Identification and Function of Human Cytomegalovirus microRNAs

Finn Grey and Jay Nelson
Vaccine & Gene Therapy Institute, 505 NW 185th Ave, Oregon Health Sciences University, Portland Oregon 97006

Abstract

microRNAs are an extensive class of non-coding genes that regulate gene expression through post-transcriptional repression. These small RNAs are evolutionarily conserved and are likely to be a basic mechanism of gene regulation present within most eukaryotic organisms. Over 100 viral miRNAs have been identified to date through a combination of bioinformatics and cloning studies. In this review we discuss the use of bioinformatics for the identification of HCMV miRNAs and also for the discovery of potential target transcripts. Such studies will enable us to define the functional role of viral miRNAs and gain a better understanding of viral gene regulation.
degraded (Lim et al., 2003). Following processing by Dicer the miRNA forms a ribonucleocomplex known as the RNA induced silencing complex or RISC (Hammond et al., 2000; Hutvagner and Zamore, 2002; Mourelatos et al., 2002) that uses the associated miRNA as a guide sequence to identify target mRNAs. This leads to post-transcriptional regulation by two distinct mechanisms, depending on the level of homology of the miRNA with the mRNA target. Perfect or near perfect homology with the target mRNA leads to RISC mediated cleavage, (Elbashir et al., 2001a; Elbashir et al., 2001b) whereas less stringent homology between the miRNA and sites within the 3’ untranslated region (UTR) of the target leads to translational repression (Wightman et al., 1993; Zeng et al., 2002). A second class of small RNAs, small interfering RNAs (siRNAs), are generated from long double stranded RNAs, following direct processing by Dicer in the cytoplasm (Bernstein et al., 2001; Elbashir et al., 2001a). It is currently thought that the targeting mechanisms of siRNAs and miRNAs are similar, with both able to direct cleavage or translational repression of the target RNA through the actions of the RISC (Doench et al., 2003; Hutvagner and Zamore, 2002; Zeng et al., 2002). In plants, cleavage of target transcripts by small RNAs is a common mechanism of post-transcriptional regulation (Rhoades et al., 2002). The perfect, or near perfect, sequence homology associated with guiding cleavage of the target mRNA has made the identification of many RNAi targets in plants a relatively straightforward process. However, in animals, cleavage of the target mRNA rarely seems to occur (Yekta et al., 2004), suggesting that post-transcriptional regulation by translational repression is likely to be a more common mechanism. The lack of direct homology of animal miRNAs to their target sequences makes it considerably harder to determine their function (Enright et al., 2003; John et al., 2004). However the miRNA targets that have been identified in animals suggests that post-transcriptional regulation by RNAi is critical in a diverse array of biological functions, particularly in developmental processes, with their disruption leading to acute abnormalities (Brennecke et al., 2003; Johnston and Hobert, 2003; Lee et al., 1993; Wightman et al., 1993; Yang et al., 2004; Yekta et al., 2004).

The initial discovery of miRNAs was achieved using forward genetic studies that mapped specific genomic regions responsible for observed developmental phenotypes (Lee et al., 1993). Subsequent strategies have utilized extensive cloning and sequencing of small RNAs as well as bioinformatics approaches.

**Identification of HCMV encoded miRNAs**

To identify miRNAs encoded by human cytomegalovirus we used a comparative bioinformatics method. A computer algorithm called Stem-loop Finder (SLF) (CombiMatrix) was used to predict potential RNA transcripts from the HCMV genome that could form stem-loop secondary structures. To further refine this method the predicted stem loop sequences were compared to the closely related chimpanzee cytomegalovirus genome as we predicted that genuine miRNA sequences would demonstrate evolutionary conservation. Additional analysis of the conserved stem loops was performed using an online algorithm called miRscan, that compares two conserved sequences and determines the likelihood that they encode an miRNA based on a number of structural aspects such as the ability to form a stem-loop, symmetry of bulge loops and conservation of the predicted miRNA sequence (Lim et al., 2003). This analysis led to the identification of 13 candidate miRNA sequences encoded throughout the viral genome. Expression of five of the 13 miRNAs, miR-UL36-1, miR-UL70-1, miR-US4-1, miR-US5-1 and miR-US5-2, was confirmed through northern blot analysis (Grey et al., 2005). Following this study, and those of two other groups, a total of 11 miRNAs have been identified, including miR-UL22A-1, miR-UL112-1, miR-UL148D-1,
miR-US25-1, US25-2, US33-1 and the five previously described (Dunn et al., 2005; Pfeffer et al., 2005).

Unlike alpha and gamma herpes viruses, where miRNA genes are clustered within regions of the genome associated with latent gene expression, miRNAs of HCMV are encoded throughout the genome (figure 1). Consistent with mammalian species many of the HCMV miRNAs are encoded within intergenic regions, and in at least one case, miR-UL36-1, encoded within the spliced intron of a coding gene. A number of the miRNAs are however encoded within open reading frames. miR-UL112-1 is encoded directly antisense to the viral uracil DNA glycosylase, a situation that could theoretically lead to the cleavage of the transcript and negative regulation of the gene. Somewhat surprisingly, however, studies in our lab have shown that the UL114 transcript is resistant to cleavage by miR-UL112-1, despite high levels of both RNAs being present during acute infection (unpublished observation). miR-UL112-1 is capable of negatively regulating an antisense sequence such as that within the UL114 transcript when placed within the 3’UTR of a reporter gene, suggesting that the UL114 transcript is resistant to the effects of miR-UL112-1, possibly achieved through protective secondary structure introduced by the sequence flanking the potential cleavage site.

**Functional role of HCMV miRNAs**

Validated target transcripts and regulatory functions for the vast majority of viral miRNAs remain unknown. Two likely scenarios for viral miRNA function include targeting of cellular gene expression to induce a more favorable environment for the virus, or the regulation of viral genes to establish precise temporal or tissue specific regulation of viral gene expression. Two recent studies would suggest that both scenarios occur and that a single viral miRNA is capable of regulating both cellular and viral transcripts (Grey et al., 2007; Stern-Ginossar et al., 2007).

In a study designed to identify potential viral targets we used a comparative bioinformatic approach similar to that used to identify miRNA genes in our previous study (Grey et al., 2007). As the miRNAs encoded by HCMV are well conserved in CCMV we predicted that genuine target sequences within the viral genomes would also be conserved. Using an online algorithm called RNAhybrid we predicted potential target sites within 3’UTR sequences within the HCMV genome. As transcripts from HCMV have not been fully mapped a putative data base of predicted 3’UTR sequences for HCMV were initially created by selecting sequence from the end of each annotated open reading frame to the first canonical poly adenylation signal (AATAAA). The analysis was repeated for the CCMV genome and the results compared to identify transcripts that contained target sites within the corresponding 3’UTRs of HCMV and CCMV. The number of predicted target transcripts identified using this approached ranged from 0 – miR-UL22A-1 to 29 – miR-UL25-2.

Of particular interest were the results for the miRNA miR-UL112-1. A total of 14 transcripts were found to contain potential target sequences, including a cluster of three targets within the major immediate early region of the virus. The major immediate early region of HCMV is known to express a number of regulatory proteins crucial for the efficient replication of the virus and coordination of viral gene expression. The most abundant of these proteins are the major trans-activators IE72 and IE86, which are expressed from 5 exons within the major immediate early region (figure 2) (Stenberg et al., 1989; Stenberg et al., 1984; Stenberg et al., 1985). Both proteins are thought to be important in promoting early and late viral transcription, although only IE86 is essential for virus replication (Greaves and Mocarski, 1998; Mocarski et al., 1996; Stenberg et al., 1990; White et al., 2004; White et al., 2007). Disruption of IE72 expression does result in a significant attenuation of viral replication following low multiplicity infections (Mocarski et al., 1996). A single target site was identified within the 3’UTR of IE72
with two further target sequences within the 3’UTRs of down stream open reading frames UL120 and UL121. Intriguingly, the corresponding site identified downstream of IE72 in CCMV maintained less homology to miR-UL112-1 than the equivalent site in HCMV, and would presumably form a less effective binding site. However, a second site was identified further upstream that was not found in the HCMV counterpart, suggesting that CCMV has evolved two less effective binding sites in comparison to the more homologous single site found in HCMV. These sites, along with a predicted site within the 3’UTR of UL112/113 demonstrated effective regulation of luciferase expression when cloned down stream of the reporter gene, following expression of miR-UL112-1. Expression of miR-UL112-1 also resulted in a decrease in IE72 protein expression following co-transfection with a construct containing the major immediate early region of HCMV and also following infection with the virus. These results demonstrate that the HCMV-encoded miRNA miR-UL112-1 is able to specifically regulate the expression of one of the major immediate early trans-activators of the virus. Given that IE72 is an important transactivator and required for efficient viral replication we hypothesized that expression of miR-UL112-1 early in infection may result in attenuation of acute replication. To test this theory we transfected cells with a synthetic miR-UL112-1 RNA duplex then infected cells with HCMV and measured DNA replication of the virus. Compared with cells transfected with a control miRNA, HCMV DNA replication was inhibited up to 5 fold indicating that expression of miR-UL112-1 has the potential to attenuate acute replication of HCMV (figure 3).

In addition to regulation of viral genes, a recent report demonstrates that miR-UL112-1 also targets a cellular gene involved in host innate immune responses to viral infection (Stern-Ginossar et al., 2007). Using a target prediction algorithm based on the identification of repeated target sites within cellular 3’UTR sequences, Stern-Ginossar et al. identified a potential target sequence for miR-UL112-1 within the 3’UTR of MHC class-I related chain B (MICB). MICB is a cellular ligand for the receptor NKG2D on Natural Killer cells that is expressed in response to viral infections. Activation of NK cells through association of MICB with NKG2D can lead to cell killing. Blocking this interaction would therefore be of obvious benefit to the virus as emphasized by the demonstration that deletion of miR-UL112-1 led to an increased susceptibility to NK killing. Interestingly, HCMV also encodes a protein, UL16, which was previously shown to inhibit MICB signaling by intracellular sequestration, suggesting a level of redundancy in targeting of MICB (Dunn et al., 2003).

Physiological relevance of miR-UL112-1 targeting of IE72 and MICB

These studies have clearly demonstrated that a viral miRNA is able to regulate both cellular and viral genes. However it is still a question as to why the virus would target these genes by miRNA expression. Targeting of MICB seems relatively understandable, although the necessity to target MICB by miRNA when the virus expresses a protein for the same function is less clear. Targeting of a viral gene required for efficient replication also seems counterintuitive. One possible explanation for both of these anomalies is that miR-UL112-1 may be specifically required for the establishment or maintenance of latent or persistent infection rather than during acute replication of the virus. Cytomegaloviruses establish lifelong infections of their hosts, during which time the virus maintains a restricted replication profile either as a persistent infection, with production of low levels of virus, or a true latent infection, in which virus is only produced following a reactivation stimulus. To achieve this, the virus must employ strict regulation of viral gene expression and mechanisms to evade the host immune system. The non-immunogenic nature of miRNAs and their ability to target multiple genes, both host and viral, would make them ideal agents of gene regulation during latent or persistent infection. Expression of miR-UL112-1 may allow the virus to restrict the expression of immunogenic viral genes involved in acute replication, while at the same time modulating the host immune response to protect the infected cell from NK induced cell death.
The major immediate early genes, including IE72, have been suggested to play pivotal roles in controlling latency and reactivation as both genes are important in driving the expression of early and late genes required for acute replication of the virus. Furthermore expression of the functional homologues of IE72 and IE86 in latent MCMV infections of mouse lungs have suggested that the regulated expression of these genes may represent a significant control step in the triggering of reactivation (Simon et al., 2006). Therefore regulation of IE72 would be an effective switch mechanism and as demonstrated by our studies showing miR-UL112-1 has the capacity to restrict acute replication.

The redundancy observed in the targeting of MICB may also be a phenomenon of HCMV latency. The expression of UL16 protein may be a more robust mechanism of targeting MICB, required during acute replication of the virus, when the host immune response is likely to be more vigorous. However targeting of MICB by miR-UL112-1 may be more suitable for long term regulation during persistent or latent infection where it is likely the less robust nature of miRNA regulation would be adequate, while avoiding the production of potentially immunogenic viral proteins.

The clinical relevance of these findings is two fold. There are currently small molecule inhibitors of miRNAs that are highly effective and specific. Targeting of miR-UL112-1 using such inhibitors could drive the virus towards acute replication, while at the same time disrupting the inhibition of MICB expression, resulting in possible immune clearance of the virus. This has potential risks of pathological effects caused by acutely replicating HCMV, especially in immunocompromised patients – which are also the patients most in need of antiviral therapy. A second option would be to mimic the expression of miR-UL112-1 using synthetic miRNAs in an attempt to attenuate acute replication of the virus. Unlike many antivirals that artificially block viral processes, delivery of endogenous viral miRNAs could exploit the virus’s own mechanisms to subdue replication. Not only might this approach be effective, but it could also be less prone to the problems of viral escape and resistance as the virus has evolved to maintain these mechanisms. Potential drawbacks include off target effects of delivering high levels of synthetic miRNAs and also the risk of immune suppression due to regulation of MICB.

Given the potential for the regulation of multiple viral and cellular genes it is likely that additional studies will demonstrate important regulatory roles for each of the HCMV miRNAs. Although challenging, it will be of particular interest to investigate the potential role of HCMV miRNAs in the establishment and maintenance of persistent and latent infections and the possible development of therapeutic agents based on the function of viral miRNAs.

References


J Clin Virol. Author manuscript; available in PMC 2009 July 6.


Figure 1.
Genomic map of HCMV miRNAs. Known HCMV miRNAs are represented by black arrows indicating genomic position and orientation in relation to predicted and known open reading frames.
Figure 2.
Genomic position of miR-UL112-1 targets sites within the major immediate early region. The major immediate early region showing IE72 and IE86 transcripts resulting from alternative splicing are shown as well as splicing to down stream exons UL120/UL121. Position of miR-UL112-1 target sites in 3’UTR region of IE72 and UL120/UL121 indicated by red boxes (Grey et al., 2007).
Figure 3.
DNA was isolated from infected cells at indicated time points and viral DNA levels determined by real time PCR and corrected against beta-actin copies. Transfection of miR-UL112-1 (syn miR-UL112-1) results in consistently lower levels of HCMV DNA at later time points compared to cells transfected with randomized control miRNA (syn negative) (Grey et al., 2007).