Molecular Phylogenetics and Evolution 83 (2015) 137–142

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

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

Divergence and codon usage bias of Betanodavirus, a neurotropic pathogen in fish

Mei He, Chun-Bo Teng *

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

ARTICLE INFO

Article history:
Received 4 September 2014
Revised 25 November 2014
Accepted 30 November 2014
Available online 10 December 2014

Keywords:
Betanodavirus
Divergence
Codon usage bias
Viral nervous necrosis

ABSTRACT

Betanodavirus is a small bipartite RNA virus of global economical significance that can cause severe neurological disorders to an increasing number of marine fish species. Herein, to further the understanding of the evolution of betanodavirus, Bayesian coalescent analyses were conducted to the time-stamped entire coding sequences of their RNA polymerase and coat protein genes. Similar moderate nucleotide substitution rates were then estimated for the two genes. According to age calculations, the divergence of the two genes into the four genotypes initiated nearly simultaneously at ~700 years ago, despite the different scenarios, whereas the seven analyzed chimeric isolates might be the outcomes of a single genetic reassortment event taking place in the early 1980s in Southern Europe. Furthermore, codon usage bias analyses indicated that each gene had influences in addition to mutational bias and codon choice of betanodavirus was not completely complied with that of fish host.

1. Introduction

Viral nervous necrosis (VNN), also termed as viral encephalopathy and retinopathy (VER), is a highly infectious disease that can cause serious damage typical of extensive vacuolation to the piscine central nervous system (Glazebrook et al., 1990; Yoshikoshi and Inoue, 1990). To date, more than 50 marine teleosts, as well as several freshwater species, have been found susceptible (Bandín and Dopazo, 2011; Vendramin et al., 2014). Infected fish, larvae and juveniles in particular, exhibit various clinical signs such as anorexia, darkened coloration and erratic swimming behavior, and suffer mass mortalities that may reach up to 100% (Munday et al., 2002). Since its first emergence in the 1980s, VNN has been a major impediment to mariculture, outbreaks of which have been reported from all over the world accompanied with significant economic losses (Nakai et al., 2009).

The aetiological agent of VNN is a small naked RNA virus assigned to the genus *Betanodavirus* in the family *Nodaviridae*, distantly related to insect-infecting *Alphanodavirus* (Ball et al., 2000). It has a single-stranded RNA genome composed of two genetic molecules designated as RNA1 and RNA2 encapsulated by the icosahedral capsid. Being positive-sense, both of them possess the methylated caps but lack the poly (A) tails (Mori et al., 1992). RNA1, the larger segment (3.1 kb), encodes the RNA-dependent RNA polymerase (RdRp, or protein A) (Nagai and Nishizawa, 1999), while RNA2, the smaller one (1.4 kb), encodes the coat protein (Cp) (Nishizawa et al., 1995). During viral replication, a subgenomic RNA3 is synthesized from the 3′ terminus of RNA1 and expresses two small proteins, B1 and B2, with regulatory functions (Chen et al., 2009; Fenner et al., 2006; Iwamoto et al., 2005; Su et al., 2009).

Based on phylogenetic analysis of the variable T4 region of the *Cp* gene, betanodaviruses can be categorized into four genotypes (Nishizawa et al., 1997): *striped jack nervous necrosis virus* (SJNNV), *tiger puffer nervous necrosis virus* (TPNNV), *barfin flounder nervous necrosis virus* (BFNNV) and *red-spotted grouper nervous necrosis virus* (RGNNV), with SJNNV being the type species. This taxonomic classification holds true for the *RdRp* gene, despite the difference in genetic relationships (Toffolo et al., 2007). Moreover, a single turbot nodavirus (TNV) has been proposed by Johansen et al. (2004) to be the fifth genotype.

The four major genotypes differ in various aspects including host range, geographic distribution and optimum temperature (Nakai et al., 2009). SJNNV is endemic to Japan affecting certain fish species; however, it also exists (mostly as reassortants) in Southern Europe (Panzarin et al., 2012). TPNNV is so far only found in Japan from Tiger puffer (*Takifugu rubripes*) and Japanese flounder (*Paralichthys olivaceus*) (Nishizawa et al., 1997). BFNNV prefers cold-water teleosts in Japan and the North Atlantic. In contrast, RGNNV possesses the broadest host spectrum covering various kinds of warm-water finfish and the widest geographical range involving

* 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.2014.11.016
1055-7903/

Asia, Australia, North America, and the Mediterranean (Bandín and Dopazo, 2011).

As is common in multipartite viruses, reassortment is employed by betanodaviruses to drive their evolution. Interestingly, all reassortants identified to date are generated from either combination of genomic segments between SJNNV and RGNNV taking place in the Mediterranean basin and the connected Iberian Atlantic region (Olveira et al., 2009; Panzarin et al., 2012; Toffolo et al., 2007; Vendramin et al., 2014). Moreover, recently, based on the dated partial sequences, the rates of nucleotide substitution and the times to the most recent common ancestor (TMRCAs) of the RGNNV group in Southern Europe have been estimated, which revealed a higher rate and a younger age for the *RdRp* gene there. However, the differences were not significant as the 95% highest probability density (HPD) ranges for the two segments overlapped (Panzarin et al., 2012).

Here, to broaden the knowledge of the molecular epidemiology and evolutionary dynamics of betanodavirus, Bayesian coalescent method was applied to the time-stamped complete coding sequences of the *RdRp* and *Cp* genes sampled worldwide. Moreover, to better understand the processes governing their evolution, codon usage bias of each gene was also examined.

2. Materials and methods

Full-length coding sequences of the *RdRp* and *Cp* genes were retrieved from GenBank and aligned with CLUSTAL W (Thompson et al., 1997). Dataset compilation, Bayesian estimates, and codon usage bias analyses were conducted as described previously (He et al., 2013). Details of the betanodavirus isolates analyzed, including collection dates supplemented via literature and Genbank Accession numbers, were listed in Table S1.

When the Bayesian Markov chain Monte Carlo (MCMC) method in BEAST v1.7.4 (Drummond et al., 2012) was applied to each panel, the four clock models (strict, exponential, lognormal and random local) as well as the five demographic models (constant, exponential, expansion, logistic and Bayesian skyline plot) were compared (He et al., 2014a). Specifically, the strict clock model was evaluated by the coefficient of variation (CoV), a scale of the degree of clocklikeness of the data (Drummond et al., 2006). Thus, it was rejected for *RdRp* in that CoV did not encompass zero. On the contrary, it was accepted for *Cp* since the lower 95% HPD of CoV approached zero. For *RdRp*, among the three relaxed clock models, the lognormal one was utilized as the other two failed to describe the evolutionary dynamics. Moreover, as most demographic models yielded similar results for *Cp*, the logistic growth one was chosen owing to its best reliability and convergence for *RdRp*. Then, independent analyses for 10–25 million MCMC iterations were combined (10% burn-in) to assure effective sample size (>200). Isolate information (name/year/origin/accession) was also included in each maximum clade credibility (MCC) tree.

Two indices of codon usage bias, the effective number of codons used by a gene (*Nc*) and the frequency of G + C at the synonymous 3rd codon position (GC_{3S}) were examined for *RdRp* and *Cp*. The overall variation in the codon usage was reflected by the relative synonymous codon usage (RSCU), the value of which higher than 1.0 indicates that the corresponding codon is adopted more frequently than expected, and vice versa. Besides, correspondence analysis, a multivariate statistical technique that yields orthogonal axes to account for variations, was conducted to identify major trends in amino acid usage. All aforementioned analyses were done by CodonW 1.4.4 (http://codonw.sourceforge.net). Moreover, an alphanodavirus *Flock house virus* (FHV) and a fish host *Dicentrarchus labrax* with more codons available (http://www.kazusa.or.jp/codon/) were selected for comparisons of RSCU and GC content.

In addition, to infer the genetic relationships of the chimeras (30 RG/SJ and 1 SJ/RG), alignments of their partial *RdRp* (867 nt, referred to CDS positions 154-1020 of SJNNV/AB056571) and *Cp* (504 nt, CDS positions 357-860 of SJNNV/AB056572) sequences were created for phylogenetic analyses, respectively. 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 (TN93 + G for *RdRp* and T92 + G for *Cp*).

3. Results

3.1. Nucleotide substitution rates of the RdRp and Cp genes

As listed in Table 1, when entire coding sequences of the *RdRp* gene from 49 betanodavirus isolates spanning 19 years, consisting of 32 RG, 14 BF, 1 TP and 2 SJ-type viruses, were subjected to Bayesian analysis, the average rate of nucleotide substitution was estimated to be 3.60×10^{-4} subs/site/year, with the 95% HPD values varying from 1.40×10^{-4} to 6.18×10^{-4}. A similar mean rate at 3.69×10^{-4} (2.31×10^{-4}–5.06×10^{-4}) subs/site/year was obtained for the *Cp* gene based on 73 isolates spanning 21 years, composed of 54 RG, 8 BF, 1 TP and 10 SJ-type viruses.

Substitution rates were also calculated for the two genes of RGNNV, the largest genotype with more complete sequence data available. The reassortants and MnNNV-12-06 (Fig. 1A) being excluded, the average rate of *RdRp* from 24 isolates was 4.28×10^{-4} (3.35×10^{-5}–8.79×10^{-4}) subs/site/year, which was insignificantly higher than that of *Cp* from 52 isolates at 3.79×10^{-4} (2.01×10^{-4}–5.60×10^{-4}) subs/site/year. In fact, rate difference was still minor when the *Cp* panel was truncated to be identical to the *RdRp* panel (data not shown).

3.2. Phylogenetic relationships of betanodaviruses

The MCC trees (Fig. 1) calculated for the *RdRp* and *Cp* genes of betanodavirus isolates collected worldwide confirmed the categorization of four major genotypes (SJ, TP, BF and RG), the difference in genetic relationships involving BFNNV, as well as the reassortments between SJNNV and RGNNV. Notably, the Atlantic cod nervous necrosis virus (ACNNV) circulating in Canada described by Gagnè et al. (2004) could be classified as a subtype of BFNNV (Fig. 1B).

Within the large RGNNV group, the isolates analyzed here clustered into 3 (a–c, Fig. 1A) and 6 (a–e, Fig. 1B) well-supported subclades (posterior probability value >0.9) based on their *RdRp* and *Cp* sequences, respectively. The subtype a of *Cp*, for example, could be further divided as well (1–7, Fig. 1B). According to the arbitrary classification, incongruent topologies were observed for GPNNV1108-P.Langkawi-1 sampled from Malaysia. In the *RdRp*

Table 1
Details of datasets and estimates of the two betanodaviral genes.

Parameter[a]	*RdRp*	*Cp*
No. of sequences	49	73
Time span	1993–2012	1991–2012
Substitution model	GTR + G + I	
Molecular clock	Lognormal	Strict
Demographic model	Logistic	
Mean substitution rate ($\times 10^{-4}$)	3.60	3.69
95% HPD rate ($\times 10^{-4}$)	1.40–6.18	2.31–5.06
Mean TMRCA	695	704
95% HPD year	267–1238	446–990

[a] HPD: highest probability density; TMRCA: time to the most recent common ancestor.