Characterising variation in five genetic loci of cytomegalovirus during treatment for congenital infection.

Kadambari S*, Atkinson C2, Luck S1,2, Macartney M2, Conibear T2, Harrison I2,3, Booth C2, Sharland M1, Griffiths PD2

1 Paediatric Infectious Diseases Research Group, St George's University of London, Cranmer Terrace London, SW17 0RE
2 Centre for Virology, University College London Medical School, Rowland Hill Street, London, NW3 2PF
3 European Reference Laboratory Network for Human Influenza, Public Health England, 61 Colindale Avenue London NW9 5HT

*Corresponding author: Dr Seilesh Kadambari, Paediatric Infectious Diseases Research Group, Room 2.215e, Level 2 Jenner Wing, St George’s University of London, Cranmer Terrace, London SW17 0RE. skadamba@sgul.ac.uk. Phone number: 0208 725 5382

Keywords: Congenital; Cytomegalovirus; Resistance; Antiviral; Vaccine

Running title: Changes in 5 genes during congenital CMV treatment
Abstract

Cytomegalovirus (CMV) is the most common congenital infection in humans and a leading cause of sensorineural hearing loss. Ganciclovir (6mg/kg twice daily for 42 days) has been shown to reduce hearing deterioration and is used in clinical practice. Vaccines and passive administration of antibody are being evaluated in randomised controlled trials in allograft candidates, women of childbearing age and pregnant women with primary CMV infection. To help define genetic variation in each of the targets of these therapeutic interventions, we amplified and sequenced genes UL97 (site utilised for ganciclovir phosphorylation), UL55 (glycoprotein B (gB) vaccine target) and UL128, UL130 and UL131a (specific monoclonal antibody targets). Serial blood, saliva and urine samples (total 120) obtained from 9 infants with symptomatic congenital CMV treated with 42 days’ ganciclovir were analysed. All samples tested were UL97 wild type at baseline and none developed mutations during treatment, showing no selection of resistance. The prevalences of UL55 genotypes were 28% gB1, 22% gB2, 1% gB3 and mixed in 20% samples. No mutations were noted in UL128-131a. Phylogenetic tree analysis showed that sequences with variations were found in multiple body sites of individual patients, so there was no evidence of body site compartmentalisation of particular strains of CMV. The significance of these results for changes in diagnostic practices and therapeutic interventions against CMV are discussed.
Introduction

Congenital CMV infection is estimated to affect 0.7 per 1000 live births worldwide (Dollard et al. 2007). Around 25% of babies born with congenital CMV develop some form of permanent neurological impairment, primarily sensorineuronal hearing loss (SNHL) and intellectual and developmental disabilities, with approximately one third having symptoms at birth (Dollard et al., 2007). Congenital CMV can be due to maternal primary infection in pregnancy, reinfection with a new strain or reactivation from latency (Gaytant et al. 2002; De Vries et al. 2013).

A phase III randomised controlled trial (RCT) demonstrated that six weeks (6mg/kg twice daily) of intravenous ganciclovir therapy started in the first month of life reduced hearing deterioration and improved neurological outcomes in newborns with congenital CMV and evidence of central nervous system (CNS) disease (Kimberlin et al 2003; Oliver et al. 2009). The protein kinase encoded by CMV gene UL97 phosphorylates ganciclovir to its active form and UL97 mutations impair ganciclovir phosphorylation with 90% of resistance mutations to ganciclovir occurring in the UL97 gene between amino acids 460-520 (Chou 2008). Resistance is suspected clinically after solid organ transplant when viral load increases despite treatment and can be confirmed by sequencing the UL97 gene (Chou 2008). Theoretical and clinical studies in solid organ transplant recipients show that therapy lasting for longer than 100 days is needed to select for resistance to ganciclovir (Lurain et al. 2002; Limaye et al. 2000; Limaye 2002; Emery & Griffiths 2000). No published studies have assessed whether short courses of treatment in neonates selects for ganciclovir-resistant strains of CMV.

In clinical studies of neonates, viral load is usually suppressed during 6 weeks of treatment but rebounds in blood, urine and saliva samples (Whitley et al. 1997; Luck et al 2009). This pattern is consistent with viral replication recovering once the antiviral drug pressure is stopped, rather than selection of resistance, but it is important to document this, especially now that longer courses of treatment with ganciclovir are being evaluated in a RCT (NCT00466817).
This cohort of symptomatic babies was used to assess genetic variation in additional genes that are currently targets for therapeutic intervention against CMV. Two phase II vaccine trials have been conducted in the last 6 years with promising results of recombinant soluble glycoprotein B (gB) vaccine in women of childbearing age and in solid organ transplant candidates (Griffiths et al. 2011; Pass et al. 2009). Five distinct genotypes (gB 1 – 5) have been identified in infants with congenital CMV and co-infection with multiple strains reported (Ross et al. 2011).

This study also examined variation in genes that are the target of a current RCT attempting to interrupt transmission of CMV from donor to recipient during renal transplantation (NCT01753167) and a third target for therapeutic intervention. Proteins encoded by the UL128-131a gene loci form a pentameric complex with the antigens glycoprotein L and glycoprotein H (Macagno et al. 2010). The UL128-131a region is a major determinant of virus entry into epithelial cells (Macagno et al. 2010). Monoclonal antibodies against these proteins that have high neutralising activity against CMV are infused at the time of transplant and the proportion of infected recipients compared to that found in recipients of placebo. As with the gB vaccine mentioned above, parallel studies of this monoclonal antibody preparation in women of childbearing age could be envisaged with the objective of reducing maternal-fetal transfer of CMV. Hyperimmune immunoglobulin has recently been evaluated during pregnancy for this purpose but did not significantly alter the course of infection during pregnancy (Revello 2014).

The final aim of this study was to address a change in contemporary diagnosis of congenital CMV infection. Although urine is still most commonly used for diagnosing suspected clinical infection at birth, the detection of CMV using PCR of dried blood spots and saliva are gaining a role in both the retrospective diagnosis of CMV and in screening programmes (Boppana et al. 2010; Boppana et al. 2011; Walter et al. 2008) . Previous studies show that multiple genotypes are present in samples taken within the first weeks of life with distinct strains found in different body compartments as shown by genotyping of gB, gH and gN (Ross et al. 2011). None of the infants were reported to be
symptomatic or to receive treatment for congenital CMV (Ross et al. 2011).

Given the above observations, this study aimed to evaluate variation in CMV genes with possible relevance to pathogenesis or treatment. This study examined different body compartments of treated infants and determined whether genetic changes segregated by patient or by body site.

**MATERIALS AND METHODS**

**Patient samples**

Blood, saliva and urine samples were analysed from 9 infants with congenital CMV recruited from 4 different paediatric units in the ethically approved Viral Load and Immunity in Congenital CMV (VI CCC) study during 2008-2010. Eight infants had been treated with 42 days ganciclovir treatment and one infant with 42 days ganciclovir followed by 39 days of valganciclovir. Samples were obtained at days 7, 28, 42 of treatment, 7 days post treatment and 3, 6 and 9 months of life.

Samples with CMV viral loads >2.5 log_{10} were selected for analysis to increase the likelihood of there being sufficient genetic material for analysis.

**DNA Extraction of CMV DNA using the Biomerieux automated extractor**

Total nucleic acid was extracted using the commercial Nuclisense Easymag system (Biomerieux, Basingstoke UK). This is a semi-automated system based on a nucleic acid purification method developed by Boom and colleagues with enhanced magnetic silica technology (Boom et al. 1990). DNA was extracted according to the manufacturer’s instructions.

**UL97 population sequencing**

The UL97 region (codons 550 – 645) was characterised using a method published previously by Castor and colleagues with the following modifications to the cycling programme (Castor et al. 2007). The PCR programme was: denaturation at 94°C for 2 min, followed by 45 cycles of 94°C for 30 seconds, 55°C for 30 seconds, and 72°C for 1 minute. Final elongation was carried out at 78°C for 4
min, with cooling at 37°C for 1 min. PCR product was visualised on agarose gel prior to cycle sequencing.

**Genotyping of cytomegalovirus glycoprotein B**

gB genotyping (sequence strains C327A (M60929), C336A (M60931), C076A (M85228), and C194A (M60926)) was performed using a method previously described by Pang et al (Pang et al. 2008). Real time PCR amplification was performed on the ABI TaqMan 7500 (Applied Biosystems, Foster City, CA) with 45 thermal cycles of 95°C for 15 seconds and 60°C for 45 seconds.

**PCR amplification of UL128-131a**

Primer sequences to cover known sequence polymorphisms within the UL128-131a regions were modified from a method published previously (Vogel et al. 2013). The primer sets were: UL128-2-F (forward) 5′-TCg gCg ATA AAC ACC ACT ATC-3′ and UL128-2-R (reverse) 5′-CCA TCC CAA TCT CAT CgT TT-3′; UL130-2-F (forward) 5′-AgA ACg gCg TCA ggT CTT T-3′ and UL130-2-R (reverse) 5′-CAA CAA AAg gAC CAC gTT CA-3′; UL131A-2-F (forward) 5′-TgA AAg Tgg TgA CgA AgC Ag-3′ and UL131A-2-R (reverse) 5′-gCT CAg AgA TCC CgA gTA Cg-3′.

**DNA sequencing and phylogenetic analysis**

Sequencing was performed using the ABI Prism BigDye terminator cycle sequencing kit (v3.1), on an Applied Biosystems 3730 DNA analyser. Sequences were analysed using Applied Biosystems SeqScape software with Genbank accession numbers of G221975, FJ527563, GQ221974, GQ466044 and AY446894 used as CMV reference sequences. Phylogenetic trees were constructed using an online phylogenetic tree maker http://www.hiv.lanl.gov/content/sequence/PHYML/interface.html.

**RESULTS**

UL97
Sequencing data were obtained from 20 blood, 24 saliva and 24 urine samples taken at four time points during and after ganciclovir treatment (day 7, day 28, day 42 and 7 days post-treatment). Not all patients had all samples available at each time point, but sequence data were obtained from all specimens tested. Results showed that all samples were wild type at day 7. No known resistance mutations were identified within the UL97 region in samples taken during treatment. Phylogenetic analysis demonstrated viral segregation with the patient and not compartmentalisation by body site (Figure 1).

**Glycoprotein B**

32 blood, 37 urine and 33 saliva samples were amplified. Genotypes were obtained for 71/102 samples (70%). gB1 28% (29/102) and gB2 22% (22/102) were the most prevalent strains with gB3 identified in only one sample and no genotype available in 30% (31/102); mixed genotypes were not uncommon (gB1/gB2 16% and gB2/gB3 4% (Table 1). In 7/9 newborns the gB genotype identified in ≥1 compartment varied at different intervals during and after treatment. No evidence of compartmentalisation by body site was noted. Mixed genotypes were not associated with sample type or time point across infants.

**UL128-131a**

It was possible to investigate this third region using the samples collected as it is not responsible for ganciclovir phosphorylation. 36 blood, 42 urine and 42 saliva samples were amplified from newborns on day 7 of life and at 6 months. Sequence data was obtained from all specimens tested. No mutations were identified in any of the samples sequenced. Results obtained from UL128-UL131a sequencing data in blood, saliva and urine show the virus segregating with the patient and not the body compartment.

**DISCUSSION**
Although the number of cases available to us was small, they have the advantage of representing a population known to benefit from therapeutic interventions; congenital CMV infection born with symptoms. Natural history studies show that 32% of pregnant women with primary CMV infection transfer virus across the placenta and that approximately 13.5% of infected babies are born with symptoms while another 12.7% develop symptoms on follow up (Kenneson & Cannon 2007; Dollard et al. 2007). It is possible that those who suffer from this infection represent a subset of infected individuals because one or more strains of CMV has above average pathogenicity. If this were true, it would be important to document the genetic composition of these proposed more pathogenic viruses.

Results from the cohort presented demonstrate that 42 days of treatment with ganciclovir does not frequently select for resistance mutations in the UL97 gene. The region sequenced covered codons 439 – 645 of the UL97 gene. This area has been shown in a previous study to include all clinically relevant ganciclovir mutations (Chou 2008). We cannot exclude, however, that as yet undefined mutations exist outside this region. These results suggest that the rebound in viral load seen at the end of treatment is more likely due to the natural dynamic nature of CMV replication rather than antiviral resistance (Emery & Griffiths 2000). It will be interesting to see whether, in contrast, the recently completed study of a longer course of 6 months’ valganciclovir in neonates selects for resistant strains (NCT00466817).

Our studies of other loci give the impression of genetic stability within CMV, with no evidence that these symptomatic babies have been infected with unusual strains of virus and no significant sequence variation observed. In addition, in contrast to Ross’s study of babies not reported to be symptomatic, this study found no evidence of body site compartmentalisation, because genetic variants segregated with individual patients rather than by body site which could be due to ganciclovir enhancing selection of strains (Ross et al. 2011). This implies that changes in diagnostic practices towards preferring saliva and blood over urine should not introduce major biases into
studies of CMV genetics. The epitopes within the UL130 complex that are targeted by a current study of passive infusion of monoclonal antibodies appear to be conformational, so interpretation is complex, but the polymorphisms seen do not suggest major changes from the wild-type virus.

Genetic variation was seen for gB; as in previous studies gB 1 was the predominant genotype in these congenitally infected infants (Ross et al. 2011). Table 1 highlights a complex relationship between genotype, body compartment and time point. A possible explanation is a combination of transplacental transfer of more than one maternal strain (cases 3, 4, 5 and 7) and postnatal reinfection from multiple CMV strains (cases 1 and 8) in congenitally infected newborns. Re-infection with another strain may induce symptoms. However, this data was unavailable for the study. No maternal samples were available for sequencing and so it was not possible to demonstrate the presence of the same genotypes present in a mother and her infected newborn as has been shown in other studies (Yamamoto et al. 2007). Future studies could consider prospectively collecting samples from other family members to determine if they are the source of CMV reinfections during the first year of life.

As regards the implications of genetic variation in gB for the potential to control CMV infection using vaccines containing gB, the variations seen mapped to antigenic domains 2 and 4 among the 5 identified by Potzsch (Pötzsch et al. 2011). Interpretation is complex, because some of the epitopes are linear whereas others are conformational. It is hoped that these and other reports of genetic variation in gB will aid future three-dimensional modelling of the structure of gBs and identify any predicted effects on the ability of antibodies to bind gB variants.

This study amplified viral DNA direct from clinical specimens giving it an advantage over studies using cell cultures which can select for different virus strains(Dargan et al. 2010). The use of population sequencing in this study, however, means that genotypes are only identifiable once they account for approximately 20% of the sequence population (Lurain & Chou 2010). Mutations present at lower levels can be detected with pyrosequencing (approximately 6%) and ultra-deep sequencing.
(≤1%) (Renzette et al. 2011). Future studies employing these newer methods could determine if even more variation is seen and consider virus evolution over longer periods of time.

**Funding:** No external funding

**Competing interests:** None

**Ethical approval:** Not required

**Acknowledgements:** We are grateful to staff at Imperial College Healthcare NHS Trust, Oxford University Hospitals NHS Trust, The Newcastle Hospitals NHS Foundation Trust and St George's Healthcare NHS Trust who contributed samples towards this analysis, and all the parents and babies involved in the VICC study. The authors are also grateful for discussions with Richard Milne, Centre for virology at UCL.
REFERENCES


Pass, R.F. et al., 2009. *Vaccine prevention of maternal cytomegalovirus infection.*,  


Figure 1 Phylogenetic tree constructed using polymorphisms obtained from sequencing of UL97 samples

Key:

P: Patient; B: Blood; S: Saliva; U: Urine

1: Sample from day 7; 2: sample from day 28; 3: sample from day 42; 4: sample 7 days post treatment
<table>
<thead>
<tr>
<th>DAY OF RX</th>
<th>BASELINE</th>
<th>DAY 28</th>
<th>DAY 42</th>
<th>D7 POST</th>
<th>M3</th>
<th>M6</th>
<th>M12</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Saliva</td>
<td>0</td>
<td>1,2</td>
<td>NA</td>
<td>1,2</td>
<td>N</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Blood</td>
<td>1</td>
<td>N</td>
<td>N</td>
<td>0</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td>Urine</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>2</td>
<td>Saliva</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td>Blood</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td>Urine</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>3</td>
<td>Saliva</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td>Blood</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>2,3</td>
<td>2,3</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td>Urine</td>
<td>2,3</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>4</td>
<td>Saliva</td>
<td>1</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Blood</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1,2</td>
<td>1,2</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Urine</td>
<td>1,2</td>
<td>1,2</td>
<td>1</td>
<td>N</td>
<td>1</td>
<td>N</td>
</tr>
<tr>
<td>5</td>
<td>Saliva</td>
<td>1</td>
<td>1</td>
<td>N</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Blood</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Urine</td>
<td>1,2</td>
<td>1</td>
<td>N</td>
<td>1</td>
<td>1</td>
<td>1,2</td>
</tr>
<tr>
<td>6</td>
<td>Saliva</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>N</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Blood</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Urine</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>N</td>
<td>1</td>
<td>N</td>
</tr>
<tr>
<td>7</td>
<td>Saliva</td>
<td>2</td>
<td>N</td>
<td>N</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Blood</td>
<td>1,2</td>
<td>2</td>
<td>N</td>
<td>N</td>
<td>1,2</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td>Urine</td>
<td>1,2</td>
<td>1,2</td>
<td>2</td>
<td>1,2</td>
<td>1,2</td>
<td>1,2</td>
</tr>
<tr>
<td>8</td>
<td>Saliva</td>
<td>2</td>
<td>2</td>
<td>N</td>
<td>2</td>
<td>3</td>
<td>2,3</td>
</tr>
<tr>
<td></td>
<td>Blood</td>
<td>2</td>
<td>0</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td>Urine</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>9</td>
<td>Saliva</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>N</td>
<td>1</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td>Blood</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td>Urine</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>N</td>
<td>1</td>
<td>N</td>
</tr>
</tbody>
</table>

Table 1 CMV gB genotype distribution in blood, saliva and urine in infants treated with ganciclovir

N= sample not available

0 = No genotype obtained

1 = genotype 1

2 = genotype 2

3 = genotype 3

4 = genotype 4

1,2 = mixed genotype 1 and 2

2,3 = mixed genotype 2 and 3