

Genomic recombination in primate bocavirus: inconsistency and alternative interpretations

Yee Ling Chong¹ · Kar Hon Ng¹

Received: 21 January 2017 / Accepted: 18 April 2017 / Published online: 29 April 2017
© Springer Science+Business Media New York 2017

Abstract Human bocavirus (HBoV) is a single-stranded DNA virus in *Parvoviridae* family, causing respiratory diseases in human. The recent identifications of genomic recombination among the four human bocavirus genotypes and related non-human primate bocaviruses have shed lights into the evolutionary processes underpinning the diversity of primate bocavirus. Among these reports, however, we found inconsistency and possible alternative interpretations of the recombination events. In this study, these recombination events were reviewed, and the related genome sequences were re-analysed, aiming to inform the research community of bocavirus with more consistent knowledge and comprehensive interpretations on the recombination history of primate bocavirus.

Keywords Primate · Bocavirus · Recombination · Evolution

Edited by Detlev H. Kruger.

Yee Ling Chong and Kar Hon Ng have contributed equally to this study.

Availability of data and materials The datasets analysed during the current study are available in the GenBank repository, with the accession numbers indicated before the strain names in the figure.

✉ Yee Ling Chong
ylchong@unimas.my
Kar Hon Ng
nkh_91@hotmail.com

¹ Faculty of Resource Science and Technology, Universiti Malaysia Sarawak, 94300 Kota Samarahan, Sarawak, Malaysia

Abbreviations

BoV Bocavirus
HBoV Human bocavirus

Main text

Human bocavirus (HBoV) is a single-stranded DNA virus in *Parvoviridae* family, which was first identified in human respiratory tract samples in 2005 [1]. HBoV was then subsequently reported in many respiratory diseases surveillances [2–5], which led to the discovery of four genotypes (HBoV1–4). Later, some non-human primate bocaviruses have been identified, raising the possibility of animal origin for the ancestral bocavirus. Brožová et al. recently reported a new primate bocavirus from chimpanzee and suggested that the genome of this virus (CPZh2-BoV) has a mosaic origin recombining between human bocavirus genotype 3 (HBoV3) and human bocavirus genotype 1 (HBoV1) or a gorilla bocavirus (GBoV) [6]. Indeed, recombination seems to play a major role in the evolution and emergence of new bocavirus variants as a number of research papers have reported inter- and intra-genotypic recombinations of bocavirus, for example, suggesting the recombinant origin of HBoV2 and HBoV3 [2, 7, 8]. However, some of these reports have presented conflicting data or have not fully elaborated all the possible interpretations on the data. Here, we reviewed and analysed the published papers and data about recombination among primate bocavirus genotypes, to provide more comprehensive interpretation on the recombination history of these viruses. After preliminary analysis of all complete genome sequences of bocaviruses aligned using MUSCLE v3.5 [9], representative strains were kept (17 HBoV and 2 non-human primate bocaviruses) to produce clear informative

phylogenetic trees. All phylogenetic trees were inferred using maximum likelihood (ML) method and generalised-time-reversible nucleotide substitution model with four categories of gamma rate variation and invariable proportion (GTR + I + Γ_4) implemented in PhyML v.3 [10] with bootstrap tests on 1000 pseudo-replicates.

Is HBoV2 a recombinant of HBoV1 and HBoV4?

Fu et al. performed recombination analysis on the complete genome sequences of bocavirus and reported that HBoV2 is a recombinant of HBoV1 and HBoV4 [7]. Such conclusion was referred by papers published later on including Babkin et al. [11]. However, it is noteworthy that those recombinant bocavirus strains labelled as HBoV2 in Fu et al. study [7] actually belonged to HBoV3 lineage [2, 8]. Such inconsistent labelling of genotype should be noted and not passed on to future literature. Although there is no solid evidence indicating that the entire HBoV2 lineage originated from ancestral recombination between other HBoV genotypes, intra-genotypic recombination among HBoV2 strains has been suggested [3].

Is HBoV3 a recombinant between HBoV1 and HBoV2/HBoV4?

There are an overwhelming number of reports of recombinant origin of HBoV3, all suggesting the breakpoint at approximately the junction between NP1 and VP1 genes [2, 3, 8]. The 5' region on the left side of the breakpoint was believed to originate from HBoV1, whereas the 3' region was discordantly suggested to originate from HBoV2 and HBoV4 by Kapoor et al. [8] and Cheng et al. [2], respectively. We found that in the 3' region, the HBoV3 viruses form monophyletic lineage with HBoV4, yet with a weak bootstrap support (bootstrap = 57%; Fig. 1E), and indeed our BOOTSCAN analysis [12] has shown that the 3' region has a mixture of phylogenetic signals for clustering with either HBoV2 or HBoV4 (Fig. 1b). Such weak clustering between HBoV3 and HBoV4 is also shown by Babkin et al. [11]; however, another study [2] suggests a good clustering. It is hard to judge and compare between studies as they used different phylogenetic methods and sequence regions for the tree analysis. Nonetheless, considering all the existing data and studies, it is not affirmative to distinguish between HBoV2- and HBoV4-origins; and indeed, more data in future may still fail to do so if the HBoV3 recombination has occurred early around the time when the ancestral bocavirus just started diverging into HBoV2 and HBoV4.

On the other hand, for the 5' region, although the HBoV1-origin was consistently proposed by different groups [2, 8], the recently published genome of a non-

human primate bocavirus (CPZh2-BoV) has suggested an alternative hypothesis. It is notable that in the SIMPLOT and BOOTSCAN analyses this HBoV3 now showed higher sequence similarity and probability of clustering with CPZh2-BoV in the right side (850-2760 nt) of this 5' region (Fig. 1a, b), which used to be clustering with HBoV1 as is in the left side (1-550 nt) of this 5' region before new CPZh2-BoV was reported. The tree topologies built from these two regions also support such clustering (Fig. 1c, d). Therefore, these results raise the possibility that the entire genome of HBoV3 might have an extra origin (from CPZh2-BoV) for the central region (around 850-2,760nt) of the genome, in addition to the two, i.e. HBoV1 and HBoV2/4, which have been previously proposed (three recombinant regions are shaded in Fig. 1a, b).

Is CPZh2-BoV a recombinant between HBoV3 and HBoV1/GBoV?

The new CPZh2-BoV genome sequence from chimpanzee reported by Brožová et al. [6] was proposed to be a recombinant between HBoV3 and HBoV1/GBoV, with high sequence identity and clustering confidence to HBoV3 at, notably, the same region (i.e. right side of 5' region; Fig. 1f) we observed when HBoV3 was used as the query in the BOOTSCAN analysis that shows high similarity to CPZh2-BoV (right side of 5' region; Fig. 1a). Such complement pattern is commonly seen when recombinant and its parent were separately used as queries in BOOTSCAN/SIMPLOT analysis [13]. In this case, either HBoV3 or CPZh2-BoV might be the parent of the other in the 5'-region: if HBoV3 has recombined from HBoV1 and CPZh2-BoV (black arrows in Fig. 1c, d), then CPZh2-BoV would not have acquired HBoV3 from in this region. Otherwise, the previous hypothesis by Brožová et al. for recombinant origin of CPZh2-BoV (green arrows in Fig. 1c, d) would be true [6]. The observed topologies are indeed compatible to both hypotheses. Considering all the evidence, in addition to the previous hypotheses, we found that the current genetic data also support that (i) HBoV3 is a recombinant between HBoV1, CPZh2-BoV and HBoV2/HBoV4, and (ii) CPZh2-BoV is a recombinant between only GBoV and HBoV1 without any involvement of HBoV3. Nevertheless, it is noteworthy that the original and our new hypotheses are equivalent and difficult to verify without additional data. This indistinguishability is particularly caused by the lack of full genome sequence data of primate bocavirus. It is expected that the more primate bocavirus genome sequences that will be available in the future would allow more insightful analyses and conclusions on the recombination between human and non-human primate bocaviruses, as well as the crossing of species barriers for

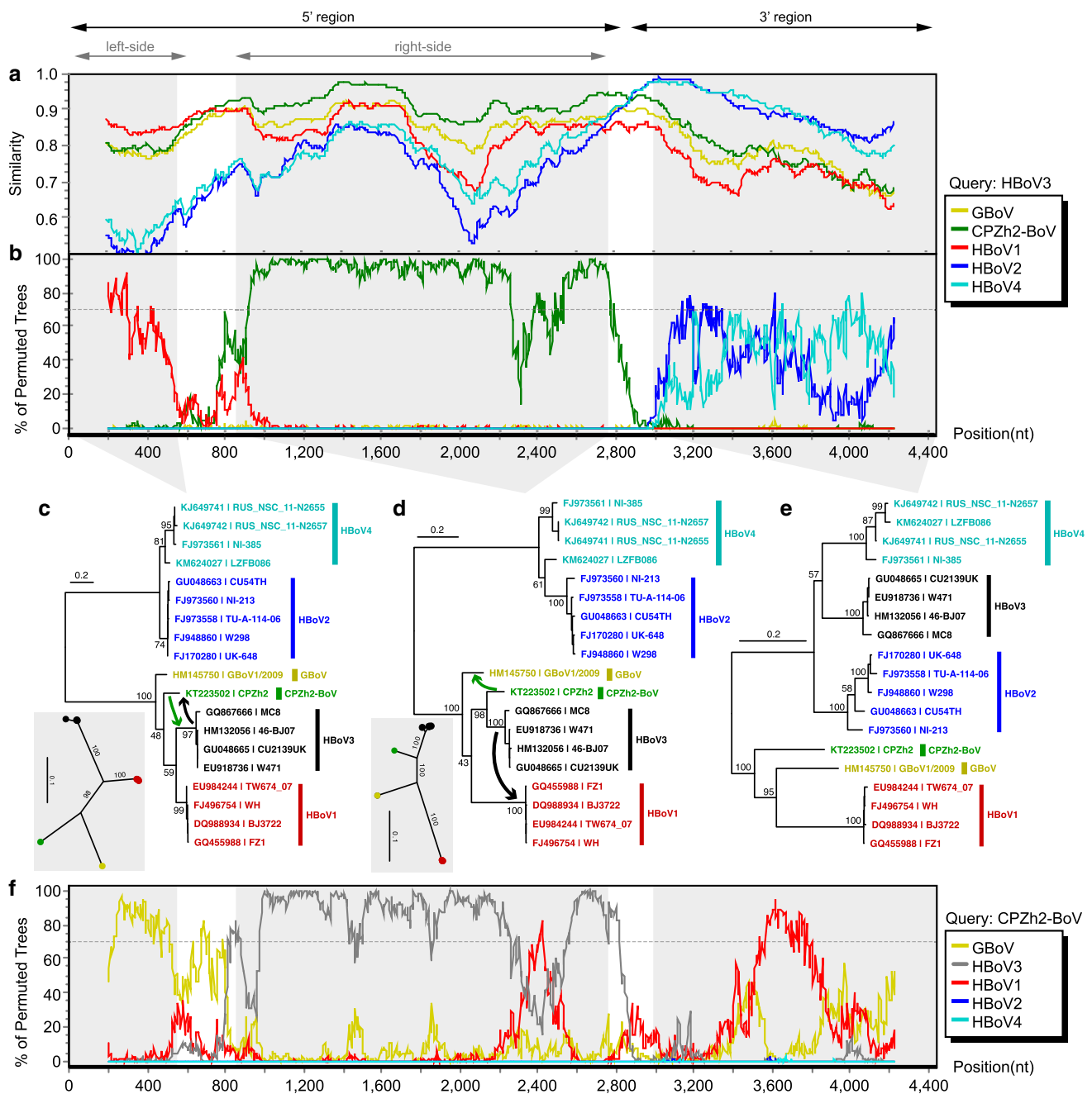


Fig. 1 Recombination and phylogenetic analyses of primate bocavirus. **a** SIMPLOT and **b** BOOTSCAN analyses using HBoV3 as query. **c–e** Maximum-likelihood (ML) phylogenetic trees built from the three shaded regions separated by the putative recombination breakpoints in two approximate areas (unshaded). The green and black arrows indicate the recombining lineages for CPZh2-BoV and HBoV3, respectively, in two alternative hypotheses. The arrows are pointing to the lineage where the recombinant virus acquired the other genomic region from. The hypothesis indicated by green arrow was

previously raised by Brožová et al. [6]. Bocavirus genotypes are indicated by different colours shown in the legend box. Small radial ML trees of only HBoV3, HBoV1, GBoV and CPZh2-BoV are displayed in the left bottom inset, showing that their relative clustering in the tree topology are indeed strongly supported (all > 95) under bootstrapping test, in the absence of influence by distant genotypes HBoV4 and HBoV2. **f** BOOTSCAN analysis using CPZh2-BoV as query (Color figure online)

interspecies transmission observed in the polyphyletic origin of chimpanzee bocavirus (Fig. 1).

In summary, we provided genomic sequence analyses of bocavirus to inform the research community with more

consistent knowledge and comprehensive interpretations on the recombination history of primate bocavirus. Moreover, it is noteworthy that the recently identified head-to-tail sequences in HBoVs were hypothesized to derive from

ancestral recombination of bovine parvovirus and canine minute virus [14–16]. A comprehensive study of the head-to-tail sequence patterns among bocaviruses/bocaparvoviruses from different host species may yield further insights into the zoonosis pathway of the virus.

Acknowledgements This paper is supported by the Malaysia Ministry of Education under Fundamental Research Grant Scheme (FRGS/STG35(02)/963/2013(04)). The authors thank the researchers for sharing their bocavirus genome sequences in GenBank.

Compliance with Ethical Standards

Conflict of interest The authors declare that they have no competing interests.

References

1. T. Allander, M.T. Tammi, M. Eriksson, A. Bjerkner, A. Tiveljung-Lindell, B. Andersson, *Proc. Natl. Acad. Sci. USA* **102**, 12891–12896 (2005)
2. W. Cheng, J. Chen, Z. Xu, J. Yu, C. Huang, M. Jin, H. Li, M. Zhang, Y. Jin, Z.J. Duan, *BMC Infect. Dis.* **11**, 50 (2011)
3. A. Kapoor, E. Slikas, P. Simmonds, T. Chieochansin, A. Naeem, S. Shaikat, M.M. Alam, S. Sharif, M. Angez, S. Zaidi, E. Delwart, *J. Infect. Dis.* **199**, 196–200 (2009)
4. N. Lasure, V. Gopalkrishna, *J. Med. Virol.* **89**, 17–23 (2017)
5. D.M. Zhang, M.M. Ma, W.T. Wen, X. Zhu, L. Xu, Z.J. He, X. He, J.H. Wu, Y.W. Hu, Y. Zheng, Y. Deng, C.J. Lin, J.H. Lu, M.F. Li, K.Y. Cao, *Epidemiol. Infect.* **143**, 2315–2329 (2015)
6. K. Brozova, K. Hrazdilova, E. Slaninkova, D. Modry, J. Cerny, V. Celer, *Infect. Genet. Evol.* **37**, 231–236 (2016)
7. X. Fu, X. Wang, B. Ni, H. Shen, H. Wang, X. Zhang, S. Chen, S. Shao, W. Zhang, *Virol. J.* **8**, 182 (2011)
8. A. Kapoor, N. Mehta, F. Esper, M. Poljsak-Prijatelj, P.L. Quan, N. Qaisar, E. Delwart, W.I. Lipkin, *PLoS ONE* **5**, e11948 (2010)
9. R.C. Edgar, *BMC Bioinform.* **5**, 113 (2004)
10. S. Guindon, F. Delsuc, J.F. Dufayard, O. Gascuel, *Methods Mol. Biol.* **537**, 113–137 (2009)
11. I.V. Babkin, A.I. Tyumentsev, A.Y. Tikunov, A.M. Kurilshikov, E.I. Ryabchikova, E.V. Zhirakovskaya, S.V. Netesov, N.V. Tikunova, *Infect. Genet. Evol.* **14**, 265–274 (2013)
12. D.P. Martin, D. Posada, K.A. Crandall, C. Williamson, *A.I.D.S. Res. Hum. Retrovir* **21**, 98–102 (2005)
13. D.P. Martin, P. Lemey, M. Lott, V. Moulton, D. Posada, P. Lefevre, *Bioinformatics* **26**, 2462–2463 (2010)
14. J. Lusebrink, V. Schildgen, R.L. Tillmann, F. Wittleben, A. Bohmer, A. Muller, O. Schildgen, *PLoS ONE* **6**, e19457 (2011)
15. A. Kapoor, M. Hornig, A. Asokan, B. Williams, J.A. Henriquez, W.I. Lipkin, *PLoS ONE* **6**, e21362 (2011)
16. O. Schildgen, J. Qiu, M. Soderlund-Venermo, *Future Virol* **7**, 31–39 (2012)