Genetic variation and dynamics of infections of equid herpesvirus 5 in individual horses

Helena Back, † Karin Ullman, † Mikael Leijon, Robert Söderlund, Johanna Penell, Karl Ståhl, John Pringle and Jean-François Valarcher

1 Department of Virology, Immunobiology and Parasitology, National Veterinary Institute, Uppsala, Sweden
2 Department of Bacteriology, National Veterinary Institute, Uppsala, Sweden
3 Department of Veterinary Epidemiology and Public Health, University of Surrey, Guildford, UK
4 Department of Disease Control and Epidemiology, National Veterinary Institute, Uppsala, Sweden
5 Department of Clinical Sciences, Swedish University of Agricultural Sciences, Uppsala, Sweden

Equid herpesvirus 5 (EHV-5) is related to the human Epstein–Barr virus (human herpesvirus 4) and has frequently been observed in equine populations worldwide. EHV-5 was previously assumed to be low to non-pathogenic; however, studies have also related the virus to the severe lung disease equine multinodular pulmonary fibrosis (EMPF). Genetic information of EHV-5 is scanty: the whole genome was recently described and only limited nucleotide sequences are available. In this study, samples were taken twice 1 year apart from eight healthy horses at the same professional training yard and samples from a ninth horse that was diagnosed with EMPF with samples taken pre- and post-mortem to analyse partial glycoprotein B (gB) gene of EHV-5 by using next-generation sequencing. The analysis resulted in 27 partial gB gene sequences, 11 unique sequence types and five amino acid sequences. These sequences could be classified within four genotypes (I–IV) of the EHV-5 gB gene based on the degree of similarity of the nucleotide and amino acid sequences, and in this work horses were shown to be identified with up to three different genotypes simultaneously. The observations showed a range of interactions between EHV-5 and the host over time, where the same virus persists in some horses, whereas others have a more dynamic infection pattern including strains from different genotypes. This study provides insight into the genetic variation and dynamics of EHV-5, and highlights that further work is needed to understand the EHV-5 interaction with its host.

INTRODUCTION

Equid herpesvirus 5 (EHV-5) belongs to the genus Perca-virus, together with the closely related Equid herpesvirus 2 (EHV-2), and is classified into the subfamily Gammaherpesvirinae in the family Herpesviridae. EHV-5 shares similarities with, and is related to, the human gammaherpesvirus Epstein–Barr virus (EBV; human herpesvirus 4) (Telford et al., 1993). EHV-5 is a large DNA virus with a genome size of 182 kbp (Wilkie et al., 2015) and can, like other known herpesviruses, establish lifelong infection (Roizmann et al., 1992), although the latency sites for the infection within the host have not been established.

EHV-5 has commonly been observed in equine populations worldwide and has in general been considered a harmless virus due to its frequent detection in healthy horses (Bell et al., 2006; Marenzoni et al., 2010; Rushton et al., 2013; Torfason et al., 2008). More recently, however, it has been linked to the occurrence of equine multinodular pulmonary fibrosis (EMPF), a severe lung disease in the horse (Poth et al., 2009; Williams et al., 2007). Our research group has recently reported that EHV-5 was frequently detected in nasal swabs from individual horses followed over 1 year, but the presence of EHV-5 was not associated with clinical signs or poor athletic performance (Back et al., 2015). It has not been established whether the observed continuous shedding is due to chronic active infection, reactivation of latent infections, frequent reinfections or shedding of multiple strains of EHV-5. As EHV-5 is associated with non-clinical as well as severe clinical indications,
Table 1. EHV-5 gB genotypes detected within each clinical sample at the two sampling occasions 1 year apart

The individual horses are named with a letter followed by '1' (first sampling occasion) or '2' (second sampling occasion); in the strain designation, the number after the dot (1–3) illustrates the most (1) and least (3) common strain within each of the clinical specimens. The proportions of the trimmed and paired reads for the genotypes are specified as percentages.

<table>
<thead>
<tr>
<th>Horse (age)/sampling occasion</th>
<th>Strain</th>
<th>GenBank accession no.</th>
<th>Genotype (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>EMPF 1 (4 years)</td>
<td>EMPF1.1 KT313065</td>
<td>- -</td>
<td>64.7 -</td>
</tr>
<tr>
<td>EMPF 2</td>
<td>EMPF2.1 KT313066</td>
<td>- -</td>
<td>53.9 -</td>
</tr>
<tr>
<td>A1 (2 years)</td>
<td>A1.1 KT313040</td>
<td>- -</td>
<td>45.3 -</td>
</tr>
<tr>
<td></td>
<td>A1.2 KT313041</td>
<td>- -</td>
<td>18.1 -</td>
</tr>
<tr>
<td>A2</td>
<td>A2.1 KT313042</td>
<td>- -</td>
<td>50.4 -</td>
</tr>
<tr>
<td></td>
<td>A2.2 KT313043</td>
<td>- -</td>
<td>8.2 -</td>
</tr>
<tr>
<td></td>
<td>A2.3 KT313044</td>
<td>- -</td>
<td>6.6 -</td>
</tr>
<tr>
<td>R1 (5 years)</td>
<td>R1.1 KT313045</td>
<td>- -</td>
<td>57.7 -</td>
</tr>
<tr>
<td>R2</td>
<td>R2.1 KT313046</td>
<td>- -</td>
<td>56.5 -</td>
</tr>
<tr>
<td>U1 (3 years)</td>
<td>U1.1 KT313047</td>
<td>- -</td>
<td>56.2 -</td>
</tr>
<tr>
<td></td>
<td>U1.2 KT313048</td>
<td>- -</td>
<td>12.7 -</td>
</tr>
<tr>
<td>U2</td>
<td>U2.1 KT313049</td>
<td>- -</td>
<td>58.4 -</td>
</tr>
<tr>
<td></td>
<td>U2.2 KT313050</td>
<td>- -</td>
<td>10.7 -</td>
</tr>
<tr>
<td>V1 (2 years)</td>
<td>V1.1 KT313051</td>
<td>- -</td>
<td>53.1 -</td>
</tr>
<tr>
<td>V2</td>
<td>V2.1 KT313052</td>
<td>- -</td>
<td>58.6 -</td>
</tr>
<tr>
<td>W1 (2 years)</td>
<td>W1.1 KT313053</td>
<td>- -</td>
<td>68.3 -</td>
</tr>
<tr>
<td>W2</td>
<td>W2.1 KT313054</td>
<td>- -</td>
<td>67.4 -</td>
</tr>
<tr>
<td>X1 (2 years)</td>
<td>X1.1 KT313055</td>
<td>- -</td>
<td>55.7 -</td>
</tr>
<tr>
<td></td>
<td>X1.2 KT313056</td>
<td>- -</td>
<td>9.7 -</td>
</tr>
<tr>
<td>X2</td>
<td>X2.1 KT313057</td>
<td>- -</td>
<td>25.9 -</td>
</tr>
<tr>
<td></td>
<td>X2.2 KT313058</td>
<td>- -</td>
<td>33.5 -</td>
</tr>
<tr>
<td>Y1 (2 years)</td>
<td>Y1.1 KT313059</td>
<td>- -</td>
<td>37.5 -</td>
</tr>
<tr>
<td></td>
<td>Y1.2 KT313060</td>
<td>- -</td>
<td>14.0 -</td>
</tr>
<tr>
<td></td>
<td>Y1.3 KT313061</td>
<td>- -</td>
<td>6.3 -</td>
</tr>
<tr>
<td>Y2</td>
<td>Y2.1 KT313062</td>
<td>- -</td>
<td>71.7 -</td>
</tr>
<tr>
<td>Z1 (5 years)</td>
<td>Z1.1 KT313063</td>
<td>- -</td>
<td>59.5 -</td>
</tr>
<tr>
<td>Z2</td>
<td>Z2.1 KT313064</td>
<td>- -</td>
<td>59.7 -</td>
</tr>
</tbody>
</table>

The virus might act similarly to EBV, which has been associated with both subclinical infection and mononucleosis (Balfour et al., 2013), as well as the development of idiopathic pulmonary fibrosis (Mora et al., 2005; Vannela & Moore, 2008), multiple sclerosis (Mechelli et al., 2015), and cancers such as Hodgkin’s lymphoma, Burkitt’s lymphoma and nasopharyngeal carcinoma (Parkin, 2006) in humans. Diversity at the whole-genome level has been identified for EBV, suggesting that genetic variability may explain the variation in pathogenicity (Kwok et al., 2014). However, information about the genetic variability of EHV-5 is limited, thus preventing detailed examination of infections and pathogenicity in horses.

The first description of the complete EHV-5 genome was published recently (Wilkie et al., 2015). Previously, only few isolates had been described, with most of the work focusing on the glycoprotein B (gB) gene. The gB protein is essential for herpesvirus cell entry as well as spread between cells (Rey, 2006), by forming the core fusion machinery together with a heterodimer of glycoprotein H/L, which with additional accessory proteins fuse with the membrane of a cell (Chesnokova et al., 2014; Garcia et al., 2013; Ryckman et al., 2006). Moreover, gB has been suggested to be a major target region for neutralizing antibodies produced by the host (Kirchmeier et al., 2014; Neubauer et al., 1997). The genetic variation in segments of the EHV-5 gB gene has been studied in nasal secretions and PBMCs in mares and their foals (Bell et al., 2006), and in isolates from young and adult horses (Dunowska et al., 2000). Heterogeneity between individual horses was identified, but it was not possible to determine the presence of multiple strains in individual horses.

The aims of this study were to identify genetic divergence in the gB gene of EHV-5 in clinical specimens from horses in a single training yard and from one horse diagnosed with EMPF, and to investigate whether individual horses were infected with multiple strains. Additionally, the dynamic pattern of EHV-5 gB strains over a period of 1 year was investigated in this population of horses.

RESULTS

Analysis of the nucleotide sequences of the partial EHV-5 gB gene

Sequencing produced ~200 000–2 000 000 paired-end reads per individual sample; after quality trimming and merging of paired-end reads, 25 000–480 000 merged reads remained per sample. The replicate run of seven samples to confirm the next-generation sequencing (NGS) results presented highly similar nucleotide results (data not shown).

From the clinical specimens, 27 nucleotide sequences in total were identified in the partial gB gene of EHV-5, amongst which 11 unique sequence types could be distinguished. Some individual horses were infected by up to three different strains at the same time, as shown in Table 1. Alignment of the nucleotide sequences and the selected references (Fig. 1) highlights regions with high similarity as well as with high variability. All positions refer to GenBank accession number GQ325593 (our gB segment starts at nt 955), as this European sequence was more similar to the investigated nucleotide sequences than other available references. The nucleotides contain a furin cleavage site motif located at nt 331–342 of the gB gene. The region close to the putative cleavage site showed high variability with insertions or deletions of multiple nucleotides, especially at nt 312–378, whereas high stability was observed in the first part and the last part of the gB segment, foremost at nt 1–58 and 385–460.
Variations of infections of EHV-5 in individual horses

http://jgv.microbiologyresearch.org
Fig. 1. Alignment of partial EHV-5 gB gene nucleotide sequences (i.e. strains) detected in clinical specimens from eight healthy horses sampled two times 1 year apart and one horse diagnosed with EMPF. The strains are clustered within four suggested genotypes of the EHV-5 gB gene; the codes are given for each strain included in the genotypes, except for genotype I, which included 13 identical sequences. All positions refer to the reference strain (GenBank accession number GQ325593), where our gB segment starts at nt 955.

Phylogenetic relationship between the strains of EHV-5

Alignment of the nucleotide sequences extracted from the NGS data with a prevalence of >5 % in each clinical specimen, together with the selected reference sequences, was used to reconstruct a maximum-likelihood phylogenetic tree (Fig. 2). Only branches with bootstrap values >80 % were accepted, which in this population of horses resulted in the detection of four clusters, here referred to as genotypes I–IV. In the genotypes, different nucleotide sequences within each clinical specimen (i.e. strains) were distinguished.

For the studied segment of ~460 nt, the diversity between strains within the four divergent genotypes (Figs 1 and 2) was <2 % [0 (I), 2 (II), 0–3 (III) and 0–8 nt (IV)]. Between the genotypes, apart from genotypes I and II, which diverged by only 9 nt, the divergence ranged from 19 to 81 nt, with genotype IV diverging most (by 81 nt between genotypes III and IV). However, despite the low diversity in nucleotide positions between genotypes I and II, the bootstrapping value was 84 and the strains classified within genotype II generated an amino acid sequence that clearly diverged from the amino acid sequence encoded by the nucleotides of the strains included in genotype I (Fig. 3). Furthermore, the EHV-5 strains detected in horse A were the sole members of the new genotype II of EHV-5. The EMPF case was infected with only one strain, classified as genotype III. Two of the four identified genotypes (genotypes I and II) did

Fig. 2. Phylogenetic trees based on DNA sequences (left) and translated amino acid (AA) sequences (right) showing four suggested genotypes (I–IV) of the EHV-5 gB gene segment identified in nine Swedish horses. The individual strains and sequences are shown within genotypes. Only strains that included >5 % of the reads in each clinical specimen were incorporated in the tree. The different clinical specimens are coded with a unique letter for each horse followed by ‘1’ (first sampling occasion) or ‘2’ (second sampling occasion); the number after the dot (1–3) illustrates the most (1) and least (3) common strain within each of the clinical specimens. In total, 1000 bootstrap replicates were performed for each tree and obtained values are shown where bootstrap support is >80 %. Bars, 2 % of the total variation.
Fig. 3. Amino acid sequences of the gB gene from different strains of EHV-5 detected in 18 samples from nine Swedish horses. The amino acid sequences from the different strains detected in all the clinical specimens are aligned with the amino acids for nucleotide sequences of the EHV-5 gB gene obtained from GenBank (named by their accession numbers on the left). The designations on the left indicate the names of each strain obtained from the clinical specimens, which are coded with a unique letter for each horse followed by ‘1’ (first sampling occasion) or ‘2’ (second sampling occasion); the number after the dot (1–3) illustrates the most (1) and least (3) common strain within each of the clinical specimens. The genotype for each strain is given in parentheses. This gB segment starts at aa 419 of GenBank accession number GQ325593. The putative cleavage site ‘R-K-R-R’ at aa 111–114 is marked with a grey box.

http://jgv.microbiologyresearch.org 173
Impact of the nucleotide variations on amino acid sequences

The diversity of the nucleotides in the sequences resulted in both non-synonymous and synonymous mutations in the gB gene segment. The start position of our gB segment corresponds to aa 320 of GenBank accession number GQ325593. The putative cleavage site (R-K-R-R) (Sorem & Longnecker, 2009) at aa 111–114 (Fig. 3) was fully conserved for all strains. However, the high variability in the nucleotides next to the furin cleavage site resulted in several non-synonymous mutations (Fig. 3). Five separate amino acid sequences were generated by translation of the nucleotide sequences of all the studied strains belonging to the four genotypes (Fig. 2). The 13 identical nucleotide sequences classified within genotype I resulted in one amino acid sequence, which did not correspond to any of the others. The strains A1.2 and A2.3 from the clinical specimens of horse A, classified as genotype II, resulted in unique amino acid sequences that differed at six positions (aa 94, 107–109, 117 and 122) compared with the amino acid sequence in genotype I. The nine strains classified as genotype III also gave rise to one unique amino acid sequence. Genotype IV encoded two different amino acid sequences, where X2.1 resulted in one amino acid sequence that differed in three positions (aa 119–121) from the amino acid sequence generated by Y1.2 and Y2.1 (Figs 2 and 3). Close to the cleavage site there are amino acid residues that might be used as characteristics for the gB genotypes. For example, the G-T at aa 118–119 might be distinctive for genotype III and D-T-G at aa 115–117 for genotype IV.

Dynamics of EHV-5 gB gene sequences over 1 year

Different patterns of EHV-5 were identified in individual horses over 1 year, based on the detection or the abundance of different strains, translated amino acids and classification into genotypes.

Phylogenetic analysis of the partial sequence of the gB gene showed that the sequence can be conserved in an individual horse over time. In horses R, V, W and Z, identical genotype I sequences were detected on both sampling occasions.

In two of the horses (U and A) the proportions between the strains changed over 1 year. Horse U shed two strains, one from genotype I and the other from genotype III, at both sampling occasions, but the proportions shifted over time (Table 1), which indicates a viral dynamic between the infecting strains and the host. In the clinical specimens from horse U, 56.2 % of the nucleotide sequences at the first sampling occasion coded for a gB strain classified as genotype I and 12.7 % of the gB strains were classified as genotype III. One year later the genotype I strain was reduced to 10.7 %, whereas 58.4 % of the gB strains were of genotype III, differing by only 3 nt from the sequence obtained in the first sampling. In horse A, the proportions of the strains (genotypes II and III), as illustrated in Table 1, also varied over time, but not as strikingly as for horse U.

Appearance and loss of different strains over time was observed in two of the infected horses (X and Y). Horse X was first identified with strains of genotypes I and III; 1 year later an identical sequence from genotype I was detected together with a new strain classified as genotype IV and the genotype III strain was absent. Horse Y first harboured three strains, which were classified in three different genotypes (I, III and IV). One year later, only genotype IV could be detected.

DISCUSSION

In a small group of cohabiting horses, we identified evidence that horses can be infected with multiple strains of EHV-5 on the same sampling occasion. Amongst the 27 partial nucleotide sequences of the gB gene obtained in this study, 11 different nucleotide sequence types could be distinguished, which by subsequent phylogenetic analysis could be classified in four different genotypes (I–IV). Amongst these strains, two identified genotypes (I and II) did not match with any previously described strains. The different strains of the EHV-5 gB gene present in nine horses were shown to be stable over time in some horses, whereas other horses had a change in strains and/or genotypes over time.

Strains and nucleotide sequences of the partial EHV-5 gene

In this group of horses, the 27 identified EHV-5 gB gene strains have 98–100 % nucleotide similarities within the same genotypes and differ by 4.4–16.2 % between different genotypes. Similarly, genetic divergence based on the gB gene has been described both within and between the four detected genotypes of the betaherpesvirus human cytomegalovirus (CMV; human herpesvirus 5) (Chantararaphonkun & Bhattacharosol, 2007; Chou & Dennison, 1991).

The segment that might contain a furin cleavage site at aa 111–114, evident in Figs 1 and 3, has been shown to be a cleavage site of the gB gene in both EHV-5 (Holloway et al., 1999) and EBV (Sorem & Longnecker, 2009). Moreover, our results of a conserved cleavage site in all 27 sequences are in agreement with another study which showed the cleavage site of the gB gene to be essential in EBV for the cell–cell fusion of the virus (Sorem & Longnecker, 2009). The alignment of the 27 nucleotide sequences in our study revealed a high stability at the 5’ and 3’ ends of the nucleotide gB segment (nt 1–58 and 385–460), whereas a high variability, including non-
synonymous mutations, was seen closer to the putative cleavage site (nt 312–378) (Fig. 1). Similar patterns have been identified in a study of the gB gene of EHV-5 and seven other herpesviruses (amongst them EHV-2), where insertions and deletions mainly occurred close to the cleavage site (Holloway et al., 1999).

**Phylogenetic relationship between the strains and amino acids of EHV-5 gB**

As illustrated in Fig. 3, the amino acid sequences within each genotype appear to be strongly conserved, but between genotypes they differ considerably. We speculate that this might affect the structural or conformational constraints on this gB segment, and thereby possibly cause genotypes to interact differently with the host.

We cannot explain why genotype II was not identified in other horses from the same training yard, but as only a limited number of horses were included in this study, prevalence of different genotypes could not be assessed. The strains of gB genotypes I and II have not been described previously and might be circulating in Sweden specifically, as differences in the geographical distribution of different genotypes have been described for EBV (Abdel-Hamid et al., 1992). It might, however, also reflect the limited genetic information of EHV-5. The selected reference nucleotide sequences of the gB gene (Figs 1–3) are strains originating from three different continents (North America, Australia and Europe). The two European isolates (genotypes III and IV) originated from the horse population in Iceland, which has been isolated from direct contact with other horse populations for >1000 years (Torfason et al., 2008); Icelandic horses, however, frequently exported from Iceland to the Scandinavian countries, but rarely in direct contact with professionally trained standardbreds.

**Dynamic pattern of the EHV-5 gB gene over time**

Three different patterns of EHV-5 infection (based on the partial gB gene) over time were detected in this study (Table 1, Fig. 2). Infection with a single identical strain on both sampling occasions (four horses) illustrates the high stability of the viral strains over time. This mimics findings in a case study on human herpesvirus 6 (a betaherpesvirus), where the same strain (based on the gB gene) was detected in an individual patient at two different time points (Achour et al., 2008).

The second most common pattern, i.e. infection with two strains from different genotypes in varying proportions (two horses), suggests, even if it needs to be further investigated, that the influence of host–virus–environment interaction may be a possible explanation for the viability of different strains.

The third pattern, i.e. emergence or disappearance of strains from different genotypes in a horse over time (two horses), highlights that horses can be infected by several strains from different genotypes and that their detection might fluctuate over time. The fact that the strains showed a high stability over time, together with the changes described in horse X, suggest that new infection or reactivation of EHV-5 strains occurs. Further studies are needed to investigate whether this stability is mainly related to host, virus and/or environmental factors.

**Clinical link**

The four different genotypes of EHV-5 based on the partial gB gene could not be associated with disease causality as the study was designed to investigate apparently healthy horses in professional training. Classification based on the gB gene has been performed previously for CMV. However, association with clinical disease and different gB genotypes has yet to be established (Humar et al., 2003; Sarcinella et al., 2002).

The strain obtained from the horse diagnosed with EMPF was classified as genotype III, with only 1 nt divergence from the reference nucleotide sequence (GenBank accession number KC715730) obtained from an EMPF case in the USA (Williams et al., 2013) (Figs 1 and 2). However, other healthy horses were also infected with strains of the same genotype. Nonetheless, it cannot be excluded that genotypes or combinations of strains might be of clinical importance. These results indicate EHV-5 genotype III to be of interest for further studies of EMPF, although conclusions based on a single case and a segment of one gene must be drawn with great caution. Moreover, as our EMPF case was infected with both EHV-5 and asinine herpesvirus 5, the latter virus may have contributed to, or even been responsible for, the clinical outcome. In comparison, human infection by multiple strains of CMV has been suggested to cause more severe clinical manifestations than infections with a single variant in immunocompromised individuals (Coaquette et al., 2004). Although EBV has been associated with idiopathic pulmonary fibrosis (Tang et al., 2003) and with several types of cancer (Parkin, 2006), the pathogenesis is still unknown. To date, genomic diversity of EBV has been identified, but whether the pathogenicity is related to strain variability remains to be clarified (Kwok et al., 2014; Palser et al., 2015).

The eight healthy horses were at the same training yard and presumably under the same infection pressure. Therefore, the different patterns of viral interaction with the host suggests that individual characteristics, such as strain and the genetic background or the immune responses to EHV-5 in the host, may play an important role in the viral dynamic over time.

**Future perspectives**

This study presents important observations to improve our understanding of the complex infection dynamics between strains of gammaherpesvirus and their hosts. In the future,
the NGS method described for the partial gB gene of EHV-5 can be used in a larger number of horses (both healthy and diagnosed with EMPF) worldwide to further investigate the geographical distribution and the possible clinical importance of the genotypes presented. This could also help define the interaction between EHV-5 strains infecting individual horses in order to appreciate the clinical consequences of simultaneous infection by different strains. Further studies should also focus on the virological aspects, such as from where and how the genotypes emerge and evolve, whether interaction such as recombination occurs between strains, and if selective latency or clearance commonly occurs for the gammaherpesviruses.

CONCLUSIONS
The introduction of NGS has created new opportunities to study viral diversity and evolution. In this study, 27 partial gB gene strains representing 11 unique sequence types were detected from eight horses from one training yard sampled twice 1 year apart and from one horse diagnosed with EMPF. These strains could be classified within four highly conserved but diverse genotypes (I–IV) of EHV-5, based on nucleotide and amino acid sequences, where two of the identified genotypes did not match with any previously described strains. Healthy horses were identified with up to three different genotypes of EHV-5 at the same time. Some of the horses, mainly those infected with strains from genotype I, maintained identical gB gene sequences over 1 year. In other horses, the presence of genotypes, the sequence variant within genotypes or the proportions of the different strains changed during the study period. The EMPF case was infected with only one strain of genotype III, whereas healthy horses in which genotype III was detected were co-infected with multiple strains. The study provides insights into the genetic variation and dynamics of EHV-5 infections in the horse and suggests areas for future research.

METHODS
Samples. Nasal swabs were sampled from actively racing standardbred trotters in a longitudinal study on a monthly basis over 13 months. The samples were analysed by quantitative (q)PCR assays to detect and quantify the viral load of EHV-2 and EHV-5, as described previously (Back et al., 2015). From this longitudinal study, eight horses were selected for further genetic studies. They were 2–5 years of age, from the same training yard, and identified with high viral load of EHV-5 at the beginning and at the end of the study (11–12 months between samples). At the two sampling time points, all horses were racing and training according to their routine schedule, and were free of clinical signs of respiratory disease, such as fever, cough or nasal discharge. In addition, two samples [pre-mortem lung biopsy (EMPF1) and post-mortem lung tissue (EMPF2)] from a horse diagnosed with EMPF associated with EHV-5 and AHV-5 (Back et al., 2012) were included in the study.

DNA extraction. The lung biopsy was homogenized in TE-buffer (10 mM Tris, 1 mM EDTA, pH 8.0). A fresh section of the lung sample from necropsy was swabbed with a sterile swab and suspended in tubes with TE-buffer. Viral DNA was extracted from both lung samples and nasal swabs (E-swab buffer) by adding 10 μl proteinase K (Sigma-Aldrich) to 90 μl buffer. Thereafter, nucleic acid extraction was performed in a Magnatrix 8000+ extraction robot (NorDiag), using the NorDiag Vet Viral NA extraction kit according to the manufacturer’s instructions.

Amplicon primers. Based on all available sequences in GenBank for EHV-5 gB, a highly variable region flanked by conserved regions that could serve as primer binding sites was selected. The primers were designed to be specific for EHV-5 and not to amplify EHV-2. Primer sequences and their positions in the gB gene of EHV-5 (GenBank accession number NC_026421) were: F 5’-AACGACAGAAAGC-TACCAT-3’ (nt 1090–1110) and R 5’-ACACGTGCTCAAGACCC-3’ (nt 1572–1599). The rationale for this design was to amplify a region that could serve as a suitable fingerprint for different strains. The variable region contains a putative protease cleavage site (Sorens & Longnecker, 2009) and an analogous region has been used for genotyping of CMV (Chou & Dennison, 1991). Primers were designed using the PrimerQuest design tool on the webpage of Integrated DNA Technologies (http://www.idtdna.com/primerquest/home/index) and purchased from Eurofins MWG Operon.

Amplicon PCR and library preparation. The amplicon PCR and library preparation were performed according to the Illumina protocol ‘16S Metagenomic Sequencing Library Preparation’ with minor modifications, and are described briefly below.

Amplicon PCR was performed using HiFi Hot Start ReadyMix (KAPA Biosystems) with primer concentrations of 200 nM. Thermal cycling started with an initial activation step at 95 °C for 3 min and proceeded with 35-cycle three-step cycling at 98 °C for 20 s, 55 °C for 30 s and 72 °C for 30 s, thereafter 72 °C for 5 min and then hold at 4 °C. The presence of amplification products of the expected size (~ 500 bp) was confirmed by gel electrophoresis.

The amplicons were thereafter purified from free primers and primer dimers by a clean-up step using AMPure XP beads. Dual indices and Illumina sequencing adapters were added in an eight-cycle PCR and a second clean-up was performed according to the manufacturer’s instructions. Quantification of the libraries was performed using a Bioanalyzer for verification of the size and a Qubit dsDNA high-sensitivity assay kit (Life Technologies) to determine the concentration of each library. Libraries were diluted with Tris (pH 8.5) to reach a concentration of 4 nM each and were then pooled.

Sequencing. The library pool was denatured with NaOH and further diluted with hybridization buffer (included in the MiSeq Reagent kit) to a final concentration of 8 pM and spiked with 5 % PhiX for diversity. Paired-end sequencing of the libraries was performed with a MiSeq Reagent kit version 3 with 600 cycles on the MiSeq instrument (Illumina). DNA extraction, amplicon PCR and sequencing of the partial gB gene were repeated for seven of the samples in order to verify the results obtained in the first run.

Data analysis. The sequence reads of the gB gene segment were quality trimmed from the 5’ end using a 4 bp Q20 sliding window in Trimmomatic 0.32 (Bolger et al., 2014) and the read pairs were connected using COPE 1.1.2 (Liu et al., 2012) requiring a minimum 50 bp overlap. To remove sequencing and PCR artefacts, sequence variants that made up <5 % of the final data were filtered away. The final sequence types were aligned using Muscle (Edgar, 2004) in MEGA 5.10 (Tamura et al., 2011) by codons. Maximum-likelihood phylogenetic trees were reconstructed in MEGA using the HKY substitution model (Hasegawa et al., 1985) with gamma distributed rate variation amongst sites for DNA and the JTT model (Jones et al., 1992) for the translated amino acid sequences. The models were
selected to minimize the Bayesian information criterion. In total, 1000 bootstrap replicates were performed for each tree.

The following selected references of the EHV-5 gB gene were used in the phylogenetic analysis: GenBank accession numbers AF9050671 (Holloway et al., 1999), GQ325592 and GQ325593 (Thorsteinsdottir et al., 2010), KC715730, KC715731 and KC715732 (Williams et al., 2013), and NC_026421 (Wilkie et al., 2015).

The different clinical specimens were coded with a unique letter for each horse followed by ‘1’ (first sampling occasion) or ‘2’ (second sampling occasion); the last number (1–3, after a dot) indicates the most (1) and least (3) common strain within each of the clinical specimens.

ACKNOWLEDGEMENTS

This study was funded by the Swedish–Norwegian Foundation for Equine Research project number H1247140.

REFERENCES


Chesnokova, L. S., Ahuja, M. K. & Hutt-Fletcher, L. M. (2014). Epstein-Barr virus glycoprotein gB and gHgL can mediate fusion and entry in trans, and heat can act as a partial surrogate for gHgL and trigger a conformational change in gB. J Virol 88, 12193–12201.


