Journal of General Virology (2003), **84**, 1117–1122 DOI 10.1099/vir.0.18952-0

Short Communication

Two novel spliced genes in human cytomegalovirus

Parvis Akter,[1] Charles Cunningham,[1] Brian P. McSharry,[3] Aidan Dolan,[1] Clare Addison,[1] Derrick J. Dargan,[1] Aycan F. Hassan-Walker,[2] Vincent C. Emery,[2] Paul D. Griffiths,[2] Gavin W. G. Wilkinson[3] and Andrew J. Davison[1]

[1]MRC Virology Unit, Institute of Virology, Church Street, Glasgow G11 5JR, UK

[2]Department of Virology, Royal Free and University College Medical School, Royal Free Campus, Rowland Hill Street, Hampstead, London NW3 2QG, UK

[3]Section of Infection and Immunity, University of Wales College of Medicine, Tenovus Building, Heath Park, Cardiff CF14 4XX, UK

Correspondence
Andrew Davison
a.davison@vir.gla.ac.uk

Received 5 November 2002
Accepted 12 February 2003

Two novel spliced genes (UL131A and UL128) flanking UL130 were predicted from sequence comparisons between human cytomegalovirus (HCMV) and its closest known relative, chimpanzee cytomegalovirus (CCMV), and the splicing patterns were confirmed by mRNA mapping experiments. Both genes were transcribed with late kinetics and shared a polyadenylation site. Comparisons with wild-type HCMV in infected human tissues showed that three of five isolates passaged in cell culture contained disruptions of UL128, one was frameshifted in UL131A and one exhibited a deletion affecting UL131A and UL130. CCMV and the Colburn strain of simian cytomegalovirus, which have been passaged in cell culture, also exhibit disruptions of UL128. These observations indicate that expression of either one of UL128 and UL131A is deleterious to growth of primate cytomegaloviruses in cell culture. Although the functions of these genes are unknown, sequence comparisons suggest that UL128 encodes a β-chemokine.

The genome of the AD169 strain of human cytomegalovirus (HCMV; human herpesvirus 5) was characterized by Chee *et al.* (1990) as containing 189 putative protein-coding open reading frames (ORFs), some duplicated in an inverted repeat. An additional genome region was subsequently discovered in the Toledo strain (Cha *et al.*, 1996). A recent comparison of these sequences with that of chimpanzee cytomegalovirus (CCMV) indicated that wild-type HCMV has 166–169 genes (Davison *et al.*, 2003a, b). The present work concerns two of the eleven newly predicted genes in this redefined set.

The upper part of Fig. 1(A) depicts the arrangement of ORFs UL131–UL128 as predicted by Chee *et al.* (1990), and the lower parts show alternative predictions based on comparisons between the AD169 and CCMV sequences. UL130 is unaltered, while spliced genes replace UL131 upstream and UL129 plus UL128 downstream. One of these genes is named UL131A because it occupies the same region as UL131 but does not share any encoded amino acid sequence, since the first exon is in a different reading frame from UL131. The other spliced gene retains the designation UL128 because it shares amino acid sequence with the original UL128 but not with UL129. Fig. 1(B, C) shows detailed alignments of the AD169 and CCMV sequences in these regions. Protein-coding regions were proposed from conservation of encoded amino acid sequences, and conceptually linked together via candidate splice donor and acceptor sites. This led to the hypothesis that UL131A and UL128 comprise two and three exons, respectively.

In order to sustain this interpretation, it is necessary to propose that AD169 has a frameshift mutation (an additional residue making a tract of eight A residues) in UL131A exon 1 (Fig. 1B) and that CCMV has a frameshift mutation (an additional residue making a tract of eight C residues) in UL128 exon 1 (Fig. 1C). Davison *et al.* (2003a) confirmed the former lesion by resequencing the relevant AD169 region and comparing it with sequence obtained directly from clinical material. The proposed mutation in CCMV UL128 remains unconfirmed, since there is no information from other strains. Nevertheless, the comparative data discussed below make it a strong probability.

A range of experiments was performed in order to investigate the expression patterns of UL131A and UL128,

The GenBank accession numbers of the HCMV sequences reported in this paper are AY169795–AY169800.

Published ahead of print on 20 February 2003 as DOI 10.1099/vir.0.18952-0

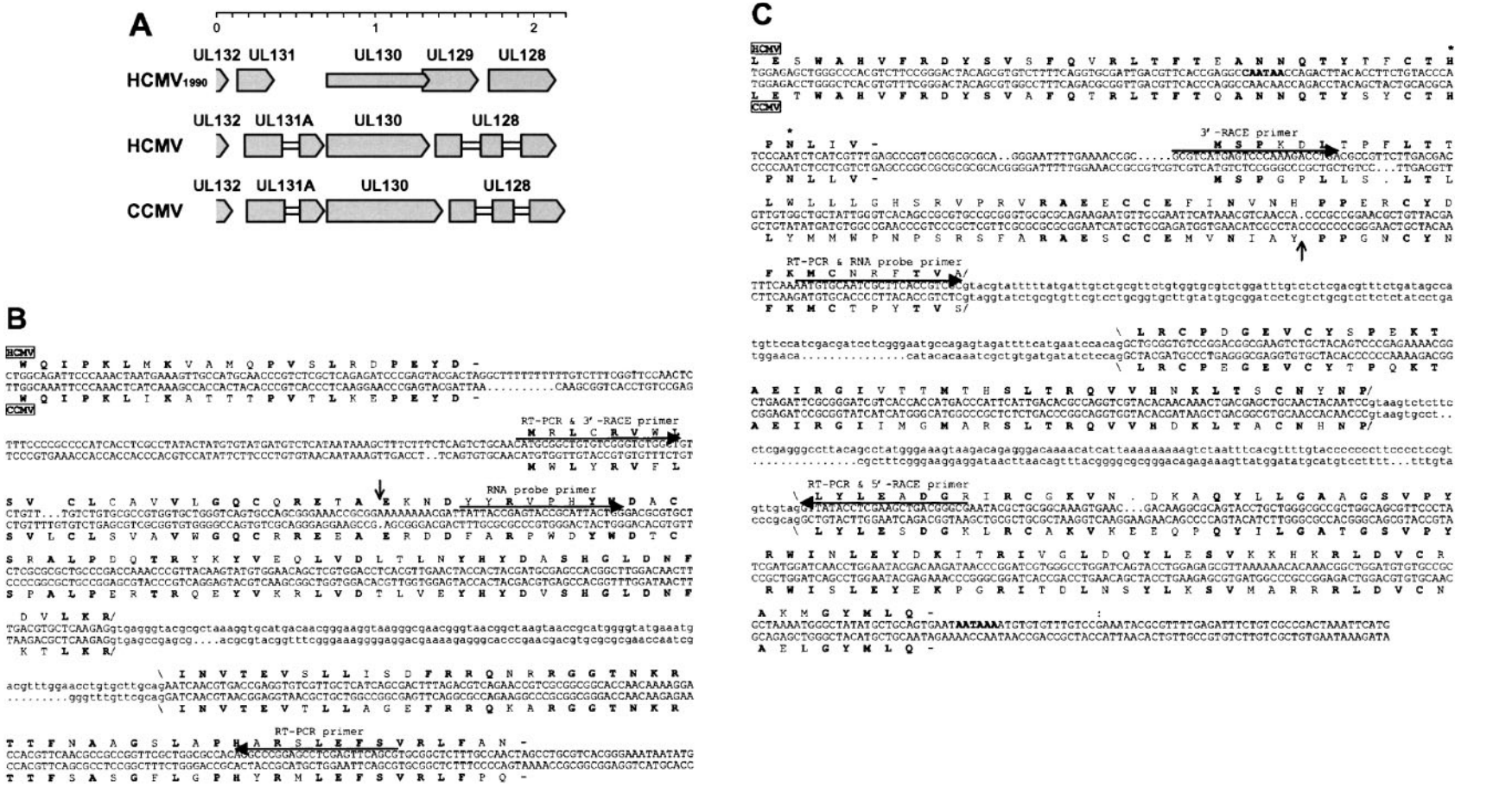

Fig. 1. (A) Arrangement of ORFs UL128–UL131 in the HCMV AD169 genome, as defined by Chee *et al.* (1990), in comparison with the HCMV and CCMV genes identified in the present study. Protein-coding regions are shaded and introns are depicted as white horizontal bars. The scale is in kbp. The locations in the AD169 and CCMV sequences are 177000–174801 (accession X17403) and 180000–177801 nucleotides (AF480884), respectively. The orientations are opposite from those of the prototype genome sequences. (B) Alignment of HCMV AD169 (upper) and CCMV (lower) sequences in the region containing UL131A and the 3′-end of UL132, oriented as in (A). The genome coordinates of the AD169 and CCMV sequences are 177000–176303 and 179974–179301 nucleotides, respectively. Intron residues are in lower case and the ends of introns are indicated by slashes. Encoded amino acid sequences for AD169 and CCMV are shown above and below the DNA sequences, respectively, with conserved residues in bold. HCMV primers used for RT-PCR, 5′- and 3′-RACE and generation of probes for Northern blotting are indicated by horizontal arrows. An additional A residue in the AD169 coding region that causes a frameshift is arrowed, and ignored in translating the sequence. (C) Alignment of HCMV AD169 (upper) and CCMV (lower) sequences in the region containing UL128 and the 3′ end of UL130, oriented as in (A). The genome coordinates of the AD169 and CCMV sequences are 175780–174801 and 178697–177745 nucleotides, respectively. The annotation scheme is the same as that in (B). The putative TATA element for HCMV UL128 and AATAAA element for UL131A, UL130 and UL128 are in bold upstream and downstream, respectively, from the UL128 coding region. The major 5′-ends of UL128 mRNAs are marked by asterisks and the 3′-end of UL128 and UL131A mRNAs by a colon. An additional C residue in the CCMV coding region that causes a frameshift is arrowed, and ignored in translating the sequence.

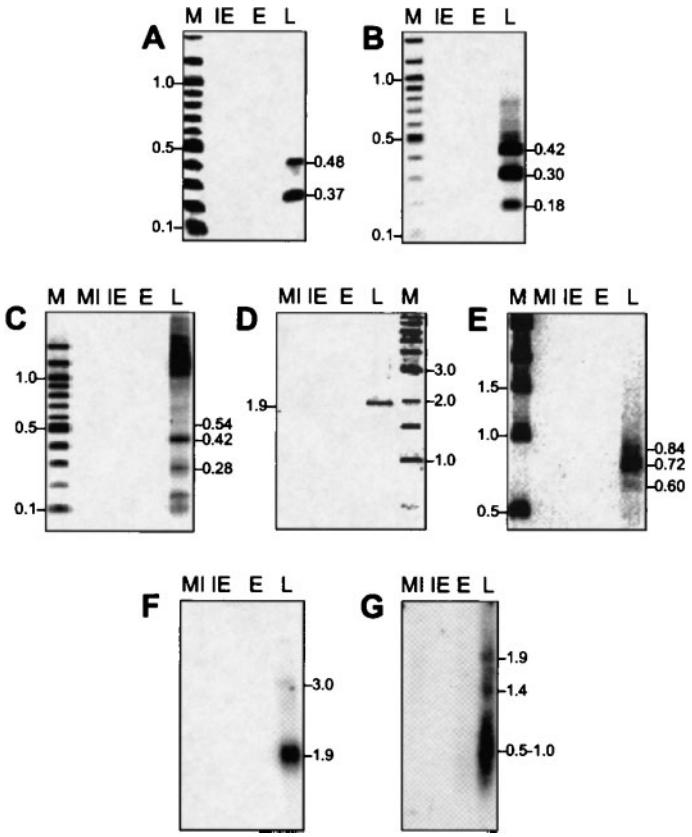

Fig. 2. Transcript mapping data for HCMV AD169 UL131A and UL128. Polyadenylated RNA purified from mock-infected (MI) or AD169-infected cells under immediate early (IE), early (E) or late (L) conditions was analysed, using primers shown in Fig. 1(B) and (C). The amounts of RNA used were 1 μg for RT-PCR and 3 μg for Northern blotting, and the equivalent of 0·2 μg for RACE. Panels (A–E) show ethidium bromide-stained agarose gels and panels (F–G) show phosphorimages of probed membranes. DNA markers (New England Biolabs; lane M) or synthetic polyadenylated RNA markers (Life Technologies; not shown) were visualized by ethidium bromide staining. Northern blot probes were ^{32}P-labelled single-stranded RNAs, extending as shown in Fig. 1(B, C) from the RNA probe primer to the appropriate RT-PCR primer. They were derived by transcription of an intron-containing PCR product for UL131A, made from DNA using the RT-PCR primers, and an intron-less RT-PCR product for UL128 generated from RNA. Sizes of PCR products (kbp) deduced from sequencing are shown in panels (A–E) and those of transcripts (kb) deduced from markers are given in panels (F–G). The smaller sizes of PCR products in panel A in comparison with the markers, as determined by sequencing, are due to uneven migration of fragments across the gel. (A) RT-PCR of UL131A. (B) RT-PCR of UL128. (C) 5′-RACE of UL128. (D) 3′-RACE of UL131A. (E) 3′-RACE of UL128. (F) Northern blot of UL131A. (G) Northern blot of UL128.

using various RNA preparations, primers and probes. A selection of results is shown in Fig. 2 and relevant primers and deduced transcriptional features are included in Fig. 1(B, C). RNA was prepared from human foetal fibroblasts mock-infected or infected with AD169 at an m.o.i. of 5. Infections were carried out under immediate early (1 h preinfection then 24 h in 200 μg cycloheximide ml^{-1}), early (48 h in 300 μg phosphonoacetic acid ml^{-1}) and late conditions (72 h with no inhibitor). RNA was extracted using TRIzol (Life Technologies), and the polyadenylated fraction was isolated using oligo(dT)–cellulose and quantified by spectrophotometry. RNA integrity was assessed by Northern blotting using a cellular mRNA probe (not shown), and the absence of detectable viral DNA was confirmed by PCR. RT-PCR was carried out using a Titan kit (Boehringer Mannheim) and RACE using a SMART RACE kit (Clontech). 5′-RACE involved extension of an oligo(dT)-containing primer by a reverse transcriptase that adds a tract of C residues at the 3′-end of the cDNA, second strand cDNA synthesis primed by the SMART oligonucleotide which has a tract of G residues at the 3′-end, and PCR using SMART-specific and gene-specific primers. 3′-RACE involved reverse transcription using an oligo(dT) primer extended at its 5′-end by the SMART sequence, followed by PCR. This approach results in 5′-RACE and 3′-RACE products that are 30 and 55 bp longer, respectively, than the cognate transcribed sequences. All RT-PCR and RACE products relevant to locating mRNA ends and introns

were cloned into pGEM-T (Promega), and several clones of each product sequenced. Northern blotting was performed using strand-specific RNA probes prepared using Lig'nScribe and MAXIscript kits (Ambion).

The 0·48 and 0·37 kbp RT-PCR products generated from late RNA using UL131A primers (Fig. 2A) correspond, respectively, to unspliced RNA and RNA spliced as predicted in Fig. 1(B). The 0·42 kbp RT-PCR product generated from late RNA using UL128 primers corresponds to unspliced RNA (Fig. 2B), the 0·18 kbp product to RNA spliced as predicted in Fig. 1(C) and the 0·30 kbp product to partially spliced RNA lacking only the first intron. 5′-RACE of UL128 (Fig. 2C) generated a 0·42 kbp product corresponding to fully spliced RNA. The 5′-ends mapped in this experiment and a second using a different primer are shown in Fig. 1(C). A 0·54 kbp product potentially corresponds to partially spliced RNA, but was very minor in amount and not analysed further. The 0·28 kbp product located additional 5′-ends in the first exon (not shown in Fig. 1C), but their significance was discounted as an equivalent product was not detected using a second primer. The two smaller products in Fig. 2(C) did not represent authentic 5′-ends as they were not primed by the SMART oligonucleotide. Larger fragments were not investigated, but presumably originated from UL130, since attempts at mapping the 5′-end of the UL131A transcript using several primers were unsuccessful. The 1·9 kbp 3′-RACE product from late RNA (Fig. 2D) mapped the 3′-end of the UL131A transcript downstream from a polyadenylation signal (AATAAA) near the 3′-end of UL128. 3′-RACE using a UL128 primer generated three fragments from late RNA (Fig. 2E). The 0·72 and 0·60 kbp products terminated at the same 3′-end as the UL131A transcript, the smaller fully spliced and the larger lacking only the first intron of UL128. The 0·84 kbp fragment was not analysed, but presumably originated from unspliced RNA. These results indicate that UL131A and UL128 (and presumably UL130) are 3′-coterminal. In Northern blotting, a UL131A probe hybridized to major 1·9 kb and minor 3·0 kb late mRNAs (Fig. 2F). The former is probably the UL131A transcript and the latter a readthrough transcript from UL132. A UL128 probe hybridized to major RNAs of 0·5–1·0 kb (Fig. 2G), a range and heterogeneity of sizes consistent with the locations of the transcript termini, inefficient splicing of the second intron and length variation in mRNA 3′-polyA tails. Minor late transcripts of approximately 1·4 and 1·9 kb probably originated from UL130 and UL131A, respectively. No evidence for splicing between UL131A and UL128 was found by 3′-RACE or Northern blotting (Fig. 2D and F) or by RT-PCR (not shown).

The transcript mapping data support the expression pattern of UL131A and UL128 anticipated in Fig. 1(A), with both genes in the late kinetic class. Although UL131A consists of two exons and UL128 of three, unspliced and partially spliced RNAs were also detected. Our results are somewhat at variance with those of Chambers *et al.* (1999), who, using microarray technology, classified transcripts from UL131, UL130 and UL128 as defined by Chee *et al.* (1990) in the late, early-late and early classes, respectively. However, kinetic class assigned from microarray data differs from that deduced from Northern blot data for a significant number of genes (Chambers *et al.*, 1999).

The sequences of UL131A and UL128 in AD169, which has been passaged many times in human fibroblast cell lines, were compared with those in six other HCMV strains. Four had been grown in human fibroblast cell lines: three of these (Merlin, 3157 and 6397) derived in Cardiff by three passages from urine samples from congenitally infected infants and one the widely used low passage Toledo strain (Quinnan *et al.*, 1984). DNA from the Cardiff strains was obtained from purified virions and from Toledo as infected-cell DNA. DNA was also prepared directly from clinical material for two strains, one (W) from the lung of an HCMV-infected AIDS patient and one (3301) from the urine of a congenitally infected child. Two overlapping fragments of about 4 kbp were PCR-amplified from five of the DNA samples and cloned into pGEM-T. The primers used were 5′-TGCTTAAGCCAATCGCAGCG-3′ (in UL147) and 5′-ATCCCGCGAATCTCAGCCGT-3′ (UL128 exon 2), and 5′-AATGTTGCGAATTCATAAACGTCA-3′ (UL128 exon 1) and 5′-ACTGGTCAGCCTTGCTTCTAGTCA-3′ (UL123). For each strain, the inserts in four plasmids were sequenced on both strands and a consensus established to exclude PCR artefacts. Corresponding data for the sixth strain (Merlin) were obtained as part of shotgun cloning the entire genome in M13. Sequences were compiled using PREGAP4 and GAP4 (Staden *et al.*, 2000) and PHRED (Ewing & Green, 1998; Ewing *et al.*, 1998). The region containing the genes of interest was analysed using the GCG suite (Accelrys), and the corresponding AD169 sequence was included.

The predicted arrangement of UL131A, UL130 and UL128 is shown for each strain in the upper part of Fig. 3(A). These genes were intact in two non-passaged strains (3301 and W), but disruptions were apparent in the five passaged strains. As explained above, AD169 has a frameshift mutation in UL131A exon 1 that would cause fusion of the N terminus of the UL131A protein to sequences encoded in another reading frame. Merlin has a C to T transition in UL128 exon 3 that introduces a stop codon and would cause premature translational termination. 3157 had a G to C transversion in the GT dinucleotide of the splice donor site at the end of UL128 exon 1, which would result in lack of splicing. 6397 has a 1 kbp deletion that would abolish expression of UL131A and UL130. In Toledo, inversion of a substantial region results in disruption of UL128 by introducing UL148A in place of UL128 exon 3. Fig. 3(A) also recapitulates observations made by Davison *et al.* (2003a) on the equivalent region in the genomes of two other primate cytomegaloviruses, CCMV and the Colburn strain of simian cytomegalovirus (SCMV), both of which have been passaged in human fibroblasts. CCMV UL128 exon 1

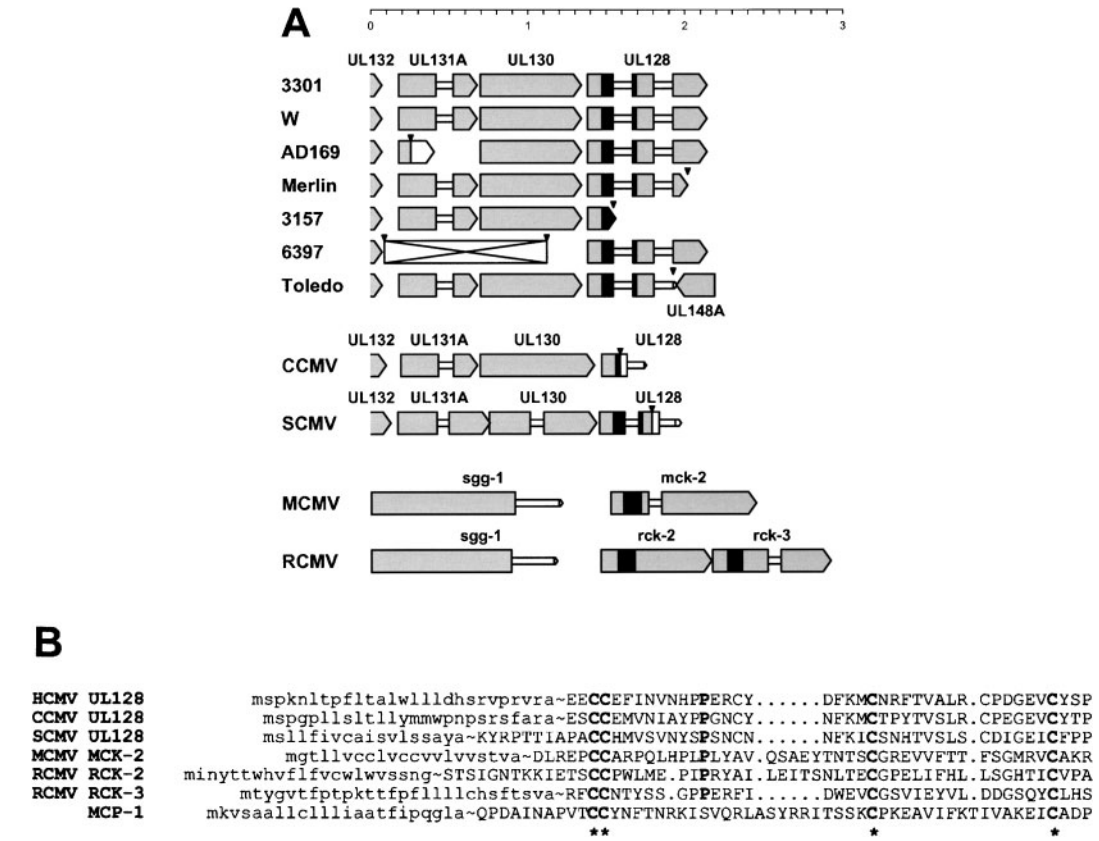

Fig. 3. (A) Coding regions in the UL131A–UL128 region of seven HCMV strains, CCMV and SCMV, in comparison with the β-chemokine-encoding region of MCMV and RCMV. The scale is in kbp, protein-coding regions are shaded and introns are depicted as white horizontal bars. Mutations are indicated by inverted filled triangles; 6397 lacks a sequence depicted by the diagonally crossed rectangle and Toledo is disrupted by an inversion about the indicated point and a point 14·4 kbp distant (not shown). The predicted effects of the mutations are illustrated in terms of loss of expression (lack of shaded coding region) or frameshift into another reading frame (unshaded coding region). Regions characteristic of β-chemokines, extending from the first to the fourth conserved cysteine residues in (B) are shaded black. The locations of the regions shown in the AD169 and CCMV sequences are 177000–174801 and 180000–177801 nucleotides, respectively. Other sequences were obtained from the following sources: SCMV, Chang *et al.* (1995), U38308; MCMV, Rawlinson *et al.* (1996), U68299; RCMV, Vink *et al.* (2000), AF232689. The splice sites mapped by Lagenaur *et al.* (1994) for MCMV *sgg-1* are shown here, but comparisons with RCMV support an alternative donor site that would extend both exons. In the alternative case, the end of exon 1 is at nucleotide 188696 and 183023 on the lower strand of the MCMV and RCMV genome sequences, respectively. The *sgg-1* gene has no homologue in primate cytomegaloviruses. (B) Amino acid sequence alignment of the predicted N termini of the HCMV, CCMV and SCMV UL128 proteins with the MCMV mck-1 β-chemokine and two related RCMV proteins. For each protein, the signal peptide predicted by SIGNALP 2.0 (Nielsen *et al.*, 1997) is in lower case with the cleavage site denoted by a tilde (~). Conserved residues are in bold, and the four cysteine residues characteristic of β-chemokines (exemplified by MCP-1, human monocyte chemoattractant protein 1; NM_002982) are asterisked.

and SCMV UL128 exon 2 appear to be frameshifted, the former by the gain of a single nucleotide, as explained above, and the latter by a single nucleotide deletion. We conclude that passage of HCMV isolates in cell culture is associated with the loss of function of UL128 or UL131A or, perhaps, UL130. Indeed, the mutated UL128 in Merlin at passage 3 was also the only form detected at the end of passage 1, during which the virus underwent many rounds of replication, but was not detected in the urine sample from which the strain was isolated. Similarly, the deletion in 6397 was not detected in unpassaged material.

Passage of HCMV strains in human fibroblasts is associated with loss of ability to grow in endothelial cells (Sinzger *et al.*, 1999; Revello *et al.*, 2001), but our study does not address whether UL131A, UL130 or UL128 are required for endotheliotropism. The functions of these genes are unknown, but each of the HCMV, CCMV and SCMV proteins commences with a predicted signal peptide, suggesting that they are secreted. Moreover, the UL128 proteins share four conserved cysteine residues near their N termini that are characteristic of β- (or CC-) chemokines (Fig. 3B). β-chemokine genes have not been predicted previously in

primate cytomegaloviruses, but have been characterized in murine cytomegalovirus (MCMV) and rat cytomegalovirus (RCMV) (Fig. 3A). An MCMV β-chemokine gene (mck-1), originally described by MacDonald *et al.* (1997), was found to be the first exon in a spliced gene designated mck-2 (MacDonald *et al.*, 1999; Fleming *et al.*, 1999). A peptide containing the conserved cysteine domain exhibits chemokine-like properties (Saederup *et al.*, 1999), and the mck-2 protein has a role in disseminating virus in the host via leukocyte recruitment (Fleming *et al.*, 1999; Saederup *et al.*, 1999, 2001). The RCMV genome contains two β-chemokine genes: rck-2, an unspliced counterpart of mck-2 (Vink *et al.*, 2000) and rck-3, an adjacent spliced gene that has not been recognized previously. Evidence that the primate cytomegalovirus UL128 and rodent cytomegalovirus β-chemokine genes represent descendants of a single gene capture event is equivocal, since, although they correspond in orientation and approximate location, adjacent genes are not homologous (Fig. 3A). Nonetheless, the hypothesis that HCMV encodes a β-chemokine, in addition to the reported α-chemokine (Penfold *et al.*, 1999), may open a new avenue of research into primate cytomegalovirus immunobiology.

ACKNOWLEDGEMENTS

P. A. was a Commonwealth Scholar during the course of these studies. This work was also supported by the Medical Research Council and the Wellcome Trust. We are grateful to our colleagues Giuseppe Gerna and Fausto Baldanti (IRCCS Policlinico San Matteo, Pavia) and Gabriele Hahn (Max von Pettenkofer Institute, Munich) for useful discussions and for generously sharing their extensive data on the genes investigated in our study. We also thank Cardiff Public Health Laboratory and Lynne Neale for primary culture of HCMV isolates and Duncan McGeoch for helpful comments on the manuscript.

REFERENCES

Cha, T. A., Tom, E., Kemble, G. W., Duke, G. M., Mocarski, E. S. & Spaete, R. R. (1996). Human cytomegalovirus clinical isolates carry at least 19 genes not found in laboratory strains. *J Virol* **70**, 78–83.

Chambers, J., Angulo, A., Amaratunga, D. & 9 other authors (1999). DNA microarrays of the complex human cytomegalovirus genome: profiling kinetic class with drug sensitivity of viral gene expression. *J Virol* **73**, 5757–5766.

Chang, Y., Jeang, K. T., Lietman, T. & Hayward, G. S. (1995). Structural organization of the spliced immediate-early gene complex that encodes the major acidic nuclear (IE1) and transactivator (IE2) proteins of African green monkey cytomegalovirus. *J Biomed Sci* **2**, 105–130.

Chee, M. S., Bankier, A. T., Beck, S. & 12 other authors (1990). Analysis of the protein coding content of the sequence of human cytomegalovirus strain AD169. *Curr Top Microbiol Immunol* **154**, 125–169.

Davison, A. J., Dolan, A., Akter, P., Addison, C., Dargan, D. J., Alcendor, D. J., McGeoch, D. J. & Hayward, G. S. (2003a). The human cytomegalovirus genome revisited: comparison with the chimpanzee cytomegalovirus genome. *J Gen Virol* **84**, 17–28.

Davison, A. J., Akter, P., Cunningham, C. & 7 other authors (2003b). Homology between the human cytomegalovirus RL11 gene family and human adenovirus E3 genes. *J Gen Virol* **84**, 657–663.

Ewing, B. & Green, P. (1998). Base-calling of automated sequencer traces using PHRED. II. Error probabilities. *Genome Res* **8**, 186–194.

Ewing, B., Hillier, L., Wendl, M. C. & Green, P. (1998). Base-calling of automated sequencer traces using PHRED. I. Accuracy assessment. *Genome Res* **8**, 175–185.

Fleming, P., Davis-Poynter, N., Degli-Esposti, M., Densley, E., Papadimitriou, J., Shellam, G. & Farrell, H. (1999). The murine cytomegalovirus chemokine homolog, m131/129, is a determinant of viral pathogenicity. *J Virol* **73**, 6800–6809.

Lagenaur, L. A., Manning, W. C., Vieira, J., Martens, C. L. & Mocarski, E. S. (1994). Structure and function of the murine cytomegalovirus *sgg*1 gene: a determinant of viral growth in salivary gland acinar cells. *J Virol* **68**, 7717–7727.

MacDonald, M. R., Li, X.-Y. & Virgin, H. W., IV (1997). Late expression of a β chemokine homolog by murine cytomegalovirus. *J Virol* **71**, 1671–1678.

MacDonald, M. R., Burney, M. W., Resnick, S. B. & Virgin, H. W., IV (1999). Spliced mRNA encoding the murine cytomegalovirus chemokine homolog predicts a β chemokine of novel structure. *J Virol* **73**, 3682–3691.

Nielsen, H., Engelbrecht, J., Brunak, S. & von Heijne, G. (1997). Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng* **10**, 1–6.

Penfold, M. E., Dairaghi, D. J., Duke, G. M., Saederup, N., Mocarski, E. S., Kemble, G. W. & Schall, T. J. (1999). Cytomegalovirus encodes a potent α chemokine. *Proc Natl Acad Sci U S A* **96**, 9839–9844.

Quinnan, G. V., Jr, Delery, M., Rook, A. H. & other authors (1984). Comparative virulence and immunogenicity of the Towne strain and a nonattenuated strain of cytomegalovirus. *Ann Intern Med* **101**, 478–483.

Rawlinson, W. D., Farrell, H. E. & Barrell, B. G. (1996). Analysis of the complete DNA sequence of murine cytomegalovirus. *J Virol* **70**, 8833–8849.

Revello, M. G., Baldanti, F., Percivalle, E., Sarasini, A., De-Giuli, L., Genini, E., Lilleri, D., Labò, N. & Gerna, G. (2001). *In vitro* selection of human cytomegalovirus variants unable to transfer virus and virus products from infected cells to polymorphonuclear leukocytes and to grow in endothelial cells. *J Gen Virol* **82**, 1429–1438.

Saederup, N., Lin, Y.-C., Dairaghi, D. J., Schall, T. J. & Mocarski, E. S. (1999). Cytomegalovirus-encoded β chemokine promotes monocyte-associated viremia in the host. *Proc Natl Acad Sci U S A* **96**, 10881–10886.

Saederup, N., Aguirre, S. A., Sparer, T. E., Bouley, D. M. & Mocarski, E. S. (2001). Murine cytomegalovirus CC chemokine homolog MCK-2 (m131–129) is a determinant of dissemination that increases inflammation at initial sites of infection. *J Virol* **75**, 9966–9976.

Sinzger, C., Schmidt, K., Knapp, J., Kahl, M., Beck, R., Waldman, J., Hebart, H., Einsele, H. & Jahn, G. (1999). Modification of human cytomegalovirus tropism through propagation *in vitro* is associated with changes in the viral genome. *J Gen Virol* **80**, 2867–2877.

Staden, R., Beal, K. F. & Bonfield, J. K. (2000). The Staden package, 1998. *Methods Mol Biol* **132**, 115–130.

Vink, C., Beuken, E. & Bruggeman, C. A. (2000). Complete DNA sequence of the rat cytomegalovirus genome. *J Virol* **74**, 7656–7665.