Analysis of the genetic diversity of ovine herpesvirus 2 in samples from livestock with malignant catarrhal fever

George C. Russell[a*], Sandra F. Scholes[b], David F. Twomey[c], Ann E. Courtenay[d], Dawn M. Grant[a], Bruce Lamond[a], David Norris[a$], Kim Willoughby[e], David M. Haig[f], James P. Stewart[g]

Author affiliations:

[a] *Moredun Research Institute, Pentlands Science Park, Penicuik, EH26 0PZ, UK*
[b] *Animal Health and Veterinary Laboratories Agency (AHVLA), Pentlands Science Park, Midlothian, UK*
[c] *AHVLA Starcross Regional Laboratory, Staplake Mount, Starcross, Devon, EX6 8PE, UK*
[d] *AHVLA Surveillance Centre, Division of Livestock Health and Welfare, School of Veterinary Science, Leahurst Campus, University of Liverpool, Wirral, CH64 7TE, UK*
[e] *Moredun Scientific, Pentlands Science Park, Penicuik, EH26 0PZ, UK*
[f] *School of Veterinary Medicine and Science, University of Nottingham, Sutton Bonington, Leicestershire LE12 5RD, UK*
[g] *Department of Infection Biology, Institute of Infection and Global Health, University of Liverpool, Liverpool Science Park IC2, L3 5RF, UK*
[$] Present address: Scottish Lowlands Forest District, Five Sisters House, Five Sisters Business Park, West Calder, EH55 8PN, UK

* Corresponding author. Tel.:+44 131 445 6149; fax: +44 131 445 6111. *E-mail address:* george.russell@moredun.ac.uk

The final publication is available at www.journals.elsevier.com

Abstract

In order to better define virus isolates from animals with malignant catarrhal fever (MCF), segments of three genes of ovine herpesvirus-2 were amplified from diagnostic samples representing MCF cases with a range of clinical presentations in cattle, including head and eye, alimentary and neurological. The variation within each gene segment was estimated by DNA sequencing, which confirmed that the newly-annotated Ov9.5 gene was significantly more polymorphic than either of the other loci tested (segments of ORF50 and ORF75), with alleles that differed at over 60% of nucleotide positions. Despite this, the nine Ov9.5 alleles characterized had identical predicted splicing patterns and could be translated into Ov9.5 polypeptides with at least 49% amino acid identity. This multi-locus approach has potential for use in epidemiological studies and in charactering chains of infection. However there was no association between specific variants of OvHV-2 and the clinical/ pathological presentation of MCF in the cattle analysed.

1. Introduction

Gammaherpesviruses (γHV) can replicate in and latently infect lymphoid cell types and are associated with lymphoproliferative diseases and tumours (Ackermann, 2006; Barton et al., 2011; Nash et al. 2001). They share similarities of size, sequence and genome organization. Evolution within the subfamily *Gammaherpesvirinae* is reflected in the colinearity of most γHV genome sequences, with approximately 50 genes considered to be conserved across the group of fully-sequenced γHV (McGeoch et al., 2005). The genes of most γHV are named with respect to the genome of the prototype virus of this family, herpesvirus saimiri (HVS), with 75 open reading frames (ORF) numbered from the left of the genome (Albrecht et al., 1992). These genes encode the components of the virus capsid, tegument and envelope; proteins involved in replication of the virus genome; regulatory proteins that control the lytic cycle and latency; and proteins that manipulate the metabolism of the host to benefit viral replication and persistence. Comparison of HVS with other γHV shows that each virus carries a complement of virus genes that do not have homologues in the HVS genome (McGeoch, 2001). These genes are termed unique genes, although they may be shared within groups of related viruses. They are annotated by a prefix letter specific to the virus and are numbered from the left of the genome. Unique genes in γHV have been identified by bioinformatic analysis, comparative genomics and by analysis of cDNA or protein sequences (Albrecht et al., 1992; Coulter et al., 2001; Hart et al., 2007; Hughes et al 2010; Mills et al., 2003; Russell et al., 2013).

Several ruminant gammaherpesviruses of the *Macavirus* genus (Davison et al., 2009) are associated with the lymphoproliferative disease malignant catarrhal fever (MCF). The best studied of these are ovine herpesvirus 2 (OvHV-2) and alcelaphine herpesvirus 1 (AlHV-1). MCF occurs when virus shed in the mucous secretions of reservoir host species (sheep for OvHV-2 and wildebeest for AlHV-1), which are infected efficiently and without obvious

clinical signs, infects susceptible species such as cattle, bison, deer and pigs (Russell et al., 2009). MCF is an often-fatal systemic disease that is generally sporadic, affecting individual animals within a group, but can occasionally cause losses of up to 40% of a herd (Russell et al., 2009; World Organisation for Animal Health (OIE), 2008). The clinical signs of MCF can be varied and several distinct disease presentations have been described, including peracute, head and eye, alimentary and neurological (World Organisation for Animal Health (OIE), 2008). Most MCF cases in cattle present with fever, depression and lymphadenopathy; while the common head and eye form is further characterized by nasal and ocular secretions, corneal opacity, skin lesions and multifocal necrotic lesions of the gums, tongue and palate (Otter et al., 2002). In the peracute form, sudden death may occur, though depression followed by diarrhoea, with death occurring within a few days has also been reported. In the alimentary form, haemorrhagic diarrhoea may also be found (Holliman et al., 2007), while nervous signs, ataxia, and blindness have been reported in the neurological form (Mitchell and Scholes, 2009). Although MCF is generally considered a disease with a case fatality rate approaching 100 %, reports of recovery from clinical MCF and chronic infection have also been published (Milne and Reid, 1990; O'Toole et al., 1997; Penny, 1998; Twomey et al., 2002).

The genomes of AlHV-1(Ensser et al., 1997) and OvHV-2 (Hart et al., 2007; Taus et al., 2007; Jayawardane et al., 2008) have been fully sequenced, demonstrating conservation of γHV genome structure and possession of a similar complement of unique genes (Russell et al., 2009). Recently, analysis of virus gene expression in AlHV-1 infected cells revealed the presence of a novel spliced gene (Russell et al., 2013) that encoded a secreted glycoprotein. This gene, termed A9.5, had not been previously predicted because of the small size of the coding exons and because no similar protein was present on any database. Predicted segments of protein sequence with similarity to these exons were found at the same position in OvHV-

2, such that the homologous gene Ov9.5 could be identified despite having only 33 % translated amino acid identity to the A9.5 polypeptide. Notably, the two published OvHV-2 genome sequences contained distinct Ov9.5 genes, which had equivalent positions and predicted splicing patterns but shared only 60 % nucleotide identity (Russell et al., 2013). The low degree of identity between the two alleles of Ov9.5 suggests distinct histories, selective regimes or functions (Russell et al., 2013). It may therefore be relevant that one Ov9.5 sequence was from a clinical case of MCF in a British cow while the other was obtained from OvHV-2 virions from sheep nasal secretions in the USA (Hart et al., 2007; Taus et al., 2007).

In order to compare sequence variation within the Ov9.5 gene with other loci in OvHV-2, two additional genes were selected for comparative sequence analysis. These were: a segment of the ORF75 gene, encoding the virion enzyme formylglycineamide ribotide amidotransferase (FGARAT), routinely used as the target of a diagnostic PCR assay for OvHV-2 (Baxter et al., 1993); and a segment of the ORF50 gene, encoding RTA, a transcription factor involved in lytic cyle activation, selected because of its important role in virus regulation. The ORF73 (latency-associated nuclear antigen) locus was also considered as a target for PCR but initial studies showed that this gene could not be reliably amplified from clinical case material despite the use of published primer sets (Coulter and Reid, 2002) or newly designed nested primers (GC Russell, unpublished data). This is likely to be a consequence of the size and repetitive nature of this gene combined with the relatively low viral load in the samples used for analysis.

In this paper we analyse genetic variation in OvHV-2 from clinical case samples by looking at three loci to address the hypothesis that the highly polymorphic Ov9.5 gene is a useful epidemiological marker of OvHV-2 strain variation. We also examine the possibility

115 that distinct strains of OvHV-2 may be responsible for different presentations of MCF in
116 cattle.
117

2 Materials and methods

2.1. Clinical samples

All DNA samples used in this work were extracted from material submitted to Moredun Research Institute for PCR-based testing in support of a diagnosis of MCF. Prior to 2006, DNA samples were purified from peripheral blood mononuclear cells (PBMC) or from tissues from MCF-suspect cases by a standard phenol-based method (Sambrook et al. 1989), while samples collected since 2006 were purified by a column-based method that did not use organic solvents (DNeasy mini, Qiagen, Crawley, UK).

OvHV-2 positive samples were selected according to reported clinical signs, representing a range of clinical disease presentations as documented in Table 1 and summarised as follows: alimentary, 11 cases, including herds A and E from the report by Holliman et al. (2007); head & eye, four recovered cases described by Twomey et al. (2002); neurological, two cases reported by Mitchell and Scholes (2009). In addition, seven DNA samples from two MCF outbreaks (head and eye form) in different parts of the UK (including one sample from an in-contact sheep) and five samples from sporadic MCF cases (where no other MCF-positive sample was submitted from the same source within at least a month) were analysed (Table 1). As a positive control for the PCR and sequencing reactions, DNA from the cell line BJ1035 was tested with all primer sets. The OvHV-2 genome from this cell line was previously sequenced (Hart et al., 2007).

2.2. Amplification and sequencing of selected OvHV-2 gene segments

Primers for nested PCR of the Ov9.5 gene were designed to target conserved areas flanking the predicted coding region, based on the available sequences of OvHV-2 (Hart et al., 2007; Taus et al., 2007; Fig. 1; supplementary Fig. S1). Primers for nested amplification of ORF50

were chosen within exon 2, to amplify a fragment of 400-500 base pairs for sequence analysis, and primers for amplification of ORF75 were as described previously (Baxter et al., 1993). Primer pairs were designed using Primer3 (www .bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi; Untergasser et al., 2007) and are detailed in Table 2.

ORF50, ORF75 and Ov9.5 gene segments were amplified from genomic DNA samples in which OvHV-2 DNA had been detected by diagnostic nested PCR (Baxter et al. 1993) or real-time PCR (Hussy et al., 2001). For each gene, first-round amplification was performed in 25 µl reactions using 1 unit KOD Hot Start DNA polymerase (Merck, Feltham, UK), 50-100 ng of genomic DNA, and 5 pmol each of the appropriate first round primer set. Amplification reactions consisted of a denaturation/activation step at 94 °C for 30s; 30 cycles of 94 °C for 30s, 55 °C for 30 s and 68 °C for 60 s; and a final extension step at 68 °C for 5 minutes. Aliquots of 2 µl from each first round PCR were then used as template in second round PCR amplifications using the same enzyme, buffer and PCR conditions but with 10 pmol per reaction of the appropriate nested primers. The nested PCR products were analysed by agarose gel electrophoresis, stained with SYBR®Safe DNA Gel Stain (Life Technologies, Paisley, UK) and visualized by UV transillumination before purification (QIAamp PCR purification system). PCR product concentrations were estimated after purification using a Nanodrop spectrophotometer (Labtech, Uckfield, UK). Approximately 300 ng of each PCR product was submitted for bidirectional nucleotide sequencing by Eurofins MWG Operon (Ebersberg, Germany), using the internal PCR primers as the sequencing primers. Electropherograms from each pair of sequencing reactions were assembled to produce sample consensus sequences for each gene segment amplified.

To confirm that direct sequencing of PCR products produced an accurate representation of the target sequence *in vivo*, PCR products of Ov9.5 from 5 samples were cloned into pGEM-T-Easy (Promega, Southampton, UK) and at least three clones

166 representing each PCR product were sequenced. The PCR products that were sequenced after
167 cloning are indicated in Table 1.

168 *2.3. Bioinformatics*

169 Unless otherwise indicated all DNA sequence analysis was done using DNASTAR
170 Lasergene software (V8.0 and above; www.DNASTAR.com). DNA sequence information
171 from each amplicon was assembled using the SEQMAN program and consensus sequences
172 representing the region flanked by the inner nested PCR primers were derived. Any DNA
173 sample that did not give good quality sequence traces on both strands for the entire region
174 was discarded.

175 The consensus sequences for each sample and locus were aligned using Lasergene
176 MEGALIGN and MAFFT (Katoh and Standley, 2013;
177 http://mafft.cbrc.jp/alignment/software/). For Ov9.5, the positions of introns and exons were
178 defined according to the annotation of Ov9.5*01 (Russell et al., 2013) and conservation of
179 splice donor and acceptor sequences was confirmed by visual inspection, while the
180 conservation of a continuous Ov9.5 open reading frame was confirmed by *in silico* generation
181 of predicted Ov9.5 cDNA sequences and their translation. Phylogenetic and evolutionary
182 analysis of all sequences was done by maximum likelihood methods using the programs
183 TOPALi (Milne et al., 2009; www.topali.org) and MEGA (version 5 or above; Tamura et al.,
184 2011; megasoftware.net).

185 *2.4 Nucleotide sequence accession numbers*

186 Nucleotide sequences of the gene fragments amplified in this work have been
187 submitted to the European Nucleotide Archive (ENA; www.ebi.ac.uk/ena) and have been
188 assigned accession numbers as follows.

For Ov9.5: Ov9.5*0101, HG813097; Ov9.5*0201, HG813098; Ov9.5*0202, HG813099; Ov9.5*0203, HG813100; Ov9.5*0301, HG813102; Ov9.5*0401, HG813101; Ov9.5*0501, HG813103; Ov9.5*0502, HG813104; Ov9.5*0503, HG813105.

For OvHV-2 ORF50: ORF50*0101, HG813085; ORF50*0102, HG813086; ORF50*0103, HG813087; ORF50*0201, HG813088; ORF50*0301, HG813089; ORF50*0401, HG813090; ORF50*0501, HG813091.

For OvHV-2 ORF75: ORF75*0101, HG813093; ORF75*0102, HG813094; ORF75*0201, HG813095; ORF75*0301, HG813096.

3. Results

3.1 Amplification and sequencing of OvHV-2 loci from MCF case samples

Amplification of the ORF75 gene segment routinely used for OvHV-2 diagnostic PCR used the published primer sequences (Baxter et al., 1993); whilst the ORF50 gene segment was amplified using novel primers specifically designed for this work. The sequences of all primers used here are given in Table 2.

To allow conserved primer sites flanking the Ov9.5 gene to be identified, the intergenic region between the annotated Ov9 and Ov10 genes in the sequenced OvHV-2 isolates from the UK and US (Hart et al., 2007; Taus et al., 2007) was aligned (Fig. 1; supplementary Fig. S1). This demonstrated that while the predicted Ov9.5 gene was divergent between the two sequenced virus isolates, with about 60 % nucleotide identity, the intergenic regions flanking Ov9.5 (approximately 150 bases at the left end and 600 bases at the right end; Fig. 1) were well-conserved, with greater than 90 % nucleotide identity. The

211 Ov9.5-specific primers in Table 2 were placed in conserved flanking regions to allow
212 amplification of the predicted Ov9.5 coding sequence for studies of genetic variation.

213 Amplification of the OvHV-2 ORF50, ORF75 and Ov9.5 gene segments was
214 attempted in DNA samples from 46 OvHV-2 infected animals, including different host
215 species, MCF presentations, sporadic cases and outbreak samples. Despite the use of nested
216 PCR, not all samples tested yielded a single clear band of the expected size for each
217 amplification reaction. This is likely to be a reflection of the age or poor quality of some of
218 the DNA samples (particularly those that were purified by phenol extraction) leading to
219 degradation of the DNA or reduced PCR performance but it is also possible that variation
220 within the primer sites could account for the lack of amplification in some cases. A number
221 of the samples selected derived from published work describing unusual clinical
222 presentations of MCF (Holliman et al., 2007; Mitchell and Scholes, 2009; Twomey et al.,
223 2002) and some of these were over ten years old. Assays of failed samples were repeated at
224 least once for confirmation. Purified PCR products were subject to DNA sequencing using
225 the inner nested primers from each locus (Table 2).

226 Samples from five animals were also analysed by cloning of the Ov9.5 PCR product
227 and subsequent sequencing of multiple clones for each sample (Table 1). This confirmed that
228 the consensus sequences obtained by direct sequencing of PCR products were identical to the
229 sequences of cloned PCR products. Among these Ov9.5 samples, we obtained ten clones that
230 had been amplified with the outer L1 and R1 primers and 5 clones amplified with the L2-R1
231 primers (Table 2). These contained the expected sequences at internal primer sites and
232 showed the same conservation of the region outside the Ov9.5 coding region as the published
233 genome sequences. These observations support the view that the degree of conservation
234 around the primer sites, even between highly divergent alleles of Ov9.5, is such that our
235 nested PCR protocol should be productive in the vast majority of cases.

3.2 Gene sequence variation in samples from MCF cases

ORF50 sequences were obtained from 25 samples, representing seven different alleles (Table 3). The majority allele was found in 11 samples and was designated ORF50*0101. Two smaller groups of samples had sequences that differed by single synonymous nucleotide substitutions from ORF50*0101 and were called ORF50*0102 and ORF50*0103. Four other ORF50 alleles encoded different polypeptide sequences: ORF50*0201 and ORF50*0301, which differed from the majority allele by single amino acid substitutions; and ORF50*0401 and ORF50*0501, each of which had 3 amino acid substitutions compared with ORF50*0101 (Table 3). The sequenced OvHV-2 genomes encoded ORF50*0401 (Hart et al., 2007), analysed here as the positive control sample BJ1035, and an allele that differed from ORF50*0101 by two amino acid substitutions, which was not found in our sample set (designated ORF50*0601 by us; Taus et al., 2007).

ORF75 sequences were obtained from 22 samples. These sequences represented 4 alleles, which differed at up to two nucleotide positions (Table 3). The majority allele was found in 18 samples and was identical to the ORF75 fragment sequence in one of the two published OvHV-2 genome sequences (Taus et al., 2007). This allele was named ORF75*0101. The remaining ORF75 alleles differed by a synonymous nucleotide substitution in allele ORF75*0102; and by single amino acid changes in alleles ORF75*0201 and ORF75*0301 (Table 3).

Nine different Ov9.5 sequences were obtained from 39 samples (Table 3) representing 5 major clades of alleles that shared 55-99% nucleotide identity (Fig. 2). All of the sequences obtained had identical patterns of predicted splice donor and acceptor sites and the predicted spliced cDNAs could encode proteins of approximately 160 residues that shared 50-100 % translated amino acid identity. The alleles were named by similarity to the previously

identified Ov9.5 gene variants encoded in the sequenced OvHV-2 genomes (Russell et al., 2013): Ov9.5*0101 (previously Ov9.5*01; Hart et al., 2007); and Ov9.5*0201 (previously Ov9.5*02; Taus et al., 2007). Of the 39 Ov9.5 amplicons sequenced, two were identical to the Ov9.5*0101 sequence and 20 were identical to the Ov9.5*0201 sequence (Table 3). The remaining 17 sequences included two alleles that were over 99% identical to Ov9.5*0201, called Ov9.5*0202 and Ov9.5*0203, and three groups of more divergent sequences. These included Ov9.5*0301, which had 78 % amino acid identity to Ov9.5*0201, and Ov9.5*0401, with 95 % identity to Ov9.5*0201. The final three alleles encoded Ov9.5 proteins that were more than 98 % identical to each other but 50-56 % identical to the proteins encoded by the other allele groups. These were termed Ov9.5*0501 to *0503 (Fig. 2; Table 3).

3.3 Analysis of Ov9.5 alleles

The Ov9.5 polypeptides encoded by the nine sequence variants in Table 3 were identical at only 57 positions in the aligned 159 residues. The conserved positions included seven asparagine residues, of which six were predicted to be potential N-linked glycosylation sites in at least one variant, and seven cysteine residues. (Fig. 3). These features suggest shared structural and potentially functional similarity within the Ov9.5 proteins. Structure-based homology searching with an alignment of the available A9.5 and Ov9.5 polypeptide sequences (http://toolkit.tuebingen.mpg.de/hhpred) showed that only two proteins of known structure gave significant alignments with the mature Ov9.5 polypeptide sequences. These were IL-4 (P = 0.001) and IL-21 (P = 0.006), which had previously been identified as potential homologues of A9.5 (Russell et al., 2013), using an alignment with fewer alleles. These cytokines share a 4-helix bundle structure stabilised by 3 disulphide bonds. Notably, among over 100 residues of each cytokine aligned with the A9.5/Ov9.5 sequences, 3

cysteines were conserved, adding further weight to the hypothesis that these viral genes encode distant homologues of 4-helix bundle cytokines.

The possibility of positive selection for polymorphism within the Ov9.5 polypeptide was investigated by analysing the ratio of synonymous to non-synonymous substitutions within the aligned codons by multiple approaches using Topali and MEGA. Within Topali, PAML/CodeML (Yang, 2007) was used to test for the presence of positive selection under three pairs of models with increasing complexity (M0 v M3, M1a v M2a and M7 v M8). In each case the likelihood ratio test statistic did not reach significance levels. Similarly, using the Hyphy software package (Kosakovsky Pond et al. 2005) within MEGA, 15 codon positions had dN-dS values greater than one and the maximum dN-dS value found was 2.0, but no positions were found to be statistically significant ($P<0.05$). Thus, there was no evidence to reject the null hypothesis of neutral selection. However, application of a specific test of neutrality within MEGA (codon-based Z-test of selection; Nei and Gojobori, 1986), suggested that the differences between the more divergent alleles (*i.e*. pairwise comparisons with less than 90 % identity) could not be explained by neutral selection ($P<0.001$).

3.4 Combined phylogenetic analysis

Phylogenetic analysis of the combined ORF50, ORF75 and Ov9.5 nucleotide sequences from the 21 animals for which data were available from all three loci showed twelve distinct viral variants (Fig. 4). The overall tree topology was driven by variation within the Ov9.5 locus, while differences at the other two loci influenced the subdivision of nodes sharing the same Ov9.5 allele. Virus genotypes have been assigned in Fig. 4 based on the Ov9.5 clades and subtypes based on allelic differences at all three loci. The most frequent genotype among the OvHV-2 strains analysed was OvHV-2 type 2a (ORF50*0101/

ORF75*0101/ Ov9.5*0201), which carries the majority allele at each locus. Samples that shared an Ov9.5 genotype but differed at other loci generally differed at a single locus. Thus the seven samples with the Ov9.5*0301 allele also shared the ORF75*0101 allele but differed at ORF50. These small differences between virus genotypes suggest the major mode of evolution is by sequence divergence whilst recombination between virus genotypes may play a minor role.

4. Discussion

In this study, we have analysed OvHV-2 strain variation by sequencing of three loci in diagnostic samples from 45 infected animals, including sheep, bison, reindeer and cattle. The analysis showed that there were seven ORF50 alleles, four ORF75 alleles and nine Ov9.5 alleles in our sample set. The degree of polymorphism varied between the genes analysed: alleles of ORF50 and ORF75 shared greater than 98 % identity whilst the most divergent alleles of Ov9.5 had around 55 % nucleotide identity (Table 3). These three gene segments comprise only about 1 % of the OvHV-2 genome. Comparative analysis of the published OvHV-2 genome sequences (Hart et al., 2007; Taus et al., 2007) showed that there were only two segments in the genome where sequence identity between the two virus strains was significantly lower than 98 % (Taus et al., 2007; G.C. Russell unpublished observations). These were the repetitive region of ORF73, which had segments of identity ranging from 70-90%, and the Ov9.5 gene, which had about 60 % identity between the two genome sequences. Thus selection of two loci with over 98 % identify between alleles, in addition to the divergent Ov9.5 region, allowed both the overall conservation between OvHV-2 strains and the diversity at Ov9.5 to be analysed.

Although much of the variation in Ov9.5 was found within the predicted intron sequences (Fig. 1; supplementary Fig. S1), the predicted Ov9.5 spliced cDNA sequences were also highly polymorphic with 62 % nucleotide and only 50 % predicted amino acid identity between the coding regions of the most divergent sequences. While no statistical support was found for positive selection within the Ov9.5 alleles, a null hypothesis of neutral selection among the more divergent alleles was also rejected. Given that polymorphic sites are found throughout the Ov9.5 gene (exons and introns), and appear to be restricted to the Ov9.5 locus, there may be multiple processes driving polymorphism within these alleles, including, but not limited to, immunological selection acting on Ov9.5 within the reservoir host. Data from a wider range of Ov9.5 alleles, from MCF cases and reservoir hosts, is required to allow a deeper analysis of polymorphism in this region.

It is also notable that the Ov9.5*0201 allele, originally identified in sheep samples from the US (Taus et al., 2007), was the most frequent allele among the samples tested here (20 of 39 samples). While this indicates that OvHV-2 strains prevalent in Europe and America may be similar, it also suggests that the characterised genomic sequence of BJ1035 (Hart et al., 2007) represents an infrequent and divergent Ov9.5 genotype. However, the two strains of virus sequenced to date are highly similar across the remainder of the genome (Taus et al., 2007)

Despite the highly divergent nature of the Ov9.5 alleles, sequencing of segments of the ORF50 and ORF75 genes of virus strains carrying the same Ov9.5 alleles showed further genetic differences. This suggests that while analysis of polymorphism at the Ov9.5 locus may be a useful tool for distinguishing MCF virus strains, a multi-locus approach may be more informative for molecular epidemiology of MCF, and for tracing specific chains of infection.

The use of Ov9.5 and other loci as genotyping tools for MCF outbreaks is illustrated in Table 1. Among five outbreaks (outbreak B and outbreak L with head & eye presentation; farm A, farm E and outbreak A with alimentary presentation; Table 1), cattle involved in four shared the same Ov9.5 genotype (Ov9.5*0201) suggesting that a single strain of virus may have been involved in each of these outbreaks and that analysis of Ov9.5 has epidemiological value. In two of these outbreaks, two cattle were analysed and found to be identical at all three loci, further supporting this view.

In the fifth outbreak (Twomey et al., 2002), four cattle samples exhibited three different Ov9.5 genotypes suggesting infection by multiple OvHV-2 strains. This outbreak was unusual in that all four animals survived MCF, and was complicated by signs of concurrent bracken poisoning. Bracken poisoning was suggested to be a predisposing or trigger factor for MCF in these cattle, and may have increased the susceptibility of the animals involved, potentially leading to infection by different virus strains.

In one MCF outbreak where a virus-positive nasal swab sample was obtained from an in-contact sheep (sample 10-530-8, outbreak L; Table 1), the genotype of this virus isolate differed at all three loci from MCF-affected cattle samples from the same outbreak. This illustrates that further research is required to define chains of infection in MCF outbreaks and to determine whether the same range of OvHV-2 strains that circulate in sheep are found in cattle affected by MCF.

An examination of potential correlations between MCF occurrence or presentation and virus strain was performed for 32 samples that had been genotyped for Ov9.5 or for all three loci tested. In the case of Ov9.5 alleles, only Ov9.5*0201, *0301 and *0401 were found in more than two samples. Ov9.5*0201 was found in 16 samples that were classed as alimentary or head & eye, whilst Ov9.5*0301 was found in six samples classed as alimentary

or head & eye, and Ov9.5*0401 was found in three samples of which two were classed as neurological and the third head & eye. Thus, there seems to be no clear association between a particular presentation of MCF and the genotype of the virus (with respect to the three genes analysed) isolated from diagnostic samples but this is limited by the low number of samples for which both clinical data and genotyping were available. Using geographical information from each case, the distribution of samples genotyped for Ov9.5 was analysed on a regional basis, in keeping with maintaining anonymity of sample providers. These data, illustrated in supplementary Fig. S2, showed that the distribution of MCF cases was in line with the distribution of sheep in the UK – mainly in the west of the country - and that the genotypes found in outbreaks appeared to reflect the main alleles in that region.

The occurrence of MCF outbreaks may reflect specific strains of virus that are more infectious or are shed in greater quantities by infected sheep with the potential to infect groups of cattle or they may indicate conditions of husbandry, environment or herd genetics that favour virus transmission. Unfortunately there are insufficient data from the samples analysed in this report to address these hypotheses and further work is required to compare MCF viruses involved in disease outbreaks with those that cause sporadic MCF.The high resolution of the molecular methods described here to discriminate between strains of OvHV-2 may facilitate the epidemiological study of MCF in a wider range of samples and locations.

Acknowledgements

This work was supported by grants from the Scottish Government's Rural and Environment Science and Analytical Services Division and from the Biotechnology and Biological Sciences Research Council (BBSRC UK) (BBS/B/00468 and BB/H008950). BL received a Wellcome Trust Vacation Studentship award.

REFERENCES

Ackermann, M., 2006. Pathogenesis of gammaherpesvirus infections. Veterinary Microbiology 113, 211-222.

Albrecht, J.C., Nicholas, J., Biller, D., Cameron, K.R., Biesinger, B., Newman, C., Wittmann, S., Craxton, M.A., Coleman, H., Fleckenstein, B., Honess, R.W., 1992. Primary Structure of the Herpesvirus Saimiri Genome. Journal of Virology 66, 5047-5058.

Barton, E., Mandal, P., Speck, S.H., 2011. Pathogenesis and Host Control of Gammaherpesviruses: Lessons from the Mouse. Annual Review of Immunology, 29, 351-397.

Baxter, S.I.F., Pow, I., Bridgen, A., Reid, H.W., 1993. PCR detection of the sheep-associated agent of malignant catarrhal fever. Archives of Virology 132, 145-159.

Coulter, L.J., Reid, H.W., 2002. Isolation and expression of three open reading frames from ovine herpesvirus-2. J Gen Virol 83, 533-543.

Coulter, L.J., Wright, H., Reid, H.W., 2001. Molecular genomic characterization of the viruses of malignant catarrhal fever. J. Comp Pathol. 124, 2-19.

Davison, A.J., Eberle, R., Ehlers, B., Hayward, G.S., McGeoch, D.J., Minson, A.C., Pellett, P.E., Roizman, B., Studdert, M.J., Thiry, E., 2009. The order Herpesvirales. Archives of Virology 154, 171-177.

Ensser, A., Pflanz, R., Fleckenstein, B., 1997. Primary structure of the alcelaphine herpesvirus 1 genome. J. Virol. 71, 6517-6525.

Hart, J., Ackermann, M., Jayawardane, G., Russe, G., Haig, D.M., Reid, H., Stewart, J.P., 2007. Complete sequence and analysis of the ovine herpesvirus 2 genome. J Gen Virol 88,
28-39.

Holliman, A., Daniel, R., Twomey, D.F., Barnett, J., Scholes, S., Willoughby, K., Russell, G., 2007. Malignant catarrhal fever in cattle in the UK. Veterinary Record 161, 494-495.

Hughes, D.J., Kipar, A., Milligan, S., Cunningham, Saunders, M., Quail, M.A., Rajandream, M-A., Efstathiou, S., Bowden, R.J., Chastel, C., Bennett, M., Sample, J.T., Barrell, B.,

Davison, A.J., Stewart, J.P., 2010. Characterization of a novel wood mouse virus related to murid herpesvirus 4. Journal of General Virology 91, 867-879.

Hussy, D., Stauber, N., Leutenegger, C.M., Rieder, S., Ackermann, M., 2001. Quantitative fluorogenic PCR assay for measuring ovine herpesvirus 2 replication in sheep. Clinical and Diagnostic Laboratory Immunology 8, 123-128.

Jayawardane, G., Russell, G.C., Thomson, J., Deane, D., Cox, H., Gatherer, D., Ackermann, M., Haig, D.M., Stewart, J.P., 2008. A captured viral interleukin 10 gene with cellular exon structure. J. Gen. Virol. 89, 2447-2455.

Katoh, K., Standley, D. 439 M., 2013. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. Molecular Biology and Evolution 30, 772-780.

Kosakovsky Pond, S.L., Frost, S.D.W., Muse, S.V. 2005. HyPhy: hypothesis testing using phylogenies. Bioinformatics 21, 676-679.

McGeoch, D.J., 2001. Molecular evolution of the gamma-Herpesvirinae. Philosophical Transactions of the Royal Society of London Series B-Biological Sciences 356, 421-435.

McGeoch, D.J., Gatherer, D., Dolan, A., 2005. On phylogenetic relationships among major lineages of the Gammaherpesvirinae. J Gen Virol 86, 307-316.

Mills, R., Rozanov, M., Lomsadze, A., Tatusova, T., Borodovsky, M., 2003. Improving gene annotation of complete viral genomes. Nucleic Acids Research 31, 7041-7055.

Milne, E.M., Reid, H.W., 1990. Recovery of a cow from malignant catarrhal fever. Veterinary Record 126, 640-641.

Milne, I., Lindner, D., Bayer, M., Husmeier, D., McGuire, G., Marshall, D.F., Wright, F., 2009. TOPALi v2: a rich graphical interface for evolutionary analyses of multiple alignments on HPC clusters and multi-core desktops. Bioinformatics 25, 126-127.

Mitchell, E., Scholes, S., 2009. Unusual presentation of malignant catarrhal fever involving neurological disease in young calves. Veterinary Record 164, 240-242.

Nash, A.A., Dutia, B.M., Stewart, J.P., Davison, A.J., 2001. Natural history of murine gammaherpesvirus infection. Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences 356, 569-579.

Nei M., Gojobori T. (1986). Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Molecular Biology and Evolution 3,418-426.

OToole, D., Li, H., Miller, D., Williams, W.R., Crawford, T.B., 1997. Chronic and recovered cases of sheep-associated malignant catarrhal fever in cattle. Veterinary Record 140, 519-524.

Otter, A., Pow, I., Reid, H.W., 2002. Outbreak of malignant catarrhal fever in Welsh Black cattle in Carmarthenshire. Veterinary Record 151, 321-324.

Penny, C., 1998. Recovery of cattle from malignant catarrhal fever. Veterinary Record 142, 227.

Russell, G.C., Stewart, J.P., Haig, D.M., 2009. Malignant catarrhal fever: A review. Veterinary Journal 179, 324-335.

Russell, G.C., Todd, H., Deane, D., Percival, A., Dagleish, M., Haig, D.M., Stewart, J.P., 2013. A novel spliced gene in alcelaphine herpesvirus encodes a glycoprotein which is secreted in vitro. J Gen Virol 94, 2515-2523.

Sambrook, J., Fritsch, E.F. & Maniatis, T. (1989) Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbour Laboratory Press, New York.

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. Molecular Biology and Evolution 28, 2731-2739.

Tamura K., Nei M., 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Molecular Biology and Evolution
10, 512-526.

Taus, N.S., Herndon, D.R., Traul, D.L., Stewart, J.P., Ackermann, M., Li, H., Knowles, D.P., Lewis, G.S., Brayton, K.A., 2007. Comparison of ovine herpesvirus 2 genomes isolated from domestic sheep (Ovis aries) and a clinically affected cow (Bos bovis). J Gen Virol 88, 40-45.

Twomey, D.F., Holt, G.J., Reid, H.W., 2002. Malignant catarrhal fever in cattle with suspected bracken poisoning. Vet Rec 151, 486-487.

Untergasser, A., Nijveen, H., Rao, X., Bisseling, T., Geurts, R., Leunissen, J.A.M., 2007. Primer3Plus, an enhanced web interface to Primer3. Nucleic Acids Research 35, W71-W74. World Organisation for Animal Health (OIE)., 2008. Malignant Catarrhal Fever. Manual of

Diagnostic Tests and Vaccines for Terrestrial Animals. OIE, Paris, pp. 779-788.
Yang, Z., 2007. PAML 4: Phylogenetic Analysis by Maximum Likelihood. Molecular Biology and Evolution 24, 1586-1591.

498 **Table 1. Details of samples analyzed**

			Genotypes[c]			
Sample[a]	**Presentation (outbreak)[b]**	**Species**	**ORF50**	**ORF75**	**Ov9.5**	**Citation/comment[d]**
09-1280		Cattle	*0501	ND	ND	
09-1334		Cattle	*0102	ND	ND	
10-441		Cattle	*0101	ND	ND	
06-1784		Cattle	*0103	*0101	*0503	
09-1058		Cattle	*0301	*0101	*0301	
09-1060		Cattle	*0201	*0101	*0301	
09-1188		Cattle	*0101	*0101	*0201	
09-1197		Cattle	*0101	*0101	*0201	
09-986		Cattle	*0201	*0101	*0301	
05-1500		Cattle	ND	*0201	ND	
US virus[e]		sheep	*0601	*0101	*0201	Taus et al., 2007
06-146	Alimentary	Cattle	ND	ND	*0201	cloned
06-202-2	Alimentary	Cattle	*0101	*0101	*0201	
05-725	Alimentary (Farm A)	Cattle	*0102	*0101	*0301	Holliman et al., 2007
05-786-4	Alimentary (Farm A)	Cattle	*0102	*0101	*0301	Holliman et al., 2007
06-442-1	Alimentary (Farm E)	Cattle	ND	ND	*0201	Holliman et al., 2007
06-442-2	Alimentary (Farm E)	Cattle	ND	ND	*0201	Holliman et al., 2007
06-443	Alimentary (Farm E)	Cattle	ND	ND	*0201	Holliman et al., 2007
06-235-1	Alimentary (outbreak A)	Cattle	*0101	ND	ND	
06-235-2	Alimentary (outbreak A)	Cattle	ND	ND	*0201	
06-235-3	Alimentary (outbreak A)	Cattle	ND	ND	*0201	

06-235-4	Alimentary (outbreak A)	Cattle	*0101	*0101	*0201	
BJ1035[f]	Head & Eye	Cattle	*0401	*0301	*0101	Hart et al., 2007 cloned
01-2021-1	Head & Eye (outbreak B)	Cattle	ND	ND	*0201	Twomey et al., 2002
01-2021-2	Head & Eye (outbreak B)	Cattle	ND	ND	*0401	Twomey et al., 2002
01-2021-3	Head & Eye (outbreak B)	Cattle	ND	ND	*0301	Twomey et al., 2002
01-2021-4	Head & Eye (outbreak B)	Cattle	ND	ND	*0301	Twomey et al., 2002
10-414	Head &Eye (outbreak L)	Cattle	ND	ND	*0201	
10-530-11	Head &Eye (outbreak L)	Cattle	ND	ND	*0201	
10-530-12	Head &Eye (outbreak L)	Cattle	ND	ND	*0201	
10-404	Head &Eye (outbreak L)	Cattle	*0101	*0101	*0201	
10-415	Head &Eye (outbreak L)	Cattle	*0101	*0101	*0201	
10-530-8	(outbreak L)	Sheep	*0103	*0102	*0502	
10-524	Head &Eye (outbreak S)	Cattle	*0101	*0101	*0201	
06-1114	Head and Eye	cattle	ND	ND	*0201	cloned
06-1116	Head and Eye	cattle	ND	ND	*0501	cloned
01-392	Neurological	Cattle	ND	ND	*0401	Mitchell and Scholes, 2009
01-393	Neurological	Cattle	*0101	*0101	*0401	Mitchell and Scholes, 2009
06-486	(outbreak)	cattle	ND	ND	*0201	cloned
07-1778-2	(Sporadic)	Cattle	ND	ND	*0101	
06-1811	(Sporadic)	Reindeer	*0103	*0101	*0203	
06-283	(Sporadic)	Bison	*0101	*0101	*0301	
08-145	(Sporadic)	Deer	*0103	*0101	ND	
06-1567	head & eye (Sporadic)	Cattle	*0102	*0101	*0301	

05-1513	head & eye (Sporadic)	Cattle	*0102	*0201	*0202	
06-387		cattle	ND	ND	*0201	cloned

[a] Sample identifiers are mainly as assigned by MRI virus surveillance unit, with the format yy-nnnn-ss, where the first two digits indicate year after 2000 and the following digits are sample (and sub-sample) number within that year. Exceptions are the samples used for genome sequencing of OvHV-2: identified as BJ1035 (Hart et al., 2007) and US virus (Taus et al., 2007).

[b] information on the presentation of MCF in a specific case is given only if recorded. Where samples were known to be submitted from an outbreak of MCF involving multiple cases with the same presentation, this is detailed in parentheses. Where no other MCF-positive sample was submitted from the same source within at least a month, samples are indicated as *sporadic*.

[c] genotypes of each locus are given as assigned in the text. ND indicates that the appropriate PCR product sequence could not be determined from that sample.

[d] where samples analysed here have been described or sequence data presented in previous publications, these are cited. Those samples for which the Ov9.5 PCR product was additionally cloned and sequenced from multiple clones are also indicated in this column.

[e] the genotype of each locus from the US virus was assigned based on the predicted PCR products from the published sequence (Taus et al., 2007).

[f] BJ1035 DNA was analysed as a positive control, confirming the published sequence for the three loci amplified.

521 **Table 2. PCR primers used in this work**

Primer name	Sequence (5'-3')	Product size (bp|)	Description/reference
Ov9.5 L1	AAAGACACATGCATCAAACTCT		3'-end of the Ov9.5 putative gene
Ov9.5 R1	GGGTAAGTACATGGTATAAAGCAG	954	5'-end of the Ov9.5 putative gene
Ov9.5 L2	TGAAAAACTGGCCACATAAA		nested primer for Ov9.5 L1-R1 product
Ov9.5 R2	AAGAACCCTGATAAACTCCAGA	893	nested primer for Ov9.5 L1-R1 product
OHVorf50_F1	CCCCAACAAGTCAGCATTTT		ORF50 exon 2 forward
OHVorf50_R1	TCAGTCGAATGCTGTTGGAG	600	ORF50 exon 2 reverse
OvHV2_orf50_F2	GGACCTCTCATCTCTTCTGCAA		nested primer for OHVorf50_F1-R1 product
OvHV2_orf50_R2	ATGGCAAAGTCACAGGGATG	444	nested primer for OHVorf50_F1-R1 product
556	TTCTGGGGTAGTGGCGAGCGAAGGCTTC		Baxter et al., 1993
755	AAGATAAGCACCAGTTATGCATCTGATAAA	422	
556	TTCTGGGGTAGTGGCGAGCGAAGGCTTC		
555	AGTCTGGGTATATGAATCCAGATGGCTCTC	238	

522

523

524 **Table 3. Sequence similarity between alleles of the three loci analysed**

ORF50			ORF75			Ov9.5		
Alleles	[a] Differences from ORF50*0101 DNA (protein)	[b] Number of samples of each type	Alleles	[a] Differences from ORF75*0101 DNA (protein)	[b] Number of samples of each type	Alleles	[c] % identity with Ov9.5*0201 DNA (protein)	[b] Number of samples of each type
ORF50*0101	0 (0)	11	ORF75*0101	0 (0)	18	Ov9.5*0101	56.6 (49.6)	2
ORF50*0102	1 (0)	5	ORF75*0102	1 (0)	1	Ov9.5*0201	100 (100)	20
ORF50*0103	1 (0)	4	ORF75*0201	2 (1)	2	Ov9.5*0202	99.7 (99.3)	1
ORF50*0201	2 (1)	2	ORF75*0301	2 (1)	1	Ov9.5*0203	99.6 (99.3)	1
ORF50*0301	2 (1)	1				Ov9.5*0301	83.2 (78.6)	9
ORF50*0401	5 (3)	1				Ov9.5*0401	98.2 (95.5)	3
ORF50*0501	4 (3)	1				Ov9.5*0501	60.5 (50.6)	1
ORF50*0601	2 (2)	0				Ov9.5*0502	60.4 (49.6)	1
						Ov9.5*0503	60.5 (51.2)	1

525

[a] For ORF50 and ORF75 the number of differences of each allele from the majority allele (*0101) in the DNA sequence is given, with differences in the translated protein sequence in parentheses.

[b] For each allele of each gene, the number of samples analysed with that allele is given.

[c] For Ov9.5, the percent identity of each allele to the majority allele (Ov9.5*0201) in the DNA and the translated protein sequences are given.

Figure Legends

Fig. 1. Schematic depiction of similarity between the two OvHV-2 genome sequences in the Ov9–Ov10 intergenic region, based on the sequence alignment in supplementary Fig. S1. Identical nucleotide positions within the alignment are indicated by vertical bars (|) and mismatched bases are depicted as dots (.). The positions of PCR primers designed to amplify the Ov9.5 region are indicated above the line, with primer names and polarities given. The predicted termination codons of the Ov9, Ov9.5 and Ov10 genes are shown as colon triplets (:::) annotated above the line, while the positions of the putative coding exons of the Ov9 5 gene are shaded dark grey within the alignment and named above it. The region encoding the Ov9.5 gene, which has less than 70% identity, is highlighted grey, whilst the intergenic flanking regions have over 90% sequence identity.

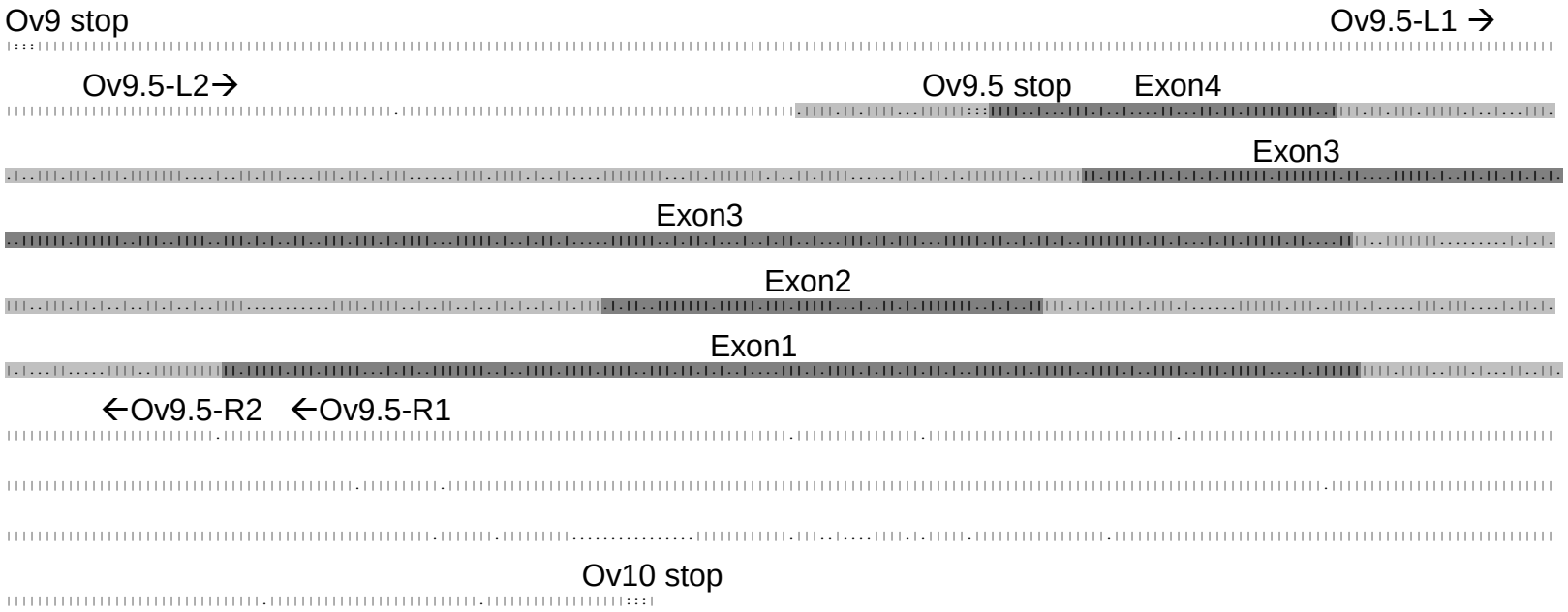
Fig. 2. Phylogenetic analysis of Ov9.5 gene sequences. Nucleotide sequences of the 9 variants Ov9.5*0101 to Ov9.5*0503 were aligned using MAFFT and phylogenetic analysis was done in MEGA version 6.0, using the maximum likelihood approach with 500 bootstrap replicates. Model selection indicated that the Tamura-Nei model (Tamura and Nei, 1993) was most appropriate to this dataset, producing tree topology that was supported by bootstrap values of least 98 % (except for the branching within the Ov9.5*05 clade). Bootstrap percentage values are given on the left of their respective nodes. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site (with scale bar) .

Fig. 3. Alignment of the Ov9.5 predicted polypeptide sequences aligned by MAFFT (http://mafft.cbrc.jp/alignment/server/index.html). Conserved residues are indicated below the alignment, with those identical in all nine sequences in uppercase, whilst residues conserved in at least seven of nine alleles are shown in lowercase. The positions of conserved cysteine residues are boxed and predicted N-linked glycosylation sites (score >0.5; NetNGlyc 1.0 Server, http://genome.cbs.dtu.dk/) are highlighted in grey. The position and length of signal peptides in

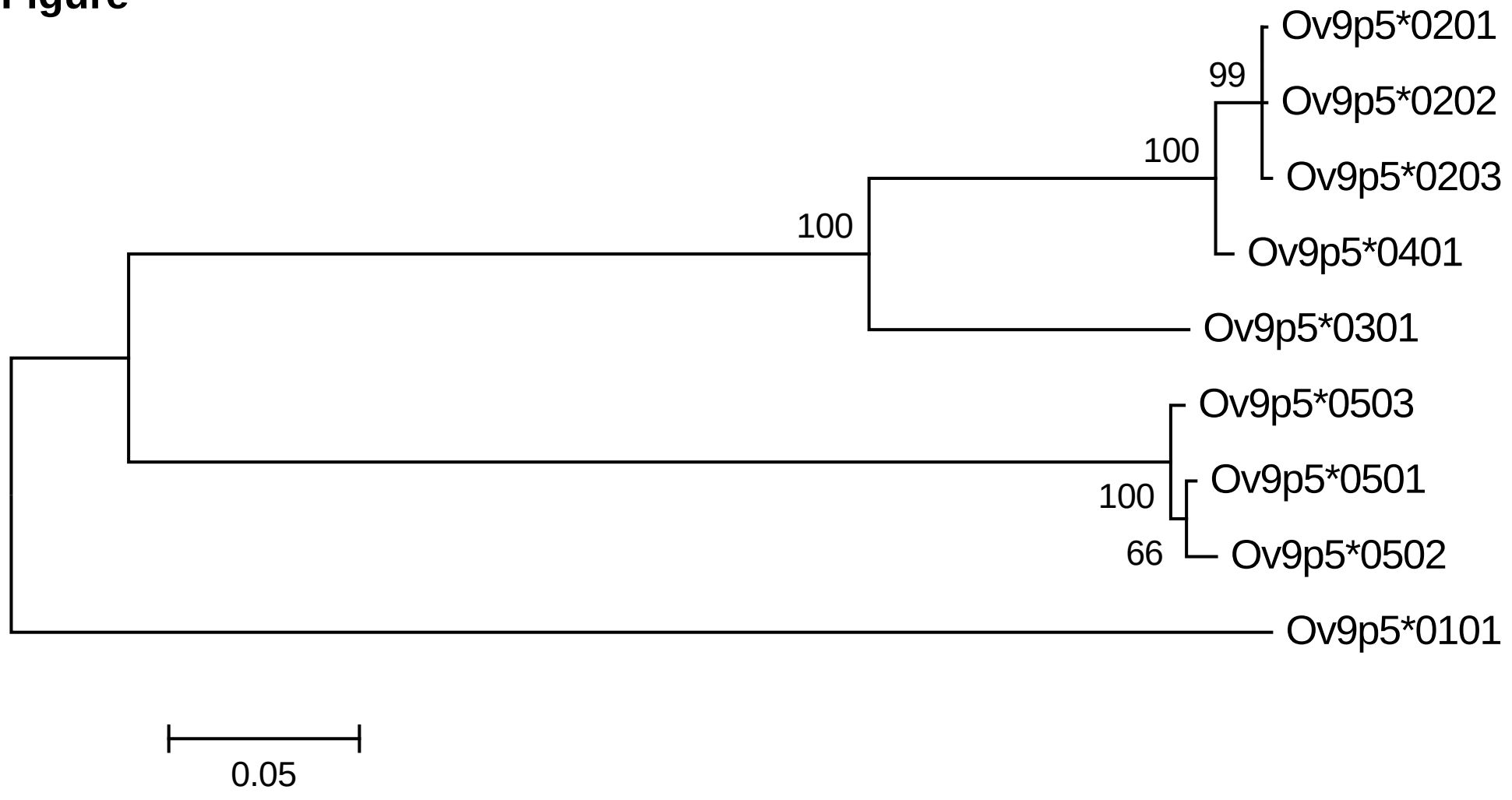
each polypeptide were predicted using SignalP Server 4.1 (http://genome.cbs.dtu.dk/) and is indicated by underlining.

Fig. 4. Phylogenetic analysis of the combined ORF50-ORF75-Ov9.5 nucleotide sequences was performed using the approach described in Fig. 2 for the 21 samples for which sequence was obtained from all three loci. Branch lengths are not proportional to genetic distance in this figure. Bootstrap values (percentage of 500 replicates) are given on the left of the respective nodes, except for two terminal nodes where they are placed to the right of the node. The combined sequences are named according to the sample designations in Table 1. The genotypes of the three loci for each sample are indicated to the right of the tree and assignment of OvHV-2 strain type is given on the right of the figure. Samples with identical genotypes are boxed whilst the Ov9.5 clades are separated by thick lines.

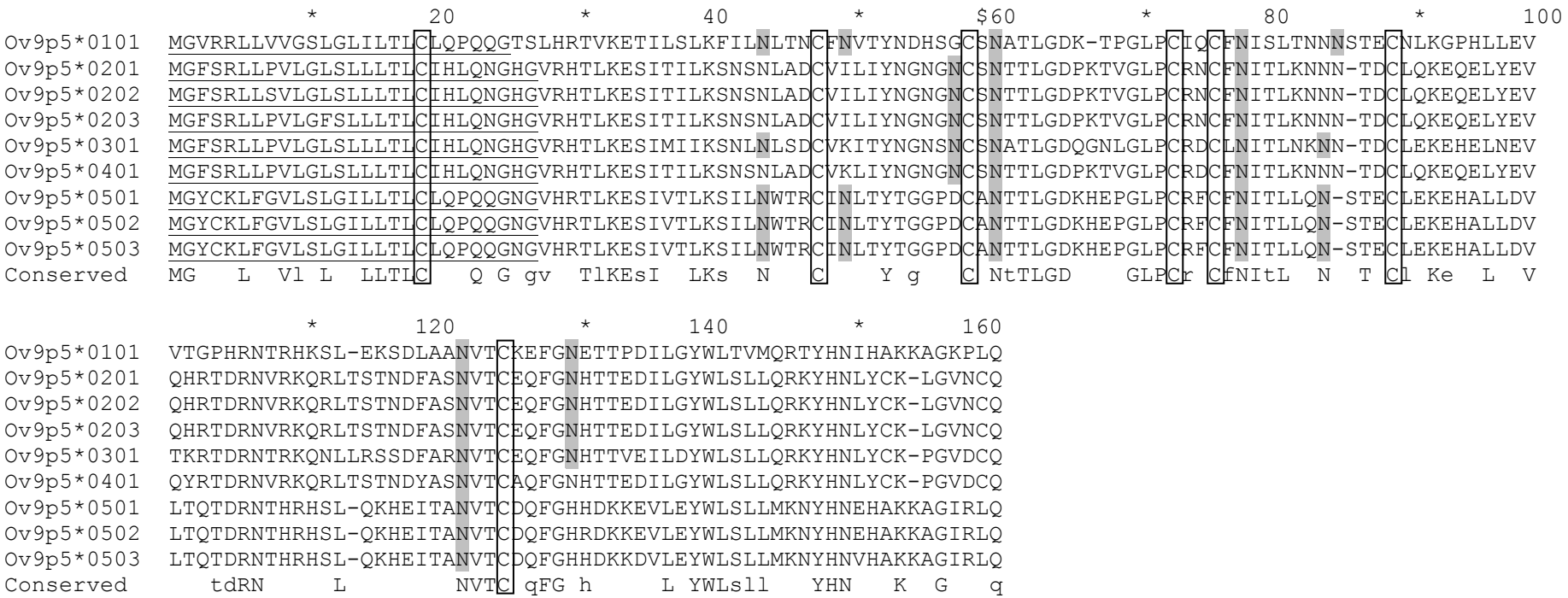
Figure


Ov9 stop
Ov9.5-L1 →
Ov9.5-L2→
Ov9.5 stop
Exon4
Exon3
Exon3
Exon2
Exon1
←Ov9.5-R2
←Ov9.5-R1
Ov10 stop

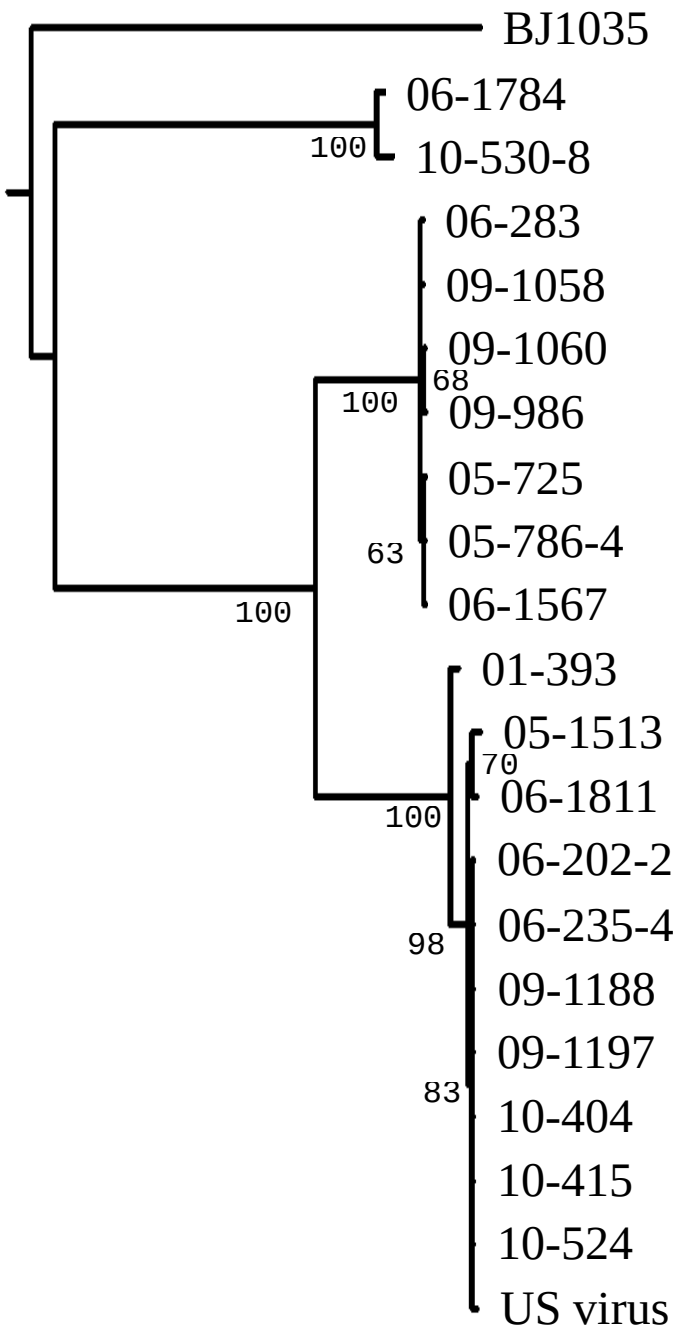
Figure


Ov9p5*0201
99
Ov9p5*0202
100
Ov9p5*0203
100
Ov9p5*0401
Ov9p5*0301
Ov9p5*0503
Ov9p5*0501
100
66
Ov9p5*0502
Ov9p5*0101
0.05

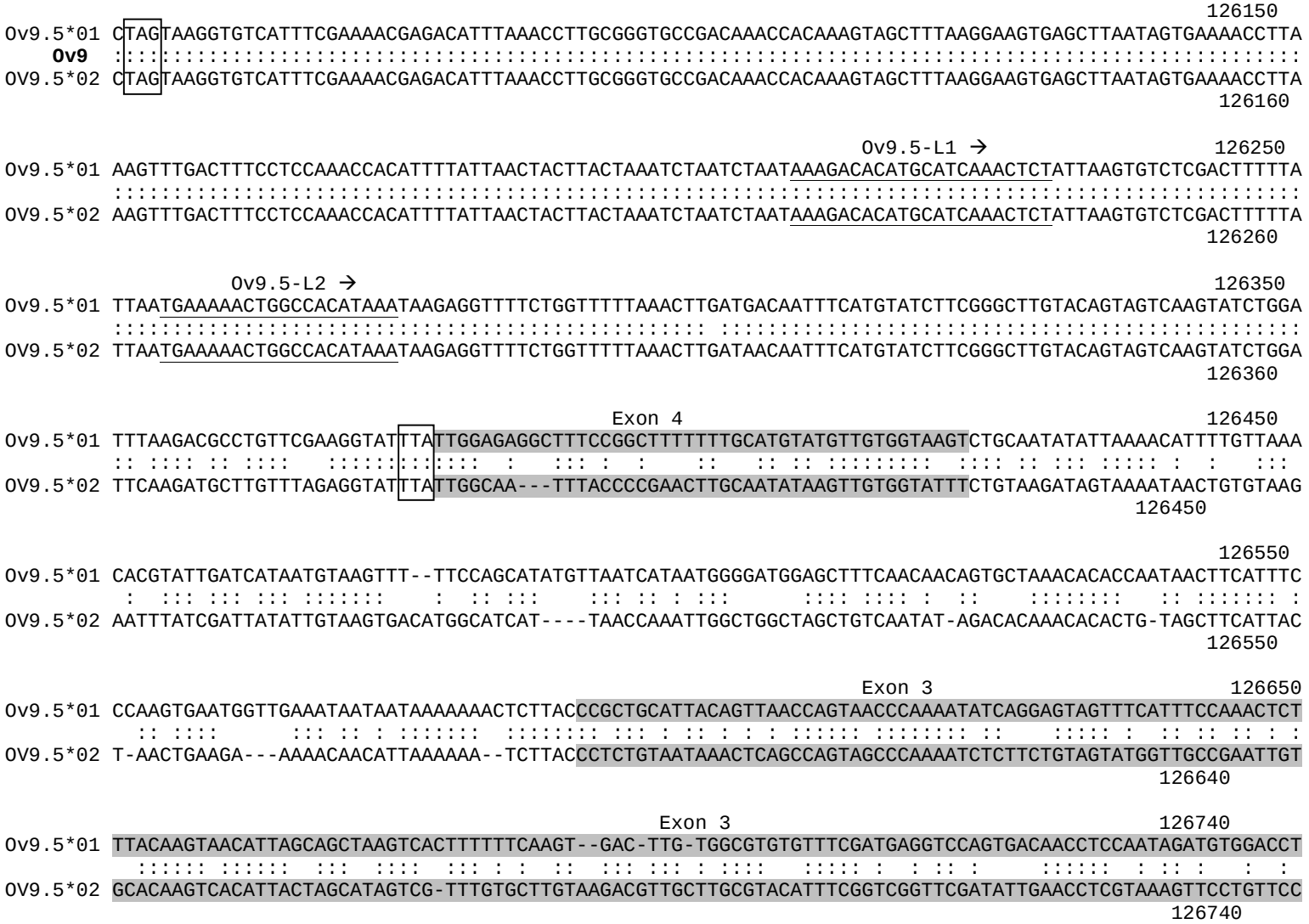
Figure


* 20 * 40 * $60 * 80 * 100
Ov9p5*0101 MGVRRLLVVGSLGLILTLCLQPQQGTSLHRTVKETILSLKFILNLTNCFNVTYNDHSGCSNATLGDK-TPGLPCIQCFNISLTNNNSTECNLKGPHLLEV
Ov9p5*0201 MGFSRLLPVLGLSLLLTLCIHLQNGHGVRHTLKESITILKSNSNLADCVILIYNGNGNCSNTTLGDPKTVGLPCRNCFNITLKNNN-TDCLQKEQELYEV
Ov9p5*0202 MGFSRLLSVLGLSLLLTLCIHLQNGHGVRHTLKESITILKSNSNLADCVILIYNGNGNCSNTTLGDPKTVGLPCRNCFNITLKNNN-TDCLQKEQELYEV
Ov9p5*0203 MGFSRLLPVLGFSLLLTLCIHLQNGHGVRHTLKESITILKSNSNLADCVILIYNGNGNCSNTTLGDPKTVGLPCRNCFNITLKNNN-TDCLQKEQELYEV
Ov9p5*0301 MGFSRLLPVLGLSLLLTLCIHLQNGHGVRHTLKESIMIIKSNLNLSDCVKITYNGNSNCSNATLGDQGNLGLPCRDCLNITLNKNN-TDCLEKEHELNEV
Ov9p5*0401 MGFSRLLPVLGLSLLLTLCIHLQNGHGVRHTLKESITILKSNSNLADCVKLIYNGNGNCSNTTLGDPKTVGLPCRDCFNITLKNNN-TDCLQKEQELYEV
Ov9p5*0501 MGYCKLFGVLSLGILLTLCLQPQQGNGVHRTLKESIVTLKSILNWTRCINLTYTGGPDCANTTLGDKHEPGLPCRFCFNITLLQN-STECLEKEHALLDV
Ov9p5*0502 MGYCKLFGVLSLGILLTLCLQPQQGNGVHRTLKESIVTLKSILNWTRCINLTYTGGPDCANTTLGDKHEPGLPCRFCFNITLLQN-STECLEKEHALLDV
Ov9p5*0503 MGYCKLFGVLSLGILLTLCLQPQQGNGVHRTLKESIVTLKSILNWTRCINLTYTGGPDCANTTLGDKHEPGLPCRFCFNITLLQN-STECLEKEHALLDV
Conserved MG L Vl L LLTLC Q G gv TlKEsI LKs N C Y g C NtTLGD GLPCr CfNItL N T Cl Ke L V
* 120 * 140 * 160
Ov9p5*0101 VTGPHRNTRHKSL-EKSDLAANVTCKEFGNETTPDILGYWLTVMQRTYHNIHAKKAGKPLQ
Ov9p5*0201 QHRTDRNVRKQRLTSTNDFASNVTCEQFGNHTTEDILGYWLSLLQRKYHNLYCK-LGVNCQ
Ov9p5*0202 QHRTDRNVRKQRLTSTNDFASNVTCEQFGNHTTEDILGYWLSLLQRKYHNLYCK-LGVNCQ
Ov9p5*0203 QHRTDRNVRKQRLTSTNDFASNVTCEQFGNHTTEDILGYWLSLLQRKYHNLYCK-LGVNCQ
Ov9p5*0301 TKRTDRNTRKQNLLRSSDFARNVTCEQFGNHTTVEILDYWLSLLQRKYHNLYCK-PGVDCQ
Ov9p5*0401 QYRTDRNVRKQRLTSTNDYASNVTCAQFGNHTTEDILGYWLSLLQRKYHNLYCK-PGVDCQ
Ov9p5*0501 LTQTDRNTHRHSL-QKHEITANVTCDQFGHHDKKEVLEYWLSLLMKNYHNEHAKKAGIRLQ
Ov9p5*0502 LTQTDRNTHRHSL-QKHEITANVTCDQFGHRDKKEVLEYWLSLLMKNYHNEHAKKAGIRLQ
Ov9p5*0503 LTQTDRNTHRHSL-QKHEITANVTCDQFGHHDKKDVLEYWLSLLMKNYHNVHAKKAGIRLQ
Conserved tdRN L NVTC qFG h L YWLsll YHN K G q

Figure



	ORF50/ORF75/Ov9.5	OvHV-2 type
BJ1035	*0401/*0301/*0101	1a
06-1784	*0103/*0101/*0503	5c
10-530-8	*0103/*0102/*0502	5b
06-283	*0101/*0101/*0301	3d
09-1058	*0301/*0101/*0301	3c
09-1060	*0201/*0101/*0301	3b
09-986	*0201/*0101/*0301	3b
05-725	*0102/*0101/*0301	3a
05-786-4	*0102/*0101/*0301	3a
06-1567	*0102/*0101/*0301	3a
01-393	*0101/*0101/*0401	4a
05-1513	*0102/*0201/*0202	2b
06-1811	*0103/*0101/*0203	2c
06-202-2	*0101/*0101/*0201	2a
06-235-4	*0101/*0101/*0201	2a
09-1188	*0101/*0101/*0201	2a
09-1197	*0101/*0101/*0201	2a
10-404	*0101/*0101/*0201	2a
10-415	*0101/*0101/*0201	2a
10-524	*0101/*0101/*0201	2a
US virus	*0601/*0101/*0201	2d


126150
Ov9.5*01 CTAGTAAGGTGTCATTTCGAAAACGAGACATTTAAACCTTGCGGGTGCCGACAAACCACAAAGTAGCTTTAAGGAAGTGAGCTTAATAGTGAAAACCTTA
Ov9
OV9.5*02 CTAGTAAGGTGTCATTTCGAAAACGAGACATTTAAACCTTGCGGGTGCCGACAAACCACAAAGTAGCTTTAAGGAAGTGAGCTTAATAGTGAAAACCTTA
126160
Ov9.5-L1 →
126250
Ov9.5*01 AAGTTTGACTTTCCTCCAAACCACATTTTATTAACTACTTACTAAATCTAATCTAATAAAGACACATGCATCAAACTCTATTAAGTGTCTCGACTTTTTA
OV9.5*02 AAGTTTGACTTTCCTCCAAACCACATTTTATTAACTACTTACTAAATCTAATCTAATAAAGACACATGCATCAAACTCTATTAAGTGTCTCGACTTTTTA
126260
Ov9.5-L2 →
126350
Ov9.5*01 TTAATGAAAAACTGGCCACATAAATAAGAGGTTTTCTGGTTTTTAAACTTGATGACAATTTCATGTATCTTCGGGCTTGTACAGTAGTCAAGTATCTGGA
OV9.5*02 TTAATGAAAAACTGGCCACATAAATAAGAGGTTTTCTGGTTTTTAAACTTGATAACAATTTCATGTATCTTCGGGCTTGTACAGTAGTCAAGTATCTGGA
126360
Exon 4
126450
Ov9.5*01 TTTAAGACGCCTGTTCGAAGGTATTTATTGGAGAGGCTTTCCGGCTTTTTTTGCATGTATGTTGTGGTAAGTCTGCAATATATTAAAACATTTTGTTAAA
OV9.5*02 TTCAAGATGCTTGTTTAGAGGTATTTATTGGCAA---TTTACCCCGAACTTGCAATATAAGTTGTGGTATTTCTGTAAGATAGTAAAATAACTGTGTAAG
126450
126550
Ov9.5*01 CACGTATTGATCATAATGTAAGTTT--TTCCAGCATATGTTAATCATAATGGGGATGGAGCTTTCAACAACAGTGCTAAACACACCAATAACTTCATTTC
OV9.5*02 AATTTATCGATTATATTGTAAGTGACATGGCATCAT----TAACCAAATTGGCTGGCTAGCTGTCAATAT-AGACACAAACACACTG-TAGCTTCATTAC
126550
Exon 3
126650
Ov9.5*01 CCAAGTGAATGGTTGAAATAATAATAAAAAAACTCTTACCCGCTGCATTACAGTTAACCAGTAACCCAAAATATCAGGAGTAGTTTCATTTCCAAACTCT
OV9.5*02 T-AACTGAAGA---AAAACAACATTAAAAAA--TCTTACCCTCTGTAATAAACTCAGCCAGTAGCCCAAAATCTCTTCTGTAGTATGGTTGCCGAATTGT
126640
Exon 3
126740
Ov9.5*01 TTACAAGTAACATTAGCAGCTAAGTCACTTTTTTCAAGT--GAC-TTG-TGGCGTGTGTTTCGATGAGGTCCAGTGACAACCTCCAATAGATGTGGACCT
OV9.5*02 GCACAAGTCACATTACTAGCATAGTCG-TTTGTGCTTGTAAGACGTTGCTTGCGTACATTTCGGTCGGTTCGATATTGAACCTCGTAAAGTTCCTGTTCC
126740


Exon 3
126840
Ov9.5*01 TTCAGATTACACTCAGTGGAGTTGTTGTTTGTTAGGGAAATGTTGAAACACTGAATGCAAGGCAATCCAGGTGTCTGAAAAACATTAGATACATATAAGT
OV9.5*02 TTTTGTAGACAATCTGTG---TTGTTATTCTTCAGAGTTATGTTGAAGCAATCCCTACAGGGCAACCCTACAGTCTAGAAAACAT---------AAAGGC
126830
Exon 2 126940
Ov9.5*01 TTTGCTATAAACACTACATA-CAATCTATCGCATACAGTAAATTACATATAATATAAAATAGAGACTATATACCTACTT-TG--TCACCTAGTGTAGCAT
OV9.5*02 TTT--TATCAAGATCAGTTAGCCCTTCATCG----------TTACTTATACAA--AAGCAATGAATGCAAACTTACCTGTGGATCACCTAATGTAGTAT
126910
Exon 2
127040
Ov9.5*01 TTGAGCAACCAGAGTGATCATTGTAGGTTACATTCTACAAAAAAA-ACATATTGGTTACTACACTTAAGCAGTGATTCAAGATATAAATTTTAATGTATT
OV9.5*02 TCGAGCAGTTACCGTTACCATTGTAAATCAATTTCTAGAAGAAAATAAATAAT------TACACTAAAGTGGTGAATAGCACTATGAATAAACAGGTGTC
127010
Exon 1
127140
Ov9.5*01 ATAGAATAAGTTGTATACATATACTTACAAAACAGTTGGTTAAATTGAGGATAAACTTTAATGATAAAATAGTTTCCTTCACAGTACGATGTAGACTTGT
OV9.5*02 ACATCTTAT----TATAGGTATACTTACAACACAGTCGGTCAAATTTGAGTTACTCTTTAATATTGTAATACTTTCTTTCAAGGTATGACGAACTCCATG
127100
Exon 1
127230
Ov9.5*01 TCCTTGCTGAGGTTGGAGACACAGGGTCAGTATAAGGCCGAGGCTTCCCACAACGAGCAACCTACGAACACCCATGCTGAGTGT--TCCTATGTCCTCTT
OV9.5*02 TCCATTCTGAAGGTGAAGGCATAATGTCAATAAAAGGCTAAGGCCTAGCACAGGGAGTAACCTGGAGAAACCCATGCTGTGTGTGTTCCCAGCGCC--TT
127200
← Ov9.5-R2
← Ov9.5-R1
127330
Ov9.5*01 CCACAATATCTGGAGTTTATCAGGGTTCTTATACTGCTTTATACCATGTACTTACCCTCAACATTTTTGTAAGTATCGAAATAAGTAAAAATGAAAGTAT
OV9.5*02 CCACAATATCTGGAGTTTATCAGGGTTTTTATACTGCTTTATACCATGTACTTACCCTCAACATTTTTGTAAGTATCGAAATAAGTAAAAATGAAAGTAT
127300
127430
Ov9.5*01 TCGCCTTTATCAGACATAATTGGATTTACTGGACCAAGGGTCAGATCAAAGCATGGGATGAGTAACTTCGGGTTAGGTTGACTTACTAAGAACTTAGGCA
OV9.5*02 TAGCCTTTATCAGACATAGTTGGATTTACTGGACCAAGGGTCAGATCAAAGTATGGGATGAGTAACTTCGGGTTAGGTTGACTTACTAAGAACTTAGGCA
127400

Supplementary Fig. S1. Alignment of the Ov9–Ov10 intergenic region between the two OvHV-2 genome sequences. The segment of the Hart et al. (2007) sequence (AY839756) is labeled Ov9.5*01 while the Taus et al. (2007) sequence (DQ198083) is labeled Ov9.5*02. The predicted termination codons of the Ov9, Ov9.5 and Ov10 genes are boxed, while the putative coding exons of the Ov9.5 alleles are shaded within the alignment and named above it. Identical positions within the alignment are indicated by colons (:), while gaps inserted to maximize similarity are shown by dashes (-). The positions of PCR primers designed to amplify the Ov9.5 region are indicated by underlining in each sequence, with primer names and polarities given above the alignment.

Russell et al., 2014.
Supplementary Fig. S2.
UK map showing regional assignment of Ov9.5 genotypes among MCF-positive samples. Ov9.5 genotypes are represented by symbols as shown in the key on the right of the map, while the genotypes found in the MCF outbreaks listed in Table 1 are shown as numbered yellow ellipses as follows: 1, Outbreak L (head & eye); 2, Farm A (alimentary); 3, Farm E (alimentary); 4, Outbreak B (head & eye); 5, Outbreak A (alimentary).
No samples from MCF cases in Scotland or Northern Ireland were used in this study.