Infection, Genetics and Evolution 9 (2009) 827–831

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

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

Relaxation of purifying selection on the SAD lineage of live attenuated oral vaccines for rabies virus

Austin L. Hughes *

Department of Biological Sciences, University of South Carolina, Coker Life Sciences Bldg., 700 Sumter St., Columbia, SC 29208, United States

ARTICLE INFO

Article history:
Received 11 March 2009
Received in revised form 17 April 2009
Accepted 22 April 2009
Available online 3 May 2009

Keywords:
Live vaccine
Purifying selection
Rabies virus
Virus evolution

ABSTRACT

Analysis of patterns of nucleotide sequence diversity in wild-type rabies virus (RABV) genomes and in the SAD live attenuated oral vaccine lineage was used to test for the relaxation of purifying selection in the latter and provide evidence regarding the genomic regions where such relaxation of selection occurs. The wild-type sequences showed evidence of strong past and ongoing purifying selection both on nonsynonymous sites in coding regions and on non-coding regions, particularly the start, end and 5′ UTR regions. SAD vaccine sequences showed a relaxation of purifying selection at nonsynonymous sites in coding regions, resulting a substantial number of amino acid sequence polymorphisms at sites that were invariant in the wild-type sequences. Moreover, SAD vaccine sequences showed high levels of mutation accumulation in the non-coding regions that were most conserved in the wild-type sequences. Understanding the biological effects of the unique mutations accumulated in the vaccine lineage is important because of their potential effects on antigenicity and effectiveness of the vaccine.

1. Introduction

Rabies virus (RABV), which causes an acute and generally fatal illness of humans and domestic animals, is the oldest recorded human viral pathogen and remains an important source of mortality in many parts of the world (Fu, 1997; Warrell and Warrell, 2004). In Europe, while control measures have greatly reduced human cases of rabies, wild mammals, particularly the red fox (*Vulpes vulpes*), continue to provide a reservoir of rabies infection (Lontai, 1997). As a consequence, the main control strategy has involved oral vaccination of foxes through baits treated with an attenuated live vaccine (Geue et al., 2008). All of the commercially available oral rabies vaccines for wildlife are derived from a single common ancestral isolate known as Street Alabama Dufferin (SAD), derived from an infected dog in North America in 1935 (Fenje, 1960; Sacramento et al., 1992).

A number of different vaccine strains have been developed from the original SAD strain by passaging under a variety of protocols (Geue et al., 2008). As in other live attenuated virus vaccines, mutations during passaging have resulted in a certain degree of nucleotide sequence polymorphism within the SAD lineage (Geue et al., 2008). Understanding the evolution of live vaccines is important because evolutionary changes may affect the biological properties of the vaccine, including its antigenicity and effectiveness, as well as its potential to reacquire virulence (Hughes, 2009). For example, attenuated live vaccines may show a pattern of relaxation of purifying selection, as was observed in three live attenuated vaccine strains of viruses belonging to the family *Paramyxoviridae* (Hughes, 2009). In the case of the SAD vaccine strains, there have been reports of SAD vaccine-associated rabies in wildlife, in spite of evidence of the stability of the vaccine in experimental passaging studies (Beckert et al., 2009).

Purifying selection – that is, selection against deleterious variants – is the most prevalent form of natural selection at the molecular level (Nei, 1987; Hughes, 1999, 2008). Evidence for the prevalence of purifying selection on coding regions is provided by the observation that the number of synonymous substitutions per synonymous site (d_S) generally exceeds the number of nonsynonymous (amino acid-altering) substitutions per nonsynonymous site (d_N). This pattern evidently occurs because most nonsynonymous mutations are deleterious to protein function and thus are eliminated by purifying selection. In fact, there are two distinct aspects of purifying selection, revealed by different methods of analysis: (1) selection against strongly deleterious mutations, which are eliminated quickly; and (2) selection against slightly deleterious mutations, which may be relatively inefficient if the effective population size is small or recombination is limited (Ohta, 1973; Hughes, 2008).

RABV is a single stranded RNA negative-strand virus belonging to the genus Lyssavirus (family *Rhabdoviridae*). As in other members of the family, there are five genes, designated *N*, *P*, *M*, *G*, and *L*, each including a 5′ untranslated region (UTR) and 3′ UTR (Conzelmann, 1998; Tordo et al., 1986). These are illustrated in Fig. 1 according to

* Tel.: +1 803 777 9186; fax: +1 803 777 4002.
E-mail address: austin@biol.sc.edu.

1567-1348/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.meegid.2009.04.016

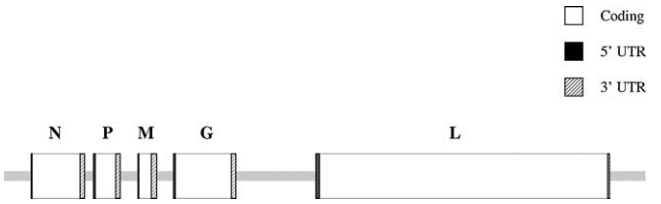

Fig. 1. Schematic diagram of the RABV genome, illustrated in the order (5′ to 3′) of the antigenome RNA (Conzelmann, 1998), showing the five genes *N*, *P*, *M*, *G*, and *L*. The lengths of regions are drawn approximately proportional to those in accession NC_001542 (M13215).

their order in the RNA antigenome (Conzelmann, 1998). There are also start, end, and intergenic regions; the longest intergenic region (between *G* and *L*) may represent the location of a remnant gene that has lost functionality (Tordo et al., 1986).

The purpose of the present paper is to use the tools of molecular evolutionary genetics to understand the evolutionary factors at work in the evolution of the SAD rabies vaccine lineage, in particular the role of purifying selection. By examining the patterns of nucleotide sequence polymorphism in wild-type genomic sequences, I provide evidence for the strength of both past and ongoing purifying selection on different genomic regions. By comparing these patterns with those seen in the SAD vaccine and in SAD vaccine-derived sequences isolated from red foxes, I both test for the relaxation of purifying selection in the latter and provide evidence regarding the genomic regions where such relaxation of selection occurs.

2. Methods

Sequence analyses were based on 35 rabies virus (RABV) sequences, 13 wild-type; 14 from SAD vaccine strains; and 6 vaccine-derived strains isolated from red foxes in Europe (for accession numbers, see Fig. 2). Genomic sequences were aligned using CLUSTALW (Thompson et al., 1997). In computing pairwise distances among a set of sequences, any site at which the alignment postulated a gap in any of the sequences was excluded from all comparisons ("complete deletion," Nei and Kumar, 2000). Complete deletion ensures that the same set of sites is compared in each pairwise comparison (Nei and Kumar, 2000).

Phylogenetic analyses were conducted including the above-mentioned 35 sequences plus two other sequences belonging to the attenuated strains Ni-CE, RC-HL, and RV-97 (for accession numbers, see Fig. 2). A number of different methods of phylogenetic analysis were applied to both complete genome sequences and to coding regions only, all of which showed similar results (not shown). The phylogenetic tree shown below was constructed by the neighbor-joining (NJ) method (Saitou and Nei, 1987) on the basis of the maximum composite likelihood (MCL) distance (Tamura et al., 2007). The reliability of clustering patterns in the tree was assessed by bootstrapping (Felsenstein, 1985); 1000 bootstrap samples were used. The tree was rooted following the phylogenetic analyses of Wu et al. (2007). Ancestral genomic sequences were reconstructed based on the NJ tree by the maximum parsimony method (Swofford, 2003).

The nucleotide usage in the RABV sequences analyzed here was remarkably even. At third positions in coding regions, the G + C content was 49.5%, while in non-coding regions it was 41.1%. By the MCL method, the transition:transversion ratio at non-coding sites and at third positions of codons was estimated at 5.4:1. Because of this strong transitional bias, the number of synonymous substitutions per synonymous site (d_S) and the number of nonsynonymous (amino acid-altering) substitutions per nonsynonymous site (d_N) were estimated by Li's (1993) method, which takes into account transitional bias in affecting the relative probabilities of synonymous and nonsynonymous mutations at twofold and threefold degenerate sites. The number of nucleotide substitutions per site (d) in non-coding regions was estimated by the MCL method. For coding regions, we computed the mean of d_S for all pairwise comparisons (termed synonymous nucleotide diversity and symbolized π_S) and the mean of d_N for all pairwise comparisons (termed nonsynonymous nucleotide diversity and symbolized π_N). Likewise, in non-coding regions, we computed mean d for all pairwise comparisons (nucleotide diversity or π). Standard errors of π_S, π_N, and π were computed by the bootstrap method (Nei and Kumar, 2000); 1000 bootstrap samples were used.

Separate estimates of π_S and π_N were obtained for each of the five protein-coding genes (Fig. 1), as well as for the combined

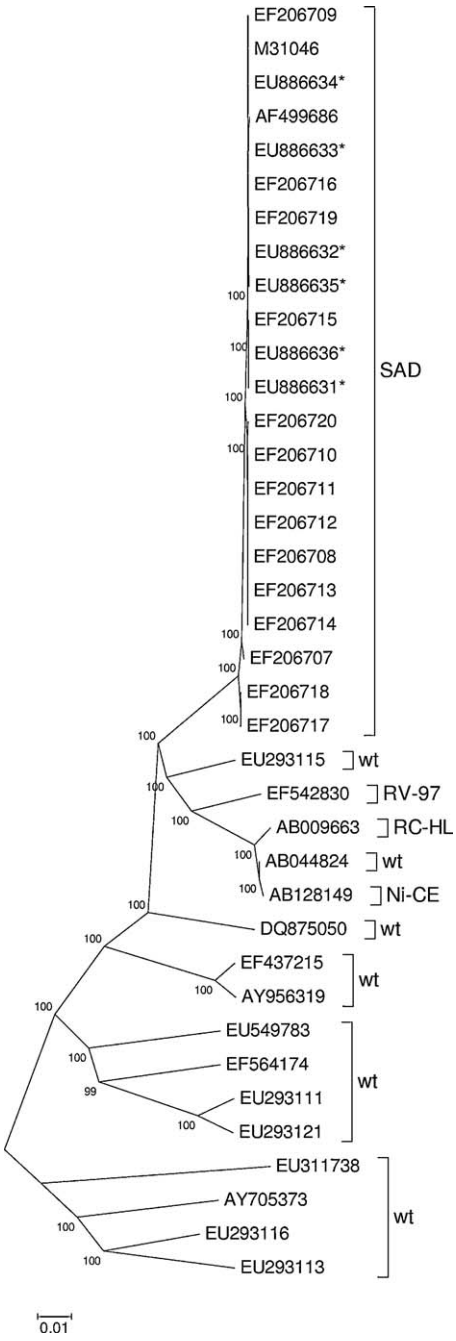

Fig. 2. Neighbor-joining tree of wild-type (wt) and SAD vaccine lineage RABV genomes based on MCL distance at 11,837 aligned nucleotide sites. Numbers on branches are the percentages of 1000 bootstrap samples supporting the branch; only values ≥95% are shown. *Asterisks* indicate SAD vaccine-derived sequences.