Molecular Phylogenetics and Evolution 77 (2014) 34–40

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

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

Evolution of the viral hemorrhagic septicemia virus: Divergence, selection and origin

Mei He [a], Xue-Chun Yan [b], Yang Liang [a], Xiao-Wen Sun [b], Chun-Bo Teng [a,*]

[a] *College of Life Science, Northeast Forestry University, Harbin 150040, China*
[b] *Heilongjiang River Fisheries Research Institute, Chinese Academy of Fishery Sciences, Harbin 150070, China*

ARTICLE INFO

Article history:
Received 14 December 2013
Revised 26 March 2014
Accepted 1 April 2014
Available online 13 April 2014

Keywords:
VHSV
Substitution rate
TMRCA
Divergence

ABSTRACT

Viral hemorrhagic septicemia virus (VHSV) is an economically significant rhabdovirus that affects an increasing number of freshwater and marine fish species. Extensive studies have been conducted on the molecular epizootiology, genetic diversity, and phylogeny of VHSV. However, there are discrepancies between the reported estimates of the nucleotide substitution rate for the *G* gene and the divergence times for the genotypes. Herein, Bayesian coalescent analyses were conducted to the time-stamped entire coding sequences of the six VHSV genes. Rate estimates based on the *G* gene indicated that the marine genotypes/subtypes might not all evolve slower than their major European freshwater counterpart. Age calculations on the six genes revealed that the first bifurcation event of the analyzed isolates might have taken place within the last 300 years, which was much younger than previously thought. Selection analyses suggested that two codons of the *G* gene might be positively selected. Surveys of codon usage bias showed that the *P*, *M* and *NV* genes exhibited genotype-specific variations. Furthermore, we proposed that VHSV originated from the Pacific Northwest of North America.

© 2014 Elsevier Inc. All rights reserved.

1. Introduction

Viral hemorrhagic septicemia, historically known as Egtved disease, is a deadly virulent disease that infects more than 80 fish species from diverse families, including rainbow trout (*Oncorhynchus mykiss*), turbot (*Scophthalmus maximus*), and yellow perch (*Perca flavescens*), which are important in commerce and recreation (Kurath, 2012; Smail and Snow, 2011). First observed in Germany by Schäperclaus (1938), the disease was believed to be restricted to only freshwater fish in continental Europe until 1988 when it was also discovered among Pacific anadromous salmonids in North America (Brunson et al., 1989; Hopper, 1989). Now, it is known to circulate within the Northern Hemisphere. With such high infectivity, broad host range and wide distribution, the disease lays a heavy burden on the global fish farming industry, especially on the European trout aquaculture which has suffered significant losses from its outbreaks for massive die-offs (Smail and Snow, 2011).

Viral hemorrhagic septicemia virus (VHSV), the causative agent identified in 1962 (Jensen, 1963), belongs to the *Novirhabdovirus* genus in the *Rhabdoviridae* family (Walker et al., 2000). Its enveloped bullet-shaped virion encapsidates a non-segmented, negative-sense, single-stranded RNA of ~11,000 nucleotides. The linear genome contains six genes encoding a non-virion protein (NV) and five structural proteins: nucleoprotein (N), phosphoprotein (P), matrix protein (M), glycoprotein (G), and RNA polymerase (L), which are organized as 3′-N–P–M–G–NV–L-5′ (Schutze et al., 1999). NV, unique to the genus, is capable of suppressing apoptosis at the early stage of viral infection (Ammayappan and Vakharia, 2011), whereas the other five proteins are common in rhabdoviruses with analogous functions (Kurath, 2012; Kuzmin et al., 2009).

So far, extensive studies have been conducted on the molecular epizootiology, genetic diversity, and phylogeny of VHSV. Phylogenetic analyses based on the *N* or *G* gene sequences of global VHSV isolates have defined four genotypes designated with Roman numerals I to IV (Einer-Jensen et al., 2004; Snow et al., 2004). Further, genotypes I and IV are divided into five (Ia–Ie) and three (IVa–IVc) subtypes, respectively. As illustrated in Fig. 1, the genotypes/subtypes of VHSV have different geographic distributions. Among the European lineages, Ia is predominantly composed of freshwater trout isolates from the mainland; Ib primarily circulates within the Baltic and North Sea water system; Ic is a small group of old isolates from freshwater rainbow trout; Id infects rainbow trout reared in fresh or brackish water; Ie prevails in the marine/estuarine Black Sea; II and III are recovered from the Baltic Sea and the North Atlantic (and connected waters), respectively. The

* 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.2014.04.002
1055-7903/© 2014 Elsevier Inc. All rights reserved.

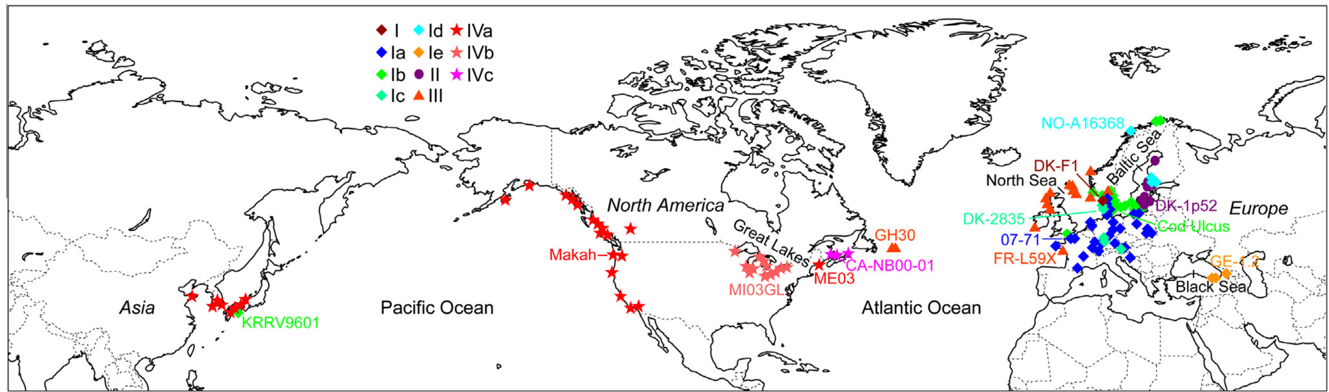

Fig. 1. Sketch of the geographical distribution of VHSV. The four major genotypes (I–IV) are depicted with different symbols and their subtypes (Ia–e and IVa–c) are distinguished with different colors. The first isolate of each lineage and the special isolates (ME03, GH30 and KRRV9601) are indicated. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

distinct genotype IV has a much wider range from North America to East Asia: IVa includes marine isolates from East Asia and the Pacific Northwest of North America; IVb contains freshwater isolates in the North American Great Lakes region; IVc comprises seaboard isolates from the Atlantic coast of Canada (Einer-Jensen et al., 2004; Kurath, 2012; Pierce and Stepien, 2012). Notably, in 2005, VHSV was identified for the first time in China from diseased flounder (*Paralichthys olivaceus*) (Zhu and Zhang, 2013). This isolate is closely related to the Korean strains, suggesting new colonization by the severe fish pathogen.

Moreover, there are differences in the level of our knowledge about these genotypes/subtypes. In Europe, both freshwater VHSV and the marine reservoir have been very well characterized owing to large-scale surveillance efforts in farmed fish and extensive ocean cruises to scrutinize wild fish. As can be seen in Fig. 1, most isolates are collected there, especially in continental Europe. Overseas from Europe, IVa is also reasonably well defined, but based on detection during surveillance of cultured fish, and smaller scale surveillance of wild fish; IVb has also been surveyed quite thoroughly but only since its outbreak in 2005; IVc, however, is less well defined, which is so far based on 5 isolations in one report.

The nucleotide substitution rates of the *N*, *G* and *NV* genes of VHSV as well as the divergence times of the genotypes according to the *G* calibration have already been assessed (Pierce and Stepien, 2012). Although the rate estimate is influenced by panel composition, their result conducted on the partial *G* sequences was 2.58×10^{-4} subs/site/year, which was much lower than the previous report as between 7.06×10^{-4} and 1.74×10^{-3} based on the full-length *G* dataset lacking IVb, IVc and later isolates (Einer-Jensen et al., 2004). Consequently, discrepancy in the divergence times of the genotypes has arisen. The primary bifurcation event was calculated to have occurred 697 years ago, nearly 200 years earlier than the former estimate (Einer-Jensen et al., 2004).

Therefore, to see whether the *G* gene of VHSV evolves at a relatively slow rate and the genotypes have diverged for such a long time, Bayesian coalescent method was applied to the time-stamped entire coding sequences of each VHSV gene, with emphasis on the divergence history of the genotypes/subtypes. In addition, to better understand the processes governing the evolution of VHSV, selection analyses and surveys of codon usage bias were also carried out.

2. Materials and methods

Complete coding sequences of the six VHSV genes were retrieved from GenBank. Dataset compilation, Bayesian estimates, selection analyses and surveys of codon usage bias were performed as previously described (He et al., 2013). Dates of isolation were supplemented via literature (Einer-Jensen et al., 2004; Elsayed et al., 2006; Gagné et al., 2007; Raja-Halli et al., 2006; Reichert et al., 2013).

When the Markov chain Monte Carlo (MCMC) method (Drummond et al., 2012) was employed, the uncorrelated exponential clock was recommended to be the best fit by the Bayes factor tests (Baele et al., 2012; Suchard et al., 2001). The exponential growth model was chosen for better performance and conformity to the known history of VHSV in Europe. Independent analyses for 5–25 million MCMC iterations (with 10% burn-in) were combined to ensure convergence in estimates of the nucleotide substitution rates and the times to the most recent common ancestor (TMRCAs). Information of the analyzed isolates was given in each of the maximum clade credibility (MCC) trees.

Moreover, representative entire *G* coding sequences of the other three recognized novirhabdoviruses, *Infectious haematopoietic necrosis virus* (IHNV), *Hirame rhabdovirus* (HIRRV), and *Snakehead rhabdovirus* (SHRV), were also retrieved from GenBank and aligned with CLUSTAL W (Thompson et al., 1997). Phylogenetic tree of the genus was constructed by MEGA 5.1 (Tamura et al., 2011) employing the Maximum Likelihood (ML) method with 1000 bootstrap replicates under the best-fit nucleotide substitution model GTR+I+G determined by MODELTEST in HyPhy (Pond et al., 2005).

3. Results

3.1. Nucleotide substitution rates of the VHSV genes and genotypes

As listed in Table 1, when the entire coding sequences of the *G* gene from 277 worldwide VHSV isolates spanning 49 years were subjected to Bayesian analysis, the average rate was 5.91×10^{-4} subs/site/year with the 95% highest probability density (HPD) values ranging from 4.59×10^{-4} to 7.22×10^{-4}, which was a little higher than that of *N* at 4.72×10^{-4} (2.25×10^{-4}–7.38×10^{-4}) calculated on 35 isolates spanning 42 years. Moreover, among the six genes, *NV* and *M* displayed the highest and lowest rates, respectively. Since the gene panels were composed of different numbers of isolates (Table 1) which might result in biased rate estimates, we compiled identical datasets consisting of only the isolates with all six gene sequences available ($n = 17$) and observed similar rate difference among the genes (data not shown).

To assess rate difference among the genotypes/subtypes, the *G* alignment was partitioned accordingly and analyzed. It could be seen that the average rates varied from 2.80×10^{-4} to

Download English Version:

https://daneshyari.com/en/article/5919447

Download Persian Version:

https://daneshyari.com/article/5919447

Daneshyari.com