The English strain of rat cytomegalovirus (CMV) contains a novel captured CD200 (vOX2) gene and a spliced CC chemokine upstream from the major immediate-early region: further evidence for a separate evolutionary lineage from that of rat CMV Maastricht

Sebastian Voigt,1,2 Gordon R. Sandford,2,3 Gary S. Hayward3 and William H. Burns3†

1Laboratory for Molecular Biology, Department of Pediatrics, Charité, 10098 Berlin, Germany
2Bone Marrow Transplant Program, Medical College of Wisconsin, Milwaukee, WI 53226, USA
3Viral Oncology Program, Sidney Kimmel Comprehensive Cancer Center, Johns Hopkins University School of Medicine, Baltimore, MD 21231, USA

INTRODUCTION

The subfamily Betaherpesvirinae is characterized by strict species specificity and is currently classified, somewhat arbitrarily, into three genera: Cytomegalovirus in primates, Muromegalovirus in rodents and Roseolovirus [human herpesvirus (HHV)-6 and HHV-7]. However, other \( \beta \)-herpesvirus species, including guinea pig cytomegalovirus (CMV), porcine herpesvirus and elephant endotheliolytic herpesvirus do not fit readily into any of these groups, based on partial phylogenetic and gene-content analyses. Human cytomegalovirus (HCMV) – HHV-5 – is an important pathogen in immunocompromised hosts and as a cause of developmental defects after congenital infections. The 220–250 kb genomes of HCMV AD169 (Chee et al., 1990), HCMV Merlin (Dolan et al., 2004), HCMV PH, TR and FIX (Murphy et al., 2003), chimpanzee CMV – pongine herpesvirus 4 – (CCMV) (Davison et al., 2003), rhesus CMV – Cercopithecine herpesvirus 8 – (Hansen et al., 2003), tupai herpesvirus – tuapiaid herpesvirus 1 – of tree shrews (Bahr & Darai, 2001), the Maastricht isolate of rat CMV – murid herpesvirus 2 – (RCMV-M) (Vink et al., 2000) and the Smith isolate of
urine CMV – murid herpesvirus 1 (MCMV) (Rawlinson et al., 1996) have all been sequenced completely. Partial sequence data for the English isolate of rat CMV (RCMV-E) have also been published (Burns et al., 1988; Sandford & Burns, 1996; Voigt et al., 2001). The two RCMV isolates differ by overall genome size, restriction-map patterns and gene content, which has raised the issue of whether these two viruses may represent different β-herpesviruses species or lineages, rather than different strains of the same virus (Beisser et al., 1998). Whilst the size of the RCMV-M genome resembles that of MCMV, at close to 230 kb, RCMV-E appears to be only 205 kb in size, based on restriction-enzyme cleavage analysis, and also contains a C-type lectin homologue that is not found in other CMV species (Voigt et al., 2001).

Identification and characterization of genes in rodent CMVs is important for fully understanding and optimizing their use as models for HCMV infection and disease. A comparison of HCMV, CCMV, MCMV and RCMV-M shows that the middle portion of their genomes, encoding genes that are necessary for DNA replication and structural proteins, is very similar in both the position and orientation of most open reading frames (ORFs) and that they share considerable similarity. However, there is divergence in the ORFs to the right of the major immediate-early (MIE) enhancer, with the presence of an immediate-early US22-type ORF (ie2) in MCMV (Cardin et al., 1995; Messerle et al., 1991) and RCMV-M (Vink et al., 2000), followed by several ORFs that are not present in the attenuated and deleted HCMV laboratory strains AD169 and Towne, compared to the clinical isolate Toledo (Cha et al., 1996).

We wondered whether these ORFs that are unique to MCMV and RCMV-M are also present in RCMV-E and, if so, whether conserved coding regions might be identified, suggesting important functional domains. This information would be particularly useful, as the production of recombinant viruses in rodent CMVs is possible and such recombinants can be used in animal experiments to elucidate the biological functions of these genes (Brune et al., 2000).

Here, we report three cellular and five viral gene homologues that are present in the genome of RCMV-E. Except for the four HindIII-M homologues e135, e136, e137 and e138, we present the expression kinetics of these genes and compare them with the corresponding homologues in RCMV-M and MCMV. By structural analysis and homology comparisons of the two RCMV isolates, we show further evidence that these two viruses are distinct species. This conclusion is supported strongly by both the results of phylogenetic analyses and the finding of a novel viral CD200 (OX2) homologue in RCMV-E. Other ORFs that we identified include a homologue for salivary gland gene 1 (SGG1), a spliced gene shown in MCMV mutation studies to be important for growth in salivary glands, but not for replication of the virus in vitro (Lagenaur et al., 1994; Vieira et al., 1994), a β (CC)-chemokine (MacDonald et al., 1997, 1999) and an Fcγ receptor homologue, which are also present in MCMV (Thäle et al., 1994). The significance of these findings in an evolutionary context is discussed.

**METHODS**

**Virus and cell culture.** RCMV-E was propagated by passage on a rat embryo fibroblast (REF) cell line. REF cells were maintained in minimum essential medium supplemented with 10% fetal bovine serum, 30 μg gentamicin ml⁻¹ and 5 mM glutamine.

**3’ and 5’ rapid amplification of cDNA ends (RACE) and DNA sequencing.** Poly(A) RNA, tagged magnetically with μMACS oligo(dT) beads (Miltenyi Biotec), was extracted from REF cells that had been infected with RCMV for 24 h. RACE experiments were performed with a FirstChoice RLM-RACE kit (Ambion). After reverse transcription, the cDNA was first amplified by PCR using the 5’-RACE outer primer 5’-GCTGATGGCGATGATTGAC-CTG-3’ and a first gene-specific outer primer (5’ outer GSP) (Table 1). The cDNA was then reamplified with a 5’-RACE inner primer, 5’-GCGGATCTGACACTCGCTTTGCTGTTGATG-3’, and a second gene-specific inner primer (5’ inner GSP) (Table 1). To map the 3’ end, cDNA was amplified with the 3’-RACE outer primer 5’-GCGACACAGAATTATACGACT-3’ and a gene-specific primer (3’ GSP) (Table 1). Products were amplified for 35 cycles on a Perkin Elmer 2400 thermocycler, visualized by electrophoresis on a 2% agarose gel and cloned into the vector TOPO TA 2.1 (Invitrogen). The 3’ and 5’ cDNA ends were determined by sequencing the inserts with forward or reverse primers provided by the manufacturer on an ABI Prism 310 analyser (Perkin Elmer). Sequence data were assembled and analysed with DNASTAR (DNASTAR Inc.). Confirming information was obtained in duplicate experiments performed with 3’- and 5’-RACE kits obtained from Life Technologies.

**Northern blot analysis.** Total RNA was harvested from RCMV-infected and uninfected REF cells with TRizol isolation solvent (Life Technologies). For infection with RCMV, an m.o.i. of 1-0 p.f.u. per cell was used. After 1 h adsorption, fresh medium was added and the cultures were harvested after 24 h. To monitor for immediate-early expression, cycloheximide (50 μg ml⁻¹) was added to the medium 30 min prior to and at the time of infection of REF cells and the cultures were harvested after 6 h. To determine whether DNA replication is necessary for expression, phosphonoformic acid (PFA) (300 μg ml⁻¹) was added at the time of infection and RNA was extracted after 9 and 24 h. Northern blotting was performed as described previously (Voigt et al., 2001). For hybridization, specific fragments for vOx2, ie2, the β-chemokine homologue and SGG1 were amplified by PCR, purified and nick-translation with [α-32P]dATP.

The following primers were used for amplification of DNA templates for Northern blot analysis: vOx2, ATGGCCACGTATAGCAAA (3228)/TAACCTTAGGACACACGGGAC (3890); ie2, CGATCCAACCATATGCTCACTGGA (4241)/CACCATCGTAGTGATG (4791); β-chemokine, AGTTTCGTACTTGTAAAGGCA (5470)/ACGGGTTCGAATAGACAGTA (6051); SGG1, ACCGTCACGAGTCCGCGGTCCCGCC (7201)/ACAGACCTCATATGGCCA (7903). Numbers in parentheses delineate the position in nucleotides of the RCMV HindIII J and M fragments.

**Identification of homologous genes and potential signal peptides.** ORFs over 300 nt in length that do not overlap with other ORFs by more than 60% of their length were sought by using the DNASTAR ORF search procedure. The BLAST program was used to search for regions homologous to previously described genes. Alignments of homologous genes were generated with the help of the CLUSTAL X program and the BOXSHADE server.
RESULTS

Identification of ORFs and transcript mapping

A 12.163 kb DNA sequence encompassing the HindIII J and M fragments subcloned in plasmids from virion DNA of cell culture-grown RCMV-E was determined. In comparison with RCMV-M and MCMV, most of the genes described here are positioned collinearly within the RCMV-E genome, but there are two loci where the gene content differs. Overall, we identified eight ORFs that are, with the exception of vOX2 and ie2, oriented leftwards in the viral genome. All ORFs have been given a numerical gene name corresponding to that in MCMV (which is itself based on a comparison with HCMV); however, the protein products of five ORFs have also been named after cellular (OX2, β-chemokine, Fc receptor) or viral (ie2 and SGG1) homologues. Candidate TATA boxes were identified upstream of all eight ORFs. A summary of the overall ORF arrangements is depicted in Fig. 1(a).

Comparison of the region upstream of the MIE region in rodent CMV

In MCMV, m128 (encoding the ie2 protein) is the first gene/ORF upstream and to the right of the MIE region and it is separated from the defined MIE promoter/enhancer by 265 bp of apparently non-coding sequence. In contrast, there are additional ORFs present between the MIE and the homologous e/r128 (ie2) genes in both rat viruses, which have been named e127 (OX2) for RCMV-E and r127 (REP) for RCMV-M, respectively (van Cleef et al., 2004; Vink et al., 2000). However, the OX2 and REP homologues are very different from each other and are arranged in opposite orientations (Fig. 1a). In addition, RCMV-M contains an ORF of 302 bp (named r126) to the left of r127 that is present in neither RCMV-E nor MCMV. The length of non-coding DNA sequence between the MIE enhancer/promoter and ORF e127 is 1243 bp in RCMV-E and 619 bp in RCMV-M. The r126 and r127 ORFs are separated by 731 bp. Therefore, the spliced m128 region in MCMV is

Table 1. Primers used in 5′-RACE experiments for the identification of TATA boxes and cap sites, and in 3′-RACE experiments for the identification of polyadenylation sites

<table>
<thead>
<tr>
<th>Gene</th>
<th>Sequence (5′→3′)</th>
<th>GSP</th>
<th>Position</th>
<th>TATA position</th>
<th>Cap site</th>
</tr>
</thead>
<tbody>
<tr>
<td>e127 (vOX2)</td>
<td>CTTCCATCAGTTGCT</td>
<td>Outer</td>
<td>3355–3335</td>
<td>2684</td>
<td>2715</td>
</tr>
<tr>
<td></td>
<td>TACGACCTGAGCAGC</td>
<td>Inner</td>
<td>3274–3254</td>
<td></td>
<td></td>
</tr>
<tr>
<td>e128 (ie2)</td>
<td>GAGTAGGACCTTCCAC</td>
<td>Outer</td>
<td>4477–4457</td>
<td>3998</td>
<td>4026</td>
</tr>
<tr>
<td></td>
<td>CAGTGATGCTGTCG</td>
<td>Inner</td>
<td>4242–4222</td>
<td></td>
<td></td>
</tr>
<tr>
<td>e131/129 (ECK-2)</td>
<td>ACGTTCCTAAGTGT</td>
<td>Outer</td>
<td>5984–6001</td>
<td>6475</td>
<td>6446</td>
</tr>
<tr>
<td></td>
<td>GAGTAGGACCTTCCAC</td>
<td>Inner</td>
<td>6215–6232</td>
<td></td>
<td></td>
</tr>
<tr>
<td>e133 (SGG1)</td>
<td>ATGAAATACGCCCAATG</td>
<td>Outer</td>
<td>7471–7491</td>
<td>7934</td>
<td>7904</td>
</tr>
<tr>
<td></td>
<td>GAGGAGTTGCTGTCG</td>
<td>Inner</td>
<td>7590–7610</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gene</th>
<th>Sequence (5′→3′)</th>
<th>Position</th>
<th>(A)ATAA(A) signal</th>
<th>Poly(A) tail</th>
</tr>
</thead>
<tbody>
<tr>
<td>e127 (vOX2)</td>
<td>CACAACTAGTCAAACCTCTACA</td>
<td>3551–3574</td>
<td>3944</td>
<td>3963</td>
</tr>
<tr>
<td>e128 (ie2)</td>
<td>GCCTTGGTAAGAAATAGGCG</td>
<td>5082–5102</td>
<td>5416</td>
<td>5439</td>
</tr>
<tr>
<td>e131/129 (ECK-2)</td>
<td>ACGGTTGGAAGACTAAACAGT</td>
<td>6051–6031</td>
<td>5884</td>
<td>5864</td>
</tr>
<tr>
<td></td>
<td>CAATTTGCAGATATTCGCGG</td>
<td>5885–5865</td>
<td>?</td>
<td>52497</td>
</tr>
<tr>
<td>e133 (SGG1)</td>
<td>AAAGAAACGATGGGCCGGGTGA</td>
<td>6724–6704</td>
<td>6437</td>
<td>6415</td>
</tr>
</tbody>
</table>
replaced by OX2 (e127) in RCMV-E and by the AAV REP homologue (r127) plus r126 in RCMV-M (Fig. 1a and b).

**ORF encoding vOX2 (e127)**

The first ORF upstream of the MIE enhancer in RCMV-E is e127, which is a highly conserved homologue of rat OX2 (Fig. 1a and b). e127, like the cellular OX2, comprises an N-terminal variable domain fused to a similarly sized constant domain; this structure is characteristic of immunoglobulin superfamily (IGSF) proteins, suggesting that it interacts with other cell-surface molecules. Data obtained from RACE studies revealed major and minor start sites for transcription initiation at positions 2715 and 3093, which are both preceded by an appropriate TATA sequence (Table 1). 3’-RACE analysis identified a poly(A) signal 53 bp after the termination codon at position 3963, followed by a poly(A) tag after 19 bp, as indicated in Table 1.
Therefore, the major transcript should consist of 1248 bp, which is consistent with the 1·5 kb poly(A) mRNA species detected by Northern blot analysis, which displayed late kinetics and complete inhibition by PFA (Fig. 2a).

The entire cellular OX2 protein, including the transmembrane portion and cytoplasmic tail, is highly conserved between humans and rats (69 % identity, 79 % similarity) (see Supplementary Fig. A in JGV Online). e127 has 56 % identity and 69 % similarity to rat OX2 and 46 % identity and 64 % similarity to human OX2. The highest level of conservation occurs within the IGSF variable region (79 % identity, 87 % similarity), e127 and the Kaposi’s sarcoma-associated herpesvirus (KSHV) ORF K14 have 38 % protein identity (see Supplementary Table in JGV Online).

ORF encoding ie2 (e128)

An RCMV-E gene (e128) with strong similarity to RCMV-M r128 and MCMV m128 and lower similarity to U95 of HHV-6 and HHV-7 maps at a similar position and orientation upstream from the MIE domain as in the other rodent CMVs (Fig. 1a and b). A TATA box and cap site, as well as an AAUAAA signal and a poly(A) tail, were identified by 5′-RACE and 3′-RACE analyses (Table 1). Surprisingly, kinetic studies revealed that e128, in contrast to m128, is expressed at late times post-infection (Fig. 2b). The unspliced e128 ORF (for a comparison with m128 and r128, see Table 2) is 1430 bp in length and encodes a putative protein of 412 aa. e128 and m128 show 47 % identity, whereas r128 and e128 have 41 % identity. The amino acid sequences of all three proteins are provided in Supplementary Fig. B in JGV Online.

ORF encoding a β-chemokine (e131/129)

The RCMV-E e131/129 gene encodes a spliced, leftwards-oriented ORF, named ECK-2, that closely resembles a β-chemokine. Northern blot analysis revealed a predominant spliced transcript of 1·5 kb, encompassing the originally defined ORFs 129 and 131 (Fig. 2c). Inhibition by PFA, shown at 9 and 24 h post-infection, indicated that RCMV-E e131/129 is transcribed at delayed-early and late times...
Table 2. Intron/exon characteristics of the ie2, β-chemokine and SGG1 homologues

Data are given as lengths (bp). Intron and exon sequence data were retrieved from GenBank and compared with data generated for RCMV-E. The rat CMVs do not contain spliced versions of the ie2 gene, in contrast to the MCMV m128 homologue. Both MCMV and RCMV-E encode similarly spliced β-chemokine genes, whereas that of RCMV-M r131 is unspliced and consists of 701 bp. Recently, however, it has been reported that RCMV-M additionally encodes a spliced chemokine gene called RCK-3 (Akter et al., 2003), shown here as r131a. Therefore, RCMV-M seems to encode two β-chemokine homologues. For the SGG1 genes, predicted data suggest that there is an alternatively spliced version for both MCMV and RCMV-E (m133B and e133B), and possibly four different versions for RCMV-M (r133A–D).

Table 3. Splice site (donor/acceptor) comparison between the β-chemokine and SGG1 homologues of MCMV, RCMV-M and RCMV-E

Donor and acceptor sites were identified by analysing sequence data. As RCMV-M r131 (RCK-2) is unspliced, donor and acceptor sites are listed only for r131a (RCK-3), the additional, spliced β-chemokine homologue.

PFA-sensitive signal at 24 h post-infection in Northern blot analysis (Fig. 2c).

The ECK-2 protein initiates with a putative signal peptide that is predicted to be cleaved from the protein between amino acid positions 18 and 19. Four conserved cysteines are located at positions 22, 23, 45 and 60 in the N terminus of the predicted 306 aa sequence, which is characteristic of the β (CC) class of chemokines. The ECK-2 protein is more similar to MCMV MCK-2 (28% identity) than to the RCMV-M r131 product (22% identity; see Supplementary Fig. C in JGV Online).

ORF encoding SGG1 (e133)

The next ORF encodes a leftwards-oriented RCMV-E gene that has been named e133 (Fig. 1a). Based on the results of our 5′- and 3′-RACE studies, RCMV-E e133 contains two TATA boxes and two transcription start sites, indicating that there is one transcript of 1·5 kb and another of 1·8 kb, with a common ATG and a single, common poly(A) tail (Fig. 3a, top illustration; Table 1). Both RCMV-E transcripts are similarly spliced and contain an intron of 265 bp (Fig. 3a, intron A). The 1·5 kb RCMV-E transcript harbours a first exon of 864 bp that is separated from a second exon of 353 bp, with an in-frame stop codon, the reading frame is extended to 305 aa (see Supplementary Fig. C in JGV Online).
In addition, two alternative e133 1.5 and 1.8 kb versions with introns of 75 bp could be predicted (Fig. 3a, intron B). This alternative splice site in the second reading frame leads into exon 2B, which encodes 97 aa, resulting in a larger variant of 443 aa. The same predictions were made for m133 and r133 (Fig. 3b and c, respectively). Two putative proteins, e133A from the experimentally proven transcripts and e133B from the additional predicted versions, are given in Supplementary Fig. D in JGV Online. Predicted splice sites, as well as experimental data, show that intron and exon sizes of the three rodent SGG1 homologues are similar in size (Table 2). Although r133 has not been reported to be spliced, three other variants of the protein can be predicted. Proven or predicted splice-junction motif sequences of the 133 variants in all three viruses are given in Table 3. The RCMV-E e133 and m133 proteins have 33% identity and 46% similarity and e133 and r133 show 30% identity and 56% similarity (see Supplementary Fig. D in JGV Online).

**ORFs encoding HJ3–HJ6 (e135–e138)**

Further, we identified ORFs that are homologous to known ORFs in MCMV and RCMV-M that have been numbered e135, e136, e137 and e138, the latter being a Fcγ receptor homologue. Protein sequence alignments of the putative Fcγ receptor protein encoded by RCMV-E e138 and MCMV m138 revealed 24% identity and 42% similarity, whereas RCMV-E e138 and RCMV-M r138 show 26% identity and 41% similarity (see Supplementary Fig. E in JGV Online).

**Phylogenetic analysis of e127 and e138**

Based on the protein alignment generated by CLUSTAL X, phylogenetic trees of OX2 homologues and viral Fcγ receptor homologues were created (Fig. 4a and b). Alignment of three herpesvirus OX2 (KSHV K14, HHV-7 U85 and RCMV-E e127) and three host OX2 (rat, human and mouse) protein sequences shows that RCMV-E OX2 is clearly most...
Fig. 4. Phylogenetic analysis of (a) OX2 homologues, (b) β-herpesviral FcR receptor homologues, (c) immediate-early IE2/ie3 exon 5 sequences and (d) complete ORFs of POL proteins. Alignments were carried out with multiple-alignment parameters in CLUSTAL X (open gap penalty, 20; extend gap penalty, 1; delay divergent, 40%; gap distance, 8; similarity matrix, PAM). GenBank accession numbers are shown in parentheses.
closely related evolutionarily to the host OX2 genes and is more distant from the other viral homologues. RCMV-E e138 was more distantly related to the cellular Fcγ receptors than MCMV m138, but more closely than r138.

Phylogenetic analysis of immediate-early (IE) and DNA polymerase (POL) proteins of β-herpesviruses

To further support our interpretation that RCMV-E is a distinct species of rodent β-herpesvirus from RCMV-M, rather than the two being strains of the same species, we also carried out phylogenetic analyses with protein sequence data from the conserved exon 5-coding region of the RCMV-E IE2 gene (Sandford et al., 1993) and the almost-complete RCMV-E POL protein (equivalent to amino acid positions 146–1069 in MCMV) (W. H. Burns, unpublished data), compared with the known exon 5 IE2/iE3 and the POL proteins from a variety of other representative β-herpesviruses. We found a significantly closer relationship for both the IE2/iE3 (Fig. 4c) and POL (Fig. 4d) proteins between RCMV-E and MCMV, compared with that between the two RCMVs.

DISCUSSION

An earlier analysis of the RCMV MIE regions revealed relatively large protein sequence differences between the two RCMVs, resulting in the hypothesis that these two viruses may actually represent different RCMVs, resulting in the hypothesis that these two relatively large protein sequence differences between the two RCMVs. An earlier analysis of the RCMV MIE regions revealed that these two viruses may actually represent different β-herpesvirus species (Beisser et al., 1998). This hypothesis was further supported by the finding of a C-type lectin homologue, located at the left terminus of the RCMV-E genome, that is not present in RCMV-M (Voigt et al., 2001). Furthermore, despite the high degree of similarity between RCMV-E, RCMV-M and MCMV, our additional sequence data upstream of the MIE region now disclose further significant genetic diversity between the three viruses. This is emphasized in particular by the presence of the OX2 homologue in RCMV-E only, the presence of the AAV REP-like gene in RCMV-M only (van Cleef et al., 2004; Vink et al., 2000) and the lack of a second unspliced chemokine gene (RCK-3/r131a) in RCMV-E and MCMV, compared with RCMV-M. Both the lack of an OX2 homologue from the analogous location of RCMV-M and the closer similarity of virtually all RCMV-E proteins analysed to MCMV than to RCMV-M (Fig. 4) suggest strongly that these two viruses need to be reclassified as distinct rodent β-herpesvirus species and not simply as two strains of RCMV; it remains to be seen which, if either, is the predominant strain of RCMV in wild rats.

The first viral conserved ORF that was identified upstream from the RCMV-E MIE region has been named e128 and is a homologue of the murine gamma-herpesvirus M128 and r128. These are all members of the paralogous US22 family of genes of unknown function that occur as multiple dispersed and highly diverged copies in all sequenced β-herpesviruses. So far, no directly matching homologue for this version of the US22 family has been identified within the HCMV or other primate CMV genomes; the positionally equivalent HCMV and CCMV UL128 genes instead appear to be related more closely to chemokines (Akter et al., 2003).

In contrast with e128, which is 1430 bp in length, the RCMV-M r128 ORF comprises 1223 bp, encoding a putative 408 aa protein, and thus parallels the size of e128. However, the e128 transcript, totalling 1-65 kb, exhibits expression characteristics distinct from those of MCMV m128. Firstly, RACE analyses of e128 (Table 1) did not reveal any splicing modifications, but rather one long transcript and a single exon that parallel the structure of r128. In contrast, the m128 gene is spliced into three exons, with the third exon giving rise to a 391 aa protein that has been shown to be dispensable for growth, as well as for latency, in mice and does not show a growth phenotype when disrupted (Cardin et al., 1995; Ménard et al., 2003). Secondly, unlike the MCMV version, which is expressed most abundantly at 2 h post-infection (Messerle et al., 1991), e128 mRNA was not expressed at immediate-early times, but rather at delayed-early and late times after infection, as shown by its partial sensitivity to the addition of PFA (Fig. 2b). Kinetic studies for r128 have not been reported.

Further, we identified a homologue (ECK-2) of a gene encoding a spliced β-chemokine (CC) protein that is present in both RCMV-M (r131) (Vink et al., 2000) and MCMV (m131/129 or MCK-2) (MacDonald et al., 1999). This leftwards-oriented gene (e131/129) of RCMV-E DNA occurs in a similar position and orientation to those in MCMV (Fig. 1a). Other homologues of cellular genes encoding α (CXC)-chemokines or chemokine receptors have been described in HCMV (UL146, UL147 and UL33, UL78, US27 or US28), β (CC)-chemokine genes have been reported for rodent CMV, KSHV (Nicholas et al., 1997), HHV-6B (French et al., 1999; Lüttichau et al., 2003) and guinea pig CMV (Haggerty & Schleiss, 2002; Penfold et al., 2003). Moreover, it has been predicted that the UL128 genes of simian CMV, CCMV and HCMV may also encode spliced β-chemokines, and HCMV UL128 is modified upon passage of the virus in cell culture (Akter et al., 2003).

RCMV-E e131/129 (ECK-2) is structurally very similar to RCMV m131/129, which is also spliced, and both genes contain a similarly sized intron of approximately 80 bp. The RCMV-M β-chemokine was initially identified as unspliced (Vink et al., 2000). However, the presence of an additional, but previously unrecognized, spliced β-chemokine gene in RCMV-M (named RCK-3), located immediately adjacent to r131, has recently been reported (Akter et al., 2003). Further analysis revealed that RCK-3 is the spliced form of r129 (Kaptein et al., 2004).

The ECK-2 gene contains a highly basic insertion that is not present in MCK-2 or RCK-2/-3 (see Supplementary Fig. C in JGV Online). In comparison with ECK-2, only the
spliced MCK-2 mRNA is expressed abundantly in MCMV, although its size in Northern blot analysis is smaller (MacDonald et al., 1999).

Another gene, RCMV-E e133, is homologous to r133 and m133 (also named SGG1). The latter has been reported to be important for MCMV replication in salivary-gland acinar cells (Lagenaur et al., 1994; Manning et al., 1992). By RACE analyses, we identified two identically spliced mRNAs of 1·5 and 1·8 kb for e133. Similarly spliced mRNAs have been documented for m133. However, our sequence analyses predict another, hitherto unreported, splice variant for both e133 and m133. Moreover, we predict four possible splice variants for RCMV-M r133, which has been reported to encode a single exon and thus an unspliced gene (Vink et al., 2000) (Fig. 3 and Table 2).

Northern blot analysis of e133 showed two similarly expressed 1·5 and 1·8 kb transcripts, corresponding to those that have also been reported for MCMV m133. However, in that case, the 1·8 kb transcript was less abundant than the 1·5 kb transcript. Both e133 transcripts were readily detectable at late times post-infection (Fig. 3a), whereas only the 1·5 kb MCMV m133 transcript has been reported to be detectable at low levels late in infection (Lagenaur et al., 1994).

Homologues of RCMV-M r135 (HJ3), r136 (HJ4), r137 (HJ5) and r138 (HJ6) (Vink et al., 2000), as well as m135–m138 of MCMV (Vieira et al., 1994), were identified within the RCMV-E sequence, with all having the same arrangement and orientation (Fig. 1a). Based on similarity to cellular Fc receptor genes (see Supplementary Fig. E in JGV Online), we predict that e138, like MCMV m138 (or FCR-1) (Thäle et al., 1994), may act as a receptor for the Fc domain of immunoglobulin G (IgG). Similar genes have been described for HCMV (Atalay et al., 2002; Lilley et al., 2001). Analysis of a MCMV m138 deletion mutant showed that this virus replicates poorly in vivo, but normally in vitro, compared with wild-type MCMV (Crnković-Mertens et al., 1998).

A surprise finding was a highly conserved homologue of the rat OX2 (CD200) gene, located in the place of the AAV REP homologue found in RCMV-M (Vink et al., 2000). The host rat CD200 gene was originally defined as the target of a monoclonal antibody against rat thymocytes (McMaster & Williams, 1979; Ragheb et al., 1999) and has been shown to bind CD200R (Wright et al., 2000). The CD200–CD200R interaction is important in negative regulation of myeloid function. CD200 is present on the cell surface of a variety of cell types, including neurons, endothelial cells, B cells, T cells and follicular dendritic cells (Barclay, 1981; Barclay & Ward, 1982; Barclay et al., 1986; Wright et al., 2001).

Another, very different viral homologue of OX2 is found in two related γ-herpesviruses in the genus Rhadinovirus: KSHV VOX2 or K14 and rhesus rhadinovirus (RRV) R15 (Alexander et al., 2000; Russo et al., 1996). KSHV K14 has been reported to target and activate myeloid-lineage cells, resulting in the production of inflammatory cytokines (Chung et al., 2002). On the contrary, another analysis of the same protein demonstrated that, similarly to the host CD200 protein, K14 downregulates myeloid function (Foster-Cuevas et al., 2004). Similarly, analyses with the host mouse OX2 gene have shown that it has inhibitory effects on macrophage function (Hoek et al., 2000). Further, an antagonistic role for cellular OX2 in antigen presentation has been described (Gorzynski et al., 1999).

Less conserved OX2 homologues are also found in both HHV-6 (U85) (Gompels et al., 1995) and HHV-7 (U85) (Nicholas, 1996), as well as in several members of the family Poxviridae (Cameron et al., 1999; Kilpatrick & Rouland-Delalande, 1985; Lee et al., 2001; Tulman et al., 2001; Willer et al., 1999). A phylogenetic analysis of OX2 protein sequences from KSHV K14, HHV-7 and RCMV-E revealed that there is no reason to suggest that there is any collinearity between the three viral OX2 genes and they appear to have been captured and evolved independently of one another, with the RCMV-E version retaining the greatest similarity to the cellular versions (therefore being the most recently captured?).

Moreover, phylogenetic analyses of the conserved exon 5 segment of the IE2/ie3 protein sequences, as well as of the viral POL proteins, were carried out to further investigate the relationship of RCMV-E to RCMV-M and MCMV. Both phylogenetic trees clearly demonstrate the close relationships among HHV-6A/B and HHV-7, the primate CMV lineages and all of the rodent CMV isolates, but, importantly, they also reveal significant sequence divergence among the rodent CMV IE2/ie3 and POL proteins, suggesting that RCMV-E, RCMV-M and MCMV have been evolving as separate and distinct species since very early in the development of the rodent lineage. In general, comparison of the ie2, β-chemokine, SGG1 and Fc receptor genes between RCMV-E, RCMV-M and MCMV revealed slightly higher identities in all cases between the RCMV-E and MCMV versions than between the RCMV-M and MCMV versions, with the exception of the Fc receptor homologue (see Fig. 4 and Supplementary Fig. E in JGV Online). Remarkably, the complex, experimentally proven splicing patterns for both the RCMV-E chemokine and SGG1 homologues were not found in RCMV-M and thus are also more similar to the MCMV versions. These findings argue strongly for RCMV-E being designated as a distinct species in the genus Muromegalovirus.

Why have the two RCMVs acquired different genes and what are the possible consequences? Both were apparently isolated from Rattus norvegicus (Bruggeman et al., 1982; Priscott & Tyrrell, 1982). However, as virus species survive by occupying distinct biological niches (Davison, 2002), it is plausible that RCMV-E and RCMV-M either evolved originally within very distinct genera of rats or that they occupy different biological niches within the same host species. The former scenario is supported by the finding that
multiple CMVs isolated from *R. norvegicus* and *Rattus rattus* in Australia show similar growth phenotypes to RCMV-E and RCMV-M, respectively. However, the *R. rattus* isolates were found to be genetically different from both RCMV-E and RCMV-M (Smith et al., 2004).

The OX2 gene has been described in *R. norvegicus* (Clark et al., 1985) and it is possible that RCMV-E has captured this gene from the host genome fairly recently. The two RCMVs differ not only by their genomic content, but also by overall genome size (as measured by restriction-enzyme digests). It is not known whether freshly isolated RCMV-E may have originally resembled the size of RCMV-M and MCMV, or whether it perhaps lost genes during passage in tissue culture, as has been demonstrated for HCMV laboratory strains AD169 and Towne (Cha et al., 1996). Possibly, RCMV-E has captured cellular genes, such as OX2 and a C-type lectin homologue, to substitute for genes mapping elsewhere along the genome that may have been lost during the evolutionary process, or only one or the other of these two viruses underwent positive selection to retain certain gene functions. The absence of approximately 25 kb from the RCMV-E genome does not appear to handicap virus replication, either *in vitro* or *in vivo*. On the contrary, RCMV-E replicates even better in tissue culture than RCMV-M. It remains to be determined whether particular RCMV-E genes, such as OX2, contribute to viral fitness *in vivo* and whether these genes are advantageous to the virus. Currently, we are investigating the role of RCMV OX2 *in vitro* and *in vivo* to understand how this viral homologue functionally affects the biology of viral infection.

**ACKNOWLEDGEMENTS**

The authors thank Neil Barclay for comments on the manuscript. S.V. was supported by the Deutsche Forschungsgemeinschaft.

**REFERENCES**


