Virology 514 (2018) 156–164

Contents lists available at ScienceDirect

Virology

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

Molecular evolution and invasion pattern of Cryphonectria hypovirus 1 in Europe: Mutation rate, and selection pressure differ between genome domains

Jelena Mlinarec[a], Lucija Nuskern[a], Marin Ježić[a], Daniel Rigling[b], Mirna Ćurković-Perica[a,*]

[a] Department of Biology, Faculty of Science, University of Zagreb, Marulićev trg 9a, HR-10 000 Zagreb, Croatia
[b] WSL Swiss Federal Research Institute, Zürcherstrasse 111, CH-8903 Birmensdorf, Switzerland

ARTICLE INFO

Keywords:
Biological control
Chestnut blight
CHV1
Evolution
Host parasite interactions
Invasive species
Mycoviruses
Recombination

ABSTRACT

Understanding virus evolution is a fundamental goal of virology, evolutionary biology, and disease epidemiology. We provide a detailed analysis of evolution and origin of Cryphonectria hypovirus 1 (CHV1) populations in Europe, based on the complete genome sequence of all European subtypes. Phylogenetic analyses divided European strains into two closely related clades. Strains of the subtype I belong to the first, while strains of the subtypes F1, D and E belong to the second clade suggesting that the subtypes F1, D and E are more closely related than previously thought. Strains of the subtype F2 appeared to be recombinant; subtypes F1/D/E contributed a larger fraction of sequence while subtype I contributed a smaller fraction. The p29 was the most variable domain, while the replication-associated large ORF B protein was the most conserved domain within the CHV1. Low sequence similarity, predominant negative selection and frequent recombination characterise the evolution of CHV1.

1. Introduction

Genetic variability allows the adaptation of populations to a changing environment. Many virus populations exist as complexes of closely related genomic variants as a result of high mutation rates, rapid replicative kinetics and large population sizes (Biebricher and Eigen, 2006). The high mutation rate of RNA viruses is a consequence of their error-prone RNA-dependent RNA polymerases (RdRps) (Holland et al., 1982). Although mutations are a predominant factor in the diversification of virus populations (Lima et al., 2017), it does not account for all genetic variation. An important other evolutionary mechanism is recombination, which also generates diversity in most viruses (Pérez-Losada et al., 2015). Because of all aforementioned reasons, RNA viruses possess great adaptive potential, which can be beneficial for those viruses that are used to control fungal diseases such as Cryphonectria hypovirus 1 (CHV1) (Nuss, 2005).

CHV1 is an unencapsidated double-stranded (ds) RNA virus of the family *Hypoviridae,* occurring in Japan, China and Europe (Allemann et al., 1999). It causes persistent infection of its fungal host, *Cryphonectria parasitica*, reduces growth, sexual and asexual reproductive ability, and virulence of the infected mycelia, a phenomenon called hypovirulence (Peever et al., 2000; Hillman and Suzuki, 2004; Bryner et al., 2012). Chestnut blight fungus is probably one of the best known invasive fungal pathogens (Rigling and Prospero, 2017). The fungus is native to Eastern Asia and has been introduced into both North America and Europe (including Turkey and the Caucasus) where it caused severe disease epidemics on the native chestnut species. Hypovirulence phenomenon has been used for the control of chestnut blight in European chestnut orchards and has been credited with reducing the severity of the chestnut blight epidemic as a result of natural spread of CHV1 in the European *C. parasitica* population, particularly in chestnut coppice forests (Nuss, 2005).

Diversity of the CHV1 in Europe has been studied thoroughly. Phylogenetic analyses based on RFLP banding pattern and partial sequences of ORFA separated CHV1 isolates into five clusters or subtypes (Allemann et al., 1999; Gobbin et al., 2003). Most viral isolates formed one large cluster comprising all viruses from Italy, Switzerland, Croatia, Bosnia and Herzegovina, Hungary, Greece, and the French island Corsica, as well as five out of 11 isolates from continental France. This cluster of viruses was assigned to the Italian subtype (I). Other eight isolates were separated into four clusters, assigned to the French (F1 and F2), German (D) and Spanish (E) subtypes. Subtype I was also found in Macedonia (Bryner et al., 2012), Slovenia (Krstin et al., 2008) and Turkey (Akilli et al., 2013), therefore making it the most

* Corresponding author.
E-mail address: mirna.curkovic-perica@biol.pmf.hr (M. Ćurković-Perica).

https://doi.org/10.1016/j.virol.2017.11.011
Received 14 September 2017; Received in revised form 13 November 2017; Accepted 14 November 2017

Table 1
CHV1 strains used in this study. Nine new complete virus genome sequences determined in this study are in bold.

CHV1 strain	CHV1 subtype[a]	Country	Reference	ORF A (aa)	ORF B (aa)	GenBank accession no.
EP713	F1	France	Shapira et al. (1991)	622	3165	9626818
Euro7	I	Italy	Chen and Nuss (1999)	622	3165	AF082191.1
EP721	I	Italy	Lin et al. (2007)	622	3165	DQ861913.1
CN280	*ND*	China	Du et al. (2017)	622	3164	KT726153.1
M56-1	I	Croatia	Krstin et al. (2008)	622	3165	MF421722
CR23	I	Croatia	Krstin et al. (2008)	622	3165	MF421721
HK27	I	Croatia	Krstin et al. (2008)	622	3165	MF421720
B11	I	Slovenia	Krstin et al. (2011)	622	3165	MF421723
SHE30	F2	Georgia	this study	622	3165	MF421719
M1147	D	Germany	Gobbin et al. (2003)	622	3164	MF431593
M1372	E	Spain	Gobbin et al. (2003)	622	3164	MF431594
M1728	F2	France	Gobbin et al. (2003)	622	3164	MF421717
M2021	F2	France	Bryner and Rigling (2011)	622	3164	MF421718

ND not determined.
[a] CHV1 subtype designation according to Allemann et al. (1999). *I* Italian, *F1* and *F2* French, *D* German, *E* Spanish.

widespread subtype in Europe. The distributions of the other subtypes are much more geographically restricted. Subtypes F1 is distributed in France and Spain (Zamora et al., 2012; Feau et al., 2014), while subtype F2 can be found in France and Turkey (Akilli et al., 2013); subtypes E and D were both detected in Spain (Gobbin et al., 2003; Trapiello et al., 2017), while in Germany, isolates belonging to subtype D were present (Peters et al., 2012).

It has been postulated that CHV1 was introduced into Europe directly from Asia with its fungal host as CHV1 has not been detected in a recent survey in North America (Heiniger and Rigling, 1994; Peever et al., 1998; Allemann et al., 1999). Furthermore, it was suggested that the presence of different CHV1 subtypes in Europe is the result of multiple introductions of the CHV1 from Asia (Allemann et al., 1999; Gobbin et al., 2003; Feau et al., 2014). This hypothesis was mainly based on the estimates of the approximate time of divergence between the different subtypes, showing that most subtypes diverged at least several hundred years ago, which is by far earlier than the date of the first report of *C. parasitica* in Europe (60–70 years ago) (Gobbin et al., 2003). However, the relatedness between the Asian and European CHV1 populations is still not completely resolved. Liu et al. (2007) have investigated genetic diversity of the CHV1 strains from China, Japan and Italy, and found that the divergence between the subtypes in China was much greater than that in Europe which is in concordance with the hypothesis about the Asian origin of CHV1 in Europe. Besides, the same authors found that most of European isolates grouped together, and were a sister group of the Chinese isolates originating from the south and north of the country. However, the authors agreed that the ultimate determination of the origin of CHV1 in Europe must involve supporting evidence from other research approaches as well as many more samples from populations throughout Asia included in the study.

The full-length nucleotide sequences of four CHV1 isolates, EP713, Euro7, EP721 and CN280 have been reported so far (Shapira et al., 1991; Chen and Nuss, 1999; Lin et al., 2007; Du et al., 2017). Isolate EP713 was originally isolated in France and belongs to the subtype F1 (Allemann et al., 1999), while isolates Euro7 and EP721 were isolated in Italy and group into subtype I (Chen and Nuss, 1999; Lin et al., 2007). CHV1 isolate CN280 was isolated in China (Du et al., 2017). CHV1 genome consists of a single large 12.712 kb positive-stranded RNA (Shapira et al., 1991) composed of a 5′-495-nt noncoding sequence, 11,406-nt coding region, 3′-851-nt noncoding sequence, and a templated poly(A) tract. The genome contains two open reading frames (ORFs), ORF A and ORF B, separated by a UAAUG pentanucleotide (Shapira et al., 1991). The 5′ proximal coding domain, ORF A (composed of 622 codons), encodes two proteins, p29 and p40 that are released from a polyprotein p69 by an autoproteolytic event mediated by a papain-like protease encoded in the p29 region. It was shown that the residues Cys^{162} and His^{215} in p29 are essential for autolytic cleavage (Choi et al., 1991). ORF B (composed of 3,165 codons) encodes the two proteins, p48 and the large ORF B protein. Although, in addition to the p48 proteinase, replication-associated RNA polymerase and helicase domains are contained within ORF B, little is known about processing events of the ORF B polyprotein and specific protein functions (Hillman and Suzuki, 2004).

We carried out full-length genome sequencing of eight CHV1 strains from Europe and one strain from Georgia followed by sequence comparisons and phylogenetic analyses of these nine viruses and four CHV1 strains previously sequenced and deposited in GenBank, in order to elucidate the genetic diversity and molecular evolution of CHV1 in Europe. In addition, we compared CHV1 strains from Europe with those from China and Japan. The aims of this study are: i) to clarify the phylogeny and ancestry of CHV1 in Europe, ii) to determine selective forces and constraints that shape CHV1 evolution and iii) to reveal patterns of recombination and their effect on genetic diversity. This work represents an important contribution toward furthering our understanding of mycovirus evolution, mycovirus-host dynamics and natural selection.

2. Results

2.1. Virus sequences

The full-length genomes of 13 CHV1 strains were analysed in this study (Table 1). The genomic sequence was determined here for the nine CHV1 strains, while four strains were taken from the GenBank. The genome sequences of nine CHV1 strains were between 12,506 and 12,511 bp long and contained the entire coding region of CHV1. The 5′ and 3′ ends within the two noncoding regions, 142 bp and 75 bp long, respectively, were excluded in this study. Thirteen CHV1 strains investigated in this study share an overall from 83.6 to 98.9% nucleotide (nt) sequence similarity, and from 87.6 to 99.6% amino acid (aa) sequence similarity.

The predominant features first identified in strains Euro7 and EP713 were mostly conserved among the other investigated CHV1 strains (Fig. 1A). In ORF A conserved predominant motifs included Phe^{25}, Cys^{70}, Cys^{72}, Gln^{73}, Cys^{162} and His^{215}. In ORF B conserved predominant features included Cys^{341}, His^{388}, Gly^{418} and Ala^{419}. Three motifs in the helicase domain, a GKST box, a DExH box and a QRxGR box were conserved among thirteen investigated CHV1 strains. However, we found two regions in ORF A which appear to be non-conserved among thirteen CHV1 strains investigated in this study: *i)* the cleavage site RIG|$G^{249}R^{250}$L between p29 and p40 in the ORF A and *ii)* region between Thr^{288} and Arg^{312} responsible for p40-mediated RNA accumulation (Suzuki and Nuss, 2002). Three types of amino acid sequences were found at the cleavage site within the thirteen strains investigated