Recombination Is Associated with an Outbreak of Novel Highly Pathogenic Porcine Reproductive and Respiratory Syndrome Viruses in China

Mang Shi,^a Edward C. Holmes,^{a,b} Manreetpal Singh Brar,^c Frederick Chi-Ching Leung^{c,d}

Sydney Emerging Infections and Biosecurity Institute, School of Biological Sciences and Sydney Medical School, The University of Sydney, NSW, Australia^a; Fogarty International Center, National Institutes of Health, Bethesda, Maryland, USA^b; School of Biological Sciences, The University of Hong Kong, Hong Kong, People's Republic of China^c; Bioinformatics Center, Nanjing Agricultural University, Nanjing, People's Republic of China^d

In 2009 to 2010, there was a marked increase in the number of infections with highly pathogenic porcine reproductive and respiratory syndrome virus (HP-PRRSV) in China. Through phylogenetic analysis, we show that viruses from this outbreak originated from a single recombination event, illustrating the potential importance of this process for disease emergence.

Highly pathogenic porcine reproductive and respiratory syndrome virus (HP-PRRSV) causes a severe reproductive and respiratory syndrome that is distinguished from other PRRSV infections by its high morbidity and motility in pigs of all ages (1). HP-PRRSV first emerged in 2006 from a less-pathogenic variant circulating in southern China and quickly spread throughout the country, causing considerable damage to the swine industry. Shortly after this initial outbreak, a second major HP-PRRSV outbreak occurred in 2009 and 2010, resulting in an increased fatality rate and affecting a wider geographical area, including several Southeast Asian countries (2, 3).

Although the genetic basis of the high pathogenicity and transmissibility of HP-PRRSV is uncertain (4), it is possible to determine what evolutionary forces might be associated with these consecutive outbreaks. To this end, we examined the role played by recombination in generating genetic diversity in HP-PRRSV and explored whether recombination can drive the appearance of disease outbreaks.

We first inferred the epidemiological history of HP-PRRSV in China using 507 complete and partial nonstructural protein 2 (Nsp2) gene sequences collected from GenBank. Sequence alignment was performed using MAFFT v7.0 (5), and a region of 1,128 nucleotides (nt) in length was utilized for phylogenetic analyses using the Bayesian Markov chain Monte Carlo (MCMC) approach available in the BEAST package (6). For this analysis, we utilized a codon-partitioned (GTR112+CP112+Gamma112) substitution model and an uncorrelated lognormal relaxed molecular clock, which resulted in a maximum clade credibility (MCC) tree. To determine how relative genetic diversity in the virus has changed through time, we inferred a Bayesian skyline plot, also utilizing the BEAST package.

The MCC tree revealed, with high resolution, the evolutionary history of HP-PRRSV and its related viruses in China (Fig. 1), from which we can infer the following chain of events. This evolutionary history began with a single viral introduction from North America, followed by (i) the emergence of the generally less pathogenic 1-amino-acid-deletion mutant virus, (ii) the appearance of highly pathogenic PRRSV (i.e., the “1 + 29” deletion mutant HP-PRRSV strain) described previously (1), and finally (iii) the emergence of the novel HP-PRRSV variant, as reported by Zhou et al. (3) and Yu et al. (2). In addition, the Bayesian skyline plot revealed two time periods in which there was a marked in-

crease in relative genetic diversity (Fig. 1) and whose timing and scale matched well with the epidemiological record of HP-PRRSV outbreaks. The first growth phase corresponded to the initial HP-PRRSV outbreak, while the second was mainly associated with the new HP-PRRSV variant (although the phylogeny revealed a number of other, smaller-scale, coincident increases in genetic diversity [Fig. 1]).

To explore the evolution of the new HP-PRRSV variant in more detail, we obtained, from GenBank, the full genome sequences ($n = 34$) of viruses from the new HP-PRRSV cluster (Table 1). Of these, 29 were labeled as highly pathogenic (2, 3), while the rest had unknown virulence. These genomes were aligned in MAFFT with all other full-genome sequences from China. The alignment showed no common indels among the new HP-PRRSV genomes, except for the characteristic “1 + 29” deletion. To test for recombination, we used SIMPLOT v3.5.1 (7), which performs similarity comparisons within a 500-bp window sliding along the genome (10-bp step size). This analysis revealed that new HP-PRRSV genomes possess a mosaic structure indicative of recombination (Fig. 2A). Notably, all 34 genomes shared the same recombination pattern, indicating that the HP-PRRSV outbreak originated from a single recombination event (although we cannot determine if recombination was responsible for the high pathogenicity). Importantly, since the recombination event involves more than one sequence from more than one laboratory, it is unlikely to be a laboratory artifact but rather a clear circulating recombinant lineage.

From the similarity plot, we identified three recombination breakpoints at bp positions 4258, 7876, and 12752 of the sequence alignment (with reference to the Ch-1a strain), which separated the genome into four regions. Phylogenetic trees were then inferred for these regions. Among these phylogenies, those from regions comprising nt 1 to 4258 and nt 7876 to 12752 had similar topologies, as did those of regions comprising nt 4258 to 7876 and

Received 10 May 2013 Accepted 16 July 2013

Published ahead of print 24 July 2013

Address correspondence to Mang Shi, rotund_zheda@hotmail.com.

Copyright © 2013, American Society for Microbiology. All Rights Reserved.

doi:10.1128/JVI.01270-13

FIG 1 Evolutionary history of HP-PRRSV-related viruses circulating in China. (Top) Time-scaled BEAST (MCC) phylogenetic tree. For clarity, posterior probability values are shown only for important nodes. The time to the most recent common ancestor (tMRCA) of specific clades with epidemic importance is shown using dotted lines with labeling. (Bottom) Bayesian skyline plot showing the periods in which relative genetic diversity has increased (highlighted with a gray box). The x axis (time scale) of the Bayesian skyline plot is the same as that of the phylogenetic tree in the upper panel. HPD, highest probability density.

nt 12752 to the end (data not shown; available on request). Hence, despite the presence of three recombination breakpoints, only two parental regions were identified, which were designated A and B (Fig. 2A). In parental region A, the recombinants had an average of 98.2% sequence similarity to sequence HLJHL as opposed to 97.3% sequence similarity to sequences 09HUB1 and -2; in contrast, in parental region B they are more closely related to sequences 09HUB1 and -2 (99.0%) than to HLJHL (98.0%). Importantly, this recombination event was confirmed by phylogenetic analyses (using MrBayes v3.0 [8], employing the GTR+CP+Gamma substitution model) of the different parental regions; in these trees, the recombinants grouped with different parental groups in different genomic regions and with high posterior probabilities (Fig. 2B and C). Because the geographic ranges of

both parental groups are wide (Table 1), it is difficult to determine where the recombination event took place. Interestingly, both parental groups contained strains with high pathogenicity, so that the original recombination event may have involved two highly pathogenic strains, although this will need to be confirmed with more data.

Recombination is not uncommon in PRRSV, and there is growing evidence of recombination from both *in vitro/in vivo* experiments (9–11) and *in silico* analyses of virus gene/genome sequences (12, 13). However, most of the natural recombination detections are represented by single sequences, whose epidemiological and evolutionary significance is likely to be modest. For the first time, our study identifies and characterizes a recombination event associated with a large-scale outbreak of HP-PRRSV. More

TABLE 1 List of sequences used in the recombination analysis

Group	Accession no.	Strain name	Yr	Place ^a of isolation
Recombinant	HM016158	JNHS	2008	Shandong
	HQ315835	BB0907	2009	Guangxi
	JF268680	09HEN2	2009	Henan
	JF268684	09HEN1	2009	Henan
	JF748718	DC	2010	NA
	JF796180	FS	2010	Guangdong
	JN626287	BH5810	2010	Laos
	JQ309822	GX1042	2010	Guangxi
	JQ309823	GX1048	2010	Guangxi
	JQ663540	1010JX	2010	Jiangxi
	JQ663541	1010BJ1	2010	Beijing
	JQ663542	1010BJ3	2010	Beijing
	JQ663544	1010BJ4	2010	Beijing
	JQ663545	1010BJ5	2010	Beijing
	JQ663546	1010FUJ1	2010	Fujian
	JQ663548	1010FUJ3	2010	Fujian
	JQ663549	1010FUJ4	2010	Fujian
	JQ663550	1010FUJ5	2010	Fujian
	JQ663551	1010HEB1	2010	Hebei
	JQ663552	1010HEB2	2010	Hebei
	JQ663553	1010HEB3	2010	Hebei
	JQ663555	1010SD	2010	Shandong
	JQ663557	10LW113	2010	Laos
	JQ663558	1010GX1	2010	Guangxi
	JQ663559	1010GX2	2010	Guangxi
	JQ663560	1010GX3	2010	Guangxi
	JQ663561	1010GX4	2010	Guangxi
	JQ663562	1010GX5	2010	Guangxi
	JQ663563	10LW26	2010	Laos
	JQ663564	10LW37	2010	Laos
	JQ663565	1010LW51	2010	Laos
	JQ663566	10LW66	2010	Laos
	JQ663567	10LW71	2010	Laos
	JQ663568	10LW81	2010	Laos
Parental group 1	HM189676	HLJHL	2009	Heilongjiang
	JF268674	09HUN2	2009	Hunan
	EU144079	SY0608	2006	Jiangsu
	JQ663554 ^b	1010JL	2010	Jilin
	JF268677 ^b	09DB1	2009	Dongbei
	JF268681 ^b	09DB2	2009	Dongbei
Parental group 2	JF268682	09HUB1	2009	Hubei
	JF268683	09HUB2	2009	Hubei
	HM016159	ZP1	2009	Shandong
	GU168568	09HUB5	2009	Hubei
	JN256115	SD0901	2009	Shandong
	JF268672	09SC	2009	Sichuan
	FJ895329	SX2009	2009	Shanxi
Other reference strains (representative)	EU200962	Henan1	2007	Henan
	EU097706	NX06	2007	Ningxia
	EU097707	BJsy06	2006	Beijing
	EF112445	JXA1	2006	Jiangxi
	EU864233	TP	2006	Guangdong
	JX317648	HV	2007	NA
	EU860248	TJ	2006	Tianjin
	EU106888	SHH	2006	Shanghai
	GU269541	GD3	2006	Guangdong
	FJ536165	NB04	2004	Zhejiang
	KF287132	HKUS1	2003	Hong Kong
	KF287136	HKUS7	2004	Hong Kong
	GQ351601	BJ0706	2007	Beijing
	KF287140	HKUS13	2005	Hong Kong
	KF287134	HKUS4	2003	Hong Kong
	KF287139	HKUS12	2004	Hong Kong
	EU864232	SHB	2005	Guangdong
	HQ233605	HB139c	2002	Beijing
	AY032626	CH-1a	1996	Beijing

^a Province of China, with the exception of Laos. NA, not available.^b The three strains are themselves recombinants. They are closely related to HP-PRRSV recombinants in bp 1 to 4258 of the major parental regions but not in bp 7876 to 12752.

FIG 2 Recombination analysis of the new HP-PRRSV cluster. (A) Genome-scale similarity comparisons of the new HP-PRRSV (query) against HLJHL (parental group 1, red), 09HUB1 and -2 (parental group 2, blue), and CH-1a (outgroup, gray). The recombination breakpoints are shown by black dashed lines, with locations indicated at the bottom. The background color of parental region A is white, while that of parental region B is shaded with gray. Phylogenies of parental region A (B) and parental region B (C) are shown below the similarity plot. For clarity, the new HP-PRRSV clusters were “collapsed.” The GenBank accession numbers and strain names of the sequences used in this analysis are shown in Table 1.

importantly, this novel recombination lineage has gradually replaced the other circulating PRRSV populations as the dominant type of HP-PRRSV in China (2). Finally, it is also of note that the recombinants appear to be highly pathogenic, such that the recombination event that generated them either preserved or increased the pathogenicity of the parental strains. In this context, it is useful to note that *in vivo* studies have shown that recombination can sometimes generate viruses of higher pathogenicity than that of the parental strains (D. Liu, R. Zhou, L. Zhou, X. Guo, X. Ge, and H. Yang, presented at the 2011 International PRRS Symposium, Chicago, IL), although it is unclear if this is the case here.

Overall, these findings highlight the important role that recombination has played in PRRSV evolution and epidemiology.

ACKNOWLEDGMENTS

This work was partly supported by grants from the Research Grant Council of the Hong Kong Government (GRF770612), HKU Strategic Research Theme of Infection and Immunology, and Bioinformatics Center, Nanjing Agricultural University, to F.C.-C.L. Mang Shi is supported by an International Postgraduate Research Scholarship from the University of Sydney, while E.C.H. is supported by an NHMRC Australia Fellowship.

No conflicts of interest are declared.

REFERENCES

- Tian KG, Yu XL, Zhao TZ, Feng YJ, Cao Z, Wang CB, Hu Y, Chen XZ, Hu DM, Tian XS, Liu D, Zhang SO, Deng XY, Ding YQ, Yang L, Zhang YX, Xiao HX, Qiao MM, Wang B, Hou LL, Wang XY, Yang XY, Kang LP, Sun M, Jin P, Wang SJ, Kitamura Y, Yan JH, Gao GF. 2007. Emergence of fatal PRRSV variants: unparalleled outbreaks of atypical PRRS in China and molecular dissection of the unique hallmark. *PLoS One* 2:e526. doi:10.1371/journal.pone.0000526.
- Yu X, Chen N, Wang L, Wu J, Zhou Z, Ni J, Li X, Zhai X, Shi J, Tian K. 2012. New genomic characteristics of highly pathogenic porcine reproductive and respiratory syndrome viruses do not lead to significant changes in pathogenicity. *Vet. Microbiol.* 158:291–299.
- Zhou Z, Ni J, Cao Z, Han X, Xia Y, Zi Z, Ning K, Liu Q, Cai L, Qiu P, Deng X, Hu D, Zhang Q, Fan Y, Wu J, Wang L, Zhang M, Yu X, Zhai X, Tian K. 2011. The epidemic status and genetic diversity of 14 highly pathogenic porcine reproductive and respiratory syndrome virus (HP-PRRSV) isolates from China in 2009. *Vet. Microbiol.* 150:257–269.
- Zhou L, Zhang J, Zeng J, Yin S, Li Y, Zheng L, Guo X, Ge X, Yang H. 2009. The 30-amino-acid deletion in the Nsp2 of highly pathogenic porcine reproductive and respiratory syndrome virus emerging in China is not related to its virulence. *J. Virol.* 83:5156–5167.
- Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30:772–780.
- Drummond AJ, Rambaut A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* 7:214. doi:10.1186/1471-2148-7-214.
- Lole KS, Bollinger RC, Paranjape RS, Gadkari D, Kulkarni SS, Novak NG, Ingersoll R, Sheppard HW, Ray SC. 1999. Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. *J. Virol.* 73:152–160.
- Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Hohna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP. 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* 61:539–542.
- Liu D, Zhou R, Zhang JL, Zhou L, Jiang QY, Guo X, Ge XN, Yang HC. 2011. Recombination analyses between two strains of porcine reproductive and respiratory syndrome virus *in vivo*. *Virus Res.* 155:473–486.
- van Vugt JJFA, Storgaard T, Oleksiewicz MB, Botner A. 2001. High frequency RNA recombination in porcine reproductive and respiratory syndrome virus occurs preferentially between parental sequences with high similarity. *J. Gen. Virol.* 82:2615–2620.
- Yuan SS, Nelsen CJ, Murtaugh MP, Schmitt BJ, Faaborg KS. 1999. Recombination between North American strains of porcine reproductive and respiratory syndrome virus. *Virus Res.* 61:87–98.
- Li B, Fang L, Xu Z, Liu S, Gao J, Jiang Y, Chen H, Xiao S. 2009. Recombination in vaccine and circulating strains of porcine reproductive and respiratory syndrome viruses. *Emerg. Infect. Dis.* 15:2032–2035.
- Shi M, Lam TT, Hon CC, Hui RK, Faaborg KS, Wennblom T, Murtaugh MP, Stadejek T, Leung FC. 2010. Molecular epidemiology of PRRSV: a phylogenetic perspective. *Virus Res.* 154:7–17.