



Figure S1. Overall sequence and structural similarity of the six Philippine SARS-CoV-2 scaffolds with respect to the NC_045512.2 reference sequence. The Mauve alignment shows a single sequence block (in red) for each of the scaffolds, suggesting highly identical genomic organizations. The sequences are also very similar at the nucleotide level, as shown by the red histogram inside the sequence blocks, except for a few breaks (depicted in white) that coincide with regions of ambiguous bases (N's) in the scaffolds.

	cov	pid	1 [:	1 100
1 NC_045512/1-932	100.0%	100.0%	SADAQSF LN RVCGVSAARLTPCGTGTSTDVVYRAFDIYNDKVAGFAKFLKTNCCRFQEKDEDDNLSYFVVKRHTFSNYQHEETIYNLLKDCPAVAKHD		
2 PGC001/1-932	100.0%	99.4%	SADAQSF LN RVCGVSAARLTPCGTGTSTDVVYRAFDIYNDKVAGFAKFLKTNCCRFQEKDEDDNLSYFVVKRHTFSNYQHEETIYNLLKDCPAVAKHD		
3 PGC002/1-932	100.0%	99.7%	SADAQSF LN RVCGVSAARLTPCGTGTSTDVVYRAFDIYNDKVAGFAKFLKTNCCRFQEKDEDDNLSYFVVKRHTFSNYQHEETIYNLLKDCPAVAKHD		
4 PGC003/1-932	100.0%	99.9%	SADAQSF LN RVCGVSAARLTPCGTGTSTDVVYRAFDIYNDKVAGFAKFLKTNCCRFQEKDEDDNLSYFVVKRHTFSNYQHEETIYNLLKDCPAVAKHD		
5 PGC004/1-932	100.0%	99.9%	SADAQSF LN RVCGVSAARLTPCGTGTSTDVVYRAFDIYNDKVAGFAKFLKTNCCRFQEKDEDDNLSYFVVKRHTFSNYQHEETIYNLLKDCPAVAKHD		
6 PGC005/1-932	100.0%	99.8%	SADAQSF LN RVCGVSAARLTPCGTGTSTDVVYRAFDIYNDKVAGFAKFLKTNCCRFQEKDEDDNLSYFVVKRHTFSNYQHEETIYNLLKDCPAVAKHD		
7 PGC006/1-932	100.0%	99.9%	SADAQSF LN RVCGVSAARLTPCGTGTSTDVVYRAFDIYNDKVAGFAKFLKTNCCRFQEKDEDDNLSYFVVKRHTFSNYQHEETIYNLLKDCPAVAKHD		
consensus/80%			SADAQSF LN RVCGVSAARLTPCGTGTSTDVVYRAFDIYNDKVAGFAKFLKTNCCRFQEKDEDDNLSYFVVKRHTFSNYQHEETIYNLLKDCPAVAKHD		
	cov	pid	401	:	5 500
1 NC_045512/1-932	100.0%	100.0%	LTNNVAFQTVKPGNFNKFYDFAVSKGFFKEGSSVELKHF FFAQDGNAAISDYDYRYRNLPTMCDIRQLLFVVEVVDKYFDCYDGGCINANQVIVNNLDK		
2 PGC001/1-932	100.0%	99.4%	LTNNVAFQTVKPGNFNKFYDFAVSKGFFKEGSSVELKHF FFAQDGNAAISDYDYRYRNLPTMCDIRQLLFVVEVVDKYFDCYDGGCINANQVIVNNLDK		
3 PGC002/1-932	100.0%	99.7%	LTNNVAFQTVKPGNFNKFYDFAVSKGFFKEGSSVELKHF FFAQDGNAAISDYDYRYRNLPTMCDIRQLLFVVEVVDKYFDCYDGGCINANQVIVNNLDK		
4 PGC003/1-932	100.0%	99.9%	LTNNVAFQTVKPGNFNKFYDFAVSKGFFKEGSSVELKHF FFAQDGNAAISDYDYRYRNLPTMCDIRQLLFVVEVVDKYFDCYDGGCINANQVIVNNLDK		
5 PGC004/1-932	100.0%	99.9%	LTNNVAFQTVKPGNFNKFYDFAVSKGFFKEGSSVELKHF FFAQDGNAAISDYDYRYRNLPTMCDIRQLLFVVEVVDKYFDCYDGGCINANQVIVNNLDK		
6 PGC005/1-932	100.0%	99.8%	LTNNVAFQTVKPGNFNKFYDFAVSKGFFKEGSSVELKHF FFAQDGNAAISDYDYRYRNLPTMCDIRQLLFVVEVVDKYFDCYDGGCINANQVIVNNLDK		
7 PGC006/1-932	100.0%	99.9%	LTNNVAFQTVKPGNFNKFYDFAVSKGFFKEGSSVELKHF FFAQDGNAAISDYDYRYRNLPTMCDIRQLLFVVEVVDKYFDCYDGGCINANQVIVNNLDK		
consensus/80%			LTNNVAFQTVKPGNFNKFYDFAVSKGFFKEGSSVELKHF FFAQDGNAAISDYDYRYRNLPTMCDIRQLLFVVEVVDKYFDCYDGGCINANQVIVNNLDK		
	cov	pid	501	:	6 600
1 NC_045512/1-932	100.0%	100.0%	SAGFPFNKWKARLYYDSMSYEDQDALFAYTKRNVIPITITQMNLYKIAISAKNRARTVAGVSICTMTNRQFHQKLLKSAATRATVIGTSKFYGGWHN		
2 PGC001/1-932	100.0%	99.4%	SAGFPFNKWKARLYYDSMSYEDQDALFAYTKRNVIPITITQMNLYKIAISAKNRARTVAGVSICTMTNRQFHQKLLKSAATRATVIGTSKFYGGWHN		
3 PGC002/1-932	100.0%	99.7%	SAGFPFNKWKARLYYDSMSYEDQDALFAYTKRNVIPITITQMNLYKIAISAKNRARTVAGVSICTMTNRQFHQKLLKSAATRATVIGTSKFYGGWHN		
4 PGC003/1-932	100.0%	99.9%	SAGFPFNKWKARLYYDSMSYEDQDALFAYTKRNVIPITITQMNLYKIAISAKNRARTVAGVSICTMTNRQFHQKLLKSAATRATVIGTSKFYGGWHN		
5 PGC004/1-932	100.0%	99.9%	SAGFPFNKWKARLYYDSMSYEDQDALFAYTKRNVIPITITQMNLYKIAISAKNRARTVAGVSICTMTNRQFHQKLLKSAATRATVIGTSKFYGGWHN		
6 PGC005/1-932	100.0%	99.8%	SAGFPFNKWKARLYYDSMSYEDQDALFAYTKRNVIPITITQMNLYKIAISAKNRARTVAGVSICTMTNRQFHQKLLKSAATRATVIGTSKFYGGWHN		
7 PGC006/1-932	100.0%	99.9%	SAGFPFNKWKARLYYDSMSYEDQDALFAYTKRNVIPITITQMNLYKIAISAKNRARTVAGVSICTMTNRQFHQKLLKSAATRATVIGTSKFYGGWHN		
consensus/80%			SAGFPFNKWKARLYYDSMSYEDQDALFAYTKRNVIPITITQMNLYKIAISAKNRARTVAGVSICTMTNRQFHQKLLKSAATRATVIGTSKFYGGWHN		
	cov	pid	801	:	9 900
1 NC_045512/1-932	100.0%	100.0%	TETDLTKGPHEFCSQHTMLVKQGDDYVYLPYPDPSRILGAGCFVDDIVKTDGTLMIERFVSLAIDAYPLTKHPNQEYADV FHL YLQYIRKLDHDELTGHML		
2 PGC001/1-932	100.0%	99.4%	TETDLTKGPHEFCSQHTMLVKQGDDYVYLPYPDPSRILGAGCFVDDIVKTDGTLMIERFVSLAIDAYPLTKHPNQEYADV FHL YLQYIRKLDHDELTGHML		
3 PGC002/1-932	100.0%	99.7%	TETDLTKGPHEFCSQHTMLVKQGDDYVYLPYPDPSRILGAGCFVDDIVKTDGTLMIERFVSLAIDAYPLTKHPNQEYADV FHL YLQYIRKLDHDELTGHML		
4 PGC003/1-932	100.0%	99.9%	TETDLTKGPHEFCSQHTMLVKQGDDYVYLPYPDPSRILGAGCFVDDIVKTDGTLMIERFVSLAIDAYPLTKHPNQEYADV FHL YLQYIRKLDHDELTGHML		
5 PGC004/1-932	100.0%	99.9%	TETDLTKGPHEFCSQHTMLVKQGDDYVYLPYPDPSRILGAGCFVDDIVKTDGTLMIERFVSLAIDAYPLTKHPNQEYADV FHL YLQYIRKLDHDELTGHML		
6 PGC005/1-932	100.0%	99.8%	TETDLTKGPHEFCSQHTMLVKQGDDYVYLPYPDPSRILGAGCFVDDIVKTDGTLMIERFVSLAIDAYPLTKHPNQEYADV FHL YLQYIRKLDHDELTGHML		
7 PGC006/1-932	100.0%	99.9%	TETDLTKGPHEFCSQHTMLVKQGDDYVYLPYPDPSRILGAGCFVDDIVKTDGTLMIERFVSLAIDAYPLTKHPNQEYADV FHL YLQYIRKLDHDELTGHML		
consensus/80%			TETDLTKGPHEFCSQHTMLVKQGDDYVYLPYPDPSRILGAGCFVDDIVKTDGTLMIERFVSLAIDAYPLTKHPNQEYADV FHL YLQYIRKLDHDELTGHML		

Figure S2. Partial amino acid sequence alignment of the ORF1ab gene RdRp region. The figure shows the alignment for residues 1-100, 401-600, and 801-900 of the RdRp region, containing all observed mismatches (total alignment length = 932 a.a.). The coverage (cov), percent identities (pid), and highlighted mismatches are relative to the NC_045512.2 reference sequence.

	cov	pid	701	.	.	.	:	.	.	8	800
1 NC_045512/1-1274	100.0%	100.0%	AENSVAYSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECNLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIKDFGGF								
2 PGC001/1-1274	100.0%	99.6%	AENSVAYSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECNLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIKDFGGF								
3 PGC002/1-1274	100.0%	100.0%	AENSVAYSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECNLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIKDFGGF								
4 PGC003/1-1274	100.0%	100.0%	AENSVAYSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECNLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIKDFGGF								
5 PGC004/1-1274	100.0%	100.0%	AENSVAYSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECNLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIKDFGGF								
6 PGC005/1-1274	100.0%	99.5%	AENSVAYSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECNLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIKDFGGF								
7 PGC006/1-1274	100.0%	100.0%	AENSVAYSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECNLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIKDFGGF								
consensus/80%			AENSVAYSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECNLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIKDFGGF								

Figure S3. Partial amino acid sequence alignment of the spike glycoprotein (S) gene. The figure shows the alignment for residues 701-800 amino acids of the S gene product, containing the single observed mismatch (total alignment length = 1,274 a.a.). The coverage (cov), percent identities (pid), and highlighted mismatch are relative to the NC_045512.2 reference sequence.

```

      cov   pid   1 [ . . . : . . . 1 100
1 NC_045512/1-235 100.0% 100.0% MADSNGTITVEELKKLLEQWNLVIGFLFTWICLLQFAYANRRNFLYIIKLI FLWLLWPVTLACFVLAAVYRINWITGGIAIAMA CLVGLMWLSYFIASF
2 PGC001/1-235 100.0% 91.1% MADSNGTITVXXXXXXXXXWNLVIGFLFTWICLLQFAYANRRNFLYIIKLI FLWLLWPVTLACFVLAAVYRINWITGGIAIAMA CLVGLMWLSYFIASF
3 PGC002/1-235 100.0% 94.9% MADSNGTITVEELKKLLEQWNLVIGFLFTWICLLQFAYANRRNFLYIIKLI FLWLLWPVTLACFVLAAVYRINWITGGIAIAMA CLVGLMWLSYFIASF
4 PGC003/1-235 100.0% 94.5% MADSNGTITVEELKKLLEQWNLVIGFLFTWICLLQFAYANRRNFLYIIKLI FLWLLWPVTLACFVLAAVYRINWITGGIAIAMA CLVGLMWLSYFIASF
5 PGC004/1-235 100.0% 94.9% MADSNGTITVEELKKLLEQWNLVIGFLFTWICLLQFAYANRRNFLYIIKLI FLWLLWPVTLACFVLAAVYRINWITGGIAIAMA CLVGLMWLSYFIASF
6 PGC005/1-235 100.0% 94.9% MADSNGTITVEELKKLLEQWNLVIGFLFTWICLLQFAYANRRNFLYIIKLI FLWLLWPVTLACFVLAAVYRINWITGGIAIAMA CLVGLMWLSYFIASF
7 PGC006/1-235 100.0% 94.9% MADSNGTITVEELKKLLEQWNLVIGFLFTWICLLQFAYANRRNFLYIIKLI FLWLLWPVTLACFVLAAVYRINWITGGIAIAMA CLVGLMWLSYFIASF
consensus/80% MADSNGTITVEELKKLLEQWNLVIGFLFTWICLLQFAYANRRNFLYIIKLI FLWLLWPVTLACFVLAAVYRINWITGGIAIAMA CLVGLMWLSYFIASF

      cov   pid 101 . . . : . . . 2 200
1 NC_045512/1-235 100.0% 100.0% RLFARTRSMWSFNPETNILLNVPLHGTILTRP LLESELVIGAVILRGHLRIAGHHLGRCDIKDLPKEITVATSRTL SYYKLGASQRVAGDSXXXXXXXXXX
2 PGC001/1-235 100.0% 91.1% RLFARTRSMWSFNPETNILLNVPLHGTILTRP LLESELVIGAVILRGHLRIAGHHLGRCDIKDLPKEITVATSRTL SYYKLGASQRVAGDSXXXXXXXXXX
3 PGC002/1-235 100.0% 94.9% RLFARTRSMWSFNPETNILLNVPLHGTILTRP LLESELVIGAVILRGHLRIAGHHLGRCDIKDLPKEITVATSRTL SYYKLGASQRVAGDSNWE LRSV*Q
4 PGC003/1-235 100.0% 94.5% RLFARTRSMWSFNPETNILLNVPLHGTILTRP LLESELVIGAVILRGHLRIAGHHLGRCDIKDLPKEITVATSRTL SYYKLGASQRVAGDSXXXXXXXXXX
5 PGC004/1-235 100.0% 94.9% RLFARTRSMWSFNPETNILLNVPLHGTILTRP LLESELVIGAVILRGHLRIAGHHLGRCDIKDLPKEITVATSRTL SYYKLGASQRVAGDSXXXXXXXXXX
6 PGC005/1-235 100.0% 94.9% RLFARTRSMWSFNPETNILLNVPLHGTILTRP LLESELVIGAVILRGHLRIAGHHLGRCDIKDLPKEITVATSRTL SYYKLGASQRVAGDSXXXXXXXXXX
7 PGC006/1-235 100.0% 94.9% RLFARTRSMWSFNPETNILLNVPLHGTILTRP LLESELVIGAVILRGHLRIAGHHLGRCDIKDLPKEITVATSRTL SYYKLGASQRVAGDSXXXXXXXXXX
consensus/80% RLFARTRSMWSFNPETNILLNVPLHGTILTRP LLESELVIGAVILRGHLRIAGHHLGRCDIKDLPKEITVATSRTL SYYKLGASQRVAGDSXXXXXXXXXX

      cov   pid 201 . . . ] 235
1 NC_045512/1-235 100.0% 100.0% XXXGFAAYSRYRIGNYKLNTDHSSSSDNIAL LVQ*
2 PGC001/1-235 100.0% 91.1% XXXGFAAYSRYRIGNYKLNTDHSSSSDNIAL LVQ*
3 PGC002/1-235 100.0% 94.9% VIXGFAAYSRYRIGNYKLNTDHSSSSDNIAL LVQ*
4 PGC003/1-235 100.0% 94.5% XXXGFAAYSRYRIGNYKLNTDHSSSSDNIAL LVQ*
5 PGC004/1-235 100.0% 94.9% XXXGFAAYSRYRIGNYKLNTDHSSSSDNIAL LVQ*
6 PGC005/1-235 100.0% 94.9% XXXGFAAYSRYRIGNYKLNTDHSSSSDNIAL LVQ*
7 PGC006/1-235 100.0% 94.9% XXXGFAAYSRYRIGNYKLNTDHSSSSDNIAL LVQ*
consensus/80% ...GFAAYSRYRIGNYKLNTDHSSSSDNIAL LVQ*

```

Figure S4. Amino acid sequence alignment of the membrane glycoprotein (M) gene (total alignment length = 235 a.a.). The coverage (cov), percent identities (pid), and highlighted mismatches are relative to the NC_045512.2 reference sequence. The symbol X corresponds to ambiguous bases (N's) or gaps in the underlying nucleotide sequence. The stretch of mismatched residues from positions 192-202 are the result of a 35 bp repeat in the nucleotide sequence assembly for sample PGC002.

	cov	pid	1 [. . . : . .] 76
1 NC_045512/1-76	100.0%	100.0%	MYSFVSEETGTLIVNSVLLFLAFVWFLLVTLAILTALRLCAYCCNIVNVSLVKPSFYVYSRVKLNLSRVPDLLV*
2 PGC001/1-76	100.0%	100.0%	MYSFVSEETGTLIVNSVLLFLAFVWFLLVTLAILTALRLCAYCCNIVNVSLVKPSFYVYSRVKLNLSRVPDLLV*
3 PGC002/1-76	100.0%	100.0%	MYSFVSEETGTLIVNSVLLFLAFVWFLLVTLAILTALRLCAYCCNIVNVSLVKPSFYVYSRVKLNLSRVPDLLV*
4 PGC003/1-76	100.0%	100.0%	MYSFVSEETGTLIVNSVLLFLAFVWFLLVTLAILTALRLCAYCCNIVNVSLVKPSFYVYSRVKLNLSRVPDLLV*
5 PGC004/1-76	100.0%	100.0%	MYSFVSEETGTLIVNSVLLFLAFVWFLLVTLAILTALRLCAYCCNIVNVSLVKPSFYVYSRVKLNLSRVPDLLV*
6 PGC005/1-76	100.0%	100.0%	MYSFVSEETGTLIVNSVLLFLAFVWFLLVTLAILTALRLCAYCCNIVNVSLVKPSFYVYSRVKLNLSRVPDLLV*
7 PGC006/1-76	100.0%	100.0%	MYSFVSEETGTLIVNSVLLFLAFVWFLLVTLAILTALRLCAYCCNIVNVSLVKPSFYVYSRVKLNLSRVPDLLV*
consensus/80%			MYSFVSEETGTLIVNSVLLFLAFVWFLLVTLAILTALRLCAYCCNIVNVSLVKPSFYVYSRVKLNLSRVPDLLV*

Figure S5. Amino acid sequence alignment of the envelope protein (E) gene (total alignment length = 76 a.a.). The coverage (cov) and percent identities (pid) shown are relative to the NC_045512.2 reference sequence. No mismatches were observed.

	cov	pid	1 [:	1 100
1 NC_045512/1-420	100.0%	100.0%	MSDNGPQNQRNAPRITFGGSPDSTGSNQNNGERSGARSKQRRPQGLPNTASWFTALTQHGKEDLKFRGQGVPIINTNSSPDDQIGYYRRATRRIRGGDGK		
2 PGC001/1-420	100.0%	99.8%	MSDNGPQNQRNALRITFGGSPDSTGSNQNNGERSGARSKQRRPQGLPNTASWFTALTQHGKEDLKFRGQGVPIINTNSSPDDQIGYYRRATRRIRGGDGK		
3 PGC002/1-420	100.0%	99.8%	MSDNGPQNQRNALRITFGGSPDSTGSNQNNGERSGARSKQRRPQGLPNTASWFTALTQHGKEDLKFRGQGVPIINTNSSPDDQIGYYRRATRRIRGGDGK		
4 PGC003/1-420	100.0%	99.8%	MSDNGPQNQRNALRITFGGSPDSTGSNQNNGERSGARSKQRRPQGLPNTASWFTALTQHGKEDLKFRGQGVPIINTNSSPDDQIGYYRRATRRIRGGDGK		
5 PGC004/1-420	100.0%	99.8%	MSDNGPQNQRNALRITFGGSPDSTGSNQNNGERSGARSKQRRPQGLPNTASWFTALTQHGKEDLKFRGQGVPIINTNSSPDDQIGYYRRATRRIRGGDGK		
6 PGC005/1-420	100.0%	99.8%	MSDNGPQNQRNALRITFGGSPDSTGSNQNNGERSGARSKQRRPQGLPNTASWFTALTQHGKEDLKFRGQGVPIINTNSSPDDQIGYYRRATRRIRGGDGK		
7 PGC006/1-420	100.0%	99.8%	MSDNGPQNQRNALRITFGGSPDSTGSNQNNGERSGARSKQRRPQGLPNTASWFTALTQHGKEDLKFRGQGVPIINTNSSPDDQIGYYRRATRRIRGGDGK		
consensus/80%			MSDNGPQNQRNALRITFGGSPDSTGSNQNNGERSGARSKQRRPQGLPNTASWFTALTQHGKEDLKFRGQGVPIINTNSSPDDQIGYYRRATRRIRGGDGK		
	cov	pid	101	:	2 200
1 NC_045512/1-420	100.0%	100.0%	MKDLSPRWYFYLLGTGPEAGLPYGANKDGIWVATEGALNTPKDHIIGTRNPANNAIVLQLPQGTTLPKGFYAEGRGGSQASSR55SRNSSRNSTPG		
2 PGC001/1-420	100.0%	99.8%	MKDLSPRWYFYLLGTGPEAGLPYGANKDGIWVATEGALNTPKDHIIGTRNPANNAIVLQLPQGTTLPKGFYAEGRGGSQASSR55SRNSSRNSTPG		
3 PGC002/1-420	100.0%	99.8%	MKDLSPRWYFYLLGTGPEAGLPYGANKDGIWVATEGALNTPKDHIIGTRNPANNAIVLQLPQGTTLPKGFYAEGRGGSQASSR55SRNSSRNSTPG		
4 PGC003/1-420	100.0%	99.8%	MKDLSPRWYFYLLGTGPEAGLPYGANKDGIWVATEGALNTPKDHIIGTRNPANNAIVLQLPQGTTLPKGFYAEGRGGSQASSR55SRNSSRNSTPG		
5 PGC004/1-420	100.0%	99.8%	MKDLSPRWYFYLLGTGPEAGLPYGANKDGIWVATEGALNTPKDHIIGTRNPANNAIVLQLPQGTTLPKGFYAEGRGGSQASSR55SRNSSRNSTPG		
6 PGC005/1-420	100.0%	99.8%	MKDLSPRWYFYLLGTGPEAGLPYGANKDGIWVATEGALNTPKDHIIGTRNPANNAIVLQLPQGTTLPKGFYAEGRGGSQASSR55SRNSSRNSTPG		
7 PGC006/1-420	100.0%	99.8%	MKDLSPRWYFYLLGTGPEAGLPYGANKDGIWVATEGALNTPKDHIIGTRNPANNAIVLQLPQGTTLPKGFYAEGRGGSQASSR55SRNSSRNSTPG		
consensus/80%			MKDLSPRWYFYLLGTGPEAGLPYGANKDGIWVATEGALNTPKDHIIGTRNPANNAIVLQLPQGTTLPKGFYAEGRGGSQASSR55SRNSSRNSTPG		
	cov	pid	201	:	3 300
1 NC_045512/1-420	100.0%	100.0%	SSRGTSPARMAGNGGDAALALLLDRLNQLLESKMSGKGGQQQQGQTVTKKSAEASKKPRQKRTATKAYNVTQAFGRRGPEQTQGNFGDQELIRQGTDYKH		
2 PGC001/1-420	100.0%	99.8%	SSRGTSPARMAGNGGDAALALLLDRLNQLLESKMSGKGGQQQQGQTVTKKSAEASKKPRQKRTATKAYNVTQAFGRRGPEQTQGNFGDQELIRQGTDYKH		
3 PGC002/1-420	100.0%	99.8%	SSRGTSPARMAGNGGDAALALLLDRLNQLLESKMSGKGGQQQQGQTVTKKSAEASKKPRQKRTATKAYNVTQAFGRRGPEQTQGNFGDQELIRQGTDYKH		
4 PGC003/1-420	100.0%	99.8%	SSRGTSPARMAGNGGDAALALLLDRLNQLLESKMSGKGGQQQQGQTVTKKSAEASKKPRQKRTATKAYNVTQAFGRRGPEQTQGNFGDQELIRQGTDYKH		
5 PGC004/1-420	100.0%	99.8%	SSRGTSPARMAGNGGDAALALLLDRLNQLLESKMSGKGGQQQQGQTVTKKSAEASKKPRQKRTATKAYNVTQAFGRRGPEQTQGNFGDQELIRQGTDYKH		
6 PGC005/1-420	100.0%	99.8%	SSRGTSPARMAGNGGDAALALLLDRLNQLLESKMSGKGGQQQQGQTVTKKSAEASKKPRQKRTATKAYNVTQAFGRRGPEQTQGNFGDQELIRQGTDYKH		
7 PGC006/1-420	100.0%	99.8%	SSRGTSPARMAGNGGDAALALLLDRLNQLLESKMSGKGGQQQQGQTVTKKSAEASKKPRQKRTATKAYNVTQAFGRRGPEQTQGNFGDQELIRQGTDYKH		
consensus/80%			SSRGTSPARMAGNGGDAALALLLDRLNQLLESKMSGKGGQQQQGQTVTKKSAEASKKPRQKRTATKAYNVTQAFGRRGPEQTQGNFGDQELIRQGTDYKH		
	cov	pid	301	:	4 400
1 NC_045512/1-420	100.0%	100.0%	WPQIAQFAPSASAFFGMSRIGMEVTPSGTWLTYTGAIKLDKDPNFKDQVILLNKHIDAYKTFPPTPEPKKDKKKKKADETQALPQRQKKQQTVTLLPAADL		
2 PGC001/1-420	100.0%	99.8%	WPQIAQFAPSASAFFGMSRIGMEVTPSGTWLTYTGAIKLDKDPNFKDQVILLNKHIDAYKTFPPTPEPKKDKKKKKADETQALPQRQKKQQTVTLLPAADL		
3 PGC002/1-420	100.0%	99.8%	WPQIAQFAPSASAFFGMSRIGMEVTPSGTWLTYTGAIKLDKDPNFKDQVILLNKHIDAYKTFPPTPEPKKDKKKKKADETQALPQRQKKQQTVTLLPAADL		
4 PGC003/1-420	100.0%	99.8%	WPQIAQFAPSASAFFGMSRIGMEVTPSGTWLTYTGAIKLDKDPNFKDQVILLNKHIDAYKTFPPTPEPKKDKKKKKADETQALPQRQKKQQTVTLLPAADL		
5 PGC004/1-420	100.0%	99.8%	WPQIAQFAPSASAFFGMSRIGMEVTPSGTWLTYTGAIKLDKDPNFKDQVILLNKHIDAYKTFPPTPEPKKDKKKKKADETQALPQRQKKQQTVTLLPAADL		
6 PGC005/1-420	100.0%	99.8%	WPQIAQFAPSASAFFGMSRIGMEVTPSGTWLTYTGAIKLDKDPNFKDQVILLNKHIDAYKTFPPTPEPKKDKKKKKADETQALPQRQKKQQTVTLLPAADL		
7 PGC006/1-420	100.0%	99.8%	WPQIAQFAPSASAFFGMSRIGMEVTPSGTWLTYTGAIKLDKDPNFKDQVILLNKHIDAYKTFPPTPEPKKDKKKKKADETQALPQRQKKQQTVTLLPAADL		
consensus/80%			WPQIAQFAPSASAFFGMSRIGMEVTPSGTWLTYTGAIKLDKDPNFKDQVILLNKHIDAYKTFPPTPEPKKDKKKKKADETQALPQRQKKQQTVTLLPAADL		
	cov	pid	401	:] 420
1 NC_045512/1-420	100.0%	100.0%	DDFSKQLQSMSSADSTQA*		
2 PGC001/1-420	100.0%	99.8%	DDFSKQLQSMSSADSTQA*		
3 PGC002/1-420	100.0%	99.8%	DDFSKQLQSMSSADSTQA*		
4 PGC003/1-420	100.0%	99.8%	DDFSKQLQSMSSADSTQA*		
5 PGC004/1-420	100.0%	99.8%	DDFSKQLQSMSSADSTQA*		
6 PGC005/1-420	100.0%	99.8%	DDFSKQLQSMSSADSTQA*		
7 PGC006/1-420	100.0%	99.8%	DDFSKQLQSMSSADSTQA*		
consensus/80%			DDFSKQLQSMSSADSTQA*		

Figure S6. Amino acid sequence alignment of the nucleocapsid phosphoprotein (N) gene (total alignment length = 420 a.a.). The coverage (cov), percent identities (pid), and highlighted mismatches are relative to the NC_045512.2 reference sequence.