### **Supporting information**

### A tick-borne segmented RNA virus contains genome segments derived from unsegmented viral ancestors

Xin-Cheng Qin<sup>a1</sup>, Mang Shi<sup>a,b1</sup>, Jun-Hua Tian<sup>c1</sup>, Xian-Dan Lin<sup>d1</sup>, Dong-Ya Gao<sup>a</sup>, Jin-Rong He<sup>a</sup>, Jian-Bo Wang<sup>a</sup>, Ci-Xiu Li<sup>a</sup>, Yan-Jun Kang<sup>a</sup>, Bin Yu<sup>c</sup>, Dun-Jin Zhou<sup>c</sup>, Jianguo Xu<sup>a</sup>, Alexander Plyusnin<sup>a,e</sup>, Edward C. Holmes<sup>a,b</sup>, Yong-Zhen Zhang<sup>a2</sup>

<sup>a</sup>State Key Laboratory for Infectious Disease Prevention and Control, Collaborative Innovation Center for Diagnosis and Treatment of Infectious Diseases, Department of Zoonoses, National Institute for Communicable Disease Control and Prevention, Chinese Center for Disease Control and Prevention, Changping, 100206, Beijing, China.

<sup>b</sup>Marie Bashir Institute for Infectious Diseases and Biosecurity, Charles Perkins Centre, School of Biological Sciences and Sydney Medical School, The University of Sydney, Sydney, NSW 2006, Australia.

<sup>c</sup>Wuhan Center for Disease Control and Prevention, Wuhan, 430015, Hubei Province, China <sup>d</sup>Wenzhou Center for Disease Control and Prevention, Wenzhou, 325001, Zhejiang Province, China. <sup>e</sup>Department of Virology, Haartman Institute, University of Helsinki, Finland.

<sup>1</sup>Contributed to this work equally.

<sup>2</sup>Correspondence to: Dr. Yong-Zhen Zhang, State Key Laboratory for Infectious Disease Prevention and Control, National Institute of Communicable Disease Control and Prevention, Chinese Center for Disease Control and Prevention, Changping Liuzi 5, Beijing, 102206, China. Tel: 086-10-58900782; Email: <u>zhangyongzhen@icdc.cn</u>

### Table S1

**Table S1.** Genetic identities between putative JMTV proteins and their homologs from major groups of flaviviruses, including West Nile Virus (WNV), Dengue Virus 1 (DENV1), Tick-borne Encephalitis Virus (TBEV), Apoi virus (APOIV), Cell Fusion Agent (CFA), Tamana bat virus (TABV) and from the cDNA library of *Toxocara canis* (TCLA). Representatives of the more divergent hepatitis C virus (HCV) and pestiviruses (Pesti) are also included.

	TCLA	WNV	DENV1	YFV	TBEV	APOIV	CFAV	TABV	HCV	Pesti
NSP1 (NS5)	$0.337^{*}$	0.245	0.246	0.234	0.242	0.249	0.252	0.253	0.140	0.176
NSP2 (NS2b-NS3)	0.248	0.215	0.236	0.230	0.229	0.227	0.225	0.209	0.169	0.185
VP1	0.222	_	_	_	_	_	_	_	_	_
VP2	0.194	_	_	_	_	_	_	_	_	_
VP3	0.218	_	_	_	_	_	_	_	_	_

<sup>\*</sup>The predicted NS5-like protein for TCLA is incomplete compared to JMTV and other members of the *Flaviviridae*. Only partial sequence is involved in similarity comparisons.

# Table S2

#### Table S2. Background information on the JMTV positive samples

Strain	Host Species	Place of sampling	GenBank Accession	Strain	Host Species	Place of sampling	GenBank Accession
JMTVX01	Haemaphysalis longicornis	Wuhan	KJ001590	JMTV60	Rhipicephalus microplus	Enshi	KJ001615
JMTVX06L	Haemaphysalis campanulata	Wuhan	KJ001591	JMTV62	Rhipicephalus microplus	Enshi	KJ001616
JMTVX08	Haemaphysalis longicornis	Wuhan	KJ001592	JMTV67	Rhipicephalus microplus	Enshi	KJ001617
JMTVX23	Haemaphysalis longicornis	Wuhan	KJ001593	JMTV69A	Rhipicephalus microplus	Enshi	KJ001618
JMTVX27	Rhipicephalus microplus	Wuhan	KJ001594	JMTV69B	Rhipicephalus microplus	Enshi	KJ001619
JMTVX35	Rhipicephalus microplus	Wuhan	KJ001595	JMTV72	Rhipicephalus microplus	Yichang	KJ001620
JMTVX37	Rhipicephalus microplus	Wuhan	KJ001596	JMTV84	Rhipicephalus microplus	Jingmen	KJ001621
JMTVX39H	Haemaphysalis campanulata	Wuhan	KJ001597	JMTV85	Rhipicephalus microplus	Jingmen	KJ001634
JMTVX39L	Haemaphysalis longicornis	Wuhan	KJ001598	JMTV90	Rhipicephalus microplus	Jingmen	KJ001622
JMTVx41	Rhipicephalus microplus	Wuhan	KJ001599	JMTV112	Rhipicephalus microplus	Xiangfan	KJ001623
JMTVX55	Haemaphysalis flava	Wuhan	KJ001600	JMTV118	Rhipicephalus microplus	Xiangfan	KJ001624
JMTVH101	Ixodide sinesis	Wuhan	KJ001630	JMTV123	Rhipicephalus microplus	Xiangfan	KJ001625
JMTV11	Rhipicephalus microplus	Shiyan	KJ001603	JMTV129	Rhipicephalus microplus	Xiangfan	KJ001626
JMTV12	Rhipicephalus microplus	Shiyan	KJ001604	JMTVZY1-7	Haemaphysalis flava	Xiangfan	KJ001601
JMTV16	Rhipicephalus microplus	Shiyan	KJ001605	JMTVZY2-4	Haemaphysalis flava	Xiangfan	KJ001602
JMTV19	Rhipicephalus microplus	Shiyan	KJ001606	JMTV138	Rhipicephalus microplus	Jingmen	KJ001627
JMTV29	Rhipicephalus microplus	Shiyan	KJ001632	JMTV141	Rhipicephalus microplus	Jingmen	KJ001628
JMTV33	Rhipicephalus microplus	Shennongjia	KJ001607	JMTV157	Rhipicephalus microplus	Jingmen	KJ001629
JMTV34	Rhipicephalus microplus	Shennongjia	KJ001608	JMTVWC3	Ixodes granulatus	Wenzhou	KJ001633
JMTV36	Haemaohysalis longicornis	Shennongjia	KJ001631	JMTVningbo	Rhipicephalus sanguineus	ningbo	KJ001589
JMTV39	Rhipicephalus microplus	Shennongjia	KJ001609	JMTVH115-1	Haemaphysalis longicornis	Wuhan	KJ001584
JMTV47	Rhipicephalus microplus	Yichang	KJ001610	JMTVH115-4	Haemaphysalis longicornis	Wuhan	KJ001585
JMTV48	Rhipicephalus microplus	Yichang	KJ001611	JMTVH115-5	Haemaphysalis longicornis	Wuhan	KJ001586
JMTV49	Rhipicephalus microplus	Yichang	KJ001612	JMTVH118-6	Haemaphysalis longicornis	Wuhan	KJ001587
JMTV52	Rhipicephalus microplus	Enshi	KJ001613	JMTVX114-1	Haemaphysalis longicornis	Wuhan	KJ001588
JMTV53	Rhipicephalus microplus	Enshi	KJ001614	JMTVDJ1-1	Armigeres sp	Wuhan	KJ001583

# Table S3

Table S3.	Background information	on de-pooled JMTV	positive samples

Strain	Host species	Place of Sampling	GenBank Accession				
Stram	nost species		Segment 1	Segment 2	Segment 3	Segment 4	
YS102-1	Rhipicephalus microplus	Wuhan	KJ001547	KJ001555	KJ001563	KJ001571	
YJ3-3	Rhipicephalus microplus	Wenzhou	KJ001548	KJ001556	KJ001564	KJ001572	
10	Rhipicephalus microplus	Shiyan	KJ001549	KJ001557	KJ001565	KJ001573	
84	Rhipicephalus microplus	Jingmen	KJ001550	KJ001558	KJ001566	KJ001574	
85	Rhipicephalus microplus	Jingmen	KJ001551	KJ001559	KJ001567	KJ001575	
164	Rhipicephalus microplus	Jingmen	KJ001552	KJ001560	KJ001568	KJ001576	
204	Rhipicephalus microplus	Jingzhou	KJ001553	KJ001561	KJ001569	KJ001577	
X84-3	Rhipicephalus microplus	Wuhan	KJ001554	KJ001562	KJ001570	KJ001578	



**Figure S1.** Detection of four JMTV segments in infected and uninfected DH82 cells. The GAPDH gene is used as a cellular control.



Figure S2. Transmission electron micrograph of uninfected DH82 cells.



**Figure S3.** Nucleotide sequence alignment of A) 5' and B) 3' UTRs of the four JMTV segments. The conserved nucleotides are indicated below the sequence and are shaded yellow.



**Figure S4.** Amino acid alignment of the NS5 proteins of JMTV, TCLA, and representatives from the genus *Flavivirus*. This shows the results of the homology search against the Conserved Domain Database in NCBI (grey lines) and the conserved motifs of the flavivirus methyltransferase and RdRp (grey boxes). Also shown are the predicated transmembrane domains (pink box), signal peptide cleavage site (pink arrows), and serine protease cleavage site (blue arrows) of JMTV.



**Figure S5.** Amino acid alignment of the NS2b-NS3 complex of JMTV, TCLA, and representatives from the genus *Flavivirus*. This shows the results of the homology search against the Conserved Domain Database in NCBI (grey lines), the transmembrane domains of the flavivirus NS2B protein (orange boxes), the conserved motifs of the flavivirus helicase (grey boxes), and the positions of the catalytic triad in serine protease (orange triangles). Also shown are the predicated transmembrane domains (pink boxes) and serine protease cleavage site (blue arrow) of JMTV.

(A)		(B)	10 20 30 40 50 60 70
JMTV VP1 T. canis ANT-30	1 1 20 30 40 50 60 70 1 1 MGHSPCGASCGEFGGGETTRDSGA2CEVRAGGEQILLGIWRKRQSSRQKNGTRSARRGKIVGGQGRSR	JMTV_VP2 T. canis ANT-3.1	1 SPNATALMAVILVCIALG 
JMTV VP1 T. canis ANT-30	6 71 GVQVGSELGRTKTGVGWLWRRLIDLFSNAQRVAGGWIWSGDILFPKRGSSREKLALDARSGCHWBTMASR	JMTV_VP2 T. canis ANT-3.1	0 90 100 110 120 130 14 48 VKYYJELIESEMPTDER TOKATEROGENER VCSAKSIVETU DELLOTAVOEVIMIERUIPOLT 68 YHEAVAYTHGTCVEKRSLOATTRSPSHGHEIGENTITESEN VANDAINAEKILVKEANGVTRGEG
JMTV VP1 T. canis ANT-30	150 160 170 100 190 210 18 WSRSENVUPLILITICICIAMARGESONETTICISS 141 SCIALWSCONDAMARGENCECTURERSCIVINGROUM CARGOWERTRASHEV- BAVIS	JMTV_VP2 T. canis ANT-3.1	150 160 170 180 190 200 21 118 MATCEDVLSD MEGGNSVRLNPSCHVPVVAARKLPCKCCVTGCKOVTDRYAAF OG VKONSVATHEN 138 MAGESKPARTHYSEASVIDLSARKGMAMIRIKSESSDCCPCEKFLSNLHELREMVOK PDD HIGDDM
JMTV VP1 T. canis ANT-30	220  230  240  250  260  270  240    79  XTAFFRAGTSHREDLARSGERHWELSSEVGSASEPOPLHDEADE  Ref	JMTV_VP2 T. canis ANT-3.1	220 230 240 250 260 270 188 QADIERMIK OPINLVADLIPHENGTIGHT VIGOTITICACITIEWRODERDEPEDPDVVFS 208 ACSETRGLHVSPROCEVVROVERROPIWRARSDIRGEIAGGE HEAVSALEYGELKKPONRO
JMTV VP1 T. canis ANT-30	290 300 310 320 330 340 350 137 GKVWTGIKDIGPWPDYSEPTCWRFCDDRAYADHUNTFHCSDBRAYGWDCDEKTOVFLSTYBYSHNDRS 271 CSCVQLKW	(C)	
JMTV VP1 T. canis ANT-30	260 207 RAIANSINTRPVCROLONCTVDRVYCTVGSHKELYPLVEFDEDPLPDGONRUSVRVLPNCRCDDP 328 DBHRAGWDHRERAINNKNSSLACL	JMTV VP3 T. canis ANT-3.2	1 RPGLPGFSVEGGHAQPVPPRSRKTTPTGREGMIPGMMTIAGDVFKRLQGKADIMGESLDIILDEARDLF
JMTV VP1 T. canis ANT-30	410 450 460 470 480 490 277 NOHASLANDASEVCEPKVK RADERVI DOW TPHOPS AGA DOVAPDE AGARGOOS PASTER 373 SEVKARLASE RAVE STOM STELRE HE EKARKET DEVISIONAL	JMTV VP3 T. canis ANT-3.2	80  90  100  110  120  130  140    71  PSSKEFALPOLIRQYLERRDVPLDSAFFLDVFLPLALDOLATTINGTR&AUSTORYUTGGYVTATULAN
JMTV VP1 T. canis ANT-30	500  510  520  530  540  550  560    344  GRNDVG  TROTHOUTE GOVERNMENTRUTGROFDDFILLDSDTHE  TROTHOUTE GOVERNMENTRUTGROFDDFILLDSDTHE  TROTHERDERSTRUCTURE    434  TOLVCGCTAHSWERDERSTRUCTURE  SHOWWERDERSTRUCTURE  SHOWWERDERSTRUCTURE	JMTV VP3 T. canis ANT-3.2	150  160  170  180  190  200  210    141  TSWSSATKATAPARAYOR COVVID: CRUSVEDCTAVA, MAAD VGTTY COPSPERT A MATLAGEA ATTACL  1
JMTV VP1 T. canis ANT-30	570.580.590.610.620.630 406 LHOTVV	JMTV VP3 T. canis ANT-3.2	220  230  240  250  260  270  280    211  Pato framework  Pato framewo
JMTV VP1 T. canis ANT-30	640  650  660  670  680  690  700    467  LH-GL ZOLZWINNIPPWZVAVWTATDIGSAH SPARKUVYTYTYTTRAKADEGTP EDUTSUPPELINNU  1000000000000000000000000000000000000	JMTV VP3 T. canis ANT-3.2	290 300 310 320 330 340 350 265 BRENAVKWD WYNLWYYAR PADLARFY 322 HHENDDG HWYV EVD MAT LUN VV
JMTV VP1 T. canis ANT-30	710 720 730 740 750 760 770 537 ALMANCOR VALPUSLANTAISHLVPMGMCLVG-HATHLLMINDON TALAC HANRGOLVG 631 GHSHAAG-BACGVVSELHIVHWWSELTHARSVRAFTETHASSVVTAASPERREISA-ACLVAMILDHTENE	JMTV VP3 T. canis ANT-3.2	340  370  380  390  400  410  422    328  AUXTMUP DISCENDRUGGINGERING TRADELIZITETIC
JMTV VP1 T. canis ANT-30	780  790  800  810  820  830  840    598  -V.M.C.M.W.S.GRLER.G.QDLARGAE.E.V.V.D.LPESTED. DIPTRCLMODLVGLOYHSOG, W.S.M.O  000  010	JMTV VP3 T. canis ANT-3.2	430 440 450 660 470 480 490 387 HTYERDOR V. WEIRAD. A GAME THE SET WAS YERDING RED A STUDY OF THE SET OF THE
JMTV VP1 T. canis ANT-30	850 860 870 880 890 900 910 664 - DLSASBARRESK GEVAN WARDERS KONSTALL POSY HERMONIC SOCIAL KONSTAND STATE SOCIAL S	JMTV VP3 T. canis ANT-3.2	500  510  520  530  540  550  540    457  HAVGTAAULERE YTHIGSENDYTWAHSVEVERE FERRE AUGUST  DOGMOUP  DOGMOUP  DOGMOUP    367  AVECVINGAUTUCLIES AT SATERIET VALUAL CELLEVER VALUE SATERIET VALUAL CELLEVER VALUE SATERIET VALUAL CELLEVER VALUE SATERIATION AND AND AND AND AND AND AND AND AND AN
JMTV VP1 T. canis ANT-30	920 930 940 733 975 PLING 816 9LG: YEEWBRTTNMKAFKWI XYLEG9 PSF	JMTV VP3 T. canis ANT-3.2	505 LIMMMAYTPISPOLALIGSIYA - KCLGRVINSRT 437 - MUMAVAIVGOLARINETREVERENTPOLVALNVVI

**Figure S6.** Amino acid alignments of the A) VP1, B) VP2 and C) VP3 of JMTV to their corresponding homologs in TCLA, which shows the predicated transmembrane domains (pink lines), serine protease cleavage site (blue arrow), signal peptides (red box), and N-linked glycosylation sites (pink triangles) of JMTV. Also shown are the predicated transmembrane domains (yellow lines), serine protease cleavage site (blue arrow), and signal peptides (red box) of TCLV.



**Figure S7.** Typical serological reactivity of cattle sera with viral antigens in JMTV infected DH82 cells (lower right). The controls are PBS and JMTV negative serum with JMTV infected DH82 cells, and JMTV positive serum with uninfected DH82 cells. The arrows indicate specific staining of cytoplasmatic antigen.



**Figure S8.** The phylogenetic positions (NSP2 gene) of JMTVs detected from cattle blood. The tree is mid-point rooted and all horizontal branches are drawn to a scale of nucleotide substitutions per site. The cattle blood samples are highlighted in red. The taxon labels contain strain names, host (species), and sample locations.



**Figure S9.** ML phylogenetic tree of JMTV NSP2 gene sequences sampled from various geographic locations across China. The tree is mid-point rooted and all horizontal branches are drawn to a scale of nucleotide substitutions per site. For clarity, bootstrap values are only shown for internal nodes. The branch and taxon labels are colored based on arthropod species. The taxon labels contain strain names and sample locations for each pool.



Figure S10. Map of Hubei and Zhejiang provinces showing the locations of sample collection sites.