UNIVERSITY OF CALIFORNIA, SAN DIEGO

Transcriptome analysis of human cytomegalovirus infection: From microRNAs to altered trends in host RNA processing.

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy

in

Biology

by

Thomas J. Stark

Committee in charge:

Professor Gene W. Yeo, Chair
Professor Amy E. Pasquinelli, Co-Chair
Professor Lawrence S. Goldstein
Professor Jens Lykke-Andersen
Professor Deborah H. Spector

2014

The Dissertation of Thomas J. Stark is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Co-Chair

Chair

University of California, San Diego

2014

DEDICATION

To my mom and dad, for your unconditional support,

and for remembering that I study RNA.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

First and foremost, I must extend my greatest thanks to my mentor and dissertation chair, Gene Yeo. My graduate career has been productive and engaging throughout thanks to Gene, and I will be forever spoiled by his quick responses, 24/7 virtual availability, and constant eagerness to hear about new results and data. Gene also pushed for me to incorporate bioinformatics into the initial stages of my training, and I am eternally grateful to have learned data analysis skills and how to apply them to my research. This bioinformatics training enabled me to author a publication early on, and my hybrid experimental-computational status has now helped me to land an exciting fellowship opportunity with the Centers for Disease Control.

I also thank my unofficial co-mentor Deborah Spector for serving on my committee, for her generous support from the very beginning of my studies, and above all, for the initial introduction to Gene. Debbie kindly allowed me to conduct many cell culture experiments in her lab, and the Spector lab has always felt like a second lab-home for me. I am grateful to have been included in several of their lab meetings as well. I additionally wish to thank the rest of my committee members: Amy Pasquinelli (Co-Chair), Jens Lykke-Andersen, and Larry Goldstein. I especially thank Amy for her mentorship and encouragement outside of committee meetings.

During my graduate school career I have been very fortunate in the personal connections I have made. I must first thank my close friends and colleagues from the Yeo lab for all of their support, especially fellow graduate students Stephanie Huelga and Melissa Wilbert. On a professional note, I credit Stephanie for my Adobe Illustrator skills and am grateful for all of her help in my early bioinformatics days. Kasey Hutt and

Michael Lovci have also been both great friends and helpful with informatics. I thank Katlin Massirer and Dave Nelles for their company in the lab well into the night, and I thank Justin Arnold for his friendship and for staying in touch since he left the lab. From the original CMM-east Yeo lab crew, I also thank Anthony Vu and Patrick Liu for good times in and out of the lab. To my lab undergraduate assistant, I thank Brett Roberts for his efforts and for working with me for two years. I would like to also thank friends from neighboring labs, Stephanie Grainger, Paulo Marinho, and especially Nisha Patel, for being one of my closest friends and for support during my more difficult stages of graduate school. I wish to additionally recognize fellow graduate student friends from Biology: Amanda Mason, Samuel Ulin, Russell DeKelver, Tiffany Dunbar, Helena Sun, Pagkapol Pongsawakul, Elizabeth Clark, and Jesse Vargas.

I wish to thank specific members of the Spector lab, especially Rebecca Routh (Sanders) for guidance when I was first starting, and for teaching me most of the virology techniques that I know to this day. I also sincerely thank Jean-Philippe Belzile for his collaboration during the second half of my studies and for our intellectual discussions. From the Bing Ren lab, I would like to specifically acknowledge Anthony Schmitt. The preliminary studies outlined in Chapter 4 have been possible thanks to considerable time and assistance from him.

Finally, I must of course thank my family for their continuous love and support throughout all of my endeavors. I am thankful to my parents for not being too upset with my relocation away from Atlanta, and I thank my younger brother Joseph for keeping them company while I have been across the country. I especially thank my parents and grandmother for taking trips out to San Diego to visit me, and for attending my final dissertation defense.

Chapter 2, in full, has been published as a manuscript entitled “High-Resolution Profiling and Analysis of Viral and Host Small RNAs during Human Cytomegalovirus Infection” in *Journal of Virology* (Stark et al., 2012). The author of this dissertation was the primary investigator and author of this paper.

Chapter 3, in full, has been submitted for publication, with the title “Induction of CPEB1 expression and host 3′UTR shortening during human cytomegalovirus infection.” The author of this dissertation was the primary investigator and author of this manuscript.

VITA

Education

2014 Doctor of Philosophy, Biology
University of California, San Diego

2008 Bachelor of Science, Biology and Music
Emory University

Professional Experience

2013 Intern (Summer), Research and Development
Life Technologies, Carlsbad, CA

2005-2008 Research Assistant, Laboratory of Dr. Guy Benian
Emory University, Atlanta, GA

2005-2008 Office Assistant
Express Employment Professionals, Atlanta, GA

2007 Visiting Researcher (Summer), Laboratory of Alex Andrianopoulos
University of Melbourne, Australia

Publications

Stark TJ, Belzile JP, Roberts BT, Huelga SC, Spector DH, Yeo GW. Induction of CPEB1 expression and host 3′UTR shortening during human cytomegalovirus infection. *Submitted.*

Van Wynsberghe PM, Finnegan EF, **Stark T**, Angelus EP, Homan KE, Yeo GW, Pasquinelli AE. The Period protein homolog LIN-42 negatively regulates microRNA biogenesis in C. elegans. *Dev Biol.* 2014.

Belzile JP, **Stark TJ**, Yeo GW, Spector DH. Human Cytomegalovirus Infection of Human Embryonic Stem Cell-Derived Primitive Neural Stem Cells is Restricted at Several Steps but Leads to the Persistence of Viral DNA. *J Virol*. 2014.

Lovci MT, Ghanem D, Marr H, Arnold J, Gee S, Parra M, Liang TY, **Stark TJ**, Gehman LT, Hoon S, Massirer KB, Pratt GA, Black DL, Gray JW, Conboy JG, Yeo GW. Rbfox proteins regulate alternative mRNA splicing through evolutionarily conserved RNA bridges. *Nat Struct Mol Biol*. 2013.

Wilbert ML, Huelga SC, Kapeli K, **Stark TJ**, Liang TY, Chen SX, Yan BY, Nathanson JL, Hutt KR, Lovci MT, Kazan H, Vu AQ, Massirer KB, Morris Q, Hoon S, Yeo GW. LIN28 binds messenger RNAs at GGAGA motifs and regulates splicing factor abundance. *Mol Cell*. 2012.

Stark TJ, Arnold JD, Spector DH, Yeo GW. High-Resolution Profiling and Analysis of Viral and Host Small RNAs during Human Cytomegalovirus Infection. *J Virol*. 2012.

Miller RK, Qadota H, **Stark TJ**, Mercer KB, Wortham TS, Anyanful A, Benian GM. CSN-5, a component of the COP9 signalosome complex, regulates the levels of UNC-96 and UNC-98, two components of M-lines in *C. elegans* muscle. *Mol Biol Cell*. 2009.

Qadota H, McGaha LA, Mercer KB, **Stark TJ**, Ferrara TM, Benian GM. A novel protein phosphatase is a binding partner for the protein kinase domains of UNC-89 (Obscurin) in *C. elegans*. *Mol Biol Cell*. 2008.

Teaching Experience

2010-2012 Research Mentor, for Brett Roberts, Yeo Laboratory
University of California, San Diego

2010-2012 Teaching Assistant, Biological Sciences
University of California, San Diego

Honors and Affiliations

2013 1st Place Poster Presentation
Life Technologies Summer Intern Poster Session

2009-2013 Graduate Research Fellowship, Genetics Training Program
University of California, San Diego

2012 Travel Fellowships, The RNA Society and CMV conferences

2010-2012 Member, The RNA Society

2012 Member, Biological Sciences Graduate Admissions Committee
University of California, San Diego

2011-2012 Member, American Association for the Advancement of Science (AAAS)

2010-2011 Peer Mentor, Biological Sciences Ph.D. Mentoring Program
University of California, San Diego

2009-2012 Student Host, Biological Sciences Graduate Recruitment
University of California, San Diego

ABSTRACT OF THE DISSERTATION

Transcriptome analysis of human cytomegalovirus infection:
From microRNAs to altered trends in host RNA processing.

by

Thomas J. Stark

Doctor of Philosophy in Biology

University of California, San Diego, 2014

Professor Gene W. Yeo, Chair
Professor Amy E. Pasquinelli, Co-Chair

Human cytomegalovirus (HCMV) is of clinical importance and is the leading viral cause of birth defects. HCMV infections are also problematic for immunocompromised individuals and organ transplant recipients. This virus infects a variety of cell types *in vivo*, and although laboratory studies have demonstrated that HCMV does not productively infect embryonic stem cells, the main symptoms in newborns constitute defects in early nervous system development. While the HCMV life cycle has been studied extensively in human fibroblasts, our understanding is limited in terms of how

different cell types in the host contribute to the overall infection. The HCMV genome not only has the capacity to encode hundreds of protein products, but also itself encodes microRNAs (miRNAs), and these viral miRNAs reach high abundance during the infection. Here, I first describe comprehensive deep-sequencing characterization of both viral and human miRNAs during lytic HCMV infection of primary fibroblasts. In addition to the identification of two new HCMV-encoded miRNAs, my results also include the finding that the genomic cluster of human miRNAs 182/96/183 is highly upregulated by late stages of infection. For the second stage of my HCMV studies, I present analysis of the full RNA transcriptomes of three different primary cell types undergoing infection. This work was initiated to compare infected neural progenitors with other more permissive cell types, namely fibroblasts and endothelial cells. In my comparative transcriptome analysis I uncovered a wealth of new findings related to conserved effects of HCMV infection on host RNA processing patterns. Finally, I report preliminary results from high-throughput chromosome conformation studies, which have suggested the existence of virus-host DNA interactions during HCMV infection.

Chapter 1: Introduction

1.1. Human cytomegalovirus (HCMV) and clinical relevance

HCMV is a member of the family *Herpesviridae*, which includes eight DNA genome viruses that infect humans. Sub-classifications of this group include alpha-, beta-, and gamma-herpesviruses. HCMV belongs to the beta-herpesvirus subgroup, and has a relatively long viral replication cycle in cells, taking 3-5 days to produce new infectious virions, depending on the strain and cell type infected. The alpha subclass includes much faster replicating viruses, such as herpes simplex virus (HSV). Further distinct from these two groups, the gamma subclass of oncogenic herpesviruses includes members such as Esptein-Barr virus (EBV).

Similar to other herpesviruses, HCMV maintains a latent infection and persists throughout the life of the host. Representing an astounding global health statistic, the majority of people in developed nations are infected with HCMV at varying stages of life. Adulthood sero-prevalence in the United States has been estimated near 100% (Staras et al., 2006). Although this latent infection does not have clear negative consequences for immunocompetent individuals, organ transplant recipients and AIDS patients suffer detrimental effects associated with the infection, and HCMV has been identified as the leading viral cause of birth defects (Mocarski et al., 2007). The risk of such defects is greatest when a mother is infected for the first time during the course of pregnancy. In newborns, the damaging effects of infection largely impact the developing central nervous system and are manifested as varying levels of neurological impairment. *In vivo* signs of infection have been observed in most cell types within the human body, and it is not known which subpopulations of cells are the most important for the pathogenesis associated with symptomatic HCMV infections.

1.2. Characteristics of the HCMV genome

Viral genomes are notoriously overcrowded with genetic information and the layout of the HCMV genome is no exception. Even though the HCMV genome is rather large at ~235 kb, there are over 200 potential open reading frames (ORFs) encoded. Possibly double this historically proposed number or more exist based on recent ribosome profiling studies on HCMV-infected cells (Stern-Ginossar et al., 2012). Similar to other viruses, HCMV transcription patterns can be subdivided into immediate-early, early, and late expression classes. One main defining characteristic, the immediate-early RNAs are expressed in the absence of new viral protein synthesis.

A number of HCMV mRNAs undergo splicing, and a well-described case of temporally regulated alternative splicing occurs for the RNAs UL122 and UL123 (also known as IE2 and IE1) that are initially expressed at an immediate-early point of infection (Sanchez et al., 2004).

1.3. Molecular impacts of lytic HCMV infection

HCMV infection has been typically studied in its lytic form in permissive human foreskin fibroblasts (HFFs). In terms of our understanding of the impact of HCMV on host gene expression, we have come a long way from initial work that demonstrated at least nine human RNAs were differential following infection (Tanaka et al., 1975). More recently microarray studies were conducted to obtain a more global view of this impact (Browne et al., 2001; Hertel and Mocarski, 2004) and we now know that hundreds of host genes are affected at the RNA level. While it might not be surprising to find this large effect on host mRNAs, interestingly HCMV infection is different from a number of other viral infections that more generally suppress and down-regulate the majority of

host genes. Therefore, HCMV both up- and down-regulates hundreds of host mRNAs (in the context of a general increase in total RNA levels) and this contributes to a diversity of changes that are even more apparent at the host protein level in HCMV-infected cells. Details of the host transcriptome response and significantly affected genes are further discussed in Chapter 3.

1.4. State of genomics in HCMV research

A major goal of these dissertation projects has been to incorporate high-throughput sequencing methodologies into studies pertaining to HCMV-host RNA biology. At the start of this investigation in 2009, no next-generation sequencing technologies had been applied to HCMV infection studies. The first fully sequenced HCMV genome became available in 1990 [15], deriving from the AD169 laboratory-adapted strain of HCMV, and the first rounds of ORF predictions were made. Since that time it has become increasingly appreciated that the HCMV genome harbors a great deal of transcriptional complexity. For example, a number of ORFs overlap, many instances of sense-antisense coding regions can be found throughout the HCMV genome, and there are examples of collections of ORFs that share a co-terminal poly(A) site.

Now, after genome sequences for a number of clinical and laboratory-adapted strains had been available for two decades, obtained through more traditional sequencing methods, new assemblies of HCMV genome sequences from Illumina DNA sequencing were reported from the Davison group in 2010 (Cunningham et al., 2010). For the purposes of HCMV genome sequencing, the Illumina-based approach enabled sequencing assembly from limited clinical specimen material, and was advantageous in that it did not require cell culture propagation of the virus.

For analysis of the combined virus-host transcriptome, the Davison laboratory also recently reported analysis of HCMV splicing through high-throughput sequencing of HFFs infected at a late time-point with the reference Merlin strain. This work presented evidence of over 200 total virus splice junctions (Gatherer et al., 2011). Before this study only 13 HCMV spliced transcripts had been identified. Additionally, another deep sequencing study, which performed ribosome profiling of infected HFFs, also detected this increased number of introns encoded in the HCMV genome (Stern-Ginossar et al., 2012).

1.5. Virus-encoded microRNAs

MicroRNAs (miRNAs) are small ~22 nt RNAs that canonically repress translation through imperfect base-pairing in the 3′UTRs of mRNAs. These small RNAs can be found throughout animal and plant species. Hundreds to over 1,000 of these regulatory molecules are expressed in human cells, each of which is predicted to target multiple gene transcripts (Bartel, 2004). In 2004 the fascinating discovery was made that certain DNA virus genomes encode their own miRNAs (Pfeffer et al., 2004), and since then the phenomenon has been extended to include several other viruses (Cai et al., 2005; Cai et al., 2006; Cui et al., 2006; Dunn et al., 2005; Grey et al., 2005; Pfeffer et al., 2005). Virus-encoded miRNAs are very interesting in that they have the potential to target not only host genes, but also their own RNAs in an auto-regulatory fashion. Further, they have the capacity to evade the host immune system, and it is unclear whether they have undergone the evolutionary pressure faced by endogenously expressed host miRNAs. Based on computational predictions and Northern blot follow-up analysis, HCMV has been shown to express 11 miRNA precursors during infection (Grey et al., 2005). Although miRNAs encoded by the murine ortholog of CMV (MCMV) are located in main

genomic clusters, HCMV miRNAs are found scattered throughout its genome (locations shown in Figure 1.1). With the exception of some chimpanzee cytomegalovirus-encoded miRNAs, HCMV miRNA sequences are not similar to any miRNAs from other herpesviruses. In terms of their processing, some examples of HCMV miRNAs such as miR-US25-2 are present in both possible mature miRNA arm forms by late stages of infection (Grey et al., 2005).

A number of studies have now investigated potential targets of HCMV miRNAs. One of the first and most convincing sets of functional targets has been described for miR-UL112, and the targets were found to reside within the viral transcriptome itself (Grey et al., 2007). Curiously, no HCMV miRNAs have been demonstrated to be essential for viral replication, although experimental modulation of the levels of miR-US25-1 and miR-US25-2 influenced viral titers (Stern-Ginossar et al., 2009). At this point, the role of HCMV miRNAs, and viral miRNAs in general, is still not well understood.

1.6. Approaches for identification of viral miRNA targets

Many computational efforts have now been employed in attempts to assign viral miRNAs to their respective gene targets. Examples have included exogenous expression in non-infected cells followed by biochemical isolation of miRNA silencing complexes and microarray profiling (Grey et al., 2010). One technique that has offered the greatest promise is the CLIP-seq methodology (crosslinking and immunoprecipitation coupled with high-throughput sequencing) (Ule et al., 2003). In this approach, Argonaute complexes, the key mediators of miRNA-mRNA target pairing, are immunoprecipitated from cells, associated mRNA material (including bound miRNAs) is fragmented down to the size protected by these complexes, and following adaptor ligation steps and cDNA

synthesis the final material is subjected to high-throughput sequencing. This technique has been proven effective in the search for miRNA targets in human cells and in nematodes (Chi et al., 2009; Zisoulis et al., 2010). Nevertheless, the bioinformatics follow-up of these experiments remains challenging in the post-sequencing assignment of miRNAs to mRNA targets as these fractions are obtained separately. There are a number of additional versions of CLIP now implemented to date, and the most recent adaptation termed CLASH is especially promising in that this strategy enables the direct isolation of miRNA-target read chimeras (Helwak et al., 2013).

1.7. HCMV infection of neural lineages

As mentioned above, HCMV infection is widespread in the population, and most individuals are infected at some stage. The main problems associated with infection are in particular suffered by newborns congenitally infected with HCMV. Most of the detrimental outcomes are neurodegenerative and severe impacts are experienced at the level of the developing central nervous system. Approximately 1% of infants are born HCMV-positive, and 10-15% of these infected newborns are unfortunately symptomatic at birth (Cheeran et al., 2009).

There exists long-standing evidence that HCMV is incapable of productively infecting embryonic stem cells (ES) cells (Gonczol et al., 1984; Penkert and Kalejta, 2013), but the pathology of disease in affected newborns implies that a class of multipotent progenitor cells must be involved that permits lytic infection. Provided the observed defects in nervous system development in symptomatic individuals, neural stem cells and related precursors could represent key infection targets in early cellular differentiation events.

To gain molecular insight into this pathology, neural precursor models have been considered in the context of HCMV infection, and studies using fetal-derived neural progenitor cells (NPCs) have demonstrated their susceptibility (Cheeran et al., 2005; Luo et al., 2008). Infection in these cells when cultured as neurospheres led to accelerated differentiation, and microarray analysis of the infected cultures identified downregulation of neural-related genes (Luo et al., 2010). Extended passages of NPCs were also shown to render the cells increasingly permissive (Pan et al., 2013). Additionally, a recent report demonstrated to some extent that rosette-differentiated NPCs from induced pluripotent stem (iPS) cells are also permissive to infection (D'Aiuto et al., 2012). Our recent studies in stem cell-derived NPCs have greatly contributed to our understanding of infection in these progenitor populations (Belzile et al., 2014), and was a foundation for global transcriptome studies described further in Chapter 3.

To summarize the detailed findings of Belzile et al., primitive neural stem cells (pNSCs) derived from human ES cells were observed to be resistant to infection with the HCMV clinical isolate TB40E. Although the virus was able to enter these cells and viral DNA could be detected in the nuclei, a number of roadblocks to further infection progression were evident. NPCs differentiated from the pNSCs were more permissive to early stages of infection but also exhibited signs of restricted late stages of infection. New virions could be produced from a sub-population of the pNSCs and NPCs, yet only cell-associated infectious particles could be found, and their tropism was altered and non-neuronal. Interestingly, viral genomes could persist over time in the pNSCs, and active infection resurfaced following neural differentiation of these populations (Belzile et al., 2014). Building on the previous studies, these findings again point to the susceptibility of neural progenitor populations to HCMV infection, but further stress the importance of the differentiation status of these cells and highlight the significance of

characterizing specific neural cell types rather than a heterogeneous mix of neural lineages.

1.8. Cellular RNA processing

The effect of lytic HCMV infection on the host cell at the molecular level is vast and highly complex. Despite the non-progressive nature of infection in hES-derived NPCs, it has become apparent that host gene expression is still significantly differential in HCMV-infected neural cells (examples first described in Luo et al., 2010 and D'Aiuto et al., 2012, with further extensive analysis featured in Chapter 3). While this aspect of host cell responses has been recognized, it is still not known whether HCMV infection of cells (neural lineages and other more permissive cell types included) affects host RNAs in other ways beyond their overall levels. Related to this open-ended question in the study of HCMV-host interactions, two key modes of RNA processing are important for further discussion.

1.8.1. Alternative splicing

In eukaryotic organisms, during the maturation of precursor messenger RNAs (pre-mRNAs) within the cell nucleus to processed mRNAs that are ready for export to the cytoplasm, select regions termed introns are removed from pre-mRNAs (Sharp, 2005). Intronic regions often constitute the bulk of the transcript, and remaining exon sequences are stitched together, altogether known as RNA splicing. These reactions are carried out as a number of distinct reaction steps by the splicesome, a very large complex composed of several ribonuclearproteins. The process of splicing is typically required to generate an RNA message capable of being translated into a protein product.

The majority of genes have multiple options in terms of their splicing patterns, and a number of configurations for generating alternatively spliced mRNAs are possible. For example, portions of intronic sequence could be retained for particular RNA isoforms, or entire exons can be skipped in the process of splicing together a final mRNA transcript. In a given cell, multiple isoforms or versions of an mRNA from the same gene are often expressed, or in other cases specific isoforms are only found in certain tissue types or developmental contexts. Regulation of alternative splicing is a crucial aspect of cellular homeostasis and these patterns are coordinated by a variety of RNA-binding proteins. Before these dissertation studies, It was not known whether alternative host splicing patterns, or splicing more generally, are suppressed or altered by HCMV infection. Although viral infections have been documented to significantly inhibit host RNA splicing activities, HCMV infection has not been thought to interfere with these processes as it encodes highly abundant and essential spliced viral mRNAs that are expressed throughout the infection. Chapter 3 features a comprehensive analysis of these RNA events in the host cell during HCMV infection.

1.8.2. Alternative polyadenylation

Mature mRNAs, including those for protein-coding genes and many non-coding transcripts, harbor a 5′ cap structure and a 3′ polyadenylated tail. To focus on the 3′ modification, in the final maturation steps of pre-mRNA processing, the 3′-end location of the RNA is determined, cleavage of the transcript occurs at this site, and a terminal poly(A) tail that confers stability is formed. (Di Giammartino et al., 2011) In eukaryotic cells, as many as 20 different factors coordinate 3′-end cleavage and polyadenylation (Mandel et al., 2008). Major regulatory complexes where these factors can be found include the Cleavage Factor (CF) I and II, the Cleavage stimulation Factor (CstF), and the Cleavage and Polyadenylation Specificity Factor (CPSF). Over half of all known

human genes have alternative polyadenylation sites (Tian et al., 2005). A classic early example of the significance of human alternative polyadenylation is exhibited by immunoglobulin heavy chain M (IgM). This case is actually mediated in conjunction with alternative splicing, and during B cell activation the IgM mRNA structure shifts to include an upstream 3′-end proximal polyadenylation site. This shift allows for a significant change in the properties of the protein product that are relevant to this cellular context (Takagaki et al., 1996).

Although not a direct subject of experimental studies presented here, it should be noted that HCMV mRNAs also harbor canonical poly(A) tails and they are processed by the host machinery. Further, returning to the example of the immediate-early UL122-123 transcripts, in this example the alternative polyadenylation is coupled with and is preceded by the alternative splicing regulation at this locus (mentioned above) (Sanchez et al., 2004).

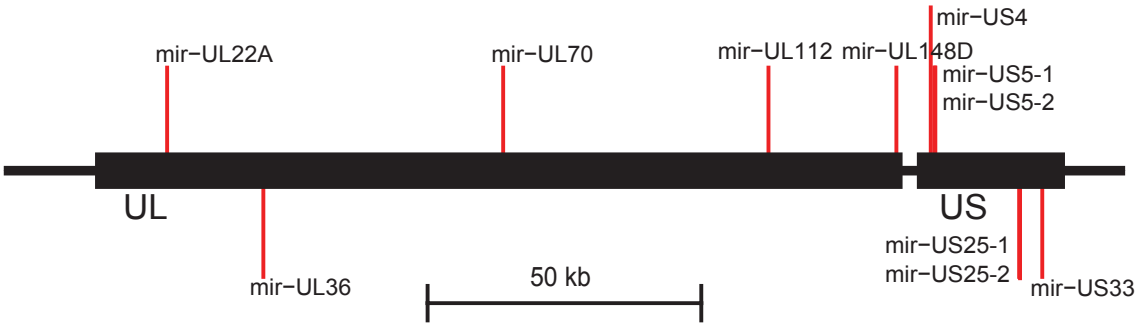

Figure 1.1: HCMV-encoded microRNAs

HCMV genome layout and the locations of microRNA genes known prior to this dissertation study. Many of these locations, which encode miRNA precursors, are processed into two individual mature miRNAs (otherwise pre-miRNAs typically yield one main mature miRNA). Two additional HCMV miRNAs discovered through this work are described in Chapter 2.

1.9. References

Bartel, D.P. (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. Cell *116*, 281-297.

Belzile, J.P., Stark, T.J., Yeo, G.W., and Spector, D.H. (2014). Human cytomegalovirus infection of human embryonic stem cell-derived primitive neural stem cells is restricted at several steps but leads to the persistence of viral DNA. J Virol *88*, 4021-4039.

Browne, E.P., Wing, B., Coleman, D., and Shenk, T. (2001). Altered cellular mRNA levels in human cytomegalovirus-infected fibroblasts: viral block to the accumulation of antiviral mRNAs. J Virol *75*, 12319-12330.

Cai, X., Lu, S., Zhang, Z., Gonzalez, C.M., Damania, B., and Cullen, B.R. (2005). Kaposi's sarcoma-associated herpesvirus expresses an array of viral microRNAs in latently infected cells. Proc Natl Acad Sci U S A *102*, 5570-5575.

Cai, X., Schafer, A., Lu, S., Bilello, J.P., Desrosiers, R.C., Edwards, R., Raab-Traub, N., and Cullen, B.R. (2006). Epstein-Barr virus microRNAs are evolutionarily conserved and differentially expressed. PLoS Pathog *2*, e23.

Cheeran, M.C., Hu, S., Ni, H.T., Sheng, W., Palmquist, J.M., Peterson, P.K., and Lokensgard, J.R. (2005). Neural precursor cell susceptibility to human cytomegalovirus diverges along glial or neuronal differentiation pathways. J Neurosci Res *82*, 839-850.

Cheeran, M.C., Lokensgard, J.R., and Schleiss, M.R. (2009). Neuropathogenesis of congenital cytomegalovirus infection: disease mechanisms and prospects for intervention. Clinical microbiology reviews *22*, 99-126, Table of Contents.

Chi, S.W., Zang, J.B., Mele, A., and Darnell, R.B. (2009). Argonaute HITS-CLIP decodes microRNA-mRNA interaction maps. Nature *460*, 479-486.

Cui, C., Griffiths, A., Li, G., Silva, L.M., Kramer, M.F., Gaasterland, T., Wang, X.J., and Coen, D.M. (2006). Prediction and identification of herpes simplex virus 1-encoded microRNAs. J Virol *80*, 5499-5508.

Cunningham, C., Gatherer, D., Hilfrich, B., Baluchova, K., Dargan, D.J., Thomson, M., Griffiths, P.D., Wilkinson, G.W., Schulz, T.F., and Davison, A.J. (2010). Sequences of complete human cytomegalovirus genomes from infected cell cultures and clinical specimens. J Gen Virol *91*, 605-615.

D'Aiuto, L., Di Maio, R., Heath, B., Raimondi, G., Milosevic, J., Watson, A.M., Bamne, M., Parks, W.T., Yang, L., Lin, B., Miki, T., Mich-Basso, J.D., Arav-Boger, R., Sibille, E., Sabunciyan, S., Yolken, R., and Nimgaonkar, V. (2012). Human induced pluripotent stem cell-derived models to investigate human cytomegalovirus infection in neural cells. PLoS One *7*, e49700.

Di Giammartino, D.C., Nishida, K., and Manley, J.L. (2011). Mechanisms and consequences of alternative polyadenylation. Mol Cell *43*, 853-866.

Dunn, W., Trang, P., Zhong, Q., Yang, E., van Belle, C., and Liu, F. (2005). Human cytomegalovirus expresses novel microRNAs during productive viral infection. Cell Microbiol *7*, 1684-1695.

Gatherer, D., Seirafian, S., Cunningham, C., Holton, M., Dargan, D.J., Baluchova, K., Hector, R.D., Galbraith, J., Herzyk, P., Wilkinson, G.W., and Davison, A.J. (2011). High-resolution human cytomegalovirus transcriptome. Proc Natl Acad Sci U S A *108*, 19755-19760.

Gonczol, E., Andrews, P.W., and Plotkin, S.A. (1984). Cytomegalovirus replicates in differentiated but not in undifferentiated human embryonal carcinoma cells. Science *224*, 159-161.

Grey, F., Antoniewicz, A., Allen, E., Saugstad, J., McShea, A., Carrington, J.C., and Nelson, J. (2005). Identification and characterization of human cytomegalovirus-encoded microRNAs. J Virol *79*, 12095-12099.

Grey, F., Meyers, H., White, E.A., Spector, D.H., and Nelson, J. (2007). A human cytomegalovirus-encoded microRNA regulates expression of multiple viral genes involved in replication. PLoS Pathog *3*, e163.

Grey, F., Tirabassi, R., Meyers, H., Wu, G., McWeeney, S., Hook, L., and Nelson, J.A. (2010). A viral microRNA down-regulates multiple cell cycle genes through mRNA 5'UTRs. PLoS Pathog *6*, e1000967.

Helwak, A., Kudla, G., Dudnakova, T., and Tollervey, D. (2013). Mapping the human miRNA interactome by CLASH reveals frequent noncanonical binding. Cell *153*, 654-665.

Hertel, L., and Mocarski, E.S. (2004). Global analysis of host cell gene expression late during cytomegalovirus infection reveals extensive dysregulation of cell cycle gene expression and induction of Pseudomitosis independent of US28 function. J Virol *78*, 11988-12011.

Luo, M.H., Hannemann, H., Kulkarni, A.S., Schwartz, P.H., O'Dowd, J.M., and Fortunato, E.A. (2010). Human cytomegalovirus infection causes premature and abnormal differentiation of human neural progenitor cells. J Virol *84*, 3528-3541.

Luo, M.H., Schwartz, P.H., and Fortunato, E.A. (2008). Neonatal neural progenitor cells and their neuronal and glial cell derivatives are fully permissive for human cytomegalovirus infection. J Virol *82*, 9994-10007.

Mandel, C.R., Bai, Y., and Tong, L. (2008). Protein factors in pre-mRNA 3'-end processing. Cellular and molecular life sciences : CMLS *65*, 1099-1122.

Mocarski, E., Shenk, T., and Pass, R. (2007). Cytomegaloviruses. Fields virology *1*, 2702-2772.

Pan, X., Li, X.J., Liu, X.J., Yuan, H., Li, J.F., Duan, Y.L., Ye, H.Q., Fu, Y.R., Qiao, G.H., Wu, C.C., Yang, B., Tian, X.H., Hu, K.H., Miao, L.F., Chen, X.L., Zheng, J., Rayner, S.,

Schwartz, P.H., Britt, W.J., Xu, J., and Luo, M.H. (2013). Later passages of neural progenitor cells from neonatal brain are more permissive for human cytomegalovirus infection. J Virol *87*, 10968-10979.

Penkert, R.R., and Kalejta, R.F. (2013). Human embryonic stem cell lines model experimental human cytomegalovirus latency. MBio *4*.

Pfeffer, S., Sewer, A., Lagos-Quintana, M., Sheridan, R., Sander, C., Grasser, F.A., van Dyk, L.F., Ho, C.K., Shuman, S., Chien, M., Russo, J.J., Ju, J., Randall, G., Lindenbach, B.D., Rice, C.M., Simon, V., Ho, D.D., Zavolan, M., and Tuschl, T. (2005). Identification of microRNAs of the herpesvirus family. Nat Methods *2*, 269-276.

Pfeffer, S., Zavolan, M., Grasser, F.A., Chien, M., Russo, J.J., Ju, J., John, B., Enright, A.J., Marks, D., Sander, C., and Tuschl, T. (2004). Identification of virus-encoded microRNAs. Science *304*, 734-736.

Sanchez, V., McElroy, A.K., Yen, J., Tamrakar, S., Clark, C.L., Schwartz, R.A., and Spector, D.H. (2004). Cyclin-dependent kinase activity is required at early times for accurate processing and accumulation of the human cytomegalovirus UL122-123 and UL37 immediate-early transcripts and at later times for virus production. J Virol *78*, 11219-11232.

Sharp, P.A. (2005). The discovery of split genes and RNA splicing. Trends in biochemical sciences *30*, 279-281.

Staras, S.A., Dollard, S.C., Radford, K.W., Flanders, W.D., Pass, R.F., and Cannon, M.J. (2006). Seroprevalence of cytomegalovirus infection in the United States, 1988-1994. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America *43*, 1143-1151.

Stern-Ginossar, N., Saleh, N., Goldberg, M.D., Prichard, M., Wolf, D.G., and Mandelboim, O. (2009). Analysis of human cytomegalovirus-encoded microRNA activity during infection. J Virol *83*, 10684-10693.

Stern-Ginossar, N., Weisburd, B., Michalski, A., Le, V.T., Hein, M.Y., Huang, S.X., Ma, M., Shen, B., Qian, S.B., Hengel, H., Mann, M., Ingolia, N.T., and Weissman, J.S. (2012). Decoding human cytomegalovirus. Science *338*, 1088-1093.

Takagaki, Y., Seipelt, R.L., Peterson, M.L., and Manley, J.L. (1996). The polyadenylation factor CstF-64 regulates alternative processing of IgM heavy chain pre-mRNA during B cell differentiation. Cell *87*, 941-952.

Tanaka, S., Furukawa, T., and Plotkin, S.A. (1975). Human cytomegalovirus stimulates host cell RNA synthesis. J Virol *15*, 297-304.

Tian, B., Hu, J., Zhang, H., and Lutz, C.S. (2005). A large-scale analysis of mRNA polyadenylation of human and mouse genes. Nucleic Acids Res *33*, 201-212.

Ule, J., Jensen, K.B., Ruggiu, M., Mele, A., Ule, A., and Darnell, R.B. (2003). CLIP identifies Nova-regulated RNA networks in the brain. Science *302*, 1212-1215.

Zisoulis, D.G., Lovci, M.T., Wilbert, M.L., Hutt, K.R., Liang, T.Y., Pasquinelli, A.E., and Yeo, G.W. (2010). Comprehensive discovery of endogenous Argonaute binding sites in Caenorhabditis elegans. Nat Struct Mol Biol *17*, 173-179.

Chapter 2: High-resolution profiling and analysis of viral and host small RNAs during human cytomegalovirus infection

2.1. Abstract

Human cytomegalovirus (HCMV) contributes its own set of microRNAs (miRNAs) during lytic infection of cells, likely fine-tuning conditions important for viral replication. To enhance our understanding of this component of the HCMV-host transcriptome, we have conducted deep-sequencing analysis of small RNAs (smRNA-seq) from infected human fibroblast cells. We found that HCMV-encoded miRNAs accumulate to ~20% of the total smRNA population at late stages of infection, and our analysis led to improvements in viral miRNA annotations and identification of two novel HCMV miRNAs, miR-US22 and miR-US33as. Both of these miRNAs were capable of functionally repressing synthetic targets in transient transfection experiments. Additionally, through cross-linking and immunoprecipitation (CLIP) of Argonaute (Ago)-bound RNAs from infected cells, followed by high-throughput sequencing, we have obtained direct evidence for incorporation of all HCMV miRNAs into the endogenous host silencing machinery. Surprisingly, three HCMV miRNA precursors exhibited differential incorporation of their mature miRNA arms between Ago2 and Ago1 complexes. Host miRNA abundances were also affected by HCMV infection, with significant upregulation observed for an miRNA cluster containing miR-96, miR-182, and miR-183. In addition to miRNAs, we also identified novel forms of virus-derived smRNAs, revealing greater complexity within the smRNA population during HCMV infection.

2.2. Introduction

Human cytomegalovirus (HCMV), a member of the herpesvirus family, is prevalent in the majority of the global population and the main viral cause of birth defects (Mocarski et al., 2007). HCMV is capable of infecting a wide variety of human cell types *in vivo* and is known to establish a latent form of infection that persists throughout the life of the host. Immunocompromised patients are especially susceptible to problems associated with infection, further motivating HCMV studies.

Expression of virus-encoded microRNAs (miRNAs) likely contributes to the molecular transformation necessary for productive infection in human cells. Since the initial discovery of viral miRNAs in 2004 (Pfeffer et al., 2004), several viruses have been documented to express their own sets of miRNAs (Cai et al., 2005; Cai et al., 2006; Cui et al., 2006; Dunn et al., 2005; Grey et al., 2005; Pfeffer et al., 2005). These small RNAs (smRNAs) are highly attractive from a viral standpoint in that they require minimal space within the genome and potentially offer a robust mechanism for specific targeting of host defense genes. HCMV is known to encode at least eleven miRNA precursors that are expressed and processed into mature miRNAs during infection (Grey et al., 2005).

High-throughput sequencing approaches have been instrumental in recent viral miRNA studies, acquiring a level of resolution that was previously unachievable through cloning efforts (Meyer et al., 2011; Riley et al., 2010; Umbach and Cullen, 2010). Deep sequencing approaches have not been extended to smRNA analysis of HCMV-infected cells and previous microarray-based studies were limited in only assessing host miRNA levels (Santhakumar et al., 2010; Wang et al., 2008). Thus, we set out to conduct a more comprehensive analysis by deep sequencing all smRNAs (smRNA-seq) from HCMV Towne-infected human foreskin fibroblast (HFF) cells. We have fully characterized both the viral and human miRNA fractions, leading us to identify novel HCMV miRNA

precursors and differentially regulated human miRNAs, including the miR-182/96/183 cluster, which was not previously reported. Additionally, the unbiased sequencing refined the 5′ ends of three HCMV mature miRNAs, with miR-US4 differing the most from its annotated boundaries. To further characterize the functional attributes of smRNAs, we performed high-throughput sequencing of Argonaute (Ago)-associated RNAs from infected HFFs to obtain direct evidence that HCMV-encoded miRNAs are incorporated into endogenous cellular silencing complexes. Finally, we also observed production of non-miRNA smRNAs from discrete regions of the HCMV genome during infection, including areas that encode long non-coding RNA transcripts. Our results revealed a greater complexity within the smRNA component of HCMV-host interactions than was previously appreciated.

2.3. Results

2.3.1. Global statistics of short cDNA libraries in HCMV-infected cells

We prepared short cDNA libraries corresponding to 18-35 nt RNAs from infected and mock-treated HFFs at 24 and 72 hours post-infection (hpi). High-throughput sequencing of the four smRNA libraries using the Illumina Genome Analyzer IIx platform generated 2.5 to 4.5 million 36 nt reads per condition. An average of 84% of smRNA-seq reads were mapped uniquely to the human and viral genomes, allowing for two base mismatches. Greater than 95% of mapped reads were classified into different categories of annotated non-coding RNAs (ncRNAs). As expected, the majority of smRNA-seq reads mapped to human and virus-encoded miRNAs, as well as snoRNAs, another class of ncRNAs (Figure 2.1A).

To validate that the cDNA libraries accurately recapitulated virus-encoded miRNA expression, the number of reads corresponding to HCMV miRNAs, as a fraction

of all smRNA-seq reads, was measured at the two time points. HCMV miRNAs increased in overall abundance more than 7-fold over the course of infection, accounting for over 30% of the total (human and HCMV) miRNAs sequenced at 72 hpi. Consistent with previous observations regarding cytomegalovirus miRNA expression kinetics (Dolken et al., 2007; Grey et al., 2005), all HCMV miRNAs were found to individually accumulate from 24 to 72 hpi (Figure 2.1B). In fact, at 72 hpi particular HCMV miRNAs reached levels comparable to the most highly expressed human miRNAs in HFFs, including miR-21 and let-7f.

Of the eleven annotated HCMV precursor miRNAs (pre-miRs), ten were well supported by the smRNA-seq data at both time points, with 416 reads corresponding to mature miR-US4-3p and 184,877 reads corresponding to miR-US5-2-3p at 72 hpi, the lowest and highest expressed viral miRNAs at this time point ("5p" and "3p" follow the miRNA name to designate arms processed from a common pre-miR hairpin structure). The annotated HCMV pre-miR-UL70 was not supported by smRNA-seq. Except for miR-UL148 and miR-US5-1, all HCMV miRNA names have been assigned a 5p or 3p designation in this study to avoid confusion with miR/miR* nomenclature. Despite the substantial presence of viral miRNAs in the infected samples, miRNAs as a fraction of the total smRNA pool decreased upon infection, while the proportion of snoRNAs relative to other smRNAs increased, as has been observed to some extent in previous studies (Pfeffer et al., 2004). Nevertheless, a similar number of individual human miRNAs were detected across all conditions (Figure 2.1A).

2.3.2. Identification of novel viral miRNAs

Since smRNA-seq reads were also detected outside annotated HCMV miRNA regions, we subjected the datasets to the MIResque miRNA prediction algorithm (G.W. Yeo., unpublished) to search for novel HCMV-encoded miRNAs. Two candidate pre-

miRs were further inspected, having satisfied established guidelines for miRNA annotation (Ambros et al., 2003) and rules for Drosha processing of primary miRNA transcripts (Han et al., 2006; Zeng et al., 2005). One candidate region, antisense to the protein-coding gene US22, and thus named pre-miR-US22, was found to contain a significant level of smRNA-seq coverage (3,410 reads for miR-US22-5p at 72 hpi) and featured a distribution of sequence alignments that mapped to a predicted stem-loop hairpin structure characteristic of a pre-miR (Figure 2.2A). As an independent method of validation, miR-US22-5p was detected through Northern blot analysis by 48 hpi (Figure 2.2C).

Substantial coverage by smRNA-seq was also detected antisense to the region encoding HCMV pre-miR-US33 (Figure 2.2B). This candidate, pre-miR-US33as, is sense to the US29 locus and again harbored a distribution of reads characteristic of a pre-miR stem-loop. While the existence of this miRNA was suggested before through computational prediction efforts and termed ‘miR-US29-1’ (Grey et al., 2005), no further experimental evidence was previously obtained. MiR-US33as-5p was also validated by Northern analysis and could be detected by 24 hpi (Figure 2.2D).

2.3.3. Characterization of miR-US22 and miR-US33as outside infection

Studies from Grey et al. have demonstrated that HCMV miRNAs such as miR-UL112 can be transiently expressed in non-infected cells (Grey et al., 2007). Placement of the miR-UL112 precursor and flanking viral genomic context within an expression construct allowed for efficient processing in the absence of viral factors, which was evident by Northern analysis of transfected cells. We followed a similar approach for miR-US22-5p, yet the precursor context of this miRNA did not permit accumulation in transfected 293Ts beyond the moderate levels present in HFFs at 72 hpi (Figure 2.2E). Nevertheless, significant repression of a fully complementary target site was observed

within the context of a luciferase assay, but miR-US22 was not able to regulate a target site bearing its 5p miRNA seed match in our system (Figure 2.2E). In the case of miR-US33as-5p, higher levels of this miRNA could be obtained in transfected 293Ts and repression was observed for both fully complementary and seed match target sites (Figure 2.2F). For comparison purposes, we have also performed these experiments for the highly expressed miR-US25-1 and obtained similar results for its 5p mature form (data not shown).

2.3.4. Refinement of HCMV miRNA annotations

MiRNA-mRNA interactions occur primarily through partial sequence complementarity between the miRNA seed region (nt 2-7 from the 5′ end) and a targeted mRNA. Consequently, accurate definition of the 5′ termini of miRNA sequences is important in the elucidation of HCMV miRNA functions. We determined the most abundant sequences associated with each miRNA at both 24 and 72 hpi, and, based on analysis of the 72 hpi dataset, 7 of the 17 existing HCMV mature miRNA annotations can be revised (Table 2.3). Notably, the sequence obtained for miR-US4-5p differs from the annotated version by 5 bp at the 5′ end, complicating the interpretation of the seed-mediated targeting described in the recent work of Kim and colleagues (Kim et al., 2011). In fact, no reads obtained from our smRNA-seq experiments matched the annotated miR-US4 mature sequence. The miR-US4 precursor and upstream genomic sequence are completely conserved between Towne, AD169 and Merlin strains of HCMV. Interestingly, five miRNAs (miR-US25-1-5p, miR-US33-5p, miR-US5-2-3p, miR-US22-5p, and miR-US33as-3p) differed in their most abundant form between 24 and 72 hpi (Table 2.3). Globally, HCMV miRNAs exhibited heterogeneous 5′ and 3′ ends in the smRNA-seq data, as shown in the depiction of read alignments across pre-miR-US22

and pre-miR-US33as/miR-US33 in Figure 2.2A and 2.2B. We have also observed comparable levels of end heterogeneity for human miRNAs.

In the process of defining mature arms within each miRNA precursor, we only found sufficient evidence for a single mature miRNA expressed from miR-UL148D and miR-US5-1, while all other pre-miRs generated two individual mature miRNA species (Table 2.3). In total, 22 mature miRNAs were identified during HCMV infection, originating from 12 HCMV miRNA precursors, including the novel miR-US22 and miR-US33as (Figure 2.1B).

2.3.5. Association of HCMV miRNAs with Argonaute proteins

Recent reports have identified that miRNAs encoded by other herpesviruses associate with host silencing complexes (Dolken et al., 2010; Riley et al., 2010). To confirm whether HCMV miRNAs are indeed incorporated into these complexes, we performed CLIP-seq (crosslinking and immunoprecipitation followed with high-throughput sequencing) experiments against Ago proteins for mock and infected HFFs at 24 and 72 hpi (Figure 2.3A). This technique has been shown to offer insight into Ago-mediated miRNA-mRNA interactions (Chi et al., 2009; Zisoulis et al., 2010). We determined that all HCMV miRNAs detected by smRNA-seq, including the novel miRNAs described above, could be found in CLIP-seq libraries generated from infected cells, using antibodies against either Ago1 or Ago2. Therefore, we conclude that all HCMV miRNAs are capable of functional activity given that they associate with endogenous Ago proteins during productive infection.

From analysis of the Ago1 and Ago2 CLIP-seq datasets, Ago-associated miRNA expression values (each miRNA abundance measured as a fraction of the total miRNA reads) correlated with abundance levels derived from smRNA-seq. Additionally, positive correlation was observed between Ago2- and Ago1-associated miRNA abundance

levels in mock and infected conditions alike (correlation coefficients for pairwise comparisons ranged from 0.77 to 0.81). To further investigate if this correlation was similarly true for both the 5p and 3p mature arms of pre-miRNAs, we used 5p over 3p enrichment as a metric to compare Ago2 versus Ago1 incorporation. At a p-value threshold of 6×10^{-5}, three HCMV miRNAs (miR-US4, miR-US33as, and miR-UL112) were found to exhibit significantly different miRNA arm incorporation at one or both time points of infection (Figure 2.3B), while others did not demonstrate this characteristic (three examples are shown in Figure 2.3C). Using the same significance threshold, only five human pre-miRNAs of the 150 expressed during infection (considering only those annotated to express two mature miRNAs) were found to display variable incorporation of their 5p and 3p arms into Ago2 and Ago1 complexes (Figure 2.3D). These included hsa-miR-214, hsa-miR-31, hsa-let-7i, hsa-miR-106b, and hsa-miR-126.

2.3.6. Altered levels of specific human miRNAs during HCMV infection

In order to explore the impact of viral infection on the host cell smRNA profile, we focused on identifying human miRNAs that were differentially expressed by comparing mock to virus-infected cells at both 24 and 72 hpi (Figure 2.4A). To avoid detecting differences at the level of transcriptional noise, we required a minimum of 50 reads in at least one of the conditions to consider the miRNA detected; 318 mature miRNAs remained for further analysis. Within each dataset, individual miRNA read counts were normalized using the total number of human and viral miRNA reads found in the library, and significantly changing miRNAs were identified as described in Materials and Methods. The large abundance of viral miRNAs at 72 hpi had the potential to complicate the statistics, however the results did not change when viral miRNA counts were removed in the normalization process. From the perspective of individual miRNA fold changes upon infection, our results largely supported previous studies that utilized

microarray technology to detect human miRNAs (Santhakumar et al., 2010; Wang et al., 2008).

Intriguingly, among the miRNAs that were most highly upregulated at both stages of infection, miR-96, miR-182 and miR-183 stood out in that they belong to the same miRNA genomic cluster (Figure 2.4B). Also of interest, these three miRNAs share nearly the same canonical seed sequence. Since this miRNA cluster was not identified as upregulated in previous studies, we also verified increased expression of miR-182 in HFFs infected with AD169, another laboratory-adapted strain of HCMV (Figure 2.4C).

2.3.7. Non-miRNA regions of concentrated HCMV smRNA production

While the majority of smRNA-seq reads aligned to human and viral smRNAs, ~3% of the mapped reads from each dataset remained unclassified (Figure 2.1A). From the infected datasets we found that this uncategorized set of smRNAs was concentrated in particular regions (Figure 2.5A and 2.5B), with the greatest coverage observed across the HCMV long non-coding RNA (lncRNA) 2.7, a large region covering UL61-UL68 (transcribed sense to UL63 and UL65), and the lncRNA5 (located between UL105 and UL111). Upon individual inspection, these areas differed from miRNA-encoding regions with respect to their smRNA-seq read distributions and predicted secondary structures.

In an assessment of the transcriptional activity within these regions, we have used preliminary data from mRNA deep sequencing (RNA-seq) experiments, which we have conducted to extend our studies beyond the smRNA fraction of the virus and host transcriptomes. The lncRNA2.7 (Figure 2.5C) has been previously described as one of the most active transcriptional units in the HCMV genome (Spector, 1996), and our RNA-seq data showed that the UL61-UL68 region was also highly transcribed during infection (Figure 2.5D). While this correlation with enrichment for non-miRNA smRNAs might suggest that these sequences were simply degradation products, other highly

transcribed regions of the viral genome were not correspondingly enriched for production of such smRNA species. Accordingly, at 24 hpi the lncRNA2.7 was the most highly transcribed region of the HCMV genome by far, but the lesser-transcribed UL61-UL68 region was more enriched for smRNA-seq alignments.

2.4. Discussion

Among the complex network of virus-host interactions present within HCMV-infected human cells, viral miRNAs expand this regulatory framework, capable of modulating both host and virus gene expression. In our smRNA-focused characterization of lytically infected HFFs we found that a total of 22 HCMV-encoded mature miRNAs are expressed, including two novel precursors, miR-US22 and miR-US33as. Characterization of these miRNAs in transfection experiments revealed that both can be processed by the host cellular machinery and are able to repress synthetic targets. Since miR-US22 lies antisense to the US22 coding region, we predict that it may directly regulate US22 transcript levels, as was previously shown to be the case for the Epstein-Barr virus (EBV) miR-BART2, which is transcribed antisense to the viral DNA polymerase BALF5 (Barth et al., 2008). Similarly, the location of the miR-US33as precursor within the putative US29 coding region has interesting implications given that Drosha processing would likely negatively impact US29 expression levels.

The set of miRNAs described in this study are conserved across different laboratory and clinical isolate strains of HCMV, despite instances of sequence variability across other genomic loci (Dolken et al., 2009; Jung et al., 2011). Our smRNA-seq data did not support expression of HCMV miR-UL70, despite the recent quantitative PCR experiments by Stern-Ginossar et al. using two clinical strains (Stern-Ginossar et al., 2009). Consistent with our smRNA-seq, Pfeffer et al. also failed to detect this particular

miRNA in their cloning efforts (Pfeffer et al., 2005). In the initial Northern characterization of miR-UL70, Grey et al. indicated that the probe used for detection sometimes hybridized to a ~22 nt species in non-infected samples, questioning its true validity as an HCMV-encoded miRNA (Grey et al., 2005).

Nevertheless, we were able to refine annotations for the other HCMV miRNAs, clarifying that all but two HCMV pre-miRNAs express two individual mature miRNAs. Except in the case of miR-US4, any discrepancies between the results from our smRNA-seq and current annotations involved only a single base difference at either the 5′ or 3′ mature HCMV miRNA boundaries. However, as stated above, sequences obtained for miR-US4-5p were shifted by 5 bp at the 5′ end from previous predictions (Grey et al., 2005). The recent studies performed by Kim et al. involved overexpression of the annotated sequence for miR-US4-5p and featured a mutant form of miR-US4-5p in which three point mutations were made within the first 7 nt of its annotated sequence (Kim et al., 2011). This third final change reaches the seed region of the form detected by smRNA-seq, which may have fortuitously reduced miRNA-mRNA targeting. While the precursor sequence of miR-US4 and more than 50 bp of upstream sequence are conserved between Towne, AD169, and several other strains of HCMV, we cannot yet rule out whether strain-specific issues account for the observed difference in mature miRNA sequence.

While it has been suggested that mammalian miRNAs associate at similar frequencies with different Ago proteins in the cell (Burroughs et al., 2011), we found, surprisingly, that three of the HCMV miRNA precursors possessed differential incorporation of their individual mature arms into Ago1 and Ago2 complexes. The differential incorporation of the 5p and 3p forms of miR-UL112 at 24 hpi is important to consider given that studies focused on identification of miR-UL112 targets have primarily

followed the 3p form (Grey et al., 2007; Stern-Ginossar et al., 2007). Furthermore, only the 3p form of HCMV miR-UL112 is listed within current miRNA annotations.

Aside from virus-encoded miRNAs, our smRNA-seq efforts also revealed human miRNAs that responded significantly in expression level during HCMV infection. In this study, we identified a genomic cluster of human miRNAs that was highly upregulated at both 24 and 72 hpi. Consisting of miR-96, miR-182, and miR-183, this miRNA cluster has been suggested to arise from processing of a single independent primary transcript (Weston et al., 2006; Xu et al., 2007). The upregulation of this cluster in multiple forms of cancer and tumorgenesis (Sarver et al., 2009; Schaefer et al., 2010) is of particular interest in the debate surrounding the potential association of HCMV infection and higher risks of tumor formation (Michaelis et al., 2011). Furthermore, studies using mice have implicated a potential role for the murine orthologs of these miRNAs in the development of the inner ear (Sacheli et al., 2009).

Our study represents the first unbiased genome-wide smRNA sequencing of HCMV-infected cells, and enabled us to identify a novel class of HCMV non-miRNA smRNAs. Interestingly, the UL61-UL68 region, where a substantial number of these smRNAs were produced, overlaps with the virus origin of lytic replication. Two previously characterized short RNA species, vRNA-1 and vRNA-2 (Prichard et al., 1998), also fall within this region (locations shown in Figure 2.5D). Further studies would have to be conducted to determine if any functional relationship exists between the smRNAs we have detected and these short RNAs. While the non-miRNA group of smRNAs did not appear to collectively increase in abundance in the way that HCMV-encoded miRNAs accumulate during infection, these elements may also harbor similar regulatory potential. Recent deep sequencing analysis of the smRNA content of host models infected with different RNA viruses has illustrated that non-miRNA viral smRNAs are found in a wide

variety of virus-infected systems (Parameswaran et al., 2010). It has been suggested that this class of smRNAs may elicit general host RNA interference responses, similar to the antiviral mechanisms known to exist in plants and lower eukaryotic organisms.

In summary, through a battery of genomic analyses, we report the following findings. First, we have improved viral miRNA annotations and identified new HCMV-encoded miRNAs, bringing the total number of known HCMV miRNA precursors to 12. Beyond computational evidence and Northern analysis of the novel HCMV miR-US22 and miR-US33as, we have further characterized and validated the processing of these miRNA precursors in transiently transfected cells and demonstrated that their mature forms are capable of target repression. In addition to showing that HCMV miRNAs account for a fifth of the total smRNA population at late stages of infection, we used CLIP-seq technology to demonstrate that all of these miRNAs associate directly with endogenous human Ago proteins. This analysis yielded the surprising result that mature miRNAs from three HCMV pre-miRs are differentially incorporated into Ago1 and Ago2 complexes. With regard to human miRNA levels, we observed significant upregulation of the entire host miR-182/96/183 cluster, previously reported to have roles in cancer and implications in hearing loss. Last, we identified novel virus-derived smRNAs emanating from coding and non-coding regions of the HCMV genome, especially across long non-coding RNAs, illustrating that a complex level of regulatory activity likely exists within the viral and host smRNA fractions of the transcriptome.

2.5. Materials and Methods

2.5.1. HCMV infections and culture conditions

Primary human foreskin fibroblasts (HFFs) were cultured as previously described (Sanders et al., 2008) and maintained within passages 16 to 20. At the time of infection,

G0-synchronized cells were trypsinized, re-suspended in the presence of virus, and plated at near confluence. Unless otherwise stated, all infections were performed with the wild-type Towne strain at an MOI of 3 PFU per cell. For mock infections, tissue culture supernatant (spent media from 3 days of culture) containing 1% DMSO was used in place of virus stock. At different times post-infection, cells were washed with 1 × PBS before harvesting. 293T cells (ATCC) were cultured in high-glucose Dulbecco's modified Eagle's medium (DMEM), supplemented with 10% heat-inactivated fetal bovine serum (FBS), 100 units/ml penicillin, 100 µg/ml streptomycin, and 2 mM L-Glutamine (Invitrogen).

2.5.2. Library preparation for smRNA-seq and CLIP-seq experiments

For smRNA-seq, total RNA from HCMV-infected and mock-treated HFFs was isolated using Trizol reagent (Invitrogen) and subsequently treated with Turbo DNase (Ambion), all according to instructions from the manufacturers. Libraries were then generated from 18-35 nt cDNAs with the Illumina Small RNA Sample Prep v1.5 kit, using 5 µg of DNase-treated total RNA. CLIP-seq libraries were prepared as described (Yeo et al., 2009; Zisoulis et al., 2010). Four 15 cm dishes (approximately 2 × 107 cells) of either infected or mock-treated cells were used for each condition. Cells were subjected to 400 mJ of UV irradiation, in 1 × PBS on ice, using a Stratalinker machine (Stratagene, Agilent Technologies). Following crosslinking, cells were manually scraped off the dishes, pelleted and snap-frozen. Cell pellets were only minimally stored at -80°C prior to proceeding with RIPA-mediated lysis and the remainder of the protocol. Wako anti-Ago1 (2A7) and Millipore anti-Ago2 (07-590) antibodies were used for immunoprecipitation steps. Illumina GAIIx sequencing for all smRNA-seq and CLIP-seq libraries was performed with 36 cycles.

2.5.3. Northern analysis of miRNAs

Total RNA was isolated using Trizol reagent (Invitrogen) according to instructions from the manufacturer up until the addition of isopropanol. Samples were precipitated overnight at -80°C and centrifuged the following day at 14,000 rpm for 30 min at 4°C, followed by one ethanol wash and a final 15 min spin. RNA isolated in this manner was subjected to polyacrylamide Northern blot analysis (10 µg per lane) as described (Pasquinelli et al., 2003). Modifications to the protocol included gel transfer to GeneScreen Plus membranes (PerkinElmer) and additional post-hybridization washes. StarFire probes were used for miRNA detection (Integrated DNA Technologies), and hybridizations were performed for 16 h at 50°C in 50 ml conical tubes. After two 30 min washes in non-stringent buffer (3 × SSC, 5% SDS, 0.025 M sodium phosphate, 10 × Denhardt's solution) at 50°C, membranes were moved to new conical tubes and washed twice for 30 min in 2 × SSC, 1% SDS at room temperature. Membranes were washed briefly once more in a tray containing 2 × SSC, 1% SDS at 42° and subjected to one final wash in 1 × SSC, 1% SDS, rotating in a conical tube at 42°C for 30 min. Exposure to film was performed for 5 h to overnight between two intensifying screens at -80°C. All membranes were re-probed with 5′ end-labeled traditional oligonucleotide for the U6 snRNA to determine loading efficiencies. Probe sequences used are listed in Table 2.1.

2.5.4. MiRNA transfection experiments and luciferase assays

24 h prior to miRNA transfection 293T cells were re-plated in media containing 5% FBS and 2 mM L-Glutamine (no antibiotics) at a density of 5 × 106 cells per 6 well dish. HCMV miRNA precursors, with approximately 100 bases of flanking viral genomic context, were PCR-amplified from HCMV Towne genomic DNA and cloned into the XhoI and BamHI restriction sites of pcDNA3.1 (Invitrogen). 200 ng of plasmid and 5 µl of Lipofectamine 2000 were diluted in Opti-MEM (Invitrogen) and added to cells according

to the manufacturer's instructions. No media changes were performed and total RNA was harvested for Northern blot analysis after 48 h. For luciferase activity measurements, 293Ts were seeded as above, at a density of 4 × 106 cells per 24 well dish. Each well was transfected with 2 μl of Lipofectamine 2000, 750 ng of miRNA expression plasmid, and 50 ng of either unmodified psiCHECK2 vector (Promega) or psiCHECK2 containing a synthetic target site. Target sites were generated by annealing complementary oligos that were compatible with the XhoI and NotI restriction sites located downstream of Renilla luciferase. Activity measurements were obtained 36 h post-transfection using the Dual-Luciferase Assay System from Promega. This assay was performed according to the manufacturer's instructions except cells were lysed in 200 μl of passive lysis buffer and further diluted 1:20 prior to acquiring measurements. All primers used for generating the constructs described above are listed in Table 2.2.

2.5.5. Computational processing of smRNA-seq and CLIP-seq data

First, 3′ adapter sequences were trimmed from sequencing reads. For barcoded libraries, which included the smRNA-seq and Ago2 CLIP-seq datasets, only reads harboring fully intact 5′ barcode sequences were considered in the analysis. Filtered datasets were then mapped against an index composed of both the human (UCSC hg19, http://genome.ucsc.edu) and HCMV Towne (Genbank FJ616285.1) genomes, using Bowtie software (version 0.12.7, with parameters -k 1 -m 10 -l 25 --best). SmRNA annotations were obtained from miRBase 16.0 (http://www.mirbase.org) and the UCSC Table Browser. For rRNAs, hg18-based coordinates were converted to hg19 using the UCSC LiftOver utility (http://genome.ucsc.edu/cgi-bin/hgLiftOver). Custom Perl scripts were used to overlap mapped smRNA-seq coordinates with these annotations.

Datasets from the smRNA-seq and CLIP-seq experiments have been deposited in the GEO database under accession number GSE33584.

2.5.6. MiRNA expression analysis

Abundance levels for human and HCMV miRNAs were determined by summing the number of smRNA-seq reads that pattern-matched to mature miRNA sequences or mapped to mature miRNA coordinates (allowing up to 2 mismatches). Normalized values were obtained by dividing each miRNA count by the total number of human and viral miRNA reads in the dataset and then multiplying the fraction by 105. Ago-CLIP miRNA levels were calculated using the same metric. To identify human miRNAs that significantly changed in abundance during infection, HCMV miRNA counts were removed from the datasets and infected libraries were compared to mock libraries from the same time point. Normalized human miRNA counts were binned by expression level into two groups. Within the low and high expression groups, Z-scores were calculated for each miRNA and values obtained that were less than -1.96 or greater than 1.96 were deemed statistically significant (Polymenidou et al., 2011). MiRNA Z-scores were required to meet this threshold in both infected vs. mock or mock vs. infected comparisons.

2.5.7. RNA-seq library preparation and HCMV transcriptome analysis

Strand-specific RNA-seq libraries were prepared as described (Parkhomchuk et al., 2009). 10 μg of total RNA, isolated with Trizol (Invitrogen) according to the manufacturer's instructions, was treated with Turbo DNase (Ambion) prior to library preparation. Illumina sequencing was performed with 76 cycles. As above, sequencing reads were adapter trimmed and mapped to the human and HCMV genomes using Bowtie (parameters -m 25 -k 1 -l 28 --best). HCMV-mapped RNA-seq reads were visualized using R software (version 2.10.1).

2.6. Publication acknowledgement

Chapter 2, in full, has been published as a manuscript entitled “High-Resolution Profiling and Analysis of Viral and Host Small RNAs during Human Cytomegalovirus Infection” in *Journal of Virology* (Stark et al., 2012). The author of this dissertation was the primary investigator and author of this paper.

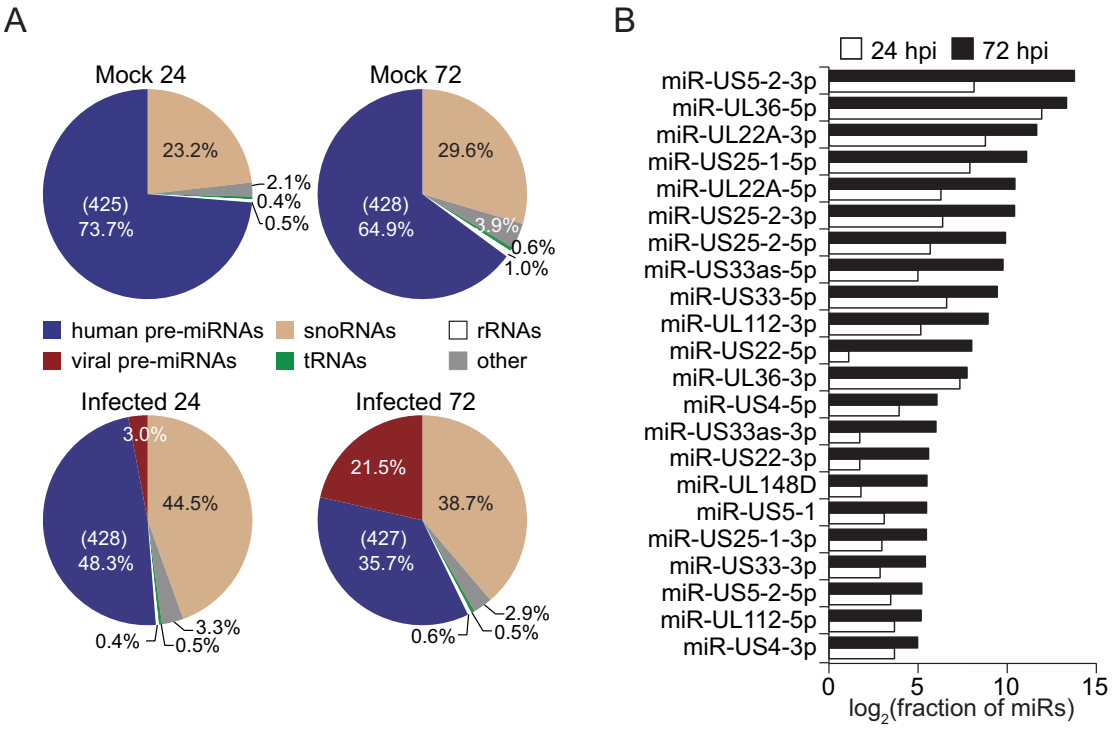

Figure 2.1: Genomic analysis of smRNAs during HCMV infection of HFFs

(A) Distribution of smRNA-seq reads across categories of annotated smRNAs (pre-miRNAs, snoRNAs, tRNAs and rRNAs) in HCMV-infected (Towne strain) or mock-treated HFFs at 24 and 72 hpi. Reads that did not match current annotations were classified as “other”. For human miRNA fractions, values listed in parentheses indicate the total number of unique pre-miRNAs detected. (B) Expression levels of HCMV mature miRNAs during infection, measured by smRNA-seq; white bars represent 24 hpi and black bars correspond to 72 hpi. Total read counts for each viral miRNA, including reads that exactly matched the mature miRNA sequence and those that were uniquely mapped to its genomic coordinates, were normalized by the total number of miRNA reads (human and viral) in the dataset.

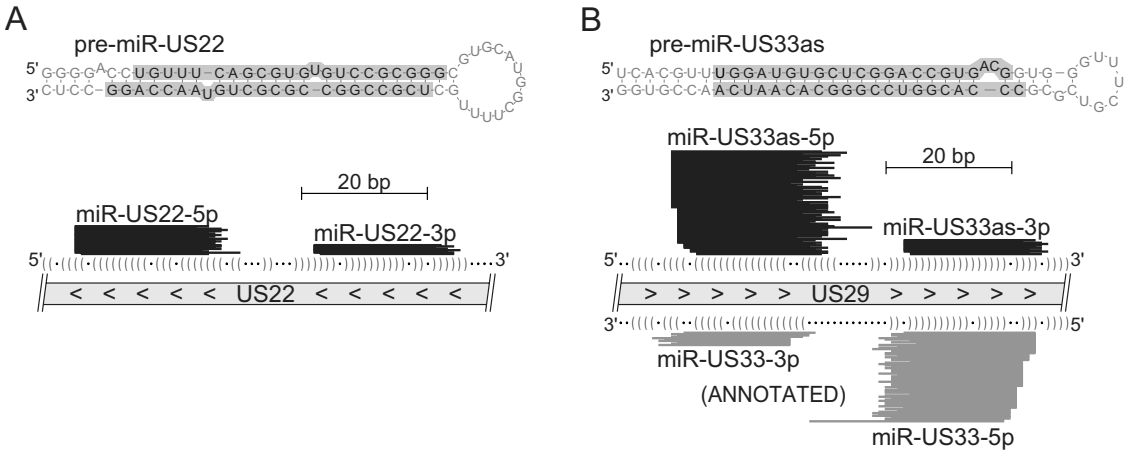

Figure 2.2: Identification of novel HCMV miRNA precursors miR-US22 and miR-US33as

(A,B) Uniquely mapped smRNA-seq reads align to predicted RNA stem-loop structures characteristic of a pre-miR hairpin. RNA stem-loop structures of pre-miR-US22 (A) and pre-miR-US33as (B) were generated using mFold software (http://mfold.rna.albany.edu). Highlighted arms indicate where the majority of smRNA-seq reads aligned at 72 hpi (percentages listed in Table 2.3). Genomic locations of the pre-miRs are depicted below the stem-loop structures (direction of transcription indicated by arrowheads). Only non-redundant sequences represented by at least five smRNA-seq reads from the 72 hpi dataset are depicted. The predicted RNA structures of the precursors are illustrated by parentheses indicating bases that are paired and dots that correspond to unpaired bases. (C,D) Northern blot analysis of miR-US22-5p (C) and miR-US33as-5p (D) in infected HFFs. Mock-treated cells were harvested at 72 h for this analysis. The U6 snRNA serves as a loading control. (E) Processing of HCMV miR-US22 in 293T cells and repression of a synthetic target. MiR-US22 expression construct or empty vector was transfected into 293Ts and Northern blot analysis was performed to compare miRNA levels in these samples to HFFs at 72 hpi. A luciferase assay was used to demonstrate repression of a perfectly complementary target sequence by miR-US22-5p (p-value < 0.01). 293Ts were transfected with miR-US22 expression vector and a construct containing a luciferase reporter with either no targeting site (control), a target consisting of the perfectly complementary target of miR-US22-5p (full target), or a canonical seed match target corresponding to only positions 2-7 of miR-US22-5p (seed target). Luciferase measurements shown are averages of three independent experiments. Within each experiment two independent reporter plasmid preparations were used to assess each type of target site. Error bars represent standard deviation and significance values were calculated using the unpaired, two-tailed Student's t-test. (F) Analysis of miR-US33as-5p in transfected 293Ts, performed as described above for panel E (p-value < 0.01 for repressed full and seed versions of synthetic targets).

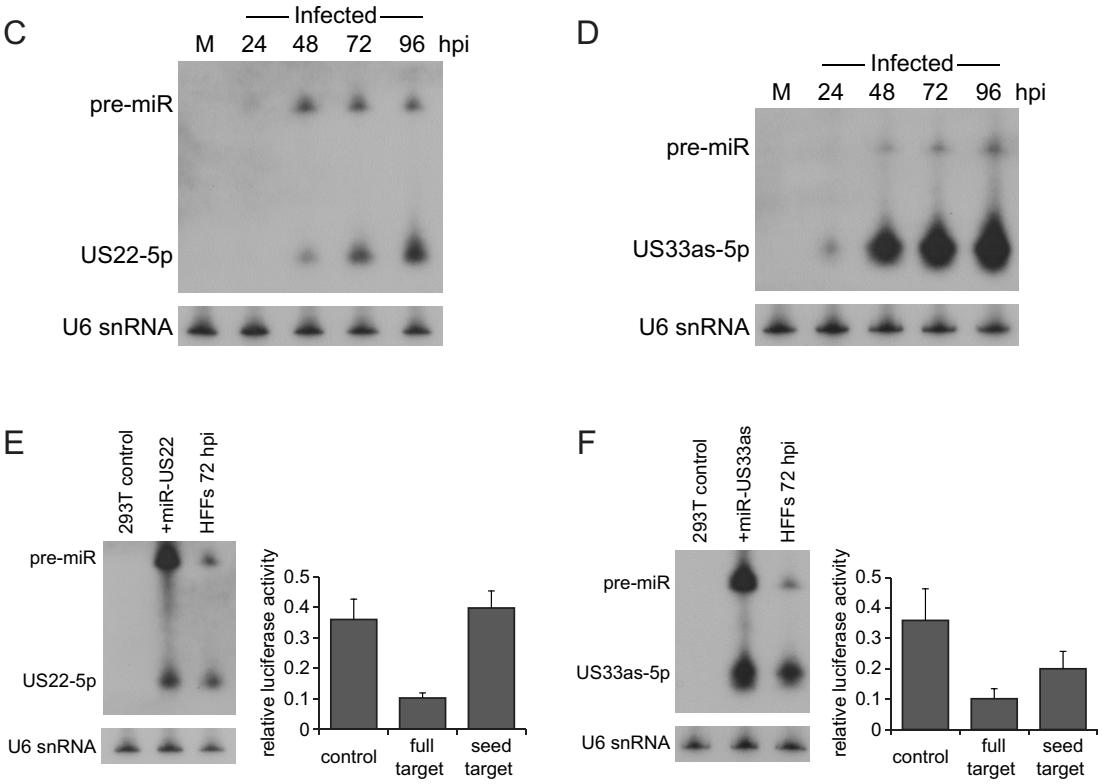

Figure 2.2 (continued): Identification of novel HCMV miRNA precursors miR-US22 and miR-US33as

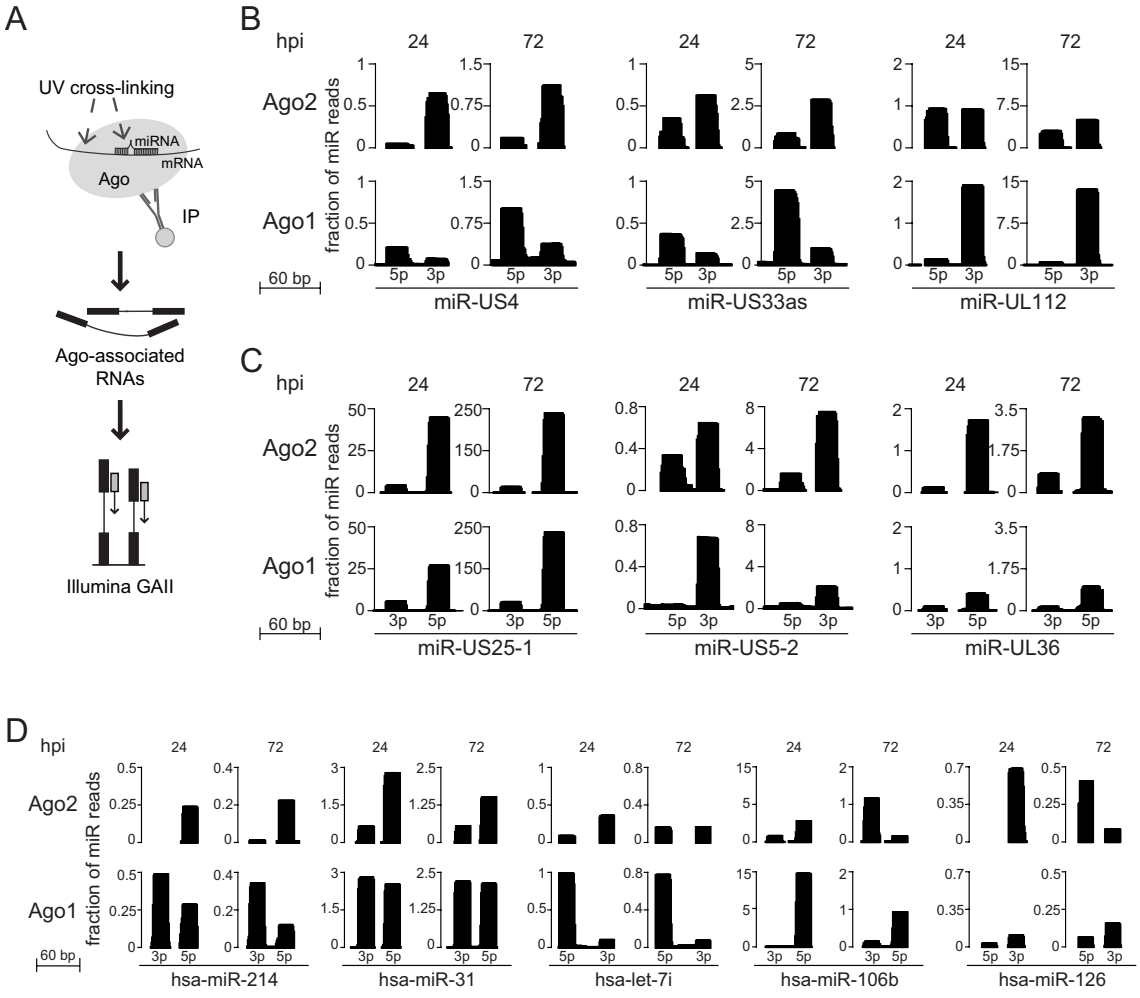

Figure 2.3: Functional association of HCMV miRNAs with endogenous Argonaute proteins

(A) Experimental strategy for generation of Ago CLIP-seq libraries. Mock and infected cells were UV crosslinked at 24 and 72 hpi. Immunoprecipitated (IP) RNA associated with Ago complexes was further processed and subjected to Illumina sequencing. (B) Differential incorporation of particular HCMV miRNAs between Ago2 and Ago1 complexes, as measured by 5p over 3p mature miRNA arm incorporation between Ago2 and Ago1 complexes. Statistical significance was assessed by chi-square analysis (p-value threshold of 6×10^{-5}). Read count heights at each position represent the total number of overlapping CLIP-seq reads, normalized by the total number of miRNA reads in the dataset. (C) Non-differential efficiencies of miRNA association with Ago2 and Ago1 demonstrated by other HCMV miRNAs. Three representative examples are shown. (D) Human miRNAs that exhibit differential 5p-3p miRNA arm incorporation into Ago2 and Ago1 complexes. CLIP-seq read counts were used from infected datasets (24 and 72 hpi) for this assessment. Similar differential incorporation of these miRNAs could be observed in Ago2-Ago1 comparisons for the mock control datasets.

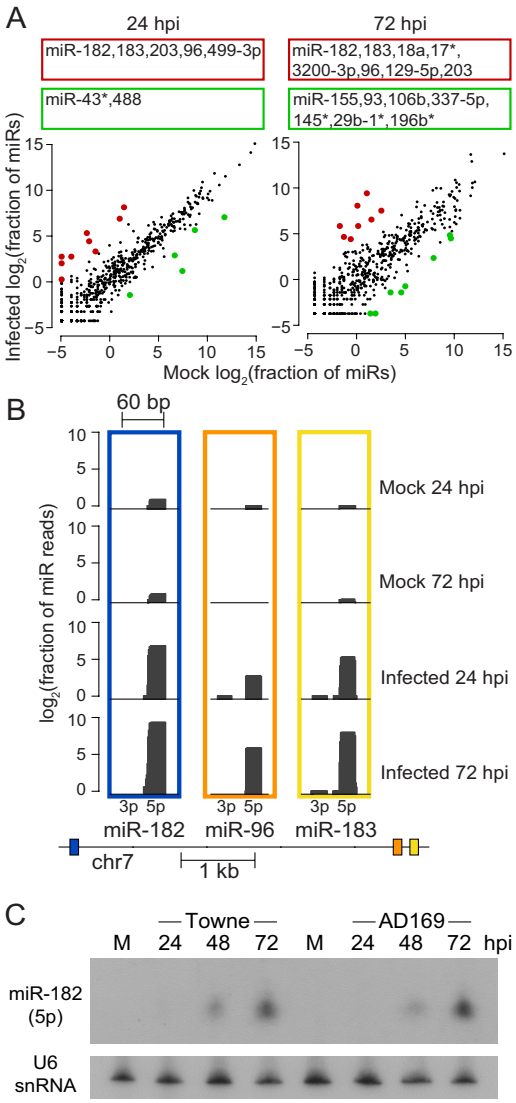

Figure 2.4: Effects of HCMV infection on human miRNA levels in HFFs

(A) Scatter plot comparisons of infected and mock human miRNA expression values at 24 and 72 hpi. Highlighted in red and green are significantly upregulated and downregulated miRNAs during infection, respectively (p-value < 0.05). Listed above the scatter plots are the names of the significantly altered miRNAs (any that changed between the two mock timepoints are not listed) (B) Expression of miR-182/96/183 during infection. The region of chromosome 7 encoding these miRNAs is depicted below the distribution of smRNA-seq reads that aligned to the 182, 96, and 183 miRNA precursors (transcription is from right to left). Read count heights at each position represent $\log_2$-transformed values of the total number of overlapping smRNA-seq reads, normalized by the total number of miRNA reads in the dataset. (C) Northern blot analysis of mature miR-182 at in both HCMV Towne and AD169-infected cells. HFFs were either mock-treated or infected with the Towne strain at an MOI of 3 or AD169 at an MOI of 5. The U6 snRNA serves as a loading control.

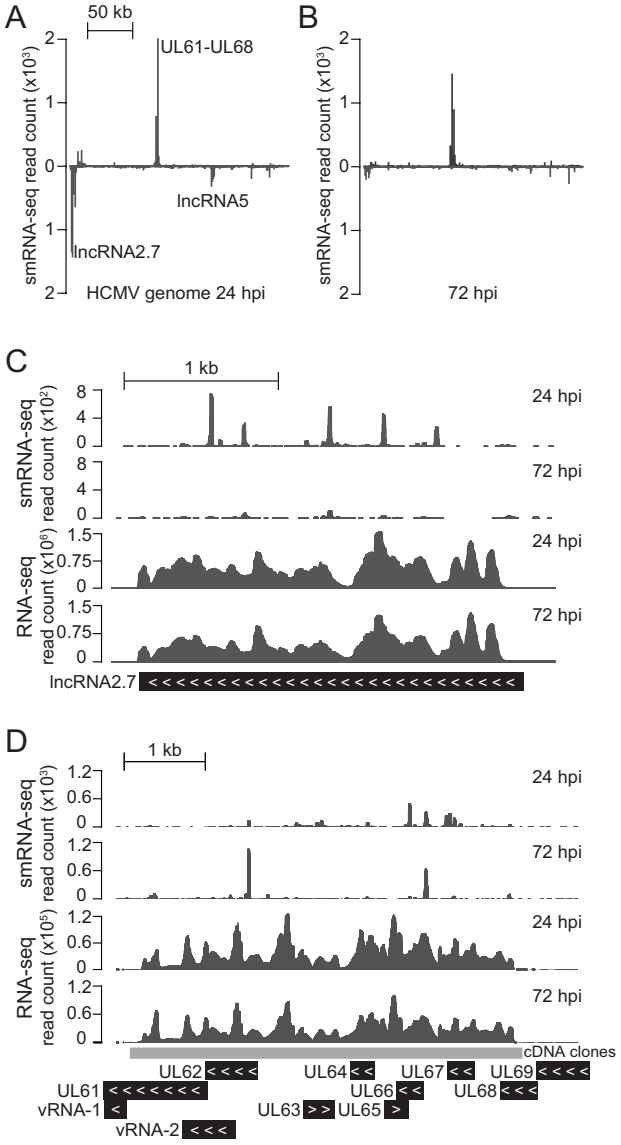

Figure 2.5: Analysis of regions of the HCMV genome enriched for production of non-miRNA smRNAs

(A,B) Distribution of uniquely aligned non-miRNA smRNA-seq reads at 24 and 72 hpi, with peaks representing stacked reads that aligned to the Watson (red) and Crick (blue) strands of the HCMV genome. Reads overlapping annotated and predicted miRNA precursor regions were excluded and the viral genome was interrogated in 1 kb segments for smRNA-seq alignments. (C) SmRNAs that are generated sense to the area encoding lncRNA2.7. (D) Representation of smRNA-seq alignments across a transcriptionally active region that falls within the UL61-UL68 region. Expression data from RNA-seq experiments is shown to illustrate transcription of the regions. Coverage by cDNA clones from Zhang et al. (Zhang et al., 2007) is shown, providing independent support of RNA-seq detected expression of the RNAs.

Table 2.1: Probe sequences used for Northern blot analysis

SmRNA Name	Probe Sequence
HCMV miR-US22-5p	ACGCCCGCGGACACACGCTGAAACA (StarFire)
HCMV miR-US33as-5p	CACCGTCACGGTCCGAGCACATCC (StarFire)
hsa-miR-182-5p	AGTGTGAGTTCTACCATTGCCAAA (StarFire)
U6 snRNA	GCGTGTCATCCTTGCGCAGGGGCC

Table 2.2: Primers used for construction of miRNA expression and luciferase reporter constructs

MiRNA Name or Target Site	Forward Primer	Reverse Primer
HCMV miR-US22	CGCTCGAGCTGTGCAGAATCATAAGTTTATGATG	GCGGATCCGAAAACGAGGACGACACGA
HCMV miR-US33as	CTCGAGCGACCATTTCCGTGCGATTAGC	GGATCCAATTGACGGTCCACTGAGGTG
miR-US22-5p-full	TCGAGCCCCGCGGACACACGCTGAAACAGC	GGCCGCTGTTTCAGCGTGTGTCCGCGGGGC
miR-US22-5p-seed	TCGAGCCCGGGAATTCGTGAAACTCTAGAGC	GGCCGCTCTAGAGTTTCACGAATTCCCGGGC
miR-US33as-5p-full	TCGAGCCGTCACGGTCCGAGCACATCCAGC	GGCCGCTGGATGTGCTCGGACCGTGACGGC
miR-US33as-5p-seed	TCGAGCCCGGGAATTCGACATCCTCTAGAGC	GGCCGCTCTAGAGGATGTCGAATTCCCGGGC

Table 2.3: Comparison of smRNA-seq results with miRBase 16.0 HCMV miRNA annotations

Only the top two most abundant mapped sequences associated with each HCMV mature miRNA are listed. Total read counts for each sequence and percentages (out of the total reads aligned to the given miRNA) are indicated. MiRBase 16.0 sequences (*) are only listed for those miRNAs that differed from annotations based on their most abundant form sequenced at 72 hpi.

Name	Most abundant sequences	Read counts	
Analysis from 24 hpi:			
miR-UL112-3p	AAGTGACGGTGAGATCCAGGCT	452	78.6%
	AAGTGACGGTGAGATCCAGGCTT	26	4.5%
miR-UL148D	TCGTCCTCCCCTTCTTCACCGT	24	47.1%
	TCGTCCTCCCCTTCTTCACCG	17	33.3%
miR-UL22A-5p	CTAACTAGCCTTCCCGTGAGA	631	46.0%
	TAACTAGCCTTCCCGTGAGA	619	45.1%
miR-UL22A-3p	TCACCAGAATGCTAGTTTGTAG	5652	73.7%
	TCACCAGAATGCTAGTTTGTAGT	1301	17.0%
miR-UL36-5p	TCGTTGAAGACACCTGGAAAGA	63877	97.3%
	TCGTTGAAGACACCTGGAAAGAA	543	0.8%
miR-UL36-3p	TTTCCAGGTGTTTTCAACGTG	926	52.6%
	TTTCCAGGTGTTTTCAACGTGT	317	18.0%
miR-US25-1-5p	AACCGCTCAGTGGCTCGGACCG	1667	38.7%
	AACCGCTCAGTGGCTCGGACC	879	20.4%
miR-US25-1-3p	GTCCGAACGCTAGGTCGGTTCT	71	93.4%
	GTCCGAACGCTAGGTCGGTTCTAA	2	2.6%
miR-US25-2-3p	ATCCACTTGGAGAGCTCCCGCGGT	1344	93.0%
	ATCCACTTGGAGAGCTCCCGCGGTT	31	2.1%
miR-US25-2-5p	AGCGGTCTGTTCAGGTGGATGA	802	87.9%
	TAGCGGTCTGTTCAGGTGGATGA	50	5.5%
miR-US33-3p	TCACGGTCCGAGCACATCCAA	74	64.3%
	TCACGGTCCGAGCACATCCA	17	14.8%
miR-US33-5p	ATTGTGCCCGGACCGTGGGCGC	554	42.6%
	GATTGTGCCCGGACCGTGGGCG	311	23.9%

Table 2.3 (continued): Comparison of smRNA-seq results with miRBase 16.0 HCMV miRNA annotations

Name	Most abundant sequences	Read counts	
miR-US4-5p	TGGACGTGCAGGGGGATGTCTG	162	68.9%
	TGGACGTGCAGGGGGATGTCTGT	29	12.3%
miR-US5-1	TGACAAGCCTGACGAGAGCGT	73	66.4%
	TGACAAGCCTGACGAGAGCGTT	33	30.0%
miR-US5-2-3p	TTATGATAGGTGTGACGATGTC	3296	68.1%
	TATGATAGGTGTGACGATGTCT	1107	22.9%
miR-UL112-5p	CCTCCGGATCACATGGTTACTCAG	107	49.5%
	CCTCCGGATCACATGGTTACTCA	82	38.0%
miR-US4-3p	TGACAGCCCGCTACACCTCTCT	60	44.8%
	TGACAGCCCGCTACACCTCTCTG	46	34.3%
miR-US5-2-5p	CTTTCGCCACACCTATCCTGAAAG	56	28.6%
	GCTTTCGCCACACCTATCCTGA	39	19.9%
miR-US22-5p	TGTTTCAGCGTGTGTCCGCGGGC	15	48.4%
	TGTTTCAGCGTGTGTCCGCGGG	14	45.2%
miR-US22-3p	TCGCCGGCCGCGCTGTAACCAGG	38	88.4%
	TCGCCGGCCGCGCTGTAACCAGGT	4	9.3%
miR-US33as-5p	TGGATGTGCTCGGACCGTGACG	94	17.7%
	GGATGTGCTCGGACCGTGACG	91	17.1%
miR-US33as-3p	CCCACGGTCCGGGCACAATCAA	11	18.6%
	CCCACGGTCCGGGCACAATCAAT	11	18.6%
Analysis from 72 hpi:			
miR-UL112-3p	AAGTGACGGTGAGATCCAGGCT	3319	67.4%
	AAGTGACGGTGAGATCCAGGCTA	215	4.4%
miR-UL148D	TCGTCCTCCCCTTCTTCACCGT	317	61.1%
	TCGTCCTCCCCTTCTTCACCG	116	22.4%
	TCGTCCTCCCCTTCTTCACCG*		
miR-UL22A-5p	TAACTAGCCTTCCCGTGAGA	9656	55.7%
	CTAACTAGCCTTCCCGTGAGA	6069	35.0%
miR-UL22A-3p	TCACCAGAATGCTAGTTTGTAG	33536	83.7%
	TCACCAGAATGCTAGTTTGTAGA	3178	7.9%

Table 2.3 (continued): Comparison of smRNA-seq results with miRBase 16.0 HCMV miRNA annotations

Name	Most abundant sequences	Read counts	
miR-UL36-5p	TCGTTGAAGACACCTGGAAAGA	115685	96.0%
	TCGTTGAAGACACCTGGAAAGAA	2010	1.7%
miR-UL36-3p	TTTCCAGGTGTTTTCAACGTG	1068	56.3%
	TTTCCAGGTGTTTTCAACGTGT	359	18.9%
	TTTCCAGGTGTTTTCAACGTGC*		
miR-US25-1-5p	AACCGCTCAGTGGCTCGGACC	11057	41.3%
	AACCGCTCAGTGGCTCGGACCG	7415	27.7%
miR-US25-1-3p	GTCCGAACGCTAGGTCGGTTCT	168	94.4%
	GTCCGAACGCTAGGTCGGTTCTA	2	1.1%
	TCCGAACGCTAGGTCGGTTCTC*		
miR-US25-2-3p	ATCCACTTGGAGAGCTCCCGCGGT	13512	86.7%
	ATCCACTTGGAGAGCTCCCGCGGTA	729	4.7%
	ATCCACTTGGAGAGCTCCCGCGG*		
miR-US25-2-5p	AGCGGTCTGTTCAGGTGGATGA	11185	94.6%
	TAGCGGTCTGTTCAGGTGGATGA	168	1.4%
miR-US33-3p	TCACGGTCCGAGCACATCCAA	316	69.9%
	TCACGGTCCGAGCACATCCA	71	15.7%
	TCACGGTCCGAGCACATCCA*		
miR-US33-5p	GATTGTGCCCGGACCGTGGGCG	2327	37.2%
	ATTGTGCCCGGACCGTGGGCGC	1706	27.3%
miR-US4-5p	TGGACGTGCAGGGGGATGTCTG	439	56.9%
	TGGACGTGCAGGGGGATGTCTGT	80	10.4%
	CGACATGGACGTGCAGGGGGAT*		
miR-US5-1	TGACAAGCCTGACGAGAGCGT	420	78.2%
	TGACAAGCCTGACGAGAGCGTT	53	9.9%
miR-US5-2-3p	TATGATAGGTGTGACGATGTCT	85959	53.1%
	TTATGATAGGTGTGACGATGTC	64215	39.7%
	TTATGATAGGTGTGACGATGTC*		
miR-UL112-5p	CCTCCGGATCACATGGTTACTCA	151	39.2%
	CCTCCGGATCACATGGTTACTCAG	150	39.0%
miR-US4-3p	TGACAGCCCGCTACACCTCTCT	102	38.5%
	TGACAGCCCGCTACACCTCTCTG	84	31.7%

Table 2.3 (continued): Comparison of smRNA-seq results with miRBase 16.0 HCMV miRNA annotations

Name	Most abundant sequences	Read counts	
miR-US5-2-5p	CTTTCGCCACACCTATCCTGAAAG	125	26.8%
	CTTTCGCCACACCTATCCTGAAAGT	78	16.7%
miR-US22-5p	TGTTTCAGCGTGTGTCCGCGGG	1362	46.3%
	TGTTTCAGCGTGTGTCCGCGGGC	836	28.4%
miR-US22-3p	TCGCCGGCCGCGCTGTAACCAGG	389	87.8%
	TCGCCGGCCGCGCTGTAACCAGGT	25	5.6%
miR-US33as-5p	TGGATGTGCTCGGACCGTGACG	2547	26.8%
	GGATGTGCTCGGACCGTGACGGTGT	1472	15.5%
miR-US33as-3p	CCCACGGTCCGGGCACAATCA	318	42.9%
	CCCACGGTCCGGGCACAATCAAC	139	18.7%

2.7. References

Ambros, V., Bartel, B., Bartel, D.P., Burge, C.B., Carrington, J.C., Chen, X., Dreyfuss, G., Eddy, S.R., Griffiths-Jones, S., Marshall, M., Matzke, M., Ruvkun, G., and Tuschl, T. (2003). A uniform system for microRNA annotation. RNA *9*, 277-279.

Barth, S., Pfuhl, T., Mamiani, A., Ehses, C., Roemer, K., Kremmer, E., Jaker, C., Hock, J., Meister, G., and Grasser, F.A. (2008). Epstein-Barr virus-encoded microRNA miR-BART2 down-regulates the viral DNA polymerase BALF5. Nucleic Acids Res *36*, 666-675.

Burroughs, A.M., Ando, Y., Hoon, M.L., Tomaru, Y., Suzuki, H., Hayashizaki, Y., and Daub, C.O. (2011). Deep-sequencing of human Argonaute-associated small RNAs provides insight into miRNA sorting and reveals Argonaute association with RNA fragments of diverse origin. RNA Biol *8*.

Chi, S.W., Zang, J.B., Mele, A., and Darnell, R.B. (2009). Argonaute HITS-CLIP decodes microRNA-mRNA interaction maps. Nature *460*, 479-486.

Dolken, L., Malterer, G., Erhard, F., Kothe, S., Friedel, C.C., Suffert, G., Marcinowski, L., Motsch, N., Barth, S., Beitzinger, M., Lieber, D., Bailer, S.M., Hoffmann, R., Ruzsics, Z., Kremmer, E., Pfeffer, S., Zimmer, R., Koszinowski, U.H., Grasser, F., Meister, G., and Haas, J. (2010). Systematic analysis of viral and cellular microRNA targets in cells latently infected with human gamma-herpesviruses by RISC immunoprecipitation assay. Cell Host Microbe *7*, 324-334.

Dolken, L., Perot, J., Cognat, V., Alioua, A., John, M., Soutschek, J., Ruzsics, Z., Koszinowski, U., Voinnet, O., and Pfeffer, S. (2007). Mouse cytomegalovirus microRNAs dominate the cellular small RNA profile during lytic infection and show features of posttranscriptional regulation. J Virol *81*, 13771-13782.

Dolken, L., Pfeffer, S., and Koszinowski, U.H. (2009). Cytomegalovirus microRNAs. Virus Genes *38*, 355-364.

Grey, F., Antoniewicz, A., Allen, E., Saugstad, J., McShea, A., Carrington, J.C., and Nelson, J. (2005). Identification and characterization of human cytomegalovirus-encoded microRNAs. J Virol *79*, 12095-12099.

Grey, F., Meyers, H., White, E.A., Spector, D.H., and Nelson, J. (2007). A human cytomegalovirus-encoded microRNA regulates expression of multiple viral genes involved in replication. PLoS Pathog *3*, e163.

Han, J., Lee, Y., Yeom, K.H., Nam, J.W., Heo, I., Rhee, J.K., Sohn, S.Y., Cho, Y., Zhang, B.T., and Kim, V.N. (2006). Molecular basis for the recognition of primary microRNAs by the Drosha-DGCR8 complex. Cell *125*, 887-901.

Jung, G.S., Kim, Y.Y., Kim, J.I., Ji, G.Y., Jeon, J.S., Yoon, H.W., Lee, G.C., Ahn, J.H., Lee, K.M., and Lee, C.H. (2011). Full genome sequencing and analysis of human cytomegalovirus strain JHC isolated from a Korean patient. Virus Res.

Kim, S., Lee, S., Shin, J., Kim, Y., Evnouchidou, I., Kim, D., Kim, Y.K., Kim, Y.E., Ahn, J.H., Riddell, S.R., Stratikos, E., Kim, V.N., and Ahn, K. (2011). Human cytomegalovirus microRNA miR-US4-1 inhibits CD8(+) T cell responses by targeting the aminopeptidase ERAP1. Nat Immunol.

Michaelis, M., Baumgarten, P., Mittelbronn, M., Driever, P.H., Doerr, H.W., and Cinatl, J., Jr. (2011). Oncomodulation by human cytomegalovirus: novel clinical findings open new roads. Med Microbiol Immunol *200*, 1-5.

Parameswaran, P., Sklan, E., Wilkins, C., Burgon, T., Samuel, M.A., Lu, R., Ansel, K.M., Heissmeyer, V., Einav, S., Jackson, W., Doukas, T., Paranjape, S., Polacek, C., dos Santos, F.B., Jalili, R., Babrzadeh, F., Gharizadeh, B., Grimm, D., Kay, M., Koike, S., Sarnow, P., Ronaghi, M., Ding, S.W., Harris, E., Chow, M., Diamond, M.S., Kirkegaard, K., Glenn, J.S., and Fire, A.Z. (2010). Six RNA viruses and forty-one hosts: viral small RNAs and modulation of small RNA repertoires in vertebrate and invertebrate systems. PLoS Pathog *6*, e1000764.

Pfeffer, S., Sewer, A., Lagos-Quintana, M., Sheridan, R., Sander, C., Grasser, F.A., van Dyk, L.F., Ho, C.K., Shuman, S., Chien, M., Russo, J.J., Ju, J., Randall, G., Lindenbach, B.D., Rice, C.M., Simon, V., Ho, D.D., Zavolan, M., and Tuschl, T. (2005). Identification of microRNAs of the herpesvirus family. Nat Methods *2*, 269-276.

Pfeffer, S., Zavolan, M., Grasser, F.A., Chien, M., Russo, J.J., Ju, J., John, B., Enright, A.J., Marks, D., Sander, C., and Tuschl, T. (2004). Identification of virus-encoded microRNAs. Science *304*, 734-736.

Prichard, M.N., Jairath, S., Penfold, M.E., St Jeor, S., Bohlman, M.C., and Pari, G.S. (1998). Identification of persistent RNA-DNA hybrid structures within the origin of replication of human cytomegalovirus. J Virol *72*, 6997-7004.

Riley, K.J., Rabinowitz, G.S., and Steitz, J.A. (2010). Comprehensive analysis of Rhesus lymphocryptovirus microRNA expression. J Virol *84*, 5148-5157.

Sacheli, R., Nguyen, L., Borgs, L., Vandenbosch, R., Bodson, M., Lefebvre, P., and Malgrange, B. (2009). Expression patterns of miR-96, miR-182 and miR-183 in the development inner ear. Gene Expr Patterns *9*, 364-370.

Santhakumar, D., Forster, T., Laqtom, N.N., Fragkoudis, R., Dickinson, P., Abreu-Goodger, C., Manakov, S.A., Choudhury, N.R., Griffiths, S.J., Vermeulen, A., Enright, A.J., Dutia, B., Kohl, A., Ghazal, P., and Buck, A.H. (2010). Combined agonist-antagonist genome-wide functional screening identifies broadly active antiviral microRNAs. Proc Natl Acad Sci U S A *107*, 13830-13835.

Sarver, A.L., French, A.J., Borralho, P.M., Thayanithy, V., Oberg, A.L., Silverstein, K.A., Morlan, B.W., Riska, S.M., Boardman, L.A., Cunningham, J.M., Subramanian, S., Wang, L., Smyrk, T.C., Rodrigues, C.M., Thibodeau, S.N., and Steer, C.J. (2009). Human colon cancer profiles show differential microRNA expression depending on mismatch repair status and are characteristic of undifferentiated proliferative states. BMC Cancer *9*, 401.

Schaefer, A., Jung, M., Mollenkopf, H.J., Wagner, I., Stephan, C., Jentzmik, F., Miller, K., Lein, M., Kristiansen, G., and Jung, K. (2010). Diagnostic and prognostic implications of microRNA profiling in prostate carcinoma. Int J Cancer *126*, 1166-1176.

Spector, D.H. (1996). Activation and regulation of human cytomegalovirus early genes. Intervirology *39*, 361-377.

Stern-Ginossar, N., Elefant, N., Zimmermann, A., Wolf, D.G., Saleh, N., Biton, M., Horwitz, E., Prokocimer, Z., Prichard, M., Hahn, G., Goldman-Wohl, D., Greenfield, C., Yagel, S., Hengel, H., Altuvia, Y., Margalit, H., and Mandelboim, O. (2007). Host immune system gene targeting by a viral miRNA. Science *317*, 376-381.

Stern-Ginossar, N., Saleh, N., Goldberg, M.D., Prichard, M., Wolf, D.G., and Mandelboim, O. (2009). Analysis of human cytomegalovirus-encoded microRNA activity during infection. J Virol *83*, 10684-10693.

Wang, F.Z., Weber, F., Croce, C., Liu, C.G., Liao, X., and Pellett, P.E. (2008). Human cytomegalovirus infection alters the expression of cellular microRNA species that affect its replication. J Virol *82*, 9065-9074.

Weston, M.D., Pierce, M.L., Rocha-Sanchez, S., Beisel, K.W., and Soukup, G.A. (2006). MicroRNA gene expression in the mouse inner ear. Brain Res *1111*, 95-104.

Xu, S., Witmer, P.D., Lumayag, S., Kovacs, B., and Valle, D. (2007). MicroRNA (miRNA) transcriptome of mouse retina and identification of a sensory organ-specific miRNA cluster. The Journal of biological chemistry *282*, 25053-25066.

Zeng, Y., Yi, R., and Cullen, B.R. (2005). Recognition and cleavage of primary microRNA precursors by the nuclear processing enzyme Drosha. EMBO J *24*, 138-148.

Zhang, G., Raghavan, B., Kotur, M., Cheatham, J., Sedmak, D., Cook, C., Waldman, J., and Trgovcich, J. (2007). Antisense transcription in the human cytomegalovirus transcriptome. J Virol *81*, 11267-11281.

Zisoulis, D.G., Lovci, M.T., Wilbert, M.L., Hutt, K.R., Liang, T.Y., Pasquinelli, A.E., and Yeo, G.W. (2010). Comprehensive discovery of endogenous Argonaute binding sites in Caenorhabditis elegans. Nat Struct Mol Biol *17*, 173-179.

Chapter 3: Induction of CPEB1 expression and host 3′UTR shortening during human cytomegalovirus infection

3.1. Abstract

Human herpesviruses persistently infect the majority of the population, significantly impacting global health. The extent to which one member of this family, human cytomegalovirus (HCMV), influences host RNA processing during infection is not well understood. We have performed comprehensive transcriptome analysis of three HCMV-infected primary cell types, and report here the identification of thousands of infection-altered host splicing and polyadenylation events. Notably the majority of the alternative polyadenylation events favored shorter 3′-untranslated regions (3′UTRs). Our comparative analysis of cell type-independent host gene expression changes revealed significant induction of cytoplasmic polyadenylation element binding protein 1 (CPEB1), an RNA binding protein with roles in alternative splicing and 3′UTR processing. CPE and polyadenylation signal co-occurrences were enriched upstream of affected proximal 3′-ends, and delivery of CPEB1 to non-infected cells caused shortening of alternative 3′UTRs tested. Prevalent 3′UTR shortening was also observed in herpes simplex virus (HSV)-infected cells, demonstrating that herpesvirus infections broadly impact host RNA 3′-end formation.

3.2. Introduction

Lytic viral infection places an enormous stress on the host cell machinery and biological processes. The cellular environment must be positioned for viral replication, and in the most dramatic cases, complete shutdown of host transcription and protein

synthesis occurs. Effects of this magnitude can arise during infection with particular members of the herpesvirus family, which collectively infect the majority of the human population and represent a global health priority. In the case of herpes simplex virus (HSV), general inhibition of RNA splicing (Hardy and Sandri-Goldin, 1994) and overall decreased cellular mRNA levels are observed (Schek and Bachenheimer, 1985). In contrast, infection with human cytomegalovirus (HCMV) does not result in an overall host mRNA reduction, although vast gene expression differences occur (Hertel and Mocarski, 2004).

HCMV has a large ~235 kb DNA genome that remarkably encodes 200+ open reading frames (ORFs), many of which are temporally expressed and alternatively spliced during infection (Mocarski, 2007; Stern-Ginossar et al., 2012). HCMV has been reported to infect nearly 100% of individuals in the United States by adulthood without readily apparent health consequences in most cases (Staras et al., 2006). Symptomatic infections are a problem, however, for immunocompromised individuals, and newborns congenitally infected with HCMV often suffer severe developmental nervous system defects (Mocarski, 2007). HCMV infects a variety of cell types, and how either the host or viral transcriptomes compare across different cellular environments remains an open question.

Alternative RNA splicing contributes immensely to the diversity of the human transcriptome, and changes in these patterns can cause dramatic effects at the protein level. Furthermore, dysregulated RNA splicing patterns have been implicated in a number of devastating human diseases (Wang and Cooper, 2007). Modifications at the 3′-terminal ends of transcripts can also greatly impact the post-transcriptional fate of RNAs. Examples of these alterations include a shift in the location of 3′-end cleavage, or even the length of the poly(A) tail itself. More than half of human genes specify

alternative sites of cleavage and polyadenylation (Tian et al., 2005), and whether this feature is affected during viral infections is also unclear.

In this study we performed deep sequencing of the transcriptomes of three different primary cell types infected with HCMV: primary human foreskin fibroblasts (HFFs), human aortic endothelial cells (ECs), and human embryonic stem (ES) cell-derived neural progenitor cells (NPCs). These cell types support varying levels of viral infection (Belzile et al., 2014), and we capitalized on this divergence to address key themes pertaining to both the viral and host RNA. Interestingly, despite differences in the progression of the infection in the cell types analyzed, we consistently observed similar widespread host alternative splicing defects and changes in polyadenylation patterns. The altered polyadenylation events primarily resulted in shorter 3′UTRs, which prompted us to search for conserved effects of HCMV infection on host factors involved in 3′-end formation and processing. In our analysis we observed upregulation of cytoplasmic polyadenylation element binding protein 1 (CPEB1), an RNA binding protein which impacts alternative splicing and 3′UTR processing, and plays an essential role in early development and neuron functionality (Bava et al., 2013; Richter, 2007). In fact, ectopic expression of CPEB1 in non-infected HFFs recapitulated 3′-end shortening for candidate targets. Finally, in extending our analysis to other viral infections, we also observed 3′UTR shortening in cells infected by herpes simplex virus 2 (HSV-2) but not by the non-related retrovirus human immunodeficiency virus type 1 (HIV-1). Our findings indicate that this important aspect of cellular RNA processing is similarly affected by infections with different members of the herpesvirus family.

3.3. Results

3.3.1. HCMV messenger RNA levels are diminished in neural progenitors versus other permissive cell types during infection

To assess for cell type-specific differences in mRNA expression patterns for the genes encoded by HCMV during infection, we generated poly(A)$^+$ RNA-seq libraries from HFFs, ECs and NPCs derived from human ES cells at both 48 and 96 hours post-infection (hpi) with a clinical isolate TB40E (Figure 3.1A). Differential HCMV expression in the NPCs was characterized by notably lower levels of the relative amount of viral vs. cellular mRNA, apparent at both 48 and 96hpi (Figure 3.1A and Table 3.1). Interestingly, viral gene expression in HFF, EC, and NPC cells grouped together at 48 hpi in hierarchical clustering analysis (Figure 3.1B, based on annotated HCMV ORFs). However, the viral mRNA profile within NPCs was relatively unchanged from 48 to 96 hpi, unlike the marked differences in HCMV gene expression observed in HFFs and ECs as the infection progressed from early to late times (Figure 3.1C). As anticipated, the HCMV transcripts in the 96hpi NPCs that showed the lowest relative abundance (Figure 3.1C, highlighted in green) predominantly belonged to the late temporal class of HCMV genes, and the levels were considerably lower than those in HFFs or ECs at 96 hpi (Figure 3.1D). The low-level late gene expression is consistent with the cellular observations of non-progressive infection that was previously observed in this NPC model (Belzile et al., 2014).

3.3.2. HCMV splicing and polyadenylation are consistent across cell types

To evaluate if differences in post-transcriptional RNA processing were also apparent, we measured splice junctions, and 3′-cleavage and polyadenylation sites of HCMV transcripts. Recent deep sequencing and ribosome profiling studies in HFFs found evidence for over 200 splice junctions (Gatherer et al., 2011; Stern-Ginossar et al.,

2012). Our RNA-seq analysis confirmed the majority (over 85%) of the previously reported splice junctions (Table 3.2). Furthermore, we found nearly complete overlap of the splice junctions detected among HFFs, ECs, and NPCs. Mapped junctions that were specific to a particular cell type had low read coverage support. We did not find evidence for spliced HCMV mRNA isoforms specific to a single cell type, demonstrating consistent viral splicing patterns across distinct cell types. Additionally, we identified a small subset of 46 splice junctions that were not detected in either the Gatherer or Stern-Ginossar et al. studies (Figure 3.2A and Table 3.2). The UL41A example shown is lowly abundant compared to non-spliced isoforms spanning this region, however the spliced isoforms for UL144 were abundant. Although we could not identify putative ORFs within the novel UL144 spliced isoform, this variant likely originates from alternative 5′ end usage at this locus.

Next, we computationally isolated poly(A)-containing RNA-seq reads that also harbored upstream HCMV-mapped sequence for analysis of 3′-end transcript termination sites in HCMV. The mapping locations of the final non-poly(A) base of each of these reads were compiled to define the 3′-ends of all HCMV transcripts detected (Figure 3.2B and Table 3.3). Filtering was performed to eliminate reads that map to the few regions in the HCMV genome that include runs of oligo(A) sequence. Using a minimum threshold of 50 raw reads to constitute a poly(A) site, we identified 137 total unique 3′-termini across all cell types, with an average of 116 sites found in each cell population at 48 hpi, and 119 sites at 96 hpi. Representative analysis is shown in Figure 3.2B for overlapping and co-terminal RNAs that in part span the UL98-99 ORFs. Previous interrogation of this region has been performed for this region, yet in these studies only one of the two termination sites was reported (Wing and Huang, 1995).

We concluded that HCMV poly(A) site choices, like splicing patterns, were consistent across the different cell types. Any detected poly(A) site specific to only one time-point and/or cell type was supported by only low sequencing coverage. In contrast, 90% of all the HCMV poly(A) sites identified were supported by 400 reads or more, with the most abundant coverage found at the 3′-end of the long non-coding RNA2.7. Examples of HCMV 3′-termination sites that are shared across cell types are shown in Figure 3.2C. The poly(A) site adjacent to UL132, along with other HCMV transcript termini defined in our analysis, was further validated in both TB40E and Towne strain infections (Figure 3.2D and Table 3.3) (Moucadel et al., 2007).

To summarize, we found similar HCMV mRNA expression patterns in the HFFs and ECs, but the 48-to-96 hpi progression was suppressed in the NPCs. However, despite the reduced accumulation of HCMV late genes, our results indicate consistent processing of HCMV RNAs in diverse cell types.

3.3.3. Host alternative RNA splicing is extensively disrupted by HCMV infection

We next centered our attention on the host cell and conducted genome-wide assessments of host RNA processing events in our infection samples. Alternative RNA splicing during HCMV infection has been documented for at least one host gene *Xbp-1*, which is involved in the cellular unfolded protein response (Isler et al., 2005b). Otherwise, the global status of cellular splicing patterns during HCMV infection has not been evaluated. We employed custom-designed splicing-sensitive Affymetrix microarrays to profile host alternative splicing events in HFFs post-infection. For infected HFFs, this analysis revealed approximately 6,000 significantly altered host RNA splicing events (Figure 3.3A and Table 3.4). Although HCMV does not completely shut down host RNA splicing, in contrast to many other viruses, these statistics highlight that a

large number of splicing defects arise by mid stages of infection and continue to accumulate throughout the rest of the viral replication cycle.

Next, we analyzed the infected NPCs and recorded many significant changes (Figure 3.3A and Table 3.4). Although we found fewer significantly changing alternative splicing events in the NPCs relative to the HFFs, this decrease is likely due to the less productive infection in the NPCs. Nevertheless, nearly a quarter of all events measured by the microarray were affected in the NPCs, and half were altered in HFFs (Figure 3.3B), demonstrating that alternative splicing is affected by HCMV infection in multiple cell types. When we examined alternative cassette exon events, where a single exon is preferentially included or excluded from a transcript, we uncovered a large degree of overlap between the NPCs and HFFs (Figure 3.3C and Table 3.5). The large set of changes in alternative splicing within infected HFFs and NPCs was not significantly enriched for particular gene ontology categories, which leads us to propose that the changes represent a more general disruption of host alternative splicing patterns.

We performed RT-PCR to confirm alternative splicing changes of candidate genes that were shared by all the infected cell types (Figure 3.3D), and Towne-HFF infections were included to rule out virus strain-specific inconsistencies. The example of *USPL1* (ubiquitin specific peptidase like 1) is notable in that the alternative exon included during infection is the first protein-coding exon of a major isoform. Without this exon, the reading frame of this isoform is disrupted and only shorter coding regions are otherwise possible. Hence the *USPL1* isoform encoding the largest possible molecular weight protein product is predominant in HCMV-infected cells (1092 amino acids, vs. 763 or less). The *CAST* gene (exon exclusion observed during infection, Figure 3.3D) encodes Calpastatin, an inhibitor of calpains, which are involved in calcium-dependent proteolysis (Hanna et al., 2008). Over 10 spliced isoforms exist for *CAST*, and the alternative exon

affected by HCMV infection is found within the first 25 bases of protein-coding sequence for most isoforms when it is included. Removal of this exon does not cause a reading frame shift (the exon is 66 bases long), but the loss of these residues could impact protein function. CAST protein levels have been reported to not change in HCMV-infected cells, although enhanced calpain activity was observed in the infected cells (Chen et al., 2001). Of note, in the previous study, the abundance levels of calpains were unchanged and could not account for this increased activity.

We validated alternative splicing changes that were specific to either the infected NPCs or HFF/EC conditions (Figure 3.3D). For example, the alternative splicing pattern observed for the *MYO18A* gene was interesting in that significant exon exclusion was observed during infection in HFFs and ECs, yet the smaller isoform of this transcript was already present in NPCs prior to infection. This alternative exon is located towards the 3′-end of the transcript and does not result in coding frame-shift when removed, although 15 residues are compromised. Importance for MYO18A has been documented in cell survival following DNA damage, particularly in the context of damage to the endoplasmic reticulum (Farber-Katz et al., 2014). The example of exon exclusion in *ATP8A1*, a P-type ATPase, appeared specific to NPCs. *ATP8A1* exon exclusion was predicted to be a common splicing change, but we could not detect the larger isoform by RT-PCR in mock-infected HFFs at this low expression level. Inclusion of this additional sequence is not predicted to impact the reading frame involved, but functionality of the protein product could be affected.

We have thus identified widespread alternative splicing changes in thousands of host genes that occur upon HCMV infection, many of which overlap in the distinct cell types analyzed. These results point to a disruptive effect of HCMV infection on host alternative splicing patterns that has not been previously appreciated.

3.3.4. HCMV infection has widespread effects on host poly(A) site preferences

To identify relevant 3′-end cleavage sites in the human transcriptome, we again employed analysis of poly(A)-containing RNA-seq reads (as performed for the virus transcriptome). Poly(A) site definitions specific to our sample conditions were used to generate custom 3′UTR annotations to analyze differential isoform coverage with the MISO algorithm (Katz et al., 2010). Our annotations were restricted to genes with more than one 3′-end detected by RNA-seq, and we used MISO to specifically query alternative tandem UTR events, or situations in which the 3′-end location preference of a transcript would change during infection. This particular class of differential 3′UTR isoforms analyzed is distinct from alternative splicing to differential 3′-terminal exons, another mechanism of 3′-end regulation.

We found hundreds of mRNAs with significantly altered cleavage and polyadenylation patterns in each of the cell types assessed. Also, the majority of these events resulted in shortened 3′UTRs for the affected genes, consistent across the HFF, EC and NPC samples (Figure 3.4A and Table 3.6). Many of these shortening events were found in at least two of the cell types, with a subset occurring in all three (Figure 3.4B). By 96 hpi, affected transcripts were generally shortened by 1 to 5 kb in length (Figure 3.4C). We considered MISO predictions for genes at 96 hpi: *ANKH*, *PCGF3* and *MARCH6*, and the 3′-end shortening was visually apparent from the RNA-seq coverage for all of these transcripts (Figure 3.4D). Furthermore, we independently confirmed these events by a qRT-PCR approach (Figure 3.4E) (Ulitsky et al., 2012).

PCGF3 encodes a component of a Polycomb-related complex, and alternative polyadenylation of this gene could have implications for the heterochromatic status of downstream genes. Genetics mutations in *ANKH* have been linked to ankylosis (disease

characterized by joint stiffness) and neurological abnormalities (Morava et al., 2011). Furthermore, the murine ortholog is expressed highly in the brain (Ho et al., 2000). *MARCH6*, also known as *TEB4*, encodes an E3-ubiquitin ligase and may be important for protein degradation in the endoplasmic reticulum.

3.3.5. 3′UTR shortening is shared among different herpesvirus infections

To determine whether these altered host polyadenylation preferences could be observed in other viral infections, we measured the distal 3′UTR levels in HSV-2-infected cells for two genes that had been altered upon HCMV infection. A strong effect could be observed in these cells, demonstrating the same trend towards 3′UTR shortening (Figure 3.4F). We subsequently investigated this finding at a transcriptome-wide scale through RNA-seq analysis of HFFs infected with HSV-2 (Figures 3.4A and 3.4D), comprehensively extending this finding of host 3′UTR shortening to infections by other members of the herpesvirus family. Additionally, we wondered if these 3′-end alterations could be observed in cells infected with a non-related virus outside of this family. We tested T-cells infected with the lentivirus HIV-1 and found only a minimal number of changing UTR events (Figure 3.4A). Our results indicate that while not all viral infections lead to this molecular phenotype, and these findings reveal a conserved consequence of herpesvirus family members on host 3′-end RNA processing.

3.3.6. The RNA binding protein CPEB1 is induced in HCMV-infected cells

Following our observation of dramatic effects of HCMV infection on both host alternative splicing and 3′UTR isoforms in all the cell types analyzed, we next searched for differences in abundance levels for host factors that regulate 3′-end formation and processing. We first performed a comparative analysis on the host gene expression changes, and determined transcript abundances (RPKMs (Mortazavi et al., 2008)) for all the infected and mock-infected conditions (Figure 3.5A-C). Significant expression

differences were determined for each infected cell type using a Z-score analysis (Polymenidou et al., 2011) (Tables 3.7-3.9). We then narrowed our search to a list of key members of the 3′-end processing machinery and factors that have been documented to affect this apparatus (Figure 3.6A). Of particular interest, and unexpectedly, the RNA-binding protein CPEB1 emerged as the most consistently affected factor in all the cell types. Additionally, while most of the critical mediators of 3′-cleavage and polyadenylation are basally expressed in these cells (CstFs, CPSFs, and CFs), CPEB1 mRNA was only minimally expressed in non-infected HFFs and NPCs, and nearly undetectable in ECs. CPEB1 expression was dramatically induced at the RNA level during infection (6-fold in HFFs, and over 20-fold in both ECs and NPCs by 96 hpi; Z-score significance reached in ECs and NPCs, p-value < 0.05).

Increased expression of CPEB1 is normally found in brain tissues, and important roles have been attributed in early development, especially in the *Xenopus* oocyte model, where CPEB stimulates translation of select mRNAs by extending their poly(A) tails (Richter, 2007). Furthermore, a recent report indicated that CPEB1 also influences RNA processing events, including alternative splicing and 3′-end cleavage patterns (Bava et al., 2013). Other CPEB family members 2, 3 and 4 were not significantly affected across all cell types (Figure 3.6A). Considering the potential implications of CPEB1 induction within infected cells, we confirmed the upregulation by an independent qPCR in two NPC lines (Figure 3.5D), and then determined by Western that this increase was also observed at the protein level in all the cell types (Figure 3.6B and Figure 3.5E).

An earlier study focusing on polyadenylation of specific HCMV transcripts demonstrated upregulation of the host 3′-end processing factor CstF-64 in HCMV-infected HFFs (Adair et al., 2004). Differential expression of this factor and other related

proteins has been shown to result in transcriptome-wide altered poly(A) site preferences (Martin et al., 2012; Yao et al., 2012), but we did not witness upregulation of CstF-64 by Western blot in the infected NPCs (Figure 3.6C). Additionally, we did not observe differences for a representative member of the cleavage factor (CF) complex (Figure 3.6C), along with other known constituents of the 3′-end processing machinery (data not shown). However, across all conditions post-infection, we did observe a reproducible higher molecular weight form of the CstF-77 protein, another subunit from the same complex as CstF-64, (Figure 3.6C). While CPEB1 upregulation is consistent with 3′UTR shortening (Bava et al., 2013), previous studies of CstF-77 do not indicate that this factor would likely shift polyadenylation preferences in predominantly this direction (Luo et al., 2013).

3.3.7. Increased CPEB1 expression affects endogenous 3′UTR targets in HFFs

To evaluate whether CPEB1 could be involved in the observed host 3′-end cleavage and polyadenylation changes in our system, we ectopically expressed the full-length isoform of CPEB1 in non-infected HFFs (Figure 3.7A). Strikingly, the increased expression of CPEB1 promoted a significant shift towards proximal poly(A) site choices for the candidates tested (Figure 3.7B; the *PCGF3 and ANKH* candidates were selected for their susceptibility to altered 3′-end regulation in both the HFFs and NPCs). Additional events investigated were indeed affected by increased CPEB1 levels (Figure 3.7C). *SYNRG*, *PPIA* (encodes cyclophilin A), and the transcription factor *YY1* all showed proximal enhancement. Regulation of the *YY1* 3′UTR is of interest since its protein product has binding recognition sites in the HCMV major immediate early promoter. In non-permissive cells YY1 has been suggested to be involved in blocking early HCMV gene expression events that are required for viral replication (Liu et al., 1994).

To consider again the well-established mechanism of how CPEB1 binds and regulates translation of targets in systems such as the *Xenopus* model, a main consensus sequence UUUUUAU, the cytoplasmic polyadenylation element (CPE) in proximity to a polyadenylation signal (PAS) is known to recruit CPEB1 to 3′UTR regions (Richter, 2007). Although we do not yet know whether these sequence elements are as significant for its recently proposed role in 3′UTR regulation (Bava et al., 2013), enrichment of these elements was reported for affected transcripts in regions surrounding alternative cleavage sites. To ask if the shortening of 3′UTRs during infection could be attributed to an enrichment of predicted CPEB1 recognition sites, we enumerated the fraction of infection-altered isoforms that contained both CPE and PAS elements within 150 to 250 bases of the proximal 3′-cleavage site. Approximately 40% to 50% of isoforms that shortened during infection contained CPE-PAS co-occurrences surrounding the proximal 3′-end, two-fold above isoforms that remained unchanged during infection (Figure 3.7D). This enrichment was statistically significant for both shortening (150 to 250 base windows around the proximal 3′-end) and extension (150 to 200 bases) events. Interestingly, the enrichment of the shortening events was significant when we limited our search to only upstream of the proximal 3′-cleavage site, while the extended events were not significant, suggesting a positional preference for CPE-PAS sites upstream of the proximal 3′-cleavage site for infection-altered shortening events (Figure 3.7D).

In conclusion, we conclude that increased CPEB1 expression in uninfected cells recapitulates shortening of 3′UTRs that are altered during HCMV infection, suggesting that CPEB1 upregulation during infection is one of the mechanisms by which HCMV could be capable of manipulating a vast network of RNA processing events in the host cell environment.

3.4. Discussion

A distinguishing feature of HCMV is its wide cellular tropism, yet not all permissive cell lineages support a full, productive infection. From our comparative transcriptome analysis of different primary cell types infected with HCMV, we recognized cell type-specific discrepancies, but also uncovered shared molecular outcomes of the infection. This unforeseen convergence was at the level of RNA processing, with massive alterations in host splicing and polyadenylation patterns observed in all the HCMV-infected cell types analyzed. We also further identified that host polyadenylation is similarly affected in HSV-2-infected cells, highlighting alterations to host 3′-end formation during infection as an emerging consequence for herpesvirus infections in general.

An initial motivation for the comparative transcriptome analysis of multiple HCMV infection models was to gain molecular insight into the neuropathological effects of HCMV infection, in part by leveraging in-depth comparisons to other permissive cell types. We assessed for gene expression inconsistencies that might account for differential infection outcomes in cells such as neural progenitors, and we observed clear differences in the viral gene expression signatures among the HFFs, ECs, and NPCs. A recent study on a separate set of cell types also reported variations in HCMV gene expression, pointing to virus transcription patterns that depend on the host cell environment (Towler et al., 2012). Beyond these gene expression differences, we were especially interested in determining whether HCMV RNAs might be processed differently between cell types. Despite the overall dampened HCMV transcription in NPCs versus HFFs and ECs, we did not observe any significant cell type-specific differences in HCMV RNA splicing or poly(A) site locations. Therefore, while HCMV genes are transcribed at

varying levels in different cellular environments, our results indicate consistent processing of the lytic HCMV transcriptome.

In terms of host gene expression, HCMV infection elicits a massive, complex network of changes, which was previously explored in HFFs through microarray profiling (Browne et al., 2001; Hertel and Mocarski, 2004). Our RNA-seq analysis demonstrated similar widespread effects in the infected ECs, which feature a similar level of infection, and also somewhat surprisingly, in the more resistant NPCs. This altered host gene expression program during HCMV infection features both marked up- and down-regulation of transcripts. In contrast, infections with certain herpesviruses from other subfamilies, such as HSV and KSHV, instead significantly down-regulate most host genes (Glaunsinger and Ganem, 2004; Hardwicke and Sandri-Goldin, 1994). Intensifying this ‘shut-down’ process, HSV infection has also been shown to cause general inhibition of host RNA splicing (Hardy and Sandri-Goldin, 1994). HCMV has considerably more introns and the splicing machinery thus needs to be maintained. Nevertheless, we found changes in a large portion of host alternative splicing events during infection. These alternative splicing defects appear to be distinct from the general HSV-mediated inhibition. In our splicing analysis of HFFs and NPCs we found many altered events common to both cell types, and this subset of modified host splicing patterns could be important for previously unrecognized aspects of HCMV infection.

In addition to this observation of altered splicing patterns throughout the host transcriptome in HCMV-infected cells, we found an extensive impact on host alternative cleavage and polyadenylation during infection. For affected transcripts this specifically meant a shift in poly(A) site choice, with the majority of cases favoring proximal poly(A) sites and resulting in decreased 3′UTR lengths. Other herpesviruses and select RNA viruses have been in fact shown to target the host polyadenylation machinery, especially

in the case of KSHV (Lee and Glaunsinger, 2009). Diminished host gene expression in KSHV-infected cells has been attributed to reduced mRNA stability during infection (Glaunsinger and Ganem, 2004), and extended poly(A) tails have been found on transcripts during infection (Lee and Glaunsinger, 2009). A KSHV-encoded protein was shown to cause hyperadenylation of the poly(A) tail, and in these cases the actual 3′-terminal modification is arguably distinct from the events profiled here in our study. Although we have not determined if the 3′UTR-shortened transcripts in HCMV-infected cells also feature differential poly(A) tail lengths, these and other host mRNAs are not down-modulated during HCMV infection.

Similar to KSHV, HSV-1 has been shown to induce hyperadenylation and degradation of reporter substrates (Ellison et al., 2000; Kumar and Glaunsinger, 2010). Although the full extent of this hyperadenylation remains to be determined in the context of HSV infection, our RNA-seq analysis of HFFs infected with HSV-2 identified altered transcript lengths, implying that both this effect and the hyperadenylation phenotype could co-exist. RNA viruses including enterovirus and Influenza A inhibit host polyadenylation altogether (Nemeroff et al., 1998; Weng et al., 2009), while SARS coronavirus infection has been suggested to cause hyperadenylation like other herpesviruses (Kumar and Glaunsinger, 2010). We attempted to extend our findings beyond herpesviruses, but did not observe comparable 3′UTR shortening in our RNA-seq analysis of T-cells infected with HIV-1.

Alternative cleavage and polyadenylation leading to 3′UTR isoforms changes is important in other biological contexts. Sweeping changes in 3′UTR lengths occur over the course of animal development (Ji et al., 2009; Ulitsky et al., 2012). Additionally, 3′UTR shortening has been observed in highly proliferative cell states, for example, in cancer cell lines (Mayr and Bartel, 2009; Sandberg et al., 2008). In line with this previous

model, a recent study of B cells infected with Epstein-Barr virus (EBV) demonstrated reduced 3′UTR lengths for a number of host transcripts (Homa et al., 2013). However, EBV establishes a latent infection and immortalizes B cells, and the altered host polyadenylation patterns were ultimately deemed a consequence of the associated proliferation rather than the infection itself (Homa et al., 2013). In our cellular contexts, HCMV infection is lytic and the cells are not proliferative, indicating potentially an alternate mechanism of global 3′UTR shortening.

A new and conserved mechanism in HCMV infections, we have shown that expression of the RNA-binding protein CPEB1 is induced in three distinct infected cell types. Association of CPEB1 with stress granules has been documented (Wilczynska et al., 2005), which would have critical implications for the accessibility of affected targets in HCMV infection. However, the infection has been demonstrated to prevent or bypass stress granule formation (Isler et al., 2005a). In oocyte models, CPEB1 has been previously shown to bind to a particular consensus sequence (UUUUUAU), termed the cytoplasmic polyadenylation element (CPE), in 3′UTRs in the proximity of poly(A) signals (Fox et al., 1989). Bioinformatics analysis revealed a striking enrichment in CPEB1 recognition sequences in the proximity of infection-altered 3′-end cleavage sites compared to unaffected 3′-ends. Our analysis was consistent with the presence of instances of this canonical binding sequence (specifically (U)UUUUAU or UUUUAAU) near the 3′-proximal ends of candidate genes that were validated in this study (Figures 3.4D, 3.7B, and 3.7C). As further support of the direct role of CPEB1 in alternative polyadenylation of genes affected during infection, ectopic expression of CPEB1 in non-infected cells led to shortening of 3′UTRs tested. Highly relevant to the pathologies associated with symptomatic HCMV disease, mis-regulated CPEB1 expression may in part explain a variety of the observed neurological abnormalities. Studies in primary rat

neural progenitor cells demonstrated altered neuronal differentiation following CPEB1 knockdown (Kim et al., 2013). Improperly timed CPEB1 upregulation in HCMV-infected progenitor cell populations could therefore have devastating effects. However, a variety of studies have indicated importance for CPEB family members in neuronal functionalities and plasticity (Richter, 2007), and the increased expression during infection may instead constitute a host protective response.

In summary, our interrogation of three distinct HCMV-infected cell populations yielded a unique opportunity to analyze critical features of RNA processing events from both the virus and host perspectives. We believe our identification of conserved induction of CPEB1 in HCMV-infected cells offers mechanistic insight into how these large-scale alterations in the host RNA landscape are in part arising in the infection program. While specific virus-encoded factors have been shown to be responsible for the above-described host shutdown phenomena, many of which converge at the theme of 3′-end regulation, HCMV does not encode such factors and must elicit the host cell to provide appropriate machinery. Induction of CPEB1 in HCMV infection could represent a new framework for further studies in how this factor participates in RNA processing events.

3.5. Materials and Methods

3.5.1. Virus infections

All TB40E HCMV infections were performed at an MOI of 5, and as recently described (Belzile et al., 2014). HUES9-derived NPCs were mainly used for this study, and H9-derived cells were used in extended follow-up comparisons (Figure 3.5D and E). Towne HCMV infections (RT-PCR comparisons shown in Figures 3.2 and 3.3) were conducted at an MOI of 3. HSV-2 infections were performed in HFFs with strain G at an

MOI of 10. Following an initial 30 min adsorption period at 4°C, mock-infected and infected cells were incubated at 37°C and harvested at 2 and 8 hpi. All HCMV and HSV-2 materials were collected at the time of harvest by trypsinization, briefly pelleted, and snap-frozen prior to subsequent analysis. HIV-1 infections of IL-2-activated CD4+ T-cells were performed as described (Noraz et al., 1997) and harvested directly into TRIzol (Life Technologies) at 5 days post-infection.

3.5.2. RNA-seq library preparation and data processing

Total RNA was isolated using TRIzol reagent, and 10 µg was first treated with Turbo DNase (Life Technologies). For all the HCMV TB40E infection and mock-infection conditions, and for the HIV-infected T-cells, we prepared strand-specific libraries using the dUTP method (Parkhomchuk et al., 2009) with adaptations described in detail previously (Huelga et al., 2012). For the TB40E- and mock-infected HFFs and ECs adaptor-containing oligod(T) was included during first-strand cDNA synthesis (cDNA Cloning Primer, ReadyMade Primers, Integrated DNA Technologies). Libraries for analysis of HSV2 infections were prepared with Illumina TruSeq Stranded mRNA Sample Preparation reagents. All samples were sequenced on the Illumina Hi-Seq platform. Each sample was run on a full lane, with the exception of the HSV2 samples, which were multiplexed and run together. After sequencing, reads were trimmed for adaptor sequences or low-quality bases and then mapped to both the human genome (hg19 build) and the HCMV Merlin genome (Genbank AY446894.2) with GSNAP. Additional filtering of reads that mapped to repetitive elements was also performed. Gene expression values (RPKM (Mortazavi et al., 2008)) were calculated within each sample, and Z-score analysis was implemented to identify significant differences in expression as previously described (Polymenidou et al., 2011). Cluster 3.0 software and

Java Treeview were used in combination to perform and visualize results from hierarchical gene expression clustering results.

3.5.3. Analysis of HCMV RNA, splicing junctions, and polyadenylation sites

For HCMV mRNA abundance measurements, we used the newly available TB40E genome sequence (Genbank KF297339.1, strain 'Lisa') for mapping and analysis of coverage of Genbank-listed HCMV ORF annotations. Otherwise all data processing was performed with the HCMV Merlin reference to facilitate comparisons to the Gatherer et al. and Stern-Ginossar et al. studies.

For HCMV splicing analysis, unmapped reads from the human+HCMV Merlin pipeline described above were queried against the Merlin reference sequence once more, with BLAT software v33 using default parameters. Resulting alignments were post-processed to remove non-spliced alignments, and only junctions that harbored canonical splice donor and acceptor sequences were considered.

For computational analysis of HCMV mRNA 3′-polyA-containing (at least 11 A bases) RNA-seq reads were extracted from the infected datasets, A-tracts were removed, and the upstream sequence was mapped to HCMV Merlin with Bowtie v0.12.7 (settings -m 10 -k 3 --best). Reads mapped to A_{11+} regions of the viral genome were filtered out. RT-PCR validations of poly(A) sites were performed on 1:10 diluted cDNA, which had been synthesized with Superscript III (Life Technologies) from 1 µg of DNase-treated total RNA. The universal oligod(T)-adaptor for RT and 3′-reverse PCR primer used for this analysis was from the GeneRacer system (Life Technologies).

3.5.4. Analysis of human alternative splicing and polyadenylation

Splicing-sensitive microarray analysis was performed as previously described (Huelga et al., 2012). RT-PCR splicing assays were performed using the equivalent of

50 ng of oligod(T)-primed cDNA (reverse transcription performed with Superscript III, Life Technologies) and 35 cycles of PCR amplification.

Tandem UTR isoform analysis was performed with the MISO algorithm v0.5.2 using default settings (Katz et al., 2010), except for use of custom 3′UTR isoform annotations. We used a Bayes-factor threshold of 10,000 and difference values (delta Psi) with an absolute value of at least 0.03 (although the cutoff selected for this latter value is low, we found significant degrees of UTR shift from this value upwards when used in combination with the high Bayes-factor for the HCMV samples). Negative difference values were interpreted as 3′-proximal-shifted, UTR shortening events (Confirmed with Y. Katz, personal communication). For the HIV-1-infected vs. non-infected T-cells we lowered the Bayes-factor threshold to 10 (7 events identified and indicated in Figure 3.4A). This threshold was also implemented for the HSV-2 infections, although the 10,000-threshold still yielded a large number of events. To generate the custom annotations we downloaded all Ensembl-defined human 3′UTR regions (http://uswest.ensembl.org, Release 75) and flattened them with the mergeBed function from Bedtools v2.16.2 (to define 3′UTR starts). Collapsed UTR regions harboring more than one cleavage site detected by RNA-seq were considered, with a minimum threshold of 5 reads required to constitute a poly(A) site. These cleavage-defining reads were based on human-mapped poly(A)$_{6+}$-containing RNA-seq reads (filtering performed for genomic regions with A-tracts). Finally, the two termination sites with the highest coverage were selected to define putative proximal and distal alternative ends of the 3′UTR. We generated one index based on poly(A) reads from the infected NPCs and another that was based on a composite of infected HFF and EC poly(A) reads (neuronal vs. non-neuronal was sufficient for this analysis and we had obtained 1.5-2X sequencing depth in the NPC samples vs. the HFF and EC libraries). Indexed annotations were

generated separately for the HSV-2 and HIV-1 conditions, again based on the samples' own poly(A)-RNA-seq reads. We additionally ran MISO for the HIV-1 infection data with the HCMV infection-derived annotations, but still recovered few candidate altered polyadenylation events.

qRT-PCR analysis of alternative 3′UTR isoforms was performed as described (Ulitsky et al., 2012). 1 µg of total RNA was first DNase-treated, and only oligod(T) was used for cDNA synthesis (generated with Superscript III, Life Technologies). All primer sequences for RT-PCRs and qPCRs are provided in the Supplementary Methods.

3.5.5. Western blot analysis

Whole cell lysates were prepared from -80°C-stored cell pellets with RIPA lysis buffer, on ice. Following brief sonication, lysates were clarified by 12,000 x g centrifugation for 10-15 min. Samples were loaded according to total protein content, determined by BCA assessment (Thermo Pierce), and within each cell type the amounts were normalized to the first mock-infected sample (corresponding to ~100,000 cells total per lane).

Antibodies and dilutions used:

CstF-64 (A301-092A), CstF-77 (A301-096A), and CFI_m-68 (A301-356A): Bethyl Laboratories, 1:2000.

CPEB1: Cell Signaling Technologies (13583), 1:1000.

β-actin: Sigma-Aldrich (clone AC-15, A1978), 1:10,000.

3.5.6. Lenviral vector production and transduction

Human CPEB1 open reading frame was PCR amplified from HFF cDNA with the following primers: Forward gcccgctgcaaaaatagtg and Reverse tcagcaagtgcaaaggtgac.

The PCR products were first cloned into the Topo/TA vector pCR2.1 (Life Technologies). CMV-turboGFP from pGIPZ (Open Biosystems) was replaced with the CMV promoter

from pCDNA3.1- (Life Technologies) and the cloned CPEB1 was inserted at SpeI and NotI restriction sites.

CPEB1 and control GFP lentiviruses were generated in 293T cells with PEI transfection reagent, in 10 cm format with 10^7 cells per dish, using the second generation packaging constructs psPAX2 and pMD2.G. Supernatants were harvested at 60 h post-transfection and 0.22 μm filtered. Freshly seeded HFFs were transduced at an MOI of 0.5 to 1, without polybrene or additional reagents. Puromycin selection was initiated in the transduced HFFs at 48 h, and cells were harvested following 3 days of drug selection.

3.6. Publication acknowledgement

Chapter 3, in full, has been submitted for publication, with the title “Induction of CPEB1 expression and host 3′UTR shortening during human cytomegalovirus infection.” The author of this dissertation was the primary investigator and author of this manuscript.

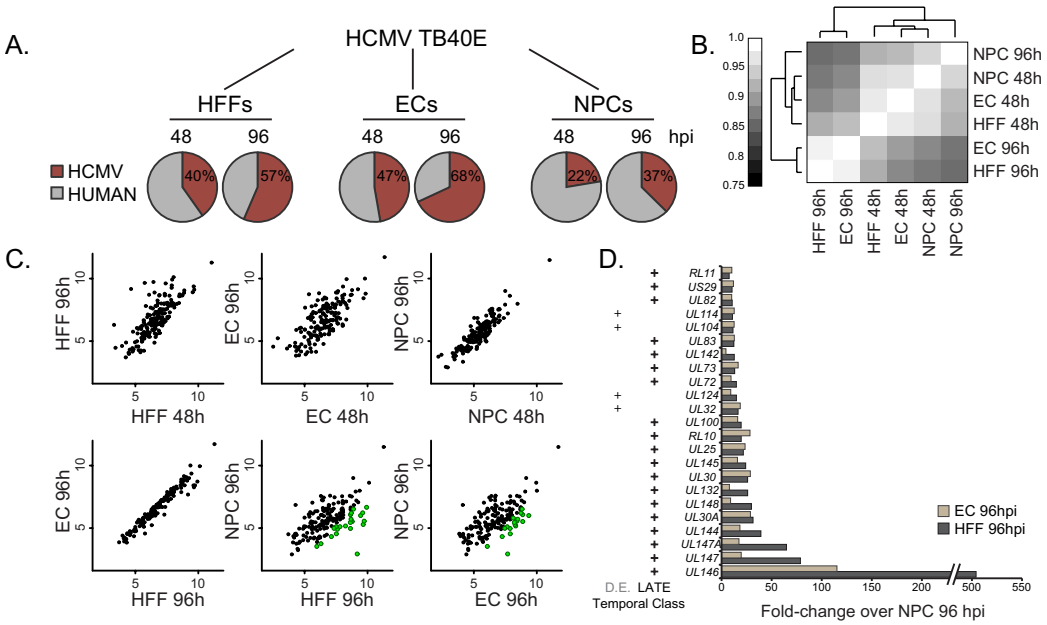

Figure 3.1: Differences in viral gene expression at late stages of HCMV infection in HFFs, ECs, and NPCs

(A) Overview of the different HCMV TB40E-infected cell types used for strand-specific mRNA-seq analysis. Control libraries were generated from mock-infected cells harvested at 48 h. The distribution of mapped sequencing reads between the human and viral genomes for each of the infection conditions is shown. (B) Correlation matrix for HCMV ORF expression values and hierarchical clustering of the correlation coefficients. (C) Scatter plot comparisons of HCMV transcription levels. Significantly under-represented genes in NPCs at 96 hpi are highlighted in green (D) Comparison of the highest fold-change differences between 96 hpi NPCs and 96 hpi HFFs/ECs. Associated temporal gene classes are indicated. D.E. is Delayed Early.

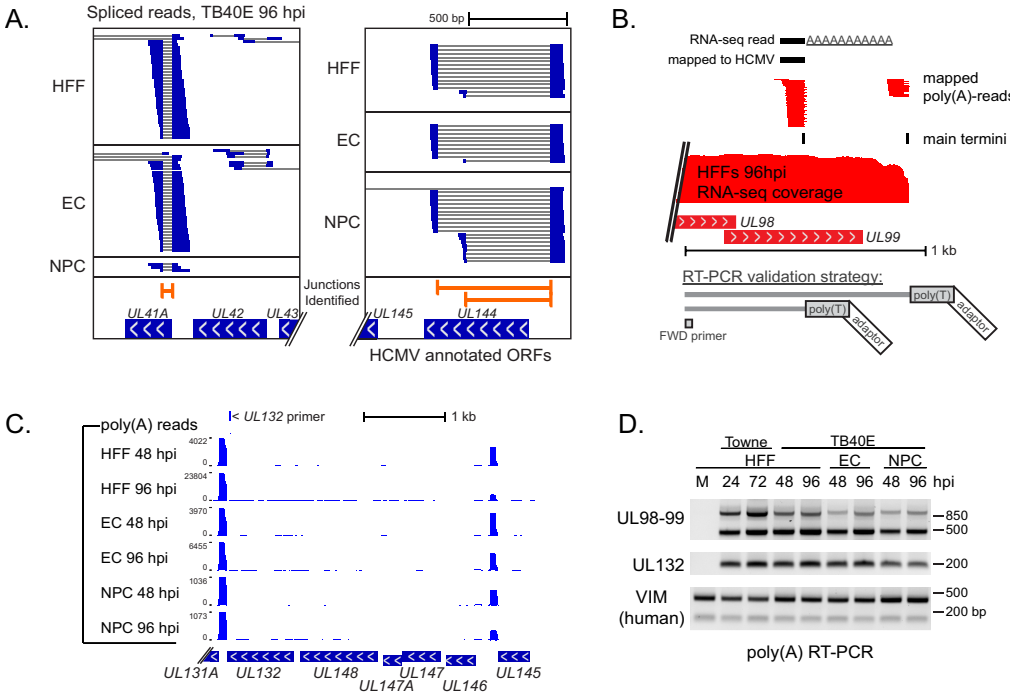

Figure 3.2: Consistent HCMV splicing and polyadenylation across cell types

(A) Examples of newly identified HCMV splicing junctions. The exon-intron junctions shown and highlighted in orange were supported by a minimum of 100 reads in at least one cell type. All spliced RNA-seq reads shown are from 96 hpi datasets. Blue blocks are the exon-aligning portion and the thin lines represent introns. Only alignments supported by more than five reads are shown for visualization purposes, with the exception of NPC coverage of UL41A, where all spliced alignments are shown due to lower expression. (B) Strategy for transcriptome-wide analysis of HCMV polyadenylation sites, using poly(A)-containing RNA-seq reads. Representative analysis from 96hpi HFFs is shown for the UL98 and UL99 ORFs. Poly(A)-read alignments with greater than 10 reads of support are shown, and the RNA-seq coverage track shown is log-transformed. Poly(T)-adaptor reverse primers are not drawn to scale. (C) Example of an HCMV poly(A) site found consistently across different infected cell types. Poly(A)-containing RNA-seq reads mapping to the negative strand of the HCMV genome are shown in blue over a region spanning the UL131A locus, defining the 3′-end of UL132 and other co-terminal RNAs. A neighboring site for UL145 is shown for comparison (D) RT-PCR assessment of HCMV 3′-termini identified from RNA-seq poly(A) reads. The human VIM gene was used to ensure equivalent loading. A universal primer (illustrated in panel B) was used in conjunction with gene-specific primers proximal to the poly(A) site of interest (the UL98-99 and UL132-specific forward primers are shown in panels B and C, respectively). ‘M’ denotes mock-infected.

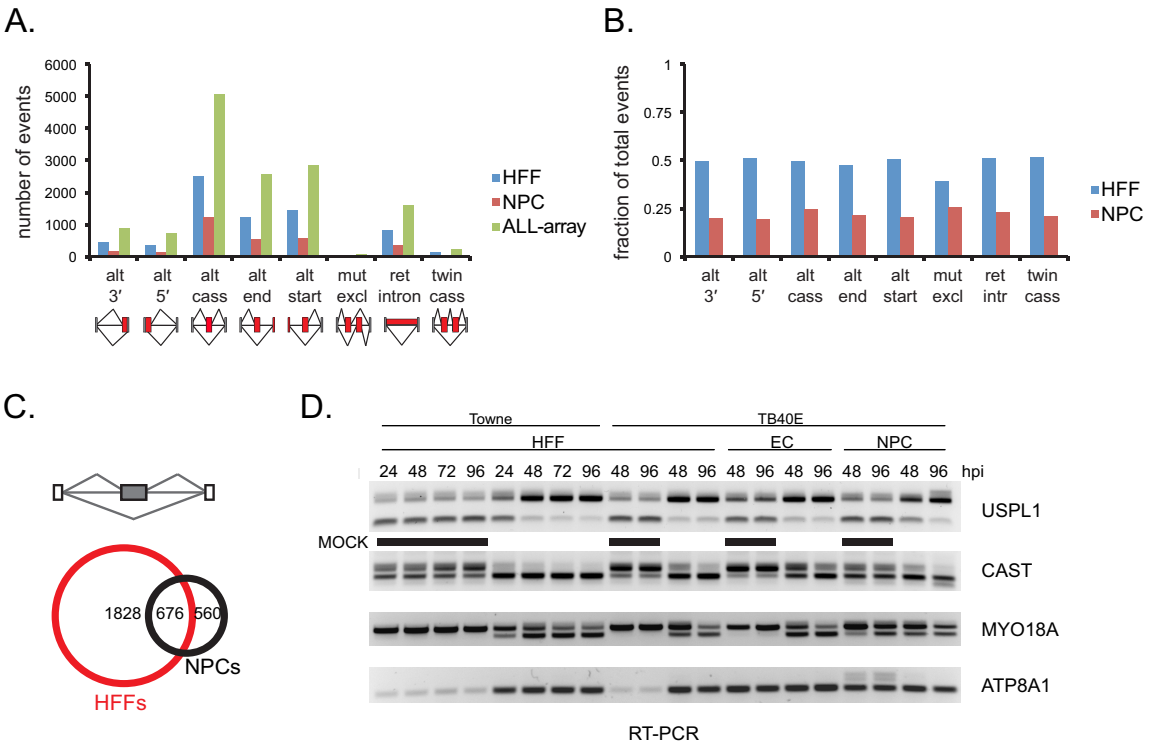

Figure 3.3: Disrupted host alternative splicing patterns during HCMV infection

(A) Numbers of infection-altered splicing events in NPCs and HFFs, and their distribution across different event types. Splicing-sensitive microarray analysis was performed for HCMV Towne-infected HFFs at 72 hpi and TB40E-infected NPCs at 96 hpi. The distribution of all splicing events measured by the microarray is also shown. Schematics of each alternative splicing event type are shown below the chart. Abbreviations used are: alt, alternative; cass, cassette; mut, mutally; ret, retained. (B) Fractions of each alternative splicing event type affected in HFFs and NPCs upon infection. (C) Overlap of alternative cassette exon splicing events between infected HFFs and NPCs. (D) RT-PCR analysis of alternative cassette events as the infection progresses. Primers were directed against exonic regions flanking alternatively included exons. *USPL1* and *CAST* were differentially spliced during infection in all conditions investigated, and the remaining examples apply to only a subset of the conditions.

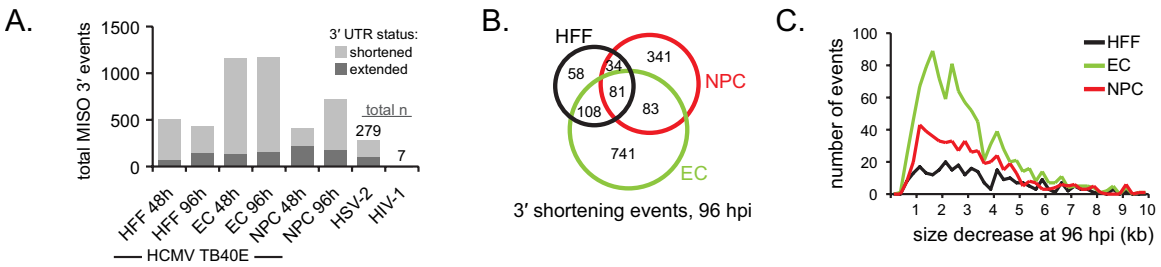

Figure 3.4: Shortening of host 3′UTRs during herpesvirus infections

(A) Identification of altered 3′UTR isoform usage in a range of viral infection conditions. RNA-seq datasets from the listed HCMV TB40E conditions were used to determine distal and proximal 3′UTR regions, and analysis of differential isoform coverage was conducted with the MISO algorithm. HSV-2 (HFFs, 8 h) and HIV-1 (T-cells, 5 d) infections were included for comparison. The total number of significant events detected from the HSV-2 and HIV-1 datasets are indicated (B) Overlap of genes exhibiting 3′UTR shortening among different HCMV-infected cell types. The comparison was made based on Ensembl gene IDs, and the circles were drawn at different sizes but not to scale. (C) Distribution of length changes for 3′UTR shortening events during HCMV infection. (D) Examples of shifted alternative poly(A) preferences towards proximal cleavage sites in both HCMV- and HSV-2-infected cells. The longer, distal-specific part of the extended isoform of each 3′UTR is shaded gray. Locations of canonical putative CPEB1 recognition sites (CPE; specifically (U)UUUUAU or UUUUAAU) are indicated by triangles below the UTR isoform annotations. Only CPEB1 sites found within 500 bp of either the proximal or distal cleavage sites are shown. PAS sites (non-canonical sequences noted for both ANKH and PCGF3 proximal 3′-ends) are also indicated and their genomic sequences are specified. (E) qRT-PCR validation of decreased expression of the distal 3′UTR regions of MISO-identified transcripts in both HCMV-infected HFFs and NPCs. Primer locations for proximal and distal measurements are shown in panel D. Expression measurements represent the ratio of distal to proximal Ct values, normalized to 48 h mock-infected cells (Ulitsky et al., 2012). All qPCR measurements were obtained in triplicate, using sample materials independent from those subjected to RNA-seq. Error bars represent the standard error of the mean.

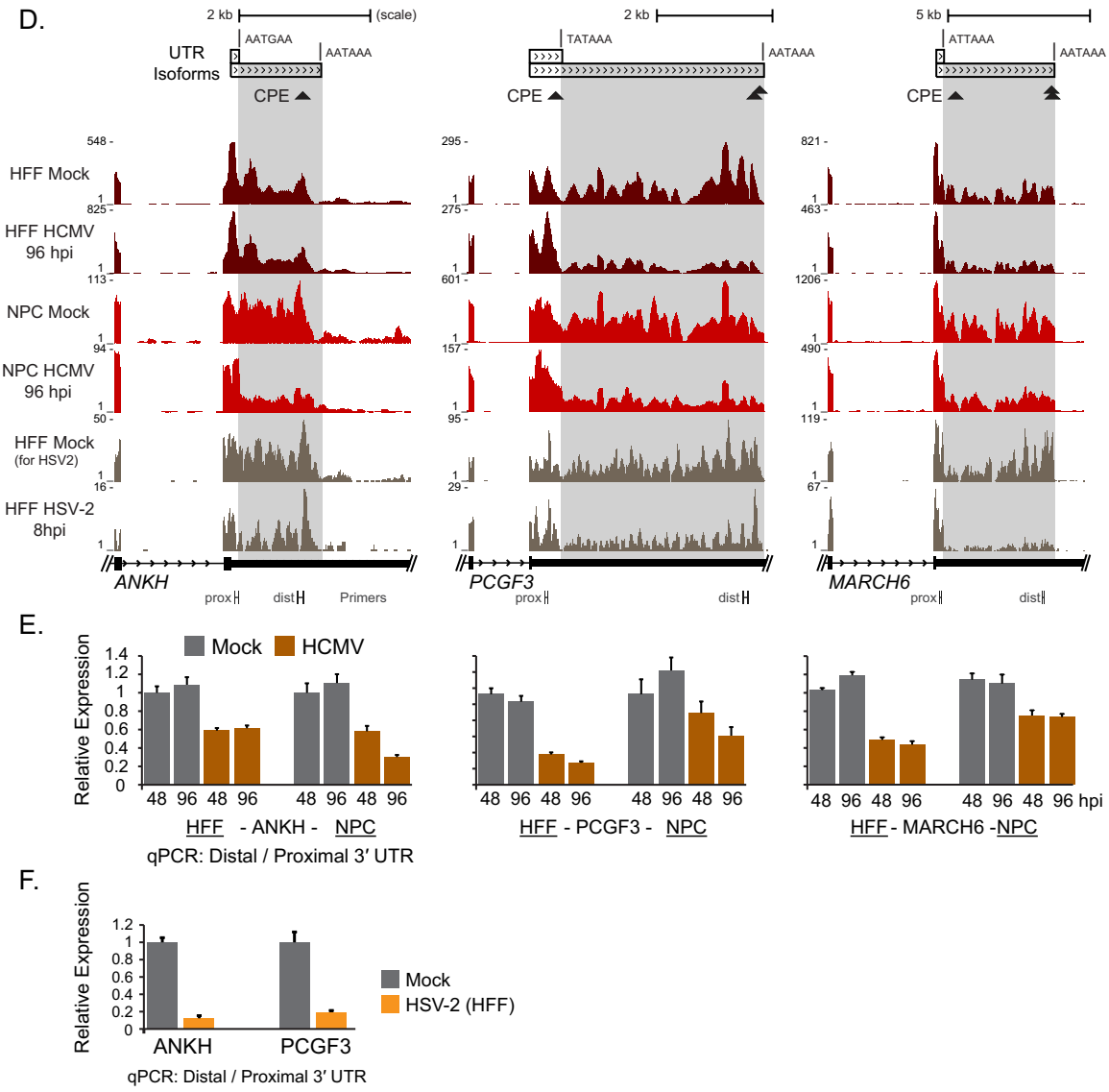

Figure 3.4 (continued): Shortening of host 3′UTRs during herpesvirus infections

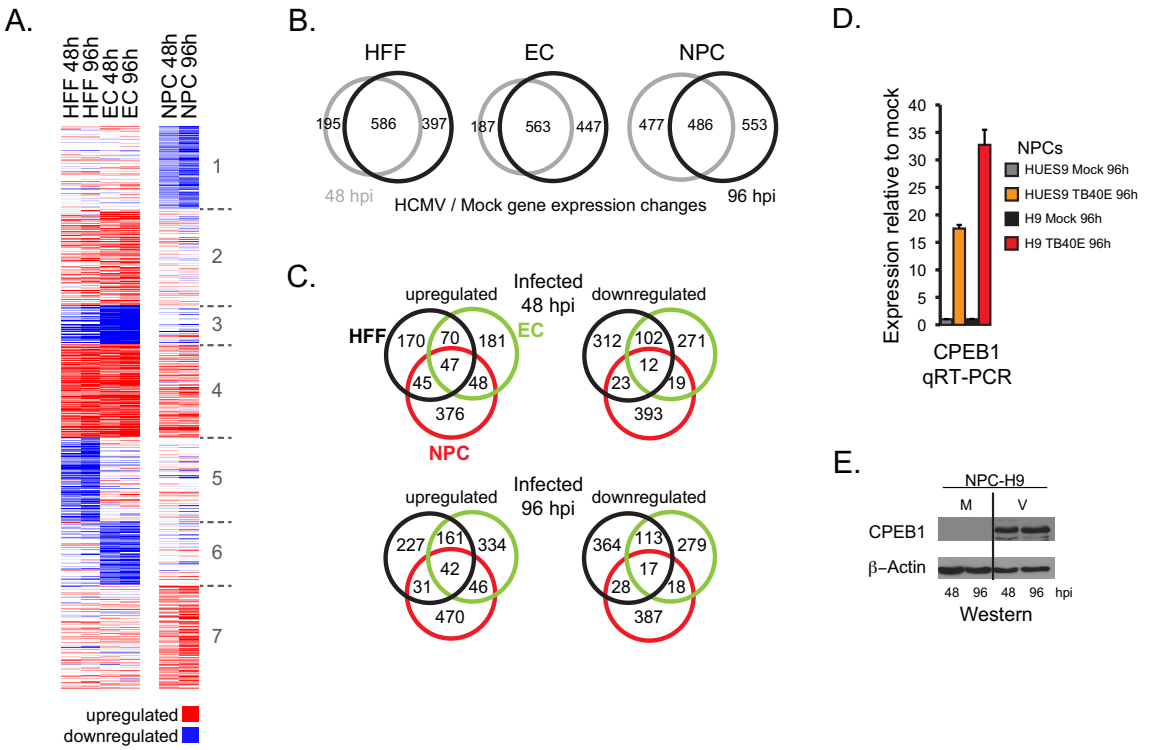

Figure 3.5: Host gene expression analysis and identification of both cell type-specific effects and conserved subsets of changes

(A) Unsupervised k-means clustering of differentially expressed host transcripts across HCMV-infected HFF, EC, and NPC (HUES9-derived) samples. Log-transformed RPKM ratios at 48 hpi and 96 hpi were normalized against mock-infected 48 hpi controls, and only genes with Z-score significance in at least one infection condition are displayed (p-value < 0.05). (B) Overlap of the most significant host expression changes between 48 and 96 hpi within the same cell type. (C) Overlap of host expression changes among different infected cell types. Up and downregulated transcript sets are depicted separately per time-point. (D) qRT-PCR validation of CPEB1 mRNA upregulation in both HUES9 and H9-dervied NPCs at 96 hpi. A PrimeTime qPCR assay probe-set from Integrated DNA Technologies was used for this analysis. (E) Western blot analysis of CPEB1 upregulation in the second NPC line derived from H9 ES cells. M, mock; V, virus-infected with HCMV TB40E, MOI 5. Samples were loaded according to total protein content, and β-actin was used to assess loading efficiency.

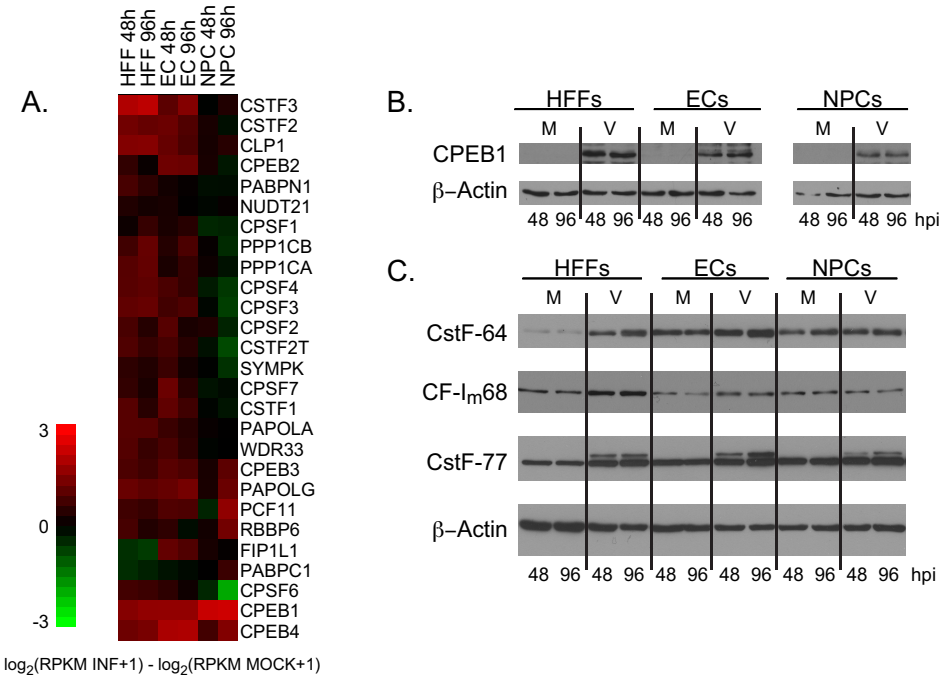

Figure 3.6: Upregulation of CPEB1 in HCMV-infected HFFs, ECs, and NPCs

(A) Hierarchical clustering of infection-over-mock transcript levels for selected host 3′-end processing factors. Fold-changes were calculated from RPKMs and $\log_2$-transformed. (B) Western blot analysis of CPEB1 upregulation across all HCMV infections. (C) Western blot analysis of host 3′-end processing factors from the CF and CstF complexes. M, mock; V, virus-infected with HCMV TB40E, MOI 5. All lanes were loaded according to total protein content. β-actin was used to evaluate consistent loading.

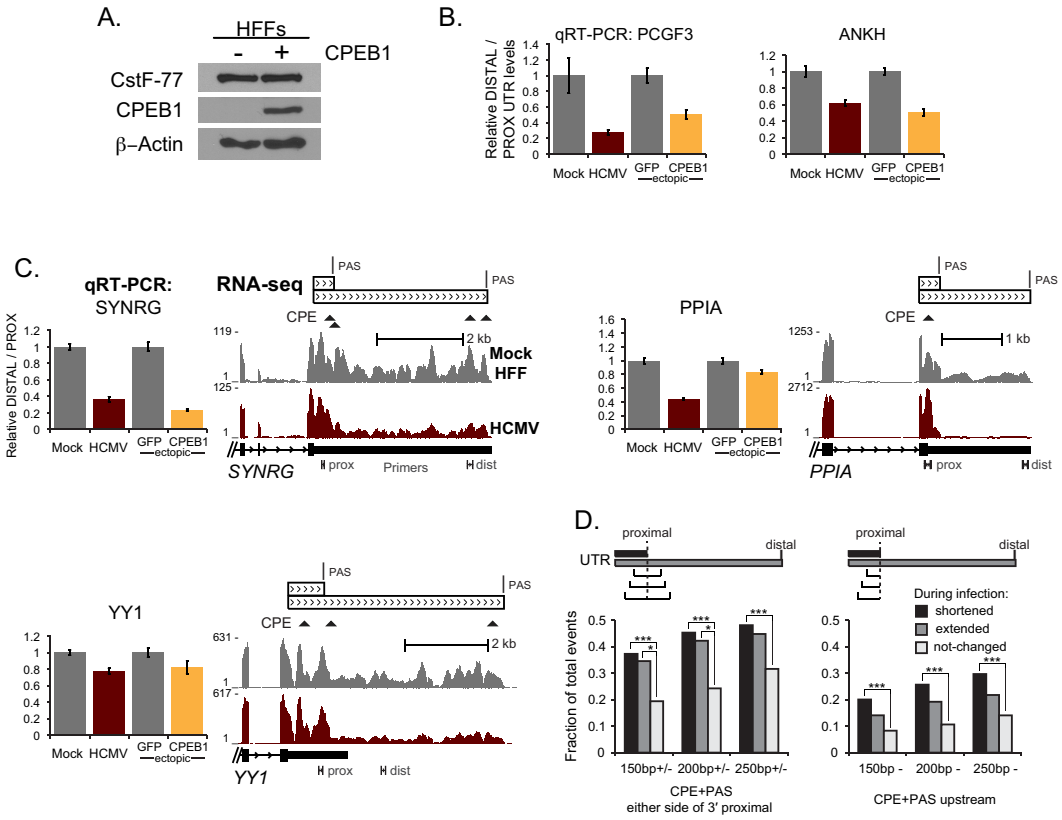

Figure 3.7: Alternative 3′UTRs shortened with ectopic expression of CPEB1 in HFFs

Western blot confirmation of CPEB1 overexpression in transduced HFFs. Cells were analyzed at 5 days post-transduction. CstF-77 abundance was determined to assess for effects on other 3′-end processing factors. (B) Parallel effects of increased CPEB1 on distal 3′UTR isoforms that are affected during HCMV infection. Measurements were obtained as described in Figure 4D. (C) qRT-PCR assessments of additional alternative 3′UTR candidates that were originally identified as shortened upon infection by RNA-seq. Sequencing coverage from mock- and HCMV-infected HFFs at 96hpi for each of the examples is shown for comparison. The location of PAS sequences nearest to 3′-cleavage sites are shown above UTR isoforms, and putative canonical CPEs are indicated by triangles. Only sites within 500 bases of 3′-proximal and distal cleavage sites are shown. (D) Global analysis of CPE-PAS motif co-occurrences surrounding or upstream of proximal 3′-ends from infection-altered events. Alternative 3′UTRs that were analyzed included those that were differential (shortened or extended, and involving the same alternative 3′-termination site in 2 or more infected cell types), or identified as not changing in any of the cell types upon infection, all based on MISO results. 249 3′UTR shortening events, 78 cases of lengthening, and 638 control transcripts were considered. Increments of 150, 200, and 250 bp, were analyzed for the presence of both a canonical CPE ((U)UUUUAU or UUUUAAT) and PAS (AAUAAA or AUUAAA) on either side or solely upstream of RNA-seq-defined 3′-proximal terminations sites. Chi-square tests were performed to evaluate the significance of the numbers of CPE-PAS co-occurrences in shortening or extension events vs. non-changing events. *** p-value < 0.0001; * p-value < 0.01.

Table 3.1: HCMV gene expression calculations

RPKM levels listed were normalized by the combined total of HCMV and human-mapping RNA-seq reads. HCMV reads were mapped to TB40E (Genbank KF297339.1). Temporal HCMV gene classes (G.C.) are listed in the final column for ORFs that were at least 10-fold lower expressed in the 96 hpi NPCs vs. 96hpi HFFs or ECs (L=Late DE=Delayed Early).

ORF	HFFTB48	HFFTB96	ECTB48	ECTB96	NPCTB48	NPCTB96	G.C.
UL74	252.8	510.4	76.6	362.9	56.6	69.7	L
UL148A	1113.9	676.1	816.4	788.7	322.2	389.2	
UL34	961.5	1591.1	593.3	805.4	516.2	1741.9	
UL43	2896.5	3060.4	1706.1	2598.2	619.4	1109.9	
UL37	339.2	233.5	403.4	218.7	146.4	222.0	
UL45	603.3	596.6	719.4	644.6	293.1	265.5	
UL91	566.9	366.7	486.3	414.7	127.2	159.7	
UL102	122.4	316.4	156.5	96.3	77.0	69.3	
UL150A	1252.1	1011.3	1560.4	1162.9	435.4	701.7	
UL105	231.8	69.9	368.5	63.2	90.3	94.0	
UL84	2535.4	4017.3	2140.7	3534.7	659.0	1247.0	
UL82	676.6	7208.6	466.7	6842.3	203.3	677.4	L
UL4	14054.3	7392.4	15956.7	4557.2	1575.3	936.2	
US27	1369.2	1876.1	1181.8	2136.4	352.7	422.6	DE
UL100	1006.0	5911.5	725.0	4853.7	237.9	301.8	L
UL93	628.5	620.2	553.1	675.2	158.1	208.2	
UL138	4000.7	2914.9	4067.3	2675.2	1136.3	1888.4	
UL16	537.9	512.6	829.9	461.6	96.3	139.0	
UL72	587.4	2008.0	260.1	1282.8	139.4	135.8	L
UL95	433.2	595.1	318.5	667.2	173.3	206.1	
US30	1013.6	1836.5	801.6	1990.1	238.8	288.3	DE
RL8A	630.6	351.2	484.7	359.5	207.2	236.0	L
UL23	382.5	408.9	600.9	410.0	235.4	576.0	
RL10	1209.1	1846.6	1537.9	2666.1	187.7	94.0	L
RNA4.9	7034.9	7021.3	10226.8	9415.0	1790.9	8156.2	
UL26	787.0	902.4	802.2	779.9	324.6	617.2	
UL21A	423.9	215.2	1042.8	257.6	345.3	369.4	
UL115	1905.8	2846.9	1352.1	3005.5	373.0	377.2	L
UL136	2277.2	1481.6	2187.0	1386.7	688.3	1036.0	
UL47	88.9	53.4	36.5	81.9	18.5	18.2	
UL145	2213.2	5611.1	3123.2	3699.2	417.2	233.2	L
UL148B	777.8	609.3	589.4	737.9	192.9	198.2	
US10	248.7	141.6	333.4	114.2	198.6	170.7	
TRS1	949.7	345.6	1037.5	415.4	435.0	567.0	
UL148	1638.8	14963.1	1879.4	4516.0	508.6	501.8	L
UL48A	2559.3	3166.6	2538.5	2824.2	758.5	527.9	L
UL31	577.0	859.7	413.1	488.4	119.5	87.8	L

Table 3.1 (continued): HCMV gene expression calculations

ORF	HFFTB48	HFFTB96	ECTB48	ECTB96	NPCTB48	NPCTB96	G.C.
UL5	15873.7	9066.2	18767.3	6594.2	2205.1	1352.0	DE
UL7	221.6	61.0	160.4	51.3	70.6	78.1	
US3	485.1	623.7	1158.4	800.2	558.6	1362.9	
UL121	30.5	87.3	12.1	95.1	4.3	25.5	
UL30A	2672.4	12714.1	3118.8	11659.9	515.1	404.7	L
US17	124.9	77.8	68.5	97.8	64.4	85.1	
RNA5.0	890.4	3089.6	603.0	1939.9	469.4	604.8	
UL147A	622.4	16821.9	826.2	4507.6	368.4	259.5	L
UL19	254.2	167.9	369.6	231.3	61.6	73.3	
UL80.5	1700.7	3056.0	731.4	2539.9	262.8	337.3	L
UL74A	471.1	1177.8	630.3	1375.5	162.8	302.8	
UL86	610.9	898.8	231.3	1037.1	130.1	217.9	L
US21	118.3	64.3	76.7	67.7	55.5	69.7	
RNA1.2	15236.0	11729.4	17858.6	22134.5	4726.4	2945.4	DE
US12	3265.7	2181.6	4423.5	2093.5	1129.4	1153.0	
UL99	2630.7	10520.0	2097.7	8081.9	659.6	1174.6	L
UL1	583.6	721.7	381.8	400.9	183.9	79.3	
UL25	861.9	3824.0	623.7	4124.9	185.6	175.8	L
US6	328.2	542.0	732.6	664.2	219.8	420.5	
UL97	699.6	928.2	522.2	1028.2	257.2	326.4	
UL52	818.3	650.7	736.0	671.8	225.2	175.9	
UL22A	3078.8	24775.9	4138.4	21009.5	2565.2	5564.0	
UL13	1619.4	729.6	1830.6	772.2	292.2	225.3	
UL79	144.3	135.7	70.0	129.6	36.7	82.4	
UL27	251.5	157.6	218.8	117.1	99.3	168.3	
UL49	1520.2	1384.8	1648.9	1293.4	517.3	319.5	
UL119	932.0	1262.9	666.6	1306.7	189.4	174.4	DE
UL29	94.9	74.9	74.9	46.7	45.5	94.6	
US15	1034.1	282.8	1432.5	344.4	506.0	474.4	
UL73	1537.1	5681.5	1996.8	7196.8	298.0	434.3	L
UL76	186.6	395.8	111.4	307.9	32.5	48.9	L
US32	804.5	1498.5	920.5	1584.4	223.2	303.9	L
UL123	193.2	72.0	184.5	61.4	197.7	941.7	
US33A	658.7	585.2	3014.0	566.8	620.8	1414.2	
UL77	246.2	487.0	149.3	373.9	47.6	76.4	DE
UL54	817.8	841.3	837.9	802.4	372.8	157.4	DE
US9	689.6	274.8	1084.5	328.4	395.5	1136.1	
UL114	1219.6	1809.3	854.0	2096.5	233.4	165.8	DE
US34A	928.1	1120.2	4051.6	833.6	637.8	1972.9	
US29	734.0	1528.4	515.8	1727.8	158.2	145.8	L
UL46	639.1	596.6	889.9	695.1	379.7	419.5	
UL36	433.1	231.1	545.5	226.3	263.9	531.6	
US19	1926.3	914.3	1857.0	956.1	417.4	469.3	
UL89	239.7	193.2	198.8	290.3	62.2	54.9	

Table 3.1 (continued): HCMV gene expression calculations

ORF	HFFTB48	HFFTB96	ECTB48	ECTB96	NPCTB48	NPCTB96	G.C.
UL141	832.0	2547.4	666.4	2219.1	197.8	402.9	
UL55	1199.9	1425.3	1234.2	1417.6	470.4	264.8	DE
UL20	398.9	268.6	518.5	269.0	113.7	138.0	
US23	454.4	289.7	413.4	164.4	152.7	242.2	
UL10	807.2	259.8	714.3	255.9	159.5	141.7	
UL33	897.2	1477.6	377.7	436.0	384.2	703.3	
RL11	2570.3	2529.5	2820.6	3400.0	534.1	332.0	L
UL133	1272.4	701.1	1024.0	734.8	423.0	619.0	
UL80	1082.5	1900.1	475.2	1611.4	178.5	231.2	L
UL71	267.3	266.4	278.4	235.8	68.6	166.4	
UL14	297.0	307.8	248.6	237.1	93.0	65.2	
UL24	250.1	268.1	352.3	230.8	102.3	218.6	
UL146	105.2	9643.2	154.0	2166.9	16.1	18.8	
RNA2.7	63362.9	79190.1	84643.7	122344.0	60354.5	97089.5	
US26	68.8	40.7	96.1	47.9	60.3	91.0	
UL104	237.1	391.8	110.2	431.9	32.0	34.4	DE
RL9A	255.5	148.0	214.1	137.0	112.9	160.1	L
UL140	737.9	2463.6	356.2	2625.3	148.0	294.7	L
UL17	1360.0	1360.6	1734.7	1434.9	561.5	655.2	
UL92	827.1	548.5	706.6	624.4	162.4	198.0	
UL150	388.5	318.4	250.5	449.3	136.6	141.7	
UL130	289.4	665.4	345.8	434.4	67.5	199.5	
UL38	688.0	444.1	777.6	435.9	246.1	332.1	
UL69	644.3	1258.3	577.0	802.0	334.1	574.8	
UL142	27.1	543.3	17.6	186.9	8.8	42.8	L
UL132	2744.1	20396.4	3461.2	6124.6	705.0	783.2	L
UL83	438.7	6052.6	208.0	6286.2	114.4	493.4	L
UL48	301.9	181.2	228.9	252.2	90.7	101.2	
US11	186.0	96.3	235.3	84.4	181.4	142.4	
US34	711.1	673.9	3042.9	503.9	641.7	1711.3	
UL8	291.9	68.8	252.8	63.0	80.0	85.8	
US31	1074.4	1907.9	1104.1	2021.8	266.7	314.8	
UL57	116.9	84.8	74.5	94.1	62.0	28.4	
UL144	987.5	5658.2	966.4	2640.7	400.2	144.1	L
UL42	4798.8	5380.9	3027.6	4557.6	1108.7	1952.8	
RL13	5673.9	5667.7	5765.3	5872.6	1042.2	666.1	
UL51	460.7	208.8	433.7	172.6	246.9	164.4	
UL32	399.1	1242.0	361.1	1406.3	106.0	75.3	DE
UL148C	302.5	251.2	248.2	199.7	127.3	253.5	
UL117	985.5	988.8	796.1	1098.2	237.3	263.2	
UL111A	176.8	273.5	145.7	207.7	54.5	101.5	
UL135	1482.3	818.8	1399.6	783.4	471.7	701.3	
UL78	1597.7	1854.9	1298.4	1485.4	376.8	295.4	DE
UL9	673.7	131.0	662.3	118.7	144.7	141.4	

Table 3.1 (continued): HCMV gene expression calculations

ORF	HFFTB48	HFFTB96	ECTB48	ECTB96	NPCTB48	NPCTB96	G.C.
UL85	1891.5	3194.7	847.1	3000.1	295.5	420.3	L
UL116	2076.4	2436.0	1684.5	2656.2	442.3	482.5	L
UL139	499.0	2149.2	259.8	2395.0	128.1	348.4	L
UL103	508.4	862.3	278.9	990.7	87.4	151.5	L
US22	1116.2	643.8	966.3	473.4	349.0	620.9	
UL35	287.2	787.6	192.4	939.6	149.2	262.5	
UL6	1072.1	584.0	1378.3	392.5	235.7	186.7	
UL15A	361.7	451.8	319.8	314.7	88.9	69.7	L
UL96	879.9	1139.9	653.2	1353.3	317.6	389.9	
UL124	340.3	2463.0	293.7	1525.2	80.6	166.1	DE
UL2	271.8	233.1	220.8	277.5	83.4	67.7	
UL122	381.9	1012.6	248.0	808.8	152.6	949.1	
US7	168.1	146.4	200.3	161.4	121.4	344.2	
UL11	1256.3	517.0	1198.6	541.5	253.3	196.7	
UL148D	1138.3	881.6	1380.0	806.4	387.2	644.2	
UL147	309.3	15590.9	416.9	3889.4	357.2	198.0	L
UL70	221.4	453.6	152.0	375.4	121.3	177.2	
RL1	167.3	421.5	98.7	409.8	37.5	63.6	DE
US24	363.7	276.0	426.4	183.6	112.8	163.2	
UL50	754.8	698.8	706.5	574.4	249.3	142.8	
UL30	1519.8	6531.6	1787.8	7217.1	247.5	251.3	L
US13	1873.6	1024.3	2483.4	894.8	839.8	863.8	
UL120	51.2	101.1	45.3	101.2	11.3	49.9	
UL18	86.4	58.7	111.5	64.7	32.9	47.1	
US8	1175.3	542.7	2072.7	599.2	587.5	1929.2	
UL131A	70.2	132.1	87.0	77.6	27.5	56.9	
UL87	127.2	77.0	95.6	92.2	29.2	42.6	
US18	4735.4	3124.6	5145.0	3743.3	1077.7	1070.6	
UL41A	4983.7	6216.7	2960.3	5577.1	1029.8	1844.5	
UL44	2966.7	3061.0	1755.1	2816.1	730.1	1306.1	
UL94	1455.0	2907.8	1392.5	2763.4	373.7	535.6	L
US28	3317.2	2798.7	3635.5	3191.7	772.5	665.6	
UL56	732.8	403.7	725.7	447.8	227.0	119.8	
US20	2148.5	1053.1	2059.4	1006.1	416.9	487.0	
UL53	1838.0	1965.0	1639.6	2106.9	382.9	319.1	DE
UL98	1162.8	3009.7	926.0	2537.7	394.9	635.2	
UL112	499.7	699.1	401.1	465.4	345.0	571.5	
UL88	161.3	96.3	114.9	116.0	34.9	45.6	
UL75	1445.7	2991.3	651.6	2485.9	418.8	720.6	
UL40	7184.5	8535.4	6420.1	8727.2	1766.9	2752.3	
RL12	6227.0	5773.2	6182.9	6110.1	1155.7	769.8	L
US14	1395.1	635.3	1998.6	625.3	682.7	746.9	
US16	737.4	206.4	980.5	260.4	412.8	446.9	

Table 3.2: HCMV splicing junction locations

Comparison of HCMV splicing events detected in Gatherer et al. 2011, Stern-Ginossar 2012, and RNA-seq datasets from this study. Novel junctions to this study are highlighted. Read counts of support are listed for each study. For this comparison HCMV mapping was performed against the Merlin reference (Genbank AY446894.2). Annotations are only listed for novel junctions detected with coverage support greater than 250 spliced reads.

Merlin Donor	Merlin Acceptor	strand	Gatherer	Stern-Ginossar	Stark	Annotations
27528	27612	+	131826	1+	216215	
193542	193654	+	3118	1+	12079	
176523	176402	-	2612	1+	9951	
162673	162831	+	316	1+	6669	
174929	174107	-	414	1+	6488	
97900	97993	+	21	0	6244	
174020	173903	-	299	1+	5262	
166437	166523	+	621	1+	5164	
168891	168802	-	831	1+	5022	
166062	166161	+	160	1+	4637	
49927	49823	-	240	1+	4139	
163109	163391	+	135	1+	3795	
168875	169009	+	160	0	3564	
173719	173548	-	222	1+	3347	
169135	**170082**	**+**	**0**	**0**	**2923**	UL119/120/121as
227915	227631	-	15	0	2328	
107496	108854	+	2574	1+	2155	
166062	166149	+	62	1+	2020	
165821	165960	+	102	1+	1897	
161185	161262	+	94	1+	1867	
27524	27612	+	0	1+	1771	
105812	108854	+	592	1+	1700	
163109	163513	+	63	1+	1587	
204854	204446	-	29	1+	1549	
192428	193172	+	286	1+	1536	
202754	203162	+	26	1+	1526	
103285	108854	+	641	1+	1520	
165821	165963	+	76	0	1493	
192428	193169	+	224	1+	1486	
177911	177802	-	200	1+	1457	
173719	172175	-	302	1+	1410	
103436	108854	+	295	1+	1174	
170157	174081	+	151	0	1173	
36278	36131	-	10	1+	1094	
101641	108854	+	203	1+	1072	
138810	134906	-	23	1+	1027	
170175	174081	+	47	1+	1007	
101294	101176	-	70	0	946	

Table 3.2 (continued): HCMV splicing junction locations

Merlin Donor	Merlin Acceptor	strand	Gatherer	Stern-Ginossar	Stark	Annotations
170157	170290	+	5	0	923	
16449	16552	+	42	0	880	
176781	176657	-	295	1+	833	
192976	193169	+	8	1+	805	
228332	228420	+	69	1+	733	
191908	192072	+	49	1+	723	
200510	200341	-	29	1+	700	
133811	135253	+	88	1+	690	
51302	51197	-	24	1+	646	
188075	187868	-	100	1+	616	
203017	**203162**	**+**	**0**	**0**	**599**	US7as
107740	**108075**	**+**	**0**	**0**	**594**	super-acceptor
170175	170290	+	3	0	588	
135555	134906	-	82	0	586	
192976	193172	+	18	1+	563	
100313	108854	+	52	1+	556	
176395	**176495**	**+**	**0**	**0**	**538**	UL128as
174242	174107	-	21	1+	468	
227722	227631	-	9	0	425	
200822	200669	-	16	1+	412	
174929	**174645**	**-**	**0**	**0**	**409**	UL123/122
187468	187368	-	17	1+	394	
108155	108854	+	788	1+	387	
103622	108854	+	90	1+	376	
100841	108854	+	30	0	374	
102291	108854	+	189	1+	373	
77623	77504	-	14	1+	370	
161462	161546	+	24	1+	368	
4732	**4439**	**-**	**0**	**0**	**365**	RNA2.7
200312	**202664**	**+**	**0**	**0**	**362**	US3/6/7as
110763	**111111**	**+**	**0**	**0**	**359**	super-acceptor
182846	**182273**	**-**	**0**	**0**	**353**	UL144
107496	**108075**	**+**	**0**	**0**	**352**	super-acceptor
149567	149833	+	0	1+	344	
43062	43361	+	17	0	331	
101416	108854	+	122	0	311	
169129	**170082**	**+**	**0**	**0**	**305**	UL119/120/121as
106577	108854	+	45	1+	304	
106352	108854	+	17	0	301	
171874	174081	+	66	1+	291	
188528	188776	+	2	0	291	
188075	187368	-	7	0	284	
9034	**8196**	**-**	**0**	**0**	**279**	RL10as
182846	**182417**	**-**	**0**	**0**	**275**	UL144
228433	228244	-	3	0	274	

Table 3.2 (continued): HCMV splicing junction locations

Merlin Donor	Merlin Acceptor	strand	Gatherer	Stern-Ginossar	Stark	Annotations
204854	194235	-	1	0	263	
54638	**54589**	**-**	**0**	**0**	**255**	UL41A
74914	74818	-	22	1+	254	
15675	**8196**	**-**	**0**	**0**	**254**	super-acceptor
202256	**202664**	**+**	**0**	**0**	**251**	US6-7as
101731	108854	+	32	1+	246	
189044	193169	+	15	0	237	
101993	108854	+	34	1+	234	
112452	**111611**	**-**	**0**	**0**	**234**	
115656	115880	+	15	1+	225	
166295	**168842**	**+**	**0**	**0**	**219**	
193464	193654	+	18	0	213	
170329	**170508**	**+**	**0**	**0**	**213**	
126321	135253	+	10	0	205	
165887	**165963**	**+**	**0**	**0**	**203**	
101186	108854	+	60	0	200	
189044	193172	+	10	1+	197	
107740	108854	+	142	1+	194	
96147	**96186**	**+**	**0**	**0**	**193**	
168875	**169022**	**+**	**0**	**0**	**192**	
125747	135253	+	4	0	191	
166054	166161	+	4	0	187	
179161	**193169**	**+**	**0**	**0**	**185**	
228502	228244	-	3	0	184	
109370	111111	+	1	0	183	
228393	228244	-	3	0	181	
101839	108854	+	28	0	179	
160957	**158166**	**-**	**0**	**0**	**178**	
100158	108854	+	18	0	175	
161529	161440	-	2	0	171	
101258	108854	+	64	0	169	
73069	72973	-	6	1+	169	
106097	**108854**	**+**	**0**	**0**	**168**	
27145	27348	+	1	0	164	
185302	**184611**	**-**	**0**	**0**	**164**	
161295	161377	+	6	1+	158	
122822	135253	+	4	0	155	
9499	8196	-	16	0	153	
18632	**8196**	**-**	**0**	**0**	**153**	
8721	**8196**	**-**	**0**	**0**	**152**	
171421	**170871**	**-**	**0**	**0**	**152**	
138810	**138622**	**-**	**0**	**0**	**145**	
111725	**115880**	**+**	**0**	**0**	**145**	
174934	**174107**	**-**	**0**	**0**	**144**	
4321	**4068**	**-**	**0**	**0**	**144**	

Table 3.2 (continued): HCMV splicing junction locations

Merlin Donor	Merlin Acceptor	strand	Gatherer	Stern-Ginossar	Stark	Annotations
76457	**8196**	**-**	**0**	**0**	**142**	
106917	108854	+	51	1+	141	
9477	**8196**	**-**	**0**	**0**	**140**	
166295	**170082**	**+**	**0**	**0**	**139**	
118090	**118558**	**+**	**0**	**0**	**139**	
170376	170508	+	18	0	138	
102272	108854	+	18	0	132	
77623	77543	-	9	0	132	
41382	43361	+	13	0	129	
101766	108854	+	5	0	129	
201867	202664	+	3	0	129	
19820	**8196**	**-**	**0**	**0**	**129**	
165887	**165960**	**+**	**0**	**0**	**129**	
138810	136518	-	1	0	125	
137959	**134906**	**-**	**0**	**0**	**123**	
168624	**168839**	**+**	**0**	**0**	**121**	
163109	163656	+	4	0	119	
134991	**135253**	**+**	**0**	**0**	**119**	
3522	**3279**	**-**	**0**	**0**	**118**	
179161	**193172**	**+**	**0**	**0**	**115**	
25037	**8196**	**-**	**0**	**0**	**114**	
77369	72973	-	6	0	108	
204854	**203315**	**-**	**0**	**0**	**106**	
101985	108854	+	12	0	105	
127769	135253	+	5	0	105	
77369	51197	-	16	0	100	
190139	190313	+	11	0	99	
76457	72973	-	13	0	93	
173683	174081	+	11	0	85	
8229	7396	-	27	0	81	
103436	135253	+	12	0	67	
102500	108854	+	11	0	67	
101282	108854	+	15	0	64	
100034	108854	+	13	0	64	
105835	108854	+	25	1+	53	
102029	108854	+	18	0	51	
107775	108854	+	14	1+	32	
170175	170508	+	365	1+	17	
3203	3309	+	0	1+	16	
4179	3821	-	20	0	12	
157658	161055	+	4	1+	12	
160957	156413	-	2432	1+	11	
224024	224100	+	1557	1+	11	
3005	3122	+	26	0	7	
3481	3748	+	13	0	7	

Table 3.2 (continued): HCMV splicing junction locations

Merlin Donor	Merlin Acceptor	strand	Gatherer	Stern-Ginossar	Stark	Annotations
183500	193169	+	11	1+	5	
179161	181563	+	0	1+	5	
170157	170508	+	154	0	4	
183500	193172	+	13	1+	4	
183179	193169	+	11	0	1	
177289	181563	+	0	1+	1	
158046	161055	+	0	1+	1	
169821	170082	+	647	1+	0	
43441	43562	+	252	1+	0	
106020	108854	+	248	1+	0	
108510	108854	+	147	1+	0	
108406	108854	+	82	1+	0	
169674	170082	+	77	1+	0	
12992	8196	-	51	1+	0	
169135	169507	+	34	0	0	
183244	193172	+	33	1+	0	
183244	193169	+	32	1+	0	
7610	7396	-	24	0	0	
27652	27761	+	22	0	0	
7610	7461	-	20	1+	0	
134991	165963	+	19	0	0	
10965	8196	-	17	0	0	
170073	169796	-	17	0	0	
11876	8196	-	16	1+	0	
183244	108854	+	15	0	0	
180803	181563	+	15	1+	0	
12654	8196	-	14	1+	0	
160077	160732	+	11	0	0	
9030	8196	-	0	1+	0	

Table 3.3: HCMV polyadenylation sites

Ranges where HCMV poly(A) sites were detected by RNA-seq, listed in order of genomic location. Most sites could be resolved to a small window size, but for some areas such as the lncRNA2.7 a large number of sites were detected; these sites were overlapped and consolidated into ranges to avoid inflation of the final number of HCMV mRNA termini reported. Coordinates are based on HCMV Merlin. The strand notation follows each coordinate range listed. 50 raw RNA-seq reads were required to constitute poly(A) sites. RNA-seq poly(A)-read counts for each infection condition are indicated.

Coordinates	EC 48h	EC 96h	HFF 48h	HFF 96h	NPC 48h	NPC 96h
2473-2531 -	20577	40001	15938	18916	13796	24215
2583-3167 -	44883	76950	40311	43597	4783	9029
3219-4150 -	187331	376904	172056	181338	102112	181445
4259-4506 -	5539	8715	5164	5005	1570	2683
4492-4526 +	86	115	84	118	183	339
5273-5282 -	367	1807	53	157	0	1
6751-6860 -	14018	22659	14998	9298	12378	7587
7025-7157 -	1937	2773	2204	1296	77	67
7309-7324 -	660	948	564	391	6	4
15101-15123 +	5096	2200	4681	2522	2349	2063
15948-15951 +	84	25	116	47	16	28
21159-21178 +	1052	729	1157	553	631	509
22547-22553 +	546	687	619	799	131	128
23974-23997 +	4564	4407	3902	3649	602	849
26748-26786 -	336	115	188	73	271	275
26802-26804 +	47	43	60	32	28	33
27557-27568 +	11	58	27	66	34	484
28128-28135 -	349	334	309	243	150	366
29102-29107 +	106	766	204	563	58	129
32217-32226 -	227	209	209	200	86	91
32474-32519 +	650	3767	748	2502	57	51
32516-32563 -	857	1144	1001	1117	389	2111
37548-37609 +	232	187	159	243	73	106
37557-37614 -	2343	11061	2207	9319	214	207
40179-40199 +	434	713	605	875	66	36
46393-46414 +	391	746	601	986	64	254
48436-48453 +	315	1316	419	849	414	866
48441-48452 -	481	233	384	174	640	1424
51346-51359 -	865	858	1053	686	791	1029
51423-51437 +	114	189	143	79	105	84
52529-52534 +	239	265	222	110	119	69
53492-53531 -	6779	10563	8140	8792	2003	3246
54671-54682 -	927	1298	960	1246	161	502
68768-68772 -	152	161	107	75	21	8
70416-70442 -	3364	5723	3740	4837	1071	1012

Table 3.3 (continued): HCMV polyadenylation sites						
Coordinates	EC 48h	EC 96h	HFF 48h	HFF 96h	NPC 48h	NPC 96h
71391-71425 +	129	129	162	96	117	138
71396-71423 -	298	206	291	146	43	45
76301-76312 +	173	170	224	137	3	1
77525-77539 -	1698	1816	1703	1816	1336	710
77530-77532 +	54	73	68	59	1	1
79323-79344 +	281	431	297	328	134	125
81684-81689 -	271	545	199	519	62	89
82043-82064 -	1571	2265	1738	2037	283	219
83415-83428 -	82	157	105	127	1	2
92453-92484 -	1	162	8	157	14	68
96393-96416 +	551	644	537	442	37	152
96850-96903 +	183	488	188	188	97	127
97087-97138 +	266	1203	384	314	219	709
97387-97388 +	52	61	36	36	0	4
97547-97569 +	9783	12604	7087	6230	1440	8019
97662-97681 +	12384	15248	9567	9183	130	503
97749-97765 +	213	220	161	158	18	129
98099-98105 +	137	163	114	121	10	23
98182-98197 +	763	611	601	591	5	60
98377-98381 +	72	80	17	36	0	4
98514-98576 +	83	107	67	28	14	234
99197-99228 +	2459	2834	1457	1671	237	1154
99207-99241 -	31	55	31	132	255	560
99448-99469 -	4	13	4	15	56	106
99520-99527 -	125	133	134	191	35	90
101059-101065 -	304	569	422	830	54	74
105836-105850 -	343	1429	693	1879	118	122
109069-109092 +	3903	13187	3313	7692	1182	2145
109070-109099 -	433	1947	1095	2202	467	817
115775-115790 +	499	613	530	509	43	44
118646-118708 -	295	5000	530	4707	386	1420
118701-118709 +	129	606	336	633	519	854
121509-121511 -	8	382	20	272	1	0
122442-122455 +	278	151	210	100	71	144
122459-122484 -	393	740	536	744	352	656
122552-122561 -	434	763	563	952	75	143
123869-123898 -	1645	3485	2251	3267	588	1598
130795-130803 -	221	403	278	238	102	127
133082-133084 -	59	130	116	70	2	2
138717-138751 +	2225	5228	2247	4244	450	802
139189-139234 +	2194	4024	2259	4091	807	1155
141708-141737 +	162	437	308	401	73	93

Table 3.3 (continued): HCMV polyadenylation sites						
Coordinates	EC 48h	EC 96h	HFF 48h	HFF 96h	NPC 48h	NPC 96h
145971-146055 +	2434	13142	2868	11156	193	575
146420-146430 -	46	413	77	441	54	115
146427-146473 +	151	741	302	1013	894	1369
148730-148731 -	23	38	37	70	1	1
149238-149250 -	1214	1939	1391	2891	264	322
150554-150566 +	593	357	434	1470	210	366
150587-150597 -	107	401	222	382	5	17
156094-156119 -	1031	5591	2245	3911	233	194
156179-156180 +	70	4	62	15	2	2
156303-156321 +	234	31	141	36	11	12
157505-157543 -	219	946	317	1543	224	409
159244-159272 -	219	797	244	1197	221	397
159630-159644 -	175	449	141	883	91	110
159722-159776 -	150	564	213	900	235	403
160590-160597 -	34	117	48	207	31	28
161753-161764 +	203	405	248	369	106	241
164154-164182 -	357	651	422	678	4	2
164251-164268 +	214	264	313	313	561	1136
164984-164989 -	946	2616	1307	2332	302	408
168158-168169 +	87	348	137	270	7	6
170583-170592 -	132	939	312	929	70	861
172207-172217 -	176	82	220	61	219	936
174647-174655 +	208	1325	262	1744	159	312
176150-176170 -	1538	3088	1539	3760	736	2402
176455-176456 -	0	0	0	0	2	69
178128-178158 -	3969	6454	4021	23798	1036	1068
178231-178239 +	90	152	86	547	24	130
181283-181290 -	20	182	27	307	14	5
181451-181482 -	3160	3823	2750	5027	562	347
182118-182153 +	191	158	161	83	113	213
184307-184414 -	860	3356	1076	3209	474	941
185435-185438 -	51	172	91	163	3	4
186137-186157 +	317	525	215	657	9	38
187085-187120 -	4350	3751	4948	3339	2378	4596
187098-187105 +	438	617	244	1333	92	110
187424-187443 -	3325	3085	3861	2689	746	1280
188773-188775 -	48	33	70	24	0	0
193394-193398 +	53	36	59	30	0	0
194335-194351 +	3255	2947	2632	2012	1330	2160
199270-199279 +	198	41	138	45	81	89
200103-200111 -	187	215	86	163	238	750
203399-203470 -	211	71	173	51	723	2259
205262-205287 +	225	205	157	145	54	51

Table 3.3 (continued): HCMV polyadenylation sites						
Coordinates	**EC 48h**	**EC 96h**	**HFF 48h**	**HFF 96h**	**NPC 48h**	**NPC 96h**
205276-205285 -	384	188	306	152	448	401
206737-206771 -	978	633	1050	541	818	867
210499-210525 +	344	404	338	239	156	257
212426-212434 -	2163	2081	2385	1508	350	377
212468-212481 +	658	567	521	435	387	474
215993-216035 -	216	142	390	165	676	1476
215998-216095 +	65	57	90	51	94	113
219775-219782 -	282	138	229	188	60	56
221875-221882 +	90	24	90	29	15	5
226523-226533 +	3532	4118	3443	2608	505	512
226594-226605 -	134	55	75	40	30	15
227432-227450 -	427	301	280	142	98	91
228244-228260 +	525	1761	718	1175	34	55
228922-228954 +	91	109	77	114	35	71
230199-230210 +	3351	6005	2939	3798	699	910
231504-231530 +	2958	883	922	995	1474	4037
232365-232375 -	195	114	193	98	224	423

Table 3.4: Significantly changing host alternative splicing events

Significant events from splicing-sensitive microarray analysis of HFFs and NPCs are listed. The type of splicing event is listed after the gene name and a numerical notation. The separation score (sepscore) threshold set was > 0.5 or < -0.5 in at least one cell type.

SPLICING_EVENT(type of event indicated) ***hg18-based***	**HFF sepscore**	**NPC sepscore**
TOP 100 events for HFFs (+sepscore)		
KCTD6_1_s0e6|alt_start|chr3:58452862-58459479:+	6.398	0
ATP5C1_(37)|alt_cassette|chr10:7884823-7889627:+	6.109	0
UTP11L_(62)|alt_cassette|chr1:38251124-38255926:+	5.777	0
RCAN2_1_s0e9|alt_start|chr6:46324592-46532674:-	5.031	1.594
ADD3_(96)|alt_cassette|chr10:111880234-111883073:+	4.852	0
CD14_1_s0e4|alt_start|chr5:139992961-139993439:-	4.756	1.117
AK310497_1_s0e3|alt_start|chr19:47593165-47681087:+	4.747	0
HEATR7A_(120)|alt_cassette|chr8:145274967-145290661:+	4.694	-0.736
PNPLA2_(140)|retained_int|chr11:813855-813997:+	4.657	-0.739
TPM1_(76_76)|mutually_exc|chr15:61140191-61141466:+	4.52	0
PPP3CB_(30)|alt_cassette|chr10:74868184-74874488:-	4.445	0
MGAT4A_1_s0e41|alt_start|chr2:98583688-98646074:-	4.363	0.732
ARID5B_1_s0e11|alt_start|chr10:63331448-63480652:+	4.345	0
HDAC6_(278)|alt_5|chrX:48545623-48546013:+	4.32	1.617
THNSL2_1_s0e6|alt_start|chr2:88250949-88253772:+	4.313	0
GPRC5C_(780)|alt_5|chr17:69947780-69951551:+	4.283	0
PPAP2C_1_s0e6|alt_start|chr19:239171-242435:-	4.274	0
TMEM108_1_s0e7|alt_start|chr3:134239942-134430799:+	4.258	2.204
FHL1_1_s0e10|alt_start|chrX:135079731-135116231:+	4.227	0
WASF3_(167_176)|mutually_exc|chr13:26144126-26153190:+	4.218	-0.586
RRAGD_(296)|alt_cassette|chr6:90146805-90178283:-	4.201	0
TMEM108_(1410)|alt_cassette|chr3:134430885-134591713:+	4.155	2.571
OLFML3_(152)|alt_cassette|chr1:114323775-114324476:+	4.128	0
PDGFRB_1_s1e9|alt_end|chr5:149473016-149515152:-	4.081	0
MBNL1_(95)|alt_cassette|chr3:153656056-153659749:+	4.076	0
SPOCD1_(116)|retained_int|chr1:32036511-32036629:-	4.065	0
EPB41L4A_1_s48e57|alt_end|chr5:111506036-111532334:-	4.041	0
ACTN1_(81)|alt_cassette|chr14:68414993-68416431:-	3.92	0
GCH1_(129)|retained_int|chr14:54379522-54380468:-	3.919	2.12
G6PC3_(119)|alt_cassette|chr17:39507664-39508203:+	3.884	0
RBM3_(269)|alt_cassette|chrX:48318999-48319645:+	3.876	0.747
C18orf32_(163)|alt_5|chr18:45264139-45267574:-	3.868	0.556

Table 3.4 (continued): Significantly changing host alternative splicing events

SPLICING_EVENT(type of event indicated) *hg18-based*	HFF sepscore	NPC sepscore
MEG3_(34)|alt_cassette|chr14:100371969-100372256:+	3.865	0
FAM46A_(54)|alt_5|chr6:82518614-82518866:-	3.857	0
DIS3L2_(72)|alt_3|chr2:232873073-232902839:+	3.845	0.526
COBL_(171)|alt_cassette|chr7:51078883-51120356:-	3.821	0
SVEP1_(92)|alt_cassette|chr9:112229859-112231243:-	3.795	0
TMEM40_(90)|alt_cassette|chr3:12759006-12765153:-	3.781	0
MBOAT7_(127)|alt_cassette|chr19:59379375-59383882:-	3.767	0.653
AP2A1_(98)|retained_int|chr19:54996690-54996790:+	3.765	0
SELENBP1_(156)|retained_int|chr1:149608654-149608812:-	3.738	0
SNCA_(84)|alt_cassette|chr4:90866834-90962419:-	3.735	1.025
MXRA8_1_s0e5|alt_start|chr1:1281947-1287020:-	3.719	0
SELENBP1_(208)|alt_3|chr1:149608869-149611737:-	3.704	0
ERCC2_(301)|retained_int|chr19:50565330-50565633:-	3.696	0
DPH2_(51)|alt_5|chr1:44208529-44208854:+	3.669	1.044
ANXA1_(121)|alt_cassette|chr9:74956660-74963257:+	3.654	0.729
ZAP70_1_s0e23|alt_start|chr2:97707986-97717414:+	3.654	0
RIN1_(153)|retained_int|chr11:65859298-65859453:-	3.639	0
RAB3IP_1_s0e11|alt_start|chr12:68435430-68464766:+	3.634	1.032
ADD2_(119)|alt_cassette|chr2:70787082-70848525:-	3.604	0
C17orf107_(89)|retained_int|chr17:4744219-4744310:+	3.594	0
SERPINB2_3_s20e38|alt_end|chr18:59721189-59753456:+	3.575	0
PTPRS_2_s8e12|alt_end|chr19:5158932-5224452:-	3.542	1.363
EXOC1_(45)|alt_cassette|chr4:56444851-56451145:+	3.506	-1.404
HIF3A_(146)|alt_cassette|chr19:51499185-51503317:+	3.502	-2.249
RASSF7_(76)|retained_int|chr11:553317-553395:+	3.502	0
PIGW_1_s0e5|alt_start|chr17:31964959-31967055:+	3.488	1.589
BMP1_(36)|alt_3|chr8:22090108-22090418:+	3.47	0
PCDHGC4_1_s27e71|alt_end|chr5:140736258-140854557:+	3.469	0
C8orf31_(130)|alt_cassette|chr8:144192361-144195969:+	3.443	0.692
MCF2L_(118)|retained_int|chr13:112798968-112799088:+	3.437	0.803
REPS1_(81)|alt_cassette|chr6:139283954-139292806:-	3.419	0
ST6GAL1_(657)|alt_cassette|chr3:188239355-188251730:+	3.402	0
ELF5_1_s0e5|alt_start|chr11:34483876-34491906:-	3.393	1.198
TRIM36_(235)|alt_cassette|chr5:114510990-114543570:-	3.367	0
GBP3_(141)|alt_5|chr1:89252550-89252817:-	3.345	0
NPEPL1_(72)|alt_cassette|chr20:56709621-56715585:+	3.334	0.991
LTBP1_1_s0e10|alt_start|chr2:33025895-33213363:+	3.322	0
IL1RAP_2_s24e29|alt_end|chr3:191846325-191858537:+	3.322	0
PHF11_1_s0e7|alt_start|chr13:48967801-48978771:+	3.315	1.076
SLC47A1_(511)|retained_int|chr17:19421419-19421932:+	3.304	2.962

Table 3.4 (continued): Significantly changing host alternative splicing events

SPLICING_EVENT(type of event indicated) *hg18-based*	HFF sepscore	NPC sepscore
FXYD5_(739)|retained_int|chr19:40337552-40338293:+	3.303	-1.156
MEGF6_1_s0e11|alt_start|chr1:3430670-3517919:-	3.302	0
LOC728661_(418)|alt_cassette|chr1:1598148-1613747:-	3.295	0
HYAL2_(828)|retained_int|chr3:50332970-50333800:-	3.287	0
ZNF721_1_s0e9|alt_start|chr4:456490-483442:-	3.286	1.408
CYTH4_(115)|alt_cassette|chr22:36029389-36035198:+	3.28	-2.184
HOXA1_(201)|retained_int|chr7:27101499-27101702:-	3.276	0.767
EMID1_(91_85)|mutually_exc|chr22:27969478-27984805:+	3.272	0
EXOC7_(153)|alt_cassette|chr17:71596996-71598818:-	3.271	2.009
EPS8L1_(92)|retained_int|chr19:60285360-60285454:+	3.27	1.587
OLFML1_1_s3e7|alt_end|chr11:7466222-7489139:+	3.266	0
KIAA0802_(57)|alt_cassette|chr18:8816230-8821604:+	3.244	0.74
CLGN_1_s0e5|alt_start|chr4:141553691-141568232:-	3.242	0
NFASC_(51)|alt_cassette|chr1:203193577-203203999:+	3.236	0.931
FAM19A3_(67)|alt_3|chr1:113067310-113068157:+	3.219	0.829
CPZ_(107)|alt_cassette|chr4:8660052-8664653:+	3.189	0
CXADR_(205)|alt_cassette|chr21:17841382-17859616:+	3.179	0
CALHM2_1_s0e3|alt_start|chr10:105200929-105202648:-	3.175	0
STRA6_(43)|alt_3|chr15:72277212-72281548:-	3.171	0
PKIG_(105)|alt_cassette|chr20:42644786-42651851:+	3.154	0
ROGDI_(55)|alt_cassette|chr16:4790580-4791504:-	3.143	0.516
EXD3_1_s48e58|alt_end|chr9:139323879-139362343:-	3.133	2.048
CASP9_(57)|alt_3|chr1:15706157-15706954:-	3.123	-0.628
TNFRSF19_1_s0e3|alt_start|chr13:23042722-23062288:+	3.098	-0.832
C1S_1_s0e9|alt_start|chr12:6966611-7039403:+	3.096	0
C12orf42_(134)|alt_5|chr12:102396355-102413809:-	3.092	0
TAZ_(42)|alt_cassette|chrX:153301156-153301564:+	3.07	0
ANKRD13D_(106)|retained_int|chr11:66814260-66814368:+	3.058	0.53
TOP 100 events for NPCs (+sepscore)		
TDRD10_(107)|retained_int|chr1:152786668-152786777:+	1.338	4.968
SCRT2_1_s1e10|alt_end|chr20:575267-604112:-	0	4.168
ACTG2_2_s7e22|alt_end|chr2:73983123-74000288:+	2.567	3.332
PPOX_(60)|alt_5|chr1:159403025-159403253:+	0	3.299
TACR1_1_s7e11|alt_end|chr2:75129737-75131885:-	0.677	3.219
SLC47A1_(613)|retained_int|chr17:19422033-19422648:+	2.681	3.127
SLC47A1_(511)|retained_int|chr17:19421419-19421932:+	3.304	2.962
STXBP2_(50)|alt_cassette|chr19:7608072-7609904:+	0.683	2.819
MICB_(128)|alt_3|chr6:31581627-31582027:+	2.223	2.769
CEBPB_(156)|retained_int|chr20:48242153-48242311:+	0	2.716
ARHGAP8_(132)|alt_cassette|chr22:43576708-43589215:+	1.466	2.694
SPAG6_(146)|alt_cassette|chr10:22730212-22745553:+	1.619	2.649

Table 3.4 (continued): Significantly changing host alternative splicing events

SPLICING_EVENT(type of event indicated) *hg18-based*	HFF sepscore	NPC sepscore
PTPN3_1_s9e59|alt_end|chr9:111191333-111256596:-	1.032	2.618
TMEM108_(1410)|alt_cassette|chr3:134430885-134591713:+	4.155	2.571
SPINT2_(171)|alt_cassette|chr19:43447478-43470355:+	1.672	2.567
TTC39A_4_s33e48|alt_end|chr1:51534338-51536050:-	2.017	2.376
BEST3_1_s0e15|alt_start|chr12:68358940-68379463:-	0	2.36
PDE4A_1_s0e24|alt_start|chr19:10426504-10431075:+	0	2.315
ZDHHC23_1_s0e15|alt_start|chr3:115149437-115159899:+	1.895	2.252
SGK3_1_s9e49|alt_end|chr8:67754813-67936811:+	2.695	2.234
FAM3A_(51)|alt_cassette|chrX:153389015-153389813:-	2.597	2.224
TMEM108_1_s0e7|alt_start|chr3:134239942-134430799:+	4.258	2.204
WSB1_(1355)|alt_3|chr17:22658035-22660252:+	-1.387	2.178
ODF3L2_(108)|alt_cassette|chr19:418762-425620:-	0.846	2.148
PRKCH_1_s0e27|alt_start|chr14:60858267-61084213:+	1.602	2.139
ATP13A3_(90)|alt_cassette|chr3:195608134-195615776:-	0	2.129
CPT1A_1_s35e41|alt_end|chr11:68278663-68283612:-	1.215	2.127
GCH1_(129)|retained_int|chr14:54379522-54380468:-	3.919	2.12
NAV2_1_s0e5|alt_start|chr11:19328846-19810605:+	0	2.119
ZDHHC15_2_s7e25|alt_end|chrX:74508785-74587361:-	1.078	2.077
PPFIBP2_1_s0e15|alt_start|chr11:7570938-7588097:+	0	2.074
SLC44A2_1_s0e5|alt_start|chr19:10574132-10597928:+	1.573	2.062
SPAG9_1_s0e11|alt_start|chr17:46474011-46553094:-	0	2.054
EXD3_1_s48e58|alt_end|chr9:139323879-139362343:-	3.133	2.048
FBLN5_1_s0e27|alt_start|chr14:91406482-91479058:-	0.536	2.034
ASPH_(129)|alt_cassette|chr8:62719745-62726162:-	0	2.022
EXOC7_(153)|alt_cassette|chr17:71596996-71598818:-	3.271	2.009
RPL32_1_s0e4|alt_start|chr3:12856840-12858081:-	2.891	1.971
C3orf63_(276)|alt_3|chr3:56632634-56633546:-	-1.138	1.971
OSTF1_(47)|alt_cassette|chr9:76893462-76932304:+	0	1.941
ARHGAP8_1_s24e53|alt_end|chr22:43511315-43561150:+	1.617	1.923
APBB2_(63)|alt_cassette|chr4:40631473-40641638:-	-1.237	1.918
SLC25A29_(368)|alt_cassette|chr14:99829467-99834931:-	-1.353	1.907
CHRNA1_(75)|alt_cassette|chr2:175330649-175332304:-	0	1.904
FAM190A_1_s0e19|alt_start|chr4:91448417-91864087:+	1.063	1.9
GSTO2_1_s0e6|alt_start|chr10:106018620-106024973:+	2.509	1.896
TTYH1_(869)|alt_cassette|chr19:59618664-59622113:+	2.053	1.879
KCND3_(57)|alt_cassette|chr1:112121418-112124369:-	-1.488	1.872
C1orf52_(127)|alt_cassette|chr1:85496993-85497628:-	-1.088	1.866
CDH11_(179)|alt_cassette|chr16:63539503-63542170:-	2.155	1.862
SNTG1_1_s0e5|alt_start|chr8:50984901-51247673:+	0	1.858
LPHN2_(39)|alt_cassette|chr1:82190414-82194148:+	0	1.857
DNASE1L1_(85)|alt_cassette|chrX:153287190-153293421:-	1.589	1.853

Table 3.4 (continued): Significantly changing host alternative splicing events

SPLICING_EVENT(type of event indicated) *hg18-based*	HFF sepscore	NPC sepscore
GPR64_(42)|alt_cassette|chrX:18959076-18963980:-	0	1.851
UBE2E3_1_s0e8|alt_start|chr2:181553913-181554989:+	-0.798	1.851
ZNF720_(552)|retained_int|chr16:31672722-31673276:+	2.94	1.846
SGK3_(96)|alt_cassette|chr8:67915899-67921981:+	2.108	1.824
NNMT_1_s0e6|alt_start|chr11:113633762-113672359:+	2.472	1.82
HMGXB4_(115)|alt_cassette|chr22:33989867-33990640:+	-0.534	1.813
ACSL6_2_s54e81|alt_end|chr5:131170738-131323127:-	0	1.796
C19orf48_1_s0e5|alt_start|chr19:55997683-55999786:-	0	1.795
NDRG1_(1540)|retained_int|chr8:134318954-134320496:-	0	1.791
TRIM58_1_s11e17|alt_end|chr1:246105962-246126765:+	0.955	1.777
MORC4_1_s5e39|alt_end|chrX:105943758-106123115:-	0	1.771
C21orf34_1_s0e9|alt_start|chr21:16364712-16525246:+	1.855	1.756
PARK2_2_s19e28|alt_end|chr6:161688579-161889875:-	0	1.755
TGM2_1_s0e26|alt_start|chr20:36194316-36227114:-	0.546	1.74
BCLAF1_(147)|alt_cassette|chr6:136624308-136630992:-	0	1.736
CYP2A7_1_s1e29|alt_end|chr19:46073183-46225287:-	-1.114	1.73
ICA1_(99)|alt_cassette|chr7:8134297-8164681:-	0	1.729
TPRA1_(61)|alt_cassette|chr3:128777038-128777472:-	1.08	1.716
MICAL2_(922)|retained_int|chr11:12240758-12241682:+	-1.192	1.705
ZNF365_1_s7e13|alt_end|chr10:63818341-63889562:+	0	1.688
C8orf46_(94)|alt_cassette|chr8:67571281-67580163:+	1.291	1.66
ZNF266_(130)|alt_cassette|chr19:9390307-9390726:-	0	1.636
ROR2_3_s21e31|alt_end|chr9:93365193-93526256:-	-0.848	1.636
TMEM54_(60)|alt_cassette|chr1:33133931-33136313:-	0	1.624
LRRFIP2_(72)|alt_cassette|chr3:37100301-37111286:-	-3.049	1.624
ANK2_1_s0e7|alt_start|chr4:113958687-114315020:+	-1.018	1.622
PPIL5_(67)|alt_cassette|chr14:49138936-49143867:+	0	1.621
RFC5_(364)|alt_cassette|chr12:116939080-116941259:+	-0.631	1.618
HDAC6_(278)|alt_5|chrX:48545623-48546013:+	4.32	1.617
ABCC6P1_1_s3e21|alt_end|chr16:18491883-18517108:+	0	1.61
FOXP2_1_s0e9|alt_start|chr7:113513617-113853792:+	0	1.607
C16orf13_(68)|alt_cassette|chr16:624798-625281:-	1.167	1.602
TMTC4_3_s30e40|alt_end|chr13:100054090-100075789:-	1.186	1.6
CCDC29_(212_223)|twin_cassett|chr9:68730129-68736496:+	0.723	1.6
RCAN2_1_s0e9|alt_start|chr6:46324592-46532674:-	5.031	1.594
PIGW_1_s0e5|alt_start|chr17:31964959-31967055:+	3.488	1.589
EPS8L1_(92)|retained_int|chr19:60285360-60285454:+	3.27	1.587
VDAC3_3_s15e23|alt_end|chr8:42378690-42382572:+	0	1.587
TPD52_1_s0e25|alt_start|chr8:81104843-81139383:-	2.633	1.586
BTAF1_(127)|alt_5|chr10:93708891-93709521:+	-2.349	1.582
CEP78_(48)|alt_cassette|chr9:80070279-80071177:+	0.968	1.58

Table 3.4 (continued): Significantly changing host alternative splicing events

SPLICING_EVENT(type of event indicated) *hg18-based*	HFF sepscore	NPC sepscore
ARAP2_1_s0e31|alt_start|chr4:35837522-35907662:-	2.328	1.573
TNRC6B_1_s0e11|alt_start|chr22:38770766-38971964:+	1.753	1.571
ZWILCH_(94)|alt_3|chr15:64584783-64588232:+	0.77	1.569
AGXT2L2_(131)|alt_cassette|chr5:177571577-177574402:-	0	1.565
PHKA2_(169)|alt_cassette|chrX:18876865-18880532:-	1.605	1.546
NAV2_(45)|alt_cassette|chr11:20027317-20032201:+	-2.12	1.544
TOP 100 events for HFFs (-sepscore)		
FAM126A_(296)|alt_cassette|chr7:22952307-22966399:-	-6.378	0
FAM111A_(928)|retained_int|chr11:58668945-58669875:+	-6.163	-4.548
PSG5_(255)|alt_cassette|chr19:48366130-48380773:-	-5.873	0
WFDC2_(129)|alt_cassette|chr20:43532586-43541995:+	-5.761	0
NELL2_1_s0e3|alt_start|chr12:43555950-43557147:-	-5.33	-0.951
XAF1_(57)|alt_cassette|chr17:6602267-6604448:+	-5.328	0
MPRIP_(63)|alt_cassette|chr17:17024127-17028861:+	-5.233	0
DTNA_1_s5e9|alt_start|chr18:30513429-30589937:+	-5.058	-0.806
PSMG4_(447)|retained_int|chr6:3208992-3209441:+	-4.955	-2.945
MOGAT3_(203)|alt_cassette|chr7:100626101-100628191:-	-4.909	-2.63
MYLK_(153)|alt_cassette|chr3:124828478-124839607:-	-4.886	-0.709
LIMA1_1_s0e8|alt_start|chr12:48902535-48963620:-	-4.827	-1.037
VASH2_2_s19e24|alt_end|chr1:211214035-211230871:+	-4.716	0
D2HGDH_(143)|alt_cassette|chr2:242338382-242343132:+	-4.703	0
L3MBTL2_(105)|alt_cassette|chr22:39953273-39953753:+	-4.635	-0.759
ZNF410_1_s0e27|alt_start|chr14:73395271-73423371:+	-4.626	-2.977
MYO18A_(45)|alt_cassette|chr17:24433582-24437581:-	-4.587	0
FLNB_(39)|alt_cassette|chr3:58099296-58103416:+	-4.578	0
TRAF3_(139)|alt_cassette|chr14:102313765-102406274:+	-4.5	-2.344
ANGPTL2_(866)|alt_cassette|chr9:128896026-128924414:-	-4.489	-0.845
TLE6_(105)|retained_int|chr19:2938788-2938895:+	-4.461	-3.373
SPAG9_(39)|alt_cassette|chr17:46407307-46409467:-	-4.393	-0.744
CDC16_(153)|alt_cassette|chr13:114034251-114042911:+	-4.352	-1.073
TCF12_1_s6e20|alt_start|chr15:55000588-55310641:+	-4.347	-1.531
VIM_(61)|alt_cassette|chr10:17311990-17315591:+	-4.318	0
B3GNT4_(163)|alt_cassette|chr12:121254301-121256817:+	-4.282	-1.025
ACTN1_(66)|alt_cassette|chr14:68413710-68415458:-	-4.245	0
ETF1_(198)|alt_5|chr5:137882455-137906420:-	-4.129	-0.97
NFASC_(117)|alt_cassette|chr1:203224557-203238346:+	-4.126	-0.549
ISLR_1_s0e5|alt_start|chr15:72253139-72254244:+	-4.086	0
CCNJL_(207)|alt_cassette|chr5:159613383-159615133:-	-4.028	0
GATA2_(41)|alt_3|chr3:129683435-129685392:-	-4.008	0
PLOD2_(63)|alt_cassette|chr3:147277372-147279592:-	-3.986	0.66
ADC_(661)|retained_int|chr1:33320542-33321205:+	-3.914	-1.219
LEF1_(84)|alt_cassette|chr4:109220219-109223960:-	-3.904	-1.341

Table 3.4 (continued): Significantly changing host alternative splicing events

SPLICING_EVENT(type of event indicated) *hg18-based*	HFF sepscore	NPC sepscore
ADD2_3_s26e29|alt_end|chr2:70755317-70757435:-	-3.878	0
KIAA0226_(75)|alt_cassette|chr3:198895485-198904982:-	-3.867	0
SYNE2_(69)|alt_cassette|chr14:63750941-63752718:+	-3.866	-0.675
PSG2_(82)|alt_3|chr19:48103846-48106571:-	-3.861	0
PTK7_(191)|alt_cassette|chr6:43205545-43206226:+	-3.817	0
EPCAM_1_s0e7|alt_start|chr2:47425800-47454105:+	-3.802	-2.834
ADAM33_(217)|retained_int|chr20:3602931-3603150:-	-3.793	0
MAP3K9_(69)|alt_cassette|chr14:70270968-70274714:-	-3.787	-0.839
DNAJB12_(1160)|retained_int|chr10:73764381-73765543:-	-3.785	0
GPR132_(233)|alt_cassette|chr14:104589484-104592768:-	-3.757	-1.253
DOCK4_(150)|retained_int|chr7:111166535-111166687:-	-3.752	-0.681
SYTL2_(144)|retained_int|chr11:85107534-85107680:-	-3.736	0
MASP1_2_s21e36|alt_end|chr3:188418631-188441962:-	-3.726	-0.561
TCOF1_(231)|alt_cassette|chr5:149729358-149733929:+	-3.711	0
MAD1L1_(176)|alt_cassette|chr7:2236304-2239032:-	-3.679	-0.55
MORN3_(185)|alt_cassette|chr12:120575097-120576562:-	-3.677	0
SLC38A11_(66)|alt_cassette|chr2:165504312-165510348:-	-3.655	0
MAP3K14_1_s0e4|alt_start|chr17:40723914-40750197:-	-3.631	-0.937
CHRNB4_(147)|alt_5|chr15:76709342-76710472:-	-3.612	0
TPD52_(70)|alt_cassette|chr8:81126418-81128071:-	-3.606	-1.082
XAF1_(196)|alt_cassette|chr17:6602267-6606197:+	-3.597	0
SEMA3B_1_s0e20|alt_start|chr3:50283324-50285732:+	-3.59	0
SUV420H2_(166)|alt_cassette|chr19:60545226-60545921:+	-3.59	0
CABP4_(81)|alt_3|chr11:66979836-66980234:+	-3.529	-2.465
LONP1_1_s0e3|alt_start|chr19:5665282-5671463:-	-3.516	0
SULF1_(143)|alt_cassette|chr8:70704468-70713523:+	-3.49	0
PCDH9_1_s3e13|alt_end|chr13:65774966-66697537:-	-3.481	0
ADM_(148)|retained_int|chr11:10283921-10284071:+	-3.457	0
NEDD4L_(132)|alt_cassette|chr18:54152104-54159889:+	-3.456	-0.775
MRPL35_(1810)|retained_int|chr2:86291247-86293059:+	-3.453	-1.005
OPALIN_(39)|alt_cassette|chr10:98101157-98105001:-	-3.429	0
LOC729678_1_s0e4|alt_start|chr5:180190562-180191364:+	-3.409	-0.585
DHFR_(50)|alt_cassette|chr5:79981069-79985978:-	-3.398	-0.999
SLC40A1_1_s0e6|alt_start|chr2:190153533-190156729:-	-3.397	0
R3HDM2_(70)|alt_cassette|chr12:55990513-56110769:-	-3.396	0
ST7OT3_(69)|alt_cassette|chr7:116559250-116563370:+	-3.387	0
PLAT_(138)|alt_cassette|chr8:42164691-42168046:-	-3.373	0
COG4_(404)|retained_int|chr16:69100228-69100634:-	-3.372	0
INADL_(84)|alt_cassette|chr1:62322908-62352312:+	-3.36	0
ARSA_(237)|alt_cassette|chr22:49412700-49413239:-	-3.348	0
LAT_(85)|retained_int|chr16:28905391-28905478:+	-3.339	0

Table 3.4 (continued): Significantly changing host alternative splicing events

SPLICING_EVENT(type of event indicated) *hg18-based*	HFF sepscore	NPC sepscore
KIAA1462_(3764)|alt_cassette|chr10:30346852-30376466:-	-3.337	0.586
ZBP1_(225)|alt_cassette|chr20:55624042-55628723:-	-3.322	0
PDLIM7_(62)|alt_cassette|chr5:176849234-176850476:-	-3.305	0
D2HGDH_(384)|alt_cassette|chr2:242322783-242329120:+	-3.301	0.547
CD3D_(132)|alt_cassette|chr11:117715419-117716299:-	-3.301	-0.52
CIRBP_(373)|retained_int|chr19:1223050-1223425:+	-3.293	0
CHODL_1_s0e9|alt_start|chr21:18195450-18550696:+	-3.265	0
INTS1_(1032)|retained_int|chr7:1488869-1489903:-	-3.264	-1.755
FAM190B_(85)|alt_cassette|chr10:86227400-86263184:+	-3.252	-1.792
KIF13A_(39)|alt_cassette|chr6:17896085-17902458:-	-3.251	0.857
CLSTN1_(30)|alt_cassette|chr1:9737954-9755916:-	-3.25	0
TTC39A_4_s36e48|alt_end|chr1:51525369-51534338:-	-3.246	0
TECR_(67)|alt_cassette|chr19:14501523-14534336:+	-3.239	0
IGSF3_(60)|alt_cassette|chr1:116948170-116952086:-	-3.216	-1.038
TERT_(189)|alt_cassette|chr5:1313715-1319578:-	-3.214	-0.578
MYO5B_1_s65e68|alt_start|chr18:45630028-45634015:-	-3.207	-1.241
SLC16A4_(133)|alt_cassette|chr1:110725940-110733381:-	-3.201	0
TAPBPL_(123)|retained_int|chr12:6436970-6437095:+	-3.2	0
TPD52_1_s23e33|alt_end|chr8:80993649-81113018:-	-3.199	-1.224
CCDC114_1_s0e6|alt_start|chr19:53515460-53517003:-	-3.19	0
PI16_(578)|alt_3|chr6:37038787-37039576:+	-3.181	0
NOVA1_(102)|alt_cassette|chr14:26019189-26134455:-	-3.164	0
TPM1_(126)|alt_cassette|chr15:61122195-61123278:+	-3.152	0
PAX6_(330)|alt_5|chr11:31778980-31779684:-	-3.137	0
TOP 100 events for NPCs (-sepscore):		
FAM111A_(928)|retained_int|chr11:58668945-58669875:+	-6.163	-4.548
CKAP2L_(756)|retained_int|chr2:113230182-113230940:-	0	-3.829
TGM2_1_s13e22|alt_end|chr20:36199928-36208532:-	0	-3.708
CPEB1_1_s6e10|alt_start|chr15:81037337-81093173:-	-0.897	-3.68
FAM193B_(179)|alt_cassette|chr5:176891128-176892049:-	-1.372	-3.423
TLE6_(105)|retained_int|chr19:2938788-2938895:+	-4.461	-3.373
ABCG5_(579)|retained_int|chr2:43918598-43919179:-	-2.8	-3.33
RBP7_(215)|alt_cassette|chr1:9979976-9990214:+	-2.454	-3.24
ZNF410_1_s0e27|alt_start|chr14:73395271-73423371:+	-4.626	-2.977
PSMG4_(447)|retained_int|chr6:3208992-3209441:+	-4.955	-2.945
SYTL2_1_s18e22|alt_start|chr11:85107534-85109650:-	-2.981	-2.864
EPCAM_1_s0e7|alt_start|chr2:47425800-47454105:+	-3.802	-2.834
PDGFA_(69)|alt_cassette|chr7:504737-507278:-	-0.608	-2.833
HLA-F_2_s11e16|alt_end|chr6:29801799-29803052:+	-0.791	-2.691
NIPAL2_(28)|alt_cassette|chr8:99274494-99276131:-	1.504	-2.648
MOGAT3_(203)|alt_cassette|chr7:100626101-100628191:-	-4.909	-2.63

Table 3.4 (continued): Significantly changing host alternative splicing events

SPLICING_EVENT(type of event indicated) *hg18-based*	HFF sepscore	NPC sepscore
PPP1R13B_(617)|alt_cassette|chr14:103293841-103314834:-	-1.253	-2.589
PDE4DIP_1_s11e15|alt_start|chr1:143706058-143727368:-	-0.928	-2.497
CABP4_(81)|alt_3|chr11:66979836-66980234:+	-3.529	-2.465
FBLN5_(123)|alt_cassette|chr14:91473298-91476661:-	2.303	-2.428
KIFC3_(29)|alt_cassette|chr16:56350322-56351140:-	0	-2.381
SLC39A6_(798)|alt_cassette|chr18:31958661-31963074:-	0	-2.38
DOCK8_(962)|alt_cassette|chr9:330321-358017:+	-1.658	-2.377
LOC645323_1_s0e8|alt_start|chr5:88004602-88016509:-	1.018	-2.37
FAM190A_2_s19e23|alt_end|chr4:91864165-92063621:+	0	-2.356
TRAF3_(139)|alt_cassette|chr14:102313765-102406274:+	-4.5	-2.344
KIAA1543_(48_33)|twin_cassett|chr19:7577457-7579011:+	-2.392	-2.295
MET_(50)|alt_cassette|chr7:116099867-116126360:+	-1.124	-2.274
MLST8_(125)|alt_cassette|chr16:2197347-2198451:+	-1.428	-2.256
HIF3A_(146)|alt_cassette|chr19:51499185-51503317:+	3.502	-2.249
BX537629_1_s0e7|alt_start|chr8:65453169-65456713:+	0	-2.241
DRD2_(87)|alt_cassette|chr11:112788815-112791352:-	1.751	-2.24
TPD52_(27_42)|twin_cassett|chr8:81117458-81126316:-	-1.42	-2.231
MCOLN3_(159)|alt_cassette|chr1:85272522-85279280:-	-1.629	-2.224
PCSK2_1_s0e4|alt_start|chr20:17154751-17155974:+	0	-2.217
CYTH4_(115)|alt_cassette|chr22:36029389-36035198:+	3.28	-2.184
CATSPER2_(120)|alt_cassette|chr15:41614966-41618959:-	0	-2.182
CACNA2D2_(102)|retained_int|chr3:50377405-50377509:-	-2.99	-2.177
TTC39A_(104)|alt_3|chr1:51547585-51549513:-	-2.185	-2.176
MDC1_(792)|alt_cassette|chr6:30783210-30787167:-	-1.344	-2.15
MAP2_1_s0e7|alt_start|chr2:209997015-210153004:+	-1.291	-2.137
KCNIP3_1_s0e10|alt_start|chr2:95326798-95403770:+	-0.658	-2.132
SLC16A7_(59_78)|twin_cassett|chr12:58385066-58451266:+	0.853	-2.129
FXYD5_1_s3e12|alt_start|chr19:40338293-40346898:+	0	-2.082
NIPAL3_(100)|alt_cassette|chr1:24655350-24657964:+	0	-2.082
PRSS1_1_s6e23|alt_end|chr7:142139198-142182363:+	-2.666	-2.075
C20orf24_1_s0e5|alt_start|chr20:34635423-34640557:+	0	-2.07
STK32A_1_s11e23|alt_end|chr5:146702799-146734974:+	0	-2.067
BOC_(1463)|retained_int|chr3:114472481-114473946:+	-1.919	-2.057
NRXN3_(90)|alt_cassette|chr14:79200045-79233725:+	-3.06	-2.055
DACH1_(162)|alt_cassette|chr13:70961281-71153927:-	0.862	-2.053
ENPP5_(62)|alt_cassette|chr6:46243993-46246503:-	0	-2.044
CPZ_(33)|alt_cassette|chr4:8645548-8653749:+	-2.261	-2.023
FAM193B_(258)|alt_3|chr5:176896093-176897479:-	-2.954	-2.004
NRXN3_1_s0e38|alt_start|chr14:78181215-79003316:+	-1.978	-1.998
ZFAND2B_(122)|retained_int|chr2:219779811-219779935:+	-1.192	-1.986

Table 3.4 (continued): Significantly changing host alternative splicing events

SPLICING_EVENT(type of event indicated) *hg18-based*	HFF sepscore	NPC sepscore
CLIC6_(54)|alt_cassette|chr21:34964931-35001447:+	0	-1.968
SLC25A24_1_s0e5|alt_start|chr1:108530099-108544497:-	0	-1.957
SAP30BP_(48)|alt_cassette|chr17:71176300-71201114:+	0	-1.952
EYA4_1_s1e11|alt_end|chr6:133604580-133894951:+	0	-1.926
WNK1_(459_279)|twin_cassett|chr12:857788-861118:+	1.612	-1.904
FN1_(273)|alt_cassette|chr2:215964782-215967495:-	-1.565	-1.882
HP1BP3_(327)|alt_5|chr1:20979620-20985593:-	0	-1.87
CRYZ_(102)|alt_cassette|chr1:74945266-74948369:-	-0.998	-1.869
RUNDC3B_(51)|alt_cassette|chr7:87118189-87167672:+	-1.389	-1.866
ZFHX4_(29)|alt_3|chr8:77917562-77923815:+	0	-1.865
SEPT8_3_s20e26|alt_end|chr5:132114407-132124392:-	-1.243	-1.853
BID_(95)|alt_cassette|chr22:16602254-16606568:-	0	-1.837
PDE1C_1_s0e15|alt_start|chr7:31886998-32305466:-	-2.441	-1.836
NDRG4_(1126)|alt_cassette|chr16:57086413-57095202:+	-1.113	-1.802
JPH4_(40)|alt_3|chr14:23101115-23101336:-	-0.854	-1.8
GRIP1_(156)|alt_cassette|chr12:65135611-65142970:-	0	-1.797
FAM190B_(85)|alt_cassette|chr10:86227400-86263184:+	-3.252	-1.792
TRIM59_1_s0e10|alt_start|chr3:161599627-161650320:-	-2.488	-1.79
SLC7A8_(167)|alt_3|chr14:22667245-22668698:-	-0.971	-1.79
C1orf201_4_s18e24|alt_end|chr1:24556075-24568750:-	0	-1.783
LRIG1_(140)|alt_5|chr3:66516515-66517245:-	-1.853	-1.779
GLS_1_s29e40|alt_end|chr2:191504608-191538021:+	0	-1.775
RIPK2_(154)|alt_cassette|chr8:90839596-90846705:+	0	-1.761
INTS1_(1032)|retained_int|chr7:1488869-1489903:-	-3.264	-1.755
CASK_(69)|alt_cassette|chrX:41299832-41303975:-	2.626	-1.752
EYA4_(69)|alt_cassette|chr6:133809585-133819386:+	-2.447	-1.739
LRRFIP1_(72)|alt_cassette|chr2:238322706-238326690:+	-1.661	-1.728
ACYP1_(79)|alt_cassette|chr14:74590115-74599925:-	-1.011	-1.725
MAP7D2_(66)|alt_cassette|chrX:19981027-19984718:-	0	-1.722
SEMA3F_(93)|alt_cassette|chr3:50186787-50189204:+	-1.401	-1.715
C16orf55_1_s0e7|alt_start|chr16:88251710-88252156:+	-2.307	-1.705
CSF3R_(27)|alt_3|chr1:36706150-36706262:-	-2.791	-1.699
TMEM117_(100)|alt_cassette|chr12:42624412-42891339:+	0	-1.692
AJAP1_1_s9e14|alt_end|chr1:4734478-4743711:+	-1.124	-1.69
SALL2_1_s0e5|alt_start|chr14:21063628-21075177:-	0.892	-1.689
ADM_(448)|retained_int|chr11:10283352-10283802:+	-1.893	-1.685
MYCBP2_(180)|alt_cassette|chr13:76571149-76593508:-	0.826	-1.678
LONRF3_(158)|alt_3|chrX:118029964-118031029:+	0	-1.676
UBE2K_(129)|alt_cassette|chr4:39452948-39456374:+	0	-1.675
NIPAL3_(110)|alt_cassette|chr1:24618717-24639248:+	0	-1.668
MOV10_(141)|alt_cassette|chr1:113019194-113033079:+	0	-1.659
NDRG1_1_s10e19|alt_start|chr8:134336020-134343592:-	-0.94	-1.638

Table 3.4 (continued): Significantly changing host alternative splicing events

SPLICING_EVENT(type of event indicated) ***hg18-based***	**HFF sepscore**	**NPC sepscore**
EPB41L3_(561)|alt_cassette|chr18:5400618-5409709:-	0.742	-1.628
HOXC4_1_s1e10|alt_start|chr12:52697465-52714328:+	-2.25	-1.617

Table 3.5: Alternative cassette exon splicing event changes common to HFFs and NPCs

Common significant changes in either direction (exon inclusion or exclusion) are listed.

SPLICING_EVENT *hg18-based-coordinates*	**HFF sepscore**	**NPC sepscore**
FAM193B_(179)|alt_cassette|chr5:176891128-176892049:-	-1.372	-3.423
RBP7_(215)|alt_cassette|chr1:9979976-9990214:+	-2.454	-3.24
PDGFA_(69)|alt_cassette|chr7:504737-507278:-	-0.608	-2.833
STXBP2_(50)|alt_cassette|chr19:7608072-7609904:+	0.683	2.819
ARHGAP8_(132)|alt_cassette|chr22:43576708-43589215:+	1.466	2.694
SPAG6_(146)|alt_cassette|chr10:22730212-22745553:+	1.619	2.649
NIPAL2_(28)|alt_cassette|chr8:99274494-99276131:-	1.504	-2.648
MOGAT3_(203)|alt_cassette|chr7:100626101-100628191:-	-4.909	-2.63
PPP1R13B_(617)|alt_cassette|chr14:103293841-103314834:-	-1.253	-2.589
TMEM108_(1410)|alt_cassette|chr3:134430885-134591713:+	4.155	2.571
SPINT2_(171)|alt_cassette|chr19:43447478-43470355:+	1.672	2.567
FBLN5_(123)|alt_cassette|chr14:91473298-91476661:-	2.303	-2.428
DOCK8_(962)|alt_cassette|chr9:330321-358017:+	-1.658	-2.377
TRAF3_(139)|alt_cassette|chr14:102313765-102406274:+	-4.5	-2.344
MET_(50)|alt_cassette|chr7:116099867-116126360:+	-1.124	-2.274
MLST8_(125)|alt_cassette|chr16:2197347-2198451:+	-1.428	-2.256
HIF3A_(146)|alt_cassette|chr19:51499185-51503317:+	3.502	-2.249
DRD2_(87)|alt_cassette|chr11:112788815-112791352:-	1.751	-2.24
FAM3A_(51)|alt_cassette|chrX:153389015-153389813:-	2.597	2.224
MCOLN3_(159)|alt_cassette|chr1:85272522-85279280:-	-1.629	-2.224
CYTH4_(115)|alt_cassette|chr22:36029389-36035198:+	3.28	-2.184
MDC1_(792)|alt_cassette|chr6:30783210-30787167:-	-1.344	-2.15
ODF3L2_(108)|alt_cassette|chr19:418762-425620:-	0.846	2.148
NRXN3_(90)|alt_cassette|chr14:79200045-79233725:+	-3.06	-2.055
DACH1_(162)|alt_cassette|chr13:70961281-71153927:-	0.862	-2.053
CPZ_(33)|alt_cassette|chr4:8645548-8653749:+	-2.261	-2.023
EXOC7_(153)|alt_cassette|chr17:71596996-71598818:-	3.271	2.009
APBB2_(63)|alt_cassette|chr4:40631473-40641638:-	-1.237	1.918
SLC25A29_(368)|alt_cassette|chr14:99829467-99834931:-	-1.353	1.907
FN1_(273)|alt_cassette|chr2:215964782-215967495:-	-1.565	-1.882
TTYH1_(869)|alt_cassette|chr19:59618664-59622113:+	2.053	1.879
KCND3_(57)|alt_cassette|chr1:112121418-112124369:-	-1.488	1.872
CRYZ_(102)|alt_cassette|chr1:74945266-74948369:-	-0.998	-1.869
RUNDC3B_(51)|alt_cassette|chr7:87118189-87167672:+	-1.389	-1.866
C1orf52_(127)|alt_cassette|chr1:85496993-85497628:-	-1.088	1.866
CDH11_(179)|alt_cassette|chr16:63539503-63542170:-	2.155	1.862

Table 3.5 (continued): Alternative cassette exon splicing event changes common to HFFs and NPCs

SPLICING_EVENT *hg18-based-coordinates*	HFF sepscore	NPC sepscore
DNASE1L1_(85)|alt_cassette|chrX:153287190-153293421:-	1.589	1.853
SGK3_(96)|alt_cassette|chr8:67915899-67921981:+	2.108	1.824
HMGXB4_(115)|alt_cassette|chr22:33989867-33990640:+	-0.534	1.813
NDRG4_(1126)|alt_cassette|chr16:57086413-57095202:+	-1.113	-1.802
FAM190B_(85)|alt_cassette|chr10:86227400-86263184:+	-3.252	-1.792
CASK_(69)|alt_cassette|chrX:41299832-41303975:-	2.626	-1.752
EYA4_(69)|alt_cassette|chr6:133809585-133819386:+	-2.447	-1.739
LRRFIP1_(72)|alt_cassette|chr2:238322706-238326690:+	-1.661	-1.728
ACYP1_(79)|alt_cassette|chr14:74590115-74599925:-	-1.011	-1.725
TPRA1_(61)|alt_cassette|chr3:128777038-128777472:-	1.08	1.716
SEMA3F_(93)|alt_cassette|chr3:50186787-50189204:+	-1.401	-1.715
MYCBP2_(180)|alt_cassette|chr13:76571149-76593508:-	0.826	-1.678
C8orf46_(94)|alt_cassette|chr8:67571281-67580163:+	1.291	1.66
EPB41L3_(561)|alt_cassette|chr18:5400618-5409709:-	0.742	-1.628
LRRFIP2_(72)|alt_cassette|chr3:37100301-37111286:-	-3.049	1.624
RFC5_(364)|alt_cassette|chr12:116939080-116941259:+	-0.631	1.618
ACADM_(101)|alt_cassette|chr1:75971195-75973063:+	1.291	-1.609
C16orf13_(68)|alt_cassette|chr16:624798-625281:-	1.167	1.602
CEP78_(48)|alt_cassette|chr9:80070279-80071177:+	0.968	1.58
GRINL1A_(533)|alt_cassette|chr15:55719035-55740939:+	1.223	-1.57
MCTP2_(110)|alt_cassette|chr15:92575897-92642433:+	1.35	-1.567
STXBP5_(60)|alt_cassette|chr6:147697051-147716151:+	-1.253	-1.563
PHKA2_(169)|alt_cassette|chrX:18876865-18880532:-	1.605	1.546
KIAA0226_(45)|alt_cassette|chr3:198908317-198911880:-	-2.212	-1.545
NAV2_(45)|alt_cassette|chr11:20027317-20032201:+	-2.12	1.544
SECISBP2L_(135)|alt_cassette|chr15:47097117-47106853:-	1.682	-1.54
SORBS1_(96)|alt_cassette|chr10:97182323-97187236:-	-1.525	-1.537
CAST_(66)|alt_cassette|chr5:96084158-96088948:+	-2.909	-1.521
TFEB_(255)|alt_cassette|chr6:41765527-41766716:-	0.878	1.514
NFATC1_(1099)|alt_cassette|chr18:75257339-75294566:+	0.6	1.509
ARHGAP9_(106)|alt_cassette|chr12:56152738-56154091:-	2.541	1.493
SMPDL3A_(214)|alt_cassette|chr6:123152302-123159667:+	1.904	-1.491
VRK3_(62)|alt_cassette|chr19:55204454-55211092:-	-0.752	-1.488
MXRA7_(81)|alt_cassette|chr17:72188556-72192748:-	-1.767	-1.463
PIP5K1A_(147)|alt_cassette|chr1:149479139-149481537:+	-0.541	1.46
DIS3_(158)|alt_cassette|chr13:72250519-72253427:-	1.352	-1.444
SRPK3_(166)|alt_cassette|chrX:152684229-152685521:+	2.573	1.438
MCCC2_(114)|alt_cassette|chr5:70936051-70963703:+	2.766	1.427
RNF146_(109)|alt_cassette|chr6:127643178-127648887:+	-0.652	1.425
NIN_(2139)|alt_cassette|chr14:50291335-50296324:-	-2.779	-1.422
ZNF507_(94)|alt_cassette|chr19:37528529-37535574:+	1.541	-1.417
BPHL_(228)|alt_cassette|chr6:3064080-3068889:+	-1.958	-1.413

Table 3.5 (continued): Alternative cassette exon splicing event changes common to HFFs and NPCs

SPLICING_EVENT *hg18-based-coordinates*	HFF sepscore	NPC sepscore
NT5C2_(281)|alt_cassette|chr10:104849766-104850791:-	-1.286	1.407
EXOC1_(45)|alt_cassette|chr4:56444851-56451145:+	3.506	-1.404
MYH14_(99)|alt_cassette|chr19:55447822-55452367:+	-1.362	-1.404
TTLL7_(108)|alt_cassette|chr1:84121407-84128591:-	-0.639	-1.402
MAPT_(87)|alt_cassette|chr17:41395672-41411576:+	-1.225	-1.398
SLC25A35_(76)|alt_cassette|chr17:8134027-8134884:-	0.935	1.387
VGLL4_(133)|alt_cassette|chr3:11618496-11719444:-	-0.584	-1.382
ABLIM1_(78)|alt_cassette|chr10:116222881-116237706:-	-0.504	1.378
MAP2_(171)|alt_cassette|chr2:210273307-210278548:+	-0.798	-1.376
EPB41_(63)|alt_cassette|chr1:29252411-29264080:+	-0.905	-1.366
PTPRD_(42)|alt_cassette|chr9:8426689-8439724:-	-1.685	-1.359
ZNF317_(96)|alt_cassette|chr19:9129070-9130501:+	-1.321	1.355
LASS2_(174)|alt_cassette|chr1:149207612-149213718:-	-0.916	-1.351
LRP8_(225)|alt_cassette|chr1:53493445-53496578:-	0.701	1.351
UBQLN1_(84)|alt_cassette|chr9:85469880-85473919:-	-0.798	1.348
ERC1_(84)|alt_cassette|chr12:1089774-1095292:+	-1.455	1.344
LEF1_(84)|alt_cassette|chr4:109220219-109223960:-	-3.904	-1.341
GUF1_(112)|alt_cassette|chr4:44375561-44377467:+	1.193	-1.34
GFPT1_(54)|alt_cassette|chr2:69430819-69435124:-	-0.532	-1.336
IL12RB2_(100)|alt_cassette|chr1:67624940-67633817:+	2.167	1.333
GPR126_(46)|alt_cassette|chr6:142801194-142806166:+	0.738	1.311
AKR7A2_(105)|alt_cassette|chr1:19506491-19507535:-	-2.318	-1.305
CENPE_(75)|alt_cassette|chr4:104303862-104315341:-	0.568	-1.294
USPL1_(167)|alt_cassette|chr13:30090193-30093900:+	2.716	1.29
RUNX2_(66)|alt_cassette|chr6:45588122-45622541:+	-0.676	1.288
RSU1_(112)|alt_cassette|chr10:16864089-16899319:-	-0.842	1.282
CSDE1_(147)|alt_cassette|chr1:115084034-115093964:-	-1.104	-1.278
ZCWPW1_(153)|alt_cassette|chr7:99836895-99838074:-	-2.034	-1.263
C13orf23_(66)|alt_cassette|chr13:38503768-38509334:-	-0.648	-1.261
GBAP_(75)|alt_cassette|chr1:153453788-153454378:-	1.116	1.26
CALCR_(71)|alt_cassette|chr7:92954255-93041536:-	-2.35	1.258
NEXN_(42)|alt_cassette|chr1:78156546-78164686:+	-0.662	-1.255
GPR132_(233)|alt_cassette|chr14:104589484-104592768:-	-3.757	-1.253
ZDHHC16_(48)|alt_cassette|chr10:99203410-99204460:+	-1.006	1.25
AHI1_(62)|alt_cassette|chr6:135855122-135860413:-	-1.03	-1.249
FAM13B_(66)|alt_cassette|chr5:137317964-137323203:-	0.919	-1.247
RAB7L1_(254)|alt_cassette|chr1:204008318-204011023:-	0.866	1.246
AK131467_(158)|alt_cassette|chr2:55045373-55047639:+	-0.845	1.244
MORN4_(97)|alt_cassette|chr10:99367069-99383093:-	0.639	1.237
SPAST_(96)|alt_cassette|chr2:32168178-32193210:+	0.927	1.225
TLR4_(167)|alt_cassette|chr9:119506664-119514487:+	-0.705	1.217
FNIP1_(84)|alt_cassette|chr5:131072864-131080156:-	-0.715	-1.217

Table 3.5 (continued): Alternative cassette exon splicing event changes common to HFFs and NPCs

SPLICING_EVENT *hg18-based-coordinates*	HFF sepscore	NPC sepscore
MAPT_(54)|alt_cassette|chr17:41424789-41429601:+	0.916	-1.211
KIF16B_(153)|alt_cassette|chr20:16300406-16307449:-	0.884	1.21
ATP8A1_(45)|alt_cassette|chr4:42252814-42271392:-	-2.556	-1.207
DST_(111)|alt_cassette|chr6:56433011-56436321:-	-2.261	1.2
KIF19_(126)|alt_cassette|chr17:69850894-69852494:+	0.627	1.197
ASNS_(278)|alt_cassette|chr7:97337061-97339598:-	0.649	1.197
IPW_(120)|alt_cassette|chr15:22861095-22866359:+	3.039	1.197
CAPRIN2_(83)|alt_cassette|chr12:30775711-30779168:-	0.967	1.196
DDB2_(192)|alt_cassette|chr11:47194599-47210940:+	0.723	1.195
TNFRSF1B_(100)|alt_cassette|chr1:12149813-12173600:+	1.024	1.194
CXorf45_(237)|alt_cassette|chrX:110866763-110882673:+	-1.099	1.189
PATZ1_(73)|alt_cassette|chr22:30053295-30061677:-	-1.297	1.188
GSTM3_(76)|alt_cassette|chr1:110083638-110084384:-	0.682	1.186
ELK1_(106)|alt_cassette|chrX:47385818-47394765:-	1.423	1.185
SC4MOL_(286)|alt_cassette|chr4:166468366-166478390:+	0.991	-1.174
EHBP1_(108)|alt_cassette|chr2:63059974-63071354:+	1.489	1.169
OBFC2A_(104)|alt_cassette|chr2:192254988-192256699:+	-0.977	1.164
SLC37A2_(57)|alt_cassette|chr11:124461124-124463224:+	-1.29	-1.156
AKAP8L_(137)|alt_cassette|chr19:15369666-15371112:-	-0.68	-1.151
CTTN_(111)|alt_cassette|chr11:69944264-69946693:+	-2.011	1.146
MAPKSP1_(46)|alt_cassette|chr4:101032214-101034514:-	0.893	1.145
DOCK9_(36)|alt_cassette|chr13:98258062-98259606:-	-2.343	-1.145
KLHDC5_(103)|alt_cassette|chr12:27836101-27841914:+	0.873	1.145
SETX_(87)|alt_cassette|chr9:134130193-134134822:-	1.395	1.144
NBN_(50)|alt_cassette|chr8:91062927-91064125:-	0.955	1.142
ITSN2_(81)|alt_cassette|chr2:24352222-24362584:-	-1.001	-1.14
EPS8L1_(876)|alt_cassette|chr19:60286824-60289020:+	1.379	-1.138
CPNE4_(876)|alt_cassette|chr3:133106978-133239090:-	0.61	-1.137
PYCR2_(71)|alt_cassette|chr1:224176706-224178366:-	0.823	1.137
ITGB4_(159)|alt_cassette|chr17:71262491-71263376:+	-1.008	-1.136
GALK2_(115)|alt_cassette|chr15:47296802-47315339:+	-0.852	-1.135
HMBS_(54)|alt_cassette|chr11:118460986-118464554:+	-2.549	-1.131
TMEM126B_(172)|alt_cassette|chr11:85017380-85020378:+	-0.995	-1.128
PREB_(174)|alt_cassette|chr2:27208454-27208974:-	0.533	1.127
FBXO22_(139)|alt_cassette|chr15:73983499-73992598:+	0.935	1.125
DAB2_(63)|alt_cassette|chr5:39419130-39424657:-	2.381	1.121
SAT1_(110)|alt_cassette|chrX:23711921-23713365:+	-0.978	-1.12
TMEM188_(51)|alt_cassette|chr16:48616752-48617807:+	0.6	1.117
CD97_(147)|alt_cassette|chr19:14368285-14369469:+	-1.05	1.114
ASXL1_(153)|alt_cassette|chr20:30480895-30482784:+	-1.169	1.112
MBD1_(75)|alt_cassette|chr18:46055613-46055967:-	1.943	-1.112

Table 3.5 (continued): Alternative cassette exon splicing event changes common to HFFs and NPCs

SPLICING_EVENT *hg18-based-coordinates*	HFF sepscore	NPC sepscore
CDON_(98)|alt_cassette|chr11:125393578-125394723:-	1.92	1.111
VCL_(204)|alt_cassette|chr10:75538920-75543947:+	-1.242	1.097
PTPRU_(30)|alt_cassette|chr1:29483968-29490967:+	-2.89	-1.09
RAP1GAP2_(45)|alt_cassette|chr17:2807838-2813473:+	-1.155	-1.083
TPD52_(70)|alt_cassette|chr8:81126418-81128071:-	-3.606	-1.082
STRAP_(136)|alt_cassette|chr12:15927020-15934128:+	1.756	1.081
ANKRD6_(105)|alt_cassette|chr6:90384471-90389849:+	1.354	-1.079
FLI1_(171)|alt_cassette|chr11:128133431-128143222:+	-1.341	-1.074
RAB43_(94)|alt_cassette|chr3:130292879-130296518:-	0.57	-1.073
CDC16_(153)|alt_cassette|chr13:114034251-114042911:+	-4.352	-1.073
ZNF519_(127)|alt_cassette|chr18:14075075-14122273:-	0.717	-1.071
RIMS1_(183)|alt_cassette|chr6:73050542-73058357:+	-0.923	-1.069
COG1_(38)|alt_cassette|chr17:68714990-68716047:+	-1.664	1.066
MTRR_(140)|alt_cassette|chr5:7924036-7928370:+	1.57	-1.063
DOCK9_(69)|alt_cassette|chr13:98274759-98277084:-	-1.755	-1.063
ZNF195_(80)|alt_cassette|chr11:3349509-3356843:-	-0.531	-1.061
CDCA7_(237)|alt_cassette|chr2:173931811-173936199:+	0.62	-1.06
MLLT4_(45)|alt_cassette|chr6:168034558-168040406:+	-1.167	-1.058
LUC7L2_(71)|alt_cassette|chr7:138695608-138711347:+	-0.811	1.058
RBM23_(48)|alt_cassette|chr14:22445417-22448531:-	-0.685	-1.056
BCAT2_(183)|alt_cassette|chr19:53990533-53991496:-	0.937	-1.055
SMAD9_(111)|alt_cassette|chr13:36337895-36344794:-	1.318	1.051
SNAP91_(42)|alt_cassette|chr6:84383477-84407533:-	2.103	1.051
DDX26B_(111)|alt_cassette|chrX:134531022-134534405:+	1.079	-1.049
AP1S3_(39)|alt_cassette|chr2:224338311-224348861:-	-0.735	-1.045
IDI1_(173)|alt_cassette|chr10:1079333-1084803:-	3.009	-1.042
TMEM146_(55)|alt_cassette|chr19:5671819-5678278:+	-3.087	-1.042
SOS1_(45)|alt_cassette|chr2:39068236-39075767:-	-2.872	1.039
IGSF3_(60)|alt_cassette|chr1:116948170-116952086:-	-3.216	-1.038
C1orf91_(141)|alt_cassette|chr1:32455539-32459318:-	0.826	1.038
ATP6V0B_(49)|alt_cassette|chr1:44213366-44214348:+	1.567	1.036
YDJC_(178)|alt_cassette|chr22:20313076-20313596:-	0.589	1.03
B3GNT4_(163)|alt_cassette|chr12:121254301-121256817:+	-4.282	-1.025
SNCA_(84)|alt_cassette|chr4:90866834-90962419:-	3.735	1.025
GOLGA4_(63)|alt_cassette|chr3:37371682-37382574:+	1.6	-1.021
MAP2_(93)|alt_cassette|chr2:210283223-210298677:+	-2.376	1.017
ZNF643_(86)|alt_cassette|chr1:40689347-40695203:+	0.835	1.014
TBL1X_(135)|alt_cassette|chrX:9429889-9568312:+	-1.483	-1.013
ISCU_(96)|alt_cassette|chr12:107480641-107482183:+	-0.578	1.012
PTPN13_(57)|alt_cassette|chr4:87891301-87898436:+	0.624	-1.011
TEX264_(292)|alt_cassette|chr3:51680344-51693468:+	-0.512	1.01
MCAM_(118)|alt_cassette|chr11:118685835-118686675:-	1.192	1.009

Table 3.5 (continued): Alternative cassette exon splicing event changes common to HFFs and NPCs

SPLICING_EVENT *hg18-based-coordinates*	HFF sepscore	NPC sepscore
HOMER3_(101)|alt_cassette|chr19:18878921-18884792:-	0.871	1.005
RCE1_(110)|alt_cassette|chr11:66283143-66320327:+	0.625	1
DHFR_(50)|alt_cassette|chr5:79981069-79985978:-	-3.398	-0.999
CIZ1_(72)|alt_cassette|chr9:129987876-129992428:-	2.511	0.997
PYCR2_(93)|alt_cassette|chr1:224175694-224176180:-	2.046	0.997
NLGN4X_(60)|alt_cassette|chrX:5957473-6079035:-	1.681	0.995
RASGRF1_(39)|alt_cassette|chr15:77085831-77094723:-	-2.587	-0.992
ANKRD12_(69)|alt_cassette|chr18:9185696-9198654:+	-0.876	-0.992
NPEPL1_(72)|alt_cassette|chr20:56709621-56715585:+	3.334	0.991
NF1_(63)|alt_cassette|chr17:26600263-26609487:+	-1.787	-0.988
ATXN2L_(72)|alt_cassette|chr16:28751221-28751638:+	0.956	0.988
CP110_(143)|alt_cassette|chr16:19442911-19446689:+	1.45	-0.986
IL4R_(133)|alt_cassette|chr16:27232842-27259007:+	0.687	0.982
C1orf112_(89)|alt_cassette|chr1:168034723-168038385:+	1.659	-0.98
CAMK2G_(687)|alt_cassette|chr10:75247318-75249295:-	1.45	-0.971
FAM195A_(128)|alt_cassette|chr16:632250-637783:+	2.559	0.968
PLD1_(114)|alt_cassette|chr3:172878178-172887854:-	-1.419	-0.965
MPP2_(64)|alt_cassette|chr17:39331274-39340367:-	2.014	0.964
LAS1L_(51)|alt_cassette|chrX:64660867-64661569:-	-0.608	0.962
SHISA6_(96)|alt_cassette|chr17:11107568-11399781:+	2.104	0.962
THEM4_(232)|alt_cassette|chr1:150128473-150134107:-	1.095	0.961
PCBP4_(129)|alt_cassette|chr3:51968707-51969244:-	-2.402	-0.958
C4B_(137)|alt_cassette|chr6:32048267-32054658:+	0.791	0.955
SLC2A8_(146)|alt_cassette|chr9:129207091-129209211:+	1.447	-0.955
ATP9B_(33)|alt_cassette|chr18:75235089-75238234:+	0.724	-0.955
DCAF11_(65)|alt_cassette|chr14:23656442-23657103:+	1.12	0.954
DMXL2_(63)|alt_cassette|chr15:49538284-49542867:-	1.35	0.942
LSR_(147)|alt_cassette|chr19:40441807-40449101:+	1.058	0.938
SNURF_(124)|alt_cassette|chr15:22758449-22764171:+	-0.587	-0.936
ADARB1_(118)|alt_cassette|chr21:45372916-45415952:+	-0.653	-0.936
TMPO_(120)|alt_cassette|chr12:97455481-97462350:+	1.371	-0.935
ANKRD36_(73)|alt_cassette|chr2:97227955-97229797:+	0.947	0.934
XPNPEP3_(116)|alt_cassette|chr22:39583195-39587567:+	2.309	0.934
FUK_(172)|alt_cassette|chr16:69060372-69061717:+	1.626	0.933
NEXN_(192)|alt_cassette|chr1:78154406-78156232:+	1.796	-0.932
ASCC2_(124)|alt_cassette|chr22:28551769-28558233:-	1.574	0.932
NFASC_(51)|alt_cassette|chr1:203193577-203203999:+	3.236	0.931
TM9SF4_(84)|alt_cassette|chr20:30202339-30209257:+	0.687	0.931
WDR85_(66)|alt_cassette|chr9:139578879-139579357:-	-0.796	0.93
PHKA1_(39)|alt_cassette|chrX:71729849-71738732:-	-1.215	-0.926
NPEPPS_(116)|alt_cassette|chr17:43001859-43011754:+	-0.542	0.923
SETD4_(94)|alt_cassette|chr21:36351372-36352983:-	1.02	0.922

Table 3.5 (continued): Alternative cassette exon splicing event changes common to HFFs and NPCs

SPLICING_EVENT *hg18-based-coordinates*	HFF sepscore	NPC sepscore
MORN4_(110)|alt_cassette|chr10:99366158-99366954:-	0.521	0.92
DYSF_(93)|alt_cassette|chr2:71592559-71594353:+	1.287	-0.915
ATP2B2_(42)|alt_cassette|chr3:10395094-10404960:-	-1.067	-0.911
CSNK2A1_(117)|alt_cassette|chr20:437304-472315:-	0.731	-0.91
EXOG_(87)|alt_cassette|chr3:38513025-38514123:+	0.589	0.909
DGAT2_(129)|alt_cassette|chr11:75157805-75178880:+	0.794	0.909
KIAA0652_(111)|alt_cassette|chr11:46637611-46643508:+	-0.959	-0.907
ELP2_(78)|alt_cassette|chr18:31972387-31975097:+	-1.056	-0.906
ISOC2_(210)|alt_cassette|chr19:60658509-60659527:-	-1.435	-0.904
RBM39_(73)|alt_cassette|chr20:33790353-33792159:-	-0.596	0.902
RABGEF1_(117)|alt_cassette|chr7:65898014-65899795:+	0.928	-0.899
SORBS1_(162)|alt_cassette|chr10:97160524-97171707:-	-2.498	-0.897
FBLN2_(141)|alt_cassette|chr3:13636332-13642945:+	-1.292	-0.897
CLEC16A_(124)|alt_cassette|chr16:11167910-11179692:+	-0.758	0.897
PRIMA1_(171)|alt_cassette|chr14:93257645-93273339:-	-1.658	-0.894
PAXIP1_(163)|alt_cassette|chr7:154405975-154413648:-	0.683	0.894
PANK1_(177)|alt_cassette|chr10:91343057-91348988:-	0.992	-0.891
SFRS2_(104)|alt_cassette|chr17:72242835-72243830:-	0.816	0.89
AHRR_(54)|alt_cassette|chr5:477092-480921:+	1.429	-0.89
AUTS2_(72)|alt_cassette|chr7:69874566-69878278:+	-0.634	0.886
ADAMTS13_(185)|alt_cassette|chr9:135311662-135313916:+	-1.018	-0.886
BRP44L_(216)|alt_cassette|chr6:166700369-166716283:-	1.413	0.886
AGPAT2_(96)|alt_cassette|chr9:138689080-138691233:-	-0.593	0.886
NUMB_(33)|alt_cassette|chr14:72833746-72859590:-	-1.362	-0.884
MAST2_(116)|alt_cassette|chr1:46120654-46197644:+	-1.673	-0.883
EML1_(57)|alt_cassette|chr14:99401736-99414574:+	1.22	-0.88
TYMP_(99)|alt_cassette|chr22:49313012-49314430:-	1.059	0.88
BCAS3_(66)|alt_cassette|chr17:56800637-56824119:+	-0.571	-0.879
AAAS_(99)|alt_cassette|chr12:51994492-51995144:-	1.701	0.876
LRP12_(57)|alt_cassette|chr8:105590478-105670222:-	-1.13	-0.875
ITGA6_(130)|alt_cassette|chr2:173071074-173077064:+	-2.965	0.875
ZNF419_(99)|alt_cassette|chr19:62694777-62696035:+	1.268	0.872
PLCB1_(118)|alt_cassette|chr20:8718908-8810268:+	2.222	0.87
GUCY1A2_(93)|alt_cassette|chr11:106084602-106152374:-	2.433	0.867
CAMTA1_(31)|alt_cassette|chr1:7735180-7749105:+	-2.191	-0.866
C18orf25_(183)|alt_cassette|chr18:42050559-42087660:+	-0.962	-0.865
ATP11B_(86)|alt_cassette|chr3:184099254-184114342:+	-1.32	-0.864
TSGA14_(216)|alt_cassette|chr7:129826116-129827783:-	0.929	0.863
CALU_(815)|alt_cassette|chr7:128166722-128175862:+	-1.049	-0.863
ARHGAP4_(78)|alt_cassette|chrX:152829482-152829776:-	0.982	-0.863
LSR_(60)|alt_cassette|chr19:40449275-40449581:+	2.438	0.862

Table 3.5 (continued): Alternative cassette exon splicing event changes common to HFFs and NPCs

SPLICING_EVENT *hg18-based-coordinates*	HFF sepscore	NPC sepscore
RBPJ_(104)|alt_cassette|chr4:26040773-26041411:+	0.852	0.862
TSEN15_(117)|alt_cassette|chr1:182290620-182307913:+	1.314	0.861
ACTN2_(86)|alt_cassette|chr1:234961237-234965557:+	2.822	0.858
KIF13A_(39)|alt_cassette|chr6:17896085-17902458:-	-3.251	0.857
QRICH1_(60)|alt_cassette|chr3:49089475-49106348:-	-1.152	-0.854
PTRH2_(1319)|alt_cassette|chr17:55130121-55139513:-	-0.705	-0.852
NDRG2_(48)|alt_cassette|chr14:20556238-20556761:-	1.822	0.85
LSR_(57)|alt_cassette|chr19:40441807-40445288:+	-1.259	-0.847
RPRD2_(78)|alt_cassette|chr1:148680123-148682330:+	0.555	-0.847
INTS3_(90)|alt_cassette|chr1:152000018-152000669:+	0.678	0.846
IL15RA_(195)|alt_cassette|chr10:6042536-6059745:-	-1.121	0.845
CSGALNACT1_(287)|alt_cassette|chr8:19341722-19360216:-	-1.743	-0.845
ANGPTL2_(866)|alt_cassette|chr9:128896026-128924414:-	-4.489	-0.845
SETD8_(122)|alt_cassette|chr12:122434708-122441129:+	-1.071	-0.845
HSD17B14_(97)|alt_cassette|chr19:54008417-54008731:-	0.73	0.843
BRSK2_(66)|alt_cassette|chr11:1423713-1427581:+	1.976	-0.843
DOLPP1_(129)|alt_cassette|chr9:130887716-130888808:+	-1.944	-0.842
MAP3K9_(69)|alt_cassette|chr14:70270968-70274714:-	-3.787	-0.839
WDFY3_(51)|alt_cassette|chr4:85864771-85873558:-	-1.238	-0.838
ZNF521_(41)|alt_cassette|chr18:21156149-21185966:-	-0.915	-0.837
BPGM_(321)|alt_cassette|chr7:133982229-133996738:+	-1.074	-0.837
SLC27A2_(159)|alt_cassette|chr15:48277198-48284727:+	1.937	0.834
CD164_(57)|alt_cassette|chr6:109796913-109803969:-	1.219	0.833
CLRN1_(86)|alt_cassette|chr3:152142238-152144296:-	-0.636	-0.828
TBC1D22A_(57)|alt_cassette|chr22:45537409-45568061:+	0.736	0.828
MTMR9L_(111)|alt_cassette|chr1:32470372-32473139:-	0.695	0.827
MAP3K7IP3_(84)|alt_cassette|chrX:30762190-30774588:-	0.501	0.825
KIF13B_(61)|alt_cassette|chr8:29099908-29158781:-	1.694	0.821
PATL1_(90)|alt_cassette|chr11:59179789-59180547:-	0.969	0.82
PBX1_(113)|alt_cassette|chr1:163048010-163057397:+	0.567	-0.819
EXO1_(105)|alt_cassette|chr1:240078952-240080333:+	0.625	0.818
MTF2_(124)|alt_cassette|chr1:93357546-93365376:+	1.2	-0.815
JPH4_(125)|alt_cassette|chr14:23105468-23106159:-	0.776	0.812
PHTF2_(102)|alt_cassette|chr7:77307553-77360860:+	-2.711	-0.812
MACF1_(327)|alt_cassette|chr1:39665873-39667500:+	-0.783	0.811
TMEM214_(135)|alt_cassette|chr2:27111657-27112341:+	0.845	0.807
DMTF1_(123)|alt_cassette|chr7:86632302-86638246:+	-0.644	0.805
PHTF2_(156)|alt_cassette|chr7:77410039-77418869:+	0.957	-0.803
DOLPP1_(85)|alt_cassette|chr9:130886868-130887306:+	2.201	0.802
FRAS1_(138)|alt_cassette|chr4:79559241-79564565:+	-1.478	-0.798
FLJ43860_(103)|alt_cassette|chr8:142513907-142514408:-	2.96	0.795

Table 3.5 (continued): Alternative cassette exon splicing event changes common to HFFs and NPCs

SPLICING_EVENT *hg18-based-coordinates*	HFF sepscore	NPC sepscore
VPS13B_(67)|alt_cassette|chr8:100584354-100589173:+	-0.702	-0.795
MCC_(114)|alt_cassette|chr5:112506956-112657924:-	1.135	0.794
CENPO_(114)|alt_cassette|chr2:24869874-24876047:+	0.61	0.792
MKL1_(150)|alt_cassette|chr22:39147050-39150155:-	-1.39	0.792
KITLG_(84)|alt_cassette|chr12:87425045-87434241:-	-1.603	0.792
CRTC2_(407)|alt_cassette|chr1:152188484-152191117:-	0.665	0.79
FAM13B_(192)|alt_cassette|chr5:137382102-137384618:-	1.363	-0.79
PPP3CA_(30)|alt_cassette|chr4:102166241-102172446:-	0.577	-0.789
BAALC_(167)|alt_cassette|chr8:104222461-104309392:+	-1.082	-0.788
RAB18_(87)|alt_cassette|chr10:27855824-27861441:+	-0.699	0.783
MFGE8_(132)|alt_cassette|chr15:87251611-87257481:-	0.544	0.782
ELOVL7_(51)|alt_cassette|chr5:60119015-60175743:-	1.615	0.781
C11orf66_(60)|alt_cassette|chr11:61009959-61010571:+	2.404	0.781
ASPM_(2218)|alt_cassette|chr1:195331917-195353541:-	-0.632	-0.779
RGS19_(98)|alt_cassette|chr20:62178443-62181090:-	-2.134	-0.778
PITPNM2_(162)|alt_cassette|chr12:122039372-122041044:-	-2.027	-0.777
ARHGAP6_(90)|alt_cassette|chrX:11072289-11084569:-	0.762	-0.777
FAM135A_(78)|alt_cassette|chr6:71248578-71257457:+	-1.797	-0.777
C19orf66_(268)|alt_cassette|chr19:10058976-10061332:+	1.324	-0.775
NEDD4L_(132)|alt_cassette|chr18:54152104-54159889:+	-3.456	-0.775
C14orf101_(115)|alt_cassette|chr14:56152498-56155064:+	-0.526	-0.774
ASNS_(198)|alt_cassette|chr7:97331744-97339598:-	-1.27	0.771
LMO7_(109)|alt_cassette|chr13:75325525-75328642:+	-1.531	0.771
RICTOR_(72)|alt_cassette|chr5:38983300-38985570:-	-0.937	0.771
WBSCR22_(51)|alt_cassette|chr7:72746316-72749870:+	-0.655	-0.768
EPB41_(42)|alt_cassette|chr1:29257744-29259520:+	-1.086	-0.765
PAPOLG_(112)|alt_cassette|chr2:60851150-60855672:+	1.492	0.764
GLCE_(90)|alt_cassette|chr15:67240150-67275807:+	0.59	0.764
NIN_(61)|alt_cassette|chr14:50266191-50271983:-	-0.731	-0.764
ABCA3_(174)|alt_cassette|chr16:2294152-2307284:-	0.581	0.763
LMBR1_(162)|alt_cassette|chr7:156169618-156209566:-	-0.543	-0.763
FAM107A_(98)|alt_cassette|chr3:58530632-58538075:-	1.918	0.762
TRPM4_(435)|alt_cassette|chr19:54385842-54395368:+	1.237	0.761
PLEKHO1_(102)|alt_cassette|chr1:148395833-148397637:+	-0.617	-0.761
L3MBTL2_(105)|alt_cassette|chr22:39953273-39953753:+	-4.635	-0.759
SMEK2_(96)|alt_cassette|chr2:55657996-55660318:-	0.726	-0.759
TMEM63B_(39)|alt_cassette|chr6:44210811-44212063:+	0.933	0.759
RARS2_(56)|alt_cassette|chr6:88312136-88321844:-	-0.673	0.758
C12orf29_(118)|alt_cassette|chr12:86953646-86960732:+	0.523	-0.756
PRC1_(42)|alt_cassette|chr15:89311436-89313680:-	-0.758	0.756
PAQR4_(105)|alt_cassette|chr16:2959842-2961516:+	-1.094	-0.756
TNIK_(87)|alt_cassette|chr3:172338862-172340881:-	0.564	-0.753

Table 3.5 (continued): Alternative cassette exon splicing event changes common to HFFs and NPCs

SPLICING_EVENT *hg18-based-coordinates*	HFF sepscore	NPC sepscore
OSBPL6_(108)|alt_cassette|chr2:178934788-178946862:+	0.865	-0.751
SEC24D_(843)|alt_cassette|chr4:119873515-119878863:-	1.136	-0.749
SSFA2_(54)|alt_cassette|chr2:182492418-182494919:+	0.901	0.748
PDCD1_(156)|alt_cassette|chr2:242442808-242443445:-	1.165	0.747
TMEM11_(292)|alt_cassette|chr17:21042745-21057995:-	1.672	0.747
RBM3_(269)|alt_cassette|chrX:48318999-48319645:+	3.876	0.747
CHCHD4_(170)|alt_cassette|chr3:14133025-14141155:-	-0.863	-0.746
VLDLR_(84)|alt_cassette|chr9:2640516-2641873:+	-1.223	-0.746
SPAG9_(39)|alt_cassette|chr17:46407307-46409467:-	-4.393	-0.744
STARD8_(82)|alt_cassette|chrX:67849552-67852934:+	0.916	0.744
TIAL1_(376)|alt_cassette|chr10:121329512-121331423:-	0.55	0.744
PIGA_(133)|alt_cassette|chrX:15253195-15259632:-	0.581	0.744
PCCA_(141)|alt_cassette|chr13:99899560-99977929:+	0.797	0.742
NPTN_(348)|alt_cassette|chr15:71671531-71712518:-	-1.329	-0.742
KIAA0802_(57)|alt_cassette|chr18:8816230-8821604:+	3.244	0.74
POMGNT1_(118)|alt_cassette|chr1:46432910-46434069:-	1.215	0.74
MRC2_(239)|alt_cassette|chr17:58120053-58120986:+	-2.014	-0.74
ACLY_(30)|alt_cassette|chr17:37302953-37307527:-	-0.733	0.74
YIPF1_(220)|alt_cassette|chr1:54127247-54127973:-	-0.699	0.739
PLAG1_(105)|alt_cassette|chr8:57246302-57286233:-	-2.433	0.738
SNX7_(153)|alt_cassette|chr1:98940047-98998197:+	1.248	0.737
HEATR7A_(120)|alt_cassette|chr8:145274967-145290661:+	4.694	-0.736
RPS24_(245)|alt_cassette|chr10:79467068-79484294:+	-0.659	0.736
ZNF549_(39)|alt_cassette|chr19:62730786-62738323:+	1.477	0.736
SCAPER_(67)|alt_cassette|chr15:74921327-74937204:-	-0.786	-0.735
SHARPIN_(125)|alt_cassette|chr8:145225885-145226167:-	1.106	0.734
TCOF1_(114)|alt_cassette|chr5:149749779-149751712:+	0.724	0.734
MAZ_(89)|alt_cassette|chr16:29725715-29727391:+	0.852	0.732
TMEM130_(306)|alt_cassette|chr7:98295897-98305335:-	-1.206	0.731
CCDC91_(86)|alt_cassette|chr12:28303642-28349848:+	-0.609	0.73
ANXA1_(121)|alt_cassette|chr9:74956660-74963257:+	3.654	0.729
MAPKAP1_(328)|alt_cassette|chr9:127472007-127509070:-	1.116	0.728
EBPL_(121)|alt_cassette|chr13:49141983-49163390:-	-1.129	-0.727
HEXA_(68)|alt_cassette|chr15:70369650-70384140:-	-0.83	-0.726
GRAMD1A_(113)|alt_cassette|chr19:40194298-40196284:+	-0.571	0.724
ZNF711_(138)|alt_cassette|chrX:84406917-84409955:+	0.568	0.721
UBE2D3_(50)|alt_cassette|chr4:103937701-103939675:-	-2.182	-0.721
C8orf31_(136)|alt_cassette|chr8:144197609-144202000:+	2.496	0.718
KAT5_(156)|alt_cassette|chr11:65237105-65237639:+	0.525	0.718
ATF2_(74)|alt_cassette|chr2:175653739-175666034:-	0.507	0.716
RAB28_(130)|alt_cassette|chr4:12992316-13085039:-	0.573	0.715
ABHD6_(65)|alt_cassette|chr3:58198683-58217328:+	2.758	0.715

Table 3.5 (continued): Alternative cassette exon splicing event changes common to HFFs and NPCs

SPLICING_EVENT *hg18-based-coordinates*	HFF sepscore	NPC sepscore
SRRT_(74)|alt_cassette|chr7:100311269-100316841:+	-0.596	0.714
MBD1_(138)|alt_cassette|chr18:46053323-46053931:-	1.466	0.714
CAMTA2_(95)|alt_cassette|chr17:4830002-4831611:-	-1.699	0.71
DPH3_(75)|alt_cassette|chr3:16277340-16281279:-	-0.657	0.71
MYLK_(153)|alt_cassette|chr3:124828478-124839607:-	-4.886	-0.709
CHID1_(93)|alt_cassette|chr11:874169-889339:-	-0.879	-0.709
ZYG11A_(752)|alt_cassette|chr1:53092890-53098990:+	1.799	0.709
TRAF6_(93)|alt_cassette|chr11:36479663-36488173:-	-1.503	-0.707
RALGAPA1_(31)|alt_cassette|chr14:35078647-35088066:-	-2.116	0.707
WDR45_(105)|alt_cassette|chrX:48821353-48822439:-	-0.98	0.706
C8orf59_(38)|alt_cassette|chr8:86316983-86319786:-	0.548	0.705
C8orf74_(192)|alt_cassette|chr8:10569758-10592518:+	1.18	0.705
C12orf73_(121)|alt_cassette|chr12:102869538-102874538:-	-2.282	-0.705
APLP2_(168)|alt_cassette|chr11:129497618-129501804:+	1.138	0.703
COL16A1_(48)|alt_cassette|chr1:31917873-31918229:-	0.886	0.703
TPRG1L_(177)|alt_cassette|chr1:3531939-3533923:+	-0.57	0.702
GTF2A1_(67)|alt_cassette|chr14:80740201-80752509:-	-0.904	-0.701
PLCH1_(60)|alt_cassette|chr3:156686193-156688510:-	-0.527	0.701
IL15RA_(99)|alt_cassette|chr10:6042536-6048113:-	2.385	0.701
USP21_(141)|alt_cassette|chr1:159396092-159397033:+	1.053	0.7
WDR51B_(172)|alt_cassette|chr12:88410023-88443027:-	1.01	-0.7
SLC24A6_(168)|alt_cassette|chr12:112241259-112242768:-	0.648	0.699
ATM_(88)|alt_cassette|chr11:107599123-107603531:+	-0.625	-0.699
MTA1_(51)|alt_cassette|chr14:104982893-104987439:+	-1.055	-0.698
ARHGAP8_(88)|alt_cassette|chr22:43561150-43582850:+	0.798	0.697
SBDSP_(122)|alt_cassette|chr7:71938298-71940117:+	0.625	0.696
C7orf63_(124)|alt_cassette|chr7:89739208-89744322:+	-2.123	0.695
PDGFRA_(45)|alt_cassette|chr4:53938847-53940148:+	-1.026	0.695
RALYL_(173)|alt_cassette|chr8:85948187-85995683:+	-2.14	-0.694
SAMD14_(84)|alt_cassette|chr17:45546669-45547926:-	1.066	0.692
C8orf31_(130)|alt_cassette|chr8:144192361-144195969:+	3.443	0.692
FGFR2_(78)|alt_cassette|chr10:123300963-123343212:-	1.69	-0.69
RNPEP_(149)|alt_cassette|chr1:200224795-200231897:+	-1.223	0.688
TRIT1_(109)|alt_cassette|chr1:40085556-40086244:-	-0.965	0.687
HDAC10_(29)|alt_cassette|chr22:49029307-49029658:-	0.554	0.686
TMEM39A_(109)|alt_cassette|chr3:120648709-120653983:-	-1.136	-0.686
TMED5_(49)|alt_cassette|chr1:93393033-93394444:-	-1.381	-0.686
ARRB2_(45)|alt_cassette|chr17:4566077-4566455:+	1.025	0.684
ACO1_(68)|alt_cassette|chr9:32374733-32395482:+	-1.113	-0.683
SLC16A10_(32)|alt_cassette|chr6:111605561-111634487:+	-0.694	-0.682
CNOT10_(81)|alt_cassette|chr3:32751472-32775953:+	0.748	-0.681
ACP1_(29)|alt_cassette|chr2:261939-262191:+	1.945	0.678

Table 3.5 (continued): Alternative cassette exon splicing event changes common to HFFs and NPCs

SPLICING_EVENT *hg18-based-coordinates*	HFF sepscore	NPC sepscore
FGFR1OP2_(114)|alt_cassette|chr12:27001943-27007541:+	-1.103	-0.676
CYP21A2_(90)|alt_cassette|chr6:32114380-32114849:+	0.895	0.676
RBM26_(72)|alt_cassette|chr13:78816930-78826574:-	-1.085	0.675
SYNE2_(69)|alt_cassette|chr14:63750941-63752718:+	-3.866	-0.675
ERMP1_(128)|alt_cassette|chr9:5814001-5820726:-	0.85	0.674
DIS3_(68)|alt_cassette|chr13:72253142-72253743:-	-1.14	0.673
ARHGAP21_(30)|alt_cassette|chr10:24950304-24958930:-	-0.867	0.672
ABCB8_(313)|alt_cassette|chr7:150356630-150362292:+	2.573	0.672
DSN1_(321)|alt_cassette|chr20:34829859-34835210:-	1.402	0.672
CDK8_(58)|alt_cassette|chr13:25826017-25857347:+	1.204	0.671
ZMYND8_(131)|alt_cassette|chr20:45345799-45349379:-	1.151	0.669
POLDIP3_(87)|alt_cassette|chr22:41325743-41328719:-	-2.335	-0.666
ALB_(68)|alt_cassette|chr4:74504220-74505672:+	1.86	0.665
RAD1_(109)|alt_cassette|chr5:34947674-34950556:-	0.93	-0.664
C17orf85_(33)|alt_cassette|chr17:3672095-3674984:-	-2.391	0.662
GFRA3_(93)|alt_cassette|chr5:137621539-137627848:-	1.6	0.661
PLOD2_(63)|alt_cassette|chr3:147277372-147279592:-	-3.986	0.66
SIL1_(88)|alt_cassette|chr5:138491441-138561856:-	1.395	-0.66
PPP1R12A_(168)|alt_cassette|chr12:78723679-78725136:-	-1.188	0.66
BEND7_(170)|alt_cassette|chr10:13582086-13584866:-	-0.684	0.659
NR1D2_(107)|alt_cassette|chr3:23979100-23981471:+	0.685	-0.657
KIAA1324L_(34)|alt_cassette|chr7:86407367-86409269:-	-2.587	0.657
MFSD9_(115)|alt_cassette|chr2:102713959-102715212:-	-1.536	-0.655
ZCWPW2_(99)|alt_cassette|chr3:28451764-28495308:+	0.916	-0.654
ADC_(174)|alt_cassette|chr1:33319696-33330237:+	-1.896	0.654
RAB5A_(170)|alt_cassette|chr3:19967557-19992096:+	-0.752	-0.654
CACNA1A_(36)|alt_cassette|chr19:13181312-13183916:-	-1.479	-0.653
MBOAT7_(127)|alt_cassette|chr19:59379375-59383882:-	3.767	0.653
SFXN4_(121)|alt_cassette|chr10:120890821-120895737:-	-0.604	-0.653
ST7OT3_(136)|alt_cassette|chr7:116616683-116618123:+	-0.652	-0.651
TNFRSF13B_(138)|alt_cassette|chr17:16793022-16816053:-	1.398	0.65
PLSCR2_(113)|alt_cassette|chr3:147660560-147696371:-	1.655	0.648
SLC30A6_(49)|alt_cassette|chr2:32272556-32276279:+	1.94	0.648
NLGN3_(60)|alt_cassette|chrX:70284781-70291788:+	2.555	-0.646
TMEM159_(72)|alt_cassette|chr16:21080100-21089258:+	0.83	-0.645
ERCC8_(218)|alt_cassette|chr5:60253739-60260447:-	0.669	0.644
F3_(495)|alt_cassette|chr1:94778512-94779680:-	1.127	-0.643
PML_(144)|alt_cassette|chr15:72104321-72112549:+	-1.348	-0.642
DMD_(39)|alt_cassette|chrX:31101642-31106706:-	0.794	0.64
SSBP3_(81)|alt_cassette|chr1:54495447-54519698:-	2.383	-0.64

Table 3.5 (continued): Alternative cassette exon splicing event changes common to HFFs and NPCs

SPLICING_EVENT *hg18-based-coordinates*	HFF sepscore	NPC sepscore
HEXA_(73)|alt_cassette|chr15:70426105-70427442:-	-1.27	0.634
BAT2L_(82)|alt_cassette|chr9:133357468-133360947:+	-0.914	0.634
RBM34_(137)|alt_cassette|chr1:233385050-233390830:-	0.822	-0.633
NRM_(197)|alt_cassette|chr6:30765208-30766597:-	-1.154	-0.632
MYBL1_(180)|alt_cassette|chr8:67639614-67641472:-	2.695	-0.632
UCP2_(181)|alt_cassette|chr11:73363814-73365333:-	1.351	0.631
GPR177_(267)|alt_cassette|chr1:68397518-68470464:-	0.8	0.631
BCAT1_(111)|alt_cassette|chr12:24922850-24938475:-	-1.153	0.629
RCOR3_(242)|alt_cassette|chr1:209544105-209553388:+	1.04	0.625
MEIS1_(96)|alt_cassette|chr2:66649392-66651881:+	-0.626	-0.624
EZH2_(117)|alt_cassette|chr7:148157873-148174494:-	0.784	0.623
SGOL1_(51)|alt_cassette|chr3:20187728-20193096:-	1.034	-0.622
ESCO1_(124)|alt_cassette|chr18:17373968-17394817:-	-1.072	-0.62
HTRA2_(96)|alt_cassette|chr2:74612560-74613454:+	1.33	0.62
RBM23_(54)|alt_cassette|chr14:22444502-22445243:-	0.828	0.619
APP_(57)|alt_cassette|chr21:26276661-26294200:-	-1.032	0.618
RIC8B_(123)|alt_cassette|chr12:105778320-105788730:+	-0.712	-0.617
INPP5K_(102)|alt_cassette|chr17:1359843-1363915:-	-0.699	0.616
IKBIP_(118)|alt_cassette|chr12:97532249-97562431:-	-0.91	-0.616
MINA_(101)|alt_cassette|chr3:99152422-99155925:-	-1.116	-0.616
GK5_(1083)|alt_cassette|chr3:143384581-143387460:-	0.706	0.614
FAM49B_(106)|alt_cassette|chr8:130937175-130952802:-	1.244	0.614
SERGEF_(167)|alt_cassette|chr11:17856367-17966719:-	0.712	0.614
C9orf6_(130)|alt_cassette|chr9:110737790-110741296:+	0.855	-0.614
FAM21C_(153)|alt_cassette|chr10:45588812-45594406:+	-0.652	0.613
STAG3L4_(76)|alt_cassette|chr7:66408505-66411356:+	0.905	-0.611
ABHD12B_(128)|alt_cassette|chr14:50408746-50415212:+	-0.539	0.611
HDAC9_(132)|alt_cassette|chr7:18600177-18640774:+	0.596	0.61
NR1H3_(180)|alt_cassette|chr11:47239576-47240045:+	-0.603	0.61
PHC2_(167)|alt_cassette|chr1:33563207-33567057:-	1.117	0.61
UBAC1_(128)|alt_cassette|chr9:137971426-137976694:-	-0.992	-0.609
URG4_(119)|alt_cassette|chr7:43885384-43887782:-	1.537	-0.607
POLI_(126)|alt_cassette|chr18:50050400-50054293:+	-1.454	0.605
USP10_(61)|alt_cassette|chr16:83324610-83335739:+	-1.012	-0.605
QDPR_(93)|alt_cassette|chr4:17115196-17122670:-	0.845	0.604
CLSTN1_(57)|alt_cassette|chr1:9718687-9723738:-	-2.943	-0.601
AP3M2_(61)|alt_cassette|chr8:42131635-42134615:+	0.72	-0.601
SLC19A2_(603)|alt_cassette|chr1:167706048-167721424:-	-1.702	0.601
EIF4EBP3_(57)|alt_cassette|chr5:139824545-139842348:+	-0.742	0.599
MST4_(231)|alt_cassette|chrX:130985368-131025141:+	2.536	0.599
CORO1A_(118)|alt_cassette|chr16:30102547-30104030:+	-0.999	0.597
NCOR1_(174)|alt_cassette|chr17:15930481-15935901:-	0.524	-0.596

Table 3.5 (continued): Alternative cassette exon splicing event changes common to HFFs and NPCs

SPLICING_EVENT *hg18-based-coordinates*	HFF sepscore	NPC sepscore
SBF1_(78)|alt_cassette|chr22:49241968-49244549:-	-0.709	-0.595
SLC25A30_(148)|alt_cassette|chr13:44878112-44883496:-	0.942	0.595
SNURF_(316)|alt_cassette|chr15:22770696-22771597:+	-0.914	-0.594
ZMYND8_(114)|alt_cassette|chr20:45354039-45360879:-	-1.754	-0.593
MADD_(70)|alt_cassette|chr11:47302704-47306784:+	-0.51	-0.593
PPM1L_(287)|alt_cassette|chr3:162162392-162265884:+	-1.901	0.592
PKP2_(132)|alt_cassette|chr12:32885406-32894966:-	-2.628	-0.592
CDK5RAP2_(237)|alt_cassette|chr9:122206212-122210445:-	-2.476	0.592
MOSPD1_(162)|alt_cassette|chrX:133850888-133858512:-	-0.686	0.59
KIF13A_(105)|alt_cassette|chr6:17873156-17880117:-	2.862	-0.589
RABL5_(90)|alt_cassette|chr7:100746543-100748956:-	1.403	0.588
UHRF2_(99)|alt_cassette|chr9:6476925-6483825:+	0.509	0.587
KIAA1462_(3764)|alt_cassette|chr10:30346852-30376466:-	-3.337	0.586
DCHS2_(36)|alt_cassette|chr4:155507043-155517854:-	-0.909	-0.585
SEC31B_(196)|alt_cassette|chr10:102258849-102265842:-	1.743	0.584
MAPT_(198)|alt_cassette|chr17:41420298-41424662:+	0.555	-0.583
KIF21A_(39)|alt_cassette|chr12:38022193-38031845:-	0.775	-0.583
RNF32_(107)|alt_cassette|chr7:156130212-156140030:+	-0.621	-0.583
NCRNA00188_(57)|alt_cassette|chr17:16284292-16285406:+	1.036	0.583
KANK1_(69)|alt_cassette|chr9:724835-728284:+	0.782	-0.582
CYP4A22_(155)|alt_cassette|chr1:47381652-47382615:+	0.655	-0.582
ALG11_(932)|alt_cassette|chr13:51484599-51500455:+	0.516	-0.582
CLTB_(54)|alt_cassette|chr5:175752552-175757213:-	0.904	-0.581
KIDINS220_(114)|alt_cassette|chr2:8794580-8807692:-	-0.502	0.578
TERT_(189)|alt_cassette|chr5:1313715-1319578:-	-3.214	-0.578
GIT2_(105)|alt_cassette|chr12:108869692-108875279:-	1.104	0.577
TSC22D2_(72)|alt_cassette|chr3:151611785-151657552:+	-0.826	-0.577
LOC145783_(62)|alt_cassette|chr15:54783757-54786752:-	-0.512	0.577
ANKRD13A_(114)|alt_cassette|chr12:108951823-108955984:+	-0.598	0.577
TRIM3_(165)|alt_cassette|chr11:6443538-6451596:-	-1.056	-0.577
ITSN1_(167)|alt_cassette|chr21:34113497-34120991:+	-1.048	-0.577
DYNC1I2_(60)|alt_cassette|chr2:172272133-172280083:+	2.171	-0.576
DCLRE1C_(56)|alt_cassette|chr10:15017569-15021814:-	-0.623	0.576
SAMD4B_(558)|alt_cassette|chr19:44525056-44539168:+	-2.074	-0.575
OTOF_(198)|alt_cassette|chr2:26534592-26537018:-	-0.518	-0.573
TBC1D2_(142)|alt_cassette|chr9:100046232-100057275:-	0.723	0.572
ITSN1_(83)|alt_cassette|chr21:34121037-34123797:+	1.887	0.569
MYEF2_(72)|alt_cassette|chr15:46228860-46230961:-	-0.503	0.568
KLHL12_(187)|alt_cassette|chr1:201128410-201129934:-	0.86	0.568

Table 3.5 (continued): Alternative cassette exon splicing event changes common to HFFs and NPCs

SPLICING_EVENT *hg18-based-coordinates*	HFF sepscore	NPC sepscore
NWD1_(106)|alt_cassette|chr19:16779947-16786838:+	-1.932	-0.567
CNN2_(117)|alt_cassette|chr19:983695-987414:+	-1.952	-0.567
FAM98C_(168)|alt_cassette|chr19:43588115-43591230:+	0.818	0.563
PTBP1_(34)|alt_cassette|chr19:757556-759359:+	1.267	-0.562
PACSIN2_(123)|alt_cassette|chr22:41602283-41604997:-	0.939	0.562
ANO4_(141)|alt_cassette|chr12:99997184-100004568:+	1.247	0.561
NFIX_(148)|alt_cassette|chr19:13053669-13062112:+	-0.564	-0.561
NAT12_(772)|alt_cassette|chr14:56927176-56933222:+	0.634	0.557
AMBRA1_(279)|alt_cassette|chr11:46490939-46520070:-	0.824	0.557
TTYH1_(46)|alt_cassette|chr19:59638676-59639082:+	1.456	-0.557
KIAA0586_(228)|alt_cassette|chr14:57995016-57997456:+	1.169	0.556
BAZ1A_(96)|alt_cassette|chr14:34324958-34331731:-	-1.194	-0.555
NCAM1_(30)|alt_cassette|chr11:112590443-112607576:+	1.142	-0.555
PCMT1_(203)|alt_cassette|chr6:150134095-150135966:+	-1.315	-0.554
TMEM168_(91)|alt_cassette|chr7:112212244-112217450:-	1.923	0.554
BCL2L13_(98)|alt_cassette|chr22:16546087-16551751:+	0.833	-0.554
ZNF142_(122)|alt_cassette|chr2:219223493-219229116:-	-1.311	-0.553
BC038559_(743)|alt_cassette|chr9:131295695-131305422:+	1.397	0.553
STXBP1_(126)|alt_cassette|chr9:129484660-129492874:+	-1.228	-0.552
ZNF846_(83)|alt_cassette|chr19:9730440-9733757:-	-0.65	0.552
ATP13A2_(111)|alt_cassette|chr1:17195382-17196101:-	0.739	0.552
PPP2R3C_(128)|alt_cassette|chr14:34649586-34660858:-	0.86	-0.551
TSC22D2_(157)|alt_cassette|chr3:151623597-151657552:+	-0.934	-0.551
MAD1L1_(176)|alt_cassette|chr7:2236304-2239032:-	-3.679	-0.55
NFASC_(117)|alt_cassette|chr1:203224557-203238346:+	-4.126	-0.549
VAPB_(104)|alt_cassette|chr20:56426825-56447406:+	-0.503	0.547
R3HDM2_(32)|alt_cassette|chr12:55964106-55975448:-	-1.401	0.547
D2HGDH_(384)|alt_cassette|chr2:242322783-242329120:+	-3.301	0.547
RAG1AP1_(162)|alt_cassette|chr1:153376051-153377078:+	1.192	0.546
SMTN_(93)|alt_cassette|chr22:29817833-29821288:+	3.012	0.546
ST6GALNAC4_(87)|alt_cassette|chr9:129716941-129718925:-	0.881	0.546
LP9056_(120)|alt_cassette|chr4:53306867-53311853:-	-1.989	-0.545
CCAR1_(211)|alt_cassette|chr10:70185299-70187056:+	-1.387	-0.542
AK125149_(45)|alt_cassette|chrX:2547538-2551425:+	-2.59	-0.542
NDEL1_(35)|alt_cassette|chr17:8304203-8310972:+	-2.808	0.542
BIN3_(41)|alt_cassette|chr8:22544036-22558212:-	-1.12	0.542
HCRTR1_(341)|alt_cassette|chr1:31857004-31859030:+	-0.846	0.542
SIRT1_(125)|alt_cassette|chr10:69321318-69336552:+	0.624	-0.541
Q53QX8_(110)|alt_cassette|chr2:160932148-161058045:-	-1.141	-0.541
HMGN2_(45)|alt_cassette|chr1:26671705-26672780:+	-0.835	-0.539

Table 3.5 (continued): Alternative cassette exon splicing event changes common to HFFs and NPCs

SPLICING_EVENT *hg18-based-coordinates*	HFF sepscore	NPC sepscore
SIPA1L2_(54)|alt_cassette|chr1:230605940-230617862:-	-2.127	0.539
COL6A3_(600)|alt_cassette|chr2:237950726-237954296:-	-1.917	-0.539
PPARGC1B_(117)|alt_cassette|chr5:149186641-149192411:+	0.825	0.538
ZDHHC20_(219)|alt_cassette|chr13:20859773-20863860:-	1.377	0.538
MFAP5_(45)|alt_cassette|chr12:8694452-8699277:-	-0.55	-0.538
SEC31A_(39)|alt_cassette|chr4:83969235-83982316:-	-1.177	0.537
ELF1_(118)|alt_cassette|chr13:40423572-40430971:-	1.356	-0.537
SMC2_(193)|alt_cassette|chr9:105936677-105940119:+	-1.004	-0.537
KIAA0182_(219)|alt_cassette|chr16:84204505-84239658:+	0.999	0.536
OSBPL9_(39)|alt_cassette|chr1:51988455-51998949:+	-1.779	0.536
ITGA7_(193)|alt_cassette|chr12:54374746-54375572:-	-0.512	-0.535
PRMT2_(109)|alt_cassette|chr21:46880103-46881235:+	-0.781	-0.533
TMEM149_(113)|alt_cassette|chr19:40927162-40927865:-	1.436	0.532
DLG1_(36)|alt_cassette|chr3:198278022-198280486:-	0.502	-0.531
EDIL3_(30)|alt_cassette|chr5:83512095-83585657:-	1.082	0.53
NELF_(90)|alt_cassette|chr9:139467453-139468663:-	1.94	0.53
TCF20_(128)|alt_cassette|chr22:40887308-40895796:-	-0.984	-0.529
SHOC2_(138)|alt_cassette|chr10:112714809-112750162:+	1.044	0.528
HKR1_(127)|alt_cassette|chr19:42507007-42530518:+	-1.29	0.528
XPNPEP1_(72)|alt_cassette|chr10:111623175-111627537:-	0.505	0.528
OBFC1_(85)|alt_cassette|chr10:105647467-105648624:-	1.762	-0.527
EDC3_(89)|alt_cassette|chr15:72754536-72775273:-	0.681	0.527
MYO5B_(78)|alt_cassette|chr18:45634015-45643594:-	-0.534	-0.526
HDLBP_(138)|alt_cassette|chr2:241843246-241844271:-	0.729	0.524
BCAP29_(79)|alt_cassette|chr7:107041113-107046008:+	0.73	-0.524
PHF19_(88)|alt_cassette|chr9:122663273-122664686:-	0.689	0.523
BTN2A1_(101)|alt_cassette|chr6:26568035-26571450:+	-1.142	-0.522
CD3D_(132)|alt_cassette|chr11:117715419-117716299:-	-3.301	-0.52
HSPH1_(123)|alt_cassette|chr13:30625088-30627650:-	1.643	0.52
TMEM218_(78)|alt_cassette|chr11:124476409-124477237:-	0.864	0.52
KHK_(117)|alt_cassette|chr2:27163723-27173100:+	1.625	0.52
CCL7_(118)|alt_cassette|chr17:29621498-29622828:+	-0.514	0.519
TGFB2_(84)|alt_cassette|chr1:216587012-216645133:+	-1.047	-0.519
TMED3_(115)|alt_cassette|chr15:77393402-77401374:+	0.614	-0.519
HAUS2_(93)|alt_cassette|chr15:40628455-40638828:+	0.699	-0.516
SCUBE2_(171)|alt_cassette|chr11:9009048-9025478:-	1.604	-0.516
ROGDI_(55)|alt_cassette|chr16:4790580-4791504:-	3.143	0.516
ZFYVE9_(177)|alt_cassette|chr1:52502132-52506722:+	0.665	0.516
IQCG_(77)|alt_cassette|chr3:199156954-199161106:-	-0.542	0.513
TPD52L2_(42)|alt_cassette|chr20:61984617-61990986:+	0.528	-0.512
DFFB_(124)|alt_cassette|chr1:3778996-3789930:+	-0.521	-0.511

Table 3.5 (continued): Alternative cassette exon splicing event changes common to HFFs and NPCs

SPLICING_EVENT *hg18-based-coordinates*	HFF sepscore	NPC sepscore
TCF12_(72)|alt_cassette|chr15:55330913-55332751:+	2.401	-0.511
RAB3IP_(100)|alt_cassette|chr12:68481768-68493010:+	1.277	-0.51
GPATCH8_(45)|alt_cassette|chr17:39869439-39897365:-	-1.183	-0.51
NCDN_(68)|alt_cassette|chr1:35796071-35797294:+	0.523	0.509
CNTLN_(115)|alt_cassette|chr9:17477064-17492548:+	1.241	0.509
FAM173B_(51)|alt_cassette|chr5:10280759-10289589:-	-0.838	-0.508
TRA2B_(276)|alt_cassette|chr3:187127216-187138306:-	0.687	0.507
DARS_(58)|alt_cassette|chr2:136453406-136459442:-	-1.612	-0.506
LOC284837_(651)|alt_cassette|chr21:44050897-44054539:-	-1.258	0.505
CATSPER2_(126)|alt_cassette|chr15:41638422-41647952:-	-0.518	0.505
RABGAP1L_(283)|alt_cassette|chr1:173193309-173205036:+	-1.111	-0.505
ZFX_(128)|alt_cassette|chrX:24080868-24100752:+	1.105	0.505
DPF2_(42)|alt_cassette|chr11:64868116-64869712:+	-0.862	-0.505
CACNA1E_(129)|alt_cassette|chr1:180026315-180030622:+	-0.975	-0.504
NUP98_(222)|alt_cassette|chr11:3661247-3669152:-	0.669	-0.504
SLC29A2_(134)|alt_cassette|chr11:65890278-65891510:-	0.925	0.504
SIK3_(144)|alt_cassette|chr11:116251377-116252178:-	0.908	0.504
CPSF7_(77)|alt_cassette|chr11:60944621-60945437:-	0.654	0.503
HDAC10_(85)|alt_cassette|chr22:49027522-49028445:-	-0.891	0.503
PGM3_(206)|alt_cassette|chr6:83955236-83959616:-	1.055	-0.502
CHEK1_(224)|alt_cassette|chr11:125001938-125004336:+	0.736	0.501

Table 3.6: Significantly changing alternative host polyadenylation events in 2 or more cell types

MISO-detected events are listed that were significant in at least 2 of the 96 hpi cell types. Events had to involve the same 3′-proximal cleavage site location for this comparison. Analysis contributing to Figure 3.7D only utilized occurrences of canonical CPE and PAS motifs, but here non-canonical (nc) PAS (or 'alt-PAS') sites are listed for some events. ncPAS includes TATAAA, AGTAAA, AAGAAA, AATATA, AATACA, CATAAA, GATAAA, GATAAA, AATGAA, ACTAAA, and AATAGA. The 3UTR column indicates the effect upon infection (S=shortened, L=lengthened). PAS+CPE co-occurrence listed are within +/- 500bp of the proximal termination site.

Gene Name	3UTR	EnsemblID:Proximal Site	Cell Types	PAS+CPE
PRDX6	S	ENSG00000117592:173457150	EC,NPC,	no-co-occur
NEDD4L	S	ENSG00000049759:56063805	EC,HFF,	CPE_PAS
SYNRG	S	ENSG00000006114:35878566	EC,HFF,NPC,	CPE_PAS
CAV2	S	ENSG00000105971:116146988	EC,HFF,	CPE_PAS
EXT2	S	ENSG00000151348:44266481	EC,NPC,	CPE_PAS
MLLT6	S	ENSG00000108292:36883448	EC,HFF,	CPE_PAS
SOS1	S	ENSG00000115904:39212252	EC,HFF,NPC,	no-co-occur
CDK16	S	ENSG00000102225:47088389	EC,NPC,	no-co-occur
C12orf29	S	ENSG00000133641:88442225	EC,HFF,	CPE_PAS
ANKH	S	ENSG00000154122:14711175	HFF,NPC,	no-co-occur
CDC42BPA	S	ENSG00000143776:227180424	EC,HFF,	no-co-occur
LRPAP1	S	ENSG00000163956:3514294	EC,HFF,NPC,	no-co-occur
GNPNAT1	S	ENSG00000100522:53244020	EC,HFF,	CPE_PAS
PPM1A	S	ENSG00000100614:60760070	HFF,NPC,	CPE_PAS
BRK1	S	ENSG00000254999:10168601	EC,NPC,	CPE_PAS
TRAK2	S	ENSG00000115993:202244184	EC,HFF,	CPE_PAS
ATP5G3	S	ENSG00000154518:176042873	HFF,NPC,	CPE_PAS
ANXA7	S	ENSG00000138279:75135539	EC,HFF,NPC,	CPE_PAS
PCGF3	S	ENSG00000185619:760386	EC,HFF,	CPE_ncPAS
OGFOD3	S	ENSG00000181396:80350099	EC,HFF,NPC,	no-co-occur
ICMT	S	ENSG00000116237:6283553	EC,HFF,	CPE_ncPAS
NAA50	S	ENSG00000121579:113439055	EC,HFF,	CPE_ncPAS
DCUN1D3	S	ENSG00000188215:20869402	HFF,NPC,	no-co-occur
MORF4L1	S	ENSG00000185787:79189547	EC,HFF,	CPE_PAS
LIN54	S	ENSG00000189308:83846873	EC,NPC,	no-co-occur
ANXA4	S	ENSG00000196975:70052909	EC,HFF,	no-co-occur
PCYT1A	S	ENSG00000161217:195965254	EC,HFF,	no-co-occur
STX3	S	ENSG00000166900:59570176	EC,HFF,	CPE_PAS
SETX	S	ENSG00000107290:135139281	EC,NPC,	no-co-occur
CMPK1	S	ENSG00000162368:47843538	EC,HFF,	CPE_PAS
TNFAIP1	S	ENSG00000109079:26672105	EC,HFF,	no-co-occur
FAM104A	S	ENSG00000133193:71205259	EC,HFF,	CPE_PAS
WHSC1	S	ENSG00000109685:1981522	EC,NPC,	CPE_ncPAS

Table 3.6 (continued): Significantly changing alternative host polyadenylation events in 2 or more cell types

Gene Name	3UTR	EnsemblID:Proximal Site	Cell Types	PAS+CPE
LARS	S	ENSG00000133706:145493473	EC,HFF,	CPE_PAS
NASP	S	ENSG00000132780:46083916	HFF,NPC,	CPE_PAS
KPNA4	S	ENSG00000186432:160217955	EC,HFF,	CPE_PAS
CYTH2	S	ENSG00000105443:48982645	EC,HFF,NPC,	no-co-occur
AKAP12	S	ENSG00000131016:151677908	HFF,NPC,	CPE_PAS
HNRNPH3	S	ENSG00000096746:70102025	HFF,NPC,	CPE_PAS
SNTB2	S	ENSG00000168807:69339185	EC,HFF,	CPE_PAS
NAP1L1	S	ENSG00000187109:76441377	EC,HFF,	no-co-occur
USP31	S	ENSG00000103404:23075438	EC,HFF,	CPE_PAS
IMPAD1	S	ENSG00000104331:57875662	EC,HFF,	CPE_PAS
NIPA1	S	ENSG00000170113:23047605	EC,HFF,	CPE_PAS
SSR1	S	ENSG00000124783:7289862	EC,HFF,NPC,	no-co-occur
PRRC2B	S	ENSG00000130723:134373516	EC,HFF,	no-co-occur
GNAQ	S	ENSG00000156052:80332438	EC,HFF,	no-co-occur
DPYSL3	S	ENSG00000113657:146772936	EC,HFF,	no-co-occur
NUS1	S	ENSG00000153989:118029748	EC,HFF,	CPE_PAS
ARPP19	S	ENSG00000128989:52843383	EC,NPC,	no-co-occur
DHX40	S	ENSG00000108406:57684920	EC,HFF,	no-co-occur
HSBP1	S	ENSG00000230989:83845270	HFF,NPC,	CPE_PAS
TM9SF3	S	ENSG00000077147:98280744	EC,HFF,NPC,	no-co-occur
CAPZB	S	ENSG00000077549:19665670	EC,NPC,	CPE_PAS
TMEM192	S	ENSG00000170088:165999571	EC,HFF,	no-co-occur
UBE2L3	S	ENSG00000185651:21976161	HFF,NPC,	CPE_PAS
HELZ	S	ENSG00000198265:65073346	EC,NPC,	no-co-occur
GPCPD1	S	ENSG00000125772:5527107	EC,HFF,	CPE_PAS
VAPB	S	ENSG00000124164:57020603	EC,HFF,NPC,	CPE_PAS
TMED7-TICAM2	S	ENSG00000251201:114916039	EC,HFF,	CPE_PAS
MEF2D	S	ENSG00000116604:156436009	EC,NPC,	no-co-occur
ARHGAP32	S	ENSG00000134909:128838733	EC,NPC,	no-co-occur
TBC1D9B	S	ENSG00000197226:179289842	EC,HFF,	no-co-occur
RP3-468K18.5	S	ENSG00000257065:144147001	EC,HFF,	CPE_ncPAS
BACE2	S	ENSG00000182240:42647746	EC,HFF,NPC,	no-co-occur
FAM129A	S	ENSG00000135842:184763251	EC,HFF,	CPE_PAS
DDI2	S	ENSG00000197312:15988524	EC,NPC,	CPE_PAS
CNIH4	S	ENSG00000143771:224563683	HFF,NPC,	no-co-occur
COMMD2	S	ENSG00000114744:149458534	EC,NPC,	CPE_PAS
TMEM50A	S	ENSG00000183726:25687458	EC,HFF,	CPE_PAS
MXD1	S	ENSG00000059728:70167772	EC,HFF,	CPE_ncPAS
ADNP	S	ENSG00000101126:49506882	EC,NPC,	CPE_PAS
PGK1	S	ENSG00000102144:77381769	EC,HFF,	CPE_PAS
PTP4A1	S	ENSG00000112245:64291360	EC,NPC,	CPE_PAS

Table 3.6 (continued): Significantly changing alternative host polyadenylation events in 2 or more cell types

Gene Name	3UTR	EnsemblID:Proximal Site	Cell Types	PAS+CPE
PSMB2	S	ENSG00000126067:36068739	HFF,NPC,	no-co-occur
CENPO	S	ENSG00000138092:25042784	EC,NPC,	no-co-occur
CLU	S	ENSG00000120885:27455513	EC,HFF,	CPE_PAS
RHNO1	S	ENSG00000171792:2998173	EC,HFF,NPC,	no-co-occur
EIF4EBP2	S	ENSG00000148730:72183580	EC,HFF,	CPE_PAS
TBC1D20	S	ENSG00000125875:416928	EC,HFF,	CPE_PAS
CRLS1	S	ENSG00000088766:6018037	EC,NPC,	CPE_PAS
ITM2B	S	ENSG00000136156:48835520	EC,HFF,	CPE_PAS
APBA2	S	ENSG00000034053:29409912	EC,HFF,	CPE_PAS
KDELR2	S	ENSG00000136240:6502502	EC,HFF,	CPE_PAS
ETNK1	S	ENSG00000139163:22838205	EC,HFF,	CPE_ncPAS
ZNF585B	S	ENSG00000245680:37675726	EC,HFF,	no-co-occur
YWHAZ	S	ENSG00000164924:101931928	EC,HFF,	CPE_PAS
DSTN	S	ENSG00000125868:17588638	EC,HFF,	no-co-occur
NXPE3	S	ENSG00000144815:101541777	EC,HFF,	CPE_ncPAS
SQSTM1	S	ENSG00000161011:179264252	EC,NPC,	no-co-occur
WAC	S	ENSG00000095787:28909458	EC,HFF,	no-co-occur
SPARC	S	ENSG00000113140:151042005	EC,HFF,	no-co-occur
RPS6KA3	S	ENSG00000177189:20172843	EC,HFF,NPC,	no-co-occur
USP3	S	ENSG00000140455:63883644	EC,NPC,	no-co-occur
DPP8	S	ENSG00000074603:65737269	EC,HFF,	no-co-occur
GSPT1	S	ENSG00000103342:11966538	HFF,NPC,	CPE_PAS
DNM1L	S	ENSG00000087470:32896543	EC,HFF,	no-co-occur
PLXNA2	S	ENSG00000076356:208200207	EC,HFF,	no-co-occur
CCDC93	S	ENSG00000125633:118677572	EC,HFF,	CPE_PAS
G3BP1	S	ENSG00000145907:151183898	EC,HFF,	no-co-occur
RSU1	S	ENSG00000148484:16634912	EC,HFF,NPC,	CPE_PAS
AKAP2	S	ENSG00000241978:112932484	EC,NPC,	CPE_PAS
PLEKHG2	S	ENSG00000090924:39916298	EC,HFF,	no-co-occur
UBXN7	S	ENSG00000163960:196081956	EC,HFF,NPC,	CPE_PAS
PPIA	S	ENSG00000196262:44841231	HFF,NPC,	CPE_PAS
SCARB2	S	ENSG00000138760:77082606	EC,NPC,	CPE_PAS
CCAR1	S	ENSG00000060339:70551307	HFF,NPC,	CPE_PAS
SMCHD1	S	ENSG00000101596:2803605	EC,HFF,	no-co-occur
OAZ2	S	ENSG00000180304:64980671	EC,HFF,	no-co-occur
SIK2	S	ENSG00000170145:111596774	EC,HFF,	no-co-occur
DNAJC5	S	ENSG00000101152:62566639	EC,HFF,	no-co-occur
AK2	S	ENSG00000004455:33475521	EC,HFF,	CPE_PAS
NAA15	S	ENSG00000164134:140310474	EC,HFF,	no-co-occur
ATG2B	S	ENSG00000066739:96751083	HFF,NPC,	CPE_PAS
EPG5	S	ENSG00000152223:43430470	EC,HFF,	CPE_PAS
RBM15B	S	ENSG00000179837:51431826	HFF,NPC,	CPE_ncPAS
MAP3K7	S	ENSG00000135341:91225413	EC,HFF,NPC,	CPE_PAS

Table 3.6 (continued): Significantly changing alternative host polyadenylation events in 2 or more cell types

Gene Name	3UTR	EnsemblID:Proximal Site	Cell Types	PAS+CPE
NUFIP2	S	ENSG00000108256:27587703	EC,HFF,	CPE_PAS
SMAD4	S	ENSG00000141646:48606146	EC,HFF,	CPE_PAS
CTDSP2	S	ENSG00000175215:58216589	EC,NPC,	no-co-occur
BBIP1	S	ENSG00000214413:112659849	EC,HFF,	CPE_PAS
RCAN3	S	ENSG00000117602:24863504	EC,HFF,	no-co-occur
RP11-613M10.9	S	ENSG00000255872:37588703	EC,HFF,	no-co-occur
ARG2	S	ENSG00000081181:68117901	EC,HFF,	no-co-occur
MTMR9	S	ENSG00000104643:11184323	EC,HFF,	CPE_ncPAS
TCF3	S	ENSG00000071564:1610970	EC,HFF,	no-co-occur
ATMIN	S	ENSG00000166454:81078779	EC,HFF,	no-co-occur
SMAD3	S	ENSG00000166949:67483849	EC,HFF,NPC,	no-co-occur
RER1	S	ENSG00000157916:2334655	EC,HFF,	no-co-occur
TM7SF3	S	ENSG00000064115:27126581	HFF,NPC,	CPE_PAS
SERINC3	S	ENSG00000132824:43127851	EC,HFF,	CPE_PAS
OSBPL2	S	ENSG00000130703:60869940	EC,HFF,	no-co-occur
HMGB3	S	ENSG00000029993:150157238	EC,NPC,	CPE_PAS
SNX1	S	ENSG00000028528:64430205	EC,HFF,	no-co-occur
AK4	S	ENSG00000162433:65693173	EC,HFF,	no-co-occur
YY1	S	ENSG00000100811:100744805	EC,HFF,	CPE_PAS
TNPO1	S	ENSG00000083312:72206156	EC,HFF,	no-co-occur
ECE1	S	ENSG00000117298:21546028	EC,NPC,	no-co-occur
PDCD6IP	S	ENSG00000170248:33908497	HFF,NPC,	no-co-occur
RMND5A	S	ENSG00000153561:87001325	EC,HFF,NPC,	CPE_PAS
FKBP14	S	ENSG00000106080:30053516	EC,HFF,	no-co-occur
SLC12A2	S	ENSG00000064651:127522714	EC,HFF,	CPE_PAS
CPD	S	ENSG00000108582:28794357	EC,HFF,	no-co-occur
CABLES2	S	ENSG00000149679:60965140	EC,HFF,NPC,	no-co-occur
SMIM14	S	ENSG00000163683:39553131	EC,HFF,	CPE_PAS
LMBR1	S	ENSG00000105983:156475633	EC,HFF,NPC,	no-co-occur
SFT2D2	S	ENSG00000213064:168212087	EC,HFF,	CPE_ncPAS
HGSNAT	S	ENSG00000165102:43056966	EC,HFF,	no-co-occur
UQCR11	S	ENSG00000127540:1598047	HFF,NPC,	CPE_PAS
BZW1	S	ENSG00000082153:201687311	EC,HFF,	no-co-occur
HSPA4L	S	ENSG00000164070:128755163	EC,HFF,	no-co-occur
CCDC149	S	ENSG00000181982:24809813	HFF,NPC,	CPE_PAS
ARIH1	S	ENSG00000166233:72876314	EC,HFF,	no-co-occur
DCAF10	S	ENSG00000122741:37862005	EC,HFF,NPC,	CPE_PAS
FAM199X	S	ENSG00000123575:103436125	EC,NPC,	no-co-occur
RLIM	S	ENSG00000131263:73810503	EC,HFF,	no-co-occur
HEXIM1	S	ENSG00000186834:43228016	EC,HFF,NPC,	CPE_PAS
SOD2	S	ENSG00000112096:160103311	EC,HFF,	no-co-occur
SRSF7	S	ENSG00000115875:38972035	HFF,NPC,	CPE_PAS

Table 3.6 (continued): Significantly changing alternative host polyadenylation events in 2 or more cell types

Gene Name	3UTR	EnsemblID:Proximal Site	Cell Types	PAS+CPE
UBE2G1	S	ENSG00000132388:4174730	EC,HFF,	CPE_ncPAS
MMP2	S	ENSG00000087245:55540149	EC,NPC,	no-co-occur
NAMPT	S	ENSG00000105835:105891003	EC,HFF,NPC,	CPE_PAS
SPAG9	S	ENSG00000008294:49042988	EC,HFF,	CPE_PAS
THBS1	S	ENSG00000137801:39887823	EC,HFF,	no-co-occur
.MARCH6	S	ENSG00000145495:10434061	EC,HFF,	no-co-occur
NADK	S	ENSG00000008130:1683958	EC,NPC,	no-co-occur
GOLT1B	S	ENSG00000111711:21669135	EC,HFF,NPC,	CPE_PAS
TMEM194A	S	ENSG00000166881:57451515	EC,HFF,	no-co-occur
LRCH3	S	ENSG00000186001:197612701	EC,HFF,	CPE_PAS
ZDHHC2	S	ENSG00000104219:17077962	EC,HFF,NPC,	no-co-occur
NRP2	S	ENSG00000118257:206660279	EC,HFF,	CPE_PAS
COMT	S	ENSG00000093010:19956464	EC,HFF,	no-co-occur
ARMCX3	S	ENSG00000102401:100881463	EC,HFF,	CPE_PAS
SAR1B	S	ENSG00000152700:133942120	EC,HFF,	CPE_PAS
FMR1	S	ENSG00000102081:147032155	EC,HFF,	CPE_PAS
KLF7	S	ENSG00000118263:207943821	EC,HFF,	CPE_PAS
FEM1B	S	ENSG00000169018:68586057	EC,HFF,NPC,	no-co-occur
RPL15	S	ENSG00000174748:23961042	EC,HFF,NPC,	no-co-occur
TMEM65	S	ENSG00000164983:125325336	EC,HFF,	CPE_ncPAS
GFPT1	S	ENSG00000198380:69552408	EC,HFF,	CPE_PAS
C9orf114	S	ENSG00000198917:131584350	EC,HFF,NPC,	no-co-occur
SFXN1	S	ENSG00000164466:174955525	EC,HFF,	CPE_ncPAS
SERTAD4	S	ENSG00000082497:210417415	EC,HFF,	CPE_PAS
SLC35E1	S	ENSG00000127526:16661668	EC,HFF,	CPE_PAS
PDS5A	S	ENSG00000121892:39825937	EC,HFF,	CPE_PAS
VPS41	S	ENSG00000006715:38765153	EC,HFF,NPC,	no-co-occur
CIRBP	S	ENSG00000099622:1273169	EC,HFF,NPC,	no-co-occur
CNOT6L	S	ENSG00000138767:78636865	EC,HFF,	CPE_PAS
GALNT10	S	ENSG00000164574:153798591	EC,NPC,	no-co-occur
GLIPR1	S	ENSG00000139278:75892789	EC,HFF,	no-co-occur
ILF3	S	ENSG00000129351:10800523	EC,HFF,NPC,	CPE_PAS
UBE2V1	S	ENSG00000244687:48699099	EC,NPC,	no-co-occur
METTL2B	S	ENSG00000165055:128142585	EC,HFF,	CPE_ncPAS
NETO2	S	ENSG00000171208:47116740	EC,HFF,NPC,	CPE_PAS
PRKCA	S	ENSG00000154229:64801433	EC,HFF,	CPE_PAS
ZBTB10	S	ENSG00000205189:81433738	EC,HFF,NPC,	no-co-occur
CSNK1G3	S	ENSG00000151292:122951135	EC,NPC,	no-co-occur
ARL5B	S	ENSG00000165997:18966612	EC,HFF,	no-co-occur
TXNDC15	S	ENSG00000113621:134235521	EC,HFF,	CPE_PAS
CTSB	S	ENSG00000164733:11701828	EC,HFF,NPC,	no-co-occur
RHEB	S	ENSG00000106615:151163795	EC,HFF,	CPE_PAS
C16orf72	S	ENSG00000182831:9212933	EC,NPC,	no-co-occur

Table 3.6 (continued): Significantly changing alternative host polyadenylation events in 2 or more cell types

Gene Name	3UTR	EnsemblID:Proximal Site	Cell Types	PAS+CPE
PREPL	S	ENSG00000138078:44547965	EC,HFF,	CPE_PAS
QSOX1	S	ENSG00000116260:180167167	EC,HFF,NPC,	CPE_PAS
TMF1	S	ENSG00000144747:69070276	EC,HFF,	no-co-occur
SRP14	S	ENSG00000140319:40328242	EC,HFF,	CPE_PAS
IRF2BP2	S	ENSG00000168264:234742494	HFF,NPC,	no-co-occur
CDK9	S	ENSG00000136807:130552363	HFF,NPC,	CPE_PAS
ZBTB44	S	ENSG00000196323:130103132	EC,NPC,	CPE_PAS
TSPAN14	S	ENSG00000108219:82279384	EC,HFF,NPC,	CPE_ncPAS
RAB8B	S	ENSG00000166128:63558244	EC,HFF,	no-co-occur
IL13RA1	S	ENSG00000131724:117926324	EC,HFF,	CPE_PAS
GTPBP4	S	ENSG00000107937:1063631	HFF,NPC,	CPE_PAS
CDIP1	S	ENSG00000089486:4561326	EC,HFF,	no-co-occur
UBE2H	S	ENSG00000186591:129472652	EC,HFF,	no-co-occur
TBC1D1	S	ENSG00000065882:38139240	EC,HFF,NPC,	no-co-occur
LSM14A	S	ENSG00000257103:34718998	EC,HFF,NPC,	CPE_PAS
METTL9	S	ENSG00000197006:21667256	HFF,NPC,	CPE_PAS
XPOT	S	ENSG00000184575:64843162	EC,HFF,	CPE_ncPAS
PPP1CB	S	ENSG00000213639:29023902	EC,HFF,	CPE_PAS
HMGCR	S	ENSG00000113161:74657240	EC,HFF,	CPE_PAS
SMAD5	S	ENSG00000113658:135514577	EC,HFF,	no-co-occur
CAV1	S	ENSG00000105974:116199494	EC,HFF,	CPE_PAS
RRAS2	S	ENSG00000133818:14300299	EC,HFF,	CPE_PAS
PSMA2	S	ENSG00000256646:42957011	EC,HFF,NPC,	CPE_PAS
TUBB	S	ENSG00000196230:30692353	EC,HFF,NPC,	no-co-occur
AVL9	S	ENSG00000105778:32623780	HFF,NPC,	CPE_PAS
RTCA	S	ENSG00000137996:100757326	HFF,NPC,	CPE_PAS
FAR1	S	ENSG00000197601:13751445	EC,HFF,	CPE_PAS
HNRNPDL	S	ENSG00000152795:83345781	HFF,NPC,	CPE_PAS
LARP4	S	ENSG00000161813:50871758	EC,HFF,	no-co-occur
CBX6	S	ENSG00000183741:39260263	EC,HFF,	CPE_ncPAS
SBNO1	S	ENSG00000139697:123779078	EC,HFF,	CPE_PAS
LEPROT	S	ENSG00000213625:65898624	EC,HFF,	no-co-occur
NR3C1	S	ENSG00000113580:142659131	EC,HFF,	no-co-occur
AZIN1	S	ENSG00000155096:103840231	EC,HFF,NPC,	CPE_PAS
METRN	S	ENSG00000103260:767478	HFF,NPC,	no-co-occur
WDR26	S	ENSG00000162923:224576838	EC,HFF,NPC,	CPE_PAS
RAB8A	S	ENSG00000167461:16244443	EC,HFF,	CPE_PAS
GPR107	S	ENSG00000148358:132900289	EC,HFF,	no-co-occur
PPP1R2	S	ENSG00000184203:195243046	EC,HFF,NPC,	CPE_PAS
STMN1	S	ENSG00000117632:26227058	EC,HFF,	no-co-occur
RAB18	S	ENSG00000099246:27827441	HFF,NPC,	CPE_PAS
ARHGAP26	S	ENSG00000145819:142603623	EC,HFF,	CPE_ncPAS
DCUN1D4	S	ENSG00000109184:52780971	EC,NPC,	CPE_PAS

Table 3.6 (continued): Significantly changing alternative host polyadenylation events in 2 or more cell types

Gene Name	3UTR	EnsemblID:Proximal Site	Cell Types	PAS+CPE
CD44	S	ENSG00000026508:35251017	EC,HFF,	no-co-occur
HN1	S	ENSG00000189159:73132083	EC,NPC,	no-co-occur
SOCS5	S	ENSG00000171150:46988752	EC,NPC,	CPE_PAS
ANAPC1	S	ENSG00000153107:112526641	EC,HFF,NPC,	CPE_PAS
C11orf57	L	ENSG00000150776:111955095	HFF,NPC,	CPE_PAS
LMO4	L	ENSG00000143013:87810715	HFF,NPC,	no-co-occur
CYFIP1	L	ENSG00000068793:23003602	EC,HFF,NPC,	CPE_PAS
ARID1A	L	ENSG00000117713:27107573	HFF,NPC,	CPE_PAS
ZNF706	L	ENSG00000120963:102211250	EC,HFF,	CPE_PAS
AKT2	L	ENSG00000105221:40738241	HFF,NPC,	no-co-occur
WIZ	L	ENSG00000011451:15532909	EC,HFF,NPC,	CPE_PAS
UBE3C	L	ENSG00000009335:157061810	HFF,NPC,	CPE_PAS
RPS23	L	ENSG00000186468:81571884	EC,HFF,	CPE_PAS
SNX3	L	ENSG00000112335:108532717	HFF,NPC,	no-co-occur
RPL23	L	ENSG00000125691:37006340	EC,HFF,	CPE_PAS
EIF1	L	ENSG00000173812:39847250	EC,HFF,	no-co-occur
CARHSP1	L	ENSG00000153048:8948770	EC,HFF,	no-co-occur
TCP1	L	ENSG00000120438:160199986	EC,HFF,	CPE_PAS
RPS28	L	ENSG00000233927:8387278	HFF,NPC,	CPE_PAS
TANK	L	ENSG00000136560:162092228	HFF,NPC,	CPE_ncPAS
IGFBP7	L	ENSG00000163453:57897245	EC,HFF,	CPE_PAS
CEBPG	L	ENSG00000153879:33870896	EC,HFF,	CPE_PAS
NSFL1C	L	ENSG00000088833:1424152	EC,HFF,	no-co-occur
LSM4	L	ENSG00000130520:18417728	HFF,NPC,	CPE_PAS
RPL31	L	ENSG00000071082:101622570	EC,HFF,	CPE_ncPAS
DBNL	L	ENSG00000136279:44101257	EC,HFF,	no-co-occur
ZNF84	L	ENSG00000198040:133636186	EC,HFF,	CPE_PAS
COPS7B	L	ENSG00000144524:232673432	EC,HFF,	CPE_PAS
RAN	L	ENSG00000132341:131360824	EC,HFF,	CPE_PAS
VTI1B	L	ENSG00000100568:68117868	EC,HFF,	no-co-occur
PSME4	L	ENSG00000068878:54092009	EC,HFF,	CPE_PAS
NARS	L	ENSG00000134440:55268832	HFF,NPC,	no-co-occur
EIF2S1	L	ENSG00000134001:67850575	EC,HFF,	CPE_PAS
UQCRB	L	ENSG00000156467:97243162	EC,HFF,	no-co-occur
STK16	L	ENSG00000115661:220113626	HFF,NPC,	no-co-occur
TSPAN6	L	ENSG00000000003:99884690	EC,HFF,	CPE_ncPAS
EIF3F	L	ENSG00000175390:8017716	EC,HFF,	no-co-occur
TM2D2	L	ENSG00000169490:38848323	HFF,NPC,	no-co-occur
LAMTOR3	L	ENSG00000109270:100802370	EC,HFF,	CPE_PAS
TMEM39A	L	ENSG00000176142:119150579	EC,HFF,	CPE_PAS
NOMO2	L	ENSG00000185164:18511419	HFF,NPC,	no-co-occur
TMEM183A	L	ENSG00000163444:202992667	EC,HFF,	CPE_PAS
SKP1	L	ENSG00000113558:133493164	EC,HFF,	no-co-occur

Table 3.6 (continued): Significantly changing alternative host polyadenylation events in 2 or more cell types

Gene Name	3UTR	EnsemblID:Proximal Site	Cell Types	PAS+CPE
FXR1	L	ENSG00000114416:180694333	HFF,NPC,	CPE_PAS
USP8	L	ENSG00000138592:50791971	EC,HFF,	CPE_PAS
TFAM	L	ENSG00000108064:60155423	EC,HFF,	CPE_PAS
ST13	L	ENSG00000100380:41222135	EC,HFF,	no-co-occur
APEX2	L	ENSG00000169188:55034201	EC,HFF,	no-co-occur
HLA-E	L	ENSG00000204592:30461070	EC,HFF,	no-co-occur
EPB41L1	L	ENSG00000088367:34818221	EC,HFF,	no-co-occur
LITAF	L	ENSG00000189067:11642676	HFF,NPC,	CPE_ncPAS
USP7	L	ENSG00000187555:8987349	EC,HFF,	CPE_PAS
RASA4	L	ENSG00000105808:102222769	EC,HFF,	no-co-occur
FAHD2A	L	ENSG00000115042:96078878	EC,HFF,	no-co-occur
TMED10	L	ENSG00000170348:75601333	HFF,NPC,	no-co-occur
C3orf17	L	ENSG00000163608:112723767	EC,HFF,	CPE_PAS
DIEXF	L	ENSG00000117597:210025094	EC,HFF,	CPE_ncPAS
LSS	L	ENSG00000160285:47609045	EC,HFF,NPC,	CPE_PAS
ZFP36L1	L	ENSG00000185650:69255359	HFF,NPC,	no-co-occur
DCAF11	L	ENSG00000100897:24593238	EC,NPC,	no-co-occur
PRNP	L	ENSG00000171867:4681585	HFF,NPC,	no-co-occur
CUL2	L	ENSG00000108094:35298796	EC,HFF,	CPE_PAS
COL4A2	L	ENSG00000134871:111165093	HFF,NPC,	no-co-occur
ZDHHC9	L	ENSG00000188706:128938914	EC,HFF,	no-co-occur
SERF2	L	ENSG00000140264:44086239	EC,HFF,	CPE_PAS
SARM1	L	ENSG00000004139:26725179	EC,HFF,	CPE_ncPAS
HNRNPU	L	ENSG00000153187:245016962	EC,NPC,	CPE_PAS
FCHSD1	L	ENSG00000197948:141020549	EC,HFF,	no-co-occur
EXT1	L	ENSG00000182197:118811605	EC,HFF,	CPE_PAS
SAP30BP	L	ENSG00000161526:73702836	EC,HFF,NPC,	no-co-occur
EIF4E	L	ENSG00000151247:99801212	EC,HFF,	CPE_PAS
APP	L	ENSG00000142192:27253118	HFF,NPC,	no-co-occur
SUPT4H1	L	ENSG00000213246:56423272	EC,HFF,NPC,	CPE_PAS
DNAJA1	L	ENSG00000086061:33039293	EC,HFF,	no-co-occur
PTP4A2	L	ENSG00000184007:32374298	EC,HFF,	no-co-occur
NOMO3	L	ENSG00000103226:16388430	HFF,NPC,	no-co-occur
HPS1	L	ENSG00000107521:100176779	EC,HFF,	no-co-occur
ABT1	L	ENSG00000146109:26599284	EC,HFF,	CPE_PAS
AK1	L	ENSG00000106992:130630120	EC,HFF,	no-co-occur
NKIRAS1	L	ENSG00000197885:23933756	EC,HFF,	no-co-occur
CCT6A	L	ENSG00000146731:56131094	EC,NPC,	CPE_PAS
HNRNPL	L	ENSG00000104824:39327281	HFF,NPC,	no-co-occur

Table 3.7: Significant host gene expression changes, HFFs

Significant changes listed are based on Z-score analysis. The fold-change estimates listed were calculated as (RPKM infected + 0.1) / (RPKM mock + 0.1).

Z-score	Gene Name	RPKM Mock	RPKM Infection	FoldChg Estimate	Ensembl ID
First 100 most downregulated (by Z-score), 48 hpi					
-5.97	KIAA1199	574.53	4.12	-136.23	ENSG00000103888
-5.93	CCRL1	45.61	1.42	-30.09	ENSG00000129048
-5.55	GRIA1	11.25	0.34	-25.56	ENSG00000155511
-5.45	MGP	19.08	0.61	-26.90	ENSG00000111341
-5.18	MRVI1	20.17	0.99	-18.58	ENSG00000072952
-5.16	MXRA5	97.56	2.47	-38.00	ENSG00000101825
-5.00	CR591448	16.68	0.78	-19.16	ENSG00000104435
-4.93	PTGIS	64.01	2.86	-21.65	ENSG00000124212
-4.82	IGFBP5	1186.81	22.16	-53.32	ENSG00000115461
-4.70	FMO1	1.86	0.00	-19.01	ENSG00000010932
-4.61	AK055621	9.45	0.39	-19.41	ENSG00000246430
-4.60	LMOD1	57.36	3.66	-15.30	ENSG00000163431
-4.52	UBL4B	11.37	0.82	-12.42	ENSG00000186150
-4.51	SEL-OB	50.65	3.44	-14.34	ENSG00000165124
-4.49	MASP1	111.41	3.98	-27.35	ENSG00000127241
-4.43	F2RL2	23.92	2.50	-9.24	ENSG00000164220
-4.41	SEMA5A	33.80	3.18	-10.34	ENSG00000112902
-4.33	POSTN	44.90	3.76	-11.66	ENSG00000133110
-4.26	OLFML1	14.12	1.18	-11.08	ENSG00000183801
-4.24	OSR2	83.95	4.89	-16.84	ENSG00000164920
-4.19	LAMA2	32.85	3.59	-8.92	ENSG00000196569
-4.18	COMP	33.48	3.60	-9.08	ENSG00000105664
-4.15	KIAA1562	37.09	3.88	-9.35	ENSG00000189184
-4.10	DKFZp686O088	68.02	4.79	-13.92	ENSG00000168398
-4.09	SSC5D	20.04	2.11	-9.11	ENSG00000179954
-4.09	SFRP2	19.85	2.10	-9.07	ENSG00000145423
-4.08	GPR133	11.15	1.09	-9.42	ENSG00000111452
-4.04	KRTAP1-5	47.17	4.79	-9.67	ENSG00000221852
-4.03	ALPK2	23.36	3.13	-7.27	ENSG00000198796
-4.03	COL8A1	25.00	3.30	-7.39	ENSG00000144810
-3.99	PTX3	69.21	4.96	-13.69	ENSG00000163661
-3.98	LRRN4CL	11.07	1.17	-8.83	ENSG00000177363
-3.98	WNT2	26.15	3.74	-6.84	ENSG00000105989
-3.97	FOSB	19.08	2.14	-8.57	ENSG00000125740
-3.97	ALDH1A1	612.87	23.23	-26.27	ENSG00000165092
-3.90	ELN	103.73	6.06	-16.87	ENSG00000049540
-3.87	CTHRC1	17.46	1.74	-9.53	ENSG00000164932
-3.86	GALNT13	7.91	0.50	-13.46	ENSG00000131386
-3.81	FAM20A	6.17	0.30	-15.68	ENSG00000108950
-3.76	DKFZp434I0916	82.04	6.79	-11.93	ENSG00000137573
-3.75	AK123947	20.32	2.88	-6.86	ENSG00000255248

Table 3.7 (continued): Significant host gene expression changes, HFFs

Z-score	Gene Name	RPKM Mock	RPKM Infection	FoldChg Estimate	Ensembl ID
-3.75	COL1A1	2321.69	105.49	-21.99	ENSG00000108821
-3.71	TNFRSF11B	208.24	9.84	-20.97	ENSG00000164761
-3.70	GAS1	20.92	3.10	-6.57	ENSG00000180447
-3.69	MAB21L1	18.34	2.53	-7.00	ENSG00000180660
-3.69	IGF1	1.14	0.00	-11.94	ENSG00000017427
-3.67	TENC1	23.26	3.97	-5.74	ENSG00000111077
-3.65	CHRM2	7.13	0.45	-13.17	ENSG00000181072
-3.65	SLC38A11	4.13	0.12	-19.69	ENSG00000169507
-3.64	ADAMTS1	85.57	7.39	-11.44	ENSG00000154734
-3.63	SEPP1	10.04	1.13	-8.27	ENSG00000250722
-3.60	ITGBL1	28.44	4.73	-5.91	ENSG00000198542
-3.60	TNC	69.75	6.80	-10.12	ENSG00000041982
-3.58	FKBP9L	26.53	4.34	-6.00	ENSG00000176826
-3.56	CEBPD	31.68	5.02	-6.21	ENSG00000221869
-3.56	THBS1	1688.37	90.18	-18.70	ENSG00000137801
-3.55	SLIT3	35.79	5.48	-6.43	ENSG00000184347
-3.54	MEGF6	37.68	5.34	-6.94	ENSG00000162591
-3.53	COL3A1	372.26	20.32	-18.23	ENSG00000168542
-3.53	C8orf4	9.45	1.02	-8.56	ENSG00000176907
-3.49	ITGA11	14.38	1.95	-7.07	ENSG00000137809
-3.47	CRISPLD2	27.72	4.94	-5.53	ENSG00000103196
-3.47	RECK	44.04	6.34	-6.86	ENSG00000122707
-3.46	C8orf84	24.03	4.47	-5.28	ENSG00000164764
-3.44	PRELP	83.60	8.64	-9.58	ENSG00000188783
-3.42	AK023507	52.23	7.05	-7.32	ENSG00000144476
-3.41	HNOEL-iso	64.66	7.52	-8.50	ENSG00000116774
-3.37	KRT34	17.11	2.48	-6.68	ENSG00000131737
-3.34	THSD4	22.06	4.52	-4.79	ENSG00000187720
-3.32	C4orf18	9.79	1.39	-6.62	ENSG00000164125
-3.30	IGFBP3	1102.18	72.89	-15.10	ENSG00000146674
-3.27	CR620010	3.39	0.11	-16.77	ENSG00000248223
-3.25	PAPPA	13.46	2.28	-5.69	ENSG00000182752
-3.24	WFDC1	42.69	7.18	-5.88	ENSG00000103175
-3.24	PDGFRA	31.00	6.05	-5.05	ENSG00000134853
-3.24	LTBP2	46.33	7.46	-6.14	ENSG00000119681
-3.24	EGR1	298.73	20.87	-14.25	ENSG00000120738
-3.23	CLDN11	32.46	6.20	-5.17	ENSG00000013297
-3.21	FOS	171.88	12.25	-13.93	ENSG00000170345
-3.21	MYLK	96.71	9.97	-9.61	ENSG00000065534
-3.21	ADAM12	12.92	2.21	-5.63	ENSG00000148848
-3.21	DKFZp434L0823	10.50	1.88	-5.35	ENSG00000149380
-3.21	GDF5	7.42	0.79	-8.41	ENSG00000125965
-3.19	BC131755	16.83	2.76	-5.92	ENSG00000108846
-3.19	CXCL12	29.00	5.96	-4.80	ENSG00000107562
-3.17	SPOCD1	19.23	3.78	-4.98	ENSG00000134668
-3.16	FAP	41.76	7.47	-5.53	ENSG00000078098
-3.16	RUNX1	17.56	3.47	-4.95	ENSG00000159216

Table 3.7 (continued): Significant host gene expression changes, HFFs

Z-score	Gene Name	RPKM Mock	RPKM Infection	FoldChg Estimate	Ensembl ID
-3.15	DIO2	6.32	0.60	-9.19	ENSG00000211448
-3.12	PIK3IP1	8.10	1.03	-7.25	ENSG00000100100
-3.10	HSD17B2	6.32	0.63	-8.80	ENSG00000086696
-3.10	PLEKHA6	7.64	0.91	-7.66	ENSG00000143850
-3.09	TMEM119	58.79	9.14	-6.38	ENSG00000183160
-3.08	THBS2	248.02	19.79	-12.48	ENSG00000186340
-3.06	FGF7	7.03	0.82	-7.75	ENSG00000140285
-3.06	ALDH3B1	24.56	5.69	-4.26	ENSG00000006534
-3.06	COL1A2	2146.02	174.19	-12.31	ENSG00000164692
-3.04	ADAMTS5	5.57	0.55	-8.79	ENSG00000154736
-3.04	FMO4	8.39	1.22	-6.44	ENSG00000076258
-3.03	TNXB	6.47	0.70	-8.22	ENSG00000248290
First 100 most upregulated (by Z-score), 48 hpi					
9.56	CCL5	0.01	16.73	153.01	ENSG00000161570
7.17	ECEL1	0.00	6.09	60.11	ENSG00000171551
7.06	RBP7	0.26	31.98	88.14	ENSG00000162444
6.64	C15orf48	0.23	24.42	74.99	ENSG00000166920
6.46	KIF1A	0.00	4.12	41.73	ENSG00000130294
6.34	EPCAM	0.08	12.26	69.83	ENSG00000119888
6.17	MAPK8IP2	0.21	21.92	70.36	ENSG00000008735
6.07	DKFZp761J191	0.03	8.65	67.82	ENSG00000126878
5.99	SEPT3	0.58	40.11	59.31	ENSG00000100167
5.96	HERC5	0.02	7.47	64.17	ENSG00000138646
5.82	CDS1	0.00	4.12	41.38	ENSG00000163624
5.71	TESC	0.00	4.70	46.13	ENSG00000088992
5.69	IFI27	0.64	36.11	48.99	ENSG00000165949
5.68	CTSU	0.18	15.75	57.21	ENSG00000136943
5.56	HOMER2	0.12	11.44	53.41	ENSG00000103942
5.55	OASL	0.02	6.40	55.05	ENSG00000135114
5.53	KIAA1319	0.01	5.45	51.88	ENSG00000143375
5.29	SPINT1	0.23	17.08	52.54	ENSG00000166145
5.23	CADM4	0.77	30.11	34.84	ENSG00000105767
5.22	BEX2	0.21	14.63	47.35	ENSG00000133134
5.00	LLGL2	0.11	10.12	48.22	ENSG00000073350
4.90	ELOVL7	0.01	4.20	40.55	ENSG00000164181
4.66	GCH1	0.32	13.64	32.95	ENSG00000131979
4.64	TMEM108	0.01	4.25	39.94	ENSG00000144868
4.62	DKFZp761H171	0.15	9.78	39.53	ENSG00000025039
4.59	TSPAN33	0.13	9.46	41.39	ENSG00000158457
4.53	TNFSF9	0.33	14.23	33.24	ENSG00000125657
4.53	PHOSPHO1	0.01	4.36	39.78	ENSG00000173868
4.41	CCDC64	0.02	5.47	44.90	ENSG00000135127
4.32	HES6	0.55	18.90	29.36	ENSG00000144485
4.30	NPTX2	0.20	10.19	34.40	ENSG00000106236
4.30	PCDH1	0.11	7.88	37.83	ENSG00000156453
4.29	KIAA1543	0.00	3.08	30.61	ENSG00000076826
4.28	CNTFR	0.02	4.62	40.38	ENSG00000122756

Table 3.7 (continued): Significant host gene expression changes, HFFs

Z-score	Gene Name	RPKM Mock	RPKM Infection	FoldChg Estimate	Ensembl ID
4.27	ANO5	0.00	2.15	22.25	ENSG00000171714
4.22	RNF157	0.65	19.77	26.35	ENSG00000141576
4.21	BC032716	1.15	24.41	19.67	ENSG00000232445
4.19	CHRDL2	0.02	4.30	37.94	ENSG00000054938
4.15	C1QL1	0.40	13.32	26.67	ENSG00000131094
4.13	FAM169A	0.06	6.07	38.81	ENSG00000198780
4.09	PLS1	0.05	5.84	38.54	ENSG00000120756
4.09	TAC1	0.02	4.46	37.68	ENSG00000006128
4.05	LMX1B	0.00	2.25	23.03	ENSG00000136944
4.04	C10orf35	0.65	19.20	25.66	ENSG00000171224
4.02	MCOLN3	0.49	13.49	23.06	ENSG00000055732
3.99	STAC2	0.01	3.37	31.54	ENSG00000141750
3.97	TMEM151A	0.06	5.86	36.32	ENSG00000179292
3.95	F11R	0.19	8.87	30.52	ENSG00000158769
3.92	HOOK1	0.00	1.81	18.70	ENSG00000134709
3.92	PLCG2	0.09	6.65	35.17	ENSG00000197943
3.88	CCK	0.03	4.41	35.79	ENSG00000187094
3.84	C9orf61	0.00	2.01	20.68	ENSG00000135063
3.83	C17orf104	0.12	6.98	32.93	ENSG00000180336
3.83	HCG26	3.53	73.44	20.29	ENSG00000206337
3.80	DKFZp686I21167	1.80	28.54	15.09	ENSG00000188211
3.80	YBX2	0.01	3.43	31.26	ENSG00000006047
3.78	SLC30A3	0.01	2.64	26.05	ENSG00000115194
3.77	DUSP2	0.48	11.94	20.76	ENSG00000158050
3.76	TEX15	0.00	1.83	18.89	ENSG00000133863
3.75	CXADR	0.16	7.61	29.30	ENSG00000154639
3.67	BST2	0.50	11.67	19.52	ENSG00000130303
3.67	TCF15	0.03	4.26	33.82	ENSG00000125878
3.61	CXorf57	0.00	2.26	22.65	ENSG00000147231
3.60	LSR	0.48	11.24	19.61	ENSG00000105699
3.56	FAM184A	0.01	3.15	29.02	ENSG00000111879
3.56	TNFAIP8L3	0.27	8.99	24.42	ENSG00000183578
3.55	SEMA6B	0.01	3.08	28.39	ENSG00000167680
3.55	ADD2	0.00	1.57	16.40	ENSG00000075340
3.53	EEF1A2	5.42	94.78	17.18	ENSG00000101210
3.51	TMEM38A	0.67	13.74	17.91	ENSG00000072954
3.47	ARG2	0.44	10.53	19.72	ENSG00000081181
3.46	AGPAT9	0.14	6.42	27.72	ENSG00000138678
3.44	CHST1	0.00	1.75	17.91	ENSG00000175264
3.44	PPP1R1A	0.01	2.24	22.29	ENSG00000135447
3.42	ABHD3	1.43	21.83	14.33	ENSG00000158201
3.42	NGFR	0.00	1.44	15.10	ENSG00000064300
3.41	RPS6KA1	0.85	16.39	17.41	ENSG00000117676
3.40	NUP210	0.28	8.44	22.78	ENSG00000132182
3.38	UNQ603	0.00	1.39	14.56	ENSG00000078596
3.36	OSBP2	0.14	6.15	26.26	ENSG00000184792
3.35	IL28RA	0.01	2.55	24.32	ENSG00000185436

Table 3.7 (continued): Significant host gene expression changes, HFFs

Z-score	Gene Name	RPKM Mock	RPKM Infection	FoldChg Estimate	Ensembl ID
3.35	c-erbB-3	0.27	8.29	22.50	ENSG00000065361
3.33	MFSD2A	0.61	11.52	16.36	ENSG00000168389
3.33	AGT	0.65	12.27	16.41	ENSG00000135744
3.31	RIMKLA	0.02	3.34	28.00	ENSG00000177181
3.27	PPP2R2C	0.01	2.72	25.16	ENSG00000074211
3.26	SVIP	2.06	24.87	11.56	ENSG00000198168
3.24	GYLTL1B	0.01	2.50	23.65	ENSG00000165905
3.21	CRTAC1	0.00	1.38	14.40	ENSG00000095713
3.19	KCNH3	0.01	1.71	17.25	ENSG00000135519
3.18	P2RY2	0.03	3.50	27.68	ENSG00000175591
3.17	SLC38A3	0.03	3.67	28.34	ENSG00000188338
3.15	SLC35F2	0.72	11.85	14.64	ENSG00000110660
3.15	NRTN	0.36	8.51	18.77	ENSG00000171119
3.08	TTC39A	0.06	4.16	27.30	ENSG00000085831
3.08	CXCR4	0.01	2.13	20.66	ENSG00000121966
3.08	CCDC88C	0.01	1.77	17.64	ENSG00000015133
3.08	JAG2	0.23	7.01	21.36	ENSG00000184916
3.08	STX3	2.99	31.62	10.26	ENSG00000166900
3.07	SLC10A4	0.14	5.70	24.39	ENSG00000145248
First 100 most downregulated (by Z-score), 96 hpi					
-6.50	RP3-509I19.1	35.97	0.19	-126.58	ENSG00000225415
-6.20	CCRL1	45.60	0.27	-124.53	ENSG00000129048
-6.15	KIAA1199	574.53	0.80	-635.65	ENSG00000103888
-6.11	MRVI1	20.17	0.15	-82.38	ENSG00000072952
-5.82	MXRA5	97.56	0.37	-207.34	ENSG00000101825
-5.54	IGFBP5	1186.81	3.19	-360.32	ENSG00000115461
-5.45	AC090953.1	13.17	0.13	-57.70	ENSG00000255701
-5.26	KRTAP1-5	47.17	0.65	-63.19	ENSG00000221852
-5.20	STMN2	16.68	0.23	-50.40	ENSG00000104435
-5.12	ELN	103.73	0.75	-122.74	ENSG00000049540
-5.09	OLFML1	14.12	0.20	-47.73	ENSG00000183801
-4.98	POSTN	44.90	0.78	-50.96	ENSG00000133110
-4.91	LOC100507632	9.45	0.10	-48.98	ENSG00000246430
-4.81	MXRA5P1	5.09	0.02	-42.92	ENSG00000235649
-4.80	MASP1	111.41	1.00	-101.28	ENSG00000127241
-4.80	GRIA1	11.25	0.18	-41.27	ENSG00000155511
-4.80	PRELP	83.60	0.98	-77.21	ENSG00000188783
-4.69	LAMA2	32.85	0.95	-31.44	ENSG00000196569
-4.68	F2RL2	23.92	0.72	-29.22	ENSG00000164220
-4.55	PTGIS	64.02	1.23	-48.13	ENSG00000124212
-4.54	COL8A1	25.00	1.03	-22.17	ENSG00000144810
-4.54	COMP	33.48	1.13	-27.37	ENSG00000105664
-4.54	SSC5D	20.04	0.57	-30.20	ENSG00000179954
-4.44	COL1A1	2321.69	20.43	-113.12	ENSG00000108821
-4.43	SVEP1	50.65	1.43	-33.10	ENSG00000165124
-4.33	TMEM119	58.79	1.45	-38.07	ENSG00000183160
-4.33	ACTG2	15.53	0.53	-24.97	ENSG00000163017

Table 3.7 (continued): Significant host gene expression changes, HFFs

Z-score	Gene Name	RPKM Mock	RPKM Infection	FoldChg Estimate	Ensembl ID
-4.27	KRT34	17.11	0.59	-25.09	ENSG00000131737
-4.23	MIR675	16.83	0.54	-26.32	ENSG00000211502
-4.21	SLC38A11	4.13	0.03	-33.34	ENSG00000169507
-4.13	BDKRB2	68.02	1.75	-36.92	ENSG00000168398
-4.10	COL3A1	372.27	4.71	-77.46	ENSG00000168542
-4.10	WFDC1	42.69	1.60	-25.20	ENSG00000103175
-4.10	SEMA5A	33.80	1.52	-20.88	ENSG00000112902
-4.08	SULF1	82.04	1.97	-39.78	ENSG00000137573
-4.07	COL1A2	2146.03	28.26	-75.67	ENSG00000164692
-4.05	TENC1	23.26	1.16	-18.62	ENSG00000111077
-4.05	GALNTL2	7.91	0.11	-37.59	ENSG00000131386
-4.03	THBS1	1688.37	23.05	-72.93	ENSG00000137801
-4.02	UNQ2565	1.29	0.01	-13.27	ENSG00000128262
-4.02	BX640700	1.27	0.01	-13.04	ENSG00000253864
-4.01	IGFBP3	1102.17	15.41	-71.06	ENSG00000146674
-3.98	OLFML3	64.66	1.96	-31.41	ENSG00000116774
-3.97	TNFRSF11B	208.24	3.04	-66.29	ENSG00000164761
-3.96	C8orf4	9.45	0.27	-25.89	ENSG00000176907
-3.95	PCDH18	37.09	1.69	-20.77	ENSG00000189184
-3.88	PSG8	16.40	0.74	-19.55	ENSG00000124467
-3.88	IL6	18.58	0.92	-18.35	ENSG00000136244
-3.84	PDCD1LG2	9.67	0.32	-23.48	ENSG00000197646
-3.84	LMOD1	57.36	2.18	-25.17	ENSG00000163431
-3.83	DIO2	6.32	0.09	-33.78	ENSG00000211448
-3.79	ALDH1A1	612.87	10.83	-56.09	ENSG00000165092
-3.78	MYLK	96.71	2.74	-34.12	ENSG00000065534
-3.78	FAM198B	9.79	0.35	-21.92	ENSG00000164125
-3.76	MAP1A	25.98	1.77	-13.93	ENSG00000166963
-3.76	ADAMTS1	85.57	2.92	-28.35	ENSG00000154734
-3.73	MEGF6	37.68	2.07	-17.44	ENSG00000162591
-3.73	MAB21L1	18.34	1.04	-16.23	ENSG00000180660
-3.72	IGF1	1.14	0.01	-11.61	ENSG00000017427
-3.71	PDGFRA	31.00	1.99	-14.90	ENSG00000134853
-3.68	ITGBL1	28.44	1.96	-13.83	ENSG00000198542
-3.67	CSDC2	9.23	0.36	-20.27	ENSG00000172346
-3.66	MIR143HG	2.90	0.03	-22.37	ENSG00000249669
-3.65	AC009469.1	25.08	2.03	-11.84	ENSG00000256309
-3.65	ISLR	17.41	1.05	-15.23	ENSG00000129009
-3.65	TNFSF18	1.14	0.01	-11.58	ENSG00000120337
-3.65	GALNT5	8.77	0.29	-22.85	ENSG00000136542
-3.64	SEPP1	10.04	0.44	-18.91	ENSG00000250722
-3.64	MUC1	12.40	0.66	-16.39	ENSG00000185499
-3.63	FOS	171.87	3.61	-46.42	ENSG00000170345
-3.63	PSG3	13.73	0.76	-16.14	ENSG00000221826
-3.61	FGF7	7.03	0.16	-27.73	ENSG00000140285
-3.60	FOSB	19.08	1.23	-14.46	ENSG00000125740
-3.56	PSG7	9.18	0.37	-19.65	ENSG00000221878

Table 3.7 (continued): Significant host gene expression changes, HFFs

Z-score	Gene Name	RPKM Mock	RPKM Infection	FoldChg Estimate	Ensembl ID
-3.56	SFRP2	19.85	1.46	-12.78	ENSG00000145423
-3.55	ALDH3B1	24.56	2.28	-10.38	ENSG00000006534
-3.55	PDGFRB	100.85	3.50	-28.01	ENSG00000113721
-3.55	CCL2	124.74	3.38	-35.91	ENSG00000108691
-3.55	EGR1	298.73	6.86	-42.92	ENSG00000120738
-3.55	THBS2	248.02	5.70	-42.78	ENSG00000186340
-3.54	OSR2	83.95	3.50	-23.34	ENSG00000164920
-3.51	HSPB6	78.13	3.37	-22.58	ENSG00000004776
-3.51	CTHRC1	17.46	1.20	-13.53	ENSG00000164932
-3.51	BX647608	20.32	1.54	-12.47	ENSG00000255248
-3.50	PTX3	69.21	3.17	-21.20	ENSG00000163661
-3.49	DCN	570.14	13.96	-40.56	ENSG00000011465
-3.48	RP11-124N14.3	32.25	2.52	-12.35	ENSG00000234961
-3.44	UBL4B	11.37	0.73	-13.78	ENSG00000186150
-3.44	GPR133	11.15	0.64	-15.26	ENSG00000111452
-3.44	SPOCD1	19.23	1.45	-12.45	ENSG00000134668
-3.43	WNT2	26.15	2.32	-10.85	ENSG00000105989
-3.43	FAP	41.76	2.78	-14.53	ENSG00000078098
-3.43	GBP1	13.41	0.87	-13.95	ENSG00000117228
-3.43	TRIM22	13.31	0.88	-13.68	ENSG00000132274
-3.42	SRPX2	14.62	1.07	-12.54	ENSG00000102359
-3.42	RNA5SP392	38.53	2.65	-14.05	ENSG00000252714
-3.41	GDF5	7.42	0.21	-24.16	ENSG00000125965
-3.41	GAS1	20.92	1.69	-11.77	ENSG00000180447
-3.39	TNS1	17.06	1.27	-12.57	ENSG00000079308
-3.38	PSG5	56.44	3.38	-16.23	ENSG00000204941
First 100 most upregulated (by Z-score), 96 hpi					
8.33	TAC1	0.02	34.25	283.92	ENSG00000006128
7.55	CCK	0.03	23.60	188.10	ENSG00000187094
7.42	CRTAC1	0.00	9.62	94.35	ENSG00000095713
6.75	EPCAM	0.08	33.50	189.82	ENSG00000119888
6.44	CDS1	0.00	7.10	70.57	ENSG00000163624
6.38	HOOK1	0.00	6.88	68.47	ENSG00000134709
6.37	CCNA1	0.02	12.12	101.87	ENSG00000133101
6.25	AIF1L	0.03	14.22	111.04	ENSG00000126878
6.02	KIF1A	0.00	4.69	47.46	ENSG00000130294
5.83	BEX2	0.21	68.59	220.88	ENSG00000133134
5.68	HERC5	0.02	9.61	82.26	ENSG00000138646
5.53	RBP7	0.26	58.05	159.75	ENSG00000162444
5.50	LAMP3	0.04	12.28	85.95	ENSG00000078081
5.36	TSPAN7	0.01	8.16	75.80	ENSG00000156298
5.36	MYO5B	0.01	7.51	70.42	ENSG00000167306
5.33	CLGN	0.01	7.62	71.44	ENSG00000153132
5.32	BC044934	0.12	18.51	86.13	ENSG00000103942
5.32	ECEL1	0.00	5.41	53.46	ENSG00000171551
5.32	RRAGD	0.15	20.80	83.60	ENSG00000025039
5.27	C15orf48	0.23	29.18	89.55	ENSG00000166920

Table 3.7 (continued): Significant host gene expression changes, HFFs

Z-score	Gene Name	RPKM Mock	RPKM Infection	FoldChg Estimate	Ensembl ID
5.14	MIR147B	0.37	47.75	102.89	ENSG00000211519
5.09	YBX2	0.01	8.04	72.04	ENSG00000006047
5.05	TESC	0.00	4.97	48.78	ENSG00000088992
5.04	DERL3	0.41	57.14	112.67	ENSG00000099958
4.99	FAM189A2	0.00	4.17	41.81	ENSG00000135063
4.99	NPTX2	0.20	22.12	74.31	ENSG00000106236
4.97	CGN	0.01	6.19	58.80	ENSG00000143375
4.93	CTSL2	0.18	19.50	70.76	ENSG00000136943
4.89	PDLIM3	0.02	8.45	72.43	ENSG00000154553
4.89	ELOVL7	0.01	5.84	56.02	ENSG00000164181
4.89	MGAT4A	0.00	4.05	40.72	ENSG00000071073
4.86	GCH1	0.32	28.62	68.87	ENSG00000131979
4.85	ZNF711	0.06	11.28	69.79	ENSG00000147180
4.85	ELAVL3	0.00	3.09	31.27	ENSG00000196361
4.84	ITM2A	0.00	3.60	36.27	ENSG00000078596
4.69	NELL2	0.01	4.59	44.70	ENSG00000184613
4.68	SPINT1	0.23	20.12	61.85	ENSG00000166145
4.68	SLC30A3	0.01	4.65	45.25	ENSG00000115194
4.61	HBA1	0.02	6.94	61.17	ENSG00000206172
4.61	LMX1B	0.00	2.84	28.83	ENSG00000136944
4.61	ARG2	0.44	26.22	48.83	ENSG00000081181
4.59	RUNDC3B	0.01	5.44	51.31	ENSG00000105784
4.58	LPL	0.01	5.81	54.18	ENSG00000175445
4.58	CHST1	0.00	3.70	36.93	ENSG00000175264
4.57	TMEM151A	0.06	10.09	62.10	ENSG00000179292
4.56	FAM71E1	0.25	20.16	57.55	ENSG00000142530
4.55	PPP1R1A	0.01	4.53	44.05	ENSG00000135447
4.54	MFAP3L	0.11	12.23	59.00	ENSG00000198948
4.52	LSR	0.48	26.22	45.54	ENSG00000105699
4.48	SEPT3	0.58	43.69	64.59	ENSG00000100167
4.43	SH3GL3	0.00	3.31	32.74	ENSG00000140600
4.41	TMEM108	0.01	5.06	47.36	ENSG00000144868
4.41	ANO5	0.00	2.17	22.46	ENSG00000171714
4.41	CXorf57	0.00	3.62	35.78	ENSG00000147231
4.40	SHE	0.02	6.79	58.38	ENSG00000169291
4.39	CXCR4	0.01	4.82	45.51	ENSG00000121966
4.38	HAP1	0.12	11.80	53.86	ENSG00000173805
4.36	HBA2	0.01	6.23	55.48	ENSG00000188536
4.29	ADD2	0.00	2.49	25.37	ENSG00000075340
4.25	AGPAT9	0.14	11.71	50.26	ENSG00000138678
4.24	PHOSPHO1	0.01	5.12	46.60	ENSG00000173868
4.22	CNTFR	0.02	6.21	53.96	ENSG00000122756
4.22	LOC401022	0.01	5.16	46.96	ENSG00000224189
4.19	BMP6	0.04	8.44	60.54	ENSG00000153162
4.15	RNF157	0.65	36.85	49.01	ENSG00000141576
4.13	CA8	0.02	6.30	53.35	ENSG00000178538
4.09	SLC35F1	0.00	2.09	21.45	ENSG00000196376

Table 3.7 (continued): Significant host gene expression changes, HFFs

Z-score	Gene Name	RPKM Mock	RPKM Infection	FoldChg Estimate	Ensembl ID
4.07	GATM	0.01	4.48	41.65	ENSG00000171766
4.07	COBL	0.00	2.09	21.51	ENSG00000106078
4.06	F11R	0.19	12.85	44.05	ENSG00000158769
4.03	BMP7	0.00	2.08	21.37	ENSG00000101144
3.99	CAMSAP3	0.00	2.61	26.06	ENSG00000076826
3.98	NGFR	0.00	2.03	20.83	ENSG00000064300
3.97	NUP62CL	0.02	4.90	43.51	ENSG00000198088
3.96	NEFM	0.30	15.48	39.35	ENSG00000104722
3.96	CADM4	0.77	35.38	40.92	ENSG00000105767
3.94	ELAVL2	0.00	1.82	18.86	ENSG00000107105
3.92	FAM184A	0.01	4.45	40.58	ENSG00000111879
3.92	CTH	0.81	36.90	40.88	ENSG00000116761
3.92	RHEBL1	0.50	21.20	35.33	ENSG00000167550
3.91	ERBB3	0.27	14.46	39.03	ENSG00000065361
3.91	PKP2	0.03	6.32	50.98	ENSG00000057294
3.89	MAL2	0.01	4.38	39.96	ENSG00000147676
3.86	MEGF10	0.00	1.10	11.92	ENSG00000145794
3.86	SCUBE1	0.01	2.87	27.97	ENSG00000159307
3.86	ABCB1	0.01	3.32	31.69	ENSG00000085563
3.84	UBE2QL1	0.00	1.82	18.79	ENSG00000215218
3.84	SLC27A2	0.04	6.79	51.05	ENSG00000140284
3.84	ERO1LB	0.77	29.75	34.47	ENSG00000086619
3.83	PPP2R2C	0.01	4.27	39.03	ENSG00000074211
3.82	UNC5D	0.00	1.08	11.69	ENSG00000156687
3.81	ADRA2C	0.33	14.92	34.60	ENSG00000184160
3.80	GYLTL1B	0.01	3.65	34.07	ENSG00000165905
3.80	SLC10A4	0.14	10.02	42.50	ENSG00000145248
3.78	AK092313	0.00	1.13	12.14	ENSG00000233863
3.77	MT1F	0.78	24.79	28.38	ENSG00000198417
3.73	CCDC64	0.02	5.34	43.85	ENSG00000135127
3.73	IKZF1	0.00	1.15	12.40	ENSG00000185811
3.72	PCDH1	0.11	9.33	44.67	ENSG00000156453
3.72	FBLL1	0.39	16.46	34.14	ENSG00000188573

Table 3.8: Significant host gene expression changes, ECs

Significant changes listed are based on Z-score analysis. The fold-change estimates listed were calculated as (RPKM infected + 0.1) / (RPKM mock + 0.1).

Z-score	Gene Name	RPKM Mock	RPKM Infection	FoldChg Estimate	Ensembl ID
First 100 most downregulated (by Z-score), 48 hpi					
-8.61	FABP4	46.95	0.05	-309.55	ENSG00000170323
-7.05	SULT1E1	31.80	0.17	-118.57	ENSG00000109193
-6.98	GJA5	54.38	0.22	-172.96	ENSG00000143140
-6.87	BC040870	36.47	0.22	-113.23	ENSG00000228495
-6.30	TNFSF18	21.12	0.15	-84.22	ENSG00000120337
-6.28	DKFZp434I0916	43.51	0.37	-92.40	ENSG00000137573
-6.19	IL33	16.00	0.12	-71.86	ENSG00000137033
-6.14	LDLRAD2	28.70	0.32	-68.41	ENSG00000187942
-6.06	C15orf54	19.72	0.14	-82.24	ENSG00000175746
-6.05	TNFSF10	29.59	0.38	-62.12	ENSG00000121858
-6.00	VCAM1	23.00	0.26	-64.52	ENSG00000162692
-5.99	AK091865	11.70	0.07	-68.22	ENSG00000235770
-5.45	CD34	21.02	0.31	-51.14	ENSG00000174059
-5.40	GJA4	60.18	1.00	-54.90	ENSG00000187513
-5.31	EMCN	43.17	0.79	-48.46	ENSG00000164035
-5.31	GIMAP4	13.01	0.21	-42.84	ENSG00000133574
-5.28	TSPAN8	19.17	0.29	-49.79	ENSG00000127324
-5.24	SELP	9.75	0.10	-48.99	ENSG00000174175
-5.24	RSPO3	18.19	0.24	-54.43	ENSG00000146374
-5.14	C8orf4	67.72	1.56	-40.90	ENSG00000176907
-5.10	CXCL2	46.17	0.91	-45.63	ENSG00000081041
-5.08	GIMAP7	56.77	1.19	-44.25	ENSG00000179144
-5.00	APLN	120.47	1.84	-62.06	ENSG00000171388
-4.89	CPA3	6.79	0.02	-57.42	ENSG00000163751
-4.85	AK125871	114.01	2.08	-52.42	ENSG00000248894
-4.84	TGFB2	16.71	0.39	-34.31	ENSG00000092969
-4.84	ART4	12.58	0.28	-33.63	ENSG00000111339
-4.80	MGP	256.07	2.77	-89.38	ENSG00000111341
-4.74	TNFSF4	20.34	0.57	-30.51	ENSG00000117586
-4.73	POSTN	46.16	1.24	-34.50	ENSG00000133110
-4.73	GIMAP1	8.94	0.09	-48.06	ENSG00000213203
-4.71	SDPR	58.79	1.68	-33.03	ENSG00000168497
-4.68	CCL2	438.60	5.32	-80.91	ENSG00000108691
-4.65	RGS4	21.91	0.68	-28.08	ENSG00000117152
-4.62	CLDN11	43.87	1.33	-30.83	ENSG00000013297
-4.62	CXCL6	20.30	0.63	-27.87	ENSG00000124875
-4.60	LYPD1	142.62	2.27	-60.29	ENSG00000150551
-4.57	C4orf18	49.45	1.59	-29.25	ENSG00000164125
-4.56	BMX	53.52	1.52	-33.20	ENSG00000102010
-4.54	TFPI2	75.09	2.58	-28.11	ENSG00000105825
-4.51	EDN1	445.94	6.28	-69.96	ENSG00000078401

Table 3.8 (continued): Significant host gene expression changes, ECs

Z-score	Gene Name	RPKM Mock	RPKM Infection	FoldChg Estimate	Ensembl ID
-4.50	LOC100216001	17.94	0.47	-31.70	ENSG00000231298
-4.48	AK023507	69.08	2.68	-24.90	ENSG00000144476
-4.46	SERPINB2	26.17	1.23	-19.78	ENSG00000197632
-4.43	TM4SF18	25.43	1.34	-17.78	ENSG00000163762
-4.43	LOC643733	13.55	0.41	-26.97	ENSG00000235505
-4.42	IGFBP5	5.91	0.03	-47.31	ENSG00000115461
-4.40	PALMD	89.51	3.11	-27.96	ENSG00000099260
-4.39	CCRL1	10.50	0.25	-30.20	ENSG00000129048
-4.39	IL6	110.73	2.95	-36.31	ENSG00000136244
-4.37	HHIP	337.92	5.41	-61.40	ENSG00000164161
-4.37	F2RL2	14.24	0.45	-26.30	ENSG00000164220
-4.37	ALDH1A1	190.45	3.06	-60.40	ENSG00000165092
-4.26	KCNAB1	11.07	0.36	-24.06	ENSG00000169282
-4.21	AK123961	9.94	0.30	-25.23	ENSG00000232759
-4.18	CTGF	1804.41	34.53	-52.10	ENSG00000118523
-4.11	RGS5	12.38	0.56	-18.86	ENSG00000143248
-4.09	MMRN1	124.98	3.60	-33.84	ENSG00000138722
-4.08	MMRN2	69.26	3.57	-18.88	ENSG00000173269
-4.07	MMP10	41.11	2.11	-18.67	ENSG00000166670
-4.04	COL8A1	37.53	1.94	-18.48	ENSG00000144810
-4.00	ABCA8	11.19	0.44	-20.78	ENSG00000141338
-4.00	SERPINE1	3478.75	78.44	-44.30	ENSG00000106366
-3.99	PTX3	382.26	8.70	-43.45	ENSG00000163661
-3.97	DKK1	179.53	4.14	-42.41	ENSG00000107984
-3.97	ANKRD1	977.89	22.65	-42.98	ENSG00000148677
-3.96	C10orf10	141.32	3.95	-34.88	ENSG00000165507
-3.93	OR56B1	18.66	0.96	-17.75	ENSG00000132274
-3.93	LXN	22.62	1.49	-14.25	ENSG00000079257
-3.91	CARD16	52.37	2.55	-19.82	ENSG00000204397
-3.91	AKR1C3	61.16	3.86	-15.46	ENSG00000196139
-3.90	I-rap	28.47	1.95	-13.94	ENSG00000164308
-3.88	TAGLN	533.07	13.35	-39.66	ENSG00000149591
-3.88	TMEM140	24.63	1.91	-12.28	ENSG00000146859
-3.86	SPOCD1	27.53	1.95	-13.47	ENSG00000134668
-3.86	PDE7B	7.98	0.19	-28.27	ENSG00000171408
-3.74	ESM1	215.68	6.14	-34.58	ENSG00000164283
-3.73	THBS1	1601.97	46.19	-34.61	ENSG00000137801
-3.73	APOL3	10.13	0.47	-17.85	ENSG00000128284
-3.70	LRRC17	7.89	0.22	-25.06	ENSG00000128606
-3.67	FAM124B	22.91	1.81	-12.04	ENSG00000124019
-3.67	LAMA4	73.35	4.97	-14.48	ENSG00000112769
-3.66	SORBS2	7.24	0.15	-28.88	ENSG00000154556
-3.65	ANGPT2	4.24	0.04	-31.88	ENSG00000091879
-3.64	EGR1	115.44	5.48	-20.70	ENSG00000120738
-3.62	HTR2B	2.97	0.02	-26.49	ENSG00000135914
-3.58	MYC	40.66	3.12	-12.64	ENSG00000136997
-3.57	PDL2	10.30	0.61	-14.67	ENSG00000197646

Table 3.8 (continued): Significant host gene expression changes, ECs

Z-score	Gene Name	RPKM Mock	RPKM Infection	FoldChg Estimate	Ensembl ID
-3.56	LTBP2	75.13	5.59	-13.21	ENSG00000119681
-3.55	SULT1B1	19.29	1.41	-12.82	ENSG00000173597
-3.54	GLIPR1	42.06	3.09	-13.21	ENSG00000139278
-3.52	CLEC1B	1.28	0.01	-13.10	ENSG00000165682
-3.52	FOS	91.84	6.10	-14.82	ENSG00000170345
-3.51	INHBA	21.41	1.88	-10.87	ENSG00000122641
-3.50	EFEMP1	1779.54	62.80	-28.30	ENSG00000115380
-3.49	GIMAP8	15.15	1.19	-11.87	ENSG00000171115
-3.49	FLJ00236	77.17	5.85	-12.98	ENSG00000154133
-3.49	CAV1	325.14	11.66	-27.65	ENSG00000105974
-3.46	IL1RL1	52.69	3.84	-13.40	ENSG00000115602
-3.46	MMP1	2261.20	83.23	-27.14	ENSG00000196611
First 100 most upregulated (by Z-score), 48 hpi					
8.92	KIF1A	0.00	10.45	104.44	ENSG00000130294
7.36	ST14	0.00	5.50	56.01	ENSG00000149418
7.32	RBP7	0.07	34.36	208.84	ENSG00000162444
7.04	ADM2	0.05	23.53	159.69	ENSG00000128165
6.89	RNF157	0.02	16.90	139.34	ENSG00000141576
6.60	SEPT3	0.11	33.76	162.78	ENSG00000100167
6.47	CRLF1	0.05	21.56	145.36	ENSG00000006016
6.15	CKB	0.67	96.66	125.99	ENSG00000166165
6.07	HCG26	0.56	73.32	110.74	ENSG00000206337
6.00	EPCAM	0.05	17.52	115.15	ENSG00000119888
5.95	PRRX2	0.04	15.09	108.50	ENSG00000167157
5.65	PHGDH	0.23	33.48	100.85	ENSG00000092621
5.54	RUNX3	0.00	7.45	73.34	ENSG00000020633
5.38	NPTX2	0.04	12.26	87.06	ENSG00000106236
5.13	FAM105A	0.03	9.61	77.02	ENSG00000145569
5.13	SPINT1	0.35	33.31	74.24	ENSG00000166145
5.12	VLDLR	0.15	21.26	84.09	ENSG00000147852
5.07	EEF1A2	1.65	112.90	64.68	ENSG00000101210
5.06	NRTN	0.01	8.88	78.78	ENSG00000171119
5.04	DUSP2	0.17	21.49	81.15	ENSG00000158050
4.96	ULBP1	0.38	28.23	59.40	ENSG00000111981
4.96	C10orf35	0.37	29.24	62.70	ENSG00000171224
4.85	IRX2	0.01	7.72	71.73	ENSG00000170561
4.74	COL6A2	0.14	16.60	68.72	ENSG00000142173
4.71	C1QL4	0.02	8.84	73.85	ENSG00000186897
4.67	DSP	0.01	7.44	68.58	ENSG00000096696
4.66	C1QL1	0.12	13.26	61.87	ENSG00000131094
4.62	DKFZp686I21167	0.57	34.13	51.09	ENSG00000188211
4.60	HOXB9	0.00	6.25	61.07	ENSG00000170689
4.59	MAPK8IP2	0.11	13.74	65.28	ENSG00000008735
4.54	DIRAS1	0.09	12.15	63.82	ENSG00000176490
4.52	INA	0.19	17.26	60.06	ENSG00000148798
4.49	TTC9	0.04	9.29	67.09	ENSG00000133985
4.34	CHAC1	1.14	53.35	42.96	ENSG00000128965

Table 3.8 (continued): Significant host gene expression changes, ECs

Z-score	Gene Name	RPKM Mock	RPKM Infection	FoldChg Estimate	Ensembl ID
4.29	HOMER2	0.12	11.80	55.35	ENSG00000103942
4.27	LITAF	1.03	42.23	37.49	ENSG00000189067
4.17	BC111371	0.01	5.64	54.17	ENSG00000242516
4.14	MAP3K9	0.00	3.76	37.83	ENSG00000006432
4.14	BEX2	0.90	38.75	38.77	ENSG00000133134
4.05	GYLTL1B	0.01	5.31	51.48	ENSG00000165905
4.04	SLC10A4	0.02	7.45	61.89	ENSG00000145248
4.02	CALB1	0.01	6.61	60.42	ENSG00000104327
4.01	ACOT4	0.24	15.07	44.62	ENSG00000177465
3.99	PHOSPHO1	0.16	12.31	47.37	ENSG00000173868
3.91	BIK	0.21	13.67	44.57	ENSG00000100290
3.91	NDRG2	0.02	7.22	60.49	ENSG00000165795
3.90	IL28RA	0.00	4.47	43.93	ENSG00000185436
3.88	TBX2	0.01	4.95	48.06	ENSG00000121068
3.88	AK024736	0.01	6.17	57.02	ENSG00000237686
3.85	L1CAM	0.12	10.63	48.54	ENSG00000198910
3.84	TMEM151A	0.06	8.87	57.14	ENSG00000179292
3.81	GAL	1.21	36.50	28.02	ENSG00000069482
3.81	IGFBPL1	0.01	5.22	49.69	ENSG00000137142
3.80	SIX2	0.02	6.68	58.48	ENSG00000170577
3.80	MA3	0.01	5.13	48.87	ENSG00000183837
3.79	EFHD1	0.05	8.49	56.89	ENSG00000115468
3.77	KIAA0845	0.33	15.74	36.66	ENSG00000100285
3.76	KIAA1319	0.24	12.86	37.66	ENSG00000143375
3.76	ITM2C	2.09	56.57	25.90	ENSG00000135916
3.71	c-erbB-3	0.01	5.28	49.36	ENSG00000065361
3.69	BEGAIN	0.01	4.60	44.30	ENSG00000183092
3.69	C15orf48	0.39	18.81	38.43	ENSG00000166920
3.63	ADRA2C	0.02	6.86	56.57	ENSG00000184160
3.61	KCNK1	0.17	10.71	40.62	ENSG00000135750
3.59	TSPAN33	0.39	15.68	32.27	ENSG00000158457
3.58	SPINT2	1.12	29.08	23.91	ENSG00000167642
3.58	KCNA7	0.00	3.67	36.26	ENSG00000104848
3.56	UNQ470	0.09	8.98	48.80	ENSG00000168913
3.54	LRRC61	0.01	4.33	41.75	ENSG00000127399
3.53	DLX3	0.00	3.38	33.83	ENSG00000064195
3.49	DKFZp761H171	0.28	12.37	32.81	ENSG00000025039
3.49	ESYT3	0.01	3.81	37.28	ENSG00000158220
3.48	ASNS	2.71	57.51	20.52	ENSG00000070669
3.47	PSAT1	3.36	61.45	17.79	ENSG00000135069
3.45	KIAA0444	0.01	4.97	46.10	ENSG00000116254
3.44	CHMP4C	0.05	7.69	51.26	ENSG00000164695
3.43	CHRDL2	0.01	4.80	44.95	ENSG00000054938
3.41	HES6	0.67	20.93	27.39	ENSG00000144485
3.40	DKFZp667I0324	0.01	3.60	35.23	ENSG00000108001
3.35	RAB11FIP4	0.01	4.00	38.36	ENSG00000131242
3.34	GPR27	0.00	3.20	32.06	ENSG00000170837

Table 3.8 (continued): Significant host gene expression changes, ECs

Z-score	Gene Name	RPKM Mock	RPKM Infection	FoldChg Estimate	Ensembl ID
3.32	LSR	0.59	19.25	28.21	ENSG00000105699
3.32	RIMKLA	0.02	5.60	47.86	ENSG00000177181
3.31	TLX3	0.01	3.62	35.11	ENSG00000164438
3.31	FAM169A	0.10	8.64	43.90	ENSG00000198780
3.30	CRB3	0.02	5.51	47.17	ENSG00000130545
3.28	LLGL2	0.49	15.38	26.46	ENSG00000073350
3.27	LMX1B	0.00	3.14	31.44	ENSG00000136944
3.27	C17orf104	0.04	6.92	50.47	ENSG00000180336
3.24	cent-b	0.07	7.84	46.40	ENSG00000184060
3.21	CEBPA	0.05	6.95	48.25	ENSG00000245848
3.20	ELOVL7	0.01	3.72	35.72	ENSG00000164181
3.20	GPRC5C	0.05	7.17	48.13	ENSG00000170412
3.20	RBP4	0.00	2.61	26.32	ENSG00000138207
3.19	DSG2	0.02	5.62	46.51	ENSG00000046604
3.18	SLC6A9	0.07	7.45	45.77	ENSG00000196517
3.18	SLC6A17	0.01	3.92	37.22	ENSG00000197106
3.17	EMILIN3	0.00	3.10	30.73	ENSG00000183798
3.17	RAP1GAP	0.01	3.96	37.25	ENSG00000076864
3.16	COBL	0.00	2.48	25.08	ENSG00000106078
First 100 most downregulated (by Z-score), 96 hpi					
-7.99	LOC100507254	36.31	0.02	-311.20	ENSG00000228495
-6.93	IL33	15.93	0.01	-141.81	ENSG00000137033
-6.50	SULT1B1	19.20	0.03	-154.40	ENSG00000173597
-6.47	POSTN	45.96	0.06	-284.30	ENSG00000133110
-6.37	GJA5	54.14	0.10	-273.94	ENSG00000143140
-6.23	MMP1	2251.17	2.03	-1056.44	ENSG00000196611
-6.16	MMRN1	124.43	0.19	-429.40	ENSG00000138722
-6.09	LDLRAD2	28.57	0.10	-142.65	ENSG00000187942
-6.08	TNFSF10	29.46	0.10	-148.55	ENSG00000121858
-5.75	VCAM1	22.90	0.07	-132.92	ENSG00000162692
-5.74	EMCN	42.98	0.13	-184.11	ENSG00000164035
-5.70	SERPINB2	26.65	0.12	-122.68	ENSG00000197632
-5.67	C15orf54	19.63	0.07	-114.72	ENSG00000175746
-5.66	GIMAP4	12.95	0.03	-101.96	ENSG00000133574
-5.55	C8orf4	67.42	0.32	-160.37	ENSG00000176907
-5.53	TAGLN	530.70	1.03	-469.73	ENSG00000149591
-5.48	ERAP2	28.34	0.18	-101.94	ENSG00000164308
-5.45	TNFSF18	21.03	0.10	-108.36	ENSG00000120337
-5.43	IL1RL1	52.45	0.24	-156.41	ENSG00000115602
-5.27	FABP4	46.74	0.22	-146.38	ENSG00000170323
-5.22	CCL15-CCL14	14.45	0.07	-87.10	ENSG00000161574
-5.18	SERPINE1	3463.31	9.90	-346.38	ENSG00000106366
-5.18	LOC100216001	17.86	0.10	-92.08	ENSG00000231298
-5.15	SULF1	43.32	0.26	-122.31	ENSG00000137573
-5.13	CD34	20.93	0.14	-89.11	ENSG00000174059
-5.07	LOC646324	11.65	0.04	-85.77	ENSG00000235770
-4.98	MGP	254.93	0.91	-253.76	ENSG00000111341

Table 3.8 (continued): Significant host gene expression changes, ECs

Z-score	Gene Name	RPKM Mock	RPKM Infection	FoldChg Estimate	Ensembl ID
-4.95	F2RL2	14.17	0.09	-76.32	ENSG00000164220
-4.94	ANKRD1	973.55	3.61	-262.65	ENSG00000148677
-4.88	ART4	12.52	0.06	-78.40	ENSG00000111339
-4.86	N/A	18.74	0.14	-78.48	ENSG00000207727
-4.85	CCRL1	10.45	0.04	-78.17	ENSG00000129048
-4.82	APOL1	17.69	0.14	-74.12	ENSG00000100342
-4.72	TSPAN8	19.08	0.17	-71.05	ENSG00000127324
-4.71	MMRN2	68.96	0.70	-86.21	ENSG00000173269
-4.69	KCNAB1	11.02	0.05	-76.67	ENSG00000169282
-4.61	SPOCD1	27.41	0.36	-59.29	ENSG00000134668
-4.59	MMP10	40.93	0.47	-71.85	ENSG00000166670
-4.58	APLN	122.53	0.98	-113.86	ENSG00000171388
-4.55	CPA3	6.76	0.02	-57.65	ENSG00000163751
-4.54	GIMAP7	56.51	0.63	-77.13	ENSG00000179144
-4.52	CLIC3	1.74	0.00	-18.20	ENSG00000169583
-4.51	CCL2	446.65	2.66	-161.69	ENSG00000108691
-4.51	EDN1	443.96	2.67	-160.43	ENSG00000078401
-4.48	GIMAP1	8.90	0.04	-66.63	ENSG00000213203
-4.47	SELP	9.70	0.04	-69.52	ENSG00000174175
-4.45	ABI3BP	36.29	0.58	-53.43	ENSG00000154175
-4.43	GGT5	11.64	0.08	-64.86	ENSG00000099998
-4.42	RP3-509I19.1	8.23	0.03	-62.63	ENSG00000225415
-4.42	AURKB	11.59	0.08	-64.20	ENSG00000178999
-4.39	AC002480.4	8.05	0.03	-61.72	ENSG00000232949
-4.38	ESM1	214.72	1.48	-135.79	ENSG00000164283
-4.31	RP11-124N14.3	42.58	0.60	-61.14	ENSG00000234961
-4.31	ABCA8	11.37	0.09	-59.73	ENSG00000141338
-4.31	RPL7P18	11.42	0.09	-59.36	ENSG00000243806
-4.28	TNFSF4	20.25	0.32	-48.34	ENSG00000117586
-4.27	AP000695.6	6.35	0.03	-51.58	ENSG00000230479
-4.27	BMX	54.44	0.83	-58.83	ENSG00000102010
-4.26	ANGPT2	4.22	0.01	-40.35	ENSG00000091879
-4.25	HHIP	336.42	2.67	-121.49	ENSG00000164161
-4.24	ALDH1A1	189.61	1.53	-116.74	ENSG00000165092
-4.23	APOL3	10.31	0.07	-59.84	ENSG00000128284
-4.21	IL6	110.24	1.59	-65.48	ENSG00000136244
-4.21	SRPX2	56.19	0.88	-57.26	ENSG00000102359
-4.17	PTX3	380.56	3.30	-112.03	ENSG00000163661
-4.17	EGR1	114.93	1.60	-67.78	ENSG00000120738
-4.14	MIR221	7.60	0.04	-53.47	ENSG00000207870
-4.05	CXCL2	45.97	0.80	-51.41	ENSG00000081041
-4.04	LYPD1	141.99	1.65	-81.15	ENSG00000150551
-4.04	TM4SF18	25.32	0.60	-36.31	ENSG00000163762
-4.02	FAM198B	49.23	0.86	-51.55	ENSG00000164125
-3.99	SDPR	58.53	1.24	-43.72	ENSG00000168497
-3.98	COL8A1	37.36	0.93	-36.30	ENSG00000144810
-3.97	IL8	194.48	2.12	-87.85	ENSG00000169429

Table 3.8 (continued): Significant host gene expression changes, ECs

Z-score	Gene Name	RPKM Mock	RPKM Infection	FoldChg Estimate	Ensembl ID
-3.97	ROBO4	76.83	1.57	-46.06	ENSG00000154133
-3.95	PALMD	89.11	1.86	-45.49	ENSG00000099260
-3.94	GBP1	12.45	0.18	-44.83	ENSG00000117228
-3.91	LOC100506178	9.90	0.10	-49.49	ENSG00000232759
-3.90	TRIM22	18.58	0.40	-37.51	ENSG00000132274
-3.88	LRRC17	7.86	0.07	-47.66	ENSG00000128606
-3.86	GIMAP2	14.09	0.29	-36.47	ENSG00000106560
-3.85	GIMAP6	19.95	0.49	-34.22	ENSG00000133561
-3.84	CARD16	52.14	1.23	-39.16	ENSG00000204397
-3.79	LOC643733	13.49	0.26	-37.33	ENSG00000235505
-3.79	THBS1	1594.86	21.18	-74.95	ENSG00000137801
-3.78	FAM124B	22.81	0.62	-32.04	ENSG00000124019
-3.77	CASP1	7.41	0.07	-44.94	ENSG00000137752
-3.77	C1orf54	26.27	0.82	-28.72	ENSG00000118292
-3.76	IGFBP5	5.88	0.04	-41.54	ENSG00000115461
-3.73	TMEM140	24.52	0.77	-28.30	ENSG00000146859
-3.72	RP11-115J23.1	7.27	0.07	-43.37	ENSG00000249867
-3.70	ANGPTL4	27.01	0.90	-27.06	ENSG00000167772
-3.66	PDE7B	7.95	0.09	-43.04	ENSG00000171408
-3.65	snoU13	10.37	0.16	-40.10	ENSG00000238541
-3.62	RP11-150O12.6	3.32	0.01	-30.77	ENSG00000253414
-3.61	GLIPR1	41.88	1.22	-31.87	ENSG00000139278
-3.60	WWTR1-IT1	8.54	0.11	-40.96	ENSG00000241985
-3.59	KRT7	208.60	3.44	-58.97	ENSG00000135480
-3.58	PDCD1LG2	10.26	0.18	-37.53	ENSG00000197646
-3.58	FERMT3	40.92	1.32	-28.85	ENSG00000149781
First 100 most upregulated (by Z-score), 96 hpi					
8.41	ST14	0.00	9.02	91.20	ENSG00000149418
7.96	CALB1	0.01	53.12	479.44	ENSG00000104327
7.40	DSP	0.01	26.61	242.80	ENSG00000096696
7.01	CRLF1	0.05	69.74	471.92	ENSG00000006016
6.88	RNF157	0.02	28.54	234.73	ENSG00000141576
6.82	IGFBPL1	0.01	17.45	164.01	ENSG00000137142
6.81	RBP4	0.00	13.02	127.33	ENSG00000138207
6.61	EPCAM	0.05	58.28	381.58	ENSG00000119888
6.43	RBP7	0.06	58.27	355.91	ENSG00000162444
6.29	NPTX2	0.04	31.84	226.55	ENSG00000106236
6.12	MAP3K9	0.00	9.24	91.54	ENSG00000006432
6.02	RAB11FIP4	0.01	12.88	121.26	ENSG00000131242
5.93	PPP1R1A	0.00	3.57	36.67	ENSG00000135447
5.82	FAM105A	0.03	19.46	155.22	ENSG00000145569
5.78	NEFM	0.02	17.91	147.63	ENSG00000104722
5.70	ADM2	0.05	24.25	164.51	ENSG00000128165
5.60	SEPT3	0.11	38.75	186.79	ENSG00000100167
5.55	RAMP1	0.04	21.39	149.26	ENSG00000132329
5.39	GYLTL1B	0.01	10.05	96.64	ENSG00000165905
5.29	DSG2	0.02	14.81	121.24	ENSG00000046604

Table 3.8 (continued): Significant host gene expression changes, ECs

Z-score	Gene Name	RPKM Mock	RPKM Infection	FoldChg Estimate	Ensembl ID
5.29	COL6A2	0.14	38.34	158.19	ENSG00000142173
5.23	VLDLR	0.15	37.77	149.09	ENSG00000147852
5.14	SLPI	0.06	20.76	130.40	ENSG00000124107
5.09	CKB	0.67	141.30	184.84	ENSG00000166165
5.01	ADRA2C	0.02	13.43	110.02	ENSG00000184160
4.84	BIK	0.22	39.58	125.58	ENSG00000100290
4.81	PHGDH	0.23	37.75	113.99	ENSG00000092621
4.81	OLFM1	0.03	12.57	100.58	ENSG00000130558
4.76	MFAP3L	0.06	17.08	106.06	ENSG00000198948
4.75	EFHD1	0.05	15.28	101.87	ENSG00000115468
4.74	SLC10A4	0.02	12.03	99.45	ENSG00000145248
4.72	ISM1	0.01	8.10	77.39	ENSG00000101230
4.72	KIF1A	0.00	4.51	45.63	ENSG00000130294
4.70	DERL3	0.14	25.14	105.15	ENSG00000099958
4.70	AK024736	0.01	9.22	84.71	ENSG00000237686
4.69	FNDC1	0.00	4.76	48.12	ENSG00000164694
4.68	FAM71E1	0.16	27.73	105.81	ENSG00000142530
4.66	LITAF	1.03	133.92	119.13	ENSG00000189067
4.65	C1QL1	0.12	22.08	102.68	ENSG00000131094
4.63	CHMP4C	0.05	14.97	99.82	ENSG00000164695
4.62	FZD7	0.06	15.25	98.42	ENSG00000155760
4.60	SLC38A3	0.04	13.98	99.17	ENSG00000188338
4.54	BC044934	0.11	19.64	92.23	ENSG00000103942
4.54	MYO5B	0.03	11.65	93.27	ENSG00000167306
4.50	TTC9	0.04	12.54	90.93	ENSG00000133985
4.50	HOOK1	0.02	10.31	88.19	ENSG00000134709
4.48	SLC27A2	0.00	5.71	56.36	ENSG00000140284
4.44	BC031953	0.04	12.36	90.26	ENSG00000254815
4.44	ENHO	0.09	15.96	86.83	ENSG00000168913
4.43	SPINT1	0.35	42.22	94.25	ENSG00000166145
4.41	LRRC61	0.01	6.44	61.70	ENSG00000127399
4.35	RUNX3	0.00	5.61	55.43	ENSG00000020633
4.34	C1QL4	0.02	10.08	84.09	ENSG00000186897
4.30	CHRDL2	0.01	7.12	66.28	ENSG00000054938
4.29	LPAR1	0.02	9.82	83.39	ENSG00000198121
4.23	SIX2	0.02	8.70	76.51	ENSG00000170577
4.22	SYT7	0.01	6.92	64.43	ENSG00000011347
4.22	GAL	1.20	92.20	70.95	ENSG00000069482
4.20	NEFH	0.34	32.21	72.76	ENSG00000100285
4.15	LSR	0.61	55.03	78.19	ENSG00000105699
4.15	CHRM4	0.01	5.73	55.48	ENSG00000180720
4.09	ULBP1	0.38	31.66	66.86	ENSG00000111981
4.09	SEMA7A	0.80	59.32	66.17	ENSG00000138623
4.09	CD24	0.38	31.96	67.06	ENSG00000185275
4.05	NMUR1	0.00	4.74	46.97	ENSG00000171596
4.02	COCH	0.12	15.63	71.48	ENSG00000100473
3.99	FAM46C	0.03	9.75	77.55	ENSG00000183508

Table 3.8 (continued): Significant host gene expression changes, ECs

Z-score	Gene Name	RPKM Mock	RPKM Infection	FoldChg Estimate	Ensembl ID
3.96	LHX2	0.01	6.80	61.05	ENSG00000106689
3.96	CDS1	0.02	7.81	67.62	ENSG00000163624
3.96	EOMES	0.00	4.46	44.30	ENSG00000163508
3.95	INA	0.19	19.12	66.74	ENSG00000148798
3.94	LINGO1	0.03	9.60	75.77	ENSG00000169783
3.94	ITM2C	2.08	127.57	58.62	ENSG00000135916
3.92	TMEM151A	0.06	11.89	76.34	ENSG00000179292
3.92	EBF3	0.01	5.26	51.09	ENSG00000108001
3.91	CGN	0.24	22.51	65.92	ENSG00000143375
3.89	CCDC64	0.04	9.92	73.12	ENSG00000135127
3.88	MIR147B	0.58	42.75	63.11	ENSG00000211519
3.87	ID4	0.01	5.49	52.76	ENSG00000172201
3.86	CPE	0.48	32.37	55.69	ENSG00000109472
3.85	DUSP2	0.17	16.99	64.24	ENSG00000158050
3.84	ERBB3	0.01	5.92	55.24	ENSG00000065361
3.83	FAM169A	0.10	13.74	69.53	ENSG00000198780
3.82	PCSK6	0.02	7.38	62.87	ENSG00000140479
3.81	NDRG2	0.02	7.99	66.87	ENSG00000165795
3.80	PDE8B	0.00	2.99	30.54	ENSG00000113231
3.79	NRTN	0.01	6.53	58.12	ENSG00000171119
3.79	C15orf48	0.39	26.83	54.95	ENSG00000166920
3.78	ELAVL2	0.00	3.43	34.60	ENSG00000107105
3.78	RRAGD	0.28	22.69	60.12	ENSG00000025039
3.74	TLX3	0.01	4.88	47.02	ENSG00000164438
3.72	GREM1	0.01	5.59	52.23	ENSG00000166923
3.71	GSG1L	0.00	3.25	32.81	ENSG00000169181
3.71	BEX2	0.94	53.64	51.72	ENSG00000133134
3.70	SLC27A6	0.03	7.82	63.33	ENSG00000113396
3.69	HAP1	0.20	16.67	56.46	ENSG00000173805
3.68	GPR27	0.00	4.10	40.79	ENSG00000170837
3.67	KCNA7	0.00	4.45	43.72	ENSG00000104848
3.67	HCP5	0.56	31.39	47.71	ENSG00000206337
3.67	L1CAM	0.12	13.81	62.39	ENSG00000198910

Table 3.9: Significant host gene expression changes, NPCs

Significant changes listed are based on Z-score analysis. The fold-change estimates listed were calculated as (RPKM infected + 0.1) / (RPKM mock + 0.1).

Z-score	Gene Name	RPKM Mock	RPKM Infection	FoldChg Estimate	Ensembl ID
First 100 most downregulated (by Z-score), 48 hpi					
-7.44	DIO3	26.40	1.57	-15.84	ENSG00000197406
-6.65	CXCR7	18.58	1.15	-14.98	ENSG00000144476
-5.84	HES5	66.55	6.84	-9.60	ENSG00000197921
-5.35	LAMP5	13.26	1.19	-10.36	ENSG00000125869
-5.30	HES3	4.44	0.10	-23.17	ENSG00000173673
-5.30	PLP1	23.14	3.14	-7.17	ENSG00000123560
-5.22	PCDH8	51.48	6.66	-7.63	ENSG00000136099
-4.89	PALMD	12.52	1.39	-8.47	ENSG00000099260
-4.75	MIAT	68.69	10.89	-6.26	ENSG00000225783
-4.74	TTYH1	85.62	14.16	-6.01	ENSG00000167614
-4.73	FABP7	513.89	85.19	-6.03	ENSG00000164434
-4.73	ASCL1	17.70	2.62	-6.53	ENSG00000139352
-4.73	DLL1	25.32	4.42	-5.63	ENSG00000198719
-4.70	SFRP2	212.52	35.66	-5.95	ENSG00000145423
-4.61	GADD45G	17.41	2.68	-6.29	ENSG00000130222
-4.55	N/A	17.97	2.99	-5.86	ENSG00000222466
-4.54	NR2F1	78.89	13.91	-5.64	ENSG00000175745
-4.52	DTX4	63.95	11.14	-5.70	ENSG00000110042
-4.48	CRABP1	1369.61	249.50	-5.49	ENSG00000166426
-4.34	DCX	8.78	1.02	-7.96	ENSG00000077279
-4.28	SEMA5B	29.45	6.11	-4.76	ENSG00000082684
-4.28	SMOC1	15.50	2.51	-5.99	ENSG00000198732
-4.24	CRB2	41.91	8.44	-4.92	ENSG00000148204
-4.16	DLL3	38.01	7.67	-4.91	ENSG00000090932
-4.04	GDPD2	11.66	1.92	-5.82	ENSG00000130055
-4.04	APLNR	17.73	3.58	-4.84	ENSG00000134817
-4.04	LAMA4	8.39	1.09	-7.14	ENSG00000112769
-4.01	JAG1	71.97	15.69	-4.57	ENSG00000101384
-3.97	TACR1	2.46	0.10	-12.52	ENSG00000115353
-3.96	A2M	109.07	24.59	-4.42	ENSG00000175899
-3.96	HNRNPA1P2	29.86	7.04	-4.19	ENSG00000237285
-3.95	EDNRB	37.50	8.25	-4.50	ENSG00000136160
-3.93	CYP26A1	6.07	0.58	-9.14	ENSG00000095596
-3.90	SNORD37	1.15	0.03	-9.31	ENSG00000212270
-3.89	TGFB2	42.00	9.60	-4.34	ENSG00000092969
-3.89	BTBD17	13.30	2.36	-5.45	ENSG00000204347
-3.86	MYCN	47.04	10.57	-4.42	ENSG00000134323
-3.86	SOX2-OT	13.64	2.58	-5.13	ENSG00000242808
-3.86	PTCH1	23.08	5.68	-4.01	ENSG00000185920
-3.85	RP11-216N14.5	5.65	0.47	-10.12	ENSG00000231827

Table 3.9 (continued): Significant host gene expression changes, NPCs

Z-score	Gene Name	RPKM Mock	RPKM Infection	FoldChg Estimate	Ensembl ID
-3.85	PCDH18	32.27	7.73	-4.13	ENSG00000189184
-3.84	FAM181A	15.32	3.00	-4.98	ENSG00000140067
-3.82	LOC100506421	37.72	9.11	-4.11	ENSG00000233639
-3.80	LINC00461	13.12	2.39	-5.30	ENSG00000245526
-3.77	LRIG1	93.35	22.26	-4.18	ENSG00000144749
-3.71	COL3A1	14.39	3.03	-4.64	ENSG00000168542
-3.69	RP11-263K19.4	19.30	4.74	-4.01	ENSG00000231064
-3.69	C3orf32	7.48	1.12	-6.22	ENSG00000125046
-3.69	CP	18.77	4.59	-4.02	ENSG00000047457
-3.66	AC010170.1	47.58	11.56	-4.09	ENSG00000178550
-3.65	ID4	38.47	9.84	-3.88	ENSG00000172201
-3.65	GAS1	15.23	3.43	-4.34	ENSG00000180447
-3.61	RGMA	54.42	13.76	-3.93	ENSG00000182175
-3.58	POU3F2	93.40	24.27	-3.84	ENSG00000184486
-3.57	APC2	10.97	2.18	-4.85	ENSG00000115266
-3.55	IGDCC3	136.03	35.28	-3.85	ENSG00000174498
-3.54	SOX21	40.81	10.71	-3.78	ENSG00000125285
-3.54	KLHDC8A	8.25	1.43	-5.46	ENSG00000162873
-3.53	GDF3	20.10	5.46	-3.64	ENSG00000184344
-3.51	TMEM132B	8.09	1.41	-5.43	ENSG00000139364
-3.47	MYCL1	23.79	6.44	-3.65	ENSG00000116990
-3.47	SLITRK2	50.31	13.28	-3.77	ENSG00000185985
-3.45	RFX4	43.09	11.61	-3.69	ENSG00000111783
-3.45	IFIT1	14.52	3.46	-4.11	ENSG00000185745
-3.43	DACH1	9.66	1.83	-5.06	ENSG00000165659
-3.43	SLIT1	11.62	2.58	-4.37	ENSG00000187122
-3.43	WNT7A	12.93	2.87	-4.39	ENSG00000154764
-3.40	TSPAN18	107.40	29.59	-3.62	ENSG00000157570
-3.40	HIST1H4K	31.16	8.94	-3.46	ENSG00000197914
-3.39	LRRN1	80.08	22.25	-3.59	ENSG00000175928
-3.38	EFNB2	59.45	16.35	-3.62	ENSG00000125266
-3.35	HIST1H4J	31.34	9.03	-3.44	ENSG00000197238
-3.35	POU3F4	10.85	2.44	-4.31	ENSG00000196767
-3.34	RNU1-60P	39.98	11.42	-3.48	ENSG00000202077
-3.34	SLC1A3	16.90	4.83	-3.45	ENSG00000079215
-3.31	ABAT	28.28	8.58	-3.27	ENSG00000183044
-3.29	SOX2	216.35	62.09	-3.48	ENSG00000181449
-3.28	AUTS2	51.68	14.74	-3.49	ENSG00000158321
-3.28	FOXJ1	9.18	1.94	-4.56	ENSG00000129654
-3.27	LUM	10.72	2.47	-4.20	ENSG00000139329
-3.23	GRIA1	29.98	9.38	-3.17	ENSG00000155511
-3.23	MIR219-2	2.54	0.21	-8.53	ENSG00000207955
-3.23	EPHA3	7.57	1.50	-4.81	ENSG00000044524
-3.23	FZD7	36.00	10.64	-3.36	ENSG00000155760
-3.20	PAX6	17.15	5.23	-3.24	ENSG00000007372
-3.19	INSM1	7.77	1.59	-4.65	ENSG00000173404

Table 3.9 (continued): Significant host gene expression changes, NPCs

Z-score	Gene Name	RPKM Mock	RPKM Infection	FoldChg Estimate	Ensembl ID
-3.19	RP11-463C14.1	2.75	0.25	-8.11	ENSG00000253668
-3.19	ITGB8	14.63	4.06	-3.54	ENSG00000105855
-3.18	AC078794.1	4.55	0.52	-7.46	ENSG00000221211
-3.17	MAP2	28.70	9.24	-3.09	ENSG00000078018
-3.16	MIR92B	1.52	0.10	-8.16	ENSG00000208011
-3.15	FGFR3	49.71	14.78	-3.35	ENSG00000068078
-3.14	SRGAP3	11.56	3.00	-3.76	ENSG00000196220
-3.13	RPRML	7.30	1.54	-4.51	ENSG00000179673
-3.13	ENC1	50.80	15.33	-3.30	ENSG00000171617
-3.13	NHLH1	6.17	1.11	-5.20	ENSG00000171786
-3.10	DBX1	2.64	0.26	-7.63	ENSG00000109851
-3.10	IGFBP5	177.43	54.74	-3.24	ENSG00000115461
-3.08	CXCL3	10.49	2.66	-3.84	ENSG00000163734
-3.08	HS6ST1P1	26.42	8.82	-2.97	ENSG00000187952
First 100 most upregulated (by Z-score), 48 hpi					
10.72	RBP7	0.46	36.54	66.02	ENSG00000162444
9.86	BARX1	0.01	5.85	56.67	ENSG00000131668
8.98	IRF5	0.00	1.68	17.82	ENSG00000128604
8.68	RBP4	0.01	4.59	44.64	ENSG00000138207
8.06	SLC30A2	0.00	2.62	26.65	ENSG00000158014
7.86	ECEL1	0.13	10.63	46.06	ENSG00000171551
7.77	TAC1	0.08	8.02	46.11	ENSG00000006128
7.57	CRLF1	0.46	17.33	30.96	ENSG00000006016
7.56	PLP2	0.31	15.07	36.73	ENSG00000102007
7.16	SHE	0.02	4.59	40.82	ENSG00000169291
7.01	HBA2	0.01	2.53	25.04	ENSG00000188536
6.97	EHD2	0.40	13.77	27.57	ENSG00000024422
6.65	SPINT1	0.38	12.67	26.88	ENSG00000166145
6.64	TBX2	0.01	3.00	28.42	ENSG00000121068
6.11	EPCAM	0.26	8.93	25.43	ENSG00000119888
6.03	INHBE	0.62	13.53	18.88	ENSG00000139269
5.94	BEX5	0.38	10.76	22.53	ENSG00000184515
5.94	INHBB	0.02	2.81	25.27	ENSG00000163083
5.69	CXCL16	0.05	4.89	33.05	ENSG00000161921
5.58	ESRP1	0.01	1.43	14.53	ENSG00000104413
5.57	TMBIM1	0.49	11.14	19.22	ENSG00000135926
5.56	EEF1A2	11.09	99.92	8.94	ENSG00000101210
5.56	GPX3	2.41	25.13	10.04	ENSG00000211445
5.44	STC2	0.82	13.14	14.45	ENSG00000113739
5.33	TRIB3	4.74	44.55	9.23	ENSG00000101255
5.28	GAL	2.67	24.82	8.99	ENSG00000069482
5.27	TMEM54	2.40	23.08	9.29	ENSG00000121900
5.27	WNT11	0.01	2.05	19.00	ENSG00000085741
5.25	KCNK6	0.16	6.23	24.24	ENSG00000099337
5.22	CD55	0.41	8.75	17.25	ENSG00000196352
5.20	NID1	37.19	268.41	7.20	ENSG00000116962

Table 3.9 (continued): Significant host gene expression changes, NPCs

Z-score	Gene Name	RPKM Mock	RPKM Infection	FoldChg Estimate	Ensembl ID
5.13	STXBP2	1.73	18.58	10.20	ENSG00000076944
5.10	KDELR3	2.70	23.67	8.49	ENSG00000100196
5.08	GYLTL1B	0.33	7.44	17.65	ENSG00000165905
5.07	N/A	0.18	6.12	22.55	ENSG00000187939
5.03	DUSP23	0.53	9.40	15.06	ENSG00000158716
4.99	KIAA1244	0.04	3.11	23.10	ENSG00000112379
4.99	PHOSPHO1	0.02	2.03	18.32	ENSG00000173868
4.96	TMEM59L	1.45	16.44	10.66	ENSG00000105696
4.78	ASCL2	0.02	1.78	16.07	ENSG00000183734
4.78	GJA3	0.01	1.12	11.36	ENSG00000121743
4.74	ENTPD2	0.00	1.05	11.08	ENSG00000054179
4.74	SLC15A1	0.01	1.18	11.71	ENSG00000088386
4.72	TFCP2L1	0.01	1.59	14.78	ENSG00000115112
4.71	TLX2	0.01	1.53	14.39	ENSG00000115297
4.68	FXYD5	3.59	27.29	7.42	ENSG00000089327
4.67	MIR147B	0.14	5.29	22.35	ENSG00000211519
4.63	TDRD10	0.01	1.48	13.88	ENSG00000163239
4.59	PAQR5	0.12	4.84	22.86	ENSG00000137819
4.53	S100A11	8.43	55.95	6.57	ENSG00000163191
4.52	ADRB1	0.02	1.49	13.70	ENSG00000043591
4.51	MCOLN3	0.18	5.49	20.05	ENSG00000055732
4.51	SLC18A2	0.02	1.67	15.03	ENSG00000165646
4.50	GIPC3	0.29	6.35	16.75	ENSG00000179855
4.49	KCNK13	0.01	1.12	11.17	ENSG00000152315
4.46	KCNA7	0.01	1.32	12.44	ENSG00000104848
4.45	PPP2R2C	0.02	1.81	15.49	ENSG00000074211
4.44	SLC16A3	5.11	34.11	6.57	ENSG00000141526
4.43	Y_RNA	0.12	4.46	21.09	ENSG00000207395
4.42	SPINT2	1.36	14.02	9.65	ENSG00000167642
4.42	COL8A2	0.03	2.11	16.98	ENSG00000171812
4.39	TCEAL2	0.68	8.95	11.68	ENSG00000184905
4.39	FBLN5	0.79	9.87	11.15	ENSG00000140092
4.33	COX6B2	0.82	9.73	10.67	ENSG00000160471
4.29	AGPAT9	0.13	4.47	19.80	ENSG00000138678
4.29	CPEB1	0.16	4.94	19.53	ENSG00000214575
4.28	FAM129A	0.78	9.16	10.58	ENSG00000135842
4.28	TGM2	0.12	4.31	19.88	ENSG00000198959
4.27	ADM2	1.69	14.93	8.38	ENSG00000128165
4.25	CD8A	0.02	1.48	13.14	ENSG00000153563
4.24	STAC2	0.76	9.00	10.56	ENSG00000141750
4.19	SNORA11	2.16	16.40	7.29	ENSG00000221164
4.16	MFSD6	1.18	11.45	9.01	ENSG00000151690
4.14	FGF19	0.03	1.94	15.70	ENSG00000162344
4.12	C15orf48	0.08	3.34	19.02	ENSG00000166920
4.11	CDS1	0.12	4.00	18.71	ENSG00000163624
4.08	PRKCH	0.21	5.20	16.98	ENSG00000027075
4.07	TUBA4A	0.70	8.02	10.20	ENSG00000127824

Table 3.9 (continued): Significant host gene expression changes, NPCs

Z-score	Gene Name	RPKM Mock	RPKM Infection	FoldChg Estimate	Ensembl ID
4.06	SCGB3A1	0.01	1.08	10.62	ENSG00000161055
4.05	LOC100132111	0.12	3.87	17.78	ENSG00000234614
4.05	TMEM151A	0.09	3.39	18.66	ENSG00000179292
4.02	NPNT	0.03	1.76	14.01	ENSG00000168743
4.01	ALDH1A2	0.06	2.54	16.09	ENSG00000128918
4.00	IL11	4.90	28.22	5.67	ENSG00000095752
4.00	CHAC1	8.75	48.01	5.44	ENSG00000128965
3.99	MKX	0.05	2.35	15.94	ENSG00000150051
3.93	DUSP2	0.12	3.59	17.08	ENSG00000158050
3.90	TRIM47	0.82	8.42	9.29	ENSG00000132481
3.87	GPR4	0.03	1.68	13.25	ENSG00000177464
3.86	GALR2	0.02	1.10	10.29	ENSG00000182687
3.86	MATN3	0.65	7.30	9.92	ENSG00000132031
3.80	CCNO	0.03	1.44	11.85	ENSG00000152669
3.79	RIMKLA	0.98	8.94	8.37	ENSG00000177181
3.78	PPIC	3.57	19.18	5.25	ENSG00000168938
3.75	TRH	0.05	2.06	14.18	ENSG00000170893
3.75	ELOVL7	0.16	3.89	15.59	ENSG00000164181
3.74	N/A	0.58	6.72	10.04	ENSG00000244202
3.71	FOXI2	0.03	1.54	12.27	ENSG00000186766
3.70	GMPR	0.18	4.10	15.00	ENSG00000137198
3.69	PHLDA2	6.43	31.50	4.84	ENSG00000181649
First 100 most downregulated (by Z-score), 96 hpi					
-6.73	AC078794.1	4.55	0.00	-45.18	ENSG00000221211
-5.74	GDF3	20.10	0.25	-58.04	ENSG00000184344
-5.37	IGFBP5	177.43	4.44	-39.15	ENSG00000115461
-5.21	MIR1247	1.78	0.00	-18.27	ENSG00000221133
-5.20	AC005863.1	1.76	0.00	-18.07	ENSG00000205325
-5.01	PALMD	12.52	0.20	-42.50	ENSG00000099260
-4.90	HES5	66.55	1.77	-35.68	ENSG00000197921
-4.88	N/A	17.97	0.40	-36.00	ENSG00000222466
-4.75	PLP1	23.14	0.73	-28.14	ENSG00000123560
-4.71	POU3F2	93.40	3.19	-28.41	ENSG00000184486
-4.66	BTBD17	13.30	0.30	-33.83	ENSG00000204347
-4.62	TGFB2	42.00	1.47	-26.78	ENSG00000092969
-4.43	LAMP5	13.26	0.36	-29.36	ENSG00000125869
-4.41	FABP7	513.89	25.56	-20.03	ENSG00000164434
-4.31	HES3	4.44	0.04	-32.44	ENSG00000173673
-4.31	SOX21	40.81	1.78	-21.74	ENSG00000125285
-4.20	FAM181A	15.32	0.54	-24.05	ENSG00000140067
-4.20	LINC00461	13.12	0.43	-25.08	ENSG00000245526
-4.14	SOX2	216.35	13.06	-16.45	ENSG00000181449
-4.08	DTX4	63.95	3.13	-19.84	ENSG00000110042
-4.06	SNORA5C	1.01	0.00	-10.62	ENSG00000201772
-4.05	EDNRB	37.50	2.01	-17.85	ENSG00000136160
-4.02	CXCR7	18.58	0.83	-20.18	ENSG00000144476
-3.89	DBX1	2.64	0.03	-21.23	ENSG00000109851

Table 3.9 (continued): Significant host gene expression changes, NPCs

Z-score	Gene Name	RPKM Mock	RPKM Infection	FoldChg Estimate	Ensembl ID
-3.79	NR2F1	78.89	5.53	-14.02	ENSG00000175745
-3.77	GDPD2	11.66	0.44	-21.86	ENSG00000130055
-3.75	MIAT	68.69	4.51	-14.91	ENSG00000225783
-3.74	MYCN	47.04	3.20	-14.28	ENSG00000134323
-3.73	RP11-82L18.4	11.36	0.42	-21.95	ENSG00000231632
-3.73	NES	560.20	45.43	-12.31	ENSG00000132688
-3.72	SLC2A1	277.19	22.55	-12.24	ENSG00000117394
-3.72	RGMA	54.42	3.77	-14.09	ENSG00000182175
-3.71	DACH1	9.66	0.33	-22.64	ENSG00000165659
-3.67	SLITRK2	50.31	3.63	-13.52	ENSG00000185985
-3.63	SOX2-OT	13.64	0.70	-17.16	ENSG00000242808
-3.63	VCAN	51.48	3.74	-13.44	ENSG00000038427
-3.60	AC010170.1	47.58	3.61	-12.85	ENSG00000178550
-3.57	SH3PXD2B	34.63	2.70	-12.41	ENSG00000174705
-3.57	HS6ST1P1	26.42	2.30	-11.07	ENSG00000187952
-3.56	LOC100506421	37.72	2.84	-12.85	ENSG00000233639
-3.54	SFRP2	212.52	19.77	-10.70	ENSG00000145423
-3.53	CRABP1	1369.61	128.33	-10.67	ENSG00000166426
-3.51	CYP26A1	6.07	0.18	-21.72	ENSG00000095596
-3.49	EPHA3	7.57	0.25	-21.90	ENSG00000044524
-3.47	LRIG1	93.35	8.24	-11.21	ENSG00000144749
-3.47	GLI2	18.47	1.28	-13.44	ENSG00000074047
-3.46	KCNJ6	1.25	0.02	-11.61	ENSG00000157542
-3.46	CRB2	41.91	3.40	-11.99	ENSG00000148204
-3.45	TMEM132B	8.09	0.30	-20.58	ENSG00000139364
-3.45	TCRDV2	2.98	0.06	-19.36	ENSG00000211829
-3.45	LAMA4	8.39	0.34	-19.16	ENSG00000112769
-3.42	PCDH18	32.27	2.81	-11.12	ENSG00000189184
-3.42	PTCH1	23.08	2.12	-10.43	ENSG00000185920
-3.41	GRIA1	29.98	2.77	-10.49	ENSG00000155511
-3.41	RFX4	43.09	3.70	-11.36	ENSG00000111783
-3.41	EDNRA	5.32	0.14	-22.47	ENSG00000151617
-3.40	C3orf32	7.48	0.27	-20.43	ENSG00000125046
-3.40	KLHDC8A	8.25	0.33	-19.55	ENSG00000162873
-3.39	DCX	8.78	0.38	-18.47	ENSG00000077279
-3.39	RP11-216N14.5	5.65	0.16	-21.86	ENSG00000231827
-3.39	N/A	1.05	0.01	-10.30	ENSG00000230270
-3.39	CP	18.77	1.39	-12.66	ENSG00000047457
-3.36	TRIM45	12.24	0.76	-14.41	ENSG00000134253
-3.33	DPYSL5	86.49	8.39	-10.20	ENSG00000157851
-3.32	AUTS2	51.68	4.70	-10.79	ENSG00000158321
-3.30	POU3F4	10.85	0.58	-16.06	ENSG00000196767
-3.29	APC2	10.97	0.58	-16.33	ENSG00000115266
-3.27	SRGAP3	11.56	0.70	-14.55	ENSG00000196220
-3.25	LRRC55	1.22	0.02	-10.93	ENSG00000183908
-3.24	ALPK2	14.42	1.07	-12.41	ENSG00000198796

Table 3.9 (continued): Significant host gene expression changes, NPCs

Z-score	Gene Name	RPKM Mock	RPKM Infection	FoldChg Estimate	Ensembl ID
-3.24	A2M	109.07	12.09	-8.96	ENSG00000175899
-3.24	AL832007	66.40	6.45	-10.15	ENSG00000150687
-3.22	SORBS2	25.10	2.82	-8.64	ENSG00000154556
-3.21	VEPH1	2.88	0.08	-16.76	ENSG00000197415
-3.19	TSPAN18	107.40	12.27	-8.69	ENSG00000157570
-3.17	SALL2	38.26	3.83	-9.77	ENSG00000165821
-3.17	FGFR2	13.80	1.07	-11.85	ENSG00000066468
-3.15	PPP1R17	1.23	0.03	-10.65	ENSG00000106341
-3.15	CXXC4	13.83	1.08	-11.83	ENSG00000168772
-3.15	LGI1	6.46	0.26	-18.16	ENSG00000108231
-3.13	CBX2	34.43	3.71	-9.06	ENSG00000173894
-3.13	SPTLC3	1.33	0.03	-11.24	ENSG00000172296
-3.13	MEX3A	31.15	3.28	-9.25	ENSG00000254726
-3.13	SALL1	38.02	3.93	-9.45	ENSG00000103449
-3.13	AL356740.1	82.03	9.20	-8.83	ENSG00000234603
-3.12	SMOC1	15.50	1.38	-10.56	ENSG00000198732
-3.11	APLNR	17.73	1.57	-10.68	ENSG00000134817
-3.09	POU3F3	20.15	2.01	-9.61	ENSG00000198914
-3.07	GAS1	15.23	1.37	-10.45	ENSG00000180447
-3.05	HS6ST1	54.28	6.28	-8.52	ENSG00000136720
-3.04	MARCKS	222.98	29.71	-7.48	ENSG00000155130
-3.03	PLEKHG1	25.76	3.22	-7.79	ENSG00000120278
-3.00	LRRN1	80.08	9.85	-8.05	ENSG00000175928
-3.00	ODZ3	33.88	4.05	-8.19	ENSG00000218336
-2.97	CDCA4	23.34	3.06	-7.41	ENSG00000170779
-2.97	ALK	2.86	0.10	-14.50	ENSG00000171094
-2.96	JAG1	71.97	9.20	-7.75	ENSG00000101384
-2.95	PTX3	49.20	5.92	-8.20	ENSG00000163661
-2.95	VCAM1	2.86	0.11	-14.38	ENSG00000162692
-2.93	RNVU1-14	24.90	3.39	-7.16	ENSG00000207501
First 100 most upregulated (by Z-score), 96 hpi					
9.04	HSPA6	0.11	101.49	495.57	ENSG00000173110
8.43	HSPA7	0.08	43.80	245.22	ENSG00000225217
7.85	HBA2	0.01	8.51	81.97	ENSG00000188536
7.52	RP11-44M6.3	0.29	92.98	240.51	ENSG00000184414
7.02	RBP4	0.01	6.82	65.90	ENSG00000138207
6.67	FOSB	0.64	105.09	142.73	ENSG00000125740
6.64	BARX1	0.01	5.83	56.44	ENSG00000131668
6.31	ECEL1	0.13	16.91	73.00	ENSG00000171551
6.19	DLX2	0.04	9.66	71.75	ENSG00000115844
6.13	RBP7	0.46	47.39	85.56	ENSG00000162444
5.83	CRLF1	0.46	26.15	46.63	ENSG00000006016
5.82	MIR22	0.65	52.71	70.13	ENSG00000199060
5.75	TAC1	0.08	11.43	65.52	ENSG00000006128
5.67	IRF5	0.00	1.29	13.90	ENSG00000128604
5.62	MIR22HG	2.72	182.69	64.80	ENSG00000186594
5.52	PLP2	0.31	20.41	49.66	ENSG00000102007

Table 3.9 (continued): Significant host gene expression changes, NPCs

Z-score	Gene Name	RPKM Mock	RPKM Infection	FoldChg Estimate	Ensembl ID
5.49	TCAM1P	0.00	1.63	17.32	ENSG00000240280
5.44	SLC30A2	0.00	2.57	26.15	ENSG00000158014
5.39	HLA-F-AS1	0.00	2.78	27.94	ENSG00000214922
5.38	AK311445	0.03	6.51	52.42	ENSG00000226334
5.31	KLF4	0.18	12.58	45.95	ENSG00000136826
5.31	BEX5	0.38	20.93	43.63	ENSG00000184515
5.23	EGR2	0.34	19.23	44.03	ENSG00000122877
5.02	SPINT1	0.38	18.82	39.83	ENSG00000166145
4.97	MYLIP	0.45	20.17	36.59	ENSG00000007944
4.90	RNU2-59P	3.45	139.09	39.20	ENSG00000222414
4.84	CD55	0.41	17.73	34.76	ENSG00000196352
4.83	RN7SL513P	0.00	2.09	21.50	ENSG00000239821
4.83	AC011385.1	1.00	34.64	31.61	ENSG00000197830
4.77	RP11-90K17.2	0.00	2.19	22.22	ENSG00000255005
4.74	TBX2	0.01	3.53	33.30	ENSG00000121068
4.74	RP1-167G20.1	0.01	4.05	36.71	ENSG00000248150
4.74	MIR147B	0.14	9.91	41.53	ENSG00000211519
4.68	CXCL16	0.05	6.68	44.88	ENSG00000161921
4.64	ITGAM	0.00	2.03	20.72	ENSG00000169896
4.63	SHE	0.02	4.05	36.09	ENSG00000169291
4.59	TRH	0.05	6.05	40.49	ENSG00000170893
4.58	KLF2	0.15	9.32	38.30	ENSG00000127528
4.56	WNT11	0.01	3.65	33.15	ENSG00000085741
4.53	CTD-2530H12.4	0.04	5.56	40.69	ENSG00000255326
4.51	CD8A	0.02	4.18	35.68	ENSG00000153563
4.48	FAM46C	0.05	5.71	39.82	ENSG00000183508
4.45	DUSP2	0.12	8.38	39.25	ENSG00000158050
4.42	AK123889	0.08	7.09	39.73	ENSG00000225914
4.41	DDIT3	7.11	199.16	27.63	ENSG00000175197
4.40	PTGER3	0.03	4.52	36.63	ENSG00000050628
4.39	SGK1	0.59	18.09	26.56	ENSG00000118515
4.37	GATA2	0.01	2.81	26.48	ENSG00000179348
4.36	GAL3ST1	0.11	7.86	38.06	ENSG00000128242
4.35	ANKRD37	0.74	20.03	23.96	ENSG00000186352
4.32	C17orf105	0.01	2.43	23.68	ENSG00000231256
4.31	SNORD15A	0.44	13.22	24.89	ENSG00000206941
4.27	GADD45B	2.80	76.64	26.47	ENSG00000099860
4.24	C15orf48	0.08	6.39	35.86	ENSG00000166920
4.22	RNU2-61P	0.06	5.74	36.70	ENSG00000223001
4.21	TMEM59L	1.45	29.48	19.06	ENSG00000105696
4.20	EGR3	0.24	10.18	30.05	ENSG00000179388
4.20	RP11-59D5__B.2	0.25	10.31	30.09	ENSG00000236345
4.17	EGR4	0.01	1.94	19.41	ENSG00000135625
4.15	TMBIM1	0.49	12.75	21.97	ENSG00000135926
4.15	TDRD10	0.01	3.04	27.53	ENSG00000163239

Table 3.9 (continued): Significant host gene expression changes, NPCs

Z-score	Gene Name	RPKM Mock	RPKM Infection	FoldChg Estimate	Ensembl ID
4.14	RNU2-64P	0.12	7.41	34.44	ENSG00000223247
4.08	HSPA1L	1.57	25.68	15.48	ENSG00000204390
4.06	DUSP23	0.53	13.00	20.76	ENSG00000158716
4.02	FAM138A	0.00	1.25	13.08	ENSG00000233630
4.02	RNU6-384P	0.04	4.18	31.68	ENSG00000201770
4.01	BAAT	0.06	5.41	34.45	ENSG00000136881
4.01	MGLL	0.01	1.88	18.72	ENSG00000074416
3.98	HBEGF	2.64	56.92	20.80	ENSG00000113070
3.96	ARC	0.50	12.13	20.35	ENSG00000198576
3.95	STXBP2	1.73	26.77	14.67	ENSG00000076944
3.94	RNU2-70P	0.28	9.76	25.68	ENSG00000222650
3.89	CRYAB	0.01	2.36	22.16	ENSG00000109846
3.84	BC031979	0.04	4.10	29.77	ENSG00000254842
3.82	PHOSPHO1	0.02	2.58	23.06	ENSG00000173868
3.81	ENTPD2	0.00	1.32	13.65	ENSG00000054179
3.80	HSPA1A	29.37	533.63	18.11	ENSG00000204389
3.79	KCNK6	0.16	7.19	27.93	ENSG00000099337
3.78	HERPUD1	11.13	198.57	17.70	ENSG00000051108
3.74	GPX3	2.41	40.89	16.31	ENSG00000211445
3.74	SLC5A5	0.02	2.87	24.11	ENSG00000105641
3.74	HSPA1B	43.17	746.18	17.25	ENSG00000204388
3.70	CCK	0.00	1.18	12.31	ENSG00000187094
3.67	MOBP	0.01	2.26	20.86	ENSG00000168314
3.67	SPINT2	1.36	21.53	14.79	ENSG00000167642
3.66	PCDH10	0.22	8.10	25.30	ENSG00000138650
3.66	N/A	0.01	1.92	18.37	ENSG00000250991
3.65	NR4A3	0.72	12.87	15.80	ENSG00000119508
3.65	SPTY2D1	5.40	94.94	17.28	ENSG00000179119
3.64	ASS1	1.04	17.40	15.41	ENSG00000130707
3.63	TMEM54	2.40	36.72	14.76	ENSG00000121900
3.59	SNAI1	5.05	84.11	16.36	ENSG00000124216
3.57	PPP1R15A	12.55	193.17	15.28	ENSG00000087074
3.56	RFPL4A	0.01	1.97	18.64	ENSG00000223638
3.55	S100A14	0.09	5.11	28.16	ENSG00000189334
3.54	BHLHE40	1.37	19.72	13.48	ENSG00000134107
3.53	ESRP1	0.01	1.27	13.00	ENSG00000104413
3.53	ULBP1	2.28	25.34	10.68	ENSG00000111981
3.53	SCGB3A1	0.01	1.86	17.68	ENSG00000161055
3.51	CALCA	0.01	1.32	13.38	ENSG00000110680

3.7. References

Adair, R., Liebisch, G.W., Su, Y., and Colberg-Poley, A.M. (2004). Alteration of cellular RNA splicing and polyadenylation machineries during productive human cytomegalovirus infection. J Gen Virol *85*, 3541-3553.

Bava, F.A., Eliscovich, C., Ferreira, P.G., Minana, B., Ben-Dov, C., Guigo, R., Valcarcel, J., and Mendez, R. (2013). CPEB1 coordinates alternative 3'-UTR formation with translational regulation. Nature *495*, 121-125.

Belzile, J.P., Stark, T.J., Yeo, G.W., and Spector, D.H. (2014). Human cytomegalovirus infection of human embryonic stem cell-derived primitive neural stem cells is restricted at several steps but leads to the persistence of viral DNA. J Virol *88*, 4021-4039.

Browne, E.P., Wing, B., Coleman, D., and Shenk, T. (2001). Altered cellular mRNA levels in human cytomegalovirus-infected fibroblasts: viral block to the accumulation of antiviral mRNAs. J Virol *75*, 12319-12330.

Chen, Z., Knutson, E., Kurosky, A., and Albrecht, T. (2001). Degradation of p21cip1 in cells productively infected with human cytomegalovirus. J Virol *75*, 3613-3625.

Ellison, K.S., Rice, S.A., Verity, R., and Smiley, J.R. (2000). Processing of alpha-globin and ICP0 mRNA in cells infected with herpes simplex virus type 1 ICP27 mutants. J Virol *74*, 7307-7319.

Farber-Katz, S.E., Dippold, H.C., Buschman, M.D., Peterman, M.C., Xing, M., Noakes, C.J., Tat, J., Ng, M.M., Rahajeng, J., Cowan, D.M., Fuchs, G.J., Zhou, H., and Field, S.J. (2014). DNA damage triggers Golgi dispersal via DNA-PK and GOLPH3. Cell *156*, 413-427.

Fox, C.A., Sheets, M.D., and Wickens, M.P. (1989). Poly(A) addition during maturation of frog oocytes: distinct nuclear and cytoplasmic activities and regulation by the sequence UUUUUAU. Genes & development *3*, 2151-2162.

Gatherer, D., Seirafian, S., Cunningham, C., Holton, M., Dargan, D.J., Baluchova, K., Hector, R.D., Galbraith, J., Herzyk, P., Wilkinson, G.W., and Davison, A.J. (2011). High-resolution human cytomegalovirus transcriptome. Proc Natl Acad Sci U S A *108*, 19755-19760.

Glaunsinger, B., and Ganem, D. (2004). Lytic KSHV infection inhibits host gene expression by accelerating global mRNA turnover. Mol Cell *13*, 713-723.

Hanna, R.A., Campbell, R.L., and Davies, P.L. (2008). Calcium-bound structure of calpain and its mechanism of inhibition by calpastatin. Nature *456*, 409-412.

Hardwicke, M.A., and Sandri-Goldin, R.M. (1994). The herpes simplex virus regulatory protein ICP27 contributes to the decrease in cellular mRNA levels during infection. J Virol *68*, 4797-4810.

Hardy, W.R., and Sandri-Goldin, R.M. (1994). Herpes simplex virus inhibits host cell splicing, and regulatory protein ICP27 is required for this effect. J Virol *68*, 7790-7799.

Hertel, L., and Mocarski, E.S. (2004). Global analysis of host cell gene expression late during cytomegalovirus infection reveals extensive dysregulation of cell cycle gene expression and induction of Pseudomitosis independent of US28 function. J Virol *78*, 11988-12011.

Ho, A.M., Johnson, M.D., and Kingsley, D.M. (2000). Role of the mouse ank gene in control of tissue calcification and arthritis. Science *289*, 265-270.

Homa, N.J., Salinas, R., Forte, E., Robinson, T.J., Garcia-Blanco, M.A., and Luftig, M.A. (2013). Epstein-Barr virus induces global changes in cellular mRNA isoform usage that are important for the maintenance of latency. J Virol *87*, 12291-12301.

Huelga, S.C., Vu, A.Q., Arnold, J.D., Liang, T.Y., Liu, P.P., Yan, B.Y., Donohue, J.P., Shiue, L., Hoon, S., Brenner, S., Ares, M., Jr., and Yeo, G.W. (2012). Integrative genome-wide analysis reveals cooperative regulation of alternative splicing by hnRNP proteins. Cell Rep *1*, 167-178.

Isler, J.A., Maguire, T.G., and Alwine, J.C. (2005a). Production of infectious human cytomegalovirus virions is inhibited by drugs that disrupt calcium homeostasis in the endoplasmic reticulum. J Virol *79*, 15388-15397.

Isler, J.A., Skalet, A.H., and Alwine, J.C. (2005b). Human cytomegalovirus infection activates and regulates the unfolded protein response. J Virol *79*, 6890-6899.

Ji, Z., Lee, J.Y., Pan, Z., Jiang, B., and Tian, B. (2009). Progressive lengthening of 3' untranslated regions of mRNAs by alternative polyadenylation during mouse embryonic development. Proc Natl Acad Sci U S A *106*, 7028-7033.

Katz, Y., Wang, E.T., Airoldi, E.M., and Burge, C.B. (2010). Analysis and design of RNA sequencing experiments for identifying isoform regulation. Nat Methods *7*, 1009-1015.

Kim, K.C., Kim, J.W., Choi, C.S., Han, S.Y., Cheong, J.H., Han, S.H., Yang, S.I., Bahn, G.H., and Shin, C.Y. (2013). A role of CPEB1 in the modulation of proliferation and neuronal maturation of rat primary neural progenitor cells. Neurochemical research *38*, 1960-1972.

Kumar, G.R., and Glaunsinger, B.A. (2010). Nuclear import of cytoplasmic poly(A) binding protein restricts gene expression via hyperadenylation and nuclear retention of mRNA. Mol Cell Biol *30*, 4996-5008.

Lee, Y.J., and Glaunsinger, B.A. (2009). Aberrant herpesvirus-induced polyadenylation correlates with cellular messenger RNA destruction. PLoS Biol *7*, e1000107.

Liu, R., Baillie, J., Sissons, J.G., and Sinclair, J.H. (1994). The transcription factor YY1 binds to negative regulatory elements in the human cytomegalovirus major immediate early enhancer/promoter and mediates repression in non-permissive cells. Nucleic Acids Res *22*, 2453-2459.

Luo, W., Ji, Z., Pan, Z., You, B., Hoque, M., Li, W., Gunderson, S.I., and Tian, B. (2013). The conserved intronic cleavage and polyadenylation site of CstF-77 gene imparts control of 3' end processing activity through feedback autoregulation and by U1 snRNP. PLoS genetics *9*, e1003613.

Martin, G., Gruber, A.R., Keller, W., and Zavolan, M. (2012). Genome-wide analysis of pre-mRNA 3' end processing reveals a decisive role of human cleavage factor I in the regulation of 3' UTR length. Cell Rep *1*, 753-763.

Mayr, C., and Bartel, D.P. (2009). Widespread shortening of 3'UTRs by alternative cleavage and polyadenylation activates oncogenes in cancer cells. Cell *138*, 673-684.

Mocarski, E.S., T. Shenk, and R. F. Pass. (2007). Cytomegaloviruses. In Fields' virology, D.M.K.a.P.M. Howley, ed. (Philadelphia, PA., Lippincott Williams & Wilkins), pp. 2701–2772.

Morava, E., Kuhnisch, J., Drijvers, J.M., Robben, J.H., Cremers, C., van Setten, P., Branten, A., Stumpp, S., de Jong, A., Voesenek, K., Vermeer, S., Heister, A., Claahsen-van der Grinten, H.L., O'Neill, C.W., Willemsen, M.A., Lefeber, D., Deen, P.M., Kornak, U., Kremer, H., and Wevers, R.A. (2011). Autosomal recessive mental retardation, deafness, ankylosis, and mild hypophosphatemia associated with a novel ANKH mutation in a consanguineous family. The Journal of clinical endocrinology and metabolism *96*, E189-198.

Mortazavi, A., Williams, B.A., McCue, K., Schaeffer, L., and Wold, B. (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat Methods *5*, 621-628.

Moucadel, V., Lopez, F., Ara, T., Benech, P., and Gautheret, D. (2007). Beyond the 3' end: experimental validation of extended transcript isoforms. Nucleic Acids Res *35*, 1947-1957.

Nemeroff, M.E., Barabino, S.M., Li, Y., Keller, W., and Krug, R.M. (1998). Influenza virus NS1 protein interacts with the cellular 30 kDa subunit of CPSF and inhibits 3'end formation of cellular pre-mRNAs. Mol Cell *1*, 991-1000.

Noraz, N., Gozlan, J., Corbeil, J., Brunner, T., and Spector, S.A. (1997). HIV-induced apoptosis of activated primary CD4+ T lymphocytes is not mediated by Fas-Fas ligand. Aids *11*, 1671-1680.

Parkhomchuk, D., Borodina, T., Amstislavskiy, V., Banaru, M., Hallen, L., Krobitsch, S., Lehrach, H., and Soldatov, A. (2009). Transcriptome analysis by strand-specific sequencing of complementary DNA. Nucleic Acids Res *37*, e123.

Polymenidou, M., Lagier-Tourenne, C., Hutt, K.R., Huelga, S.C., Moran, J., Liang, T.Y., Ling, S.C., Sun, E., Wancewicz, E., Mazur, C., Kordasiewicz, H., Sedaghat, Y., Donohue, J.P., Shiue, L., Bennett, C.F., Yeo, G.W., and Cleveland, D.W. (2011). Long pre-mRNA depletion and RNA missplicing contribute to neuronal vulnerability from loss of TDP-43. Nat Neurosci *14*, 459-468.

Richter, J.D. (2007). CPEB: a life in translation. Trends in biochemical sciences *32*, 279-285.

Sandberg, R., Neilson, J.R., Sarma, A., Sharp, P.A., and Burge, C.B. (2008). Proliferating cells express mRNAs with shortened 3' untranslated regions and fewer microRNA target sites. Science *320*, 1643-1647.

Schek, N., and Bachenheimer, S.L. (1985). Degradation of cellular mRNAs induced by a virion-associated factor during herpes simplex virus infection of Vero cells. J Virol *55*, 601-610.

Staras, S.A., Dollard, S.C., Radford, K.W., Flanders, W.D., Pass, R.F., and Cannon, M.J. (2006). Seroprevalence of cytomegalovirus infection in the United States, 1988-1994. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America *43*, 1143-1151.

Stern-Ginossar, N., Weisburd, B., Michalski, A., Le, V.T., Hein, M.Y., Huang, S.X., Ma, M., Shen, B., Qian, S.B., Hengel, H., Mann, M., Ingolia, N.T., and Weissman, J.S. (2012). Decoding human cytomegalovirus. Science *338*, 1088-1093.

Tian, B., Hu, J., Zhang, H., and Lutz, C.S. (2005). A large-scale analysis of mRNA polyadenylation of human and mouse genes. Nucleic Acids Res *33*, 201-212.

Towler, J.C., Ebrahimi, B., Lane, B., Davison, A.J., and Dargan, D.J. (2012). Human cytomegalovirus transcriptome activity differs during replication in human fibroblast, epithelial and astrocyte cell lines. J Gen Virol *93*, 1046-1058.

Ulitsky, I., Shkumatava, A., Jan, C.H., Subtelny, A.O., Koppstein, D., Bell, G.W., Sive, H., and Bartel, D.P. (2012). Extensive alternative polyadenylation during zebrafish development. Genome Res *22*, 2054-2066.

Wang, G.S., and Cooper, T.A. (2007). Splicing in disease: disruption of the splicing code and the decoding machinery. Nature reviews. Genetics *8*, 749-761.

Weng, K.F., Li, M.L., Hung, C.T., and Shih, S.R. (2009). Enterovirus 71 3C protease cleaves a novel target CstF-64 and inhibits cellular polyadenylation. PLoS Pathog *5*, e1000593.

Wilczynska, A., Aigueperse, C., Kress, M., Dautry, F., and Weil, D. (2005). The translational regulator CPEB1 provides a link between dcp1 bodies and stress granules. Journal of cell science *118*, 981-992.

Wing, B.A., and Huang, E.S. (1995). Analysis and mapping of a family of 3'-coterminal transcripts containing coding sequences for human cytomegalovirus open reading frames UL93 through UL99. J Virol *69*, 1521-1531.

Yao, C., Biesinger, J., Wan, J., Weng, L., Xing, Y., Xie, X., and Shi, Y. (2012). Transcriptome-wide analyses of CstF64-RNA interactions in global regulation of mRNA alternative polyadenylation. Proc Natl Acad Sci U S A *109*, 18773-18778.

Chapter 4: Analysis of higher-order chromatin structure in cells infected with human cytomegalovirus.

4.1. Introduction

Human cytomegalovirus (HCMV) infection results in numerous, significant host gene expression changes, affecting a complex set of cellular pathways. These events have been catalogued in previous gene expression studies of HCMV infection as described in Chapter 1, and in multiple cell types from this dissertation study as presented in Chapter 3. Our knowledge is limited in terms of how most of these gene expression cascades are influenced during infection, and it is unclear which transcription factors from the host are responsible, or which gene expression changes are attributable to factors encoded by HCMV.

An alternative way of approaching this problem is to consider the spatial conformation of DNA, both host and viral, within HCMV-infected cells. A subset of infection-altered gene expression events could conceivably be explained by chromatin organization, especially if this higher-order structure in any way changes during infection. Furthermore, if we could find evidence for interaction between the HCMV genome and host chromosomes, this level of interaction could greatly impact *cis*-acting elements found in the host DNA.

Methods have been devised for analysis of the spatial organization of DNA within cells, and a genome-wide approach has emerged known as Hi-C (Lieberman-Aiden et al., 2009). Similar to previous techniques, this approach implements a restriction digest of formaldehyde-crosslinked DNA followed by ligation of proximal fragments. The main advancement of Hi-C is the strategy for purifying ligation products and generation of

libraries that are amenable to high-throughput sequencing. The centers of purified sequences represent original interaction points between DNA elements in the cells analyzed, and this information is resolved through paired-end deep sequencing.

Hi-C experiments have been implemented to study higher-order chromatin structure in human and mouse embryonic stem (ES) cells versus differentiated cells (Dixon et al., 2012). Striking similarity of the interaction domains could be observed between pluripotent and differentiated cell types and even across species. Additionally, in another related analysis of human fibroblasts before and after TNF-α treatment, interaction domains were determined to be constant despite the signaling events incurred by these cells (Jin et al., 2013). Furthermore, pre-existing interaction domains were identified that could explain gene expression events encountered during the treatment, suggesting that a poised chromatin landscape exists and facilitates gene activation or repression as needed in cells.

An interesting example of interactions between viral and host DNA has been documented for human papillomaviruses (Knipe et al., 2013; McBride, 2008). In this case the viral genome interacts with host mitotic chromosomes in order to persist and maintain its genome in dividing cells. Reportedly, gamma-herpesviruses also utilize similar mechanisms of viral genome maintenance in dividing cells (McBride, 2008).

This pilot study presented here has been conducted in part to determine whether HCMV-host DNA interactions occur during infection. Although the initial analysis has been limited to this question, current analysis efforts are in place to determine whether changes in host DNA architecture are evident post-infection, or if instead pre-existing interaction domains predominate and account for gene regulation events during infection.

4.2. Preliminary Results

Here for this study an early infection timepoint, 8 h post-infection (hpi) with the laboratory strain Towne, was chosen and subjected to Hi-C analysis. The 8 hpi early infection timepoint was selected in order to conduct measurements prior to the onset of viral DNA replication, reasoning that the high copy number of viral genomes present at later stages of infection could complicate the analysis. Downstream computational analysis of Hi-C experiments requires large amounts of sequencing coverage, and in this case Illumina deep sequencing was performed to generate an average of 1 billion reads per sample (Table 4.1).

The probability of obtaining HCMV-mapping material in this DNA analysis was hypothesized to be very low, considering the small sampling size of the HCMV genome versus all human chromosomes combined. However, a significant number of HCMV-containing reads were in fact obtained in the infected Hi-C dataset, and my analysis so far has focused on this subset of reads. Paired-end Hi-C sequencing data yields mate-pair read chimeras that represent interactions between proximal DNA fragments, and interestingly both virus-virus pairs and virus-human pairs could be observed in this experiment (numbers listed in Table 4.1), in addition to the expected human-human chimeras which constituted the majority of the sequencing. The virus-virus interactions obtained could be useful to investigate the conformation of the viral genome at this early stage of infection, and the virus-human chimeras suggest potential interactions between the HCMV genome and host chromosomes. The distribution of the HCMV-mapping reads from these virus-host interaction pairs are shown in Figure 4.1. The highest coverage points of these reads were centered around HindIII restriction sites in the HCMV genome, which is consistent with the manner in which interacting DNA fragments were isolated during the library generation. Reminiscent of a more general interaction

between the HCMV genome and the host cell DNA, all of the HCMV HindIII sites had significant sequencing coverage, rather than enrichment at only specific loci. However, a select few HCMV HindIII sites did harbor greater coverage than others (Figure 4.1), and further analysis testing will need to be performed to determine whether this differential is significant. For the human-mapping portions of the virus-human chimeras, these reads were well distributed across the human genome, and not concentrated at particular locations. In summary, these first rounds of analysis suggest that there are general, potentially widespread interactions between viral and host DNA in HCMV-infected cells. For future follow-up studies, targeted versions of chromosome conformation experiments could be performed to independently validate this observation, and DNA-capture strategies could be employed to increase the resolution of HCMV-specific analysis.

4.3. Materials and Methods

HFF cells were mock-infected (spent media containing 1% DMSO) or infected with HCMV Towne at MOI 3, and at 8hpi cells were trypsinized and crosslinked by resuspension in 1% formaldehyde for 10 mins at room temperature (RT). Samples were quenched with glycine (0.125 M final concentration) for 5 mins at RT and 15 additional mins on ice. Following a final spin, cell pellets were snap-frozen in liquid nitrogen and stored at -80°C. For this experiment, sequencing libraries were generated from 10 million cells per sample.

Hi-C libraries were prepared as outlined previously (Lieberman-Aiden et al., 2009). Reagent volumes and incubation times were adhered to as indicated in the Lieberman-Aiden supplementary methods. The parameters used for sonication on the Covaris S2 machine differed slightly, and were as follows: Duty cycle 10, Intensity 4, Cycles/burst 200, Time 55 seconds, 1 cycle. Final sequencing library concentrations

were determined by qPCR. Illumina paired-end sequencing was performed at 36 cycles on the Hi-Seq platform. Following an initial small-scale quality-control sequencing run (in part to confirm the presence of HCMV-mapping reads), a full flow-cell of sequencing was obtained for the mock and infected samples.

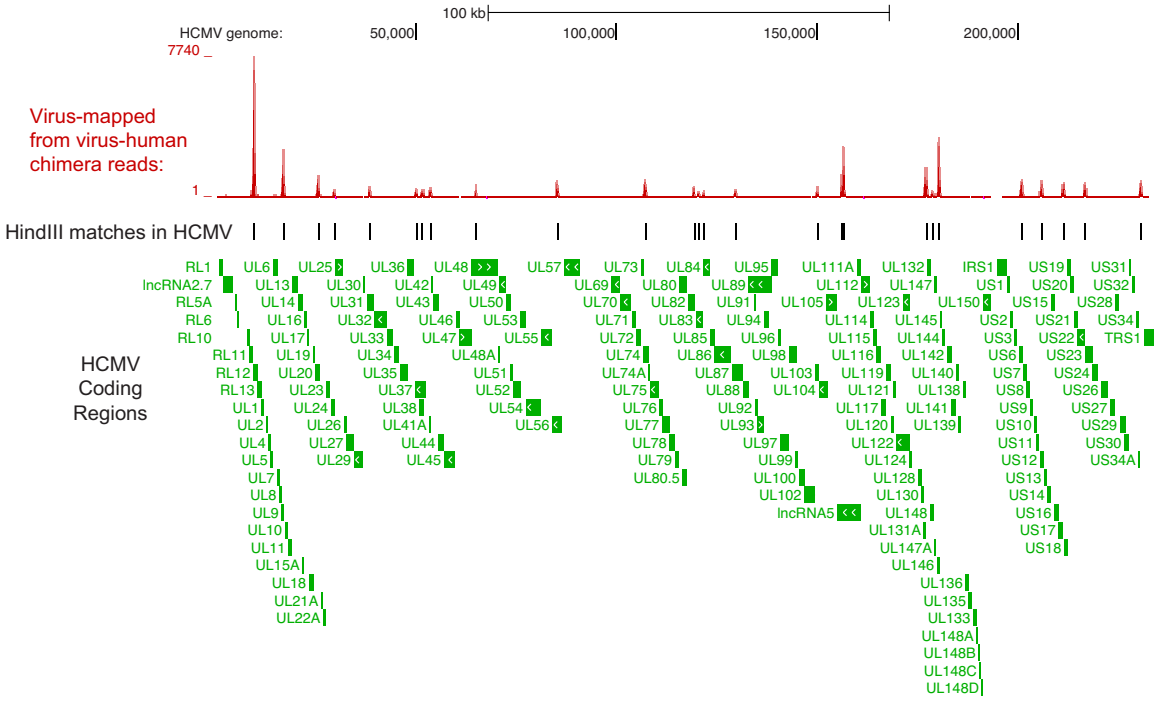

Figure 4.1: Visualization of HCMV-mapped Hi-C data from virus-human chimeras

Distribution of the HCMV-aligning portions of virus-human read pairs across the HCMV genome. The relative sequencing coverage of alignments is graphed in red and demonstrates enrichment at all HindIII restriction site locations. Genbank annotations for HCMV coding regions and non-coding RNAs are listed in green from the bottom.

Table 4.1: Summary statistics from Hi-C sequencing libraries

Samples	Raw Reads	Uniquely Mapped	Non-redundant
Mock	822,479,749	481,263,253	381,633,111
Infected	1,043,252,689	596,840,002	466,667,960
Chimeras	**Reads, uniquely mapped, non-redundant**		
Mock virus-virus	1,058		
Mock virus-human	284		
Mock human-human	381,631,769		
Virus virus-virus	893,070		
Virus virus-human	228,413		
Virus human-human	465,546,477		

4.4. References

Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., and Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature *485*, 376-380.

Jin, F., Li, Y., Dixon, J.R., Selvaraj, S., Ye, Z., Lee, A.Y., Yen, C.A., Schmitt, A.D., Espinoza, C.A., and Ren, B. (2013). A high-resolution map of the three-dimensional chromatin interactome in human cells. Nature *503*, 290-294.

Knipe, D.M., Lieberman, P.M., Jung, J.U., McBride, A.A., Morris, K.V., Ott, M., Margolis, D., Nieto, A., Nevels, M., Parks, R.J., and Kristie, T.M. (2013). Snapshots: chromatin control of viral infection. Virology *435*, 141-156.

Lieberman-Aiden, E., van Berkum, N.L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B.R., Sabo, P.J., Dorschner, M.O., Sandstrom, R., Bernstein, B., Bender, M.A., Groudine, M., Gnirke, A., Stamatoyannopoulos, J., Mirny, L.A., Lander, E.S., and Dekker, J. (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science *326*, 289-293.

McBride, A.A. (2008). Replication and partitioning of papillomavirus genomes. Adv Virus Res *72*, 155-205.

Chapter 5: Overall Summary and Perspectives

The studies presented here have contributed to our understanding of interactions between human cytomegalovirus (HCMV) and the host cell at a genomics level. My work has primarily focused on themes in RNA biology within the infection context, and a priority has been maintained to address both the host and virus profiles throughout each layer of analysis.

5.1. Small RNA profiling of HCMV infection

In the first stages of this work, I completed a next-generation sequencing characterization of the small RNA pool (smRNA-seq) in HCMV-infected fibroblasts. In this analysis I found evidence for two new HCMV miRNAs that had not been previously identified. The expression of these miRNAs was confirmed through Northern analysis, and I also confirmed their ability to regulate synthetic targets in cells. Furthermore, I used Argonaute CLIP-seq experiments to demonstrate that all HCMV-encoded miRNAs could be found within host silencing complexes during infection. In analysis of host miRNA levels, I found a particular miRNA cluster to be highly induced during infection. This cluster includes miR-96, miR-182, and miR-183, and expression of these miRNAs is not detectable in fibroblasts prior to infection. Finally, I also detected small RNAs across HCMV non-coding RNA regions, which represent an interesting and distinct class of non-miRNA virus-derived small RNAs.

5.2. Implications for upregulated host miRNAs in neurological contexts

Following the initial small RNA studies in fibroblasts, I have also conducted preliminary analysis of HCMV-infected neural progenitor cells (NPCs) and observed that the miR-182/-96/-183 cluster upregulation is also conserved in these cells (unpublished results from smRNA-seq data and Northern blot analysis). Further experimentation and

analysis will be needed to determine functional importance of this miRNA upregulation in neural precursors, but the consistent effect in multiple cell types is interesting. Among the potential roles that have been assigned to miRNAs 96, 182, and 183, importance has been suggested for these miRNAs in early development of the inner ear (Sacheli et al., 2009), which could be relevant to symptomatic HCMV infections in newborns. In this context, infection-altered miRNA levels in neural lineages could be therefore especially important. A recent publication has proposed that the miR-96 family is important for neural differentiation, and target genes included neural identity factors such as PAX6 (Du et al., 2013). From this point, another interesting connection to neural cell types includes the fact that in HCMV-infected NPCs a general downregulation occurs for genes related to neural identity and differentiation status (Chapter 3 and (Luo et al., 2010)). While a complex network of events is likely responsible for the downregulation of this class of neural-related genes, the increased levels of miR-182/-96/-183 could in part contribute to some of these gene expression changes during infection.

5.3. Comparative transcriptome analysis of different infected cell types

Moving beyond the small RNA fraction of infected cells, the next major stage of this work involved analysis of the mRNA component of the host-virus transcriptome. RNA sequencing (RNA-seq) libraries were prepared from three different primary cell types, including human foreskin fibroblasts (HFFs), human aortic endothelial cells (ECs), and human embryonic stem (ES) cell-derived NPCs. In addition to traditional gene expression measurements, this work included in-depth analysis of RNA processing events, including both splicing (splicing-sensitive microarrays) and polyadenylation patterns (RNA-seq analysis).

For analysis specific to the viral transcriptome, I found that the NPCs harbored lower levels of HCMV late genes, which was consistent with a suppressed infection

program in these cells in comparison with the HFFs and ECs. However, in terms of the splicing and polyadenylation of HCMV mRNAs, I did not observe significant differences between the cell types. For the infected host transcriptome, consistent across all the cell types, I uncovered widespread effects on the RNA processing landscape. These changes included thousands of alternative splicing and polyadenylation events, and the altered polyadenylation patterns were interesting in that they primarily resulted in shorter 3′-untranslated regions (3′UTRs). Potentially contributing to the observed 3′UTR shifts, I identified significant upregulation of a particular RNA binding protein in all the cell types, cytoplasmic polyadenylation element binding protein 1 (CPEB1). Finally, I also extended the analysis to HFFs infected with another virus from the herpesvirus family, herpes simplex virus 2 (HSV-2), and found similar global effects on polyadenylation patterns in these cells.

Changes in alternative host polyadenylation patterns and 3′UTR shortening have also been documented for B cells infected with the herpesvirus Epstein-Barr virus (EBV) (Homa et al., 2013). Infection with Kaposi's sarcoma-associated herpesvirus (KSHV) has also been shown to affect host polyadenylation. However, in this case the type of effect is distinct in that hyperadenylation of poly(A) tails occurs for many transcripts during infection (Lee and Glaunsinger, 2009). Combined with my observations related to HCMV and HSV-2, these findings illustrate an interesting, general convergence of multiple herpesvirus family members at the level of 3′-end regulation. The molecular consequences of these events could be very important for infected cells, and further studies will need to be conducted to determine whether the transcripts affected are impacted in terms of their stability or translational status.

Particular herpesviruses such as KSHV and HSV encode factors that are responsible for RNA-level effects such as the hyperadenylation phenomenon. Since

such factors are not encoded by HCMV, the search for potentially responsible host factors was important, and the induction of CPEB1 during HCMV infection may in part explain the polyadenylation events that I have uncovered. The CPEB1 upregulation could also be of more general interest related to the neurological pathologies associated with symptomatic HCMV infections. CPEB1 expression has been shown to be important in neurological contexts (Richter, 2007) and the levels induced by HCMV infection could be problematic in early cellular differentiation events or maintenance of neural progenitor populations.

5.4. Higher-order structure of host and viral DNA within infected cells

In the final work presented here, I have initiated a preliminary study of the three-dimensional conformation of DNA in HCMV-infected cells. This experiment has utilized an intensive approach known as Hi-C, a high-throughput sequencing version of chromosome conformation capture (3C). In analysis of 8 h-infected HFFs I have observed ligation products that suggest a general interaction between HCMV DNA and host chromosomes. Further normalization techniques must still be implemented, and further analysis is underway to determine whether host DNA conformational changes are evident following infection. The initial results from this pilot study have been encouraging, and future results from the host DNA conformation analysis could be integrated with the above-described RNA datasets to assign particular long-range interactions (or theoretically, changes in these interactions) with gene expression events that arise during HCMV infection.

5.5. References

Du, Z.W., Ma, L.X., Phillips, C., and Zhang, S.C. (2013). miR-200 and miR-96 families repress neural induction from human embryonic stem cells. Development *140*, 2611-2618.

Homa, N.J., Salinas, R., Forte, E., Robinson, T.J., Garcia-Blanco, M.A., and Luftig, M.A. (2013). Epstein-Barr virus induces global changes in cellular mRNA isoform usage that are important for the maintenance of latency. J Virol *87*, 12291-12301.

Lee, Y.J., and Glaunsinger, B.A. (2009). Aberrant herpesvirus-induced polyadenylation correlates with cellular messenger RNA destruction. PLoS Biol *7*, e1000107.

Luo, M.H., Hannemann, H., Kulkarni, A.S., Schwartz, P.H., O'Dowd, J.M., and Fortunato, E.A. (2010). Human cytomegalovirus infection causes premature and abnormal differentiation of human neural progenitor cells. J Virol *84*, 3528-3541.

Richter, J.D. (2007). CPEB: a life in translation. Trends in biochemical sciences *32*, 279-285.

Sacheli, R., Nguyen, L., Borgs, L., Vandenbosch, R., Bodson, M., Lefebvre, P., and Malgrange, B. (2009). Expression patterns of miR-96, miR-182 and miR-183 in the development inner ear. Gene Expr Patterns *9*, 364-370.