VIROLOGICA SINICA

Electronic Supplementary Material

Detection and characterization of diverse alpha and betacoronaviruses from bats in China

Lin Xu[1#], Fuqiang Zhang[2#], Weihong Yang[3#], Tinglei Jiang[4#], Guanjun Lu[4], Biao He[1], Xingyu Li[1], Tingsong Hu[2], Gang Chen[2], Yun Feng[3], Yuzhen Zhang[3], Quanshui Fan[2], Jiang Feng[4], Hailin Zhang[3], Changchun Tu[1,5]✉

1. Military Veterinary Institute, Academy of Military Medical Sciences, Changchun 130122, China
2. Center for Disease Control and Prevention of Chengdu Military Region, Kunming 650118, China
3. Yunnan Institute of Endemic Diseases Control and Prevention, Dali 671000, China
4. School of Urban and Environmental Sciences, Northeast Normal University, Changchun 130024, China
5. Jiangsu Co-innovation Center for Prevention and Control of Important Animal Infectious Diseases and Zoonoses, Yangzhou 225009, China

Supporting information to DOI: 10.1007/s12250-016-3727-3

Table S1. Primers used in this study.

Primers	Position of first nucleotide (nt)	Sequences (5′–3′)
Pan-coronavirus nested primers		
CoVOF	14615	ATGGGWTGGGAYTAYCCIAARTG
CoVOR	15200	TGYTGIGARCAAAAYTCRTG
CoVF	14618	GGITGGGAYTAYCCIAARTGYGA
CoVR	15035	CCRTCATCWGAIARWATCATCAT
JTMC15		
F3	2811	GCGTGTAGAYAARGTGCTTAA
R3	4799	CCACGCTTRAGAAATTCAA
F4	4648	GTKTCAGTDTCWTCACCAGA
R4	6919	AATRCTTAACAAYAAYAGCCACAT
F5	6806	CACTWCCTACRACTATWGCTAAAAAT
R5	8939	GCAGARGTRGMAAARTCACTATACT
F6	8827	CCTGGHTTACCDGGTACTGT
R6	11329	CGYCTAGCAGCATCATCATA
F10	17350	TGAGTGTYGTCAATGCTAGAC
R10	19665	CTACYTTDGTGTAAACAGCATTATT
F12	21283	GCTATACCATGCATGCTAACT
R12	22615	CGAAAAAGARGTTGAGTTGTAG
5’OR	252	ATTGGCTGAAACGACACCACTTC
5’IR	161	GTCGATTAAAGCACTTGGCTCCA
3’OF	27341	GACATCCCAGAGTGGAGGAG
3’IF	27448	AGGTGTTGATGCCTCAGGCTAT
JPDB144		
F1	1	GATTTAAGWGAATAGCYTRGCTATC
R1	1749	GTVGTWCCAGAVAGWARTGC
F2	1572	GGTACTATGYACTTTRTKCCT
R2	3846	CWGCDATRCCACCRCCAT
F3	3759	GTKACHHTAGTHTTWGGTGA
R3	5978	ACTAATAGYATCACYGCCA
F4	5940	TAYWCTAATAGYTGCCTTG
R4	8244	ACATCAGAYTCCACACC
F5	796	TGGCCAGGAAARTTTAGC
R5	10071	TCACTACCAGTYTCRCTGTA
F6	9854	TACTGATGGTAARCTKAATTGTAG
R6	12438	CATAGTTTGCATAGCACT
F7	12559	TCWATGTATAAGCAAGCACGT
R7	14645	GGATCWGCKGCATACATCAT
F8	14556	TATCTTGTGGTTATCACTAC
R8	17048	ATACCTCTCTTGATTCAC
F9	16931	CGYATWGAYTATAGTGATGCTG
R9	18983	ATCCCAMTCMACACGTTC

To be continued

Continued

Primers	Position of first nucleotide (nt)	Sequences (5′–3′)
F10	18791	TATGCCTGCTGGASTCATTC
R10	20854	ATACTGRCACAATTGCATATATT
F11	20639	CCTATTGAYTTAACWATGATTG
R11	22365	GARWAGAGRTGAACRCCTTG
F12	22338	GAGTGGTTYGGYATTACMCA
R12	24688	GAAATAGCACCRAAAGTRTTAG
F13	24267	GCWGATCCYGGYTATATGC
R13	26250	CATAACGRTTKTGYYCGAAG
F14	26140	ACTAAAGYATYAGCAAAACAAGA
R14	27969	CGTTAAACCCASTCSTCAG
F15	27842	GCTAYTMGATTATGTGTGC
R15	30232	GCCTAATCTAATTGAATAATAGC
5’OR	268	GTCACACTAGCCTTGGAAAGCA
5’IR	83	CAGACCACAACACAACACGCACACAACA
3’OF	30061	ATCATGTTARACTTACAGTGCAAG
3’IF	30151	AAAGACTGTCACCTCTGCGTGATT
JTAC2		
F1	4010	CCACTATGTSACCAATWTYTATGAT
R1	6107	CTTATCAATAAGCTTAGTAGCGTCT
F2	5961	TGTYGGMCAYTATACTGTTTTTGA
R2	8460	ACACGGCAATARGTCATAGC
F3	8172	TGGTAAAACWCTTGTKTTTGC
R3	9704	ACAAGCGCCATTAATGAA
F4	9615	TTAAYATTYTGGCRTGCTATGAT
R4	11457	CYTGTTCMGCCATTCTATCAA
F5	11364	GTTCTCCACCTCAGTTGGT
R5	13349	TCCTCACCAAAWATATCACTCTT
F6	13199	GATAAYCAGGATCTTAATGGTGA
R6	15458	TGACATGRTCATAAGCRCACTT
F7	15309	ATTCWACTGCTAARTTTTGGGA
R7	17318	CCATAAASGAKATWACATGCTCATA
F8	17174	GAKGGTTGYGGTCTYTTTAAAG
R8	18608	GGTGTTGTARGCATTARCATAGC
F9	18470	TGCCMTTYTTYTTCTATGATG
R9	20451	TCRAGCACACTRTTGTAAGACATAG
F10	20301	GGACAATGTTYTGTACCAGTG
R10	22672	ACATTCTTRAAGGCKARCAACTG
F11	22567	AAYGTGTGCACCCAGTATACTAT
R11	24958	TGAMGCTTTAAACAGTGCAA
F12	24352	ATCCCAGAKTATGTYGATGTTAA
R12	26127	ACCTTATAGCCYTCKACAAGCA

Continued

Primers	Position of first nucleotide (nt)	Sequences (5′–3′)
F13	25921	CAGCATCCTTATGGCTTG
R13	27429	ACTTTGGCACAGTCATYTTATAG
Genome walking		
R1	4161	TGGCTGTAAAGTTGGCTGAGGT
R2	4274	GCCACCACCATGAGACAAATTCT
R3	4368	CAGAGCCAACCTTAAGTTTGCCA
R4	2642	ACTTACARCTAACACCGGCCAGT
R5	2745	TAGTCAAACCGTTCTCTACWGGAAT
R6	2889	CGTCATAGAATGCATAACCATCAAC

Table S2. Putative nonstructural proteins (nsps) of ORF1a and ORF1b (replicase) in BatCoV JTAC2, JTMC15 and JPDB144.

nsp	JTAC2$^{\$}$		JTMC15		JPDB144	
	Length (aa)	First - last residue^{Position}	Length (aa)	First - last residue^{Position}	Length (aa)	First - last residue^{Position}
1	-	-	179	M^{1}-G^{179}	731	M^{1}-G^{731}
2	467	Y^{1}-G^{467}	639	G^{180}-G^{818}	489	R^{732}-G^{1220}
3 (ADRP/PL^{Pro})	1637	G^{468}-A^{2104}	1724	A^{819}-G^{2542}	1572	G^{1221}-A^{2792}
4	480	G^{2105}-Q^{2584}	500	K^{2543}-Q^{3042}	512	T^{2793}-Q^{3304}
5 (3CL^{Pro})	302	S^{2585}-Q^{2886}	306	S^{3043}-Q^{3348}	306	S^{3305}-Q^{3610}
6	276	S^{2887}-Q^{3162}	290	G^{3349}-Q^{3638}	292	S^{3611}-Q^{3902}
7	83	S^{3163}-Q^{3245}	83	S^{3639}-Q^{3721}	83	S^{3903}-Q^{3985}
8	195	T^{3246}-Q^{3440}	198	A^{3722}-Q^{3919}	199	A^{3986}-Q^{4184}
9	108	N^{3441}-Q^{3548}	113	N^{3920}-Q^{4032}	110	N^{4185}-Q^{4294}
10	135	A^{3549}-Q^{3683}	139	A^{4033}-Q^{4171}	139	A^{4295}-Q^{4433}
11	17	S^{3684}-D^{3700}	13	S^{4172}-V^{4184}	14	S^{4434}-V^{4447}
12 (RdRP)	927	S^{3684}-Q^{4610}	932	$S^{4172-}Q^{5103}$	934	S^{4434}-Q^{5367}
13 (Hel)	597	S^{4611}-Q^{5207}	601	A^{5104}-Q^{5704}	598	$A^{5368-}Q^{5965}$
14 (ExoN)	517	A^{5208}-Q^{5724}	527	A^{5705}-Q^{6231}	523	S^{5966}-Q^{6488}
15 (XendoU)	339	N^{5725}-Q^{6063}	346	S^{6232}-Q^{6577}	342	G^{6489}-Q^{6830}
16 (2'-O-MT)	301	S^{6064}-K^{6364}	298	A^{6578}-N^{6875}	302	A^{6831}-L^{7132}

Note: $^{\$}$The nsp2 of JTAC2 was partial sequence, lacking the 5' terminal.

Table S3. Comparison of ORF amino acid identities of JTAC2 with other three representative *Alphacoronavirus* strains[#].

ORF	JTAC2	Neixiang-14		512		PEDV	
	Length	Length	% identity	Length	% identity	Length	% identity
1a	3700[§]	2030	87.9	4128	68.9	4117	76.4
1b	2680	2679	92.8	2681	85.3	2680	88.8
S	1365	-	-	1371	58.6	1386	57.1
3a	224	-	-	224	56.9	224	63.1
E	76	-	-	76	83.1	76	83.1
M	226	-	-	227	79.7	226	84.6
N	307[§]	-	-	394	62.6	441	58.8

Note: [#]Abbreviation and accession numbers: Neixiang-14: MIBtCoV Neixiang-14, KF294377; 512: BatCoV/512/2005, NC_009657; PEDV: PEDV-1C, KM609203. [§]incomplete sequences.

Table S4. Comparison of ORF amino acid identities of JPDB144 with other two representative strains of *Betacoronavirus* lineage 3 (β3)[#].

ORF[§]	JPDB144	HKU4		HKU5		MERS	
	Length	Length	% identity	Length	% identity	Length	% identity
1a	4447	4445	93.8	4481	71.1	4391	64.7
1b	2699	2699	97.7	2715	89.4	2701	87.4
S	1352	1352	94.5	1352	69.6	1353	67.1
3a (3)	91	91	89.3	121	44.8	103	46.8
3b (4a)	119	119	92.5	119	53.0	109	38.3
3c (4b)	285	285	88.8	256	39.2	246	27.8
3d (5)	227	227	93.0	223	46.6	224	46.9
E	82	82	98.8	82	80.7	82	69.9
M	219	219	97.3	220	82.3	219	84.2
N	423	423	97.9	427	74.4	413	70.8

Note: [#]Abbreviation and accession numbers: HKU4: BatCoV HKU4-4, EF065508; HKU5: BatCoV HKU5-1, EF065509; MERS: MERS-CoV ChinaGD01, KT006149. [§]ORF3a, 3b, 3c, 3d in HKU4 and HKU5 are described in MERS-CoV as ORF3, 4a, 4b and 5 respectively.

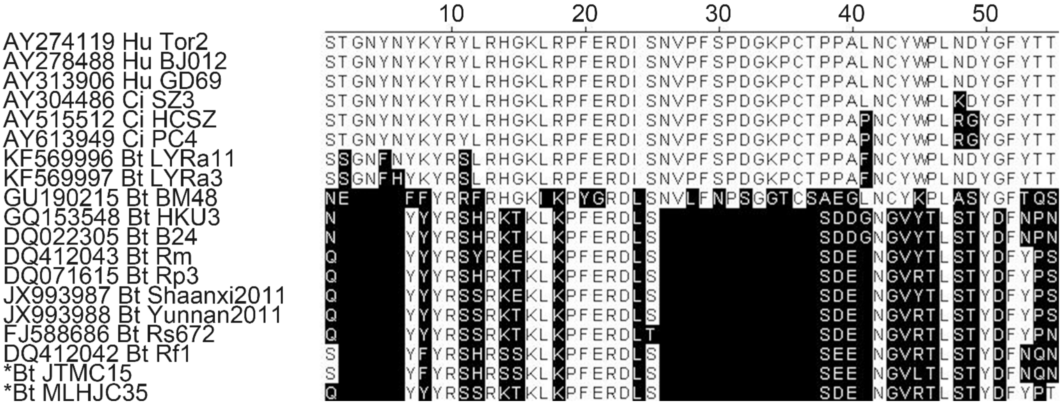

Figure S1. Amino acid sequences comparison of entire RBM of JTMC15, MLHJC35 (marked by asterisks) with other SARS or SARS-related coronaviruses. Middle letters: Hu: human SARS-CoV; Ci: civet SARS-CoV; Bt: Bat SARS-CoV.