J. Vet. Malaysia (2007) 19(1): 21 - 32

MOLECULAR CHARACTERISATION AND IDENTIFICATION OF THE MAJOR IMMEDIATE-EARLY (MIE) GENES OF RAT CYTOMEGALOVIRUS (RCMV)

H.S.Loh[1], M.L. Mohd-Azmi[2], A.R. Sheikh-Omar[3], M. Zamri-Saad[3] and N.A. Zeenathul[3]

[1]School of Biosciences, Faculty of Health and Biological Sciences, University of Nottingham Malaysia Campus Jalan Broga 43500 Semenyih, Selangor, Malaysia
[2]Institute of Bioscience, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia
[3]Department of Veterinary Pathology and Microbiology, Faculty of Veterinary Medicine Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia

SUMMARY

The paper first describes the characterisation and identification of two transcripts of major immediate-early (MIE) genes of rat cytomegalovirus (RCMV) strain ALL-03 using molecular approach. In the absence of any known sequences in the strain ALL-03, IE genes, the first recognised viral genes expressed during infection were the most important genetic source for the exploration of the viral genome. The steps taken were the construction of cDNA libraries from rat embryonic fibroblast (REF) infected with RCMV strain ALL-03 under 'IE' condition by using several techniques such as RT-PCR, PCR, cloning, sequencing and bioinformatics. Two transcripts were confirmed to be of RCMV strain ALL-03 origin by specific PCR and hybridisation analyses. Based on the bioinformatics analyses, it is most likely, that these transcripts are exon 4 (IE1) and exon 5 (IE2) genes. The considerable divergence in the nucleotide sequence of the two IE transcripts from those of MIE regions of both Maastricht and English strains further supports the claim that the ALL-03 is a distinct RCMV strain indeed.

Keywords: Molecular characterisation, MIE, RCMV, strain ALL-03, molecular approach

INTRODUCTION

Cytomegaloviruses (CMVs) are species-specific and belong to subfamily *Betaherpesvirinae* of the *Herpesviridae* family. They cause acute, latent, and persisting infections in human and animals. Infection in immuno-compromised patients can be very severe and sometimes fatal. Congenital CMV infection is a leading cause of mental retardation and other birth defects in the human.

In this paper, we describe the generation of immediate-early (IE) transcripts of RCMV strain ALL-03, which were originally isolated from the placenta and uterus of house rats (Loh *et al.*, 2003) and evidenced to possess the capabilities of crossing the placenta and infecting the offsprings congenitally (Loh *et al.*, 2006).

Since this is a novel strain, there is no sequence data available. Primers for generation of IE genes in both Maastricht (Beisser *et al.*, 1998) and English (Sandford *et al.*, 1993) RCMV could not amplify the genes in ALL-03. However, the IE mRNA can be transcribed under conditions that inhibit protein synthesis as described in previous studies (Sandford *et al.*, 1993; Beisser *et al.*, 1998).

The major immediate-early (MIE) genes are the first viral genes expressed during infection and believed to regulate expression of both viral and some cellular genes (Boldogh *et al.*, 1990; 1991). The MIE gene products may also be important in initiating or maintaining the latent state (Honess *et al.* 1989). They are also of great interest because the MIE upstream regulatory sequences function as strong enhancer/promoters, which have found widespread use as constitutive promoters for heterologous genes in a wide variety of cell types (Sandford and Burns, 1996).

To date, the whole genome of RCMV strain Maastricht had been completely sequenced (Vink *et al.*, 2000), and for the MIE gene locus of RCMV strain English, only the sequence information is available (Sandford *et al.*, 1993; Sandford and Burns, 1996). The MIE locus of RCMV consists of a promoter/enhancer region that directs transcription of a region consisting of five exons. These exons are differentially spliced to generate transcripts encoding two separate proteins, IE1 and IE2. The IE1 and IE2 transcripts share the first three exons, the first of which is noncoded. Splicing of these exons to either exon 4 or 5 generates IE1 and IE2 transcripts, respectively (Standford *et al.*, 1993). The architecture of the MIE region of RCMV resembles that of HCMV and MCMV and there are striking homologies in exon 5 of IE2s of HCMV and RCMV and the analogous IE3 of MCMV (Messerle *et al.*, 1992).

Molecular characterisation of MIE region in this study might lead to the generation of rat-specific

Corresponding Author: Dr. Loh Hwei San: Email: *Sandy.Loh@nottingham.edu.my*

expression vectors driven by the RCMV MIE gene promoter-enhancer (Sanford and Burns, 1996), and candidate 'viral-vectored immuno-contraception' for rat control (Mohd-Azmi *et al.*,1999). Additionally, possibly it will provide animal model facilitate and intervention studies for the prevention of congenital CMV infection in the human.

MATERIALS AND METHODS

In Vitro mRNA generation and first strand cDNA synthesis

Following infection with ALL-03 using multiplicity of infection (MOI) of 4 $TCID_{50}$/cell and a 5- or 10-hr incubation in the presence of cycloheximide (50μg/ml), PolyA$^+$ RNA of RCMV-infected rat embryonic fibroblast (REF) cells were directly isolated by using Micro-FastTrack™ mRNA Isolation Kit (Invitrogen) according to manufacturer's instruction. First strand cDNA was synthesised using the SMART™ PCR cDNA Synthesis Kit (Clontech) and treated with RNase H (Invitrogen) prior to PCR amplification.

PCR1 amplification of suspected IE cDNAs and purification

The reaction mixture in volume of 50μl contained 1X Advantage 2 PCR buffer, 0.2 mM dNTP mix, 0.4μM PCR1 primer (5'-AAGCAGTGGTAACAACGCAGAGT-3'), 0.05-0.1μg cDNA, 1X Advantage 2 polymerase mix and nuclease-free water. The reaction conditions include an initial denaturation step at 95°C for 1 min, 28 cycles of 15-sec denaturation at 95°C, 30-sec annealing at 65°C, 3-min extension at 68°C. A final extension was done at 68°C for 3 min. Amplified PCR products were electrophoresed in 1.2% agarose gel and purified using the QIAquick® Gel Extraction Kit (Qiagen).

Cloning of the suspected IE amplicons and plasmid characterisation

The purified PCR product and pCR®2.1-TOPO vector were ligated and transformed in competent cells (TOP 10 *E. coli*) according to the instruction manual of the TOPO TA Cloning® Kit (Invitrogen). Plasmid DNA was isolated using the standard method (Sambrook *et al.*, 2001) from individual white *E. coli* colonies. Plasmids were analysed by restriction endonuclease analysis (digest with *Eco*RI), PCR amplification using PCR1 primer and universal M13 primers.

DNA sequence assembly and bioinformatics

The purified positive clones were sequenced using both M13 forward and reverse primers. Each sequencing data was assembled and edited to remove the sequences originated from plasmid, PCR1 primer and poly-As at 3' end using bioinformatics tools such as DNAsis and BioEdit analyses. Therefore, the original nucleotide of the insert was retained. The net sequences of the original cDNAs were checked for identity and origin by comparing to the other sequences available from databanks such as GenBank, EMBL, DDBJ and PDB by BLAST (Basic Local Alignment Search Tool) search program of National Centre for Biotechnology Information (*http://www.ncbi.nlm.nih.gov/*). Sequences of REF cDNAs identified by BLAST search (best matched) were excluded from the subsequent studies since these were not of interest. The unpronounced sequences (not significantly identical to either sequences from databanks) or so called suspected RCMV IE genes were compared to other strains of RCMV for homology search. The comparison of sequences was conducted by using Pairwise Alignment (Optimal Global Alignment) of BioEdit software as well as DNA Homology Search of DNAsis software. The nucleotide databases from GenBank used for the comparison were of English strain RCMV with accession number: U62396 and Maastricht strain RCMV with accession number: AF232689. The MIE cDNA nucleotides were of greatest interest for comparison purposes.

Design of IE-specific primers

Two sets of primers were designed from the known suspected IE sequences, i.e. BIE flanking the position 259 to 827 in transcript [IE05]; whereas DPC flanking the position 119 to 878 in transcript [IE10]. The sequences of BIE primers are 5'-CACAGAGATCTCACTAA CCTGCCACCTATAACCAC-3' (forward) and 5'-TCCAGCAGACTTCTGTATCCTGATTCAAG-3' (reverse). The sequences of DPC primers are 5'-AGTGAATCTCTGCCACAGGAGTGCT-3' (forward) and 5'-GAGGATGTGTGGAAACCCTCAC-3' (reverse). These primers were used for PCR to check the origin of the cDNAs from either REF or RCMV.

Isolation of cellular and viral genomic DNA

Extracellular virus was collected from REF cell culture supernatant 5-10 days after RCMV infection. Viral DNA was extracted as described by Lai *et al.* (1999). Meanwhile, the DNAzol® Genomic DNA Isolation Reagent (Molecular Research Centre) was carried out using REF genomic DNA extraction according to manufacturer's instructions.

PCR2 amplification

The PCR2 cocktail in volume of 50μl consisted of 1X optimised buffer, 0.2mM dNTP mix, 0.5μM each of BIE or DPC primer set, 2 units of DyNAzme II polymerase

(Finzymes), denatured template either recombinant plasmid (0.1 ng) or genomic DNA (100 ng) and nuclease-free water. The amplification of DNA was achieved in 40 cycles with 1-min denaturation at 94°C, 30-sec annealing at 69°C (BIE) or 65°C (DPC) and 1-min extension at 72°C. This was followed by a final extension step of 1 min at 72°C. The amplified PCR products were analysed in 1.2% agarose gel.

Dot blot hybridisation

The extracted genomic DNA from both REF and RCMV were digested individually with *Hin*dIII (Fermentas) to facilitate binding of DNA onto the BioDyne™ A membranes. Approximately 10ng plasmid and digested genomic DNA ranging 2.5-10μg were used per blot. Amplicons generated from recombinant plasmids, BIE-[IE05] and DPC-[IE10] were purified separately by QIAquick® Gel Extraction Kit (Qiagen) prior to preparation of biotinylated probes by using BIOPRIME® DNA Labeling System (GibcoBRL®, Life Technologies). Following the removal of unincorporated nucleotides by using ProbeQuant™ G-50 Micro Columns (Amersham Pharmacia Biotech), the purified biotinylated probes were hybridised to DNA blotted membranes under specified conditions. Membranes were first hydrated in 2x SSC solution. A total of 100μl prehybridisation (5X SSC, 5X Denhardt's, 20 mM Na_2HPO_4, 0.5mg/ml freshly denatured sheared herring sperm DNA, 10% PEG and 1% SDS) and hybridisation (5X SSC, 1X Denhardt's, 20mM Na_2HPO_4, 0.2mg/ml freshly heat-denatured sheared herring sperm DNA, 300ng/ml freshly heat-denatured probe DNA, 10% PEG and 1% SDS) solution per cm^2 of membrane were used for the 4-h pre-hybridisation and overnight hybridisation at 55°C, respectively. A large amount of stringency washings was carried out as follows: (i) washed twice with 2X SSC, 0.1% SDS for 5 min each, (ii) washed twice with 0.2X SSC, 0.1% SDS for 5 min each, and (iii) washed twice with 0.16X SSC, 0.1% SDS for 15 min each at 50°C. After briefly rinsing with 2X SSC, the membranes were proceeded to detection step by using BluGene® Nonradioactive Nucleic Acid Detection System (GibcoBRL®, Life Technologies) according to the instruction manual.

RESULTS

PCR1 amplification of suspected IE cDNAs

Following two separate attempts of *in vitro* transcription of RCMV in REF, there were sixteen amplicons ranging from 450 bp to 2.1 kbp in total generated by PCR1 amplification on the suspected IE cDNAs using PCR1 primer. These amplicons were cloned and sequenced.

DNA sequence identification and comparison analyses

Based on the homology search using BLAST program, the identities of the sixteen cDNA nucleotide sequences were revealed. Fourteen cDNAs originated from REF cells, that is, of *Rattus norvegicus* homologue with the identity ranging from 81 to 99%. Some were of the same identity, flanking within the larger fragments. These cDNAs were *Rattus norvegicus* procollagen (Colla2), phospholipids scramblase 1 (Plscr), LOC293580, similar to ribosomal protein L18a, proteasome (Psmb3), calpactin (S100a10), ADP-ribosylation factor 2 (Arf2), gro (Gro1), annexin 1 (Anxa1), intestinal type II membrane glycoprotein 4F2hc and Ca^{2+}-ATPase plasma membrane 4 (Atp2b4). The other two cDNA sequences, [IE05] and [IE10] showed no significant matches throughout the databanks with any subjects including the entries of *Rattus norvegicus* and viruses. When compared to the *R. norvegicus* entry, the matching sequences were very short, ranging from 17 to 23bp with 95-100% identities and high expect values of 0.092-5.6. The matching sequences were almost the same when compared to virus entries, that is, 18-22bp with 95-100% identities and high expect values of 0.043-6.8. The nucleotide sequences of [IE05] and [IE10] with the sizes of 943bp and 914bp respectively are shown in Figures 1 and 2.

The [IE05] and [IE10] were compared to the MIE gene regions of RCMV strains, English and Maastricht using DNAsis and BioEdit analyses. Tables 1 and 2 depict the comparison of [IE05] and [IE10] independently to the exons of MIE regions of RCMV strains by DNAsis and BioEdit, respectively. Based on the size of the sequences (which exceeded 400bp), [IE05] and [IE10] were most probably either exon 4 or exon 5 because the sizes of exon 1, exon 2 and exon 3 of RCMV strains, English and Maastricht were short fragments, that is, 104-396 bp. Based on the DNA sequence comparisons, none of any two sequences were shown with significant homology to each other including the two RCMV strains, English and Maastricht. The identity values of less than 0.5 for comparison of two sequences of any exon are shown, obtained from BioEdit program (Table 2). The DNAsis program also revealed a low homology over size of sequence, that is, less than 52% of any two exon sequences (Table 1). The aligned homology and identity between the [IE05] and [IE10] were 45% and 0.44, respectively. In order to obtain the maximum nucleotide homology and to which exons that the sequences fitted in mostly, the [IE05] and [IE10] were probed individually to all exons of MIE regions. From the findings of both DNAsis and BioEdit analyses, the [IE05] and [IE10] were discriminatingly probed into exon 4 and exon 5 of MIE regions of RCMV strains, English and Maastricht. Overall, [IE05] had homology of 43.2% (over 876bp) to MIE of RCMV strain English and 44.1% (over 488bp) to MIE of RCMV strain Maastricht; meanwhile, [IE10] had

Legends for Figures

```
Alignment:    C:\BioEdit\[IE05]

                ....|....| ....|....| ....|....| ....|....| ....|....|
                      10         20         30         40         50
[IE05]          CTTTTTTCAT CTCTTAGTGG TCTTACATTT ATCAGTTGAA GTGCCACAGA
                ....|....| ....|....| ....|....| ....|....| ....|....|
                      60         70         80         90        100
[IE05]          TTCCCTTTTA AACTTATTTT GCTTTAAAAA AAAGAAAAAA GAAAAAGAAA
                ....|....| ....|....| ....|....| ....|....| ....|....|
                     110        120        130        140        150
[IE05]          AAAGAGAGAA AAGAAAATTT AACTTAAAAT ATTTTAGCAT TAGTTGCACA
                ....|....| ....|....| ....|....| ....|....| ....|....|
                     160        170        180        190        200
[IE05]          AAATGTTTGG ATAAAATTCT AGTTTTTATA AGTCTTTTTA TATATTTAGC
                ....|....| ....|....| ....|....| ....|....| ....|....|
                     210        220        230        240        250
[IE05]          CTGTCACAAC AGTGCTCCAG ATCATATTGA TGTTTTTCTG CATTATACAA
                ....|....| ....|....| ....|....| ....|....| ....|....|
                     260        270        280        290        300
[IE05]          CCCTAAAACA CAGAGATCTC ACTAACCTGC CACCTATAAC CACTTTCTTT
                ....|....| ....|....| ....|....| ....|....| ....|....|
                     310        320        330        340        350
[IE05]          CCTCCTTTTG CCTAATACAG CTCAGCTAAG TCCTTCATAA AGATGATAAT
                ....|....| ....|....| ....|....| ....|....| ....|....|
                     360        370        380        390        400
[IE05]          CTCCTTCGTC TACATGTCTT CATAAACCAA GGACTATTAG TGAATTAAAA
                ....|....| ....|....| ....|....| ....|....| ....|....|
                     410        420        430        440        450
[IE05]          CAAAGGTGTG GAGTTTAGTA CTCCAGGTGA ACTCTTTAAA AACTAATCCT
                ....|....| ....|....| ....|....| ....|....| ....|....|
                     460        470        480        490        500
[IE05]          GACATCCTAT CACTATGTGA TATTTTGTAC ATCTTACTGT ATTTACAAGT
                ....|....| ....|....| ....|....| ....|....| ....|....|
                     510        520        530        540        550
[IE05]          TTATTTATCA AGGATTATCC ATCTTGCTAA TGGCCTATTA TTTATTTTAC
                ....|....| ....|....| ....|....| ....|....| ....|....|
                     560        570        580        590        600
[IE05]          TTTTTGTTGG TCTTTATATC CTTTTATTAA TGTTCTAAAG GAACCTGTAT
                ....|....| ....|....| ....|....| ....|....| ....|....|
                     610        620        630        640        650
[IE05]          ATCTATTTTG GAATCTATTT TGTTCGTGTT GAACAGTCTA AAATAGACTA
                ....|....| ....|....| ....|....| ....|....| ....|....|
                     660        670        680        690        700
[IE05]          AAAATGGAAT ATTTTATAGT TAAGTTACAA ACCAAGGTGT TTCTCCCCAA
                ....|....| ....|....| ....|....| ....|....| ....|....|
                     710        720        730        740        750
[IE05]          GATACTATCT TGATTTACCA TGATTCCAAA ACAAAATGAT CTAGTTAAAA
                ....|....| ....|....| ....|....| ....|....| ....|....|
                     760        770        780        790        800
[IE05]          AAAATTGGAG GATATTTTAT TACCATTACA AATAAGATAT TTAAAAAACT
                ....|....| ....|....| ....|....| ....|....| ....|....|
                     810        820        830        840        850
[IE05]          TGAATCAGGA TACAGAAGTC TGCTGGATTT GGACTAGTCT TCCCTGTTTT
                ....|....| ....|....| ....|....| ....|....| ....|....|
                     860        870        880        890        900
[IE05]          ATTTTTTTAG ATGTGCTTAC TTTTATGGTG TAACTGAAGA TTTTCTTTGT
                ....|....| ....|....| ....|....| ....|....| ...
                     910        920        930        940
[IE05]          GTAATGCATG ATTAAGACAA TAAAGTATTT TTTCTAGTCT TCC
```

Figure 1: Nucleotide sequence of [IE05] cDNA in 5' to 3'direction

```
Alignment: C:\BioEdit\[IE10]

                ....|....| ....|....| ....|....| ....|....| ....|....|
                       10         20         30         40         50
[IE10]          GAGACCGAAG GTCACTCCTG TCATAGTCCT TTTGAAAATA GTAGTTGCAG
                ....|....| ....|....| ....|....| ....|....| ....|....|
                       60         70         80         90        100
[IE10]          TTTTTAAAAA CCGTCTTGGG CCTCCGCCCC ACAACCTAAT GATCCCTGTG
                ....|....| ....|....| ....|....| ....|....| ....|....|
                      110        120        130        140        150
[IE10]          TGTGGCCTTG TTTAAAGCAG TGAATCTCTG CCACAGGAGT GCTCCCTTCT
                ....|....| ....|....| ....|....| ....|....| ....|....|
                      160        170        180        190        200
[IE10]          CTCTCTCGCA TATTTTCATG GATCAGGTTT TTAAATCGTT TTGAATTTTA
                ....|....| ....|....| ....|....| ....|....| ....|....|
                      210        220        230        240        250
[IE10]          AAAATAACTT TGAGTTCTAT CAATAAACTT TATGTCCATT TGATTGGCTC
                ....|....| ....|....| ....|....| ....|....| ....|....|
                      260        270        280        290        300
[IE10]          AAGTATTGAA GAACCTTTAA ATTCCTTAAG AAAGCCCTAT CCTCTGCTTC
                ....|....| ....|....| ....|....| ....|....| ....|....|
                      310        320        330        340        350
[IE10]          CTGCATACCG GTTACTACTA GACACCGCGA TACCTTTGAT TATTAATCCT
                ....|....| ....|....| ....|....| ....|....| ....|....|
                      360        370        380        390        400
[IE10]          ACTTTAGCAG CACAGGGAAT TCCTAACACG TGGGCTTGTG ATTGGTCTTG
                ....|....| ....|....| ....|....| ....|....| ....|....|
                      410        420        430        440        450
[IE10]          CACCGACACC CATGTTGAGT TCCGTGCAGA GGTGAAGCTT TTGTTTGATT
                ....|....| ....|....| ....|....| ....|....| ....|....|
                      460        470        480        490        500
[IE10]          AGTTTTGTGG GTTGTTTCTC ATATAAAACA TTGATACCAA CAGAGGAAGT
                ....|....| ....|....| ....|....| ....|....| ....|....|
                      510        520        530        540        550
[IE10]          GTTCCACACC ACTGGACATG GAACGCACAC CCTGTTCTTC AGCCCTGTGC
                ....|....| ....|....| ....|....| ....|....| ....|....|
                      560        570        580        590        600
[IE10]          CCAGCTGGGT GTTTAGAGAT GTGTAGAGCG AGTCCACCCA GACGCAAGTT
                ....|....| ....|....| ....|....| ....|....| ....|....|
                      610        620        630        640        650
[IE10]          GCATTTACAT TTTGAATTTT AAATGATGGT GTTATCTGTT TGGAGAAGTG
                ....|....| ....|....| ....|....| ....|....| ....|....|
                      660        670        680        690        700
[IE10]          GCTCACCCTA CACTGAGGAC CAAGAAACCC TGTGTCCTGG GGAAAACACC
                ....|....| ....|....| ....|....| ....|....| ....|....|
                      710        720        730        740        750
[IE10]          TCGGAGGTAG GAAACAGTGC CCTTGGTTTT TTAAATTTCC CGTCCAGACA
                ....|....| ....|....| ....|....| ....|....| ....|....|
                      760        770        780        790        800
[IE10]          CCAACTGGCC TTCGCTTTTG TGCCTTGTGA ATTAGTAATC CAGTAGGCTA
                ....|....| ....|....| ....|....| ....|....| ....|....|
                      810        820        830        840        850
[IE10]          AAACTCAGTC CCAGGGGAGG GGGAAGAAAA GCAAGTCCGT AGGAAGGGGT
                ....|....| ....|....| ....|....| ....|....| ....|....|
                      860        870        880        890        900
[IE10]          GCTGGTGTGA GGGTTTCCAC ACATCCTCCC CTTGGGCTGT GTACAATAAA
                ....|....| ....
                      910
[IE10]          GTGTGTGCCC TGTC
```

Figure 2: Nucleotide sequence of [IE10] cDNA in 5' to 3' direction

Table 1: DNA sequence comparison between [IE05], [IE10] and MIE region of English and Maastricht RCMVs by using DNA Homology Search of DNAsis software. The homology between two responsive sequences was presented in percentage over the size of range as stated in []. The exact size of the two exons, [IE05] and [IE10] was stated in ().

Strain of RCMV	Maastricht					ALL-03	
	EXON 1 (152 bp)	EXON 2 (104 bp)	EXON 3 (194 bp)	EXON 4 (1413 bp)	EXON 5 (1569 bp)	[IE05] (943 bp)	[IE10] (914 bp)
ENGLISH							
EXON 1 (396 bp)	49.3% [152 bp]	NA	NA	NA	NA	45.0% [362 bp]	43.5% [310 bp]
EXON 2 (110 bp)	NA	44.9% [89 bp]	NA	NA	NA	51.9% [108 bp]	50.5% [91 bp]
EXON 3 (191 bp)	NA	NA	50.6% [170 bp]	NA	NA	44.3% [192 bp]	44.2% [172 bp]
EXON 4 (1401 bp)	NA	NA	NA	45.4% [1395 bp]	NA	44.5% [922 bp]	42.8% [869 bp]
EXON 5 (1375 bp)	NA	NA	NA	NA	50.1% [1281 bp]	42.6% [935 bp]	42.9% [927bp]
ALL-03							
[IE05] (943 bp)	45.8% [144 bp]	48.3% [87 bp]	44.2% [190 bp]	45.4% [934 bp]	43.1% [908 bp]	NA	45.0% [820 bp]
[IE10] (914 bp)	46.6% [146 bp]	44.8% [96 bp]	48.6% [173 bp]	44.5% [880 bp]	41.9% [919 bp]	45.0% [820 bp]	NA

Note: NA = Not applicable

Table 2: DNA sequence comparison between [IE05], [IE10] and MIE region of English and Maastricht RCMVs by using Pairwise Alignment (Optimal Global Alignment) of BioEdit software. The values indicate the identities between two responsive nucleotide sequences.

Strain of RCMV	Maastricht					ALL-03	
	EXON 1	EXON 2	EXON 3	EXON 4	EXON 5	[IE05]	[IE10]
ENGLISH							
EXON 1	0.25	NA	NA	NA	NA	0.10	0.26
EXON 2	NA	0.48	NA	NA	NA	0.07	0.09
EXON 3	NA	NA	0.49	NA	NA	0.14	0.15
EXON 4	NA	NA	NA	0.46	NA	0.37	0.36
EXON 5	NA	NA	NA	NA	0.48	0.33	0.37
ALL-03							
[IE05]	0.23	0.08	0.14	0.36	0.39	NA	0.44
[IE10]	0.11	0.08	0.15	0.38	0.34	0.44	NA

Note: NA = Not applicable

Table 3: Comparison of G+C ratio of the nucleotides between [IE05], [IE10] and MIE region of English and Maastricht RCMVs

Strain of RCMV	Ratio between G+C Bases (%)
English	
EXON 1	62.4
EXON 2	44.5
EXON 3	49.7
EXON 4	47.5
EXON 5	42.5
MIE EXONS	47.2
Maastricht	
EXON 1	53.3
EXON 2	52.9
EXON 3	50.5
EXON 4	44.2
EXON 5	51.8
MIE EXONS	48.7
ALL-03	
[IE05]	29.6
[IE10]	45.2

41.9% (over 887 bp) and 41.9% (over 919 bp) homology to that of RCMV strains, English and Maastricht, respectively. RCMV strains, English and Maastricht had aligned homology of 45.4% over 2894bp in their whole MIE exons and identity of 0.48 analysed by DNAsis and BioEdit programs, respectively. The G+C contents of the base composition of the three RCMV strains, English, Maastricht and ALL-03 are illustrated in Table 3.

PCR2 identification of the [IE05] and [IE10] transcripts

Two sets of primers, BIE and DPC synthesised from [IE05] and [IE10] sequences, were applied to PCR2 amplification on plasmids, RCMV and REF genomic DNA. The PCR products of 569bp and 760bp in molecular size were successfully amplified in pCR®2.1-[IE05] and pCR®2.1-[IE10] using BIE primers and DPC primers, respectively. Figure 3 illustrates a distinct fragment of 569 bp and 76 bp individually amplified on RCMV DNA using primer set BIE and primer set DPC (Lane 4 and 5), respectively as conforming to their plasmids. There were no fragments amplified on REF DNA (Lane 6 and 7) by using these two sets of primers. The primers designed within the suspected IE transcripts, [IE05] and [IE10] (Figure 3, Lane 2 and 3, respectively) were specific only to RCMV but not REF.

Dot Blot hybridisation for identification of the [IE05] and [IE10] transcripts

The dot blot hybridisation produced reliable results as the test strip performed well in the detection system. No colour developed on blot of 0 pg biotinylated probe; however, colour developed with ascending intensity of biotinylated probe from 10 pg to 100 pg in concentration. Biotinylated probe BIE-[IE05] developed colour only in pCR®2.1-[IE05] as well as RCMV DNA blots but not in the negative control pCR®2.1 and REF DNA blots (Figure 4a). Meanwhile, only pCR®2.1-[IE10] and RCMV DNA blots were stained purplish blue by biotinylated probe DPC-[IE10]. However, negative control pCR®2.1 and REF DNA blots were not stained by the probe (Figure 4b). Figures 5a and 5b show the positive stainings with different intensities resulting from the dilution of RCMV DNA concentration blotted on membranes (Dot 1-4). Similarly, there were no stains on REF DNA dilution blots (Dot 5 and 6). On the other hand, a finding was obtained when the BIE-[IE05] was probed to purified PCR product DPC-[IE10] or vice versa; resulting in a higher background than that of negative plasmid control.

DISCUSSION

Generation of IE Transcripts

The generation of IE transcripts of RCMV strain ALL-03, originally isolated from the placenta and uterus of the house rat (*Rattus rattus diardii*)(Loh *et al.*, 2003) is described. Since this is a novel strain, the sequence data are not available. Primers for generation of MIE genes in both Maastricht (Beisser *et al.*, 1998) and English (Sandford *et al.*, 1993) strains of RCMV could not amplify the genes in ALL-03 strain (data not shown). Therefore, the IE mRNAs were isolated under the 'IE' conditions. Following RCMV infection of a susceptible cell, the

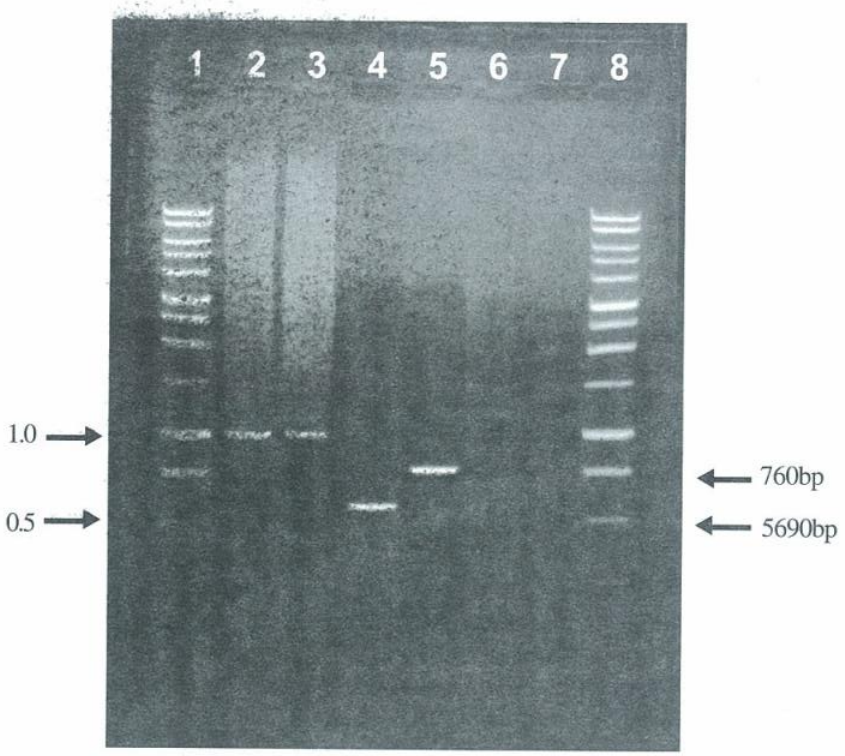

Figure 3: RT-PCR and PCR profiles detected on a 1.2% TAE agarose gel stained with EtBr. Lane 1 and 8: GeneRuler™ 1 kb DNA ladder (Fermentas). (i) RT-PCR amplification using PCR1 primers in [IE05] and [IE10] mRNAs, Lane 2: PCR1 primer in [IE05] mRNA. Lane 3: PCR1 primer in [IE10] mRNA. (ii) PCR amplification using primer sets, BIE or DPC in genomic DNA, Lane 4: BIE primers in RCMV DNA. Lane 5: DPC primers in RCMV DNA. Lane 6: BIE primers in REF DNA. Lane 7: DPC primers in REF DNA.

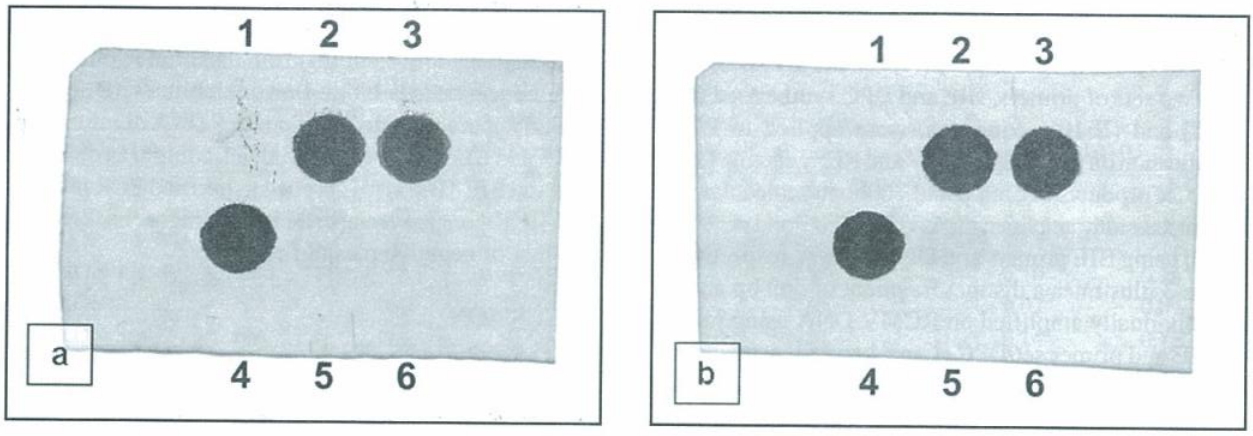

Figure 4: Dot blot hybridisation profiles employed biotinylated probes, prepared from gel-purified PCR amplicons of pCR®2.1-[IE05] and pCR®2.1-[IE10] using primer sets, BIE and DPC on positive and negative plasmid control as well as genomic DNA blots. Panel a: BIE/[IE05] biotinylated probes. Dot 1: negative control of pCR®2.1 (10 μg). Dot 2 and 3: RCMV genomic DNA (10μg). Dot 4: positive control of pCR®2.1-[IE05] (10ng). Dot 5 and 6: REF genomic DNA (10 μg). Panel b: DPC/[IE10] biotinylated probes. Dot 1: negative control of pCR®2.1 (10ng). Dot 2 and 3: RCMV genomic DNA (10μg). Dot 4: positive control of pCR®2.1-[IE10] (10 μg). Dot 5 and 6: REF genomic DNA (10μg).

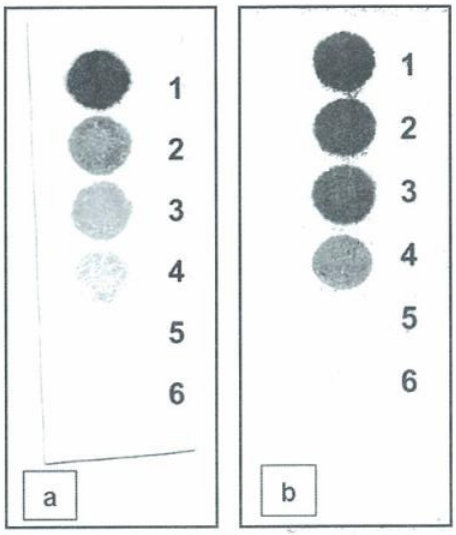

Figure 5: Dot blot hybridisation profiles employed biotinylated probes on genomic DNA blots of different concentrations. Panel a: BIE/[IE05] biotinylated probes. Panel b: DPC/[IE10] biotinylated probes. Dot 1: RCMV genomic DNA (10 μg). Dot 2: RCMV genomic DNA (7.5μg). Dot 3: RCMV genomic DNA (5μg). Dot 4: RCMV genomic DNA (2.5μg). Dot 5: REF genomic DNA (10μg). Dot 6: REF genomic DNA (5μg).

temporal expression of the genome of the input virus was closely controlled. Expression proceeded by a cascade synthesis of mRNAs and proteins termed IE, E and L (Wathern and Stinski, 1982). The cellular presence of cycloheximide at the time of inoculation of RCMV prevented translation of IE mRNAs. The transcripts obtained during this 'IE' phase were so-called suspected IE mRNAs until identification.

Unlike other herpesviruses, CMV does not shut off host macromolecular synthesis, but stimulates cellular transcription and translation (Fortunato *et al.*, 2000). Since cycloheximide was applied to inhibit translation and the IE-gene-specific primers were not employed in the procedure, it was predicted that both transcripts originated from REF and RCMV would be isolated. There were 16 bands generated by RT-PCR in total from the attempts on isolation of IE mRNAs. However, only two bands were of ALL-03 virus origin. The efficiency of the viral transcription was low (12.5%). This finding was thought to be analogous to the previous study of virus growth curve in Loh *et al.* (2003). However, the transcription levels of MIE do not ultimately determine virus production (Sandford *et al.*, 2001). This was seen in the extreme when infection of non-permissive cells allows for MIE expression without subsequent production of infectious virus (Lafemina and Hayward, 1988). The other possibilities may be due to the low initial MOI (4 $TCID_{50}$/cell) applied in the inoculum or the REF cells were still in active growth stage during inoculation period especially in the transcription process by which the REF-produced transcripts had boosted up the ratio.

Sequence assembly and bioinformatics

The bioinformatics, BioEdit, DNAsis and BLAST analyses, revealed that ALL-03 is a novel virus that has not been published elsewhere. According to these analyses, the [IE05] and [IE10] were confirmed to be of distinct sequence from each other because their identities or homology values were low. As could be expected from the very different restriction patterns conferred previously in Loh *et al.* (2003), the nucleotide sequences of the [IE05] and [IE10] of ALL-03 virus showed a low level of homology to the MIE regions of the Maastricht and English strains of RCMV. In fact, interstrain discrepancy has been reported within this region for different clinical strains of HCMV (Chou, 1992) and between the English strain and the Maastricht strain of RCMV (Beisser *et al.*, 1998). This divergence may reflect the unique pathogenicity of RCMV strain ALL-03 infection *in vivo*.

The [IE05] and [IE10] are most probably exon 4 or exon 5 of the MIE transcripts, based on the size and the random probing towards MIE sequences though the homology values were less than 45%. However, the size of the two sequences was slightly short compared to exon 4 and exon 5 of other RCMVs. Their exon 4 and exon 5 have nucleotide sizes of more than 1.3 kbp compared to the sequences in the study, which is less than 1kbp. It was previously reported that the MIE regions of HCMV and RCMV strain Maastricht produced alternative transcripts that have a part of exon 5 deleted (Stenberg *et al.*, 1989; Beisser *et al.*, 1998). RCMV (Maastricht) were

concluded to express at least two exon 5- or IE2-specific transcripts, which are composed of three exons (1, 2 and 3), alternatively spliced to either 'full-length' exon 5 (resulting in a transcript termed IE2) or a deleted version of exon 5 (resulting in transcript IE2A). The larger fragment (approximately 950bp) represents 'full-length' exon 5, whereas the smaller fragment (466bp) represents the alternative splicing product, as confirmed by DNA sequencing (Beisser *et al.*, 1998). A similar alternative splicing event had been deleted from exon 5 of HCMV resulting in an-in-frame deletion (Stenberg *et al.*, 1989), while alternative splicing of IE2A of RCMV (Maastricht) resulted in an out-of-frame deletion (Beisser *et al.*, 1998). However, the existence of minor IE mRNA transcripts was not found in RCMV strain English (Sandford *et al.*, 1993).

Besides, G+C ratio is a characteristic of a particular species, and tends to be similar between different species within a genus, while varying more widely between less related organisms. Therefore, an analysis of the G+C ratio was applied to the present study. Base composition can therefore be used as a guide to establish the taxonomic relationship between species (Dale and Schantz, 2002). [IE10] but not [IE05], was found to conform to the MIE regions of RCMV strains English and Maastricht for the G+C content. It may be a notable exception for [IE05]. There is probably an 'island' located in this sequence with a different base composition from that of the remainder of the genome. The dinucleotide CG (usually written as CpG) occurs much less commonly than would be expected from a random distribution of bases, forming a lower G+C content than that of the whole genome (Dale and Schantz, 2002). Nevertheless, the IE coding regions of HCMV, MCMV and simian CMV have been shown to have low frequency of CpG and increased frequencies of TpG and CpA (Honess *et al.*, 1989). This has been attributed to the activity of host methylating systems on cytosine in the CpG dinucleotide, with a subsequent deamination of methyl-cytosine to thymidine. This results in a decreased frequency of CpG and increased frequency of TpG and its replicated CpA (Honess *et al.*, 1989). A crude analysis of dinucleotide frequencies using [IE05] and [IE10] full nucleotides was carried out. The CpG deficiency was seen throughout the two regions, which associated with relatively increased frequencies of TpG and CpA (data not shown). This finding conformed to the studies of Honess *et al.* (1989). The probability of [IE05] and [IE10] that are parts of MIE region is considerably enhanced.

Identification of IE transcripts

The sequences were initially cloned into non-directional vector (pCR®2.1-TOPO), however, this had not brought about any difficulty to identify the 5' to 3' direction of mRNA since every single mRNA carries a poly-A tail in different sizes of nucleotides. With the help of bioinformatics tools, the primers and poly-As were discarded. The original sequence nucleotide, which represents the most likely form as in the viral genome, would be attained. However, at this stage, it is still not known in which direction of genome that the mRNAs transcribed. It is because the transcription (from a genomic sequence) probably occurs in the opposite direction. Therefore, the specific primers must be designed based on the obtained sequence data to confirm the origin and to acquire more information about the transcription direction. In fact, it was assumed at the beginning that the two transcripts, [IE05] and [IE10] contain more than one exon. However, the [IE05] and [IE10] mRNAs were confirmed to consist of an exon on its own. This was because the PCR products generated in ALL-03 genomic DNA were similar in molecular size with that generated in recombinant plasmid by using few primer sets flanking different locations within the sequences, for example, primer sets of BIE, DPC and some other primer sets that were not shown in the study.

In contrast, the primer set designed within the transcript of *R. norvegicus* small inducible cytokine A2 which originally flanked 475 bp in its recombinant plasmid successfully generated a 1.4 kbp amplicon in REF genomic DNA. This implies that the mRNA may contain two exons that are separated by an intron as revealed in its genome. Virtually, very few RNA molecules are transcribed directly into the final mature RNA product. Most newly transcribed RNA molecules undergo various alterations to yield the mature product, that is, RNA processing (Turner *et al.*, 2000). Commercial mRNA extraction kits isolate mRNA in both nucleus and cytoplasm whereby the pre-mRNA in nucleus is captured. This could be explained by the phenomenon that some of the *R. norvegicus*-derivative mRNAs have the same identities but with different size of nucleotides. Interestingly, the two similar mRNAs isolated at 5h p.i. were longer in nucleotide size than that at 10h p.i. This may imply that the mRNAs are instable when incubated over long duration or these mRNAs are captured during an immature period as mentioned previously. If the former happened, another argument can be brought into the explanation of the relative small size of [IE05] and [IE10] sequences.

The probe amplification-hybridisation system employed in the study was amplified during PCR and labeling, following hybridisation of target for further confirmation of the identity. The probe sequences can be amplified 10^6-fold or more, making final detection relatively simple (Lizardi *et al.*, 1988; Saiki *et al.*, 1988). The hybridisation analysis detects a biotin-labeled probe, which has been hybridised to picogram amounts of the target DNA. Its sensitivity level is comparable to PCR. The hybridisation analysis has a higher specificity because the probe being longer than that of PCR primers, made it more useful for identification of amplified PCR

product (Stoler and Michael, 1995). Since dot blot was applied in the study, the sensitivity was poorer than that of Southern blot. In addition, the total genomic DNA was probed for a single gene, that is, either [IE05] or [IE10]; a larger amount of sample DNA (range 2.5-10μg) was blotted onto the filter membrane. This range of DNA template produced satisfactory results. The biotinylated probes prepared by introducing the biotin-14-dCTP into the base composition, [IE05] was expected to have fewer signals than that of [IE10] even with the same concentration of template as demonstrated in Figure 5. It can be explained by its G+C content as well as the length of probe prepared which were lesser than that of [IE10]. Another phenomenon observed is the higher background obtained from the target/probe hybrids of [IE05] and [IE10] compared to probing with control plasmids. This reflects that [IE05] and [IE10] are more closely related with each other because they originate from the same species.

Nevertheless, the identity of [IE05] and [IE10] was confirmed to be of RCMV origin without doubt by PCR with subsequent hybridisation analysis. This indicates that the two primer sets, BIE and DPC, were single exon-specific to ALL-03, generate amplicons by PCR flanking the genome. These primers which are ALL-03-specific would be useful for the development of molecular tools in particular PCR or RT-PCR for detection of RCMV strain ALL-03 in pathogenicity studies modeling for CMV infection and disease. Hence, the exploration of these IE genes paves the way for understanding the molecular mechanism involved in transplacental infection.

CONCLUSION

Two different IE transcripts, [IE05] and [IE10] were successfully isolated, characterised and identified in the current study by inhibiting protein synthesis *in vitro*, gene cloning, bioinformatics, PCR and hybridisation analyses. The results suggest that [IE05] and [IE10] are IE1 and IE2, respectively. A considerable divergence in MIE sequence between three RCMV strains, ALL-03, Maastricht and English was observed. However, additional sequence data from the genome of strain ALL-03 (in particular the complete MIE gene locus) is essential for a comprehensive comparison to the other strains of RCMV. Further experiments, particularly Southern blot hybridisation, cloning and sequencing of the other three exons, are recommended to generate the complete MIE genes and to localise the MIE genes within the RCMV genome.

REFERENCES

Beisser, P.S., Kaptein, S.J.F., Beuken, E., Bruggeman, C.A. and Vink, C. (1998). The Maastricht strain and English strain of rat cytomegalovirus represent different Betaherpesvirus species rather than strains. *Virol.* **246**: 341-351.

Boldogh, I., Abubakar, S., and Albrecht, T. (1990). Activation of proto-oncogenes: an immediate early event in human cytomegalovirus infection. *Science* **247**: 561-564.

Boldogh, I., Abubakar, S., Deng, C.Z., and Albrecht, T. (1991). Transcriptional activation of cellular oncogenes fos, jun, and myc by human cytomegalovirus. *J. Virol.* **65**: 1568-1571.

Chou, S. (1992). Effect of interstrain variation on diagnostic DNA amplification of the cytomegalovirus major immediate-early gene region. *J. Clin. Microb.* **30**: 2307-2310.

Dale, J.W. and Schantz, M. von (2002). From Genes to Genomes: Concepts and Applications of DNA Technology. John Wiley and Sons, London.

Fortunato, E.A., McElroy, A.K., Sanchez, V. and Spector, D.H. (2000). Exploitation of cellular signalling and regulatory pathways by human cytomegalovirus. *Trends Microbiol.* **8**: 111-119.

Honess, R.W., Gompels, U.A., Barrell, B.G., Craxton, M., Cameron, K.R., Staden, R., Chang, Y.N. and Hayward, G. S. (1989). Deviations From expected frequencies of CpG dinucelotides in herpesvirus DNAs may be diagnostic of differences in the states of their latent genomes. *J. Gen. Virol.* **70**: 837-855.

LaFemina, R.L. and Hayward, G. S. (1988). Differences in cell-type-specific blocks to immediate early gene expression and DNA replication of human, simian and murine cytomegalovirus. *J. Gen. Virol.* **69**: 355-374.

Lai, K.Y., Mohd-Azmi, M.L., Khoo K.K. and Sheikh Omar, A.R. (1999). Random amplified polymorphic DNA (RAPD) for identification of new rat cytomegaloviruses isolated from rice-field rats. *J Vet Malaysia* **11**: 23-26.

Lizardi, P.M., Guera, C.E., Lomeli, H., Tussie-Luna, I. and Kramer, F.R. (1988). Exponential amplification of recombinant-RNA hybridization probes. *Biotech.* **6**: 1197-1202.

Loh, H.S., Mohd-Azmi, M.L., Lai, K.Y., Sheikh-Omar, A.R. and Zamri-Saad, M. (2003). Characterization of a novel rat cytomegalovirus (RCMV) infecting placenta-uterus of *Rattus rattus diardii*. *Arch. Virol.* **148**: 2353-2367.

Loh, H.S., Mohd-Azmi, M.L., Sheikh-Omar, A.R. and Kiew, L.J. (2006). Pathogenicity and vertical transmission of a transplacental rat cytomegalovirus. *J. Virol.* **3**: 42.

Messerle, M., Buhler, B., Keil, G.M. and Koszinowski, U.H. (1992). Structural organisation, expression, and functional characterisation of the murine cytomegalovirus immediate-early gene 3. *J. Virol.* **66**: 27-36.

Mohd-Azmi, M.L., Lai, K.Y., Sheikh-Omar, A.R., Smith, A.L. and Lam, Y.M. (1999). Infectious viruses for the biological control of rice-field rats: Rat cytomegalovirus (RCMV) as a suitable candidate for the development of immunocontraceptive vaccine. In: Symposium on Biological Control in the Topics, 18-19 March, Mardi Training Centre, Serdang, Selangor, Malaysia. pp. 89-91.

Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis, K.B. and Erlich, H.A. (1988). Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Sci.* **239**: 487-491.

Sandford, G.R. and Burns, W.H. (1996). Rat cytomegalovirus has a unique immediate-early gene enhancer. *Virol.* **222**: 310-317.

Sandford, G.R., Ho, K. and Burns, W.H. (1993). Characterisation of the major locus of immediate-early genes of rat cytomegalovirus. *J. Virol.* **67**: 4093-4103.

Stenberg, R.M., Depto, A.S., Fortney, J. and Nelson, J.A. (1989). Regulated expression of early and late RNA and protein from the human cytomegalovirus immediate-early gene region. *J. Virol.* **63**: 2699-2708.

Stoler, D.L. and Michael, N.L. (1995). Nucleic acid blotting techniques for virus detection. In: Molecular Methods for Virus Detection. Wiedbrauk, D.L. and Farkas(Eds.) D.H. Academic Press, San Diego. pp. 39-74.

Turner, P.C., McLennan, Bates, A.D. and White, M.R.H. (2000). Molecular Biology, 2nd ed. UK, BIOS Scientific Publishers Ltd.

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

Wathern, M.W. and Stinski, M.F. (1982). Temporal patterns of human cytomegalovirus transcription: mapping the viral RNAs synthesized at immediate early, early, and late times after infection. *J. Virol.* **41**: 462-477.