Investigation of the deregulated miRNome identified during acute viral infections in a murine model of HSV-1 encephalitis

By

Kyle A. Caligiuri

A Thesis
Submitted to the Faculty of Graduate Studies of
The University of Manitoba
In Partial Fulfillment of the Requirements for the Degree of

MASTER OF SCIENCE

Department of Medical Microbiology
University of Manitoba
Winnipeg, Manitoba

© Kyle Caligiuri, October 2014

ABSTRACT

Herpes simplex virus type 1 (HSV-1) is a double stranded DNA virus that causes epithelial skin infections and persists through the life of the host by infecting neurons, where it can switch to a latent state to evade an immune response. In rare cases during primary infection or after reactivation, instead of undergoing lytic infection at the epithelial surface, it instead travels to the brain and causes herpes simplex virus encephalitis (HSVE) which can have a ≥70% mortality rate if untreated. As the virus takes over its host cell, it gains control of the host cell machinery and manipulates host gene expression in order to evade the immune system and to pool its resources into the replication of the virus. One aspect of the dysregulated gene expression involves microRNAs (miRNAs). MiRNAs are short, non-coding RNAs that bind to the 3' untranslated region (3'UTR) of messenger RNAs (mRNAs), leading to translational repression of the target. Dysregulated miRNAs are often down-regulated during infection as the virus takes over, but many miRNAs have also been found to be up-regulated as well[1–5]. The aim of this study is to observe the full cellular miRNA changes in the context of an acute viral encephalitic infection using HSV-1, and to further characterize selected up-regulated miRNAs to determine their function in the context of the disease state. Of particular note were miR-141 and miR-200c which showed anti-apoptotic effects on neuronal cell culture and did not impact cell viability during an over-expression of the miRNAs. MiR-141, miR-183 and miR-200a expression was enriched within specific areas of the brain during infection. In addition, the potential for miR-150 to bind to a bioinformatically predicted target site within the shared 3'UTR of the HSV-1 U_L18, U_L19 and U_L20 genes was explored. Examining the changes in expression of this class of regulatory RNAs and investigating their potential functions may yield new insight into the relationship between host and virus during infection.

ACKNOWLEDGEMENTS

I would first like to thank my supervisor, Dr. Stephanie Booth, for giving me the opportunity to pursue a Master's degree. I would also like to thank her for hiring me for 3 co-op terms before that, which gave me the experience and direction I needed to solidify my choice of continuing in science. I would also like to thank my committee members, Dr. Xiaojian Yao and Dr. Sabine Hombach-Klonisch for their help and support throughout my degree.

Next I would like to thank the members of the Molecular PathoBiology department at the National Microbiology Laboratory. I would especially like to thank Yulian Niu for her histology work, which contributed to this thesis. Special thanks as well to Shane Thiessen from the Bioinformatics department, for assisting with the local Galaxy setup and maintenance. I would also like to thank my fellow graduate students, Anna Majer, Amrit Boese and Shantel Gushue, for their support, expertise and encouragement. Extra gratitude is extended to Anna for her consistent help on many of my projects and for her work on the animal inoculations from which my project is based.

I would also like to thank Dr. Dirk Motzkus for his help with transfection calculations, Sergio Pereira for completing the next generation sequencing, James Huang and Qi Zhu for completing the ACGT101-miR analysis of the sequencing data and Hugh Arnold for completing the GeneSifter analysis. Thanks also to Malachi Griffith and Adam Majer for helping with the Unix commands and programs, respectively, as part of the NGS workflow.

I gratefully acknowledge the University of Manitoba, the Public Health Agency of Canada and the Biotechnology Initiative of Government Laboratories for their financial support.

Finally, I would like to thank Candita Kyca for accommodating my work schedule while finishing my thesis, Richard Vieweg for helping me with image formatting, and my friends and family for their support and encouragement throughout my studies.

TABLE OF CONTENTS

ABBREVIATIONS 7
LIST OF FIGURES 11
LIST OF TABLES 12
1.0 INTRODUCTION 14
1.1 HERPES SIMPLEX VIRUS TYPE 1 (HSV-1) 14
1.1.1 HSV-1 LIFE CYCLE 15
1.1.2 HSV-1 DISEASES 20
1.2 HOST IMMUNE RESPONSES 24
1.2.1 INNATE IMMUNE RESPONSE 24
1.2.2 ACQUIRED IMMUNE RESPONSE 27
1.2.3 CENTRAL NERVOUS SYSTEM (CNS) IMMUNE RESPONSE 28
1.3 MicroRNAs (miRNAs) 29
1.3.1 miRNA BIOGENESIS 30
1.3.2 HSV-1 VIRAL-ENCODED miRNAs 32
1.3.3 THE ROLE OF HOST miRNAs DURING INFECTION 33
1.4 RATIONALE, HYPOTHESIS AND PROJECT OBJECTIVES 36
2.0 MATERIALS AND METHODS 39
2.1 TISSUE CULTURE CELL LINES AND VIRUSES 39
2.1.1 PREPARATION AND TITERING OF VIRAL STOCKS 39
2.2 ANIMAL EXPERIMENTS 40
2.2.1 INTRACEREBRAL INOCULATION 41
2.2.2 RNA EXTRACTION 42
2.2.3 SEQUENCE ANALYSIS 43
2.2.4 QUANTITATIVE PCR (qPCR) VERIFICATION OF miRNA EXPRESSION PROFILE 46
2.3 PRIMERS AND PLASMID DESIGN 48
2.4 LUCIFERASE REPORTER ASSAY 51
2.5 APOPTOSIS MEDIATION WITH SELECTED miRNAS 52
2.6 CELL VIABILITY ASSAYS 53
2.8 IN SITU HYBRIDIZATION (ISH) AND FLUORESCENT LABELING OF PARAFFIN-EMBEDDED BRAIN SECTIONS 54
3.0 RESULTS 58
3.1 IN VIVO MIRNA PROFILE DURING HSVE IN MICE 58

3.1.1 INTRACEREBRAL INOCULATION OF MICE WITH HSV-1 58
3.1.2 RNA PREPARATION 60
3.1.3 ANALYSIS OF HIGH THROUGHPUT SEQUENCING DATA 60
3.1.4 qPCR VERIFICATION OF *IN VIVO* NGS miRNA ANALYSIS 105
3.2 MiR-150 SHOWS WEAK BINDING POTENTIAL FOR TARGETING THE U_L18/19/20 3' UTR. 111
3.2.1 TARGET PREDICTION 111
3.2.2 VALIDATION OF TARGET BINDING VIA LUCIFERASE REPORTER ASSAY 115
3.4 TARGET PREDICTION FOR CANDIDATE miRNAS 117
3.5 APOPTOSIS ASSAY REVEALS ANTI-APOPTOTIC EFFECT OF miR-141 AND miR-200c 120
3.6 CANDIDATE miRNAS DO NOT AFFECT CELL VIABILITY OF NEURONAL CELL CULTURE IN VITRO 126
3.7 ISH REVEALS GLOBAL UP-REGULATION OF CANDIDATE miRNAs IN THE BRAIN 129
4.0 DISCUSSION 131
4.1 NGS IDENTIFIES HUNDREDS OF KNOWN AND NOVEL miRNAs WITHIN MOCK- AND HSV-1-INFECTED MOUSE BRAINS 131
4.1.1 COMPARISON OF miRNA PROFILE ANALYSIS TECHNIQUES 132
4.1.2 CANDIDATE miRNAs ARE INVOLVED IN OTHER VIRAL INFECTIONS 135
4.1.3 OTHER DEREGULATED miRNAs MAY HAVE ROLES DURING HSV-1 INFECTIONS 139
4.3 APOPTOSIS MEDIATION BY miR-141 AND miR-200c 142
4.4 miR-150 AS A POTENTIAL REGULATOR OF HSV-1 REPLICATION 143
4.4.1 HOST miRNAs HAVE THE POTENTIAL TO BE USED THERAPEUTICALLY 145
4.5 MiRNA EXPRESSION INCREASED IN SPECIFIC AREAS DURING INFECTION 147
4.6 FUTURE DIRECTIONS 147
REFERENCES 149
APPENDIX 160
1.0 DEEP SEQUENCING ANALYSIS PUBLISHED CHAPTER 160

ABBREVIATIONS

µg	micrograms
µL	microlitres
µM	micromolar
ADP	adenosine diphosphate
AP	alkaline phosphatase
APC	antigen-presenting cell
ARN	acute retinal necrosis
ATCC	American Type Culture Collection
ATP	adenosine triphosphate
ATP5B	ATP synthase subunit 5B
BBB	blood-brain barrier
BCIP	5-bromo-4-chloro-3-indolyl-phosphate
BD	Becton-Dickinson
Bid	B cell lymphoma 2 interacting domain
bp	base pairs
BSA	bovine serum albumin
CCL	C-C motif ligand
$CD8^+$	cluster of differentiation 8-expressing
cDNA	complementary DNA
CMV	cytomegalovirus
CNS	central nervous system
CO_2	carbon dioxide
CREB	cyclic adenosine monophosphate-responsive element-binding protein
CSCHAH	Canadian Science Centre for Human and Animal Health
CSCHAH-ACC	CSCHAH - Animal Care Committee
CSF	cerebrospinal fluid
CXCL	C-X-C motif ligand
DAB	diaminobenzidine
DC	dendritic cell
DCD	Dead Cell Discriminator
DEPC	diethyl pyrocarbonate
DFA	direct fluorescence assay
DMEM	Dulbecco's Modified Eagle's Medium
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleotide triphosphate
DPBS	Dulbecco's Phosphate Buffered Saline
dpi	days post-infection
E	early
EBV	Epstein-Barr virus
EDTA	ethylenediaminetetraacetic acid
EEEV	eastern equine encephalitis virus
eIF4E	eukaryotic translation initiation factor 4E

ELISA	enzyme-linked immunosorbent assay
EMEM	Eagle's Minimum Essential Medium
EtBr	ethidium bromide
EtOH	ethanol
FAM	fluorescein amidite
FAP-1	first apoptosis signal-associated phosphatase 1
FBS	fetal bovine serum
Fc	fragment, crystallizable
FFPE	formalin-fixed, paraffin-embedded
FGF	fibroblast growth factor
FGFR	FGF receptor
FSC	forward scatter
FTP	file transfer protocol
GAII	Genome Analyzer II
GENSAT	Gene Expression Nervous System Atlas
H_2O_2	hydrogen peroxide
HCMV	human cytomegalovirus
HCV	hepatitis C virus
HHV	human herpesvirus
HIV	human immunodeficiency virus
hpi	hours post-infection
HSV-1	herpes simplex virus type 1
HSV-2	herpes simplex virus type 2
HSVE	herpes simplex virus encephalitis
ICTV	International Committee on Taxonomy of Viruses
IE	immediate early
IFN	interferon
IHC	immunohistochemistry
IKKε	inhibitor of NF-κB kinase, subunit ε
IL	interleukin
iNOS	inducible nitric oxide synthase
ISH	*in situ* hybridization
JEV	Japanese encephalitis virus
kbp	kilobase pairs
KEGG	Kyoto Encyclopedia of Genes and Genomes
KSHV	Kaposi's sarcoma-associated herpesvirus
L	late
LAT	latency-associated transcript
LB	Luria-Bertani
LCM	laser capture microdissection
MAPK	mitogen-activated protein kinase
MAQ	Mapping and Assembly with Quality
MCS	multiple cloning site
MHC	major histocompatibility complex

MKNK1	MAPK-interacting kinase 1
miRNA	microRNA
mL	milliliters
MOI	multiplicity of infection
MRI	magnetic resonance imaging
mRNA	messenger RNA
NaCl	sodium chloride
NBT	nitro blue tetrazolium
NEB	New England Biolabs
NF-κB	nuclear factor κB
ng	nanograms
NGS	next generation sequencing
NK	natural killer cells
NS	not significant
NTF3	neurotrophic factor 3
P	processing
PBS	phosphate buffered saline
PBST	phosphate buffered saline with Tween 20
PCR	polymerase chain reaction
PEI	polyethylenimine
PFU	plaque-forming units
pmol	picomoles
pre-miR	pre-miRNA
pre-miRNA	precursor miRNA
pri-miRNA	primary miRNA transcript
PRKCA	protein kinase C-α
qPCR	quantitative PCR
RER	rough endoplasmic reticulum
RISC	RNA-induced silencing complex
RLU	relative light units
RNA	ribonucleic acid
ROS	reactive oxygen species
RPM	reads per million
RR	ribosomal reductase
RRR	relative response ratio
RT-PCR	reverse transcriptase-PCR
sed	stream editor
SHIP1	Src homology 2-containing inositol phosphatase 1
SIP1	Smad interacting protein 1
siRNA	small interfering RNA
SIV	simian immunodeficiency virus
SIVE	SIV encephalitis
snoRNA	small nucleolar RNA
SOCS1	suppressor of cytokine signaling 1

SSA	single strand annealing
SSC	saline-sodium citrate
SSC	side scatter
TAE	Tris-acetate-EDTA
TBS	Tris buffered saline
TBST	Tris buffered saline with Tween 20
TE	Tris EDTA
TGF	transforming growth factor
TLDA	TaqMan Low Density Array
TNF	tumor necrosis factor
TNFAIP3	TNF-α-induced protein 3
TRBP	trans-activation response RNA binding protein
Tris	Tris(hydroxymethyl)aminomethane
TrkB	tyrosine kinase receptor B
U	units
UTR	untranslated region
U_L	unique long region
VEEV	Venezuelan equine encephalitis virus
VWR	Van Waters & Rogers
VZV	varicella-zoster virus
WNV	West Nile virus
ZEB	zinc finger E-box-binding homeobox

LIST OF FIGURES

Figure 1. The HSV-1 life cycle. HSV-1 enters the cell by fusion of the cell membrane with its envelope (1) 16
Figure 2. The latency-associated life cycle of HSV-1 19
Figure 3. The immune response to HSV-1 infection 25
Figure 4. The miRNA biogenesis pathway 31
Figure 5. The workflow for secondary analysis of NGS data as developed in-house 44
Figure 6. Summary of cloning procedure for U_L18/19/20 pmirGLO Vector 49
Figure 7. Histology and immunohistochemistry (IHC) staining of HSV-1 infected and PBS mock-infected mouse brains 59
Figure 8. Overview of the deep sequencing process 62
Figure 9. Flow diagrams of the three different methods used for secondary analysis 63
Figure 10. Various comparisons between the 3 bioinformatic analyses used to identify known and novel miRNAs from PBS mock-infected and HSV-1 infected samples 99
Figure 11. Venn diagrams comparing the identification of known miRNAs by 3 different analysis programs 104
Figure 12. Differential expression between mock-infection and HSV-1 infection in mouse brains 106
Figure 13. Comparison of the sequences of the miR-200 miRNA family 110
Figure 14. QPCR validation of differential expression for select miRNAs 112
Figure 15. RNA22[102] output showing the target prediction for miR-150 within the HSV-1 genome 114
Figure 16 116
Figure 17. Partial MAPK signaling pathway by KEGG 121
Figure 18. Apoptosis pathway by KEGG 124
Figure 19. ADP/ATP ratio assay demonstrates an anti-apoptotic role of miR-141 and miR-200c 125
Figure 20. Cell counts via Trypan blue staining report little change when transfected with miR-141 or miR-150 127
Figure 21. ADP/ATP ratio assay reports no significant changes to cell viability when transfecting miR-141, miR-150 or miR-200c 128
Figure 22. An increased expression of selected miRNAs is observed in the HSV-1 infection compared to mock-infected samples 130

LIST OF TABLES

Table 1. Exact conditions for each probe during ISH procedure 56
Table 2. Quality and quantity of total RNA sent used for the generation of an in vivo miRNA profile of acute viral encephalitis caused by HSV-1 strain F..... 61
Table 3. Summary of characteristics for the various software used to predict miRNAs from deep sequencing data..... 64
Table 4. List of known mature miRNAs identified by miRanalyzer in the mock-infected sample 67
Table 5. List of known mature miRNAs identified by miRanalyzer in the HSV-1 infected sample70
Table 6. List of miRNAs identified with miRanalyzer by mapping to Rfam for both the mock-infected and HSV-1 infected sample..... 73
Table 7. Potential novel miRNAs found in the mock-infected sample using miRanalyzer..... 75
Table 8. Potential novel miRNAs found in the HSV-1 infected sample using miRanalyzer..... 77
Table 9. Differential expression results between the mock-infected sample and the HSV-1 infected sample using miRanalyze..... 80
Table 10. List of the top 50 known miRNAs present in the mock-infected sample as found by GeneSifter, and the comparative positions found in the HSV-1 sample..... 83
Table 11. Explanation of groupings by ACGT101-miR for known and novel miRNAs 87
Table 12. List of the top 50 known miRNAs found in the mock-infected sample by ACGT101-miR v3.5..... 88
Table 13. List of the top 50 known miRNAs found in the HSV-1 infected sample by ACGT101-miR v3.5..... 90
Table 14. List of the top 50 novel miRNAs found in the mock-infected sample by ACGT101-miR v3.5..... 92
Table 15. List of the top 50 novel miRNAs found in the HSV-1 infected sample by ACGT101-miR v3.5..... 94
Table 16. Differential expression results using ACGT101-miR between the mock-infected sample and the HSV-1 infected sample for known and novel miRNAs 96
Table 17. Comparison of the top 10 mature miRNAs as predicted by miRanalyzer, for both the mock-infected (A) and HSV-1 infected (B) samples..... 101
Table 18. Comparison of the top 10 mature miRNAs as predicted by ACGT101-miR, for both the mock-infected (A) and HSV-1 infected (B) samples..... 102
Table 19. Differential expression results between the HSV-1 infected sample and the mock-infected sample for known rodent miRNAs using TLDA technology..... 107
Table 20. Comparison of expression levels of select miRNA families as determined by NGS, TLDA and individual qPCR assays 113
Table 21. Pathways predicted by miRPath to be potentially regulated by at least one of 10 miRNAs (miR-141, miR-200c, miR-96, miR-182, miR-183, miR-200a, miR-200b, miR-429, miR-150, miR-155)..... 118
Table 22. Number of genes within the MAPK pathway predicted to be regulated by the candidate miRNAs..... 122
Table 23. Genes within the MAPK pathway that are predicted by microT to be regulated by miR-141 and miR-200c 123

Table 24. Dysregulated miRNAs identified during HSVE have been previously reported as dysregulated during other viral infections 136

1.0 INTRODUCTION

1.1 HERPES SIMPLEX VIRUS TYPE 1 (HSV-1)

HSV-1 is a ~152 kbp double-stranded DNA virus consisting of a linear genome[6] within a 162-capsomer, icosapentahedral capsid[7–9]. In a mature virion, the capsid is distributed within an envelope obtained by budding out of either the nucleus[7,9], rough endoplasmic reticulum (RER)[10,11] or outer membrane[11] of the infected cell. Before budding, approximately 20 different viral proteins are deposited in varying amounts in and around the membrane in order to be included within the mature virion. These proteins are collectively known as the tegument, and serve a variety of roles such as structural support, facilitating entry to a new host cell, or directing the capsid to the nucleus so as to release the genome[12].

HSV-1 is a member of the *Herpesviridae* family and is classified under the *alphaherpesvirinae* subfamily, along with two other human-infecting viruses: herpes simplex virus type 2 (HSV-2) and varicella-zoster virus (VZV)[7–9,13,14]. HSV-1 can also be categorized within another unofficial 'family' of human-infecting herpesviruses, of which there are 8[13]. This family is derived from the true taxonomical names for the viruses according to the International Committee on Taxonomy of Viruses (ICTV): human herpesvirus (HHV) 1-8[15] (of which HSV-1 is HHV-1). These herpesviruses have several characteristics in common: of note are the ability to create persistent infections throughout the life of the host by establishing latent infections to evade the immune system[7,14] and the ability of a given species to infect a wide variety of cell types[13]. In regards to the alphaherpesviruses, they are characteristically neurotropic, usually establishing a latent infection within neurons. Their lytic life cycle, however, is normally

established within epithelial cells near the initial site of infection, which results in a vesicular rash[7,13].

1.1.1 HSV-1 LIFE CYCLE

HSV-1 is generally spread through oral secretions and so the main area of infection is around the nose and mouth, usually at the mucocutaneous boundary[7–9,13,14]. In contrast, HSV-2 is usually spread by sexual contact or by perinatal transmission during child birth. HSV-2 infections are therefore normally within the anogenital area, but can be disseminated throughout the body of a newborn via an active vaginal infection during delivery. However, given the similarities between HSV-1 and HSV-2 – they are morphologically and structurally indistinguishable, and require antigen specificity, restriction endonuclease pattern analysis[7] or PCR analysis[8,9] to differentiate between the two viruses – and the prevalence of orogenital sexual contact, it is possible for both viruses to establish infection in any of the areas described above.

Infection first occurs at the epithelial surface, where the virus carries out a lytic life cycle, which causes the characteristic vesicular rash known as gingiostomatitis in the mouth or cold sores upon recurring infections, due to the reduced severity because of an acquired immune response[7]. As mentioned, HSV-1 can also present as genital herpes in the anogenital region. The lytic cycle is the phase of the virus life cycle that causes the most damage and therefore the disease symptoms are associated with this phase. During this time, HSV-1 is undergoing an active infection. A summary of the life cycle is presented in Figure 1.

The virus first enters the host cell by recognizing one of at least 5 known receptors. HSV-1 glycoproteins gC and gD are found inserted within its envelope and can bind heparin sulphate,

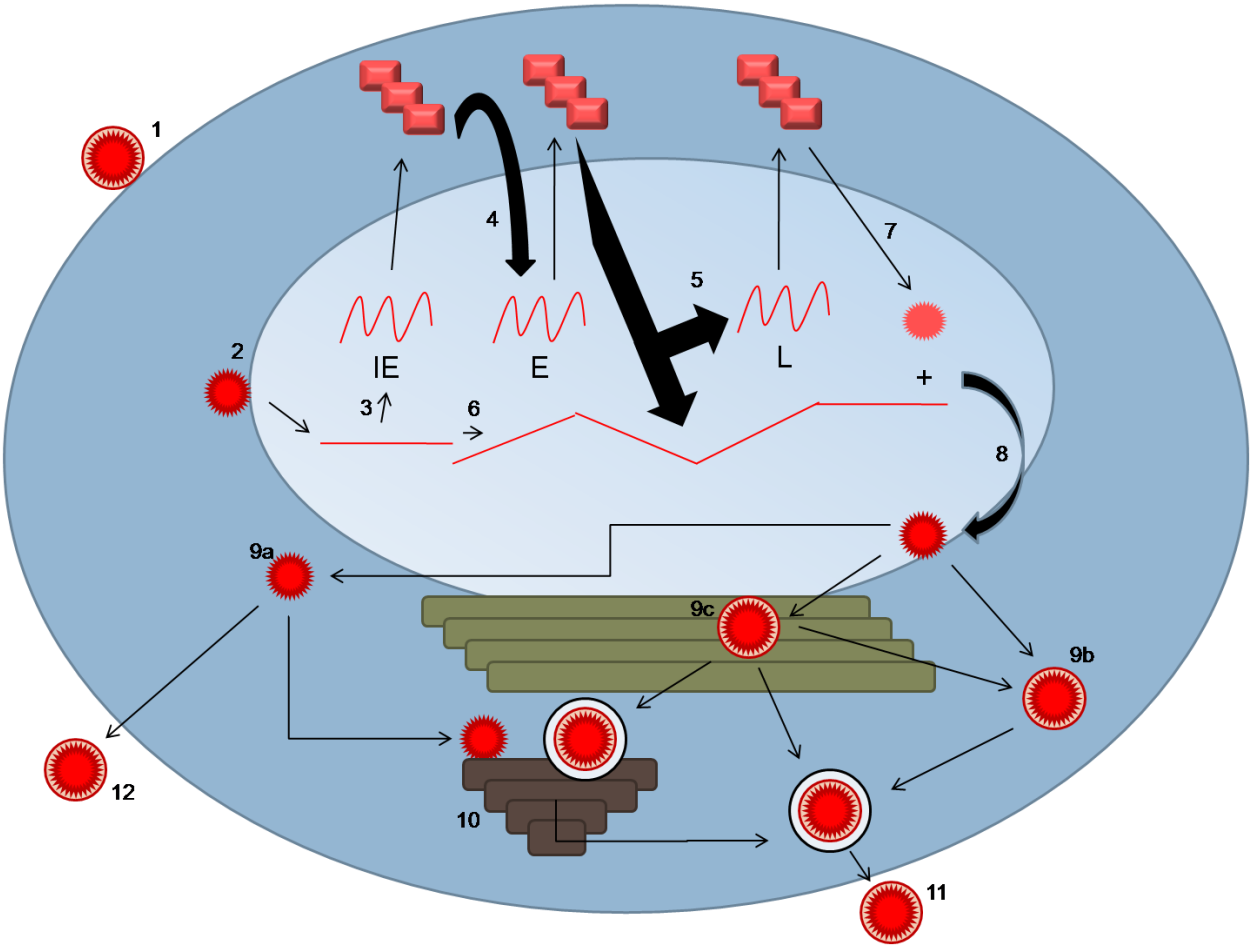


Figure 1. The HSV-1 life cycle. HSV-1 enters the cell by fusion of the cell membrane with its envelope (1). Once inside, the virus travels to the nucleus and injects its DNA genome inside (2). Transcription of IE genes is initiated by viral tegument proteins using host machinery (3). IE proteins initiate E phase transcription (4). E proteins continue the transcription cascade by initiating L transcription (5). E proteins are also involved in DNA replication (6) which proceeds by concatemerization. L proteins are structural proteins that are assembled in the nucleus (7). The DNA genome is inserted into assembled capsids (8) and cleaved at unit lengths. These virions then bud through the inner nuclear membrane, where they either bud out and lose their envelope (9a), are transported out to retain their envelope (9b) or they could move through the perinuclear space to the RER (9c). Non-enveloped, cytoplasmic virions can regain an envelope by budding through the Golgi, and virions leaving the RER are also transferred to the Golgi via vesicles (10). Enveloped, mature virions leave the cell by vacuolization (11) whereas remaining non-enveloped virions can leave the cell by budding through the cell membrane (12).

the Herpesvirus Entry Mediator (HVEM)[16,17], fibroblast growth factor receptor (FGFR)[7,18] nectin-1[16,17,19], or sometimes nectin-2[19], all found on the host cell membrane. The large variety of receptors are expressed by a multitude of different cell types, which gives the herpesviruses their pleiotropic trait. Glycoprotein gB then mediates membrane fusion to allow the capsid core into the cell[16,17]. Once in the cell, the DNA genome is injected through a nuclear pore, for which the process is mediated by tegument viral protein VP1/2. Next, the first phase of transcription can begin. This is known as the Immediate Early (IE) phase, and uses the host machinery to produce 5 proteins which mediate further transcription in the Early (E) phase[7,16]. The last phase of transcription is known as the Late (L) phase. Following the systematic transcription cascade, E transcription produces proteins which allow L transcription to proceed. However, IE and E transcription do not cease when the next phase starts; by the time L phase proceeds, transcripts from all phases may be actively transcribed and subsequently translated[16].

Once the HSV-1 genome is inserted into the nucleus, replication can begin. Replication requires viral proteins from E phase, whereas the majority of L phase transcripts produce structural proteins. Replication of the HSV-1 DNA occurs as a long concatemer of full-length genomes which are cleaved upon packaging[16].

Capsids are comprised mainly of the major capsid protein VP5, in association with minor capsid proteins VP19C, VP23 and VP26[16,17]. In addition, the U_L6 protein – named after its location in the genome within the unique long (U_L) region – is found at one vertex only, and acts as a pore to allow entry of the genome into the capsid. A complex of U_L15, U_L28 and U_L33 proteins work to mediate viral genome entry into the capsids and cleave the genome into unit lengths to

sever the entering genome from the concatemer. Finally, U_L31 and U_L33 proteins facilitate budding through the inner nuclear membrane. From here, the mechanisms are poorly understood, but the viruses are believed to be either transferred out of the perinuclear space by nuclear pores or tubules to retain the envelope[7,10], transferred to the Golgi via the RER[10], or bud out of the outer nuclear membrane and lose the envelope[16]. They could then bud through the Golgi apparatus[16], RER or the cell membrane[10,11] to regain an envelope for mature virion production. Mature, enveloped virions still within the Golgi or cytoplasm are thought to be released via vacuolization[7,10].

1.1.1.1 LATENCY

During the primary lytic infection at the epithelial surface, the virus will eventually find axons of neurons supporting the area. The virus can enter normally and travel up the axon to the cell body[8,9,16]. A summary of the latency life cycle of HSV-1 is presented in Figure 2. Within neurons, HSV-1 switches to a latent state, and the DNA genome circularizes within the nucleus. It isn't well understood why HSV-1 establishes latency within neurons but acts in a lytic fashion in epithelial cells. The current idea is that IE transcription fails, possibly due to a loss of the transcriptional activator VP16 from the tegument during the travel up the axon[16]. DNA replication may continue, as several hundred copies of the genome are often found within the cytoplasm of latently infected cells. However, translation of viral proteins is halted, and instead transcription of only a single latency-associated transcript (LAT) proceeds[8,9,16]. The LATs are poorly understood as well, but possible functions include establishment and/or maintenance of latency, or establishment of reactivation[16,17]. They have also been thought to be involved in

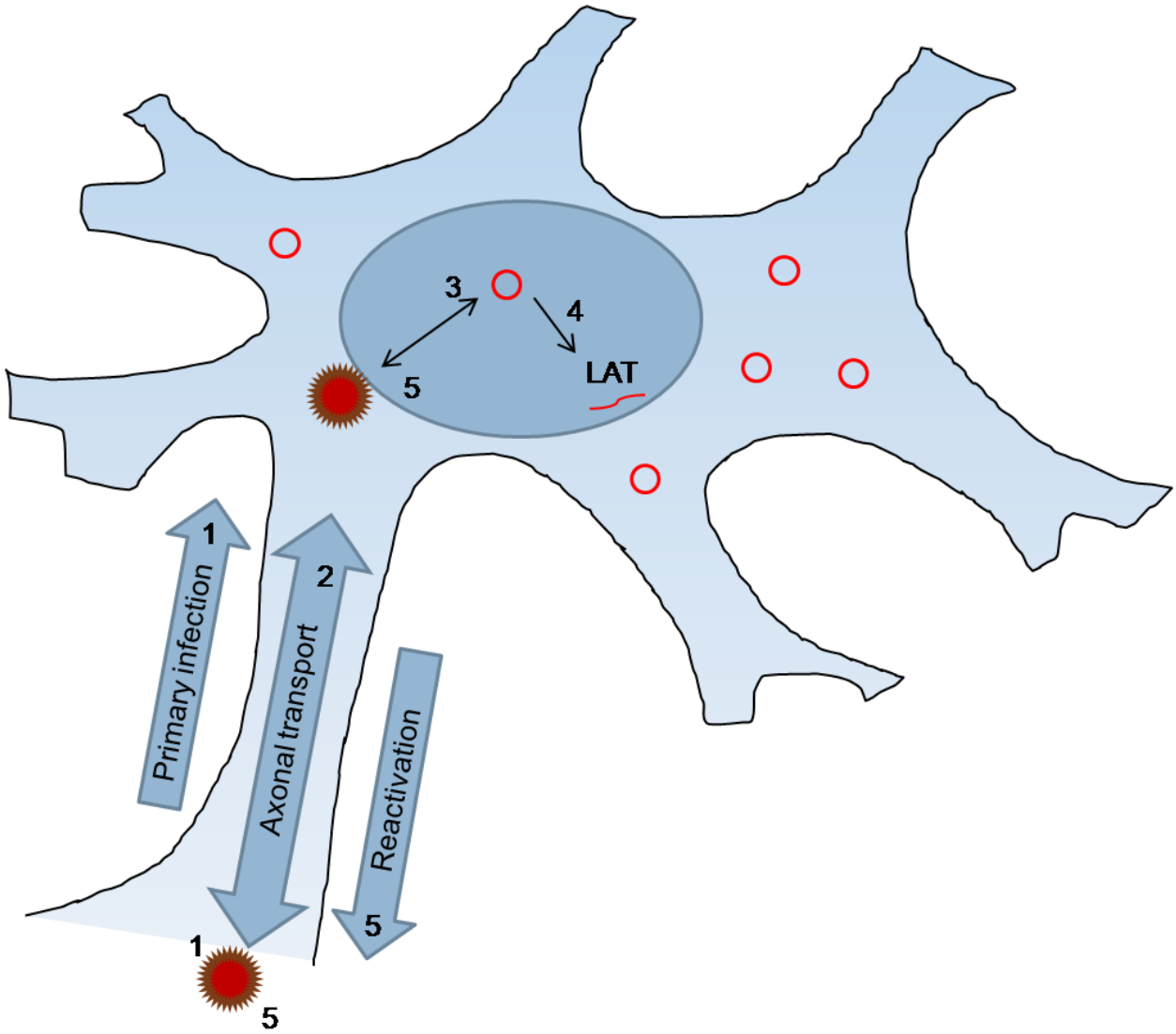


Figure 2. The latency-associated life cycle of HSV-1. The virus enters the cell normally during primary infection (1) and travels to the cell body via axonal transport mechanisms (2). The virus then injects the DNA genome into the nucleus where it circularizes into an episome (3). HSV-1 may still initiate the lytic cycle before establishing latency, or it may undergo latency immediately. The presence of viral genomes within the cytoplasm is an indicator that the virus has gone to E phase before switching to latency. Latency shuts down transcription of all genes except for the LAT (4). Latent virus may periodically reactivate (5) leading to normal formation of the virus. However, to avoid the immune response, virions are often transported back down the axon for release and productive infection at the initial site within the epithelium. Adapted from Schleiss (2009)[14].

immune evasion via anti-apoptotic functions[8,9,16,17], to allow prolonged infection throughout the life of the host.

Reactivation from a latent state still holds many mysteries as to why and how it happens. The trigger is often thought to be stress-induced[8,9], but there is also evidence of sub-clinical reactivation, allowing virus shedding asymptomatically[16,20–22]. It is possible that latency is not tightly maintained as was previously thought, and instead the virus may be in a constant flux where latency is a struggle to maintain and virus is periodically released. Regardless, when virus is released as part of a clinical manifestation of the lytic cycle, the originally-infected neurons are not lysed and destroyed[8,9,16]. Instead, virions travel down the axonal path to the original site of infection where they can then infect epithelial cells. A recurrence of the vesicular rash erupts due to epithelial infection, which also allows for spread of the virus to others during contact.

1.1.2 HSV-1 DISEASES

HSV-1 is able to infect a wide variety of cell types, and therefore different symptoms may arise depending on the area affected. The most common forms of disease are herpetic gingivostomatitis and cold sores[7]. Both are vesicular or ulcerative rashes around the mucocutaneous regions of the nose and mouth, and are differentiated based on the severity and the order of presence. Primary infection presents as gingivostomatitis and recurring infections present as less-severe cold sores due to an immune response. A similar disease is herpetic pharyngitis, when the rash presents at the back of the throat[14].

Other skin rashes caused by HSV-1 infection include herpetic whitlow, herpes gladiatorum, eczema herpeticum and genital herpes[7,14]. Herpetic whitlows usually occur on the fingers or

hands and the primary cause is due to laboratory infection when working with HSV-1. Herpes gladiatorum is often an athletically-acquired disease spread by close skin contact in sports such as amateur wrestling or rugby. Eczema herpeticum is a combination of herpesvirus infection within the area of atopic dermatitis or other skin breakages and is often observed on children. Finally, HSV-1 infection can present as genital herpes during orogenital sexual contact. However, genital herpes are more-commonly associated with HSV-2, the primary infectious agent responsible for the disease.

Besides skin infections, several serious diseases caused by HSV-1 are manifested when the virus infects the eye. This can occur through shedding of the virus to other areas of the body or, for neonates, can occur due to contact with an active infection site during delivery. Depending on the severity and location of the infection within the eye, disease can present as herpetic keratitis, retinitis or acute retinal necrosis (ARN), and can even lead to blindness[7,14,19]. Herpetic keratitis, or keratoconjunctivitis, is inflammation of the cornea caused by an immune response to HSV-1 infection. Similarly, retinitis is inflammation in the retina. HSV-1 is the second most common cause of ARN and usually occurs upon recurrent infections after an initial case of retinitis[19].

1.1.2.1 HERPES SIMPLEX VIRUS ENCEPHALITIS (HSVE)

HSVE is the most severe form of disease caused by HSV-1 and occurs when the virus reactivates in the trigeminal ganglia (the sensory ganglia of the trigeminal nerve, where the perikaryons for the sensory nerves of the face are located) and instead of travelling to the initial site of infection, travels to other neurons, eventually making its way into the brain. It is unclear why this happens[23], but occurs in <0.001% of the population[24]. HSV-1 is the major cause of acute

viral encephalitis (which accounts for 5-20% of all encephalitis cases) and mortality rates can reach 70-75% if untreated[9,14,24]. It is interesting that while it is estimated that 53-89% of the population is infected with HSV-1 (depending on age, country and other demographics)[25–27], only 10-25% of HSVE patients record a history of gingivostomatitis/cold sores, indicating that HSVE may also present as a primary infection[23,24]. Another scenario involved is that the initial infection was asymptomatic and went by unnoticed[23].

HSVE can cause a variety of symptoms, initially starting with malaise, fever, headache and nausea, and possibly including olfactory and/or gustatory hallucinations[7,9,14]. If left untreated, the disease can progress to obtundation, lethargy, confusion, delirium, aphasia or other focal deficits, hemiparesis, seizures or even death[9,14]. Depending on the severity of the infection and the speed of recovery, long-lasting brain defects may persist after recovery, including aphasia, personality changes, motor deficits, amnesic syndrome, global cognitive decline and epilepsy[7,23,24].

Pathologically, HSVE presents as asymetric localized necrosis within the brain[7,9,14,23,24]. The main affected areas in humans are the temporal lobes of the cerebral cortex. Depending on the severity, the damage can spread to the hippocampal complex, basal forebrain and cingulate cortex. Tissue damage is characteristic of a typical HSV-1 lytic infection: intranuclear inclusion bodies are observed along with perivascular cuffing and petechial hemorrhaging, followed by necrosis and liquefaction as disease progresses[9,23]. In addition, a glial response to the infection is observed within the brain, including astrogliosis, glial nodules and neuronophagia[23].

1.1.2.2 DIAGNOSIS AND TREATMENT

Skin lesions caused by HSV-1 can typically be diagnosed simply by observation by a clinician, but would be confirmed by either culturing a sample, a direct fluorescence antibody assay (DFA) or the Tzanck smear[7,13,28]. Standard viral culturing or centrifugation-enhanced culturing (for faster diagnosis) are the gold standards for confirming an HSV-1 infection. An immunofluorescent assay such as DFA or enzyme-linked immunosorbent assay (ELISA)[7] can also confirm viral presence as well as specify between herpesvirus types. DFA can be combined into a culture, whereas ELISA must be performed separately. The Tzanck smear and other histological staining techniques are relied on less-so in recent years as they are often subjective and would need to be combined with another test such as electron microscopy for a stronger diagnosis[28,29]. In regards to HSVE, diagnosis is often performed by virus DNA detection in the cerebrospinal fluid (CSF)[7,9,13,28]. The only true confirmation is by biopsy since mature virions cannot be isolated within the CSF, but the presence of HSV-1 DNA is a strong indicator and allows clinicians to pursue the use of HSV-1-specific antivirals. Magnetic resonance imaging (MRI) and/or electroencephalograms can also be used to determine the severity of the infection within the brain[7,9,23].

The principle antiviral drug used to combat HSV-1 infection is acyclovir[7,13], which inhibits viral but not cellular thymidine kinase and DNA polymerase[9]. Both viral proteins are needed for efficient replication of HSV-1 DNA. Acyclovir can be administered topically for skin infections, intravenously for more severe internal infections, or orally for both. Other stronger antivirals used include valacyclovir and famciclovir, which work faster and have simpler dosage requirements[13]. These drugs are usually favoured for treating genital herpes[7]. Ganciclovir,

foscarnet and cidofovir can be used to treat coinfections with HSV and cytomegalovirus (CMV)[13] or for cases of antiviral resistance[7]. Penciclovir or docosanol are milder topical agents that can be used for recurrent cold sores. For infection of the eye, trifluridane is applied topically, and for HSVE the standard treatment is intravenously-administered acyclovir. It is important to note that these treatments do not eradicate the virus, but simply shorten the duration of active infections and reduce the extent of shedding the virus[7]. Unfortunately there is currently no vaccine or drug that can target latently-infected cells, so the systemic infection is never truly cleared and recurrences are entirely possible.

1.2 HOST IMMUNE RESPONSES

HSV-1 triggers many of the classic immunological defense mechanisms upon infection of the host organism. The innate immune defense cells, such as macrophages, dendritic cells (DCs) and other antigen-presenting cells (APCs) are the first line of response to infection and help to recruit other immune cells while also working to limit viral spread via phagocytosis of dead cells and extracellular virions[16,25]. Other innate immune cells are recruited first, such as natural killer cells (NKs) and neutrophils, followed by the principle members of the acquired immune response, cluster of differentiation 8-expressing ($CD8^+$) T cells, which induce apoptosis of infected cells in an effort to reduce the viral load produced and therefore decrease the spread of infection. A summary of the innate and acquired responses are presented in Figure 3.

1.2.1 INNATE IMMUNE RESPONSE

Innate immunity is the first line of defense during an infection and can start with the infected cell itself. Before the virus even enters the cell, pattern recognition receptors (PRRs) can bind and detect gD to create a signal cascade that results in the expression of interferons (IFNs) and

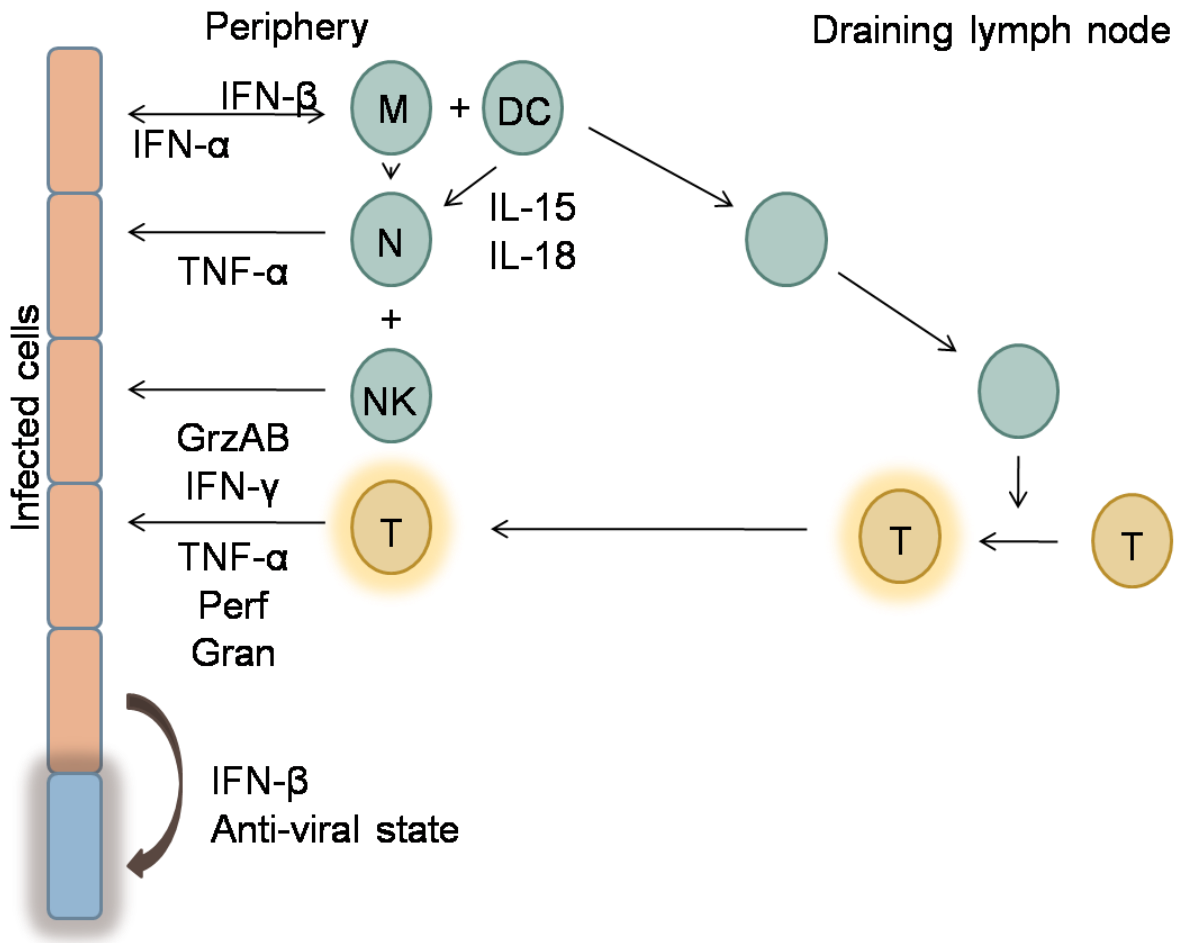


Figure 3. The immune response to HSV-1 infection. Infected cells recognize PAMPs and respond via the release of cytokines such as IFN-β. IFN-β activates innate immune cells such as macrophages (M) and DCs, and they respond by releasing their own IFNs, such as IFN-α to induce an antiviral state. They also release cytokines such as IL-15 and IL-18 to recruit neutrophils (N) and NKs. Ns and NKs help clear a viral infection by inducing apoptosis of infected cells via TNF-α, granzymes A and B (GrzAB) and IFN-γ. In addition, Ms and DCs can travel to the nearest draining lymph node to present antigen which activates $CD8^+$ T cells. These T cells can return to the infection to help induce apoptosis and phagocytose extracellular virus. They induce apoptosis with GrzAB, IFN-γ, TNF-α, perforin (Perf) and granulysin (Gran). Finally, infected cells that recognize PAMPs can also release cytokines to neighbouring uninfected cells to induce an antiviral state so that they are not as easily infected. Adapted from Egan *et al.* (2013)[25].

other cytokines[25]. These molecular signals are released from the cell to warn neighbouring cells of infection and to induce an antiviral state. Once the virus enters the cell, Toll-like receptors (TLR) -3 and -9 can detect several viral signatures known as pathogen-associated molecular patterns (PAMPs). These PAMPs are associated with DNA replication. TLR-3 can recognize replication intermediates in the form of double-stranded RNA (dsRNA), which occurs erroneously due to bidirectional transcription of two polymerases within the same region[25,30]. TLR-9 can recognize CpG motifs within the DNA genome that are characteristic of microbial but not vertebrate genome structures. Recognition of PAMPs by TLRs is another way of inducing interferon and cytokine expression.

Outside the infected cell, macrophages and DCs are activated by the release of IFNs and cytokines, which are then stimulated to secrete their own IFNs and cytokines[25]. This increased production helps to recruit other innate immune cells such as neutrophils and NKs. Neutrophils act by inducing apoptosis via secretion of tumor necrosis factor (TNF) -α to infected cells and subsequently phagocytosing dead or dying cells. NKs also induce apoptosis by releasing granzymes A and B. They also release IFN-γ which helps induce the antiviral state of surrounding cells by promoting antigen presentation and inhibiting viral replication. Replication is inhibited by stimulating the expression of genes such as protein kinase RNA-activated (PKR) and ribonuclease L (RNase L). PKR phosphorylates eukaryotic initiation factor (EIF) -2α which inhibits protein translation of viral transcripts. Viral messenger RNA (mRNA) transcripts then accumulate, and RNase L actively degrades them[25].

HSV-1 has a few mechanisms to reduce the antiviral response. Infected cell polypeptide (ICP) 0 is an IE protein produced primarily as a promoter for E transcription[16,17]. However, it also helps

to inhibit IFN regulatory factor (IRF) 3, which is a promoter for IFN production[31]. ICP0 works in conjunction with the viral host shutoff factor (vhs), an E protein, to further reduce the antiviral response by indiscriminately degrading host mRNAs. Lastly, the $U_L34.5$ protein inhibits the action of PKR to proceed with protein translation and subsequent replication[16,17].

1.2.2 ACQUIRED IMMUNE RESPONSE

The aquired immune response is created by two main cell types: T cells and B cells. T cells actively clear infection by inducing apoptosis of infected cells[25]. B cells are responsible for producing antibodies that bind and neutralize extracellular virions. However, in a primary HSV-1 infection B cells appear to work primarily in an innate fashion by presenting antigens and secreting cytokines. B cells switch to their normal acquired response role only after the infection is cleared and produce antibodies so that recurring infections are milder. T cells are recruited to the infection and activated by macrophages and DCs which travel to the nearest lymph node to present antigen[25].

$CD8^+$ T cells respond to infection by releasing IFN-γ and induce apoptosis with the release of TNF-α, perforin, granzymes A and B, and granulysin[25]. Perforin forms a pore on the infected cell's membrane, allowing entry of granzymes and granulysin. Granzyme A cleaves single-stranded DNA and activates interleukin (IL) -1β, an inflammatory cytokine, by cleaving the precursor. Granzyme B induces apoptosis in several different ways: it can activate procaspase 3, the precursor of caspase-3 which is an integral protein in the apoptosis pathway; it can disrupt the membrane potential of mitochondria by increasing its permeability; and it can activate the B cell lymphoma 2 interacting domain (Bid) protein, which also has an apoptotic function. Finally, granulysin can lyse cells by interacting with negatively-charged cell membrane

proteins, and can induce apoptosis by releasing cytochrome C from the mitochondria, which destabilizes the membrane potential[25].

HSV-1 has several proteins that can impede the acquired immune response. Glycoproteins gE and gI mimic crystallizable fragment (Fc) receptors which normally bind antibodies bound to virus to initiate endocytosis to help clear infection[16,17,32]. Instead, gE and gI can bind unbound antibodies for endocytosis to help clear the extracellular immune response. In addition, IE protein ICP47 inhibits antigen presentation by infected cells by disrupting the loading of antigen onto major histocompatibility complex (MHC) class I molecules[16,17]. This will inhibit T cells from recognizing an infected cell, and therefore prevents apoptosis.

1.2.3 CENTRAL NERVOUS SYSTEM (CNS) IMMUNE RESPONSE

The immune response to HSV-1 within the CNS relies on many of the same characteristics as the classic responses described above. However, because of the blood brain barrier, immune cells have a harder time reaching the infection than if the infection were at the epithelial surface[33]. The primary responder to infection in the brain is its own innate immune cell, the microglial cell. Microglia release many of the same cytokines to promote an antiviral response, such as TNF-α and IL-1β[34]. They also increase production of reactive oxygen species (ROS) and inducible nitric oxide synthase (iNOS) in response to infection, in order to induce oxidative stress to damage virions in the extracellular space. Microglia are responsible for promoting the expression of MHC class I and II glycoproteins as well, which are normally poorly expressed within the brain cell types[33].

An important role of the microglia is to recruit outside immune cells. Microglia secrete many chemokines to efficiently attract immune cells to the brain. Chemokines such as C-C motif

ligand (CCL) 2 and 3, as well as C-X-C motif ligand (CXCL) 9 and 10 are secreted to the blood and CSF in order to recruit granulocytes, macrophages and T cells[34,35]. However, an increased immune response can both help and hinder the host. With an increased response comes an increase in apoptosis of infected cells, which adds to the destruction caused by HSV-1. Oxidative stress molecules are non-specific so they will also damage host cells inadvertently. It is debated whether the viral infection or the immune response itself is the larger cause of tissue damage within the brain, but both are certainly to blame[35].

1.3 MicroRNAs (miRNAs)

MiRNAs are short (~22 nucleotides), non-coding, single-stranded RNAs that have reverse complementarity to mRNAs, particularly within the 3' untranslated region (UTR)[36]. MiRNAs are loaded into the RNA-induced silencing complex (RISC) which then loads mRNAs to bind them for translational repression and sequestration within processing (P) bodies[36,37]. While many mammalian miRNAs show only partial homology, allowing a smaller number of miRNAs to target a large range of transcripts, plant miRNAs tend to show perfect homology which results in cleavage of the transcripts. It is estimated that 30-50% of the vertebrate genome is influenced directly by miRNA regulation[38]. MiRNAs may play an even larger role due to feedback loops caused by the direct regulation of certain transcripts, causing downstream differential expression of other genes and thereby heavily impacting the gene expression of a given cell[39]. In addition, miRNAs could be used to affect neighbouring cells by secreting miRNAs extracellularly either on their own or enclosed within exosomes[40].

1.3.1 miRNA BIOGENESIS

MiRNA genes are found throughout the genome[36]. They can be isolated to their own transcription units, or can be found within introns or exons. If grouped into a transcription unit, these miRNAs often have similar targets, are highly conserved[41] and are phylogenetically classified into a miRNA family[42]. All family members are not necessarily contained within a single transcription unit, however. MiRNAs are particularly conserved between species within the seed sequence region (nucleotides 2-8), the main area of homology in which miRNAs bind their targets[36]. A summary of the biogenesis pathway of miRNAs is presented in Figure 4.

MiRNAs are first transcribed into a long stem-loop called a primary miRNA transcript (pri-miRNA) by RNA polymerase II or, less-commonly, RNA polymerase III[36,43]. If they are part of a family, different miRNAs may be connected to the transcript. They are then cleaved by the Drosha RNase III endonuclease, with help from Pasha, to separate the miRNAs into precursor miRNAs (pre-miRNAs or pre-miRs). Pre-miRs are transported outside the nucleus by Ran-GTP and Exportin-5, where they are processed by the Dicer RNase III endonuclease, with help from the trans-activation response RNA binding protein (TRBP), to cleave off the loop section. This leaves a dsRNA segment called a miRNA:miRNA* duplex, of which the reverse complement of the mature miRNA is known as the miRNA*[36]. Recently, however, it has been found that miRNA*s can bind and target mRNAs as well, leading to a change of annotation. MiRNAs are now simply known as -3' (3p) or -5' (5p) transcripts based on their location within the hairpin structure. The duplex associates with Argonaute proteins and GW182 to form the RISC and the dsRNA strands are separated by Argonaute[44]. One strand is left within the RISC to be used for transcriptional repression and the other is removed and subsequently degraded[43].

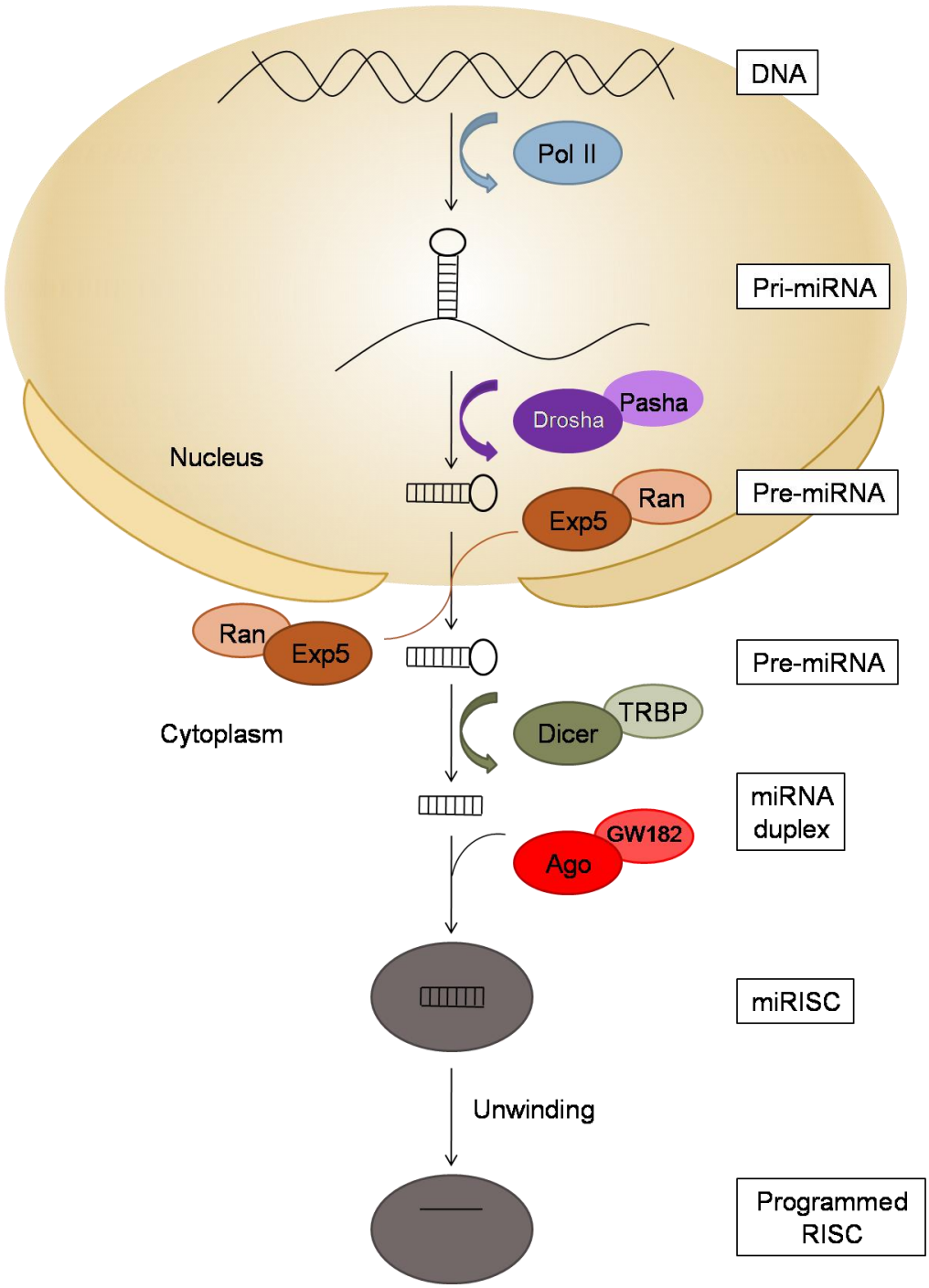


Figure 4. The miRNA biogenesis pathway. MiRNAs are first transcribed into long transcripts with hairpin loop segments (pri-miRNAs) by RNA polymerase II (Pol II). Next, Drosha and Pasha cleave the stem-loop segments off the transcript to form pre-miRNAs. These pre-miRNAs are transported out of the nucleus by exportin 5 (Exp5) and Ran-GTP (Ran). The loop segments are subsequently cleaved off by Dicer and TRBP to leave a miRNA duplex. This duplex is loaded into Argonaute proteins (Ago) and associates with GW182 to form the miRNA RISC (miRISC). Ago then unwinds the duplex to remove the unwanted strand, which leaves a programmed RISC ready to bind and repress mRNA target transcripts. Adapted from Cullen (2005)[45].

1.3.2 HSV-1 VIRAL-ENCODED miRNAs

In recent years it has been discovered that many viruses also encode miRNAs. To date, 21 mature miRNAs have been identified within the HSV-1 genome, of which 17 have been confirmed to be expressed during infection[46,47]. The majority of the miRNAs are contained within the LAT region and so their function is related to maintaining latency. Only 3 miRNAs, miR-H1, miR-H6-3p and miR-92944, have been identified as being expressed during the lytic life cycle of HSV-1[46]. Viral miRNAs are still being investigated, and so limited information is available. Of the lytically-expressed miRNAs, MiR-H6-3p is the only miRNA with a known target[46,48]. MiR-H6-3p targets and reduces expression of IE protein ICP4, one of the proteins responsible for driving transcription into E and then L phase. MiR-H6-3p is therefore related to latency, even though it is produced during the lytic cycle. It is thought that miR-H6 helps establish latency by halting translation of viral proteins from E and L phases, but it cannot initiate latency on its own. Additionally, miR-H6-3p has been linked to a reduction of IL-6 concentrations due to the suppression of viral replication, which would aid the virus in immune evasion during latency[46,48].

The miRNAs produced by cleavage of the LAT have been better characterized than the lytic-associated viral miRNAs. MiR-H2-3p targets ICP0 to hinder productive replication and miR-H4-3p targets U_L34.5, which would increase levels of host PKR and therefore reduce levels of replication[46]. MiR-H3 is also predicted to bind U_L34.5 transcripts, but to date has not been confirmed. Finally, miR-H4-5p helps protect the latent cell from apoptosis[46] by inhibiting host cytokine transforming growth factor (TGF) -β and the downstream effector SMAD3[47,49], which induces transcription of pro-apoptotic genes[50].

1.3.3 THE ROLE OF HOST miRNAs DURING INFECTION

Cellular miRNAs have increasingly been found to have the potential to impact the viral replication cycle, either by direct targeting of viral transcripts, or by indirect targeting of cellular transcripts needed by the virus for efficient replication[41,51]. These miRNAs may even be up-regulated as a response to infection; for example, IFN-α, -β and -γ increase expression of 8 miRNAs that are complementary to the hepatitis C virus (HCV) RNA genome[51,52]. However, viruses tend to co-evolve with their hosts[53] and so it is thought that they would undergo selective pressure to reduce transcripts that are complementary to host miRNAs[41,54]. They may have also adapted to miRNA repression tactics by decreasing expression of harmful miRNAs during infection, while selectively up-regulating miRNAs that are advantageous to a productive infection[54]. There are currently very few examples of host miRNAs directly targeting viral transcripts. Currently, only one miRNA, miR-138, has been implicated in affecting the HSV-1 life cycle directly. MiR-138 has cell-type specificity and is expressed mainly in neurons[55,56]. This host miRNA works in conjunction with the viral miRNA miR-H2-3p to inhibit ICP0 expression in order to establish and maintain latency. Other examples of host miRNAs directly targeting viral transcripts include liver-specific miR-122 which binds 2 locations within the 5'UTR of the RNA genome of HCV to produce a positive regulation of replication by unknown means[54,57–59], and miR-32 which limits primate foamy virus (PFV) replication *in vitro* by binding to the 3' region of its genome[54,60].

Host miRNAs impacting virus infection is a new and emerging area of research and so little is known regarding HSV-1 interactions with host miRNAs. It is difficult to study host miRNAs directly related to infection because complete knockout of a miRNA will have its own

implications on the cell regardless of viral infection. It is also challenging to determine if specific miRNA-related changes are truly caused by virus-mediated regulation or if the changes are a result of the host responding to an invading organism. However, recent studies have shown that mmu-miR-101a and mmu-miR-155 are important host miRNAs indirectly involved in HSV-1 infection. MiR-101a targets adenosine triphosphate synthase subunit 5B (ATP5B)[61] and has been shown to be down-regulated by HSV-1 during infection. Upon over-expression of miR-101a during infection, viral replication decreased significantly, indicating ATP5B was an essential cellular protein for productive replication of HSV-1. MiR-155 is known to increase inflammation caused by T cells and macrophages, and miR-155 knockout mice infected with HSV-1 were more susceptible to infection[62]. Ocular infections of HSV-1 in miR-155 knockout mice developed HSVE where wild-type mice did not, and latent reactivations occured more often in knockout mice as well. Another herpesvirus, human cytomegalovirus (HCMV), has been shown to be indirectly affected by host miRNAs as well. Hsa-miR-100 and hsa-miR-101 repress the mammalian target of rapamycin (mTOR) gene and hsa-miR-100 additionally represses raptor, involved in the same pathway[54,63]. The mTOR pathway is important for a productive HCMV infection and so the virus represses miR-100 and miR-101 during its lytic life cycle. Addition of miR-100 and miR-101 mimics reduced viral titres *in vitro* and showed increased inhibition in combination rather than individually.

MiR-155 is known to have inflammatory-related functions and has been implicated in other viral encephalitis diseases besides HSVE, such as those caused by Japanese encephalitis virus (JEV)[64], Venezuelan equine encephalitis virus (VEEV) and West Nile virus (WNV)[5]. In JEV infection, miR-155 was found up-regulated in brain samples of both human and mouse origin[64].

JEV infection of primary microglia and microglial cell culture were analyzed and miR-155 over-expression was observed as well. Thounaojam *et al.* (2014)[64] concluded that miR-155 regulates Src homology 2-containing inositol phosphatase 1 (SHIP1) during JEV infection of microglia. SHIP1 is a negative regulator of inflammation, so miR-155-mediated reduction of SHIP1 results in an increased immune response to infection. However, due to the damage caused by inflammation in the CNS, miR-155 knockout using complementary anti-miRs resulted in increased survival *in vivo*[64]. MiR-29b was also up-regulated during both *in vitro* and *ex vivo* infection of microglia with JEV[65]. MiR-29b was implicated in regulating tumor necrosis factor α-induced protein 3 (TNFAIP3), another negative regulator of inflammation. Similar to miR-155, knockout of miR-29b reduced the expression of inflammatory cytokines as well, which would suggest a similar increase of survivability. However, Thounaojam *et al.* (2014)[65] did not pursue viability studies using anti-miR-29b.

Another set of miRNAs found during viral encephalitis are from the miR-142 pre-miRNA. MiR-142-3p and miR-142-5p have been found to be over-expressed in rhesus monkey brains infected with simian immunodeficiency virus (SIV)[66]. MiR-142 miRNAs are normally expressed only in hematopoietic cell lines[67], but during SIV encephalitis (SIVE) both neurons and myeloid cells were expressing miR-142 *in vivo*[66]. MiR-142-5p was predicted and later confirmed to bind and inhibit sirtuin1, a neuroprotective protein involved in synaptic plasticity and a negative regulator of transcription. The increased inflammation observed during SIVE is thought to be caused in part by infection of immune cells and the subsequent repression of sirtuin1 by miR-142[66]. Conversely, miR-142-3p has been implicated in encephalitic conditions caused by eastern equine encephalitis virus (EEEV)[67]. MiR-142-3p is able to bind the 3'UTR of the EEEV

RNA genome to inhibit transcription and subsequent replication. It is thought that repression of viral transcription by this miRNA inhibits the formation of a proper innate immune response. This allows for the production of viremia and eventual access to the CNS, where it causes encephalitis[67].

Research of host miRNAs is an important area of focus, since they can be harnessed for therapeutics. Understanding the roles and differential expression patterns of miRNAs normally and under infectious conditions can yield insights into the pathogenesis of the virus, and potential uses for host miRNAs. For example, anti-miR-122 oligonucleotides have recently completed phase 2a trials in humans to determine their effectiveness at reducing or eliminating HCV infection[68]. Results showed a reduction in HCV RNA levels that persisted after treatment, and no adverse health effects caused by treatment.

1.4 RATIONALE, HYPOTHESIS AND PROJECT OBJECTIVES

The purpose of this project was to study the miRNAs involved in acute viral encephalitis, specifically using HSV-1 to induce HSVE in a murine model. To my knowledge, no other comprehensive study such as this has been described to date. To start, a miRNA profile (miRNome) was created using next generation sequencing (NGS) techniques. NGS technology was new at the start of this work and a significant objective was to develop a bioinformatic protocol for the analysis of miRNA NGS data. The miRNome comprises all miRNAs, both novel and known, detected within HSV-1 and mock-infected samples. Additionally, differential expression can be analyzed to determine the extent of dysregulation of the miRNAs during infection. Based on these results, the majority of miRNAs were repressed and only a relatively small portion of miRNAs were positively-regulated. Verifications of differential expression were

performed by TaqMan Low-Density Arrays (TLDAs) and individual quantitative polymerase chain reaction (qPCR) assays of 10 significantly up-regulated miRNAs, of which 8 miRNAs comprise 2 miRNA families and 2 miRNAs are found individually. These miRNAs were further characterized using *in situ* hybridization to determine where expression occurs in the brain during HSVE. Additionally, potential targets of the miRNAs were bioinformatically predicted using target prediction software. Multiple signaling pathways were among the top results for the 10 up-regulated miRNAs, of which cellular processes such as proliferation and apoptosis could be impacted. To investigate the potential role of the selected miRNAs during these processes, multiple experiments were performed in order to investigate the viability of the cells during over-expression of the miRNAs and the effects of miRNAs during apoptotic stress. Finally, the ability to repress specific genes of the HSV-1 genome based on bioinformatically predicted targets was confirmed via luciferase assays.

Hypothesis

The overall hypothesis of this project is that the host response to an acute viral encephalitis (and the virus itself) will cause a range of gene expression changes, including changes at the post-transcriptional level of regulation where miRNAs are involved. MiRNAs will therefore contribute to the disease progression in the brain during HSVE.

Objectives:

1. To generate an *in vivo* miRNome of an acute viral encephalitis using HSVE in a murine model.
2. To determine putative functions of the miRNAs during infection by analyzing target predictions and subsequently testing the putative targets with various functional assays.
3. To determine the regions of the brain that differentially express selected miRNAs.

2.0 MATERIALS AND METHODS

2.1 TISSUE CULTURE CELL LINES AND VIRUSES

All cell lines were purchased from American Type Culture Collection (ATCC, Manassas, VA). HeLa cells were used to perform luciferase assays and to test conditions for functional assays, and Neuro-2A cells (a neuroblastoma tissue culture cell line) were used as a biologically relevant cell line for functional assays. Vero E6 cells were used to propagate and titre stocks of HSV-1 strain F (ATCC, Manassas, VA). All cell lines were maintained and passaged in Eagle's Minimum Essential Medium (EMEM, ATCC, Manassas, VA) supplemented with 10% fetal bovine serum (FBS, Life Technologies, Burlington, ON) and 1% antibiotic-antimycotic (Life Technologies, Burlington, ON), as recommended by ATCC.

2.1.1 PREPARATION AND TITERING OF VIRAL STOCKS

Prior to adding virus, two passages of Vero E6 cells were reduced to medium containing 5% FBS (Life Technologies, Burlington, ON) and 1% antibiotic-antimycotic (Life Technologies, Burlington, ON) in 150 cm^2 tissue culture flasks (Corning, Tewksbury, MA). Viral stocks were diluted by adding 100 µL of HSV-1 strain F (ATCC, Manassas, VA) to 20 mL serum-free EMEM (ATCC, Manassas, VA). Cells were washed with Dulbecco's Phosphate Buffered Saline (DPBS, Life Technologies, Burlington, ON) and 3 mL of the viral dilution were added to the flasks. Cells were incubated for 1 hour at 37°C before the addition of 20 mL of medium containing 2% FBS and were incubated for a further 48 hours. Cells were scraped off the flasks and pooled twice to maximize the amount of virus recovered. The cells were pelleted by centrifuging at 2000 x g for 5 minutes and the supernatant was discarded. Ten millilitres of fresh DPBS was added and

the cells were then subjected to 3 freeze-thaw cycles consisting of an acetone (Thermo Fisher Scientific, Nepean, ON)-dry ice bath followed by an incubation at 37°C in a bead bath comprising of Lab Armor Beads (Cornelius, OR) in a VWR model 1235 PC water bath (Mississauga, ON). Finally, the viral stock was centrifuged at 2000 x g for 5 minutes to pellet cell debris and the supernatant was aliquoted into 2 mL tubes (Sarstedt, Montreal, QC) and stored at -80°C.

Viral titres were determined by plaque assay. Vero E6 cells were plated on 6-well plates and incubated until ~90% confluency. The virus stock was serially diluted (0.1 mL into 0.9 mL EMEM containing 2% FBS and 1% antibiotic-antimycotic) to 10^{-7}. Cells were washed once with 1X PBS and 300 µL of each virus dilution was added to the cell monolayers in duplicate. Virus was absorbed for 1 hour at 37°C prior to removal and rinsing with 1X PBS. Two millilitres of an agar overlay, consisting of 20 mL 5% SeaPlaque agarose (Lonza, Basel, Switzerland) diluted in PBS, 18.4 mL 2X EMEM (Lonza, Basel, Switzerland), 0.8 mL 1X FBS, 0.4 mL antibiotic-antimycotic and 0.4 mL GlutaMAX supplement (Life Technologies, Burlington, ON) was added to each dish. Plaques were evident after a 2-3 day incubation at 37°C and these were counted. Plaque-forming units (PFU)/mL were determined with the following calculation:

$$\frac{PFU}{mL} = \frac{number\ of\ plaques}{(dilution\ factor)(volume\ of\ virus\ added)}$$

2.2 ANIMAL EXPERIMENTS

All procedures involving live animals were approved by the Canadian Science Centre for Human and Animal Health - Animal Care Committee (CSCHAH-ACC) according to the guidelines set by the Canadian Council on Animal Care. All protocols were designed to minimize animal discomfort.

2.2.1 INTRACEREBRAL INOCULATION

Mice were first acclimatized for 1 week and a baseline weight was recorded the day before inoculation in order to monitor weight loss of sick animals. Ten female SJL mice were intracerebrally inoculated and monitored for infection. Mice were anesthetized with Isoflurane prior to inoculation and a 3 mm stepper needle was used to inject into the left hemisphere of the mouse brain. A total of 10 µL was injected into each animal: 6 mice were infected with HSV-1 strain F (3 with $1x10^3$ PFU virus concentration and 3 with $1x10^6$ PFU), 2 mice underwent mock infection with 1X PBS, and 2 mice were inoculated with Vero cell supernatant as a specific control for debris present in the viral stocks. Mice were examined daily and scored based on the appearance of clinical signs of disease. Animals were sacrificed and brains were extracted 2 days post infection (dpi) after the infected animals all suffered multiple symptoms including seizures. Samples were either fixed in 10% formalin and subsequently paraffin-embedded, or were cut on the sagittal midline and placed in optimal cutting temperature (OCT) compound (Sakura, Alphen aan den Rijn, The Netherlands). OCT samples were frozen via a methanol/dry ice bath and were stored at -80°C, whereas formalin-fixed, paraffin-embedded (FFPE) samples were stored at room temperature[69].

2.2.1.1 HISTOLOGY AND IMMUNOHISTOCHEMISTRY STAINING

FFPE samples were stained to observe the effects of the inoculations[69]. They were first de-paraffinized by incubating overnight at 60°C and then washed in xylene (Surgipath, Winnipeg, MB) twice for 5 min each. They were then rinsed before being placed in haematoxylin (Gill III Formula, Surgipath, Winnipeg, MB) for 3 min. Slides were rinsed with water and then de-

stained in 1% acid alcohol for 30 sec. Slides were counter-stained by placing in eosin (Surgipath, Winnipeg, MB) for 3 min and were then rinsed with water. Slides were then washed with ethanol and mounted with Permount (Thermo Fisher Scientific, Nepean, ON)[69].

For HSV-1 detection, slides were de-paraffinized at 65°C overnight and then washed in 2 xylene baths for 3 min each. They were then rinsed with water and were submerged in Antigen Unmasking Solution (Vector, Burlington ON) where they were incubated for 10 min at 100-115°C. Slides were rinsed with Tris buffered saline (TBS) and Tris buffered saline with Tween 20 (TBST) before blocking with 1X Universal Blocking Reagent (Biogenex, Markham ON) at room temperature for 10 minutes. The rabbit anti-HSV-1 antibody (Dako, Burlington, ON) was diluted 1:1000 in EnVision FLEX antibody diluent (Dako, Burlington, ON) and added, followed by UltraVision One AP Polymer (Thermo Fisher Scientific, Nepean ON). Slides were incubated for a further 30 minutes at room temperature. The slides were then washed with 10x TBST and the substrate was added. Slides were incubated for 2.5 min at room temperature and then rinsed with water. They were counterstained with haematoxylin for 1 min, rinsed again and washed with 1% acid alcohol for 10-20 sec. Finally, slides were rinsed with STWS for 10-20 sec[69].

2.2.2 RNA EXTRACTION

A half of a brain (embedded in OCT compound, Sakura, Alphen aan den Rijn, The Netherlands) from both a 10^3 PFU HSV-1-infected mouse and a mock infected mouse were thawed and OCT solution was removed by washing in 1X PBS. The *mir*Vana miRNA isolation kit (Life Technologies, Burlington, ON) was used and the recommended protocols for frozen tissue and total RNA isolation were followed. Briefly, 0.25 g of tissue (the entire half-brain) was homogenized on ice in 2.5 mL of Lysis/Binding Buffer (Life Technologies, Burlington, ON) Next,

250 µL miRNA Homogenate Additive (Life Technologies, Burlington, ON) was added, mixed and incubated on ice for 10 minutes. Following incubation, 2.5 mL of Acid-Phenol:Chloroform (Life Technologies, Burlington, ON) was added and samples were vortexed for 45 seconds. Samples were then centrifuged at 10,000 x g for 5 minutes and the upper phase was transferred to fresh tubes. Subsequently, 1.25 volumes of 100% ethanol (EtOH) was added, and the mixture was loaded onto a Filter Cartridge (Life Technologies, Burlington, ON) centrifuged and washed as per the manufacturer's instructions. Finally, the RNA was eluted into a fresh tube by adding 100 µL nuclease-free water (Life Technologies, Burlington, ON) pre-heated to 95°C to the Filter Cartridge and centrifuged at 14,000 rpm for 30 seconds. RNA quantity and quality was assessed using an Agilent 2100 Bioanalyzer (Mississauga, ON) and the RNA 6000 Nano Kit (Agilent, Mississauga, ON), following the recommended protocol.

2.2.3 SEQUENCE ANALYSIS

Sequencing and primary analysis (the generation of data in .FASTA and/or .FASTQ file format from image files produced during sequencing) was performed by The Centre for Applied Genomics using the Illumina Pipeline software (San Diego, CA). Secondary analysis (the filtering and processing of FASTA/FASTQ files to gain meaningful results) was performed in-house by creating a protocol using two cloud-based software programs: pre-processing of reads was performed in Galaxy[70–72], and mature and novel reads were identified with miRanalyzer[73]. The protocol has recently been published as a book chapter by Ying (ed., 2013)[74] and can also be found in Appendix 2.0. An outline of the protocol is shown in Figure 5.

.FASTQ files were first uploaded to Galaxy via file transfer protocol (FTP) server and were identified as 'fastqsanger' format for downstream processing. They were then processed via a

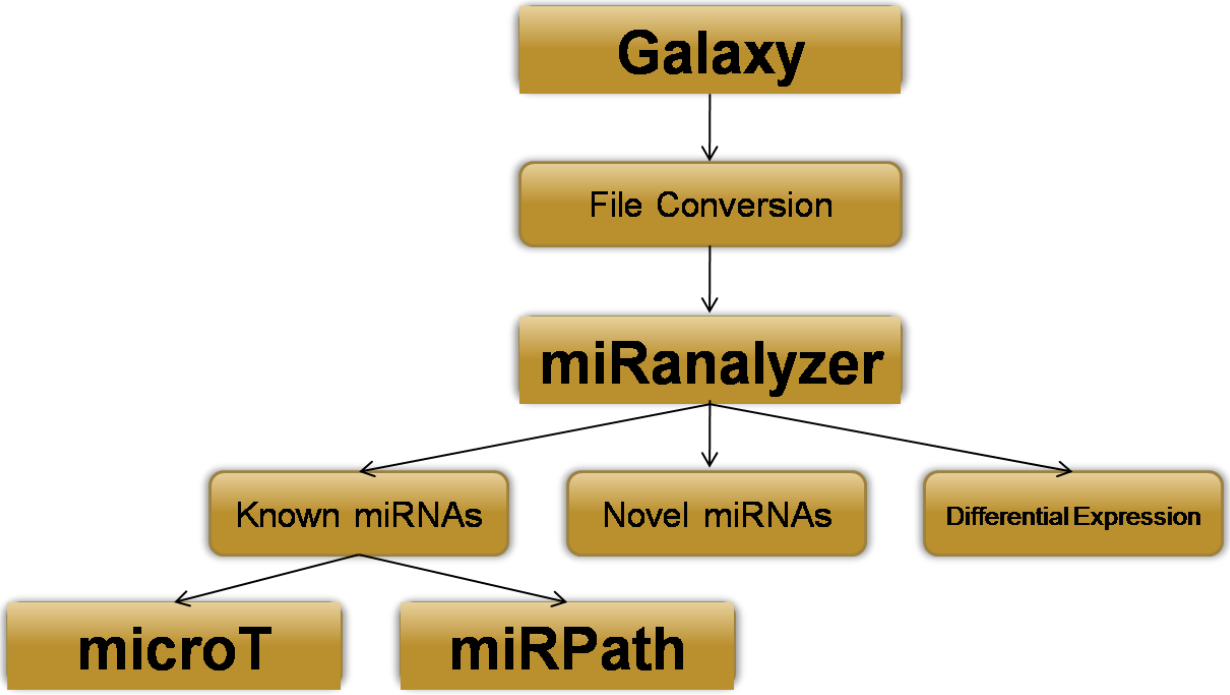


Figure 5. The workflow for secondary analysis of NGS data as developed in-house. Pre-processing of data is completed with Galaxy, which includes trimming out the adapter sequence, and filtering out low-quality data. Next, file conversion is performed on the command line terminal of a Unix operating system in order to create the proper format to be accepted by miRanalyzer. Known miRNAs are identified by miRanalyzer, along with predicted novel miRNAs. In addition, miRanalyzer compares samples to output differential expression patterns. Finally, known miRNAs were further analyzed for portential function using DIANA LAB microT (for target prediction) and miRPath (for pathway analysis). Bolded text represents the programs used. Adapted from Ying (2013)[74].

workflow (http://main.g2.bx.psu.edu/u/kyle-caligiuri/w/mirna-secondary-analysis) that combines several Galaxy tools: Clip, Filter FASTQ and FASTQ to FASTA. Clip first trims the 3' adaptor sequences from the reads and removes reads less than 15 bp long. Filter FASTQ removes low quality reads that have >3 bp lower than a 99% accurate base-call. FASTQ to FASTA removes the quality data that is no longer needed, which converts the file to .FASTA format.

Converted files were downloaded to a Unix computer and further formatted at the command line terminal. Unique reads were grouped and counted using the following command:

```
grep -d ">" input.fasta | sort | uniq -c > sorted.txt
```

where input.fasta is the original .FASTA file, and sorted.txt is the new tab-delimited, read count text file. The text file columns need to be reversed to be accepted by miRanalyzer, so a stream editor (sed) program was created to achieve this, by entering the following into the Text Editor:

```
#!/usr/bin/sed -f
s/^\s\+(\w\+\)\s*\(\w*\)$/\2\1/
```

The sed program was then initialized with the following command entered in the terminal:

```
cat sorted.txt | sed -f flip_script > output.txt
```

where flip_script is the above sed program and output.txt is the newly-sorted read count table. Read count files were uploaded to miRanalyzer (http://bioinfo5.ugr.es/miRanalyzer/miRanalyzer.php) and the following parameters were used to analyze the data: 0 mismatches were allowed for the discovery of known miRNAs, 1 mismatch was allowed when mapping to the libraries (ie. Repbase[75], the database of known repetitive DNA and RFam[76], the database of known non-coding RNAs), 2 mismatches were

allowed when mapping to the genome, the threshold of the posterior probability was set to 0.95 and the minimum number of models which predict a novel miRNA was set to 3. The original .FASTQ files were also sent to LC Sciences (Houston, TX) to be analyzed with their proprietary ACGT101-miR v3.5 software (LC Sciences, Houston, TX) and to Geospiza (Seattle, WA) to be analyzed on their Genesifter beta software (Geospiza, Seattle, WA) to confirm the results.

Tertiary analysis (the analysis of the newly-discovered miRNAs to gain insight into their biological functions) was performed by looking at target predictions of selected miRNAs using tools such as microT[77] and miRPath[78] on the DIANA LAB website (http://diana.cslab.ece.ntua.gr/?sec=home, Alexander Fleming Biomedical Sciences Research Center, Athens, Greece).

2.2.4 QUANTITATIVE PCR (qPCR) VERIFICATION OF miRNA EXPRESSION PROFILE

To verify the results of the deep sequencing analysis, two subsets of miRNAs were validated by qPCR. A larger subset was first validated using TaqMan Low Density Arrays (TLDAs). Using the rodent Megaplex Pools of primers (Life Technologies, Burlington, ON) on the samples, a total of 518 mouse-specific miRNAs were validated along with 303 rat-specific miRNAs. The TLDAs were run following the manufacturer's recommended instructions, without a preamplification step (as recommended based on the quantity of total RNA available for each sample). In brief, samples were first reverse transcribed into cDNA by combining the recommended reagents with the sample and the Megaplex RT primers. Samples were amplified on an MJ-Research PTC-200 Peltier Thermal Cycler (Bio-Rad, Mississauga, ON) using the following conditions: 40

cycles consisting of 16°C for 2 minutes, 42°C for 1 minute and 50°C for 1 second; followed by an incubation at 85°C for 5 minutes, then cooled to 4°C. Samples were then loaded along with TaqMan Universal PCR Master Mix, No AmpErase UNG (Life Technologies, Burlington, ON) into 384-well plates, which were pre-loaded with TaqMan primers and probes (Life Technologies, Burlington, ON) corresponding to the amplified products along with appropriate controls. Samples were run on an Applied Biosystems 7900HT Fast Real-Time PCR System (Life Technologies, Burlington, ON) with the following conditions: a 2 minute incubation at 50°C and a 10 minute incubation at 95°C followed by 40 cycles of 15 seconds at 95°C and 1 minute at 60°C. The small nucleolar RNA (snoRNA) U6 was used as an endogenous control to compare differential expression rates of the miRNAs tested. The samples were analyzed using Integromics' RealTime StatMiner software (Madison, WI).

The expression of a smaller subset of miRNAs were also verified via individual miRNA qPCR assays. Total RNA was first reverse transcribed into cDNA following the recommended protocol for the High-Capacity cDNA Archive Kit (Life Technologies, Burlington, ON). Samples were added to the recommended amounts of supplied reagents and loaded onto the MJ-Research PTC-200 Peltier Thermal Cycler (Bio-Rad, Mississauga, ON) with the following settings: 25°C for 10 minutes, followed by 37°C for 120 minutes. The qPCR reactions were then prepared using the TaqMan Gene Expression Assays (Life Technologies, Burlington, ON) specific to the miRNAs tested, in triplicate, following the manufacturer's protocol. The samples were run on an Applied Biosystems 7900HT Fast Real-Time PCR System (Life Technologies, Burlington, ON) using the following conditions: a 2 minute incubation at 50°C, a 20 second incubation at 95°C, then 40 cycles of 95°C for 1 second and 60°C for 20 seconds. The endogenous control used was

again snoRNA U6 and a negative control was run alongside the samples using a plant miRNA ath-miR-163. Samples were analyzed using the $2^{-\Delta\Delta CT}$ method to determine relative fold change, as described by Livak and Schmittgen (2001)[79]. The following calculation is therefore used:

$$Relative\ Fold\ Change = 2^{-[(C_{T,sample\ gene} - avgC_{T,sample\ endog\ ctrl}) - average(C_{T,reference\ gene} - avgC_{T,reference\ endog\ ctrl})]}$$

where $C_{T,\ sample\ gene}$ is the threshold cycle (C_T) of the miRNA of interest in the infected sample, avg $C_{T,\ sample\ endog\ ctrl}$ is the average C_T of snoRNA U6 in the infected sample, $C_{T,\ reference\ gene}$ is the C_T of the same miRNA in the mock-infected sample, and avg $C_{T,\ reference\ endog\ ctrl}$ is the average C_T of snoRNA U6 in the mock-infected sample. Mock-infected ΔC_Ts were combined for an average first before comparing to infected ΔC_Ts to form the final $\Delta\Delta C_T$. Relative fold changes were also averaged for a final fold change per sample.

2.3 PRIMERS AND PLASMID DESIGN

Two plasmids were made for luciferase reporter assays in order to test the binding potential of miR-150 to the putative target U_L18/19/20 3'UTR. Both plasmids use the pmirGLO Vector (Promega, Madison, WI) and contain a 66 bp insert containing the entire 3'UTR. However, one plasmid was used as a control vector by flipping the insert so that the proper miRNA target would not be transcribed. This control vector would still transcribe luciferase but would not be able to be repressed by miR-150 due to the lack of a complementary target sequence. Vector information and a summary of the cloning process is outlined in Figure 6.

For the U_L18/19/20 functional plasmid, DNA was extracted from a sample of Vero E6 (ATCC, Manassas, VA) tissue culture cells infected with HSV-1 strain F (ATCC, Manassas, VA) for 8

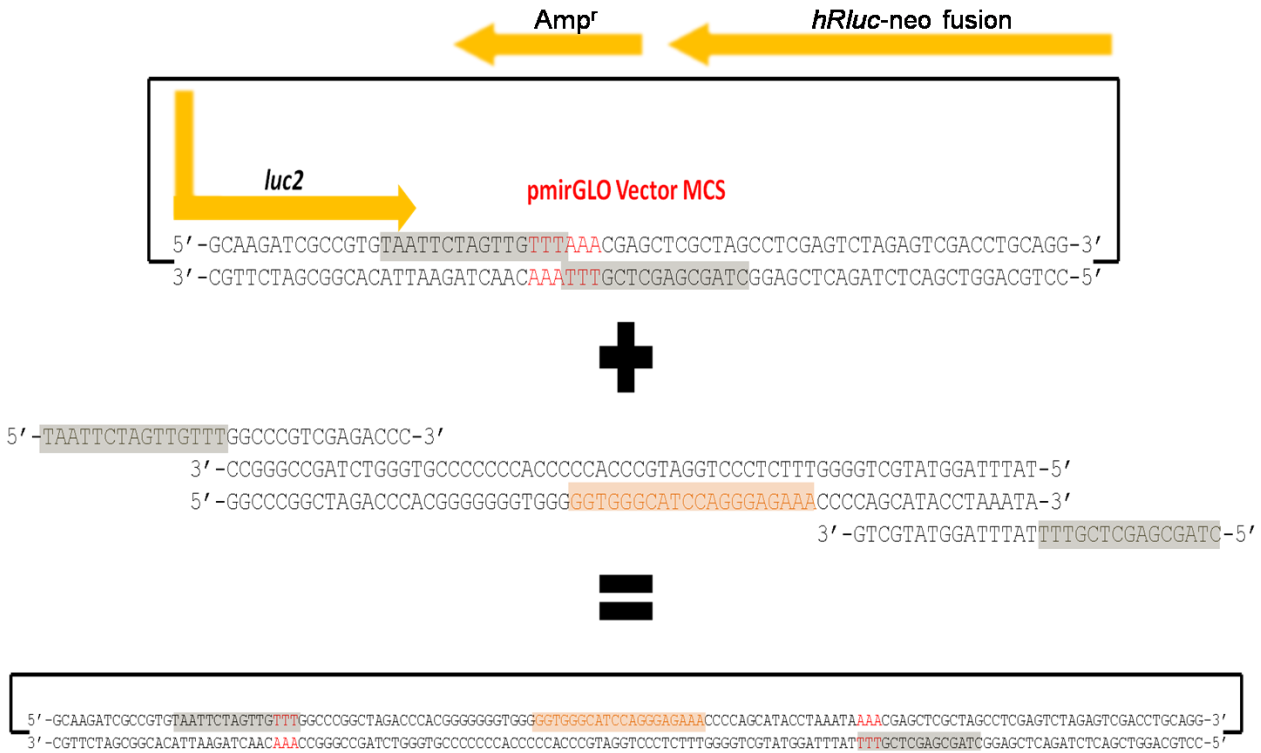


Figure 6. Summary of cloning procedure for U_L18/19/20 pmirGLO Vector. Vector was cut at the PmeI restriction site (red) within the MCS, which flanks the luc2 luciferase gene. Primers that were complimentary to both the vector and insert were used to isolate the 3'UTR segment of the U_L18/19/20 genes that contains the putative binding site for mmu-miR-150 (orange). Poxvirus DNA polymerase favours blunt ends and excises segments on the 3' side, creating sticky ends on both the vector and insert[80]. The complimentary portion of the insert that matches the vector (grey) is annealed onto the vector and the insert is incorporated, creating the final product. Amp^r: ampicillin resistance cassette. *hRluc*-neo fusion: humanized *Renilla* luciferase gene fusion with neomycin resistance cassette.

hours[69]. The sample was subsequently fragmented with Fermentas' (Burlington, ON) FastDigest EcoR1 kit, following the recommended protocol. The pmirGLO Vector (Promega, Madison, WI) was cut with the PmeI restriction enzyme (New England Biolabs, NEB, Pickering, ON) using general conditions outlined in the In-Fusion 2.0 CF Liquid PCR Cloning Kit, creating a single blunt cut within the multiple cloning site (MCS). Life Technologies' (Burlington, ON) PCR SuperMix High Fidelity was used to amplify the inserts following the suggested protocol, using the primers synthesized in-house by the DNA Core facility. Next, a proprietary Cloning Enhancer was added to the PCR products in accordance with the recommended In-Fusion protocol, and the samples were incubated at 37°C for 20 minutes. The treated insert was then added to 200 ng of the linearized pmirGLO Vector (yielding a 2:1 ratio of insert to vector) along with the recommended reagents according to the manufacturer's protocol and was incubated at 37°C for 15 minutes, followed by 50°C for 15 minutes and finally transferred to ice. The equation to determine molar ratios is:

$$insert{:}\ vector\ ratio\ sought \times \frac{[amount\ of\ vector\ (ng) \times insert\ size\ (bp)]}{vector\ size\ (bp)} = amount\ of\ insert\ (ng)$$

The reaction mixture was diluted in 30 µL TE buffer (Qiagen, Toronto, ON) and used to transform Fusion-Blue Competent Cells according to the recommended protocol.

For the control U_L18/19/20 plasmid with the flipped insert, a small oligonucleotide was synthesized by Integrated DNA Technologies (IDT, Coralville, IA). The insert was then ligated onto the vector at a ratio of 100:1 following the procedure outlined by Worthington *et al.* (2001)[81] and incubated at 16°C overnight. The reaction mixture was used to transform JM109 Competent Cells (Promega, Madison, WI) as recommended by the pmirGLO protocol. The

transformation followed recommended procedure outlined by Promega Corporation (Madison, WI).

All transformants were analyzed by both PCR screening and DNA sequencing to ensure the entire sequence had been cloned in the proper orientation. The following sequencing primers were synthesized by Life Technologies (Burlington, ON), which flank the MCS and will capture the entire insert during the sequencing run: 5'-AAGAAGCTGCGCGGTGGTGTTGT-3' and 5'-TGTAGAGGGGGACTTGGACTTTGT-3'.

Upon confirmation of an accurate sequence, plasmid stocks were made by following recommended procedure outlined by Invitrogen Corporation (Burlington, ON). Plasmids were amplified following recommended protocol for use of Qiagen's EndoFree Plasmid Purification Maxi Kit (Toronto, ON). Final plasmid DNA concentration was determined via a Nanodrop ND-1000 spectrophotometer (Thermo Fisher Scientific, Nepean, ON).

2.4 LUCIFERASE REPORTER ASSAY

To test the binding capabilities of miR-150 to the putative U_L18/19/20 3'UTR target, co-transfections of the pre-miR (Life Technologies, Burlington ON) and the pmirGLO reporter vector mentioned above were performed, followed by luciferase assays after 24 hour incubations to allow ample time for the transfection reactions to occur. The Effectene Transfection Reagent (Qiagen, Toronto, ON) was used with its recommended protocol. As such, $1x10^4$ HeLa cells were first plated into 96-well white-walled plates (Corning, Tewksbury, MA) and supplemented with 125 uL EMEM to incubate for 24 hours. Prior to transfection, varying amounts of pre-miR-150 and 15.6 ng plasmid DNA (previously optimized) were diluted in a proprietary buffer (Qiagen, Toronto, ON). The various reagents were mixed with the diluted

plasmid DNA and pre-miRs and the complexes were incubated for 10 minutes at room temperature. Cells were washed with DPBS (Life Technologies, Burlington, ON) and fresh growth medium was replaced. Complexes were then diluted with 25 µL growth medium per well, mixed briefly by pipetting, and added drop-wise to the wells. Plates were swirled gently to mix and then incubated at 37°C for 24 hours. Each sample was performed in triplicate.

Luciferase assays were performed according to instructions. Three readings were taken per well for both the firefly and *Renilla* luciferase readings. The luminometer measures light output from luciferase production and records arbitrary relative light units (RLU). In addition to the conditions mentioned above, four controls were also included: a control containing only fresh growth medium is used to determine baseline RLU; a negative control to the pre-miRs used corresponding concentrations of a 'scrambled' pre-miR (Life Technologies, Burlington, ON) that shouldn't bind to the reporter vector; a positive control contained the reporter vector only, without adding pre-miR; and the control reporter vector with the flipped insert acted as a negative control to the proper reporter vector, since luciferase could not be repressed by miR-150 due to the lack of a valid binding site.

2.5 APOPTOSIS MEDIATION WITH SELECTED miRNAS

To determine the effects of miRNAs from the miR-141/200c family on apoptosis, the ADP/ATP Ratio Assay (Abcam, Toronto, ON) was used. First, $1x10^4$ Neuro-2A cells were plated in 96-well plates (Corning, Tewksbury, MA) and transfected as described above using Effectene Transfection Reagent (Qiagen, Toronto, ON). Briefly, varying amounts of pre-miRs were diluted in the supplied buffer (Qiagen, Toronto, ON). The various reagents were mixed with the diluted pre-miRs and the complexes were incubated for 10 minutes at room temperature. Cells were

washed with PBS (Life Technologies, Burlington, ON) and fresh growth medium was replaced. Complexes were diluted with 25 µL growth medium per well, mixed briefly by pipetting, and added drop-wise to the wells. Plates were swirled gently to mix and incubated at 37°C. Twenty four hours after transfection, cells were induced for apoptosis using a UVC 500 crosslinker (GE Healthcare Life Sciences, Baie d'Urfe, QC). Cells were subjected to 100 mJ/cm^2 UVC and then prepared using the ADP/ATP Ratio Assay as described above so that readings could be taken at 30 minutes, 1- 2- and 3 hours after induction of apoptosis. Conditions were performed in triplicate and a negative control was created by transfecting a 'scrambled' pre-miR (Life Technologies, Burlington, ON) that should not bind to any transcripts within the cells. Untransfected cells were also induced along with the transfected cells to compare the levels of apoptosis between endogenous and ectopic miRNA expression. Data from Section 2.7 were used as the control for cells that were not stimulated for apoptosis. Data was analyzed using the calculation shown in Section 2.7.

2.6 CELL VIABILITY ASSAYS

MiRNAs were tested for their effect on cell viability *in vitro* by transfecting increasing amounts of pre-miRs (Life Technologies, Burlington, ON) into Neuro-2A cells (ATCC, Manassas, VA) and measuring cell vitality by both standard cell counting with Trypan Blue (Life Technologies, Burlington, ON) and with an ADP-ATP Ratio Assay (Abcam, Toronto, ON). For the Trypan Blue staining, 3x10^5 cells were plated in 6-well plates (Corning, Tewksbury, MA) and were transfected following the recommended protocol as described above using the Effectene Transfection Reagent (Qiagen, Toronto, ON). MiRNAs were tested in duplicate and two negative controls were also tested: a scrambled miRNA (Life Technologies, Burlington, ON) was

transfected for comparison and non-transfected cells were used to normalize results. After a 24 hour incubation following transfection, cells were trypsinized with TrypLE, stained with Trypan Blue and counted on a BrightLine hemacytometer (Sigma-Aldrich, St. Louis, MO).

To confirm our results, $1x10^4$ cells were also plated in 96-well plates (Corning, Tewksbury, MA) and transfected as before. Twenty-four hours after transfection, media was removed and proprietary reagents from the ADP/ATP Ratio Assay (Abcam, Toronto, ON) were added according to the manufacturer's recommended protocol, before reading on a plate reader to measure ATP activity. The plates were read again after 10 minutes and then 1 µL proprietary enzyme was added to each well. Samples were incubated for 1 minute before reading on the palte reader to measure ADP activity. To analyze the data, the following calculation was used to determine the ADP/ATP ratio:

$$\frac{ADP}{ATP} ratio = \frac{(3rd\ reading - 2nd\ reading)}{1st\ reading}$$

Conditions were performed in triplicate and samples were compared to untransfected cells. A negative control using a scrambled pre-miR (Life Technologies, Burlington, ON) was used again. However, no mock-transfection was performed.

2.8 IN SITU *HYBRIDIZATION (ISH) AND FLUORESCENT LABELING OF PARAFFIN-EMBEDDED BRAIN SECTIONS*

To identify the areas of the brain that were expressing certain miRNAs, ISH was performed on HSV-1 infected and mock-infected mouse brains by the Molecular PathoBiology department of the CSCHAH. ISH of select miRNAs was performed using the IsHyb In Situ Hybridization (ISH) Kit (BioChain, Newark, CA). Briefly, slides were deparaffinized, then fixed in 4% paraformaldehyde

and washed twice in DEPC-PBS for 5 minutes each. Next, slides were permeabilized with proteinase K (Life Technologies, Burlington, ON) and fixed again with 4% paraformaldehyde. Slides were pre-hybridized with proprietary buffer overnight in a Boekel Scientific InSlide Out hybridization oven (Feasterville, PA) at 57-63°C depending on the miRNA (previously optimized; see Table 1 for exact conditions of each probe). Hybridization of slides was performed using 40 nM concentrations of the linearized double-end digoxygenin labeled miRNA probes (Exiqon, Woburn, MA) in proprietary hybridization buffer. Slides were incubated overnight in the hybridization oven at 52-58°C. Following hybridization, slides were washed in TBST and blocked with proprietary blocking solution for 1 hour at room temperature. To visualize the miRNAs, samples were incubated overnight at 4°C with a 1:500 concentration of alkaline phosphatase (AP)-conjugated anti-digoxygenin 1° antibody (BioChain, Newark, CA) diluted in blocking solution. Next, slides were washed and submerged in proprietary AP buffer twice for 5 minutes each at room temperature. Slides were then incubated in nitro blue tetrazolium (NBT, BioChain, Newark, CA) and 5-bromo-4-chloro-3-indolyl-phosphate (BCIP, BioChain, Newark, CA) overnight in the dark at room temperature. The slides were then rinsed with distilled water and counterstained by submerging in Nuclear Fast Red (Vector Laboratories, Burlington, ON) for 2 minutes followed by a 2 minute rinse with distilled water. Slides were air dried and then immersed in xylene for 2 minutes before coverslips were mounted with Permount (Thermo Fisher Scientific, Toronto, ON). Negative controls were run alongside the samples that lacked the probes for the specific miRNAs. Slides were imaged on a Mirax MIDI digital slide scanner (Zeiss, Toronto, ON).

Table 1. Exact conditions for each probe during ISH procedure. Conditions were individually troubleshot to reduce background staining of samples.

Probe Target	Pre-hybridization Temperature	Probe Concentration	Hybridization Temperature
miR-141	61°C	40 nM	56°C
miR-183	63°C	40 nM	58°C
miR-200a	57°C	40 nM	52°C

Some slides were instead stained to target areas of infection via staining for HSV-1. Slides were deparaffinized and blocked with 3% hydrogen peroxide (H_2O_2) for 10 minutes and rinsed with water. Antigen retrieval was performed by submerging slides in 10 nM sodium citrate buffer (pH 6.0) for 10 minutes at 120°C. Slides were blocked again for 1 hour at room temperature with 1:20 normal goat serum (Cedarlane, Burlington, ON) diluted in EnVision FLEX antibody diluent (Dako, Burlington, ON). Next, slides were incubated overnight at 4°C with rabbit anti-HSV-1 antibody (Dako, Burlington, ON) diluted 1:1000 in EnVision FLEX antibody diluent. Slides were rinsed twice with TBST, then incubated at room temperature for 1 hour in Alexa Fluor 594 goat anti-rabbit IgG 2° antibody (Life Technologies, Burlington, ON). The fluorescent antibody was diluted as above for the primary antibody. Finally, slides were rinsed twice with water, submerged in DAPI (diluted identically to the antibodies) for 20 minutes at room temperature, and rinsed again. Slides were air dried overnight and mounted with Permount.

3.0 RESULTS

3.1 IN VIVO MIRNA PROFILE DURING HSVE IN MICE

3.1.1 INTRACEREBRAL INOCULATION OF MICE WITH HSV-1

Mice were inoculated with HSV-1 to determine the transcriptional response in acute encephalitis. A total of 9 mice were used for the experiment, of which 4 were used as controls. Mice were intracerebrally infected with either $1x10^3$ PFU or $1x10^6$ PFU HSV-1 strain F propagated in Vero E6 cells, or mock-infected with either 1x PBS or Vero cell supernatant. While the mock-infected mice displayed no symptoms from infection, the HSV-1 infected mice showed signs of illness such as hunched posture, sunken eyes and ruffled appearance[69]. The infected mice also lost between 1-2 grams of body mass. All mice were sacrificed and brains were extracted at 2 dpi after at least 3 of the infected mice were afflicted with seizures, and samples were stored in either OCT medium and frozen at -85°C or were formalin fixed and paraffin embedded.

Histology performed in the Molecular PathoBiology core laboratory at the CSCHAH confirmed the characteristic neuropathology of HSVE in the infected mice with extensive lesions in the hippocampus (Figure 7). An obvious absence of these lesions were observed in the mock-infected samples. Immunohistochemistry (IHC) staining for HSV-1 confirmed viral presence at the hippocampal lesions in the infected mice and spread throughout the brain to areas such as the cerebellum and hypothalamus. A lack of presence in the control samples was both expected and observed.

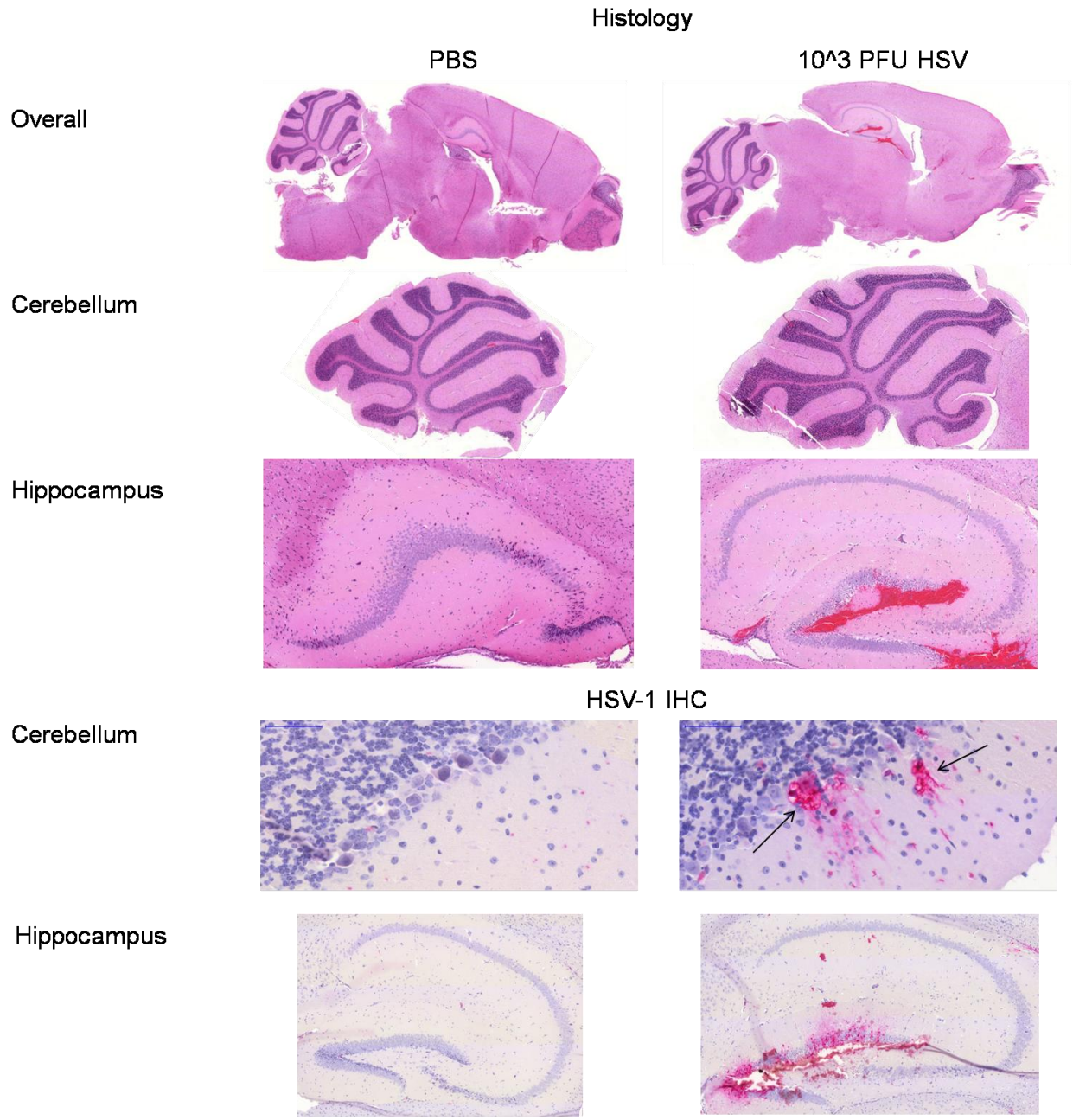


Figure 7. Histology and immunohistochemistry (IHC) staining of HSV-1 infected and PBS mock-infected mouse brains. In the histology sections, hemorrhaging is observed in the HSV-1 infected hippocampus but not in the mock-infected control. In the IHC sections, HSV-1 (red) was found spread throughout the brain, such as at the cerebellum where no extensive hemorrhaging was observed, as well as at the hippocampus were damage occurred.

3.1.2 RNA PREPARATION

RNA was prepared from half brains infected or mock infected female SJL mice[69]. Total RNA was extracted as described in Section 2.2.2 and samples were analyzed for quantity and quality via Nanodrop and Bioanalyzer (in duplicate), respectively. The results are described in Table 2. Samples were then sent to The Centre for Applied Genomics in Toronto, ON to generate a small RNA library and run on the high throughput sequencer as described in Section 2.2.2. A brief outline of the library preparation and sequencing is described in Figure 8.

3.1.3 ANALYSIS OF HIGH THROUGHPUT SEQUENCING DATA

Next generation sequencing and primary data analysis were performed by The Centre for Applied Genomics and the data was sent back for further downstream analysis. Using current cloud-based software, a protocol was developed (Appendix 1.0) in order to first process the raw reads and then map them against multiple databases of small RNAs to identify known miRNAs present in the samples, as well as to filter out other non-coding RNA sequences. Additionally, the protocol enables the discovery of novel miRNAs by mapping to the genome of interest. Finally, the protocol covers the prediction of potential miRNA targets and their possible ontological relevance. The focus of the protocol was to minimize computing requirements and the need for extensive computer programming knowledge while still obtaining data comparable to proprietary analyses. Therefore, in an effort to validate the protocol, data was also submitted to geospiza (Seattle, WA) for use with their beta version of GeneSifter and to LC Sciences in Houston, TX to be analyzed with their proprietary software. The various methods are outlined in Figure 9 and their different characteristics are compared in Table 3.

Table 2. Quality and quantity of total RNA sent used for the generation of an in vivo miRNA profile of acute viral encephalitis caused by HSV-1 strain F. Quality is given as an average RNA integrity number (RIN), an arbitrary quality score produced by the Bioanalyzer with a maximum of 10, and quantity is given in ng/µL as an average between Nanodrop and Bioanalyzer results.

Sample	Concentration (ng/µL)	Quality (RIN value)
PBS (mock-infected)	1441.57	7.90
HSV-1	597.69	6.70

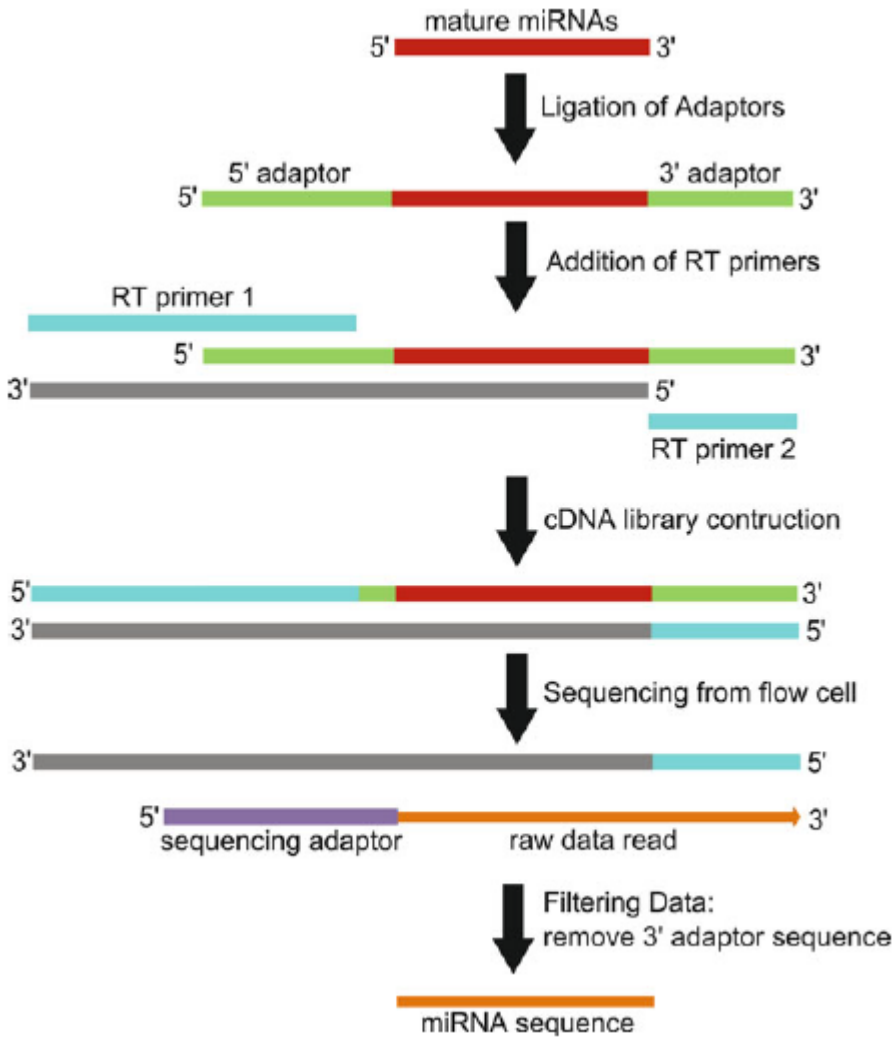


Figure 8. Overview of the deep sequencing process. Small RNAs are size selected from total RNA via an agarose gel and adaptors are ligated onto the RNAs. These adaptors are important for the sequencing process: the 3' adaptor allows complementarity to the primers attached to the flow cell and are important for capturing the sample, while the 5' adaptors allow complementarity to the sequencing adaptor. Next, RT-PCR is performed to generated a cDNA library. The cDNA is separated into single strands and loaded onto the flow cell, where only the strands complementary to the bound primers (the antisense strands) will anneal. The sequencing process is then started and the sequencing adaptors are added, which anneal to the sample so a new strand can be created with the addition of dNTPs. Four images are captured with every additional nucleotide (1 of each colour corresponding to each possible nucleotide) and these are compiled during primary analysis to create a raw data read. The raw read obtained will still contain the 3' adaptor sequence and therefore it needs to be removed in order to obtain the proper miRNA sequence. This step is performed during secondary analysis. Figure reprinted with permission from Majer *et al.* (2013)[74].

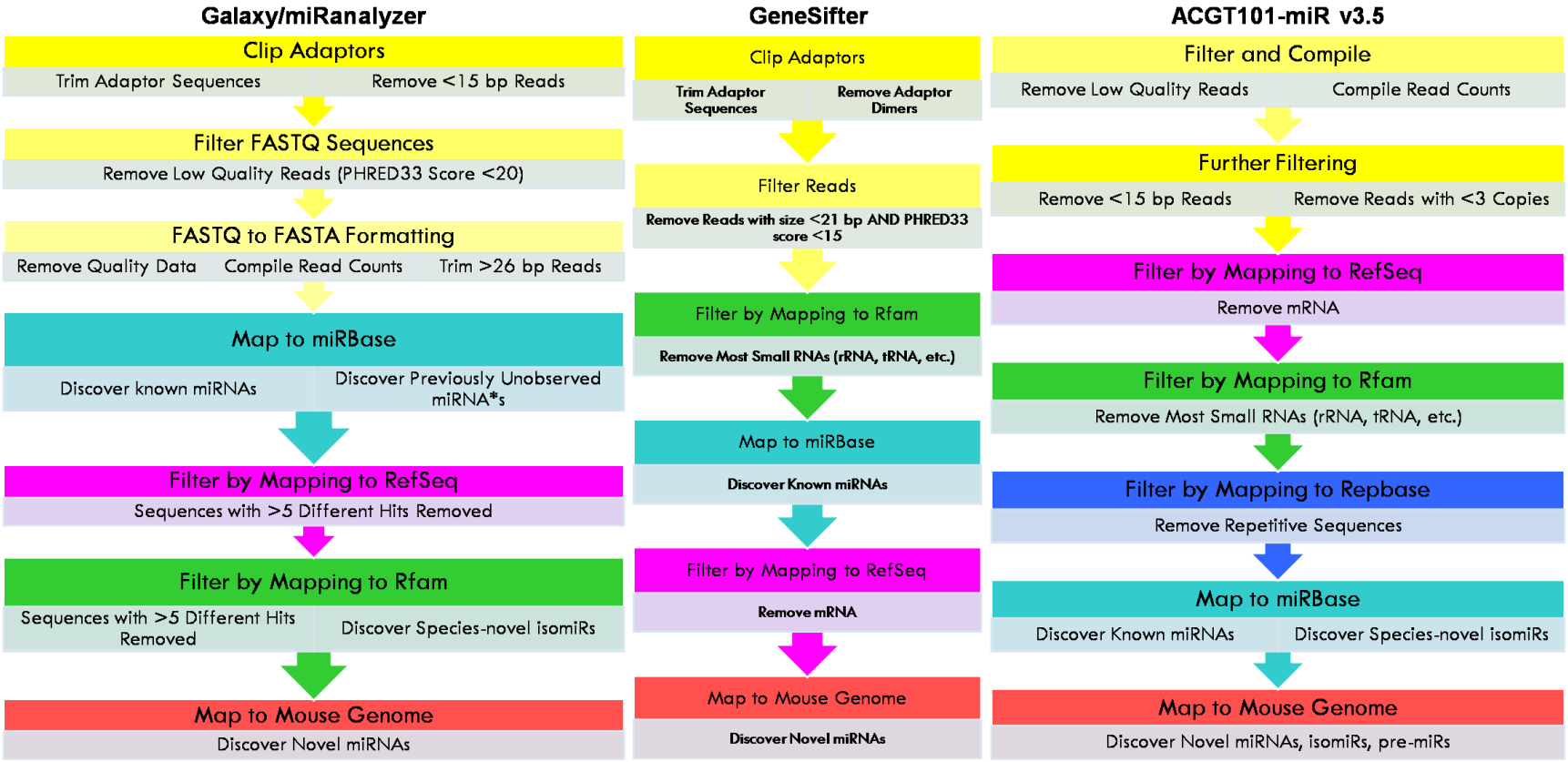


Figure 9. Flow diagrams of the three different methods used for secondary analysis. Pre-processing of the reads (yellow) is an important first step to allow downstream applications to properly interpret the data. All programs can capitalize on FASTQ quality data to remove reads; each program has its own slightly different standards to comply with. Reads too small to be miRNAs are also removed and files can be compiled and condensed by summing the total read counts for each unique sequence. After pre-processing, the files are mapped against various databases. Mapping to miRBase (teal) and the genome (red) are the most important steps, as they generate the lists of known and novel miRNAs, respectively. However, mapping to additional sets of known sequences before mapping to the genome can further filter the data so that known sequences such as rRNAs, tRNAs, mRNAs and repetitive sequences aren't predicted as miRNAs. Only miRanalyzer maps to Rfam after mapping to miRBase to include known miRNAs that are present in the Rfam database (green).

Table 3. Summary of characteristics for the various software used to predict miRNAs from deep sequencing data.

Program	Format/ Accessibility	Location	Read Pre-Processing	Target Genomes	Other Databases Incorporated	Mapping Algorithm	Functions
miRanalyzer	Web-based	http://bioinfo5.ugr.es/miRanalyzer/miRanalyzer.php	Accepts FASTA or tab-delimited read count summary files. Adaptor sequences must be removed prior to use.	Seven genomes (dog, fruit fly, human, mouse, nematode, rat, zebra fish) Multiple versions available.	miRBase, RefSeq, Rfam	Fixed: Bowtie. Certain parameters are variable.	Novel, known miRNA prediction, normalization, some species distribution via Rfam, differential expression
GeneSifter	Proprietary service, web-based	http://www.geospiza.com/Products/AnalysisEdition.shtml	Accepts FASTA/FASTQ file formats. Adaptor sequences can be left intact.	Flexible	miRBase, RefSeq, Rfam	Flexible: Bowtie, Mapreads, MAQ	Novel, known miRNA prediction at the pre-miR level, normalization, genomic visualization
ACGT101-miR	Proprietary service, large file transfer method required	http://www.lcsciences.com/	Accepts FASTA/FASTQ file formats. Adaptor sequences can be left intact.	Flexible	miRBase, RefSeq, Repbase, Rfam	nd	Novel, known miRNA prediction, species distribution, differential expression

nd: no data available.

3.1.3.1 DATA ANALYSIS USING CLOUD-BASED COMPUTER PROGRAMS

To perform the secondary analysis of the deep sequencing data, the protocol developed in-house used 3 Galaxy[70–72] tools to remove unwanted data. First, Clip trimmed the 3' adaptor sequences from the reads (Figure 9). These 3' adaptor sequences are an artifact from the sequencing process; they are used to bind the samples to the flow cell. Next, Filter FASTQ removed low quality reads by filtering reads less than 15 bp long and reads with a Phred33[82] score lower than 20 on 3 or more bases. Phred33 scoring is contained within the FASTQ file along with the FASTA-formatted sequence, and a score of 20 or higher ensures a minimum 99% accuracy in basecalling[83,84]. The third Galaxy tool, FASTQ to FASTA, converted the filtered data to .FASTA format by removing the quality data. .FASTA files were then downloaded and sorted via basic commands on a Unix terminal, and by the creation of a simple 2-line sed program to format the data. This created a tab-delimited list of unique reads and a count of each unique read. While miRanalyzer is able to take FASTA files, the read count still needs to be incorporated into the file, so converting it to a tab-delimited format was the easiest method found that incorporates all necessary data. The tab-delimited read count file was then uploaded to miRanalyzer[73,85], another cloud-based analysis program used to gain meaningful results from that data including lists of known and possible novel miRNAs. The samples were also compared to each other to provide differential expression data.

MiRanalyzer identified 642 known mature miRNAs in the PBS mock-infected sample and 521 mature miRNAs in the HSV-1 infected sample using the same cutoffs as GeneSifter (only 1 mismatch allowed and at least 3 reads detected per miRNA). MiRanalyzer found additional miRNAs previously known and identified in miRBase (a database of all known miRNAs) which

were classified into several categories: ambiguous mature miRNAs matched several different miRNA sequences and so they could not be definitively recognized as one or the other; unobserved mature miRNAs contain sequences included in the hairpin structures of known miRNAs and are therefore documented in miRBase - however these miR* sequences haven't yet been observed as mature miRNAs[85]; and unobserved ambiguous miRNAs are miR* sequences previously unobserved that match more than one sequence and so they can be related to any combination of the matches. The top 50 mature miRNAs are compiled into Tables 4 and 5 for the mock-infected and HSV-1 infected samples, respectively. All miRNAs are reported for the other 3 categories of known miRNAs mentioned.

In addition, miRanalyzer is able to find and predict possible novel miRNAs. First, miRNAs that did not map to the reference genome of the species were found by mapping against Rfam[76], a database of many RNA families such as structural RNA and non-coding RNA. 131 unique sequences were identified in the PBS mock-infected sample and 21 unique sequences in the HSV-1 infected sample that match miRNA sequences previously unidentified in the mouse model (Table 6). MiRNAs were also predicted based on the thermodynamic potential to form hairpin structures, and a total of 587 candidates were identified in the mock-infected sample, 182 of which were found outside of the current transcriptome. A total of 389 miRNA candidates were found in the HSV-1 infected sample, with only 89 candidates found outside the transcriptome. Tables 7 and 8 list the top 50 candidates for the PBS mock- infected and HSV-1 infected samples, respectively. Again, all predictions were represented by at least 3 reads in a given sample. While the genome location is given, other programs such as GeneSifter have built-in genome viewers to view the mapped reads in relation to the genome. There is no way

Table 4. List of known mature miRNAs identified by miRanalyzer in the mock-infected sample. MiRNAs are categorized as mature, ambiguous mature, previously unobserved mature or previously unobserved ambiguous mature miRNAs. The list includes the number of unique reads, the total number of reads (Read Count), and two normalized expression values based on the total number of filtered reads (Norm. Expr. All) and the total number of mapped reads (Norm. Expr. Mapped). The top 50 mature miRNAs are displayed, along with all results for the remaining categories.

Known Mature MiRNA	# of Unique Reads	Read Count	Norm. Expr. All (%)	Norm. Expr. Mapped (%)
mmu-miR-143-3p	463	326033	2.572	8.642
mmu-miR-30c-5p	249	195363	1.541	5.178
mmu-miR-378a-3p	262	190264	1.501	5.043
mmu-miR-30a-5p	275	189651	1.496	5.027
mmu-miR-124-3p	679	184968	1.459	4.903
mmu-miR-29a-3p	243	178042	1.404	4.719
mmu-miR-9-5p	354	174988	1.38	4.638
mmu-miR-127-3p	418	155248	1.225	4.115
mmu-miR-101a-3p	369	104071	0.821	2.758
mmu-miR-30e-5p	206	80087	0.632	2.123
mmu-let-7c-5p	243	71817	0.567	1.904
mmu-let-7f-5p	210	70375	0.555	1.865
mmu-miR-29b-3p	128	61719	0.487	1.636
mmu-miR-218-5p	269	60239	0.475	1.597
mmu-miR-99b-5p	197	59046	0.466	1.565
mmu-miR-30b-5p	164	53992	0.426	1.431
mmu-miR-9-3p	196	52945	0.418	1.403
mmu-miR-434-3p	143	48306	0.381	1.28
mmu-miR-125b-5p	237	46937	0.37	1.244
mmu-let-7g-5p	180	46485	0.367	1.232
mmu-miR-30d-5p	168	44778	0.353	1.187
mmu-miR-26a-5p	251	44162	0.348	1.171
mmu-miR-128-3p	291	43822	0.346	1.162
mmu-miR-24-3p	235	40796	0.322	1.081
mmu-miR-101b-3p	226	40284	0.318	1.068
mmu-miR-132-3p	217	39828	0.314	1.056
mmu-miR-181a-5p	192	39742	0.313	1.053
mmu-let-7a-5p	162	35928	0.283	0.952
mmu-let-7b-5p	268	35620	0.281	0.944
mmu-miR-29c-3p	130	32549	0.257	0.863
mmu-let-7i-5p	229	32436	0.256	0.86
mmu-miR-16-5p	160	31127	0.246	0.825

mmu-miR-99a-5p	148	31112	0.245	0.825
mmu-miR-140-3p	252	30501	0.241	0.808
mmu-miR-328-3p	222	29103	0.23	0.771
mmu-miR-34c-5p	52	27594	0.218	0.731
mmu-miR-103-3p	80	27293	0.215	0.723
mmu-miR-191-5p	159	25153	0.198	0.667
mmu-miR-153-3p	93	24733	0.195	0.656
mmu-miR-23b-3p	159	20267	0.16	0.537
mmu-miR-27b-3p	154	20032	0.158	0.531
mmu-miR-451a	103	19823	0.156	0.525
mmu-miR-379-5p	104	19598	0.155	0.519
mmu-miR-21a-5p	110	19499	0.154	0.517
mmu-miR-129-2-3p	95	19007	0.15	0.504
mmu-miR-138-5p	190	18113	0.143	0.48
mmu-miR-192-5p	87	17033	0.134	0.451
mmu-miR-125a-5p	105	16319	0.129	0.433
mmu-miR-137-3p	156	14565	0.115	0.386
mmu-miR-126-3p	128	14288	0.113	0.379
Ambiguous Mature MiRNAs				
mmu-miR-378c, mmu-miR-378a-3p	96	8963	0.071	0.238
mmu-let-7a-1-3p, mmu-let-7c-2-3p	36	1236	0.01	0.033
mmu-miR-107-3p, mmu-miR-103-3p	67	1142	0.009	0.03
mmu-miR-486-5p, mmu-miR-3107-5p	42	953	0.008	0.025
mmu-miR-30c-5p, mmu-miR-30b-5p	45	334	0.003	0.009
mmu-miR-129-2-3p, mmu-miR-129-1-3p	32	326	0.003	0.009
mmu-miR-669a-3p, mmu-miR-669o-3p	13	264	0.002	0.007
mmu-miR-27a-3p, mmu-miR-27b-3p	26	236	0.002	0.006
mmu-miR-199a-3p, mmu-miR-199b-3p	17	224	0.002	0.006
mmu-miR-133a-3p, mmu-miR-133b-3p	17	199	0.002	0.005
mmu-let-7a-5p, mmu-let-7c-5p	44	175	0.001	0.005
mmu-miR-1843b-5p, mmu-miR-1843a-5p	10	104	0.001	0.003
mmu-miR-467d-3p, mmu-miR-467a-3p	6	52	0	0.001
mmu-miR-466i-3p, mmu-miR-466f-3p	7	45	0	0.001
mmu-miR-101a-3p, mmu-miR-101c	17	40	0	0.001
mmu-miR-23b-3p, mmu-miR-23a-3p	18	33	0	0.001
mmu-miR-378c, mmu-miR-378a-3p, mmu-miR-378b	17	25	0	0.001
mmu-miR-669a-5p, mmu-miR-669p-5p	5	21	0	0.001
mmu-miR-466a-3p, mmu-miR-466e-3p	5	19	0	0.001
mmu-miR-466a-3p, mmu-miR-466e-3p, mmu-miR-466b-3p, mmu-miR-466c-3p, mmu-miR-466p-3p	5	17	0	0
mmu-miR-29b-3p, mmu-miR-29c-3p	8	14	0	0

mmu-miR-1982.2-3p, mmu-miR-1982.1-3p	4	10	0	0
mmu-miR-328-3p, mmu-miR-6395	5	7	0	0
mmu-miR-467a-5p, mmu-miR-467b-5p	1	5	0	0
mmu-miR-297b-3p, mmu-miR-297a-3p, mmu-miR-297c-3p	3	4	0	0
mmu-miR-466p-5p, mmu-miR-466a-5p	3	4	0	0
mmu-miR-145a-5p, mmu-miR-145b	3	3	0	0
mmu-miR-3070a-5p, mmu-miR-3070b-5p	3	3	0	0
mmu-miR-3096b-3p, mmu-miR-3096a-3p	2	3	0	0
mmu-miR-467d-3p, mmu-miR-467c-3p, mmu-miR-467a-3p, mmu-miR-467e-3p	1	3	0	0
Previously Unobserved Mature MiRNAs				
mmu-mir-382*	19	488	0.004	45.311
mmu-mir-299*	9	225	0.002	20.891
mmu-mir-425*	15	90	0.001	8.357
mmu-mir-1187*	34	51	0	4.735
mmu-mir-543*	2	31	0	2.878
mmu-mir-153*	5	28	0	2.6
mmu-mir-322*	5	24	0	2.228
mmu-mir-145*	5	20	0	1.857
mmu-mir-1954*	11	16	0	1.486
mmu-mir-1983*	8	13	0	1.207
mmu-mir-3470b*	10	12	0	1.114
mmu-mir-487b*	4	12	0	1.114
mmu-mir-466f-1*	3	8	0	0.743
mmu-mir-503*	3	7	0	0.65
mmu-mir-715*	3	6	0	0.557
mmu-mir-135a-1*	1	5	0	0.464
mmu-mir-712*	1	5	0	0.464
mmu-let-7e*	4	4	0	0.371
Previously Unobserved Ambiguous Mature MiRNAs				
mmu-mir-9-2*, mmu-mir-9-1*, mmu-mir-9-3*	2	7	0	0.65

Table 5. List of known mature miRNAs identified by miRanalyzer in the HSV-1 infected sample. MiRNAs are categorized as mature, ambiguous mature, previously unobserved mature or previously unobserved ambiguous mature miRNAs. The list includes the number of unique reads, the total number of reads (Read Count), and two normalized expression values based on the total number of filtered reads (Norm. Expr. All) and the total number of mapped reads (Norm. Expr. Mapped). The top 50 mature miRNAs are displayed, along with all results for the remaining categories.

Known Mature MiRNA	# of Unique Reads	Read Count	Norm. Expr. All (%)	Norm. Expr. Mapped (%)
mmu-let-7f-5p	464	471739	3.583	21.901
mmu-miR-143-3p	502	211114	1.604	9.801
mmu-let-7g-5p	322	110246	0.837	5.118
mmu-miR-30e-5p	303	109641	0.833	5.09
mmu-let-7c-5p	388	106675	0.81	4.952
mmu-miR-30a-5p	342	90885	0.69	4.219
mmu-let-7a-5p	337	88357	0.671	4.102
mmu-miR-9-5p	417	76174	0.579	3.536
mmu-miR-26a-5p	339	61833	0.47	2.871
mmu-miR-127-3p	354	51799	0.393	2.405
mmu-miR-24-3p	312	43840	0.333	2.035
mmu-miR-378a-3p	210	42682	0.324	1.982
mmu-miR-30d-5p	230	41789	0.317	1.94
mmu-miR-9-3p	252	35224	0.268	1.635
mmu-miR-21a-5p	157	30437	0.231	1.413
mmu-miR-101a-3p	296	28710	0.218	1.333
mmu-miR-103-3p	91	28629	0.217	1.329
mmu-miR-30c-5p	177	27158	0.206	1.261
mmu-miR-29a-3p	179	22475	0.171	1.043
mmu-miR-182-5p	189	20760	0.158	0.964
mmu-miR-379-5p	156	20627	0.157	0.958
mmu-let-7i-5p	260	20238	0.154	0.94
mmu-miR-124-3p	305	19388	0.147	0.9
mmu-miR-101b-3p	187	18895	0.144	0.877
mmu-miR-138-5p	183	18687	0.142	0.868
mmu-miR-146b-5p	144	16529	0.126	0.767
mmu-let-7b-5p	292	15449	0.117	0.717
mmu-let-7d-5p	166	15064	0.114	0.699
mmu-miR-29b-3p	94	10936	0.083	0.508
mmu-miR-34c-5p	50	10812	0.082	0.502
mmu-let-7e-5p	211	10654	0.081	0.495
mmu-miR-137-3p	150	9988	0.076	0.464

mmu-miR-126-3p	158	9961	0.076	0.462
mmu-miR-434-3p	125	9008	0.068	0.418
mmu-miR-181a-5p	142	8824	0.067	0.41
mmu-miR-128-3p	200	7745	0.059	0.36
mmu-miR-30b-5p	121	7717	0.059	0.358
mmu-miR-140-3p	159	6827	0.052	0.317
mmu-miR-340-5p	83	6170	0.047	0.286
mmu-miR-183-5p	151	5968	0.045	0.277
mmu-miR-153-3p	79	5694	0.043	0.264
mmu-miR-16-5p	112	5672	0.043	0.263
mmu-miR-1a-3p	60	5435	0.041	0.252
mmu-miR-381-3p	48	5187	0.039	0.241
mmu-miR-148a-3p	66	5062	0.038	0.235
mmu-miR-218-5p	149	5037	0.038	0.234
mmu-miR-27b-3p	91	4372	0.033	0.203
mmu-miR-127-5p	156	4349	0.033	0.202
mmu-miR-148b-3p	59	4122	0.031	0.191
mmu-miR-191-5p	103	3707	0.028	0.172
Ambiguous Mature MiRNAs				
mmu-miR-378c, mmu-miR-378a-3p	86	2429	0.018	0.113
mmu-miR-107-3p, mmu-miR-103-3p	69	620	0.005	0.029
mmu-let-7a-5p, mmu-let-7c-5p	88	519	0.004	0.024
mmu-let-7a-1-3p, mmu-let-7c-2-3p	14	173	0.001	0.008
mmu-miR-30c-5p, mmu-miR-30b-5p	29	87	0.001	0.004
mmu-miR-1843a-5p, mmu-miR-1843b-5p	7	82	0.001	0.004
mmu-miR-199a-3p, mmu-miR-199b-3p	16	69	0.001	0.003
mmu-miR-486-5p, mmu-miR-3107-5p	8	62	0	0.003
mmu-miR-129-2-3p, mmu-miR-129-1-3p	10	35	0	0.002
mmu-miR-27a-3p, mmu-miR-27b-3p	17	33	0	0.002
mmu-miR-133a-3p, mmu-miR-133b-3p	5	28	0	0.001
mmu-miR-378c, mmu-miR-378a-3p, mmu-miR-378b	15	19	0	0.001
mmu-miR-669a-3p, mmu-miR-669o-3p	6	16	0	0.001
mmu-miR-23b-3p, mmu-miR-23a-	7	11	0	0.001

3p				
mmu-miR-146a-5p, mmu-miR-146b-5p	4	10	0	0
mmu-miR-29b-3p, mmu-miR-29c-3p	2	9	0	0
mmu-miR-467d-3p, mmu-miR-467a-3p	3	8	0	0
mmu-miR-101a-3p, mmu-miR-101c	7	7	0	0
mmu-miR-669a-5p, mmu-miR-669p-5p	3	6	0	0
mmu-miR-466a-3p, mmu-miR-466e-3p	3	4	0	0
Previously Unobserved Mature MiRNAs				
mmu-mir-382*	13	78	0.001	37.681
mmu-mir-153*	7	20	0	9.662
mmu-mir-3472*	10	16	0	7.729
mmu-mir-1954*	12	15	0	7.246
mmu-mir-3470b*	7	10	0	4.831
mmu-mir-145*	4	9	0	4.348
mmu-mir-1187*	9	9	0	4.348
mmu-mir-299*	2	8	0	3.865
mmu-mir-543*	2	7	0	3.382
mmu-mir-140*	3	6	0	2.899
mmu-mir-135a-2*	2	4	0	1.932
mmu-mir-503*	1	4	0	1.932
mmu-mir-322*	3	3	0	1.449
Previously Unobserved Ambiguous Mature MiRNAs				
mmu-mir-9-2*, mmu-mir-9-1*, mmu-mir-9-3*	3	3	0	1.449

Table 6. List of miRNAs identified with miRanalyzer by mapping to Rfam for both the mock-infected and HSV-1 infected samples. These miRNAs did not match the mouse version of the miRNA but did match the human version of the miRNA mentioned. The list includes the amount of unique sequences that match a given miRNA (Unique Reads), the total number of reads (Read Count) and the Rfam accession number along with the name of the miRNA. Capitalized names indicate a match to the miRNA gene but not necessarily to the mature miRNA. Highlighted reads indicate at least one unique read also matches a non-miRNA sequence.

PBS			
Unique Reads	**Read Count**	**Rfam Accession Number**	**Name**
42	165	RF01059	mir-598
6	59	RF00782	MIR480
13	45	RF00998	mir-562
9	33	RF01061	mir-548
8	30	RF00906	MIR1122
6	30	RF01029	mir-649
5	24	RF00994	mir-1255
7	23	RF01045	mir-544
4	19	RF00638	MIR159
6	19	RF01062	MIR812
4	13	RF01058	MIR806
4	12	RF00885	MIR821
3	10	RF00951	mir-1302
2	8	RF01063	mir-324
1	7	RF00639	mir-515
1	7	RF00917	mir-708
1	7	RF01062;RF01058	MIR812;MIR806
1	6	RF00885;RF00754	MIR821;mir-279
1	4	RF00678	mir-140
1	4	RF00876	mir-684
1	4	RF01062;F00885	MIR812R;MIR821
1	3	RF00640	MIR167_1
1	3	RF00764	mir-191
1	3	RF01016	mir-584
1	3	RF00872	mir-652
1	3	RF01062;RF01061	MIR812;mir-548
		HSV-1	
6	75	RF01061	mir-548
3	16	RF01059	mir-598
1	11	RF00782	MIR480
1	9	RF01045	mir-544
2	7	RF00994	mir-1255

1	5	RF01058;RF01062	MIR806;MIR812
1	5	RF00885;RF00904	MIR821;mir-392
1	4	RF00951	mir-1302
1	4	RF01029	mir-649
1	4	RF01058	MIR806
1	4	RF01062	MIR812
1	3	RF00998	mir-562
1	3	RF00885;RF01062	MIR821;MIR812

Table 7. Potential novel miRNAs found in the mock-infected sample using miRanalyzer. The chromosomal location is included with the number of unique reads representing the miRNA, the total read count (minimum 3) and the number of prediction models that verify the miRNA as a possible novel candidate. A minimum of 3 models showing a positive prediction was used as a cutoff to determine novel candidates. Highlighted miRNAs have a chromosomal location outside the current sequenced transcriptome.

Putative Mature MiRNA	Chr.	Chrom. Start	Chrom. End	Strand	Unique Reads	Read Count	No. Positive Prediction Models (min 3, max 5)
GATAGTAGACCGTATAGCGTACGCT	chr12	110948358	110948502	+	184	28196	5
CACTGGACTTGGAGTCAGAAGGCCT	chr18	61557449	61557593	-	223	15085	5
CACATTGATCATTGACACTTCA	chr18	73693084	73693192	-	42	6597	3
TCTGAGATGAAGCACTGTAGCTCAGGAAG	chr18	61808807	61808959	-	289	4896	4
ATGGTAGACTATGGAACGTAGGC	chr12	110947259	110947349	+	137	4527	5
ATTAAGGCACGCGGTGAATGCCAAGAAT	chr14	65209476	65209596	+	584	3485	5
ATTAAGGCACGCGGTGAATGCCAAGA	chr2	180628725	180628831	+	581	3480	5
ATTAAGGCACGCGGTGAATGCCAAGA	chr3	17695650	17695776	+	580	3479	5
ACTGTAAACATCCTCGACTGGAAGCT	chr1	23279091	23279197	+	280	2959	5
TTGTAAACATCCTACACTCTCAGCTGT	chr1	23298532	23298640	+	217	2521	5
TGTGTAAACATCCTACACTCTCAGCTGT	chr4	120442118	120442238	-	214	2506	5
CATCGGATCCGTCTGAGCTTGGCTG	chr12	110831021	110831155	+	302	2476	5
AAACCCGTAGATCCGATCTTGTGG	chr16	77599154	77599276	+	143	2407	5
ATCATAGAGGAACATCCACTTTT	chr12	110961640	110961780	+	44	1865	5
GGTACAGTACTGTGATAACTGAAG	chr4	101019531	101019653	-	316	1800	5
GGTGGATATTCCTTCTATGGTTAC	chr12	110961640	110961782	+	40	1612	5
TCGGGGCACGCCGGCCTCGCGGGCC	chr3	5860630	5860734	+	53	1401	4
TCTAGCACCATCTGAAATCGGTTATA	chr6	31012651	31012757	-	230	1335	5
TGAGGTAGTAGGTTGTATGGTTTA	chr16	77599877	77600019	+	145	976	5
TCTTTGGTTATCTAGCTGTATGAG	chr13	83878391	83878523	+	171	936	5
TCTTTGGTTATCTAGCTGTATGAG	chr7	86650135	86650257	+	171	935	5
TCTTTGGTTATCTAGCTGTATGAG	chr3	88019500	88019632	+	171	935	5
TTTGTGCTTGATCTAACCATGTGGT	chr11	35430311	35430435	+	160	782	5
GTTGTGCTTGATCTAACCATGTGCTT	chr5	48615164	48615310	+	143	731	5
ATGTAAACATCCTACACTCAGCTGTC	chr15	68168967	68169073	-	82	718	5
CTGTAAACATCCTTGACTGGAAGCTG	chr4	120445201	120445307	-	156	710	5
CTGTAAACATCCCCGACTGGAAGCT	chr15	68172737	68172881	-	154	697	5
CACCCGTAGAACCGACCTTGCG	chr17	17967143	17967231	+	150	679	5
GTTTGAACCATCACTCGACTCCT	chr12	110832696	110832836	+	119	609	5
TAGGCAGTGTAGTTAGCTGATTGCTA	chr9	50911103	50911239	-	45	595	3
AAGCAGCATTGTACAGGGCTATGAAA	chr2	131113757	131113903	+	67	576	5
AAGCAGCATTGTACAGGGCTATGAA	chr11	35595868	35596012	+	66	575	5

TGAGGTAGTAGGTTGTATAGTTTA	chr9	41344775	41344917	+	97	519	5
TGAGGTAGTAGATTGTATAGTTG	chr13	48633189	48633309	-	86	445	5
CATAAAGCTAGATAACCGAAAGTAAA	chr13	83878381	83878527	+	140	440	5
TGAGGTAGTAGGTTGTGTGGTTTCAG	chr15	85537730	85537846	+	132	440	5
CATAAAGCTAGATAACCGAAAGTAAA	chr3	88019502	88019628	+	140	440	5
GAACATTCAACGCTGTCGGTGAGTTT	chr1	139863009	139863145	+	196	416	5
GAACATTCAACGCTGTCGGTGAGTTT	chr2	38708245	38708341	+	196	416	5
TTCAAGTAATCCAGGATAGGCT	chr10	126432559	126432697	+	119	410	5
TTCAAGTAATCCAGGATAGGCT	chr9	118940894	118941022	+	119	410	5
GGTACAGTACTGTGATAGCTGAAGAAT	chr19	29209750	29209888	+	133	356	5
CTCACAGTGAACCGGTCTCTTTTTC	chr1	130098915	130099029	+	140	356	5
CATAAAGCTAGATAACCGAAAGTAGA	chr7	86650127	86650263	+	96	349	5
GACATCTGTCACCCCATTGATCGCCAGG	chr15	93064792	93064912	+	17	338	3
GTCCCTGAGACCCTAACTTGTGAT	chr9	41389982	41390124	+	96	336	5
CCCGGCGGATCTTTCCCGCACCC	chr5	147072845	147072965	+	27	317	3
TACCACAGGGTAGAACCACGGACA	chr8	110075124	110075236	+	140	310	5
AATCACATTGCCAGGGATTACCAC	chr13	63401767	63401889	+	104	292	5
GTAACAGTCTACAGCCATGGTCGC	chr11	74987165	74987267	+	84	285	5

Table 8. Potential novel miRNAs found in the HSV-1 infected sample using miRanalyzer. The chromosomal location is included with the number of unique reads representing the miRNA, the total read count (minimum 3) and the number of prediction models that verify the miRNA as a possible novel candidate. A minimum of 3 models showing a positive prediction was used as a cutoff to determine novel candidates. Highlighted miRNAs have a chromosomal location outside the current sequenced transcriptome.

Putative Mature MiRNA	Chr.	Chrom. Start	Chrom. End	Strand	Unique Reads	Read Count	No. Positive Prediction Models (min 3, max 5)
GATAGTAGACCGTATAGCGTACGCT	chr12	110948358	110948502	+	211	12914	5
AGTGAGGTAGTAGATTGTATAGTTGTGG	chr13	48633181	48633311	-	349	12697	4
CTGAGATGAAGCACTGTAGCTCAG	chr18	61808811	61808953	-	340	4026	5
TTGAGGTAGTAGGTTGTATGGTTTA	chr16	77599876	77600020	+	294	3521	5
TTGAGGTAGTAGGTTGTATAGTTTA	chr9	41344774	41344918	+	263	3247	5
GCTGAGGTAGTAGTTTGTACAGTTTG	chr9	106081150	106081266	+	227	2928	5
ACTGTAAACATCCTCGACTGGAAGCT	chr1	23279091	23279197	+	349	2675	5
ACTGGACTTGGAGTCAGAAGGCCT	chr18	61557449	61557591	-	169	2636	5
CTGTAAACATCCTTGACTGGAAGCT	chr4	120445203	120445307	-	295	2618	5
ATGGTAGACTATGGAACGTAGGC	chr12	110947259	110947349	+	196	2608	5
ATCGGATCCGTCTGAGCTTGGCTG	chr12	110831018	110831160	+	344	2078	5
TCTCTTTGGTTATCTAGCTGTATGAGT	chr7	86650138	86650256	+	237	2024	5
TATCTTTGGTTATCTAGCTGTATGAGT	chr13	83878389	83878527	+	236	2023	5
TATCTTTGGTTATCTAGCTGTATGAGT	chr3	88019503	88019631	+	236	2023	5
GGGGGTGTAGCTCAGTGGTAGAGCGCG TGC	chr13	21334419	21334533	+	13	1794	3
TTCAAGTAATCCAGGATAGGCT	chr9	118940894	118941022	+	217	1519	5
TTCAAGTAATCCAGGATAGGCT	chr10	126432559	126432697	+	216	1516	5
TGTAAACATCCCCGACTGGAAGCTGT	chr15	68172744	68172870	-	264	1402	5
GTACAGTACTGTGATAACTGAAGG	chr4	101019530	101019652	-	299	1155	5
TTAAGGCACGCGGTGAATGCCAAGAAT	chr14	65209478	65209596	+	312	1120	5
TTAAGGCACGCGGTGAATGCCAAGA	chr3	17695652	17695776	+	308	1116	5
TTAAGGCACGCGGTGAATGCCAAGA	chr2	180628727	180628831	+	308	1115	5
CATAAAGCTAGATAACCGAAAGTAAAA	chr13	83878380	83878528	+	226	1111	5
CATAAAGCTAGATAACCGAAAGTAAAA	chr3	88019491	88019639	+	226	1111	5
ATAGCTTATCAGACTGATGTTGACTGT	chr11	86397546	86397674	-	162	851	5
GGTGGATATTCCTTCTATGGTTACG	chr12	110961640	110961784	+	76	850	5
CATAAAGCTAGATAACCGAAAGTAGAA	chr7	86650131	86650259	+	164	850	5
TGAGGTAGGAGGTTGTATAGTTGA	chr17	17967295	17967427	+	126	833	5
ACATTGATCATTGACACTTCAAA	chr18	73693087	73693187	-	21	828	3
TCTAGCACCATCTGAAATCGGTTATAA	chr6	31012650	31012758	-	146	680	5
GTACAGTACTGTGATAGCTGAAGAAT	chr19	29209752	29209888	+	180	669	5
AAGAGGTAGTAGGTTGCATAGTTTTAG	chr13	48631361	48631509	-	144	601	5
CTGGCTCAGTTCAGCAGGAACAGGA	chr13	63402493	63402607	+	116	581	4

CTGGCTCAGTTCAGCAGGAACAGGA	chr8	86732693	86732837	+	115	580	5
TTGTAAACATCCTACACTCTCAGCTG	chr1	23298532	23298638	+	131	574	5
GTGTAAACATCCTACACTCTCAGCTG	chr4	120442116	120442242	-	130	573	5
AGCAGCATTGTACAGGGCTATGAAA	chr2	131113759	131113903	+	93	567	5
AGCAGCATTGTACAGGGCTATGAA	chr11	35595870	35596012	+	89	562	5
GTTTGAACCATCACTCGACTCCTG	chr12	110832695	110832837	+	95	481	5
TTTGGCAATGGTAGAACTCACACCG	chr6	30115892	30116016	-	133	469	5
TGAGGTAGTAGGTTGTGTGGTTTC	chr15	85537730	85537842	+	171	469	5
CAGCTGGTGTTGTGAATCAGGCCGA	chr8	96848181	96848305	+	86	458	5
CAGCTGGTGTTGTGAATCAGGCCGTT	chr9	122592000	122592096	+	81	444	5
CTCGTACCGTGAGTAATAATGCG	chr2	26446863	26446963	+	155	398	5
TCCATTGTGCTTTGCACGCGTGGGTTCGA	chr19	5038279	5038431	-	69	350	3
AACATTCAACGCTGTCGGTGAGTTT	chr1	139863010	139863144	+	132	289	5
AACATTCAACGCTGTCGGTGAGTTT	chr2	38708246	38708340	+	132	289	5
TCACAGTGAACCGGTCTCTTTTT	chr1	130098913	130099033	+	118	266	5
AAACCCGTAGATCCGATCTTGTGGT	chr16	77599154	77599278	+	51	257	5
CTAGCACCATTTGAAATCAGTGTT	chr6	31012985	31013127	-	52	255	5

to visualize how the reads overlap at the miRNA location with miRanalyzer. Lastly, miRanalyzer is able to compare the known miRNAs from each sample against each other to identify differential expression between control and infected samples (Table 9). This allows for the identification of miRNAs that either increase or decrease in abundance during infection and by how large of a change, or if there is no change at all.

3.1.3.2 PRELIMINARY ANALYSIS WITH GEOSPIZA GeneSifter BETA SOFTWARE

.FASTQ files were sent to geospiza to validate the protocol developed in-house by using a different analysis program: GeneSifter. GeneSifter is a web-based application for secondary analysis of next generation sequencing data. A beta version was employed to first trim adaptor sequences and remove adaptor dimers (Figure 9). Next, reads were filtered based on size (reads less than 21 bp were removed) and quality (reads with a Phred33 score less than 15 were removed). Reads were subsequently filtered by mapping to Rfam and removing known small RNAs before detection of known miRNAs by mapping to miRBase, and normalization in the reads per million (RPM) format. RPM is calculated as: (# reads for a given miRNA/total usable reads)x1,000,000, where usable reads are the reads remaining after filtering based on quality and read length (pre-processing). RPM is the most common method of normalization, however, no clear standard has yet been decided[86]. For example, Tables 4 and 5 show that miRanalyzer normalizes its data in two different ways: as a percentage of all processed, filtered reads and as a percentage of reads mapped to miRBase.

GeneSifter is very flexible in that it can take both FASTQ and FASTA formats of data, which are both obtained from an Illumina deep sequencing run. It is able to use one of three alignment programs to map reads to the various databases; Bowtie[87], Mapreads and Mapping and

Table 9. Differential expression results between the mock-infected sample and the HSV-1 infected sample using miRanalyzer. Total read counts are compared between PBS and HSV-1 samples, followed by the fold change and log_2 fold change. The log_2 fold change allows for actual positive and negative expression values whereas a down-regulation with normal fold change is represented as a value less than 1. The top 25 largest changes for both up- and down-regulation are displayed.

Name	Mock	HSV-1	Fold Change	Log_2 Fold Change
Up-regulated				
mmu-miR-378d	3	175	58.3333	5.86625
mmu-miR-3471	2	33	16.5	4.04439
mmu-miR-706	1	16	16	4
mmu-mir-3472*	1	16	16	4
mmu-miR-183-5p	441	5968	13.5329	3.7584
mmu-miR-200b-5p	2	22	11	3.45943
mmu-miR-5615-5p	1	8	8	3
mmu-miR-1a-3p	749	5435	7.25634	2.85924
mmu-miR-141-3p	169	1208	7.14793	2.83753
mmu-miR-5117-3p	2	14	7	2.80735
mmu-let-7f-5p	70375	471739	6.70322	2.74485
mmu-miR-200c-3p	70	443	6.32857	2.66188
mmu-miR-200b-3p	291	1649	5.66667	2.5025
mmu-miR-155-5p	5	26	5.2	2.37851
mmu-miR-344d-3-5p	2	10	5	2.32193
mmu-let-7k	7	34	4.85714	2.28011
mmu-miR-5117-5p	16	76	4.75	2.24793
mmu-miR-3096a-3p	3	14	4.66667	2.22239
mmu-miR-96-5p	297	1317	4.43434	2.14872
mmu-miR-1186a	1	4	4	2
mmu-miR-3057-5p	1	4	4	2
mmu-miR-193b-5p	1	4	4	2
mmu-miR-135a-2-3p	6	23	3.83333	1.9386
mmu-miR-539-3p	8	28	3.5	1.80735
mmu-miR-144-5p	587	1803	3.07155	1.61897
Down-regulated				
mmu-miR-485-3p	4194	76	0.01812	-5.7862
mmu-miR-339-5p	567	10	0.01764	-5.8253
mmu-miR-323-3p	993	17	0.01712	-5.8682
mmu-miR-34b-3p	767	13	0.01695	-5.8826
mmu-miR-92b-3p	4201	67	0.01595	-5.9704
mmu-let-7g-3p	1401	22	0.0157	-5.9928

mmu-miR-331-3p	7633	112	0.01467	-6.0907
mmu-miR-504-5p	146	2	0.0137	-6.1898
mmu-miR-702-3p	147	2	0.01361	-6.1997
mmu-miR-877-3p	888	12	0.01351	-6.2095
mmu-miR-323-5p	78	1	0.01282	-6.2854
mmu-miR-1306-5p	173	2	0.01156	-6.4346
mmu-miR-328-3p	29103	316	0.01086	-6.5251
mmu-miR-134-3p	389	4	0.01028	-6.6036
mmu-miR-7a-2-3p	389	4	0.01028	-6.6036
mmu-miR-874-3p	1503	15	0.00998	-6.6467
mmu-miR-3085-3p	102	1	0.0098	-6.6724
mmu-miR-532-3p	750	7	0.00933	-6.7434
mmu-let-7i-3p	3944	34	0.00862	-6.858
mmu-miR-574-3p	388	3	0.00773	-7.015
mmu-miR-3102-3p.2-3p	142	1	0.00704	-7.1497
mmu-miR-193b-3p	3835	24	0.00626	-7.32
mmu-miR-326-3p	5349	23	0.0043	-7.8615
mmu-miR-665-3p	1785	7	0.00392	-7.9944
mmu-miR-1249-3p	3956	2	0.00051	-10.95

Assembly with Quality (MAQ, http://maq.sourceforge.net/) are all built into the application. This analysis used Bowtie with the -n 1 option, meaning only 1 mismatch was allowed within the first 5 bases of the read, since the seed sequences are usually highly conserved[41]. In contrast, miRanalyzer used more stringent settings than GeneSifter when comparing to miRBase (0 mismatches). Despite this, miRanalyzer also identified more known miRNAs than GeneSifter.

Table 10 lists the 50 most abundant miRNAs identified in the PBS mock-infected sample, and a comparison to the same miRNAs found within the HSV-1 infected sample. In summary, 415 known miRNAs were identified in the mock-infected sample, and 385 known miRNAs were identified in the HSV-1 infected sample. These miRNAs were represented by at least 3 reads present in a given sample. Three reads is a typical low-end cutoff[88–92] that was used with all 3 analyses.

3.1.3.3 LC SCIENCES DATA ANALYSIS WITH ACGT101-miR v3.5

LC Sciences (Houston, TX) performed an analysis of the same small RNA deep-sequencing data using their proprietary analysis software, ACGT101-miR v3.5. Similar to the other programs, ACGT101-miR also filters out reads based on size and quality before mapping to databases (Figure 9). This analysis was the most stringent, by removing all reads found in other RNA databases before mapping to miRBase, and by filtering against an additional database (Repbase, a database of known repetitive sequences[75]) that was not used by the other programs mentioned. ACGT101-miR allowed no mismatches within the entire read when mapping, which resulted in a lower abundance of both known and novel miRNAs. In contrast, miRanalyzer allowed 0 mismatches only within the seed sequence when mapping to miRBase

Table 10. List of the top 50 known miRNAs present in the mock-infected sample as found by GeneSifter, and the comparative positions found in the HSV-1 sample. Included is the accession number, chromosomal location and the number of reads present in each sample for each miRNA. Read count is also given as reads per million (RPM), a common method of normalization.

		Sample					
		Mock			HSV-1		
MiRNA Hairpin	Accession Number and Chromosomal Location	Position	# of Reads	RPM	Position	# of Reads	RPM
mmu-mir-30a	MI0000144 1(23279107-23279178)	1	363550	77971.02	5	227524	100190.1
mmu-mir-143	MI0000257 18(61808849-61808912)	2	348481	74739.15	10	197056	86773.57
mmu-mir-30e	MI0000259 4(120445210-120445302)	3	320739	68789.29	8	203045	89410.82
mmu-mir-9-2	MI0000157 13(83878418-83878490)	4	269380	57774.26	14	115001	50640.66
mmu-mir-9-1	MI0000720 3(88019519-88019608)	5	269379	57774.05	13	115002	50641.11
mmu-mir-9-3	MI0000721 7(86650149-86650239)	6	269361	57770.19	15	114986	50634.06
mmu-mir-29a	MI0000576 6(31012659-31012747)	7	268998	57692.33	29	29981	13202.13
mmu-mir-30d	MI0000549 15(68172769-68172851)	8	265406	56921.95	11	129579	57060.08
mmu-mir-101a	MI0000148 4(101019549-101019632)	9	264008	56622.12	18	75079	33061.02
mmu-mir-124-1	MI0000716 14(65209493-65209578)	10	261594	56104.39	36	23918	10532.29
mmu-mir-124-3	MI0000150 2(180628744-180628812)	11	261585	56102.46	35	23920	10533.17
mmu-mir-124-2	MI0000717 3(17695661-17695770)	12	261572	56099.67	34	23931	10538.01
mmu-mir-29c	MI0000577 1(196863740-196863828)	13	258252	55387.62	31	28164	12402.01
mmu-mir-378	MI0000795 18(61557488-61557554)	14	252603	54176.08	22	48280	21260.09
mmu-mir-30c-1	MI0000547 4(120442138-120442227)	15	226247	48523.47	33	27903	12287.08
mmu-mir-30c-2	MI0000548 1(23298539-23298623)	16	225784	48424.17	32	28142	12392.32
mmu-let-7a-1	MI0000556 13(48633547-48633641)	17	212183	45507.15	1	667807	294068.7
mmu-let-7a-2	MI0000557 9(41344798-41344894)	18	206743	44340.43	2	647766	285243.6
mmu-mir-127	MI0000154 12(110831055-110831125)	19	197226	42299.3	21	56609	24927.76
mmu-let-7c-2	MI0000560 15(85537032-85537127)	20	154387	33111.57	6	207363	91312.25

mmu-let-7b	MI0000558 15(85537748-85537833)	21	152271	32657.75	9	198268	87307.27
mmu-let-7c-1	MI0000559 16(77599901-77599995)	22	151646	32523.7	7	203474	89599.73
mmu-mir-101b	MI0000649 19(29209768-29209865)	23	126730	27179.94	25	40985	18047.74
mmu-let-7f-2	MI0000563 X(148346888-148346971)	24	118835	25486.69	3	565957	249219
mmu-let-7f-1	MI0000562 13(48633197-48633286)	25	117375	25173.56	4	565443	248992.7
mmu-mir-29b-1	MI0000143 6(31013022-31013093)	26	82872	17773.66	47	11878	5230.475
mmu-mir-29b-2	MI0000712 1(196863233-196863314)	27	82133	17615.17	48	11819	5204.494
mmu-mir-218-2	MI0000701 11(35430317-35430427)	28	72240	15493.4	67	5007	2204.831
mmu-mir-218-1	MI0000700 5(48615180-48615290)	29	71493	15333.19	68	4843	2132.614
mmu-mir-99b	MI0000147 17(17967151-17967221)	30	69350	14873.58	82	2860	1259.4
mmu-mir-30b	MI0000145 15(68168976-68169072)	31	63319	13580.1	53	8134	3581.805
mmu-mir-434	MI0001526 12(110832715-110832809)	32	61290	13144.94	45	12867	5665.981
mmu-let-7g	MI0000137 9(106081170-106081258)	33	55675	11940.69	12	115253	50751.63
mmu-mir-411	MI0001163 12(110948384-110948466)	34	53931	11566.65	39	19711	8679.734
mmu-let-7d	MI0000405 13(48631380-48631483)	35	53154	11400	16	97569	42964.49
mmu-mir-125b-2	MI0000152 16(77646517-77646588)	36	52539	11268.1	75	3533	1555.756
mmu-mir-125b-1	MI0000725 9(41390008-41390085)	37	51657	11078.94	76	3336	1469.007
mmu-mir-26a-2	MI0000706 10(126432585-126432669)	38	51593	11065.21	19	62097	27344.4
mmu-mir-26a-1	MI0000573 9(118940913-118941003)	39	51458	11036.26	20	62034	27316.66
mmu-mir-181a-1	MI0000697 1(139863031-139863118)	40	50717	10877.34	51	9426	4150.737
mmu-mir-128-1	MI0000155 1(130098937-130099007)	41	50434	10816.64	55	7486	3296.458
mmu-mir-132	MI0000158 11(74987183-74987249)	42	49967	10716.48	93	2357	1037.904
mmu-mir-181a-2	MI0000223 2(38708254-38708330)	43	49666	10651.93	52	9273	4083.363
mmu-mir-99a	MI0000146 16(77599180-77599245)	44	49434	10602.17	77	3335	1468.567
mmu-mir-128-2	MI0000726 9(112021139-112021215)	45	48739	10453.11	56	7024	3093.017
mmu-mir-24-2	MI0000572 8(86732713-86732820)	46	47366	10158.64	23	43522	19164.9

mmu-let-7e	MI0000561 17(17967315-17967408)	47	47131	10108.24	17	83789	36896.47
mmu-mir-24-1	MI0000231 13(63402515-63402583)	48	47046	10090.01	24	43488	19149.93
mmu-mir-100	MI0000692 9(41339507-41339587)	49	44593	9563.916	81	3186	1402.954
mmu-let-7i	MI0000138 10(122422695-122422780)	50	40272	8637.185	42	18009	7930.259

for known miRNA discovery, but was less stringent when mapping for novel miRNAs by allowing 2 mismatches.

This analysis was the most extensive: known and novel miRNAs were grouped based on the amount of criteria met through the analysis. The grouping criteria are explained in Table 11 and show that ACGT101-miR is able to report when isomiRs more closely match a miRNA previously identified in another species. The higher stringency noted above, coupled with the elaborate grouping process allowed for easier detection of isomiRs, which are often masked by the relatively simpler grouping processes of other programs.

The top 50 known miRNAs for the PBS mock-infected and HSV-1 infected samples are presented in Tables 12 and 13, respectively. Additionally, the top 50 novel miRNA predictions are presented in Tables 14 and 15 for the PBS and HSV-1 samples, respectively. LC Sciences also compiled a table highlighting differential expression between the two samples, which included both known and novel miRNAs (Table 16). In summary, a total of 354 known and 444 predicted miRNAs were found in the mock-infected sample, and 248 known and 339 predicted miRNAs were found in the HSV-1 infected sample. All miRNAs were represented by at least 3 reads per sample.

3.1.3.4 COMPARISON OF SECONDARY ANALYSIS PROGRAMS

In order to validate the accuracy of the results from the in-house protocol using Galaxy and miRanalyzer, the programs were compared to each other in a variety of ways. The first was to compare the amount of useable reads for each comparison (Figure 10A). Useable reads comprise the reads left after pre-processing that can be used for further downstream analysis.

Table 11. Explanation of groupings by ACGT101-miR for known and novel miRNAs.

Group	MiRNA Type	Description
gp1a	Known/novel	Mapped to known mouse miRNAs in miRBase. If novel, the miRNA is a previously unknown isomiR.
gp1b	Known	Mapped to known mammalian miRNAs in miRBase and a mouse miRNA equivalent is identified. However, the specific miRNA more closely matches an isomiR of another species.
gp1c	Novel	Mapped to known mouse or other mammalian miRNAs in miRBase, but the specific miRNA does not map to the annotated chromosomal location in the mouse genome. These mature miRNAs belong to novel pre-miRs.
gp2	Novel	Mapped to known mammalian miRNAs in miRBase, but there are currently no known mouse equivalents. However, these miRNAs successfully map to the mouse genome and are predicted to properly form hairpins.
gp3	Novel	Mapped to known mammalian miRNAs in miRBase only.
gp4	Novel	Mapped to the mouse genome only, but are also predicted to properly form hairpins.

Table 12. List of the top 50 known miRNAs found in the mock-infected sample by ACGT101-miR v3.5. The mature miRNA and miRNA gene with the closest match (including species specificity) are given. The group column classifies the criteria met for the given miRNA: see Table 11 for information pertaining to group classifications. Finally, a count of the reads is given for each specific miRNA. The copy number is the sum total of reads for all isomiRs matching either the 5p or 3p version of a miRNA.

Mature MiRNA	MiRNA Gene	Group	Copy # (All IsomiRs in either 5p or 3p)
miR-29a-3p	mmu-mir-29a	gp1a	102114
miR-128-3p	mmu-mir-128-1	gp1a	25300
miR-153-3p	mmu-mir-153	gp1a	21845
miR-29c-3p	mmu-mir-29c	gp1a	15039
miR-27b-3p	mmu-mir-27b	gp1a	14134
miR-125b-5p	mmu-mir-125b-1	gp1a	10836
miR-27a-3p	cfa-mir-27a	gp1b	10239
miR-30a-5p	mmu-mir-30a	gp1a	7574
miR-129-3p	mmu-mir-129-2	gp1a	7050
miR-148a-3p	mmu-mir-148a	gp1a	5946
miR-7b-5p	mmu-let-7b	gp1a	5362
miR-30b-5p	mmu-mir-30c-2	gp1a	4813
miR-195-5p	mmu-mir-195	gp1a	4438
miR-331-3p	mmu-mir-331	gp1a	3993
miR-193b-3p	cfa-mir-193b	gp1b	3951
miR-410-3p	cfa-mir-410	gp1b	3919
miR-151-5p	mmu-mir-151	gp1a	3777
miR-32-5p	mmu-mir-32	gp1a	3734
miR-23a-3p	mmu-mir-23a	gp1a	3570
miR-30d-5p	mmu-mir-30d	gp1a	3292
miR-495-3p	mmu-mir-495	gp1a	2928
miR-340-5p	mmu-mir-340	gp1a	2888
miR-1249-3p	mmu-mir-1249	gp1a	2765
miR-137-3p	mmu-mir-137	gp1a	2267
miR-212-3p	cfa-mir-212	gp1b	2051
miR-376a-3p	mmu-mir-376a	gp1a	1994
miR-101-3p	mmu-mir-101b	gp1a	1992
miR-7i-3p	mmu-let-7i	gp1a	1973
miR-433-3p	cfa-mir-433	gp1b	1946
miR-30e-5p	mmu-mir-30e	gp1a	1932
miR-187-3p	mmu-mir-187	gp1a	1826
miR-143-3p	mmu-mir-143	gp1a	1800
miR-652-3p	cfa-mir-652	gp1b	1784
miR-148b-3p	mmu-mir-148b	gp1a	1564

miR-25-3p	cfa-mir-25	gp1b	1497
miR-136-3p	mmu-mir-136	gp1a	1479
miR-301a-3p	mmu-mir-301a	gp1a	1471
miR-598-3p	mmu-mir-598	gp1a	1443
miR-541-5p	mmu-mir-541	gp1a	1415
miR-3102-3p	mmu-mir-3102	gp1a	1403
miR-329-3p	mmu-mir-329	gp1a	1387
miR-487b-3p	cfa-mir-487b	gp1b	1347
miR-9-5p	mmu-mir-9-1; mmu-mir-9-2	gp1a	1279
miR-19b-3p	cfa-mir-19b-2	gp1b	1253
miR-484-5p	mmu-mir-484	gp1a	1113
miR-33-5p	mmu-mir-33	gp1a	1047
miR-129-3p	mmu-mir-129-1	gp1a	1041
miR-29b-3p	mmu-mir-29b-1	gp1a	1015
miR-142-3p	mmu-mir-142	gp1a	990
miR-1-3p	cfa-mir-1-2	gp1b	930

Table 13. List of the top 50 known miRNAs found in the HSV-1 infected sample by ACGT101-miR v3.5. The mature miRNA and miRNA gene with the closest match (including species specificity) are given. The group column classifies the criteria met for the given miRNA: see Table 11 for information pertaining to group classifications. Finally, a count of the reads is given for each specific miRNA. The copy number is the sum total of reads for all isomiRs matching either the 5p or 3p version of a miRNA.

Mature MiRNA	MiRNA Gene	Group	Copy # (All Isoforms in 5p or 3p)
miR-29a-3p	mmu-mir-29a	gp1a	17712
miR-381-3p	rno-mir-381	gp1b	8874
miR-1-3p	cfa-mir-1-2	gp1b	5725
miR-27b-3p	cfa-mir-27b	gp1b	4467
miR-30a-5p	mmu-mir-30a	gp1a	4411
miR-148a-3p	mmu-mir-148a	gp1a	4399
miR-340-5p	mmu-mir-340	gp1a	4318
miR-30d-5p	mmu-mir-30d	gp1a	4073
miR-7b-5p	mmu-let-7b	gp1a	4011
miR-153-3p	mmu-mir-153	gp1a	3722
miR-128-3p	mmu-mir-128-1	gp1a	3575
miR-182-5p	mmu-mir-182	gp1a	2996
miR-30e-5p	mmu-mir-30e	gp1a	2712
miR-148b-3p	mmu-mir-148b	gp1a	2273
miR-7f-5p	mmu-let-7f-2	gp1a	1843
miR-27a-3p	cfa-mir-27a	gp1b	1805
miR-101-3p	mmu-mir-101b	gp1a	1235
miR-410-3p	cfa-mir-410	gp1b	1133
miR-152-3p	cfa-mir-152	gp1b	1104
miR-29c-3p	mmu-mir-29c	gp1a	1057
miR-142-3p	mmu-mir-142	gp1a	1046
miR-143-3p	mmu-mir-143	gp1a	1033
miR-23b-3p	bta-mir-23b	gp1b	995
miR-137-3p	mmu-mir-137	gp1a	926
miR-872-5p	mmu-mir-872	gp1a	887
miR-376a-3p	mmu-mir-376a	gp1a	808
miR-9-5p	mmu-mir-9-1, mmu-mir-9-2	gp1a	658
miR-151-5p	mmu-mir-151	gp1a	637
miR-195-5p	mmu-mir-195	gp1a	617
miR-7a-5p	mmu-let-7a-1	gp1a	571
miR-7a-5p	mmu-let-7c-2	gp1a	566
miR-7a-5p	mmu-let-7a-2	gp1a	480
miR-186-5p	mmu-mir-186	gp1a	465

miR-129-3p	mmu-mir-129-2	gp1a	443
miR-7c-5p	mmu-let-7c-1	gp1a	436
miR-200c-3p	cfa-mir-200c	gp1b	402
miR-203-3p	mmu-mir-203	gp1a	372
miR-30b-5p	mmu-mir-30c-2	gp1a	365
miR-200a-3p	mmu-mir-200a	gp1a	349
miR-744-5p	mmu-mir-744	gp1a	316
miR-26a-5p	mmu-mir-26a-1	gp1a	302
miR-33-5p	mmu-mir-33	gp1a	299
miR-30c-3p	mmu-mir-30c-2	gp1a	289
miR-103-3p	mmu-mir-103-2	gp1a	278
miR-125b-5p	mmu-mir-125b-1	gp1a	273
miR-344d-3p	mmu-mir-344d-2	gp1a	267
miR-101-3p	mmu-mir-101a	gp1a	265
miR-328-3p	cfa-mir-328	gp1b	264
miR-7g-5p	mmu-let-7g	gp1a	253
miR-222-3p	mmu-mir-222	gp1a	248

Table 14. List of the top 50 novel miRNAs found in the mock-infected sample by ACGT101-miR v3.5. The miRNA sequence column lists the highest isomiR found for a given miRNA. The miRNA ID(s) column indicates whether a novel miRNA matches one or a number of known miRNAs. The group column classifies the criteria met for the given miRNA: see Table 11 for information pertaining to group classifications. Finally, a count of the reads is given for each specific miRNA. The copy number is the sum total of reads for all isomiRs matching either the 5p or 3p version of a miRNA. All miRNAs are represented by at least 3 reads.

MiRNA Sequence	MiRNA ID(s)	Group	Copy # (All IsomiRs in either 5p or 3p)
AGACTGTGAATCTGACATC		gp4	2856
AATGACACGATCACTCCCGTTGATC	oan-mir-425	gp3	1318
TGTAACAGCAACTCCATGTGGATC	hsa-mir-194-1	gp3	800
TGATTGGAAGACACTCTGCAATC		gp4	701
ACTCTGTAGACCAGGCATC		gp4	535
TGAGGTAGTAGTTTGTGCTG	mmu-let-7i	gp1a	461
CTATGAGGCTAAGAGAT		gp4	426
TTAGTAGATTCTGAATATC		gp4	250
CGTGCAAAGGTAGCATAATCATC	hsa-mir-1973	gp3	250
CACGTTGGGCGCCAATC	ssc-mir-4335	gp3	248
TAATACTGTCTGGTAATGCCGATC	cfa-mir-429	gp3	222
TCCCACATGGTCTAGCGGTTAGGATC		gp4	219
TGGAAGACTAGTGATTTTGTTGATC	oan-mir-7-2	gp3	205
TGAGGCAGTAGATTGTATAGTTATC	hsa-mir-1827	gp3	179
TTTTTCATTATTGCTCCTGACATC	hsa-mir-335	gp3	165
TAGATTCTGAATAAAATC		gp4	164
GCCAATCTAGCCAATC		gp4	150
TGGAATGTAAAGAAGTATGTATC	hsa-mir-1-1	gp1c	148
TACAGTATAGATGATGTACTATC	oan-mir-144	gp3	138
CTCAGTGGTAGAGCATGCATC		gp4	132
TTTGGCACTAGCACATTTTTGCATC	mdo-mir-96	gp3	115
CACTAGATTGTGAGCTGCTGGAATC	cfa-mir-28	gp3	105
TTAAAGCATCTGGCCATC		gp4	96
GAGAACTTTGAAGGCATC		gp4	88
ATCTCGCTGGGGCCTCCAATC	mmu-mir-720	gp1a	86
CCCTGAACTAGGGGTCTGGAGTATC	cfa-mir-345	gp3	82
AATAACAAAGCAAAGCATC		gp4	75
GGTATCCTGACCGTGCAAAGGTATC	hsa-mir-1973	gp3	70
TAAAGTTGTAATTACTCTTGTT		gp4	60
GAGAAATCTCTACTAAGATC		gp4	60
GTTTCTGTTGTTGAAATC		gp4	58
TAATCCATTGGTCTTAATC		gp4	54

TCTTATTTCTACCAATC		gp4	51
TGAGGTAGTAGATTGTATAGTTATT	oan-let-7f-1	gp3	48
CAAAGTGCTTACAGTGCAGGTATC	rno-mir-17-2	gp3	47
CACTTTGTAGACCAGGCATC		gp4	45
TACTGACTATGAAGACCATC		gp4	45
GTTCCCAGCACCCATC		gp4	44
CCCCCATCGCCTCTCCCATC		gp4	43
TTGGGTGCTGGTGGATC		gp4	42
GCTCGGCGGGGTCCCCGCGTCATC		gp4	42
TAGCAGCACGTAAATATTGGTC	lca-mir-16	gp3	41
TAACACTGTCTGGTAAAGATGGATC	hsa-mir-141	gp3	40
AGATGTAGAACTCTCAGCATC		gp4	39
TGTAAACATCCTACACTCAGCTATA	bta-mir-30b	gp3	32
AACTCACTCTGTAGACCAGGCATC		gp4	31
GTGACTGAGAGAAATC		gp4	30
CATTACTCTGGTGTTGTAAACATC	ssc-mir-30d	gp3	29
GTAGATTCTCCTTCTATGAGTACATC	bta-mir-376a	gp3	29
TGGGGATTGTGGGTATC		gp4	28

Table 15. List of the top 50 novel miRNAs found in the HSV-1 infected sample by ACGT101-miR v3.5. The miRNA sequence column lists the highest isomiR found for a given miRNA. The miRNA ID(s) column indicates whether a novel miRNA matches one or a number of known miRNAs. The group column classifies the criteria met for the given miRNA: see Table 11 for information pertaining to group classifications. Finally, a count of the reads is given for each specific miRNA. The copy number is the sum total of reads for all isomiRs matching either the 5p or 3p version of a miRNA. All miRNAs are represented by at least 3 reads.

Sequence	MiRNA ID(s)	Group	Copy # (All IsomiRs in either 5p or 3p)
TATTGCACATTACTAAGTTGCATC	hsa-mir-1827, ppy-mir-1827	gp3	2976
TTCACAGTGGCTAAGTTCCGCATC		gp4	1510
TGTAAACATCCCCGACTGGAATC		gp4	1429
AATATAACACAGATGGCCTGTATC		gp4	1156
AACTGGCCCACAAAGTCCCGCTATC	hsa-mir-1-1, cfa-mir-1-2, ptr-mir-1-1, mml-mir-1-1, cfa-mir-1-1	gp1c	1007
TTATAAAGCAATGAGACTGATTATC	cfa-mir-130a	gp3	695
ATCGTAGAGGAAAATCCACGTTATC	hsa-mir-194-1, oan-mir-194-1, ppy-mir-194, bta-mir-194-1, cfa-mir-194, oan-mir-194-2, mml-mir-194-2	gp3	622
CTGCGCAAGCTACTGCCTTGCTATC		gp4	486
ATCATGATGGGCTCCTCGGTGTATC		gp4	450
TCAGTGCATCACAGAACTTTGTATC		gp4	279
AAGGGATTCTGATGTTGGTCACAATC	oan-let-7f-1	gp3	268
AAACAAACATGGTGCACTTCTTATC	cfa-mir-429, bta-mir-429, mdo-mir-429	gp3	262
AACACACCCAGCTAACCTTTTTATC	hsa-mir-138-1, oan-mir-138-1, bta-mir-138-1	gp3	244
AATGACACGATCACTCCCGTTGATC		gp4	242
TACGTCATCGTCGTCATCGTTATC		gp4	237
CAGTGCAATAGTATTGTCAAAGCATC		gp4	202
CATTGCACTTGTCTCGGTCTGAATC		gp4	177
AGACTGTGAATCTGACATC	hsa-mir-141	gp3	155
TGCGGGGCTAGGGCTAACAGCAATC	oan-mir-425	gp3	141
TGTAACAGCAACTCCATGTGGATC	hsa-mir-1973	gp3	140
TGAACTATTGCAGTAGCCTCCTATC		gp4	125
AAGCCCTTACCCCAAAAAGTATC	bta-mir-495, cfa-mir-495	gp3	119
AGCTACATCTGGCTACTGGGTATC		gp4	118
TAACAGTCTCCAGTCACGGCCAATC		gp4	110

AATCGTACAGGGTCATCCACTTATC		gp4	107
TGTGCAAATCCATGCAAAACTGAATC		gp4	104
TTTATTGAGCACCTCCTATCAATC	oan-mir-7-2, oan-mir-7-3	gp3	98
TTCACAAAGCCCATACACTTTCATC		gp4	87
TGATTGGAAGACACTCTGCAATC		gp4	76
AAGGTTACTTGTTAGTTCAGGATC	cfa-mir-28	gp3	70
TATTGCACTTGTCCCGGCCTGTATC		gp4	69
TATTCATTTACTCCCCAGCCTACATC		gp4	61
TATGGCTTTTCATTCCTATGTGATC		gp4	54
TTTGGCAATGGTAGAACTCACATC		gp4	54
GATATAACCACTGCCAGACTGAATC		gp4	46
TCCTGTACTGAGCTGCCCCGAGATC	lca-mir-19b-2, cfa-mir-19b-2, cfa-mir-19b-1	gp1c	42
TCGTGTCTTGTGTTGCAGCCGGTATC	hsa-mir-1973	gp3	40
TTGGTCCCCTTCAACCAGCTGTATC	hsa-mir-335, ptr-mir-335, ppy-mir-335, mml-mir-335, bta-mir-335	gp3	37
CAAAGAATTCTCCTTTTGGGCTATC		gp4	35
CTGGGAGAAGGCTGTTTACTCTATC		gp4	35
TATGTAACACGGTCCACTAACCATC		gp4	33
TGGAATGTAAAGAAGTATGTACATC		gp4	32
TCATTCCTAGAAATTGTTCACAATC		gp4	31
AAAAGCTGGGTTGAGAGGGCGAATC		gp4	30
CTTTCAGTCGGATGTTTGCATC		gp4	28
TCTCTGGGCCTGTGTCTTAGGATC		gp4	28
CAACAAATCACAGTCTGCCATATC		gp4	28
TTAGTAGATTCTGAATATC		gp4	26
TCATCGTCTCAAATGAGTCTTCATC	cfa-mir-345	gp3	25
TAGCAGCACATCATGGTTTACATC	hsa-mir-182, mml-mir-182, ppy-mir-182, ptr-mir-182, oan-mir-182, bta-mir-182, cfa-mir-182, mdo-mir-182, rno-mir-182	gp3	24

Table 16. Differential expression results using ACGT101-miR between the mock-infected sample and the HSV-1 infected sample for known and novel miRNAs. Both the sequence and miRNA name (if known) are given. Total read counts are compared between PBS and HSV-1 samples, followed by the fold change and $\log_2$ fold change. The $\log_2$ fold change allows for actual positive and negative expression values whereas a down-regulation with normal fold change is represented as a value less than 1. The 25 largest changes for both up- and down-regulated miRNAs are displayed.

Sequence	MiRNA Name	Copy # (PBS)	Copy # (HSV-1)	Fold Change	$\log_2$ Fold Change
	Up-regulated				
TGAGGCAGTAGATTGTATAGTTATC	hsa-mir-1827, ppy-mir-1827	101	1752	17.346 53465	4.116576
AGCAGCATTGTACAGGGCTATGAA	mmu-mir-103-2, hsa-mir-103-2, ssc-mir-103-2, ptr-mir-103-2, cfa-mir-103-2, mdo-mir-103-2, bta-mir-103-2, cfa-mir-103-1, cfa-mir-107	5	59	11.8	3.560715
TAGATTCTGAATAAAATC		164	1510	9.2073 17073	3.202781
TAACACTGTCTGGTAACGATGATC	mmu-mir-200a, hsa-mir-200a, rno-mir-200a, bta-mir-200a, mml-mir-200a, ppy-mir-200a, ptr-mir-200a	50	345	6.9	2.786596
TGGAATGTAAAGAAGTATGTATC	hsa-mir-1-1, cfa-mir-1-2, ptr-mir-1-1, mml-mir-1-1, cfa-mir-1-1	148	991	6.6959 45946	2.743288
GTTTCTGTTGTTGAAATC		58	279	4.8103 44828	2.26614
TGAGGTAGTAGATTGTATAGTTATT	oan-let-7f-1	48	226	4.7083 33333	2.235216
CTAGTGATTTTGTTGTTATC	hsa-mir-7-2	3	14	4.6666 66667	2.222392
TAACACTGTCTGGTAAAGATGGATC	hsa-mir-141	40	155	3.875	1.954196
TTTGGCAATGGTAGAACTCACAAT	hsa-mir-182, mml-mir-182, ppy-mir-182, ptr-mir-182, oan-mir-182, bta-mir-182, cfa-mir-182, mdo-mir-182, rno-mir-182	3	11	3.6666 66667	1.874469
TGGAATGTAAGGAAGTGTGTGATC	ssc-mir-206, cfa-mir-206	6	21	3.5	1.807355
GAAAATGATGATTGGAATC		3	10	3.3333 33333	1.736966
TAGCTTATCAGACTGATGTTGAAT	mmu-mir-21, hsa-mir-21, ssc-mir-21, bta-mir-21, mdo-mir-21, cfa-mir-21	8	26	3.25	1.70044
CAGAGTGGTTGTACAAGCATC		3	9	3	1.584963
CACTTTGTAGACCAGGCATC		21	57	2.7142 85714	1.440573
GAGAACTTTGAAGGCATC		88	202	2.2954	1.19878

				54545	
TGCTTAGATGGATAAATC		5	11	2.2	1.137504
TTCAAGTAATCCAGGATAGGC	mmu-mir-26a-1, hsa-mir-26a-1, rno-mir-26a, bta-mir-26a-1, mne-mir-26a, cfa-mir-26a-1	3	6	2	1
TGGAAGACTAGTGATTTTGTTGTTAT	mmu-mir-7a-2, cfa-mir-7-2	3	6	2	1
AGCTCGACTCATGGTTTGAATC	mmu-mir-434	12	23	1.916666667	0.938599
ATGGATAAGGCACTGATC		3	5	1.666666667	0.736966
TGATATGTTTGATATATTAGGTC	mmu-mir-190, hsa-mir-190, rno-mir-190, mdo-mir-190a, oan-mir-190a	3	5	1.666666667	0.736966
CCAGGACAGCCAGGGCATC		5	8	1.6	0.678072
TTATAAAGCAATGAGACTGATTATC	mmu-mir-340, hsa-mir-340, ptr-mir-340, rno-mir-340, ppy-mir-340, bta-mir-340, mml-mir-340, ssc-mir-340, cfa-mir-340	2843	4297	1.511431586	0.595916
TCAGTGCATCACAGAACTTTGTATC	mmu-mir-148b, hsa-mir-148b, rno-mir-148b, ptr-mir-148b, bta-mir-148b, ssc-mir-148b, cfa-mir-148b	1552	2244	1.445876289	0.531944
	Down-regulated				
TAGCACCATTTGAAATCAGTGTTC	mmu-mir-29b-1, hsa-mir-29b-1, rno-mir-29b-1, mdo-mir-29b, ssc-mir-29b, cfa-mir-29b-1	40	3	0.075	-3.736965594
CATTGCACTTGTCTCGGTCTGAATC	cfa-mir-25	1168	86	0.07363	-3.763559804
TGTCACTCGGCTCGGCCCACTATC	mmu-mir-668, hsa-mir-668, rno-mir-668, ptr-mir-668, ppy-mir-668	221	15	0.067873	-3.881011964
TGGTTTACCGTCCCACATACATC	mmu-mir-299, hsa-mir-299, ssc-mir-299, ptr-mir-299, ppy-mir-299, bta-mir-299, mml-mir-299	237	16	0.067511	-3.888743249
CTCACTGAACAATGAATGCATC	mmu-mir-181b-1, hsa-mir-181b-1, rno-mir-181b-1, mml-mir-181b-1, ppy-mir-181b, ptr-mir-181b-1, ssc-mir-181b-1, mdo-mir-181b, cfa-mir-181b-1	136	9	0.066176	-3.91753784
TATTCATTTACTCCCCAGCCTACATC	rno-mir-664-2, rno-mir-664-1	443	29	0.065463	-3.933181893
AAGCCCTTACCCCAAAAAGCATTC	mmu-mir-129-2, rno-mir-129-2, bta-mir-129-2, mml-mir-129-2, ssc-mir-129, cfa-mir-129-2	47	3	0.06383	-3.969626351
ATCTCGCTGGGGCCTCCAATC	mmu-mir-720	54	3	0.055556	-4.169925001
TTGGTCCCCTTCAACCAGCTGTATC	mmu-mir-133a-2, mmu-mir-133a-1, bta-mir-133a-1, cfa-mir-133c, mdo-mir-133a	316	17	0.053797	-4.216317907
AACACACCCAGCTAACCTTTTTATC	mmu-mir-329	1340	72	0.053731	-4.218092284

TTCACAAAGCCCATACACTTTCATC	mmu-mir-350, rno-mir-350, cfa-mir-350	525	28	0.053333	-4.22881869
AATCGTACAGGGTCATCCACTTATC	cfa-mir-487b	607	32	0.052718	-4.245552706
TTTGTGACCTGGTCCACTAACCATC	mmu-mir-758, rno-mir-758, cfa-mir-758, ssc-mir-758	96	5	0.052083	-4.263034406
TAACAGTCTCCAGTCACGGCCAATC	cfa-mir-212	643	33	0.051322	-4.284280808
ACATAACATACACACACACGTATC	mmu-mir-669a-1, mmu-mir-669a-2, mmu-mir-669a-4, mmu-mir-669o	81	4	0.049383	-4.339850003
TTTATTGAGCACCTCCTATCAATC	mmu-mir-325, cfa-mir-325, ssc-mir-325	529	26	0.049149	-4.346684194
TATTGCACTTGTCCCGGCCTGTATC	cfa-mir-92a-1	497	24	0.04829	-4.372139541
TAACAGTCTACAGCCATGG	mmu-mir-132, hsa-mir-132, rno-mir-132, bta-mir-132, mdo-mir-132, cfa-mir-132	65	3	0.046154	-4.437405312
TATTGCACATTACTAAGTTGCATC	mmu-mir-32, ssc-mir-32, rno-mir-32, cfa-mir-32, mdo-mir-32	3620	166	0.045856	-4.446734551
TCCTGTACTGAGCTGCCCCGAGATC	mmu-mir-486	343	15	0.043732	-4.515174171
CAACAAATCACAGTCTGCCATATC	mmu-mir-7a-1, hsa-mir-7-1, ptr-mir-7-1, mml-mir-7-1, ssc-mir-7, sla-mir-7, bta-mir-7-2, rno-mir-7a-1, mdo-mir-7, cfa-mir-7-1	255	11	0.043137	-4.534921818
AGACCTACTTATCTACCAACATC	mmu-mir-1839, bta-mir-1839, cfa-mir-1839	220	9	0.040909	-4.611434712
ATCATGATGGGCTCCTCGGTGTATC	cfa-mir-433	1673	46	0.027496	-5.184659774
CTGCGCAAGCTACTGCCTTGCTATC	mmu-let-7i, hsa-let-7i, ptr-let-7i	1922	13	0.006764	-7.207952903
AACTGGCCCACAAAGTCCCGCTATC	cfa-mir-193b	2912	16	0.005495	-7.50779464

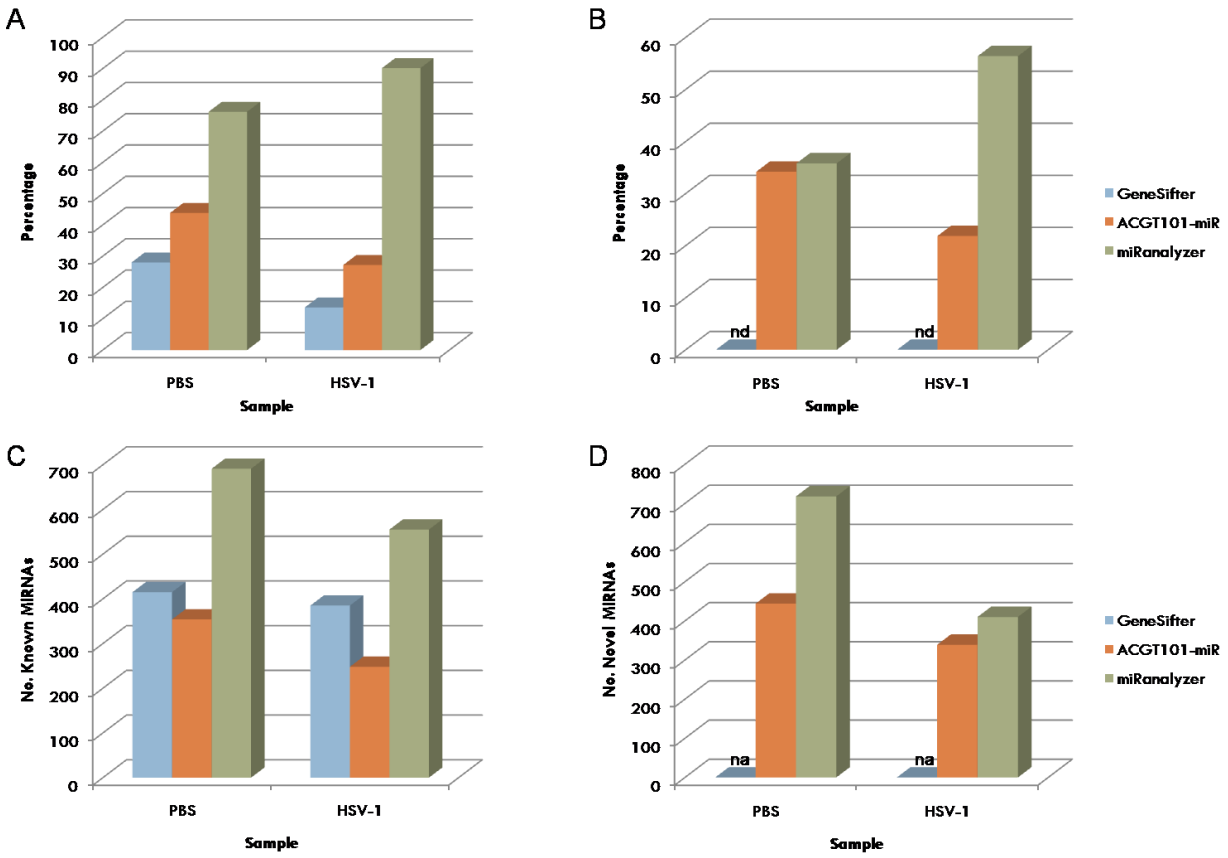


Figure 10. Various comparisons between the 3 bioinformatic analyses used to identify known and novel miRNAs from PBS mock-infected and HSV-1 infected samples. A) The percentage of useable reads left over after filtering out poor quality/small reads during pre-processing. B) The percentage of reads mapped to either miRBase to identify known miRNAs, or to the mouse genome to discover potential novel miRNAs. C) The number of known miRNAs identified by each analysis. D) The number of novel miRNAs identified by each analysis. nd: no data available. na: not applicable - the GeneSifter analysis did not predict novel miRNAs.

These reads have been filtered based on quality, size and other factors. A comparison of the results shows that the miRanalyzer analysis (via Galaxy) retains the most reads after filtering out the poor data (76% and 90% of the reads are kept for the PBS mock-infected and HSV-1 infected samples, respectively). In contrast, ACGT101-miR and GeneSifter both remove more than 50% of their reads during the pre-processing steps. The amount of mapped reads were also compared (Figure 10B). These reads were recognized as either known or potential novel miRNAs when mapped to miRBase or the mouse genome, respectively. While the percentage of mapped reads in the mock-infected sample is quite similar for both ACGT101-miR and miRanalyzer, there is a large difference in the number of mapped reads in the HSV-1 infected sample. However, given that only 44% of the mock-infected reads were useable by ACGT101-miR (32% less useable reads than miRanalyzer), overall miRanalyzer was able to map more reads for this sample as well.

Next, the amount of data generated from each analysis was observed, and it was found that miRanalyzer identified the most known and novel miRNAs within the samples (Figure 10C and D, respectively). GeneSifter discovered slightly more known miRNAs than ACGT101-miR but it is important to note that GeneSifter only reported to the pre-miR level of the transcripts and therefore reported several hits for a given miRNA, since one miRNA may have come from different pre-miRs based on the genome location. Also, GeneSifter doesn't make the distinction between the -3p or -5p version of a miRNA since they are both generated from the same pre-miR, so the data is harder to correlate with the other analyses in this case. When comparing the top miRNAs of each analysis, the similarities and differences between the 3 analyses can be observed quite well. Table 17 shows the top 10 known miRNAs as identified

Table 17. Comparison of the top 10 mature miRNAs as predicted by miRanalyzer, for both the mock-infected (A) and HSV-1 infected (B) samples. For miRanalyzer, () brackets denote ambiguous reads that may match other miRNAs as well. For GeneSifter, multiple positions and their corresponding number of reads occur when a given miRNA is found on multiple pre-miRs. For ACGT101-miR, [] brackets denote reads classified as novel because they match a non-mouse, mammalian isomiR but did not successfully map to the annotated genome. Lastly, {} brackets denote a known, non-mouse, mammalian isomiR that successfully maps to the mouse genome.

A

miRNA	miRanalyzer		GeneSifter		ACGT101-miR	
	Position	# of Reads	Position	# of Reads	Position	# of Reads
mmu-miR-143-3p	1	326033	2	348481	32	1800
mmu-miR-30c-5p	2	195363 (+334)	15,16	226247,225784	12	4813
mmu-miR-378a-3p	3	190264 (+8988)	14	252603	80	316 [+21]
mmu-miR-30a-5p	4	189651	1	363550	8	7574
mmu-miR-124-3p	5	184968	10,11,12	261594,261585,261572	61	574
mmu-miR-29a-3p	6	178042	7	268998	1	102114
mmu-miR-9-5p	7	174988	4,5,6	269380,269379,269361	43	1279
mmu-miR-127-3p	8	155248	19	197226	223	16
mmu-miR-101a-3p	9	104071 (+40)	9	264008	58	661
mmu-miR-30e-5p	10	80087	3	320739	30	1932

B

miRNA	miRanalyzer		GeneSifter		ACGT101-miR	
	Position	# of Reads	Position	# of Reads	Position	# of Reads
mmu-let-7f-5p	1	471739	3,4	565957,565443	15, {61}	1843, {185} [+268]
mmu-miR-143-3p	2	211114	10	197056	22	1033
mmu-let-7g-5p	3	110246	12	115253	49	253
mmu-miR-30e-5p	4	109641	8	203045	13	2712
mmu-let-7c-5p	5	106675 (+519)	6,7	207363,203474	31, 35	566,436
mmu-miR-30a-5p	6	90885	5	227524	5	4411
mmu-let-7a-5p	7	88357 (+519)	1,2	667807,647766	30, 32	571,480 [+3]
mmu-miR-9-5p	8	76174	13,14,15	115002,115001,114986	27	658
mmu-miR-26a-5p	9	61833	19,20	62097,62034	41	302
mmu-miR-127-3p	10	51799	21	56609	152	18

by miRanalyzer, which correlate rather well with GeneSifter's results. Again, GeneSifter only identifies miRNAs to the pre-miR level, so one mature miRNA can appear at multiple positions within its list due to the miRNA matching multiple pre-miRs at different chromosomal locations. However, when comparing to ACGT101-miR, we find that almost all the top miRNAs predicted by miRanalyzer and GeneSifter are found much lower on the list according to ACGT101-miR, and at lower quantities as well. The inverse it true if we look at the top miRNAs predicted by ACGT101-miR (Table 18): these miRNAs are predicted at a lower position by both miRanalyzer and GeneSifter. However, the copy numbers match more closely with ACGT101-miR in this case.

Given the discrepancies between analyses, it is worth investigating whether a particular analysis is identifying miRNAs that other analyses are not. Given that miRanalyzer found more miRNAs than the other two programs, there will inevitably be some miRNAs discovered by miRanalyzer that ACGT101-miR and/or GeneSifter did not identify. It is also possible that the opposite is true, where GeneSifter and/or ACGT101-miR uncovered some miRNAs that miRanalyzer may have missed. Figure 11 shows that this is in fact the case. At the pre-miR level, the majority of the miRNAs are discovered by all three programs (Figure 11A and B). However, there are still some known miRNAs that were identified exclusively by one of the programs, as well as some identified by only two of the three programs. When focusing instead on the mature miRNAs (Figure 11C and D), we again find that there are a small number of miRNAs found exclusively by one program.

Table 18. Comparison of the top 10 mature miRNAs as predicted by ACGT101-miR, for both the mock-infected (A) and HSV-1 infected (B) samples. Multiple positions and their corresponding number of reads occur when a given miRNA is found on multiple pre-miRs.

A

miRNA	ACGT101-miR		miRanalyzer		GeneSifter	
	Position	# of Reads	Position	# of Reads	Position	# of Reads
miR-29a-3p	1	102114	6	178042	7	268998
miR-128-3p	2	25300	23	43822	41, 45	50434, 48739
miR-153-3p	3	21845	39	24733	60	30431
miR-29c-3p	4	15039	30	32549	13	258252
miR-27b-3p	5	14134	41	20032	63	29225
miR-125b-5p	6	10836	19	46937	36, 37	52539, 51657
miR-27a-3p	7	10239	74	5885	70	21573
miR-30a-5p	8	7574	4	189651	1	363550
miR-129-3p	9	7050	45, 125	19007, 2270	62, 65	29225, 28250
miR-148a-3p	10	5946	72	6681	94	8171

B

miRNA	ACGT101-miR		miRanalyzer		GeneSifter	
	Position	# of Reads	Position	# of Reads	Position	# of Reads
miR-29a-3p	1	17712	19	22475	29	29981
miR-381-3p	2	8874	44	5187	57	6639
miR-1-3p	3	5725	43	5435	63, 64	5479, 5456
miR-27b-3p	4	4467	47	4372	66	5133
miR-30a-5p	5	4411	6	90885	5	227524
miR-148a-3p	6	4399	45	5062	65	5351
miR-340-5p	7	4318	39	6170	58	6451
miR-30d-5p	8	4073	13	41789	11	129579
miR-7b-5p	9	4011	127	611	102	2029
miR-153-3p	10	3722	41	5694	60	5953

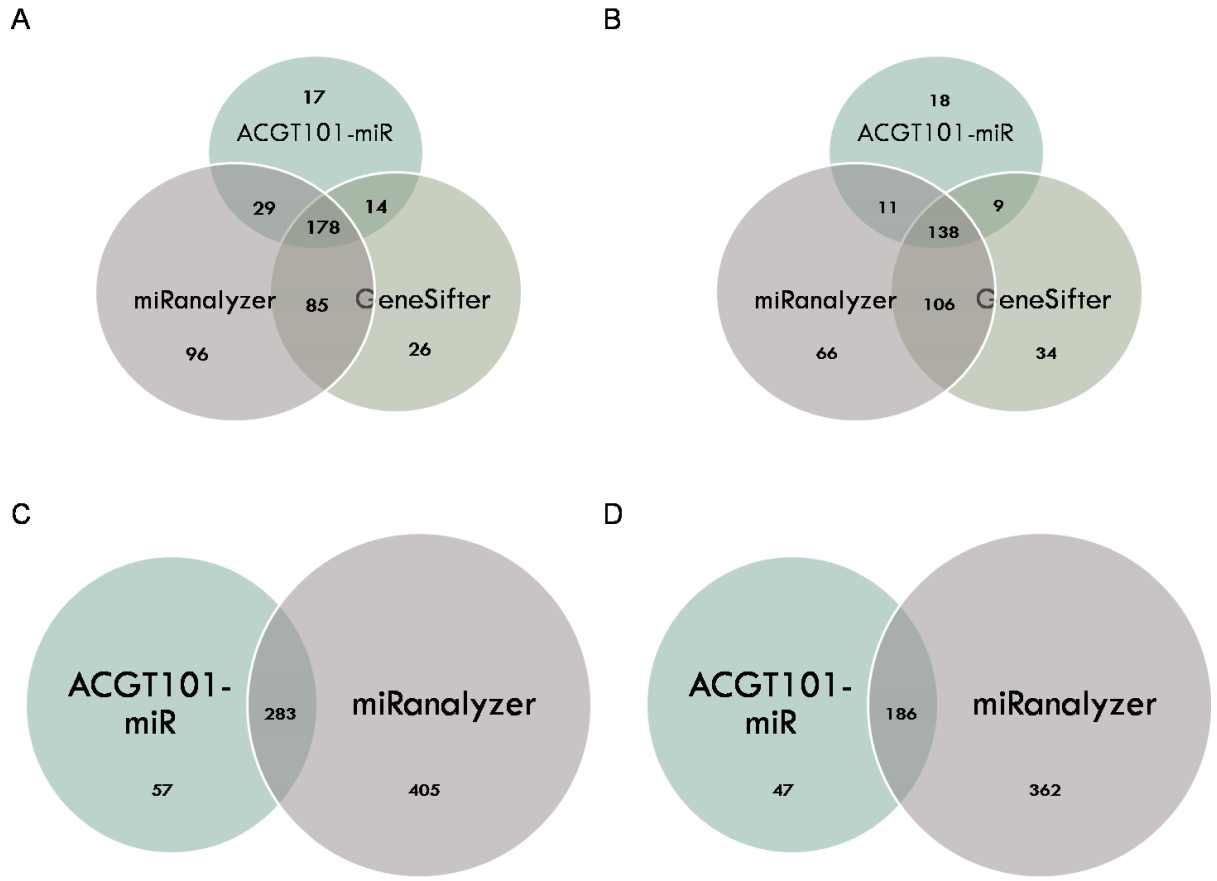


Figure 11. Venn diagrams comparing the identification of known miRNAs by 3 different analysis programs. Known miRNAs identified at the pre-miR level are given by all three programs for both mock-infected (A) and HSV-1 infected (B) samples. Only ACGT101-miR and miRanalyzer are able to identify mature miRNAs for the PBS mock-infected (C) and HSV-1 infected (D) samples.

3.1.3.5 DIFFERENTIAL EXPRESSION BETWEEN MOCK-INFECTED AND HSV-1 INFECTED SAMPLES

Tables 9 and 16 detailed the results of differential expression analyses by miRanalyzer and ACGT101-miR, respectively. Figure 12A and B show that the majority of the identified miRNAs are present in both samples, but there are some miRNAs expressed in only one of the two samples. In both analyses, mock-infection showed expression of more exclusive miRNAs than the HSV-1 infection.

Further, when comparing the miRNAs of each sample against each other (Figure 12C), the majority of the miRNAs identified (82.7% and 57.7% according to miRanalyzer and ACGT101-miR, respectively) were down-regulated and therefore less copies were found in the HSV-1 sample compared to the mock-infected sample. However, it was interesting to find that many miRNAs were instead up-regulated during infection, since host miRNAs are often repressed during infection[93,94]. MiRanalyzer revealed that 15.2% of the known miRNAs discovered were up- regulated, while ACGT101-miR calculated 41.6% of both known and novel miRNAs were up-regulated. Lastly, only a small number of miRNAs (2.1% and 0.7% of the miRNAs in the miRanalyzer and ACGT101-miR analyses, respectively) were left unchanged between samples.

3.1.4 qPCR VERIFICATION OF *IN VIVO* NGS miRNA ANALYSIS

3.1.4.1 TLDA SCREENING AND ANALYSIS OF SMALL RNAS

To validate the screening of small RNAs via NGS technology, the samples were re-analyzed via qPCR using TLDA cards. These pre-loaded 384-well plates contain primers for a sampling of known rodent miRNAs; two cards are available for the screen and the majority are mouse

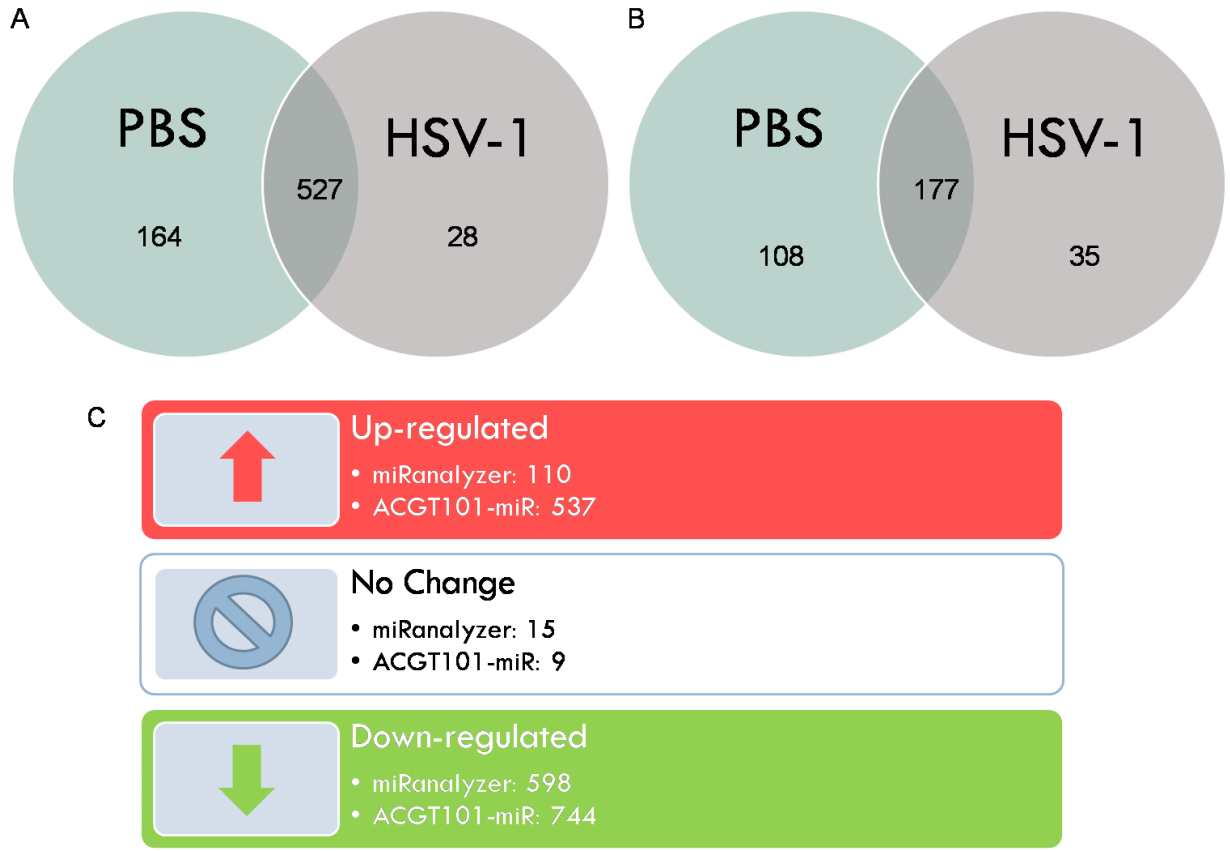


Figure 12. Differential expression between mock-infection and HSV-1 infection in mouse brains. MiRanalyzer (A) and ACGT101-miR (B) show that the majority of the miRNAs identified were found in both samples. There were more miRNAs found exclusively in the PBS mock-infected sample than the HSV-1 infected sample, as shown by both analyses. C) When comparing the expression of the miRNAs, both analyses found that the majority of the miRNAs were down-regulated during HSV-1 infection. A large number of miRNAs were up-regulated, and only a few miRNAs remained unchanged. In A and B, the numbers reflect only the mature miRNAs, whereas in C, ACGT101-miR analyzed each individual isomiR for both mature and novel miRNAs, resulting in higher quantities than previously reported. MiRanalyzer only compared mature miRNAs.

miRNAs. Out of a total of 518 mouse miRNAs, 366 miRNAs were detected in both samples. The samples also contained isomiRs that matched 31 rat miRNAs. The PBS mock-infected sample and HSV-1 infected sample were run separately and the results were compared with the use of Integromics' RealTime StatMiner software. Using a 5 fold change minimum, 58 miRNAs increased in abundance and 5 miRNAs decreased in HSV-1 infected brain tissue versus control. The relatively high threshold was selected in order to be somewhat conservative given the fact that biological replicates were not available for analysis. The results are displayed in Table 19.

3.1.4.2 SELECTION OF POSITIVELY-EXPRESSED CANDIDATE miRNAS FOR TERTIARY ANALYSIS

The top 5 up-regulated miRNAs according to the TLDA data (Table 19), along with their corresponding family members, were chosen to pursue further analysis of possible functional roles for the miRNAs. Interestingly, miR-200c and -141 were the top two over-expressed miRNAs according to the TLDA data, and are part of the same miRNA family cluster. It was observed that the third highest miRNA, miR-183, is a member of the same family as miR-96, along with miR-182 which is also up-regulated, albeit lower on the list (position 32). MiR-155 is not a member of a miRNA family but the fifth miRNA, miR-429, is a member of a miRNA cluster with miR-200a (position 9, with its star sequence at position 7) and miR-200b (position 8). MiR-200a also shares an identical seed sequence as miR-141 while miR-200b, miR-200c and miR-429 match closely as well (Figure 13). MiR-200b, miR-200c and miR-429 all include the same C → U substitution compared to the miR-141/200a seed sequence.

Table 19. Differential expression results between the HSV-1 infected sample and the mock-infected sample for known rodent miRNAs using TLDA technology. The log2 fold change allows for actual positive and negative expression values whereas a down-regulation with normal fold change is represented as a value less than 1.

MiRNA	Fold Change	Log_2 Fold Change
mmu-miR-200c	119.45	6.90
mmu-miR-141	93.83	6.55
mmu-miR-183	57.82	5.85
mmu-miR-155	56.13	5.81
mmu-miR-429	53.12	5.73
mmu-miR-152	49.48	5.63
mmu-miR-200a*	39.83	5.32
mmu-miR-200b	36.26	5.18
mmu-miR-200a	32.34	5.02
mmu-miR-183*	30.72	4.94
mmu-miR-702	21.96	4.46
rno-miR-207	19.49	4.28
mmu-miR-362-3p	18.39	4.20
mmu-miR-218-1*	16.47	4.04
mmu-miR-411*	15.99	4.00
mmu-miR-540-3p	15.97	4.00
mmu-miR-148b	15.14	3.92
mmu-miR-494	14.16	3.82
mmu-miR-15b	13.33	3.74
mmu-miR-331-5p	13.19	3.72
mmu-miR-210	13.14	3.72
mmu-miR-467a	12.04	3.59
mmu-miR-125b-3p	11.97	3.58
mmu-miR-451	10.55	3.40
mmu-miR-673-3p	9.78	3.29
mmu-miR-130a	9.24	3.21
mmu-let-7g*	9.00	3.17
mmu-miR-23b	8.89	3.15
mmu-miR-361	8.21	3.04
mmu-miR-27a*	8.17	3.03
mmu-miR-547	8.14	3.03
mmu-miR-182	8.00	3.00
mmu-miR-450a-5p	7.96	2.99
mmu-miR-699	7.93	2.99
mmu-miR-223	7.69	2.94
mmu-miR-501-5p	7.55	2.92
mmu-miR-339-3p	7.49	2.91
mmu-miR-721	7.42	2.89
mmu-miR-680	7.14	2.84
mmu-miR-26b*	7.09	2.83
mmu-miR-448	7.04	2.82
mmu-miR-425	6.67	2.74
mmu-miR-873	6.51	2.70
mmu-miR-141*	6.51	2.70

mmu-miR-804	6.38	2.67
mmu-miR-17*	6.09	2.61
mmu-miR-431	6.00	2.58
mmu-miR-488	5.97	2.58
mmu-miR-369-3p	5.87	2.55
rno-miR-598-5p	5.80	2.54
rno-miR-463	5.66	2.50
mmu-miR-99b*	5.62	2.49
mmu-miR-136*	5.36	2.42
mmu-miR-715	5.28	2.40
mmu-miR-188-5p	5.18	2.37
mmu-miR-27a	5.13	2.36
rno-miR-345-3p	5.02	2.33
mmu-miR-190b	5.01	2.33
mmu-miR-384-3p	0.10	-3.37
mmu-miR-674	0.14	-2.81
mmu-miR-409-3p	0.16	-2.63
mmu-miR-29b*	0.18	-2.46
mmu-miR-382	0.20	-2.33

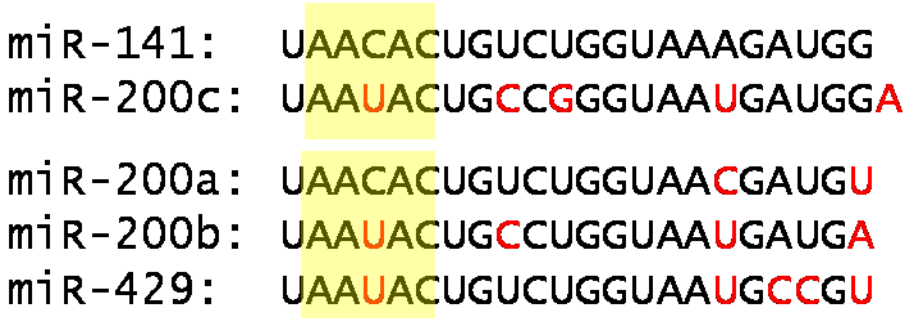


Figure 13. Comparison of the sequences of the miR-200 miRNA family. Nucleotides that differ from miR-141 are highlighted in red. The seed sequence is highlighted by the yellow box.

3.1.4.3. qPCR VERIFICATION OF SELECTED miRNAS

Individual qPCR analysis was performed for a number of the highly related co-transcribed miRNAs and for miR-155. Interestingly, miR-141 was significantly upregulated in a previous profiling experiment to identify miRNAs with altered expression in Vero E6 cells during HSV-1 infection *in vitro*[69]. The only other miRNA to be similarly dysregulated in the same *in vitro* assay was miR-150. Contrary to this result NGS data appeared to show a decrease in abundance of miR-150 in infected brain tissue. Therefore miR-150 was also chosen for validation by qPCR. The results of qPCR analysis are displayed in Figure 14 and the NGS, TLDA and qPCR data are all compared in Table 20. All samples were up-regulated as expected except for miR-96 which was not detactable in either sample. Contrary to the NGS screen, miR-150 correlated instead with the previous *in vitro* screen[69] which showed a positive fold change of this miRNA.

3.2 MiR-150 SHOWS WEAK BINDING POTENTIAL FOR TARGETING THE U_L18/19/20 3' UTR

3.2.1 TARGET PREDICTION

Target prediction using the RNA22[95] program determined a putative binding site for miR-150 within the HSV-1 genome (Figure 15). This site was within the U_L18/19/20 3'UTR, which could block production of several proteins important for completing a mature virion. U_L18 encodes VP23 which helps capsid formation and cleavage/packaging of the genome[97]; U_L19 encodes the major capsid protein[97,98]; and U_L20 encodes a membrane protein essential for exocytosis and envelope development[97].

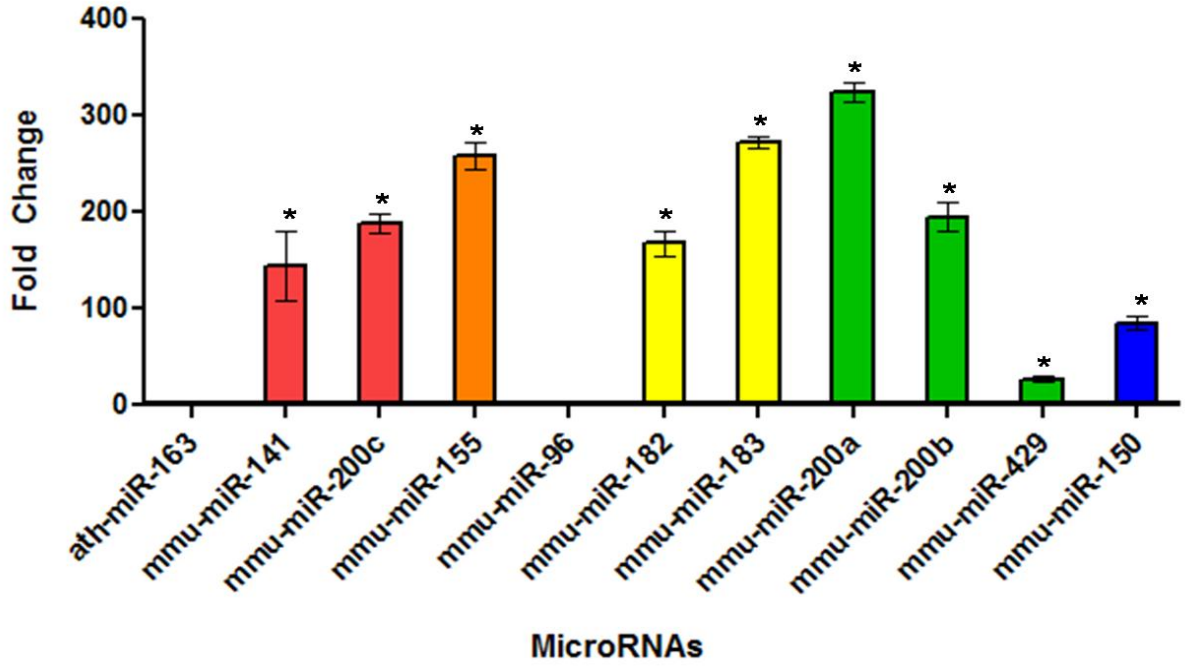


Figure 14. QPCR validation of differential expression for select miRNAs. PBS mock-infected and HSV-1 infected samples were run separately in triplicate and analyzed by the ΔΔCt method. The samples were then compared to determine fold change. MiRNAs are colour-coded by family. Ath-miR-163 is specifically expressed in plants and therefore acts as a negative control. Error bars represent standard error and * = $p < 0.001$ for Student's t-test versus the mock-infected group.

Table 20. Comparison of expression levels of select miRNA families as determined by NGS, TLDA and individual qPCR assays. NGS data taken from miRanalyzer results.

MiRNA family	MiRNA	Abundance in brain		Fold Change	Validated by TLDA	Validated by qPCR Assay
		HSV-1	Mock			
miR-200	mmu-miR-141	✓✓✓	✓✓	7.1	93.8	144.1
	mmu-miR-200c	✓✓	✓	6.3	119.5	188.2
	mmu-miR-200a	✓✓✓	✓✓✓	2.7	32.3	324.3
	mmu-miR-200b	✓✓✓	✓✓	5.7	36.3	195.3
	mmu-miR-429	✓✓	✓✓	1.3	53.1	26.9
miR-182	mmu-miR-96	✓✓✓	✓✓	4.4	nd	nd
	mmu-miR-182	✓✓✓✓	✓✓✓	2.6	8	167.4
	mmu-miR-183	✓✓✓	✓✓	13.5	57.8	272
miR-150	mmu-miR-150	✓✓	✓✓✓	0.02	0.24	85
miR-155	mmu-miR-155	✓	✓	5.2	56.1	258

The abundance of the miRNAs *in vivo* as determined by NGS read counts are: ✓ is ≤100 read counts; ✓✓ is 101-1000 read counts; ✓✓✓ is 1001-10000 read counts; ✓✓✓✓ is ≥10001 read counts. nd: no data.



Figure 15. RNA22[95] output showing the target prediction for miR-150 within the HSV-1 genome. MiR-150 is predicted to bind to the U_L18/19/20 3'UTR. RNA22 uses the Vienna RNA Package[96] to experimentally fold the target and miRNA together into a heteroduplex with the use of a standard linker. Brackets show areas of binding for each side of the heteroduplex. Beneath the heteroduplex, a diagram displaying binding between the target and miRNA is given. The top strand is the target displayed 5' → 3' and the bottom strand is the miRNA displayed 3' → 5'. Lines join the strands to show specific interactions between the target and miRNA.

3.2.2 VALIDATION OF TARGET BINDING VIA LUCIFERASE REPORTER ASSAY

In order to validate the binding of miR-150 to the putative U_L18/19/20 target, the 66 bp DNA segment containing the putative binding site was inserted into the pmiRGLO luciferase reporter vector downstream of the luciferase gene. When transcribed, the putative binding site will be located in the 3'UTR of the luciferase transcript. The miRNA is transfected and if it binds to the putative binding site in the transcript, luciferase translation will be inhibited and therefore light output will diminish. Pre-miR-150 was transfected at a wide variety of different concentrations, alongside a fixed concentration of reporter vector.

After transfection, HeLa cells were incubated for 24 hours before measuring luciferase output via the luminometer. Cells were plated in triplicate, and each well was measured 3 times for a total of 9 readings per sample. Samples were compared to non-transfected cells for background subtraction of basal luminescence values, and were normalized to *Renilla* luciferase, which is also translated by the reporter vector via a separate promotor, but does not contain the MCS so it is constitutively expressed (Figure 6). A negative control was also included, which uses a scramble miRNA that does not bind the target.

Figure 16 shows the representative results of several trials to observe binding between miR-150 and the U_L18/19/20 3'UTR target. Figure 16A shows that transfection of a number of concentrations successfully inhibited translation of the target with miR-150 over-expression. However, only concentrations from 30-100 nM were confirmed by at least two separate experiments (Figure 16B). Additionally, a reporter plasmid containing a flipped insert was included as a negative control. Luciferase expression with this plasmid was similar when transfecting either a scramble pre-miR or pre-miR-150, since the reverse compliment of the

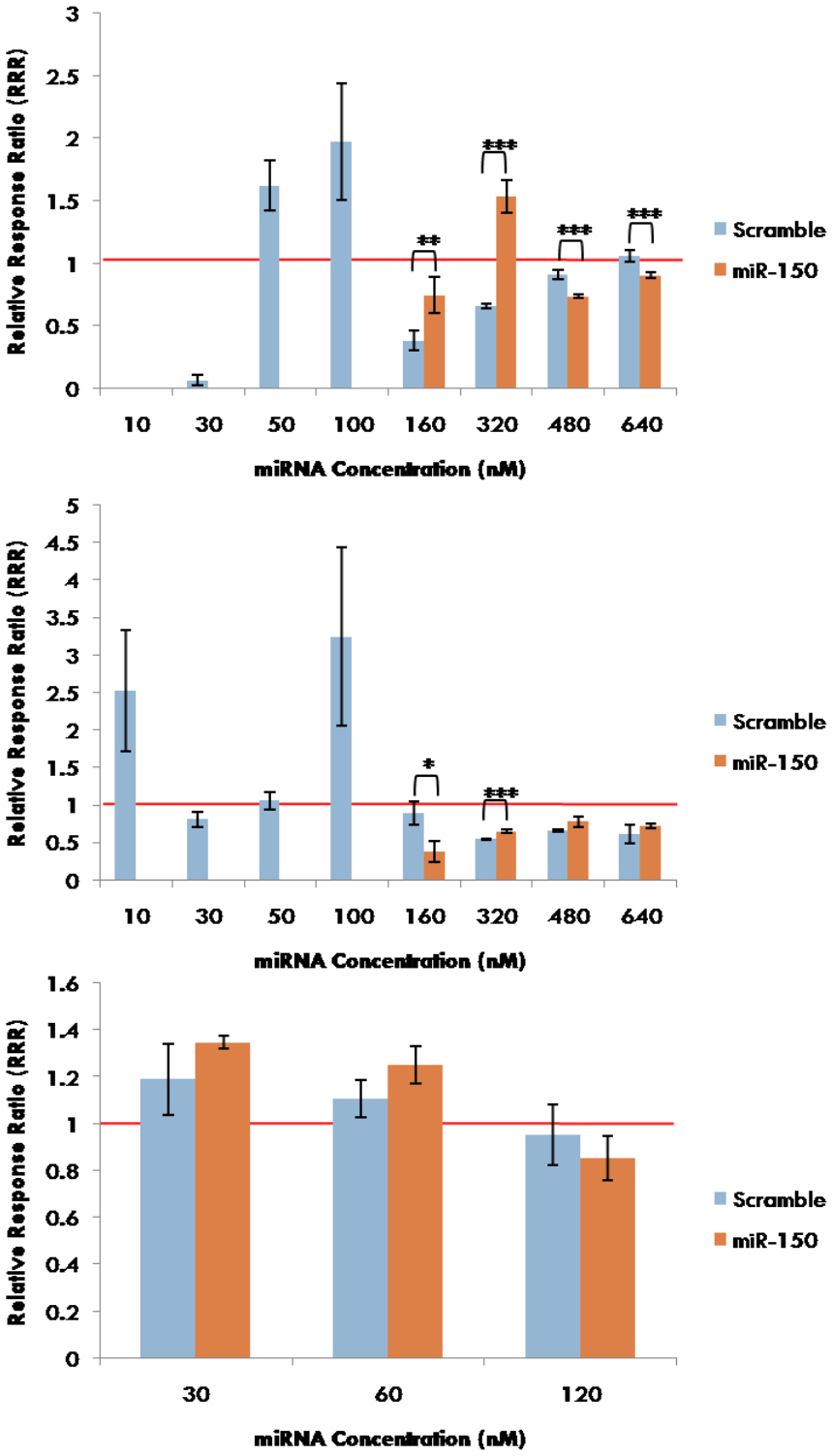


Figure 16. Luciferase assay results showing the relative response ratio (RRR) between samples co-transfected with a reporter vector and pre-miRNAs, and samples transfected only with the reporter vector. A red line is drawn at the ratio of 1, which represents the arbitrary ratio of cells with no miRNAs transfected. Each experiment was performed in triplicate, and 3 readings per well were captured. Cells were transfected with varying concentrations of either miR-150 or a negative control scramble miRNA. A and B) Results of two separate experiments confirming a knockout of luciferase activity using up to 100 nM concentrations of pre-miRs . C) A reporter vector containing a flipped insert was used as another negative control. Results show that miR-150 does not affect luciferase activity due to the reverse compliment of the putative target site being transcribed. * = $p<0.1$, ** = $p<0.05$, and *** = $p<0.01$ for student's t test versus scramble negative control group.

putative binding site would be transcribed and therefore miR-150 cannot bind to inhibit translation of the enzyme (Figure 16C).

3.4 TARGET PREDICTION FOR CANDIDATE miRNAS

After validating a subset of miRNAs to further characterize, the selected families and single miRNAs were analyzed with the target prediction web tools microT and miRPath, from the DIANA LAB website (http://diana.cslab.ece.ntua.gr/). The 10 miRNAs were analyzed by miRPath to determine which biological processes could be impacted by regulation (Table 21). MiRPath reported pathways with a p value <0.05, and the specific targets potentially regulated by the miRNAs were based on microT predictions where at least one miRNA target showed a probability of correct prediction of ≥ 0.8. MiRPath's results contained links to pathways created by the Kyoto Encyclopedia of Genes and Genomes (KEGG) to show where the pathways may be disrupted by the knockout of a target. In total, 65 pathways were predicted to be regulated by at least one of the selected miRNAs. Many of the top results were predicted to be regulated by all 10 miRNAs.

Multiple signaling pathways were predicted to be disrupted by the selected miRNAs, which are responsible for regulation of cellular functions that could impact the viability of the cells, including proliferation, cell survival, cell cycle arrest and apoptosis. In addition, the potentially affected signaling pathways, including MAPK, ErbB, Neurotrophin, Wnt and TGF-β, interact with each other as part of their signaling cascades, according to the KEGG pathway diagrams. From these results, cell viability was chosen for further study, with a focus on apoptosis, which is induced in infected cells by cytokine signalling from immune cells[25]. The signaling pathway with the lowest p value (and thus ranked higher on the list) was the mitogen-activated protein

Table 21. Pathways predicted by miRPath to be potentially regulated by at least one of 10 miRNAs (miR-141, miR-200c, miR-96, miR-182, miR-183, miR-200a, miR-200b, miR-429, miR-150, miR-155). All pathways were predicted with a $p<0.05$. MiRNAs targeting the genes within the pathways had a precision (probability of correct prediction) ≥0.8 for at least one of the putative genes involved.

KEGG Pathway	p-Value	# Genes Targeted	# miRNAs Targeting
Axon guidance	5.88E-21	42	10
MAPK signaling pathway	5.96E-18	61	10
ErbB signaling pathway	4.45E-16	28	9
Neurotrophin signaling pathway	2.7E-15	35	10
PI3K-Akt signaling pathway	1.57E-12	67	10
Regulation of actin cytoskeleton	2.85E-12	48	10
Gap junction	3.69E-12	22	9
T cell receptor signaling pathway	2.06E-11	28	10
mTOR signaling pathway	3.5E-11	20	10
Prostate cancer	3.5E-11	25	10
Acute myeloid leukemia	3.5E-11	18	10
Transcriptional misregulation in cancer	2.59E-09	36	10
Focal adhesion	4.36E-09	41	10
Adherens junction	1.81E-08	20	10
Renal cell carcinoma	4.47E-08	21	9
Salmonella infection	9.42E-08	20	10
Bacterial invasion of epithelial cells	1.35E-07	17	8
Fc gamma R-mediated phagocytosis	1.35E-07	22	9
Dopaminergic synapse	2.60E-07	30	10
Insulin signaling pathway	2.97E-07	29	10
GnRH signaling pathway	4.02E-07	21	10
Long-term depression	1.47E-06	16	9
Tight junction	2.78E-06	31	10
Endometrial cancer	2.78E-06	14	10
Glycosphingolipid biosynthesis - ganglio series	3.00E-06	6	7
Pathways in cancer	5.15E-06	58	10
Inositol phosphate metabolism	9.01E-06	14	9
Hypertrophic cardiomyopathy (HCM)	1.05E-05	19	8
Dilated cardiomyopathy	1.22E-05	20	8
Colorectal cancer	1.24E-05	16	10
Pancreatic cancer	1.33E-05	17	10
Ubiquitin mediated proteolysis	1.59E-05	28	10
Chronic myeloid leukemia	4.32E-05	17	10
Cholinergic synapse	6.12E-05	25	10
Phosphatidylinositol signaling system	6.93E-05	17	9
Osteoclast differentiation	6.93E-05	24	10

TGF-beta signaling pathway	9.30E-05	19	10
Dorso-ventral axis formation	9.73E-05	7	8
Glioma	0.000103	16	9
Non-small cell lung cancer	0.000103	13	9
Hepatitis B	0.000301	28	9
Protein processing in endoplasmic reticulum	0.000301	30	10
Serotonergic synapse	0.000336	22	9
Wnt signaling pathway	0.000409	30	10
Chemokine signaling pathway	0.000447	32	10
Long-term potentiation	0.000734	14	10
Hedgehog signaling pathway	0.000883	11	9
Melanoma	0.000883	15	10
Endocytosis	0.001113	35	10
Thyroid cancer	0.001565	7	4
VEGF signaling pathway	0.001957	13	9
HTLV-I infection	0.002818	43	10
Fc epsilon RI signaling pathway	0.003214	14	9
Lysine degradation	0.009271	10	9
Circadian rhythm	0.009945	8	10
Maturity onset diabetes of the young	0.010568	6	7
Bladder cancer	0.010568	9	8
B cell receptor signaling pathway	0.011234	14	9
Progesterone-mediated oocyte maturation	0.012348	15	10
Glycosphingolipid biosynthesis - globo series	0.012653	4	6
Oocyte meiosis	0.021287	20	10
Glutamatergic synapse	0.021287	21	10
Glycosaminoglycan biosynthesis - keratan sulfate	0.021364	4	6
Gastric acid secretion	0.04028	12	9
RNA transport	0.044019	24	10

kinase (MAPK) pathway, which was targeted by all 10 miRNAs used in the analysis (Figure 17). A summary of the number of genes each miRNA is predicted to target within the MAPK pathway is given in Table 22. Further, the potential target genes of miR-141 and miR-200c are displayed in Table 23. A total of 61 targets were predicted within the MAPK pathways, which could affect cellular processes including proliferation, differentiation, inflammation and apoptosis. The miR-141/200c family of miRNAs were also predicted to bind 3 targets within the specific apoptosis pathway (Figure 18).

3.5 APOPTOSIS ASSAY REVEALS ANTI-APOPTOTIC EFFECT OF miR-141 AND miR-200c

To determine whether the miRNAs impact the process of apoptosis, an ADP/ATP ratio assay was used to monitor the induction of apoptosis with an over-expression of either miRNAs -141 or -200c. A range of concentrations of the above miRNAs were transfected into Neuro-2A cells and incubated for 24 hours, then apoptosis was induced via UV exposure. The strength of UV required to induce apoptosis was previously optimized with a bioluminescent assay measuring caspase 3 and 7 induction. The levels of ADP and ATP were monitored over 3 hours. A high level of ADP and low level of ATP (and therefore a high ADP/ATP ratio) indicates increased apoptosis because the mitochondrial membrane potential is disrupted during apoptosis, which halts the production of ATP and therefore ADP accumulates[99]. The results for both miRNAs are displayed in Figure 19, and contrary to what was hypothesized, it appears that the miRNAs instead have a protective role during apoptosis. At most concentrations tested (10-75 nM) both miR-141 and miR-200c appear to inhibit apoptosis, reducing ADP/ATP ratios close to 0. At the 100 nM concentration, miR-200c appears to have no effect on apoptosis mediation,

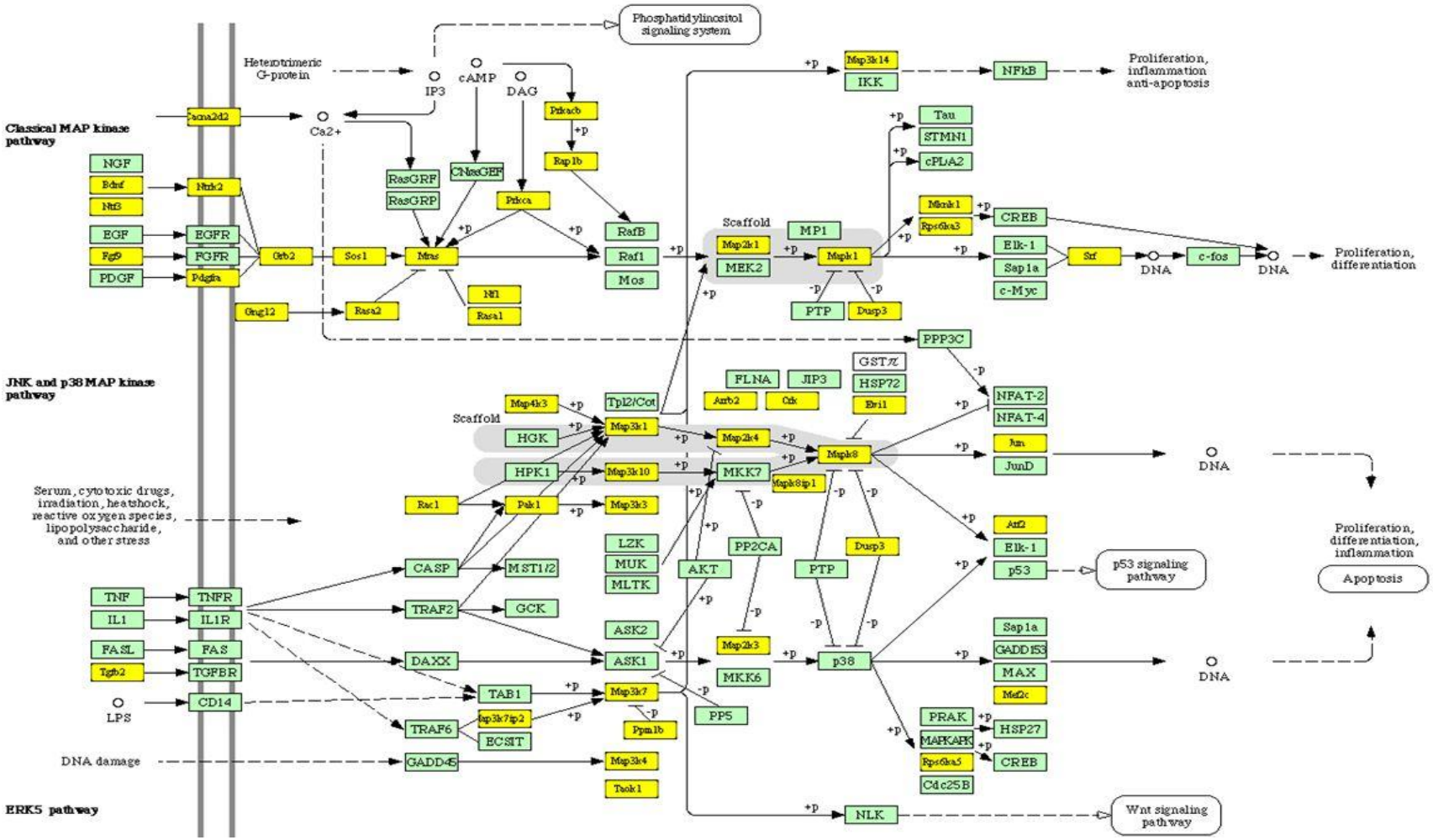


Figure 17. Partial MAPK signaling pathway by KEGG. Yellow genes are potentially targeted by three selected miRNA clusters (miR-141/200c, miR-96/182/183 and miR-200a/200b/429) and two individual miRNAs (miR-150 and miR-155) that are differentially expressed to a higher degree during HSV-1 infection *in vivo*. Data compiled by the microT and mirPath web tools from DIANA LAB (http://diana.cslab.ece.ntua.gr/).

Table 22. Number of genes within the MAPK pathway predicted to be regulated by the candidate miRNAs.

miRNA	# Genes
mmu-miR-141-3p	14
mmu-miR-200c-3p	14
mmu-miR-96-5p	23
mmu-miR-182-5p	17
mmu-miR-183-5p	8
mmu-miR-200a-3p	13
mmu-miR-200b-3p	14
mmu-miR-429-3p	18
mmu-miR-150-5p	4
mmu-miR-155-5p	11

Table 23. Genes within the MAPK pathway that are predicted by microT to be regulated by miR-141 and miR-200c. Precision is the probability of a correct prediction.

miR-141	Gene Name	Precision	Number of Predicted Binding Sites	miR-200c	Gene Name	Precision	Number of Predicted Binding Sites
1	Dusp3	0.8	2	**1**	Dusp1	0.9	1
2	Map2k4	0.8	2	**2**	Jun	0.9	1
3	Map3k3	0.8	2	**3**	Map4k3	0.9	3
4	Ntrk2	0.8	2	**4**	Map4k4	0.9	1
5	Prkacb	0.8	4	**5**	Ntf3	0.9	2
6	Rapgef2	0.8	1	**6**	Rap1b	0.9	2
7	Tgfb2	0.8	3	**7**	Rapgef2	0.9	1
8	Cacna2d1	0.7	2	**8**	Rasa2	0.9	2
9	Cdc42	0.7	1	**9**	Rps6ka	0.9	3
10	Grb2	0.7	1	**10**	Srf	0.9	3
11	Ppp3r1	0.7	1	**11**	Mknk1	0.8	1
12	Map3k7	0.6	1	**12**	Taok1	0.8	2
13	Mapk8	0.5	3	**13**	Egfr	0.7	2
14	Pdgfra	0.5	1	**14**	Rras2	0.7	2

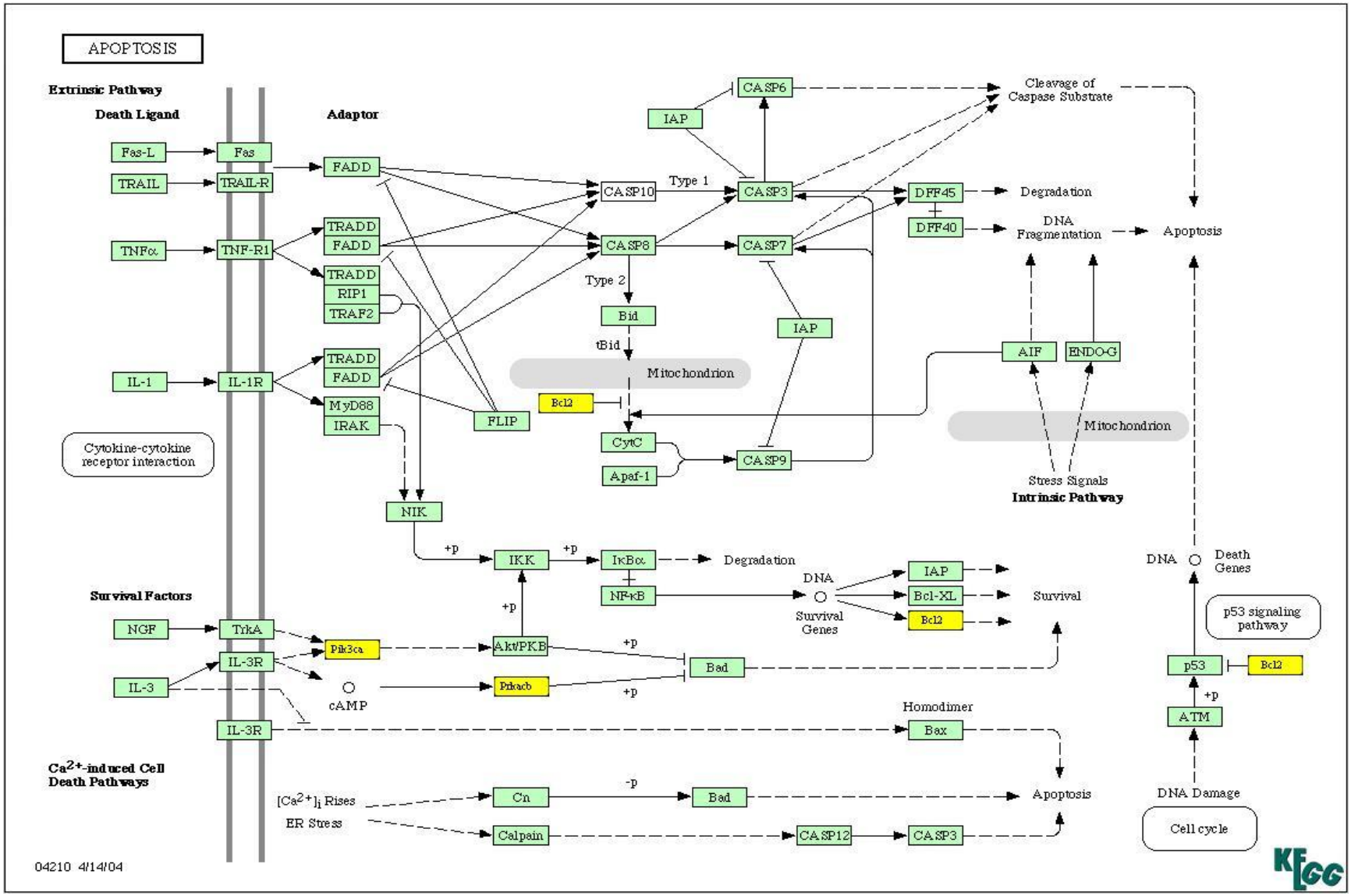


Figure 18. Apoptosis pathway by KEGG. Yellow genes are potentially targeted by the miR-141/200c miRNA family that is differentially expressed to a higher degree during HSV-1 infection *in vivo*. Data compiled by the microT and mirPath web tools from DIANA LAB (http://diana.cslab.ece.ntua.gr/).

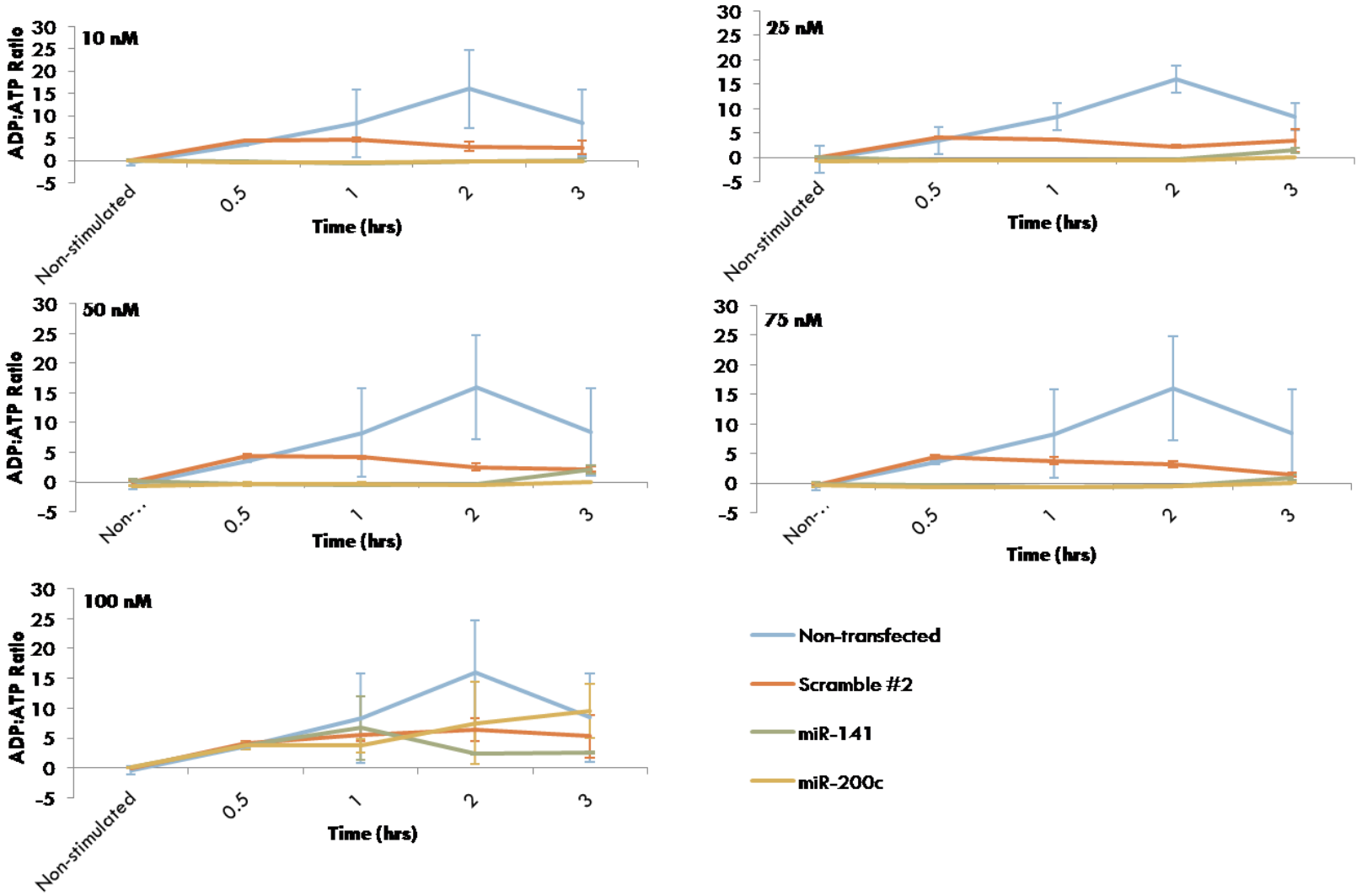


Figure 19. ADP/ATP ratio assay demonstrates an anti-apoptotic role of miR-141 and miR-200c. ADP and ATP levels were measured by luminescence at 0- 0.5-, 1-, 2- and 3 hours after UV-induced apoptosis and the ratios were plotted. The results of the transfected samples were measured against an untransfected positive control to show the apoptotic ratios over time, as well as a scrambled control to account for any variation caused by transfection. The successive graphs show the changes to ADP:ATP levels caused by an increase in transfected miRNA concentration from 25 - 100 nM.

although miR-141 continues to lower apoptosis levels. A positive control of non-transfected cells was used to monitor the ADP/ATP ratios during apoptosis induction by UV exposure.

3.6 CANDIDATE miRNAS DO NOT AFFECT CELL VIABILITY OF NEURONAL CELL CULTURE IN VITRO

Previous target predictions with the miRPath program yielded many potential targets that could affect cell viability with regulation by candidate miRNAs such as miR-141, miR-150 and miR-200c. It is therefore important to observe whether or not these miRNAs have an impact on the host cells by analyzing whether there is any change in viability with an over-expression of the miRNAs. A reduction of cell number could either mean the miRNAs are cytotoxic, in that they trigger a response that leads to cell death, or they may just slow the replication rate. Conversely, the miRNAs may cause an increase in replication rate, possibly by causing a feedback loop where the miRNAs down-regulate replication repressors. Given the number of potential targets within the MAPK signaling pathway alone, (Figure 17), an effect on proliferation is certainly possible. Trypan blue staining was used to count Neuro-2A cells after transfection with miR-141 or miR-150 (Figure 20), and these results were confirmed by ADP:ATP ratio assay (Figure 21). The ADP:ATP ratio assay was additionally used to test proliferation changes caused by miR-200c. Both figures show no significant changes to cell viability when transfected with miR-150, and miR-200c neither enhances nor decreases cell viability as shown in Figure 21 alone. However, a significant increase in cell count is reported when transfecting miR-141 at the 25 nM concentration, and a significant decrease is observed at the 75 nM concentration. These results were not confirmed by the ADP/ATP ratio assay and may only be a product of different sensitivities between the tests used.

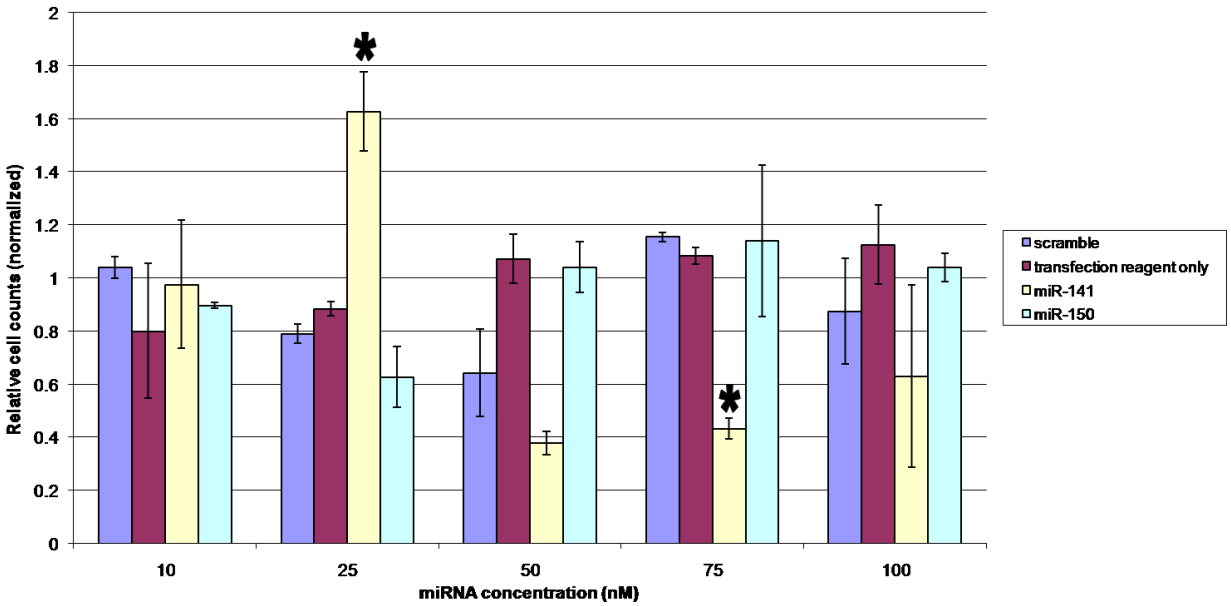


Figure 20. Neuro-2A cell counts via Trypan blue staining report little change when transfected with miR-141 or miR-150. Non-stained (live) cells from transfected samples were counted and normalized against untransfected cells to account for variations between plates. A 'scramble' negative control miRNA was transfected at the same concentrations (10-100 nM) as the sample miRNAs to show non-specific changes in cell counts due to the transfection process. Additionally, cells were plated with the transfection reagents alone to account for any toxicity caused by the chemicals themselves. * = $p<0.05$ for Student's t-test versus scramble negative control group.

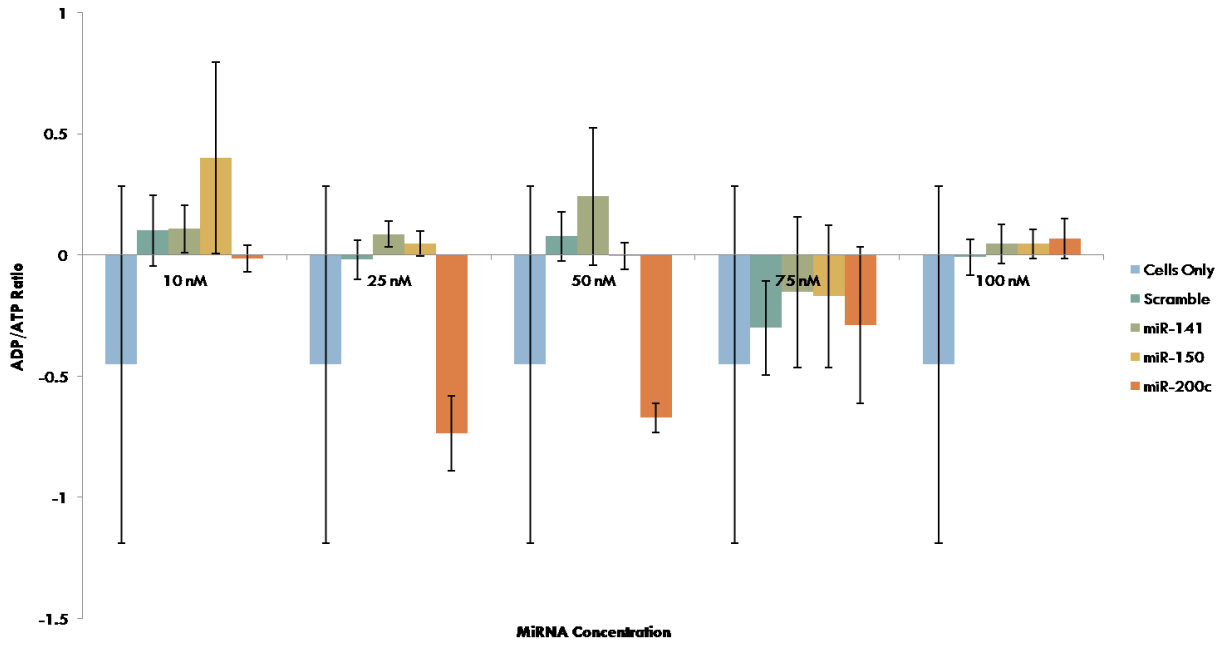


Figure 21. ADP/ATP ratio assay reports no significant changes to cell viability when transfecting miR-141, miR-150 or miR-200c. ADP and ATP levels were measured by luminescence 24 hours after transfection of an increasing concentration of miRNAs (10-100 nM), and the ratios were plotted. The results of the transfected samples were measured against an untransfected control as well as a scrambled negative control to account for any variation caused by transfection.

3.7 ISH *REVEALS GLOBAL UP-REGULATION OF CANDIDATE miRNAs IN THE BRAIN*

In order to determine the patterns of expression of the selected miRNAs within the brain, histology was performed on the mock-infected and HSV-1 infected samples. Slides were stained via ISH using double-digoxygenin-labeled probes to target specific miRNAs and were counterstained with nuclear fast red. One miRNA from each of the three clusters (miR-141, miR-183 and miR-200a) were stained. It was observed that in addition to an increase in overall miRNA expression, resulting in a darker, stronger and more abundant signal, expression was specifically increased in certain areas (Figure 22). Specific increases of expression were observed in the hypothalamus and in the pons region connecting to the cerebellum. Additionally, slides were stained for HSV-1 using fluorescent-labeled antibodies to compare viral infection in the regions of specific increases of miRNA expression. HSV-1 presence correlated with these areas of increased expression.

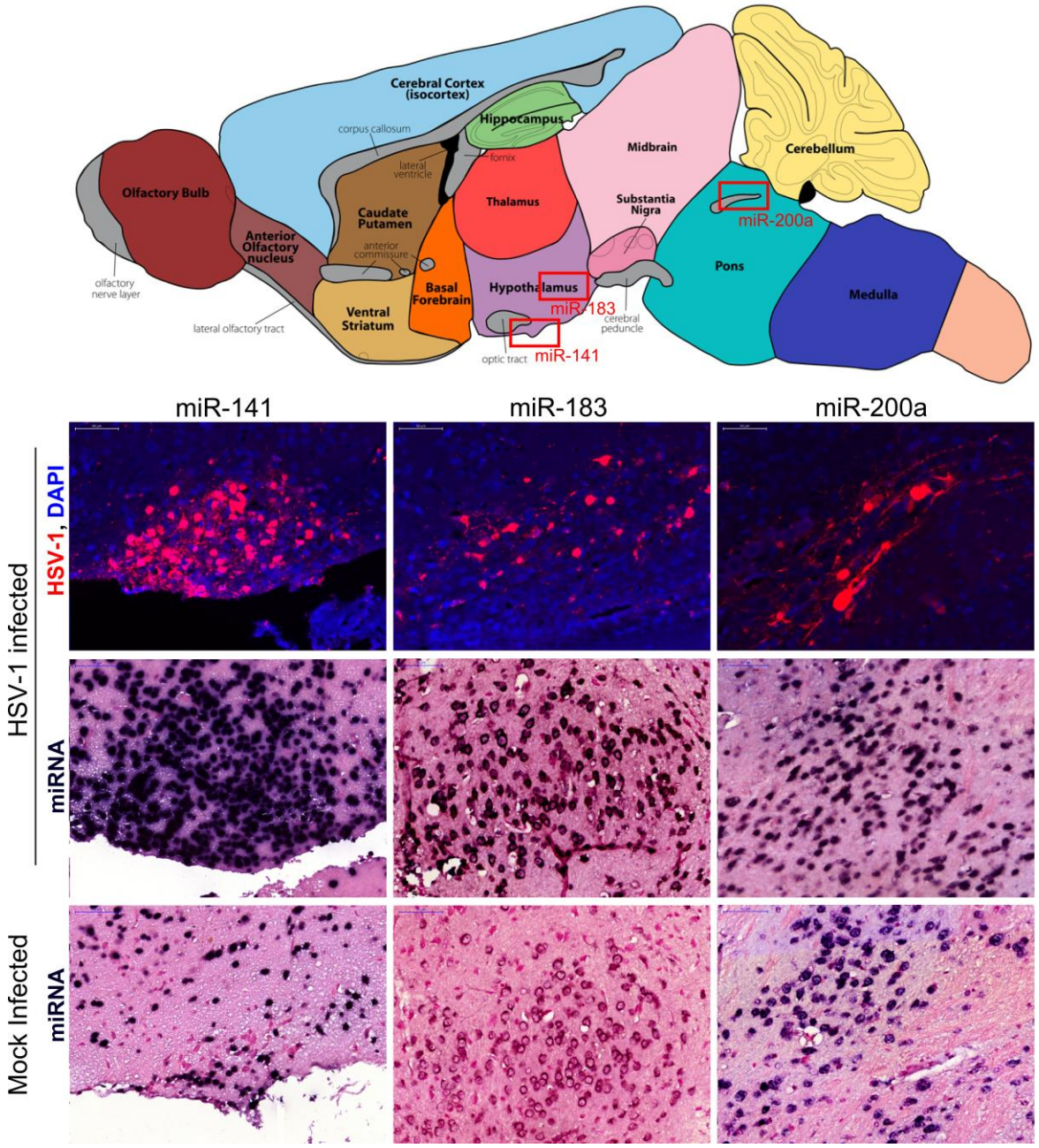


Figure 22. An increased expression of selected miRNAs is observed in the HSV-1 infection compared to mock-infected samples. Above, areas of increased candidate miRNA expression are highlighted on a map of the mouse brain. Below, microscopy magnifications compare the highlighted sections for the given miRNA stained with double-digoxygenin-labeled miRNA probes (blue-black). Fluorescence images of HSV-1 infection (red) are included to show a correlation between areas of increased expression and areas of infection (counterstained with DAPI, blue). The scale bars indicate a 50 µM segment. Mouse brain map obtained from the Gene Expression Nervous System Atlas (GENSAT)[100]. Reprinted with permission from Majer *et al.* [manuscript in preparation][101].

4.0 DISCUSSION

The focus of this project was to determine the brain-specific changes in miRNA expression during acute viral encephalitic infection using a murine model of HSVE. MiRNAs were profiled using NGS technology and were analyzed with several different programs to validate the miRNome. Differential expression during infection was determined as compared to a mock-infection, and a subset of miRNAs were verified for dysregulation using TLDA assays. A smaller subset of 10 miRNAs were further validated by qPCR in order to pursue further characterization by exploring potential targets of the miRNAs. The candidate miRNAs were predicted to target genes within many different signaling pathways that were involved in such cellular processes as proliferation, cell cycle arrest and apoptosis. ADP/ATP ratio assays were used to determine the effects of miR-141 and miR-200c on apoptosis and cell viability. Additionally, miR-150 was predicted to bind a target within the HSV-1 genome. Luciferase assays were used to experimentally validate the prediction. Finally, ISH was used to identify the regions of the brain expressing several of the selected miRNAs.

4.1 NGS IDENTIFIES HUNDREDS OF KNOWN AND NOVEL miRNAs WITHIN MOCK- AND HSV-1-INFECTED MOUSE BRAINS

Complete transcriptome analysis is a fairly new area of research and has advanced significantly with next generation sequencing technology. NGS can easily capture millions of reads per sample. The limiting factor is currently the availability of programs to properly analyze the samples and the requirement for powerful computing resources to handle the large amount of data. To address this, an analysis protocol was developed using cloud-based programs, Galaxy and miRanalyzer, to capitalize on server-side computing power. This protocol was validated by

two additional programs, GeneSifter and ACGT101-miR. MiRanalyzer identified the most known and novel miRNAs of the three programs. The program found 642 known and 587 novel miRNAs in the mock-infected sample, and 521 known and 389 novel miRNAs in the HSV-1 infected sample. Additionally, mapping to Rfam led to the identification of novel isomiRs of 26 and 13 human miRNAs in the mock-infected and HSV-1 samples, respectively. The majority of the known miRNAs were identified by all three programs, but miRanalyzer was able to identify significantly more novel miRNAs than ACGT101-miR (Figure 11). 405 (59%) of the unique sequences identified as novel miRNAs from the mock-infected sample and 362 (66%) from the HSV-1 infected sample that were identified by miRanalyzer were not identified by ACGT101-miR. In contrast, only 57 (17%) and 47 (20%) of the novel miRNA unique sequences identified by ACGT101-miR in the mock and HSV-1 infected samples, respectively, were not identified by miRanalyzer.

4.1.1 COMPARISON OF miRNA PROFILE ANALYSIS TECHNIQUES

After comparing the different NGS analyses, it is clear the trends are similar between programs, with slight variations in copy number and therefore the order of the individual miRNA abundances (Table 17 and 18). This can likely be attributed to the variations during analysis between the different methods. GeneSifter and ACGTmiR-101 remove any matches to Rfam[76], but ACGTmiR-101 additionally removes matches to Repbase (a database of repetitive sequences)[102] and RefSeq[73,103] before searching for known miRNAs. Both programs then map reads to miRBase[104] to identify known miRNAs. GeneSifter maps to and filters reads matching to RefSeq only after mapping to miRBase. This could be beneficial as it could capture any miRNAs that may be formed by post-transcriptional splicing of exons[85,105,106]. In contrast,

miRanalyzer first maps to miRBase and then maps to RefSeq and Rfam after all known miRNAs are found. While the filtered reads are removed from the rest of the downstream analysis to determine novel miRNAs, they are still viewable with miRanalyzer. This allows for the ability to analyze any matches to known miRNAs from the Rfam database. As seen in Table 6, there were many unique reads that matched several human miRNAs in Rfam, which could be reported as novel isomiRs within the mouse genome. These isomiRs would not be identified by GeneSifter or ACGT101-miR since Rfam hits were discarded. With regards to novel miRNA prediction, miRanalyzer and ACGT101-miR appear to differ quite extensively given the information available. The complex grouping of miRNAs with ACGT101-miR make it is possible to differentiate between completely novel miRNAs, new isomiRs within the species, new chromosomal locations for existing miRNAs, and known miRNAs not previously observed in the mouse genome (Table 11). With miRanalyzer, these same sequences may be discovered, however less information about them would be available without further analyzing the novel sequences by BLAST, or by re-analyzing them with miRanalyzer against different genomes. Although ACGT101-miR has a more complex classification scheme, miRanalyzer predicted more candidates than ACGT101-miR for both the mock-infected and HSV-1 infected samples.

Many of the 'novel' predictions by ACGT101-miR that are previously reported in another species were annotated as unique reads for a known miRNA within the mouse species by the other programs, due to the high conservation of miRNA sequences[41]. An example is miR-141, which was identifed by GeneSifter and miRanalyzer as a known miRNA in the PBS sample, but not by ACGT101-miR. However, hsa-miR-141 was identified as a novel miRNA with ACGT101-miR. Both mmu-miR-141 and hsa-miR-141 have the following sequence: 5'-

uaacacugucugguaaagaugg-3', but the isomiR that matched to hsa-miR-141 by ACGT101-miR has 3 additional bases at the 3' end, which better match the human pre-miR than the mouse pre-miR. Further, hsa-miR-141 was grouped into gp3 (novel) instead of gp1b (known) (Table 11) because of the high stringency for mapping to the mouse genome. Zero mismatches to the genome were allowed with both ACGT101-miR and miRanalyzer. However, the mapping algorithm of ACGT101-miR allows 0 mismatches within the entire read, whereas miRanalyzer allows 0 mismatches only within the annotated mature miRNA sequence. Thus, miR-141 was recognized as a known miRNA with miRanalyzer but was classified as novel using ACGT101-miR. Due to this difference in classification schemes and the higher stringency for recognizing known miRNAs by ACGT101-miR, this program reported less miRNAs overall than both miRanalyzer and GeneSifter (Figures 10 and 11).

The last major difference between the programs is the decision to use a normalization strategy. ACGT101-miR does not normalize its data and only reports copy numbers. GeneSifter used the RPM technique which is currently the best procedure for normalization, although no technique has been been accepted as a gold standard as of yet[86]. RPM normalization divides the number of reads for a given miRNA by the total number of reads that align to miRBase. The number is then multiplied by 1 million to obtain the reads per million[86,107]. In contrast, miRanalyzer normalized with two methods: by dividing the read count by the total number of reads (normalized expression: all) and by the total number of unique reads mapped to miRBase (normalized expression: mapped). Both methods then multiplied by 100 to obtain a percentage. While normalization may be helpful between samples analyzed in the same fashion, it is not necessarily useful when comparing to other published experiments and

therefore the copy number would have to be included. This is because of the lack of a standard analysis procedure. As mentioned above, the order of mapping to databases and filtering reads will change the final total of reads on which the normalization is based. The stringency of acceptable mismatches will also change the amount of mapped reads. Therefore, even if samples were normalized the same way, they could only be compared to other samples analyzed with the same procedure.

4.1.2 CANDIDATE miRNAs ARE INVOLVED IN OTHER VIRAL INFECTIONS

The known miRNAs identified by NGS analysis were validated by TLDA assays, which found 58 up-regulated and 5 down-regulated miRNAs with ≥5 fold change. A smaller subset consisting of the highest positively regulated miRNAs and their respective family members was further validated using individual qPCR assays. These miRNAs have been previously reported as dysregulated in at least one other viral infection (Table 24). MiR-141 is the most extensively studied of the candidate miRNAs, with dysregulation reported in 6 different viral infections, and 3 verified targets associated with infection. Infection with all viruses studied so far show positive regulation of miR-141 during infection. The only exception is that during infection with influenza viruses, miR-141 is up-regulated early on during infection (3 hpi), but expression decreases over time to show a net down-regulation at later time points (18-24 hpi)[108]. Of all the miR-141 targets identified, DLC-1 appears to be the least relatable to HSV-1 infection. DLC-1 is a tumour supressor and repression by miR-141 during HCV infection is related to the formation of hepatocellular carcinoma[109]. To date, no studies have investigated DLC-1 expression during HSV-1 infection. DLC-1 expression may be more important for oncogenic

Table 24. Dysregulated miRNAs identified during HSVE have been previously reported as dysregulated during other viral infections. Targets are listed if referenced authors identified any in addition to miRNA expression profiling.

MiRNA	HSV-1 Dysregulation	Other Viral Infections	Dysregulation	Experimentally Validated	Target(s)	Reference
miR-141	↑	enterovirus 71	↑	*in vitro*	eIF4E	[110]
		poliovirus 3	↑	*in vitro*		[110]
		coxsackievirus B3	↑	*in vitro*		[110]
		avian influenza A H5N1	↑/↓ (time course)	*in vitro*	TGF-β2	[108]
		human influenza A H1N1	↑/↓ (time course)	*in vitro*		[108]
		hepatitis C virus	↑	*in vivo*	DLC-1	[109,111]
miR-200c	↑	hepatitis C virus	↑	*in vivo*	FAP1	[111]
		Epstein-Barr virus	↓ (latency)/↑ (lytic)	*in vitro*		[112]
miR-96	↑	human cytomegalovirus	↑	in vitro		[113]
miR-182	↑	hepatitis C virus	↑	*in vivo*		[111]
		human cytomegalovirus	↑	*in vitro*		[113]
miR-183	↑	human cytomegalovirus	↑	*in vitro*		[113]
miR-200a	↑	Epstein-Barr virus	↓ (latency)/↑ (lytic)	*in vitro*		[112,114]
		hepatitis C virus	↑	*in vitro*	DLC-1	[109]
miR-200b	↑	Epstein-Barr virus	↓ (latency)/↑ (lytic)	*in vitro*	ZEB1, ZEB2	[112,114,115]
		hepatitis C virus	↑	*in vitro*		[109]
miR-429	↑	Epstein-Barr virus	↓ (latency)/↑ (lytic)	*in vitro*	ZEB1, ZEB2	[112,115]
miR-150	↓	dengue virus	↑	*in vivo*	SOCS1	[116]
		hepatitis C virus	↑	*in vivo*		[117]
miR-155	↑	Epstein-Barr virus	↑	*in vitro*	IKKε	[118]
		human cytomegalovirus	↓	*in vitro*		[113]

herpesviruses such as Epstein-Barr virus (EBV) and Kaposi's sarcoma-associated herpesvirus (KSHV).

During infection with enterovirus 71, miR-141 represses eukaryotic translation initiation factor 4E (eIF4E) to shut off host cap-dependent translation and switch to primarily cap-independent translation, which is beneficial to the virus[110]. However, it has been shown that during HSV-1 infection, eIF4E activation by MAPK interacting kinase 1 (MKNK1) is necessary for a productive infection and is stimulated by ICP0[119]. Thus, eIF4E is likely not a target during HSV-1 infection. Interestingly, MKNK1 was predicted by microT as a target of miR-200c within the MAPK signaling pathway (Table 23). The work of Walsh and Mohr (2004)[119] suggests MKNK1 is not a target during HSV-1 infection, but miR-200c could function by fine-tuning the expression of the kinase rather than completely repressing the gene.

Studies of influenza A infection led to the identification of TGF-β2 as a target of miR-141. TGF-β2 is a context-sensitive cytokine that can be either pro- or anti-inflammatory. During influenza A infection, it was theorized that TGF-β2 repression may contribute to the excessive inflammation observed during disease progression[108]. TGF-β2 is involved in the MAPK signaling pathway and was one of the predicted targets of miR-141 (Table 23). Increased TGF-β2 during HSV-1 infection has shown to result in increased levels of latency[120], but too much or too little TGF-β2 has led to apoptosis[121], so miR-141 could be used to try to regulate TGF-β2 levels during a lytic infection.

HCV infection resulted in an increase in expression of miR-200c, which subsequently repressed the first apoptosis signal-associated phosphatase 1 (FAP-1) enzyme. FAP-1 repression led to increased expression of fibroblast growth factor (FGF) and increased fibrosis *in vitro*[111]. FAP-1 is

a regulator of the MAPK signaling pathway, and FGF is a signaling molecule that initiates a signal cascade through the MAPK pathway to stimulate proliferation and differentiation. Additionally, the miR-96/182/183 family of miRNAs were predicted to target FGF9 by microT (Figure 17). FGF has been shown to reduce HSV-1 lytic infection[122] and the FGF receptor (FGFR) also functions as a receptor for HSV-1 viral entry[18]. Regulation of FGF during HSV-1 infection has not been studied, but repression of FGF by HSV-1 could allow increased viral entry to uninfected cells by reducing competition for the receptor, and would therefore lead to increased levels of infection.

Other herpesviruses such as EBV and human cytomegalovirus (HCMV) have shown a positive regulation of many of the selected miRNAs during a productive infection. However, HCMV shows a down-regulation of miR-155, whereas HSV-1 and EBV show a positive regulation. During HSVE, miR-155 is important for a proper immune response, as miR-155-knockout mice showed increased susceptability to HSVE when inoculated ocularly[62]. Conversely, over-expression of miR-155 during EBV infection is thought to be involved in immune evasion by repressing the inhibitor of nuclear factor κB (NF-κB) kinase, subunit epsilon (IKKε), which in turn represses NF-κB activation and inhibits the immune response[118]. Interestingly, in EBV, miR-200a -200b, -200c and -429 have been identified as down-regulated during latency. MiR-200b and miR-429 were found to target the zinc finger E-box-binding homeobox (ZEB) proteins ZEB1 and ZEB2, which interact with the EBV genome at the Zta promotor to inhibit transcription of lytic genes[112,115]. ZEB1 and ZEB2 were also predicted targets of miR-141 and miR-200c according to microT, however the expression of host miRNAs during HSV-1 latency has not yet

been investigated, nor has there been any established interactions between the HSV-1 genome and ZEB1 or ZEB2.

The last known target of the candidate miRNAs associated with viral infections is the suppressor of cytokine signaling 1 (SOCS1), which is repressed by miR-150. SOCS1 was repressed during dengue virus infection but has been reported as induced during HSV-1 infection[123,124]. MiR-150 expression during HSVE was down-regulated according to NGS and TLDA data, which would coincide with an induction of SOCS1, but conflicting expression data for miR-150 was reported by qPCR (Table 20). SOCS1 induction during HSV-1 infection is thought to suppress the IFN immune response to infected cells as part of the HSV-1 immune evasion strategy[124].

4.1.3 OTHER DEREGULATED miRNAs MAY HAVE ROLES DURING HSV-1 INFECTIONS

The miRNomes created during other viral infections, as well as bacterial and fungal infections, are increasingly being published due to the advent of NGS technology, but the functional roles of the miRNAs involve further investigation and so they take more time. Table 24 shows that out of the 10 miRNAs studied, only 7 have been analyzed for functional roles, whereas the other 3 were simply found to be over-expressed during infection. Many other miRNAs were deregulated during infection as well, but to date have not been further investigated. One miRNA that has been deregulated in infections by 3 different viruses in addition to HSV-1 is miR-188-5p. MiR-188 was confirmed to be up-regulated during HSV-1 infection (Table 19) and was also up-regulated during infections with avian influenza A H5N1 and H1N1[108]. In contrast, miR-188 was significantly down-regulated during enterovirus 71 infection[110]. MiR-188 has not been significantly studied in the context of infection and only one target was predicted using microT's default settings: CDH20, a gene encoding cadherin 20, a cell-adhesion molecule.

Cadherins associate with nectin-1, an HSV-1 entry receptor, and it was shown that disassociation of the proteins allowed for greater rates of infection by allowing better access of viral glycoproteins to nectin-1[125]. MiRNAs have been known to be used as signaling molecules by packaging in microvesicles[40,126,127], so it is possible that miR-188 is modulated by HSV-1 and secreted to facilitate infection. Neighbouring cells may decrease cadherin upon receiving miR-188 to expose nectin-1 for increased rates of viral entry.

A miRNA that was highly up-regulated during HSV-1 infection but was not studied as part of this project was miR-152. MiR-152 was the 6th highest up-regulated miRNA as shown by TLDA (Table 19). Conversely, miR-152 was shown to be down-regulated in HCV infection[128]. Huang *et al.* (2014)[128] showed that a decrease of miR-152 creates an over-expression of its target, Wnt1. Wnt1 is involved in the Wnt signaling pathway, which is responsible for regulating cellular factors such as proliferation. The increased expression of Wnt1 has been implicated as a contributor to HCV-induced hepatocellular carcinoma[128]. MiR-152 has not been investigated for its function during HSV-1 infection, but a repression of the Wnt signaling cascade follows the trend of repression of signaling pathways as predicted by microT for the other highly up-regulated miRNAs, such as miR-141 and miR-200c. HSV-1 is known to decrease host gene expression upon infection[129], so repression of signaling pathways could be part of the mechanism for achieving this cellular state. There are many more miRNAs that were deregulated during infection with HSV-1. Exploring their function may yield new insight into the cellular processes that are invoked or repressed during infection.

4.1.3.1 POTENTIAL FUNCTION OF NOVEL miRNAs DURING INFECTION

The miRNome of HSVE also includes novel miRNAs that have yet been authenticated. These novel miRNAs may be new isomiRs of previously identified mature miRNAs, or they could be entirely new mature miRNAs that to date have not been discovered. In order to properly catalogue them, the novel miRNAs must first be verified to be expressed within the samples. This is achieved by performing Northern blot analysis on the samples, wherein the probes match the sequence found with NGS. Nevertheless, it is already possible to investigate the potential roles of novel miRNAs using microT v3.0. This tool allows the input of novel miRNAs in FASTA format in order to identify possible targets.

This project did not focus on novel miRNAs, but they were reported by both miRanalyzer and ACGT101-miR. Additionally, ACGT101-miR was able to provide differential expression data for novel miRNAs in addition to known mature miRNAs (Table 16). The novel miRNA with the highest up-regulated fold change was an ortholog of miR-1827 with the following sequence: 5'-TGAGGCAGTAGATTGTATAGTTATC-3'. To date, miR-1827 has only been identified in 3 species (human, orangutan and chimpanzee), so this novel miRNA will herein be referred to as 'mmu-miR-1827.' The top putative targets for mmu-miR-1827 are related to cellular functions such as metabolism, cycle arrest and DNA repair. An increase in mmu-miR-1827 may therefore be involved in the general reduction of host processes as the virus takes over[130]. However, mmu-miR-1827 was also predicted to target PKR with 7 putative target sites within the transcript. As mentioned, PKR inhibits viral translation by phosphorylating EIF2α[25], so an increase of mmu-miR-1827 caused by viral infection would be beneficial to HSV-1 replication and/or latency (if expressed during this life cycle) by helping to avoid the immune system.

Another novel miRNA that is up-regulated during HSV-1 infection has the following sequence: 5'-TAGATTCTGAATAAAATC-3' and will herein be referred to as 'Candidate 1'. Both mmu-miR-1827 and Candidate 1 were predicted to target zinc finger protein 587 (ZNF587), a transcriptional regulator. Information is scarce regarding this protein, but it was shown to be important for successful replication of human immunodeficiency virus (HIV) *in vitro*[131]. It is unknown if ZNF587 is required for HSV-1 replication. Other putative targets include signaling molecules such as protein kinase C-α (PRKCA), involved in the MAPK signaling pathway and cyclic adenosine monophosphate-responsive element-binding protein (CREB) 5, part of the CREB signal cascade which mediates cellular functions such as proliferation and cell survival. Potential repression of these targets continues the trend observed by the target prediction for the highly up-regulated known miRNAs mentioned above. The potential repression of multiple overarching signaling cascades responsible for multiple cellular processes may contribute to the restricted host translation as the virus takes over the cell[129,130].

4.3 APOPTOSIS MEDIATION BY miR-141 AND miR-200c

The MAPK signaling pathway was an intriguing system to pursue investigation with the selected miRNAs, which were predicted to regulate up to 61 targets within the pathway (Table 22). Targets such as MKNK1 and FGF9 have been previously implicated to be involved in HSV-1 infection[119,122] and TGF-β2 has already been experimentally validated to be regulated by miR-141[117]. The MAPK pathway is involved in a wide range of cellular processes including apoptosis and TGF-β2 is involved in regulating apoptosis[121].

ADP/ATP ratio assays were used to determine the effects of miR-141 and miR-200c on cells undergoing apoptosis. From the apoptosis it appears that both miRNAs have an inhibitory

effect on apoptosis. As shown in Figure 19, miR-141 and miR-200c reduce ADP:ATP levels below the untransfected control for a minimum of 2 hours after induction of apoptosis. This effect was observed with a miRNA concentration range of 10-75 nM.

In addition to the targets predicted for MAPK and other signaling pathways, miR-141 and miR-200c were also predicted to target 3 genes involved in anti-apoptosis (Figure 18). Based on the apoptosis assays, it appears that these genes are not targeted under UV-induced apoptotic conditions. The apoptosis pathway is no longer predicted by newer versions of miRPath, suggesting the prediction was not a strong candidate. However, target predictions shown in the MAPK pathway may still be involved in producing the observed results. Apoptosis is only one aspect of the MAPK pathway, which is also responsible for proliferation, differentiation and inflammation. If the predictions are correct, the results of the miRNA-mediated post-transcriptional repression would more likely be a context sensitive matter. In the case of the induction of apoptosis, perhaps these miRNAs target one or a number of the predicted gene transcripts when the signaling cascade is initiated. It would be interesting to see if other outcomes of the MAPK signaling pathways are affected by these miRNAs as well. Proliferation does not seem to be affected by these miRNAs overall, as observed by the cell viability assays (Figures 20 and 21). However, a more targeted approach could be taken to determine whether proliferation is truly affected, and whether the MAPK pathway is manipulated during proliferation and other processes with an over-expression of miR-141 and/or miR-200c.

4.4 miR-150 AS A POTENTIAL REGULATOR OF HSV-1 REPLICATION

Research of direct regulation of viral genes by host cellular miRNAs is still in its infancy, but promising results have been obtained through the investigation of interactions between miR-

122 and HCV. Anti-miR-122 has completed phase 2a clinical trials to negate the positive regulation of miR-122 on HCV replication and therefore decrease infection[68]. In this regard, target prediction using RNA22 was performed to determine potential binding sites for cellular miRNAs within the HSV-1 genome. The shared 3'UTR of the U_L18, U_L19 and U_L20 genes was predicted to be targeted by miR-150 (Figure 15). Luciferase reporter assays confirmed the binding of miR-150 to the putative target (Figure 16). A complete knockout of luciferase translation occured with 30-100 nM concentrations of miR-150. Additionally, a control plasmid containing the reverse compliment of the putative binding site confirmed that the knockout of luciferase expression was due to miR-150 binding, since luciferase expression was restored when co-transfecting pre-miR-150 with the control plasmid.

The three genes that could potentially be regulated by miR-150 are important for a productive infection of HSV-1. U_L19 encodes the major capsid protein VP5, responsible for the formation of the hexons and pentons that make up the majority of the capsid[98]. U_L18 encodes VP23, one of two minor capsid proteins that make up the triplexes that link capsomer complexes together[132]. Knockout of VP23 was shown to result in smaller, misformed capsids[132], and knockout of either VP5 or VP23 inhibited plaque formation *in vitro*[133]. U_L20 is responsible for depositing glycoprotein gK into nuclear and/or cell membranes to be included in mature virions during budding, and also facilitates budding of the virus[134]. Knockout of U_L20 resulted in an accumulation of cytoplasmic and vesicular virions that were not released out of the infected cell[134].

Recent work by Bogerd *et al.* (2014)[135] suggested that most human-infecting viruses are resistant to RNA interference by endogenous miRNAs and given that most viruses evolve

alongside their natural hosts, it is logical that viruses would change areas complementary to host miRNAs due to selective pressure over time[41,54,135]. The authors tested this theory by creating a tissue culture cell line lacking the Dicer enzyme so that pre-miRNAs could not be processed in to mature, functional miRNAs[135]. They infected this cell line with a wide variety of viruses, and most replication rates of the viruses did not change. The authors concluded that miRNA-mediated repression of viral genes is not a major factor during infection[135].

In regards to HSV-1 infection, replication rates were reduced due to the lack of Dicer functionality. The authors contribute this to the fact that the viral-encoded miRNAs were inhibited, rather than a need for cellular miRNAs during infection[135]. However, the functions of the viral miRNAs that have been discovered to date have been to promote latency, which would further reduce viral replication[46,47]. Further, only 3 miRNAs have been confirmed to be expressed during the lytic replication of HSV-1. HSV-1 was the only DNA virus tested in the study[135], so it is possible that HSV-1 and other DNA viruses require cellular miRNAs for positive regulation, whereas RNA viruses do not. In regards to miR-150 specifically, the confirmation of binding to the target site could help explain why the miRNA was down-regulated *in vivo*. HSV-1 could have selectively evolved for a mechanism to repress this miRNA during infection to evade the consequent negative regulation of the U_L18/19/20 genes caused by miR-150.

4.4.1 HOST miRNAs HAVE THE POTENTIAL TO BE USED THERAPEUTICALLY

Bogerd *et al*. (2014)[135] addressed the fact that many viral infections occur unhindered by miRNA-mediated repression under normal circumstances. However, the introduction of exogenous miRNAs in concentrations that are not normally encountered by viruses may still have therapeutic function. Therefore whenever a miRNA is discovered that confers a negative

effect on the virus, there is always the hope that it could one day be administered to combat viral infection with minimal harm to the host cells. There are several strategies to pursue when looking for a potential therapeutic candidate: miRNAs could be used to inhibit the virus directly by binding to viral transcripts, they could bind to cellular factors to confer resistance to infection and therefore decrease replication, or they could be used to prime and/or boost the immune system to increase the rate at which infection is cleared[136]. Additionally, anti-miRs could bind to inhibit cellular or viral miRNAs to impede any positive functions that may have otherwise occurred.

MiR-150 may follow the first strategy, as it has the potential to bind HSV-1 transcripts directly (Figure 16). Additionally, miR-150 was found to repress SOCS1[116] which has been shown to be induced during HSV-1 infection[123]. SOCS1 is a negative regulator of the IFN immune response pathway[124], so an increase of miR-150 could help strengthen the immune response to HSV-1 infection in addition to potentially inhibiting viral replication directly. Another miRNA that could be used therapeutically is miR-155. MiR-155 expression has been shown to increase the immune response during HSVE and a knockout of miR-155 led to decreased survival rates *in vivo*[62], so therapeutically-administered miR-155 shows potential to benefit the host during infection.

With regards to HSV-1 infection specifically causing HSVE, therapeutic agents must be able to cross the blood-brain barrier (BBB) to target infected cells in the CNS. Several compounds working together were able to successfully carry miRNAs across the BBB[136,137]. Mannitol was used to permeabilize the BBB and polyethylenimine (PEI) was used to complex with the miRNAs to create a net positive charge, allowing the complexes to interact with the cell surface for

endocytosis. The PEI complexes were attached to the rabies virus glycoprotein (RVG) via disulfide bonds to target neurons. It was found that much higher accumulation of labeled miRNA was found in the brain when attached to RVG than the PEI complexes alone[136,137].

4.5 MiRNA EXPRESSION INCREASED IN SPECIFIC AREAS DURING INFECTION

To visualize the expression of miRNAs within the brain, ISH staining was performed for 1 miRNA from each of the 3 clusters of selected miRNAs. MiR-141, miR-183 and miR-200a were stained with double-digoxygenin-labeled probes and the expression patterns were compared between mock- and HSV-1-infected samples. Overall, expression was increased globally during infection when compared to mock across the brain for the 3 miRNAs that were stained. Figure 22 highlighted several areas where specific increases in miRNA expression were observed during HSV-1 infection. These regions correlated with viral infection. The fact that specific increases in miRNA expression were observed in regions where viral replication was evident suggests the expression may be a result of infiltrating immune cells in addition to infected neurons present in the region.

4.6 FUTURE DIRECTIONS

Analysis of the deregulated miRNome created during acute viral encephalitis with HSV-1 revealed a wealth of new information regarding the differential expression of known and novel miRNAs. Each change in miRNA expression affects its own set of targets, which will cause its own unique changes to the cell during infection. Therefore any significantly deregulated miRNA could warrant further investigation to determine the potential implications caused by the change of expression of the given miRNA. For example, miR-188 has been deregulated in several viral infections[108,110] including HSV-1, yet little is known about the miRNA in this context.

This project focused on the miRNAs with the largest positive fold changes, and more research into these miRNAs could continue. To date, 8 targets have been discovered for 7 of the 10 miRNAs investigated (Table 24). The role of these targets and their miRNA-mediated repression (if any) could be investigated specifically during HSV-1 infection. Additionally, predicted targets for these miRNAs could be confirmed and further investigated in regards to the consequences of repression under infectious conditions. Finally, many novel miRNAs have been discovered by NGS sequencing and these miRNAs should be confirmed by Northern blot analysis. Target prediction software could then be used to discover potential genes repressed by these newly discovered miRNAs.

Further research could be completed regarding the functional studies of the selected miRNAs as well. MiR-141 and miR-200c were determined to have anti-apoptotic properties, but it is unclear which genes are targeted to confer the apoptotic resistance. Western blots could be used to determine which genes are repressed during apoptosis induction in cells over-expressing miR-141 or miR-200c. The investigation of the effectiveness of miR-150 against HSV-1 infection could be continued by performing Western blots to determine if miR-150 can truly inhibit translation of the target genes during infection. Additionally, $TCID_{50}$ assays could be performed to observe whether miR-150 could effectively inhibit a productive infection. Finally, cell-specific co-staining with ISH of selected miRNAs could be performed to further investigate whether microglia and other immune cells co-localize with miRNAs in infected areas of the brain to confirm their relationship. This correlation would confer an innate immune-related role for the miRNAs that has not previously been recognized.

REFERENCES

1. Corsten, M. F. *et al.* MicroRNA profiling identifies microRNA-155 as an adverse mediator of cardiac injury and dysfunction during acute viral myocarditis. *Circ. Res.* **111,** 415–25 (2012).

2. Duskova, K. *et al.* MicroRNA regulation and its effects on cellular transcriptome in human immunodeficiency virus-1 (HIV-1) infected individuals with distinct viral load and CD4 cell counts. *BMC Infect. Dis.* **13,** 250 (2013).

3. Gupta, P., Cairns, M. J. & Saksena, N. K. Regulation of gene expression by microRNA in HCV infection and HCV-mediated hepatocellular carcinoma. *Virol. J.* **11,** 64 (2014).

4. Mehrabadi, M., Hussain, M. & Asgari, S. MicroRNAome of Spodoptera frugiperda cells (Sf9) and its alteration following baculovirus infection. *J. Gen. Virol.* **94,** 1385–97 (2013).

5. Kumar, M. & Nerurkar, V. R. Integrated analysis of microRNAs and their disease related targets in the brain of mice infected with West Nile virus. *Virology* **452-453,** 143–51 (2014).

6. Baines, J. D. & Pellett, P. E. in *Hum. Herpesviruses Biol. Ther. Immunoprophyl.* (Arvin, A. et al.) (Cambridge University Press, 2007). at <http://www.ncbi.nlm.nih.gov/books/NBK47393/>

7. Levinson, W. *Review of Medical Microbiology and Immunology*. 659 (The McGraw-Hill Companies, 2006).

8. Steiner, I., Kennedy, P. G. E. & Pachner, A. R. The neurotropic herpes viruses: herpes simplex and varicella-zoster. *Lancet Neurol.* **6,** 1015–28 (2007).

9. Baringer, J. R. Herpes simplex infections of the nervous system. *Neurol. Clin.* **26,** 657–74, viii (2008).

10. Leuzinger, H. *et al.* Herpes simplex virus 1 envelopment follows two diverse pathways. *J. Virol.* **79,** 13047–13059 (2005).

11. Crump, C. M., Yates, C. & Minson, T. Herpes simplex virus type 1 cytoplasmic envelopment requires functional Vps4. *J. Virol.* **81,** 7380–7387 (2007).

12. Kelly, B. J., Fraefel, C., Cunningham, A. L. & Diefenbach, R. J. Functional roles of the tegument proteins of herpes simplex virus type 1. *Virus Res.* **145,** 173–86 (2009).

13. Zamora, M. R. DNA viruses (CMV, EBV, and the herpesviruses). *Semin. Respir. Crit. Care Med.* **32,** 454–70 (2011).

14. Schleiss, M. R. Persistent and recurring viral infections: the human herpesviruses. *Curr. Probl. Pediatr. Adolesc. Health Care* **39,** 7–23 (2009).

15. *Varicella-zoster Virus: Virology and Clinical Management*. 531 (Cambridge University Press, 2000). at <http://books.google.com/books?hl=en&lr=&id=_lKRzO1MJhkC&pgis=1>

16. Everett, R. D. HSV-1 biology and life cycle. *Methods Mol. Biol.* **1144,** 1–17 (2014).

17. *Alphaherpesviruses: Molecular Virology | Book*. 448 (Caister Academic Press, 2011). at <http://www.horizonpress.com/alphaherpesviruses>

18. Kaner, R. J. *et al.* Fibroblast growth factor receptor is a portal of cellular entry for herpes simplex virus type 1. *Science* **248,** 1410–3 (1990).

19. Tiwari, V. *et al.* Role for nectin-1 in herpes simplex virus 1 entry and spread in human retinal pigment epithelial cells. *FEBS J.* **275,** 5272–5285 (2008).

20. Phipps, W. *et al.* Persistent genital herpes simplex virus-2 shedding years following the first clinical episode. *J. Infect. Dis.* **203,** 180–187 (2011).

21. Schiffer, J. T. & Corey, L. Rapid host immune response and viral dynamics in herpes simplex virus-2 infection. *Nat. Med.* **19,** 280–90 (2013).

22. Mertz, G. J. Asymptomatic shedding of herpes simplex virus 1 and 2: implications for prevention of transmission. *J. Infect. Dis.* **198,** 1098–100 (2008).

23. Whitley, R. J. Herpes simplex encephalitis: adolescents and adults. *Antiviral Res.* **71,** 141–8 (2006).

24. Hokkanen, L. & Launes, J. Neuropsychological sequelae of acute-onset sporadic viral encephalitis. *Neuropsychol. Rehabil.* **17,** 450–77

25. Egan, K. P., Wu, S., Wigdahl, B. & Jennings, S. R. Immunological control of herpes simplex virus infections. *J. Neurovirol.* **19,** 328–45 (2013).

26. Xu, F. *et al.* Trends in herpes simplex virus type 1 and type 2 seroprevalence in the United States. *JAMA* **296,** 964–73 (2006).

27. Howard, M. *et al.* Regional Distribution of Antibodies to Herpes Simplex Virus Type 1 (HSV-1) and HSV-2 in Men and Women in Ontario, Canada. *J. Clin. Microbiol.* **41,** 84–89 (2003).

28. Singh, A., Preiksaitis, J., Ferenczy, A. & Romanowski, B. The laboratory diagnosis of herpes simplex virus infections. *Can. J. Infect. Dis. Med. Microbiol.* **16,** 92–8 (2005).

29. Chisholm, C. & Lopez, L. Cutaneous Infections Caused by Herpesviridae: A Review. *Arch. Pathol. Lab. Med.* **135,** 1357–1362 (2011).

30. Weber, F., Wagner, V., Rasmussen, S. B., Hartmann, R. & Paludan, S. R. Double-stranded RNA is produced by positive-strand RNA viruses and DNA viruses but not in detectable amounts by negative-strand RNA viruses. *J. Virol.* **80,** 5059–64 (2006).

31. Orzalli, M. H., DeLuca, N. A. & Knipe, D. M. Nuclear IFI16 induction of IRF-3 signaling during herpesviral infection and degradation of IFI16 by the viral ICP0 protein. *Proc. Natl. Acad. Sci. U. S. A.* **109,** E3008–17 (2012).

32. Ndjamen, B., Farley, A. H., Lee, T., Fraser, S. E. & Bjorkman, P. J. The herpes virus Fc receptor gE-gI mediates antibody bipolar bridging to clear viral antigens from the cell surface. *PLoS Pathog.* **10,** e1003961 (2014).

33. T'Hart, B. a. & Den Dunnen, W. F. Commentary on special issue: CNS diseases and the immune system. *J. Neuroimmune Pharmacol.* **8,** 757–759 (2013).

34. Kaushik, D. K., Gupta, M. & Basu, A. Microglial response to viral challenges: every silver lining comes with a cloud. *Front. Biosci. (Landmark Ed.* **16,** 2187–205 (2011).

35. Conrady, C. D., Drevets, D. A. & Carr, D. J. J. Herpes simplex type I (HSV-1) infection of the nervous system: is an immune response a good thing? *J. Neuroimmunol.* **220,** 1–9 (2010).

36. Bartel, D. P. MicroRNAs: Genomics, Biogenesis, Mechanism, and Function. *Cell* **116,** 281–297 (2004).

37. Chan, E. K. L., Yao, B. & Fritzler, M. J. Reflections on ten years of history of, and future prospects for, GW182 and GW/P body research. *Adv. Exp. Med. Biol.* **768,** 261–70 (2013).

38. Lewis, B. P., Burge, C. B. & Bartel, D. P. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* **120,** 15–20 (2005).

39. Park, S.-Y., Lee, J. H., Ha, M., Nam, J.-W. & Kim, V. N. miR-29 miRNAs activate p53 by targeting p85 alpha and CDC42. *Nat. Struct. Mol. Biol.* **16,** 23–9 (2009).

40. Zhang, Q. *et al.* Selective secretion of microRNA in CNS system. *Protein Cell* **4,** 243–7 (2013).

41. Gottwein, E. & Cullen, B. R. Viral and cellular microRNAs as determinants of viral pathogenesis and immunity. *Cell Host Microbe* **3,** 375–87 (2008).

42. Zou, Q., Mao, Y., Hu, L., Wu, Y. & Ji, Z. miRClassify: an advanced web server for miRNA family classification and annotation. *Comput. Biol. Med.* **45,** 157–60 (2014).

43. Jurak, I., Griffiths, A. & Coen, D. M. Mammalian alphaherpesvirus miRNAs. *Biochim. Biophys. Acta* **1809,** 641–53 (2011).

44. Betancur, J. G., Yoda, M. & Tomari, Y. miRNA-like duplexes as RNAi triggers with improved specificity. *Front. Genet.* **3,** 127 (2012).

45. Cullen, B. R. RNAi the natural way. *Nat. Genet.* **37,** 1163–5 (2005).

46. Sun, L. & Li, Q. The miRNAs of herpes simplex virus (HSV). *Virol. Sin.* **27,** 333–8 (2012).

47. Grey, F., Hook, L. & Nelson, J. The functions of herpesvirus-encoded microRNAs. *Med. Microbiol. Immunol.* **197,** 261–7 (2008).

48. Duan, F., Liao, J., Huang, Q., Nie, Y. & Wu, K. HSV-1 miR-H6 inhibits HSV-1 replication and IL-6 expression in human corneal epithelial cells in vitro. *Clin. Dev. Immunol.* **2012,** 192791 (2012).

49. Gupta, A., Gartner, J. J., Sethupathy, P., Hatzigeorgiou, A. G. & Fraser, N. W. Anti-apoptotic function of a microRNA encoded by the HSV-1 latency-associated transcript. *Nature* **442,** 82–5 (2006).

50. Conery, A. R. *et al.* Akt interacts directly with Smad3 to regulate the sensitivity to TGF-beta induced apoptosis. *Nat. Cell Biol.* **6,** 366–72 (2004).

51. Ghosh, Z., Mallick, B. & Chakrabarti, J. Cellular versus viral microRNAs in host-virus interaction. *Nucleic Acids Res.* **37,** 1035–1048 (2009).

52. Pedersen, I. M. *et al.* Interferon modulation of cellular microRNAs as an antiviral mechanism. *Nature* **449,** 919–922 (2007).

53. Umene, K. & Sakaoka, H. Evolution of herpes simplex virus type 1 under herpesviral evolutionary processes. *Arch. Virol.* **144,** 637–56 (1999).

54. Roberts, A. P. E., Lewis, A. P. & Jopling, C. L. The role of microRNAs in viral infection. *Prog. Mol. Biol. Transl. Sci.* **102,** 101–39 (2011).

55. Pare, J. M. & Sullivan, C. S. A host MicroRNA brokers truce with HSV-1. *Cell Host Microbe* **15,** 395–7 (2014).

56. Pan, D. *et al.* A neuron-specific host microRNA targets herpes simplex virus-1 ICP0 expression and promotes latency. *Cell Host Microbe* **15,** 446–56 (2014).

57. Jopling, C. L., Schütz, S. & Sarnow, P. Position-dependent function for a tandem microRNA miR-122-binding site located in the hepatitis C virus RNA genome. *Cell Host Microbe* **4,** 77–85 (2008).

58. Jopling, C. L., Yi, M., Lancaster, A. M., Lemon, S. M. & Sarnow, P. Modulation of hepatitis C virus RNA abundance by a liver-specific MicroRNA. *Science* **309,** 1577–1581 (2005).

59. Randall, G. *et al.* Cellular cofactors affecting hepatitis C virus infection and replication. *Proc. Natl. Acad. Sci. U. S. A.* **104,** 12884–9 (2007).

60. Lecellier, C.-H. *et al.* A cellular microRNA mediates antiviral defense in human cells. *Science* **308,** 557–60 (2005).

61. Zheng, S., Li, Y., Zhang, Y., Li, X. & Tang, H. MiR-101 regulates HSV-1 replication by targeting ATP5B. *Antiviral Res.* **89,** 219–26 (2011).

62. Bhela, S. *et al.* Critical role of microRNA-155 in herpes simplex encephalitis. *J. Immunol.* **192,** 2734–43 (2014).

63. Wang, F.-Z. *et al.* Human cytomegalovirus infection alters the expression of cellular microRNA species that affect its replication. *J. Virol.* **82,** 9065–74 (2008).

64. Thounaojam, M. C. *et al.* MicroRNA 155 regulates Japanese encephalitis virus-induced inflammatory response by targeting Src homology 2-containing inositol phosphatase 1. *J. Virol.* **88,** 4798–810 (2014).

65. Thounaojam, M. C., Kaushik, D. K., Kundu, K. & Basu, A. MicroRNA-29b modulates Japanese encephalitis virus-induced microglia activation by targeting tumor necrosis factor alpha-induced protein 3. *J. Neurochem.* **129,** 143–54 (2014).

66. Chaudhuri, A. D., Yelamanchili, S. V, Marcondes, M. C. G. & Fox, H. S. Up-regulation of microRNA-142 in simian immunodeficiency virus encephalitis leads to repression of sirtuin1. *FASEB J.* **27,** 3720–9 (2013).

67. Trobaugh, D. W. *et al.* RNA viruses can hijack vertebrate microRNAs to suppress innate immunity. *Nature* **506,** 245–8 (2014).

68. Janssen, H. L. A. *et al.* Treatment of HCV infection by targeting microRNA. *N. Engl. J. Med.* **368,** 1685–94 (2013).

69. Majer, A. Host Response to a Herpes Simplex Virus Type-1 Infection. 256 (2009).

70. Blankenberg, D. *et al.* Galaxy: a web-based genome analysis tool for experimentalists. *Curr. Protoc. Mol. Biol.* **Chapter 19,** Unit 19.10.1–21 (2010).

71. Goecks, J., Nekrutenko, A. & Taylor, J. Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. *Genome Biol.* **11,** R86 (2010).

72. Giardine, B. *et al.* Galaxy: a platform for interactive large-scale genome analysis. *Genome Res.* **15,** 1451–5 (2005).

73. Hackenberg, M., Rodríguez-Ezpeleta, N. & Aransay, A. M. miRanalyzer: an update on the detection and analysis of microRNAs in high-throughput sequencing experiments. *Nucleic Acids Res.* **39,** W132–8 (2011).

74. Ying, S.-Y. & Walker, J. M. *MicroRNA Protocols*. *Life Sci.* **531,** 375 (Humana Press, 2013).

75. Jurka, J. *et al.* Repbase Update, a database of eukaryotic repetitive elements. *Cytogenet. Genome Res.* **110,** 462–7 (2005).

76. Burge, S. W. *et al.* Rfam 11.0: 10 years of RNA families. *Nucleic Acids Res.* **41,** D226–32 (2013).

77. Paraskevopoulou, M. D. *et al.* DIANA-microT web server v5.0: service integration into miRNA functional analysis workflows. *Nucleic Acids Res.* **41,** W169–73 (2013).

78. Vlachos, I. S. *et al.* DIANA miRPath v.2.0: investigating the combinatorial effect of microRNAs in pathways. *Nucleic Acids Res.* **40,** W498–504 (2012).

79. Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* **25,** 402–8 (2001).

80. Hamilton, M. D., Nuara, A. A., Gammon, D. B., Buller, R. M. & Evans, D. H. Duplex strand joining reactions catalyzed by vaccinia virus DNA polymerase. *Nucleic Acids Res.* **35,** 143–51 (2007).

81. Worthington, M. T., Pelo, J. & Luo, R. Q. Cloning of random oligonucleotides to create single-insert plasmid libraries. *Anal. Biochem.* **294,** 169–75 (2001).

82. Cock, P. J. A., Fields, C. J., Goto, N., Heuer, M. L. & Rice, P. M. The Sanger FASTQ file format for sequences with quality scores, and the Solexa/Illumina FASTQ variants. *Nucleic Acids Res.* **38,** 1767–71 (2010).

83. Ewing, B., Hillier, L., Wendl, M. C. & Green, P. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res.* **8,** 175–85 (1998).

84. Ewing, B. & Green, P. Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Res.* **8,** 186–94 (1998).

85. Hackenberg, M., Sturm, M., Langenberger, D., Falcón-Pérez, J. M. & Aransay, A. M. miRanalyzer: a microRNA detection and analysis tool for next-generation sequencing experiments. *Nucleic Acids Res.* **37,** W68–76 (2009).

86. Meyer, S. U., Pfaffl, M. W. & Ulbrich, S. E. Normalization strategies for microRNA profiling experiments: a "normal" way to a hidden layer of complexity? *Biotechnol. Lett.* **32,** 1777–88 (2010).

87. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10,** R25 (2009).

88. Friedländer, M. R. *et al.* Discovering microRNAs from deep sequencing data using miRDeep. *Nat. Biotechnol.* **26,** 407–15 (2008).

89. Mathelier, A. & Carbone, A. MIReNA: finding microRNAs with high accuracy and no learning at genome scale and from deep sequencing data. *Bioinformatics* **26,** 2226–34 (2010).

90. Yang, C.-X. *et al.* Small RNA profile of the cumulus-oocyte complex and early embryos in the pig. *Biol. Reprod.* **87,** 117 (2012).

91. Nygaard, S. *et al.* Identification and analysis of miRNAs in human breast cancer and teratoma samples using deep sequencing. *BMC Med. Genomics* **2,** 35 (2009).

92. Zhang, J., Xu, Y., Huan, Q. & Chong, K. Deep sequencing of Brachypodium small RNAs at the global genome level identifies microRNAs involved in cold stress response. *BMC Genomics* **10,** 449 (2009).

93. Jackson, A. C. *Advances in Virus Research: Research Advances in Rabies: Vol 79 by Jackson, Alan C. (Editor): Academic Pr*. 486 (Academic Press, 2011). at <http://www.abebooks.com/Advances-Virus-Research-Rabies-Vol-79/5064417621/bd>

94. Skalsky, R. L. & Cullen, B. R. Viruses, microRNAs, and host interactions. *Annu. Rev. Microbiol.* **64,** 123–41 (2010).

95. Miranda, K. C. *et al.* A pattern-based method for the identification of MicroRNA binding sites and their corresponding heteroduplexes. *Cell* **126,** 1203–17 (2006).

96. Hofacker, I. L. *et al.* Fast folding and comparison of RNA secondary structures. *Monatshefte f??r Chemie Chem. Mon.* **125,** 167–188 (1994).

97. Roizman, B. The function of herpes simplex virus genes: a primer for genetic engineering of novel vectors. *Proc. Natl. Acad. Sci. U. S. A.* **93,** 11307–12 (1996).

98. McGeoch, D. J. *et al.* The complete DNA sequence of the long unique region in the genome of herpes simplex virus type 1. *J. Gen. Virol.* **69 (Pt 7),** 1531–74 (1988).

99. Bradbury, D. A., Simmons, T. D., Slater, K. J. & Crouch, S. P. Measurement of the ADP:ATP ratio in human leukaemic cell lines can be used as an indicator of cell viability, necrosis and apoptosis. *J. Immunol. Methods* **240,** 79–92 (2000).

100. The Gene Expression Nervous System Atlas (GENSAT) Project, NINDS Contracts N01NS02331 & HHSN271200723701C to The Rockefeller University (New York, NY). at <http://www.gensat.org/index.html>

101. Majer, A., Caligiuri, K. A., Niu, Y., Booth, T. F. & Booth, S. A. Profound induction of miR-200 family and miR-182 cluster microRNAs in acute herpes simplex virus 1 encephalitis. 29 [manuscript in preparation]

102. Jurka, J. Repbase update: a database and an electronic journal of repetitive elements. *Trends Genet.* **16,** 418–20 (2000).

103. Pruitt, K., Brown, G., Tatusova, T. & Maglott, D. in *NCBI Handb. [Internet]* (National Center for Biotechnology Information (US), 2012). at <http://www.ncbi.nlm.nih.gov/books/NBK21091/>

104. Kozomara, A. & Griffiths-Jones, S. miRBase: integrating microRNA annotation and deep-sequencing data. *Nucleic Acids Res.* **39,** D152–7 (2011).

105. Kim, V. N. & Nam, J.-W. Genomics of microRNA. *Trends Genet.* **22,** 165–73 (2006).

106. Kim, V. N., Han, J. & Siomi, M. C. Biogenesis of small RNAs in animals. *Nat. Rev. Mol. Cell Biol.* **10,** 126–39 (2009).

107. Motameny, S., Wolters, S., Nürnberg, P. & Schumacher, B. Next Generation Sequencing of miRNAs - Strategies, Resources and Methods. *Genes (Basel).* **1,** 70–84 (2010).

108. Lam, W.-Y. *et al.* Effect of avian influenza A H5N1 infection on the expression of microRNA-141 in human respiratory epithelial cells. *BMC Microbiol.* **13,** 104 (2013).

109. Banaudha, K. *et al.* MicroRNA silencing of tumor suppressor DLC-1 promotes efficient hepatitis C virus replication in primary human hepatocytes. *Hepatology* **53,** 53–61 (2011).

110. Ho, B.-C. *et al.* Enterovirus-induced miR-141 contributes to shutoff of host protein translation by targeting the translation initiation factor eIF4E. *Cell Host Microbe* **9,** 58–69 (2011).

111. Ramachandran, S. *et al.* Hepatitis C virus induced miR200c down modulates FAP-1, a negative regulator of Src signaling and promotes hepatic fibrosis. *PLoS One* **8,** e70744 (2013).

112. Lin, Z. *et al.* Differential expression of the miR-200 family microRNAs in epithelial and B cells and regulation of Epstein-Barr virus reactivation by the miR-200 family member miR-429. *J. Virol.* **84,** 7892–7 (2010).

113. Stark, T. J., Arnold, J. D., Spector, D. H. & Yeo, G. W. High-resolution profiling and analysis of viral and host small RNAs during human cytomegalovirus infection. *J. Virol.* **86,** 226–35 (2012).

114. Shinozaki, A. *et al.* Downregulation of microRNA-200 in EBV-associated gastric carcinoma. *Cancer Res.* **70,** 4719–27 (2010).

115. Ellis-Connell, A. L., Iempridee, T., Xu, I. & Mertz, J. E. Cellular microRNAs 200b and 429 regulate the Epstein-Barr virus switch between latency and lytic replication. *J. Virol.* **84,** 10329–43 (2010).

116. Chen, R.-F. *et al.* Augmented miR-150 expression associated with depressed SOCS1 expression involved in dengue haemorrhagic fever. *J. Infect.* (2014). doi:10.1016/j.jinf.2014.05.013

117. Bandyopadhyay, S. *et al.* Hepatitis C virus infection and hepatic stellate cell activation downregulate miR-29: miR-29 overexpression reduces hepatitis C viral abundance in culture. *J. Infect. Dis.* **203,** 1753–62 (2011).

118. Lu, F. *et al.* Epstein-Barr virus-induced miR-155 attenuates NF-kappaB signaling and stabilizes latent virus persistence. *J. Virol.* **82,** 10436–43 (2008).

119. Walsh, D. & Mohr, I. Phosphorylation of eIF4E by Mnk-1 enhances HSV-1 translation and replication in quiescent cells. *Genes Dev.* **18,** 660–72 (2004).

120. Allen, S. J. *et al.* Adaptive and innate transforming growth factor beta signaling impact herpes simplex virus 1 latency and reactivation. *J. Virol.* **85,** 11448–56 (2011).

121. Gao, Y. *et al.* DNA sequencing of TGFβ2 in sporadic patients with tetralogy of Fallot. *Exp. Ther. Med.* **3,** 878–880 (2012).

122. Rieck, P. *et al.* Fibroblast growth factor 2, heparin and suramin reduce epithelial ulcer development in experimental HSV-1 keratitis. *Graefe's Arch. Clin. Exp. Ophthalmol.* **235,** 733–740 (1997).

123. Frey, K. G. *et al.* HSV-1-induced SOCS-1 expression in keratinocytes: use of a SOCS-1 antagonist to block a novel mechanism of viral immune evasion. *J. Immunol.* **183,** 1253–62 (2009).

124. Yokota, S. *et al.* Induction of suppressor of cytokine signaling-3 by herpes simplex virus type 1 contributes to inhibition of the interferon signaling pathway. *J. Virol.* **78,** 6282–6 (2004).

125. Yoon, M. & Spear, P. G. Disruption of Adherens Junctions Liberates Nectin-1 To Serve as Receptor for Herpes Simplex Virus and Pseudorabies Virus Entry. *J. Virol.* **76,** 7203–7208 (2002).

126. Zhang, Y. *et al.* Secreted monocytic miR-150 enhances targeted endothelial cell migration. *Mol. Cell* **39,** 133–44 (2010).

127. Chen, X., Liang, H., Zhang, J., Zen, K. & Zhang, C.-Y. Secreted microRNAs: a new form of intercellular communication. *Trends Cell Biol.* **22,** 125–32 (2012).

128. Huang, S., Xie, Y., Yang, P., Chen, P. & Zhang, L. HCV core protein-induced down-regulation of microRNA-152 promoted aberrant proliferation by regulating Wnt1 in HepG2 cells. *PLoS One* **9,** e81730 (2014).

129. Phelan, A., Carmo-Fonseca, M., McLaughlan, J., Lamond, A. I. & Clements, J. B. A herpes simplex virus type 1 immediate-early gene product, IE63, regulates small nuclear ribonucleoprotein distribution. *Proc. Natl. Acad. Sci. U. S. A.* **90,** 9056–60 (1993).

130. Roizman, B. & Taddeo, B. The strategy of herpes simplex virus replication and takeover of the host cell. (2007). at <http://www.ncbi.nlm.nih.gov/books/NBK47362/>

131. König, R. *et al.* Global analysis of host-pathogen interactions that regulate early-stage HIV-1 replication. *Cell* **135,** 49–60 (2008).

132. Saad, A., Zhou, Z. H., Jakana, J., Chiu, W. & Rixon, F. J. Roles of Triplex and Scaffolding Proteins in Herpes Simplex Virus Type 1 Capsid Formation Suggested by Structures of Recombinant Particles. *J. Virol.* **73,** 6821–6830 (1999).

133. Desai, P., DeLuca, N. A., Glorioso, J. C. & Person, S. Mutations in herpes simplex virus type 1 genes encoding VP5 and VP23 abrogate capsid formation and cleavage of replicated DNA. *J. Virol.* **67,** 1357–64 (1993).

134. Foster, T. P., Melancon, J. M., Baines, J. D. & Kousoulas, K. G. The herpes simplex virus type 1 UL20 protein modulates membrane fusion events during cytoplasmic virion morphogenesis and virus-induced cell fusion. *J. Virol.* **78,** 5347–57 (2004).

135. Bogerd, H. P. *et al.* Replication of many human viruses is refractory to inhibition by endogenous cellular microRNAs. *J. Virol.* (2014). doi:10.1128/JVI.00985-14

136. Zhang, Y., Wang, Z. & Gemeinhart, R. A. Progress in microRNA delivery. *J. Control. Release* **172,** 962–74 (2013).

137. Hwang, D. W. *et al.* A brain-targeted rabies virus glycoprotein-disulfide linked PEI nanocarrier for delivery of neurogenic microRNA. *Biomaterials* **32,** 4968–75 (2011).

APPENDIX

1.0 DEEP SEQUENCING ANALYSIS PUBLISHED CHAPTER

Chapter 3

A User-Friendly Computational Workflow for the Analysis of MicroRNA Deep Sequencing Data

Anna Majer, Kyle A. Caligiuri, and Stephanie A. Booth

Abstract

Second-generation high-throughput sequencing is a robust and inexpensive methodology that is becoming an increasingly common technique for the study of microRNA (miRNA) expression levels in the central nervous system. This method allows for the identification of both known and novel miRNAs, reporting on the qualitative and quantitative levels these RNA species represent in any given sample. Numerous bioinformatic programs are currently available to analyze deep sequencing data but many require at least a partial understanding of the command line interface. In this chapter, we describe a user-friendly computational workflow guiding the user through the process from the initial FASTQ deep sequencing file to the identification of known and potentially novel miRNAs in a given experiment, as well as the assessment of the differential expression of these miRNAs between experimental samples. Furthermore, programs that can predict potential targets for these miRNAs are also highlighted.

Key words: MicroRNA, High-throughput sequencing, Central nervous system, Galaxy, miRanalyzer, Ingenuity Pathway Analysis, Program, Workflow, Target prediction, DIANA LAB microT

1. Introduction

The advent of "next generation sequencing" technologies has provided the tools with which to study miRNA expression changes on a global scale and at an unprecedented rate. This platform generates overarching data that provides a detailed glimpse into the global transcriptional events that occur under specific conditions. Nevertheless, this extensive data output, which typically spans millions of sequencing reads, can be challenging to analyze. This bottleneck is the focus of intense activity to design sophisticated software that is specifically tailored to analyze miRNAs. A rapidly increasing knowledge base relating to miRNA biogenesis and function is also driving the refinement of bioinformatic tools. The latest

Shao-Yao Ying (ed.), *MicroRNA Protocols*, Methods in Molecular Biology, vol. 936,
DOI 10.1007/978-1-62703-083-0_3, © Springer Science+Business Media, LLC 2013

advancements allow us to not only identify known miRNA sequences within a sample type and provide both a quantitative and qualitative assessment, but also to predict potentially novel miRNAs. Furthermore, the identification of isomiRs, sequence variants of mature miRNAs generated through RNA processing, or along with the identification of other related small RNA species are also obtainable by deep sequencing methods.

The complete process for miRNA deep sequencing data analysis can be divided into three phases: Primary, secondary, and tertiary analysis (1). Primary analysis covers the initial processing of the image files that are generated during the sequencing run. In this step, each nucleotide base is represented by a different color, resulting in four unique image files. All four images are then aligned to each other and the light intensities are normalized, analyzed, and filtered for base-calling and quality scoring to yield FASTQ and/or FASTA files (2). This analysis is commonly performed by the software that accompanies the sequencer and is beyond the scope of this chapter. One of the most important considerations for miRNA deep sequencing data is the proper exclusion of all sequences that are not miRNAs, such as RNA degradation products, open reading frames, and other noncoding RNA sequences. These steps are performed in both secondary and tertiary phases of analysis (1).

Typically, the secondary analysis encompasses trimming and quality filtering of raw reads and alignment with the reference genome. This part of the analysis can be done using the Galaxy program; an open source, Web-based interface encompassing many different bioinformatic tools (3, 4). In contrast to some protocols, we do not perform the genome alignment during the initial stage of analysis using Galaxy, but rather this is part of the tertiary phase where it forms an important component of the novel miRNA prediction module. Considering that the initial step is to look for known miRNAs, the entire dataset is compared to the annotated sequences in miRBase (5) so as to include as much information as possible. Thus, starting with a FASTQ file, we first remove the 3¢ adaptor motif, an artifact of the sequencing process, using Galaxy (Fig. 1). This is executed by the Clip tool which also limits poorly sequenced reads from subsequent filtering steps, and thus, increases the processing speed of the analysis. These imprecise sequencing reads may represent a lack of the 3¢ adaptor sequence or consist of less than 15 base pairs in length, too short to be mature miRNAs. Additional filtering performed by the Filter FASTQ tool is used to select for the highest quality reads in the data. More specifically, we employ a Phred33 (6) score of 20; this threshold ensures 99% accuracy in base-calling (7, 8), removing transcripts that contain more than three bases which fall below this threshold. In addition, the filtered output file is converted to a FASTA format, which is readily usable for analysis by many programs, and is easily accomplished using the FASTQ to FASTA tool.

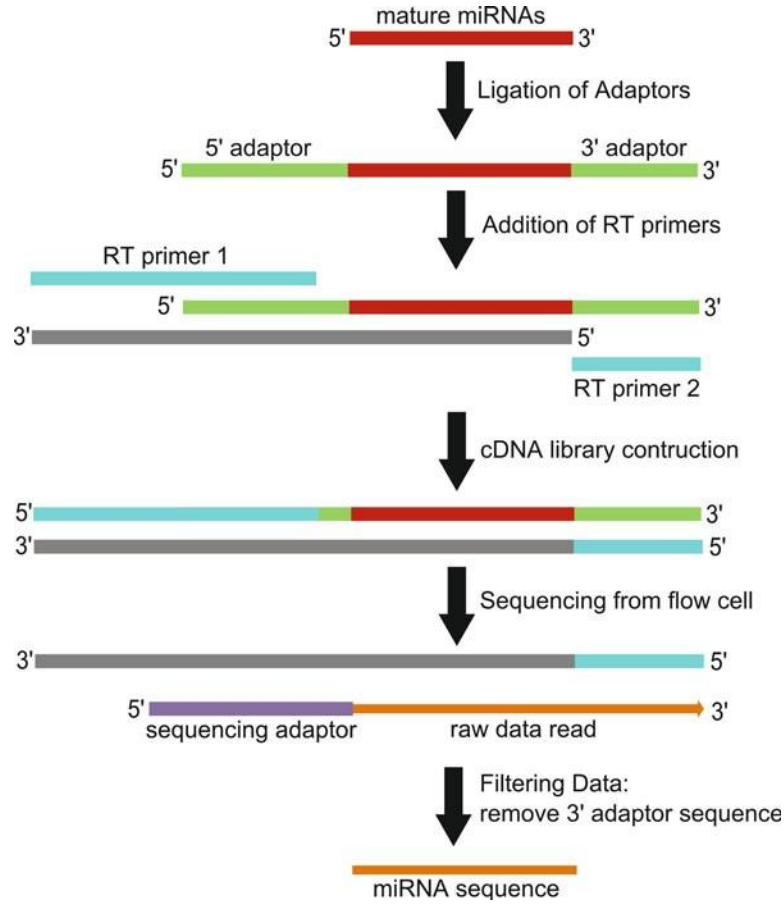


Fig. 1. A brief outline highlighting the sample library preparation. Small RNAs (*red*) are isolated from total RNA followed by the ligation of 3¢ and 5¢ adaptors (*green*). RT-PCR primers (*blue*) are used to amplify the sample and create a stable cDNA library. The DNA is then denatured and loaded onto the flow cell. The RT primer 2 corresponds to a surface bound primer to attach the antisense strand. The sequencing adaptor (*purple*) binds the antisense strand for read synthesis. The raw data read (*orange*) starts after the sequencing adaptor and incorporates the small RNA and 3¢ adaptor sequence. This adaptor sequence must be removed during secondary analysis before aligning to a reference such as miRBase.

The tertiary analysis starts with miRanalyzer (9, 10), a program that aligns all input sequencing reads to libraries of known mature miRNAs such as miRBase. Differential expression analysis of multiple replicates can also be performed via miRanalyzer. The remaining, unmapped sequences are filtered for potential open reading frames and other noncoding RNAs using databases such as RefSeq (11) and Rfam (12). The sequences that do not fall within any of these databases are aligned to the reference genome and further analyzed to predict potentially novel miRNAs identified in at least three different prediction models. The output file containing both known and novel miRNAs can be subjected to additional

bioinformatic analyses to identify the predicted genes of these miRNAs together with any ontological significance.

Prediction of miRNA targets can be done in two ways, depending on the nature of the candidate miRNA. If the miRNA has been annotated previously, target prediction can be done using Ingenuity Pathway Analysis (IPA), which incorporates both predictions from TargetScan (13–15) and experimentally obser ved miRNA:mRNA associations from TarBase (16, 17). IPA also provides a sophisticated ontological analysis of gene interactions that can be used to further annotate candidate miRNA targets for roles in biological processes involved in central ner vous system (CNS) function and dysfunction. If the miRNA of interest is a novel miRNA, the DIANA LAB microT v3.0 program (18, 19) is a suitable starting point to identify potential gene targets.

The protocol described in this chapter encompasses a comprehensive workflow starting with the initial high-throughput sequencing data file and resulting in the identification of known and novel miRNAs. Included in the analysis are the differential expression of these miRNAs between different sample types as well as the prediction of potential target transcripts and their respective functional pathways (Fig. 2).

2. Materials

This protocol was designed to be as user-friendly as possible (advanced computer literacy is not required), and most steps can be performed online. Almost all segments of this protocol are performed explicitly using Windows, although the use of Ubuntu, a version of Unix (which can be found at http://www.ubuntu.com/download), is also required. Freeware programs for processing miRNA deep sequencing data that are highlighted in this chapter include Galaxy (available at http://main.g2.bx.psu.edu/) and miRanalyzer (available at http://bioinfo2.ugr.es/miRanalyzer/miRanalyzer.php). The use of IPA (Ingenuity Systems®, http://www.ingenuity.com) for target pathway exploration of known miRNAs will also be briefly described. Target prediction for novel miRNAs can be done using the DIANA LAB microT v3.0 program (available at http://diana.cslab.ece.ntua.gr/microT/).

3. Methods

3.1. Galaxy: Trimming and Filtering Sequence Reads

1. Large files (>2 GB) must be uploaded via an FTP server by opening any directory or folder within Windows and entering the following into the address bar: ftp://main.g2.bx.psu.

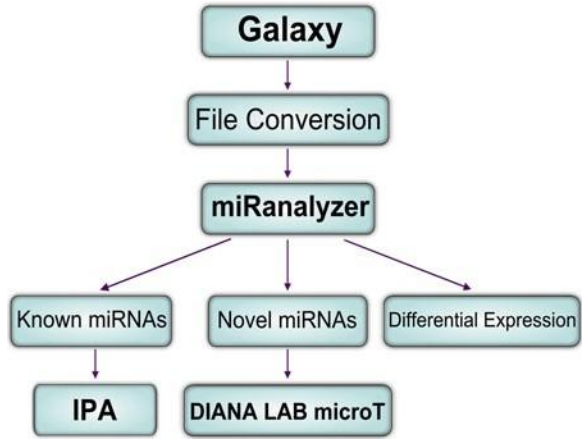


Fig. 2. An overall workflow designed to process miRNA deep sequencing data for both known and novel miRNA sequences, which are then used by additional programs for potential mRNA target identification. Furthermore, differential expression for these high-throughput platforms can also be performed using this workflow. The bolded text represents programs discussed in this chapter.

edu/. A login window will appear prompting you to enter your Galaxy credentials (i.e., user name and password). Press "Log On" and copy and paste your data files into this window.

2. On the Galaxy Web site, go to "User", then "Login". Enter your user name and password then click "Login". In the Tools menu on the left, go to "Get Data" then "Upload File". Under "File Format" select fastqsanger (see Note 1). Choose your files from the "Files uploaded via FTP" section. Under "Genome", the organism from which your sample originated from may be selected by choosing the appropriate descriptor from the pull-down menu (see Note 2). Click "Execute", and you will see your files show up in the History on the right hand side.
3. To trim the 3¢ adaptor sequences and to filter your data, import the miRNA workflow that is located at http://main.g2.bx.psu.edu/u/kyle-caligiuri/w/mirna-secondary-analysis by clicking "Import workflow" found on the top right hand side of the main pane. Next, click "Workflow" located at the top of the page and choose the workflow entitled "imported: miRNA Secondary Analysis". Select "Run" from the pop-up menu (see Note 3).
4. Select your data using the drop-down menu below "Library to clip" in "Step 1: Clip" then click "Run workflow".
5. Once all your files have been trimmed and filtered, you can download them by clicking on the name of the file in the green box to expand the box, then clicking the floppy disk icon to Download.

6. To use these files in the miRanalyzer program, the output file(s) need to be converted to a space-separated format which is performed using Ubuntu. Open a terminal in Ubuntu (located under Applications > Accessories) and change to the directory where your files are located. For example, if your files are located in a folder labeled Control in your Documents, type the following and then press Enter:

   ```
   cd Documents/Control
   ```

7. To convert the output data file from Galaxy to the space-separated format, type the following command and press Enter (see Note 4):

   ```
   grep –d ">" input.fasta | sort | uniq –c > sorted.txt
   ```

8. The output file from step 7 needs to be flipped for performing miRanalyzer. This can be achieved by opening the Text Editor (also located under Applications > Accessories) and enter the following two lines:

   ```
   #!/usr/bin/sed –f
   s/^\s\+(\w\+\)\s*\(\w*\)$/\2\1/
   ```

9. Save the file as "flip_script" and ensure it is in the same folder as your datasets (although you can copy and paste the file to other folders if needed).
10. In the terminal, enter the following command for each newly sorted file:

    ```
    cat sorted.txt | sed –f flip_script > output.txt
    ```

3.2. miRanalyzer: Identifies Known miRNAs, Predicts Novel miRNAs, and Compares Datasets to Give you Differential Expression Changes

1. Upload your "output.txt" file by clicking "Browse", and choose the species your sample originated from using the pull-down menu.
2. The following settings should be selected: Zero for the number of mismatches (known miRNAs); one for the number of mismatches (library); select two for the number of mismatches (genome); the threshold of the posterior probability should be set to 0.95; and the minimum number of models which predict the miRNA should be left at the default 3 (see Note 5).
3. Click on the "Launch Program" button for the analysis to commence at which point you will be redirected to an individual page with the following URL:

   http://bioinfo2.ugr.es/miRanalyzer/miRanalyzer.php?launched=true&id=__

   where the "__" represents a unique miRanalyzer job ID number for the submitted analysis (see Note 6).

4. The resulting output analysis, represented by each unique job ID number, will contain sequence information for known and novel miRNAs (see Note 7), which can then be submitted for differential expression analysis by pressing the "Differential Expression" button on the left hand side (see Note 8).
5. Enter the unique miRanalyzer job ID corresponding to the control sample into the text box labeled Group 1 and the treatment job ID into Group 2 and click "Run" (see Note 9).

3.3. Target Prediction for Known and Novel miRNAs Identified from Deep Sequencing Data

3.3.1. IPA: Exploring miRNA Targets and Functional Pathways

1. Make an Excel spreadsheet of the candidate miRNA(s) by listing each unique miRNA in individual rows under the same column with each miRNA being annotated as per miRBase (i.e., mmu-miR-1). Fold change values for each miRNA may also be included in subsequent columns if desired (see Note 10).
2. Upload your miRNA data file by selecting File > Upload Dataset and select the appropriate file and press "Open". Provide information regarding the column headers such that the first column is the miRNA designation or "ID" and if applicable, an additional column contains fold change data that should be designated as "observation 1" and "fold change" in the pull-down menus. Save the file under the appropriate project folder or make a new folder if needed.
3. Once the file is uploaded, the user sees an overall summary of the candidate miRNAs and all additional information. Select "analyze/filter dataset" located in the bottom right corner, then select "miRNA Target Filter" from the pop-up menu to open the analysis window.
4. The "miRNA Target Filter" provides options for filtering candidate miRNA targets and further exploration of the mRNA's functional associations. One can filter mRNA targets for each miRNA by editing the "Relationship > Confidence" column via pressing the "Filter" icon next to that column's title and selecting for moderate, high, or experimentally observed relationships, or a combination of the three. Furthermore, additional columns can be added or removed by using the buttons found on the right hand side of the columns and the data can be sorted accordingly.
5. One of the most important and useful options offered by the miRNA target filter is the inclusion of mRNA data from your own experiments. To do this, select "add/replace mRNA dataset" found at the top, underneath the toolbar > select the mRNA analysis in "My Projects" and press "Next". One may also upload a desired dataset not found in IPA.
6. The "expression pairing" option can be used to filter for the inverse relationship between the miRNA and potential target (i.e., if miRNA is up-regulated the target mRNA should be

down-regulated). This can be achieved by pressing the "expression pairing" button (located next to "add/replace mRNA dataset") and selecting the appropriate data for the miRNA and mRNA expression data columns to use via the pull-down menus and press "OK". Use the "filter" option in the newly-incorporated Expression Pairing column located within the Relationship column such that the fold change of the miRNA is inversely proportional to the mRNA fold change (see Note 11). One can save the analysis by pressing File > Save and/or exporting the data by pressing the "Export Data" icon found at the top of the pane, underneath the Details tab.

7. Once the target gene list has been filtered in the desired manner, pathway associations and functions can be further explored. Select all the genes chosen to be included in the pathway by checking the left hand column boxes and selecting "add to my pathway" above this column, then select "New My Pathway" > "Both" > "Ok". Highlight all the genes by making a box around the newly-created figure, then select "Build". From the "Tool" menu, select "Connect" then press "Apply". Make sure to keep the direct and indirect interactions checked, unless deemed otherwise.
8. For doing a core analysis on the newly selected mRNA candidate list, return to your dataset by pressing "Minimize" at the top right hand corner of the "My Pathways" window. Press "add to my list" > "New My List" > "Both" > "Ok". At this point name and save this file using the "Save My List" button, then press the right arrow icon found on the top right corner of the Lists window, select "Run Core Analysis". Select the appropriate General Settings and other preferences (see Note 12) and then run the analysis by pressing "Run Analysis". You will be prompted to save this analysis, and press "Ok" to continue.

3.3.2. DIANA LAB microT v3.0: Target Prediction for Novel miRNAs

1. On the home page, select the "OR predict for your miRNA sequence" option highlighted in red which is located under the "search for targets of a specific miRNA" panel.
2. Type in your email address to receive the results of the analysis, select the species the novel miRNA originated from and input the sequence in the "file content" window (see Note 13) and select "run job".

4. Notes

1. The data must be in FASTQ Sanger format for this protocol (i.e., Phred33 quality scoring is essential). If your data is already in FASTQ Sanger format and you choose Auto-detect, Galaxy will only establish that it is a FASTQ format but it cannot tell

what scoring method was used. Furthermore, the tools in Galaxy only recognize FASTQ Sanger format, so if your data is not formatted properly, you will have to convert your data to this format after uploading. If this is the case, convert the file format to FASTQ Sanger via "NGS: QC and manipulation" in the Tools pane on the left, then select "FASTQ Groomer" and upload your file.

2. The Genome section is mostly for organizational purposes and is more useful when performing tertiary analysis on Galaxy, which is not the case for this protocol.
3. If desired, the specific settings employed in the workflow can be changed to increase or decrease stringency. In addition, if your data is not from an Illumina sequencer, you will have to change the sequence of the 3′ adaptor that is being trimmed. To do this, select "Edit" from the pop-up menu which will bring you to a workflow of the secondary analysis steps performed by Galaxy. You can click an individual module and its options will display on the right in the Details pane. When you are finished, click Options > Save at the top right of the main pane.
4. The "input.fasta" is the file that was downloaded from Galaxy and "sorted.txt" is the file that will be generated using this command. Sorted.txt is a file containing the number of reads for each transcript and can be renamed if necessary.
5. We did not see any significant differences in our results when setting parameters to the most stringent options (i.e., zero mismatches allowed in "known miRNA", "library" or "genome" and a 0.99 threshold of the posterior probability with three models which predict a novel miRNA). Also, a stand-alone version of miRanalyzer is downloadable where the user can change additional parameters if desired.
6. Avoid moving away from this page until you are able to bookmark the page and/or mark down the jobID so that you can return to your data later.
7. The output analysis produced similar results to an independent analysis performed by LC Sciences (http://www.lcsciences.com/) where the miRNAs that were deemed abundantly present were also the top contenders identified in the analysis done by miRanalyzer.
8. The prediction of novel miRNAs is a very slow process and may take up to one week. Considering that the differential expression analyses take into account novel miRNA reads, it is best not to perform this analysis until the initial job is fully complete.
9. miRanalyzer is able to perform differential expression on multiple replicates of the sequencing runs for the same sample.

This can be performed by separating each job ID using ":". For example, for Group 1 enter jobID1:jobID2:jobID3.

10. It will be advantageous to include fold changes in the upload file because options in IPA exist to pair miRNA expression levels with potential mRNA targets, which further helps to refine the candidate mRNA target list.
11. An inverse relationship between the miRNA and candidate target mRNA is not necessarily true in all cases, hence, caution should be taken when filtering potential miRNA targets in this manner.
12. The General Settings reflect the stringency of the analysis that is being performed on the select list of genes. For CNS work, choose under the "Tissues and Cell Lines" preferences the "Tissues and Primary Cells" > "Nervous System" and the "Cell Line" > "CNS Cell Lines" options. Also, by selecting the species the data originated from in the "Species" preferences, it will make the analysis more relevant to your particular experiment.
13. Up to five novel miRNAs can be analyzed by this program at the same time. The input for all miRNAs must be in the FASTA format which consists of the miRNA name or identifier and the nucleotide sequence as shown below following two lines separated by the "Enter" key:

    >mmu-miR-124

    UAAGGCACGCGGUGAAUGCC

Acknowledgments

We would like to extend our sincere gratitude to Malachi Griffith for providing the grep command, Adam Majer for writing the "flip_script" and Bernard Abrenica for the technical expertise with Ubuntu and the command line.

References

1. Baskin D, Olson EN, Lucas L, Smith T (2009) GeneSifter: next generation data management and analysis for next generation sequencing. Geospiza. http://www.geospiza.com/Products/WhitePaper_06102009.pdf. Accessed 30 June 2011
2. Werner VB (2008) Primary Analysis: from sample to wig file in 8 easy steps. Presented at 1st workshop on high throughput sequencing, Switzerland, October 2008. http://werner.yellowcouch.org/Papers/pippres08/index.html. Accessed 30 June 2011
3. Blankenberg D, Von Kuster G, Coraor N, Ananda G, Lazarus R, Mangan M, Nekrutenko A, Taylor J (2010) Galaxy: a web-based genome analysis tool for experimentalists. Curr Protoc Mol Biol. In: Frederick M. Ausubel [et al.] Chapter 19:Unit 19.10.11–21

4. Goecks J, Nekrutenko A, Taylor J, Galaxy T (2010) Galaxy: a comprehensive approach for supporting accessible, reproducible, and trans- parent computational research in the life sci- ences. Genome Biol 11(8):R86
5. Kozomara A, Griffiths-Jones S (2011) miR- Base: integrating microRNA annotation and deep-sequencing data. Nucleic Acids Res 39 (Database issue):D152–157
6. Cock PJ, Fields CJ, Goto N, Heuer ML, Rice PM (2010) The Sanger FASTQ file format for sequences with quality scores, and the Solexa/ Illumina FASTQ variants. Nucleic Acids Res 38(6):1767–1771
7. Ewing B, Green P (1998) Base-calling of auto- mated sequencer traces using phred. II. Error probabilities. Genome Res 8(3):186–194
8. Ewing B, Hillier L, Wendl MC, Green P (1998) Base-calling of automated sequencer traces using phred. I. Accuracy assessment. Genome Res 8(3):175–185
9. Hackenberg M, Rodriguez-Ezpeleta N, Aransay AM (2011) miRanalyzer: an update on the detection and analysis of microRNAs in high-throughput sequencing experiments. Nucleic Acids Res 39(Web Server issue):W132–138
10. Hackenberg M, Sturm M, Langenberger D, Falcon-Perez JM, Aransay AM (2009) miRan- alyzer: a microRNA detection and analysis tool for next-generation sequencing experiments. Nucleic Acids Res 37 (Web Server issue):W68–76
11. Pruitt KD, Tatusova T, Klimke W, Maglott DR (2009) NCBI Reference Sequences: current status, policy and new initiatives. Nucleic Acids Res 37 (Database issue):D32–36
12. Gardner PP, Daub J, Tate J, Moore BL, Osuch IH, Griffiths-Jones S, Finn RD, Nawrocki EP, Kolbe DL, Eddy SR, Bateman A (2011) Rfam: wikipedia, clans and the "decimal" release. Nucleic Acids Res 39 (Database issue):D141–145
13. Lewis BP, Burge CB, Bartel DP (2005) Conserved seed pairing, often flanked by ade- nosines, indicates that thousands of human genes are microRNA targets. Cell 120(1):15–20
14. Grimson A, Farh KK, Johnston WK, Garrett- Engele P, Lim LP, Bartel DP (2007) MicroRNA targeting specificity in mammals: determinants beyond seed pairing. Mol Cell 27(1):91–105
15. Friedman RC, Farh KK, Burge CB, Bartel DP (2009) Most mammalian mRNAs are con- served targets of microRNAs. Genome Res19(1):92–105
16. Sethupathy P, Corda B, Hatzigeorgiou AG (2006) TarBase: a comprehensive database of experimentally supported animal microRNA targets. RNA 12(2):192–197
17. Papadopoulos GL, Reczko M, Simossis VA, Sethupathy P, Hatzigeorgiou AG (2009) The database of experimentally supported targets: a functional update of TarBase. Nucleic Acids Res 37 (Database issue):D155–158
18. Maragkakis M, Reczko M, Simossis VA, Alexiou P, Papadopoulos GL, Dalamagas T, Giannopoulos G, Goumas G, Koukis E, Kourtis K, Vergoulis T, Koziris N, Sellis T, Tsanakas P, Hatzigeorgiou AG (2009) DIANA-microT web server: elucidating microRNA functions through target prediction. Nucleic Acids Res 37 (Web Server issue):W273–276
19. Maragkakis M, Alexiou P, Papadopoulos GL, Reczko M, Dalamagas T, Giannopoulos G, Goumas G, Koukis E, Kourtis K, Simossis VA, Sethupathy P, Vergoulis T, Koziris N, Sellis T, Tsanakas P, Hatzigeorgiou AG (2009) Accurate microRNA target prediction correlates with protein repression levels. BMC bioinformatics 10 (Journal Article):295