Molecular Phylogenetics and Evolution 79 (2014) 385–391

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Evolution of mammalian and avian bornaviruses

Mei He, Tie-Zhu An, Chun-Bo Teng *

College of Life Science, Northeast Forestry University, Harbin 150040, China

ARTICLE INFO

Article history:
Received 8 April 2014
Revised 29 June 2014
Accepted 8 July 2014
Available online 19 July 2014

Keywords:
BDV
ABV
Divergence
Codon usage bias

ABSTRACT

Recently, *Avian Bornavirus* (ABV) was identified to be a new member of the *Bornaviridae* family consisting solely of the mammal-infecting *Borna disease virus* (BDV). Here, to gain more insights into the evolution of these bornaviruses, the time-stamped *N* gene sequences of BDV genotype 1 (BDV1) and ABV were subjected to Bayesian coalescent analyses. The nucleotide substitution rates and the divergence times were estimated. Age calculations suggested that the first diversification event of the analyzed BDV1 isolates might have taken place about 300 years ago, and revealed that ABV was an old virus newly recognized. Great differences were observed in the rate of nucleotide substitution and the pattern of codon usage bias between BDV1 and ABV. Moreover, the analyzed bornaviruses might be descended from an AT-rich ancestor.

1. Introduction

Borna disease (BD) is an infectious neurologic syndrome of warm-blooded animals that may lead to profound behavioral abnormalities and even fatalities (Lipkin et al., 2011; Lipkin et al., 1990). It is named after the devastating epidemic that killed a large number of horses during the 1890s in Borna, a town in Saxony, Germany. However, its history can even be dated back to the early 18th century (Kolodziejek et al., 2005; Lipkin et al., 2011). To date, it is still endemic yet sporadic in Central Europe predominantly afflicting horses and sheep as nonpurulent meningoencephalitis (Dürrwald et al., 2006). Due to the restricted pattern, BD did not draw international attention until the 1980s when the zoonotic potential and a global distribution were suggested (Dürrwald et al., 2006; Rott et al., 1985). This prompted the identification and characterization of the etiologic agent *Borna disease virus* (BDV) in the 1990s (Briese et al., 1994; Cubitt et al., 1994; Lipkin et al., 1990), ~70 years after the viral nature was proven (Zwick et al., 1926, 1929).

Now, it is well acknowledged that BDV is an enveloped RNA virus with its non-segmented, negative-sense, single-stranded genome of ~8.9 kb packaged in a spherical virion (Lipkin et al., 2011; Richt and Rott, 2001). By virtue of alternative splicing and overlapping translation, six open reading frames (ORFs) are skillfully compacted in the small genetic material (de la Torre, 1994; Schneemann et al., 1995). In addition to the five proteins common in the order *Mononegavirales*: nucleoprotein (N), phosphoprotein (P), matrix protein (M), glycoprotein (G) and RNA polymerase (L) (Lipkin et al., 2011), a unique non-structural protein (X) is encoded by a short ORF upstream and overlapping that of P (Wehner et al., 1997). Another unusual feature is that the virus accomplishes replication and transcription in the nuclei of the host cells (Briese et al., 1994). Therefore, BDV is solely endowed with a new family *Bornaviridae* (Pringle, 1996), in special relation to the family *Rhabdoviridae* (Cubitt et al., 1994).

Surprisingly, despite the inherent error-prone RNA polymerase, all but one BDV isolates from various hosts over several decades so far exhibit remarkable sequence homology, composing a large group designated genotype 1 (BDV1) (Kolodziejek et al., 2005). No/98, the single BDV2-type virus isolated from an Austrian pony, differs from all others by over 15% at the nucleotide level (Nowotny et al., 2000). Due to such great conservation leading to suspicion of contamination, as well as several unrepeatable serological tests, the involvement of BDV in human neuropsychiatric disorders remains controversial (Dürrwald et al., 2007; Lipkin et al., 2011). In fact, our genomes, in which BDV-like segments (EBLNs) are endogenized (Belyi et al., 2010; Horie et al., 2010), have witnessed that ancient bornaviruses did infect us human beings once upon a time.

BDV was long thought to be the singular member of the family until 2008 when *Avian Bornavirus* (ABV) was isolated during the investigation on Proventricular Dilatation Disease (PDD), a fatal neurologic condition of pet parrots (Honkavuori et al., 2008; Kistler et al., 2008). Unlike the mammalian relative, the psittacine bornavirus (ABV-P) is clearly widespread and has had seven genotypes (1–7) identified (Payne et al., 2012). Later, distinct ABV strains were detected in captive finches and wild waterfowl, including ABV-C in canaries and Bengalese finches (Rubbenstroth

* Corresponding author. Fax: +86 451 8219 1784.
E-mail address: chunboteng@nefu.edu.cn (C.-B. Teng).

http://dx.doi.org/10.1016/j.ympev.2014.07.006
1055-7903/

et al., 2013; Weissenböck et al., 2009), ABV-EF in Estrildid finches (Rubbenstroth et al., 2014), and ABV-CG in Canada geese and swans (Delnatte et al., 2013; Guo et al., 2012; Payne et al., 2011). Recently, a distantly related *Reptile Bornavirus* (RBV) was reported from a Gaboon viper venom gland (Fujino et al., 2012).

Here, to gain more insights into the evolution of the seemingly emerging bornaviruses, Bayesian coalescent method was applied to the dated complete sequences of the *N* gene, with emphasis on the divergence scenario. Moreover, to better understand the processes governing their evolution, selection on the *N* gene and codon usage biases of the six genes were analyzed.

2. Materials and methods

Full-length ORF sequences of the six bornaviral genes were retrieved from GenBank and aligned with CLUSTAL W (Thompson et al., 1997). Dataset compilation, Bayesian estimates, selection analyses, and surveys of codon usage bias were performed as previously described (He et al., 2013). In Bayesian analyses, panels of the *N* gene sequences of BDV1 and ABV (No/98 included) were separately compiled with only time-stamped field isolates.

To estimate the nucleotide substitution rates and the times to the most recent common ancestor (TMRCAs), the Bayesian Markov chain Monte Carlo (MCMC) method (Drummond et al., 2012) was employed. For each dataset, the 20 kinds of combinations of 4 clock models (strict, exponential, lognormal and random local) and 5 demographic models (constant, exponential, expansion, logistic and Bayesian skyline plot) were compared, and the better one (Table 1) was chosen according to convergence and performance. Independent analyses for 10–15 million MCMC iterations (with 10% burn-in) were combined. Isolate information (name/year/country/host/accession) was given in each maximum clade credibility (MCC) tree.

Moreover, alignments of representative partial *N* (382 nt, referred to nucleotide positions 632-1013 of strain V (Accession No. U04068)), entire *X–P* and partial *M* (308 nt, 1950-2257 of strain V) sequences were respectively created for phylogenetic analyses. Each Maximum Likelihood (ML) tree was drawn by MEGA 5.1 (Tamura et al., 2011) with 1000 bootstrap replicates under the best-fit nucleotide substitution model (GTR + G + I for *N* and *M*; GTR + G for *X–P*) determined by MODELTEST in HyPhy (Pond et al., 2005).

3. Results and discussion

3.1. Great difference in the nucleotide substitution rate of the N gene between BDV1 and ABV

As with the strong dissimilarity in genetic diversity, great difference in nucleotide substitution rate between BDV1 and ABV was observed (Table 1). When ORF sequences of the *N* gene from 56 natural BDV1 isolates spanning 27 years were subjected to Bayesian analysis, the average rate was calculated to be 1.06×10^{-4} subs/site/year, with the 95% highest probability density (HPD) values ranging from 4.20×10^{-5} to 1.78×10^{-4}. It was only about one seventeenth of the mean rate at 1.79×10^{-3} (4.87×10^{-4} to 3.16×10^{-3}) subs/site/year estimated for the *N* gene conducted on 32 field isolates consisting of one BDV2 isolate in 1998 and 31 ABV isolates spanning 7 years. Notably, there was no intersection between their HPD values.

Such evolutionary difference may be related to host difference. Each of the two relatives possesses a wide host range, with various kinds of mammals for BDV and diverse bird species for ABV; however, BDV may employ a special species, the bicolored white-toothed shrew (*Crocidura leucodon*), as an indigenous viral reservoir and other animals including horses and sheep as spill-over hosts (Dürrwald et al., 2014). This may impose more constraints on virus evolution and thus result in a remarkably homogenous virus group.

Table 1
Details of datasets and estimates of bornaviruses based on the *N* gene.

Parameter[a]	BDV1	ABV[b]
No. of sequences	56	32
Time span	1985–2012	1998, 2006–2013
Substitution model	GTR + G	GTR + G + I
Molecular clock	Lognormal	Random local
Demographic model	Exponential	Logistic
Mean substitution rate	1.06×10^{-4}	1.79×10^{-3}
95% HPD rate	4.20×10^{-5}–1.78×10^{-4}	4.87×10^{-4}–3.16×10^{-3}
Mean TMRCA	302	772
95% HPD year	132–525	262–1593
d_N/d_S	0.034	0.046

[a] HPD: highest probability density; TMRCA: time to the most recent common ancestor; d_N/d_S ratio: mean ratio of nonsynonymous to synonymous substitution per site.
[b] No/98 (BDV2 isolate in 1998) included.

3.2. Purifying selection on the bornaviral N gene

However, their nucleotide substitution patterns were much similar. As revealed by selection analyses using the ML-based single likelihood ancestor counting (SLAC) method implemented in HyPhy (Pond et al., 2005), the synonymous substitution was predominant over the nonsynonymous one in the evolution of the *N* gene ($d_N/d_S < 0.05$, Table 1), reflecting intense purifying selection on bornaviruses. Actually, isolate No/98 was a good example. It shared 93–98% similarity (except X sacrificing for conservation of P) to BDV1 isolates at the amino acid level in contrast to the over 15% variability at the nucleotide level (Nowotny et al., 2000).

3.3. Divergence of BDV1 based on the N gene

Due to the uncommon evolutionary mode of BDV1, the divergence times of BDV1 and ABV were separately estimated. Based on the *N* gene, the mean TMRCA calculated for BDV1 was 302 (95% HPD: 132-525) years before 2012 (Table 1), that is, the first diversification event of the analyzed BDV1 isolates might have taken place in the early 18th century, around the time the meningoencephalitis of German horses was observed (Lipkin et al., 2011).

The MCC tree of BDV1 (Fig. 1A) confirmed the finding of Kolodziejek et al. (2005) that there were five different clusters corresponding to the geographical origins: a Saxony-Anhalt and bordering northern Saxony group (termed G1 here), a mixed group mainly from Thuringia and Lower Saxony (G2), a Bavaria I group (G3), a Baden-Wurttemberg and Bavaria II group (G4), as well as a Swiss and Liechtenstein group (SL). The former three were ancestral to G4 and SL, the geographically adjacent sister groups diverging from each other ~200 years ago, which could be inferred from the MCC tree of the *X–P* gene (Fig. S1). In addition, according to the fine regional correlation, the etiologic agent of the epidemic in Borna of Saxony, after which the disease was named, was most likely to be a G1-type virus.

Based on the first reliable description of BD in the Swabian Alb in the 1820s, Dürrwald et al. (2014) speculated that BDV1 might have occurred first in Southern Germany (G3 and G4 region) and spread from there northward to Central Germany (G1 and G2) and southward to Switzerland and Liechtenstein (SL). However, tip calibration demonstrated that BDV1 indeed had emerged to be the cause of horse meningoencephalitis in the early 1700s. In fact, judging from the time-scaled MCC trees (Fig. 1A and S1), either Southern Germany (G3) or Central Germany (G1) could be the origin of BDV1. Here, considering that a G1-type virus population has been established in the natural viral reservoir *C. leucodon*