Phylogenetic evidence for homologous recombination within the family Birnaviridae

Chung-Chau Hon, Tommy Tsan-Yuk Lam, Chi-Wai Yip, Raymond Tsz-Yeung Wong, Mang Shi, Jingwei Jiang, Fanya Zeng and Frederick Chi-Ching Leung

School of Biological Sciences, The University of Hong Kong, Hong Kong SAR

Birnaviruses are bi-segmented double-stranded RNA (dsRNA) viruses infecting insects, avian species and a wide range of aquatic species. Although homologous recombination is a common phenomenon in positive-sense RNA viruses, recombination in dsRNA viruses is rarely reported. Here we performed a comprehensive survey on homologous recombination in all available sequences (>1800) of the family Birnaviridae based on phylogenetic incongruence. Although inter-species recombination was not evident, potential intra-species recombination events were detected in aquabirnaviruses and infectious bursal disease virus (IBDV). Eight potential recombination events were identified and the possibility that these events were non-naturally occurring was assessed case by case. Five of the eight events were identified in IBDVs and all of these five events involved live attenuated vaccine strains. This finding suggests that homologous recombination between vaccine and wild-type IBDV strains may have occurred; the potential risk of mass vaccination using live vaccines is discussed. This is the first report of evidence for homologous recombination within the family Birnaviridae.

INTRODUCTION

Members of the family Birnaviridae are non-enveloped, bi-segmented double-stranded RNA (dsRNA) viruses infecting insects, avian species and a wide range of aquatic species (Delmas et al., 2005). Birnaviruses can be classified into five phylogenetic clusters, namely tellina virus 1 (TV-1), blotched snakehead virus (BSNV), Drosophila X virus (DXV), infectious bursal disease virus (IBDV) and aquabirnaviruses (AQBV) (Nobiron et al., 2008). The genome segment A of birnaviruses (~3 kb) mainly encodes a pVP2-4-3 polyprotein and the genome segment B (~2.8 kb) encodes an RNA-dependent RNA polymerase (RdRp) named VP1 (Delmas et al., 2005). Homologous proteins of birnaviruses share only ~20 to 50 % amino acid identity (Da Costa et al., 2003; Nobiron et al., 2008). Intraspecies diversity of IBDV and AQBV has been characterized extensively, while the diversity of DXV, BSNV and TV-1 is currently unknown. AQBV can be classified into two serotypes and ten subserotypes (A1–A9 and B1) (Hill & Way, 1995), which are correlated with the seven genogroups that are classified based on the genome segment A (Nishizawa et al., 2005). IBDV can be classified into two serotypes. Serotype-I strains can be phenotypically classified into classical, variant, very virulent, attenuated and Australian lineages (Van den Berg, 2000). Serotype-II strains are non-pathogenic and less characterized (Van den Berg, 2000). Evidence for genome segment reassortment among IBDV strains has been documented (Gao et al., 2007; Le Nouen et al., 2006; Wei et al., 2006, 2008). However, the possibility of homologous recombination among IBDV or AQBV strains was not addressed previously.

Recombination is a relatively common phenomenon in positive-sense RNA viruses (Chare & Holmes, 2006) but it appears to be rare in negative-sense RNA viruses. It was explained by the unavailability of their negative-sense RNA genomes for template-switching due to encapsidation by ribonucleoprotein complex (Chare et al., 2003). In birnaviruses, the genomic materials are associated with unique threadlike ribonucleoprotein complexes independent of the capsid (Hjalmarsson et al., 1999), in contrast with the observation that the dsRNA segments are free within virions in viruses of the family Reoviridae. Nonetheless, recombination in dsRNA viruses has not been studied systematically, except for some sporadic reports of recombination events in rotavirus (Phan et al., 2007). Recombination not only complicates the genotyping systems of viruses, it may also affect analyses of the molecular clock, demographic processes or natural selection. Therefore, the extent of recombination in dsRNA viruses deserves further investigations. Here we performed
a conservative yet comprehensive survey on potential homologous recombination within the family **Birnaviridae**, at both inter- and intra-species level.

**METHOD**

**Datasets.** All sequences of the family **Birnaviridae** \((n=1881)\) were downloaded from GenBank in December, 2007. Alignments of pVP2-4-3 polyprotein and RdRp coding sequences were generated using **CLUSTAL W** based on their amino acid sequences. For inter-species alignments, full-length coding sequences of all birnavirus reference strains were used. For intra-species alignments of IBDV and AQBV, alignments were generated from different genomic regions depending on the availability of the GenBank sequences. With these overlapping alignments, we aimed to maximize the number of sites and taxa being analysed. All datasets are summarized in Table 1.

**Classification of IBDV and AQPV.** Neighbour-joining (NJ) phylogenies were constructed for datasets specified in Table 1. These phylogenies were constructed under the maximum composite likelihood substitution model using **MEGA v4.0** (1000 bootstrap replicates) (Tamura et al., 2007). Maximum-likelihood (ML) phylogenies were then constructed for selected strains using **PHYML v2.4.4** (1000 bootstrap replicates) (Guindon & Gascuel, 2003) under the best-fit substitution model suggested by **MODELTEST v3.7** (Posada & Crandall, 1998).

**Screening for potential recombinants.** Four methods implemented in the recombination detection program (RDP) v2.0, RDP, GENECONV, MaxChi and Chimaera (Martin et al., 2005), were used for screening of potential recombinants as described by Ma et al. (2007). Phylogenetic incongruence between different genome regions (sequences in Supplementary Material S1, available in JGV Online) was further analysed using Bootscan implemented in Simplot v3.5.1 (Lole et al., 1999).

**Identification of breakpoints.** Breakpoint locations were preliminarily estimated based on the distribution of informative sites supporting the two incongruent topologies which maximizes the \(\chi^2\) value (Robertson et al., 1995); this method was implemented in Simplot. To locate the breakpoint with statistical significance, we used a likelihood method implemented in **LARD** (Holmes et al., 1999). Exhaustive searches for a single breakpoint, rather than heuristic searches for multiple breakpoints, were used to enhance the precision of our estimation. Therefore, for recombinants with multiple breakpoints, the alignments were trimmed so that only one breakpoint was estimated at a time. Significance of the estimated breakpoints was assessed using a likelihood ration test (LRT) (Holmes et al., 1999). To investigate whether the mosaic genome structure was a result of convergent evolution, the proportion of informative sites that are third-codons was calculated (Suzuki et al., 1998).

**Identification of the closest parental strains.** For recombinants with multiple breakpoints, genome regions with the same parental ancestry (i.e. parental regions) were concatenated manually. The closest parental strains are defined as strains that share the highest nucleotide sequence identity with the potential recombinant. For each of the parental regions, a ML phylogeny was constructed for selected taxa using **PHYML** as described above.

**RESULTS AND DISCUSSION**

**Lack of evidence for recent inter-species recombination**

RDP results suggest that there is no significant evidence of recombination in datasets ALL-A and ALL-B. ML phylogenies were then constructed for different genome regions. Although discordant topologies were barely observed,
most of these phylogenies were poorly resolved (data not shown). This observation is likely to be caused by the relatively high divergence among birnaviruses, e.g. ~19% amino acid similarity for VP3 (Nobiron et al., 2008), which could seriously affect the accuracy of alignment. Such high divergence prevents us from confidently drawing a conclusion about the possibility of any ancient inter-species recombination events among these birnaviruses. Nonetheless, our results suggest that these birnaviruses were not originated from recent inter-species recombination events.

Phylogenies of AQBV and IBDV

AQBV genome segment A (Fig. 1a) can be classified into seven lineages (designated A-I to A-VII) as described previously (Nishizawa et al., 2005). The phylogeny of AQBV genome segment B (Fig. 1b) is compatible to that of its genome segment A (i.e. no apparent reassortment), thus the lineages were named after those in the segment A phylogeny as B-I, B-II, B-III and B-VII.

In Fig. 1(c), IBDV genome segment A can be classified into serotype-II (A-SEII), Australian (A-AUS), very virulent (A-VV), variant (A-VAR), attenuated (A-ATT) and four classical lineages (A-CLA1 to A-CLA4), according to the typical classification (Van den Berg, 2000). On the other hand, the phylogeny in Fig. 1(d) suggest the genome segment B of very virulent IBDV strains (vvIBDV) form an independent lineage (designated B-VV), except for some sporadic reassortants reported previously (Gao et al., 2007; Le Nouen et al., 2006; Wei et al., 2006, 2008). Genome segment B of the non–very-virulent (B-NVV) IBDVs can be further divided into two lineages designated B-NVVI and B-NVVII (Hon et al., 2006). Incongruence between

---

**Fig. 1.** Classification of AQBV (a and b) and IBDV (c and d) segments A (a and c) and B (b and d). ML phylogenies were constructed for selected taxa from datasets specified in Table 1. Lineage names are indicated on the right. For segment B phylogenies, the segment A genotype of the taxa are also indicated. The lineages that are well-supported with bootstrap values >80% are indicated with asterisks. The serotype-II IBDV strains in (c) were omitted for simplicity.
Table 2. Details of recombination events

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Recombinant (genotype of another genome segment)</th>
<th>Major parental lineage: closest parental strain; length (bp), identity (%)</th>
<th>Minor parental lineage: closest parental strain; length (bp), identity (%)</th>
<th>Simplot breakpoints† (3rd codon‡, (\chi^2))</th>
<th>LARD breakpoints‡‡; (likelihood ratio of A to B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AQBV-A1</td>
<td>AF342729_Ab (lineage B-II)</td>
<td>A-II; AJ489223_2284; 2563, 99.7</td>
<td>A-I; AF342727_WestBuxton; 339, 97.6</td>
<td>A:363–384; B:705–733 (72, 174.7)</td>
<td>A:384, B:725 (48.1 : 140.3)</td>
</tr>
<tr>
<td>AQBV-B1</td>
<td>AY780931_20G1d (lineage A-I)</td>
<td>B-II; AY780926_6B1a; 2225, 99.8</td>
<td>B-III; AY379743_NVI-016; 310, 99.7</td>
<td>A:1236–1284; B:1524–1557 (71, 108.2)</td>
<td>A:1240, B:1549 (69.0 : 76.6)</td>
</tr>
<tr>
<td>IBDV-A1</td>
<td>AF165150_KK1 (NA)</td>
<td>A-ATT; AF499929_D78; 1857, 99.1</td>
<td>A-VV; AF240686_D6948; 1062, 99</td>
<td>A:510–607; B:1647–1747 (81, 59.5)</td>
<td>A:607, B:1668 (27.5 : 56.7)</td>
</tr>
<tr>
<td>IBDV-A1</td>
<td>AF165151_KSH (NA)</td>
<td>A-ATT; AF499929_D78; 1858, 99</td>
<td>A-VV; AF533670_SH.92; 1061, 99.1</td>
<td>A:510–607; B:1646–1755 (81, 46.1)</td>
<td>A:607, B:1667 (26.4 : 52.4)</td>
</tr>
<tr>
<td>IBDV-A2</td>
<td>EU328333_SSh (NA)</td>
<td>A-VV; EU042141_HLJ-5; 624, 99.4</td>
<td>A-ATT; AF499929_D78; 634, 99.8</td>
<td>A:507–636; B:868–961 (86, 16.7)</td>
<td>636 (29.2)</td>
</tr>
<tr>
<td>IBDV-A2</td>
<td>X95883_849VB (lineage B-VV)</td>
<td>A-ATT; AF499929_D78; 703, 100</td>
<td>A-VV; AF240686_D6948; 559, 96.6</td>
<td>A:282–405; B:903–989 (83, 25.1)</td>
<td>A:405, B:963 (20.6 : 17.4)</td>
</tr>
<tr>
<td>IBDV-B1</td>
<td>EF517529_Harbin-1 (lineage A-VV)</td>
<td>B-NVV2; Y705393_Gx; 1303, 96.5</td>
<td>B-NVV2; J878671_HENAN; 525, 97.7</td>
<td>993–1185 (91, 22.9)</td>
<td>1063 (57.6)</td>
</tr>
</tbody>
</table>

*GenBank accession number of the closest strain and the percentage nucleotide identity between the corresponding parental region of the closest parental strain and that of the recombinant are given.

†Breakpoint locations were numbered according to the corresponding datasets in Table 1.

‡Proportion of informative sites that are third codons as a percentage.

§Maximum \(\chi^2\) value calculated using Simplot; for all values, \(P<0.001\).

‖Likelihood ratio (LR) of the single (A) or double (B) breakpoints calculated using LARD; for all values, \(P<0.001\).
these two phylogenies has been explained as a result of reassortment (Brown & Skinner, 1996).

**Identification of intra-species recombination**

Eight recombinants were identified from both genome segments of AQBV and IBDV (Table 2). Only recombination events identified by both Bootscan and LARD analyses with statistical significance ($P<0.001$ in $\chi^2$ test and LRT) are included in Table 2. For all recombinants, the majority (71 to 94 %) of the informative sites are third-codons (Table 2), suggesting that the detected mosaic genome structures were likely to be a result of recombination rather than convergent evolution due to selection such as substitutions in VP2 during tissue culture adaptation.

**Recombination in both genome segments of AQBV**

Two potential AQBV recombinants (Ab and 20G1d) were identified (Table 2). Strain Ab (isolated from trout in Denmark) was previously classified as A3 serotype (or lineage A-II in this study) (Blake et al., 2001). Our results suggest that its genome segment A is a recombinant between lineage A-II and A-I. Its closest parental strains are 2284 (trout/Spain) and WestBuxton (trout/USA) (Fig. 2a). In contrast, the genome segment B of Ab (AM114033) is clustered within lineage B-II (Table 2; Supplementary Fig. S1b, available in JGV Online). These results suggest that the genesis of Ab may involve recombination between a lineage I and a lineage II strain.

The genome segment B of 20G1d (deepwater redfish/Spain) was identified as a recombinant between lineage B-II and B-III. Its closest parental strains are 6B1a (Atlantic cod/Spain) and NVI-016 (trout/Norway) (Fig. 2b). Interestingly, the genome segment A of 20G1d is clustered within lineage A-I (Table 2; Supplementary Fig. S1a). This finding suggests that a reassortment event may have been involved in the genesis of 20G1d, in addition to the recombination events within its genome segment B.

**Recombination in both genome segments of IBDV**

Six potential IBDV recombinants were identified (Table 2). Interestingly, all potential recombination events in genome segment A involved the A-VATT lineage, which mainly consists of live attenuated vaccine strains. In particular, strain KK1, KSH, SHh and 849VB were identified as potential recombinants between lineage A-VV and A-VATT. Their closest A-VATT parental strain was identified as D78, which is a widely distributed live attenuated vaccine (Nobilis Gumboro D78 Live, Intervet). The two Korean potential recombinant strains, KK1 and KSH, were previously identified as divergent vvIBDV strains (Kwon et al., 2000; Mardassi et al., 2004). Breakpoints of these two potential recombinants were estimated at almost identical locations (Table 2), suggesting that they were likely to have originated from the same recombination events. Strain KSH and KK1 were isolated in 1992 and 1997, respectively (Kwon et al., 2000). The divergence between them could be explained by the 5-year period between their isolation. The closest A-VV parental strain of KSH is SH.92, which is a Korean vvIBDV isolated in 1992 (Kim et al., 2004). Strain SH.92 is geographically and temporally correlated with KSH, suggesting that KSH and KK1 may be originated from recombination events between a Korean vvIBDV and a live vaccine strain during the early 1990s.

Potential recombinant 849VB is a European vvIBDV that was isolated from Belgium in 1987 (Van den Berg et al., 1996). It is geographically and temporally consistent with its closest A-VV parental strain, D6948, which was isolated from the Netherlands in 1989. Moreover, its genome segment B is almost identical to other European vvIBDVs that were isolated during the late 1980s (Le Nouen et al., 2006), suggesting that 849VB might have originated from recombination between early European vvIBDVs and live vaccine strains (Fig. 2d). In contrast, potential recombinant SHh and its closest A-VV parental strain HJL-5 were both isolated from China (Fig. 2e), suggesting that recombination between Chinese vvIBDVs and live vaccine strains may have occurred.

Potential recombinant ViBursaCE is a tissue culture attenuated vaccine strain derived from a non-cloned master seed (Lohmann Animal Health). Our results suggest that it may be a potential recombinant between lineage A-VAR and A-VATT (Fig. 2f). Its closest A-VAR parental strain is Variant E, which is an American antigenic variant. Its closest A-VATT parental strain is strain-CT, which is a French vaccine strain (Rhone-Merieux). Its genome segment B (GenBank accession number EU162092) shares over 99.8 % nucleotide sequence identity with strain-CT (GenBank accession number AJ310186) (Table 2; Supplementary Fig. S1d). Taken together, ViBursaCE may have originated from a vaccine strain, and its genome segment A may have recombined with a variant strain. Since the genome sequence strains of ViBursaCE were obtained from samples directly taken from the commercial

**Fig. 2.** Phylogenetic incongruence between the parental regions of recombinants (indicated by ‘Query’) in (a) to (g). The left column shows Bootscan (bottom plot) and LARD (upper bar) results for the recombinants. Major and minor parental strains are indicated with grey and black lines or bars, respectively. Outgroups in Bootscan are represented by dotted lines. Four taxa were used in Bootscan (indicated on top of the plot), while the outgroups were excluded in LARD. The middle and right columns refer to the ML phylogenies based on the major and minor parental regions, and the recombinants are highlighted with black circles. All phylogenies were rooted with the indicated outgroups except (g), in which HENAN was absent from the phylogeny since the sequence of its 5’ region was not available. Bars, number of substitutions per site.
vaccine vials (personal communication, Dr D. J. Jackwood, Ohio State University), the potential recombination events may have occurred during propagation of the vaccine.

The genome segment B of Harbin-1 was identified as a potential recombinant between strains within the B-NVV2 lineage. Its closest parental strains were identified as Gx
and HENAN (Fig. 2g). It should be noted that the full-length sequence of HENAN is not available; therefore, the closest parental strain in dataset IB DV-B1, 2.73, was used in the Bootscan and LARD analyses (Fig. 2g). The genome segment A of Harbin-1, Gx and HENAN (GenBank accession numbers EF517528, AY444873 and AJ878901) were clustered within the A-VV lineage and thus were identified as Chinese vvIBDVs (Supplementary Fig. S1c). In particular, the corresponding parental region of Harbin-1 and Gx shares over 99.8% nucleotide identity (Table 2). Taken together, these data suggest that Harbin-1 may be originated from strains closely related to Gx and its genome segment B may have recombined with strains closely related to HENAN. This result suggests that recombination between the genome segment B of Chinese vvIBDVs has occurred.

**Possibilities that these potential recombination events were non-naturally occurring**

Despite the apparent phylogenetically incongruent signals within the genomes of potential recombinants, care has to be taken to interpret whether these signals were non-naturally occurring, i.e. as a result of recombination in cell culture coinfected with multiple strains or PCR recombination in the presence of multiple templates. In both scenarios, these *in vitro* recombinants might not be picked up by direct sequencing of the PCR product, since they only represent a minor portion of the PCR amplicon and the resulting sequence is a consensus sequence of the overall population of amplicons (assuming the recombinants have no selection advantage in replication and PCR amplification). However, if the amplicons were cloned, these *in vitro* recombinants could be picked up as rare genotypes among the parental genotypes. In the cases of Ab and 20Gld, the sequences were obtained using a direct PCR sequencing strategy (Blake et al., 2001; Romero-Brey et al., 2004). In addition, for 849VB, six independent clones were sequenced although the amplicons were cloned before sequencing (Van den Berg et al., 1996). Therefore, these three potential recombinants are not likely to have been derived *in vitro*. Nonetheless, we recognize the uncertainty in these published sequences and confess that our interpretation may not be conclusive if non-random errors were introduced during data processing, before the sequences were submitted to GenBank, e.g. accidental mixing of files belonging to different viruses during sequence assembly, leading to *in silico* recombination (de Silva & Messer, 2004). However, we consider such non-random sequencing errors as individual events and these should not be overstated.

In the cases of KK1 and KSH, a single recombination pattern was observed for two strains collected at different time points. In addition, for ViBursaCE, two independent but almost identical sequences were deposited in GenBank (EU162089 and EU162088). Considering that the possibility of having two independent artefacts for an identical recombination pattern is negligible, we can confidently accept that KK1, KSH and ViBursaCE are naturally occurring recombinants. However, due to the lack of information for Harbin-1 and SHh, we cannot determine whether these recombinant signals were artefacts. Nonetheless, given the conservativeness of our methods, the phylogenetic incongruence within these genomes is definitive. Even if these incongruent phylogenetic signals were non-naturally occurring, their possible biased effects on further phylogenetic analyses should not be overlooked.

**Exchange of genetic materials between IBDV vaccine and field strains**

Our findings suggest that homologous recombination between live vaccines and wild-type IBDV strains may have occurred. Mass vaccination of farmed animals may increase the probability of emergence of chimeric viruses resulting from recombination or reassortment. Despite the promising efficacy of live vaccines, there are considerable controversies over the potential risks of reversion or generation of chimeric viruses (Seligman & Gould, 2004). Such concerns have been highlighted by the fact that a high proportion (over 50%) of polioviruses involved in vaccine-associated paralytic poliomyelitis are recombinant (Guillot et al., 2000). While the vast majority of these recombinants did not show dramatically altered phenotypes, a few exceptions with altered virulence were documented (Li et al., 1996; Martin et al., 2002). In the case of IBDV, virulence and tropism are primarily determined by critical residues on VP2 and partially by VP1 (Brandt et al., 2001; Hon et al., 2006). Vaccine strains may acquire these virulence factors through intra-serotypic recombination with wild-type strains. Although we cannot exclude the possibility that chimeric viruses with dramatically altered phenotypes may have resulted from recombination, such theoretical risk may be remote at this point, based on the fact that no inter-species or inter-serotypic recombination was detected in this study. However, continuous surveillance of field strains is recommended for early detection of any recombinants with dramatically altered phenotypes.

**Implications and conclusions**

This is the first report of homologous recombination within the family *Birnaviridae*. Homologous recombination in segmented dsRNA viruses is thought to be rare, since replication of their genomes occurs within the capsid after packaging of plus strands (Patton & Spencer, 2000). A prerequisite of recombination is co-packaging of different molecules of the same genome segment into a capsid, which could be rare, if not impossible (Worobey & Holmes, 1999). However, the non-virion-associated form of RdRp of infectious pancreatic necrosis virus (IPNV) was shown to behave as a functional replicase (Xu et al., 2004). Therefore, we cannot exclude the possibility that replication, and thus recombination, may happen outside the capsid before packaging. In addition, previous studies showed that birnaviruses can be recovered from mRNA.
Recombination in birnaviruses

(Thrush & Vakharia, 1996; Yao & Vakharia, 1998), implying that full-length positive strand RNA is recognizable by viral RNA polymerase; this is similar to the observation in positive strand RNA viruses. This observation suggests that replication of birnaviruses may be different from other dsRNA viruses, while its role in homologous recombination of birnaviruses deserves further investigation. In summary, this study demonstrated that not all birnaviruses are clonal. These recombinants may mislead the phylogenetic-based genotyping system. Here, we suggest a routine screening for recombination before any phylogenetic analysis was performed on the newly sequenced birnaviruses.

ACKNOWLEDGEMENTS

We sincerely thank Dr D. J. Jackwood of The Ohio State University for useful information about their GenBank sequences.

REFERENCES


