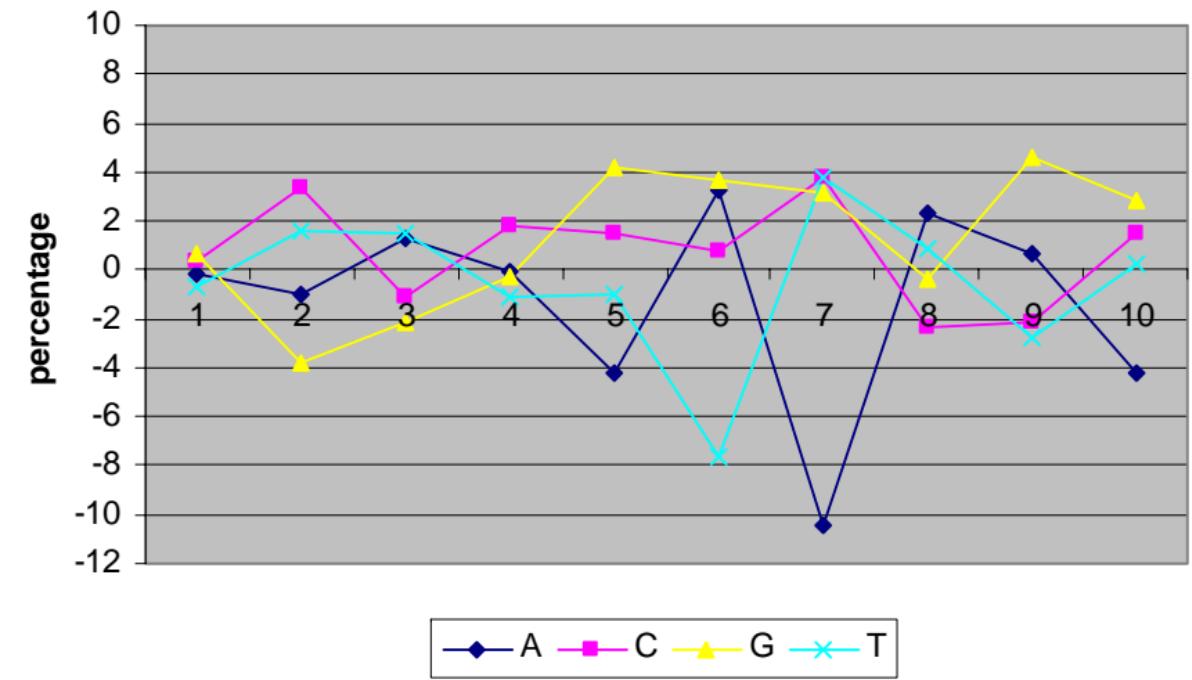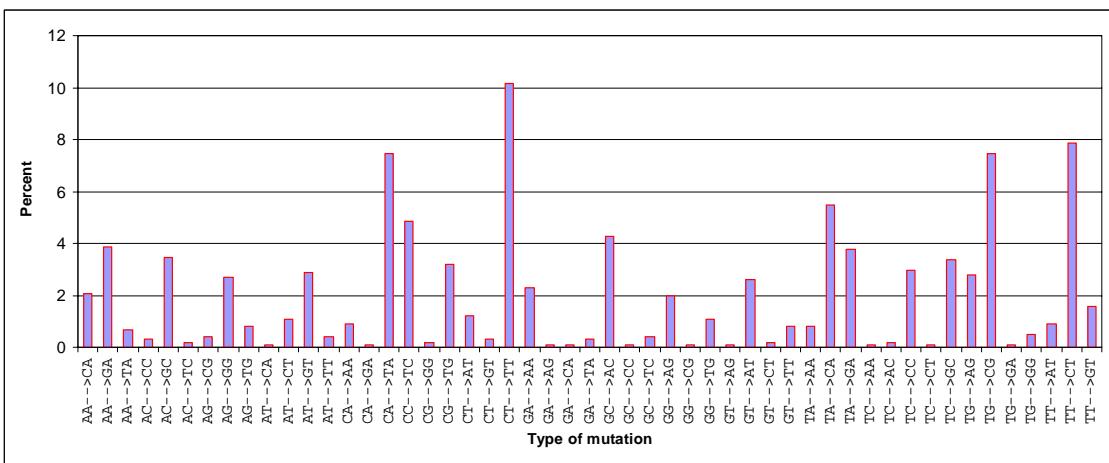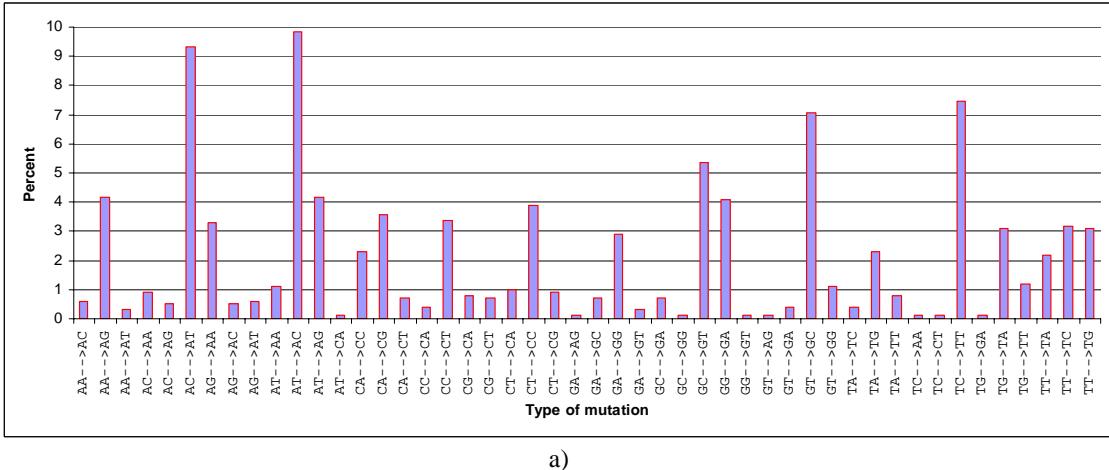
Figure S7. Differences between the percentage of nucleotides at a given position and in the whole genome, for up to the distance 10 left and right from SNV sites

a)

b)

Figure S8. Distribution of substitutions preceded by different nucleotide bases (a) and followed by different nucleotide bases (b)

Envelope protein

Membrane protein

Unknown (putative) orfs

Nucleocapside protein

orf1ab

Spike protein

Total in proteins
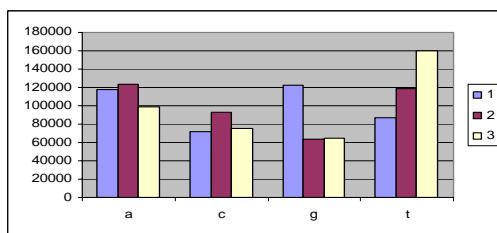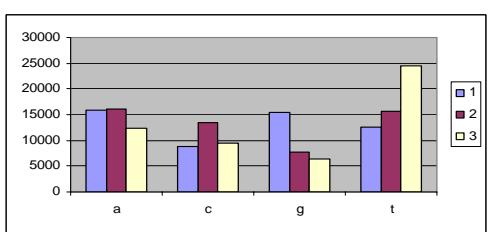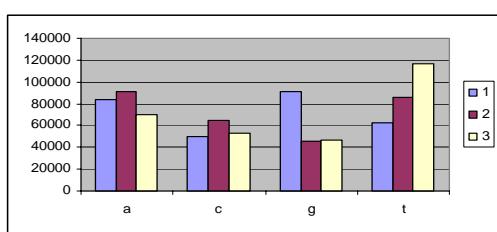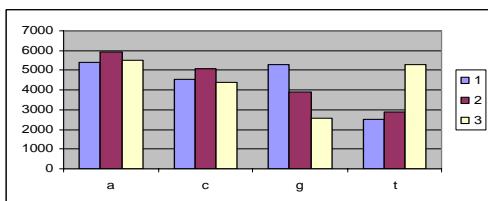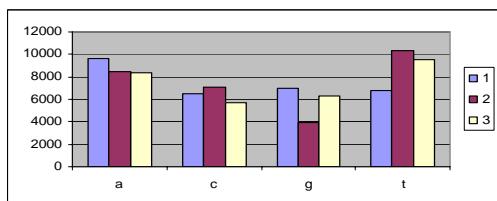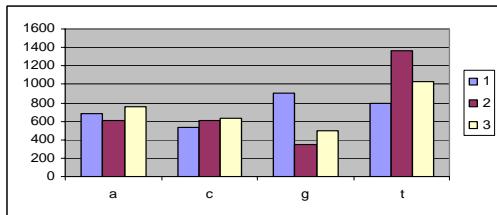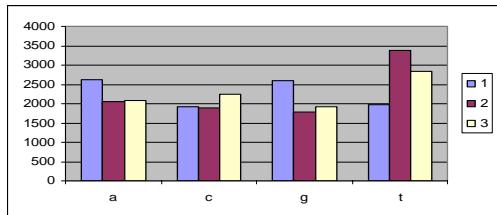
Figure S9. Distribution of nucleotides over the three codon positions in specific proteins and in total
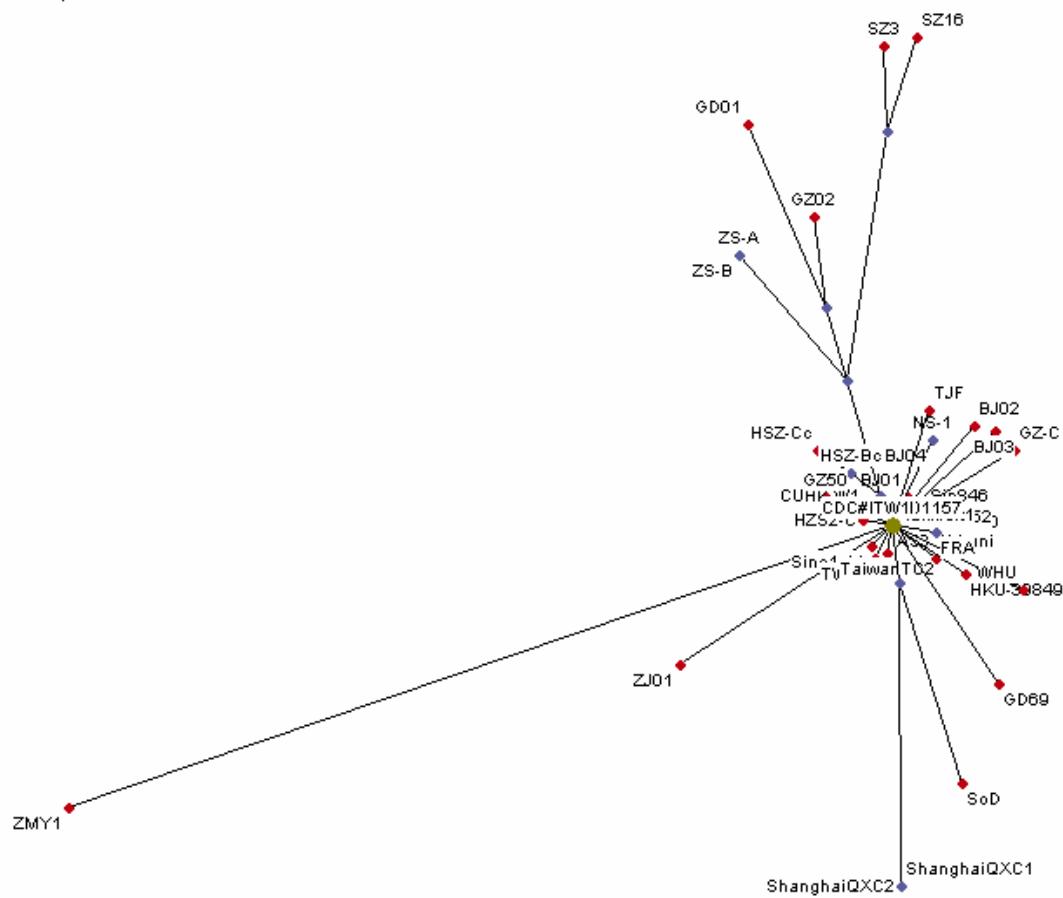
Figure S10. Root distances for isolates from the dataset

The isolates ZMY1, ZJ01, and the two Shanghai isolates have the largest root distance in the A group (the south side of the picture), and thus may not belong to it. (The isolates SoD and GD69 also have large root distances, but thay have been moved into the A group on the basis of their SNV sites)

Table S1. List of all the isolates considered. It includes identifiers, accession numbers, revision dates, country and source of the isolates considered, labels to be referred to in this paper, as well as the length of isolates and the ambiguous nucleotide codes. The labels are assigned in the (approximate) order of the first submission of isolates

| Label	ID	Accession No.	Length	Revision date	Country/Source	GI	FirstSubmitted	No. ambigBases
1	Tor2	NC_004718.3	29751	30-SEP-2004	Canada: Toronto, patient #2	30271926	13-APR-2003	
2		AY274119.3		24-MAR-2004	Canada: Toronto, patient #2	30248028	23-APR-2003	
3	Urbani	AY278741.1	29727	12-AUG-2003	USA: Atlanta cell_line="Vero"	30027617	17-APR-2003	
4	CUHK-W1	AY278554.2	29736	31-JUL-2003	China: Hong Kong	30027610	17-APR-2003	
5	BJ01	AY278488.2	29725	01-MAY-2003	China: Beijing	30275666	17-APR-2003	
6	BJ02	AY278487.3	29745	05-JUN-2003	China: Beijing	31416292	17-APR-2003	
7	BJ03	AY278490.3	29740	05-JUN-2003	China: Beijing	31416305	17-APR-2003	
8	BJ04	AY279354.2	29732	05-JUN-2003	China: Beijing	31416306	19-APR-2003	
9	NS-1	AY508724.1		17-JAN-2004		40795744	18-DEC-2003	
10	GD01	AY278489.2	29757	18-AUG-2003	China: Beijing	31416290	17-APR-2003	
11	HKU-39849	AY278491.2	29742	29-AUG-2003	China: Hong Kong	30023963	18-APR-2003	
12	CUHK-Su10	AY282752.2	29736	17-NOV-2003	China: Hong Kong	30421451	24-APR-2003	
13	Sin2500	AY283794.1	29711	12-AUG-2003	Singapore	30468042	27-APR-2003	
14	Sin2679	AY283796.1	29711	12-AUG-2003	Singapore	30468044	27-APR-2003	
15	Sin2774	AY283798.2	29711	02-OCT-2003	Singapore	37361915	27-APR-2003	
16	Sin2677	AY283795.1	29705	12-AUG-2003	Singapore	30468043	27-APR-2003	
17	Sin2748	AY283797.1	29706	12-AUG-2003	Singapore	30468045	27-APR-2003	N=1
18	Frankfurt 1	AY291315.1	29727	16-MAR-2004	Germany: Frankfurt	31581502	06-MAY-2003	
19	FRA	AY310120.1	29740	12-DEC-2003	Germany: patient from Frankfurt	33578015	29-MAY-2003	
20	ZJ01	AY297028.1	29715	19-MAY-2003	China: Beijing	30910859	12-MAY-2003	M=1
21	SZ3	AY304486.1	29741	05-NOV-2003	China: Hong Kong	34482137	26-MAY-2003	
22	SZ16	AY304488.1	29731	05-NOV-2003	China: Hong Kong	34482142	27-MAY-2003	
23	GZ50	AY304495.1	29720	05-NOV-2003	China: Hong Kong	34482146	27-MAY-2003	
24	GD69	AY313906.1	29754	15-DEC-2003	China: Jiangmen, Guangdong	37960831	03-JUN-2003 isolated in May 2003	
25	TWC	AY321118.1	29725	26-JUN-2003	Taiwan, first fatal case	31873092	11-JUN-2003	
26	HSR 1	AY323977.2	29751	15-OCT-2003	Italy	33115118	16-JUN-2003	R=1
27	Taiwan TC1	AY338174.1	29573	28-JUL-2003	Taiwan	32493129	08-JUL-2003	
28	Taiwan TC2	AY338175.1	29573	28-JUL-2003	Taiwan	32493130	09-JUL-2003	
29	Taiwan TC3	AY348314.1	29573	29-JUL-2003	Taiwan	33188324	23-JUL-2003	
30	CUHK-AG01	AY345986.1	29736	29-NOV-2003	Hong Kong: patient #1 of the Amoy Gardens cohort	33114190	18-JUL-2003	
31	CUHK-AG02	AY345987.1		29-NOV-2003	patient #2 of the Amoy Gardens cohort	33114202	18-JUL-2003	
32	CUHK-AG03	AY345988.1	29736	29-NOV-2003	Hong Kong: patient #3 of the Amoy Gardens cohort	33114214	18-JUL-2003	
33	PUMC01	AY350750.1	29738	17-NOV-2003	China: Beijing	38231927	24-JUL-2003	
34	PUMC02	AY357075.1	29738	17-NOV-2003	China: Beijing	38231932	31-JUL-2003	
35	PUMC03	AY357076.1	29745	17-NOV-2003	China: Beijing	38231937	31-JUL-2003	
36	ZMY 1	AY351680.1	29749	03-AUG-2003	China: Guangdong	33304219	28-JUL-2003	
37	TWH	AP006557.1	29727	02-AUG-2003	Taiwan: patient #01 – cultured	33411399	30-JUL-2003	
38	WC2	AY362698.1		13-AUG-2003	Taiwan: Hoping Hospital	33518724	05-AUG-2003	
39	TWK	AP006559.1	29727	02-AUG-2003	Taiwan: patient #06 - primary	33411429	30-JUL-2003	
40	TWS	AP006560.1	29727	02-AUG-2003	Taiwan: patient #04 - primary	33411444	30-JUL-2003	
41	TWY	AP006561.1	29727	02-AUG-2003	Taiwan: patient #02 - primary	33411459	30-JUL-2003	
42	TWC3	AY362699.1	29727	13-AUG-2003	Taiwan: Hoping Hospital – throat swab	33518725	05-AUG-2003	
43	TWJ	AP006558.1	29725	02-AUG-2003	Taiwan: patient #043 Primary	33411414	30-JUL-2003	
44	GZ02	AY390556.1	29760	31-JAN-2004	China: Guangzhou	41323719	15-SEP-2003	
45	WHU	AY394850.2	29728	07-JUN-2004	China: Wuhan	40795428	18-SEP-2003	
46	HZS2-D	AY394989.1	29736	29-JAN-2004	China: Guangzhou	37624332	19-SEP-2003	
47	HZS2-E	AY394990.1	29736	29-JAN-2004	China: Guangzhou	37624333	19-SEP-2003	
48	HZS2-Fc	AY394991.1	29736	29-JAN-2004	China: Guangzhou	37624334	19-SEP-2003	
49	HZS2-C	AY394992.1	29736	29-JAN-2004	China: Guangzhou	37624335	19-SEP-2003	
50	HGZ8L2	AY394993.1	29736	29-JAN-2004	China: Guangzhou	37624336	19-SEP-2003	
51	LC1	AY394998.1	29736	29-JAN-2004	China: Guangzhou	37624341	19-SEP-2003	
52	GZ-B	AY394978.1	29640	29-JAN-2004	China: Guangzhou	37624321	19-SEP-2003	
53	GZ-C	AY394979.1	29645	29-JAN-2004	China: Guangzhou	37624322	19-SEP-2003	
54	HSZ2-A	AY394983.1	29699	29-JAN-2004	China: Guangzhou	37624326	19-SEP-2003	
55	HZS2-Fb	AY394987.1	29709	29-JAN-2004	China: Guangzhou	37624330	19-SEP-2003	
56	HSZ-Bb	AY394985.1	29530	29-JAN-2004	China: Guangzhou	37624328	19-SEP-2003	
57	HSZ-Cb	AY394986.1	29729	29-JAN-2004	China: Guangzhou	37624329	19-SEP-2003	
58	HSZ-Bc	AY394994.1	29765	29-JAN-2004	China: Guangzhou	37624337	19-SEP-2003	
59	HSZ-Cc	AY394995.1	29765	29-JAN-2004	China: Guangzhou	37624338	19-SEP-2003	
60	ZS-B	AY394996.1	29683	29-JAN-2004	China: Guangzhou	37624339	19-SEP-2003	
61	ZS-A	AY394997.1	29683	29-JAN-2004		37624340	19-SEP-2003	
62	ZS-C	AY395003.1	29647	29-JAN-2004	China: Guangzhou	37624346	19-SEP-2003	
63	LC2	AY394999.1	29350	29-JAN-2004	China: Guangzhou	37624342	19-SEP-2003	
64	LC3	AY395000.1	29350	29-JAN-2004	China: Guangzhou	37624343	19-SEP-2003	
65	LC4	AY395001.1	29350	29-JAN-2004		37624344	19-SEP-2003	
66	LC5	AY395002.1	29350	29-JAN-2004	China: Guangzhou	37624345	19-SEP-2003	
67	AS	AY427439.1	29711	21-OCT-2003	Italy: Milan	37576845	02-OCT-2003	
68	SoD	AY461660.1	29715	23-NOV-2003	Russia	38385714	31-OCT-2003	
69	ShanghaiQXC1	AY463059.1	29592	05-JAN-2004	Shanghai	40457433	11-NOV-2003	M=1, R=1, Y=1
70	ShanghaiQXC2	AY463060.1	29013	05-JAN-2004	Shanghai	40457448	11-NOV-2003	R=1, Y=1
71	Sino1-11	AY485277.1	29741	30-NOV-2003	China: Beijing	38505482	21-NOV-2003	
72	Sino3-11	AY485278.1	29740	30-NOV-2003	China: Beijing	38505491	21-NOV-2003	
73	TW1	AY291451.1	29729	25-FEB-2004	Taiwan	30698326	06-MAY-2003	
74	TW2	AY502925.1		25-FEB-2004	Taiwan	40548897	15-DEC-2003	
75	TW3	AY502926.1	29729	25-FEB-2004	Taiwan	40548909	15-DEC-2003	M=1, Y=1
76	TW4	AY502927.1	29729	25-FEB-2004	Taiwan	40548921	15-DEC-2003	R=1, W=1
77	TW5	AY502928.1	29729	25-FEB-2004	Taiwan	40548933	15-DEC-2003	
78	TW6	AY502929.1	29729	25-FEB-2004	Taiwan	40548945	15-DEC-2003	
79	TW7	AY502930.1	29729	25-FEB-2004	Taiwan	40548957	15-DEC-2003	
80	TW8	AY502931.1	29729	25-FEB-2004	Taiwan	40548969	15-DEC-2003	M=1
81	TW9	AY502932.1	29729	25-FEB-2004	Taiwan	40548981	15-DEC-2003	R=1
82	TW10	AY502923.1	29729	25-FEB-2004	Taiwan	40548873	15-DEC-2003	

Label	ID	Accession No.	Length	Revision date	Country/Source	GI	FirstSubmitted	No. ambigBases
83	TW11	AY502924.1	29727	25-FEB-2004	Taiwan	40548885	15-DEC-2003	W=1, Y=2
84	Sin842	AY559081.1	29716	24-SEP-2004	Singapore	45644994	24-FEB-2004	
85	Sin852	AY559082.1	29670	24-SEP-2004	Singapore	45644996	24-FEB-2004	
86	Sin3765V	AY559084.1	29722	24-SEP-2004	Singapore	45645000	24-FEB-2004	
87	Sin848	AY559085.1	29713	24-SEP-2004	Singapore	45645001	24-FEB-2004	
88	Sin849	AY559086.1	29661	24-SEP-2004	Singapore	45645003	24-FEB-2004	
89	Sin846	AY559094.1	29577	24-SEP-2004	Singapore	45645021	24-FEB-2004	
90	Sin3725V	AY559087.1	29716	24-SEP-2004	Singapore	45645004	24-FEB-2004	Y=5
91	SinP1	AY559088.1	29714	24-SEP-2004	Singapore	45645007	24-FEB-2004	R=1
92	SinP3	AY559090.1	29725	24-SEP-2004	Singapore	45645013	24-FEB-2004	R=1
93	SinP5	AY559092.1	29713	24-SEP-2004	Singapore	45645017	24-FEB-2004	R=1
94	SinP4,	AY559091.1	29710	24-SEP-2004	Singapore	45645016	24-FEB-2004	R=1
95	Sin845	AY559093.1	29712	24-SEP-2004	Singapore	45645019	24-FEB-2004	
96	Sin847	AY559095.1	29719	24-SEP-2004	Singapore	45645022	24-FEB-2004	
97	Sin850	AY559096.1	29720	24-SEP-2004	Singapore	45645023	24-FEB-2004	
98	LLJ-2004	AY595412.1	29716	29-JUN-2004	"aerosol sample" China: Beijing	49176846	08-APR-2004	K=3, M=1, R=5, S=1, W=1
99	TJF	AY654624.1	29745	24-JUL-2004	China	50365700	16-JUN-2004	
100	CDC#200301157	AY714217.1	29727	28-SEP-2004	USA	52546959	10-AUG-2004	
101	Sin3408	AY559083.1	29767	24-SEP-2004	Singapore	45644998	24-FEB-2004	N=263, Y=1
102	SinP2	AY559089.1	29717	24-SEP-2004	Singapore	45645010	24-FEB-2004	N=30, R=1, S=1
103	Sin3408L	AY559097.1	29715	24-SEP-2004	Singapore	45645024	24-FEB-2004	N=105, W=3, Y=1

Table S2. SARS-CoV genome polymorphism (extended table). All the 103 isolates are represented. Shaded entries correspond to annotated isolates. Last three isolates contain large number of ambiguous nucleotide codes (N).

IDENTIFICATION of isolates is given in accordance to the Labels, IDs and Accession numbers from the table S1. The four **SNVs** columns correspond to the total number of SNVs, number of SNVs in genes, number of SNVs in 5' and 3' UTRs, and the number of SNVs in intergenic regions (IGR). The eight columns named **INDELs** include number of deletions at the 5' end (at **first 15** genome positions, where most of the isolates are empty, and at the **following 5'UTR** positions, where most of the isolates are non-empty); the length and position of long insertions and long deletions (**longIns**, **longDel**), both absolute for the isolate in question and relative to the CLUSTAL output; number and length of short insertions and short deletions (**shortIns**, **shortDel**) in the form *axb* where *b* denotes the length and *a* the number of occurrences, along with their absolute and relative positions; number of deletions at the 3'end (**3'del**) and the length of a poly-a sequence at the 3' end (**3'poly-a**). **CLASSIFICATION** includes two columns. The **Type** column corresponds to the nine loci nucleotides that are given in the form NNNN/NNNNN and represent nucleotides at (relative to CLUSTAL output) positions 9420, 17604, 222274, 27891 and 3861, 9495, 11514, 21773, 26534, respectively (absolute Tor2 positions 9404, 17564, 22222, 27827 / 3852, 9479, 11493, 21721, 26477). The last column, **Group,** reflects the proposed grouping of isolates.

IDENTIFICATION			SNVs				INDELs								CLASSIFICATION	
Label	ID	Accession No.	Total	Genes	5'/3' UTR	IGR	5' Del -first 15	5' Del -follow.	longIns	longDel	shortIns	shortDel	3'del	3'poly-a	Type	Group
1,2	Tor2	Nc_004718.3	2	2	-/-	-	-	-	-	-	-	-	-	24	tttt/ttcgt	A1
3	Urbani	Ay278741.1	5	5	-/-	-	-	-	-	-	-	-	-	-	tttt/ttcgt	A1
4	CUHK-W1	Ay278554.2	9	8	-/-	1	15	-	-	-	-	-	-	24	cgcc/ttcat	B1
5	BJ01	Ay278488.2	12	11	-/-	1	15	4	-	-	-	-	-	17	cgcc/ttcat	B1
6	BJ02	Ay278487.3	22	22	-/-	-	-	-	-	-	-	-	-	18	cgcc/ttcat	B1
7	BJ03	Ay278490.3	22	22	-/-	-	4	-	-	-	-	-	-	17	cgcc/ttcat	B1
8,9	NS-1(BJ04)	AY508724.1	15	14	-/-	1	15	1	-	-	-	-	-	21	tgcc/ttcgt	B1
10	GD01	Ay278489.2	49	49	-/-	-	15	1	29 (27868/ 27995)	-	-	-	-	17	cgcc/tccat	B2
11	HKU-39849	Ay278491.2	9	9	-/-	-	-	-	-	-	-	-	-	15	tttt/ttcgt	A1
12	CUHK-Su10	Ay282752.2	2	1	-/-	1	15	-	-	-	-	-	-	24	tttt/ttcgg	A1
13	Sin2500	Ay283794.1	2	2	-/-	-	15	1	-	-	-	-	-	-	tttt/ttcgt	A1
14	Sin2679	Ay283796.1	2	2	-/-	-	15	1	-	-	-	-	-	-	tttt/ttcgt	A1
15	Sin2774	Ay283798.2	4	4	-/-	-	15	1	-	-	-	-	-	-	tttt/ttcgt	A1
16	Sin2677	Ay283795.1	3	3	-/-	-	15	1	-	-	-	1x6 (27766 / 27893)	-	-	tttt/ttcgt	A1
17	Sin2748	Ay283797.1	1	1	-/-	-	15	1	-	-	-	1x5 (27794 / 27921)	-	-	tttt/ttcgt	A1
18	Frankfurt 1	Ay291315.1	7	7	-/-	-	-	-	-	-	-	-	-	-	tttt/ttcgt	A1
19	FRA	Ay310120.1	7	7	-/-	-	-	-	-	-	-	-	-	13	tttt/ttcgt	A1
20	ZJ01	Ay297028.1	23	23	-/-	-	14	-	-	-	7x1 (8548,12975, 13374,13450, 13462,15558, 27734 / 8612,13051, 13451,13527, 13539,15644, 27853)	2x1 (15526, 28465 / 15611, 28613 /)	3	-	tttt/ttcgt	B4
21	SZ3	Ay304486.1	54	53	-/1	-	15	-	29 (27869 / 27995)	-	-	-	-	-	cgcc/tccat	B2
22	SZ16	Ay304488.1	55	55	-/-	-	15	-	29 (27869 / 27995)	-	-	-	10	-	cgcc/tccat	B2
23	GZ50	Ay304495.1	11	10	-/1	-	15	-	-	-	-	-	-	8	tgcc/ttcat	B1
24	GD69	Ay313906.1	21	21	-/-	-	-	-	-	-	1x1 (14295 / 14370), 1x10 (20374/ 20456)	-	-	16	tttt/cttgg	A1
25	TWC	Ay321118.1	2	2	-/-	-	-	-	-	-	-	1x2 (27806 / 27919)	-	-	tttt/ttcgt	A1
26	HSR 1	Ay323977.2	0	0	-/-	-	-	-	-	-	-	-	-	24	tttt/ttcgt	A1
27	Taiwan TC1	Ay338174.1	4	4	-/-	-	15	54	-	-	-	-	85	-	tttt/cttgg	A1
28	Taiwan TC2	Ay338175.1	9	9	-/-	-	15	54	-	-	-	-	85	-	tttt/cttgg	A1
29	Taiwan TC3	Ay348314.1	7	6	-/-	1	15	54	-	-	-	-	85	-	tttt/cttgg	A1

IDENTIFICATION			SNVs				INDELs								CLASSIFICATION	
Label	ID	Accession No.	Total	Genes	5'/3' UTR	IGR	5' Del -first 15	5' Del -follow.	longIns	longDel	shortIns	shortDel	3'del	3'poly-a	Type	Group
30,31	CUHK-AG01(02)	AY345986.1	3	3	-/-	-	15	-	-	-	-	-	-	24	tttt/cttgg	A1
32	CUHK-AG03	AY345988.1	5	4	-/-	1	15	-	-	-	-	-	-	24	tttt/cttgg	A1
33	PUMC01	AY350750.1	3	2	-/-	1	13	-	-	-	-	-	-	24	tttt/ttcgg	A1
34	PUMC02	AY357075.1	2	1	-/-	1	14	-	-	-	-	1x2 (27869 / 27994 (1), 28024 (1))	-	27	tttt/ttcgg	A1
35	PUMC03	AY357076.1	4	3	-/-	1	14	-	-	-	-	1x3 (26142 / 26260)	-	35	tttt/ttcgg	A1
36	ZMY 1	Ay351680.1	78	77	1/-	-	-	-	-	-	24x1 (1031, 1041, 1053, 2521, 2528, 3850, 4220, 4826, 6339, 6426, 7798, 8084, 11729 11782, 14063, 14113, 14134,14177, 14261, 14271, 19118, 20692, 22750, 25301 / 1067, 1077, 1089, 2558, 2565, 3887, 4257, 4863, 6377, 6464, 7836, 8122, 11777,11830, 14120, 14170, 14191,14234, 14318, 14328, 19178,20766, 22825,25381)	2x1 (10601, 10609 / 10642, 10651) 1x2 (28734 / 28852)	-	2	tttt/ttcgt	B4
37,38	TWH (TWC2)	Ap006557.1	4	4	-/-	-	-	-	-	-	-	-	-	-	tttt/cttgg	A1
39	TWK	Ap006559.1	7	7	-/-	-	-	-	-	-	-	-	-	-	tttt/cttgg	A1
40	TWS	Ap006560.1	6	6	-/-	-	-	-	-	-	-	-	-	-	tttt/cttgg	A1
41	TWY	Ap006561.1	6	6	-/-	-	-	-	-	-	-	-	-	-	tttt/cttgg	A1
42	TWC3	Ay362699.1	3	3	-/-	-	-	-	-	-	-	-	-	-	tttt/cttgg	A1
43	TWJ	Ap006558.1	6	6	-/-	-	-	-	-	-	-	1x2 (27167 / 27271)	-	-	tttt/cttgg	A1
44	GZ02	AY390556.1	39	39	-/-	-	-	-	29 (27884 / 27995)	-	-	-	-	4	cgcc/tccat	B2
45	WHU	Ay394850.2	15	15	-/-	-	-	-	-	-	-	1x2 (27808 / 27919)	-	3	tttt/ttcgt	A1
46	HZS2-D	AY394989.1	5	5	-/-	-	15	-	-	-	-	-	-	24	tgcc/ttcat	A1
47	HZS2-E	AY394990.1	5	5	-/-	-	15	-	-	-	-	-	-	24	tgcc/ttcat	A1
48	HZS2-Fc	AY394991.1	6	6	-/-	-	15	-	-	-	-	-	-	24	tgcc/ttcgt	A1
49	HZS2-C	AY394992.1	7	7	-/	-	15	-	-	-	-	-	-	24	tgcc/ttcat	A1
50	HGZ8L2	AY394993.1	7	7	-/-	-	15	-	-	-	-	-	-	24	tgcc/ttcat	A1
51	LC1	AY394998.1	1	1	-/-	-	15	-	-	-	-	-	-	24	tttt/ttcgg	A1
52	GZ-B	AY394978.1	3	3	-/-	-	15	57	-	39 (27699 / 27882)	-	-	-	24	tttt/ttcgt	A3

	IDENTIFICATION			SNVs				INDELs								CLASSIFICATION	
Label	ID	Accession No.	Total	Genes	5'/3' UTR	IGR	5' Del -first 15	5' Del -follow.	longIns	longDel	shortIns	shortDel	3'del	3'poly-a	Type	Group
53	GZ-C	AY394979.1	14	14	-/-	-	15	37	-	39 (27719 / 27882), 12 (28039 / 28273)	-	1x3 (27791 / 27993 (2), 28024 (1))	-	24	cttt/ttcgt	A3
54	HSZ2-A	AY394983.1	5	5	-/-	-	15	37	-	-	-	-	-	24	tgcc/ttcat	A1
55	HZS2-Fb	AY394987.1	5	5	-/-	-	15	27	-	-	-	-	-	24	tgcc/ttcgt	A1
56	HSZ-Bb	AY394985.1	14	14	-/-	-	15	235	29 (27634 / 27995)	-	-	-	-	24	cgcc/tccat	B2
57	HSZ-Cb	AY394986.1	16	16	-/-	-	15	36	29 (27833 / 27995)	-	-	-	-	24	cgcc/tccat	B2
58	HSZ-Bc	AY394994.1	13	13	-/-	-	15	-	29 (27869 / 27995)	-	-	-	-	24	cgcc/tccat	B2
59	HSZ-Cc	AY394995.1	19	19	-/-	-	15	-	29 (27869 / 27995)	-	-	-	-	24	cgcc/tccat	B2
60,61	ZS-A (ZS-B)	AY394997.1	38	38	-/-	-	15	-	-	53 (27843 / 27969 (26), 28024 (27))	-	-	-	24	cgcc/tccat	B3
62	ZS-C	AY395003.1	38	38	-/-	-	15	36	-	53 (27807 / 27969 (26), 28024 (27))	-	-	-	24	cgcc/tccat	B3
63	LC2	AY394999.1	4	4	-/-	-	15	-	-	386 (27704 / 27829)	-	-	-	24	ttt-/ttcgt	A3
64,65	LC4 (LC3)	AY395001.1	3	3	-/-	-	15	-	-	386 (27704 / 27829)	-	-	-	24	ttt-/ttcgt	A3
66	LC5	AY395002.1	4	4	-/-	-	15	-	-	386 (27704 / 27829)	-	-	-	24	ttt-/ttcgt	A3
67	AS	Ay427439.1	0	0	-/-	-	15	1	-	-	-	-	-	-	tttt/ttcgt	A1
68	SoD	AY461660.1	30	10	1/19	-	15	-	-	-	-	-	-	-	tttt/ttcgt	A1
69	ShanghaiQXC1	AY463059.1	39	39	-/-	-	15	64	-	-	-	-	56	-	cgtt/ttcgt	B1
70	ShanghaiQXC2	AY463060.1	39	39	-/-	-	15	64	-	579 (5834 / 5959 (418), 6378 (86), 6465 (75))	-	-	56	-	cgtt/ttcgt	B1
71	Sino1-11	AY485277.1	6	6	-/-	-	-	-	-	-	-	1x3 (26156 / 26260)	-	17	tttt/ttcgg	A1
72	Sino3-11	AY485278.1	3	3	-/-	-	-	-	-	-	-	1x2 (27883 / 27994 (1), 28024 (1))	-	15	tttt/ttcgg	A1
73,74	TW2 (TW1)	Ay502925.1	1	1	-/-	-	-	-	-	-	-	-	-	2	tttt/ttcgt	A1
75	TW3	AY502926.1	2	2	-/-	-	-	-	-	-	-	-	-	2	tttt/ttcgt	A1
76	TW4	AY502927.1	2	2	-/-	-	-	-	-	-	-	-	-	2	tttt/ttcgt	A1
77	TW5	AY502928.1	1	1	-/-	-	-	-	-	-	-	-	-	2	tttt/ttcgg	A1
78	TW6	AY502929.1	3	3	-/-	-	-	-	-	-	-	-	-	2	tttt/cttgg	A1
79	TW7	AY502930.1	4	4	-/-	-	-	-	-	-	-	-	-	2	tttt/cttgg	A1
80	TW8	AY502931.1	3	3	-/-	-	-	-	-	-	-	-	-	2	tttt/cttgg	A1
81	TW9	AY502932.1	5	4	-/-	1	-	-	-	-	-	-	-	2	tttt/cttgg	A1
82	TW10	AY502923.1	6	5	-/-	1	-	-	-	-	-	-	-	2	tttt/cttgg	A1
83	TW11	AY502924.1	9	8	-/-	1	-	-	-	-	-	1x2 (27068 / 21172)	-	2	tttt/cttgg	A1
84	Sin842	AY559081.1	4	4	-/-	-	13	-	-	-	1x1 (13953 / 14034)	-	-	1	tttt/ttcgt	A1
85	Sin852	AY559082.1	19	9	10/-	-	1	-	-	57 (27797 /	-	-	-	1	ttt-/ttcgt	A3

	IDENTIFICATION		SNVs				INDELs								CLASSIFICATION	
Label	ID	Accession No.	Total	Genes	5'/3' UTR	IGR	5' Del -first 15	5' Del -follow.	longIns	longDel	shortIns	shortDel	3'del	3'poly-a	Type	Group
										27909)						
86	Sin3765V	AY559084.1	9	9	-/-	-	15	1	-	-	-	-	-	11	tttt/ttcgt	A1
87	Sin848	AY559085.1	11	11	-/-	-	15	1	-	-	-	-	-	2	tttt/ttcgt	A1
88	Sin849	AY559086.1	4	4	-/-	-	15	1	-	49 (27745 / 27872)	-	-	1	-	tttt/ttcgt	A3
89	Sin846	AY559094.1	7	7	-/-	-	15	1	-	137 (27663 / 27787 (66), 27854 (71)	2x1 (9567, 9667 / 9634, 9734)	-	-	1	tttt/ttcgt	A3
90	Sin3725V	AY559087.1	4	4	-/-	-	15	1	-	-	-	-	-	5	tttt/ttcgt	A1
91	SinP1	AY559088.1	4	4	-/-	-	15	1	-	-	2x1 (2488, 25277 / 2543, 25394)	-	-	1	tttt/ttcgt	A1
92	SinP3	AY559090.1	9	4	1/4	-	15	1	-	-	2x2, (10, 25286 / 60, 25394) 9x1 (10651,10697,11112,11567, 11782,11933,11988,22506, 29716 / 10718,10764,11179,11634, 11852,12003,12058,22607, 29860)	-	-	1	tttt/ttcgt	A4
93	SinP5	AY559092.1	4	4	-/-	-	15	1	-	-	1x2 (25276 / 25394)	1x1(28088 / 28242)	-	1	tttt/ttcgt	A1
94	SinP4	AY559091.1	7	4	-/3	-	15	1	-	-	1x2 (25274 / 25394)	2x1 (24085, 24106 / 24201, 24223)	1	-	tttt/ttcgt	A1
95	Sin845	AY559093.1	10	10	-/-	-	15	1	-	-	-	-	-	1	tttt/ttcgt	A1
96	Sin847	AY559095.1	12	10	2/-	-	10	-	-	-	-	-	-	2	tttt/ttcgt	A1
97	Sin850	AY559096.1	11	6	5/-	-	8	-	-	-	-	-	-	1	tttt/ttcgt	A1
98	LLJ-2004	AY595412.1	11	10	-/-	1	15	6	-	-	1x6 (27637 / 27763)	1x1 (2919 / 2982)	-	5	cgcc/ttcat	B1
99	TJF	AY654624.1	17	10	2/1	4	1	-	-	-	-	-	-	19	tgcc/ttcgt	B1
100	CDC#200301157	AY714217.1	2	2	-	-	-	-	-	-	-	-	-	-	tttt/ttcgt	A1
101	Sin3408L	AY559097.1	4	4	-	-	15	1	-	-	1x1 (5120 / 5181)	2x1 (24524, 25029 / 24639, 25145)	-	5	tttt/ttcgt	A1
102	SinP2	AY559089.1	4	4	-	-	15	1	-	-	13x1 (10649,11638,13060,17980, 19282,19497,19928,20029, 23499,23543,23589,23607, 25127 / 10718,11710,13138,18071, 19374,19589,200	5x1 (16584, 25141, 25144, 25169, 25171 / 16674, 25246, 25250, 25276, 25279), 1x6 (25189 / 25298)	-	2	tttt/ttcgt	A4

IDENTIFICATION			SNVs				INDELs								CLASSIFICATION	
Label	ID	Accession No.	Total	Genes	5'/3' UTR	IGR	5' Del -first 15	5' Del -follow.	longIns	longDel	shortIns	shortDel	3'del	3'poly-a	Type	Group
											20,20121, 23604,23648,23694,23712, 25232), 1x2 (25278 / 25394)					
103	Sin3408	AY559083.1	14	4	10/-	-	-	-	5'end (34)	-	-	-	-	6	tttn/ttcgt	A1

Table S3. RSCU for all the annotated proteins and in total, for all the annotated isolates.
The highest RSCU synonumous codons are in bold

a.a.	codon	TOTAL	Spike	Envel.	Membr.	Nucleo.	1AB	Other
Ala	gca	1.11	0.86	0.00	0.63	1.06	1.16	1.32
	gcc	0.59	0.57	1.00	0.84	0.94	0.55	0.54
	gcg	0.24	0.24	**2.00**	0.42	0.35	0.2	0.40
	gct	**2.07**	**2.33**	1.00	**2.10**	**1.65**	2.09	**1.74**
Arg	aga	**2.09**	**1.85**	0.00	0.80	**1.94**	2.22	2.20
	agg	0.97	**1.85**	0.00	0.80	0.39	0.85	1.68
	cga	0.47	0.61	**3.00**	0.80	1.16	0.26	1.13
	cgc	0.76	0.46	0.00	1.20	1.55	0.76	0.15
	cgg	0.12	0.15	0.00	0.80	0.00	0.07	0.24
	cgt	1.59	1.07	**3.00**	1.60	0.97	1.85	0.60
Asn	aac	0.75	0.69	0.80	**1.23**	0.72	0.74	0.93
	aat	**1.25**	**1.31**	**1.20**	0.77	**1.28**	1.26	**1.07**
Asp	gac	0.75	0.63	0.00	**1.67**	**1.09**	0.72	**1.21**
	gat	**1.25**	**1.37**	**2.00**	0.33	0.91	1.28	0.79
Cys	tgc	0.75	0.93	**1.33**	0.32	0.00	0.69	**1.06**
	tgt	**1.25**	**1.07**	0.67	**1.68**	**2.00**	1.31	0.94
Gln	caa	**1.21**	**1.60**	0.00	**1.20**	**1.41**	1.31	**1.13**
	cag	0.79	0.40	0.00	0.80	0.59	1	0.87
Glu	gaa	**1.05**	1.00	**2.00**	**1.15**	1.00	0.69	0.79
	gag	0.95	1.00	0.00	0.85	1.00	1.1	**1.21**
Gly	gga	0.91	1.01	2.00	0.80	1.42	0.9	0.71
	ggc	1.01	1.26	0.00	0.80	1.42	1.07	1.38
	ggg	0.18	0.20	0.00	0.80	0.27	0.14	0.18
	ggt	**1.9**	**1.52**	2.00	**1.60**	0.89	2.11	**1.73**
His	cac	0.71	0.27	0.00	**2.00**	0.80	0.68	**1.05**
	cat	**1.29**	**1.73**	0.00	0.00	**1.20**	1.32	0.95
Ile	ata	0.66	0.54	1.00	0.50	0.28	0.65	1.05
	atc	0.63	0.35	1.00	0.33	0.81	0.72	0.51
	att	**1.7**	**2.11**	1.00	**2.17**	**1.91**	1.62	**1.44**
Leu	cta	0.69	0.61	0.86	1.16	1.15	0.6	1.15
	ctc	0.82	1.15	0.00	1.15	0.69	0.8	0.69
	ctg	0.59	0.24	0.86	0.58	0.92	0.62	0.60
	ctt	**1.76**	**2.06**	**2.57**	**1.55**	**1.62**	1.76	**1.44**
	tta	1.06	1.15	0.86	0.58	0.23	1.11	0.98
	ttg	1.08	0.79	0.86	0.97	1.38	1.11	1.13
Lys	aaa	**1.06**	**1.13**	**2.00**	**1.33**	**1.38**	1.03	**1.03**
	aag	0.94	0.87	0.00	0.67	0.62	0.97	0. 97
Met	atg	1	1.00	1.00	1.00	1.00	1	1.00
Phe	ttc	0.76	0.58	**1.49**	**1.04**	**1.39**	0.78	0.68
	ttt	**1.24**	**1.42**	0.51	0.96	0.61	1.22	**1.32**
Pro	cca	1.7	1.33	2.00	**2.32**	1.29	1.79	**1.91**
	ccc	0.41	0.22	0.00	0.06	1.16	0.37	0.51
	ccg	0.17	0.14	0.00	0.83	0.26	0.12	0.49
	cct	**1.72**	**2.31**	2.00	0.79	1.29	1.72	1.08
Ser	agc	0.53	0.38	0.86	1.00	0.86	0.47	1.06
	agt	1.13	0.69	0.86	0.50	1.20	1.28	0.69
	tca	1.76	1.75	0.86	**2.00**	1.54	1.75	**2.24**
	tcc	0.41	0.44	0.00	1.00	0.51	0.4	0.25
	tcg	0.24	0.19	**1.71**	0.50	0.17	0.18	0.76
	tct	**1.92**	**2.56**	**1.71**	1.00	**1.71**	1.92	1.01
Thr	aca	1.56	1.49	1.60	**1.54**	1.34	1.63	1.30
	acc	0.55	0.48	0.00	0.92	0.61	0.54	0.65
	acg	0.19	0.16	1.60	0.31	0.00	0.16	0.44
	act	**1.69**	**1.86**	0.80	1.24	**2.06**	1.67	**1.60**
Trp	tgg	1	1.00	0.00	1.00	1.00	1	1.00
Tyr	tac	0.88	0.67	**1.99**	**1.33**	**1.64**	0.86	0.88
	tat	**1.12**	**1.33**	0.01	0.67	0.36	1.14	**1.12**
Val	gta	0.85	0.53	0.86	1.25	0.36	0.9	0.83
	gtc	0.69	0.83	0.79	0.50	1.46	0.67	0.40
	gtg	0.79	0.57	0.57	**1.50**	0.73	0.79	1.01
	gtt	**1.68**	**2.07**	**1.78**	0.76	1.46	1.64	**1.75**
End	taa	**1.98**	**3.00**	**3.00**	**3.00**	**3.00**	0.93	**1.85**
	tag	0.42	0.00	0.00	0.00	0.00	0.85	0.22
	tga	0.6	0.00	0.00	0.00	0.00	0.9	0.94

Table S4. Positions of multiple SNVs in annotated isolates.
Positions of two or more SNVs are presented along with nucleotides and ORFs (based on HSR 1 annotation),
type of mutation (transition / transversion), a.a. position in ORF, a.a. change, a.a. properties change,
nucleotide position in codon and the number of SNVs. SNVs are in red.

Table row labels (leftmost column, isolate profiles):

- Nucleotide position on HSR 1 scale
- Nucleotide position on relative scale
- PROFILE
- gi|30275666|gb|AY278488.2|BJ01
- gi|49176846|gb|AY595412.1|LLJ-2004
- gi|41323719|gb|AY390556.1|GZ02
- gi|31416290|gb|AY278489.2|GD01
- gi|30027610|gb|AY278554.2|CUHK-W1
- gi|40795744|gb|AY508724.1|NS-1
- gi|10365700|gb|AY654624.1|TJF
- gi|30271926|ref|NC_004718.3|Tor2
- gi|38505491|gb|AY485278.1|Sino3-11
- gi|38505482|gb|AY485277.1|Sino1-11
- gi|33115118|gb|AY323977.2|HSR 1
- gi|38231932|gb|AY357075.1|PUMC02
- gi|38231937|gb|AY357076.1|PUMC03
- gi|38304867|gb|AY282752.2|CUHK-Su10
- gi|38231927|gb|AY350750.1|PUMC01
- gi|33114214|gb|AY345988.1|CUHK-AG03
- gi|33114190|gb|AY345986.1|CUHK-AG01
- gi|31581502|gb|AY291315.1|Frankfurt 1
- gi|33578015|gb|AY310120.1|FRA
- gi|40457433|gb|AY463059.1|ShanghaiQXC1
- gi|40457448|gb|AY463060.1|ShanghaiQXC2
- gi|37576845|gb|AY427439.1|AS
- gi|40548981|gb|AY502932.1|TW9
- gi|40548885|gb|AY502924.1|TW11
- gi|40548873|gb|AY502923.1|TW10
- gi|33411459|gb|AP006561.1|TWY
- gi|33411429|gb|AP006559.1|TWK
- gi|33411414|gb|AP006558.1|TWJ
- gi|33411444|gb|AP006560.1|TWS
- gi|33188324|gb|AY348314.1|Taiwan TC3
- gi|32493130|gb|AY338175.1|Taiwan TC2
- gi|32493129|gb|AY338174.1|Taiwan TC1
- gi|40548957|gb|AY502930.1|TW7
- gi|33411399|gb|AP006557.1|TWH
- gi|40548945|gb|AY502931.1|TW8
- gi|40548965|gb|AY502929.1|TW6
- gi|40548933|gb|AY502928.1|TW5
- gi|30027617|gb|AY278741.1|Urbani
- gi|40548921|gb|AY502927.1|TW4
- gi|40548909|gb|AY502926.1|TW3
- gi|40548897|gb|AY502925.1|TW2
- gi|52546959|AY714217.1|CDC#200301157
- a.a. changes
- a.a. position
- a.a. properties changes
- transition(Ts)/transversion(Tv)
- position in codon
- Number of SNVs

Table S5. Categories of nucleotide substitutions in the whole genome (annotated isolates)
Only sites where two or more isolates have an SNV are considered

			1.pos	2.pos	3.pos	TotalNo		1.pos %	2.pos %	3.pos %	Total %	silent
Transitions	A-G	A→G	5/10	6/12	6/20	17/42	30/71	3.47%	4.17%	6.94%	14.58%	6/20
		G→A	4/14	6/9	3/6	13/29		4.86%	3.12%	2.08%	10.07%	3/6
	C-T	C→T	6/19	8/24	7/35	21/78	45/156	6.60%	8.33%	12.15%	27.08%	9/46
		T→C	2/9	14/39	8/30	24/78		3.12%	13.54%	10.42%	27.08%	8/30
	Total		17/52	34/84	24/91		75/227	18.06%	29.17%	31.60%	78.82%	26/102
Transversions	A-C	A→C	0	1 / 2	1 / 2	2/4	2/4	0	0.69%	0.69%	1.39%	1/2
		C→A	0	0	0	0		0	0	0	0	0
	A-T	A→T	3/6	0	0	3/6	7/14	2.08%	0	0	2.08%	0
		T→A	0	0	4/8	4/8		0	0	2.78%	2.78%	4/8
	G-C	G→C	0	0	0	0	0	0	0	0	0	0
		C→G	0	0	0	0		0	0	0	0	0
	G-T	G→T	2/4	0	1/2	3/6	7/43	1.39%	0	0.69%	2.08%	0
		T→G	3/14	1/23	0	4/37		4.86%	7.99%	0	12.85%	0
	Total		8/24	2/25	6/12		16/61	8.33%	8.68%	4.17%	21.18%	5/10
Total			25/76	36/109	30/103		91/288	26.39%	37.85%	35.77%	100%	31/112

Table S6. Mutation analysis of S-protein.
Nonsynonumous mutation sites are in gray.

The table is a large matrix (≈90 GenBank accession rows × 62 nucleotide-position columns). The column/position annotation rows are transcribed below; the strain-by-position nucleotide matrix is a dense single-letter grid that is reproduced in summary form by the PROFILE row.

Rel. nt position	PROFILE	a.a. position	a.a. change	position in codon	number of SNVs
81	t	27	Ala (silent)	3	3
90	c	30	Tyr (silent)	3	1
146	c	49	Ser→Leu	2	1
147	a	49	Ser (silent)	3	1
183	a	61	Pro (silent)	3	1
224	c	75	Thr→Arg	2	1
230	g	77	Gly→Asp	2	2
411	c	137	Phe (silent)	3	21
430	a	144	Met→Leu	1	1
654	t	218	Pro (silent)	3	1
681	c	227	Asn→Lys	2	2
716	c	239	Ser→Leu	2	2
729	c	243	Asp (silent)	3	6
731	t	244	Ile→Thr	2	1
782	c	261	Thr→Lys	1	25
931	g	311	Gly→Arg	3	2
1026	a	342	Arg (silent)	3	5
1031	a	344	Lys→Arg	2	8
1058	c	353	Ser→Phe	1	6
1079	t	360	Phe→Ser	2	11
1098	c	366	Cys (silent)	3	2
1170	g	390	Lys (silent)	3	1
1306	t	436	Tyr→His	1	1
1325	a	442	Tyr→Ser	2	12
1437	t	479	Asn→Lys	3	3
1460	c	487	Thr→Ser	1	2
1502	t	501	Phe→Tyr	1	1
1672	t	558	Phe→Ile	1	2
1683	c	561	Ser (silent)	3	1
1729	t	577	Ser→Ala	1	11
1788	t	596	Val (silent)	3	2
1813	g	605	Asp→Asn	1	3
1819	t	607	Ser→Pro	1	11
1822	a	608	Thr→Ala	2	2
1954	a	652	Ile→Val	1	2
1994	t	665	Val→Leu	1	1
2003	g	668	Ser→Arg	1	1
2102	c	701	Thr→Ala	2	1
2227	a	743	Ser→Leu	2	6
2244	a	748	Thr→Ala	2	1
2261	c	754	Ala (silent)	3	1
2301	c	767	Ala→Val	2	2
2332	t	778	Val (silent)	1	7
2380	c	794	Tyr→Asp	1	1
2578	g	860	Pro→Ser	2	1
2581	a	861	Val→Leu	2	1
2680	a	894	Ser→Arg	2	1
2892	g	964	Thr→Ala	1	3
3002	g	1001	Ser (silent)	3	1
3075	t	1025	Arg→Met	2	1
3204	a	1068	Cys (silent)	3	6
3321	t	1107	Lys (silent)	3	1
3381	t	1127	Leu (silent)	1	1
3395	a	1132	Glu→Gly	2	2
3442	c	1148	Leu→Phe	1	4
3487	a	1163	Lys→Glu	1	7
3505	a	1169	Asn→Asp	1	1
3546	t	1182	Leu (silent)	3	1
3547	c	1183	Gln→Lys	1	1
3623	c	1208	Ala→Val	2	3
3667	t	1223	Cys→Ser	1	1
3739	c	1247	Leu→Phe	1	1

Additional annotation rows in the table: nucleotide change (e.g. t→g, c→t, c→t, a→c, a→c, c→g, g→a, …), transition(Ts)/transversion(Tv), and a.a. properties (e.g. "silent", "P+S+T→Hp+Ap", "Hp+P→S→P+PCh", "Hp+S+T→P+NCh+S", etc.).