Study on a novel equine infectious anemia virus from feral Misaki horse

野生馬（御崎馬）由来の馬伝染性貧血ウイルスに関する研究

The United Graduate School of Veterinary Science

Yamaguchi University

Jianbao DONG

March 2013

Contents

1. **General Introduction** ... 1

1.1 Equine infectious anemia ... 1

1.1.1 Overview ... 1

1.1.2 Disease course and clinical signs ... 2

1.1.3 Transmission ... 3

1.1.4 Necropsy lesions ... 4

1.1.5 Diagnosis ... 5

1.1.6 Prevention and control ... 9

1.1.7 Vaccine development ... 10

1.2 Equine infectious anemia virus ... 11

1.2.1 Overview ... 11

1.2.2 Long terminal repeat ... 13

1.2.3 Structural proteins ... 13

1.2.4 Accessory proteins ... 16

1.3 EIA in Japan ... 16

1.2 Feral Misaki horse ... 17

2. **Establishment of a novel nested PCR detection method** ... 19

2.1 Introduction ... 20

2.2 Materials and methods 23

2.2.1 Primer design 23

2.2.2 Positive control 23

2.2.3 Optimization for annealing temperature 24

2.2.4 Sensitivity 24

2.2.5 Application for field samples 25

2.2.6 Serological diagnositic techniques 26

2.3 Results 26

2.4 Discussion 28

2.5 Tables and figures 33

3. Identification of a novel EIAV field strain 37

3.1 Introduction 38

3.2 Materials and methods 40

3.2.1 Diagnosis and histopathology 40

3.2.2 Sample collection and DNA extraction 41

3.2.3 EIAV diagnostic PCR 41

3.2.4 Complete EIAV proviral genome amplification 42

3.2.5 Direct nucleotide sequencing 42

3.2.6 Cloning and sequencing 43

3.2.7 Genomic analysis 43

3.2.8 A 910 bp EIAV genomic fragment Amplification 44

3.2.9 Nucleotide sequence accession numbers 45

3.3 Results 45

3.3.1 Diagnosis of EIA in feral Misaki horses 45

3.3.2 Full-length EIAV proviral genome 46

3.3.3 Sequence identity among complete EIAV genomes 47

3.3.4 Phylogenetic analysis using LTR, *gag*, *pol* and *env* genes 49

3.3.5 Population phylogenetic analysis 49

3.3.6 Characterization of LTR and structural genes 50

3.4 Discussion 54

3.5 Tables and figures 62

4. Conclusion 75

5. Acknowledgement 78

6. References 79

1. General Introduction

1.1 Equine infectious anemia

1.1.1 Overview

Equine infectious anemia (EIA), also known as swamp fever, is an infectious viral disease that affects only members of the equine species, including horses, ponies, donkeys, and mules. EIA is caused by equine infectious anemia virus (EIAV), a lentivirus in the family *Retroviridae* (subfamily *Orthoretrovirinae*). The disease is characterized by acute and/or chronic recurring clinical signs including depression, fever, anemia, edema and cachexia. EIA was first reported in France in 1843 (MacLachlan & Dubovi, 2011) and was diagnosed as a nutritional disease by the scientists who noted the symptoms at that time. The weak horses were failed to be cured. In 1861, the disease was reported again and forced the scientists to search the cause of the disease. Two years later, EIA was documented in Switzerland. And two years after that, it spread through the German cavalry horses in Germany. Subsequently, almost all of the European countries, several African countries and much of Asia have reported cases of EIA. Japan was also subjected to the first epidemics and more than 300 horses a year died from this disease during the first decade of 19th century (Communication of Florida Horse.Com). In 1904, the disease was confirmed to be associated with the infection by a "filterable agent", which became one of the first animal diseases to be assigned a viral etiology (MacLachlan & Dubovi, 2011). There is no cure for this disease. Since it was first

reported, EIA has been reported worldwide and posed a significant challenge to veterinary medicine worldwide and caused significant losses to the equine industry (Issel & Coggins, 1979; MacLachlan & Dubovi, 2011). EIA threaten not only domestic equid but also wild animals. Recent report from Brazil on living wild horses have shown that about 30% of domesticated and about 5.5% of the wild horses are chronically infected with EIA in the Pantanal (Silva *et al.*, 1999).

1.1.2 Disease course and clinical signs

The incubation period of EIA is a week to 45 days or longer. Some horses remain asymptomatic until they are stressed. Although the course of disease following primary infection can be highly variable, in some cases resulting in the loss of life whereas in others the equid remains asymptomatic, horses commonly experience an acute clinical episode, lasting one to three days in which the main disease signs are fever and thrombocytopenia. This is often followed by a chronic disease phase characterized by episodic bouts of fever, thrombocytopenia, depressed neurologic signs, anemia, edema and cachexia. In some particularly severe cases jaundice, bloodstained feces, tachypnea, and petechial hemorrhages of the mucosae are also apparent (Clabough *et al.*, 1991; Issel & Coggins, 1979; Lairmore, 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 (Craigo *et al.*, 2002; Kono *et al.*, 1976).

For donkeys and mules, the infection is less likely to cause severe clinical signs. Mules can be infected asymptomatically, but typical EIA signs have been reported in some naturally or experimentally infected animals. It was reported donkeys inoculated with two horse-adapted strains showed infection without clinical signs. However, donkeys inoculated with a serially-passaged, donkey-adapted strain are reported to have developed clinical signs.

1.1.3 Transmission

EIA is a thought to be a blood-born infectious disease. It can be transmitted through blood, saliva, milk, and body secretions. The virus can to be transmitted mechanically on the mouthparts of biting insects. EIAV persists in blood leukocytes for life in the infected horses, and also can be detected in plasma during febrile episodes. Moreover, the Symptomatic horses are more likely to transmit the disease than animals with asymptomatical infections. After biting an carrier, the insect is likely to become a vector.

Biting flies in the family *Tabanidae*, especially horse flies (*Tabanus spp. and Hybomitra spp.*) and deer flies (*Chrysops spp.*), are the most effective vectors for EIAV transmission, although other insects including stable flies (*Stomoxys calcitrans*) are also have the possibility for transmission. In the biting course, the equid reacts with the biting because of pain, and interrupts the feeding. Then the fly attempts to resume feeding on the same animal or on another nearby host, which result to transfer the infectious blood. EIAV survives for a limited time on the mouthparts of insects, and it is less likely to be spread to more distant hosts.

EIAV can also be transmitted by contaminated surgical equipment, recycled needles and syringes in blood transfusions and teeth floats. It is reported to persist for up to 96 hours on hypodermic needles. EIAV may also be passed by the vertical transmission from a mare to her foal in utero. There are also some minor routes for EIAV transmission. For example, the virus can be found in milk and semen, and horses can be infected by inoculating these secretions subcutaneously. In some nursing foals, possible transmission through milk has been reported. The venereal transmission is thought to be one of the minor routes for EIAV transmission, but there was one has been reported to have transmitted the virus to a mare with a vaginal tear during breeding. In 2006, the possibility of aerosol transmission by infectious material during close contact was raised during the EIA outbreak in Ireland (Cullinane *et al.*, 2007; More *et al.*, 2008).

1.1.4 Necropsy

In chronic cases, the gross necropsy findings include emaciation, an enlarged

meaty spleen, an enlarged liver with a prominent lobular pattern, mucosal and visceral hemorrhages, generalized lymph node enlargement, ventral subcutaneous edema, the pale mucous membranes and vascular thrombosis. Edema is often found in the limbs and along the ventral abdominal wall. Petechiae may be observed on internal organs such as the spleen and kidney. And blood vessel thrombosis have also been reported. Histopathology usually shows accumulations of lymphocytes and macrophages in sinusoids and portal areas of the liver, in medullary sinus of lymph nodes, adrenal glands, spleen, meninges, and lung. This lymphoproliferation may be due to an attempt to control the infection by the T-lymphocytes. There is also marked extramedullary hematopoiesis. Other liver lesions include fatty degeneration and hepatocellular necrosis. Kupffer cells are swollen with hemosiderin accumulation. In chronically infected horses, gross lesions are generally unremarkable, although some animals may have proliferative glomerulonephritis or ocular lesions.

1.1.5 Diagnosis

EIA diagnosis is usually carried out via clinical signs, differential diagnosis and laboratory tests. For clinical signs, EIA should caused differentials in individual horses with depression, edema, weight loss, and intermittent fever. Meanwhile, we also need to pay attention to horses when several of them showed clinical signs such as fever, anemia, edema, progressive weakness or weight loss, particularly when new animals have been introduced into the herd or a member of the herd has died. During the clinical diagnosis, other febrile illnesses, including

equine viral arteritis, purpura hemorrhagica, leptospirosis, babesiosis, severe strongyliasis or fascioliasis, phenothiazine toxicity, autoimmune hemolytic anemia and other diseases that cause fever, edema and/or anemia, are needed to pay attention for differential diagnosis.

For laboratory tests, there are several methods including virus isolation and identification, agar gel immunodiffusion (AGID or Coggins) test (Coggins *et al.*, 1972), enzyme-linked immunosorbent assays (ELISAs) (Suzuki *et al.*, 1982), western blotting (Cook *et al.*, 2005), polymerase chain reaction (PCR), reverse-transcriptase PCR (RT-PCR), and real-time PCR assays (Cook *et al.*, 2002; Nagarajan & Simard, 2001; 2007). Virus isolation is usually thought to be not necessary for making a diagnosis. EIAV may be found in both plasma and blood leukocytes during febrile episodes, between these periods, the virus is cell-associated. To isolate EIAV from suspect horses usually need to inoculate their blood on to leukocyte cultures prepared from horses free of infection. Then the virus production in cultures can be confirmed by detection of specific EIA antigen by ELISA (Shane *et al.*, 1984), by imunofluorescence assay (Weiland *et al.*, 1982), by molecular tests or by subinoculation into susceptible horses. However, the disadvantage of virus isolation is the difficulty of growing horse leukocyte cultures.

EIA is often confirmed by serological methods. Once an animal is infected, it becomes a carrier for life. The two most commonly used serological tests are the AGID (also called Coggins) test (Coggins *et al.*, 1972) and enzyme-linked ELISAs (Suzuki *et al.*, 1982). Precipitating antibody is rapidly produced as a result of EIA

infection, which can be detected by the AGID test. Specific reactions are indicated by precipitin lines between the EIA antigen and the test serum. At present, there are four ELISAs including a competitive ELISA and three non-competitive ELISAs that are approved by the United States Department of Agriculture for the diagnosis of equine infectious anemia and are available internationally. The competitive ELISA and two non-competitive ELISAs detect antibody produced against the p26 core protein antigen. The third non-competitive ELISA incorporates both p26 core protein and gp45 (viral transmembrane protein) antigens (OIE, 2008). In 2005, the immunoblot test (western blot) has been used as a research tool to detect EIA infection and was showed to be a sensitive serological method (Cook *et al.*, 2005). According to OIE Terrestrial Manual 2008, all positive samples by ELISA should be confirmed with AGID because some false-positive results have been noted with the ELISA (OIE, 2008), and also can be confirmed with immunoblotting test (Western blotting).

Although serologic tests have been well established, all of them have recognized limitations because some false-positive results and indeterminate results occurred in practice. Therefore, direct detection of EIAV need to be developed to supplement or confirm serological tests. Based on this, a nested polymerase chain reaction (PCR) assay to detect EIA proviral DNA from the peripheral blood of horses has been described (Nagarajan & Simard, 2001). The nested PCR method is based on primer sequences from the *gag* region of the proviral genome. It has proven to be a sensitive technique to detect field strains of EIAV in white blood cells of EIA infected horses; the lower limit of detection is

typically around 10 genomic copies of the target DNA (Nagarajan & Simard, 2001; 2007). Meanwhile, a real-time reverse-transcriptase PCR assay has also been established to detect the virus and showed to be an efficient detection tool (Cook *et al.*, 2002).

Comparing above diagnostic methods, extensive variation between individual cases coupled with the lack of specific disease characteristics precludes the diagnosis of EIA on the basis of clinical signs or pathological lesions. In addition, conventional virus isolation attempts in equine monocyte derived macrophage cultures are not sufficiently sensitive or reliable enough for routine detection purposes. Therefore, the diagnosis of EIA is currently exclusively dependent on indirect serological detection techniques with the most commonly used method being the AGID or Coggins test (Coggins *et al.*, 1972). Although ELISA-based tests are approved for use in some countries (Centaur, 1993; Matsushita *et al.*, 1989) these nations follow the OIE Terrestrial Manual 2008 guidelines requiring that ELISA positive samples be confirmed by AGID. This is because false-positive results have occasionally been noted with the potentially more sensitive ELISA tests (OIE). In addition to these commercially available assays the immunoblot test (Western blot) has been used both as a research tool (Cook *et al.*, 2005) and very recently in an official capacity to resolve difficult diagnostic cases in Italy (C. J. Issel personal communication). While serological detection techniques have been used successfully in EIA control for more than four decades they all suffer from the disadvantage they cannot detect cases from the time of exposure until the development of sufficient virus specific antibodies. This

problem may be particularly evident in the case of the relatively insensitive AGID test. Although most equids become seropositive in AGID within 45 days of exposure to EIAV, delays of 157 days have been reported (Cullinane *et al.*, 2007). Furthermore, AGID failed to detect the presence of EIAV infected horses following a compulsory 90 day quarantine period on a farm at the center of the 2006 Italian outbreak resulting in eight new clinical cases (Cappelli *et al.*, 2011). Detection of viral genetic material by sensitive PCR or quantitative RT-PCR has the potential to overcome the shortfalls associated with indirect serological diagnostic techniques and such methods have been successfully applied to the detection of laboratory adapted strains along with some field EIAV isolates (Cappelli *et al.*, 2011; Cullinane *et al.*, 2007; Nagarajan & Simard, 2001; OIE, 2008; Quinlivan *et al.*, 2007; Zheng *et al.*, 2000). Since these PCR-based diagnostic assays showed to be very sensitive, OIE recommended that duplicate samples of each diagnostic specimen be processed, PCR positive samples need to be confirmed by AGID test. Because of the risk of cross contamination, it is also important that proper procedures are followed (OIE, 2008).

1.1.6 Prevention and control

For EIA prevention, it is very important to make the prevention programs to test EIA before entry of the horse, and require the local horses for test in regular. Regular voluntary testing of the equids on a farm, as well as testing of new animals before introduction, is very helpful in maintaining an EIA-free herd. It is thought to be a hard work to control this disease. Once equids are infected by EIAV, they

become lifelong carriers. Since there is no efficacious vaccine against EIA, the infected horse must be permanently isolated from other susceptible animals or euthanized. When a horse is confirmed as infection, it must be marked and isolated as soon as possible. If an outbreak happened, we need to control insect vectors by spraying, as well as the use of insect repellents and insect-proof stabling, which may aid in interrupting transmission. Meanwhile, it is better to place animals in small groups separated by at least 200 yards might be beneficial when the virus is being transmitted within a farm. Moreover, we need to pay attention to our handling to avoid iatrogenic transmission. Most of common disinfectants can be used to destroy the virus particle in farm since EIAV is an enveloped virus.

1.1.7 Vaccine development

At present, it is a great challenge to develop an efficacious vaccine against EIA and other lentiviral infections in human and veterinary medicine field. Although some encouraging results were reported, so far there has no actually available vaccine for EIA. For most naturally or experimentally EIAV-infected animals, they usually be able to control EIAV replication and disease processing within a certain period, which suggested it may be effective to develop a vaccine to prevent this disease. However, to date, all international efforts, including the classical vaccines based on inactivated or attenuated whole virus and vaccines with viral recombinant protein, failed to induce a broadly protective immune response. More than 20 years ago, one effective vaccine, using a live attenuated EIAV strain produced by serial passages on donkey leukocyte cells, was reported in China

(Shen & Wang, 1985). Unfortunately, since there were no independent investigations, the vaccine were not accepted by the international scientific community. Another research on inactivated vaccine showed that the prototype cell-adapted Wyoming strain (Malmquist *et al.*, 1973) efficiently protected ponies against the challenge with the homologous strain. But for the challenge of heterologous strain, although the virus replication was significantly reduced, it failed to prevent infection (Issel *et al.*, 1992). The subunit vaccine was also tried, but the result show the subunit vaccine enriched in envelope glycoproteins not only failed to protect the ponies against the attacks by homologous or heterologous strain, but also possessed a high potential to enhance EIAV replication and exacerbate the disease. Another live attenuated EIAV vaccine was constructed by using a molecular clone $EIAV_{UK}$ which carrying two inactivating stop codons in the S2 ORF (Li *et al.*, 2003). Experimental result showed this vaccine protected the animal against the challenge with a homologous virus strain (Li *et al.*, 2003). However, the protection experiment was not done for challenge with a heterologous EIAV strain. In conclusion, so far there have been some encouraging results reported from the EIA vaccine trails, but still have many problems need to be solved.

1.2 Equine infectious anemia virus

1.2.1 Overview

The EIAV is a lentivirus in the family *Retroviridae*, subfamily *Orthoretrovirinae*. Other members of the lentivirus genus include: bovine

immunodeficiency virus; caprine arthritis encephalitis virus; feline immunodeficiency virus; human immunodeficiency virus 1; human immunodeficiency virus 2; and maedi/visna virus. Nucleic acid sequence comparisons have demonstrated a marked relatedness among these viruses. 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 open reading frames (ORFs). These encode the *tat* and *rev* proteins that are present in all lentiviruses and a protein designated as S2 (Leroux *et al.*, 2004).

Although EIA has a worldwide distribution and was described more than 150 years ago, so far only two independent EIAV strains ($EIAV_{Wyoming}$ [North America] and $EIAV_{Liaoning}$ [China]) have been identified based on their complete proviral sequences. All other reported strains with complete proviral sequences are laboratory-generated derivatives of these two strains. 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 sequence 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 which is often happened in other lentivirus (Lole *et al.*, 1999; Sa Filho *et al.*, 2005; Sherefa *et al.*, 1998; Su *et al.*, 2000).

1.2.2 Long terminal repeat

In lentivirus, long terminal repeats (LTR) consist of three segments namely U3 (unique, 3' end), R (Repeated) and U5 (unique, 5' end) and serve as the site of transcriptional initiation. LTR contains important elements for virus replication and viral transcription. The lentivirus promoters are regulated by various cellular DNA-binding factors interacting with specific motifs on the LTR sequence. In common with all reported EIAV strains possesses a consensus TATA-box, TATATAA along with three consecutive G residues that serve as the initiation site for RNA synthesis and define the R-U5 border. They also contain a consensus poly(A) addition signal (AATAAA). Additionally, LTR also contains *cis*-acting DNA elements corresponding to the methylated DNA-binding protein site MDBP, two PEA2 elements, a PEA3/ets motif and an AP-1 site (Carvalho & Derse, 1993a). The PU.1 (purine-rich element 1), which recognize the *ets* element, was showed to positively regulate the promoter of EIAV (Carvalho & Derse, 1993b). The generation of B-lymphocytes and macrophages, the natural target of EIAV infection, require the expression of PU.1 which is normally restricted to the haematopoietic system (Lloveras *et al.*, 1999).

1.2.3 Structural proteins

The structural proteins of EIAV are encoded by *gag*, *pol* and *env* genes, and produced from polyprotein precursor molecules. The products of *gag* and *pol* genes are translated from the full-length viral messenger RNA (Noiman *et al.*, 1990). The Gag-Pol polyprotein was synthesized with the ribosomes shifting their

translational reading frame to read through the stop codon in *gag*. This processing is regulated by several critical motifs including an AAAAAAC slippery sequence, a 5-base GC-paired segment downstream of the slippery sequence and a pseudoknot structure (Chen & Montelaro, 2003). When the next generation virus bud from the host cells, the Gag precursor proteins is necessary to be expressed on the plasma membrane. The EIAV Gag-precursor (Pr55gag) polyprotein is cleaved by viral encoding protease into four major internal structural proteins of the mature virion. These comprise the the membrane-interacting matrix (MA) p15, the capsid (CA) p26, the RNA-binding nucleocapsid (NC) proteins p11 and p9 (Hussain *et al.*, 1988; Stephens *et al.*, 1986). It is reported that conservation of amino acid sequences encoded by EIAV *gag* conforms to a hierarchy in which p26 > p11 > p15 > p9. The MA or p15 in all $EIAV_{Wyoming}$-like viruses contains a potential leucine-rich type nuclear export signal (LKKLEKVTV) (Hatanaka *et al.*, 2002). Although the actual function of this motif is unknown the fact that it is likely to be important within the EIAV replication cycle. The CA contains a motif present in many retroviruses termed the Major Homology Region (MHR) that is believed to be essential for virus assembly and budding from the host cell plasma membrane (Grund *et al.*, 1994). The NC proteins of all lentiviruses share a high percentage of basic amino acid residues and one or two zinc-binding motifs. EIAV contains two of these zinc binding regions (C-X2-C-X4-H-X4-C where C represents cysteine, H represents histidine and X indicates a variable amino acid residue) separated by an usually short five basic residue-rich (R-A-P-K-V) linker (Amodeo *et al.*, 2006). Meanwhile, the p9 YPDL late domain that has been shown to be essential for

budding from the host-cell plasma membrane in EIAV (Chen *et al.*, 2001; Li *et al.*, 2002; Puffer *et al.*, 1997). The Pol protein is yielded from the Cleavage of the EIAV Gag-Pol precursor (Pr180gag/pol) and it produce various enzymatic activities including the reverse transcriptase-RNaseH (p66) essential for the conversion of viral RNA into DNA, a viral protease (p12) for the processing of the polyproteins, a dUTPase (p15) essential for EIAV replication in non dividing monocyte-derived macrophages, an integrase and the natural host cells for EIAV expression and replication (Lichtenstein *et al.*, 1995; Steagall *et al.*, 1995; Threadgill *et al.*, 1993).

The surface unit (SU, gp90) and transmembrane envelope glycoproteins (TM, gp45) were encoded by *env* gene and produced by the action of host-cell specified proteases. The gp90 may interact with the cellular receptor, still unknown for EIAV. It is highly glycosylated; based on the deduced amino-acid consensus sequence of $EIAV_{PV}$ strain, 17 and 5 potential N-linked glycosylation sites (NX[S/T]) are present in gp90 and gp45 respectively (Leroux *et al.*, 1997). The gp90-surface glycoprotein is showed for rapid evolution during the course of disease and the mutations are restricted to defined variable regions, V1 to V8, in gp90 (Leroux *et al.*, 1997). The cysteine residue sites are conserved in gp90 between different EIAV strains and that this may contribute via the formation of disulfide bridges to maintaining structural or functional integrity in this otherwise highly variable envelope glycoprotein (Craigo *et al.*, 2009). Lentiviruses utilize dense glycosylation to shield the envelope proteins from immune recognition.

1.2.4 Accessory proteins

In addition to the structural proteins encoded by *gag*, *pol* and *env*, EIAV has three additional ORFs encoding the small accessary proteins Tat, S2 and Rev, This is in contrast with the seven additional genes contained in the genome of the primate lentiviruses (Thomas & Furman, 1991). The Tat or trans-activator protein regulates transcription of EIAV proviral DNA. It consists of the amino terminus, core region, basic region and carboxy terminus, but lacks a cysteine-rich region found in other lentiviruses (Carroll *et al.*, 1991). EIAV Tat interacts with the viral LTR through the trans-activation responsive (TAR) element, suggesting that highly divergent Tat and TAR elements control viral gene expression by the same mechanism (Albrecht *et al.*, 2000; Bieniasz *et al.*, 1999; Sune *et al.*, 2000). The *S2* gene is located in the *pol-env* intergenic region following the second exon of *tat* and overlaps with the 5'- terminal of the *env* gene while Rev is encoded by two exons comprising alternate reading frames within the *env* gene.

1.3 EIA in Japan

EIA was first reported in Japan in 1883. Since then, it has threaten the horse breeding in Japan. Because of EIAV infection, about 10,000 horses were killed in 1952, and more than 50,000 horses were killed in 1950s (Sentsui, 2011). The kinetic of infection amount illustrated that infection peak was showed in 1952, and then it has been decreased because the amount of breeding horses have been reduced. However, one small peak was showed from 1973 to 1975 and total of 591 horse were infected. AGID method was introduced into Japan and accepted by the

government to detect EIAV infection in 1978. The diagnosis became standard and healthy checking for EIAV has been held every year. Therefore, the disease was controlled step by step. In 1983, 4 horses were diagnosed as positive, and then there had been no positive horse reported during 1984 to 1992. EIA was thought to disappear in Japan. Unfortunately, 2 horses were diagnosed as EIA positive in 1993 in the same farm where EIA was reported 10 years ago. From 1993 to 2010, there had been no reports for EIA in Japan. However, after the 18-years silence, the disease was reported again in 2011.

In Japan, it is a duty to accept regular health examination for equine. If the animal is diagnosed as EIA positive, it will lose the qualification not only for the horse racing but also for reproduction. Moreover, according to Act on Domestic Animal Infectious Diseases Control in Japan, the infected horses must be euthanized, and incinerated or buried as soon as possible. It is thought to possess a good system and experience to control EIA in Japan.

1.4 Feral Misaki horse

The Misaki horses were first documented more than 300 years ago and has been designated as a natural monument in Japan in 1958. It is considered as one of eight native breeds to Japan, also including Hokkaido Horse, Kiso Horse, Noma Horse, Tsushima Horse, Tokara Horse, Miyako Horse, and Yonaguni Horse. The Misaki horse inhabits in the Toi-Cape area, at the south end of Miyazaki Prefecture on the island of Kyusyu, in southern Japan. It was documented in the historical record in 1697 when the Akizuki family of the Takanabe Clan rounded up feral

horses and developed a pool of breeding stock. At present, the Misaki horse is classified as an endangered breed, with only about 100 living animals. This population has remained relatively stable for the past 30 years, up from a low of 53 individuals recorded in 1973 [Country Report (For FAO State of the World's Animal Genetic Resources Process)].

The Misaki horse average 12.2 to 13 hands (130 – 135 cm) in height. It is exactly a breed of pony, under the official definition of a pony is a horse that measures less than 14.2 hands (58 inches, 147 cm) at the withers. They are mostly bay or black, with a few chestnuts. Their average life of male is 14 years while female is 16 years. The female horse can give birth after 3 years old and can give 5 to 6 birth during the whole life. The Misaki horse is thought to be one of pure horse species. However, One male hybrid male horse named Komatsugo was introduced into Misaki horse herd in 1913. The father of Komatsugo is Standardbred horse from America and his mother is a hybrid from Hokkaido Wasyu and Nanbu horses. But the hybridization contribution of Komatsugo is thought to be very limited to Misaki horse.

2. Establishment of a nested PCR diagnosis test

Development of a nested PCR assay to detect equine infectious anemia proviral DNA from peripheral blood of naturally infected horses

2.1 Introduction

Equine infectious anemia (EIA) is a disease affecting all members of the *Equidae*. The causative agent, EIA virus (EIAV) is a lentivirus in the family *Retroviridae* (subfamily *Orthoretrovirinae*) (Lairmore, 2011; Montelaro *et al.*, 1993). Although the course of disease following primary infection can be highly variable, in some cases resulting in the loss of life whereas in others the equid remains asymptomatic, horses commonly experience an acute clinical episode, lasting one to three days in which the main disease signs are fever and thrombocytopenia. This is often followed by a chronic disease phase characterized by episodic bouts of fever, thrombocytopenia, depressed neurologic signs, anemia, edema and cachexia. In some particularly severe cases jaundice, bloodstained feces, tachypnea, and petechial hemorrhages of the mucosae are also apparent (Clabough *et al.*, 1991; Issel & Coggins, 1979; Lairmore, 2011; Sellon *et al.*, 1994). After 12 to 24 months the frequency of disease episodes diminishes and the equid enters a prolonged phase, the inapparent carrier state where it is devoid of overt clinical signs. The transition to this state is dependent on the development of broadly reactive adaptive immune responses that although effectively restrict viral replication, do not eliminate EIAV and so the equid remains a potential reservoir of the infection (Cheevers & McGuire, 1985). Therefore, in order to prevent transmission, infected animals must either be euthanized or remain permanently isolated from other equids. Since it was first described in France in 1843, EIA has been reported worldwide and even today poses a considerable threat to the equine industry in some countries (Issel & Coggins, 1979; Lairmore, 2011).

Extensive variation between individual cases coupled with the lack of specific disease characteristics precludes the diagnosis of EIA on the basis of clinical signs or pathological lesions. In addition, conventional virus isolation attempts in equine monocyte derived macrophage cultures are not sufficiently sensitive or reliable enough for routine detection purposes. Therefore, the diagnosis of EIA is currently exclusively dependent on indirect serological detection techniques with the most commonly used method being the agar gel immunodiffusion (AGID) or Coggins test (Coggins *et al.*, 1972). Although enzyme-linked immunosorbent assay (ELISA)-based tests are approved for use in some countries (Centaur, 1993; Matsushita *et al.*, 1989) these nations follow the OIE Terrestrial Manual 2008 guidelines requiring that ELISA positive samples be confirmed by AGID. This is because false-positive results have occasionally been noted with the potentially more sensitive ELISA tests (OIE). In addition to these commercially available assays the immunoblot test (Western blot) has been used both as a research tool (Cook *et al.*, 2005) and very recently in an official capacity to resolve difficult diagnostic cases in Italy (C. J. Issel personal communication). While serological detection techniques have been used successfully in EIA control for more than four decades they all suffer from the disadvantage they cannot detect cases from the time of exposure until the development of sufficient virus specific antibodies. This problem may be particularly evident in the case of the relatively insensitive AGID test. Although most equids become seropositive in AGID within 45 days of exposure to EIAV, delays of 157 days have been reported (Cullinane *et al.*, 2007). Furthermore, AGID failed to detect the presence of EIAV infected

horses following a compulsory 90 day quarantine period on a farm at the center of the 2006 Italian outbreak resulting in eight new clinical cases (Cappelli *et al.*, 2011). Detection of viral genetic material by sensitive PCR or quantitative RT-PCR has the potential to overcome the shortfalls associated with indirect serological diagnostic techniques and such methods have been successfully applied to the detection of laboratory adapted strains along with some field EIAV isolates (Cappelli *et al.*, 2011; Cullinane *et al.*, 2007; Nagarajan & Simard, 2001; OIE, 2008; Quinlivan *et al.*, 2007; Zheng *et al.*, 2000).

In 2011, EIA was detected in the Misaki horses, using both AGID and Western blotting techniques. These are feral horses located in the Toi-Cape area, Miyazaki prefecture, in southern Japan. Despite clear evidence of positive serological reactions proviral DNA was not demonstrable in peripheral blood of any of the infected horses using PCR primers recommended by OIE (OIE Terrestrial Manual 2008) (Nagarajan & Simard, 2001; OIE, 2008), previously published primer sets (Quinlivan *et al.*, 2007; Zheng *et al.*, 2000) or unpublished primer pairs currently being evaluated in other laboratories. In the present study, a novel nested PCR assay was established using primers designed within conserved 5' regions of the EIAV genome extending from within the long terminal repeat (LTR) to the trans-activator (*tat*) gene. This assay successfully detected proviral DNA from peripheral blood cells in all seropositive horses proving it to be a rapid, sensitive and specific technique for the diagnosis of EIAV in naturally infected horses.

2.2 Materials and methods

2.2.1 Primer design

A total 11 whole genome sequences of EIAV were downloaded from GenBank (Wyoming (AF033820), USA; Liaoning (AF327877), China; V26 (AB008197), Japan; DLV18-8 (HM141923), China; DV35-20 (HM141911), China; EIAVuk (AF016316), USA; DV117 (HM141912), China; FDDV-10 (GU385360), China; V70 (AB008196), Japan; Vaccine Strain (AF327878), China; WSU5 (AF247394), USA) and aligned using the Laser Gene sequence analysis package (DNASTAR, Inc., Madison, WI) (Thompson *et al.*, 1997). Subsequently, specific primers were designed and analyzed using the Oligo 6.31 program, (Molecular Biology Insights, Inc., CO) based on highly conserved 5' EIAV genomic sequences extending from the LTR to the *tat* gene (Fig. 1). Sequences of the oligonucleotide primers are shown in Table 1.

2.2.2 Positive control

DNA was extracted from peripheral blood cells of a feral horse Miyazaki-A, diagnosed as EIAV positive in both AGID and Western blotting (Cook *et al.*, 2005), using a DNeasy blood and tissue kit (Qiagen, Maryland, MD) according to the manufacturer's instructions. PCR was performed with outer primer set (EIAVltr-1F and EIAVltr-1R) using Takara Ex Taq (Takara, Shiga, Japan) in a 50 μl reaction volume with cycling conditions consisting of an initial denaturation step at 94°C for 5 min followed by 35 cycles of 94°C for 45 s, 56°C (recommended by Oligo 6.31 program) for 45s, 72°C for 60 s, and a final

extension of 72°C for 10 min. The product was analyzed by electrophoresis in a 1.5% agarose gel. A fragment of the predicted size (246 bp) was purified with QIAquick gel extraction kit (Qiagen, Hilden, Germany) and cloned into pMD20-T vector (Takara, Shiga, Japan). The cloned insert was sequenced and shown to have significant homology to EIAV following a Blast search in GenBank. The recombinant plasmid, named pMD-A246, was used as positive control to optimize the amplification conditions of the nested PCR assay.

2.2.3 Optimization for annealing temperature

Optimal annealing temperature was determined by gradient PCR with temperatures of 50, 52, 54, 56, 58 and 60 °C. Using pMD-A246 (10^5 copies/μl) as template, the amplification for two primer sets was carried out using an initial denaturation at 94°C for 5 min followed by 35 cycles of 94°C for 30 s, 30 s at the different gradient annealing temperatures, 72°C for 40 s, and a final extension of 72°C for 10 min. Ten microliter of the PCR products were electrophoresed in 1.5% agarose gels and the intensity of GelRed nucleic acid stain (Biotium, Hayward, CA) analyzed using software associated with a CS Analyzer Ver 3.0 (ATTO, Tokyo, Japan) to determine the optimal annealing temperature.

2.2.4 Sensitivity

Plasmid pMD-A246 was serially diluted 10-fold with DNase-free sterile distilled water from 10^6 to 10^{-1} copies/μl and 5 μl of each dilution used as template to determine the sensitivity of the individual outer and inner primer set PCR in

addition to the full nested reaction. Individual first and second stage reactions were conducted using Takara Ex Taq (Takara, Shiga, Japan) in a final 50 μl volume according to the manufacturer's instructions. Cycle amplification was carried out using an initial denaturation at 94°C for 5 min followed by 35 cycles of 94°C for 30 s, 30 s at the optimized annealing temperature described in the section of *optimization for annealing temperature*, 72°C for 40 s, with a final extension of 72°C for 10 min. For nested PCR, 2.5 μl of the first round reaction product was used as template with cycle amplification conditions being the same as those used for single outer and inner primer set PCRs. Ten microliter of the product from each PCR was analyzed by electrophoresis in 1.5% agarose gels.

2.2.5 Application for field samples

To evaluate the diagnostic efficacy of the nested PCR assay, a total of 14 peripheral blood samples were collected from the Misaki feral horse population in Toi-cape area, Miyazaki, in southern Japan. DNA was extracted using a DNeasy blood and tissue kit (Qiagen, Maryland, MD) according to the manufacturer's instructions and EIAV proviral sequences amplified using the optimized version of the nested PCR assay as described above. The resultant PCR-generated products were purified with QIAquick gel extraction kit (Qiagen, Hilden, Germany) prior to cloning into the pMD20-T vector (Takara, Shiga, Japan). The nucleotide sequences of the cloned inserts were determined using a Bigdye Terminator v3.1 cycle sequencing kit (Applied Biosystems, Foster City, CA) in conjunction with an ABI Prism 3130 genetic analyzer (Applied Biosystems, Foster City, CA). Sequencing

results were analyzed with the LaserGene sequence analysis package (DNASTAR, Inc., Madison, WI) (Thompson *et al.*, 1997) and GENETYX Ver. 7.0.9 (Genetyx Corp., Tokyo, Japan).

2.2.6 Serological diagnostic techniques

Serum samples from each of the 14 feral horses were tested for the presence of antibodies against EIAV using AGID (Nisseiken, Tokyo, Japan) as specified by the manufacturer or by Western Blotting as described previously (Cook *et al.*, 2005).

2.3 Results

A nested PCR for the amplification of EIAV proviral sequences has been developed using oligonucleotide primers (*outer primer pair:* EIAVltr-1F/EIAVltr-1R and *inner primer pair:* EIAVltr-2F/EIAVltr-2R) designed based on highly conserved 5' viral genomic sequences extending from the LTR to the *tat* gene. The predicted DNA fragment product sizes are 246 bp and 198-203 bp (dependent on EIAV strain) for the first and second round PCR amplification steps respectively. The optimum annealing temperatures for both reaction steps were determined by gradient PCR using cloned sequences (pMD-A246) obtained from an EIAV infected horse (Miyazaki-A) and amplified with the EIAVltr-1F/EIAVltr-1R primer pair. Optimal temperatures for the first and second stage reactions were shown to be 52 °C and 58°C respectively. Sensitivity of the first, second and nested PCR were determined under optimal amplification

conditions using 10-fold serial dilutions of plasmid pMD-A246 as the template. Both the individual first and second stage reactions have detection limits of 100 genomic copies of the target DNA, while the combined or nested PCR is capable of detecting just 10 template molecules (Fig. 2).

As a result of a testing program some feral horses living in the Toi-Cape area of Japan were discovered to have been exposed to EIAV. Consequently, serum and peripheral blood samples were collected from 14 of these horses. It should be noted that none showed disease signs consistent with clinical EIA at the time of collection. In the conventional AGID test 10 horses possessed convincing serological reactivity against EIAV antigen whereas 2 (74, A4) were judged negative while the remaining animals (53, 56) produced what can only be described as equivocal reactions. However, when samples were re-tested using the more sensitive Western blotting technique serum from both 53 and 56 along with all horses seropositive in AGID were found to possess antibodies that reacted against multiple EIAV structural proteins (Fig. 3). All initial attempts to amplify EIAV proviral sequences from peripheral blood samples collected from these horses using OIE recommended primers, previously published primers (Quinlivan *et al.*, 2007; Zheng *et al.*, 2000) or unpublished primer sets under investigation in other laboratories were unsuccessful. In contrast, the optimized nested PCR described above, produced a single predominant fragment following amplification of the peripheral blood cell template DNA from all horses (Fig. 3) with the exception of the two seronegative animals (74, A4). Although the length of this fragment was consistent with the amplicon predicted for the

EIAVltr-2F/EIAVltr-2R primer pair (198-203 bp), the identity of each product was confirmed by molecular cloning in a pMD20-T vector (TaKaRa) and nucleotide sequencing. All resultant sequences had significant homology to EIAV when submitted to BLAST searches in GenBank and shared 83.1 to 88.3 % nucleotide identity with the international reference strain Wyoming (AF033820). Taken together these results demonstrate the optimized nested assay has an equivalent sensitivity and specificity to more conventional serological diagnostic techniques in horses that are not experiencing overt clinical signs of EIA. Indeed, while all infected horses were successfully detected using the optimized nested PCR assay only 83% produced clear evidence of seropositivity in the internationally accepted AGID test. In addition, to straightforward diagnosis, analysis of the nucleotide sequence data from each PCR product suggested all 12 horses had been infected with very closely related EIAV strains in that they possessed between 97.0 and 100 % identity with the prototype Miyazaki-A virus. Furthermore, molecular characterization of the PCR amplicon produced by the EIAVltr-2F/EIAVltr-2R primer pair with the Miyazaki-A virus demonstrated this region is extensively conserved in different Asian and North American EIAV strains as seen by the relatively large blocks of identical sequences interspersed with generally very short regions containing nucleotide substitutions, deletions and insertions (Fig. 4).

2.4 Discussion

In the absence of effective vaccines control of EIA is limited to identification of infected equids that at present is entirely reliant on the indirect serological

techniques of AGID and ELISA. Furthermore, in almost every country regulatory actions can only be taken once an equid becomes seropositive in the immunodiffusion assay. Although, AGID has earned a reputation for high specificity it is relatively insensitive and sometimes produces difficult to interpret or even false negative results as seen with serum samples from horses 53 and 56. On the other hand the higher potential sensitivity of ELISA-based tests can result in false positive readings. Although AGID negative/ELISA positive samples can often be resolved by Western blotting this assay is not currently available commercially. Another obvious disadvantage with serological diagnostic assays is that they cannot detect cases of recent exposure. Although PCR-based tests for the detection of EIAV genetic material have the necessary sensitivity and specificity to overcome many of these disadvantages none of the primer pairs developed to date (Cappelli *et al.*, 2011; Nagarajan & Simard, 2001; OIE, 2008; Quinlivan *et al.*, 2007; Zheng *et al.*, 2000) permitted amplification of viral sequences from Toi-Cape area feral horses suggesting the existence of sequence mismatches. In the case of lentiviruses where there is substantial genetic diversity between geographically distinct isolates the possibility of mismatches between primers and their target sequences is seen as a major obstacle to the routine implementation of PCR-based diagnostic tests. However, it has been demonstrated the LTR's contain elements that are highly conserved during persistent infection (Reis *et al.*, 2003). Furthermore, structural and functional studies at both the sequence and the expressed protein levels suggest the viral *tat* gene is also likely to be extensively conserved (Dorn *et al.*, 1990; Stephens *et al.*, 1990; Sticht *et al.*, 1993). These

predictions were confirmed in the nucleotide sequence alignments between different EIAV strains conducted as part of these studies and enabled the design of nested primer sets within 5' regions of the viral genome for the successful amplification of proviral sequences from viral strains that have not been previously characterized at the molecular level. Furthermore, the nested PCR was shown to be potentially more sensitive than the current internationally accepted gold-standard AGID technique (Cullinane *et al.*, 2007). Another factor that could limit the widespread adoption of PCR-based assays against EIA is the generally low tissue and plasma associated viral burdens that exist between clinical episodes and particularly once inapparent carrier status is attained (Harrold *et al.*, 2000). However, results reported by Cappelli *et al.* 2011 (Cappelli *et al.*, 2011) suggest that while plasma associated viral RNA burdens may fall below the detection limits of most nested reverse transcriptase (RT) PCR-based assays in asymptomatic EIAV infected equids, amounts of proviral DNA in peripheral blood mononuclear cells remain at significantly higher and in most cases detectable levels. None of the horses tested as part of this study had overt clinical signs of EIA at the time the samples were collected and therefore it is highly likely that most if not all had already entered the inapparent stage. Amplification of DNA fragments with significant sequence identity to known EIAV strains from peripheral blood samples of all 12 seropositive horses demonstrates the efficacy of the optimized nested PCR assay even in cases where clinical signs are absent. In addition, the results offer some support for the suggestions of Cappelli *et al* 2011 (Cappelli *et al.*, 2011) concerning the relative abundance of proviral DNA. However, more extensive

comparative experiments are required before the use of peripheral blood cell proviral DNA rather than plasma associated viral RNA is recommended as the template of choice for PCR-based methods of EIA diagnosis in inapparent carriers.

A significant advantage of PCR for EIA diagnosis is that in contrast to conventional serological detection techniques it will yield important epidemiological information when combined with nucleotide sequencing. The finding that viruses isolated from all 12 horses were genetically distinguishable from previously published sequences and yet possessed 97-100% nucleotide identity with the original Miyazaki-A isolate reflects the known history of the Toi-cape region and is consistent with a segregated horse population exposed, potentially for long periods of time, to a family of closely related EIAV strains. Another specific advantage of the nested PCR protocol described here is that because primer binding sequences in the 5' LTR are repeated in 3' LTR it is possible to modify the technique and amplify the entire proviral genome for subsequent characterization at the nucleotide level. Based on this modified protocol, we have successfully obtained a whole proviral genome sequence (Miyazaki2011-A) from one of 12 EIA positive horses. Although this will be described fully elsewhere the Miyazaki2011-A isolate differs significantly from all previously published EIAV genomic sequences. Analysis of this whole proviral genomic data has confirmed suspicions that mismatches within primer binding sites are responsible for the failure of previously published PCR assays (Quinlivan *et al.*, 2007; Zheng *et al.*, 2000) including that recommended by OIE (Nagarajan & Simard, 2001) to generate detectable products with any of the Toi-cape area horses.

In summary, a nested PCR assay has been developed to detect EIAV proviral DNA in peripheral blood cells of naturally infected horses using primer sequences located within conserved 5' regions of the viral genome extending from the LTR to the *tat* gene. This assay is rapid, demonstrated equivalent specificity to Western blotting and is potentially more sensitive than AGID, the current international standard for the detection of EIAV infected equids. Furthermore, in contrast to serological diagnostic techniques the nested PCR assay can be used in cases of recent exposure and when combined with nucleotide sequencing delivers additional valuable molecular epidemiological information. The nested PCR therefore provides a powerful new assay that effectively compliments conventional serological tests for the diagnosis and control of EIA.

2.5 Tables and figures

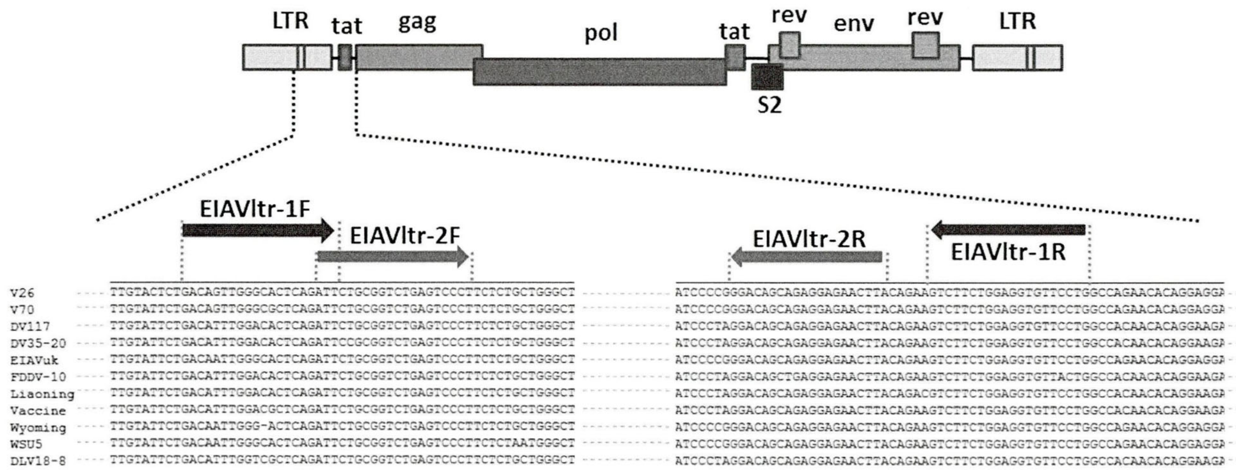


Fig. 1. Primer designing for nested PCR. Sequence alignments were conducted with the LaserGene sequence analysis package. Primers were designed by using the Oligo 6.31 program. International sequences are as follows: V26 (AB008197), V70 (AB008196), DV117 (HM141912), DV35-20 (HM141911), EIAVuk (AF016316), FDDV-10 (GU385360), Liaoning (AF327877), Vaccine Strain (AF327878), Wyoming (AF033820), WSU5 (AF247394), DLV18-8 (HM141923).

Table 1. Primer sets and reaction conditions for nested PCR

	Primer name	Oligonucleotide sequences	Annealing temperature	Amplicon length (bp)
Outer Primers	EIAVltr-1F	5'- GACAGTTGGGCACTCAGATT -3'	52 °C	246
	EIAVltr-1R	5'- CAGGAACACCTCCAGAAGAC -3'		
Inner Primers	EIAVltr-2F	5'- ATTCTGCGGTCTGAGTCCCT -3'	58 °C	198 to 203
	EIAVltr-2R	5'- TAAGTTCTCCTCTGCTGTCC -3'		

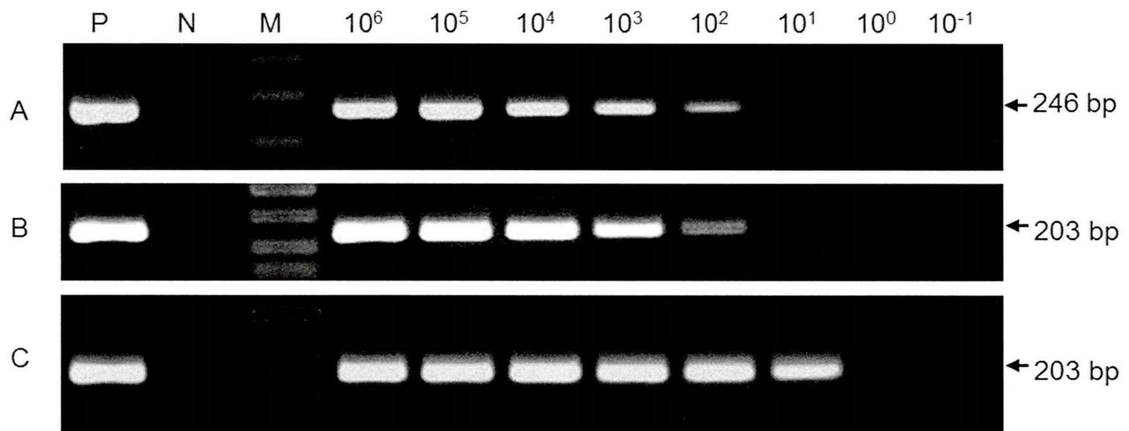


Fig. 2. Sensitivity determination for individual and nested PCR. (A): First stage PCR using EIAVltr-1F and EIAVltr-1R primers. (B): Second stage PCR with primer pair EIAVltr-2F and EIAVltr-2R. (C): Nested PCR. Positive control consisting of plasmid pMD-A246 at 10^7 copies/μl (Lane P) with sterilized distilled water as a negative control (Lane N); DNA base-pair size marker (Lane M); templates comprised 10-fold serial dilutions ranging from 10^6 to 10^{-1} DNA copies per reaction of plasmid pMD-A246 (Lanes 10^6-10^{-1}). PCR products were electrophoresed in 1.5% agarose gels and visualized using 1×GelRed (BIOTIUM, USA).

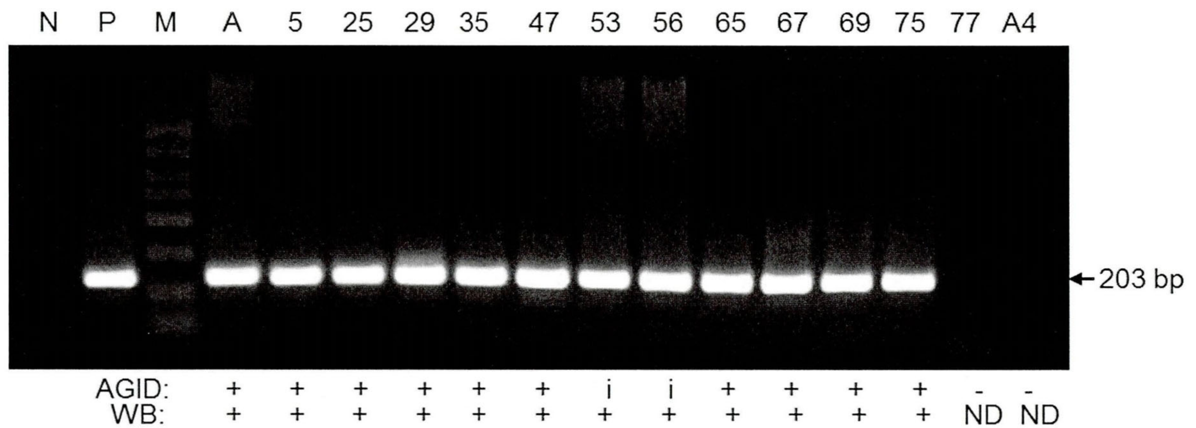


Fig. 3. Detection of EIAV proviral DNA in feral horses. Horse peripheral blood cell DNA was used as a template in an optimized EIAV-nested PCR assay. Amplified products were resolved using 1.5% agarose gels and visualized with 1×GelRed (BIOTIUM, USA). In addition to samples from feral horses (designated A,5,25,29,35,47,53,65,67,69,75,77,A4) a positive control consisting of plasmid pMD-A246 at 10^5 copies/µl (Lane P) and a negative control of sterilized distilled water (Lane N) are also shown. Lane M is a DNA size marker. Serum samples from each feral horse were also tested for the presence of EIAV specific antibodies by both AGID and Western blotting (WB) Results of these serological assays are designated positive (+), negative (-) or inconclusive (i). Not Done (ND).

```
Miyazaki2011-A 1   ATTCTGCGGTCTGAGTCCCTTCTCTGCTGGGCTAAA---TAGCCTTGACAATAAATATATAATTCTCTGTTCAGTCCCTGTGTTCAACTTGTTTTGCTTACAGCAACCCA 107
DLV18-8        1   ATTCTGCGGTCTGAGTCCCTTCTCTGCTGGGCTAAACTCTAGCCTTGGTAATAAATATA--ATTCTCTGCTAAGTCCCTGTTCTTAGTTTGTCTTGTTTTCAAGA-TCTA 107
DV117          1   ATTCTGCGGTCTGAGTCCCTTCTCTGCTGGGCTAGACTCTAGCCTTGGTAATAAATATA--ATTCTCTGCTAAGTCCCTGTTCTTAGTTTGTCTTGTTTTCAAGA-TCTA 107
DV35-20        1   ATTCCGCGGTCTGAGTCCCTTCTCTGCTGGGCTAGACTCTAGCCTTGGTAATAAATATA--ATTCTCTGCTAAGTCCCTGTTCTTAGTTTGTCTTGTTTTCAAGA-TCTA 107
EIAVuk         1   ATTCTGCGGTCTGAGTCCCTTCTCTGCTGGGCTGAA--AAGGCCTTTGTAATAAATATA--ATTCTCTACTCAGTCCCTGTCTCTAGTTTGTCT--GTTCGA-GA-TCCT 102
FDDV-10        1   ATTCTGCGGTCTGAGTCCCTTCTCTGCTGGGCTAGACTCTAGCCTTGGTAATAAATATA--ATTCTCTGCTAAGTCCCTGTTCTTAGTTTGTCTTGTTTTCAAGA-CCTA 107
Liaoning       1   ATTCTGCGGTCTGAGTCCCTTCTCTGCTGGGCTA-ACTCTAGCCTTGGTAATAAATATA--ATTCTCTGCTCAGTCCCTGTTCTTAGTCTGTCTTGTTTTCAAGG-TCTA 106
V26            1   ATTCTGCGGTCTGAGTCCCTTCTCTGCTGGGCTGAA--AAGGCCTTTGTAATAAATATA--ATTCTCTACTCAGTCCCTGTCTCTAGTTTGTCT--GTTCGAAGA-TCCT 102
V70            1   ATTCTGCGGTCTGAGTCCCTTCTCTGCTGGGCTGAA--AAGGCCTTTGTAATAAATATA--ATTCTCTACTCAGTCCCTGTCTCTAGTTTGTCT--GTTCGA-GA-TCCT 102
Vaccine        1   ATTCTGCGGTCTGAGTCCCTTCTCTGCTGGGCTAGACTCTAGCCTTGGTAATAAATATA--ATTCTCTGCTGGGTCCCTGTTCTTAGTTTGTCTTGTTTTCAAGA-TCTA 107
WSU5           1   ATTCTGCGGTCTGAGTCCCTTCTCTAATGGGCTGAA--AAGGCCTTTGTAATAAATATA--ATTCTCTACTCAGTCCCTGTCTCTAGTTTGTCT--GTTCGA-GA-TCCT 102
Wyoming        1   ATTCTGCGGTCTGAGTCCCTTCTCTGCTGGGCTGAA--AAGGCCTTTGTAATAAATATA--ATTCTCTACTCAGTCCCTGTCTCTAGTTTGTCT--GTTCGA-GA-TCCT 102

Miyazaki2011-A 108 ACAGTTGGCGCCCGAACAGGGACCTGAGAGGGGCGCAGACCCTACCTGCTGAACCTGGCTGATTGTAGAATCCCTAGGACAGCAGAGGAGAACTTA 203
DLV18-8        108 ACAGCTGGCGCCCGAACAGGGACCTGAG---GGCGCAGACCCTGCCTGCTGAACCTGGCTGATCATAGGATCCCTAGGACAGCAGAGGAGAACTTA 200
DV117          108 ACAGTTGGCGCCCGAACAGGGACCTGAG---GGCGCAGACCCTGCCTGCTGAACCTGGCTGATCGTAGAATCCCTAGGACAGCAGAGGAGAACTTA 200
DV35-20        108 ACAGTTGGCGCCCGAACAGGGACCTGAG---GGCGCAGACCCTGCCTGCTGAACCTGGCTGATCGTAGGATCCCTAGGACAGCAGAGGAGAACTTA 200
EIAVuk         103 ACAGTTGGCGCCCGAACAGGGACCTGAGAGGGGCGCAGACCCTACCTGTTGAACCTGGCTGATCGTAGGATCCCCGGACAGCAGAGGAGAACTTA 198
FDDV-10        108 ACAGCTGGCGCCCGAACAGGGACCTGAG---GGCGCAGACCCTGCCTGCTGAACCTGGCTGATCATAGGATCCCTAGGACAGCTGAGGAGAACTTA 200
Liaoning       107 ACAGTTGGCGCCCGAACAGGGACCAGAG---GGCGCAGACCCTGCCTGCTGAACCTGGCTGATCGTAGGATCCCTAGGACAGCAGAGGAGAACTTA 199
V26            103 ACAG-TGGCGCCCGAACAGGGACCTGAGAGGGGCGCAGACCCTACCTGTTGAACCTGGCTGATCGTAGGATCCCCGGACAGCAGAGGAGAACTTA 198
V70            103 ACAGTTGGCGCCCGAACAGGGACCTGAGAGGGGCGCAGACCCTACCTGTTGAACCTGGCTGATCGTAGGATCCCCGGACAGCAGAGGAGAACTTA 198
Vaccine        108 ACAGCTGGCGCCCGAACAGGGACCTGAG---GGCGCAGACCCTGCCTGCTGAACCTGGCTGATCATAGGATCCCTAGGACAGCAGAGGAGAACTTA 200
WSU5           103 ACAGCTGGCGCCCGAACAGGGACCTGAGAGGGGCGCAGACCCTACCTGTTGAACCTGGCTGATCGTAGGATCCCCGGACAGCAGAGGAGAACTTA 198
Wyoming        103 ACAGTTGGCGCCCGAACAGGGACCTGAGAGGGGCGCAGACCCTACCTGTTGAACCTGGCTGATCGTAGGATCCCCGGACAGCAGAGGAGAACTTA 198
```

Fig. 4. Alignment between Miyazaki-A and published EIAV nucleotide sequences within the predicted amplicon produced by the EIAVltr-2F/EIAVltr-2R PCR primer pair. Alignment was performed using GENETYX software (Genetyx Corp., Tokyo, Japan). GenBank accession numbers for published EIAV sequences are listed in Material and methods. Identical residues are boxed with nucleotide deletions represented by (-).

3. Identification of a novel EIAV field strain

Complete genomic analysis of a novel EIAV field strain isolated from feral horses in Southern Japan

3.1 Introduction

Equine infectious anemia (EIA) is a persistent lentiviral disease restricted to members of the *Equidae* (horses, zebras, 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 anemia, jaundice, tachypnea, petechial hemorrhages 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 (Craigo *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).

Equine infectious anemia 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 open reading frames (ORFs). These encode the *tat* and *rev* proteins that are present in all lentiviruses and a protein designated as S2 (Leroux *et al.*, 2004). Therefore, EIAV is unique among replication competent lentiviruses in that it does not encode an equivalent of *vif*. 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 sequence 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; Sa Filho *et al.*, 2005; 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 $\text{EIAV}_{\text{Wyoming}}$ (AF033820, isolated in North America) (Petropoulos, 1997) and $\text{EIAV}_{\text{Liaoning}}$ (AF327877, isolated in China) (Tu *et al.*, 2007). All other reported complete EIAV proviral sequences are laboratory-generated derivatives of these

two strains.

In 2011, twelve 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). The Misaki horses were first documented more than 300 years ago and is one of eight breeds considered native to Japan. These feral horses are located in the Toi-Cape area, Miyazaki, in southern Japan. In the present study, the whole proviral genome (Miyazaki2011-A) was amplified, sequenced and subsequently molecularly cloned directly from peripheral blood of one naturally infected Misaki horse. Selected regions of the EIAV genome were also isolated from the other seropositive Misaki horses. Subsequent analysis of the nucleotide sequences demonstrated the Miyazaki2011-A provirus possessed significant differences throughout its entire length from both $EIAV_{Wyoming}$ and $EIAV_{Liaoning}$. Furthermore its *gag* and *env* genes differed considerably from those recently reported for novel field strains isolated in Europe and North America (Cappelli *et al.*, 2011; Craigo *et al.*, 2009). Phylogenetic studies suggested that while all viruses infecting all twelve Misaki horses were closely related, showing extensive nucleotide sequence identity, they comprise a distinct monophyletic group compared to other previously characterized EIAV strains.

3.2 Materials and methods

3.2.1 Diagnosis and histopathology

Twelve feral Misaki horses were found to be EIA seropositive using either the

agar gel immunodiffusion (AGID) (Coggins *et al.*, 1972) or the immunoblot test (Cook *et al.*, 2005). Following diagnosis all horses were euthanized and submitted for post mortem examination. Samples of liver, spleen, lung, heart and kidney were collected for detection histopathological lesions by H&E staining.

3.2.2 Sample collection and DNA extraction

Peripheral blood was collected from all EIA seropositive horses followed by extraction and purification (starting from 200 μl peripheral blood) of cellular DNA using a DNeasy blood and tissue kit (Qiagen, Maryland, MD) according to the manufacturer's instructions.

3.2.3 EIAV diagnostic PCR

Serological diagnosis was confirmed in Misaki horses by nested PCR assay which was established in our lab (Dong *et al.*). Briefly, primer pairs EIAVltr-1F (5'-GACAGTTGGGCACTCAGATT-3')/EIAVltr-1R (5'-CAGGAACACCTCCAGAAGAC -3') plus EIAVltr-2F (5'-ATTCTGCGGTCTGAGTCCCT -3')/EIAVltr-2R (5'-TAAGTTCTCCTCTGCTGTCC-3') were used to perform the first and second stage reactions respectively. All reactions were conducted using Takara Ex Taq (Takara, Shiga, Japan) according to the manufacturer's instructions. Conditions for the first round PCR consisted of an initial denaturation step at 94°C for 5 min followed by 35 cycles of 94°C for 30 s, 52°C for 30 s, 72°C for 40 s, and a final extension of 72°C for 10 min. For the second stage 2.5 μl of the first round PCR

product was used as a template with cycle conditions consisting of initial denaturation at 94°C for 5 min followed by 35 cycles of 94°C for 30 s, 58°C for 30 s, 72°C for 40 s, a final extension at 72°C for 10 min. All PCR products were sequenced using EIAVltr2F and EIAVltr2R as primers.

3.2.4 Complete EIAV proviral genome amplification

The full length EIAV proviral genome was amplified using the primer pair EIAVcom1F (5'-CGCAGACCCTACCTGCTGAACCTGGCTGAT-3') and EIAVcom1R (5'-ATTTAGCCCAGCAGAGAAGGGACTCAGACC-3'). These primers were designed on the basis of results produced by sequencing the 910 bp EIAV genomic fragments described above. Long range PCR amplification was performed using a LA Taq kit (Takara, Shiga, Japan) according to protocols recommended by the manufacturer. The cycle conditions consisted of an initial denaturation at 94°C for 1 min followed by 30 cycles of 98°C for 10 s, 68°C for 15 min, and a final extension of 72°C for 10 min. After electrophoresis, the expected 8.2 kbp product was purified using a QIAquick gel extraction kit (Qiagen, Hilden, Germany).

3.2.5 Direct nucleotide sequencing

A primer walking strategy was employed to sequence the purified products resulting from long-range PCR. Both forward (F) and reverse (R) primers were used to generate a complete proviral genomic sequence and primer sequences were presented in Table S1 (Additional file 3). The genome was sequenced at least three

times using a Bigdye Terminator v3.1 cycle sequencing kit (Applied Biosystems, Foster City, CA) in conjunction with an ABI Prism 3130 genetic analyzer (Applied Biosystems, Foster City, CA). To reconstitute the full-length genomic sequence, overlapping sequences were assembled by Seqman DNAStar software (Lasergene; DNAStar, Inc., Madison, WI).

3.2.6 Cloning and sequencing

To confirm results produced by direct PCR product sequencing, the full-length EIAV proviral genome was PCR amplified to generate 6 fragments with specific primers presented in Table S1 (Additional file 3). Conditions for PCR consisted of an initial denaturation step at 94°C for 5 min followed by 35 cycles of 94°C for 50 s, 55°C for 50 s, 72°C for 90s, and a final extension of 72°C for 10 min. PCR products were purified using the QIAquick gel extraction kit (Qiagen, Hilden, Germany), cloned into the pMD20-T vector (Takara, Shiga, Japan) and sequenced with pMD20-T specific primers (Takara).

3.2.7 Genomic analysis

Sequences were assembled and checked for errors using Seqman DNAStar software (Lasergene; DNAStar, Inc., Madison, WI). For genome organization analysis, putative ORFs and their corresponding amino acids were predicted using the ORF finder tool (http://www.ncbi.nlm.nih.gov/gorf/gorf.html), with similarity analysis performed using BLAST (http://blast.ncbi.nlm.nih.gov) available via the National Center for Biotechnology Information (NCBI). Nucleotide and deduced

amino acid sequences were aligned using the ClustalW multiple sequence alignment program of MEGA version 5.05 (Tamura *et al.*, 2011) and Seqman DNAStar software. Nucleotide sequence alignment results were analyzed with the SimPlot algorithm (SimPlot for Windows Version 3.5.1, Stuart C. Ray) using a sliding window of 200 nucleotides moving in 20-nucleotide steps (Lole *et al.*, 1999). Phylogenetic analysis and phylogenetic trees were constructed by the neighbor-joining method with MEGA version 5.05 (Tamura *et al.*, 2011) with bootstrap values determined over 1,000 iterations. The evolutionary distances were computed using the Maximum Composite Likelihood method. Potential N-linked glycosylation sites were predicted as described previously (Craigo *et al.*, 2009) and analyzed with NetNGlyc 1.0 Server.

3.2.8 A 910 bp EIAV genomic fragment Amplification

Phylogenetic analysis of the EIAV strains infecting all Misaki horses was conducted using a 910 bp fragment incorporating 90 bp of the LTR, all the non-coding region (47bp), exon 1 of *tat* (96 bp) and the first 677 bp of *gag*. This was generated from the 11 feral horses (designated: 5, 25, 29, 35, 47, 53, 56, 65, 67, 69 and 75) that were not used in the production of full length EIAV proviral DNA by semi-nested PCR. Cellular DNA obtained from peripheral blood as described above was used as the template in first and second stage reactions involving primer pairs EIAVltr-1F (5'-GACAGTTGGGCACTCAGATT-3')/ EIAVcan-1R(5'-CCTCTAATAAATCTTGCTGTC-3') and EIAVltr-2F (5'-ATTCTGCGGTCTGAGTCCCT-3')/EIAVcan-1R(5'-CCTCTAATAAATCTT

GCTGTC-3’) respectively. All reactions were conducted using Takara Ex Taq (Takara, Shiga, Japan) with identical cycle conditions for both stages consisting of an initial denaturation step at 94°C for 5 min followed by 35 cycles of 94°C for 30 s, 52°C for 30 s and 72°C for 45 s, with a final extension of 72°C for 10 min. The PCR products were purified, cloned, sequenced and analyzed as described above.

3.2.9 Nucleotide sequence accession numbers

The whole genome sequence of EIAV Miyazaki2011-A and partial sequences of other 11 strians have been deposited into GenBank. The Nucleotide Accession number are from JX003252 to JX003263.

3.3 Results

3.3.1 Diagnosis of EIA in feral Misaki horses

Misaki horses were gathered in order to monitor their overall health. At this time serum was collected and checked for antibodies to EIAV using the AGID test (Coggins *et al.*, 1972) as recommended in the OIE Terrestrial Manual 2008 (OIE, 2008). Ten horses were found to be clearly seropositive in AGID with two more producing equivocal results. Subsequently all twelve horses were shown to be infected with this virus in both the immunoblot test (Western blotting) (Cook *et al.*, 2005) and in a nested PCR assay developed in this laboratory (in Chapter 2). At the time the samples were collected none of the horses displayed signs of disease consistent with clinical EIA. As a result of the positive EIA diagnosis all 12 horses were euthanized followed by a post-mortem examination. No gross macroscopic

lesions were observed and there was no obvious histopathology (liver, spleen, lung, heart and kidney) in any of the horses with the exception of one designated Misaki-A where phagocytic macrophages were observed in the spleen (Fig. 5).

3.3.2 Full-length EIAV proviral genome

Peripheral blood was collected from horse Misaki-A for extraction of the cellular DNA used as a template in long range PCR with the EIAVcom1F and EIAVcom1R primer pair. 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. Analysis to identify putative ORFs and their predicted amino acid sequences were conducted using the ORF finder tool (http://www.ncbi.nlm.nih.gov/gorf/gorf.html), with similarity analysis performed using BLAST (http://blast.ncbi.nlm.nih.gov). The Miyazaki2011-A proviral genome comprises 8208 bp and possesses a prototypical EIAV genomic organization with the *gag*, *pol*, *env* genes bounded at both ends by long terminal repeats (LTR). The proviral genome also contains three additional ORFs encoding *tat*, S2 and *rev* (Fig. 6). 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 amino acid (aa) Gag-precursor poly-protein while the *pol* gene ORF is 3408 bp and extends from nucleotide position 1691 to 5098 in the proviral genome. A probable 2601bp *env* gene ORF occupies genomic nucleotide

positions 5285 to 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 nucleotide positions 5259 to 5492 and as such all but the extreme 5' 26 nucleotides form an alternative open reading frame within the 5' terminal region of *env*. Two additional alternative open reading frames located within *env* at positions 5426 to 5525 and 7232 to 7635 are predicated to encode *rev* (Fig. 6).

3.3.3 Sequence identity among complete EIAV genomes

Multiple nucleotide sequence alignments were performed to compare the full-length Miyazaki2011-A provirus against complete genomic sequences for $EIAV_{Wyoming}$ including strains that were derived from it (V26, V70, WSU5, $EIAV_{UK}$) plus $EIAV_{Liaoning}$ and its derivative strains (The Chinese Vaccine strain, DV35-20, DLV18-8, DV117, FDDV-10). The results demonstrated that despite in some cases very extensive passage both *in vitro* and *in vivo,* viral strains derived from $EIAV_{Wyoming}$ or $EIAV_{Liaoning}$ retained significant similarities to the progenitor isolates with nucleotide sequence identities of 97.5 to 98.9% and 97.1 to 97.9% respectively. In contrast, Miyazaki2011-A shared only 77.2% nucleotide sequence identity with $EIAV_{Wyoming}$, and 78.7% with $EIAV_{Liaoning}$. This is similar to the 80.2% nucleotide sequence identity between $EIAV_{Wyoming}$ and $EIAV_{Liaoning}$ 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.*, 2005; 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 nucleotides was utilized moving in 20-nucleotide steps with percent identity calculated for each window and plotted as a line chart. In contrast to $EIAV_{Wyoming}$ (Fig. 7B) 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. 7A). Furthermore, as suggested from the percent nucleotide sequence identity this analysis also demonstrated the genomes of the North American $EIAV_{Wyoming}$ and Asian $EIAV_{Liaoning}$ are not closely related. Therefore it is highly unlikely that Miyazaki2011-A, $EIAV_{Wyoming}$ or $EIAV_{Liaoning}$-like viruses were derived as a result of recombination events between each other and as 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. 7A), there was significant conservation with the Wyoming and its derived strains (Fig. 7B).

3.3.4 Phylogenetic analysis using LTR, *gag, pol* and *env* genes

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 [11] and the New World [26], information for the LTR and *pol* is restricted to strains that are derived from either $EIAV_{Wyoming}$ or $EIAV_{Liaoning}$. 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. 8A, 8B, 8C and 8D). The analysis also demonstrated that Miyazaki2011-A shared 71.6% to 80.4%, 77.7% to 81.8%, 80.9% to 81.3% and 71.0% to 73.3% nucleotide sequence identity with LTR, *gag, pol* and *env* sequences respectively.

3.3.5 Population phylogenetic analysis

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. A 910-bp segment, located at the 5'-end of the EIAV genome and consisting of 3' LTR sequences (90 bp), the non-coding region (47 bp), *tat* exon 1 (96 bp), and the first 677 bp of the *gag* gene, was amplified from all 11 additional horses. Alignment of the Japanese sequences revealed 95.1% to 98.2% nucleotide

sequence identity suggesting the Misaki horses were infected with related virus strains. In contrast the Misaki horse EIAV isolates possessed only 80.7% to 81.9% and 81.2% to 82.5% nucleotide 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 to reference strains related to either $EIAV_{Wyoming}$ or $EIAV_{Liaoning}$ (Fig. 9). This provides strong supporting evidence that all 12 Misaki horses were infected with EIAV strains derived from a common ancestor.

3.3.6 Characterization of LTR and structural genes

The proviral genome of Miyazaki2011-A contains a 306 bp LTR comprising a Unique, 3' region (U3) a Repeat region (R) and Unique, 5' end (U5) each having lengths of 186 bp, 80 bp and 40 bp respectively. Transcriptional control elements were predicted based on published information for Wyoming-like EIAV strains (Carvalho & Derse, 1993a; Derse *et al.*, 1987). In common with $EIAV_{Wyoming}$ and $EIAV_{Liaoning}$ Miyazaki2011-A possesses a consensus TATA-box, TATATAA (nucleotide position 160-166) along with three consecutive G residues that serve as the initiation site for RNA synthesis and define the R-U5 border (Fig. 10). All three virus strains also contain a consensus poly(A) addition signal (AATAAA), that in Miyazaki2011-A is located between nucleotide position numbers 243 and 248 (Fig. 10). Additional conserved motifs include a methylated DNA-binding protein site (MDBP, also called EF-C or EP), a PEA1/AP-1 site, a PU.1 or *ets*/PEA3 binding

motif and two additional PU.1 or ets binding sites (Hines *et al.*, 2004) (Fig. 10).

The putative *gag* gene of Miyazaki2011-A is 1452 bp in length, and encodes a 484 amino acid predicted *Gag*-precursor polyprotein that is cleaved by the virally encoded protease into four major internal structural proteins of mature virion. These comprise the membrane-interacting matrix (MA/p15) protein, the capsid antigen (CA/p26), the RNA-binding nucleocapsid (NC/p11) protein and p9 (Hussain *et al.*, 1988; Stephens *et al.*, 1986). The predicted amino acid identity between Miyazaki2011-A and Wyoming is 84.9% for p26, 84.8% for p11, 80.6% for p15, and 42.9% for p9 confirming as reported previously that conservation of amino acid sequences encoded by EIAV *gag* conforms to a hierarchy in which p26 > p11 > p15 > p9. Almost all investigations to discover the structural and functional properties of EIAV Gag proteins have been conducted using viral strains derived from $EIAV_{Wyoming}$. Therefore predicted Miyazaki2011-A Gag polyprotein amino acid sequences were compared to those found in $EIAV_{Wyoming}$-like virus strains to determine if previously identified important structural and functional motifs are conserved (Fig. 11). The MA or p15 in all $EIAV_{Wyoming}$-like viruses contains a potential leucine-rich type nuclear export signal (LKKLEKVTV) (Hatanaka *et al.*, 2002). Although the actual function of this motif is unknown the fact that it is likely to be important within the EIAV replication cycle is reinforced by the fact that it is also conserved in Miyazaki2011-A (Fig. 11). Similarly, the CA contains a motif present in many retroviruses termed the Major Homology Region (MHR) that is believed to be essential for virus assembly and budding from the host cell plasma membrane. In $EIAV_{Wyoming}$-like viruses this sequence is

represented by IRQG(AorV)KEPYPEF(VorI)D (Grund *et al.*, 1994). This sequence also appears to be conserved in Miyazaki2011-A where the only changes are the substitution of arginine (A)/ valine (V) by proline (P) at position 5 within the motif and substitution of glutamic acid (E) by aspartic acid (D) at position 11 (Fig. 11). It can be seen from the $EIAV_{Wyoming}$-like strains in Figure 11 that some amino acid variation at position 5 within the MHD is tolerated. Furthermore substitution of A or V by P can be considered conservative in that they are all non-polar amino acids. The E to D substitution at MHD position 11 is also conservative as both are acidic amino acids suggesting that a negative charge is important at this location. The NC proteins of all lentiviruses share a high percentage of basic amino acid residues and one or two zinc-binding motifs. All $EIAV_{Wyoming}$-like viruses contain two of these zinc binding regions (C-X2-C-X4-H-X4-C where C represents cysteine, H represents histidine and X indicates a variable amino acid residue) separated by an usually short five basic residue-rich (R-A-P-K-V) linker (Amodeo *et al.*, 2006). The importance of this dual motif is again confirmed by the fact it is completely conserved in Miyazaki2011-A (Fig. 11). Similarly, the p9 YPDL late domain that has been shown to be essential for budding from the host-cell plasma membrane in $EIAV_{Wyoming}$-like viruses (Chen *et al.*, 2001; Li *et al.*, 2002; Puffer *et al.*, 1997) is also present in Miyazaki2011-A (Fig. 11) despite the fact much of the remainder of the protein shows significant variation in predicted amino acid sequence between these strains.

The putative *env* gene of Miyazaki2011-A is 2601 bp and encodes both the surface unit (SU, gp90) and transmembrane envelope glycoproteins (TM, gp45). It

is predicted that gp90 contains 452 amino acids while 415 amino acid residues comprise gp45. The deduced amino acid sequence of Miayzaki2011-A gp90 shares 60.1%, 59.1% and 56.2% identity with $EIAV_{Wyoming}$, $EIAV_{Liaoning}$ and a recently reported Pennsylvania EIAV ($EIAV_{PA}$, GQ855755) field isolate (Craigo *et al.*, 2009), respectively. Amino acid sequence variation between Miyazaki2011-A and the reference viral strains (Fig. 13) is distributed throughout gp90 and is not confined to the eight hypervariable regions (V1 to V8) that undergo the most genetic substitutions, deletions or insertions during the prolonged course of an infection in individual horses or ponies (Leroux *et al.*, 1997; Zheng *et al.*, 1997). A total of 16 potential N-linked glycosylation sites were observed in gp90 of Miyazaki2011-A (Fig. 13) and this is equivalent to the numbers detected in the other strains analyzed although the actual locations within gp90 were not always conserved (Fig. 13). It has also been suggested that cysteine residue sites are conserved in gp90 between different EIAV strains and that this may contribute via the formation of disulfide bridges to maintaining structural or functional integrity in this otherwise highly variable envelope glycoprotein (Craigo *et al.*, 2009). Support for this hypothesis is provided by the fact that 16 of 18 cysteine residues seen in the reference viruses were conserved in Miyazaki2011-A (Fig. 13). In the case of the two exceptions both involved C-to-Y substitutions (C242Y and C333Y). Meanwhile, comparison of the predicted gp45 amino acid sequence demonstrated that Miyazaki2011-A shared just 62.7%, 68.2% and 65.5% amino acid identities with the $EIAV_{Wyoming}$, $EIAV_{Liaoning}$, $EIAV_{PA}$ strains respectively. However, total numbers of extracellular potential N-linked glycosylation sites are maintained

although the actual locations differ (Fig. 14). Although the predicted transmembrane spanning domain of Miyazaki2011-A showed significant differences from $EIAV_{Wyoming}$ all substitutions with the exception of an I to T transition involved the exchange of one hydrophobic amino acid for another and therefore can probably be regarded as conservative (Fig. 14).

EIAV contains three additional ORFs encoding the small accessary proteins Tat, S2 and Rev. The Tat or trans-activator protein regulates transcription of EIAV proviral DNA. It consists of the amino terminus, core region, basic region and carboxy terminus, but lacks a cysteine-rich region found in other lentiviruses (Carroll *et al.*, 1991). The *S2* gene is located in the *pol-env* intergenic region following the second exon of *tat* and overlaps with the 5'- terminal of the *env* gene while Rev is encoded by two exons comprising alternate reading frames within the *env* gene (Fig. 15). Comparison of the predicted amino acid sequences between Miyazaki2011-A and the Wyoming strain demonstrate that Tat is the most conserved accessary protein, particularly within the core domain, with 81.7% sequence identity while S2 and Rev show relatively high degrees of divergence with just 43.9%, and 55.2% identity respectively (Fig. 15B and 15C).

3.4 Discussion

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 as a National Natural Treasure. In March 2011, 12 horses were found to be seropositive for EIA although none displayed obvious

clinical signs and there were no gross lesions at the time of necropsy. The fact that an infiltration of phagocytic macrophages was only observed in the spleen of one horse is consistent with these animals being either between disease episodes or having entered the inapparent carrier state. Misaki horses are restricted to the Toi-cape area of Southern Japan and since the introduction of a single horse in 1913 have not been exposed to any other equids. Therefore, it is likely that EIAV has been circulating within this rare breed population for a minimum of almost 100 years.

A new PCR-based method has been developed to amplify complete EIAV proviral genomes directly from infected equids without the need to passage the virus either *in vivo* or *in vitro* which is an advantage since passage *in vitro* can result genetic changes that produce attenuation *in vivo* (Carpenter & Chesebro, 1989). This enabled amplification of proviral DNA from one EIAV infected Misaki horse Miyazaki2011-A. Although pathogenicity or even infectivity cannot be inferred or assumed based on sequence information alone this proviral genome contained no obvious defects such as frame-shift mutations that would disrupt any of the recognized open reading frames. Until this time only two complete EIAV proviral sequences ($EIAV_{Wyoming}$ [North America] and $EIAV_{Liaoning}$ [China]) have been described. Comparison of nucleotide sequence alignments demonstrated that Miyazaki2011-A was not closely related to $EIAV_{Wyoming}$, $EIAV_{Liaoning}$ 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 $EIAV_{Wyoming}$-like or $EIAV_{Liaoning}$-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 as 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 $EIAV_{Goshun}$ (Kono *et al.*, 1971), $EIAV_{Tokyo}$ (Tabuchi *et al.*, 1967) and $EIAV_{Tsukiboshi}$ (Tabuchi *et al.*, 1965) that were originally isolated in the 1940s ($EIAV_{Goshun}$) and 1960s ($EIAV_{Tokyo}$ and $EIAV_{Tsukiboshi}$), 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 $EIAV_{Wyoming}$ strongly suggesting these virus isolates have a North American rather than Japanese ancestry.

The comparative nucleotide sequencing studies provide considerable amounts of new information about the extent of variation between different geographically distinct EIAV strains. These studies confirmed that as reported previously [11] variation in *gag* is not uniformly distributed across the entire gene but is lowest in sequences encoding p26 (sequence identity 84.8%) and highest in those for p9 (sequence identity 42.9%). The high degree of conservation in p26 between different isolates is important since this antigen is used in all common EIA serological diagnostic techniques such as AGID [19] and commercial ELISA tests [32, 33]. Almost all studies to determine the structure and function of EIAV proteins has been conducted in virus strains derived from $EIAV_{Wyoming}$. Although there is significant divergence in overall genomic sequence between these North American isolates and Miyazaki2011-A there appears to be general conservation of structural and functional motifs identified as important in the $EIAV_{Wyoming}$-like viruses. This is certainly true in the case of the LTR where all the nuclear transcription-factor binding motifs discovered in the highly pathogenic $EIAV_{Wyoming}$ isolate are found in both the horse virulent $EIAV_{Liaoning}$ strain and in Miyazaki2011-A. These motifs include three PU.1 or *ets* binding sites, a transcription factor expressed in equine macrophages. In the case of the $EIAV_{wyoming}$ LTR all three sites are capable of binding recombinant horse PU.1 although the 5' site can be deleted without significant transcriptional effects *in vitro*, However, the middle site is essential for maintenance of basal transcription

rates in horse macrophage cells while deletion of the 3' site reduces Tat-transactivated expression by 40-fold (Hines *et al.*, 2004). Therefore, it appears that at least within the U3 enhancer region of the LTR, Miyazaki2011-A closely resembles EIAV strains that are highly pathogenic in horses. Similarly, in the case of deduced amino acid sequences for Miyazaki2011-A encoded proteins many of the structural or functional motifs previously identified in $EIAV_{Wyoming}$ are completely conserved whereas others such as the MHD in p26 or the transmembrane spanning domain in gp45, the predicted amino acid substitutions are highly conservative. For example examination of the sequences of the two virus strains demonstrates that positively or negatively charged amino acids are usually replaced by similarly charged residues. This is seen in the nuclear localization signal in Rev that has also been shown to be essential in RNA binding (Lee *et al.*, 2006) where arginine in $EIAV_{Wyoming}$ is replaced by histidine in Miyazaki2011-A. However, in some cases non-conservative substitutions are observed. An example occurs in the critical core domain of Tat that serves as a scaffold for the flexible NH_2- and COOH terminal domains (Willbold *et al.*, 1994) where polar uncharged glutamine in $EIAV_{Wyoming}$ is predicted to be replaced by positively charged histidine in Miyazaki2011-A. Additional examples of non-conservative amino acid substitutions are seen in the COOH terminus of Tat. Furthermore, in some cases such as S2 and Rev there are substantial differences in overall predicted amino acid sequences between $EIAV_{Wyoming}$ and Miyazaki2011-A. Although regions of the EIAV genome encoding Rev have been shown to change rapidly over time in individual infected horses (Baccam *et al.*, 2003; Belshan *et al.*,

2001; Sparks *et al.*, 2008), it is reported that S2 is highly conserved during infection especially the putative nucleoporin (GLFG), SH3 domain binding (PXXP) and nuclear localization sequence (RRKQETKK) motifs (Li *et al.*, 2000; Li *et al.*, 1998). Interestingly, none of these motifs are completely conserved in Miyazaki2011-A or with the exception of the putative nucleoporin motif, in the Liaoning and $EIAV_{PA}$ strains (Fig. 15B). However the substitutions that occur appear to be conservative in that within the putative nucleoporin and SH3 domain binding motifs non-polar amino acids are replaced with ones having similar properties (Fig. 15B). Furthermore, while there is extensive variation within the putative nuclear localization domain, in all cases the carboxyl terminus of S2 comprises a number of basic amino acids suggesting that positive charges in the region are essential to the function of this viral protein (Fig. 15B). While S2 is important for replication *in vivo* its mode of action has not been conclusively determined. However, it is suggested based on the protein encoded by a derivative of the $EIAV_{Wyoming}$ strain, that it may act as an adapter molecule interacting with cellular proteins (OS-9, PSMC3) to stimulate the production of inflammatory cytokines in macrophages (Covaleda *et al.*; Covaleda *et al.*, 2010). This hypothesis would be supported if these interactions were maintained with the S2 from Miyazaki2011-A as this protein is both larger and possesses less than 50% nucleotide sequence identity with the equivalent molecule from $EIAV_{Wyoming}$ derived strains.

Although, it is possible to make reasonable predictions concerning the biological and/or molecular consequences of potential amino acid substitutions

between Miyazaki2011-A and strains such as those derived from $EIAV_{Wyoming}$, determining the actual effects of these changes will require considerable additional experimental effort. Future studies could include the use of full-length molecular clones to study infectivity and pathogenicity. In cases such as p15 and Tat where the secondary and tertiary structures are known for $EIAV_{Wyoming}$ derived viruses (Hatanaka *et al.*, 2002; Willbold *et al.*, 1994) the experimental studies could be duplicated following expression of the equivalent proteins from Miyazaki2011-A, to ascertain the precise relationships between amino acid substitutions and conformation. However, obtaining complete sequence information for a novel EIAV field isolate represents a significant step forward in increasing understanding of the molecular epidemiology of this equine lentivirus.

In conclusion, 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 artifact 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 plots analysis, Miyazaki2011-A is not closely related to $EIAV_{Wyoming}$, $EIAV_{Liaoning}$ or any other

strain reported to date. Moreover, phylogenetic analysis of LTR, *gag*, *pol* and *env* revealed Miyazaki2011-A forms a separate monophyletic group and as such can be classified as a novel EIAV isolate.

3.5 Tables and figures

Table 2. Oligos for whole genome sequencing and cloning

Forward primer, 5'→3'		Reverse primer, 5'→3'	
Amplification and sequencing of whole genome			
EIAVcom1F:	TACCTGCTGAACCTGGCTGAT	EIAVcom1R:	GCAGAGAAGGGACTCAGACC
EIAVcom2F:	GAACATGGTGGGCGATCTCT	EIAVcom2R:	ACAACATGGAATAGTAAGAG
EIAVcom3F:	ACAAGGACCTATCCCTATGA	EIAVcom3R:	ATGTCTCCTTCAGGTAACGA
EIAVcom4F:	ACAAGAAAACTTTCCCTGTG	EIAVcom4R:	TTAAGGCAATATAGGACATC
EIAVcom5F:	GGCACTGAAATATCTAGGGG	EIAVcom5R:	CTAAGACCCCAAGAAACACC
EIAVcom6F:	ACAAGGAATATTATGGGCAGG	EIAVcom6R:	CTTCTGGGGGTATTTGCTTCC
Whole genome cloning			
EIAVclone1F:	ATATGTATTTGTATAAAACTGC	EIAVclone1R:	GTATCTTCTGGTCTTAGATGTC
EIAVclone2F:	CCTAAAGAACCTTATCCAGA	EIAVclone2R:	TTCCTAGATTCATTGCTGCC
EIAVclone3F:	CTTTCAGAGATAGGCATCCA	EIAVclone3R:	ACCCATGTTCTCCTCTTTCA
EIAVclone4F:	GTCACTGGCAAGTCATCAAT	EIAVclone4R:	GCCTACATCCAGGAAATGAA
EIAVclone5F:	GGATTCCTTTGGTGGTATGA	EIAVclone5R:	CCTCGTGCTCTAGCCCATTC
EIAVclone6F:	AATGACAGTACACAATGGGA	EIAVclone6R:	TGTTGGGTTGCTGTAAGC
Amplification for the 910 bp genomic fragment			
EIAVltr-1F:	GACAGTTGGGCACTCAGATT	EIAVcan-1R:	CCTCTAATAAATCTTGCTGTC
EIAVltr-2F:	ATTCTGCGGTCTGAGTCCCT	EIAVcan-1R:	CCTCTAATAAATCTTGCTGTC

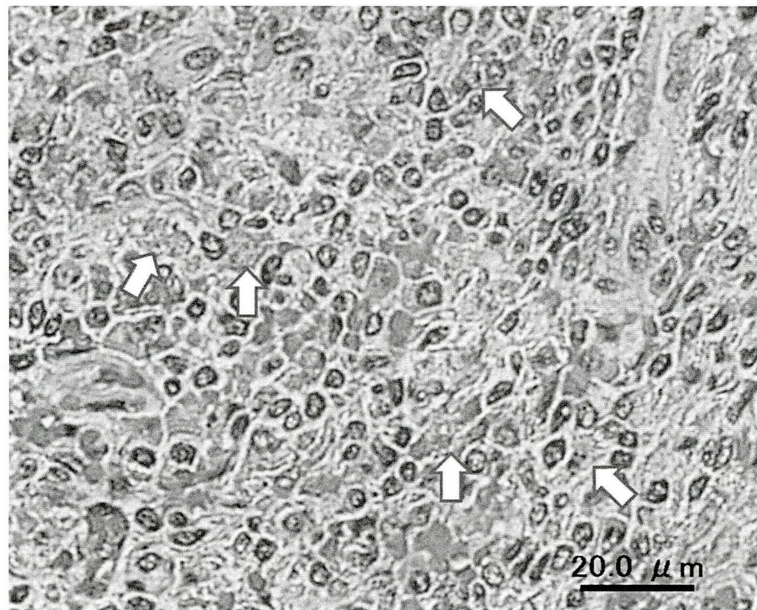


Fig. 5. H&E staining of spleen from horse Misaki-A. Arrows indicate phagocytic macrophages in spleen tissue (400×).

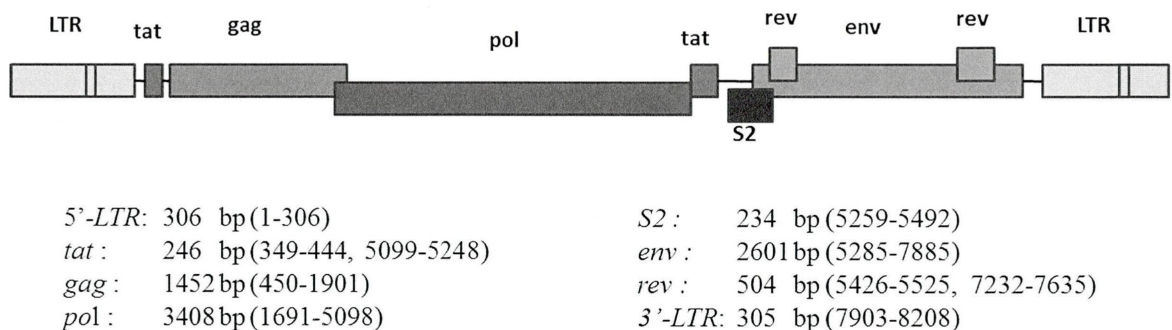
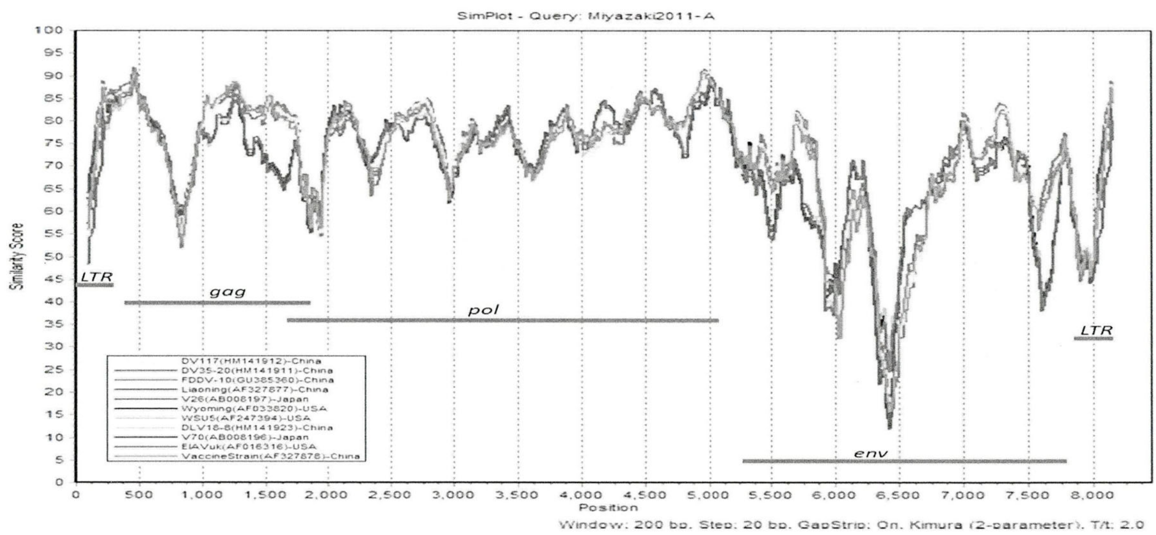


Fig. 6. Schematic representation showing the proviral genome organization of EIAV Miyazaki2011-A. ORFs were predicted using the ORF finder tool (NCBI), while the identity of potential viral genes was determined by BLAST (NCBI) searches against previously published EIAV sequences. The length and location of each ORF are represented as nucleotide position numbers within the Miyazaki2011-A genome.

(A)



(B)

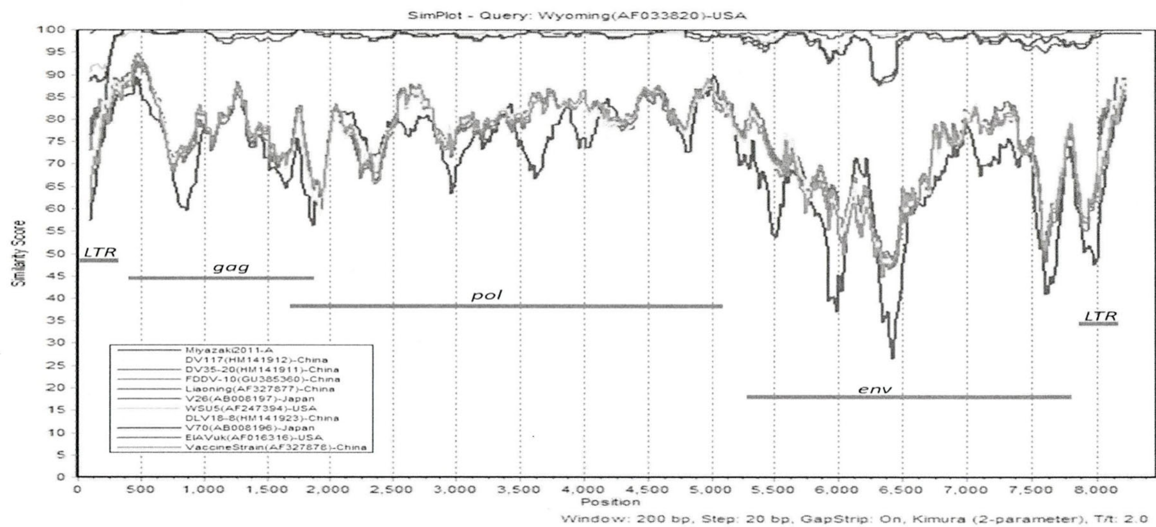
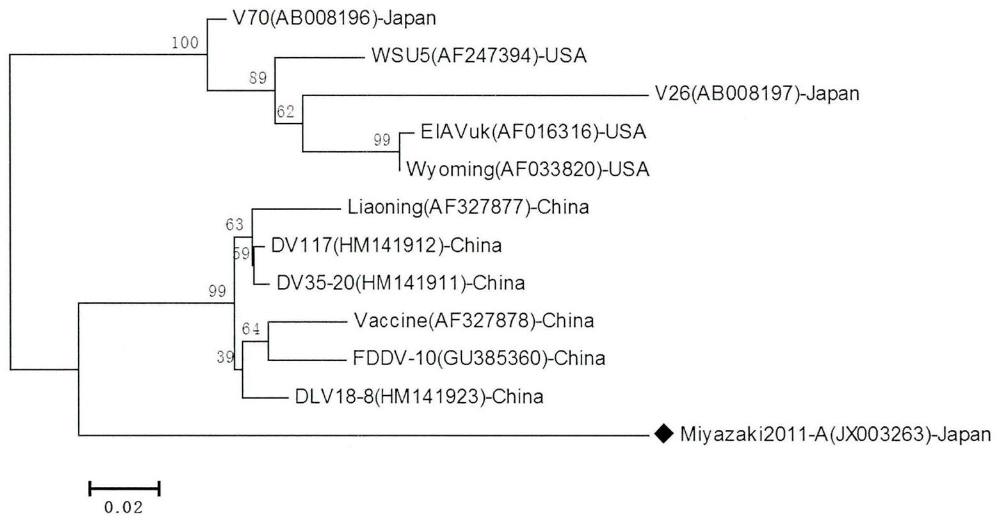
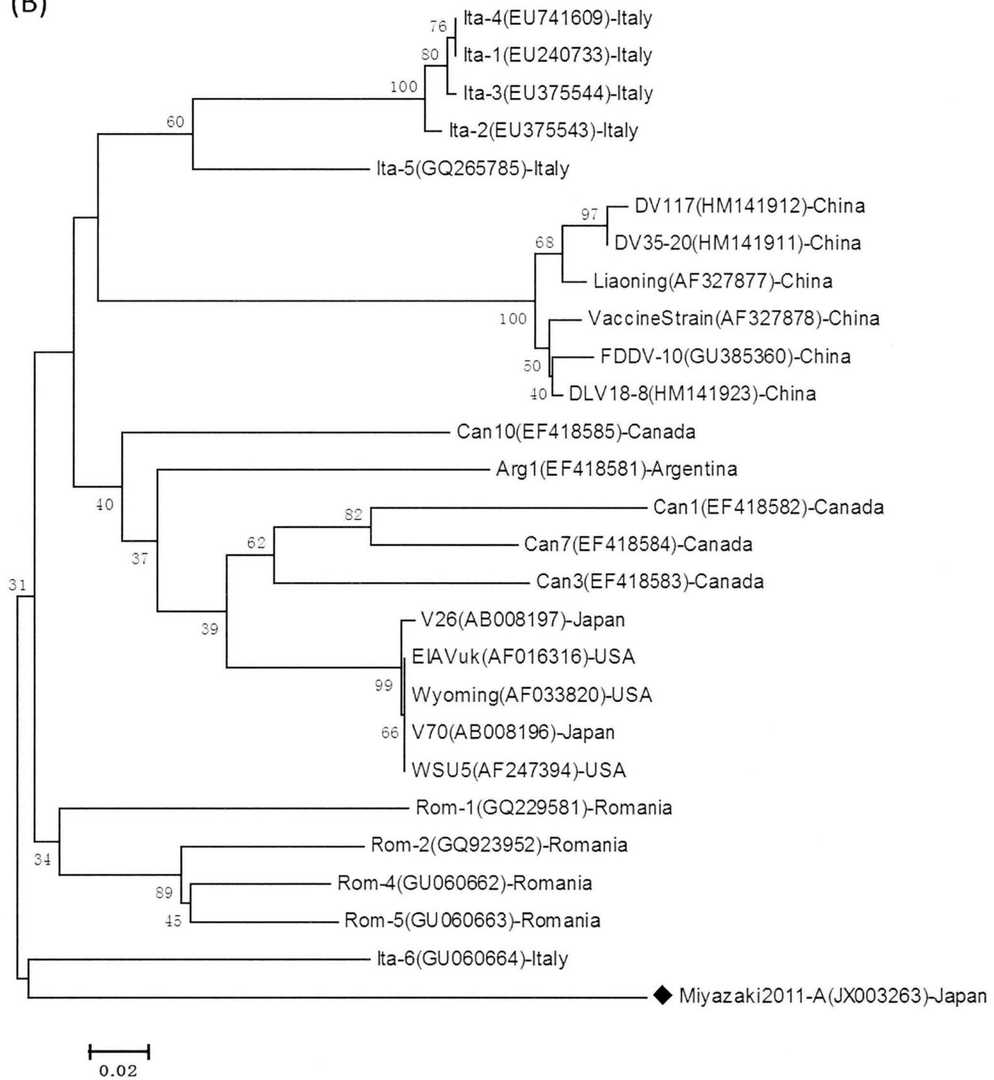
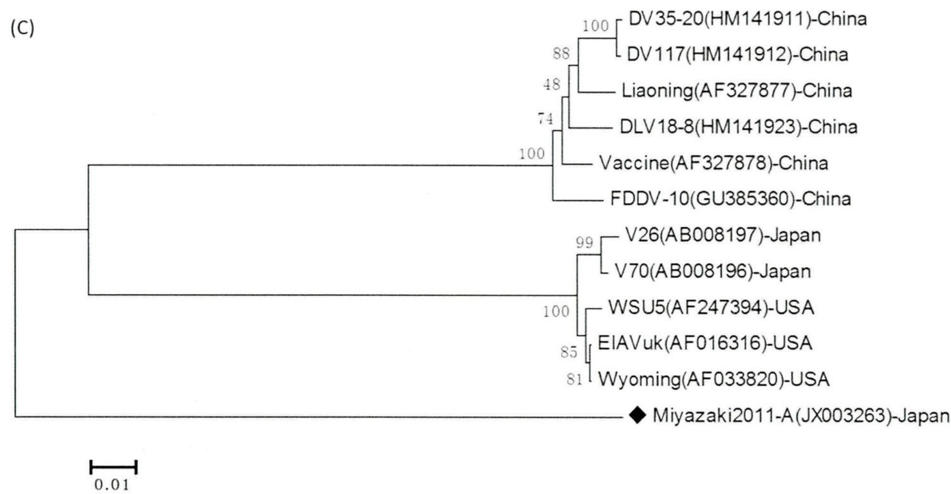


Fig. 7. Complete EIAV genomic plots of nucleotide similarity (generated by SimPlot). (A) Similarity plots of reference EIAV sequences against Miyazaki2011-A .while (B) shows a similar analysis with $\mathrm{EIAV}_{\mathrm{Wyoming}}$ as the query genome. Each curve is a comparison between the genome analyzed and a reference genome. Each point represents the percent 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 open reading frames. Each strain is designated by a different color.


(A)
V70(AB008196)-Japan
WSU5(AF247394)-USA
V26(AB008197)-Japan
EIAVuk(AF016316)-USA
Wyoming(AF033820)-USA
Liaoning(AF327877)-China
DV117(HM141912)-China
DV35-20(HM141911)-China
Vaccine(AF327878)-China
FDDV-10(GU385360)-China
DLV18-8(HM141923)-China
Miyazaki2011-A(JX003263)-Japan
100
89
62
99
63
59
99
64
39
0.02


(B)
Ita-4(EU741609)-Italy
Ita-1(EU240733)-Italy
Ita-3(EU375544)-Italy
Ita-2(EU375543)-Italy
Ita-5(GQ265785)-Italy
DV117(HM141912)-China
DV35-20(HM141911)-China
Liaoning(AF327877)-China
VaccineStrain(AF327878)-China
FDDV-10(GU385360)-China
DLV18-8(HM141923)-China
Can10(EF418585)-Canada
Arg1(EF418581)-Argentina
Can1(EF418582)-Canada
Can7(EF418584)-Canada
Can3(EF418583)-Canada
V26(AB008197)-Japan
EIAVuk(AF016316)-USA
Wyoming(AF033820)-USA
V70(AB008196)-Japan
WSU5(AF247394)-USA
Rom-1(GQ229581)-Romania
Rom-2(GQ923952)-Romania
Rom-4(GU060662)-Romania
Rom-5(GU060663)-Romania
Ita-6(GU060664)-Italy
Miyazaki2011-A(JX003263)-Japan
76
80
100
60
97
68
100
50
40
40
37
82
62
39
99
66
31
34
89
45
0.02



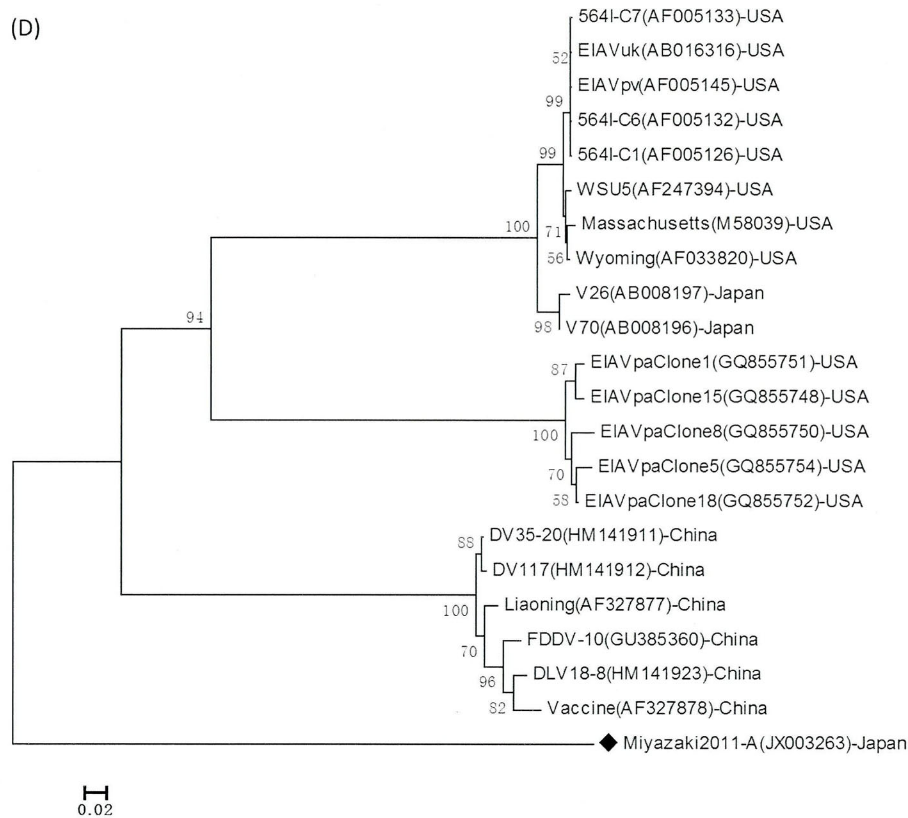


Fig. 8. Phylogenetic relationships between Miyazaki2011-A and available EIAV sequences within the LTR, *gag*, *pol* and *env* genes. (A) LTR. (B) *gag* gene. (C) *pol* gene. (D) *env* gene. Trees were constructed based on LTR, gag, pol and *env* genes from Miyazaki2011-A, Wyoming, Liaoning (including derivative strains) along with available EIAV sequences from North American and European isolates (*gag* and *env*). Sequences were aligned using the ClustalW multiple sequence alignment program of MEGA version 5.05 (Tamura *et al.*, 2011). This software was also used for phylogenetic analysis with phylogenetic trees constructed by the Neighbor-Joining method in conjunction with the Maximum Composite Likelihood algorithm for computation of evolutionary distance. Bootstrap values were determined over 1,000 iterations. EIAV strain information including name, GenBank accession number and country of origin is included in the figure.

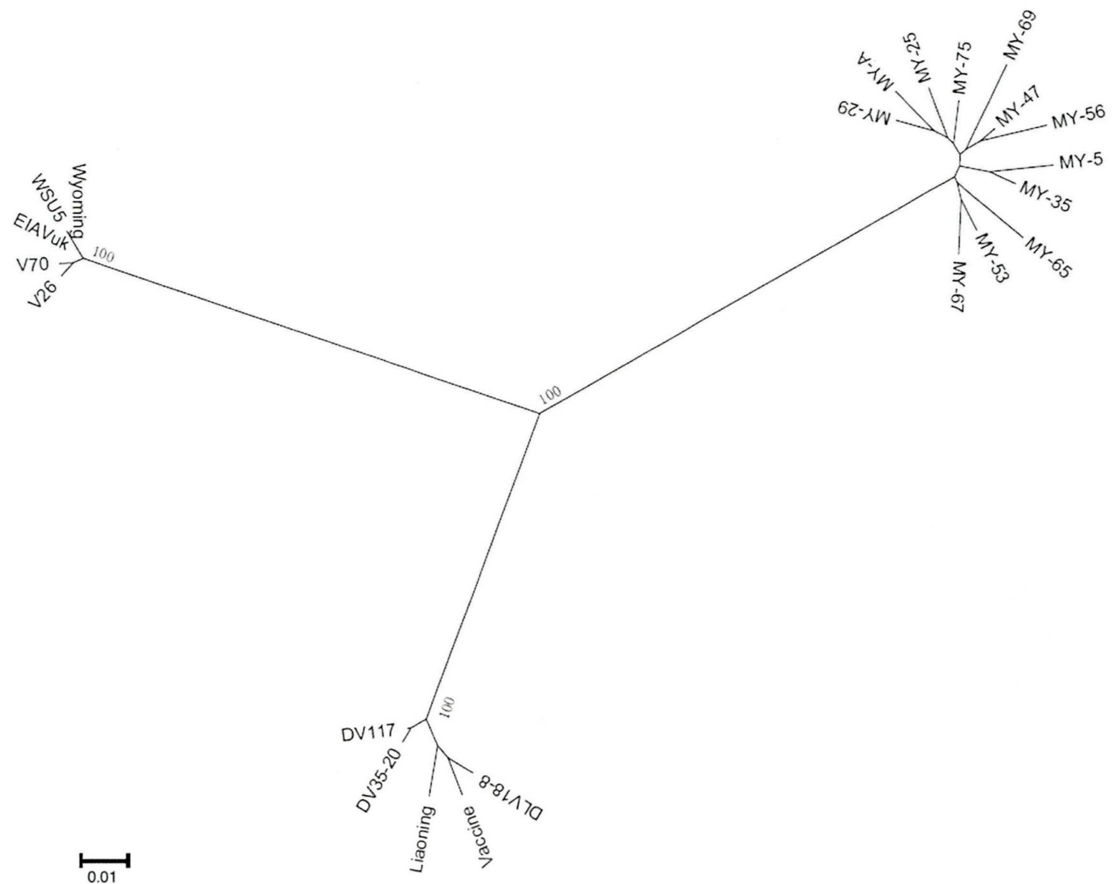


Fig. 9. Phylogenetic analysis involving sequences from all EIAV infected Misaki horses. The analysis was performed based on the nucleotide sequence of a 910 bp of 5'-end genomic fragment amplified from all 12 Misaki horses. The phylogenetic tree was constructed with MEGA version 5.05 (Tamura *et al.*, 2011) software employing the Neighbor-Joining method in conjunction with the Maximum Composite Likelihood algorithm to compute evolutionary distance. Bootstrap values were determined over 1,000 iterations using MEGA version 5.05 (Tamura *et al.*, 2011). EIAV sequences from Misaki horses are designated MY-A (Miyazaki2011-A), -5, -25, -29, -35, -47, -53, -56, -65, -67, -69 and -75. Comparison sequences included: Wyoming (AF033820); EIAVuk (AF016316); V26 (AB008197); V70 (AB008196); WSU5 (AF247394); Liaoning (AF327877); Vaccine (AF327878); DLV18-8 (HM141923); DV35-20 (HM141911); DV117 (HM141912).

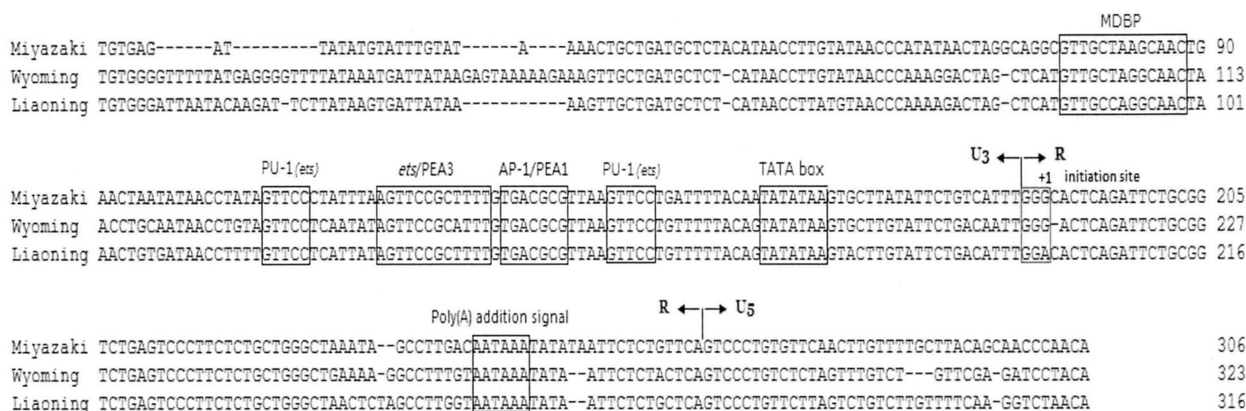


Fig. 10. Comparison of LTR nucleotide sequences. Structural and putative transcriptional control elements were predicted based on previous characterizations of the Wyoming strain LTR (Carvalho & Derse, 1993a; Derse *et al.*, 1987). The U3, R and U5 structural domains are indicated by arrows. Consensus transcriptional control elements are designated using black boxes whereas the initiation site of transcription (GGG) is highlighted in red. Dashes indicate deletions. Miyazaki, Miyazaki2011-A strain (JX003263); Wyoming, Wyoming strain (AF028232); Liaoning, Liaoning strain (AF327877).

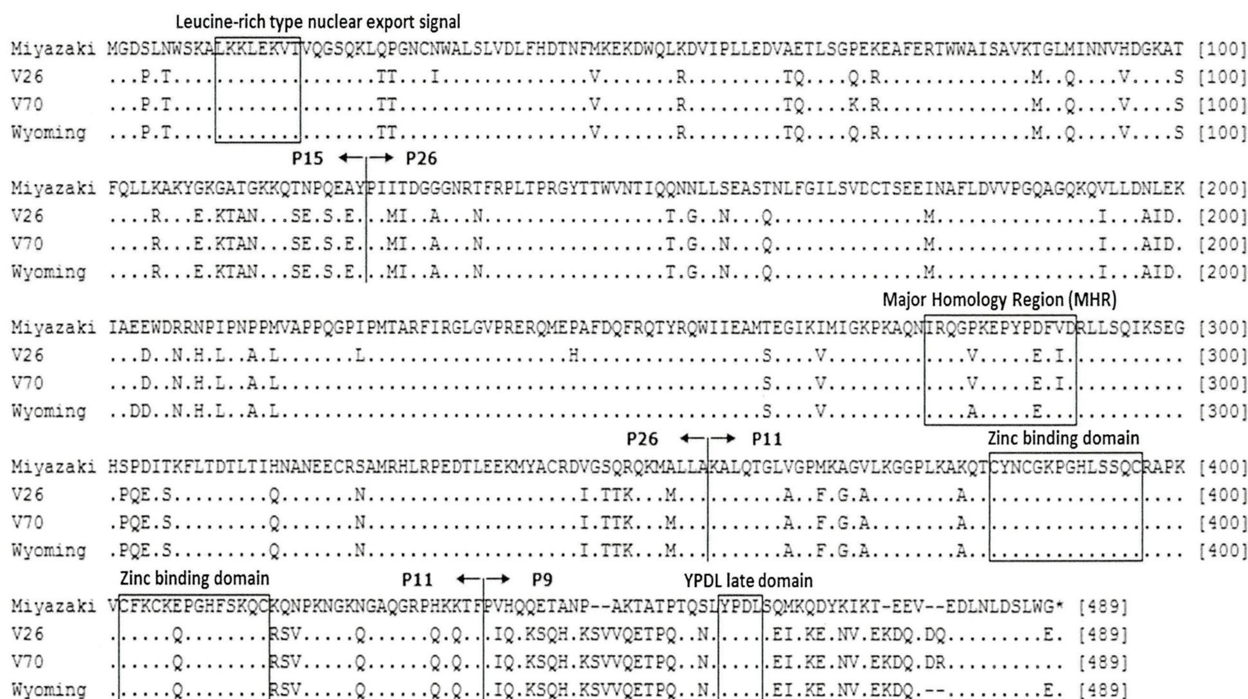


Fig. 11. Comparison of predicted amino acid sequences comprising the Gag polyprotein. Locations of the membrane-interacting matrix (MA) or p15, the capsid (CA) antigen p26, the RNA-binding nucleocapsid (NC) protein p11 and p9 are indicated by arrows. Previously identified functional motifs are indicated within boxes. Dots indicate amino acids identical to that of Miyzaki2011-A strain. Dashes indicate deletions. Miyazaki: Miyazaki2011-A strain. Sequence information of other strains are as follows: V26 (AB008197), Japan; V70 (AB008196), Japan; Wyoming (AF033820), USA.

```
Miyazaki MQTKSKKREEWGSGEAPQENFP--CASAGDCQSCKDSNPDSEFIPGSESNETGLQNKDR-RGR--GSQSGQFVGVTYKLDRRPTTVVIINDTPLNVLLDTGADTSVLTTAHYSKLKYRGKRYQGTGILGVGGNVETFSTPVTIKKKGKQIKTRMLVADIPVTI
Wyoming  ..KC.....AR..R...ET...DTTEESAQQIC.TRDSS..KSV.R..R.KK.I.CQGEGSS.--...P........N.EK...I.L.........................NR.....RK......I...................RH..............
V70      ..KC.....AR..R...ET...DTTEESAQQIC.TRDSS..KSVSR..R.KK.I.CQGEGSS.SR...P........N.EK...I.L.........................NR.....RK......I...................RH..............
WSU5     ..KC.....AR..R...ET...DTTEESAQQIC.TRDSS..KSV.R..R.KK.I.CQGEGSS.--...P........N.EK...I.L.........................NR......K......I...................RH..............
Liaoning --KC.....TR.....KT..--..EGINE.NT.GGKTARDL.SRF..D...V.DQG.GKS.--............N.EK....I.L...................I...NR.....RK......V.............V...................
Vaccine  --KC.....TR.....ET..--..EGVNG.NT.RGETARDL.SRF..D...I.DQG.GKSG--....E.......N.EK....I.L...................I...NR.....RK......V.............V...................
DV117    --KC.....TR.....KT..--..EGVNE.NT.GGETARDL.SRF..D...V.DQG.GKSG--............N.EK....I.L...................I...NR.....RK......V.............V...................

Miyazaki IGRDILQELGAQLVMAQLSTEISPRRIKLKPGMTGPKIPQWPLTKEKLTGAKEIIQRLLSEGKISEASDDNPFNSPIFVIKKRSGKWRLLQDLRELNKAVQVGTEISRGLPHPGGLIKSKNMTVMDIGDAYFTIPLDPLFRRYTAFTVPSINHQEPDKRYVWN
Wyoming  L......D...K..L....K..KF.K.E..E.TM..............E....V..............N..Y..........................T..................C.H...L.............E..P.....I...............K
V70      L......D...K..L....K..KF.K.E..E.TM..............E....V..............N..Y..........................T..................C.H...L.............E..P.....I...Y..........
WSU5     L......D...K..L....K..KF.K.E..E.TM..............E....V..............N..Y..........S...............T..................C.H...L.............E..P.....I..............
Liaoning L............I.....K..T..E....T.TV...V..........L....VKK..D............Y.........K...............V..................CNH...L.............K..Q...................I..
Vaccine  L............I.....K..T..E....T.TV...V..........L....VKK..D............Y.........K...............V..................CNH...L.............K..Q.....................
DV117    L............I.....K..T..E....T.TV...V..........L....VKK..D.D..........Y.........K...............V..................CNH...L.............K..Q.....................

Miyazaki CLPQGFILSPYIYQKTLQEILQPFRDRHPEVQLYQYMDDLFIGSNESRKQHNELVKELRMILLEKGFETPEEKLQDEAPYGWLGYQLFPNNWKMPTVQLELAKEPTLNDVQKLMGNITWISSGIPGLTVKHIAATTKGCLDLNEKVNWTAEAQQELEQNNEKV
Wyoming  .....V.................E.Y.............M...G.K...K..II...A...........DD...EVP..S......C.E...VQKM..DMV.N................M.....................E..Q..I..E...K...E....I
V70      .....V.................E.Y.............V...G.K...K..II...A...........DD...EVP..S......C.E...VQKM..DMV.N................M...V.................E..Q..I..E...K...E..K.I
WSU5     .....V.................E.Y.............V...G.K...K..II...A...........DD...EVP..S......C.E...VQKM..DMV.N................M...V.................E..Q..I..E...K...E....I
Liaoning .....V...........N...A..E...D...............G.KR..K...E...A...........D...E....N......S.G...VQKM....V................M...V.....Q............Q..V..E...K...E..K.I
Vaccine  .....V...........D...A..E...D...............K.KR..K...E...AT..........D...E....N......S.G...VQKM....V................M...V.....Q............Q..V..E...K...E..K.I
DV117    .....V...........D...A..E...D.................KR..K...E...A...........D...E....N......S.G...VQKM....V................M...V.....Q............Q..V..E...K...E..K.I

Miyazaki QKAQGLQYYNPDEEVICEINLTKNCEATYIVKQSQGILWAGKKIMKANKGRSTVKNLMLLLQHVATESITRIGICPTFKVPFTKEQVSWEMKKGWYYSWLPEIIYSSQVVHDEWKLKLVEEPTSGITIYTDGGKQNGRGVAAYVTSTGKTKQKQLGPVTHQRA
Wyoming  KN.........E..ML..VEI...Y....VI....................W......................V.K.............M...Q...........V.TH.....D.RM......................E.I......N.R....R.......V.
V70      KN.........E..ML..VEI...Y....VI....................W......................V.K.............M...Q...........V.TH.....D.RM......................E.I......N......R.......V.
WSU5     KN.........E..ML..VEI...Y....VI....................W......................V.K.............M...Q...........V.TH.....D..M......................E.I......N.R....R.......V.
Liaoning .E.........E.......EI...Y.....I...............R....W.AA...............V...T..K..........K...E.........DMV..H.....D.......Q..................EE.........N......R.......T.
Vaccine  .E.........E.......EI...Y.....I...............R....W.AA...............V...T..K........R.K...E.........DM...H.....D.R.....Q....V..........DEE.........N......R.......T.
DV117    .E.........E.......EI...Y.....I...............R....W.AA...............V...T..K..........K...E.........DM...H.....D.......Q..................EE.........N......R.......T.

Miyazaki EMIAIQMALEDYNDKQLNIVTDSYYCWKNITEGLGLEGPDSPWWPIIQNIYNKETIYFAWVPGHKGIYGNQLADEATKITEEIMLAYQGTQIRNKRKEDAGFDLCSPYDITLHVSETKIIPTDVKIQVPPQCFGWVTGKSSMAKQGLLVNGGIIDEGYTGEIQ
Wyoming  .RM........TR...V.........................Q...........RE..IV......................A..K.............KE..D........V....MIP..D................NS.....................I................
V70      .RM........TR...I.........................Q...........RE..IV......................A..K.............KE..D........V....MIP..D................NS.....................I................
WSU5     .RM........TR...V.........................Q...........RE..IV...........C..........A..K.............KE..D........V....MIP..D............R...NS.....................I................
Liaoning .R.........TEETLV..........................................RA..MV..............................D..........KE..D........I....IIP.....V...........HK.....................I................
Vaccine  .R.........TEETLV..........................................RA..MV.........................................KE..D....Y...I....MIP.....V...........HK.....................I................
DV117    .R.........TEETLV..........................................RA..MV.........................................KE..D....Y...I...MIIP.....V...........HK.....................I................

Miyazaki VICTNIGQGIVKLLEGQKFAQLIILQHQSNDTHQWDENKKSERGEHGFGSTGIFWVDNIQDAQDEHENWHTSPKILAKKYGLPLTVAKQITQECPHCTKQGSGPAGCVMRSPNHWQADCTRLENKIIMTFVDSNSGYIHATLLPKENALHTSLAILEWVRIFS
Wyoming  .......KSNI..I...............H..SRQP.....I.Q..DK......V...E...E..................RN.KI.....................................................H.D....L...E............S.....C........A.L..
V70      .......K.NI..I...............H..SRQP.....I.Q..DK......V...E...E..................RN.KI............L........................................H.D....L...E............S.....C........A.L..
WSU5     .......KSNI..I...............H..SRQP.....I.Q..DK......V...E...E..................RN.KI.....................................................H.D....L...E............S.....C........A.L..
Liaoning .......KSNI..R...............R...KQI.....T.Q..DK......V...E...E...................R.........................P.............................H...R.....E............S.....C...........L..
Vaccine  .......KSNM..R...............R...KQI.....T.Q..DK..........E...E...................R.......................................................H...RV....E............S.....C...........L..
DV117    .....V.KSNI..R...............R.K.KQI.....T.Q..DE......V...E...E..G................R.......................................................H...RV....E............S.....C...........L..

Miyazaki PKSLHTDNGTNFTAESVQNLLKFLKVTHTTGIPYHPESQGIIERANRTLKEKIQSHRENTQTLEAALQLALITCNKGRESMGGQTPWEVFVTNQAQTIHEELLLQQAQSSKKFCFYKVPGEHNWKGPTRVLWKGDGAVVVNDEEKGIIAIPLTRTKLLIKPN*
Wyoming  ............V..P.V.......IA..............V..............D.............................I.....V...K.................I...D....................G.....V.............
V70      ............V..P.V.......IA..............V..............D.....................R.......I.....V...K.................I...D....................G.....V.............
WSU5     ............V..P.V.......IA..............V..............D.............................I.....V...K.................I...D....................G.....V.........V...
Liaoning ............V....A.......................V..........K..................................I.....................................................V.........R...
Vaccine  ............V....A.......................V..........K..................................I.....................................................V.........R...
DV117    ............V....A.......................V..........K..................................I.....................................................V.........R...
```

Fig. 12. Comparison of deduced amino acid sequences in EIAV Pol. The deduced amino acid sequences of Miyazaki2011-A along with reference strains were aligned by using the MEGA version 5.05 (Tamura *et al.*, 2011) ClustalW multiple sequence alignment program. Residues that differ from Miyazaki2011-A are indicated by their single letter amino acid code designations. Regions of amino acid identity are indicated by dots. Miyazaki, Miyazaki2011-A (JX003263); Wyoming (AF033820); V70 (AB008196); WSU5 (AF247394); Liaoning (AF327877); DV117 (HM141912); Vaccine, Chinese vaccine strain (AF327878).

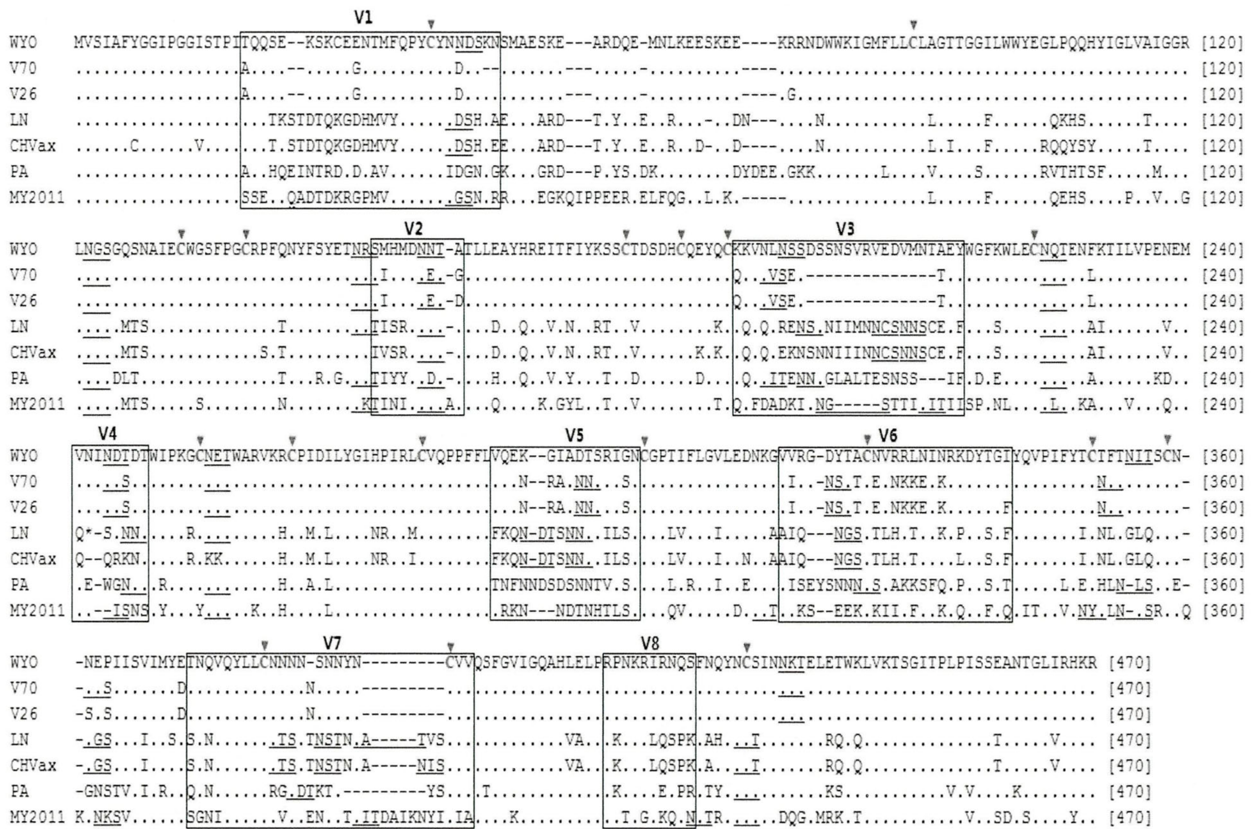


Fig. 13. Comparison of deduced amino acid sequences in EIAV gp90. The deduced amino acid sequences of Miyazaki2011-A along with reference strains were aligned relative to the Wyoming strain using MEGA version 5.05 (Tamura *et al.*, 2011) ClustalW multiple sequence alignment program. Only the residues that differ from Wyoming are indicated by their single letter amino acid code designations. Reported variable regions V1 to V8 for the gp90 sequence are outlined using boxes. Regions of amino acid identity are indicated by dots. Potential N- linked glycosylation sites are underlined in blue while cysteine residues are indicated by red arrows. WYO, Wyoming(AF033820); V26 (AB008197); V70 (AB008196); LN, Liaoning (AF327877); CHVax, Chinese vaccine strain (AF327878); PA, $EIAV_{PA}$ (GQ855755); MY2011, Miyazaki2011-A (JX003263).

```
WYO    DFGISAIVAAIVAATAIAASATMSYVALTEVNKIMEVQNHTFEVENSTLNGMDLIERQIKILYAMILQTHADVQLLKERQQVEETFNLIGCIERTHVFCHTGHPWNMSWGHLNESTQWDDWVSKMEDLNQEILTTLHGARNNLAQSMITF
V70    ......V...................................K...N...............................................................................N...D......V............
V26    ...M..V.....T.................................N...............................................................................N...D......V..S.........
LN     .......I.................I.......LDS..........N.I..LE.V.E..H....V.........S...Q.KI............S.T.........E...Q..D........D...N..HD......T.....E......
CHVax  .......I.................I.......LDS..........N.I.SIE.T.E..H....V.............Q.KI............S.T.........E...Q..D........D...N..HD......T.....E......
PA     .......I............G..A.I....A..LAD..........N.....EI..K..H....V.........I...K.HI...........Q..T.........D...Q..DT....E...R..SY..D......V.......A....
MY2011 .......I.................I...QT.HL.D.T........N.IS..E...K..HM.................Q.KI.....M...A..S.I.........ET..S..D........TR.NG.EH.V.....A.....E..I...

                      Transmembrane Domain
WYO    NTPDSIAQFGKDLWSHIGNWIPGLGASIIKYIVMFLLIYLLLTSSPKILRALWKVTSGAGSSGSRYLKKKFHHKHASREDTWDQAQHNIHLAGVTGGSGDKYYKQKYSRNDWNGESEEYNRRPKSWVKSIEAFGESYISEKTKGEISQPG
V70    .................A....................................................NR...........................N...R..C...........................................
V26    .................A....................................................NR...........................NN..R..C................................K..........
LN     ...........NI....A...............LL..V.V....A.....G.LTTM......A....R.RY..R....G.I.A.V.YHAY..DE.H....SNMR.L...N............Q.N.K.L.KRS..N.NTHEDNMGTMGRL
CHVax  ...........NI....A..............MLL..V.M....A.....G.LTTM......A....R.RY..R....GNI.A.V.YHAY..DE.H....SNMR.L...N............Q.N.K.L.KRS..N.NTHEDNMGTMGRF
PA     .....V.....NI.D..A..........T...............A.......LTTI......A.HF...RY.RR.VWQ..NL..G.YS....D..A.LE.E.KGR..F.SN......GS.KNMR.SKGLMREY.G..T.LR..K.TIH..
MY2011 ....T......NI...VA.........LT..VIL...V.VV...M..F..H.LTTI...R..A.....ETY.RRRVWQ.GH...D.Y.....N..V..EEEFNT.NNF..N...D.KV.DKLQ.H.RRLTKRS..DWHIPGVNEGT.HH.

WYO    AAINEHKNGSGGNNPHQGSLDLEIRSEGGNIYDCCIKAQEGTLAIPCCGFPLWLFWGLVIIVGRIAGYGLRGLAVIIRICIRGLNLIFEIIRKMLDYIGRALNPGTSHVSMPQY-V
V70    .........A.................................V..............L..L....................................................-.
V26    ..V......A................................................L..L....................................................-.
LN     VTTAAE.KNV.-V.......N...Q.............................L...I..L..LL.....I.K..M.LGK...V.ITGL..LC....KM...A...T......D.
CHVax  VTTAAE.KNV.-V.......N...Q.............................L...I..L..LL.....I.K..M.LAK...V.ITRL..LC....KM...A...T......D.
PA     GVT.TK.KVT.-E.......N...QND...............I.......L.......L..L..LI..V....K.LMNVGK..SMLL.V..RV.....K....ASQ........-.
MY2011 TTTGAK.KHI.-G..Q........Q.T................T..............L..M..LL....K...IMLQ.IGK..YGLLNM.KQIF..M.N.FS.PK..I.....I-
```

Fig. 14. Comparison of predicted amino acid sequences in EIAV gp45. The deduced amino acid sequences of Miyazaki2011-A along with reference strains were aligned relative to the Wyoming strain using MEGA version 5.05 (Tamura *et al.*, 2011) ClustalW multiple sequence alignment program. Residues that differ from Wyoming are indicated by their single letter amino acid code designations. Regions of amino acid identity are indicated by dots. Potential N- linked glycosylation sites are underlined in blue while amino acids comprising the transmembrane spanning domain are boxed. WYO, Wyoming(AF033820); V26 (AB008197); V70 (AB008196); LN, Liaoning (AF327877); CHVax, Chinese vaccine strain (AF327878); PA, $EIAV_{PA}$ (GQ855755); MY2011, Miyazaki2011-A (JX003263).

Fig. 15. Comparison of predicted amino acid sequences of EIAV Tat, S2 and Rev. The deduced Tat (a), S2 (b) and Rev (c) amino acid sequences of Miyazaki2011-A along with reference strains were aligned relative to the Wyoming strain using MEGA version 5.05 (Tamura *et al.*, 2011) ClustalW multiple sequence alignment. Residues that differ from Wyoming are indicated by their single letter amino acid code designations while regions of amino acid identity are represented by dots. Putative functional or structural motifs and domains identified in EIAV strains derived from EIAVWyoming are labeled. WYO, Wyoming(AF033820); V26 (AB008197); V70 (AB008196); LN, Liaoning (AF327877); CHVax, Chinese vaccine strain (AF327878); PA, $EIAV_{PA}$ (GQ855755); Miyazaki, Miyazaki2011-A (JX003263).

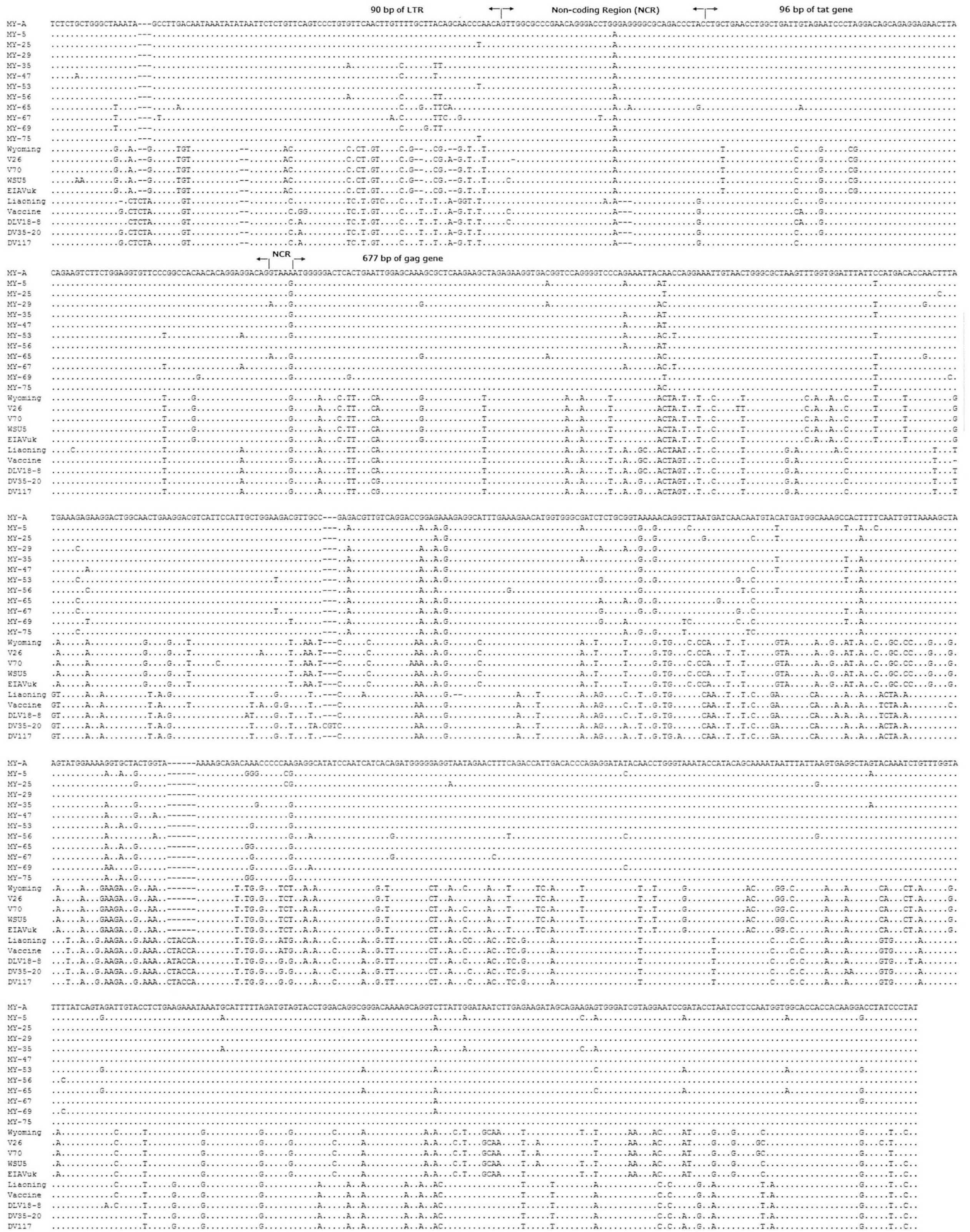


Fig. 16. Nucleotide sequences alignment involving sequences from all EIAV infected Misaki horses. The alignment was performed based on the nucleotide sequence of a 910 bp 5' genomic fragment amplified from all 12 Misaki horses with ClustalW method by using MEGA version 5.05 (Tamura *et al.*, 2011). EIAV sequences from Misaki horses are designated MY-A (Miyazaki2011-A), -5, -25, -29, -35, -47, -53, -56, -65, -67, -69 and -75 (JX003252 to JX003263).

4. Conclusion

Equine infectious anemia (EIA) is a persistent retroviral-based disease with a worldwide distribution. Since it was first described in France in 1843, EIA has posed a major challenge and caused significant losses to the equine industry worldwide. Equine infectious anemia 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. Although EIA was described more than 150 years ago, complete genomic sequences have only been obtained from two strains of EIAV strain, $EIAV_{Wyoming}$ (North America) and $EIAV_{Liaoning}$ (China), with a passage history restricted to horses in order to retain the virulent phenotype of the original field isolate. All other reported complete EIAV proviral sequences are laboratory-generated derivatives of these two strains. Although novel field EIAV isolates have been reported based on the nucleotide sequences of *gag* and *env* genes in USA and Europe, the whole proviral genome sequence 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.

In 2011, twelve Misaki horses were found to be seropositive for EIA in both the agar gel immunodiffusion (AGID or Coggins) and immunoblot tests (Western blotting). The Misaki horses were first documented more than 300 years ago and is one of eight breeds considered native to Japan. These feral horses are located in the Toi-Cape area, Miyazaki, in southern Japan. Despite clear evidence

of positive serological reactions (both AGID and Western blotting), proviral DNA was not demonstrable in peripheral blood of any of the infected horses using PCR primers recommended by OIE (OIE Terrestrial Manual 2008), previously published primer sets or unpublished primer pairs currently being evaluated in other laboratories. In the present study, a nested PCR assay for detection of EIAV proviral DNA in peripheral blood cells of naturally infected horses was developed. Primer sets were designed based on conserved 5' regions of the viral genome extending from the long terminal repeat (LTR) to the trans-activator (*tat*) gene. Preliminary studies demonstrated the method has a detection limit of 10 genomic copies and when applied to a naturally EIAV infected feral horse population possessed a 100% correlation with conventional serological diagnostic techniques. 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 artifact hybrid molecules that inevitably occur when different regions of the viral genome are amplified separately. It provides a powerful new tool in the control of EIAV.

EIA was first reported in Japan in 1883. During the first decade of 19th century, Japan was subjected to the first EIA epidemics and more than 300 horses a year died from this disease. Since then, EIA has continually threaten the horse breeding and caused a huge economic loss in Japan. However, although 130 years passed after EIA was first reported, so far there have been no genetic information described for Japanese EIAV strain. Therefore, it is still unclear for the relationship

not only between Japanese ancient and current strains, but also between Japanese strain and other reported strains. This epidemiology information is very important to disease prevention and control. In the present study, base on the nucleotide sequence result from the nested PCR assay, complete proviral sequences comprising a novel field strain (Miyazakai2011-A) 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 $EIAV_{Wyoming}$ and $EIAV_{Liaoning}$ strains respectively while similarity plot analysis suggested all three strains arose independently. Moreover, phylogenetic analysis of LTR, *gag*, *pol* and *env* revealed Miyazaki2011-A forms a separate monophyletic group and as such can be classified as a novel EIAV isolate. 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. Additionally, despite considerable nucleotide divergence almost all functional elements or motifs identified in $EIAV_{Wyoming}$ or its derivatives are retained in Miyazaki2011-A. In conclusion, the complete proviral sequence of a phylogenetically distinct EIAV field strain (Miyazaki2011-A) has been successfully obtained from an ancient breed of horses (Misaki) in southern Japan. Here, we first report genetic information of Japanese EIAV local strain which is important for EIAV epidemiology research and disease control.

5. Acknowledgements

First and foremost, I would like to gratefully acknowledge the enthusiastic supervision of Prof. GOTO Yoshitaka and Prof. HAGA Takeshi during my study. I am really appreciate their valuable guidance and advice, and their various supports. Meanwhile, I would like to give my sincerely appreciation to Prof. HORII Yoichiro for his help on sample collection and valuable comments. In particular, I would like to give my great thanks to Prof. Frank R. Cook in Gluck Equine Research Center, University of Kentucky, who taught me a lot of basic knowledge of this study and help me a lot to modify the manuscript. I would also like to thank Prof. IWATA Hiroyuki for his helpful comments.

I would like to thanks to the staffs working in the United Graduate School of Veterinary Sciences of Yamguchi University. Their support help me to complete this study on time. I would also like to take this opportunity to thank all members of Veterinary Microbiology Lab, Their love and friendly help was one of the main sources of inspiration for me to focus on my study. In addition, I am really appreciate the help from the staffs working in Veterinary Department.

Finally, I would like to give my great thanks to my parents for their support, although words can not express it. Meanwhile, I thank Dr. ZHU Wei, who is my partner in lab and my wife in life. Her love gave me power to continue my study. I would also like to thank my daughter Kayo and my son Kako, Their love encourage me and their smile give me confidence.

6. References

Albrecht, T. R., Lund, L. H. & Garcia-Blanco, M. A. (2000). Canine cyclin T1 rescues equine infectious anemia virus tat trans-activation in human cells. *Virology* 268, 7-11.

Amodeo, P., Castiglione Morelli, M. A., Ostuni, A., Battistuzzi, G. & Bavoso, A. (2006). Structural features in EIAV NCp11: a lentivirus nucleocapsid protein with a short linker. *Biochemistry* 45, 5517-5526.

Baccam, P., Thompson, R. J., Li, Y., Sparks, W. O., Belshan, M., Dorman, K. S., Wannemuehler, Y., Oaks, J. L., Cornette, J. L. & Carpenter, S. (2003). Subpopulations of equine infectious anemia virus Rev coexist in vivo and differ in phenotype. *J Virol* 77, 12122-12131.

Belshan, M., Baccam, P., Oaks, J. L., Sponseller, B. A., Murphy, S. C., Cornette, J. & Carpenter, S. (2001). Genetic and biological variation in equine infectious anemia virus Rev correlates with variable stages of clinical disease in an experimentally infected pony. *Virology* 279, 185-200.

Bieniasz, P. D., Grdina, T. A., Bogerd, H. P. & Cullen, B. R. (1999). Highly divergent lentiviral Tat proteins activate viral gene expression by a common mechanism. *Mol Cell Biol* 19, 4592-4599.

Cappelli, K., Capomaccio, S., Cook, F. R., Felicetti, M., Marenzoni, M. L., Coppola, G., Verini-Supplizi, A., Coletti, M. & Passamonti, F. (2011). Molecular detection, epidemiology, and genetic characterization of novel European field isolates of equine infectious anemia virus. *J Clin Microbiol* 49, 27-33.

Carpenter, S. & Chesebro, B. (1989). Change in host cell tropism associated with in vitro replication of equine infectious anemia virus. *J Virol* 63, 2492-2496.

Carroll, R., Martarano, L. & Derse, D. (1991). Identification of lentivirus tat functional domains through generation of equine infectious anemia virus/human immunodeficiency virus type 1 tat gene chimeras. *J Virol* 65, 3460-3467.

Carvalho, M. & Derse, D. (1993a). Physical and functional characterization of transcriptional control elements in the equine infectious anemia virus promoter. *J Virol* 67, 2064-2074.

Carvalho, M. & Derse, D. (1993b). The PU.1/Spi-1 proto-oncogene is a transcriptional regulator of a lentivirus promoter. *J Virol* 67, 3885-3890.

Centaur (1993). *Equine infectious anemia antiboday test protocol*. Overland Park, Kansas: Centaur Inc.

Cheevers, W. P. & McGuire, T. C. (1985). Equine infectious anemia virus: immunopathogenesis and persistence. *Rev Infect Dis* 7, 83-88.

Chen, C., Li, F. & Montelaro, R. C. (2001). Functional roles of equine infectious anemia virus Gag p9 in viral budding and infection. *J Virol* 75, 9762-9770.

Chen, C. & Montelaro, R. C. (2003). Characterization of RNA elements that regulate gag-pol ribosomal frameshifting in equine infectious anemia virus. *J Virol* 77, 10280-10287.

Clabough, D. L., Gebhard, D., Flaherty, M. T., Whetter, L. E., Perry, S. T., Coggins, L. & Fuller, F. J. (1991). Immune-mediated thrombocytopenia in horses infected with equine infectious anemia virus. *J Virol* 65, 6242-6251.

Coggins, L., Norcross, N. L. & Nusbaum, S. R. (1972). Diagnosis of equine infectious anemia by immunodiffusion test. *Am J Vet Res* 33, 11-18.

Cook, R. F., Cook, S. J., Bolin, P. S., Howe, L. J., Zhou, W., Montelaro, R. C. & Issel, C. J. (2005). Genetic immunization with codon-optimized equine infectious anemia virus (EIAV) surface unit (SU) envelope protein gene sequences stimulates immune responses in ponies. *Vet Microbiol* 108, 23-37.

Cook, R. F., Cook, S. J., Li, F. L., Montelaro, R. C. & Issel, C. J. (2002). Development of a multiplex real-time reverse transcriptase-polymerase chain reaction for equine infectious anemia virus (EIAV). *J Virol Methods* 105, 171-179.

Covaleda, L., Fuller, F. J. & Payne, S. L. EIAV S2 enhances pro-inflammatory cytokine and chemokine response in infected macrophages. *Virology* 397, 217-223.

Covaleda, L., Gno, B. T., Fuller, F. J. & Payne, S. L. (2010). Identification of cellular proteins interacting with equine infectious anemia virus S2 protein. *Virus Res* 151, 235-239.

Craigo, J. K., Barnes, S., Zhang, B., Cook, S. J., Howe, L., Issel, C. J. & Montelaro, R. C. (2009). An EIAV field isolate reveals much higher levels of subtype variability than currently reported for the equine lentivirus family. *Retrovirology* 6, 95.

Craigo, J. K., Leroux, C., Howe, L., Steckbeck, J. D., Cook, S. J., Issel, C. J. & Montelaro, R. C. (2002). Transient immune suppression of inapparent carriers infected with a principal neutralizing domain-deficient equine

infectious anaemia virus induces neutralizing antibodies and lowers steady-state virus replication. *J Gen Virol* 83, 1353-1359.

Cullinane, A., Quinlivan, M., Nelly, M., Patterson, H., Kenna, R., Garvey, M., Gildea, S., Lyons, P., Flynn, M., Galvin, P., Neylon, M. & Jankowska, K. (2007). Diagnosis of equine infectious anaemia during the 2006 outbreak in Ireland. *Vet Rec* 161, 647-652.

Derse, D., Dorn, P. L., Levy, L., Stephens, R. M., Rice, N. R. & Casey, J. W. (1987). Characterization of equine infectious anemia virus long terminal repeat. *J Virol* 61, 743-747.

Dong, J. B., Zhu, W., Cook, F. R., Goto, Y., Horii, Y. & Haga, T. (,in press). Development of a nested PCR assay to detect equine infectious anemia proviral DNA from peripheral blood of naturally infected horses. *Arch Virol*.

Dorn, P., DaSilva, L., Martarano, L. & Derse, D. (1990). Equine infectious anemia virus tat: insights into the structure, function, and evolution of lentivirus trans-activator proteins. *J Virol* 64, 1616-1624.

Grund, C. H., Lechman, E. R., Issel, C. J., Montelaro, R. C. & Rushlow, K. E. (1994). Lentivirus cross-reactive determinants present in the capsid protein of equine infectious anaemia virus. *J Gen Virol* 75 (Pt 3), 657-662.

Hammond, S. A., Li, F., McKeon, B. M., Sr., Cook, S. J., Issel, C. J. & Montelaro, R. C. (2000). Immune responses and viral replication in long-term inapparent carrier ponies inoculated with equine infectious anemia virus. *J Virol* 74, 5968-5981.

Harrold, S. M., Cook, S. J., Cook, R. F., Rushlow, K. E., Issel, C. J. & Montelaro,

R. C. (2000). Tissue sites of persistent infection and active replication of equine infectious anemia virus during acute disease and asymptomatic infection in experimentally infected equids. *J Virol* 74, 3112-3121.

Hatanaka, H., Iourin, O., Rao, Z., Fry, E., Kingsman, A. & Stuart, D. I. (2002). Structure of equine infectious anemia virus matrix protein. *J Virol* 76, 1876-1883.

Hines, R., Sorensen, B. R., Shea, M. A. & Maury, W. (2004). PU.1 binding to ets motifs within the equine infectious anemia virus long terminal repeat (LTR) enhancer: regulation of LTR activity and virus replication in macrophages. *J Virol* 78, 3407-3418.

Hussain, K. A., Issel, C. J., Rwambo, P. M., Arnizaut, A. B., Ball, J. M., Schnorr, K. L. & Montelaro, R. C. (1988). Identification of gag precursor of equine infectious anaemia virus with monoclonal antibodies to the major viral core protein, p26. *J Gen Virol* 69 (Pt 7), 1719-1724.

Issel, C. J. & Coggins, L. (1979). Equine infectious anemia: current knowledge. *J Am Vet Med Assoc* 174, 727-733.

Issel, C. J., Horohov, D. W., Lea, D. F., Adams, W. V., Jr., Hagius, S. D., McManus, J. M., Allison, A. C. & Montelaro, R. C. (1992). Efficacy of inactivated whole-virus and subunit vaccines in preventing infection and disease caused by equine infectious anemia virus. *J Virol* 66, 3398-3408.

Kobayashi, K. & Kono, Y. (1967). Serial passages of equine infectious anemia virus in horse leukocyte culture. *Natl Inst Anim Health Q (Tokyo)* 7, 1-7.

Kono, Y., Hirasawa, K., Fukunaga, Y. & Taniguchi, T. (1976). Recrudescence of

equine infectious anemia by treatment with immunosuppressive drugs. *Natl Inst Anim Health Q (Tokyo)* 16, 8-15.

Kono, Y., Kobayashi, K. & Fukunaga, Y. (1970). Immunization of horses against equine infectious anemia (EIA) with an attenuated EIA virus. *Natl Inst Anim Health Q (Tokyo)* 10, 113-122.

Kono, Y., Kobayashi, K. & Fukunaga, Y. (1971). Serological comparison among various strains of equine infectious anemia virus. *Arch Gesamte Virusforsch* 34, 202-208.

Lairmore, M. D. (2011). Equine Infectious Anemia Virus (Retroviridae). In *Fenner's Veterinary Virology*, Fourth edn, pp. 269-270. Edited by N. J. MacLachlan & E. J. Dubovi. London: Elsevier.

Lee, J. H., Murphy, S. C., Belshan, M., Sparks, W. O., Wannemuehler, Y., Liu, S., Hope, T. J., Dobbs, D. & Carpenter, S. (2006). Characterization of functional domains of equine infectious anemia virus Rev suggests a bipartite RNA-binding domain. *J Virol* 80, 3844-3852.

Leroux, C., Cadore, J. L. & Montelaro, R. C. (2004). Equine Infectious Anemia Virus (EIAV): what has HIV's country cousin got to tell us? *Vet Res* 35, 485-512.

Leroux, C., Issel, C. J. & Montelaro, R. C. (1997). Novel and dynamic evolution of equine infectious anemia virus genomic quasispecies associated with sequential disease cycles in an experimentally infected pony. *J Virol* 71, 9627-9639.

Li, F., Chen, C., Puffer, B. A. & Montelaro, R. C. (2002). Functional replacement

and positional dependence of homologous and heterologous L domains in equine infectious anemia virus replication. *J Virol* 76, 1569-1577.

Li, F., Craigo, J. K., Howe, L., Steckbeck, J. D., Cook, S., Issel, C. & Montelaro, R. C. (2003). A live attenuated equine infectious anemia virus proviral vaccine with a modified S2 gene provides protection from detectable infection by intravenous virulent virus challenge of experimentally inoculated horses. *J Virol* 77, 7244-7253.

Li, F., Leroux, C., Craigo, J. K., Cook, S. J., Issel, C. J. & Montelaro, R. C. (2000). The S2 gene of equine infectious anemia virus is a highly conserved determinant of viral replication and virulence properties in experimentally infected ponies. *J Virol* 74, 573-579.

Li, F., Puffer, B. A. & Montelaro, R. C. (1998). The S2 gene of equine infectious anemia virus is dispensable for viral replication in vitro. *J Virol* 72, 8344-8348.

Lichtenstein, D. L., Rushlow, K. E., Cook, R. F., Raabe, M. L., Swardson, C. J., Kociba, G. J., Issel, C. J. & Montelaro, R. C. (1995). Replication in vitro and in vivo of an equine infectious anemia virus mutant deficient in dUTPase activity. *J Virol* 69, 2881-2888.

Lloveras, J., Soler, C. & Celada, A. (1999). The key role of PU.1/SPI-1 in B cells, myeloid cells and macrophages. *Immunol Today* 20, 184-189.

Lole, K. S., Bollinger, R. C., Paranjape, R. S., Gadkari, D., Kulkarni, S. S., Novak, N. G., Ingersoll, R., Sheppard, H. W. & Ray, S. C. (1999). Full-length human immunodeficiency virus type 1 genomes from subtype C-infected

seroconverters in India, with evidence of intersubtype recombination. *J Virol* 73, 152-160.

MacLachlan, N. J. & Dubovi, E. J. (2011). Equine Infectious Anemia Virus (Retroviridae). In *Fenner's Veterinary Virology*, Fourth edn, pp. 269-270: Elsevier.

Malmquist, W. A., Barnett, D. & Becvar, C. S. (1973). Production of equine infectious anemia antigen in a persistently infected cell line. *Arch Gesamte Virusforsch* 42, 361-370.

Matsushita, T., Hesterberg, L. K., Porter, J. P., Smith, B. J. & Newman, L. E. (1989). Comparison of diagnostic tests for the detection of equine infectious anemia antibody. *J Vet Diagn Invest* 1, 50-52.

Montelaro, R. C., Ball, J. M. & E., R. (1993). Equine retroviruses. In *The Retroviridae* J. A. Levy edn, pp. 257-360. Edited by J. A. Levy. New York: Plenum Press.

More, S. J., Aznar, I., Bailey, D. C., Larkin, J. F., Leadon, D. P., Lenihan, P., Flaherty, B., Fogarty, U. & Brangan, P. (2008). An outbreak of equine infectious anaemia in Ireland during 2006: investigation methodology, initial source of infection, diagnosis and clinical presentation, modes of transmission and spread in the Meath cluster. *Equine Vet J* 40, 706-708.

Nagarajan, M. M. & Simard, C. (2001). Detection of horses infected naturally with equine infectious anemia virus by nested polymerase chain reaction. *J Virol Methods* 94, 97-109.

Nagarajan, M. M. & Simard, C. (2007). Gag genetic heterogeneity of equine

infectious anemia virus (EIAV) in naturally infected horses in Canada. *Virus Res* 129, 228-235.

Noiman, S., Yaniv, A., Sherman, L., Tronick, S. R. & Gazit, A. (1990). Pattern of transcription of the genome of equine infectious anemia virus. *J Virol* 64, 1839-1843.

OIE (2008). Equine infectious anaemia. In *Manual of diagnostic tests and vaccines for terrestrial animals*, Sixth edn, pp. 866-870.

Petropoulos, C. J. (1997). Retroviral Taxonomy, Protein Structures, Sequences, and Genetic Maps. In *Retroviruses*, p. 757. Edited by J. M. Coffin. New York: Cold Spring Harbor Laboratory Press.

Puffer, B. A., Parent, L. J., Wills, J. W. & Montelaro, R. C. (1997). Equine infectious anemia virus utilizes a YXXL motif within the late assembly domain of the Gag p9 protein. *J Virol* 71, 6541-6546.

Quinlivan, M., Cook, R. F. & Cullinane, A. (2007). Real-time quantitative RT-PCR and PCR assays for a novel European field isolate of equine infectious anaemia virus based on sequence determination of the gag gene. *Vet Rec* 160, 611-618.

Reis, J. K., Craigo, J. K., Cook, S. J., Issel, C. J. & Montelaro, R. C. (2003). Characterization of EIAV LTR variability and compartmentalization in various reservoir tissues of long-term inapparent carrier ponies. *Virology* 311, 169-180.

Sa Filho, D. J., Sanabani, S., Diaz, R. S., Munerato, P., Brunstein, A., Fusuma, E., Sabino, E. C. & Janini, L. M. (2005). Analysis of full-length human

immunodeficiency virus type 1 genome reveals a variable spectrum of subtypes B and f recombinants in Sao Paulo, Brazil. *AIDS Res Hum Retroviruses* 21, 145-151.

Sellon, D. C., Fuller, F. J. & McGuire, T. C. (1994). The immunopathogenesis of equine infectious anemia virus. *Virus Res* 32, 111-138.

Sentsui, H. (2011). Equine infectious anemia. In *Infectious Disease of Animals*, 3rd edn, p. 157. Edited by H. Akashi. Tokyo: Kindai publisher (in Japanese).

Shane, B. S., Issel, C. J. & Montelaro, R. C. (1984). Enzyme-linked immunosorbent assay for detection of equine infectious anemia virus p26 antigen and antibody. *J Clin Microbiol* 19, 351-355.

Shen, R. & Wang, Z. M. (1985). Development and use of an equine infectious aneamia donkey leukocyte attenuated vaccine. In *Equine Infectious Anemia: a national review of policies, programs, and future objectives*, pp. 135-148. Edited by R. Tashjian. Texas: Amarillo.

Sherefa, K., Johansson, B., Salminen, M. & Sonnerborg, A. (1998). Full-length sequence of human immunodeficiency virus type 1 subtype A, recombined with subtype C in the env V3 domain. *AIDS Res Hum Retroviruses* 14, 289-292.

Silva, R. A. M. S., Abreu, U. G. P. d., Davila, A. M. R. & Ramirez, L. (1999). Swamp fever in wild horses from the Pantanal, Brazil *Revue d'Élevage et de Médecine Vétérinaire des Pays Tropicaux* 52, 99-101

Sparks, W. O., Dorman, K. S., Liu, S. & Carpenter, S. (2008). Naturally arising point mutations in non-essential domains of equine infectious anemia virus

Rev alter Rev-dependent nuclear-export activity. *J Gen Virol* 89, 1043-1048.

Steagall, W. K., Robek, M. D., Perry, S. T., Fuller, F. J. & Payne, S. L. (1995). Incorporation of uracil into viral DNA correlates with reduced replication of EIAV in macrophages. *Virology* 210, 302-313.

Stephens, R. M., Casey, J. W. & Rice, N. R. (1986). Equine infectious anemia virus gag and pol genes: relatedness to visna and AIDS virus. *Science* 231, 589-594.

Stephens, R. M., Derse, D. & Rice, N. R. (1990). Cloning and characterization of cDNAs encoding equine infectious anemia virus tat and putative Rev proteins. *J Virol* 64, 3716-3725.

Sticht, H., Willbold, D., Bayer, P., Ejchart, A., Herrmann, F., Rosin-Arbesfeld, R., Gazit, A., Yaniv, A., Frank, R. & Rosch, P. (1993). Equine infectious anemia virus Tat is a predominantly helical protein. *Eur J Biochem* 218, 973-976.

Su, L., Graf, M., Zhang, Y., von Briesen, H., Xing, H., Kostler, J., Melzl, H., Wolf, H., Shao, Y. & Wagner, R. (2000). Characterization of a virtually full-length human immunodeficiency virus type 1 genome of a prevalent intersubtype (C/B') recombinant strain in China. *J Virol* 74, 11367-11376.

Sune, C., Goldstrohm, A. C., Peng, J., Price, D. H. & Garcia-Blanco, M. A. (2000). An in vitro transcription system that recapitulates equine infectious anemia virus tat-mediated inhibition of human immunodeficiency virus type 1 Tat activity demonstrates a role for positive transcription elongation factor b and associated proteins in the mechanism of Tat activation. *Virology* 274, 356-366.

Suzuki, T., Ueda, S. & Samejima, T. (1982). Enzyme-linked immunosorbent assay for diagnosis of equine infectious anemia. *Vet Microbiol* 7, 307-315.

Tabuchi, E., Katada, M., Hirasawa, M. & Fujino, M. (1967). An outbreak of equine infectious anemia in race horses - epizootical, clinical, and hematological observations. *Bull Nat Inst Anim Hlth* 55, 45-51.

Tabuchi, E., Tanaka, K., Akiyama, Y., Katada, M., Yamamoto, H., Sakaki, K. & Yoshino, T. (1965). Long-term reservation of equine infectious anemia virus in a horse. *J Jap vet med Ass* 18, 437-442.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. & Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* 28, 2731-2739.

Thomas, D. A. & Furman, P. A. (1991). Purification and kinetic characterization of equine infectious anemia virus reverse transcriptase. *Biochem Biophys Res Commun* 180, 1365-1371.

Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. & Higgins, D. G. (1997). The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* 25, 4876-4882.

Threadgill, D. S., Steagall, W. K., Flaherty, M. T., Fuller, F. J., Perry, S. T., Rushlow, K. E., Le Grice, S. F. & Payne, S. L. (1993). Characterization of equine infectious anemia virus dUTPase: growth properties of a dUTPase-deficient mutant. *J Virol* 67, 2592-2600.

Tu, Y. B., Zhou, T., Yuan, X. F., Qiu, H. J., Xue, F., Sun, C. Q., Wang, L., Wu, D. L., Peng, J. M., Kong, X. G. & Tong, G. Z. (2007). Long terminal repeats are not the sole determinants of virulence for equine infectious anemia virus. *Arch Virol* 152, 209-218.

Weiland, F., Matheka, H. D. & Bohm, H. O. (1982). Equine infectious anaemia: detection of antibodies using an immunofluorescence test. *Res Vet Sci* 33, 347-350.

Willbold, D., Rosin-Arbesfeld, R., Sticht, H., Frank, R. & Rosch, P. (1994). Structure of the equine infectious anemia virus Tat protein. *Science* 264, 1584-1587.

Zheng, Y. H., Nakaya, T., Sentsui, H., Kameoka, M., Kishi, M., Hagiwara, K., Takahashi, H., Kono, Y. & Ikuta, K. (1997). Insertions, duplications and substitutions in restricted gp90 regions of equine infectious anaemia virus during febrile episodes in an experimentally infected horse. *J Gen Virol* 78 (Pt 4), 807-820.

Zheng, Y. H., Sentsui, H., Kono, Y. & Ikuta, K. (2000). Mutations occurring during serial passage of Japanese equine infectious anemia virus in primary horse macrophages. *Virus Res* 68, 93-98.