Identification of a novel equine infectious anemia virus field strain isolated from feral horses in southern Japan

Jian-Bao Dong,1,2,3 Wei Zhu,1,2 Frank R. Cook,4 Yoshitaka Goto,1 Yoichiro Horii5 and Takeshi Haga6

Correspondence
Takeshi Haga
ahaga@mail.ecc.u-tokyo.ac.jp

1Departments of Veterinary Microbiology, University of Miyazaki, Miyazaki 889-2192, Japan
2Research Fellow DC of the Japan Society for the Promotion of Science, Tokyo 102-8472, Japan
3The United Graduate School of Veterinary Science, Yamaguchi University, Yamaguchi 753-8511, Japan
4Gluck Equine Research Center, Department of Veterinary Science, University of Kentucky, Lexington, KY 40506, USA
5Veterinary Parasitic Diseases, University of Miyazaki, Miyazaki 889-2192, Japan
6Division of Infection Control and Disease Prevention, Department of Veterinary Medical Science, University of Tokyo, Tokyo 113-8654, Japan

Although equine infectious anemia (EIA) was described more than 150 years ago, complete genomic sequences have only been obtained from two field strains of EIA virus (EIAV), EIAWYoming and EIAVLiaoning. In 2011, EIA was detected within the distinctive feral Misaki horse population that inhabits the Toi-Cape area of southern Japan. Complete proviral sequences comprising a novel field strain were amplified directly from peripheral blood of one of these EIAV-infected horses and characterized by nucleotide sequencing. The complete provirus of Miyazaki2011-A strain is 8208 bp in length with an overall genomic organization typical of EIAV. However, this field isolate possesses just 77.2 and 78.7 % nucleotide sequence identity with the EIAWYoming and EIAVLiaoning strains, respectively, while similarity plot analysis suggested all three strains arose independently. Furthermore, phylogenetic studies using sequences obtained from all EIAV-infected Misaki horses against known viral strains strongly suggests these Japanese isolates comprise a separate monophyletic group.

Equine infectious anemia (EIA) is a persistent lentiviral disease restricted to members of the family Equidae (horses, zebras and donkeys). Although there is considerable variation in clinical responses infection of horses or ponies (Equus caballus) frequently results in an initial or acute febrile episode with an associated thrombocytopenia, followed by a chronic phase characterized by recurring cycles of disease in which the clinical signs include fever, pronounced thrombocytopenia, severe anaemia, jaundice, tachypnea, petechial haemorrhages of the mucosae and cachexia (Clabough et al., 1991; Issel & Coggins, 1979; MacLachlan & Dubovi, 2011; Sellon et al., 1994). If the horse survives, the frequency of disease episodes will gradually diminish and after 12–24 months it will progress to a long-term clinically quiescent phase termed the inapparent carrier state. It is in this state that most infected equids are discovered (Clabough et al., 1991; Issel & Coggins, 1979; MacLachlan & Dubovi, 2011; Sellon et al., 1994). Although inapparent carrier horses appear clinically healthy they remain infected for life with viral replication continuing in macrophage-rich tissues even in the presence of strong virus-specific cellular and humoral immune responses (Hammond et al., 2000; Harrold et al., 2000; MacLachlan & Dubovi, 2011). Therefore, inapparent carrier horses remain potential reservoirs for transmission to other horses (Cheevers & McGuire, 1985) and may experience recrudescence of disease if subjected to environmental stress or immune suppression (Craig et al., 2002; Kono et al., 1976). As a result all infected horses must either be destroyed or remain permanently isolated from all other equids not previously exposed to the virus. Since it was first reported in France in 1843, EIA has posed a significant challenge to veterinary medicine worldwide and caused significant losses to the equine industry (Issel & Coggins, 1979; MacLachlan & Dubovi, 2011).
EIA virus (EIAV), the causative agent of EIA, is a lentivirus in the family Retroviridae (subfamily Orthoretrovirinae). The single-stranded EIAV RNA is approximately 8.2 kb in length and contains the simplest known genomic organization of any extant lentivirus. In addition to the structural proteins encoded by the retroviral prototypical gag, pol and env genes, EIAV possesses just three additional ORFs. These encode the tat and rev proteins that are present in all lentiviruses and a protein designated S2 (Leroux et al., 2004). Although novel field EIAV isolates have been reported based on the nucleotide sequences of gag (Cappelli et al., 2011) and env (Craigo et al., 2009) genes in USA and Europe, the whole proviral genome sequences for these viruses have not been determined. Therefore, it is not known if these European and North American isolates are truly novel because their actual phylogeny might be obscured by genome mosaicism (Lole et al., 1999; Sherefa et al., 1998; Su et al., 2000). Although EIA has a worldwide distribution and was described more than 150 years ago, complete genomic sequences have only been obtained from two isolates EIAVWyoming (GenBank accession no. AF033820, isolated in North America) (Petropoulos, 1997) and EIAVLiaoning (GenBank accession no. AF327877, isolated in China) (Tu et al., 2007). All other reported complete EIAV proviral sequences are laboratory generated derivatives of these two strains.

The Misaki Horse (although technically a pony based on size and conformation) is one of eight breeds considered native to Japan and since the end of World War II has been designated a National Natural Treasure. These feral horses are located in the Toi-Cape area, Miyazaki, in southern Japan. In 2011, 12 Misaki horses were found to be seropositive for EIA in both the agar gel immunodiffusion (AGID or Coggins) (Coggins et al., 1972) and immunoblot tests (Western blotting) (Cook et al., 2005). This serological diagnosis was confirmed in Misaki horses by a nested-PCR assay established in our laboratory (Dong et al., 2012). The full-length proviral genome was amplified from peripheral blood of horse Misaki-A by using long-range PCR. The resultant product was purified and sequenced by primer walking prior to the assembly of overlapping sequences to recreate the full-length EIAV proviral genome designated Miyazaki2011-A. Primers for amplification and sequencing are shown in Table S1 (available in JGV Online). Analysis to identify putative ORFs and their predicted amino acid sequences was conducted using the ORF finder tool (http://www.ncbi.nlm.nih.gov/orf/gorf/gorf.html), with similarity analysis performed using BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi) available via the National Center for Biotechnology Information.

The Miyazaki2011-A proviral genome comprises 8208 bp and possesses a prototypical EIAV genomic organization with the gag, pol and env genes bounded at both ends by long-terminal repeats (LTR). The proviral genome also contains three additional ORFs encoding tat, S2 and rev (Table S2). The LTR is 306 bp in length and as such shorter than both the Wyoming (323 bp) and Liaoning (316 bp) strains. It consists of a 186 bp unique, 3’ region (U3), an 80 bp repeat region (R) along with short 40 bp unique, 5’ region (U5). The putative ORF of the gag gene is 1452 bp in length encoding a predicted 484 aa Gag-precursor polypeptide, while the pol gene ORF is 3408 bp and extends from nt position 1691–5098 in the proviral genome. A probable 2601 bp env gene ORF occupies genomic nt positions 5285–7885 and is predicted to encode a 1356 bp (425 aa) surface unit (SU) glycoprotein (gp90) along with a 1245 bp (415 aa) transmembrane (TM) glycoprotein (gp45). The predicted ORF of tat (246 bp) consists of two exons, one (96 bp) located between the LTR and the start of gag (349–444), while the other (150 bp) is located immediately downstream of the pol gene (5099–5248). The sequences predicted to encode S2 (234 bp) extend over genomic nt positions 5285–5492 and as such all but the extreme 5’ 26 nt form an alternative ORF within the 5’ terminal region of env. Two additional alternative ORFs located within env at positions 5426–5525 and 7232–7635 are predicted to encode rev (Table S2).

Multiple nucleotide sequence alignments were performed to compare the full-length Miyazaki2011-A provirus against complete genomic sequences for EIAVWyoming including some strains derived from it (V26, V70, WSU5 and EIAVUK) plus EIAVLiaoning and its derivative strains (the Chinese Vaccine strain, DV35-20, DLV18-8, DV117 and FDDV-10). The results demonstrated that despite in some cases very extensive passage both in vitro and in vivo, viral strains derived from EIAVWyoming or EIAVLiaoning retained significant similarities to the progenitor isolates with nucleotide sequence identities of 97.5–98.9 % and 97.1–97.9 %, respectively. In contrast, Miyazaki2011-A shared only 77.2 % nucleotide sequence identity with EIAVWyoming and 78.7 % with EIAVLiaoning. This is similar to the 80.2 % nucleotide sequence identity between EIAVWyoming and EIAVLiaoning, suggesting all three viruses arose independently after diverging from a common ancestor. However, actual relationships can be obscured by natural mosaicism and genomic recombination events (Lole et al., 1999; Sa Filho et al., 1998; Sherefa et al., 1998; Su et al., 2000).

The possibility of recombination and/or mosaicism along with potential breakpoints was investigated using the SimPlot algorithm (Lole et al., 1999). For this analysis a sliding window of 200 nt was utilized moving in 20 nt steps with per cent identity calculated for each window and plotted as a line chart. In contrast to EIAVWyoming (Fig. 1b) where there was a very high degree of similarity across the entire genome with the reference strains that were derived from it, the Miyazaki2011-A proviral genome showed evidence of significant variation against all viruses used in this analysis. (Fig. 1a). Furthermore, as suggested from the per cent nucleotide sequence identity, this analysis also demonstrated that the genomes of the North American EIAVWyoming and Asian EIAVLiaoning are not closely related. Therefore, it is highly unlikely that Miyazaki2011-A, EIAVWyoming or EIAVLiaoning-like viruses were derived as a result of recombination events between each other and as
Fig. 1. Complete EIAV genomic plots of nucleotide similarity (generated by SimPlot). (a) Similarity plots of reference EIAV sequences against Miyazaki2011-A, while (b) shows a similar analysis with EIAV\textsubscript{Wyoming} as the query genome. Each curve is a comparison between the genome analysed and a reference genome. Each point represents the per cent identity within a sliding window of 200 bp moving in increments of 20 bp. Horizontal bars indicate the position of the LTRs and each of the major ORFs. Each strain is designated by a different colour.
such they probably constitute separate lineages. Finally, the EIAV strains V70 and V26 have been described as Japanese virulent and attenuated strains, respectively (Zheng et al., 2000). Although they showed relatively low levels of nucleotide similarity with Miyazaki2011-A (Fig. 1a) there was significant conservation with the Wyoming strain (Fig. 1b).

To determine potential relationships between Miyazaki2011-A and other reported EIAV strains, published sequences comprising the LTR, gag, pol and env genes were aligned for phylogenetic analysis. Although several EIAV gag gene sequences have been published for field isolates from Europe (Cappelli et al., 2011) and the New World (Nagarajan & Simard, 2007), information for the LTR and pol is restricted to strains that are derived from either EIAVWyoming or EIAVLiaoning. A similar situation also exists for EIAV env with the exception of a single sequence from a virus strain isolated from an EIA case in Pennsylvania, USA. While the relative paucity of information suggests caution must be exercised in the interpretation of results it was observed that in each case, including the analysis with gag where the most sequence information is available, Miyazaki2011-A comprises a separate monophyletic group (Fig. S1). The analysis also demonstrated that compared with EIAVWyoming and EIAVLiaoning Miyazaki2011-A shared 71.6–80.4, 77.7–81.8, 80.9–81.3 and 71.0–73.3 % nt sequence identity with LTR, gag, pol and env sequences, respectively.

Analysis based on comparative alignments of predicted amino acid sequences for all viral structural proteins along with nucleotide sequences of the LTR demonstrated that almost all functional elements or motifs identified in EIAVWyoming or its derivatives are retained in Miyazaki2011-A (Figs S2, S3, S4, S5, S6 and S7).

In addition to Miyazaki2011-A, 11 other Misaki horses [numbered: Misaki (MY)-5, -25, -29, -35, -47, 53, -56, -65, -67, -69 and -75] were diagnosed as EIA positive. Alignment of a 910 bp 5' fragment (Fig. S8), consisting of 5' LTR sequences (90 bp), the non-coding region (47 bp), tat exon 1 (96 bp), and the first 677 bp of the gag gene amplified from all infected animals revealed 95.1–98.2 % nt sequence identity, suggesting the Misaki horses were infected with related virus strains. In contrast the Misaki horse EIAV isolates possessed only 80.7–81.9 and 81.2–82.5 % nt sequence identity to equivalent sequences within the Wyoming and Liaoning strains (including their derivative viruses), respectively. Phylogenetic analysis based on the 910 bp fragment demonstrated that all viral sequences derived from Misaki horses formed a separate cluster compared with the reference strains related to either EIAVWyoming or EIAVLiaoning (Fig. 2). This provides strong supporting evidence that all 12 Misaki horses were infected with EIAV strains derived from a common ancestor.

Until this time only two complete EIAV proviral sequences [EIAVWyoming (North America) and EIAVLiaoning (China)] have been described. Comparison of nucleotide sequence alignments demonstrated that Miyazaki2011-A was not closely related to EIAVWyoming, EIAVLiaoning or EIAV field isolates from Europe and North America that had been identified based on gag or env (Cappelli et al., 2011; Craigo et al., 2009). Furthermore, similarity plot analysis showed that differences were maintained throughout the entire genome and that it is highly unlikely that Miyazaki2011-A could have arisen recently by recombination with either EIAVWyoming-like or EIAVLiaoning-like isolates. In fact phylogenetic analysis conducted on sequences from the LTR and each of the major structural genes demonstrated that Miyazaki2011-A comprises a separate monophyletic group and can therefore be designated a novel EIAV strain. These results, imply that at least two EIAV subtypes or clades are or were at some time circulating in Asia. Furthermore, phylogenetic analysis performed with the 910 bp EIAV genomic fragments showed that all Misaki horses had been infected with closely related strains that almost certainly shared a common ancestor, an observation entirely consistent with the fact these horses have had no recent contacts with other equids. Additional studies are required to determine the distribution of Miyazaki2011-A-like viruses and the evolutionary relationships between these and other Japanese strains such as EIAVGoshun (Kono et al., 1971), EIAVTokyo (Tabuchi et al., 1967) and EIAVTsukiboshi (Tabuchi et al., 1965) that were originally isolated in the 1940s (EIAVGoshun) and 1960s (EIAVTokyo and EIAVTsukiboshi), respectively.

This study also clarified the situation regarding EIAV strains V70 and V26. These have been described as of Japanese origin with either a virulent (V70) or attenuated (V26) phenotype (Zheng et al., 2000). However, previous studies published in Japanese scientific journals indicate they were derived from a horse that had been inoculated with passaged variants of the Wyoming strain (Kobayashi & Kono, 1967; Kono et al., 1970). The results of the complete genome analysis demonstrated V70 and V26 possessed 98.0 and 97.5 %, respectively, nucleotide sequence identity with EIAVWyoming strongly, suggesting these virus isolates have a North American rather than Japanese ancestry.

A full-length EIAV proviral genome (Miyazaki2011-A) has been successfully amplified from peripheral blood of a naturally infected Misaki horse. The PCR-based strategy developed to achieve this goal avoids passage in vivo or in vitro and results in the amplification of complete proviral sequences. Therefore, it eliminates the selection pressure that can occur when viruses are subjected to a replicative phase and prevents problems such as the assembly of artefact hybrid molecules that inevitably occur when different regions of the viral genome are amplified separately. Consequently, Miyazaki2011-A is predicted to represent an actual provirus and while it is not known if infectious progeny viruses can be derived from it, no obvious fatal defects were detectable by nucleotide sequencing. Based on sequence alignments and similarity plot analysis, Miyazaki2011-A is not closely related to EIAVWyoming, EIAVLiaoning or any other strain reported to date. Moreover, phylogenetic analysis of LTR, gag, pol and env revealed that Miyazaki2011-A forms a separate
monophyletic group and as such can be classified as a novel EIAV isolate.

Acknowledgements

We are grateful to Dr Charles Issel at University of Kentucky, USA, Dr Terrance M. Wilson at California, USA, and Dr Hiroshi Sentusi at Nippon University, Japan, for their valuable comments and suggestions.

References


levels of subtype variability than currently reported for the equine lentivirus family. Retrovirology 6, 95.


