Cytomegalovirus Glycoprotein B Genotypes in Mexican Children and Women

Hilda M. Gonzalez-Sanchez, Diana L. Alvarado-Hernandez, Sandra Guerra-Palomares, Christian A. Garcia-Sepulveda, Daniel E. Noyola

Departamento de Bioquímica, Laboratorio de Genómica Viral y Humana and Departamento de Microbiología, Facultad de Medicina, Universidad Autónoma de San Luis Potosí, San Luis Potosí, México

Key Words
Cytomegalovirus · Glycoprotein B · Genotypes · Mexico · Latin America · Congenital infection · Children

Abstract
Objective: Cytomegalovirus (CMV) is widely distributed and constitutes the main cause of congenital infections worldwide. CMV transmission during pregnancy represents one of the major impacts of this virus on public health. This study aimed at assessing glycoprotein B (gB) CMV genotypes in Mexican children and pregnant women, since there is limited information regarding CMV genomic diversity in Mexico.

Methods: We analyzed CMV strains detected in Mexican children (n = 38) and women (n = 38) between 2001 and 2012. A fragment of the gB gene was amplified and sequenced, and genotypes were defined based on prototype sequences.

Results: The gB1 genotype was detected more frequently in children (68.4%) compared to women (31.6%; p = 0.0028), while genotype 2 was more common in women (65.8%) compared to children (26.3%, p = 0.0012). Genotype 3 was uncommon in both groups (5.3 and 2.6%). Nucleotide sequences exhibited a high degree of similarity to prototype strains. However, we identified 17 distinct sequences that resulted in changes in the encoded amino acid sequence in four strains. Conclusions: gB1 and gB2 are the most common strains associated with CMV infection in Mexican children and women. In addition, we found that the frequency for each genotype differed amongst them, possibly due to variability in transmission or reactivation dynamics.

Introduction

Human cytomegalovirus (CMV; family Herpesviridae, genus Cytomegalovirus, species Human herpesvirus 5) is widely distributed throughout the world and may cause infection at any age. The lifetime risk of acquiring CMV is close to 100%. Infection during pregnancy is of particular concern due to the risk of transplacental transmission leading to congenital infection [1]. CMV is the most common cause of congenital infection and a leading cause of neurological and hearing disabilities in childhood [2]. Nevertheless, most infections caused by this virus, even in the developing fetus, are self-limiting and do not lead to serious clinical consequences. Both host and viral factors have been associated with the risk of acquiring the infection and the development of symptoms. Among viral factors, genetic features have been suggested...
to explain some of the diversity in clinical outcomes following infection. Multiple genes have been studied to determine their role in pathogenicity. Among these, the glycoprotein B (gB) gene (UL55) has attracted much attention because its gene product is expressed in the viral surface, and neutralizing antibodies with potential protective effect are generated against it. Several studies have suggested that some gB genotypes may be associated with severe infection, while other studies have failed to find such associations [3–6]. These differences may be the result of study design, the populations included in the studies and the diverse outcomes that have been analyzed.

In addition, because gB has potential usefulness for vaccine development, molecular epidemiological studies are warranted to determine the distribution of genotypes in different regions and populations. There is currently little information regarding CMV genomic diversity in Mexico. We performed a DNA-based molecular analysis to determine the CMV gB genotypes associated with infections in Mexican children and women during pregnancy and puerperium. CMV strains from children included those detected in infants with congenital infection and in children with postnatal infection.

**Materials and Methods**

**Samples**

Samples from children (urine, saliva, blood and pharyngeal exudates) and from pregnant and puerperium women (milk and vaginal secretion) obtained between 2001 and 2012 which had been identified as CMV positive in previous studies were selected. All samples were obtained following approval by the corresponding research and ethics committees.

**CMV Subtyping**

Selected samples were tested to confirm the presence of CMV using a PCR-based protocol, as described previously [7]. Subsequently, genotyping of CMV strains focused on the gB gene of the virus which encompasses the protein cleavage site according to previous protocols [8]. A nested PCR assay that amplifies a fragment of the CMV gB gene was used; the amplified region encompasses the cleavage site of this protein, which has been preferentially used to characterize the different CMV genotypes. Viral DNA was extracted from clinical samples or supernatants from first-passage cultures (0.2 μl). Outer and inner primers targeted the UL55 gene of CMV and amplified a final 719-bp fragment. For both reactions: 94° for 3 min followed by 30 cycles of 94° for 30s, 30 s annealing at 56.5° or 50° (for the first and second PCR, respectively) and a 60 s extension at 72°, followed by a final extension of 10 min at 72°. Nested PCR products were observed after electrophoresis in a 1.5% agarose gel stained with ethidium bromide and photographed using a Kodak Gel Logic 100 digital imaging system. PCR products were sequenced at Laboratorio Nacional de Genómica para la Biodiversidad (CINVESTAV-Irapuato). Nucleotide sequences were translated into the corresponding amino acid sequences and the genotypes assigned by comparing them to prototype strains using Clustal Omega and a local unanimity reformatting tool. Prototype strains were Merlin (AY446894) and Towne (FJ616285) for gB1, AD169 (FJ527563) for gB2, Toledo (GU937742) for gB3 and C194A (M69926) for gB4. Traditional phylogenetic trees were inferred using the neighbor-joining method with the ClustalW2-Phylogeny Tool. In order to detect possible recombination events [9, 10], SimPlot and Bootscan analyses of the nucleotide sequences were performed. The sequences generated in this work have been submitted to GenBank (GenBank accession Nos. KM980094–KM980169).

**Statistical Analysis**

Genotype frequencies are reported as percentages of the isolates assigned to each class. Comparisons between populations (children and women) were done using Fisher’s exact test or the χ² test. p < 0.05 was considered significant. Statistical analyses were performed using Epi Info v7 (Centers for Disease Control and Prevention, Atlanta, Ga., USA).

**Results**

Genotyping based on the gB sequence was determined in samples from 76 subjects with CMV infection (38 children and 38 adult women). Samples from children corresponded to 8 neonates with congenital CMV infection (from whom CMV was detected within the first 3 weeks of life) and 30 children with probable postnatal infection. Young infants in the latter group had urine/saliva samples obtained during the first weeks of life that tested negative for CMV, while older children were considered to have postnatal infection based on the lack of history of congenital or neonatal abnormalities. Samples from adults participating in the study corresponded to women enrolled during pregnancy or early puerperium. Serological testing indicated that most infections were likely to be CMV reactivations or reinfections during pregnancy (presence of CMV-specific IgG antibodies and negative CMV-specific IgM antibodies). All samples were from unrelated subjects, with the exception of two pairs of related samples: samples from 2 children (one sibling had congenital infection and the other had postnatal infection) and samples from a mother and her child (with postnatal infection). The general characteristics of subjects included in this analysis are shown in table 1.
**Genotype Frequency**

All samples could be assigned to one of the previously described gB genotypes on analysis of the amino acid sequence of the cleavage site region of gB. Most sequences corresponded to gB1 (n = 38) and gB2 (n = 35), while three corresponded to gB3 (fig. 1). None of the CMV strains detected in children and adult women in this study exhibited homology to genotypes 4 or 5. The genotype distribution within the different study groups is shown in figure 2. The most common genotype observed in children was genotype 1 (68.4% of all children; 50% of children with congenital infection and 73.3% of children with postnatal infection). The proportion of women with genotype 1 infection was significantly lower compared to children (31.6%; p = 0.0028 when compared to all children). In contrast, genotype 2 was more commonly detected in women (65.8%) compared to all children (26.3%; p = 0.0012). Detection of genotype 3 strains was rare in both children and adult women (5.3 and 2.6%, respectively; p = 1.0). Comparisons between children with postnatal infection (excluding those with congenital infection) and adult women showed that the statistically significant differences between groups persisted.

Genotype distribution in children with congenital infection appeared to differ somewhat from that observed in children with postnatal infection and adults. As shown in figure 2, the equal detection of infections caused by gB1 and gB2 strains appears to be closer to the distribution observed in pregnant women than that seen in children. Statistical comparisons between children with congenital infection and either women or children with postnatal infection failed to demonstrate significant differences, probably due to the small number of children with congenital infection that could be analyzed. Among infants with congenital infection, there were 2 cases with symptomatic infection, while the other 6 were asymptomatic. Both children with symptomatic infection had gB2 strains, while gB1 and gB2 strains were detected in those without symptoms. The small number of cases did not allow for statistical comparisons between the two groups.

**Nucleotidic Analysis**

Nucleotide sequences exhibited a high degree of similarity to prototype strains (98–100%). The 613- to 619-bp fragment that was included in the analysis displayed 100% similarity to prototype strains in half of the identified sequences (7 were identical to Merlin strain and 31 identical to AD169 strain). However, we identified 17 additional distinct sequences (6 of them were seen in two or more instances, while 11 were seen in a single instance); in spite of this, these resulted in changes in the encoded amino acid sequence in only four strains. As previously described, analysis at the cleavage site region allowed classification of the study strains in previously identified genotypes. Nevertheless, in three gB1 strains some variability was observed towards the 5′-terminal region of the sequence; the nucleotidic changes registered resulted in amino acid substitutions similar to those observed in genotype 3 strains. However, SimPlot analysis did not indicate the presence of recombination events. Among all study sequences, only one strain (UASLP/24037/10) showed evidence of a possible recombination; this strain had a nucleotidic sequence very similar (two nucleotide changes) to the Towne strain, which has previously been reported to possess evidence of recombination [10].

**Amino Acid Sequence Analysis**

Overall, 6 different amino acid sequences were identified (fig. 3). Most amino acid sequences inferred from the nucleotide sequences were identical to reference strains: 33 sequences were identical to the Merlin strain sequence and 34 to the AD169 strain. In addition, 4 novel and distinct
Fig. 1. Phylogenetic tree based on the analysis of CMV gB amino acid sequences of 76 strains identified in San Luis Potosí, Mexico.
sequences were identified: sequence UASLP/24076/09 was similar to AD169 but showed an amino acid substitution, 2 sequences were similar to the Merlin strain but also had a single amino acid substitution, and the three samples classified as genotype 3 showed differences in one amino acid compared to the Toledo strain. In addition, 3 sequences were identical to the Merlin strain at the cleavage site but showed distinct aminoacidic changes associated with genotype 3 strains (such as the Toledo strain) at the carboxy-terminal end of the sequence (fig. 3). Interestingly, the amino acid substitution present in UASLP/24050/11 corresponds to the change of an aspartic acid normally present in genotype 1 for an asparagine, and the genotype 1 sequence UASLP/24076/09 changes a serine for a proline; both substitutions correspond to amino acids having different physicochemical properties and, as such, might impact on overall tridimensional protein structure, folding or function.

Discussion

CMV infections have a high prevalence worldwide. Most infections caused by this virus are asymptomatic and not associated with a clinically significant impact. However, certain subjects, such as those suffering from immunosuppression may experience severe infections. In addition, pregnant women are at risk of transmitting CMV to the fetus with the potential for long-term sequelae in the newborn. Nevertheless, even in these at-risk populations there is wide variability in the clinical features of CMV infection. Both host and viral factors play an important role at determining the evolution of CMV infections. Among viral factors, genetic variability has been analyzed in order to determine the possible association between viral genetic features and infection outcome; several genes have been targeted for evaluation including those encoding for gB, glycoprotein N, glycoprotein O, and the UL144 gene, among others [6, 11–14]. The gB gene has been studied in different risk groups, including both solid organ and hemopoietic transplant patients [5, 15–27], HIV-infected patients [4, 6, 15, 16, 28–33], infants with congenital infection [3] and healthy subjects [34]. So far, studies directed towards assessing the impact of genotype features on disease have yielded conflicting results. Nevertheless, it appears that in certain circumstances strain type may have a role in the occurrence of CMV disease; for instance, most studies in HIV-positive patients report gB2 as the most common genotype [4, 6, 15, 16, 28–33], while in transplant patients gB1 is the pre-
dominant genotype [5, 15–27]. Also, the interpretation of these results is affected by the frequent detection of co-infections with two or more CMV strains in immunosuppressed patients [22, 33]. Although infection with multiple strains in healthy individuals has also been reported [34], the frequency and clinical implications of these findings have not been defined. For example, while the great majority of CMV seropositive women in which CMV DNA was detected had multiple strains [34], the occurrence of infection with more than one CMV strain was infrequent in women with primary infections [35]. Thus, although single sequence analyses may underestimate the frequency of infections caused by multiple strains, it is likely that this approach may identify the predominant strain in immunocompetent individuals, and this may provide clinically and epidemiologically useful information.

In addition to differences among patient populations, there could also be significant regional differences regarding the predominant genotypes [3, 36]. While infections caused by gB4 strains are not unusual in studies performed in Europe [10, 37], most reports from Asian countries fail to detect infections caused by this gB subtype [26, 27, 38, 39]. Little information about CMV strains circulating in Mexico is available. A recent report analyzing Mexican children with bone marrow transplantation found predominance of gB1 and gB2 infections (30 and 27%, respectively) [21]. Consistent with these results, the majority of CMV strains we analyzed (96%) corresponded to these genotypes; in addition, we did not identify infections caused by gB4 or gB5 genotypes among the children and women included in our study. In the light of these results, it would be interesting to pursue a larger study to allow for samples from populations other than women and children to be analyzed, including subjects with underlying conditions (e.g. HIV infection and malignant disorders) to confirm the predominance of gB1 and gB2 genotypes in Mexico.

An interesting finding in our study is the difference in the predominant strain in children compared to women during pregnancy and puerperium. In this regard, it is likely that infections in children may represent primary infections, while viruses detected in women can include primary infections as well as reactivations or reinfections. Unfortunately, we did not have serum available to further assess the presence of reinfections or co-infections. A possible explanation for these findings is that certain strains may be more readily transmitted during primary infections, while others may be more commonly associated with reactivations. Alternatively, this finding may reflect the existence of preferential viral excretion routes (saliva, urine, breast milk, etc.) for different viral genotypes. Also, transmission dynamics may differ among viral strains and within risk populations (children vs. adults). Although the task of obtaining paired samples is complex, given that most women with CMV infection during pregnancy will not transmit the infection to their offspring before birth, we consider that future studies that include strains isolated from mothers and their infants would be valuable in assessing the transmission dynamics of different viral strains.

In summary, we identified that gB genotype 1 and 2 strains predominate in Mexican children and women with CMV infections. These results provide additional information regarding CMV variability in Latin American countries.

Acknowledgements

This work was supported by a grant provided by the Mexican government through Fondo Sectorial de Investigación en Salud y Seguridad Social SSA/IMSS/ISSSTE-CONACYT (Proyecto CONACYT-SALUD-2009-01-115487).

Disclosure Statement

There is no conflict of interest for any author.

References