

## SUPPLEMENTARY TABLES

**Table 1:** The GCD decoded SARS-CoV-2 key antigens.

SARS-CoV-2	SARS-CoV		1918 H1N1		2009 H1N1		<i>S. aureus</i>		<i>S. pneumoniae</i>	
	S	E	H	N	H	N	SPA	ClfA	SpsA	PspA
S	15	3	1	1	15	3	57	7809	3	1
E	1	256	64	4	256	4	1	4	16	16
M	1	1	1	1	1	1	19	19	1	1
N	1	1	1	17	1	1	1	1	1	1
ORF1ab	1	1	1	1	1	1	1	1	1	1
ORF1a	1	8	8	4	8	4	1	4	8	8
ORF3a	3	3	1	1	3	3	3	3	3	1
ORF6	5	2	2	14	10	2	1	2	2	14
ORF7a	1	1	1	1	1	1	1	1	1	1
ORF7b	3	3	1	7	3	7	3	3	3	7
ORF8	1	104	8	1156	8	4	1	4	104	104
ORF10	5	1	1	1	5	1	1	1	1	1

**Note:** The GCD value greater than radix 20 as the significant criterion.

**Table 2:** The GCD decoded the SARS-CoV-2 AQPs channel.

GCD	SARS-CoV-2	
	S protein	ORF8 protein
AQP0	57	1
AQP1	1	334568
AQP2	1	1
AQP3	5	1
AQP4	19	52
AQP5	1	1
AQP6	3	4
AQP7	3	52
AQP8	1	1
AQP9	5	1
AQP10	1	1
AQP11	1	1
AQP12	1	2

**Table 3:** The GCD decoded pathogens SIg JC/SC-hijacked invasion.

GCD	SARS-CoV-2		SARS-CoV		1918 H1N1		2009 H1N1		<i>S. aureus</i>		<i>S. pneumoniae</i>	
	S	ORF8	S	ORF8ab	H	N	H	N	SPA	ClfA	SpsA	PspA
SIg-JC	285	2	15	1	2	2	2	18	513	114	6	2
SIg-SC	1	1768	1	1	16	476	16	4	1	8	208	1456
B cell CD4	3	884	3	1	4	476	4	36	9	12	156	2548
B cell CD22	285	1	15	1	1	1	1	3	57	57	3	1

**Table 4:** GCD decoded the relationship of antigens and enzymes

GCD	1918 H1N1		2009 H1N1		SARS-CoV-2		SARS-CoV		
	H	N	H	N	S	ORF8	S	ORF8a	ORF8b
ADAR 1	32	476	32	4	1	136	1	1	1
ADAR 2	1	289	1	3	3	3757	3	1	13
ADAR 3	1	1	23	9	3	1	3	1	1

GlcNAcase	1	119	1	1	1	221	1	1	13
HexNAcase	2	578	2	2	1	578	1	1	1

**Table 5:** The GCD decoded the relationship between NSPs and enzyme complexes.

GCD		Mutation enzyme			Modification enzyme			
		ADAR 1	ADAR 2	ADAR 3	GlcNAcase a	GlcNAcase b	HexNAcase $\alpha$	HexNAcase $\beta$
nsp1	SARS-CoV-2	16	39	3	13	1	1975	2
	SARS-CoV	16	1	1	1	1	25	2
nsp2	SARS-CoV-2	32	13	1	13	1	25	2
	SARS-CoV	32	1	1	1	1	25	2
nsp3	SARS-CoV-2	32	1	1	1	1	25	2
	SARS-CoV	32	1	1	1	1	25	2
nsp4	SARS-CoV-2	1	1	1	1	1	1	1
	SARS-CoV	1	1	1	1	1	1	1
nsp5	SARS-CoV-2	1	3	3	1	1	1	1
	SARS-CoV	1	1	1	1	1	1	1
nsp6	SARS-CoV-2	1	1	23	1	1	1	1
	SARS-CoV	1	1	1	1	1	1	1
nsp7	SARS-CoV-2	1	39	9	13	1	1	1
	SARS-CoV	1	1	1	1	1	1	1
nsp8	SARS-CoV-2	7	1	1	7	1	1	1
	SARS-CoV	1	1	1	1	1	1	1
nsp9	SARS-CoV-2	1	1	1	1	1	1	1
	SARS-CoV	1	3	9	1	1	1	1
nsp10	SARS-CoV-2	7	13	23	91	1	1	1
	SARS-CoV	49	1	1	49	1	1	1
nsp11	SARS-CoV-2	6664	17	1	833	1	1	34
	SARS-CoV	8	1	1	1	1	1	2
nsp12	SARS-CoV-2	1	3	3	1	1	1	1
	SARS-CoV	1	1	1	1	1	1	1
nsp13	SARS-CoV-2	17	663	3	221	1	1	17
	SARS-CoV	1	3	3	1	1	1	1
nsp14	SARS-CoV-2	1	3	3	1	1	1	1
	SARS-CoV	1	1	1	1	1	1	1
nsp15	SARS-CoV-2	1	1	1	1	1	1	1
	SARS-CoV	17	51	3	17	1	1	17
nsp16	SARS-CoV-2	1	1	1	1	1	1	1
	SARS-CoV	1	1	1	1	1	1	1

**Table 6:** The GCD decoded the relationship of SARS-CoV-2 NSPs and structural-accessory proteins.

GCD	S	E	M	N	ORF1ab	ORF1a	ORF3a	ORF6	ORF7a	ORF7b	ORF8	ORF10
nsp1	15	16	1	1	1	8	3	10	1	3	104	5
nsp2	5	128	1	1	1	8	1	10	1	1	104	5
nsp3	5	128	1	1	1	8	1	10	1	1	3	5
nsp4	1	1	1	1	1	1	1	1	1	1	1	1
nsp5	3	1	1	1	1	1	3	1	1	3	1	1
nsp6	1	1	1	1	1	1	1	1	1	1	1	1
nsp7	3	1	1	1	1	1	3	1	1	3	13	1
nsp8	1	1	1	1	1	1	1	7	1	7	1	1
nsp9	1	1	1	1	1	1	1	1	1	1	1	1
nsp10	1	1	1	1	1	1	1	7	1	7	13	1
nsp11	1	8	1	17	1	8	1	14	1	7	136	1

nsp12	3	1	1	1	1	1	3	1	1	3	1	1
nsp13	3	1	1	17	1	1	3	1	1	3	221	1
nsp14	3	1	1	1	1	1	3	1	1	3	1	1
nsp15	1	1	1	1	1	1	1	1	1	1	1	1
nsp16	1	1	1	1	1	1	1	1	1	1	1	1

**Table 7:** Immune@enzyme decode SARS-CoV-2 VOCs.

GCD	SARS-CoV-2		Alpha		Beta		Gamma		Delta	
	(NC_045512.2)		(MZ314997.2)		(MZ314998.1)		(MZ427312.1)		(MZ724507.1)	
	S	ORF8	S	ORF8	S	ORF8	S	ORF8	S	ORF8
SIg-JC	285	2	15	5	30	2	10	2	5	18
SIg-SC	1	1768	1	1	16	1768	4	8	7	8
CD4	3	884	3	1	12	884	4	4	7	36
CD22	285	1	15	5	15	1	5	1	5	3
ADAR1	1	136	1	1	32	136	4	8	7	8
ADAR2	3	3757	3	1	3	3757	1	1	1	3
ADAR3	3	1	3	1	3	1	23	1	1	9
GlcNAcase	1	221	1	1	1	221	1	1	7	1
HexNAcase	1	578	1	1	2	578	2	2	1	2

**Table 8:** Immune@enzyme decode SARS-CoV-2 Omicron.

GCD	SARS-CoV-2		Omicron		Omicron		Omicron		Omicron	
	(NC_045512.2)		(OQ151515)		(OQ151513.1)		(OQ151397)		(OQ151387)	
	S	ORF8	S	ORF8	S	ORF8	S	ORF8	S	ORF8
SIg-JC	285	2	95	2	15	2	135	2	15	2
SIg-SC	1	1768	1	1768	7	1768	1	1768	7	1768
CD4	3	884	1	884	147	884	9	884	21	884
CD22	285	1	95	1	15	1	15	1	15	1
ADAR1	1	136	1	136	49	136	1	136	7	136
ADAR2	3	3757	1	3757	3	3757	3	3757	3	3757
ADAR3	3	1	1	1	3	1	9	1	3	1
GlcNAcase	1	221	1	221	49	221	1	221	1	221
HexNAcase	1	578	1	578	1	578	1	578	1	578

## SUPPLEMENTARY INFORMATION

protein@integer data

SARS-CoV-2 (NC\_045512.2)

nsp1

>YP\_009725297.1 leader protein [Severe acute respiratory syndrome coronavirus 2]

MESLVPGFNEKTHVQLSLPVLQVRDVLVRGFGDSVEEVLSEARQHLLKDGTCGLVEVEKGVLPQLEQPYVF  
IKRSDARTAPHGHVMVELVAELEGIQYGRSGETLGLVPHVGEIPVAYRKVLLRKNNGKAGGHSYGADL  
KSFDLGDELGTDPYEDFQENWNTKHSSGVTTRELMRELNGG

YP\_009725297.1 nsp1 protein@integer

1123005142691333116512868833540164916836257849622557724441601929087350957440799258910507476880923279  
523265159495165799711040992947449017591588899269375646987216948156566375503273670815537540319137960  
667154498974982311281519446712539600

nsp2

>YP\_009725298.1 nsp2 [Severe acute respiratory syndrome coronavirus 2]

AYTRYVDNFCGPDGYPLECIKDLLARAGKASCTLSEQLDFIDTKRQVYCCREHEHEIAWYTERSEKSYE  
LQTPFEIKLAKKFDTFNGECPNFVFLPLNSIIKTIQPRVEKKKLDGFMGRIRSVYPVAVSPNECNQMCLSTL  
MKCDHCGETSWQTGDFVKATCEFCGTENLTKEGATTCGYLPQNAVVKIYCPACHNSEVGPESHSLAEYHNE  
SGLKTILRKGGRITAFGGCVFSYVVGCHNKCAWVPRASANIGCNHTGVVGESEGLNDNLLEILQKEKVN  
INIVGDFKLNIEIAIILASFSASTSAFVETVKGLDYKAFKQIVESCNGFKVTKGKAKKGAWNIGEQKSIL  
SPLYAFASEAARVRSIFSRTLETAQNSVRVLQKAAITILDGISQYSLRLIDAMMFTSDLATNNLVMAY  
ITGGVVQLTSQWLTNIFGTVYEKLPVLDWLEEFKEGVEFLRDGWEIVKFISTCACEIVGGQIVTCAKE  
IKESVQTFKLVNKFALCADSIIIGGAKLKALNLGETFVTHSKGLYRKC VKSREETGLLMPLKAPKEII  
FLEGETLPTEVLTEEVVLKTGDLQPLEQPTSEAVEAPLVGTPVCINGLMLEIKDTEKYCALAPNMMVTN  
NTFTLKGG

YP\_009725298.1 nsp2 protein@integer

109198532398067420580044526301830372437666832956283289638108601624213341948515090272246530028752937  
3103190154029314020755430377879838821652426394191353647911763114545196069800891225755735496984578912  
108095143432268952930543058833800890545104656629326100776101563401749990799063432295423703171268911  
598145536004084738430142150835852941261990159667195300516600444398361261169769068375082132873230764  
72159395313533263535529263927665052196361637407651990282784868476189126784100497779994964528853812  
376903325191132360028347607832733574087137796522533707512053087430050170773215466234549908016602981  
2381116365748577527178349240659949401271440180138999914329687760970123974422049926609517110089700876  
0774949883414724000522659013521836704100984536501883590055597808347182811155415419127268233628565411  
74586647292429637200734962292060800

nsp3

>YP\_009725299.1 nsp3 [Severe acute respiratory syndrome coronavirus 2]

APTQVTFGDDTVIEVQGYKSVNITFELDERIDKVLNEKCSAYTVELGTEVNEFACVVADAVIKTLQPVSE  
LLTPLGIDLDEWSMATYYLFDESGEFKLASHMYCSFYPPDEDEEEGDCEEEEFEPSTQYIEYGTEDDYQGG  
PLEFGATSAAALQPEEEQEEDWLDDDSQQTGQDQSGEDNQTITTIQTIVEVQPQLEMELTPVVQTIEVNSF  
SGYLKLTDNVYIKNADIVEEAKKVKPTVVVNAANVYLKHGGGVAGALNKATNNAMQVESDDYIATNGPLK  
VGGSCVLSGHNLAHCLHVVGPNVNGEDIQLLKSAYENFNQHEVLLAPLLSAGIFGADPIHSLRVCVDT  
VRTNVYLAVFDKNLYDKLVSSFLEMKSEKQVEQKIAEIPKEEVKPFITESKPSVEQRKQDDKKIKACVEE  
VTTTLEETKFLTENLLLIDINGNLHPDSATLVSDIDITFLKKDAPYIVGDVQEGVLTAVVIPTKKAGG  
TTEMLAKALRKVPTDNYITTPGQGLNGYTVVEAKTVLKKCKSAFYILPSIISNEKQEILGTVSWNLREM  
LAHAETRKLMPVCVETKAIVSTIQRYKGIKIQEGVVDYGARFYFYSKTTVASLINTLNDLNETLVTM  
PLGYVTHGLNLEEAARYMRSLKVPATVSVSPDAVTAINGYLTSSSKTPEEHFIETISLAGSYKDWSYSG  
QSTQLGIEFLKRGDKSVYYSNPTTFHLDGEVITFDNLKTLSSLREVRTIKVFTTVDNINLHTQVVDMSM  
TYGQQFGPTYLDGADVTKIKPHNSHEGKTFYVLPNDDTLRVEAFEYHHTDPSFLGRYMSALNHTKKWKY  
PQVNGLTSIKWADNNCYLATALTLQQIELKFNPPALQDAYRRARAGEAANFCALILAYCNKTVGELGDV  
RETMSYLFQHANLDSCKRVLNVCKTCGQQQTTLKGVEAVMYMGTLSEYQFKKGVQIPCTCGKQATKYL  
QQESPFVMSAPPAQYELKHGTFTCASEYTGNYQCGHYKHITSKETLYCIDGALLTKSSEYKGPITDVFY  
KENSYTTTTIKPVYKLDGVVCTEIDPKLDNYYKDNSYFTEQPIDLVPNQYPNASFDNFKFVCDNIKFA  
DDLNLQLTGYKPPASRELKVTFFPDNLNGDVVAIDYKHYTPSFKKGAKLLHKPIVWHVNNATNKATYKPNW  
CIRCLWSTKPVETSNSFDVLKSEDAQGMDNLACEDLKPVSEEVENPTIQKDVLECNVKTTEVVGDIIK  
PANNLSLKITEEVGHTDLMAAYVDNSSLTIKPNELSRVLGLKTLATHGLAAVNSVPWDTIANYAKPFLNK

VVSTTTNIVTRCLNRVCTNYMPYFFTLQLCTFTRSTNSRIKASMPPTIAKNTVKSVMKFCLEASFNYL  
KSPNFSKLINIIWFLLLSVCLGLIYSTAALGVLMNSLGMPSYCTGYREGYLNSTNVTIATYCTGSIPC  
SVCLSGLDSDTYPSTLETIQITISSFKWDLTAFGLVAEWFLAYILFTRFFYVLGLAAIMQLFFSYFAVHF  
ISNSWLMWLIINLVQMAPISAMVRMYIFFASFYVWKSYPVHVDGCNSSTCMCYKRNRATRVECTTIVN  
GVRRSFYVYANGGKGFCKLHNWNCVNCDFCAGSTFISDEVARDLSLQFKRPINPTDQSSYIVDSVTVKN  
GSIHLYFDKAGQKTYERHSLSHFVNLDNLRANNTKGSPLINVIVFDGKSKCEESSAKSASVYYSQLMCQP  
ILLLDQALVSDVGDSEAVVVMFDAYVNTFSSTFNVPMEKLLTLVATAEAEELAKNVSLDNVSTFISAAR  
QGFVDSVETKDVVECLKLSHQSDIEVTGDSCNNYMLTYNKVENMTPRDLGACIDCSARHINAQVAKSHN  
IALIWNVKDFMSEQLRKQIRSAAKNNLFPKLTCAATTRQVNVVTTKIALKGG

YP\_009725299.1 nsp3 protein@integer

185470377115498318092166266114613830307353365322888489977706485254475466928092548434396174326007516  
078003704959471225877391665064121097440237271214626116492636884488630028474594721528327788348342272  
989865073208870842427377587497216409298291601740023565457457979561359415234651089399376648037281899  
849445314268841564112569331734532301266600259325207473805261906502751864763315007706857015008823992  
410088230754485843777683242798315004128855618553595929778673927089963226630367413892996313679405849  
290192340181192733588229101747320327832603358050836485560192941087054399918527868963223018942387919  
4333931548963643819811609727994174923186551722577839793675461223757237441806495397655528018372882121  
60180574030504587748509480344205841355970092803252828055243650877328218261643395917829130599059799  
7965676788225204705743445370429032783997188743516394499161599166927646721359467181647949932241365353  
026973199165049935747741857586780419915910843509548140522332989889887326550519419680222052224518005  
5929333414370112772345177110064487927968822829884706413287241791008949304309738070929404875183280911  
2857420926730231058568517680893763035015103526677717262534701749112616467587838319154780930339584364  
395773660413480296866709486161045442395555641514349702047833236080126927000713324705172011926997095  
8435924294118084111429278204130452003495585457478365142366526611000373794281986927316439951171069540  
709525771929914236608373448204245797526718065488759186864504435697367955650261145338632481759455567  
590342095755105709627684260839361505717335230355962707982816303288150546629111441193105544858367359  
25639180380313447219275946136318555223402432720149988190684893990640397293291852385652250754259502  
28848810018647233598109645428276503493654948848583617373132004508138656929300864688156854894702797  
330098807118481132042400476491686763285595842371537346699337597861892486052147175559225601584997162  
92823077307027159590662676697670026286577822527229940825218652200923102250430792723268516593997169  
6791215797603801888521414189928547891145218791493098904608786133368018698556526933140093927715146984  
013034873253359256545590634085467119838867038699702294585395127230604184445158412772120815365369187  
1382360436122960833242508996251542184996115062288591408574711899667720171012336794544036482874465186  
015694213281370604954383824723166468012138877442400048233245325310848819281849513314255432488468581  
869616981172342298562257178809973660800192172387929071381895268412860876237241529407891864396887040  
5538636069732719349623731813420812224350343193820800

nsp4

>YP\_009725300.1 nsp4 [Severe acute respiratory syndrome coronavirus 2]

KIVNNWLKQLIKVTLVFLFVAAIFYLITPVHVMKHTDFSSSEIIGYKAIDGGVTRDIASDTDCFANKHAD  
FDTWFSQRGGSYTNDKACPLIAAVITREVGFFVPGPLGTILRTTNGDFLHFLPRVFSAVGNICYTPSKLI  
EYTDFAVSACVLAAECTIFKDASGKVPYCYDTNVLEGSVAYESLRPDTRYVLMDSIIQFPNTYLEGSV  
RVVTTDFSEYCRHGTCESEAGVCVSTSGRWVLNNDYRSLPGVFCGVDAVNLLTNMFTPLIQPIGALDI  
SASIVAGGIVAIVVTCLAYYFMRFRRAFGEYSHVAFNTLLFLMSFTVLCLTPVYSFLPGVYSVIYLYLT  
FYLTNDVSLAHIQWVMFTPLVPFWITAIYIICISTKHFWFFSNYLKRRVVFNGVFSSTFEEAALCTF

LLNKEMYLKLRSVLLPLTQYNRYLALYNKYKYFSGAMDTTSYREAACCHLAKALNDFSNVSGSDVLYQPP  
QTSITSAVLQ

YP\_009725300.1 nsp4 protein@integer

203133319812863039785385993068228514647341216169840905292087432759351548236476027705769297273310075  
5934795481569442180641426651894740318735972969462478741746255340813449013673204595676330679198191935  
500852622942542590055236992598477319416870850091087400004018627116783170768677803082902074498537608  
6224185615649171796443800194879578919535485689895767168714292981983605006074101950047745298530916466  
6011723819800737256380549601744954988674819883860193208490418528282236472337017918824741212530778294  
482008717190859401125409574267633378995246316016591022301183356222834393568686243486624464791233641  
067436225221129092469665452429181072061066883088329751

nsp5

>YP\_009725301.1 3C-like proteinase [Severe acute respiratory syndrome coronavirus 2]

SGFRKMAFPSGKVEGCMVQVTCGTTTTLNLWLDDVVYCPRHVICTSEDMLNPNYEDLLIRKSNHNFLVQA  
GNVQLRVIGHSMQNCVLLKLVDTANPKTPKYKFVRIQPGQTFVSLACYNGSPSGVYQCAMRPNFTIKGSF  
LNGSCGSVGFNIDYDCVFCYMHMELPTGVHAGTDLEGNFYGPFVDRQTAQAAGTDTTITVNLAWLYA  
AVINGDRWFLNRFITTLNDFNLVAMKYNIEPLTQDHVDILGPLSAQTGIAVLDMCASLKELLQNGMNGRT  
ILGSALLEDEFTPFVVRQCSGVTFQ

YP\_009725301.1 nsp5 protein@integer

1331214028941156729269672708323180702278796690142791813257137683169021220975300635616332377346205699  
001769309309751565728580770023556670694774647879430482130911516299452578244479669279172210884055951  
841643499197304432295810854459603188793857085880648451704531389158527225275956626138883970252846910  
1016410543315738290510013023372331353756099812178297242607988977445814489829483356473387457670434331

nsp6

>YP\_009725302.1 nsp6 [Severe acute respiratory syndrome coronavirus 2]

SAVKRTIKGTHHWWLLLTSLVLVQSTQWLSFFLYENAFLPFAMGIIAMSAFAMMFVKHKHAFCLF  
LLPSLATVAYFNMVYMPASWVMRIMTWLDMVDTSLSGFKLKDCVMYASAVVLLILMTARTVYDDGARRVW  
TLMNVLTLYKVVYGNALDQAISMWALISVTSNYSVGVTTVMFLARGIVFMCVEYCIFFITGNTLQCI  
MLVYCFLGYFCTCYFGLFCLLNRYFRLTLGVYDYLSTQEFRYMNSQGLLPPKNSIDAFKLNKLLGVGG  
KPCIKVATVQ

YP\_009725302.1 nsp6 protein@integer

20505710727072688704959460978233570982560907324714648629326546696024843170812780948942407635437153  
5222532331999991817670166552568375569265108603723166717575380613406970557212076408604047936515303651  
636663995311029320914582152328836731296599038855092409049742712848770979454220989919087510016190782  
43549548049254984162685531849879320708401270226289765370935618781190229031050091

nsp7

>YP\_009725303.1 nsp7 [Severe acute respiratory syndrome coronavirus 2]

SKMSDVKCTSVVLLSVLQQLRVESSSKLWAQCVQLHNDILLAKDTTEAFEKMSVLLSVLLSMQGAVDINK  
LCEEMLDNRATLQ

YP\_009725303.1 nsp7 protein@integer

127435962709688865300893277401241546623400041889338055465541234671424254488291793017563004302680944  
031530151

nsp8

>YP\_009725304.1 nsp8 [Severe acute respiratory syndrome coronavirus 2]

AIASEFSSLPSYAAFATAQEAYEQAVANGDSEVVLKCLKKSLNVAKSEFDRDAAMQRKLEKQAMTQM  
YKQARSEDKRAKVTSAMQTMFLTMLRKLNDALNNIINNARDGCVPLNIPLTTAAKLMVVIPDYNTYKN  
TCDGTTFTYASALWEIQVVDADSKIVQLSEISMDNSPNLAWPLIVTALRANSAVKLQ

YP\_009725304.1 nsp8 protein@integer

2817838741459423519776105041095872621464330279357538801723915815608116343806776077292094323158947725  
375465328331833799823803412278064508209537758247817307038901685580533513207018015396501434453994038  
8480716424112277662570104563815517283397485582261062596951

nsp9

>YP\_009725305.1 nsp9 [Severe acute respiratory syndrome coronavirus 2]

NNELSPVALRQMSCAAGTTQTACTDDNALAYYNTTKGGRFVLALLSDLQDLKWARFPKSDGTGTIYTELE  
PPCRFVTDTPKGPVKYLYFIKGLNNLNRMVGLSLAATVRLQ

YP\_009725305.1 nsp9 protein@integer

492405683224698977802404933489923142696369231645975037235796429726794104111877409626206676616028  
409131776548429673257913850126331143034940228038951

nsp10

>YP\_009725306.1 nsp10 [Severe acute respiratory syndrome coronavirus 2]

AGNATEVPANSTVLSFCAFAVDAKAYKDYLASGGQPITNCVKMLCTHTGTGQAITVTPEANMDQESFGG  
ASCCLYCRCHIDHPNPKGFCDLKGKYVQIPTTCANDPVGFTLKNVCTVCGMWKGYGCSCDQLREPMLQ

YP\_009725306.1 nsp10 protein@integer

356345132828815625062945761998285926168214786226183190581879798702483302627505099736859255160679570  
685921575092919674742688778874593128114386688253114135798284027331812622584509751

nsp11

>YP\_009725312.1 nsp11 [Severe acute respiratory syndrome coronavirus 2]

SADAQSFLNGFAV

YP\_009725312.1 nsp11 protein@integer

8499997207846424

nsp12

>YP\_009725307.1 RNA-dependent RNA polymerase [Severe acute respiratory syndrome coronavirus 2]

SADAQSFLNRVCGVSAARLTPCGTGTSTDVVYRAFDIYNDKVAGFAKFLKTNCCRFQEKDEDDNLIDSYF  
VVKRHTFSNYQHEETIYNLLKDCPAVAKHDFKFRIDGDMVPHISRQRLTKYTMADLVYALRHFDEGNCD  
TLKEILVTYNCCDDDYFNKKDWYDFVENPDILRVYANLGERVRQALLKTVQFCDAMRNAGIVGVLTLDNQ  
DLNGNWYDFGDFIQTPGSGVPVVDSSYLLMPILTLTRALTAESHVDTDLTKPYIKWDLKDYDFTEERL  
KLFDRYFKYWDQTYHPNCVNCLDDRCILHCANFNVLVSTVFPPTSFGPLVRKIFVDGVPFVSTGYHFRE  
LGVVHNQDVNLHSSRSLFKELLVYAADPAMHAASGNLLDKR TTCFSVAALTNNVAFQTVKPGNFNKDFY  
DFAVSKGFFKEGSSVELKHFFFAQDGNAAISDYDYRYNLPMTCDIRQLLFVVEVVDKYFDCYDGGCINA  
NQVIVNNLDKSAGFPFNKWGKARLYYDSMSYEDQDALFAYTKRNVIPITITQMNLYAISAKNRARTVAGV  
SICSTMTNRQFHQKLLKSIAATR GATV VIGTSK FYGGWHNMLKTVYSDVENPHLMGWDPKCDRAMPNML  
RIMASLV LARKHTTCCSLSHRFYRLANECAQVLSEMMCGGSLYVKPGGTSSGDATTAYANSVFNICQAV  
TANVNALLSTDGNKIADKYVRNLQHRLYECLYRNRD VDTDFVNEFYAYLRKHFSMMILSDDAVVCFNSTY

ASQGLVASIKNFKSVLYYQNNVFMSEAKCWTETDLTKGPHEFCSQHTMLVKQGDDYVYLPYPDPSPRILGA  
GCFVDDIVKTDGTLMIERFVSLAIDAYPLTKHPNQEYADVHFLYLQYIRKLHDELTDGHMLDMYSVMLTND  
NTSRYWEPEFYEAMYPHTVLQ

YP\_009725307.1 nsp12 protein@integer

376690613337660346392811552369903863188860175516130654743240796591022120642038418796488069755371002  
52901878685550219618329244246239779690115281323680347509762664110902901551125265313036348673443253  
270703986064996661214257063819940615666843830983994461373632976114867426593091172266506316938488899  
7826784110107485753766005392211593533705767564274314616930230963913922282259237407652459430081097425  
41791452094837285184739152480141214475616464640922156358116631386200470314934470412163360189118156399  
894247913231440050065730993861587838284286960560780024813984185896752219799101220179267726514533802  
4277170626351371911919200881923159958407857193880601950597609172815533338112694900658962299733022376  
76603182874782298252358124121491741757651757375598049593310400691161203152459463756445041530768141026  
9088201016025043017693871412913849407784755133857079310810210773392346997047057272567519830310951610  
8537077481151558697881943515086167433605859317134702073664936623459907726214626819733709964944017628  
508312933668838997278879439841213366765395801411418948943655923542964443183212246178545385008459090  
8929450213610399240514621329927694811827686134209399093932130004834946903310612115150726511447021473  
516413772041751

nsp13

>YP\_009725308.1 helicase [Severe acute respiratory syndrome coronavirus 2]

AVGACVLCNSQTSRLRCGACIRRPFLCCKCCYDHSVISTSHKLVLSVNPYVCNAPGCDVTDVTQLYLGGMSY  
YCKSHKPPISFPLCANGQVFGLYKNTCVGSDNVTFDFNAIATCDWTNAGDYILANTCTERLKLFAAETLKA  
TEETFKLSYGIATVREVLSRELHLSWEVGKPRPPLNRNYVFTGYRVTKNSKVQIGEYTFEKGDYGDVV  
YRGTTTTYKLVNGDYFVLTSHVTMPLSAPTLVPEHYVRITGLYPTLNISDEFSSNVANYQKVGMMQKYSTL  
QGPPGTGKSHFAIGLALYPSARIVYTACSHAARDALCEKALKYLPIDKCSRIIPARARVECFDKFKVNS  
TLEQYVFCTVNALPETTADIVVDFEISMATNYDLSVNVNARLRAKHVYVIGDPAQLPAPRTLLTKGTLEPE  
YFNVCRLMKTIGPDMFLGTCRRCPAEIVDVTVSALVYDNKLKAHKDKSAQCFKMFYKGVITHDVSSAINR  
PQIGVVREFLTRNPAWRKAVFISPYNSQNAVASKILGLPTQTVDSSQGSEYDYVIFTQTTETAHSCNVNR  
FNVAITRAKVGILCIMSARDLYDKLQFTSLEIPRRNVATLQ

YP\_009725308.1 nsp13 protein@integer

498009864061529971915472919950973027458866337542065919670337689380601070593890740464003312453676981  
645050271261504179539927805179394090662621489978345693036106634052891414598340841196007400929393553  
2432161310131711449542310756673687979077630935000644560473664720037497281705737098789276584229076892  
1660094112411402260261650708199977305184035714847207238954675355540665552230191508721129712147249891  
2508471209875812892044916679430666217973713690037176920549170727795681563882198307359958010038061184  
647582411728570760026197087409385613320889690178390253267357248926366683646214430218769449607961138  
46892857533683249842308870619063846636398232320267136883385236425171240869955047322310916318882084  
64025544143991277296831549646127389434827285955397124065277554358741001762915677450151

nsp14

>YP\_009725309.1 3'-to-5' exonuclease [Severe acute respiratory syndrome coronavirus 2]

AENVTGLFKDCSKVITGLHPTQAPTHLSVDTKFKTEGLCVDIPGIPKDMTYRRLISMMGFKMNYQVNGYP  
NMFITREEAIRHVRAWIGFDVEGCHATREAVGTNLPLQLGFSTGVNLVAVPTGYVDTPNNTDFSRVSAKP  
PPGDQFKHLIPLMYKGLPWNVVRIVQMLSDTLKNSDRVVFVLWAHGFELTSMKYFVKIGPERTCCLC  
DRRATCFSTASDITYACWHHSIGFDYVYNPFMIDVQQWGFNGLQSNHDLQVHGNHVAASCAIMTRCL



AVHECFVKRVDWTIEYPIIGDELKINAACRKVQHMMVKAALLADKFPVLHDIGNPKAIKCVQADVEWKF  
YDAQPCSDKAYKIEELFYSYATHSDKFTDGVCLFWNCNVDRYPANSIVCRFDTRVLSNLNLPDGGSLY  
VNKHAFHTPAFDKSAFVNLKQLPFFYSDSPCESHGKQVSDIDYVPLKSATCITRCNLGGAVCRHHANE  
YRLYLDAYNMMISAGFSLWVYKQFDTYNLWNTFTRLQ

YP\_009725309.1 nsp14 protein@integer

36752069921051730587260208894382308833829369266390668447446679828425722419204253579194997381810279  
587030586842173348834977428747801617885359283875748385771237090650479509835557685568933982051826667  
663453941954135370965219653997751682001213403897657232192018262809104676840413027436068887632966796  
0484152650610762537993749814543941636650717466062682775984898410713647042033982167435978742971775623  
544892824326707925799958298912901208254591100260231818936306724782082174772166497947625648469673623  
6957783215238390395314799820583556578136921247977298721812728924787701793221561615599040414936907930  
773449310105456332976116233726309449326228556859058390852376023258692733871293132114606951

nsp15

>YP\_009725310.1 endoRNase [Severe acute respiratory syndrome coronavirus 2]

SLENVAFNVVNGHFDGQQGEVPSIINNTVYTKVDGVDVELFENKTTLPVNVAFELWAKRNIKPVPEVK  
ILNNLGVDAANTVIWDYKRDAPAHISTIGVCSMTDIAKKPTETICAPLTVFFDGRVDGQVDFLRNARNG  
VLITEGSVKGLQPSVGPQASLNGVTLIGEAVKTQFNYYKKVDGVVQQLPETYFTQSRNLQEFKPRSQME  
IDFLELAMDEFIERYKLEGYAFEHIVYGDFSHSQLGGLHLLIGLAKRFKESPFLEDFIPMDSTVKNYFI  
TDAQTGSSKCVCSVIDLLDDFVEIISQDLSVSVKVVKTIDYTEISFMLWCKDGHVETFPKLQ

YP\_009725310.1 protein@integer

1708406025813217284823887107566232696343907537647665388357721374202770034061575869788134418416677578  
522896545077391072826075245410146112857282897136675300261741593158477682129949098758042644690628354  
498484184976988003605060813178560902737243013068493108095000263067717445111838855547073267908666952  
259177596817953182410568981444426650257527894733601049894489725803881149434470727673928893923706616  
71660328194922108266308416402315279288327799414108951

nsp16

>YP\_009725311.1 2'-O-ribose methyltransferase [Severe acute respiratory syndrome coronavirus 2]

SSQAWQPGVAMPNLYKMQRMLLEKCDLQNYGDSATLPKGIMMNVAKYTQLCQYLNTLTLAVPYNMRVIHF  
GAGSDKGVAPGTAVLRQWLPTGTLLVDSLDLNDVSDADSTLIGDCATVHTANKWDLIISDMYDPKTKNVT  
KENDSKEGFFTYICGFIQQKALGGSVAIKITEHSWNADLYKLMGHFAWWTAFVTNVNASSEAFILGCN  
YLGKPREQIDGYVMHANYIFWRNTNPIQLSSYSFLDMSKFLKLRGTAVMSLKEGQINDMILLSLLSKGRL  
IIRENNRVVISSDVLVNN

YP\_009725311.1 nsp16 protein@integer

5417872081616457124811722568288080766492612897726132932005704744954412579336603251684081583047324439  
74589164561771294133042970296182571774186514937590306403092142687041081164717741572136799394865832224  
6201152831776123409357671966555077719941162011888269761166192325535497530790982560258807352821680112  
36893056763161661018357953034766767975600224415729487244379120578647612868815520697789

SARS coronavirus MA15 (FJ882957.1)

FJ882957.1 nsp1

MESLVLGVNEKTHVQLSLPVLQVRDVLVRGFGDSVEEALSEAREHLKNGTCGLVELEKGVLPQLEQPYVFIKRSDA  
LSTNHGHKVVVELVAEMDGIQYGRSGITLGLVPHVGETPIAYRNVLLRKNNGKAGGHSYGIDLKSYDLGDELGT  
DPIEDYEQNWNTKHGSGALRELTRELNGG

FJ882957.1 nsp1 protein@integer

112300523775394713585003707858137894475387439597230164180592766968790677438970089675202917708911400  
41335884487002510218591874100631006311619618833186866412046366174054184745752737016018658668262074006  
00047392388358308004125526136539600

FJ882957.1 nsp2

AVTRYVDNFCGPDGYPLDCIKDFLARAGKSMCTLSEQLDYIESKRGVYCCRDHEHEIAWFTERSDKSYEHQTPF  
EIKSAKKFDTFKGECPKFVPLNSKVVIQPRVEKKKTEGFMGRIRSVYPVAVSPQECNNMHLSTLMKCNHCDEVSW  
QTCDFLKATCEHCCTENLVIEGPTTCGYLPTNAVVKMPCPACQDPEIGPEHSVADYHNHSNIETRLRKGGRTCRF  
GGCVFAYVGCYNKRAYWVPRASADIGSGHTGITGDNVETLNEDLLEILSRERNINIVGDFHLNEEVAILASFSAS  
TSAFIDTIKSLDYKSFKTIVESCNGYKVTGKPKVKGAWNIGQQRSVLTPLCGFPSQAAGVIRSIFARTLDAANHSIPD  
LQRAAVTILDGISEQSLRLVDAMVYTSDLLNSVIIIMAYVTGGLVQQTSQWLSNLLGTTVEKLRPIFEWIEAKLSAG  
VEFLKDAWEILKFLITGVFDIVKGQIQVASDNIKDCVKCFIDVVNKALEMCIDQVTIAGAKLRSLNLGEVFIAQSKG  
LYRQCIRGKEQLQLLMPKAPKEVTFLEGSDHDTVLTSEEVVLKNGELEALETPVDSFTNGAIVGTPVCVNGMLMLL  
EIKDKEQYCALSPGLLATNNVFRKGG

FJ882957.1 nsp2 protein@integer

69277176993873212416933622090199121069250111528411277332575727698128648333410277283777005588903218519  
33863843590342050318594444596257328619459373003330666570129463005853427434134970207168949426348257  
1543701918697926617063126743603029165836861890000763155662236142419748165606246088962783923752915374  
120888108114134550220501342659042441466711514363843406362502928650293867201558602505943948192051393  
932103684426215732586729365163764482849142997876844165978617072153164817290723580504572252532974859  
448591863733673815322579973569315775812889880563047211461968920255635264144531530659796972215711238  
6702586104333707511194838755342985581895669741221328195673742471541982050720597524594130354290958661  
5928459175265270837571295962692081902202136333034478199224048754811642415657626181750170671856802806  
586940788852069201318642229980800

FJ882957.1 nsp3

APIKGVTFGEDTVWEVQGYKNVRITFELDERVDKVLNEKCSVYTVESGTEVTEFACVVAEAVVKTLQPVSDLLTN  
MGIDLDEWSVATFYLFDDAGEENFSSRMYSFYPPDEEEEDDAECEEEIDETCEHEYGTEDDYQGLPLEFGASAE  
TVRVEEEEEEDWLDDETTQSEIEPEPEPTPEEPVNQFTGYLKLTDNVAIKCVDIVKEAQSANPMVIVNAANIHLKH  
GGGVAGALNKATNGAMQKESDDYIKLNGPLTVGGSCLLSGHNLAKKCLHVVGPNLNAGEDIQLLKAAYENFNS  
QDILLAPLLSAGIFGAKPLQSLQVCVQTVRTQVYIAVNDKALYEQVVM DYLDNLKPRVEAPKQEEPPNTEDSKTEE  
KSVVQKPVDPKPKIKACIDEVTTTLEETKFLTNKLLLFADINGKLYHDSQNMLRGEDMSFLEKDAPYMGVDVITSG  
DITCVVIPSCKAGGTTEMLSRALKKVPVDEYITTYPGQGCAGYTLLEEAKTALKKCKSAFYVLPSEAPNAKEEILGTVS  
WNLREMLAHAEETRKLMPICMDVRAIMATIQRKYKGIKIQEGIVDYGVRRFFYTSKEPVASIIKLNLSNEPLVTMPI  
GYVTHGFNLEEAARCMRSLKAPAVSVSSPDAVTTYNGYLTSSSKTSEEHFVETVSLAGSYRDWSYSGQRTELGVE  
FLKRGDKIVYHTLESPVEFHLDGEVLSLDKLSLLSLREVKTIKVFTTVDNLTNLHTQLVDMSMTYGGQFGPTYLDG  
ADVTKIKPHVNHGKTFVLPSSDTRLSEAFYHTLDESFLGRYMSALNHTKKWKFPQVGGGLTSIKWADNNCYL  
SSVLLALQQLEVKFNAPALQEAYYRARAGDAANFCALILAYSNTKTVGELGDVRETMTLLQHANLESKRVLNVV  
CKHCGQKTTTLTGVEAVMYMGTLSDNLKTGVSIPCVCGRDATQYLVQQESSFVMMMSAPPAEYKLLQGGTFLCAN  
EYTGNYQCGHYTHITAKETLYRIDGAHLTKMSEYKGPVTDVFKETSYYTTTIKPVSYKLDGVTYTEIEPKLDGYKK  
DNAYYTEQPIDLVPQPLPNASFDNFKLTCSENTKFADDLNQMTGFTKPASRELSVTFPPDLNGDVVAIDYRHYSASF

KKGAKLLHKPIVWHINQATTKTTFKPNTWCLRCLWSTKPVDTNSFEVLAVEDTQGMDNLACESQQPTSEEVVE  
NPTIQKEVIECDVKTTEVVGNVILKPSDEGVKVTQELGHEDLMAAYVENTSITIKKPNELSLALGLKTIATHGIAAIN  
SVPWSKILAYVKPFLGQAAITTSNCAKRLAQRVFNMYMPYVFTLLFQLCTFTKSTNSRIRASLPTTIKNSVKSVAKL  
CLDAGINYVKSFKFLFTIAMWLLLLSICLGLSICVTAAGVLLSNFNGAPSVCNGVRELYLNSSNVTTMDFCEGSFP  
CSICLSGLDSLDSYPALETIQVTISSYKLDLTILGLAAEWVLAYMLFTKFFYLLGLSAIMQVFFGYFASHFISNSWLMW  
FIISIVQMAPVSAMVRMYIFFASFYIWKSYVHIMDGCTSSTCMMCCKYKRNRASTRVECTTIVNGMKRSFYVYANGGR  
GFCKTHNWNCLNCDTFCSTGTFISDEVARDSLQFKRPINPTDQSSYIVDSVAVKNGALHLYFDKAGQKTYERHPL  
SHFVNLDNLRANNTKGLPINVIVFDGKSKCDESASKSASVYYSQLMCQPILLDDQVLVSDVDGSTEVSVMFDAY  
VDTFSATFVPMKALKALVATAHSELAKGVALDGVLTSTFVSAARQGVVDTVDVTDKDVIECLKLSHSDLEVTGDS  
CNNFMLTYNKVENMTPRDLGACIDCNARHINAQVAKSHNVSLIWNVVDYMSLSEQLRKQIRSAKKNIPFRLTC  
ATTRQVVNVITTKISLKG

FJ882957.1 nsp3 protein@integer

22251764780498508662900706704881061791370592392233054498561599527526035926994290162583575727803797  
0079456369683716182942188467312201263441094042626345056791810005179646738910203764989246586178515616  
036576435279955790960775742024846364044712882586044283349705189368939288718612718352394602569006177  
8577820452310625792256810749857850057798332389927514158138383958560759411921140695431479845011940275  
3269305441989944209790575018093177015010490165148498953061560601892172004117228992829920312022576151  
74021550271041027874115958277711168626022235465631711413434316880549699838916690812452164273268783614  
83268186967164334232430919308913620448325775082285532352800690293774558351970380013906244406243665  
5310695754290276100785791514632888009080879155532108633065362744796171690917693420925716091218499303  
48390559981603440229206505784202779460931051496397561591899339770537373688629817924265888253346578  
6371958612500201490594621103009656466974958120120242341997773428449712086390317784595013995393725  
752631783013908836927385733813863467418470812239484576044571665283249370817388186961884553679735836  
97058994459035589962088592552706988502136810347823544897005692998923373766995357475062599856818123  
6568142064832563979405408132531095177855559789484227056127851319175073537145094925376430507119897837  
1233001268130752040561131006708078736608688812404298324291632142415476038860414827264812731647793310  
2072547035424445861691012263260657709316093242616782110940119168019746492923760891129780256900982222  
837519480664598570246052280033287444136293044775943634690019454138359108619126109772992722614686613  
477033835607094640171096749637079223351972988775583109526758008203227006744543288584970162101963624  
516196007391597934428543049850008552477005562304788742255418789562064103457499707083179826128853777  
33772620215801354158020168509011297733184212599341742474744682731025342741760831884179534482362885693  
142552372453304978815355057732855994234334772201481634784945724202411041657740559683766703446347805  
60888757205183884339143523878902040885485029296244402180526250573816195573180932262748932846398691  
2901367053188830238802905147850315802611198108893717533988567900533212261075337832366108121643721056  
71125044741641210120736075452752374441347015835298190746816807192372121822590585048203857781735475425  
343197119202637566603401129028362805706605328331535334184067756621229305439717183403235012275345795  
724390184138498091557050489073214238791939332212204453158540234971630006973271934995948584862081222  
4352391193980800

FJ882957.1 nsp4

KIVSTCFKMLKATLLCVLAALVCYVMPVHTLSIHDGYTNEIIGYKAIQDGVTRDIISTDDCFANKHAGFDAWFSQ  
RGGSYKNDKSCPVAIIITREIGFIVPGLPGTVLRAINDFLHFLPRVFSAVGNICYTPSKLIEYSDFATSACVLAEC  
TIFKDAMGKVPYCYDTNLLEGSISYSELRPDTRYVLMDGSIQFPNTYLEGSRVVTTFDAEYCRHGTCESEVVIC  
LSTSGRWVLNNEHYRALSQVFCGVDAMNLIANIFTPLVQPVGALDVSASVVAGGIIAILVTCAYYFMKFRRVFGE  
YNHVVAANALLFLMSFTILCLVPAYSFLPGVYSVFYLYLTFYFTNDVSFLAHLQWFAMFSPVFPWITAIYVFCISLKH

CHWFFNNYLKRVMFNGVTFSTFEEAALCTFLLNKEMYLKLRSETLLPLTQYNRYLALYNKYKYFSGALDTTSYREA  
ACCHLAKALNDFSNSGADVLYQPPQTSITSAVLQ

FJ882957.1 nsp4 protein@integer

203118525363648375635703927997417666467954986267195069483795421697623567169311200802037841575572  
977015863360127591610455864438100274555564638297329326112035907440430582693582593295941512160328  
906209346916665061042384772887308871439376075265804740883403372223329862054323036845582833447196  
216836420903369399818973585141054330888684056255133275268844252320448108351010243568512956113346  
651069866125521065188270373043569983136493086610524074243052454358297810675294516263336828349943  
818100680314466430471058281194726139171331982414079181067901489450818168450252156227335399122671  
374080928464791233610281110647493129092469665452429167964861066883088329751

FJ882957.1 nsp5

SGFRKMAFPSGKVEGCMVQVTCGTTTTLNLWLDDTVYCPRHVICTAEDMLNPNYEDLLIRKSNHSFLVQAGNV  
QLRVIGHSMQNCLLRKVDTSNPKTPKYKFVRIQPGQTFSVLACYNGSPSGVYQCAMRPNYTIKGSFLNGSCGSVG  
FNIDYDCVSFCYMHMELPTGVHAGTDLEGKFGPFVDRQTAQAAGTDTTITLNLVLAWLYAAVINGDRWFLNRF  
TTTTLNDFNLVAMKYNYEPLTQDHVDILGPLSAQTGIAVLDMCAALKALLQNGMNGRTILGSTILEDEFTPFVVR  
QCSGVTFQ

FJ882957.1 nsp5 protein@integer

133121402894115672926967270832318070227879669393706699338549624807431728714871822876994598589776  
17976350027545854778779346785394768811269711127466640684288541729289653571056822143505314696758  
998053799125528969432583168787024645960321431503460495103321145762694714961537258550673885730150  
883332691010164105433157382905100130233723313537560998121777466621603801454458144983229489564733  
87457670434331

FJ882957.1 nsp6

GKFKKIVKGTHHWMLLFTLSTLLILVQSTQWSLFFFVYENAFPLPFTLGIMAIAACAMLLVKHKHAFCLFLLPSLATV  
AYFNMVYMPASWVMRIMTWLELADTSLSGYRLKDCVMYASALVLLILMTARTVYDDAARRVWTLMNVTILVYKV  
YYGNALDQAISMWALVISVTSNYSVGVVTTIMFLARAIVFCVEYYPFLFITGNTLQCIMLVYCFLGYCCCCYFGLFC  
LLNRYFRLTLGVYDYLSTQEFYRMSQGLLPKSSIDAFKLNKLLGIGGKPCIKVATVQ

FJ882957.1 nsp6 protein@integer

638142868916862997228520692524021230041040388191798355873592896055472141822345788987818368520406  
542321137273725973520614056013319828368107835726222525473550087008123706808904944745885824413104  
719935450148825707507515117661194560822373431188680876468197888856225932881782148709884281013697  
1971797020354488234854984162685531849879320708401270179313560570935618797574229031050091

FJ882957.1 nsp7

SKMSDVKCTSVVLLSVLQQLRVESSEKSLWAQCQVLHNDILLAKDTTEAFKMSVLLSVLLSMQGAVDINRLCEEM  
LDNRATLQ

FJ882957.1 nsp7 protein@integer

127435962709688865300893277401241546623400041889338055465541234671424254488291793017563004712280944  
031530151

FJ882957.1 nsp8

AIASEFSSLPSYAAAYATAQEAYEQAVANGDSEVVLKLLKSLNVAKSEFDRDAAMQRKLEKMADQAMTQMYKQA  
RSEDKRAKVTSAMQTMFTMLRKLNDALNNIINNARDGCVPLNIPLTTAAKLMVVVPDYGTYNKTCDGNTFTY  
ASALWEIQVVDADSKIVQLSEINMDNSPNLAWPLIVTALRANSVAVKLQ

FJ882957.1 nsp8 protein@integer

281783874145942352222809790648129479483768552279803349551389841110187982540677607729209432315894772  
537546532833183379982380341227806450820953775824781730703890168558053348369181244571956074423060422  
84877516424112277662570106912625757283397485582261062596951

FJ882957.1 nsp9

NNELSPVALRQMSCAAGTTQTACTDDNALAYYNNNSKGGRFVLALLSDHQLKWARFPKSDGTGTIYAELEPPCRF  
VTDTPKGPVKYLYFIKGLNNLNRGMVLGSLAATVRLQ

FJ882957.1 nsp9 protein@integer

492405683224698977802404933489923142696369254917797064817408072335472622047159995226203861866261  
302571776548429673257913850126331143034940228038951

FJ882957.1 nsp10

AGNATEVPANSTVLSFCAFAVDPAKAYKDYLASGGQPITNCVKMLCTHTGTGQAITVTPEANMDQESFGGASCC  
LYCRCHIDHPNPKGFCDLKGKQVQIPTTCANDPVGFTLRNTVCTVCGMWKGYGCSQDQLREPLMQ

FJ882957.1 nsp10 protein@integer

356345132828815625062945762014901276115526234637480778835384105902483302627505099736859255160679570  
685921575092919674742688778874593128114386688286668567798284027331812622584507091

FJ882957.1 nsp11

SADASTFLNGFAV

FJ882957.1 nsp11 protein@integer

8499770647846424

FJ882957.1 nsp12

SADASTFLNRVCGVSAARLTPCGTGTSTDVVYRAFDIYNEKVAGFAKFLKTNCCRFQEKDEEGNLLDSYFVVKRH  
TMSNYQHEETIYNLVKDCPAVAVHDFKFRVDGDMVPHISRQRLTKYTMADLVYALRHFDEGNCDTLKEILVTYN  
CCDDDYFNKDWYDFVENPDILRVYANLGERVRQSLKTVQFCDAMRDAGIVGVLTLDNQLNGNWDYDFGDF  
VQVAPGCGVPIVDSYSLLPILTALRALAAESHMDADLAKPLIKWDLKYDFTEERLCLFDYFYWDQTYHPNC  
INCLDDRCILHCANFNVLSTVFPPTSFGPLVRKIFVDGVPFVSTGYHFRELGVVHNQDVNLHSSRSLFKELLVYA  
ADPAMHAASGNLLDKRTTCFSVAALTNVAFQTVKPGNFNKDFYDFAVSKGFFKEGSSVELKHFFFAQDGNAAI  
SDYDYRYNLPTMCDIRQLLFVVEVDKYFDCYDGGCINANQVIVNNLDKSAGFPFNKWKARLYYDSMSYEDQ  
DALFAYTKRNVIPITITQMNLYAISAKNRARTVAGVSIKSTMTNRQFHQKLLKSIAATRATVIGTSKFGYGGWHN  
MLKTVYSDVETPHLMGWDPKCDRAMPNMLRIMASLVLARKHNTCCNLSHRFYRLANCAQVLSEMVMCGGSL  
YVPGGTSSGDATTAYANSVFNICQAVTANVNALLSTDGNKIADKYVRNLQHRLYECLYRNRDVEHFVDEFYAY  
LRKHFSMMILSDDAVVCYNSNYAAQGLVASIKNFKAVLYYQNNVFMSEAKCWTETDLTKGPHFCSQHTMLVKQ  
GDDYVYLPYPDPSRILGAGCFVDDIVKTDGTLMIERFVSLAIDAYPLTKHPNQEYADVFLYLYQYIRKLHDELTHG  
MLDMYSVMLTNDNTSRYWEPEFYEAMYTPHTVLQ

FJ882957.1 nsp12 protein@integer

376680572978437912631134182029092227889210241393828876099195596571150973055551272074131822814569  
008335062293805327477199330175702138058874694365932824765620428655699712294809328273836392213583  
443352030591858625845791011903587069294950450868994192562274399282349189249451048195888538005224

774186432419003396673326632646584104021429713992298983834284258358055481258288247088244267429570  
019367206333196210648902759596664289029687459125756529482105181304492862792745024369356150338362  
479762145463148616511164291432953909840432375710915101394719073505580532120542977586842038306807  
588179267726514533802427717062635137191191920088192315995840785719388060195059760917281553333811  
269490065896229973302237676603182874782298252358124121491741757651757375598049593310400691161203  
152459463756445041530768123939012532734321845141595737784874468119182082954234163144683900965392  
641250256577123522976255497585039642803316087078142988167506574861674336058593171347020736649367  
508851514183359386591097832034274361583353059049803205020499921265643517331004669162897455991595  
602368869644431832122461785453850084590908929450213610399240514621329927694811827686134209399093  
932130004834946903310612115150726511447021473516413772041751

FJ882957.1 nsp13

AVGVCVLCNSQTSRLRCGACIRRPFLCCKCCYDHSVISTSHKLVLSVNPYVCNAPGCDVTDVTQLYLGGMSYYCKSH  
KPPISFPLCANGQVFLYKNTCVGSDNVTDFNAIATCDWTNAGDYILANTCTERLKLFAAETLKATEETFKLSYGLA  
TVREVLSDRELHLSWEVGKPRPPLNRNVFTGYRVTKNSKVQIGEYTFEKGDYGDVVYRGTTTYKLVNGDYFVL  
TSHTVMPLSAPTLVPQEHYVRITGLYPTLNISDEFSSNVANYQKVGMQKYSTLQGPPTGKSHFAIGLALYPSARIV  
YTACSHAAVDALCEKALKYLPIDKCSRIIPARARVECFDKFKVNSTLEQYVFCTVNALPETTADIVVFDEISMATNYD  
LSVVNARLRKHYVYIGDPAQLPAPRTLLTKGTLEPEYFNSVCRLMKTIGPDMFLGTCCRCPAEIVDTSALVYDNK  
LKAHKDKSAQCCKMFYKGVITHDVSSAINRPQIGVVREFLTRNPAWRKAVFISPYNSQNAVASKILGLPTQTVDSSQ  
GSEYDYVIFTQTTETAHSCNVNRFNVAITRAKIGILCIMS DRDLYDKLQ

FJ882957.1 nsp13 protein@integer

152028036772266543320134623793292677612923982017753604447658546914814596580297462316739899129642  
485973308129480667756139990888976305974915831157146042472018354766493899243821700433165773742959  
409653699711954044546858201364464009817138033374967498798967299611350098800576223929965173012591  
071443996953308392677036750061833650842895438553418077743063986935286484849672441260468113735110  
869482871246006014548098716983914808629357736409368278110867320915808273062355894678211633971651  
416002810680560664260126468423025744636900966616633793127712918145260697408251113563954676405705  
087347036575197742046651894706828867339606655488949779805408930152344565237070877796707675905975  
204612189669544430182143067185795409890054224595631290508710779641297646986999062425364951

FJ882957.1 nsp14

FTSLEIPRRNVATLQAENVTLGFKDCSKIITGLHPTQAPTHLSVDIKFKTEGLCVDIPGIPKDMTYRRLISM MGFKMN  
YQVNGYPNMFITREEAIRHVRAWIGFDVEGCHATRDAVGTNLPLQLGFSTGVNLVAVPTGYVDTENNTFTRVN  
AKPPPGDQFKHLIPLMYKGLPWNVRIKIVQMLSDTLKGLSDRVFVLWAHGFELTSMKYFVKIGPERTCCLCDKR  
ATCFSTSSDYACWNHSGFDYVYNPFMIDVQQWGFTGNLQSNHDQHCQVHGNAHVASCDAIMTRCLAVHE  
CFVKRVDWSVEYPIIGDELRVNSACRQVQHMVVKALLADKFPVLHDIGNPKAIKCVPAEVEWKFYDAQPCSDK  
AYKIEELFYSYATHHDKFTDGVCLFWNCNVDRYPANAI VCRFDTRVLSNLNLP GCDGGSLYV NKHAFHTPAFDKS  
AFTNLKQLPFFYSDSPCESHGKQVVS DIDYVPLKSATCITRCNLGGAVCRHHANEYRQYLDAYNMMISAGFSLWI  
YKQFDTYNLWNTFTRLQ

FJ882957.1 nsp14 protein@integer

117014651740170217788952223182358954049557526164915959360008345728327200740960515904182742319055  
942740199235100335350645768367187855705168010873407116786719436179487537314383971038926170782577  
967412358837432784531757517176364257963426409664490675554879140939443602777496315391840081869338  
661867388197207909080662295786791081641090075150471352613694199003220580839273629392601658384888  
505108191019694509948019646560396919304521540395690565799916048876967648974300516889270630189168

856116143610962011984727697561744289870022204530013121779006811085245470336622156203188415264792  
892478770179445950165488442068983603215477344931010545633297611623372630944932622855754625315821  
2376023258745162671293132114606951

FJ882957.1 nsp15

SLENVAYNVVVKGHFDGHAGEAPVSIINNAVYTKVDGIDVEIFENKTTLPVNVAFELWAKRNIKPVPEIKILNNLGV  
DIAANTVIWDYKREAPAHVSTIGVCTMTDIAKKPTESACSSLTVLFDGRVEGQVDFLRNARNGVLITEGSVKGLTP  
SKGPAQASVNGVTLIGESVKTQFNFKKVDGIIQLPETYFTQSRDLEDFKPRSQMETDFLELAMDEFIQRYKLEGY  
AFEHIVYGDVSHGQLGGLHLMIGLAKRSQDSPLKLEDFIPMDSTVKNYFITDAQTGSSKVCVSVIDLLDDFVEIIS  
QDLSVISKVVKTIDYAEISFMLWCKDGHVETFYPKLQ

FJ882957.1 nsp15 protein@integer

170840604821066470660212059655380023827629639736344590965144864375184385260586519857623031056044  
890793458975789590172288633839662472981644135768594888519762980210851696441418015660653308745502  
547954041727764005587638796172998704104589270241470299584928398977169307770765963214816446113543  
514510535008708698722135301073747288919655817622766930762681970561147650760473874943447072767392  
889392370661671660328196640095184708408013707279288327799414108951

FJ882957.1 nsp16

ASQAWQPGVAMPNLYKMQRMLLEKCDLQNYGENAVIPKGINNVAKYTQLCQYLNTLTLAVPYNMRVIHFGA  
GSDKGVAPGTAVLRQWLPTGTLLVDSLDNDFVSDADSTLIGDCATVHTANKWDLIISDMYDPRTKHVTKENDSK  
EGFFTYLCGFIKQKLALGGSIAVKITEHSWNADLYKLMGHFSWWTAFVTNVNASSSEAFILGANLYLGGPKKEQIDGY  
TMHANYIFWRNTNPIQLSSYSLFDMSKFLKLRGTAVMSLKENQINDMIYSLLEKGRLLIRENNRNVVSSDILVNN

FJ882957.1 nsp16 protein@integer

287157711119834951697616545777635806517867460474631425692966083766217798057508263135260174109631  
280784403017545539321356861210296182571774186514937590306403092142687041081165302429917167400416  
322397932781826877066193785940330113563057079736955630800379407039309945815957674096488192895648  
125354208932958854766331161661018357953034766767975600301725140889064590894018647612868713121337  
789

Table1 GCD between SARS-CoV-2 and pulmonary microbes

SARS-CoV-2 (NC\_045512.2)

S

>YP\_009724390.1 surface glycoprotein [Severe acute respiratory syndrome coronavirus 2]

MFVFLVLLPLVSSQCVNLTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIHV  
SGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTLLDSTKQSLIVNNTNVIKVFCEQFCNDPF  
LGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFNIDGYFKIYSKHTPI  
NLVRDLPQGFSALEPLVDLPIGINITRFQTLALHRSYLTGPDSSSGWTAGAAAYVGYLQPRTFLLKYN  
ENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASV  
YAWNKRISNCVADYSVLYNSASFSTFKCYGVSPTKLDLNCFTNVYADSFVIRGDEVRQIAPGQTGKIAD  
YNYKLPDDFTGCVIAWNSNNLDSKVGGNYNLYRLFRKSNLKPFERDISTEIQAGSTPCNGVEGFNCYF

PLQSYGFQPTNGVGYQPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFL  
PFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLT  
PTWRVYSTGSNVFQTRAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPRRARSVASQSIIAYTMSLG  
AENSVAYSNNIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECNLLLQYGSFCTQLNRALTGI  
AVEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSPKRSFIEDLLFNKVTLADAGFIKQYGDC  
LGDIAARDLICAQKFNGLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIG  
VTQNVLYENQKLIANQFNSAIGKIQDLSSTASALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDI  
LSRLDKVEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLM  
SFPQSAPHGVVFLHVTVPAQEKNFHTAPAICHDKAHFPREGVFSVNGTHWFVTQRNFYEPQIITTDNT  
FVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVLGDISGINASVVNIQKEIDRLNEVA  
KNLNESLIDLQELGKYEYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTSCCCLKGCCSCGSCCKFDEDD  
SEPVKGVKLYHT

S (YP\_009724390.1) protein@integer

120439657203882774484390066932584916898659202256493842498812594333330854483743286377701022981821  
031063561692048351038006678781157706612192808938101284633541757413763423983981946276212481608696  
731979217004898006232586357136905969133445289743811349130465003480541971088280609707714138848849  
356346229448627993666287684759527276245742621823038178561145607504573064271778156159732239979547  
95310161681466684119254997050797204245145540333247660267773364120436910809810160629419636986439  
400059623223648039943751471783080005224864420933818793839229405778501033919706311547286517248699  
280660555138463160733657264613809405911099826282678260572708345848386266480053934338769028869491  
682487296020832270652414675645379801524346344827439716273440292210098642926332847447598106390278  
392018333564757733948358245199818529845955360710788566225821301158140373667773769964218277122658  
538876619372594136703042431242205399939662841495858073511758707053788106830686332759998721251141  
861138473854187844090884482272637578852431406234905684004159968044423963658425114348706579790572  
643646934697766217879929783983663618477090509370407971890783344790313957920989990494728270760509  
776015806896251653606269697319739925895690127204512148606560243611151759314322589794035924203986  
164656715532677666727366646529148537937143920231248630014874030906841467792393310564445574308321  
487320097392295346624975526588885932282027694981568319839419111253496862159322806300095755234862  
898857225696691104469675189065953809567050618882114384887323695443751776468558174083709579375183  
617358537746254154588352095630880476686902050661133141940384240661597477754999913797979771429677  
5069319998087362574782365

E

>YP\_009724392.1 envelope protein [Severe acute respiratory syndrome coronavirus 2]

MYSFVSEETGLIVNSVLLFLAFVVFLLVTLAILTALRLCAYCCNIVNVSLVKPSFYVYSRVKNLNSSRV  
PDLLV

E (YP\_009724392.1) protein@integer

281585756715583860382754061271694289415355617329836193393526618386393321015201910475515428613629  
44

M

>YP\_009724393.1 membrane glycoprotein [Severe acute respiratory syndrome coronavirus 2]

MADSNGTITVEELKKLLEQWNLVIGFLFTWICLLQFAYANRNRFLYIIKLIFLWLLWPVTLACFVLA  
YRINWITGGIAIAMAACLVGLMWLSYFIASFRLFARTRSMWSFNPETNILLNVPLHGTILTRP  
LLESELVI  
GAVILRGHLRIAGHHLGRCDIKDLPKEITVATSRTLSYYKLGASQRVAGDSGFAAYSRYRIGNYKLNTDH



SSSSDNIALLVQ

M (YP\_009724393.1) protein@integer

474336885130087246559227181474873626505021258393791115212814809842932891146570677880164818373755  
020183216971990705471677016821370632918533688487363182856097368741814388171137448805653084419936  
514188608412885771729978230649964328609523546978953420569604994275957386736268990507116860181889  
1

N

>YP\_009724397.2 nucleocapsid phosphoprotein [Severe acute respiratory syndrome coronavirus 2]

MSDNGPQNQRNAPRITFGGPSSTGSGNQNNGERSGARSKQRRPQGLPNNTASWFTALTQHGKEDLKFPRGQ  
GVPINTNSSPDDQIGYYRRATRIRGGDGMKDLSPRWYFYLLGTGPEAGLPYGANKDGIIWVATEGALN  
TPKDHIHTRNPANNAIVLQLPQGTTLPKGFYAEGSRGGSQASSRSSSRNSSRNSTPGSSRGTSPARM  
AGNGGDAALALLLLDRLNQLKMSGKGGQQGQTVTKKSAEASKKPRQKRTATKAYNVTQAFGRRGPE  
QTQGNFGDQELIRQGTDYKHWPQIAQFAPSASAFFGMSRIGMEVTPSGTWLTYTGAIKLDDKDPNFKDQV  
ILLNKHIDAYKTFPTEPKKDKKKKADEQALPQRQKKQTVTLLPAADLDDFSKQLQQSMSSADSTQA

N (YP\_009724397.2) protein@integer

956227582509132504718849355695730693961321517736652145117075131168374450978223364454488348970423  
118447512373703160229961005598661705320883242954575894810810193574332949149747712379816770960182  
246055380168976963162519036258142944758612842438988879411378433758788132669340431425716988813342  
262680761690591744928934324948193050865092822456307472028570449876463718569780832663406116693176  
099354459884651197599680303878827228346254307907969244760770453400593740764587442418141468590139  
63729877390314018379371238969097694156462631243888958825092818221

ORF1ab

>YP\_009724389.1 ORF1ab polyprotein [Severe acute respiratory syndrome coronavirus 2]

MESLVPGFNEKTHVQLSLPVLQVRDVLVRGFGDSVEEVLSEARQHLLKDGTCGLVEVEKGVLPQLEQPYVF  
IKRSDARTAPHGHVMVELVAELEGIQYGRSGETLGVLPVHVGEPVAYRKVLLRKNNGKAGGHSYGADL  
KSFDLGDELGTDPYEDFQENWNTKHSSGV TRELMRELNGGAYTRYVDN NFCGPDGYPLECIKDLLARAGK  
ASCTLSEQLDFIDTKRGVYCCREHEHEIAWYTERSEKSYELQTPFEIKLAKKFDTFNGECPNFVFP LNSI  
IKTIQPRVEKKKLDGFMGRIRSVYPVAVSPNECNQMCLSTLMKCDHCGETSWQTGDFVKATCFEFCGTENLT  
KEGATTCGYLPQNAVVKIYCPACHNSEVGPESHSLAEYHNESGLKTILRKGRTIAFGGCVFSYVGCHNKC  
AYWVPRASANIGCNHTGVVGESEGLNDNLLEILQKEKVNINIVGDFKLN EEAIIILASFSASTSAFVET  
VKGLDYKAFKQIVESCGNFVKTKGKAKKGAWNIGEQKSILSPYAFASEAARVRSIFSRTLETAQNSVR  
VLQKAAITILDGISQYSLRLIDAMMFTSDLATNNLVVMAYITGGVVQLTSQWLTNIFGTVYEKLPVLDW  
LEEKFKEGVEFLRDGWEIVKFISTCACEIVGGQIVTCAKEIKESVQTFKLVNKFLALCADSIIIGGAKL  
KALNLGETFVTHSKGLYRKCVKSREETGLLMPLKAPKEIIFLEGETLPTEVLTEE VVLKTGDLQPLEQPT  
SEAVEAPLVGTPVCINGMLLEIKDTEKYCALAPNMMVTNNTFTLKG GAPTQVTFGDDTVIEVQGYKSVN  
ITFELDERIDKVLNEKCSAYTVELGTEVNEFACVVADAVIKTLQPVSELLTPLGIDLDEWSMATYYLFDE  
SGEFKLASHMYCSFYPPDEDEEEGDCEEEEFEPSTQY EYGTEDDYQGKPLEFGATSAAALQPEEEQEEDWL  
DDDSQQTVGQQDGS EDNQT TTIQTIVEVQPQLEMELTPVVQTIEVNSFSGYLKLTDNVYIKNADIVEEAK  
KVKPTVVVNAANVYLKHGGGVAGALNKATNNAMQVESDDYIATNGPLKVG GSCVLSGHNLAKHCLHV VGP  
NVNKGEDIQLLSAYENFNQHEVLLAPLLSAGIFGADPIHSLRVCVDTVR TNVYLA VFDKNLYDKLVSSF  
LEMKSEKQVEQKIAEIPKEEVKPFITESKPSVEQRKQDDK KIKACVEEVT TTTLEETKFLTENLLLYIDIN  
GNLHPDSATLVSDIDITFLKKDAPYIVGDVVQEGVLTAVVIPTKKAGGTTEMLAKALRKVPTDNYITTP  
GQGLNGYTVEEAKTVLKKCKSAFYILPSIISNEKQEILGTVSWNLREMLAHAEETRKLMPVCVETKAIVS

TIQRKYKGIKIQEGVVDYGARFYFYTSKTTVASLINTLNDLNETLVTMPLGYVTHGLNLEEAARYMRSK  
VPATVSVSSPDAVTAYNGYLTSSSKTPEEHFIETISLAGSYKDWSYSGQSTQLGIEFLKRGDKSVYYTSN  
PTTFHLDGEVITFDNLKTLTSLREVRTIKVFTTVDNINLHTQVVDMSMTYQQFGPTYLDGADVTKIKPH  
NSHEGKTFYVLPNDLTLRVEAFEYHTTDPFSLGRYMSALNHTKKWKYPQVNGLTSIKWADNNCYLATAL  
LTLQQIELKFNPPALQDAYRRARAGEAANFCALILAYCNKTVGELGDVRETMSYLFQHANLDSCKRVLNV  
VCKTCGQQQTTLKGVEAVMYMGTLSEYQFKKGVQIPCTCGKQATKYLQVESPFVMMMSAPPAQYELKHGT  
FTCASEYTGNYQCQGHYKHITSKETLYCIDGALLTKSSEYKGPITDVFYKENSYTTTTIKPVTYKLDGVVCT  
EIDPKLDNYYKKNDSYFTEQPIDLVPNQYPNASFDNFKVCNFKADDLNQLTGYKKPASRELKVTFE  
PDLNGDVVAIDYKHYTPSFKKGAKLLHKPIVWHVNNATNKATYKPNWTCIRCLWSTKPVETSNSFDVLKS  
EDAQGMNDLACEDLKPVEEVVENPTIQKDVLECNVKTTEVVGDIIKLPANNSLKITEEVGHTDLMAAYV  
DNSSLTIKKNELSRVLGLKTLATHGLAAVNSVPWDTIANYAKPFLNKVVSTTTNIVTRCLNRVCTNYMP  
YFFTLQLCTFTRSTNSRIKASMPPTIAKNTVKSVMGKFCLEASFNYLKSPNFSKLINIIWFLLSVCL  
GSLIYSTAALGVLMSNLGMPYCTGYREGYLNSTNVTIATYCTGSIPCSVCLSGLDSDTYPSELETIQIT  
ISSFKWDLTAFGLVAEWFVFLAYILFTRFFYVGLAAIMQLFFSYFAVHFISNSWLMWLIINLVQMAPISAM  
VRMYIFFASFYVWKSIVHVDGCSNSTCMMCYNRNRATRVECTTIVNGVRRSFYVYANGGKGFCKLHNW  
NCVNCDFCAGSTFISDEVARDLSLQFKRPINPTDQSSYIVDSVTVKNGSIHLYFDKAGQKTYERHSLSH  
FVNLNLRANNTKGLPINVIVFDGKSKCEESSAKSASVYYSQLMCQPILLDQALVSDVGDSDAEVAVKM  
FDAYVNTFSSTFNVPMEKLTAVATAEAEAKNVSLDNVLSVFISAARQGFVDSVETKDVVECKLSHQ  
SDIEVTGDSNNYMLTYNKVENMTPRDLGACIDCSARHINAQVAKSHNIALIWNVKDFMSLSEQLRKQIR  
SAAKNNLPPFKLTCAATTRQVVNVVTTKIALKGGKIVNNWLKQLIKVTLVFLFVAAIFYLITPVHVMKHT  
DFSSEIIGYKAIDGGVTRDIASDTCFANKHADFDTWFSQRGGSYTNDKACPLIAAVITREVGFFVPLP  
GTILRTTNGDFLHFLPRVFSVAVGNICYTPSKLIEYDFATSACVLAECTIFKDASGKVPYCYDTNVLE  
GSVAYESLRPDTRYVLMGSIQFPNTYLEGSRVVTTFDSEYCRHGTCEERSEAGVCVSTSGRWVLNNDY  
YRSLPGVFCGVDAVNLLTNMFTPLIQPIGALDISASIVAGGIVAVVTCLAYYFMRFRRAFGEYSHVVA  
NTLLFLMSFTVLCLTPVYSFLPGVYSVIYLYLTFYLTNDVSFLAHIQWMVMFTPLVPFWITIAIICIST  
KHFYWFFSNYKRRVVFNGVSFSTFEEAALCTFLLNKEMYLKLRSVLLPLTQYNRYLALYNKYKYFSGA  
MDTTSYREAACCHLAKALNDFSNSGSDVLYQPPQTSITSAVLQSGFRKMAFPPSGKVEGCMVQVTCGTTTT  
NGLWLDVVYCPRHVICTSEDMLNPNYEDLLIRKSNHNFLVQAGNVQLRVIGHSMQNCVLLKLVDTANPK  
TPKYKFVRIQPGQTFVLACYNGSPSGVYQCAMRPNFTIKGSFLNGSCGSGVFNIDYDCVFCYMHMEL  
PTGVHAGTDLEGNFYGPFVDRQTAQAAGDTTITVNLAWLYAAVINGDRWFLNRFTTTLNDFNLVAMKY  
NYEPLTQDHDVILGPLSAQTGIAVLDMCASLKELLQNGMNGRTILGSALLEDEFTPFVVRQCSGVTFQS  
AVKRTIKGTHHWLLLTIITSLLVLQSTQWSLFFFLYENAFLPFAMGIIAMSAFAMMFVKHKAFLCLFL  
LPSLATVAYFNMVYMPASWVMRIMTWLDMVDTSLSGFKLKDCVMYASAVLLILMTARTVYDDGARRVWT  
LMNVLTLVYKVYYGNALDQAISMWALISVTSNYSQVTTVMFLARGIVFMCVEYCPFITGNTLQCIM  
LVYCFGLYFCTCYFGLFCLLNRYFRLTLGVYDYLSTQEFYRMYNSQGLLPPKNSIDAFKLNKLLGVGGK  
PCIKVATVQSKMSDVKCTSVLLSVLQQLRVESKSLWAQCVQLHNDILLAKDTTEAFEKMSVLLSVLLS  
MQGAVDINKLCEEMLDNRATLQAIASEFSSLPSYAAFATAQEAYEQAVANGDSEVVLLKLLKSLNVAKSE  
FDRDAAMQRKLEKMAQAMTQMYKQARSEDKRAKVTSAMQTMFTMLRKLNDALNIIINNARDGCVPLN  
IIPLTAAKLMVIPDYNTYKNTCDGTTFTYASALWEIQQVVDADSKIVQLSEISMDNSPNLAWPLIVTA  
LRANSVAVKLQNNELSPVALRQMSCAAGTTQACTDDNALAYYNTTKGGRFVLALLSDLQDLKWARFPKSD  
GTGTIYTELEPPCRFVTDTPKGPVKYLYFIKGLNLRGMVLSLAATVRLQAGNATEVPANSTVLSFC  
AFAVDAKAYKDYLASGGQPITNCVKMLCTHTGTGQAITVTPEANMDQESFGGASCCLYCRCHIDHPNPK  
GFCDLKGYVQIPTTCANDPVGFTLKNVCTVCGMWKGYGCSQDLREPMLQSADAQSFLNRVCGVSAAR  
LTPCGTGTSTDVYRAFDIYNDKVAGFAKFLKTNCCRFQEKDEDDNLIDSYFVVKRHTFSNYQHEETIYN  
LLKDCPAVAKHDFKFRIDGMVPHISRQRLTKYTMADLVYALRHDFDEGNCDTLKEILVTYNCCDDDYFN

KKDWYDFVENPDILRVYANLGERVRQALLKTVQFCDAMRNAGIVGVLTLDNQDLNGNWDYDFGDFIQTPG  
SGVPVVDSSYLLMPILTLTRALTAESHVDTDLTKPYIKWDLKDYDFTEERLKLFDYFYWDQTYHPNC  
VNCLDDRCILHCANFNVLVSTVFPPTSFGPLVRKIFVDGVPFVSTGYHFRELGVVHNQDVNLHSSRLSF  
KELLYAADPAMHAASGNLLLDKRTTCFSVAALTNNAFQTVKPGNFNKDFYDFAVSKGFFKEGSSVELK  
HFFFAQDGNAAISDYDYRYNLPMTCDIRQLLFVVEVDKYFDCYDGGCINANQVIVNNLDSAGFFPNK  
WGKARLYYDSMSYEDQDALFAYTKRNVIPITITQMNLYAISAKNRARTVAGVSIKSTMTNRQFHQKLLKS  
IAATRGATVVIGTSKFYGGWHNMLKTVYSDVENPHLMGWDPKCDRAMPNMLRIMASLVLARKHTTCCSL  
SHRFYRLANCAQVLESEMVMCGGSLYVKGPGTSSGDATTAYANSVFNICQAVTANVNALLSTDGNKIADK  
YVRNLQHRLYECLYRNRDVTDFVNEFYAYLRKHFSMMILSDDAVVCFNSTYASQGLVASIKNFKSVLYY  
QNNVFMSEAKCWTETDLTKGPHEFCSQHTMLVKQGDDYVYLPYDPSRILGAGCFVDDIVKTDGTLMIER  
FVSLAIDAYPLTKHPNQEYADVHLYLQYIRKLHDELTDGMLDMYSVMLTNDNTSRYWEPEFYEAMYTPH  
TVLQAVGACVLCNSQTSRLCGACIRRPFLCCKCCYDHVISTSHKLVLSVNPYVCNAPGCDVTDVTQLYLG  
GMSYYCKSHKPPISFPLCANGQVFGLYKNTCVGSDNVTDFNAIATCDWTNAGDYILANTCTERLKLFAAE  
TLKATEETFKLSYGIATVREVLSRELHLSWEVGGKPRPPLNRNYVFTGYRVTKNSKVQIGEYTFEKGDYG  
DAVVYRGTTTTYKLVNGDYFVLTSHVMPLSAPTLVPEHYVRITGLYPTLNISDEFSSNVANYQKVGMMQK  
YSTLQPPGTGKSHFAIGLALYPSARIVYTACSHAAVDALCEKALKYLPIDKCSRIIPARARVECFDKF  
KVNSTLEQYVFCVNALPETTADIVVFEISMATNYDLSVNVNARLRAKHVYVIGDPAQLPAPRTLLTKGT  
LEPEYFNSVCRMLKTIGPDMFLGTCRRCPAEIVDTVSALVYDNKLAHKDKSAQCCKMFYKGVITHDVSS  
AINRPQIGVVREFLTRNPAWRKAVFISPYNSQNAVASKILGLPTQTVDSQGEYDYVIFTQTTETAHSC  
NVNRFNVAITRAKVGILCIMS DRDLYDKLQFTSLEIPRRNVATLQAENV TGLFKDCSKVITGLHPTQAPT  
HLSVDTKFKTEGLCVDIPGIPKDMTYRRLISMGMFKMNYQVNGYPNMFITREEAIRHVRAWIGFDVEGCH  
ATREAVGTNLPLQLGFSTGVNLVAVPTGYVDTPNNTDFSRVSAKPPPGDQFKHLIPLMYKGLPWNVVRK  
IVQMLSDTLKNLSDRVVFVLWAHGFELTSMKYFVKIGPERTCCLCDRRATCFSTASDTYACWHHSIGFDY  
VYNPFMIDVQQWGFTGNLQSNHDLYCQVHGNHVASCDAIMTRCLAVHECFVKRVDWTIEYPIIGDELKI  
NAACRKVQHMVKAALLADKFPVLHDIGNPKAIKCVQADVEWKFYDAQPCSDKAYKIEELFYSYATHSD  
KFTDGVCLFWNCNVDRYPANSIVCRFDTRVLSNLNLPDGGSLYVKNHAFHTPAFDKSAFVNLKQLPFF  
YYS DSPESHGKQVVS DIDYVPLKSATCITRCNLGGAVCRHHANEYRLYLDAYNMMISAGFSLWVYKQFD  
TYNLWNTFTRLQSLNVAFNVVNKGHFDGQQGEVPSIINNTVYTKVDGVDVELFENKTTLPVNVAFELW  
AKRNIKPVPEVKILNNGVDIAANTVIW DYKRDAPAHISTIGVCSMTDIAKKPTETICAPLTVFFDGRVD  
GQVDLFRNARNGVLITEGSVKGLQPSVGPQASLNGVTLIGEAVKTQFNYYKKVDGVVQQLPETYFTQSR  
NLQEFKPRSQMEIDFLELAMDEFIERYKLEGYAFEHIVYGDFSHSQLGGLHLLIGLAKRFKESPFLEDF  
IPMDSTVKNYFITDAQTGSSKVCVIDLLDDFVEIISQDLSVSVKVVKTIDYTEISFMLWCKDGHV  
ETFYPKLQSSQAWQPGVAMPNLYKMQRMLLEKCDLQNYGDSATLPKGIMMNVAKYTQLCQYLNTLTLAVP  
YNMRVIHFAGSDKGVAPGTAVLRQWLPTGTLVSDLNDFVSDADSTLIGDCATVHTANKWDLIISDMY  
DPKTKNVTKENDSKEGFFTYICGFIQQKLALGGSVAIKITEHSWNADLYKLMGHFAWWTAFVTNVNASSS  
EAFLIGCNYLGKPREQIDGYVMHANYIFWRNTNPIQLSSYSLFDMSKFPLKLRGTAVMSLKEGQINDMIL  
SLLSKGRLLIENNRVVISSDVLVNN

ORF1ab (YP\_009724389.1) protein@integer

941524785343653806957528235584704781411137876785039813941253212594796324131944527862663381320676  
636329188498351499935963744263109888968217308277742834052238246674102743184535799225022030847381  
809199665505653404795977281795544393689725727202274000391690575315869949860104462420624224892250  
667350226709191492557491377040644386459023911752471294249308928210305186111699841746485425181548  
289025209902584251076025597264067715917248279467054643569730821486053038037971394558822036660617  
375226847016073637650369132813704141823920052583542734387115751514313758624482351780038430669787  
657636963654609712727735920362828671078720571755509503218043094866835151571365567786755542465780

382033319891688766711374532992449134995104795637308130788346672717739566773494588783040070559823  
924351627426367740512778378754259489825495666821199268705961820867095074757593746311650872108904  
054901096904983812376865245588712519715532388210330644387509989560816561467573474687529257686827  
187010492142464354983805200175776387193045218365022941656972311624143663908071251173787590181334  
665328479461974705801858016364908102063369889417553624601400539727413277264258468546205943905724  
223930540020870050325492725616076797112734937387151106893077488470645161071355989549106936773336  
140148832227742146341448469866892351202895329475115009253427848443406163763621172288037351111813  
132535055401421113969005108845101509073508439571968136546771296753076837633817477993451486496342  
781299094754284598310860439250093924090145165880814335311622652391347867814806816520974285078856  
85993313934633609490726926475268311441651710996087159693411761548825668114404335550757998104778  
597940990529127394945729362463500665095219094692397904906306451923913114354704671266151601370868  
323193963883118171503041276207160084713361011466026056009766814884066459079003823279985359474924  
159639341489375033929116842949824905703009596916962056119088709563582459636812397159626203311326  
759886598307506975072660402485030143439098651221031212535171378178220326684447835081274437398328  
151667247900534588793060830605730250066420200916783642623640871289381774324922915301718931753240  
494346076894165627304761542038210883746783448632074854396730686773253905633884766124315182876590  
62309806035754052116453768826397509170456688854049390054667433281234449929859691302359550448178  
533804676508364691486831230625270249538037937322580240667303379436885517301803545402952856034076  
626622104957195120369961480880995389107080937468492297578100447108370210992402507359908525831118  
426832465369040099701504063914718057734952570819742096880107697449648198675769852235541601491996  
491220543196147887501892739496416116735625251920790156091314021154290418503831175995425750325581  
577886345184450371003882156323970981354475446338747285625414039317060114438081046268470423392881  
92096025236833232238557613206635062839501676860725454544823749116561455429218480288276667973034  
189408621773261668740884807521260256458693225346276958177570771166517880881754931776328123588109  
830415092367126321667364577604687803924553254476026235837168795468753535999955107401255792899273  
820364801252707901659046269343804721279115350124236281831634517967882053820923374920692786117358  
246605068338919121086907087355707156834458536417305709387785712234959678575468349666786676708142  
014185838400265263018600788894818250735251034623239730101331382760202988962658165403999565255051  
920418518016920221493455801275186630270480043303619255204048760630666062027530838213514757210016  
806859143951781705081318340859415104877609515411024974006582292139072103971409705049662947489693  
525177398413066748078890924016820892983563520910702446938561607960063974879581974421618399668394  
646616619979396526664955883493070707189599042651633625251751580996519313475795802655101222869593  
531255608988947067421086165030866082122683962375971396006308587127302683872912367734775751005548  
455813519918395417405875582856099454778718485935356175325173358111094277433118127674571325943580  
812371211488257849389887808142328359102884990907684279384506709314569261410755028318516715131879  
684443790141282596724572099775039157969538631304712385258312842896109839892748700169976904241499  
283337768822351434266200321776269211111883684421434900357908105858935812748495352340286056285532  
138911985156565300626122216325938078871651041716620214115003661907497854699702332578781872243909  
673382116941836812654837966731433484406860213580000456269950517967045340467522131896451865419602  
997032669400101462098342388796083129990812847577489338743699814698813006807614838653531906253417  
628686245685084637161242099675638275770604629754479155251032044016510898577209208023102341845352  
434770536713127248214078341363834565644824652190966594775865093724685559279411244718795139515562  
034768654817180969536721189261019999195385279595501755281301620544779621892285192951325344671538  
422329438226742387899404584335776055750936794797911666189420428739034516340941641412447915544574  
764558231174697788922682467758447456459008392902092651487023373042562844866632648422304242331084

506237043608273156815815088072903479127491617325547316197508323337203264223739641776031552456703  
176304998630309211722615871273298338971536692307272276214032710501019225298863509966930551734533  
218149930789518007962130230486292114908154019880775655912174314047400017693208295669167344055363  
603604989032422109785196456691492024349310778905412394950907641332051841022950703406074879199113  
339974821854432949749273150514371141790752995157053857483464271753431346473313785714867087952045  
914884395972459377167370145148233452295514598388293324035015059786467677300636591796464290795276  
035850768533584351864371863435872004673909244103100528343482475547078455513188609982122827807323  
876242040853547932244596666880184003582379742435176708393796619044523599787575513617515127411568  
21538810746116375534796761806603735728757891854655046202870977100224908209986891393385480662023  
879652376784189703883747169775452707754996425809135203826353170444163378339844954959065501935540  
657025810419089127051872206980345913957710942404884859310770399041405483429447094001733522478836  
456220806483341261096471975217171840960642499455207498824927762256547225205432353612698934453911  
123813354510205090595819603215504162731804689304331861885772271438384647764013342277296290681835  
070913955961397242225893044236832884439576793652928927809914158551703963132604868804879256869346  
990766482456788752809364297871187696121559180146192264201191096125587010996903923332309456704179  
374718328274363789763803852486645299951887342751083791875016682196350120179586279355249365010148  
835533454837912696835964444684764469398633697100732661942147150164894647102758242119252038127122  
252046360039219422520838240633120761108822651994435594664499551451632744916359756671818668360132  
400126094008469030397365844289290979018680730018995566772791708057074698293386203101012004142222  
224771550698554079033717214845210349456194010494848565014114938516456286901651960447848719935015  
047195237262810827313080851589969723665786264506379136352779393706697202853715144911159168631030  
962727205802659049059584206175962943336407627977120958494826509839455512022817815731046534175358  
281998007480533067588269421708196009590553212709799928227590243792544797597632588205171956907961  
878247717401779143315902083662790807211141890849309002533080862541555006290232043637793384528370  
526099755690682734904272165408793045578835799439640125946216097606984513481877982959949277775219  
510776490656138060319184042325553151984081377460248107132802603147392968947971667333721053221757  
423359686323164780046093233939613854118140026262923310047931619734623889968866349344343555428400  
750412840202638000447869309641495918020162180033232620574970058595980596371772661299141919573810  
307493649921823616174293810141516656505162645417143735702469157715064547340185029640631999199694  
066624633811973169008469577240882921057930664516803521558248942047312242159427487478862883699655  
586431128484670286917448261470940220449006263580054665328072241613361362913871570607754065500666  
496906970397604571101604560365556193748209848465888422526624270993024589585651067428974348219358  
480793235271979600871326306733903969601415622472180716871533320885631452100349787647290908547374  
121635839137942432202451986182280349631893060004871015908606416744451351087806183974345616329364  
939717879266582512282007946658534542559168418128684735826966589207693716376054201071926804773709  
185661359108160309171034906998204602300886226262179258422794211935702007341152113765324663496060  
47094999987309504463354542693545490284510662718644042459159827063945499668900259435058235222764  
468315490799302225640386116989771226707767964874414478724429312485135162403385888313806977074078  
330240689932523600283890464002033527421080371402463512917046050170636852208661842369180877554339  
504814696128004187660793186455132509640971116478175004871548459065875338936202888467775646145723  
201744012649285061434160348022589700816551155605289330895356671565559698176745066243648830363764  
718959830473244397458916456177129413304297029618257177418651493759030640309214268704108116471774  
157213679939486583222462011528317761234093576719665550777199411620118882697611661923255354975307  
909825602588073528216801123689305676316166101835795303476676797560022441572948724437912057864761  
2868815520697789

ORF1a

>YP\_009725295.1 ORF1a polyprotein [Severe acute respiratory syndrome coronavirus 2]

MESLVPGFNEKTHVQLSLPLVLRDVLVRGFGDSVEEVLSEARQHLKDGTCGLVEVEKGVLPQLEQPYVF  
IKRSDARTAPHGHVMVELVAELEGIQYGRSGETLGLVPHVGEIPVAYRKVLLRKNNGKAGGHSYGADL  
KSFDLGDELGTDPYEDFQENWNTKHSSGVTRELMRELNGGAYTRYVDNNFCGPDGYPLECIKDLLARAGK  
ASCTLSEQLDFIDTKRGVYCCREHEHEIAWYTERSEKSYELQTPFEIKLAKKFDTFNGECPNFVPLNSI  
IKTIQPRVEKKKLDGFMGRIRSVYPVASPNECNQMCLSTLMKCDHCGETSWQTGDFVKATCEFCGTENLT  
KEGATTCGYLPQNAVVKIYCPACHNSEVGPESHSLAEYHNESGLKTILRKGGRTIAFGGCVFSYVGCHNKC  
AYWVPRASANIGCNHTGVVGESEGLNDNLLEILQKEKVNINIVGDFKLNEEIAIILASFSASTSAFVET  
VKGLDYKAFKQIVESCGNFVTKGKAKKGAWNIGEQKSILSPLYAFASEAARVRSIFSRTLETAQNSVR  
VLQKAAITILDGISQYSLRLIDAMMFTSDLATNNLVVMAYITGGVVQLTSQWLTNIFGTVYEKLPVLDW  
LEEKFKEGVEFLRDGWEIVKFISTCACEIVGGQIVTCAKEIKESVQTFFKLVNKFLALCADSIIIGGAKL  
KALNLGETFVTHSKGLYRKCVKSREETGLLMPLKAPKEIIFLEGETLPTEVLTEEVVLTGDLQPLEQPT  
SEAVEAPLVGTPVCINGLMLLEIKDTEKYCALAPNMMVTNNTFTLKGGAFTKVTFGDDTVIEVQGYKSVN  
ITFELDERIDKVLNEKCSAYTVELGTEVNEFACVVADAVIKTLQPVSELLTPLGIDLDEWSMATYYLFDE  
SGEFKLASHMYCSFYPPDEDEEEEGDCEEEFEPSTQYEYGTEDDYQGKPLEFGATSAALQPEEEQEEDWL  
DDDSQQTVGQQDGSSEDNQTITTIQTIIVEVQPQLEMELTPVVQTIEVNSFSGYLKLTDNVYIKNADIVEEAK  
KVKPTVVVNAANVYLKHGGGVAGALNKATNNAMQVESDDYIATNGPLKVGGSVCVLSGHNLAKHCLHVVG  
NVNKGEDIQLLSAYENFNQHEVLLAPLLSAGIFGADPIHSLRVCVDTVRTNVYLAVFDKNLYDKLVSSF  
LEMKSEKQVEQKIAEIPKEEVKPFITESKPSVEQRKQDDKIKACVEEVTTTLEETKFLTENLLLYIDIN  
GNLHPDSATLVSDIDITFLKADAPYIVGDVVQEGVLTAVVIPTKKAGGTTEMLAKALRKVPTDNYITTPY  
GQGLNGYTVEEAKTVLKKCKSAFYILPSIISNEKQEILGTVSWNLREMLAHAEETRKLMPVCVETKAIVS  
TIQRKYKGIKIQEGVVDYGARFYFYTSKTTVASLINTLNDLNETLVTMPLGYVTHGLNLEEAAARYMRSK  
VPATVSVSSPDAVTAYNGYLTSSSKTPEEHFIETISLAGSYKDWSSYSGQSTQLGIEFLKRGDKSVYYTSN  
PTTFHLDGEVITFDNLKTLTSLREVRTIKVFTTVDNINLHTQVVDMSMTYQQFGPTYLDGADVTKIKPH  
NSHEGKTFYVLPNDDTLRVEAFEYHHTDPSFLGRYMSALNHTKKWKYPQVNGLTSIKWADNNCYLATAL  
LTLQQIELKFNPPALQDAYRRARAGEAANFCALILAYCNKTVGELGDVRETMSYLFQHANLDSCKRVLNV  
VCKTCGQQQTTLKGV EAVMYMGTLSYEQFKKGVQIPCTCGKQATKYL VQQESPVMMSAPPAQYELKHGT  
FTCASEYTGNYQCGHYKHITSKETLYCIDGALLTKSSEYKGPITDVFYKENSYTTTIKPVYKLDGVVCT  
EIDPKLDNYKKDNSYFTEQPIDLVPNQYPNASFDNFKVCNDNIKFADDLNQLTGYYKPPASRELKVTF  
PDLNGDVVAIDYKHYTPSFKKGAKLLHKPIVWHVNNATNKATYKPNWTCIRCLWSTKPVETSNSFDVLKS  
EDAQGMNDLACEDLKPVSEEVVENPTIQKDVLECNVKTTEVVGDIILKPANNSLKITEEVGHTDLMAAYV  
DNSSLTIKKNELSRVLGLKTLATHGLAAVNSVPWDTIANYAKPFLNKVVSTTTNIVTRCLNRVCTNYMP  
YFFLLLQLCTFTRSTNSRIKASMPPTIAKNTVKSVMGKFCLEASFNYLKSPNFSKLINIIIFLLLSVCL  
GSLIYSTAALGVLMSNLGMPSTYCTGYREGYLNSTNVTIATYCTGSIPCSVCLSGLDSLDTYPSLETIQIT  
ISSFKWDLTAFGLVAEWFLAYILFTRFFYVLGLAAIMQLFFSYFAVHFISNSWLMWLIINLVQMAPISAM  
VRMYIFFASFYVWKSIVHVDGCNSSTCMCYKRNRATRVECTIVNGVRRSFYVYANGGKGFCKLHNW  
NCVNCDTFCAGSTFISDEVARDLSLQFKRPINPTDQSSYIVDSVTVKNGSIHLYFDKAGQKTYERHLSH  
FVNLDNLRANNTKGLPINVIVFDGKSKCEESSAKSASVYYSQLMCQPILLDQALVSDVGDSDAEVAVKM  
FDAYVNTFSSTFNVPMEKLTVAATAEAEAKNVSLDNVLSSTFISAARQGFVDSVETKDVVECKLQSHQ  
SDIEVTGDSNNYMLTYNKVENMTPRDLGACIDCSARHINAQVAKSHNIALIWNVKDFMSLSEQLRKQIR  
SAAKNNLPFKLTCATTRQVVNVVTTKIALKGGKIVNNWLKQLIKVTLVFLFVAIFYLITPVHVMKHT  
DFSSEIIGYKAIDGGVTRDIASDTCTCFANKHADFDTWFSQRGGSYTNDKACPLIAAVITREVGFVVPGLP  
GTILRTTNGDFLHFLPRVFSAVGNICYTPSKLIEYDFATSACVLAAECTIFKDASGKPVPCYDTNVLE

GSVAYESLRPDTRYVLM DGSIIQFPNTYLEGSRVVTTFDSEYCRHGT CERSEAGVCVSTSGRWV LNN DY  
YRSLPGVFCGVDAVNLLTNMFTPLIQPIGALDISASIVAGGIVAIVVTTCLAYYFMRFRRAFGEYSHV VAF  
NTLLFLMSFTVLCLTPVYSFLPGVYSVIYLYLTFYLTNDVSFLAHIQWMVMFTPLVPFWITIAIYIICIST  
KHFYWFSSNYLKRRVVFNGVSFSTFEEAALCTFLLNKEMYLKLRSDVLLPLTQYNRYLALYNKYKYFSGA  
MDTTSYREAACCHLAKALNDFNSGSDVLYQPPQTSITSAVLQSGFRKMAFPGKVEGCMVQVTCGTTTTL  
NGLWLDDVVYCPRHVICTSEDMLNPNYEDLLIRKSNHNFLVQAGNVQLRVIGHSMQNCVLKLVDTANPK  
TPKYKFVRIQPGQTFSVLACYNGSPSGVYQCAMRPNFTIKGSFLNGSCGSGVFNIDYDCVSEFCYMHMEL  
PTGVHAGTDLEGNFYGPFVDRQTAQAAGTDTTITVNVLAWLYAAVINGDRWFLNRFTTTLNDFNLVAMKY  
NYEPLTQDHDVLDILGPLSAQTGIAVLDMCASLKELLQNGMNGRILGSALLEDEFTPFVVRQCSGVTFQS  
AVKRTIKGTHHWLLL TILTSLLVLVQSTQWSLFFFLYENAF LPFAMGIIAMSAFAMMFVKHKAFLCLFL  
LPSLATVAYFNMVYMPASWVMRIMTWLDMVDTSLSGFKLKDCVMYASAVVLLILMTARTVYDDGARRVWT  
LMNVLTLVYKVYYGNALDQAISMWALIISVTSNYSGVVTTVMFLARGIVFMCVEYCPFITGNTLQCIM  
LVYCFGLGYFCTCYFGLFCLLNRYFRLTLGVYDYL VSTQEFRYMNSQGLLPPKNSIDAFKLNKLLGVGGK  
PCIKVATVQSKMSDVKCTSVVLLSVLQQLRVES SKLWAQCVQLHNDILLAKDTTEAFEKMSLLSVLLS  
MQGAVDINKLCEEMLDNRATLQAIASEFSSLPYAAFATAQEAYEQAVANGDSEVVLKLLKSLNVAKSE  
FDRDAAMQRKLEK MADQAMTQMYKQARSEDKRAKVTSAMQTMLFTMLRKLNDALNNIINNARDGCVPLN  
IPLTTAAKLMVVIPDYNTYKNTCDGTTFTYASALWEIQQVVDADSKIVQLSEISMDNSPNLAWPLIVTA  
LRANSAVKLQNNELSPVALRQMSCAAGTTQACTDDNALAYYNTTKGGRFVLALLSDLQDLKWARFPKSD  
GTGTIYTELEPPCRFVTDTPKGPVKYLYFIKGLN LNRGMVLGSLAATVRLQAGNATEVPANSTVLSFC  
AFAVDAAKAYKDY LASGGQPITNCVKMLCTHTGTGQAITVTPEANMDQESFGGASCCLYCRCHIDHPNPK  
GFCDLKGKYVQIPTTCANDPVGFTLKNVCTVCGM WKGYGCS CDQLREPMLQSADAQSFLNGFAV

ORF1a (YP\_009725295.1) protein@integer

798203128054969237687825709562533401070266615537192640433475764738580089302229375670162438980344  
681059221613835023710283700233189836924220901574581176725369543338530443347698866138914352058204  
083729051467688732393191262234555861515621911789295618931852014083597328218236654669240925998832  
162338519513353733529976499777174203341082512812652945787474730395562297390275795433313509101777  
072058729457783741881760110815022058243229789256590510032251263494625558865998362688069383271559  
397396258738638632781532793815067730293576761596042638392245868741748653525897534456419752139159  
719370138058086995053199221694984011738639583005166000387027615874152030231453975093760003077094  
396239351539879177887365224659547065010479969822168229633015915150176319175369366688858466808624  
016638708330414067002839800700974198648498370319008394974671947206689704293160290628956597824721  
222049419536567959613802712819328313012855650971698955914789394529528546253432921103957277706297  
713291809224244792176479645936824897703419405455979221212831659021804856574483838575563973317480  
806106973898622293291592893270956702623063182228266458578042445425040846160855656209300335240595  
911514539412532117745248443928709171215563484545039732148924960765110030106216574189568527593792  
024695486533164145081629621538261413907349374110050648750442723749681063916824481016582154360771  
183593415075727473237931784485227206180570821395239139504704539991752815371469433809652288253369  
234645547527723122132638017378366396102555637464229469843514843372801911132269814639592031506628  
458511942985543974970640947978504006921770241824763057457448136064081575172449004965856764394907  
164812065869983851039064768968961596733422268724917486450799587323105075895364638631060556862828  
329321107715598113320196973833926307840233953785621783908489410584477805330422337421297111090071  
799633305820553907008226140077825672991014582218541765040413036726739063742033795435705668685616  
831505522119575764256984789631551510244574159619905341575638307557220667302768303544307121555896  
134284270507809680276680044692768232503904969526647727979414682236362749746812137411965492105047  
769866186402494422446411555615556147542072597786393167818848962217658863708212980175556362088332

122584000120056185781752372395702030496660038968953636571063662502434803549810468213332273934318  
119351942362857766635998969187993362147310528598831252937062638868968090189715129705770006485593  
876575233095172657132484202930403761743289223583313329873448932641337998664702782131945027552949  
448119171138414119472843089178898265853251377031935366402705217014488879926190086854705085553498  
470457385011183279583177503088082013675618643275163578339480534953454323715754946394191690613689  
999526432040100552030634289658740781506931530850181341138074413134408096054789396363596813403792  
452423874602970203628387199977854674488440293857284919513575242792283827818589098721394294617629  
616382826388366845178083531703667371449455566801410729993811437157937234203469595957821445781991  
919257268195806688555380614281193734086493529745503167034445890498656073117447090843587403124927  
869162311791631385723805160921256203081869002929941184799613849698190014884405848396798889463586  
852399117306635449083562975441132407436113250156228600440355016010402326461525702348379949165373  
980914147982367617337293157448232360988467096637585253445987209115992181640720844304252617645485  
659369583861074168801930570743598829777429284428714237065799919544132251542119653031563870522549  
384573610033322115248054563986578290891209402041728288133829727743373251861448109728620721594533  
366759262299520756575591570011185669841159405638239066274834200537627949889643736002687899073242  
622679917295801659043521680809249117197857345599481960288126783837341835506358221159516909161515  
576577566636282933024532109177550378864069211110370057958098371311991360961199245258855434548414  
817936355983925646997215541567341084418012321598369340012335890689020700424287883033623885360140  
110486528314312056755450208607790165617441632467600225709658378774701172305907606642234312172723  
46081301877795457950396640668084724916581549204250747455899233616270103466481813417759616031478  
529523063237119547454363778280254779549762032284580874564938983772939444475248749088460323486440  
854317959484779029242459218240279678414806571880336625308227591593843851209620943854465825251123  
282460821374038169167373817224794788832259789279631523240049018360475775874757442520737507939699  
786117260048377107338816513820357637057922743851905338952173740050069887746054816135713789263194  
622442732383896668196630275579044946502815161680376225465501311672726140941165681399324129827216  
682060282111228353285833740310997521900880595419731350146942371814361523685211225201951175322515  
742909632432832826303423561967348872842790249065849588272760690818486333189000366842980213511112  
042868854132542640270744215957932422570977601980004771956980457753202759470173645267570938962657  
167029695886316022346689129217574076959174742152395923833914869034821078777944445920630892407702  
526598137549292235426936646321121001419363460694184560629644477406780592144355235365818158491052  
073793698159970610408401572913876248193668483393964464695084787359817054161789696229114331149003  
854014212394780796721373273864026816966486523240241687869215085344848722316490826763721667081007  
755771351636470897122654014489290665484369232455043076472122076250756058181608717193914426723626  
679921206048833889635780153969267312962521095449160911335463556448982980768723190474372243240180  
461490472193215009610643891423999334626377042854456362600630252729017737938112934829069393464599  
692600072488764892697246759310970743215124521777044351018276287043059069543161197975492106476540  
6669055130557501695110004595427519022090042123038810419997207846424

ORF3a

>YP\_009724391.1 ORF3a protein [Severe acute respiratory syndrome coronavirus 2]

MDLFMRIFTIGTVTLKQGEIKDATPSDFVRATATIPIQASLPFGWLIVGVALLAVFQSASKIITLKKRWQ  
LALSKGVHFCVNLNLLLVTVYSHLLLVAAGLEAPFLYLYALVYFLQSINFVRIIMRLWLCWKCRSKNPLL  
YDANYFLCWHTNCYDYCIPYNSVTSSIVITSGDGTTPISEHDYQIGGYTEKWESGVKDCVVLHSYFTSD  
YYQLYSTQLSTDTGVEHVTFIYNKIVDEPEEHVQIHTIDGSSGVNPNVMEPIYDEPTTTTTSVPL

ORF3a (YP\_009724391.1) protein@integer



440730649252419324990291953656119715554958146091212954455265484056914955855139776456842379438167  
803836832849161018449104532199800804798435447215386591632975236871201767344794167552144756617263  
529173978825932684312764194236872610752513534745877761018703658550827309107751584294503471415012  
7957711845732644249970727023860653956164936450483639373523539536817667

#### ORF6

>YP\_009724394.1 ORF6 protein [Severe acute respiratory syndrome coronavirus 2]  
MFHLVDFQVTIAEILLIIMRTFKVSIWNLDYIINLIKNSKSLTENKYSQLDEEQPMEID  
ORF6 (YP\_009724394.1) protein@integer  
17107514355168548507541923023543897063349407742162867858273516986894097507797370

#### ORF7a

>YP\_009724395.1 ORF7a protein [Severe acute respiratory syndrome coronavirus 2]  
MKIILFLALITLATCELYHYQECVRGTTVLLKEPCSSGTYEGNSPFHPLADNKFALTCFSTQFAFACPDG  
VKHVYQLRARSVSPKLFIRQEEVQELYSPIFLIVAAIVFITLCFTLKRKTE  
ORF7a (YP\_009724395.1) protein@integer  
194347074730437457513533293928686167159068966553058627192805787175963161728733970686606463547877  
67544544880647990174053532260084993372508042430109358424460913

#### ORF7b

>YP\_009725318.1 ORF7b [Severe acute respiratory syndrome coronavirus 2]  
MIELSLIDFYLCFLAFLFLVLMILIFWFSLELQDHNETCHA  
ORF7b (YP\_009725318.1) protein@integer  
63478719745535930874782340113888030785218082599390922701

#### ORF8

>YP\_009724396.1 ORF8 protein [Severe acute respiratory syndrome coronavirus 2]  
MKFLVFLGIITVAAFHQECSLQSCTQHQPYYVDDPCPIHFYSKWYIRVGARKSAPLIELCVDEAGSKSP  
IQYIDIGNYTVSCLPFTINCQEPKLGSLVVRCSFYEDFLEYHDVRRVLDLFI  
ORF8 (YP\_009724396.1) protein@integer  
194611009554485015080235775478493357293746189205074064273910773374230324575542845088267908932795  
32250367336043068643456811124221825233150723390246662221500328

#### ORF10

>YP\_009725255.1 ORF10 protein [Severe acute respiratory syndrome coronavirus 2]  
MGYINVFAFPFTIYSLLLCRMNSRNYIAQVDVNFNLT  
ORF10 (YP\_009725255.1) protein@integer  
19304754777615722453291221050421241462387470371745

#### SARS-CoV (NC\_004718.3)

##### S

>YP\_009825051.1 spike glycoprotein [SARS coronavirus Tor2]  
MFIFLLFLTSGSDLRCTTFDDVQAPNYTQHTSSMRGVYYPDEIFRSDTLYLTQDLFLPFYSNVTGFH  
TINHTEFGNPVIFPKDGIYFAATEKSNVVRGWVFGSTMNNKSQSVIIINNSTNVVIRACNFELCDNPFPAV  
SKPMGTQHTMIFDNAFNCTFEYISDAFSLDVSEKSGNFKHLREFVFNKDGFLYVYKGYQPIDVVRDLP

SGFNTLKIPIFKLPLGINITNFRAILTAFSPAQDIWGTSAAYFVGYLKPTTFMLKYDENGTTITDAVDCSQ  
NPLAELKCSVKSFEDKGIYQTSNFRVVPSPGDVVRFPNITNLCPFGEVFNATKFPSVYAWERKKISNCVA  
DYSVLYNSTFFSTFKCYGVSATKLNLDLCSNVYADSFVVKGDDVRQIAPGQTGVIADYNYKLPDDFMGCV  
LAWNTRNIDATSTGNYNKYRYLRHGKLRPFERDISNVPFSPDGKPCPPALNCYWPLNDYGFYTTTTGIG  
YQPYRVVLSFELLNAPATVCGPKLSTDLIKNQCVNFENGLTGTGVLTPSSKRFQPFQFGRDVSDFTD  
SVRDPKTSEILDISPCAFGGVSVITPGTNASSEVAVLYQDVNCTDVSTAIHADQLTPAWRIYSTGNNVFQ  
TQAGCLIGAEHVDTSECDIPIGAGICASYHTVSLLRSTSQKSIVAYTMSLGADSSIAYSNNTIAIPTNF  
SISITTEVMPVSMAKTSVDCNMYICGDSTECANLLLQYGSFCTQLNRALSGIAAEQDRNTREVFQVQKQM  
YKPTLKYFGGFNFSQILPDPLKPTKRSFIEDLLFNKVTLADAGFMKQYGECLGDINARDLICAQKFNGL  
TVLPPLLTDDMIAAYTAALVSGTATAGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKQIANQFN  
KAISQIQESLTTTSTALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLIT  
GRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGGQSKRVDFCGKGYHLMSFPQAAPHGVVFLHVTVY  
PSQERNFTTAPAICHEGKAYFPREGVVFVNGTSWFITQRNFFSPQIITDNTFVSGNCDVVIGIINNTVY  
DPLQPELDSFKEELDKYFKNHTSPDVLGDISGINASVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQ  
YIKWPWYVWLGFIAGLIAIVMVTILLCCMTSCCCLKGACSCGSCCKFDEDDSEPVLKGVKLHYT

S (YP\_009825051.1) protein@integer

459751045002885528226701443152844149038387483475285677758778632738790035612657588145099489054591  
287681959227194984221498241347839332415060527822021573191587589840981159749058543286921921061975  
175948573722983620448384450081843216484072462444436140003997903154763264297097396159109145779729  
485276510543779194540762537954023276300587563720887615320556227879914521359915362278520710368991  
868051667329224968686403550185732203765358837750985627084821351988153755050895536478459895734780  
626956549845463397666712305258623346887139611715806855994370733418047937159370310179416740307538  
757275243958601581170415557987616725102188162890313927352404797393596193656506798362977114952482  
616868465482523867986344034464628150963921508937545815040466682047715713721621227840324697695626  
297777630041199821864342678475804753248357192547333794306822437076270494661819142506335163767792  
185876954060706075956670454614678684265240481251149104651999608145845880163111859507682653813678  
357232568947154682104162034599843896128077364254667194933511270876876852157768841945030719601517  
226439162139651502405594308730155191795529592939712298205651094061254839082136704164078524220717  
767668597107065912600327715156622109448422724026607472176853125781534048192947128087896545589613  
183001952244442933361217542553447200405077042949545788510564445574308321487320097392295346624975  
526584673440615285407513780234122515226132533585195065471689561622773989143216748230195747388151  
754268630718965814601032795528326619377064655541419598174083709579375183617358537746254154588352  
095630880476686902050603487066710041891861589781173605481797976415986477506931999808736257478236  
5

E

>YP\_009825054.1 small envelope protein [SARS coronavirus Tor2]

MYSFVSEETGLIVNSVLLFLAFVVFLLVTLAILTALRLCAYCCNIVNVSLVKPTVYVYSRVKLNLSSEG  
VPDLLV

E (YP\_009825054.1) protein@integer

563171513431167720765508122543388578830711234659672386787053236772786647063568620951030851853362  
944

Influenza A virus (A/Brevig Mission/1/1918(H1N1)) (NCBI:txid88776)

H

>sp|Q9WFX3.2|HEMA\_I18A0 RecName: Full=Hemagglutinin; Contains: RecName: Full=Hemagglutinin HA1 chain; Contains: RecName: Full=Hemagglutinin HA2 chain; Flags: Precursor

MEARLLVLLCAFAATNADTICIGYHANNSTDTVDTVLEKNVTVTHSVNLLLED SHNGKLCCKLKGIAPLQLG  
KCNIAGWLLGNPECDLLLTASSWSYIVETSNSENGTCPYGFDFIDYEELREQLSSVSSFEKFEIFPKTSSW  
PNHETTKGVTAACSYAGASSFYRNLLWLTCKGSSYPKLSKSYVNNKGKEVLVLWGVHHPPTGTDQQSLYQ  
NADAYVSVGSSKYNRRFTPEIAARPKVRDQAGRMNYYWTLLEPGDTITFEATGNLIAPWYAFALNRGSGS  
GIITSDAPVHDCNTKCQTPHGA INSSLPFQNIHPVTIGECPKYVRSTKLRMATGLRNIPSIQSRGLFGAI  
AGFIEGGWTGMIDGWYGYHHQNEQSGSYAADQKSTQNAIDGITNKVNSVIEKMNTQFTAVGKEFNLER  
IENLNKKVDDGFLDIWYTNAELLVLENERLDFHDSNVRNLYEKVKSQ LKNNAKEIGNGCFEFYHKCDD  
ACMESVRNGTYDYPKYSEESKLNREEIDGVKLESMGVYQILAIYSTVASSLVLLVSLGAISFWMCSNGSL  
QCRICI

H (Q9WFX3.2) protein@integer

1769797687162803602729033994009434465558381408802707470343146032774789961066187255601899244183691809  
412473403279225888529463413854640866735761852793554636941313398480125009389665799662193160419709514  
376303101087309483108642281882915869952370088040500620476707220288564166342966773296535770621347477  
573554235502634937160490489744454953588351626142315976637899772982751031718282442950968374921106983  
90124326680228106322677855186800623672822088932643932332265974648419142179760702873206391282875956  
714213466762082277887564223738419727747302146855707352998247335844286446297373693671278727191516691  
706110644625456709935153325921254768153964871373238475701728236396244877230143973666401698651598049  
661511407878682165921803445477891044299328

N

>sp|Q9IGQ6.1|NRAM\_I18A0 RecName: Full=Neuraminidase

MNPNQKIITIGSICMVVGIISLILQIGNIISIWVSHSIQTGNQNH PETCNQSIITYENNTWVNQTYVNI  
NTNVVAGQDATSVILTGNSSLCPISGWAIYSKDN GIRIGSKGDV FVIREPFISCSHLECRTFFLTQGALL  
NDKHSNGTVKDRSPYRTLMSCPVGEAPSPYNSRFESVAWSASACHDGMGWLTIGISGPDNGAVAVLKYNG  
IITDTIKSWRNILRTQESECACVNGSCFTIMTDGPSNGQASYKILKIEKGKVT KSIELNAPNYHYEES  
CYPDTGKVMCVCRDNWHGSNRPWV SFDQNL DYQIGYICSGVFGDNPRPNDGTGSCGPVSSNGANGIKGFS  
FRYDNGVWIGRTKSTSSRSGFEMIWD PNGWTETDSSFSVRQDIVAITDWSGYSGSFVQHPELTGLDCMRP  
CFWVELIRGQPKENTIWTSGSSISFCGVNSDTV GWSWPDGAELPFSIDK

N (Q9IGQ6.1) protein@integer

110196333912281467256866753169858814391628996857105456326272748008981501392647241814172877654448  
981702101432594278955083720319073200040795200007449524171552454905156852947491905757413992106686  
575111796911474915204511930770177314662797405761562921555949884027511565481309244015284853410740  
809255241474669738911597351305613982861217330434812994838850366544821333029998376698696477698206  
259293597886076159397066052589355300568366300034679244257399960844808170533876421585713900041050  
466396311880867768155248136739412187984149968489854224439506900934113302451033439604018426128403  
56224233806301788811980842700179412

Influenza A virus (A/California/07/2009(H1N1)) (NCBI:txid641809)

H

>AMV49023.1 hemagglutinin [Influenza A virus (A/California/07/2009(H1N1))]

MKAILVLLYTFATANADTLCIGYHANNSTDTVDTVLEKNVTVTHSVNLLLEDKHNGKLCCKLRGVAPLHLG  
KCNIAGWILGNPECESLSTASSWSYIVETPSSDNGT CYPGFDFIDYEELREQLSSVSSFERFEIFPKTSSW  
PNHDSNKGVTAACPHAGAKSFYKNLIWL VCKGNSYPKLSKSYINDKGKEVLVLWGIHHPSTSADQQSLYQ

NADAYVFGSSRYSKKFKPEIAIRPKVRDQEGRMNYWTLVEPGDKITFEATGNLVVPRYAFAMERNAGS  
GIIISDTPVHDCNTTCQTPKGAINSLPFQNIHPITIGKCPKYVKSTKLRLATGLRNIPSIQSRGLFGAI  
AGFIEGGWTGMVDGWYGYHHQNEQGSYAADLKSTQNAIDEITNKVNSVIEKMNTQFTAVGKEFNHLEKR  
IENLNKKVDDGFLDIWTYNAELLVLENERTLDYHDSNVKNLYEKVRSQKNNAKEIGNGCFEFYHKCDN  
TCMESVKNGTDYDPKYSEEAKLNREEIDGVKLESTRIYQILAIYSTVASSLVLVVSLGAISFWMCSNGSL  
QCRICI

H (AMV49023.1) protein@integer

176362337141980046838142812005942123004517025437229818934225506295086730065270409351942471488730  
650674402713706411075160742664578112503357239864503238269620158959601166331925444626845719977974  
212455280804657866292852581097176969563548891416026146007787538132311109374290182457012165566124  
595291741536854470673643165886822768110639595212304690698880271673517138771621786901741373697742  
421500830773224512609363658536728387331127106809944677456624787675522893726593861614211501417427  
883434016385310646457431149114586810589690720357921210291072482792472144699362811083408342634385  
775075192545516205376392248592604661122775188510877933261674042669103257169080890897799791114707  
40180872799908646668030333511407878556336801803445477891044299328

N

>AMV49025.1 neuraminidase [Influenza A virus (A/California/07/2009(H1N1))]

MNPNQKIITIGSVCMTIGMANLILQIGNIISIWISHSIQLGNQNQIETCNQSVITYENNTWVNQTYVNI  
NTNFAAGQSVSVKLAGNSSLCPVSGWAIYSKDNSVRIGSKGDVVFVIREPFISCSPLECRTFFLTQGALL  
NDKHSNGTIKDRSPYRMLMSPIGEVSPYNSRFESVAWSASACHDGINWLTIGISGPDNGAVAVLKYNG  
IITDTIKSWRNILRTQESECACVNGSCFTVMTDGPSNGQASYKIFRIEKGKIVKSVEMNAPNYHYEECS  
CYPDSSEITCVCRDNWHGSRNPWVSNFNQLEYQIGYICSGIFGDNPRPNDKTGSCGPVSSNGANGVKGFS  
FKYGNVWIGRTKSISSRNGFEMIWDPNGWTGTDNNSFIKQDIVGINEWSGYSGSFVQHPELTSGLDCIRP  
CFWVELIRGRPKENTIWTSGSSISFCGVNSDVTGWSWPDGAELPFTIDK

N (AMV49025.1) protein@integer

110196333912281459814317487227312566026606235506892502698128666180874047476749058599041037813883  
084736327809493635218753099276479598530404179742254021202784311333940781301838584784376250448848  
064928900768203084657325218459265024062705330642710392638739822785978722355133312698371899206677  
284945131770072551403495496485470301555054263065596711970210584348094334781970441070739703958782  
211164229634219822599148981475971225905064634605697673360682054551507927602754112981573959015931  
834200981137532170841694491412521329644207474174523646023890900799005313629918592589367259408403  
56224233806301788811980842700203412

*S. aureus*

ClfA

>RYV20394.1 MSCRAMM family adhesin clumping factor ClfA [Staphylococcus aureus]

MNMKKKEKHAIRKKSIGVASVLVGLTIGFGLSSKEADASENSVTQSDSASNESKSNDSVVNAAPKTDN  
TNVSDTKTTSNTNGETSVAQYLAQQETTQSVSTNATTEETPVTGEATTATNQANTPATTQSSNTNAEEL  
VNQTSNETTSNDTNTVSSVNSPQNSTNAENVSTTQDTSTEATPSNNESAPQSTDASNKDVVNQAVNTSAP  
RMRAFSLAAVAADAPAAGTDITNQLTNVTVGIDSGTTVYPHQAGYVKLNYGFSVPNSAVKGDTFKITVPK  
ELNLNGVTSTAKVPPIMAGDQVLANGVIDNDGNVIYFTFDYVNTKDDVKATLTMPAYIDPENVTKTGNVT  
LATGIGSTTANKTVLVDEKYGKFYNLSIKGTIDQIDKTNNYRQTIYVNPNGDNVIAPVLTGNLKPNTD  
SNALIDQQNTSIKVYKVDNAADLSESYFVNPENFEDVTVNSVNITFPNPQYKVEFPTDDDQITTPYIVVV  
NGHIDPNSKGDALRSTLYGYSNIIWRMSWDNEVAFNNGSGSGDGIDKPVVPEQPDEPGEIEPIPEDS

DSDPGSDSGSDSGSDSNSDSGSDSGSDSTSDSDSDSDSDSDSASDSDSASDSDSASDSTSDSDSDSASDS  
DSASDSDSASDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSASDSDSDSDSDSDSDSDSDS  
DSESDSDSDSVSDSDSDSDSDSDSNSESDSDSDSDSDSDSASDSDSGSDSDSDSDSDSDSDSDSDSDSDS  
TSDSDSDSTSDTGSNDSDSESDSNSDSDSGSNNNVVPPNSPKNGTNASNKNEAKDSKEPLPDTGSEGEE  
NTSLIWGLLASLGSLLLFRRKKENKDKK

ClfA (RYV20394.1) protein@integer

142551594247385244852272037694088919547548512006016242151854647688412116338759186626186070977958  
689198371715962257172153602408233878034594540795713744790617430951534645704391033360102622439933  
47235753617074905943848372060354638213164996534342222743533927073994852461690809110673683427143  
437594379463063119207020929822088927604946025366307277842742694706953432828479491164951807835689  
639774786992604577881129892883940476060565422277784330351465181924007593357298961328007678988503  
680792759389583477905851488215806178203047166987587193850668619183239331250066272963644319828349  
854084999375947616341189016807294176421624061351313532489684470593057847169780249455376276868047  
989827633894342758536698959705674973883109836645265160047849825011472396975095394107271532990362  
108710984406632371587139125670572825871188651858788119925913297863530898972062013353864641883716  
977997233017230797279001121617339891460158722432498750080260453317328187405686935298616673859834  
652634857078882013906538893303504731788294056682494124245015764821386707291352024441718009634128  
75476017818400133480265299591101191175604544901727156309800518595371140252

SPA

>sp|P02976.3|SPA\_STAA8 RecName: Full=Immunoglobulin G-binding protein A; Short=IgG-binding protein A;  
AltName: Full=Staphylococcal protein A; Short=SpA; Flags: Precursor

MKKKNIYSIRKLGVGIASVTLGTLISGGVTPAANAAQHDEAQQNAFYQVLNMPNLNADQRNGFIQSLKD  
DPSQSANVLGEAQKLNDSPQPKADAQQNNFNKDQQSAFYEILNMPNLNEAQRNGFIQSLKDDPSQSTNVL  
GEAKKLNESQAPKADNNFNKEQQNAFYEILNMPNLNEEQRNGFIQSLKDDPSQSANLLSEAKKLNESQAP  
KADNKFNKEQQNAFYEILHLPNLNEEQRNGFIQSLKDDPSQSANLLAEAKKLNDAPKADNKFNKEQQN  
AFYEILHLPNLTEEQRNGFIQSLKDDPSVSKEILAEAKKLNDAPKKEEDNNKPGKEDNNKPGKEDNNK  
GKEDNNKPGKEDNNKPGKEDGNKPGKEDNKKPGKEDGNKPGKEDNKKPGKEDGNKPGKEDGNKPGKEDGN  
GVHVVKPGDVTNDIAKANGTTADKIAADNKLADKNMIKPGQELVVDKKQPANHADANKAALPETGEENP  
FIGTTVFGGLSLALGAALLAGRREL

SPA (P02976.3) protein@integer

156941706585774372868196694244164565713212393772694157852978879536875219910831286911720511868904  
219973985478767911074284050514925647124376034986319672097812025477859165816285495225437881427364  
883789886448652227806856960295688260852534361223494577629662412980239379060862720173456714211563  
656594152772907866005753174656422355945367157197485241062150933129699119446046363176754739526527  
179650519227645204589005857633569221419869720894139205611351098970568862963075676464874567209539  
547240132086987970977420432411277692380920534132718104164247127236117834478985549871326046141194  
808514632749183271338363667613850271842143824329531320889665185590238752281144635402940226863067

*S. pneumoniae*

SpsA

>VPQ17379.1 pspA [Streptococcus pneumoniae]

MNKKMILTSLASVAILGAGFVTSQPTFVRAEEAPVASQSKAEKDYDTAKRDAENAKKALEEAKRAQKKY  
EDDQKKTEEKAKEEKQASEAEQKANLQYQLKLREYIQTGDRSKIQKEMEEAEKHKHNAKAEFDKVRGKV

IPSAEELKETRRKAEEAKAKEAELTKKVEEAEEKKVTEAKQKLD AERAKEVALQAKIAELENQVHRLETEL  
KEIDESDSEDIKEGLRAPLQSKLDAKKAKLSKLEELSDKIDELDAEIAKLEKDVDFKNSDGEQAEQYL  
VAAKKDLDAKKAELENTEADLKKAVDEPETPAPAPKPAPAPAPTPTPEAPAPAPKPAPAPKPAPAPAPT  
EAPAPAPKPAPAPKPAPAPKPAPAPKPETPKTGWKQENGMWYFYNTDGSMATGWLQNNGSWYYLNANGAM  
ATGWAKVNGLWYYLNSNGAMATGWLQYNDWSYYLNTNGAMATGWLEYNGSWYYLNANGAMATGWLEYNG  
S  
WYYLNSNGAMATGWLQYNGSWYYLNANGAMATGWAKVNGSWYYLNANGSMATGWVKDGD TWYYLEASGA  
M  
KANQWFKVSDKWYYVNGLGALAVNTTVDGYTVNENGEWV

SpsA (VPQ17379.1) protein@integer

150228849051900023040605651647315364361455349107216843854124495314359445746799551397367594448928467  
1839405448635574381390458226711579831077838801740030972187835724975074973209684942582714855710871015  
19036080947868062982698144465434882768827054393001222401869052269838020436091526600339702752786204  
5964920312942518111524815306582235597274250629139827419175619435992435893782663922914227465119272040  
3139734988854688887746360333444199440927878976788669205962190767957522171451737483028011448826379403  
52918143230071518745194819133677280719137055143038454088158726426323312463354703768966169627914168575  
281590359936401986815541361920153562805140782229457746643108124482971693553500553197940536189786283  
37563064680336054331203200277406160314315919026576981978258357920172041349739045584

SpsA

>CNZ85632.1 choline binding protein A [Streptococcus pneumoniae]

MFASKSERKVHYSIRKFSIGVASVVVASLFLGGVVHAEGVRSKNNPTVTSSGQDISKKYADEVKSHLEKI  
LREIQLD RRRKHTQNADFSKLSRIQTEYFYELSVLKEKSEAELMSKTKEELTSKTKEKLTAAFEQFKKDT  
LKPGEKVAEAQKKVAEAKKKAEDQKEEDRRNYPTNTYKTLLEIEAFDVKVKEAELELLKEKAKESQDDE  
KIKQAEAKVESKKA EATRLENIKTDRKKA EEEAKRKADAKLKEANVATSDQGKPKGRAKRGVPGELATPD  
KKENDAKSSDSSVGEETLPSSSLKSGKVAEAEKKVEEA EKKAKDQKEEDRRNYPTNTYKTLDL EIAESD  
VKVKEAELELVKEEAKEPRDEEKIKQAKAKVESKKA EATRLEKIKTDRKKA EEA KRKAAEEDKVKEKPAE  
QPQPAPAPQPEKPAPKPEKPAEQPKAEKPADQQA EEDYARRSEEEYNRLTQQQPPKTEKPAQPSTPKTGW  
KQENGMWYFYNTDGSMATGWLQNNGSWYYLNANGAMATGWLQNNGSWYYLNANGAMATGWLQYNGSWY  
YL  
NSNGAMATGWLQYNGSWYYLNANGDMATGWLQNNGSWYYLNSNGAMVTGWLQNNGSWYYLNANGDMAT  
GW  
VKDGD TWYYLEASGAMKASQWFKVSDKWYYVNSNGAMATGWLQNNSSWYYLNANGAMATGWAKVNGSWY  
Y  
LNANGSMATGWVKDGD TWYYLEASGAMKASQWFKVSDKWYYVNGLGSLAVNTTVDGYTVNENGEWV

SpsA (CNZ85632.1) protein@integer

287270686708774391510075610717792133413879492418991875325929780956206826552255833422803381008363158  
0142073496428234659500953123937481170010140388604566193438153865270258147997303346757014875316251714  
659024623964883640822688864006068766742823159937952751201068371004740202470204932816518334840156544  
511484731943830775962226062039812284199994836818338653162803120472683190258290456488079611811672926  
8523880144239259474837101077590551618201573050790028123666941588296148497070015876502825467911392109  
107177882192637119903881914205712999857772318196576286123839043825529580162946659077831537302134564  
0591817219244801647430040158248881622639869875124405751272520426137589468230040381772402494618076541  
5983455609506827722448590650087916399505909236089156222955609532796866828929605013772346755272270

9394476033158792927935609397778682274170104884218413779176444327742553632351771249844687869949045582  
623274509566697862833756306468033605433120320027740615983327958187457698197826360080017204134973904  
5584

#### CD4

>NP\_001181943.1 T-cell surface glycoprotein CD4 isoform 2 [Homo sapiens]

MPTPLVHPHLPISSPRVSPFPPPAFQKASSIVYKKEGEQVEFSPLAFTVEKLTGSGELWWQAERASSK  
SWITFDLKNKEVSVKRVTDPKLQMGKKLPLHLTLPQALPQYAGSGNLTALAEAKTGKLGHEVNLVVMRA  
TQLQKNLTCEVWGPTSPKMLSLKLENKEAKVSKREKAVWVLNPEAGMWQCLLSDSGQVLLSNIKVLPT  
WSTPVQPMALIVLGGVAGLLLFIGLGIFFCVRCRHRRRQAERMSQIKRLLSEKKTCCQCPHRFQKTCSPI

CD4 (NP\_001181943.1) protein@integer

687846614202831577808010159849001106156983221397499915565083649090615553463228685512583913799174032  
740995574679739791288737164428022478312155532518457834650504473629260028882370154173180278306459233  
11375895403817013087714718631660543090222741845704659026078951647466568757913782528655334802997009  
908245911502378364493486659166007641154241870307865843656423248868

#### CD22

>sp|P20273.2|CD22\_HUMAN RecName: Full=B-cell receptor CD22; AltName: Full=B-lymphocyte cell adhesion  
molecule; Short=BL-CAM; AltName: Full=Sialic acid-binding Ig-like lectin 2; Short=Siglec-2; AltName: Full=T-cell surface  
antigen Leu-14; AltName: CD\_antigen=CD22; Flags: Precursor

MHLLGPWLLLLVLEYLAFSDSSKVVFEHPETLYAWEGACVWIPCTYRALDGDLESFILFHNPEYNKNTSK  
FDGTRLYESTKDGKVPSEQKRVQFLGDKNKNCTLSIHPVHLNDSGQLGLRMESKTEKWMERIHLNVSERP  
FPPHIQLPPEIQESQEVTLTCLLNFSCYGYPIQLQWLLEGVPMRQAAVTSTSLTIKSVFTRSELKFSPQW  
SHHGKIVTCQLQDADGKFLSNDTVQLNVKHTPKLEIKVTPSDAIVREGDSVTMTCEVSSSNPEYTTVSWL  
KDGTSLKKQNTFTLNLREVTKDQSGKYCCQVSNVGVGPRSEEVFLQVQYAPEPSTVQILHSPAVEGSQVE  
FLCMLANPLPTNYTWYHNGKEMQGRTEEKVHIPKILPWHAGTYSCVAENILGTGQRGPGAELDVQYPPK  
KVTTVIQNPMPIREGDTVTLSCNYNSSNPSVTRYEWKPHGAWEEPSLGVLKIQNVGWDNTTIACAACNSW  
CSWASPVALNVQYAPRDVVRKIKPLSEIHSGNSVSLQCDFSSHPKEVQFFWEKNGRLLGKESQLNFDS  
ISPEDAGSYSCWVNNSIGQTASKAWTLEVLYAPRRLRVSMSPGDQVMEGKSATLTCESDANPPVSHYTW  
DWNQSLPYHSQKLRLEPVKVQHSGAYWCQGTNSVGKGRSPLSTLTVYYPETIGRRVAVGLGSCLAILI  
LAICGLKLQRRWKRTQSQQGLQENSSGQSFVVRNKKVRRAPLSEGPLSLGCYNPMMEDGISYTTLRFPEM  
NIPRTGDAESSEMQRPPDCDDTVTYSALHKRQVGDYENVIPDFPEDEGIHYSELIQFGVGERPQAQENV  
DYVILKH

CD22 (P20273.2) protein@integer

692961457230670026657062948252478776701300424999986876932293167596623453192699758336590075958404911  
629513797785596208968340815608160306086764795817851000016218501585718082099455083165188585050791320  
047706172760478343960274893452761545097326755788101471259834704808381526788503309135055696422363185  
311023398707702073946872541611793928228200382135564343430321083661081450392972565360126672122468713  
676303491797393563679981463001212838568960099796623077495631381089457716905644547088777918375062559  
918940149017668643473927974740560125663843033603246301975806980727701296256483671072858180538885064  
15057922143349730216099280750061681888928459846798666772840241099045143422929478683765246518806132  
16603528035463255999864296992325192994549103789443721579315140182296013498016630048065986072085872  
8594145294429579697449207211421272861248748423597485430135204041896524232941731094861095489568478561  
5621744791343135422359988341869773678012077323782101587387443372235955214257015047092618096127686840  
74190891955563052122273743697543267954979588973509566003758903919364735284986595528899826856292794  
0332218307055

SIg-JC

>NP\_653247.1 immunoglobulin J chain precursor [Homo sapiens]

MKNHLLFWGVLAVFIKAVHVKAQEDERIVLVDNKCKCARITSRIIRSEDPNEDIVERNIRIIVPLNNRE  
NISDPTSPLRTRFVYHLSDLCKKCDPTEVELDNQIVTATQSNICDEDSATETCYTYDRNKCYTAVVPLVY  
GGETKMOVETALTPDACYPD

SIg-JC (NP\_653247.1) protein@integer

534340388306364710653764801926507336803883698971891575359281672249831540827372391836159231652200  
406042338674631107179131806415860547374157012037929234735324688960946972926249677445650194944346  
778297824215270

SIg-SC

>pdb|5D4K|B Chain B, Polymeric immunoglobulin receptor

KSPIFGPEEVNSVEGNSVSITCYPPPTSVNRHTRKYWCRQGARGGCITLISSEGYVSSKYAGRANLTNFP  
ENGTFFVNNIAQLSQDDSGRYKCGLGINSRGLSFDVSLEVSQGPGLLNDTKVYTVDLGRTVTINCPFKTEN  
AQKRKSLYKQIGLYPVLVIDSSGYVNPNYTGRIRLDIQGTGQLLFSVVINQLRLSDAGQYLCQAGDDNS  
NKKNADLQVLKPEPELVYEDLRGSVTFHCALGPEVANVAKFLCRQSSGENCDVVVNTLTKRAPAFEGRIL  
LNPQDKDGSFVITGLRKEDAGRYLCAHSDGQLQEGSPIQAWQLFVNEESTIPRSPTVVKGVAGSSVA  
VLCPNRKEKSIKYWCLWEGAQNGRCPLLVDSEGWVKAQYEGRLSLLEEPGNGTFTVILNQLTSRDAGF  
YWCLTNGDTLWRTTVEIKIIEGEPNLKVPGNVTAVLGETLKVPCHPCKFSSYEKYWCKWNNTGCCALPS  
QDEGPSKAFVNCDENSRLVSLTLNLVTRADEGWYWCGVKQGHFYGETAAVYVAVEERGSWSHPQFEK

SIg-SC (5D4K\_B) protein@integer

285608760214857610208583182830139398261811843013444667420485365201411381191110595975577949608647  
521139318801565858202459108790630469167646158214350898329812295793064029491668299145642261254053  
386833590982136265713414449507351113955726470659961126918537785555211701800292132360419410163767  
577084064463328543040095036695100451587552936519622535372938177233047086753074419815220057928933  
406593228880373928223452394715965932384729534353760295796835291297128740582294268850369007500817  
436012621149233554127083665325782986559489670488514931395371121670641504567703250805888027233578



662438874015299820683985873973756077527191157414975799430711093205766750280330251729951638829730  
64111768560018670939611255543233382374164555696574672

Human GlcNAcase a

>NP\_036347.1 protein O-GlcNAcase isoform a [Homo sapiens]

MVQKESQATLEERESELSSNPAASAGASLEPPAAPAGEDNPAGAGGA AVAGAAGGARRFLCGVVEGFYGRPWVMEQRKELFRRLQKWELNTYLYAPKDDYKHRMFWREMYSV EEAQLM TLISAAREYEIEFIYAISPGLDITFSNPKEVSTLKRKLDQVSQFGCRSFALLFDDIDHNMCAADKEVFSSFAHAQVSITNEIYQYLGEPE TFLFCPT EYCGTFCYPNVSQSPYLRTVGEKLLPGIEVLWTGPKVVSKEIPVESIEEVSKIIKRAPVIWDN IHANDYDQKRLFLGYPYKGRSTELIPRLKGVLTNPNCEFEANYVAIH TLATWYKSNMNGVRKDVVMTDSED STVSIQIKLENEGSDEDIETDVL YSPQMALKLALTEWLQFEFGVPHQYSSRQVAHSGAKASVVDGTPLVAA PSLNATTVVTTVYQEPIMSQGAALS GEPTTLTKEEEKKQPDEEPM DMVVEKQEETHDKNDNQILSEIVEA KMAEELKPMDDKESIAESKSP ESMQEDCISDIAPMQTDEQTNKEQFVPGPNEKPLYTAEPVTLEDLQL LADLFYLPYEHGPKGAQMLREFQWLRANSSVSVNCKGKDSEKIEE WRSRAAKFEEMCGLVMGMFTRLSN CANRTILYDMYSYVWDIKS IMSMVKSFVQWLGCRSHSSAQFLIGDQEPWAFRGLAGEFQRLLPIDGAND LFFQPPPLTPTSKVYTIRPYFPKDEASVYKICREMYDDGVGLPFQSQPDLIGDKLVGGLLS LSLDYCFVL EDEDGICGYALGTVDPFPIKCKISWIPFMQEKYTKPNGDKELSEAEKIMLSFH EEQEVLPE TFLANFP SLIKMDIHKKVTDP SVAKSMMA CLSSLKANGSRGAFCEVRPDDKRILEFY SKLGC FEIAKMEGF PKD VV ILGRSL

GlcNAcase a (NP\_036347.1) protein@integer

3941149562093479901160527872139939869953948328717448309602075904617639852765945916537772170887588913  
777760741172225419058993174618449700819722020154518702683811512865205056993341462772546223338698372  
1402964375598157546637116156924279597212332019651564117548382671604539961725503706897735714899984118  
06014935596262763747945043157656864735710326130352938246306646449830806350943992779426250490629331  
05717204828412582988248168711408930001109074970612171397190589610378189764664617608879750771893149325  
703234667852859231589215026272173420215399339907210129333224063854880356386375621030134935672120130  
1182158461415936423120637434347353003426842559697796327071647986461830586067918017754414457439239556  
105794613393123607246824115000032711726489102914303601308009259048205472279983560823820077297582358  
179664677858758610884688114892380227174055044911625343181269465663312612797038045228606892503534909  
04053200471496813891995342035335150071662609939061752070884770494552358391253502368356686068894419  
247600561381489785637526973435030429756393323513367039788577539639326487769367168014443728381308829  
742027488935828240337784123405635224582039088107227197931414720130267879239425482181873925402726847

Human GlcNAcase b

>NP\_001135906.1 protein O-GlcNAcase isoform b [Homo sapiens]

MVQKESQATLEERESELSSNPAASAGASLEPPAAPAGEDNPAGAGGA AVAGAAGGARRFLCGVVEGFYGRPWVMEQRKELFRRLQKWELNTYLYAPKDDYKHRMFWREMYSV EEAQLM TLISAAREYEIEFIYAISPGLDITFSNPKEVSTLKRKLDQVSQFGCRSFALLFDDIDHNMCAADKEVFSSFAHAQVSITNEIYQYLGEPE TFLFCPT EYCGTFCYPNVSQSPYLRTVGEKLLPGIEVLWTGPKVVSKEIPVESIEEVSKIIKRAPVIWDN IHANDYDQKRLFLGYPYKGRSTELIPRLKGVLTNPNCEFEANYVAIH TLATWYKSNMNGVRKDVVMSRQVA HSGAKASVVDGTPLVAA PSLNATTVVTTVYQEPIMSQGAALS GEPTTLTKEEEKKQPDEEPM DMVVEKQE ETDHKNNDNQILSEIVEAKMAEELKPMDDKESIAESKSP ESMQEDCISDIAPMQTDEQTNKEQFVPGPN EKPLYTAEPVTLEDLQLLADLFYLPYEHGPKGAQMLREFQWLRANSSVSVNCKGKDSEKIEE WRSRAAK FEEMCGLVMGMFTRLSN CANRTILYDMYSYVWDIKS IMSMVKSFVQWLGCRSHSSAQFLIGDQEPWAFRGLAGEFQRLLPIDGANDLFFQPPPLTPTSKVYTIRPYFPKDEASVYKICREMYDDGVGLPFQSQPDLIGD

KLVGGLLSLSDYCFVLEDEDGICGYALGTVDVTPFIKKCKISWIPFMQEKYTKPNGDKELSEAEKIMLS  
FHEEQEVLPETFLANFPSLIKMDIHKKVTDPVSAKSMACLLSSLKANGSRGAFCEVRPDDKRILEFYSK  
LGCFEIAKMEGFPKDVVILGRSL

GlcNAcase b (NP\_001135906.1) protein@integer

4375554987327534452080133178397613880010684016227893510910891166762626106352336204416722166152352484  
2449370155475191678305916436884592469546941839902441609626361787982690217938920821141908471905932889  
334574428863229028267802362619973654790631825101284060058713118080369049081646405030292047977356530  
770193172137652885820219331900090403415182567370803543845066581552217595375502404463825823478095123  
971933820033048767464391656372245997820392415473957484090758654655863667348075629108767065720867143  
071403822400027893319969468905246284007908415772092132766524888754999267580019678891720588021632129  
5421682759180177544144574392395561057946133931236072468241150000327117264891029143036013080092590482  
054722799835608238200772975823581796646778587586108846881148923802271740550449116253431812694656633  
12612797038045228606892503534909040532004714968138919953420353351500716626099390617520708847704945  
52358391253502368356686068894419247600561381489785637526973435030429756393323513367039788577539639  
3264877693671680144437283813088297420274889358282403377841234056352245820390881072271979314147201302  
67879239425482181873925402726847

Human HexNAcase  $\alpha$

>NP\_001305754.1 beta-hexosaminidase subunit alpha isoform 1 precursor [Homo sapiens]

MTSSRLWFSLLLAAAFAGRATALWPWPQNFQTSQRYVLYPNNFQFQYDVSSAAQPGCSVLDEAFQRYRD  
LLFGSGSWPRPYLTGWPHQAYPVFLGKRHTLEKNVLVSVVTPGCNQLPTLESVENYTLTINDDQCLLS  
ETVW GALRGLETFSQLVWKS AEGTFFINKTEIEDFPRFPHRGLLLDTSRHYP LSSILD TLDVMAYNKNL N  
VFHWHLVDDPSFPYESFTPELMRKGSYNPVTHIYTAQDVKEVIEYARLRGIRVLA EFDTPGHTLSWGP G  
IPGLLTPCYSGSEPSGTFGPVNP SLNNTYEFMSTFFLEVSSVFPDFYLHLGGDEVDFTCWKSNPEIQDFM  
RKKGFGE DFKQLESFYIQTLLDIVSSYGKGYVWVQEVFDNKVKIQPDTIIQVWREDIPVNYMKELELVTK  
AGFRALLSAPWYLNRSYGPDWKDFYIVEPLAFEGTPEQKALVIGGEACMWGEYVDNTNLVPRLWPRAGA  
VAERLWSNKL TSDLT FAYERLSHFRCELLRRGVQAQPLNVGFCEQEFEQT

HexNAcase  $\alpha$  (NP\_001305754.1) protein@integer

2565345204515766711480306911006177947198821532684813532447810423646865764573107554662772808381526600  
642117917154057226835486085540609036627326462010905381722827152902018450244952460665787127848557329  
395690420680769261634047149495299639641226982165094725020646629229474509472842655874683072561673703  
041028821950141244569110159970542189743393425436133850860020642423772374665885844339307607242289110  
53322711285023631702493108185183827081831884917104527304308889323868957612012185652176729187617388637  
972721110986911289289797936885194513457758542882302568596633212217798783484552402181106893688337939  
0864129453307575172981681286669387475194084804497616538662258630500090954851754479511848826716578149  
413425

Human HexNAcase  $\beta$

>sp|P07686.3|HEXB\_HUMAN RecName: Full=Beta-hexosaminidase subunit beta; AltName: Full=Beta-N-  
acetylhexosaminidase subunit beta; Short=Hexosaminidase subunit B; AltName: Full=Cervical cancer proto-oncogene 7  
protein; Short=HCC-7; AltName: Full=N-acetyl-beta-glucosaminidase subunit beta; Contains: RecName: Full=Beta-  
hexosaminidase subunit beta chain B; Contains: RecName: Full=Beta-hexosaminidase subunit beta chain A; Flags:  
Precursor

MELCGLGLPRPPMLLALLLATLLAAMLALLTQVALVVQVAEAAARAPSVSAKPGPALWPLPLSVKMTPNLL  
HLAPENFYISHSPNSTAGPSC TLL EAFRRYHGYIFGFYKWHHEPAEFQAKTQVQQLLV SITLQSECD AF

PNISSDESYTLLVKEPVAVLKANRVW GALRGLTFSQLVYQDSYGTFTINESTIIDS PRFSSHGILIDTS  
RHYLPVKIILKTL DAMAFNKFNVLHWHIVDDQSPYQSITPELSNKGSYLSHVYTPNDVRMVIEYARL  
RGIRVLPEFDTPGHTLSWGKGQKDLLTPCYSRQNKLDSFGPINPTLNTTYSFLTTFKKEISEVFPDQFIH  
LGGDEVEFKCWESNPKIQDFMRQKGGFTDFKLESFYIQKVLDIATINKGSIVWQEVFDDKAKLAPGTI  
VEVWKDSAYPEELSRVTASGFPVILSAPWYLDLISYGQDWRKYYYKVEPLDFGGTQKQKQLFIGGEACLWG  
EYVDATNLTPRLWPRASAVGERLWSSKDVRDMDDAYDRLTRHRCRMVERGIAAQPLYAGYCNHENM  
HexNAcase  $\beta$  (P07686.3) protein@integer

17299197692074680621697954965867057756855656305894249295852967670385768790849930607954400754469209  
4993310878154614525498879223813549001532872627888979778546253278344887769315106755115166581387496703  
867005445080043051604580041003354372585648597096644308577458390303472599411303874044310161459794305  
350132925167052700990365647705266288463161406649682481589793289802622986356639421417858654610613287  
342959339086659692526024189943642369145506371129106857266351112035308598106219460224555090955722541  
1851816066991752507131424936031594793807515502910897524376809113505995652198694804260215575138256880  
1601576729129151835697809537156753771636462818975800183987475805561661786249489263227199191446493005  
09116032878310071322772765394

SARS-CoV (NC\_004718.3)

S

>YP\_009825051.1 spike glycoprotein [SARS coronavirus Tor2]

MFIFLLFLTSGSDLRCTTFDDVQAPNYTQHTSSMRGVYYPDEIFRSDTLYLTQDLFLPFYSNVTGFH  
TINHTFGNPVIPFKDGIYFAATEKSNVVRGWVFGSTMNNKSQSVIIINNSTNVVIRACNFELCDNPFPAV  
SKPMGTQHTMIFDNAFNCTFEYISDAFSLDVSEKSGNFKHLREFVFNKNDGFLYVYKGYQPIDVVRDLP  
SGFNTLKPFIKPLGINITNFRAILTAFAAQAQDIWGTSAAYFVGYLKPPTFMLKYDENGITDAVDCSQ  
NPLAELKCSVKSFEIDKGIYQTSNFRVPSGDVVRFPNITNLCPFGEVFNATKFPVYAWERKKISNCVA  
DYSVLYNSTFFSTFKCYGVSATKLNLDLCSNVYADSFVVKGDDVRQIAPGQTGVIADYNYKLPDDFMGCV  
LAWNTRNIDATSTGNYNKYRYLRHGKLRPFERDISNVPFSPDGKPCPPALNCYWPLNDYGFYTTTIGIG  
YQPYRVVLSFELLNAPATVCGPKLSTDLIKNQCVNFENGLTGTGVLTPSSKRFQPFQFGRDVSDFTD  
SVRDPKTSEILDISPCAFGGVSVITPGTNASSEVAVLYQDVNCTDVSTAIHADQLTPAWRIYSTGNNVFQ  
TQAGCLIGAEHVDTSYECDIPIGAGICASYHTVSLRSTSQKSIVAYTMSLGADSSIAYSNNTIAIPTNF  
SISITTEVMPVSMAKTSVDCNMYICGDSTECANLLQYGSFCTQLNRALSIAAEQDRNTREVFQVQKQM  
YKTPTLKYFGGFNFSQILPDLKPTKRSFIEDLLFNKVTLADAGFMKQYGECLGDINARDLICAQKFNGL  
TVLPPLLTDDMIAAYTAALVSGTATAGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKQIANQFN  
KAISQIQESLTTTSTALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLIT  
GRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLQSKRVDFCGKGYHLSFPQAAPHGVVFLHVTVY  
PSQERNFTTAPAICHEGKAYFPREGVVFVNGTSWFITQRNFFSPQIITDNTFVSGNCDVVIGIINNTVY  
DPLQPELDSFKEELDKYFKNHTSPDVLGDISGINASVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQ  
YIKWPWYVWLGFIAGLIAIVMTILLCCMTSCCCLKGACSCGSCCKFDEDDSEPVLKGVKLHYT

S (YP\_009825051.1) protein@integer

459751045002885528226701443152844149038387483475285677758778632738790035612657588145099489054591  
287681959227194984221498241347839332415060527822021573191587589840981159749058543286921921061975  
175948573722983620448384450081843216484072462444436140003997903154763264297097396159109145779729  
485276510543779194540762537954023276300587563720887615320556227879914521359915362278520710368991  
868051667329224968686403550185732203765358837750985627084821351988153755050895536478459895734780  
626956549845463397666712305258623346887139611715806855994370733418047937159370310179416740307538

757275243958601581170415557987616725102188162890313927352404797393596193656506798362977114952482  
616868465482523867986344034464628150963921508937545815040466682047715713721621227840324697695626  
297777630041199821864342678475804753248357192547333794306822437076270494661819142506335163767792  
185876954060706075956670454614678684265240481251149104651999608145845880163111859507682653813678  
357232568947154682104162034599843896128077364254667194933511270876876852157768841945030719601517  
226439162139651502405594308730155191795529592939712298205651094061254839082136704164078524220717  
767668597107065912600327715156622109448422724026607472176853125781534048192947128087896545589613  
183001952244442933361217542553447200405077042949545788510564445574308321487320097392295346624975  
526584673440615285407513780234122515226132533585195065471689561622773989143216748230195747388151  
754268630718965814601032795528326619377064655541419598174083709579375183617358537746254154588352  
095630880476686902050603487066710041891861589781173605481797976415986477506931999808736257478236  
5

#### ORF8a

>YP\_009825059.1 ORF8a protein [SARS coronavirus Tor2]  
MKLLIVLTCISLCSCICTVVQRCASNKPHVLEDPCKVQH  
ORF8a (YP\_009825059.1) protein@integer  
401828244162252274550859809912295392002244490657835

#### ORF8b

>YP\_009825060.1 ORF8b protein [SARS coronavirus Tor2]  
MCLKILVRYNTRGNTYSTAWLALGKVLPHRWHTMVQTCTPNVTINCQDPAGGALIARCWYLHEGHQTA  
AFRDVLVVLNKRRTN  
ORF8b (YP\_009825060.1) protein@integer  
138485375954911707664884992178500691168706800448724085743814025911774065932694474567872094225894422  
32599942909

#### Human CoV

Spike protein [Human coronavirus NL63]  
>YP\_003767.1 Spike protein [Human coronavirus NL63]  
MKLFLILLVPLASCFFTCNSNANLSMLQLGVPDSSSTIVTGLLPHTWFCANQSTSVYSANGFFYIDVGN  
HRSAFALHTGYDANQYYIYVTNEIGLNASVTLKICKFSRNTTFDFLSNASSSFDICVNULLFTEQLGAPL  
GITISGETVRLHLYNVTRTFYVPAAYKLTKLSVKCYFNYSVFNATVTVNVTTTHNGRVVNYTVCDDC  
NGYTDNIFSVQQDGRIPNGFPFNWFLLTNGSTLVDGVSRLYQPLRLTCLWPVPGLKSSSTGFVYFNATGS  
DVNCGYQHNSVVDVMRYNLNFSANSLDNLKSGVIVFKTLQYDVLFYCSNSSSGVLDTTIPFGPSSQPYY  
CFINSTINTTHVSTFVGIPLPTVREIVVARTGQFYINGFKYFDLGFIEAVNFVTTASATDFWTVAFATF  
VDVLVNVSATNIQNLLYCDSPFEKLQCEHLQFGLQDGFYSANFLDDNVLPETYVALPIYYQHTDINFTAT  
ASFGGSCYVCKPHQVNISLNGNTSVCVVRTSHFSIRYIYNRVKSGSPGDSSWHIYLKSGTFCPFSFKLNNF  
QKFKTICFSTVEVPGSCNFPLEATWHYTSYITVIGALYVTWSEGNISITGVPYPVSGIREFSNLVNNCTKY  
NIYDYVGTGIIRSSNQSLAGGITYVNSGNLLGFKNVSTGNIFIVTPCNQPDQVAVYQQSIIGAMTAVNE  
SRYGLQNLQLPNFYVSNNGNCTTAVMTYSNFGICADGSLIPVRPRNSSDNGISAIITANLSIPSNWT  
TSVQVEYLQITSTPIVDCATYVCNGNPRCKNLLKQYTSACKTIEDALRLSAHLETNDVSSMLTFDSNAF  
SLANVTSFGDYNLSSVLPQRNIRSSRIAGRSALEDLLFSKVVTSGLGTVDVDYKSCTKGLSIADLACAQY

YNGIMVLPGVADAERMAMYTGSLIGGMVLGGLTSAAIIPFSLALQARLNYVALQTDVLQENQKILAASFN  
KAINNIVASFSSVNDAITQTAEAIHTVTIALNKIQDVVNQQGSALNHLTSQLRHNFAISNSIQAIYDRL  
DSIQADQQVDRLITGRLAALNAFVSQVLNKYTEVRGSRRLAQKINCEVKSQSNRYGFCGNGTHIFSIVN  
SAPDGLLFLHTVLLPTDYKNVKAWSGICVDGIYGYVLRQPNLVLYSDNGVFRVTSRVMFQPRLPVLSDFV  
QIYNCNVTFVNISRVELHTVIPDYVDVNKTLQEFANLPKYVKPNFDLTPFNLTYNLNSSELKQLEAKTA  
SLFQTTVELQGLIDQINSTYVDLKLNRFENYIKWPWWVWLIISVVFVLLSLLVFCCLSTGCCGCCNCL  
TSSMRGCCDCGSTKLPYYEFKVVHVQ

S (YP\_003767.1) protein@integer

114968276709084265349745519407587441802494720943939955159293714251431855182474105257636543865350  
623245093805820100291794178506917388546145205574003003155513025053799657539099286097718632987324  
139872940532698458078894512537696468218976986030930306216519300313045905802605004195950367138735  
287143452421171110916031730745174466130910756848145715453841526985779113432544865153910762974548  
243812217302481332779289564410124456371611399930971268218754271048086268475326024031422691886833  
855094975831456772709439611873582441790522651814894305641095862421720346515677999271727543146862  
802289917603813062210502357135172255806765050010996707761802865379040125365716867480186865201764  
949843688980050677295949082808480907388892100808789199193906563119928409456540545041442044385144  
874102833408915426451626419966345732471327657327969033343095102596697576715587192972195584002234  
435016094396628624425281171111256696905761088234939449877034848372490726540499902862801579336142  
823151243250812973319868702398206165518304661015917160904209873445999633201961990510892058828789  
028488820909817383258829863208255405917417462273963733254093111976081695366782099148323052131155  
701382420005804502586184462036038696528603428380535085260841037309775690718812787677627799641655  
535550138629025493490583018220266935857432702217157451382607291242120229132811482145379783871600  
670234102370272955753303935785326619122386186512815991008257680026049475898507840855585583302427  
952639624672835054992692010623200043773701115754842423606121647465793001286014664182989145676404  
451209753261528831496918727993573125060134756006468294661779642437951429690368142011324037866426  
929451924797291273163086159393195326321944744716764620505077076786868557138143186238391918014347  
3516278389440909151897026014507558091

surface glycoprotein [Human coronavirus 229E]

>NP\_073551.1 surface glycoprotein [Human coronavirus 229E]

MFVLLVAYALLHIAGCQTTNGLNTSYSVCNGCVGYSENVFAVESGGYIPSDFAFNNWFLLTNTSSVVDGV  
VRSFQPLLNCLWSVSGLRFTTGFVYFNGTGRGDCKGFSSDVLSDVIRYNLNFEENLRRGTLIFKTSYGV  
VVFYCTNNTLVSGDAHIFPGTVLGNFYCFVNTTIGNETTSFAVFGALPKTVREFVISRTGHFYINGYRYFT  
LGNVEAVNFNVTTAETTDFACTVALASYADVLVNVQSITSIANIIYCNSVINRLRCDQLSFDVDPDGFYSTSP  
IQSVELPVSIVSLPVYHKHTFIVLYVDFKPQSGGGKCFNCYPAGVNITLANFNETKGPLCVDTSHFTTKY  
VAVYANVGRWSASINTGNCPFSFGKVN NFVKFGSVCFLKDIPGGCAMP IVANWAYS KYTTIGSLYVSW  
DGDGITGVPQPVEGVSSFMNVTLDKCTKYNIYDVSGVGVIRVSNDFLNGITYTSTSGNLLGFKDVTKGT  
IYSITPCNPPDQLVVYQAVVGAMLSNFSTYGFNSVVELPKFFYASNGTYNCTDAVLTYSSFGVCADGS  
IIAVQPRNVSYDSVSAIVTANLSIPSNWTTSVQVEYLQITSTPIVDCSTYVCNGNVRCVELLKQYTSAC  
KTIEDALRNSARLESADVSEMLTFDKKAFTLANVSSFGDYNLSSVIPSLPTSGSRVAGRSAIEDILFSKL  
VTSGLGTVDADYKCTKGLSIADLACAQYYNGIMVLPGVADAERMAMYTGSLIGGIALGGLTSAVSIPFS  
LAIQARLNYVALQTDVLQENQKILAASFNKAMTNIVDAFTGVNDAITQTSQALQTVATALNKIQDVVNQQ  
GNSLNHLTSQLRQNFQAISNSIQAIYDRLDTIQADQQVDRLITGRLAALNVFVSHTLT KYTEVRASRQLA  
QQKVNCEVKSQSKRYGFCGNGTHIFSIVNAPEGLVFLHTVLLPTQYKDVEAWSGLCVDGTNGYVLRQPN

LALYKEGNYRITSRIMFEPRIPTMADFVQIENCNVTFVNISRSELQTIIVPEYIDVNKTLQELSYKLPNY  
TVPDLVVEQYNQNTILNLTSEISTLENKSAELNYTVQKLQTLIDNINSTLVDLKWLNRVETYIKWPWVWVWL  
CISVVLIFVSMLLLCCCSTGCCGFFSCFASSIRGCCSTKLPYYDVEKIHQ

S (NP\_073551.1) protein@integer

950029202782614236942309637699421941713240880498828877156345960485208534158605460696944977035594  
106187293164802402816171801500765320159522361439719359287321188343615666189330106581639865357806  
372063366191340104826405043293041231964660780669787318094741126692504794054988291476660210718971  
469921527430101953699274669106146455498277338252184137256409252605365872026543731034928170666480  
879471370942460599114658937153958696659711210895141166848317882081633842702595059579660598930068  
424058390027089642627386074396746628720930006733431023039579053439768663907160589908399550260265  
106473559537730537314079943271741027314717163024861800232006850093934230493685379609057007000886  
516452215738745518331328496926537855064888838495906133812665694278126090448490995741377174176719  
367938087989539618466444565466903922301206088667832109728549588135465987765786997839896215352217  
039465249659781810267455070637408265526696365117045073859990050792707669760815904841932710332094  
627791250032314812364445062630274217997745448441174977891493354221402159890938839683976579214029  
215031179551946680750800586412253778032618417887544127558659275973287958024128756841211957210900  
290149633882709964526759845060010574518683583466116403314211101419956582250363795717928121836237  
143178930204294798975438998858871833933934864743360906379106401556748885285654596852032262336085  
020347655185402688453447750290560142934669325396827213368106457870601480654516362050470612073200  
53833640452997127847967597948726374147465533790950022685010045256765081026009899590171

spike protein [Human coronavirus OC43]

>CAA79896.1 surface protein [Human coronavirus OC43]

MFLILLISLPTAFVIGDLKCTSDTSYINDKDTGPPPISTDTVDVTNGLGTYVLDREVYLNNTFFLLNGYY  
PTSGSTYRNMALKGSVLLSRLWFKPPFLSDFINGIFAKVKNTKVIKDGVMYSEFPAITIGSTFVNTSYSV  
VVQPRINSTQDGYNKLGLEVSVCQYNMCEYPQTICHPNLGNHRKELWHLDTGVVSCLYKRNFTYDVN  
ADYLYFHFYQEGGTFYAYFTDTGVVTKFLFNVLGMALSHYYVMPLTCNSKVKNFTLEYWVTPLSRQY  
LLAFNQDGIIFNAVDCMSDFMSEIKCKTQSIAPPTGVYELNGYTVQPIADVRRKPNLPCNIEAWLNDK  
SVPSPLNWERKTFSNCFNMSSLMSEFIQADSFTCNNIDA AKIYGMCFSSITIDKFAIPNGRKVDLQLGNL  
GYLQSFNYRIDTTATSCQLYLNLPAAANVSVRFPSTWNRKRFIEDSVFKPRPAGVLTNHDVVYAQHCF  
KAPKNFCPCKLNGSCVSGPGKNNIGITCPAGTNYLTCNLCPTDPITFKATGTYKCPQTKSLVGIGEHC  
SGLAVKSDYCGGNPCTCRPQAFGLWSADSCLQGDKCNIFANFILHDVNSGLTCSTDLQKANTDIILGVCV  
NYDLYGILGQGIFVEVNATYYNSWQNLLYDSNGNLYGFRDYITNRTFMIRSCYGRVSAAFHANSSEPAL  
LSRNIKCNYVFNNLSTRQLQPINYFDSYLGCVVNAYNSTAISVQTCDLTVGSGYCVDYSKNRRSRGAIT  
GYRFTNFEPFTVNSVNDLSEPVGGLYEIQIPSEFTIGNMEEFIQTSSPKVTIDCAAFVCGDYAACKSQLV  
GYGSFCDNINAILTEVNELLDTTQLQVANSLMNGVTLSTKLKDGVNFNVDINFPVLGCLGSECKASS  
RSAIEDLLFDKVKLSDVGFVEAYNNCTGGAEIRDICVQSYKGIKVLPLLENQISGYTLAATSASLFP  
PWTAAGVPFYLNQYRINGLGVTMDVLSQNKLIANAFNNALDAIQEGFDATNSALVKIQAVVNANAEA  
LNNLLQQLSNRFGAISASLQEILSRLDALEAEAQIDRLINGRLTALNAYVSQQLSDSTLVKFSAAQAMEK  
VNECVKSQSSRINFCGNGNHIISLVQNAPYGLYFIHFSYVPTKYVTAKVSPGLCIAGDRGIAPKSGYFVN  
VNNTWMYTGSGYYYLEPITENNVVMSTWAVNYTKAPYVMLNTSIPNLPDFKEELDQWFKNQTSVAPDLS  
LDYINVTFLDLQVEMNRLQEAIKVLNQSYINLKDIGTYEYVYKWPWYVWLLICLAGVAMLVLLFFICCT  
GCGTSCFKKCGGCCDDYTG YQELVIKTSHDD

S (CAA79896.1) protein@integer

37290628025036892085413292035589341255146982048744292943683003868577258958820004756285199169916094

3495214442793021200019051719003611586548360682978415698090994426725620408739011295647156177013197208  
8113541062868267847616111416821460153387667346152744250552794549155400552662172087015008327175551044  
169310368475025321211779561326740395026410450594965706280550496787990799170601985292872303304228909  
051057744924788579651656815555810579600702919893356255771249034360070842480798035733903850058901964  
6649877566411789074382943083882227403860808412721618530434105539377477360734110733287251895822619664  
457324322275876010570278255241724403503440208758930855651952526201919056907996246495164322551881487  
397134270535933381706517196026768914680206879024073449079339750198170216575888016442339945078754137  
073241084941988759514090333628906311603396811008485217334002821618030438664648630474446212783995098  
260851325936357958547929446152771578866245303147420126391980243963885038568248210273096745289042497  
076866306520829760694624284753652738091161223876771464819208178898683100237291948551934004782354625  
605079223903015840963370997446664472030029269171196832771677301919987529906602778136586302743234541  
843454999173856346246547564619899961008849047858722843262860589299318542383058346154631412963958171  
919984914421834941765228647563464320372541242620889879804307286413120939509837267059741932391003601  
784291624502720024011630229640449852228544175849727542913992231341829488119260985801933635929840364  
879931590244404352448021990624370907836724026678272533326845169329050328469987902350121611433955812  
9957019466985549751865569851377044274507467548585952606165384458765129119504633142982243998161074143  
264574845295002366771530059060867561951482507226361349896406650418384605880871222210

spike glycoprotein [Human coronavirus HKU1]

>ADN03339.1 spike glycoprotein [Human coronavirus HKU1]

MLLIIFILPTTLAVIGDFNCTNFAINDLNTTIPRISEYVVDVSYGLGTYIILDRVYLNTTILFTGYFPKS  
GANFRDLSLKGTTLKSTLWYQKPFSLDFNNGIFSRVKNKLYVNKTLYSEFSTIVIGSVFINNSYTIIVVQ  
PHNGVLEITACQYTMCEYPHTICKSIGSSRNESWHFDKSEPLCLFKKNFTYNVSTDWLYFHFYQERGTGF  
AYYADSGMPTTFLSLYLGTLLSHYYVLPLTCNAISSNTDNETLQYWVTPLSKRQYLLKFDDRGVITNAV  
DCSSFFSEIQCKTKSLLPNTGVYDLSGFTVKPVATVHRRIPDLPDCDIDKWLNNFNVPSPLNWERKIFS  
NCNFNLSLTLRLVHTDSFSCNNFDESKIYGSCFKSIVLDFKFAIPNSRRSDLQLGSSGFLQSSNYKIDTTS  
SSCQLYYSLPAINVTINNYNPSSWNRRYGFNNFNLSHVSRYCFSVNNFTFCPCAKPSFASCKSHKP  
PSASCPIGTNYRSCESTTVDHTDWCRCCLPDPITAYDPRSCSQKSLVGVGEHCAGFGVDEEKCGVLD  
GSYNVSCLCSTDAFLGWSYDTCVSNRNCNIFSNFILNGINS GTTCSNDLLQPNTTEVFTDVCVDYDLYGIT  
GQGIFKEVSAVYNSWQNLLYDFNGNIIGFKDFVTNKTYNIFPCYAGRVSAAFHQNASLALLYRNKCS  
YVLNNISLATQPYFDSYLGCVFNADNLT DYSVSSCALRMGSGFCVDYNSPSSSSRRKRRSISASYRFVT  
FEPFNVSFVNDIESVGGLYEIKIPTNFTIVGQEEFIQTNSPKVTIDCSLFVCSNYAACHDLLSEYGTFC  
DNINSILDEVNGLD TTQLHVADTLMQGVTLSSNLNTNLHFDVDNINFKSLVGCLGPHCGSSRSFFEDL  
LFDKVKLSDVGFVEAYNCTGGSEIRDLLCVQSFNGIKVLPILSESQISGYTTAATVAAMFPPWSAAAG  
IPFSLNVQYRINGLGVTMDVLNKNQKLIATAFNALLSIQNGFSATNSALAKIQSVVNSNAQALNSLLQQ  
LFNKFGAIISSSLQEILSRLEAQAQVQIDRLINGRLTALNAYVSQQLSDISLVKLGAAALAMEKVNCEVKS  
QSPRINFCGNGNHILSLVQNPYGLLFMHFSYKPIFKTVLVSPGLCISGDVGIAPKQGYFIKHNDHWMF  
TGSSYYYPEPISDKNVFMNTCSVNFTKAPLVYLNHSPKLSDFESELHWFKNQTSIAPNLTNLHTIN  
ATFLDLYEMNLIQESIKLSNNSYINLKDIGTYEMYVKWPWYVWLLISFSFIIFLVLLFFICCTGCGSA  
CFSKCHNCCDEYGGHHDFVIKTSHDD

S (ADN03339.1) protein@integer

11299447882441443029750953527793836986789178709660859794845058610670706770040029160903906989406720  
619052096161917746860879868309743222084856360946044466780410608859496961496656170631437044455043648  
753388490839320374900658586687280903243745222060118516920118124398054907021348464204989380634141375  
66345257450799453304531555086754395004193876875567735855258329631122282093165885949513290203925205

805945282358207365857195263181400298298630589778535324819168632715080343702741600773652848769563964  
868806537150959774145937956372738241427004814257739522227550381433896557969241336055759630858901836  
055051844081967344256293017908456468571116152799484908632921038439510016620421978064462970595342863  
198410447386562721645724938093871833581847623046372313734140954295102234010692406308485588969931290  
499355611412383530418823555106288827794023432661221937852029994328926162492201802810880734995924517  
4066703802770877831553157310525896918363375419491620109827786572931072572928654506984301112750641767  
10544905726536800328034687680579474931494189481810274984522037121525491017477466714931017323593417964  
4752478966468182893745134295759526426655376862254727031105062277719363354671835327406228742617785197  
4576628822299643349270746231865103703148973146091872615381813016725013445963701005362058535186104662  
3734670561375738661803597586931635274794415451317365714965745104477328688336079298966126223103474416  
456142190884651314183956774713812834598538643206574993103546307827992695520864477662329509429573796  
757572521682975302147492655731199335227095738932272524795679486056459227941604904282908749540519969  
7766564682621021055717648531464774357572103375798415312987363596148595471112680358344534563695982572  
54774633819012452097192336953446650507225951977439018700436278431135271222210

S protein [Middle East respiratory syndrome-related coronavirus]

>AHN10812.1 S protein [Middle East respiratory syndrome-related coronavirus]

MIHSVFLMFLLTPTESYVDVGPDSVKSACIEVDIQQTFDFKTWPRPIDVSKADGIIYPQGRYTSNITIT  
YQGLFPYQGDHGDYVYSAGHATGTPPQKLFVANYSQDVKQFANGFVVRIGAAANSTGTVIISPSTSATI  
RKIYPAFMLGSSVGNFSDGKMGRFFNHTLVLLPDGCGTLLRAFYCILEPRSGNHCPAGNSYTSFATYHTP  
ATDCSDGNYNRNASLNSFKEYFNLRNCTFMYTYNITEDEILEWFGITQTAQQVHLFSSRYVDLYGGNMFQ  
FATLPVYDTIKYYSIIPHSIRSISQSDRKAWAAFVYKQLPLTFLDFSVDGYIRRAIDCGFNDLSQLHCS  
YESFDVESGVYSVSSFEAKPSGSVVEQAEGVECDFSPLLSGTPPQVYNFKRLVFTNCNYNLTKLLSLFSV  
NDFTCSQISPAAIASNCYSSLILDYFSYPLSMKSDLSVSSAGPISQFNYKQSFNPTCLILATVPHNLTT  
ITKPLKYSYINKCSRLLSDDRTEVPQLVNAVQYSPCVSIVPSTVWEDGDYRQKQLSPLEGGGWLVASGST  
VAMTEQLQMFGGITVQYGTDTNSVCPKLEFANDTKIASQLGNCVEYSLYGVSGRGVFNCTAVGVRQQRF  
VYDAYQNLVGYSSDDGNYCLRACVSPVSVIYDKETKTHATLFGSVACEHISSTMSQYSRSTRSMLKRR  
DSTYGPLQTPVGCVLGLVNSSLFVEDCKLPLGQSLCALPDTPTSTLTPRSVRSVPGEMRLASIAFNHPIQV  
DQLNSSYFKLSIPTNFSFGVTQEYIQTTIQKVTVDCQYVCNGFQKCEQLLREYQGFCSKINQALHGANL  
RQDDSVRNLFASVKSSQSSPIIPGFGGDFNLTLLEPVSISTGSRSAIAEDLLFDKVTIADPGYMQGYD  
DCMQQGPASARDLICAQYVAGYKVLPLMDVNMEAAYSLLGSIAGVGVWTAGLSSFAAIPFAQSIFYRL  
NGVGITQQVLSNQKLIANKFNQALGAMQTGFTTTNEAFHKVQDAVNNNAQALSKLASELSNTFGAISAS  
IGDIIQRDLVLEQDAQIDRLINGRLTTLNAFVAQQQLVRSESAALSAQLAKDKVNECVKAQSKRSGFCGQG  
THIVSFVNAPNGLYFMHVGYYPSNHIEVVSAAYGLCDAANPTNCIAPVNGYFIKTNTRIVDEWSYTGSS  
FYAPEPITSLNTKYVAPQVTYQNISTNLPPLLGNSTGIDFQDELDEFFKNVSTSIPNFGSLTQINTLL  
DLTYEMLSLQVVKALNESYIDLKELGNYTYNKWPWYIWLGFIAGLVALALCVFFILCCTGCGTNCMGK  
LKCNRCCDRYEEYDLEPHKVHVH

S (AHN10812.1) protein@integer

1419235811026163370452109220630288651437396241999409776109063741404495107353594163837109176386979408  
537725299229059581596864188452252546996194697262212740545168325252745662956092167819159834898446399  
573234373630685522746577080014959023770685824670254200488798912723584287336874067928417806590412901  
85965874302780218929718880152039856860430160464053297361924365435870442956934662205878582336801001  
421115721393792522752954240658200212369087632397357740277689753904922926770585460450961390530741904  
9634184021191155682644999394401702199357104548573757739116600824023519489577698432181592731479213389



937820891005602159424664283979033044774394199143749115821144448884156380306643358006901968898689730  
2979893303508208932687965365471028602827423590057269565098339259580022893042989036308138779872855  
2818042446389699801030683088861816181171088034027716405142250688126721671494455875880920961308941893  
5962409793721449321636161281820850195128391560583861565005750019355701624138190761768933406227106460  
0435527359389451050017265489273849168269915542369543084114098581171645824571622142198774026993593971  
00682480525420802933759892653808649901630350137715940639083167782330944128335804429406518673014815  
3523768320474184715460871977453349459811326639657397878784117086423336062756473104254519035141287570  
2293821701133215637392058160600888114121241321375510732315015882850483743364379600134275794361246065  
479771347869836809371909385280388744671328668609983249989233082198614108566581578849242855596527736  
99421627160966062455195914709565955928216517878966797705744968218582919035869038825963341241549442  
031283917120273172613563428163908346114258229245571784722730988864913215578296434983946566550008258  
948873018743201339460665918684848294364965649676657574900423460036081958095

spike glycoprotein [SARS coronavirus TjF]

>AAT76147.1 spike glycoprotein [SARS coronavirus TjF]

MFIFLLFLTLTSGSDLRCTTFDDVQAPNYTQHTSSMRGVYYPDEIFRSDTLYLTLQDLFLPFYSNVTGFH  
TINHTFGNPVIPFKDGIYFAATEKSNVVRGWVFGSTMNNKSQSVIIINNSTNVIRACNFELCDNPFPAV  
SKPMGTQHTMIFDNAFNCTFEYISDAFSLDVSEKSGNFKHLREFVFNKDGFLYVYKGYQPIDVVRDLP  
SGFNTLKPFIKPLGINITNFRAILTAFAFPAQDTWGTSAAYFVGYLKPPTTFMLKYDENGITITDAVDCSQ  
NPLAELKCSVKSFEIDKGIYQTSNFRVVPSPGDVVRFPNITNLCPFGEVFNATKFPVYAWERKKISNCVA  
DYSVLYNSTFFSTFKCYGVSATKLNLDLFCFSNVYADSFVVKGDDVRQIAPGQTGVIADYNYKLPDDFMGCV  
LAWNTRNIDATSTGNYNKYRYLRHGKLRPFERDISNVPFSPDGKPCPPALNCYWPLNDYGFYTTTIGIG  
YQPYRVVLSFELLNAPATVCGPKLSTDLIKNQCVNFENGLTGTGVLTPSSKRFQPFQFGRDVSDFTD  
SVRDPKTSEILDISPCSFGGVSVITPGTNASSEVAVLYQDVNCTDVSTAIHADQLTPAWRIYSTGNNVFQ  
TQAGCLIGAEHVDTSYECDIPIGAGICASYHTVSLRNTS QKSIVAYTMSLGADSSIAYSNNTIAIPTNF  
SISITTEVMPVSMAKTSVDCNMYICGDSTECANLLLQYGSFCTQLNRALSGIAAEQDRNTREVFQVQKQM  
YKPTPLKYFGGFNFSQILPDPLKPTKRSFIEDLLFNKVTLADAGFMKQYGECLGDINARDLICAQKFNGL  
TVLPPLLTDDMIAAYTAALVSGTATAGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKQIANQFN  
KAISQIQESLTTTSTALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLIT  
GRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQAAPHGVVFLHVTVY  
PSQERNFTTAPAICHEGKAYFPREGVVFVNGTSWFITQRNFFSPQIITDNTFVSGKCDVVIGIINNTVY  
DPLQPELDSFKEELDKYFKNHTSPDVLGDISGINASVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQ  
YIKWPWYVWLGFIAGLIAIVMTILLCCMTSCCCLKGACSCGSCCKFEDEDDSEPVLKGVKLHYT

S (AAT76147.1) protein@integer

459751045002885528226701443152844149038387483475285677758778632738790035612657588145099489054591287  
6819592271949842214982413478393324150605278220215731915875898409811597490585432869219210619751759485  
737229836204483844500818432164840724624444361400039979031547632642970973961591091457797294852765105  
4377919454076253794743992741803513730098049181913157675582525019679656379438630798284443119355572168  
837047556410655037986782368843230445153313500071669457627118197107458365553619738218610454232802068  
942936198442905353183762477996157694324965604361525958584916984043196202776442286666373614803349260  
304313471483756264036354849031392735240479739359619365650679836297711495248261686846548252386798634  
4034464628150963921508937545815040466682047715713721621353251737450523592683642600317081795740111812  
11092452185593263064311159568452612336399943696732652707423923006475535829364597936855563578274534  
975934518150627447818526649981885623207708835141503585286224477293449703549865587685773783577028925  
55874462932177156146117289186754294015990130063500819907196015172264391621396515024055943087301551917

955295929397122982056510940612548390821367041640785242207177676685971070659126003277151566221094484  
2272402660747217685312578153404819294712808789654558961318300195224444293336121754255344720040507704  
29495457885105644455743083214873200973922953466249755265846734406152854075137802341225152261325335  
85195065471689561622773989143216748230195747388152824704400248435725394747272615447971664124623541419  
5981740837095793751836173585377462541545883520956308804766869020506034870667100418918615897811736054  
817979764159864775069319998087362574782365

Bat coronavirus

Bat coronavirus RaTG13 (MN996532.2)

S

>QHR63300.2 spike glycoprotein [Bat coronavirus RaTG13]

MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSSTRGVYYPDKVFRSSVLHLLTQDLFLPFFSNVTWFHAIHV  
SGTNGIKRFDNPLPFNDGVYFASTKSNIRGWIFGTLLDSTKQSLNATNVVIVKCEVFCNDPF  
LGVYYHKNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFNIDGYFKIYSKHTPI  
NLVRDLPPGSALEPLVDLPIGINITRFQTLALHRSYLTGPDSSSGWTAGAAAYVGYLQPRTFLLKYN  
ENGTITDAVDCALDPLSETKCTLSFTVEKGIYQTSNFRVQPTDSIVRFPNITNLCPFGEVFNATTFASV  
YAWNKRISNCVADYSVLYNSTSFSTFKCYGVSPTKLNLDLCTNVYADSFVITGDEVRQIAPGQTGKIAD  
YNYKLPDDFTGCVIAWNSKHIDAKEGGNFNYLRLFRKANLKPFERDISTEIQAGSKPCNGQTGLNCYY  
PLYRYGFYPTDGVGHQPYRVVLSFELLNAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFL  
PFQQFGRDIADTTDAVRDPQLEILDITPCSFGGVSVITPGTNASNQVAVLYQDVNCTEVPVAIHADQLT  
PTWRVYSTGSNVFQTRAGCLIGAEHVNNSYECDIPIGAGICASYQTQTSRVSASQSIAYTMSLGAENS  
VAYSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECNLLLQYGSFCTQLNRALTGIAVEQ  
DKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQYGDCLGDI  
AARDLICAQKFNGLTVLPLLTDEMAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQN  
VLYENQKLIANQFNSAIGKIQDLSSTASALGKLQDVVNQNAQALNTLVKQLSSNFGAIVLNDILSRL  
DKVEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSPFQ  
SAPHGVVFLHVTVPAQEKNFTTAPAICHGKAHFPREGVFSNGTHWFVTQRNFYEPQIITDNTFVSG  
SCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVLGDISGINASVNIQKEIDRLNEVAKNLN  
ESLIDLQELGKYEQYIKWPWYIWLGFIAGLIAIIMVTIMLCCMTSCCSCCLKGCCSCGCKFDEDDSEPV  
LKGVKLHYT

S (QHR63300.2) protein@integer

752747857524267340527437918328655730616586883245930408571123989622534847707375331915898778448790  
838649727782194024123600195168038216507005128605295529985960432046213882598576178276311940202762  
884596965594756163177821687329835913579093416980311381321970268841831777077077331743586645908881  
497676592388240703158873214470043811143519742139368917649108880667763551690265177959925861868676  
662638442173861960300996521982048152275873154741390970828929642094255580596153492688341969839342  
288632781713308341567645585139229810952798765270578926305574612725262371833491429951530091814533  
168200091853906202336912011556278993032867083739330020294989098628156091191553664426692283401074  
091256407721907109542647294573688152243901853385770692485064939355877179695806842841289242863749  
892096894356131737953609141908900791795183196538611070826236968754425296189556447849109524675850  
613744850659083278048954177382742712203148883934288472060919150973195650411378771965204855097704  
873461345655476565128691183005821362497848017060376680985304314236600064249390720423979178561802  
309289766217879929783983663618477090509370407971890783344790313957920989990494728270760509776015  
806896251653606269697319739925895690127204512148606560243611151759314322589794035924203986164656  
715532677666727366646529148537937143920231248630014874030906841467792393310564445574308321487320

097392295346624975526588885932282027694981568319839419111253496862159322806300095755234862898857  
225696691104469672691382491574137258766881667848270835025637159776468558174083709579375183617358  
537746254154588352095630880476686902050661133141940384254735346313287799913797979771429677506931  
9998087362574782365

E

>QHR63302.1 envelope protein [Bat coronavirus RaTG13]

MYSFVSEETGLIVNSVLLFLAFVVFLLVTLAILTALRLCAYCCNIVNVSLVKPSFYVYSRVKNLNSSRV  
PDLV

E (QHR63302.1) protein@integer

281585756715583860382754061271694289415355617329836193393526618386393321015201910475515428613629  
44

ORF8

>QHR63307.1 nonstructural protein NS8 [Bat coronavirus RaTG13]

MKLLVFLGILTTVTAHFHQECSLQSCAQHQPYVDDPCPIHFYSKWYIRVGARKSAPLIELCVDEVGSKSP  
IQYIDIGNYTVSCSPFTINCQEPKLGSLVVRCSFYEDFLEYHDVRRVLDLFI

ORF8 (QHR63307.1) protein@integer

194311933255407448173576852068640922392487101348910877186010101374230324575542845088484081714909  
10631167336036196695783211124221825233150723390246662221500328

betacoronavirus sp. RpYN06 (MZ081381.1)

S

>QWN56252.1 surface glycoprotein [Betacoronavirus sp. RpYN06]

MLFFLFLQFTLVNSQCVNLTGRTPLNPNYTNSSQRGVYYPDTIYRSDTLVLSQGYFLPFYSNVSWYYSLT  
ANNAATKRTDNPILDFKDGIFYAATEHSNIVRGWIFGTTLDNTSQSLIVNNATNVIKVCNFDYDPY  
LSGYYHNNKTWSIREFAVYSSYANCTFEYVSKSFM LNISGNGGLFNTLREFVFRNVDGHHFKIYSKFTPN  
LNRGLPTGLSVLQPLVELPVSINITKFRLLTIHRGDPMPNNGWTAFAAAYFVGYLKPRTFMLKYDENG  
ITDAVDCALDPLSEKCTKSLTVQKGIYQTSNFRVQPTQSVVRFNITNVC PFHNVFNATRFPSVYAW  
RTKISDCIADYTVFYNSTSFSTFKCYGVSPSKLIDLCFTSVYADTFLIRFSEVRQVAPGQTGVIADYNYK  
LPDDFTGCVIAWNTAKQDVGNYFYRSHRSTKLPFERDLSSDENGVRTLSTYDFNPVPLEYQATR VVLL  
SFELLNAPATVCGPKLSTQLVKNQCVNFNFENGLKGTGVLTDSSKRFQSFQQFGKDASDFIDSVRDPQTL  
ILDITPCSFGGVSVITPGTNTSLEVAVLYQDVNCTDVPTTIHADQLTPAWRIYATGANVFQTQAGCLIGA  
EHNASYECDIPIGAGICASYHAASILRSTSQKAI VAYTMSLGAENSIAYANNSIAIPTNFSISVTTEVM  
PVSMAKTSVDCTMYICGDSIECSNLLLQYGSFCTQLNRALSIAIEQDKNTQEVFAQVKQIYKTPPIKDF  
GGFNFSQILPDPSPKRSFIEDLLFNKVTLADAGFIKQYGDCLGGVSARDLICAQKFNGLTVLPPLLTD  
EMIAAYTSALVSGTATAGWTFGIGAALQIPFAMQMAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQES  
LTSTASALGKLQDVVNQNAQALNTLVKQLSSNFGAIVSLNDILSRDLKVEAEVQIDRLITGRLQSLQTY  
VTQQLIRAAEIRASANLAATKMSECVLQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTT  
APAICHNGKAHFPREGV FVSNNGTHWFVTQRNFYEPQIITDNTFVSGNCDVVIGIVNNTVYDPLQPELDS  
FKEELDKYFKNHTSPDVLGDISGINASVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYIW  
LGFIAGLIAIVMVTIMLCCMTSCCCLKGCCSCGSCCKFDEDDSEPVLKGVKLHYT

S (QWN56252.1) protein@integer

871901102170401784955832427277985446353123721479493372451774093695733462561257218583802680412925  
428121626377217613118946026245822336501040740056232896133848696249550976140129115268243252941572

182990709782763623286014968464191133201644245714669829907701640084596737079470284497136744123767  
405670804659171098445829195667823194774155415218739169876143563600577748446842588551342382564367  
303478024396828559086899913468804870704183637310580483666392947752993105682647159995835495709535  
128439251715189592597856818766385255303190010313386152718067576359803452189958287652605092968175  
472602806334748288570832134972874510682356188151437172566456980148187517557295233649136948736248  
914612749687988599664958438754719162969926468233411571921987631529568739747885588138827017545011  
569977662225338986585989095840514061473830786021686111496535104505710874314286675922831665471483  
741827964438568859381381988659192703620275434603383994744150900372251801655469231344608492807105  
689274391564972016221522936080014275563063807254941665996080756892296288192827260701670302347596  
712519860332397393927686741030667390817162656172578142184318877105019859607329559734545198536425  
272734756559188706989658839078254167435648959182501458241515250744507247661924944554990973498388  
466773356753173192158662640256616467344393310564445574308321487320097392295346624975526588885932  
282027694981568319839419111253418398387630472790547807867072803027024417248048669243777865953809  
567050618882114384887323695443751776468558174083709579375183617358537746254154588352095630880476  
6869020506611331419403842406615974777549999137979797714296775069319998087362574782365

E

>QWN56254.1 envelope protein [Betacoronavirus sp. RpYN06]

MYSFVSEETGTLIVNSVLLFLAFVVFLLVTLAILTALRLCAYCCNIVNVSLVKPSFYVYSRVKLNLSRV  
PDLV

E (QWN56254.1) protein@integer

281585756715583860382754061271694289415355617329836193393526618386393321015201910475515428613629  
44

ORF8

>QWN56259.1 ORF8 [Betacoronavirus sp. RpYN06]

MKFLVFLGILTTVAFAHQECSLQSCAQHQPYVDDPCPIHFYSKWWYIRVGARKSAPLIELCVDEAGSKSP  
IQYIDIGNYTVSCSPFTINCQEPKLGSLVVRCSFYQDFLEYHDVRRVLDLFI

ORF8 (QWN56259.1) protein@integer

194611009554459053595943101340350546355775060196398877186010101374230324575542845088267908932795  
32250367336036196695783211124221825233150657854246662221500328

Bat SARS coronavirus HKU3-1 (DQ022305.2)

S

>AAY88866.1 spike glycoprotein [Bat SARS coronavirus HKU3-1]

MKILIFAFLANLAKAQEGCGIISRKPQPKMAQVSSRRGVYNDIFRSDVLHLLTQDYFLPFDSNLTYF  
SLNVDSDRYTYFDNPILDFGDGVYFAATEKSNVIRGWIFGSSFDNTTQSAVIVNNSTHIIIRVCNFNLC  
EPMYTVSRGTQQNAWVYQSAFNCTYDRVEKSFQLDTPKTGNFKDLREYVFKNRDGLSVYQTYTAVNLP  
RGLPTGFSVLKPKLKPFGINITSYRVVMAMFSQTTSNFLPESAAAYVGNLKYSTFMLRFNENGTITDAV  
DCSQNPLAELKCTIKNFNVDKGIYQTSNFRVSPTQEVIRFPNITNRCPFDKVFNATRFPNVYAWERTKIS  
DCVADYTVLYNSTSFSTFKCYGVSPSKLIDLCFTSVYADTFLIRSSSEVRQVAPGETGVIADYNYKLPDDF  
TGCVIAWNTAKHDTGNYYRSHRKTCLKPFERDLSSDDGNGVYTLSTYDFNPVNPVAYQATRNVVLSFEL  
LNAPATVCGPKLSTELVKNQCVNFNFNGLKGTGVLTSSSKRFQSFQQFGRDTSDFDTSVRDPQTLEILDI  
SPCSFGGVSIVITPGTNASSEVAVLYQDVNCTDVPTAIRADQLTPAWRVYSTGVNVFQTQAGCLIGAEHVN  
ASYECDIPIGAGICASYHTASVLRSTGQKSIVAYTMSLGAENSIAYANNSIAIPTNFSIVTTEVMPVSM

AKTAVDCTMYICGDSLECSNLLLQYGSFCTQLNRALTGIAIEQDKNTQEVEFAQVKQMYKTPAIKDFGGFN  
FSQILPDPSPKPTRSFIEDLLFNKVTLADAGFMKQYGDCLGDVVSARDLICAQKFNGLTVLPPLLTDEMVA  
AYTAALVSGTATAGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQESLSST  
ASALGKLQDVVNQNAQALNLTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLITGRLQSLQTYVTQQ  
LIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLSFPQSAPHGVVFLHVTVYVPSQEKNFPTAPAI  
CHEGKAYFPREGVVFVSNGTSWFITQRNFYSPQLITTDNTFVSGNCDVVIGIINNTVYDPLQPELDSFKEE  
LDKYFKNHTSPDVLGDISGINASVUNIQKEIDRLNEVAKNLNESLIDLQELGKYEYIKWPWYVWLGF  
AGLIAIVMTILLCCMTSCCCLKGACSCGSCCKFEDEDDSEPVLKGVKLHYT

S (AAY88866.1) protein@integer

553604507726171140559209236056232394069104112935630531887077596391251144914466033173086661386769721  
9589051450640049340560683386040366001264372331862543549616675701016710367242214516563064151035691764  
772930562005169186038789437604247103488619210451101578937052677229755600299973505592201261095901834  
0776153246463196279365772294625284976061132471599827100242472641388707456783779441773317709356680964  
37699315395645998922168725317184384079968219153506829043696232100255876847864908009785968394606687  
702908219707863640175558245686027235288832585215684713651749268832892312867995546796686411661788543  
7283059124190253596371204489499645843283242835194764765347746513583151878571991581736384863289045171  
445846452591871965230531542989240983871071142329102528726567228915868995512821011530117207897300392  
0412487078713274182787064232626417903373775947257351153474809746457838916049576899752979063973255292  
988333161860320599227694534669420779509685529045787979246530907024190106350939167137997068539018465  
1387402354623527193057307519745153640287984820676688262332139315023783319741107738181301439799364798  
515755637699984333790013645034931009910819502222112252504588752248837527637007442540886130095583609  
02077595976239708225900842477812600942234025061399359095612791603147843934007759509883792393310564  
445574308321487320097392295346624975526588885932282027694981696874882959830475775611912982819068194  
5997827029265898548817615947629412033484508313392749449706865171778193770646555414195981740837095793  
7518361735853774625415458835209563088047668690205060348706671004189186158978117360548179797641598647  
75069319998087362574782365

E

>AAY88868.1 small membrane protein [Bat SARS coronavirus HKU3-1]

MYSFVSEETGTLIVNSVLLFLAFVVFLLVTLAILTALRLCAYCCNIVNVSLVKPTVYVYSRVKLNLSSEG  
VPDLLV

E (AAY88868.1) protein@integer

563171513431167720765508122543388578830711234659672386787053236772786647063568620951030851853362944

ORF8

>AAY88873.1 hypothetical protein orf8 [Bat SARS coronavirus HKU3-1]

MKLLIVFGLLASVYCFHRECSIQECCENQPYQIEDPCPIHYSDWFIKIGSRKSARLVQLCEGDYGRIP  
IHYEMFGNYTISCEPLEINCQAPPVGLSIVRCSYDYDFVEHHDVRRVLDLFI

ORF8 (AAY88873.1) protein@integer

194312217585058806996835080468573177297896488686076903395837346035516233674327126417627376306294  
45631158913992966542329489171035381805331858737927942221500328

Synthetic coronavirus (Bat-SRBD)

S

>ACJ60703.1 spike glycoprotein [recombinant coronavirus]

MKILIFAFLANLAKAQEGCGIISRKPQPKMAQVSSRRGVYYNDDIFRSDVLHLTQDYFLPFDSNLTKYF  
SLNVDSDRYTYFDNPILDFGDGVYFAATEKSNVIRGWIFGSSFDNTTQSAVIVNNSTHIIIRVCNFNLC  
EPMYTVSRGTQQNAWVYQSAFNCTYDRVEKSFQLDTPKTNFKDLREYVFKNRDGFSLVYQTYTAVNLP  
RGLPTGFSVLKPKILKLPFGINITSYRVVMAMFSQTTSNFLPESAAYYVGNLKYSTFMLRFNENGTITDAV  
DCSQNPLAELKCTIKNFNVKGIYQTSNFRVSPTQEVIRFPNITNLCPFGEVFNATKFPSVYAWERKKIS  
NCVADYSVLYNSTFFSTFKCYGVSATKLNLDLCSNVYADSFVVKGDDVRQIAPGQTGVIADYNYKLPDDF  
MGCVLAWNTRNIDATSTGNYNYKYRYLRHGKLRPFERDISNVPFSPDGKPCPPALNCYWPLNDYGFYTT  
TGIGYQPYRVVLSFELLNAPATVCGPKLSTDLVKNQCVNFNFNGLKGTGVLTSSSKRFQSFQQFGRDTS  
DFTDSVRDPQTLEILDISPCSFSGGVSVITPGTNASSEVAVLYQDVNCTDVPTAIRADQLTPAWRVYSTGV  
NVFQTQAGCLIGAEHVNASYECDIPIGAGICASYHTASVLRSTGQKSIVAYTMSLGAENSIAYANNSIAI  
PTNFSISVTTEVMPVSMAKTAVDCTMYICGDSLECSNLLLQYGSFCTQLNRALTGIAIEQDKNTQEVFAQ  
VKQMYKTPAIKDFGGFNFSQILPDPSPKTKRSFIEDLLFNKVTLADAGFMKQYGDCLGDVSARDLICAQK  
FNGLTVLPLLTDENVAAAYTAALVSGTATAGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKLLA  
NQFNSAIGKIQESLSTASALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRDKVEAEVQID  
RLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLH  
VTYVPSQEKNFHTTAPAICHEGKAYFPREGVVFVSNGTSWFITQRNFYSPQLITTDNTFVSGNCDVVIGIIN  
NTVYDPLQPELDSFKEELDKYFKNHTSPDVLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELG  
KYEQYIKWPWYVWLGFIAGLIAIVMTILLCCMTSCCCLKGACSCGSCCKFDEDDSEPVLKGVKLHYT  
S (ACJ60703.1) protein@integer

72562050036684703735376672988362492355425614290698965075503034714194066078622889190006281488128668  
0996597615173829254716596989277508285251772381028188730813535691748366226125517154111495394440455019  
099231755462314153555247620916566387634846029715224678615443736850985852604251812732498100369436204  
519622119583204240627290305062011253523822847553175325376829813740521006637755635389921122988007988  
8736282124667538112370726491903904089909532474991295420096387702257298875718098618146314749170655775  
689504130227736916839076390228038887394522906626950786183751792424319570889178371151357008838975634  
15265835596378373700230453198397038644388328190294029195096375252033907661588271733385315866178062  
636999665199117639048009071615933234195362390215798590122646292957478883016714675267551872693039906  
713977444971290757989946169474033340234918680731987805159130607316167322444990354709132256603897228  
40732775592979063973255292988333161860320599227694534669420779509685529045787979246530907024190106  
350939167137997068539018465138740235462352719305730751974515364028798482067668826233213931502378331  
9741107738181301439799364798515755637699984333790013645034931009910819502222112252504588752248837527  
63700744254088613009558360902077595976239708225900842477812600942234025061399359095612791603147843  
93400775950988379239331056444557430832148732009739229534662497552658888593228202769498169687488295  
9830475775611912982819068194599782702926589854881761594762941203348450831339274944970686517177819377  
0646555414195981740837095793751836173585377462541545883520956308804766869020506034870667100418918615  
897811736054817979764159864775069319998087362574782365

E

>ACJ60696.1 envelope protein [recombinant coronavirus]

MYSFVSEETGTLIVNSVLLFLAFVVFLLVTLAILTALRLCAYCCNIVNVSLVKPTVYVYSRVKLNLSSEG  
VPDLLV

E (ACJ60696.1) protein@integer

563171513431167720765508122543388578830711234659672386787053236772786647063568620951030851853362  
944

ORF8

>ACJ60700.1 ORF 8 [recombinant coronavirus]

MKLLIVFGLLASVYCFHRECSIQECCENQPYQIEDPCPIHYYSDFWFIKIGSRKSARLVQLCEGDYGRRI  
IHYEMFGNYTISCEPLEINCQAPPVGLIVRCSYDYDFVEHHDVRVVLDFI

ORF8 (ACJ60700.1) protein@integer

194312217585058806996835080468573177297896488686076903395837346035516233674327126417627376306294  
45631158913992966542329489171035381805331858737927942221500328

Mutation enzyme

ADAR 1

>NP\_056656.3 double-stranded RNA-specific adenosine deaminase isoform c [Homo sapiens]

MNPRQGYLSGYTHPFQGYEHRQLRYQQPGGSSPSSFLKQIEFLKQQLPEAPVIGKQTPSLPPSLPG  
LRPRFPVLLASSTRGRQVDIRGVPRGVHLRSQGLQRGFQHPSPRGRSLPQRGVDCLSSHFQELSIYQDQE  
QRILKFLEELGEGKATTAHDLGKLGTPKKEINRVLYSLAKKGLQKEAGTPPLWKIAVSTQAWNQHSGV  
VRPDGHSQGAPNSDPSLEPEDRNSTSVSEDLLEPFIAVSAQAWNQHSGVVRPDSHSQGSNSDPGLEPED  
SNSTSALEDPLEFLDMAEIKEKICDYLFNVSDSSALNLAKNIGLTKARDINAVLIDMERQGDVYRQGTTP  
PIWHLTDKKRERMQIKRNTNSVPETAPAAIPETKRNAEFLTCNIPTSNASNMMVTTEKVENGEQEPVIPLE  
NRQEARPEPARLKPPVHYNGPSKAGYVDFENGQWATDDIPDDLNSIRAAPGEFRAIMEMPSFYSHGLPRC  
SPYKKLTECQLKNPISGLLEAQAQFASQTCEFNMIQSGPPHEPRFKFQVINGREFPPAEAGSKKVAQD  
AAMKAMTILLEEAKAKDSGKSEESHYTEKESEKTAESQTPPTSATSFSGKSPVTTLLECMHKLGNCS  
EFRLLSKEGPAHEPKFYCVAVGAQTFPSVSAPSKKQAKMAAEEAMKALHGEATNSMASDNQVRKIGEL  
VRYLNTNPVGGLEAYARSHGFAAEFKLVDQSGPPHEPKFVYQAKVGGRWFPVCAHKKQKQEAADAAL  
RVLIGENEKAERMGFTLPLTGSTFHDQIAMLSHRCFNTLNSFQPSLLGRKILAAIIMKKDSEDMGVVV  
SLGTGNRCVKGDLSLKGTVNDCHAEIISRRGFIRFLYSELMKYNSQTAKDSIFEPAKGGEKLQIKKTV  
SFHLYISTAPCGDGFALDKSCSDRAMESTESRHYPVFENPKQKLRKTKVENGEQIPVESSDIVPTWDGI  
RLGERLRTMCSKILRWVNLGLQGALLTHFLQPIYLSVTLGYLFSQGHLLTRAICCRVTRDGSFEDGL  
RHPFVNHVPKVRVSIYDSKRQSGKTKETSVNWCLADGYDLEILDGTRGTVDGPRNELSRVSKKNIFLLF  
KKLCSFRYRRDLLRLSYGEAKKAARDYETAKNYFKKGLKDMGYGNWISKPQEEKNFYLCVP

NP\_056656.3 ADAR 1 protein@integer

23744005647123592848072562765264854870948459081648557052726664187985203094297030146053384826158736  
00426517767398622324043326342185694370234760302124797800356479078220486833456962057978747805392030  
854702546467939267737458116825946591158344565456272571498499832912545625954495188252200804313780226  
179152791565982744244671306769607540332746330985738082113182440055515113071372178602150806293862549  
56202249550322703840949284968829375509461527149583651956692675469244305598567556696020055312332264  
380359249618555148717871973503940574614595490364942350571336265148316262258156058931592976252298158  
7331544626307310609115661862502262618894236811303309741601061642494748596324647760951684005410460777  
46562425148576748944759844501872304212340905047364882463731252533024497252908960760662270443501965  
652877460200967422702298994899310254406338667027713822170277034091833471731077129084922702038825852  
236883882539185748901559289463901630824823746999165071723173433434074739241985594566839635822058266  
4980578192849165946552863213493982249112512058712464186792853327312259718227121900207203346092619162  
863256581218837351785633812302360571622316043572924187582252750488946516599944821384409463648786953  
7408776985167218778866036451837664067949551695508288918235179508904463665619679740310290877033321788  
463631542650744178930311511818442882229002066280128038952727255086696092559113574739370913455605310  
48452270207769912556292571555424655982591105399072983951835467423613842216617186326785196895728073

5059195784275979936618094426331526152926044790138464

### ADAR 2

>NP\_001333616.1 double-stranded RNA-specific editase 1 isoform 8 [Homo sapiens]

MASSTTGPLHMGTFFSIMGRRYKRRRKKRSERKDRNSLRQSRNPQKYFAMDIEEENMSSSSTDVKENRN  
LDNVSPKDGSTPGPGEGSQLSNGGGGGPGRKRPLEEGSNHGHYRLKRRRTPGPVLPKNALMQLNEIKP  
GLQYTLLSQTGPVHAPLFFVMSVEVNGQVFEGSGPTKKKAKLHAAEKALRSFVQFPNASEAHLAMGRTLV  
NTDFTSDQADFPDITLNGFETPKAEPPFYVGSNGDDSFSSGDLSLSASPVPASLAQPPLPVLPPFP  
SGKNPVMILNELRPLKYDFLSESGESHAKSFVMSVVVDGQFFEGSGRNKKLAKARAAQSALAAIFNLHL  
DQTPSRQPIPSEGLQLHLPQVLADAVSRLVLGKFGDLTDNFSSPHARRKVLAVGVMTTGTDVKDAKVISV  
STGKTCINGEYMSDRGLALNDCHAEIISRRSLLRFLYTQLELYLNNKDDQKRSIFQKSERGGFRLKENVQ  
FHLIYSTPCGDARIFSPHEPILEGSRSYTQAGVQWCNHGSLQPRPPGLSDPSTSTFQAGTTEPADRH  
PNRKARGQLRTKIESGEGTIPVRSNASIQTWGVLQGERLLTMSCSDKIARWNVVGIQGSLLSIFVEPIY  
FSSILGSLYHGDHLSRAMYQRISNIEDLPPLYTLNKPLLSGISNAEARQPGKAPNFSVNWTVGDSAIEV  
INATTGKDELGRASRLCSLPLTTLQDYQAQRVP

NP\_001333616.1 ADAR 2 protein@integer

317538518908554143096831679314873308199868992704285981811172758055985325197118903769159477207902  
612467835883195482617140874003067287866081406333469331130040557412292015588462546672275373354318  
479602131151467815371262165011049467316597245314638089638893859812707651746057216151607757671345  
618312719074156453025856552584164276973645050131205229842631047448021116329291073471782316057028  
582189741573016150110571351254566922669326084847153530204430931568576536482765305728148342529872  
193849856792047155298285964072749930122433848618531135463694716501630738583134561996268931720940  
189082290249296421392382717327899327156283905429439720486775500717927054713329163882525515828732  
014394710117824888224535866651619717448614914076433657666793033956708296856814301064165164275624  
313573580742284535738283405927722125797726376949860549853312905975601885418297401171948146995830  
420190685095150920676585526236432110135730398147694205736282891605079567840311079587454883

### ADAR 3

>NP\_061172.1 double-stranded RNA-specific editase B2 [Homo sapiens]

MASVLGSGRSGGGLSSQLKCKSKRRRRRSKRKDKVSILSTFLAPFKHLSPGITNTEDDDLSTSSAEVK  
ENRNVGNLAARPPSGDRARGGAPGAKRKRPLEEGNGGHLCCKLQLVWKKLSWSVAPKNALVQLHELRLPGL  
QYRTVSQTGPVHAPVFAVAVEVNGLTFTGTPTKKKAKMRAAELALRSFVQFPNACQAHLAMGGGPGPGT  
DFTSDQADFPDITLQFEFPPAPRPLAGGRPGDAALLSAAYGRRLLCRALDLVGPTPATPAAPGERNPV  
VLLNRLRAGLRYVCLAEPERRARSFVMAVSVDGRTFEGSGRSKKLARGQAAQAALQELFDIQMPGHAPG  
RARRTPMPQEFADSISQLVTQKFREVTTDLTPMHARHKALAGIVMTKGLDARQAQVVALLSSGTKCISGEH  
LSDQGLVVNDCHAEVVARRAFLHFLYTQLELHLSKRREDSERSIFVRLKEGGYRLRENILFHLVSTSPC  
GDARLHSPYEITDLHSSKHLVRKFRGHLRRTKIESGEGTVPVRGPSAVQTDGVLGELQITMSCTDKIA  
RWNVLGLQGALLSHFVEPVYLQSIIVGSLHHTGHLARVMHRMEGVGQLPASVYRHRPRLSGVSDAEARQ  
PGKSPFMSMNWVVGSADEIINATTGRRSCGGPSRLCKHVLSARWARLYGRLSTRTPSPGDTPSMYCEAK  
LGAHTYQSVKQQLFKAFQKAGLGTWVRKPPEQQQLLTL

NP\_061172.1 ADAR 3 protein@integer

203228425407285014027135733646961593929231197872690715686822387801082769232486094050207564926619  
006124207770946690488348033320064452874240401897098757061564178747684281590537050167973200924261  
414509598056125331655447415220113329498390293478534345330371119272421488109719131837309666311255  
236501398897602462747291335522199630150760450160714679642878087050772100821961867620919380278577



441143564568862829069963023849253583124424855398817531606725145214795915113370978110215166181027  
703298148781795759258654305444739677384763764495870230599181884442025022031641063723429175610050  
491691987338156841199176669775310422918439821318970883139084171688136577214481693618583623025869  
10266178857505727156722156305934079236531350705928701226389231733427792908468603093375445546232  
132644410149016312602853039585002103855747804190939402014842366270468125149255930605286127235276  
214545283573397776715046063058518140636805461389317845383457538237933427937380257694022036218189  
07

Modification enzyme

GlcNAcase

>sp|O60502.2|OGA\_HUMAN RecName: Full=Protein O-GlcNAcase; Short=OGA; AltName: Full=Beta-N-acetylglucosaminidase; AltName: Full=Beta-N-acetylhexosaminidase; AltName: Full=Beta-hexosaminidase; AltName: Full=Meningioma-expressed antigen 5; AltName: Full=N-acetyl-beta-D-glucosaminidase; AltName: Full=N-acetyl-beta-glucosaminidase; AltName: Full=Nuclear cytoplasmic O-GlcNAcase and acetyltransferase; Short=NCOAT  
MVQKESQATLEERESELSSNPAASAGASLEPPAAPAGEDNPAGAGGA AVAGAAGGARRFLCGVVEGFYG  
RPWVMEQRKELFRRLQKWELNTYLYAPKDDYKHRMFWREMYSVEEAEQLMTLISAAREYEIEFIYAI SPG  
LDITFSNPKEVSTLKRKLDQVSQFGCRSFALLFDDIDHNMCAADKEVFSSFAHAQVSITNEIYQYLGEPE  
TFLFCPTEYCGTFCYPNVSQSPYLRTVGEKLLPGIEVLWTGPKVVSKEIPVESIEEVSKIIKRAPVIWDN  
IHANDYDQKRLFLGPYKGRSTELIPRLKGVLTNPNCFEANYVAIHTLATWYKSNMNGVRKDVVMTDSED  
STVSIQIKLENEGSDEDIETDVLVYSPQMALKLALTEWLQEFVPHQYSSRQVAHSGAKASVVDGTPLVAA  
PSLNATTVVTTVYQEPIMSQGAALSGEPTTLTKEEEKQKPDEEPMDMVVEKQEETHKNDNQILSEIVEA  
KMAEELKPMDDTKESIAESKSPMSMQEDCISDIAPMQTDEQTNKEQFVPGPNEKPLYTAEPVTLEDLQL  
LADLFYLPYEHGPKGAQMLREFQWLRANSSVSVNCKGKDKSEKIEE WRSRAAKFEEMCGLVMGMFTRLSN  
CANRTILYDMYSVVDIKSIMSMVKSQVWLGCRSHSSAQFLIGDQEPWAFRGGLAGEFQRLPIDGAND  
LFFQPPPLTPTSKVYTIRPYFPKDEASVYKICREMYDDGVLPFQSQPD LIGDKLVGGLLSLSDYCFVL  
EDEDGICGYALGTVDVTPFIKKCKISWIPFMQEKYTKPNGDKELSEAEKIMLSFHEEQEVLPETFLANFP  
SLIKMDIHKKVTDPVAKSMMACLLSSLKANGSRGAFCEVRPDDKRILEFYSKLGCFEIAKMEGFPKD VV  
ILGRSL

sp|O60502.2|GlcNAcase protein@integer

3941149562093479901160527872139939869953948328717448309602075904617639852765945916537772170887588913  
7777607411722225419058993174618449700819722020154518702683811512865205056993341462772546223338698372  
1402964375598157546637116156924279597212332019651564117548382671604539961725503706897735714899984118  
06014935596262763747945043157656864735710326130352938246306646449830806350943992779426250490629331  
05717204828412582988248168711408930001109074970612171397190589610378189764664617608879750771893149325  
703234667852859231589215026272173420215399339907210129333224063854880356386375621030134935672120130  
1182158461415936423120637434347353003426842559697796327071647986461830586067918017754414457439239556  
105794613393123607246824115000032711726489102914303601308009259048205472279983560823820077297582358  
179664677858758610884688114892380227174055044911625343181269465663312612797038045228606892503534909  
04053200471496813891995342035335150071662609939061752070884770494552358391253502368356686068894419  
247600561381489785637526973435030429756393323513367039788577539639326487769367168014443728381308829  
742027488935828240337784123405635224582039088107227197931414720130267879239425482181873925402726847

HexNAcase

>NP\_001305754.1 beta-hexosaminidase subunit alpha isoform 1 precursor [Homo sapiens]  
MTSSRLWFSLLLAFAFRATALWPWPQNFQTSQRYVLYPNNFQFQYDVSSAAQPGCSVLDEAFQRYRD  
LLFGSGSWPRPYLTGWPHQAYPVFLGKRHTLEKNVLVSVVTPGCNQLPTLESVENYTLTINDDQCLLS  
ETVW GALRGLETFSQLVWKS AEGTFFINKTEIEDFPRFPHRGLLLDTSRH YLPLSSILDTLDVMAYNKLN  
VFHWHLVDDPSFPYESFTPELMRKGSYNPVTHIYTAQDVKEVIEYARLRGIRVLA EFDTPGHTLSWGP  
IPGLLTPCYSGSEPSGTFGPVNP SLNNTYEFMSTFFLEVSSVFPDFYLHLGGDEVDFTCWKSNPEIQDFM  
RKKGFGE DFKQLESFYIQTLLDIVSSYGKGYVVWQEVFDNKVKIQPDTIIQVWREDIPVNYMKELELVTK  
AGFRALLSAPWYLN RISYGPDWKDFYIVEPLAFEGTPEQKALVIGGEACMWGEYVDNTNLVPRLWPRAGA  
VAERLWSNKLTSDLTFAYERLSHFRCELLRRGVQAQPLNVGFCEQEFEQT

NP\_001305754.1 HexNAcase protein@integer

2565345204515766711480306911006177947198821532684813532447810423646865764573107554662772808381526600  
642117917154057226835486085540609036627326462010905381722827152902018450244952460665787127848557329  
395690420680769261634047149495299639641226982165094725020646629229474509472842655874683072561673703  
041028821950141244569110159970542189743393425436133850860020642423772374665885844339307607242289110  
53322711285023631702493108185183827081831884917104527304308889323868957612012185652176729187617388637  
972721110986911289289797936885194513457758542882302568596633212217798783484552402181106893688337939  
0864129453307575172981681286669387475194084804497616538662258630500090954851754479511848826716578149  
413425

Human Aquaporin AQP0 - AQP12

AQP0

>NP\_036196.1 lens fiber major intrinsic protein [Homo sapiens]  
MWELRSASFWR AIFAEFFATLFYVFFGLGSSLRWAPGPHLVQVAMAFGLALATLVQSVGHISGAHVNPA  
VTFAFLVGSQMSLLRAFCYMAAQLLGAVAGAAVLYSVTPPAVRGNLALNTLHPAVSVGQATTVEIFLTLQ  
FVLCIFATYDERRNGQLGSVALAVGFSLALGHLFGMYT GAGMNPARSFAPAILTGNFTNHVWVYVWGPII  
GGGLGSLLYDFLLFPR LKSISERLSVLK GAKPDVSNGQPEVTGEPVELNTQAL

AQP0 (NP\_036196.1) protein@integer

111037995035055769383528042980331735926051568239605743221159145388851008050517558019578550099766718  
9264336433815218052099848690552461425019301590017923709478771079030714150560267666491748748143010416  
254905397164557971709156646102387026909394563139392604332326764711391629454573555205664662998074787  
792598295726670518178926139245574738775884427

AQP1

>sp|P29972.3|AQP1\_HUMAN RecName: Full=Aquaporin-1; Short=AQP-1; AltName: Full=Aquaporin-CHIP; AltName:  
Full=Urine water channel; AltName: Full=Water channel protein for red blood cells and kidney proximal tubule  
MASEFKKKLFWRAVVAEFLATTLFVVISIGSALGFKYPVGNNTAVQDQDNVKS LAFGLSIATLAQSVGHI  
SGAHLNPAVTLG LLLSCQISIFRALMYIIAQCVGAI VATAILSGITSSLTGNSLGRNDLADGVNSGQGLG  
IEIIGTLQLVLCVLATTD RRRRDLGGSAPLAIGLSVALGHLLAIDYTGCGINPARSFGSAVITHNFSNH  
WIFWVGPFIGGALAVLIYDFILAPRSSDLTDRVKVWTS GQVEEYDL DADDINSRVEMKPK

AQP1 (P29972.3) protein@integer

6666887219760417187022819852151957153203742336443106809655687871878132608894478649018695455972138883  
39639343332660812503043765593381738742434365146256137672807340479093832524464910544438602313844408  
341237821839449781986241696763570397669739853103526362332569326396054112942869579321782676984689221

95564852078873449448486292552075522054250730062996872

#### AQP2

>NP\_000477.1 aquaporin-2 [Homo sapiens]

MWELRSIAFSRAVFAEFLATLLFVFFGLGSALNWPQALPSVLQIAMAFLGIGTLVQALGHISGAHINPA  
VTVACLVGCHVSVLRAAFYVAAQLLGAVAGAALLHEITPADIRGDLAVNALSNTTAGQAVTVLFLTLQ  
LVLCIFASTDERRGENPGTPALSIGFSVALGHLLGIHYTGCSMNPAPVAVVTGKFDDHWVFWIGPLV  
GAILGSLLYNYVLFPPAKLSERLAVLKGLEPDTDWEEREVRRRQSVLHSPQSLPRGTKA

AQP2 (NP\_000477.1) protein@integer

284257269349285693839963234239768578801304751615519144743271451597725088424519829381134543887694643  
865501804258558819600494462066052036706443844193904529578642240246036176244761688605418638407209974  
1341640172657506767020166308746193003741559270355050969836866505366687894013439119974178206005020051  
1387183678728218452332693795886301912546326300620322241

#### AQP3

>NP\_001305073.1 aquaporin-3 isoform 2 [Homo sapiens]

MGRQKELVSRCEMHLIRYLLRQALAECLGTLILVMFGCGSVAQVVLSRGTHGGFLTINLAFGFAVTLG  
ILIAGQVSGAHLNPAVTFAMCFLAREPWIKLPIYTLAQTLAGFLGAGIVFGLYDAIWHFADNQLFVSGP  
NGTAGIFATYPSGHLDMINGFFDQDRPALVVGHRVPTPGLHCGCLRVPADDRLPPGAAPTLQRGRECEA  
GPCEAQGADLSGQGPSHSAALALSIH

AQP3 (NP\_001305073.1) protein@integer

155084495076116418787770976441926939614892058882457847166868891680037416500434720042440062875895799  
18230682558805622127228555075249365964621107349569832792294108544639592172938692099981513560385581  
938169153175808237870225752655747244880284070296534545803321265082819618098392236447446606939915513  
6052566616975

#### AQP4

>NP\_001641.1 aquaporin-4 isoform M1 [Homo sapiens]

MSDRPTARRWGKCGPLCTRENIMVAFKGVWTQAFWKAVTAEFLAMLIFVLLSLGSTINWGGTEKPLPVD  
VLISLCFGLSIATMVQCFGHISGGHINPAVTVAMVCTRKISIAKSVFYIAAQCLGAIIGAGILYLVTPPS  
VVGGLGVTMVHGNLTAGHGLLVELIITFQLVFTIFASCDSKRTDVTGSIALAIGFSVAIGHLFAINYTGA  
SMNPARSFGPAVIMGNWENHWIYWVGPIIGAVLAGGLYEYVFCPDVEFKRRFKEAFSKAAQTKGSYMEV  
EDNRSQVETDDLILKPGVVHVIDVDRGEEKKGDQSGEVLSSV

AQP4 (NP\_001641.1) protein@integer

12070159722124820628819019501449720865413581317694353998241315021747744525192257186066711458798706689  
77495302963920846559328957938264768822527697873498339460961104074663841542652762500787230122541590  
384445533602203404379061399503257637448883409545857565352790891094589866419574365628926027061136181  
1787359656568794557685227062703420339130301835275594109140347762503013573076067211206370319203586777  
68230715520284202296844

#### AQP5

>NP\_001642.1 aquaporin-5 [Homo sapiens]

MKKEVCSVAFLKAVFAEFLATLIFVFFGLGSALKWPSALPTILQIALAFGLAIGTLAQLGVPVSGGHINP  
AITLALLVGNQISLLRAFFYVAAQLVGAIAGAGILYGVAPLNARGNLAVNALNNNTTQGGQAMVVELILTF  
QLALCIFASTDSRRTSPVGSVALSIGLSVTLGHLVGIYFTGCSMNPARSFGPAVVMNRFSPAHWVFWVGP

IVGAVLAAAILYFYLLFPNSLSLSEVAIIKGTYPEDWEEQREERKKTMELTTR

AQP5 (NP\_001642.1) protein@integer

433722763158508929035738491793919700222859410536375049267195823621645318817517984829534579646987820  
758290558414159020632063201829824335330528097820474603021294015501095092501583785584596179184353918  
278589057719625685453096020974926098248603519506106435408140829339294549696936157109234050140374237  
626578477573932627783952249444096756546926938117

AQP6

>NP\_001643.2 aquaporin-6 [Homo sapiens]

MDAVEPGGRGWASMLACRLWKAISRALFAEFLATGLYVFFGVGSMRWPTALPSVLQIAITFNLVTAMAV  
QVTWKASGAHANPAVTLAFLVGSISHLPRAVAYVAAQLVGATVGAALLYGVMPGDIRETLGINVVRNSVS  
TGQAVAVELLTLQLVLCVFASTDSRQTSGPSATMIGISVALGHLIGHFTGCSMNPARSFGPAIIIGKF  
TVHWVFWVGPLMGALLASLIYNFVLPDTKTLAQLRLAILTGTVEVGTGAGAGAEPLKESQPGSGAVEME  
SV

AQP6 (NP\_001643.2) protein@integer

563493732166859948299063859144595069672296375651404606475203075479277370274245271362739024691151121  
893158283432932673036021495013185276974467823887184282651758201754715412285407045920812850488276398  
587287414461495263777723856014983246440178927672989290106075727568935170809698884036895629544805345  
6237851833565009136173409410919808204643554566647075072235571278997244

AQP7

>NP\_001363120.1 aquaporin-7 isoform 1 [Homo sapiens]

MVQASGHRRSTRGSKMVSWSVIAKIQEILQRKMOVREFLAEFMSTYVMMVFGLGSAHMLNKKYGSYLGV  
NLGFGFGVTMGVHVAGRISGAHMNAAVTFANCALGRVPWRKFPVYVLGQFLGSFLAAATIYSLFYTAILH  
FSGGQLMVTGPVATAGIFATYLPDHMTLWRGFLNEAWLTGMLQLCLFAITDQENNPALPGTEALVIGILV  
VIIGVSLGMNTGYAINPSRDLPRIFFIAGWGKQVFSNGENWWWVPPVAPLLGAYLGGIHYLVFIGSTI  
PREPLKLEDSVAYEDHGITVLPKMGSHIPTISPLTPVSVSPANRSSVHPAPPLHESMALEHF

AQP7 (NP\_001363120.1) protein@integer

637325536808418035364782200803797782150449656810281911066755126876023138472436508389661772435133649  
2201835360401268608752638091128718582145558250634761424597390810514494418438831490132256626716267985  
9176775567602997070820261517796346422020154573825547565525559107741555537903114128319414496633315184  
536878212136738757264472612560920760843180143150679171437889035823904978597581587468720982249560220  
05412093032063843996689263649018505104813021516

AQP8

>NP\_001160.2 aquaporin-8 [Homo sapiens]

MSGEIAMCEPEFGNDKAREPSVGGRWRVSWYERFVQCLVELLGSALFIFIGCLSVIENGTDGTLQPAL  
AHGLALGLVIATLGNISGGHFNPVSLAAMLIGLNLVMLLPYVWSQLLGGMLGAALAKAVSPEERFWNA  
SGAAFVTVQEQQVAGALVAEIIITLLALAVCMGAINKTKGPLAPFSIGFAVTVDILAGGPVSGGCMN  
PARAFGPAVVANHWNFHWIYWLGPLLAGLLVGLLIRCFIGDGKTRLILKAR

AQP8 (NP\_001160.2) protein@integer

2612579955380142218745710898134712385719666605469470378281282211752020296918638454694551212858876219  
6235516946606609043322143186076393413250661249912467117925928456323441395861321310352693131395259492  
02442842191150471276872913802231482153494179294117620921331314196262142620436590279154962351958576032  
125897729684958706969509990714711740837

AQP9

>NP\_066190.2 aquaporin-9 isoform 1 [Homo sapiens]

MQPEGAEKGKSFQRLVLKSSLAKETLSEFLGTFILIVLGCVCVAQAILSRGRFGGVITINVGFSMAVAM  
AIYVAGGVSGGHINPAVSLAMCLFGRMKWFKLPFYVGAQFLGAFVGAATVFGIYYDGLMSFAGGKLLIVG  
ENATAHIFATYPAPYLSLANAFADQVVATMILLIIVFAIFDSRNLGAPRGLEPIAIGLLIIVASSLGLN  
SGCAMNPARDLSPRLFTALAGWGFVFRAGNNFWWIPVVGPLVGA VIGGLIYVLVIEIHHPEPDSVFKTE  
QSEDKPEKYELSVIM

AQP9 (NP\_066190.2) protein@integer

463397851225672165813437897990257360505365795687815271562912638383187438987954802339137537681378166  
86329963010258995305275532700705944015598577873644664828121982194345797326800966241589573813529564  
464423502013068132961507219616742716376526724937839663320863240345828574725732940356603617244520717  
3434650553773542855735542825661508791690751217565486289894548674082031425660765354737774

AQP10

>NP\_536354.2 aquaporin-10 [Homo sapiens]

MVFTQAPAEIMGHLRIRSLARQCLAEFLGVFVLMMLTQGAVAQAVTSGETKGNFFTMFLAGSLAVTIAI  
YVGGNVSGAHLNPAFSLAMCIVGRPLWVKLPYIYLQLLSAFCASGATYVLYHDALQNYTGGNLTVTGPK  
ETASIFATYPAPYLSLNNGLDQVLGTGMLIVGLLAILDRRNKGVPAGLEPVVVGMLLALGLSMGANCG  
IPLNPARDLGPRLFTYVAGWGPVFSAGNGWWWVPVAPLVGATVGTATYQLLVALHHPEGPPEPAQDLVS  
AQHKASELETPASAQMLECKL

AQP10 (NP\_536354.2) protein@integer

290088062314319559544871954566394232198816315466048037555728743944798933889456254176315664124430058  
278481410939569922240876118980416595571255403775232065563945904400364593357380479609416724197912212  
046352439539299064352662332539429986288100103961805961474419963229208831909731776055905044249791711  
88128174705292813588412650823126828536584648445982917866529514151754238683469222943285230026647

AQP11

>NP\_766627.1 aquaporin-11 isoform 1 [Homo sapiens]

MSPLLGLRSELQDTCTSLGLMLSVLLMGLARVVARQQLHRPVAHAHVLEFLATFQLCCCTHELQLLSEQ  
HPAHPTWTLTLVYFFSLVHGLTLVGTSSNPCGVMMQMMGLGMSPETGAVRLLAQLVSALCSRYCTSALWS  
LGLTQYHVSERSFACKNPIRVDLLKAVITEAVCSFLFHSALLHFQEVRTKLRIHLLAALITFLVYAGGSL  
TGAVFNPALALSLHFMCFDEAFPQFFIVYWLAPSLGILLMILMFSFFLPWLHNNHTINKKE

AQP11 (NP\_766627.1) protein@integer

2676561177988498082119368854665787972283642740217080367323401800441415863553706172032660786801569913  
3141147879483403639619385931829555650616002214961779877026804719157038527790304606347877375081350266  
831912136044120766726341366025967298856316992982062240692178400945098648941455532442983826062160908  
761889380998638099162394954394450553012098802897357053

AQP12

>NP\_945349.1 aquaporin-12A precursor [Homo sapiens]

MAGLNVLSFFFATFALCEAARRASKALLPVGAYEVFAREAMRTLVELGPWAGDFGPDLLLTLFLFLA  
HGVTLDGASANPTVSLQEFLMAEQSLPGTLLKLAQGLGMQAACTLMRLCWAWELSDHLHLQSLMAQSCS  
SALRTSVPHGALVEAACAFCHLTLHLRHSPAYS GPAVALLVTAYTAGPFTSAFFNPALAAASVTA  
CSGHTLLEYVQVYWLGPLTGMVLAVLLHQGRPLHLFQRNLFYQKKNKYRAPRGKPAPASGDTQTPAKGSS

VREPGRSGVEGPHSS

AQP12 (NP\_945349.1) protein@integer

44722273560267218187611263437137040741848710218546225908817154728625918913829949902141552515486642142  
786376712999040774269547776710304641596473364939674023865399301172548895858927766531980089331904087  
10113212379958172581838411996833052778813758201516724357146127554946132896397983507417506364773746201  
18367792982296982508716043438199190790634034053535198819446162792256111155497630042

VOCs

Alpha

MZ314997.2

>QWA53375.2 surface glycoprotein [Severe acute respiratory syndrome coronavirus 2]

MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAISG  
TNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLIVNNATNVVIKVFCEQFCNDPFLG  
VYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFNIDGYFKIYSKHTPINLV  
RDLPPQGFSALEPLVDLPIGINITRFQTLALHRSYLTGDSGWTAGAAAYVGYLQPRTFLLKYNENG  
TITDAVDCALDPLSEKTKLSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAW  
NRKRISNCVADYSVLYNSASFSTFKCYGVSPKLNLDLCTNVYADSFVIRGDEVQRQIAPGQTGKIADYNY  
KLPDDFTGCVIAWNSNNLDSKVGNYNYLYRFRKSNLKPFERDISTEIQAGSTPCNGVEGFNCYFPLQ  
SYGFQPTYGVGYQPYRVVLSFELLHAPATVCGPKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQ  
QFGRDIDDDTTDAVRDPQTLEILDITPCSFGGVSVITPGTNTSNQVAVLYQGVNCTEVPVAIHADQLTPTW  
RVYSTGSNVFQTRAGCLIGAEHVNSYECDIPIGAGICASYQTQTNHRRARSVASQSIIAYTMSLGAEN  
SVAYSNSIAIPINFITISVTTEILPVSMTKTSVDCTMYICGDSTECNLLLQYGSFCTQLNRALTGIAVE  
QDKNTQEVEFAQVKQIYKTPPIKDFGGFNFSQILPDPSPKSKRSFIEDLLFNKVTADAGFIKQYGDCLGD  
IAARDLCAQKFNGLTVLPLLDEMIQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQ  
NVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILAR  
LDKVEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSPF  
QSAPHGVVFLHVTVPAQEKNFTTAPAICHGDKAHFPREGVFSNGTHWFVTQRNFYEPQIITHTNTFVS  
GNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVLGDISGINASVVNIQKEIDRLNEVAKNL  
NESLIDLQELGKYEQYIKWPWYIWLGFIAGLIAVMVTIMLCCMTSCCCKGCCSCGSCCKFDEDDSE  
VLKGVKLHYT

S (QWA53375.2) protein@integer

150549571504853468105487583665731146123324002820617303123515742916663568104679107972126274184556  
775884808689963618511045166966530769527155935549737165265168758440118680647078765683705402638344  
668093591356030994356599927194248395746390459701421936838579811821540197532576503125078315952563  
463197006737381012158729125832652423194999564163518161471601448573721689925566807240169801402143  
422139080975763924325906504310655491918046103979399478557836534791713138969267258813474961865485  
004246038976474935285870968894582050754448058187938392294057785010339197063115472865172486992806  
605551384631607336572646138094059110998262826782605727083458483862664802774969440603309561531355  
336346212002276440011404496918055478398205787769031117119238565898929812105838007238481284245010  
763050107347330239947404510614816642963267880462935465235905374431302334105482948586536008136290  
506930056796515286988247237162485462155116904132044798670210461257032297917051542059251983293688  
330778198954350535759872026359908782465081302365859870621550178415997880235685450638951991580949

319046701383232583726565938273037095666693704079718907833447903139579209899904947282707605097760  
158068962516536062696973197399258956901272045121486065602436111517593143225897940359242039861646  
567155326776667273666465287506793480160909175243242163565742347611028717554323904932031268069597  
224873655949201801755265888859322820276949815683198394191112534968621593228063000957552348628988  
572256966933880659835186017635028246017380743325972819220954437517764685581740837095793751836173  
585377462541545883520956308804766869020506611331419403842406615974777549999137979797714296775069  
319998087362574782365

>QYC97631.1 truncated ORF8 protein [Severe acute respiratory syndrome coronavirus 2]  
MKFLVFLGIITTVAAAFHQECSLQSCT

ORF8 (QYC97631.1) protein@integer  
4912672549220032844152618631608925

Beta  
MZ314998.1

>QWA53387.1 surface glycoprotein [Severe acute respiratory syndrome coronavirus 2]  
MFVFLVLLPLVSSQCVNLTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFSNVTWFXAIHV  
SGTNGTKRFANPVLFPNDGVYFASTEKSNIIRGWIFGTTLDSTQSLIVNATNVVIKVFCEQFCNDPF  
LGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFNIDGYFKIYSKHTPI  
NLVRGLPQGFSALEPLVDLPIGINITRFQXXXLHRSYLTGPDSSSGWTAGAAAYVGYLQPRTFLLKYNE  
NGTITDAVDCALDPLSEKTKLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASVY  
AWNRKRISNCVADYSVLYNSASFSTFKCYGVSPKLNLDLCFTNVYADSFVIRGDEVQRQIAPGQTGNIADY  
NYKLPDDFTGCVIAWNSNNLDSKVGGNYNLYRFLRKSNNLKPFERDISTEIQAGSTPCNGVKGFNCYFP  
LQSYGFQPTYGVGYQPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLP  
FQQFGRDIADTTDAVRDPQTEILDITPCFSGGVSVITPGTNTSNQVAVLYQGVNCTEVPVAIHADQLTP  
TWRVYSTGSNVFQTRAGCLIGAEHVNSYECDIPIGAGICASYQTQTNSPRRARSVASQSIIAYTMSLGV  
ENSVAYSNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIA  
VEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSPKRSFIEDLLFNKVTLADAGFIKQYGDCL  
GDIAARDLICAQKFNGLTVLPPLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIGV  
TQNVLYENQKLIANQFNSAIGKIQDLSSTASALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDIL  
SRLDKVEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMS  
FPQSAPHGVVFLHVTVPAQEKNFHTTAPAICHGDKAHFPREGVSVNGTHWFVTQRNFYEPQIITDNTF  
VSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVLGDISGINASVVNIQKEIDRLNEVAK  
NLNESLIDLQELGKYEQYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTSCCCLKGCCSCGSCCKFDEDDSD  
EPVLKGVKLHYT

S (QWA53387.1) protein@integer  
602198286019413872421950334662924584493296011282469212494062971666654272418716431888354871823538  
534054472672677906566129036259896758818242859904900195927055006565373804707800102421072881571903  
621460370527542155795332172966201756877959201280335937896187933618882118250682714546719718105687  
063057935977359271434622238398434984233296953505778722904709653077556826777971090017714585266156  
644937961927648534078312456703463826519522407690981445722982746232658476774187705644423914226111  
856848963922125708663263668420988508943509155036940551323850088326924991676119672392525457405138

900104816616869285887101109712181139400276997719823908479884827354112620324422816512053380848737  
550557686160690788941460982757510308292775377257138293161065803261259915538912906146135206761650  
360680441701975747520834712232494281875989898882849929679951146780122888495691142665423897899658  
876348824101213453267936060083602894741188085696062157055345647765356944663069801108762887279283  
412214449812459697572477397990330934337049840631996702269194963599075752858781841575675546951407  
750201757689243337517818617895633448149926527550252533516645023326735839847915652332168156416252  
910340026457700315157115838814331042035272194377704963897778428149029161436704574787263778634507  
448522842667637866344466376606994302723699978080237984494509463484678292969031129188933143797121  
558276725545999608425422174916512443119705093117430705780055949794549164900801532083757806381715  
611147057671514803025055260953072809902113830158197179127100028423496930785339353270002937877736  
603940066473413633530094087626990432810578130271046147850585559644079998620767676342874840110911  
99693978011965178400000

>QWA53394.1 ORF8 protein [Severe acute respiratory syndrome coronavirus 2]  
MKFLVFLGIITTVAAAFHQECSLQSCTQHQPYYVDDPCPIHFYSKWYIRVGARKSAPLIELCVDEAGSKSP  
IQYIDIGNYTVSCLPFTINCQEPKLGSLVVRCSFYEDFLEYHDVRRVLDLFI

ORF8 (QWA53394.1) protein@integer

194611009554485015080235775478493357293746189205074064273910773374230324575542845088267908932795322  
50367336043068643456811124221825233150723390246662221500328

Gamma

MZ427312.1

>QWW27582.1 surface glycoprotein [Severe acute respiratory syndrome coronavirus 2]  
MFVFLVLLPLVSSQCVNFTNRTQLPSAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIHV  
SGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSTKQSLIVNNAATNVVIKVECFQFCNYPF  
LGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNFKNLSEFVFNIDGYFKIYSKHTPI  
NLVRDLPQGFSALEPLVDLPIGINITRFQTLALHRSYLTGDSGGWTAGAAAYVGYLQPRTFLLKYN  
ENGTITDAVDCALDPLSETKCTLSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASV  
YAWNRRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLNLDLCTNRYADSFVIRGDEVQRQIAPGQTGXIA  
YNYKLPDDFTGCVIAWNSNNLDSKVGNYNYLRLFRKSNLKPFRDISTEIYQAGSTPCNGVKGFNCYF  
PLQSYGFQPTYGVGYQPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFL  
PFQQFGRDIADTTDAVRDPQTLTILEITPCSFSGVSVITPGTNTSNQVAVLYQGVNCTEVPVAIHADQLT  
PTWRVYSTGSNVFQTRAGCLIGAEVNNNSYECDIPIGAGICASYQTQTNPRRARSVASQSIIAYTMSLG  
AENSVAYSNNIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECNLLLQYGSFCTQLNRALTGI  
AVEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSPKPSKRSFIEDLLFNKVTLADAGFIKQYDC  
LGDIAARDLICAQKFNGLTVLPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIG  
VTQNVLYENQKLIANQFNSAIGKIQDSLSTASALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDI  
LSRLDKVEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAAIKMSECVLGQSKRVDFCGKGYHLM  
SFPQSAPHGVVFLHVTVPAQEKNTTAPAICHGKAHFPREGVFSVNGTHWFVTQRNFYEPQIITTDNT  
FVSGNCDVVIGVNNNTVYDPLQPELDSFKEELDKYFKNHTSPDVLGDISGINASFVNIQKEIDRLNEVA  
KNLNESLIDLQELGKYEYIKWPWYIWLGFIAGLIAIVMTIMLCCMTSCCSCLKGCCSCGSCCKFDEDD  
SEPVKGVKLHYT



S (QWW27582.1) protein@integer

120439657203882774484395656386107689668341418444046080766830520302291170715686320987651844305859028  
150524466997826169466976022222602495852218574674887192544603789810881629468446839524213417476600501  
2540959582370102712674511109586688633860658119997817780908334451999186276976567616465231757189062024  
7020968325589955103272977798042724357964590857315091116657802134214793875519156773501060617175298699  
7268751010294811526361853966134403580382759315228379844610878139157907603922979858169672142031973078  
950606627305608063803745996436792849120129877059541121172652466988700542510396007346732501597919651  
2350446742309574539215175111072956734654571011103018178855413900615320472727941725829424810957226589  
84477331990585527999720793124989015510794643254026409969435783344924505494723536575190636448836908  
414910290617852344987373603562412099938933475153610619620454955435604808682293095912243251309210960  
305734271890701245402436581398469307149640986771894348396259674394382104492380037548829756937209592  
782547711950078799478739224672763941148405366286974131595811452872938693955324357598595679673272369  
5418101874081594378156668958062791584197998098945654152101955203161379250330721253939463947985179138  
025440902429721312048722230351862864517958807184840797232931343106535533345473329305829707587428784  
0462497260029748061813682935584786621128891216463570262130153040589317278479124185547246204079157955  
00631780847482190539372431864561260019151046972579771445139338220893935037813190761913410123776422876  
9774647390887503552937116348167419162553319035401763696998418792832191261760953373804101322266283880  
76848132319495550999982759595954285935501386399961747251495647300

>QWW27589.1 ORF8 protein [Severe acute respiratory syndrome coronavirus 2]

MKFLVFLGIITVAAAFHQECSLQSQCTQHQPYYVDDPCPIHFYSKWYIRVVGARKSAPLIELCVDEAGSKSP  
IQYIDIGNYTVSCLPFTINCQKPKLGLSVVRCFSFYEDFLEYHDVRRVLDLFI

ORF8 (QWW27589.1) protein@integer

194611009554485015080235775478493357293746189205074064273910773374230324575542845088267908932795322  
50367336043068643456757437130625233150723390246662221500328

Delta

MZ724507.1

>QYM89633.1 surface glycoprotein [Severe acute respiratory syndrome coronavirus 2]

MFVFLVLLPLVSSQCVNLTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIHV  
SGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLIVNNAATNVVIKVCFCNDPFC  
LDVYYHKNNKSWMKSEFRVYSSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFNIDGYFKIYSKHTPI  
NLVRDLPHGFSALEPLVDLPIGINITRFQTLALHRSYLTGPDSSSGWTAGAAAYVGYLQPRFTLLKYN  
ENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASV  
YAWNRRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLNLCFTNVYADSFVIRGDEVQRQIAPGQTGKIAD  
YNYKLPDDFTGCVIAWNSNNLDSKVGGNYNRYRFLFRKSNLKPFRDISTEIQAGSTPCNGVQGFNCFY  
PLQSYGFQPTNGVGYQPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFL  
PFQQFGRDIADTTDAVRDPQTLEILDITPCSFSGVSVITPGTNTSNQVAVLYQGVNCTEVPVAIHADQLT  
PTWRVYSTGSNVFQTRAGCLIGAEHVNSYECDIPIGAGICASYQTQTSNRRRARSVASQSIAYTMSLG  
AENSVAYSNNIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECNLLLQYGSFCTQLNRALTGI  
AVEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQYGDC  
LGDIAARDLICAQKFNGLTVLPLLTDEMAIQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIG

VTQNVLYENQKLIANQFNSAIGKIQDSLSTASALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDI  
LSRLDKVEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLM  
SFPQSAPHGVVFLHVTVVPAHEKNFTTAPAICHGDKAHFPREGVFSVNGTDWVFTQRNFYEPQIITTDNT  
FVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRLNEVA  
KNLNESLIDLQELGKYEQYIKWPWYIWLGFIAGLIAIVMTIMLCCMTSCCSCLKGCCSCGSCCKFDEDD  
SEPVKGVKLHYT

S (QYM89633.1) protein@integer

120439657203882774484390066932584916898659202256493842498812594333330854483743286377701022981821  
031063561692048351038006678781157706612192808938101284633541757413763423983981946276212510777935  
534339307250415435125851637754188336581367392850323452139782909256890507013671415649895080244227  
753892453255854226036164684746090924262825067480191810544515793202652778418327264373501335424897  
507486307101064484516216011846993518476845229353148389231843118778962105966301237713197273396673  
154650698605869838224010935738274836873907675590238256662954469282435478525221952968070707021215  
722665407912635875443810269828868916845361293005858032342051206530502130813189065323282569390479  
409031925332725967612179379967844880150021046618046825251152215146787742244300744438609037227028  
418129997339883952103529953511075753005955010287108852306279055182505635594238449552295116081010  
878958732822289175062067755878676934527599568628837221876568738111065301375865715470764155593739  
826169630629152178391172773668408354107943314647140508440741430194487431149032503484864959664092  
695723478697766217879929783983663618477090509370407971890783344790313957920989990494728270760509  
776015806896251653606269697319739925895690127204512148606560243611151759314322589794035924203986  
164656715532677666727366646529148537937143920231248630014874030906841467792393310564445574308321  
487320097392295346624975526588885932282027694981594030848127255097905533552800264901736080551209  
891005698753895639095645712823400311086902138882114384887323695443751776468558174083709579375183  
6173585377462541545883520956308804766869020506611331419403842406615974777549999137979771429677  
5069319998087362574782365

>QYM89640.1 ORF8 protein [Severe acute respiratory syndrome coronavirus 2]

MKFLVFLGIITVAAAFHQECSLQSQCTQHQPYYVDDPCPIHFYSKWYIRVGARKSAPLIELCVDEAGSKLP  
IQYIDIGNYTVSCLPFTINCQEPKLGSLVVRCSFYEDFLEYHDVRRVVLDFI

ORF8 (QYM89640.1) protein@integer

1946110095544850150802357754784933572937461892050740642739107733742303245755428450882679111845951361  
8892136043068643456811124221825233150723390246662221500328

Omicron

OQ151515.1

>WBE23649.1 surface glycoprotein [Severe acute respiratory syndrome coronavirus 2]

MFVFLVLLPLVSSQCVNLITRTQSYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAISGTNG  
TKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLIVNNAATNVVIVKVEFCQFCNDPFLDVYY  
HKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNFKNREFVFNIDGYFKIYSKHTPINLGRD  
LPQGFSALEPLVDLPIGINITRFQTLALHRSYLTGPDSSSGWTAGAAAYVGYLQPRTFLLKYNENGTI  
TDAVDCALDPLSETKCTKLSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFDEVFNATTFASVYAWNR  
KRISNCVADYSVLYNFAPFFAFKCYGVSPTKLDLFCFTNVYADSFVIRGNEVSQIAPGQTGNIADYNYKL  
PDDFTGCVIAWNSNKLDSTVGGNYRYRRLFRKSKLKPFERDISTEIQAGNKPCNGVAGVNCYFPLQSY

GFRPTYGVGHQPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQF  
GRDIADTTDAVRDPQTLEILDITPCSFSGGVSIVITPGTNTSNQVAVLYQGVNCTEVPVAIHADQLTPTWRV  
YSTGSNVFQTRAGCLIGAEYVNNSECDIPIGAGICASYQTQTKSHRRARSVASQSIAYTMSLGAENSV  
AYSNNIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLQYGSFCTQLKRALTGIAVEQD  
KNTQEVFAQVKQIYKTPPIKYFGGFNFSQILPDPSPKRSFIEDLLFNKVTLADAGFIKQYGDCLGDIA  
ARDLICAQKFNGLTVLPLLTDemiaQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQNV  
LYENQKLIANQFNSAIGKIQDLSSTASALGKLQDVVNHNAQALNTLVKQLSSKFGAISSVLNDILSRD  
KVEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSPQ  
APHGVVFLHVTVPAQEKNFTTAPAICHGDKAHFPREGVVFVSNGTHWFVTQRNFYEPQIITDNTFVSGN  
CDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVLGDISGINASVVNIQKEIDRLNEVAKNLNE  
SLIDLQELGKYEQYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTSCCSCLKGCCSCGSCCKFDEDDSEPV  
L  
KGVKLHYT

S (WBE23649.1) protein@integer

376373928762133670263719249958465545828514680166800659114254914881567080025955061053595967223533  
095174984722289315381625923428720430113735781163447970309617070644064250251730920330901229955549  
907789093284529474361641318329898881519884217698018728495002838227195093015096300043836317551758  
788950535612778722034804606862461192315962215677876945612164045438102549625523109554644190170306  
384292235342207016930988798873096461076593319125635347478872831418472231151700872255233495589932  
872756309817610697145945148133281031136684550143383534003640950984560008550860971729578790551413  
174005965382031887484299623882841208788030852850315478661233479804291398961455885338952921269470  
324235363992021975135253784180597226066743205983149794664744647063600116431270930765173749561217  
78383004285380887878597703267802603227633657477027407189003526498304374338914595213225478046565  
775331728753648540233194362377515517160419329629033944735961302692477555348876739973877527061108  
357320617363860139300473491147838391201907224699921787001870430782264991473554959459560878685244  
811271725939027510790989045922466240913757811765580469338814690639074862326611941982611720655606  
044641804821368987149197399258956901272045121486065602436111517593143225897940359284759602364985  
356975654878665698060904948323749008759228731837134144352984332836340082450407073664702182903941  
453922953466249755265888859322820276949815683198394191112534968621593228063000957552348628988572  
256966911044696751890659538095670506188821143848873236954437517764685581740837095793751836173585  
377462541545883520956308804766869020506611331419403842406615974777549999137979797714296775069319  
998087362574782365

>WBE23656.1 ORF8 protein [Severe acute respiratory syndrome coronavirus 2]

MKFLVFLGIITTVAAAFHQECSLQSQCTQHQPYYVDDPCPIHFYSKWYIRVVGARKSAPLIELCVDEAGSKSP  
IQYIDIGNYTVSCLPFTINCQEPKLGSLVVRCSFYEDFLEYHDVRRVLDLFI  
ORF8 (WBE23656.1) protein@integer

194611009554485015080235775478493357293746189205074064273910773374230324575542845088267908932795  
32250367336043068643456811124221825233150723390246662221500328

Omicron

OQ151513.1

>WBE23626.1 surface glycoprotein [Severe acute respiratory syndrome coronavirus 2]

MFVFLVLLPLVSSQCVNLITRTQSYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFSNVTFWHAISGTNG  
TKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLIVNNAATNVVIVKVEFCQFCNDPFLDVYH

KNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFNIDGYFKIYSKHTPINLGRDL  
PQGFSALEPLVDLPIGINITRFQTLALHRSYLTPGDSSSGWTAGAAAYVGYLQPRTFLLKYNENGTIT  
DAVDCALDPLSETKCTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFDEVFNATTFASVYAWNRK  
RISNCVADYSVLNFAPFFAFKCYGVSPTKLNLDLFTNVYADSFVIRGNEVSQIAPGQTGNIADYNYKLP  
DDFTGCVIAWNSNKLDSTVGGNYNYRYRLFRKSKLKPFERDISTEIQAGNKPCNGVAGVNCYFPLQSYG  
FRPTYGVGHQPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNGLTGTGVLTESNKKFLPFQFG  
RDIADTTDAVRDPQTLILDITPCSFGGVSITPGTNTSNQVAVLYQGVNCTEVPVAIHADQLTPTWRVY  
STGSNVFQTRAGCLIGAEYVNNSECDIPIGAGICASYQTQTKSHRRARSVASQSIIAYTMSLGAENSVA  
YSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLLQYGSFCTQLKRALTGIAVEQDK  
NTQEVFAQVKQIYKTPPIKYFGGFNFSQILPDPSPKSKRSFIEDLLFNKVTLADAGFIKQYGDCLGDIAA  
RDLICAQKFNGLTVLPLLTDEMAIQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIGVTVQNVL  
YENQKLIANQFNSAIGKIQDLSSTASALGKLQDVVNHNAQALNTLVKQLSSKFGAISSVLNDILSRDK  
VEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQSA  
PHGVVFLHVTVPAQEKNTTAPAICHGKAHFPREGVFSNGTHWFVTQRNFYEPQIITDNTFVSGNC  
DVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVLGDISGINASVVNIQKEIDRLNEVAKNLNES  
LIDLQELGKYEQYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTSCCCLKGCCSCGSCCKFDEDDSEPVLK  
GVKLHYT

S (WBE23626.1) protein@integer

188186964381066835131859624979232772914257340083400329557127457440783540012977530526797983611766  
547587492361144657690812961714360215056867890581723985154808535322032125125865460165444858639767  
660561757998336009789392030938318395174223245764345995411697719567893222538614170863203269615457  
824028888238845893264829338648877135547854913593695747734452704157451945126385367238768130616664  
500951236234752087141965929390166132369345077926227567424170663476513904583382960648835071790596  
881076734930230533358309642755893107194045501433835340036409509845600085508609717295787905514131  
740059653820318874842996238828412087880308528503154786612334798042913989614558853389529212694703  
242353639920219751352537841805972260667432059831497946647446470636001164312709307651737495612177  
838300428538088787785977032678026032276336574770274071890035264983043743389145952132254780465657  
753317287536485402331943623775155171604193296290339447359613026924775553488767399738775270611083  
573206173638601393004734911478383912019072246999217870018704307822649914735549594595608786852448  
112717259390275107909890459224662409137578117655804693388146906390748623266119419826117206556060  
446418048213689871491973992589569012720451214860656024361115175931432258979403592847596023649853  
569756548786656980609049483237490087592287318371341443529843328363400824504070736647021829039414  
539229534662497552658888593228202769498156831983941911125349686215932280630009575523486289885722  
569669110446967518906595380956705061888211438488732369544375177646855817408370957937518361735853  
774625415458835209563088047668690205066113314194038424066159747775499991379797977142967750693199  
98087362574782365

>WBE23633.1 ORF8 protein [Severe acute respiratory syndrome coronavirus 2]

MKFLVFLGIITVAAFHQECSLQSCTQHQPYYVDDPCPIHFYSKWYIRVGARKSAPLIELCVDEAGSKSP  
IQYIDIGNYTVSCLPFTINCQEPKLGSLVVRCSFYEDFLEYHDVRRVLDLFI

ORF8 (WBE23633.1) protein@integer

194611009554485015080235775478493357293746189205074064273910773374230324575542845088267908932795322  
50367336043068643456811124221825233150723390246662221500328

Omicron

OQ151397.1

>WBE22247.1 surface glycoprotein [Severe acute respiratory syndrome coronavirus 2]

MFVFLVLLPLVSSQCVNLITRTQSYTNSFTRGVVYYPDKVFRSSVLHSTQDLFLPFFSNVTFWHAISGTNG  
TKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLIVN NATNVVIVKVFCEQFCNDPFLDVYH  
KNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFNIDGYFKIYSKHTPINLGRDL  
PQGFSALEPLVDLPIGINITRFQTLALHRSYLT PGDSSSGWTAGAAAYVGYLQPRTFLLKYNENGTIT  
DAVDCALDPLSETKCTLSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFDEVFNATTFASVYAWNRK  
RISNCVADYSVLYNFAPFFAFKCYGVSPTKLNLCFTNVYADSFVIRGNEVSQIAPGQTGNIADYNYKLP  
DDFTGCVIAWNSNKLDSNVGGNYNYRYLFRKSKLPPFERDISTEIQAGNKPCNGVAGVNCYFPLQSYG  
FRPTYGVGHQPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNGLTGTGVLTESNKKFLPFQQFG  
RDIADTTDAVRDPQTL EILDITPCSFGGVSVITPGTNTSNQVAVLYQGVNCTEVPVAIHADQLTPTWRVY  
STGSNVFQTRAGCLIGAEYVNSSYECDIPIGAGICASYQTQTKSHRRARSVASQSIAYTMSLGAENSV  
YSNNSIAIPTNFTISVTEILPVSMTKTSVDCTMYICGDSTECNLLLQYGSFCTQLKRALTGIAVEQDK  
NTQEVFAQVKQIYKTPPIKYFGGFNSQILPDPSPKSKRSFIEDLLFNKVTLADAGFIKQYGDCLGDIAA  
RDLICAQKFNGLTVLPPLLTDEMIQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVL  
YENQKLIANQFNSAIGKIQDLSLSTASALGKLQDVVNHNAQALNTLVKQLSSKFGAISSVLNDILSRDLK  
VEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQSA  
PHGVVFLHVTYVPAQEKNFTTAPAICHGKAHFPREGVFSVNGTHWFVTQRNFYEPQIITDNTFVSGNC  
DVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRLNEVAKNLNES  
LIDLQELGKYEQYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTSCCCLKGCSCGSCCKFDEDDSEPVLK  
GVKLHYT

S (WBE22247.1) protein@integer

188186964381066835131859624979232772914257340083400329557127457440783540012977530526797983611766  
547587492361144657690812961714360215056867890581723985154808535322032125125865460165444858639767  
660561757998336009789392030938318395174223245764345995411697719567893222538614170863203269615457  
824028888238845893264829338648877135547854913593695747734452704157451945126385367238768130616664  
500951236234752087141965929390166132369345077926227567424170663476513904583382960648835071790596  
881076734930230533358309642755893107194045501433835340036409509845600085508609717295787906946076  
935983401880539814039679980752948174838917061129626859097450351243181634822526117602820444113474  
710329100870205755072380879659641095689683361909726927673090186034961769426049370725686434424429  
838712015800462584313657369557206800215691898485038871890035264983043743389145942614261949189191  
344799866875065164530197315105486759540459257333788215566906554734889761448174490939588773826690  
035907978627899662268221181259785892521826900165761496648457822774081860975549594595608786852448  
112717259390275107909890459224662409137578117655804693388146906390748623266119419826117206556060  
446418048213689871491973992589569012720451214860656024361115175931432258979403592847596023649853  
569756548786656980609049483237490087592287318371341443529843328363400824504070736647021829039414  
539229534662497552658888593228202769498156831983941911125349686215932280630009575523486289885722  
569669110446967518906595380956705061888211438488732369544375177646855817408370957937518361735853  
774625415458835209563088047668690205066113314194038424066159747775499991379797977142967750693199  
98087362574782365

>WBE22254.1 ORF8 protein [Severe acute respiratory syndrome coronavirus 2]

MKFLVFLGIITTVAAAFHQECSLQSCTQHQPYYVDDPCPIHFYSKWYIRVGARKSAPLIELCVDEAGSKSP  
IQYIDIGNYTVSCLPFTINCQEPKLGSLVVRCSFYEDFLEYHDVRRVLDLFI

ORF8 (WBE22254.1) protein@integer

194611009554485015080235775478493357293746189205074064273910773374230324575542845088267908932795322  
50367336043068643456811124221825233150723390246662221500328

Omicron

OQ151387.1

>WBE22128.1 surface glycoprotein [Severe acute respiratory syndrome coronavirus 2]

MFVFLVLLPLVSSQCVNLITRTQSYTNSFTRGVVYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIHVSGT  
NGTKRFDNPVLPFNDGVYFASTEKSNIRGWIFGTTLDSTQSLIVNNAATNVVIKVFCEFCNDPFLDV  
YHENNKSMESELRVYSSANNCTFEYVSQPFLMDLEGGKQGNFKNLREFVFNIDGYFKIYKHTPVNLG  
RDLPPQGFSALEPLVDLPIGINITRFQTLALHRSYLTGPDSSSWTAGAAAAYVGYLQPRFTLLKYNENG  
TITDAVDCALDPLSETKCTLSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFHEVFNATTFASVYAW  
NRTRISNCVADYSVLYNFAPFFAFKCYGVSPTKLDLCFTNVYADSFVIRGNEVSQIAPGQTGNADYNY  
KLPDDFTGCVIAWNSNKLDKSVGNYNLYRLFRKSKLKPFERDISTEIQAGNKPCNGVAGFNCYSPLQ  
SYGFRPTYGVGHQPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQ  
QFGRDIADTTDAVRDPQTLEILDITPCSFSGGVSVITPGTNTSNQVAVLYQGVNCTEVPVAIHADQLTPTW  
RVYSTGSNVFQTRAGCLIGAEVNNSECDIPIGAGICASYQTQTKSHRRARSVASQSIIAYTMSLGAEN  
SVAYSNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLQYGSFCTQLKRALTGIAVE  
QDKNTQEVFAQVKQIYKTPPIKYFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTADAGFIKQYGDCLGD  
IAARDLCAQKFNGLTVLPLLTDEMIQYTSALLAGTITSGWTFGAGAALQIPFAMQMAFRFNGIGVTD  
NVLYENQKLIANQFNSAIGKIQDLSSTASALGKLQDVVNHNAQALNTLVKQLSSKFGAISSVLNDILSR  
LDKVEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSPF  
QSAPHGVVFLHVTYVPAQEKNFHTTAPAICHGDKAHFPREGVFVSNNGTHWFVTQRNFYEPQIITDNTFVS  
GNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVLGDISGINASVUNIQKEIDRLNEVAKNL  
NESLIDLQELGKYEQYIKWPWYIWLGFIAGLIAIVMTIMLCCMTSCCSCLKGCCSCGSCCKFDEDDSEP  
VLKGVKLHYT

S (WBE22128.1) protein@integer

150549571504853468105487699983386218331405872066720263645701965952626832010382024421474728645516795  
21739466347801785810244870239807672609929558090340219135344494242036329002966049756560430948779039  
616807215875381961385924535332057205419545600024330111187680955909502146321202728564742500084075641  
8697286264339385433361592916893438429912402279890290451679894536679886500423986481286696965391129  
4108379915914791490235999264215924947578250045059681650027354407215685127521169316478523294096766619  
656175081133362910618789709705196574439599075288988742494794380831030905245009647280913559320673554  
270215000430349707852877125964978153909381399391651900379307339771343298861412431940828440883129049  
488382946994739567719143606756805998823323701789107650331353040000726483507731027736967256825198755  
118263632017030755051099487985364189385102811858477303957195515494459865637006333172875364854023319  
4362377515517160419329629033944735961302692477555348876739973877527061108357320617363860139300473491  
147838391201907224699921787001870430782264991473554959459560878685244811271725939027510790989045922  
4662409137578117655804693388146906390748623266119419826117206556060446418048213689871491973992589569  
012720451214860656024361115175931432258979403592847596023649853569756548786656980609049483237490087

59228731837134144352984332836340082450407073664702182903941453922953466249755265888859322820276949  
815683198394191112534968621593228063000957552348628988572256966911044696751890659538095670506188821  
1438488732369544375177646855817408370957937518361735853774625415458835209563088047668690205066113314  
19403842406615974777549999137979797714296775069319998087362574782365

>WBE22135.1 ORF8 protein [Severe acute respiratory syndrome coronavirus 2]  
MKFLVFLGIITTVAAAFHQECSLQSCTQHQPYYVDDPCPIHFYSKWYIRVGARKSAPLIELCVDEAGSKSP  
IQYIDIGNYTVSCLPFTINCQEPKLGSLVVRCFSFYEDFLEYHDVRRVVLDFI

ORF8 (WBE22135.1) protein@integer

194611009554485015080235775478493357293746189205074064273910773374230324575542845088267908932795322  
50367336043068643456811124221825233150723390246662221500328