Polish Journal of Veterinary Sciences Vol. 18, No. 2 (2015), 255–259

DOI 10.1515/pjvs-2015-0033

Original article

Variability of non-structural proteins of equine arteritis virus during persistent infection of the stallion

W. Socha, J. Rola, J.F. Żmudziński

National Veterinary Research Institute, Partyzantow 57, 24-100 Pulawy, Poland

Abstract

The genetic stability of ORF1a encoding non-structural proteins nsp1, nsp2, nsp3 and nsp4 of equine arteritis virus (EAV) has been analysed for nearly seven years in a persistently infected stallion of the Malopolska breed. Between November 2004 and June 2011, 11 semen samples were collected. Viral RNA extracted from semen of this carrier stallion was amplified, sequenced and compared with the sequences of the other known strains of EAV. Sequence analysis of ORF1a showed 84 synonymous and 16 non-synonymous mutations. The most variable part of ORF1a was the region encoding nsp2 protein with 13 non-synonymous substitutions. The degree of amino acid identity between isolates ranged from 98.91 to 100%. Only single non-synonymous mutations were detected in nsp1 (one substitution) and nsp4 (two substitutions). The most stable was nsp3 in which no amino acid substitutions were observed during the whole period of observation.

Key words: equine arteritis virus, nsp, stallion, persistent infection

Introduction

Equine arteritis virus (EAV) is the causative agent of equine viral arteritis (EVA), a respiratory and reproductive disease of horses and other equid species that occurs in many countries all over the world (Balasuriya et al. 2013).

Exposure to EAV usually results in a mild or subclinical infection but some strains of the virus can cause outbreaks of EVA characterized by an influenza-like illness in adult horses, abortion in pregnant mares, interstitial pneumonia in young foals and death in newborn foals (Holyoak et al. 2008). Most infected horses recover without complications but 10 to 70% of stallions infected with EAV can subsequently become carriers and constantly shed the virus in their semen even over their whole lifetime (Timoney et al. 1986). Apart from being a natural reservoir of the virus in equine populations, carrier stallions could also be a source of genetic and phenotypic diversity of EAV (Balasuriya et al. 2004).

The genome of EAV is composed of positive-sense RNA of approximately 12.7 kb in length which includes 10 known open reading frames (ORFs). The two most 5'-proximal ORFs (1a and 1b) occupy approximately three-quarters of the genome and encode two replicase polyproteins. These polyproteins are further processed during viral replication

Correspondence to: e-mail: jrola@piwet.pulawy.pl, tel.: +48 81 889 30 69

resulting in the formation of 13 non-structural proteins (Balasuriya et al. 2013). The remaining part of the EAV genome is composed of 8 ORFs (ORF2a, ORF2b, ORF3, ORF4, ORF5, 5a, ORF6 and ORF7) encoding structural proteins of the virus (de Vries et al. 1992, Wieringa et al. 2002, Firth et al. 2011). Studies by Balasuriya et al. (2004) showed that within ORF1a and ORF1b, the most variable regions are located at the 5' part of the EAV genome. Proteins nsp1, nsp2 and nsp4 are three ORF1a encoded proteases engaged in the posttranslational processing of the two viral replicases (van Aken et al. 2006, Zhang et al. 2008, Miszczak et al. 2012). It was also shown that during persistent infection regions encoding nsp1 and nsp2 are prone to mutations which could affect the virulence of EAV (Zhang et al. 2008, Miszczak et al. 2012). The majority of studies on the genetic variation of EAV have focused on the genes encoding structural proteins of the virus (Patton et al. 1999, Liu et al. 2008). Papers published by research teams from Poland also concerned these proteins (Rola et al. 2011, Surma-Kurusiewicz et al. 2013). So far there is no information on the genetic variability of non-structural proteins of Polish strains of EAV.

The aim of the study was to characterize genetic changes in selected non-structural proteins of EAV isolates during persistent infection in the shedder stallion.

Materials and Methods

Sample collection

A total of 11 blood and semen samples were collected from a persistently infected, 18-year old stallion of the Malopolska breed between 2004 and 2011 as indicated below (11/2004 [month/year], 08/2005, 11/2005, 03/2006, 10/2006, 12/2006, 01/2007, 11/2009, 02/2010, 04/2010 and 06/2011).

RNA extraction and RT-PCR

Viral RNA was extracted directly from the semen samples using TRI Reagent (Sigma Aldrich) according to the manufacturer's instructions. PCR were performed with Titan One-Tube RT-PCR kit (Roche) containing AMV Reverse Transcriptase and the Expand High Fidelity PCR System enzyme mix. For partial amplification of ORF1a five sets of primers were used: 1P:1006N, 833P:1540N, 1891P:2508N, 2296P:3118N and 3047P:4029N (Zhang et al. 2005) (Table 1).

Sequence analysis

PCR products were purified with QIAquick Gel extraction kit (Qiagen) and suspended in a total volume of 20 µl of DEPC water. Sequencing reactions were performed using Applied Biosystems BigDye® Terminator v3.1 kit (Life Technologies) using the Sanger method by Genomed (Poland). The nucleotide sequences reported in this study were submitted to GenBank under accession numbers KP009899--KP009909.

For each of the analysed non-structural proteins, the average dN/dS ratio was calculated using MEGA 5.05, based on the Nei – Gojobori algorithm. Additionally, based on the sequence of the ORF1a region encoding nsp2 protein, a phylogenetic tree of all the EAV isolates was constructed using the Neighbour-Joining method (Fig. 1).

Results

Sequencing of genome fragments encoding selected non-structural proteins of the EAV revealed 84 synonymous and 16 non-synonymous mutations which appeared during the nearly seven-year observation period of persistent infection; however, no deletions or insertions were identified (Table 2). Additionally, 2 variable sites were found in the 5'UTR Leader sequence in nucleotide positions 84 and 134.

In the region encoding nsp1 protein, 14 synonymous and 1 non-synonymous mutations were found. The degree of amino acid identity between isolates ranged from 99.12 to 100% with a dN/dS ratio equal to 0.009.

The most variable part of ORF1a was the region encoding nsp2 protein with 31 synonymous and 13 non-synonymous substitutions. It was also characterized by the highest average dN/dS ratio – 0.1382. The degree of amino acid identity between isolates ranged from 98.91 to 100%.

The Nsp3 coding region was the most conservative part of ORF1a, with only 18 synonymous mutations.

In the region encoding nsp4 protein, 21 synonymous and 2 non-synonymous substitutions were found. The dN/dS ratio was 0.0269, whereas the degree of amino acid substitutions ranged from 98.38 to 100%.

Discussion

Several studies have shown that genes encoding non-structural proteins of EAV undergo mutations during persistent infection (Balasuriya et al. 2004,

Table 1. Primers used for amplification and sequencing of ORF1a of EAV strains detected in the semen of the carrier stallion.

Name	Forward Primer (5'-3' sequence)	Name	Reverse Primer (5'-3' sequence)	Product size	Amplified region
1P	GCTCGAAGTGTGTATGGTGCCATATACGGC	1006N	TCCATCCCCTGGAGGGTTGTA	1005 nt	leader/nsp1
833P	AGACCTGGGTTTGGGCATCA	1540N	CAGGCGCCCATCCCAGCACC	707 nt	nsp1/nsp2
1891P	TGGGCAATAATGTTGTTCTG	2508N	TCACAACAGCTCCTGTTTGTTC	617 nt	nsp 2
2296P	TCGATGTTGTGGGCATGGC	3118N	GAACAGATTCACAAAAAGCG	822 nt	nsp 2/3
3047P	GTATTGCTTGTGTTCCC	4029N	GCTTGTCCAACAGGGA	982 nt	nsp 3/4

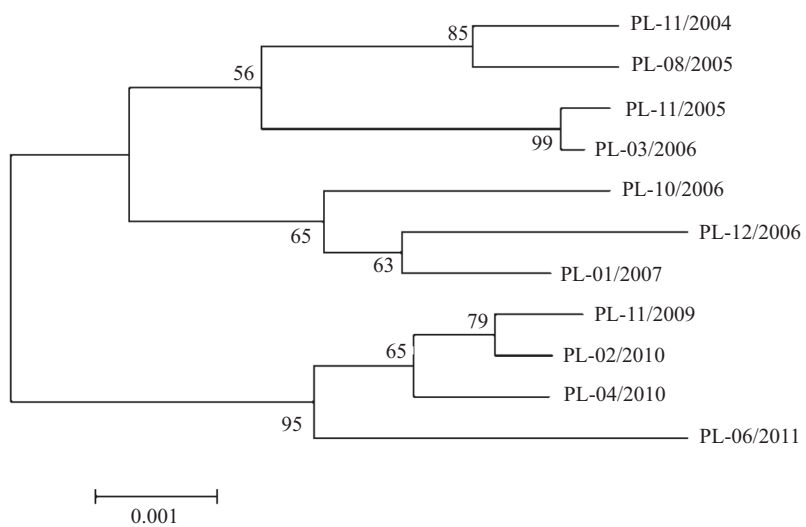

Fig. 1. Phylogenetic tree of ORF1a fragment encoding nsp2. The tree was constructed using the Neighbour-Joining method using MEGA 5.05 software.

Table 2. Nucleotide and amino acid differences between EAV isolates, sampled between 2004 and 2011.

Region	Nucleotide substitutions	Amino Acid substitiutions	N	dS	dN/dS
5'UTR (Leader)	2	NA	NA	NA	NA
nsp1 (ORF1a)	15	1	0.0003	0.0327	0.009
nsp2 (ORF1a)	44	13	0.0030	0.0217	0.1382
nsp3 (ORF1a)	18	0	0	0.0271	0
nsp4 (ORF1a)	23	2	0.021	0.0269	0.078
nsp1 – nsp4 (ORF1a)	100	16	0.0013	0.0331	0.039

NA, Not applicable; dN/dS – average ratio of non-synonymous substitutions per non-synonymous site (dN) to the number of synonymous substitutions per synonymous site (dS)

Miszczak et al. 2012). However, the intervals between samplings in these studies were usually longer compared to our study, therefore it seems that our sampling scheme was more reliable in identifying fixed mutations in the analysed region of the EAV genome.

The 5'UTR region of EAV is involved in regulating translation, replication and transcription of the EAV genome. Although this region does not undergo translation, variation in its nucleotide sequence could lead to changes in its secondary structure, affecting the viability of the virus (van den Born et al. 2004). However, in our study nucleotide variation was observed only at positions with no importance for the functionality of the 5'UTR region.

Table 3. Amino acid substitutions in non-structural proteins of EAV strains detected in the semen of the carrier stallion. Substitutions in respect to 11/2004 isolate are shaded.

Position / Strain	Nsp1	Nsp2													Nsp4	
	153	146	170	177	192	216	221	223	240	265	309	434	449	471	39	87
11/2004	Ile	Ser	Pro	Thr	Pro	Ala	Ala	Leu	Leu	Ala	Ser	Val	Pro	Iso	His	Pro
08/2005	**Met**	Ser	Pro	Thr	Pro	Ala	Ala	**Val**	Leu	Ala	**Cys**	Val	Pro	Iso	His	Pro
11/2005	Ile	Ser	Pro	**Iso**	**Ser**	Ala	Ala	**Val**	Leu	Ala	**Cys**	Val	Pro	**Leu**	His	Pro
03/2006	Ile	Ser	Pro	Thr	**Ser**	Ala	Ala	Leu	Leu	Ala	**Cys**	Val	**Thr**	**Leu**	His	**Leu**
10/2006	Ile	**Gly**	**Leu**	Thr	**Ser**	**Ser**	**Thr**	Leu	Leu	Ala	**Cys**	Val	**Thr**	**Leu**	His	**Leu**
12/2006	Ile	**Gly**	**Leu**	Thr	**Ser**	**Ser**	**Thr**	Leu	Leu	Ala	**Cys**	Val	**Thr/Pro**	**Thr/Pro**	His	**Leu**
01/2007	Ile	**Gly**	**Leu**	**Iso**	**Ser**	Ala	**Thr**	Leu	Leu	Ala	**Cys**	Val	**Thr**	**Leu**	His	**Leu**
11/2009	Ile	Ser	Pro	Thr	**Ser**	Ala	**Thr**	Val/Leu	**Iso**	**Thr**	**Cys**	Val	**Thr**	**Leu**	His	Pro
02/2010	Ile	Ser	Pro	Thr	**Ser**	Ala	**Thr**	Leu	**Iso**	**Thr**	**Cys**	Val	**Thr**	**Leu**	His	Pro
04/2010	Ile	Ser	Pro	**Iso**	**Ser**	**Ser**	**Thr**	Leu	**Iso**	**Thr**	**Cys**	Val	**Thr**	**Leu**	Pro	Pro
06/2011	Ile	Ser	Pro	**Iso**	**Ser**	**Ser**	**Thr**	Leu	**Iso**	**Thr**	**Cys**	**Iso**	**Thr**	**Leu**	His	Pro

Nsp1 protein is a papain-like cysteine proteinase (PCP) which cleaves ORF1a transcript into functional non-structural replicase proteins (Ziebuhr et al. 2000). Analysis of the nsp1 protein showed that the minimal domain required for activity of this proteinase spans from 123 to 263 amino acid and mutation in one of two residue positions: either Cys-164 or His-230 could abolish its function. The significance of this domain for the activity of the nsp1 was also confirmed by Zhang et al. (2008), who showed that mutations at Ser141>Asn and Ile156>Thr were associated with the attenuation of the EAV strain HK. In our study we found only a single mutation, Ile153>Met in the 08/2005 isolate (Table 3). The lack of similar changes in the other isolates indicate that nsp1 is stable and has rather a low potential for changing during persistence as evidenced by a low dN/dS ratio that was 0.009. Balasuriya et al. (2004), also detected one amino acid substitution in nsp1 of EAV; however, the dN/dS ratio was much higher (0.043) than in our study. This may indicate that the American strain was characterized by a higher rate of genetic change than this Polish one. It is hard to deduce whether this difference is due to different environmental conditions or is caused by a longer observation period or a higher number of isolates used in our study than in Balasuriya et al. (2004) studies.

Nsp2 together with nsp3 are transmembrane proteins, functioning as elements of viral RNA replication and transcription complex (Snijder et al. 2013). We found, similarly to the results of Balasuriya et al. (2004), that the amino acid sequence of nsp3 remained unchanged during persistent infection, but nsp2 was characterized by high variability. This is in accordance with already available data on the stability of nsp2 (Balasuriya et al. 2004, Miszczak et al. 2012, Snijder et al 2013). However, contrary to those results, we did not find any deletions or substitutions. Nevertheless, nsp2 protein remained the most variable part of the analysed ORF1a region with 13 amino acid substitutions observed in the 7 year study and dN/dS ratio of 0.1382. Mutations in this protein were not only numerous but also in most cases remained stable till the end of the study, which was visible on the phylogenetic tree created based on the sequence encoding nsp2 (Fig 2.). On this dendrogram isolates from this stallion grouped together relative to the time of their isolation. This supports the previous assumptions that this protein undergoes molecular evolution as a result of selective pressure during persistant infection (Balasuriya et al. 2004).

The last of the analysed non-structural proteins was nsp4, which functions as a main 3C-like serine proteinase (Ziebuhr et al. 2000). It is responsible for the cleavage of nsp3-8 and nsp3-12 intermediates into individual polyproteins (Barrette-Ng et al. 2002). Previous analysis showed that the catalytic activity of nsp4 is dependent on the 3 residues on position His-39, Asp-65 and Ser-120, whereas the function of its unique C-terminal domain remains unknown (van Aken et al. 2006). In our study, two variable amino acid positions were found, both in the N-terminal part of the nsp4. Unexpectedly, in one of the Polish isolates (04/2010) His-39 was replaced with Pro. This could have greatly affected the catalytic capacity of nsp4 and as a result alter the fitness of the virus. Nevertheless, the fact that this variant of nsp4 coding region was present in only one isolate may suggest that it is not beneficial for the virus. Balasuriya et al. (2004) also analysed the variability of nsp4, finding

one amino acid substitution; however, they did not specify the position where it was found so it is not possible to compare if it is located in the same part of this protein.

Conclusion

Among the analysed non-structural proteins only nsp-2 possesses a high number of variable sites with stable mutations accumulating during the course of persistent infection. Changes in nsp1, nsp3 and nsp4 were observed rarely and were not fixed, which indicate that these regions of ORF1a are less prone to molecular changes.

Acknowledgements

This study was supported by the Polish Ministry of Science and Higher Education grant N N 308 101937.

References

Balasuriya UB, Go YY, MacLachlan NJ (**2013**) Equine arteritis virus. Vet Microbiol 167: 93-122.

Balasuriya UB, Hedges JF, Smalley VL, Navarrette A, McCollum WH, Timoney PJ, Snijder EJ, MacLachlan NJ (**2004**) Genetic characterization of equine arteritis virus during persistent infection of stallions. J Gen Virology 85: 379-390.

Barrette-Ng IH, Ng KK, Mark BL, Van Aken D, Cherney MM, Garen C, Kolodenko Y, Gorbalenya AE, Snijder EJ, James MN (**2002**) Structure of arterivirus nsp4. The smallest chymotrypsin-like proteinase with an alpha/beta C-terminal extension and alternate conformations of the oxyanion hole. J Biol Chem 277: 39960-39966. de Vries AA, Chirnside ED, Horzinek MC, Rottier PJ (**1992**) Structural proteins of equine arteritis virus. J Virol 66: 6294-6303.

Firth AE, Zevenhoven-Dobbe JC, Wills NM, Go YY, Balasuriya UB, Atkins JF, Snijder EJ, Posthuma CC (**2011**) Discovery of a small arterivirus gene that overlaps the GP5 coding sequence and is important for virus production. J Gen Virology 92: 1097-1106.

Holyoak GR, Balasuriya UB, Broaddus CC, Timoney PJ (**2008**) Equine viral arteritis: current status and prevention. Theriogenology 70: 403-414.

Liu L, Castillo-Olivares J, Davis-Poynter NJ, Baule C, Xia H, Belhk S (**2008**) Analysis of ORFs 2b, 3, 4, and partial ORF5 of sequential isolates of equine arteritis virus shows genetic variation following experimental infection of horses. Vet Microbiol 129: 262-268.

Miszczak F, Legrand L, Balasuriya UB, Ferry-Abitbol B, Zhang J, Hans A, Fortier G, Pronost S, Vabret A (**2012**) Emergence of novel equine arteritis virus (EAV) variants during persistent infection in the stallion: origin of the 2007 French EAV outbreak was linked to an EAV strain present in the semen of a persistently infected carrier stallion. Virology 423: 165-174.

Patton JF, Balasuriya UB, Hedges JF, Schweidler TM, Hullinger PJ, MacLachlan NJ (**1999**) Phylogenetic characterization of a highly attenuated strain of equine arteritis virus from the semen of a persistently infected standardbred stallion. Arch Virol 144: 817-827.

Rola J, Larska M, Rola JG, Belhk S, Autorino GL (**2011**) Epizotiology and phylogeny of equine arteritis virus in hucul horses. Vet Microbiol 148: 402-407.

Snijder EJ, Kikkert M, Fang Y (**2013**) Arterivirus molecular biology and pathogenesis. J Gen Virol 94: 2141-2163.

Surma-Kurusiewicz K, Winiarczyk S, Adaszek Ł (**2013**) Comparative analysis of ORF5 nucleotide sequences and amino acid sequences of the GP5 protein of equine arteritis virus (EAV) detected in the semen of stallions from Eastern Poland. Res Vet Sci 94: 361-367.

Timoney PJ, McCollum WH, Roberts AW, Murphy TW (**1986**) Demonstration of the carrier state in naturally acquired equine arteritis virus infection in the stallion. Res Vet Sci 41: 279-280.

van Aken D, Zevenhoven-Dobbe J, Gorbalenya AE, Snijder EJ (**2006**) Proteolytic maturation of replicase polyprotein pp1a by the nsp4 main proteinase is essential for equine arteritis virus replication and includes internal cleavage of nsp7. J Gen Virol 87: 3473-3482.

van den Born E, Gultyaev AP, Snijder EJ (**2004**) Secondary structure and function of the 5'-proximal region of the equine arteritis virus RNA genome. RNA 10: 424-437.

Wieringa R, de Vries AA, Raamsman MJ, Rottier PJ (**2002**) Characterization of Two New Structural Glycoproteins, GP(3) and GP(4), of Equine Arteritis Virus. J Virol 76: 10829-10840.

Zhang J (**2005**) Permissiveness of selected cell lines to equine arteritis virus: establishment, characterization, and significance of persistent infection in HeLa cells. University of Kentucky doctoral dissertations.

Zhang J, Go YY, MacLachlan NJ, Meade BJ, Timoney PJ, Balasuriya UB (**2008**) Amino acid substitutions in the structural or nonstructural proteins of a vaccine strain of equine arteritis virus are associated with its attenuation. Virology 378: 355-362.

Ziebuhr J, Snijder EJ, Gorbalenya AE (**2000**) Virus-encoded proteinases and proteolytic processing in the Nidovirales. J Gen Virol 81: 853-879.