Developing an *in vitro* model for investigating the effects of the HSV-1 latency-associated ncRNAs in human neurons

Amy Isabella Jacobs

Thesis for the degree of Doctor of Philosophy (PhD)

Imperial College London
Department of Medicine
Section of Virology

September 2020

Declaration of originality

I declare that this thesis is entirely the result of my own work, except where appropriately acknowledged or referenced.

Amy Jacobs

Copyright declaration

Acknowledgements

First and foremost, I would like to thank my supervisors Mike, Stacey and Peter. Stacey, I can't thank you enough for all the advice and support you've given me throughout my PhD and the warmth with which you provided it. I especially want to thank you for your help and encouragement whilst I was writing up my thesis, over countless (and sometimes very long!) zoom chats – I couldn't have done it without you! Mike, thank you for all your help during my PhD and for your friendship. This thesis would not be the same and I would not be the same scientist that I am now without all the work you did and the training and encouragement you gave me. And thanks for teaching me Hive! Peter, thank you for all your guidance, even over a drink at the pub, I am grateful to have heard some of your great scientific musings! Thank you for your voice comments on my thesis – they were extremely helpful and entertaining!

Nicola, thank you for being such a supportive line-manager, especially considering you simultaneously had an entire division to run amidst a pandemic! Martin, Tom, Mark and Ryan, thank you for all the sequencing and bioinformatics work you did for this thesis and Martin and Tom also for answering my many questions! Thank you to NIBSC for funding my PhD.

Thanks to my family and friends for their love and encouragement. Thank you, Mum and Dad, for supporting me through my PhD. Mum thanks for using your editing eye over some of the pages of this thesis, despite not really understanding the content! Also thank you for appreciating how hard I worked for it. Dad thank you for discussing writing techniques with me and for always being proud of me. Joseph, although not always appreciated in the moment, thanks providing distractions from all things herpes, with jokes and TV suggestions.

A special thank you to my best friends Maia and Lia. You are always there for me, there to celebrate the successes and cheer me on (often believing in me more than I do!) but also there to pick up the pieces when things don't go to plan! I love and appreciate you both so much. Mya, thank you for our spontaneous hangouts when I was too busy to plan in anything and reminding me to have fun! Thank you to Mitch for all our conversations.

Thank you to all my friends at NIBSC who made my PhD an enjoyable time. Emma, thanks for being my "best work friend!" and for checking in on me when I was left alone in the lab! Tom for all the little favours you did, no questions asked. And to Machaela for brightening up the start of my PhD. Thank you to my fellow PhD student friends who understood the unique experience that is a PhD! And thanks to all my lunchtime/games/office friends for the banter!

Finally, thank you to anyone who reads this thesis, while I find the content interesting, I appreciate that for some of you Herpes isn't as much of a topic of fascination as it is to me!

Abstract

Herpes simplex virus 1 (HSV-1) persists for the lifetime of the host due to establishing latency in sensory neurons. During latency, the only abundantly transcribed HSV-1 gene is the latency-associated transcript (LAT), which is processed into the 1.5kb or 2.0kb major LAT intron and several microRNAs. These latency-associated non-coding RNAs (ncRNAs) have been reported to impact the establishment, maintenance and reactivation from latency. However, the molecular mechanisms of these ncRNAs are not fully characterised, especially in the context of human neurons.

In this study an *in vitro* model was developed to investigate the roles of the latency-associated ncRNAs in human neurons by establishing a method to differentiate SH-SY5Y cells into neurons and deliver the latency-associated ncRNAs. To achieve this, the cells were infected with a replication-defective HSV-1 mutant, in1382, that establishes a quiescent infection in which LAT is strongly expressed. Alternatively, lentiviruses were engineered to express the first 3.1kb of LAT, without or with mutations in splice sites that prevents splicing of the major LAT intron, or five HSV-1 microRNAs, shown to be abundant in latently infected human ganglia.

To investigate how the latency-associated ncRNAs affects the human neuronal transcriptome, RNA-Seq was performed on uninfected versus infected/transduced differentiated SH-SY5Y cells. Over 1500 significantly differentially expressed genes (DEGs) were identified, including 179 DEGs that overlapped following quiescent infection or lentivirus delivery of the latency-associated ncRNAs. A subset of these gene changes was validated by PCR.

Various cellular pathways and functions were found to be associated with the DEGs found. This included neuroprotection, control of the Wnt pathway, regulation of the cell-cycle and links to neurodegeneration. Some of these functions have been implicated in previous HSV-1 latency data. This work provides insight into putative roles of the latency-associated ncRNAs that could have implications on how HSV-1 latency affects human neurons.

Table of Contents

List of figures

List of Tables

Abbreviations

β-gal – β-galactosidase

3’ – 3 prime

5’ – 5 prime

ABS – analytical and biological sciences

ACV – acyclovir

AGO – Argonaute

ATRA – *all-trans* retinoic acid

BARTs- BamH1 A rightward transcripts

BDNF – brain-derived neurotropic factor

bp – base pair

CN – catalogue number

CNS – central nervous system

CPE – cytopathic effect

Ct – cycle threshold

CXA-D - contextual analysis of DNA

d – day(s)

DAVID - The Database for Annotation, Visualization and Integration Discovery

DEGs – differentially expressed genes

DMEM – Dulbecco’s Modified Eagle Medium

DNA – deoxyribonucleic acid

dp – decimal places

dpi – days post infection

dpt – days post transduction

DRG – dorsal root ganglia

dsDNA – double-stranded DNA

EBV – Epstein-Barr virus

ECM – extracellular matrix

eGFP – enhanced green fluorescent protein

EtBr – ethidium bromide

FCS – foetal calf serum

FDR – false discovery rate

FU – functional units

gX – glycoprotein X (where x = B/C/D/H/L)

GFP – green fluorescent protein

GO – gene ontology

gpc – genomes per cell

h – hour(s)

HCF – host cell factor

HCMV – human cytomegalovirus

HDAC – histone deacetylase

HHV – human herpesvirus

hiFCS – heat inactivated foetal calf serum

hPGK – human phosphoglycerate kinase

hpi – hours post infection

hpt – hours post transduction

HS – hyperthermic stress

HSPGs – heparin-sulfate proteoglycans

HSV – herpes simplex virus

HSV-1 – herpes simplex virus 1

HSV-2 – herpes simplex virus 2

ICP – infected cell protein

IE – immediate early

IF - immunofluorescence

IFN - interferon

IFNa – interferon alpha

IPSCs – induced pluripotent stem cells

IR – inverted repeat

ISH – *in situ* hybridisation

Kb - kilobase

KDa - Kilodaltons

KEGG - Kyoto Encyclopedia of Genes and Genomes

KSHV – Kaposi's sarcoma-associated herpesvirus

$_{L}$ – long

LAT – latency-associated transcript

LB – luria broth

mM – millimolar

MAP2 – microtubule associated protein 2

MEM – minimum essential medium

MIEP – major immediate-early promoter

miRNA – microRNA

mRISC – miRNA-RISC

mRNA – messenger RNA

MOI – multiplicity of infection

MS – multiple sclerosis

MTOR– mammalian target of rapamycin

MTORC1 – MTOR-containing complex

ncRNA – non-coding RNA

NF-M – neurofilament medium

NGF – nerve growth factor

NIBSC – national institute for biological standards and control

nm - nanometres

NSLAT – non-splicing LAT

nt – nucleotides

Oct – octamer binding protein

padj – adusted p-value

PBS – phosphate-buffered saline

pc – per cell

PCR – polymerase chain reaction

pfu – plaque forming units

PI3-K – phosphatidylinositol 3-kinase

PIGT – phosphatidylinositol glycan anchor biosynthesis class T

PML – promyelocytic leukaemia

pNF-H – phosphorylated neurofilament Heavy subunit

PNS – peripheral nervous system

pre-miRNA – precursor miRNA

qPCR – quantitative polymerase chain reaction

RA – retinoic acid

RING – really interesting new gene

RISC – RNA-induced silencing complex

RNA – ribonucleic acid

RNA-Seq – RNA-Sequencing

RS – RNAscope

RT – reverse transcription

RTemp – room temperature

$_S$ – short

SG – SYBR Green

SNAP – sensory neuron-associated phenotype

sncRNA – small non-coding RNA

SNP – single nucleotide polymorphism

SYP – synaptophysin

TG – trigeminal ganglia

TH – tyrosine hydroxylase

TK – thymidine kinase

TM - TaqMan

TR – terminal repeat

TUNEL - terminal deoxynucleotidyl transferase-mediated nick-end labelling

UL – unique long

US – unique short

UTR – untranslated region

VSV-G – vesicular stomatitis virus glycoprotein

VZV – varicella zoster virus

WB – western blot

WPRE – woodchuck hepatitis virus post-transcriptional regulatory element

WT – wildtype

X-gal – 5-bromo-4-chloro-3-indolyl-β-D-galactosidase

Chapter 1: Introduction

1.1 The Herpesvirales

The order *Herpesvirales* contains double-stranded DNA (dsDNA) viruses with genomes ranging from 125-290 kilobase (kb) pairs with a shared virion morphology. The linear dsDNA genome resides within an icosahedral capsid with a diameter of approximately 125 nanometres (nm). The nucleocapsid is surrounded by a proteinaceous matrix known as the tegument, which is encased by a lipid envelope that contains a number of membrane-associated proteins including embedded viral glycoproteins (Davison et al., 2009; Roizmann et al., 1992)(figure 1.1).

The order *Herpesvirales* can be divided into three families; the *Alloherpesviridae, Malacoherpesviridae* and *Herpesviridae.* Members of the *Herpesviridae* family share common biological features. For example, all *Herpesviridae* members encode many enzymes that contribute to viral DNA synthesis, nucleic acid metabolism and protein processing. Additionally, viral DNA synthesis and viral assembly occurs in the cell nucleus for these viruses. Each virus also can undergo a lytic infection, whereby infectious progeny virions are produced, resulting in cell death. Alternatively, the herpesviruses establish a latent infection, during which the genome is present, but no infectious virions are produced, this type of infection can persist for the lifetime of the host.

The *Herpesviridae* family contains all nine human herpesviruses and can be further divided into three sub-families: *Alphaherpesvirinae, Betaherpesvirinae* and *Gammaherpesvirinae.* Classification into each sub-family was originally decided by a number of biological properties such as cell tropism but given the advancement of sequencing techniques can now also be determined by amino acid sequence similarity across core viral genes (Davison, 2010; McGeoch et al., 1995). The human herpesviruses are spread across these three sub-families as highlighted in table 1.1.

As indicated in table 1.1, herpes simplex virus 1 (HSV-1) is a member of the *Alphaherpesvirinae* subfamily and members of this group have generally been shown to have a wide host range, relatively short replication cycle which leads to rapid cytopathic effect (CPE) and exhibit neurotropism that results in latent infection of sensory neurons (Rechenchoski et al., 2017).

Figure 1.1 Schematic diagram of the structure of an HSV-1 virion

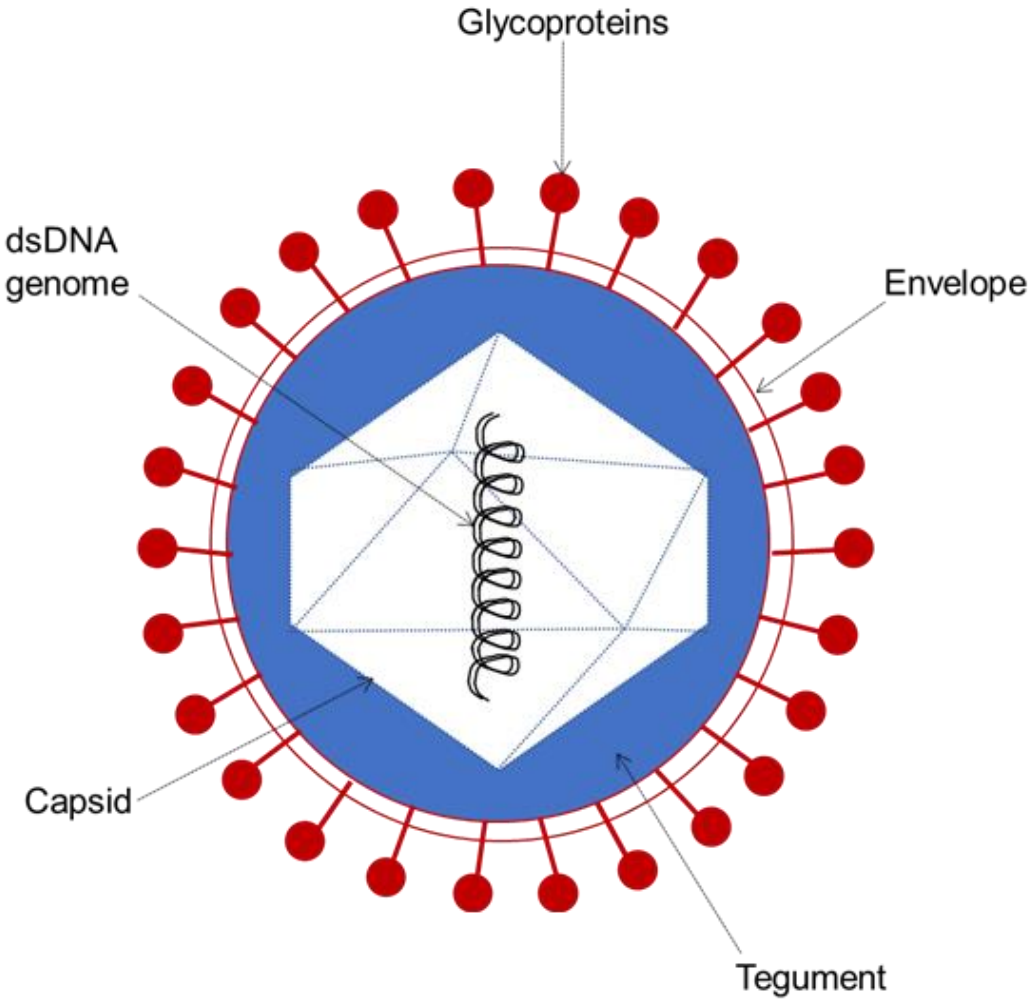


Table 1.1 showing the classification of the 9 human herpesviruses into their relevant subfamilies

Common name	Species name	Subfamily
Herpes simplex virus 1 (HSV-1)	*Human alphaherpesvirus 1* (HHV1)	*Alphaherpesvirinae*
Herpes simplex virus 2 (HSV-2)	*Human alphaherpesvirus 2* (HHV2)	*Alphaherpesvirinae*
Varicella zoster virus (VZV)	*Human alphaherpesvirus 3* (HHV3)	*Alphaherpesvirinae*
Human cytomegalovirus (HCMV)	*Human betaherpesvirus 5* (HHV5)	*Betaherpesvirinae*
Human herpesvirus 6A	*Human betaherpesvirus 6A* (HHV6A)	*Betaherpesvirinae*
Human herpesvirus 6B	*Human betaherpesvirus 6B* (HHV6B)	*Betaherpesvirinae*
Human herpesvirus 7	*Human betaherpesvirus 7* (HHV7)	*Betaherpesvirinae*
Epstein-Barr virus (EBV)	*Human gammaherpesvirus 4* (HHV4)	*Gammaherpesvirinae*
Kaposi's sarcoma-associated herpesvirus (KSHV)	*Human gammaherpesvirus 8* (HHV8)	*Gammaherpesvirinae*

1.2 Herpes simplex virus 1

HSV-1 is the prototypical alphaherpesvirus with a short replication cycle that rapidly causes CPE in tissue culture. Despite humans being the natural host for HSV-1, the virus has a wide host range. HSV-1 is also characteristic of an alphaherpesvirus in that it establishes latency in sensory neurons (McLennan & Darby, 1980; Roizmann et al., 1992; Stevens & Cook, 1971).

HSV-1 is also one of the most extensively studied alphaherpesviruses, this is largely owing to the ease with which the virus replicates in a number of *in vitro* and *in vivo* systems (Wagner & Bloom, 1997).

1.2.1 Genome Structure

HSV-1 has a dsDNA genome of 152kb with a G+C content of 68.3% (McGeoch et al., 1988). The HSV-1 genome comprises a unique long (U_L) and unique short (U_S) region, each flanked by inverted repeat sequences termed terminal and internal long or short repeats (TR_L, IR_L, IR_S TR_S), depending on their position along the genome (figure 1.2A) (Nicoll et al., 2012a)**.** The orientation of the U_L and U_S regions can be inverted relative to each other, therefore resulting in 4 equally likely HSV-1 isomers (Hayward et al., 1975).

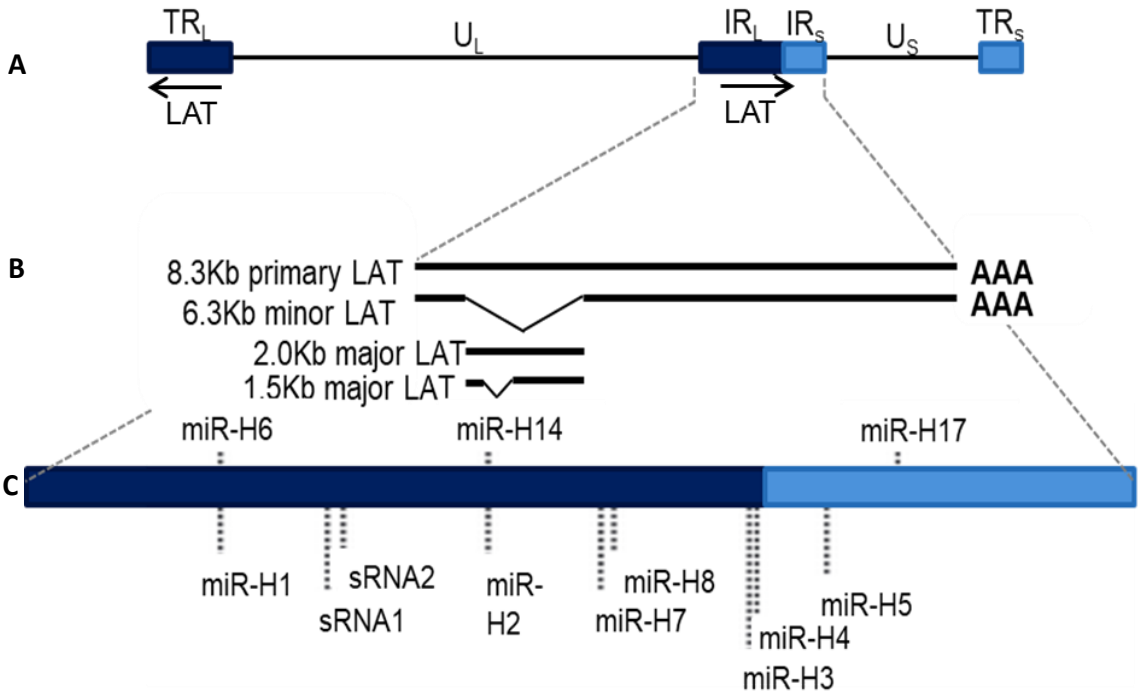
As already mentioned, a defining feature of the herpesviruses is their virion structure. As such, the HSV-1 dsDNA genome is contained within an icosahedral capsid surrounded by a tegument protein layer and finally the lipid envelope (figure 1.1).

The HSV-1 genome has at least 94 open reading frames (ORFs) and encodes approximately 84 proteins (Rajč & Uchova, 2004). A recent study by Whisnant *et al.* utilised a combination of functional genomics techniques to identify 284 ORFs present in the HSV-1 genome (Whisnant et al., 2020). However, the functional significance of these ORFs is yet to be elucidated.

1.2.2 Pathogenesis

HSV-1 is a highly prevalent neurotropic virus with more than 3.7 billion people infected worldwide (Looker et al., 2015). Primary lytic HSV-1 infection occurs following contact with the mucosal surfaces or abraded epithelial tissue. This may cause pathogenesis in the form of lesions. While HSV-2 is usually associated with genital lesions and spread following adolescence through sexual contact, HSV-1 is usually spread in early childhood, through non-sexual contact and is associated mainly with oral-facial lesions

Figure 1.2 Schematic diagram of the HSV-1 genome and layout of the latency-associated ncRNAs



Adapted from (Nicoll, Proença, & Efstathiou, 2012)

Schematic diagram of **(A)** the layout of the HSV-1 genome with particular focus on the region of the genome encoding **(B)** the Latency associated transcript (LAT) highlighting the genomic organisation of LAT and it's products, including **(C)** the microRNAs (miRNAs) that are processed from this region of the HSV-1 genome. Adapted from (Nicoll et al., 2012a).

such as cold-sores. However, HSV-1 can also be spread through sexual contact and is increasingly detected in genital lesions (Groves, 2016; Yousuf et al., 2020).

At the site of primary infection, the virus produces infectious progeny, which can then disseminate and enter the nerve termini of the sensory neurons innervating this site. These consist of the autonomic and sensory ganglia, most often the trigeminal ganglia following primary infection of oral mucosa. The virus travels via retrograde transport up the axon to the neuronal cell body where latency can be established (Efstathiou & Preston, 2005)(figure 1.3). Likely as a result of this initial viral input load, the viral-genome copy number within individual latently infected neurons is hugely variable. Accordingly, Sawtell demonstrated that HSV copy numbers differed over 3 orders of magnitude (from less than 10 to more than 1000) in latently-infected neurons of the murine trigeminal ganglia (Sawtell, 1997).

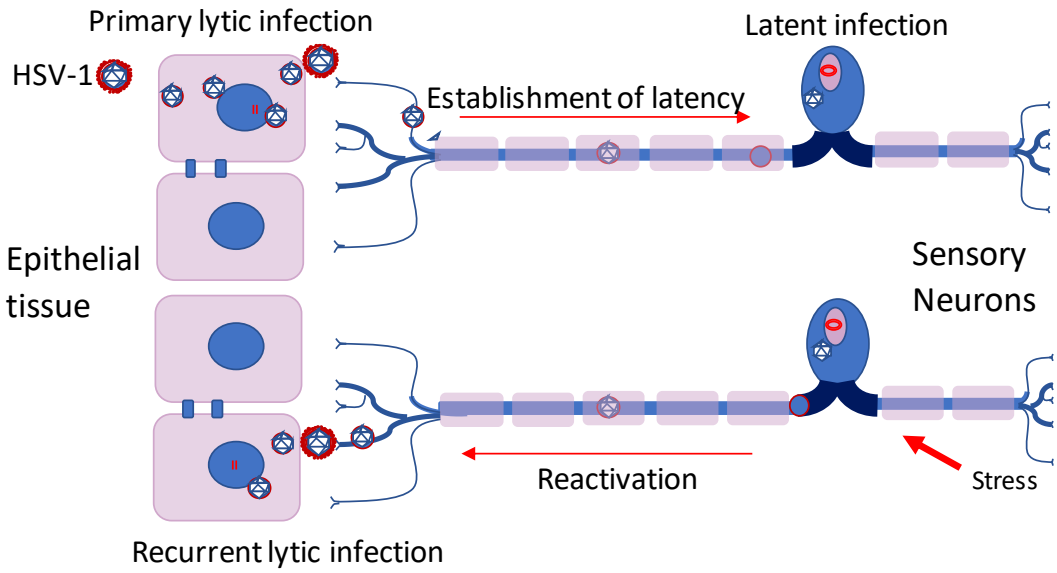
During latency, the virus genome circularises to form an episome. Additionally, although there is potential for transient lytic transcription during latency (Kim et al., 2012), which will be discussed later, acute stage lytic gene expression is generally repressed during latency so that the only abundantly transcribed gene is the latency associated transcript (LAT). To achieve this, HSV-1 lytic genes are largely associated with repressive heterochromatin (Knipe & Cliffe, 2008).

When transient lytic gene expression during latency goes further to initiate a fully lytic infection, capable of producing nascent virions, this is known as reactivation. There are various stress stimuli that have been shown to trigger reactivation from latency in neurons, such as, exposure to ultraviolet light, emotional stress, fever and immune suppression (Nicoll et al., 2012a). Stress can also be simulated *in vitro* to trigger reactivation, for example by use of heat-shock or histone-deacetylase (HDAC) inhibitors (Roizman & Whitley, 2013). Cliffe *et al.* were able to uncover some of the molecular mechanisms involved in neuronal stress leading to HSV-1 reactivation, which is described in more detail below (section 1.4.2) (Cliffe et al., 2015).

Following reactivation of the virus in neurons, the viral progeny travel by anterograde axonal transport back to peripheral tissue where the virus can cause pathogenesis and shed and transmit to a new host (figure 1.3).

In many cases HSV-1 infection is asymptomatic, due to latency or immune suppression of the virus. However, HSV-1 can cause mild to severe pathogenesis depending on the site of infection and the immunity of the host. For example, in rare cases the virus can enter the central nervous system (CNS) and cause encephalitis (Kennedy et al., 2015). Alternatively, if the host is immunocompromised or the immune system isn't fully

Figure 1.3 Schematic diagram of establishment of and reactivation from latency within sensory neurons



developed, as in neo-nates, severe disseminated infection can occur. HSV-1 can also cause blindness if primary infection occurs in the cornea or if reactivation occurs in the ophthalmic branch of the trigeminal ganglia (Webre et al., 2012).

1.3 Lytic replication cycle

Primary lytic infection occurs chiefly in the mucosal epithelia following contact of the mucosal surfaces with virus but can also occur at micro-lesions in skin epithelia.

For HSV-1 to undergo productive replication, the virus must first bind to the cell surface, through its glycoproteins. Next the capsid must enter the cytoplasm of the cell and migrate towards the nuclear envelope to allow for nuclear entry of the viral genome that follows. Once the viral genome is in the cell nucleus, viral gene expression can occur that in turn leads to genome replication. This provides the final components necessary for assembly of the virion particle, which enables viral egress. This entire process can take approximately 18-20 hours (h) and allows the virus to propagate and disseminate (Ibáñez et al., 2018). This is a brief overview of the lytic replication cycle, but each step requires complex molecular interactions between host and viral components as described below and represented in figure 1.4.

1.3.1 Entry

To initiate entry of the viral components into a host cell, the virus must first attach to the cell surface. This binding is mediated by viral glycoproteins binding to host-cell receptors. High-affinity binding of HSV is mediated by glycoprotein D (gD), which binds to one of 3 target receptors: nectin-1 (HveC), nectin-2 (HveB) or TNFRSF14 (HVEM – herpes virus entry mediator) (Eisenberg et al., 2012; Ibáñez et al., 2018). Alternatively, low-affinity binding can occur between glycoprotein C (gC) or glycoprotein B (gB) and cell-surface glycosaminoglycans, such as heparan-sulfate proteoglycans (HSPGs) (Kukhanova et al., 2014). Next, fusion must occur between the viral envelope and the cellular membrane, which takes place following a conformational change in gD that occurs after it has bound to one of the receptors mentioned above, this then activates the heterodimer glycoprotein H/glycoprotein L (gH/gL) which promotes the activation gB. Once activated, gB acts as the fusion protein, bringing together the cell membrane and viral envelope, allowing the nucleocapsid and associated tegument proteins to enter the cell (Eisenberg et al., 2012; Ibáñez et al., 2018). Alternatively, in certain cells the virus may enter through pH-dependent endocytosis and then fusion takes place with the intracellular vesicle membrane (Kukhanova et al., 2014; Milne et al., 2005).

For the viral genome to enter the nucleus, which is the site of HSV-1 replication, the capsid must migrate to the nucleus, which has been suggested to occur via active transport along the microtubule network, mediated by the cytoplasmic motor protein dynein (Döhner et al., 2002). Once the capsid reaches the nucleus, it docks to nuclear pore complexes, where the viral genome can then be injected into the nucleus.

1.3.2 Replication

Once the genome has entered the nucleus, it has been suggested that the genome circularises (Strang & Stow, 2005), and then the process of viral replication can begin. HSV-1 genes are expressed in a temporal cascade that is self-regulated by viral proteins. The immediate-early (IE) or α genes are the first to be transcribed, followed by the early (E/β) genes and then the Late (L/γ) genes, which can be broken up into the leaky-late (L_1/γ-1) genes first followed by the true late (L_2/γ-2) genes.

VP16 is an L_1 protein packaged in viral particle as a tegument protein. VP16 triggers the onward cascade of HSV-1 gene expression starting with the IE genes. VP16 (also known as Vmw65 and UL48) is a transcriptional activator that binds with host factors octamer binding protein 1 (Oct-1) and host cell factor (HCF) to initiate the transcription of IE genes (Ace et al., 1988; Coding & O'Hare, 1989; Wysocka & Herr, 2003). HCF is found in the nucleus in most cell types, where, as part of the complex with Oct-1 and VP16, it can interact with IE gene promotors. However, in sensory neurons HCF has been shown to accumulate in the cytoplasm. Nevertheless, neuronal stresses associated with reactivation can trigger relocalisation of HCF into the nucleus (Kim et al., 2012). Oct-1 – as a component of the VP16-Oct-1-HCF complex – binds to the TAATGARAT consensus sequences in the promotor region of the IE genes. This recruitment of VP16 from the viral tegument layer to activate transcription means that *de novo* protein synthesis is not required for expression of the IE genes.

Six IE genes have been identified to-date (infected cell protein 0 (ICP0), ICP4, ICP22, ICP27, ICP47 and US1.5). The proteins that the IE genes encode either subvert host gene expression or activate transcription of E genes (Kukhanova et al., 2014). ICP0 and ICP4 proteins are both major transactivators of gene expression. ICP0 is a really interesting new gene (RING)-finger-containing protein that possesses E3 ubiquitin ligase activity that enables it to transactivate all three kinetic classes of HSV genes (Smith et al., 2011). ICP4 regulates gene expression, reducing levels of IE genes, and increasing transcription of E and L genes by directly binding to DNA. ICP4 is essential for HSV-1 replication, whereby deletant mutants can only replicate on complementing ICP4 expressing cell lines (DeLuca & Schaffer, 1985). ICP27 is another essential IE gene.

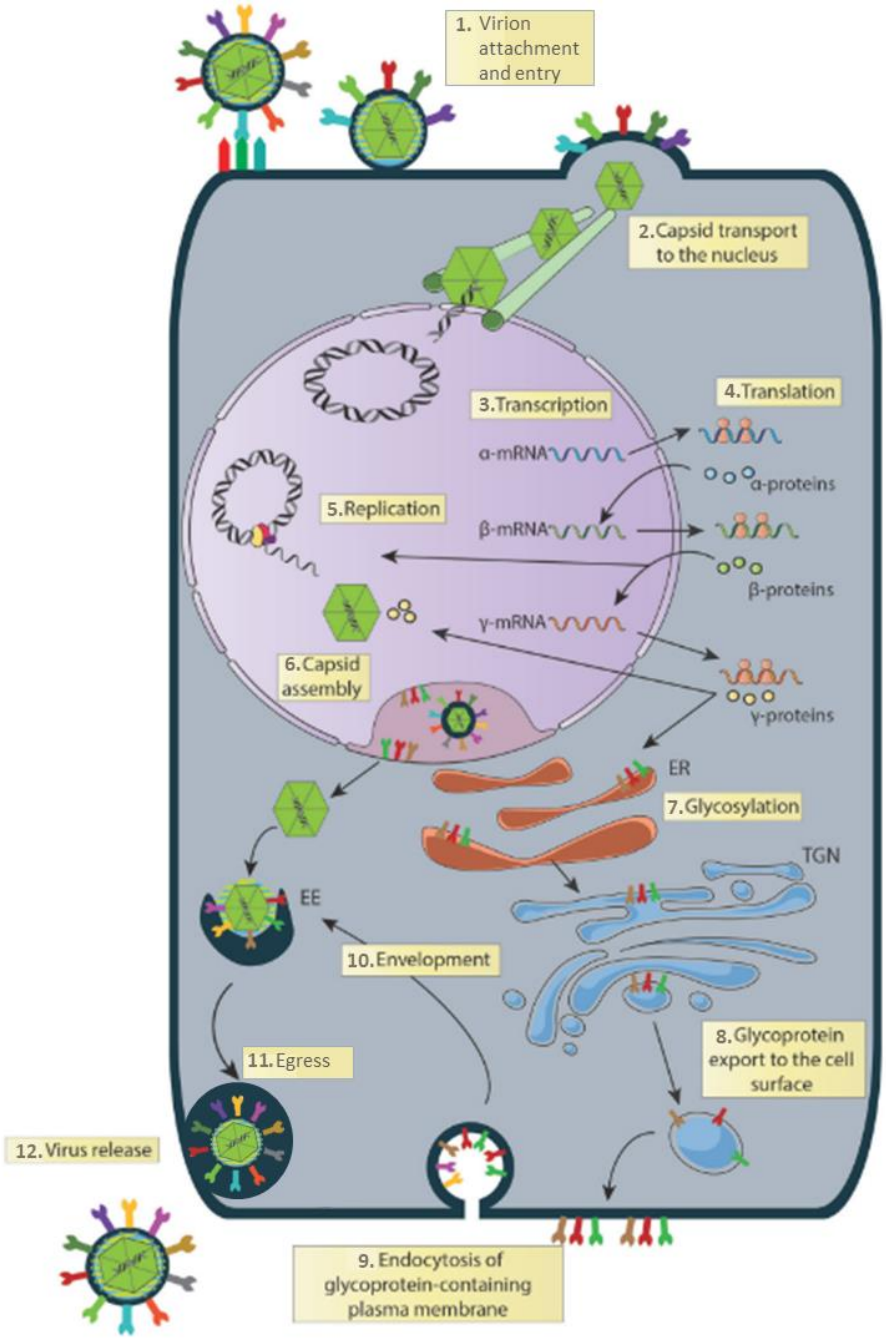
ICP27 disrupts cellular RNA splicing machinery and promotes the transport of viral intron-less mRNAs therefore promoting viral protein synthesis over cellular protein synthesis (Sandri-Goldin, 1998).

The proteins encoded by the E genes contribute to viral genome replication and activation of the final kinetic class of HSV-1 genes – L genes. An example of an E gene is UL30, which encodes the viral DNA dependent DNA polymerase, and is essential for the viral DNA replication (Nishiyama, 1996). The late genes mainly encode structural proteins that assemble into new virions, for example the major capsid protein VP5 (Wagner & Bloom, 1997). A summary of the HSV-1 genes and their known functions is shown in table 1.2.

1.3.3 Virion assembly and egress

Once there are enough L genes present, the newly replicated HSV-1 genomes can be packaged into capsids. These nucleocapsids exit the nucleus by budding through the nuclear envelope into the cytoplasm (Browne et al., 1996). During nuclear egress the virion acquires its tegument proteins and finally through budding in vesicles, the virus acquires a lipid envelope and glycoproteins. The vesicles fuse with the cell membrane and the mature virion is released (Ibáñez et al., 2018). This virion can go on to infect a new cell and repeat the lytic replication cycle described (figure 1.4).

Figure 1.4 Schematic representation of the steps of the HSV-1 lytic life cycle, adapted from (Ibáñez et al., 2018)



(1) Virion attachment and entry into the cell, mediated through binding of glycoproteins to their host-cell receptors **(2)** Capsid transport to the nucleus, aided by microtubules **(3)** Sequential viral gene transcription, starting with immediate early (α), then early (β), and finally late (γ) genes **(4)** Translation: viral mRNAs are translated sequentially into α, β and γ proteins **(5)** Viral DNA replication **(6)** Nucleocapsid Assembly **(7)** Glycosylation: glycoproteins are translated and glycosylated in the endoplasmic reticulum (ER). **(8)** Glycoprotein export to the cell surface: glycoproteins are processed in the trans-Golgi network (TGN) and multivesicular bodies. Then, they are exported to the plasma membrane **(9)** Glycoprotein-containing plasma membrane endocytosis **(10)** Envelopment: glycoproteins within early endosomes (EE) fuse with viral nucleocapsids in the cytoplasm and the virus acquires its lipid envelope and glycoproteins **(11)** Viral egress: the vesicles containing the complete virion fuse with the cell membrane **(12)** Virus is released out of the cell and can go on to infect new cells.

Table 1.2 Outlining HSV genes and their functions taken from (Nishiyama, 1996)

The table outlines lytic HSV-1 genes whereby the genes are divided based on whether they are essential (E) or dispensable (D) to virus replication in cell cultures. The genes are expressed at the immediate early (IE), early (E) and late (L) phases of infection. ICP = Infected cell protein/infected cell-specific polypeptides. MW = Predicted molecular weight. '?' Represents unknown functions at the time of publication.

Gene	Essential (E) or dispensable (D)	Period of expression	MW	Function
RL1	D	IE	78K	Transcriptional regulator, IE110, ICP0, α0
RL2	D	L		ICP34.5, neurovirulence
RL3	D			LAT transcription unit
UL1	E	L	25K	Glycoprotein L, fusion
UL2	D	E	36K	Uracil DNA glycosylase
UL3	D		26K	? Nuclear phosphoprotein
UL4	D	L	22K	? Virion protein
UL5	E	E	99K	Helicase/primase complex, DNA-dependent ATPase
UL6	E		74K	Capsid protein, VP11, DNA cleavage-packaging
UL7	D		33K	?
UL8	E	E	80K	Helicase/primase complex,
UL9	E	E	94K	Replication origin-binding protein
UL10	D	L	51K	Glycoprotein M
UL11	D	L	10K	Myristylated virion protein, nucleocapsid envelopment
UL12	D	E	67K	Alkaline exonuclease
UL13	D	L	57K	Protein kinase, virion protein
UL14	E		23K	?
UL15	E	L	81K	Capsid formation, DNA packaging
UL16	D		40K	?
UL17	E	L	75K	?
UL18	E	L	34K	Capsid protein, VP23, Capsid formation
UL19	E	L	150K	Major capsid protein, VP5
UL20	D	L	24K	Virion transport, Intrinsic membrane protein
UL21	D		58K	? Nucleotidylylated phosphoprotein
UL22	E	L	90K	Glycoprotein H
UL23	D	E	41K	Thymidine kinase
UL24	D	L	30K	Membrane protein
UL25	E	L	63K	Capsid formation, DNA cleavage-packaging
UL26	E	L	62K	Proteinase
UL26.5	E	L		Substrate of UL26 protease, DNA packaging
UL27	E	L	100K	Glycoprotein B
UL28	E	L	86K	Capsid formation, DNA cleavage-packaging
UL29	E	E	128K	Major DNA-binding protein
UL30	E	E	136K	DNA polymerase
UL31	E	L	34K	? Nucleotidylylated phosphoprotein, Nuclear matrix
UL32	E	L	64K	Capsid formation, DNA packaging
UL33	E	L	14K	Capsid formation, DNA packaging
UL34	E	L	30K	Nonglycosylated, membrane-associated protein
UL35	E	L	12K	Capsid protein
UL36	E	L	336K	Tegument protein, DNA release?
UL37	E	L	121K	? Cytoplasmic phosphoprotein
UL38	E	L	50K	DNA-binding protein, Anchoring DNA in the capsid
UL39	D	E	124K	Ribonucleotide reductase (large subunit)
UL40	D	E	38K	Ribonucleotide reductase (small subunit)
UL41	D	L	55K	Host shut-off protein, virion protein
UL42	E	E	51K	DNA polymerase accessory protein
UL43	D	L	45K	? Membrane protein
UL44	D	L	55K	Glycoprotein C, C3b-binding activity
UL45	D	L	18K	? Membrane protein
UL46	D	L	78K	Tegument protein, α-TIF modifier
UL47	D	L	74K	Tegument protein, α-TIF modifier
UL48	E	L	54K	Major tegument protein, VP-16, α-TIF
UL49	E	L	32K	Nucleotidylylated tegument protein
UL49.5	E	L		? Membrane protein
UL50	D	E	39K	Deoxyuridine triphosphatase
UL51	D	L	25K	?
UL52	E	E	114K	Helicase/primase complex
UL53	E	L	38K	Glycoprotein K, Fusion, Viral exocytosis
UL54	E	IE	55K	Posttranscriptional gene regulator, IE63, ICP27, α27
UL55	D		20K	?
UL56	D		21K	? Virion-associated protein
RS1	E	IE	133K	Major transcriptional regulator, IE175, ICP4, α4
US1	D	IE	46K	Regulatory protein, IE68, ICP22, α22
US2	D		32K	?
US3	D	E	53K	Protein kinase
US4	D	L	25K	Glycoprotein G
US5	D		10K	Glycoprotein J
US6	E	L	43K	Glycoprotein D
US7	D	L	41K	Glycoprotein I
US8	D	L	59K	Glycoprotein E, Fc-binding activity
US9	D	L	10K	Tegument phosphoprotein
US10	D		34K	Virion protein
US11	D	L	18K	RNA-binding protein, Tegument protein
US12	D	IE	10K	Down-regulation of MHC class I, ICP47

Table 1.2 Table outlining HSV genes and their functions taken from (Nishiyama, 1996)

1.4 Latency

Latency is a feature common to all the members of the *Herpesviridae* family and lasts the lifetime of the host, often with periodic reactivation events. Latency is a state whereby viral genomes are present, but viral replication does not occur and therefore no infectious virions are produced. For HSV-1, latency occurs in sensory neurons of the CNS and peripheral nervous system (PNS), primarily in the trigeminal or dorsal root ganglia (TG and DRG respectively) .

During HSV-1 latency, the genome is present as a circular episome (Efstathiou et al., 1986; Rock & Fraser, 1983) and gene expression is greatly restricted so that the only abundantly transcribed gene is the latency associated transcript (LAT) (Nicoll et al., 2012a; Wagner & Bloom, 1997).

1.4.1 Latency associated ncRNAs

The Latency associated Transcript (LAT)

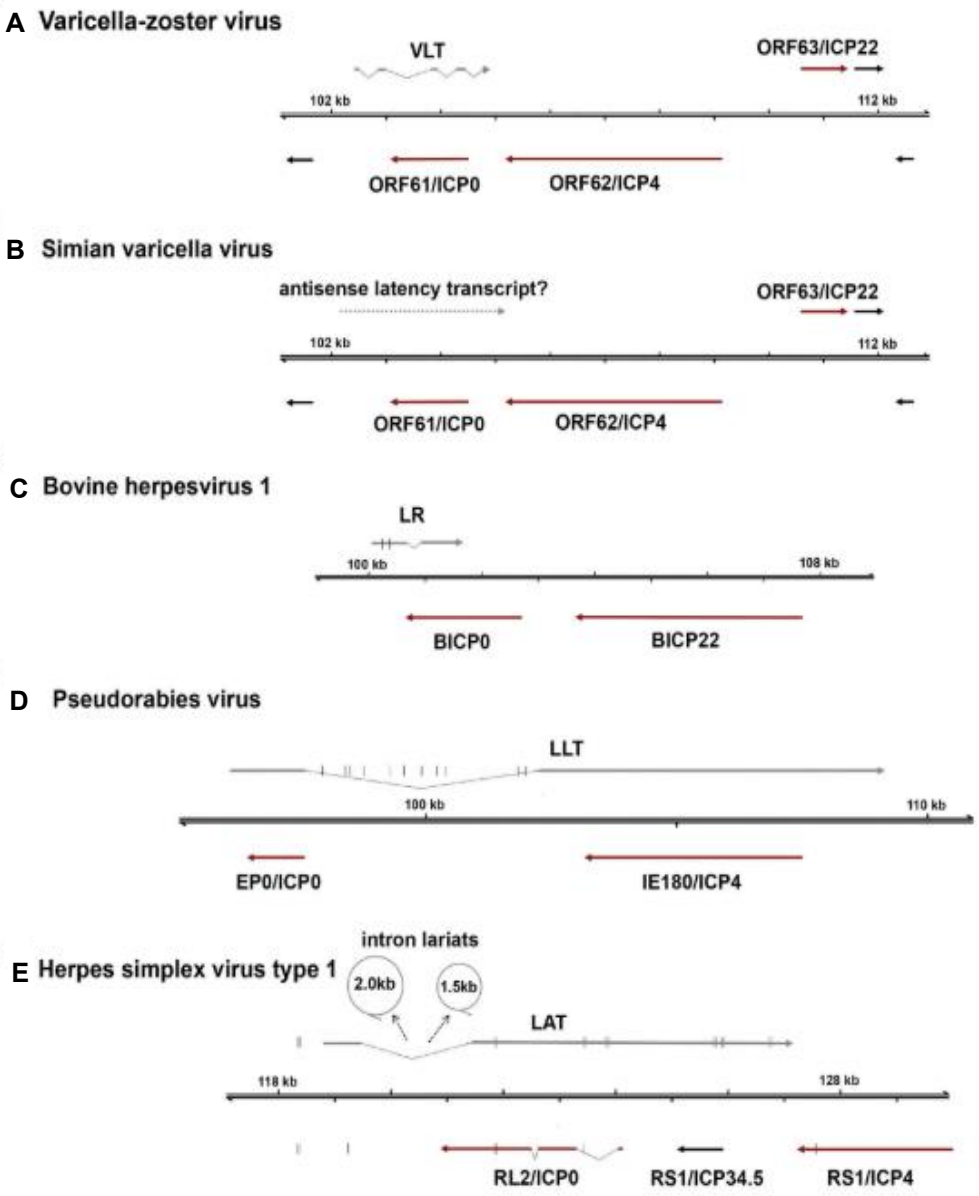
A feature that is conserved among the *alphaherpesviruses* is that they all produce latency associated transcripts (LATs). In addition to this, all of the *alphaherpesvirus* LATs are produced from the region of the genome encoding the ICP0 homologs (figure 1.5) (Depledge et al., 2018).

HSV-1 LAT is an 8.3kb polyadenylated primary transcript transcribed from the repeat regions of the genome (figure 1.2A and B). This primary transcript is spliced to produce the major LAT intron and the minor LAT exon. The major LAT intron is highly abundant and stable during HSV-1 latency and exists either as a 2.0kb or 1.5kb intron (figure 1.2B)(Brinkman et al., 2013; Rødahl & Haarr, 1997). Additionally, ~8 microRNAs (miRNAs) can be processed from the minor LAT exon and the LAT promotor region (figure 1.6) (Igor Jurak et al., 2010). Together these transcripts can be referred to as the latency-associated ncRNAs.

In addition to this, 2 small non-coding RNAs (sncRNAs) have also been detected to be expressed from the first 1.5kb of LAT. They are 62 and 36 nucleotides long and have the potential to form complex secondary structures (da Silva & Jones, 2013; Peng et al., 2008).

The major LAT intron is the most abundant and stable product during HSV-1 latency. This stability owes to the structure of major LAT as a lariat, due to a debranching defect during splicing (Block & Taylor, 1996; Farrell et al., 1991; Wu et al., 1998). In support of

Figure 1.5 Schematic diagram comparing the LATs across *alphaherpesviruses* taken from (Depledge et al., 2018)



Depledge, Sadaoka, & Ouwendijk, 2018

(A) The VZV latency-associated transcript (VLT) is a 496-nucleotides long, multi-exon mRNA found partially antisense to the ORF61 coding region **(B)** Transcripts expressed during simian varicella virus latency have been mapped antisense to ORF61 (dotted arrow) but their identity is currently undefined **(C)** Bovine herpesvirus 1 encodes a 2.2 kb latency-related (LR) RNA plus two miRNAs in exon 1 **(D)** The pseudorabies virus large latency transcript (LLT) encodes eleven miRNAs within the spliced intron **(E)** The 8.3 kb herpes simplex virus type 1 (HSV-1) LAT is spliced to yield two stable intron lariats (shown as circles) along with the minor LAT exon, which gets processed into ~6 miRNAs. The latency transcripts are shown as grey arrows, immediate early viral transactivators as dark red arrows, and encoded miRNAs as short vertical lines.

this, it has been shown that circular RNAs, are less likely to become degraded (Salzman, 2016).

The 2.0kb is the most abundant and well-studied form of major LAT, but during latency in neurons a 1.5kb intron has also been observed and been shown to be as stable as the 2.0kb major LAT intron (Brinkman et al., 2013). The nested intron structure of major LAT has been described as a twintron, where removal of a 0.5kb intron from the major LAT region of the primary LAT transcript results in splicing of a 1.5kb intron instead of a 2.0kb intron (Brinkman et al., 2013).

As the only abundantly transcribed gene during HSV-1 latency, there has been lots of interest as to what the functions of LAT may be. Initial studies, using LAT-negative HSV-1 mutants, showed that latent infection could still be established with these viruses, and therefore LAT is not essential for latency (Izumi et al., 1989; Sedarati et al., 1989; Steiner et al., 1989). Sawtell and colleagues' use of LAT-negative HSV-1 reporter viruses (that express Beta galactosidase during latency) revealed that there were approximately 80% fewer latently infected neurons in the trigeminal ganglia using LAT-negative HSV-1 mutants (KOS/29 and KOS/62) compared to the KOS/1 LAT-positive reporter virus (Sawtell & Thompson, 1992a). This suggested a role for LAT in the establishment of latency. In support of this, Nicoll *et al.* demonstrated that during high-dose infection of the mouse TG, latency is established in a higher number of neurons in the presence of LAT, compared to LAT-negative recombinant viruses. This supports previous data suggesting that LAT has a role in promoting the establishment of latency. However, the reverse observation was made at low-dose infection, where LAT-negative mutants established a higher latent load than the LAT-positive strains. This highlights the importance of experimental details such as inoculum dose and the difficulty of confirming a role for LAT not only across different latency models but also within the same model. To demonstrate their observations on the role of LAT on the establishment of latency, Nicoll *et al.* used a ROSA26R reporter mouse model that expresses a Cre recombinase activated *lacZ* gene, in conjunction with Cre recombinase-encoding HSV-1 recombinants (to enable marking of infected cells), with or without deletions in the LAT promotor (Nicoll et al., 2012b). This work also revealed that neurons latently infected with the LAT-positive virus remained infected over the study period of 140 days, whereas the LAT-negative mutants showed a decrease in the number of latently infected TG mouse neurons over time (Nicoll et al., 2012b). This indicated a role for LAT in the maintenance of latency.

Elucidating the role for LAT in reactivation has been more complex. Initial studies comparing LAT-positive HSV-1 to LAT-null mutants suggested a role for LAT in promoting reactivation (Block et al., 1993; Perng et al., 1994). For example, Perng *et al.* used a rabbit ocular model of HSV-1 infection to show that spontaneous reactivation was reduced by 33% when using LAT-null mutants compared to LAT-positive strains, suggesting that LAT is required for efficient reactivation *in vivo* (Perng et al., 1994). Perng *et al.* were also successful in mapping this spontaneous reactivation function to the first 1.5kb of LAT. HSV-1 mutants lacking the first 1.5kb of LAT were impaired for spontaneous reactivation, whereas the development of an HSV-1 mutant that can only transcribe the first 1.5kb of LAT rescued the spontaneous reactivation phenotype observed in wild-type HSV-1 strains (Perng et al., 1996). However, it was unclear from these studies whether LAT directly promoted reactivation or whether, as described above, LAT has a role in enhancing the establishment of latency, which in turn increased the number of neurons from which reactivation could occur. To address this question, single-cell approaches have been utilised; for example, the ROSA26R mouse model described above (Nicoll et al., 2012b) and contextual analysis of DNA (CXA-D), which is a method whereby single cell suspensions are isolated from the latently-infected TG of perfuse-fixed mice, for PCR amplification of DNA, to compare against HSV-1 DNA standards to mark the neuron as HSV-1 infected or not (Sawtell, 1997). This latter technique allows quantification of the number of infected neurons as well as the amount of viral DNA within each neuron. Thompson and Sawtell used CXA-D to demonstrate that latency was established in approximately 30% of neurons while using LAT-positive viruses but only in about 10% of neurons when using LAT-negative viruses, which supports previous data showing that LAT has an effect on the establishment of latency (Sawtell & Thompson, 1992a; Thompson & Sawtell, 1997). This reduced establishment of latency in LAT-null mutants correlated with a lower rate of reactivation following hyperthermic stress. However, when the latency establishment phenotype was rescued (by inducing hyperthermic stress during acute infection), the level of reactivation was rescued too. This suggests that the lower levels of reactivation observed in LAT-negative mutants were as a result of lower initial establishment of latency rather than a direct effect on reactivation. Moreover, it was later shown, again using single-cell analysis and therefore removing any residual effect from the establishment of latency, that LAT has a suppressive effect on reactivation. Nicoll *et al.* used a fluorescent mouse model of infection to isolate and culture single latently infected neurons. This enabled the observation that there was a higher frequency of reactivation from individual neurons infected with LAT-negative HSV-1, compared to those latently infected with LAT-positive HSV-1 (Nicoll et al., 2016). This suggests a role for LAT in suppressing reactivation. The

complex nature of the role of LAT in regulating reactivation is one of the reasons why further study is needed to fully understand the functions of this latency-associated ncRNA.

The above functions of LAT describe the direct impact that LAT has on latency, however, two further functions have been attributed to LAT. LAT has been suggested to block apoptosis and supress lytic gene expression.

Blocking apoptosis during latency has been observed in other *alphaherpesviruses,* for example, Bovine herpesvirus 1 (BHV-1) latency-related (LR) protein has demonstrated an anti-apoptosis function (You et al., 2017). HSV-1 might also have a method of blocking apoptosis during latency.

There is evidence that HSV-1 LAT promotes cell-survival, Perng *et al.* used terminal deoxynucleotidyl transferase-mediated nick-end labelling (TUNEL) to demonstrate that infection of rabbit TG using LAT-negative virus revealed extensive apoptosis, which was not observed when using a LAT-positive virus strain (Perng et al., 2000). In addition to this, this anti-apoptosis function was mapped to the first 1.5kb of LAT, using HSV-1 mutants (Peng et al., 2004). Carpenter *et al.* took this further by introducing 8 point-mutations across the first 1.5kb of LAT, and all 8 had some impact on the anti-apoptotic function of LAT (Carpenter et al., 2008). This confirms the idea that the first 1.5kb of LAT are important for its anti-apoptotic function and moreover it seems that even slight sequence changes alters this function.

Thompson and Sawtell used LAT mutants with a 1.9kb deletion, including the entire LAT promotor, to infect a mouse ocular model of latency. Their work revealed that LAT promotes neuronal survival, with significantly less destruction of neurons of the mouse TG in the presence of LAT in either wild-type or revertant HSV-1 strains compared to the LAT-negative mutant (Thompson & Sawtell, 2001). However, unlike the data by Perng and colleagues, this cell survival function did not seem to be mediated by a block in apoptosis, as tested by TUNEL (Thompson & Sawtell, 2001).

The evidence suggests that LAT promotes cell-survival, however, there are some contradictions about the mechanism involved. Da Silva and Jones suggested that this cell-survival function is at least partially mediated by the sncRNAs encoded in the first 1.5kb of LAT. Co-transfection of mouse Neuro-2A neuroblastoma cells with either LAT sncRNA1 or LAT sncRNA2 and retinoic acid inducible gene I (RIG-I) stimulated NF-κB transcription and reduced the levels of cold-shock induced apoptosis (da Silva & Jones, 2013). The uncertainty in the exact mechanisms involved in LAT promoting cell-survival further highlights the need to better understand the function of LAT.

Another role attributed to LAT is to suppress lytic gene expression. The first evidence of this was in a paper by Farrell *et al.* that used transient transfection assays to show that ICP0 trans-activation of the thymidine kinase (TK) promotor was reduced by 50-80% in the presence of the LAT intron (Farrell et al., 1991). This suppressive role of LAT was confirmed *in vivo* during acute HSV-1 infection by *in situ* hybridisation (ISH). A greater abundance of lytic transcripts was observed in mouse TG neurons following infection with LAT-negative mutant viruses compared to a wildtype HSV-1 strain KOS infection (Garber et al., 1997). Additionally, Mador *et al.* used northern blot analysis to reveal that the levels of ICP0, ICP4 and ICP27 mRNA were reduced following infection with a LAT-negative HSV-1 mutant in a mouse neuroblastoma cell-line stably expressing LAT, compared to mouse neuroblastoma cells in which LAT wasn't expressed (Mador et al., 1998). This suggests a role for LAT in the suppression of HSV-1 IE gene expression. In support of this, Nicoll *et al.* used firefly luciferase or beta-galactosidase expressing reporter viruses to show that the expression of these reporter genes, driven by the CMV promotor, were significantly upregulated in LAT-negative HSV-1 recombinants compared to those expressing LAT, during neuronal latency of the mouse TG *in vivo* (Nicoll et al., 2016). Both these data suggest a role for LAT in HSV-1 gene suppression. It has been suggested that LAT mediates this gene expression repression by promoting the formation of facultative heterochromatin, associated with gene silencing (Cliffe et al., 2009). Given LAT mediates silencing of viral genes, it would be interesting to examine the impact on host gene transcription.

Remarkably, despite these important roles during HSV-1 latency, data has revealed that only a proportion of latently infected neurons express LAT at any given time (Chen et al., 2002; Edwards & Bloom, 2019; Mehta et al., 1995). Mehta *et al.* used *in situ* PCR of HSV-1 infected murine TG neurons to first show that not all latently infected neurons express LAT, in fact, only approximately 30% of the number of latently-infected neurons express LAT at any given time (Mehta et al., 1995). Nonetheless, Proenca *et al.* used the ROSA26R reporter mouse with Cre HSV-1 recombinants, that as described previously enables historical marking of infected cells, to demonstrate that by 30dpi there has been LAT promoter activity in almost all latently-infected neurons despite individual timepoints showing a lower proportion of cells expressing LAT (Proença et al., 2008).

The HSV-1 miRNAs

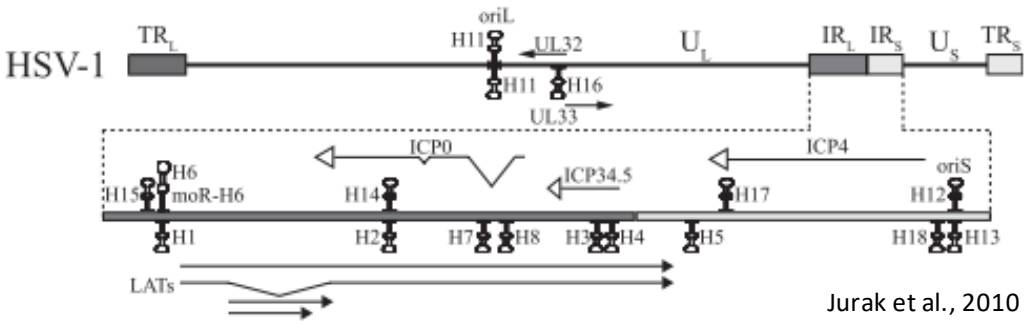
As already described, the primary LAT transcript is processed into miRNAs as well as the major LAT intron. The HSV-1 miRNAs were first discovered by Cui et al. who used computational methods to identify 11 predicted miRNA precursors (Cui et al., 2006).

Since then at least 16 miRNAs have been identified across the HSV-1 genome (Igor Jurak et al., 2010) (figure 1.6) with further suggestions that there are 20 HSV-1 miRNAs (Brdovčak et al., 2018). The 10 miRNAs that are encoded within or proximal to the LAT region are highlighted in figure 1.6 and 1.2. Five of these 10 miRNAs encoded in the LAT-encoding region of the HSV-1 genome and expressed during latency, were shown to be produced in post-mortem human latently infected trigeminal ganglia both by deep-sequencing and qRT-PCR. These were miR-H2-4 and miR-H6-7 (Umbach et al., 2009). The position and/or sequence of 9 of the HSV miRNAs are conserved between HSV-1 and HSV-2, which suggests that the these miRNAs are important to HSV function.

To understand the function of these HSV-1 latency associated miRNAs it is important to consider the known functions of miRNAs more generally. MiRNAs are approximately 22nt ncRNAs. Frequently miRNAs are transcribed as long primary miRNAs (pri-miRNAs), which are capped and have a poly-A tail, a double-stranded stem of approximately 30 base pairs, a terminal loop and two flanking unstructured single-stranded tails (Catalanotto et al., 2016; O'Brien et al., 2018). These pri-miRNAs can then processed into precursor miRNAs (pre-miRNAs) which are 70-nt long stem-loop structures. Alternatively, miRNAs are processed from within other gene introns/exons as with LAT, where LAT can be folded into the expected precursor stem-loop structure for each miRNA (Umbach et al., 2008). Regardless, the host proteins responsible for processing of miRNAs are Drosha, an RNAse III enzyme, and Di George syndrome critical region 8 gene (DGCR8), which is a double-stranded RNA-binding protein. Another RNAse III enzyme, Dicer, produces 22nt miRNA duplexes from pre-miRNAs, which contain a 5 prime (5') phosphate and a 3' hydroxyl group, with a 2-nt overhang. The miRNA duplex is loaded onto an Argonaute (AGO) protein by the RNA-induced silencing complex (RISC), which comprises Dicer, *trans*-activating response RNA-binding protein (TRBP) and an AGO protein. When loaded onto the AGO protein the 5p or 3p strand (arising from the 5' or 3' end of the pre-miRNA hairpin respectively) are each designated as the guide or passenger strand depending on the cellular environment and the thermodynamic stability of the 5' end of each strand (O'Brien et al., 2018). The AGO protein then unwinds the duplex, removing the passenger strand and leaving the single-stranded mature miRNA associated with the protein (Catalanotto et al., 2016).

As part of this miRNA-RISC (mRISC), the miRNA acts as a guide for the RISC to target the 3' UTR of mRNA to inhibit its translation and promote its degradation (Gosline et al., 2016). As such miRNAs (as part of the RISC complex) are known regulators of gene expression. There are multiple ways in which mRISCs have been shown to regulate

Figure 1.6 Schematic diagram of the position of the HSV-1 miRNAs along the genome, taken from (Igor Jurak et al., 2010)



mRNA targets. Most commonly, miRNAs are considered post-transcriptional regulators of gene expression by targeting mRNA antisense to their seed region, either for degradation or translational repression. However, recent data suggests that miRNAs can also control the transcription of their gene targets as well as promoting epigenetic modifications (Catalanotto et al., 2016; Gosline et al., 2016). Accordingly, Gosline et al. used RNA-Seq to show significant transcriptional changes in murine fibroblasts following global loss of miRNAs, due to knockout of DICER, which as discussed is key in the production of miRNAs (Gosline et al., 2016). This is evidence for miRNAs impacting the transcriptome of a cell.

The known functions of the latency-associated HSV-1 miRNAs are in keeping with these functions generally attributed to miRNAs, whereby their seed sequence is complementary to specific mRNAs, which leads to the downregulation of these mRNA targets. Viral mRNA have been found to be some of targets of the HSV-1 miRNAs, with certain HSV-1 miRNAs shown to be antisense to lytic genes and target them for degradation. For example, HSV-1 miR-H2-3p has been shown to be antisense to ICP0 and cause a reduction in ICP0 protein levels, although it does not significantly impact ICP0 mRNA levels (Umbach et al., 2008). This suggests that the miR-H2-3p targeting of ICP0 occurs at the translational level. Similarly, Umbach *et al.* found HSV-1 miR-H6 to be complementary to lytic HSV-1 gene ICP4 and was shown by again to impact protein but not mRNA levels (Umbach et al., 2008). These are examples of latency-associated HSV-1 miRNAs targeting viral mRNA at the translational level, however, as already discussed miRNAs can function at different stages from transcription to translation (Catalanotto et al., 2016; Gosline et al., 2016). There have also been examples found of HSV-1 miRNAs targeting host genes. For example, miR-H8, which although is processed from the LAT region (figures 1.2C and 1.6), was not present in latently infected post-mortem human trigeminal ganglia (Umbach et al., 2009), was shown to target host genes. Enk *et al.* demonstrated that miR-H8 targets phosphatidylinositol glycan anchor biosynthesis class T (PIGT), which is part of the glycosylphosphatidylinositol (GPI) anchoring pathway, crucial for the expression of several cell-surface proteins. This down-modulation of GPI-anchored proteins included ligands for natural killer activation, therefore suggesting that this microRNA may provide an immune evasion technique in targeting PIGT during lytic infection (Brdovčak et al., 2018; Enk et al., 2016). The role of this microRNA seems to fit with its expression during lytic infection. Accordingly, it will be interesting to elucidate what physiological implications the effects of expression of the latency-associated miRNAs has during latency. There is already some evidence of the effects of latency-associated miRNAs on

host cells with Zhao *et al.* demonstrating that miR-H4-5p targets cyclin-dependent kinase inhibitor 2A (CDKN2A) mRNA, that encodes the p16 protein, in undifferentiated SH-SY5Y cells (Zhao et al., 2015). This miR-H4-5p targeting of p16 leads to cell cycle progression. It will be interesting to see what the impact of this latency-associated ncRNA is in differentiated SH-SY5Y cells that no longer divide and are neuron-like, given that neurons are the site of HSV-1 latency. Clearly there is already information being discovered about the targets of the HSV-1 miRNAs. For a more extensive list of the targets already discovered for the HSV-1 miRNAs see table 1.3. Nevertheless, there are still unknown targets (table 1.3) and more to understand about the physiological implications of these targets, especially regarding the role of the miRNAs on the host cell during latency.

In addition to the functions described, recent work by Barrozo *et al.* revealed that a recombinant HSV-1 mutant lacking HSV-1 miR-H1-5p and miR-H6-3p was impaired for epinephrine-induced reactivation in the rabbit ocular model. Additionally, explant of latently infected DRG neurons using this HSV-1 mutant showed delayed and reduced reactivation when compared to wildtype HSV-1 infection (Barrozoa et al., 2020). This suggests that like LAT, the HSV-1 miRNAs have a role in regulating reactivation. Therefore, perhaps the roles of the latency-associated ncRNAs are conserved. The functions that are unique to or overlap between major LAT and the latency-associated miRNAs will be examined in this thesis.

The evidence that miRNAs can target at different stages of mRNA processing, including transcription (Catalanotto et al., 2016; Gosline et al., 2016), along with data showing HSV-1 latency-associated miRNAs targeting host mRNAs, not just viral mRNAs, is indicative that there may be scope for the HSV-1 miRNAs having an impact on the neuronal transcriptome. This hypothesis will be explored further in this PhD project.

1.4.2 Establishment, maintenance and reactivation

As already described (section 1.4.1), LAT has been suggested to have a role in every stage of latency, from establishment to reactivation. However, LAT is not essential for latency to occur nor is it expressed in every latently infected cell at any given time, therefore there must be other determinants to whether a virus establishes latency or not.

Establishment

The original view of latency was that it is a failure of the virus to undergo lytic infection. This idea came about from the observation that a lack of IE gene expression, for example with HSV-1 mutants, leads to the establishment of a quiescent infection, whereby viral

Table 1.3 Outlining HSV-1 miRNAs and their known targets

MiRNA	Proposed target	Reference
miR-H1	*ATRX *Ubr1 Antisense to miR-H6	(Brdovčak et al., 2018; Jurak et al. 2012)
miR-H2-3p	ICP0 *DDX41	(Duan et al. 2019; Umbach et al. 2008)
miR-H3-3p	ICP34.5	(Umbach et al. 2008)
miR-H4-3p	ICP34.5	(Umbach et al. 2008)
miR-H4-5p	ICP34.5 *P16	(Tang, Patel, and Krause 2009; Umbach et al. 2008; Zhao et al. 2015)
miR-H5-5p	?	-
miR-H6-3p	ICP4	(Umbach et al. 2008)
miR-H7-5p	Antisense to the first intron of ICP0	(Umbach et al. 2009)
miR-H8	*PIGT *CD48 Antisense to the first intron of ICP0	(Brdovčak et al. 2018; Enk et al. 2016; Umbach et al. 2009)
miR-H11-H13	?	-
miR-H14	Antisense to miR-H2	(Jurak et al. 2010)
miR-H15-H18	?	-
miR-H28-H29	?	-

The table lists HSV-1 miRNAs and their known targets. The latency-associated miRNAs found to be expressed in human trigeminal ganglia, both by deep sequencing and qRT-PCR by (Umbach et al. 2009) and are studied in this thesis are highlighted in bold.

ATRX - alpha- thalassemia/mental retardation syndrome X-linked (ND10

DDX41 - Asp-Glu-Ala-Asp (DEAD)-box helicase 41 (DDX41)

ICP – infected cell polypeptide

PIGT – phosphatidylinositol glycan anchor biosynthesis class T

*- Host target **?** – unknown **-** – Not applicable

lytic replication is repressed (Harris & Preston, 1991). Additionally, Margolis *et al.* observed that inhibition of viral replication by treatment with nucleotide analogues, had no bearing on DNA content in HSV-1 infected neurons, suggesting that there is no DNA replication during the establishment of latency (Margolis et al., 1992). However, more recent data has revealed that latency is a more dynamic process than once assumed, whereby subpopulations of latently infected neurons undergo lytic promotor activation prior to the establishment of latency and there can be low-level lytic gene expression even during latent infection (Arthur et al., 2001; Harkness et al., 2014; Nicoll et al., 2012b; Singh & Tscharke, 2019). Arthur *et al.* used β-galactosidase or GFP-expressing reporter viruses to infect neonatal rat DRG-derived neurons and demonstrated that IE promotor activity was transiently detected in the majority of neurons prior to the establishment of latency (Arthur et al., 2001). Harkness *et al.* used RNA-Seq to examine the viral gene expression profile during quiescent infection of adult murine TG neurons and found that in addition to LAT expression there was low-level lytic gene expression across all kinetic classes observed (Harkness et al., 2014). Therefore, although there is low-level lytic expression during HSV-1 latency, it differs from full lytic replication as it does not follow the established cascade paradigm for lytic gene expression (and does not produce virions). Nevertheless, work by Russel and Tscharke, using a Cre reporter mice system, has demonstrated that the lytic promoters found to be active during latency are capable of generating protein during latent infection (Russell & Tscharke, 2016). Our understanding of latency is evolving to incorporate these dynamic molecular interactions but to be able to appreciate the implications this may have, the relationship with the host cell needs to be considered.

Chromatin configuration surrounding the HSV-1 genome is a major factor in determining whether HSV-1 undergoes a latent or lytic infection. It has been established that post-translational modifications of the amino terminal tails of histones help regulate transcription (Grewal & Moazed, 2003; Strahl & Allis, 2000). Open chromatin configuration is associated with active transcription and is known as euchromatin, while condensed chromatin leads to gene silencing and is known as heterochromatin (Knipe & Cliffe, 2008). During lytic HSV-1 infection the genome is associated with euchromatin that encourages gene transcription as the nucleosome can be accessed by transcriptional activators and the RNA polymerase II complex to initiate transcription (Danaher et al., 2005). It is thought that VP16 plays a role in recruiting histone acetyltransferase to IE gene promoters along the viral genome to mediate this euchromatin formation (Knipe & Cliffe, 2008). During latency however, lytic gene promoters are associated with chromatin that contains repressive histone modifications

suggestive of heterochromatin, such as H3K9me2, H3K9me3 and H3K27me3 (Knipe & Cliffe, 2008). Conversely, the LAT region has been shown to be associated with comparatively high levels of acetylated H3 histone, indicative of active chromatin, which explains why this is the only abundantly transcribed gene during HSV-1 latency (Knipe & Cliffe, 2008).

Cell-type is an important determinant in whether HSV-1 establishes a lytic or latent infection, where neurons are the only natural site of latency. Therefore, it is logical that there may be components of neurons that aid in the establishment of latency. Accordingly, it has been suggested that the polarised state of neurons, with long cellular axonal projections, contributes to the establishment of latency (Koyuncu et al., 2018). Upon viral entry into the nerve termini, a subset of the tegument proteins separates from the viral nucleocapsid. This outer tegument layer contains the viral transcriptional activator VP16, which as previously described, binds with host factors Oct-1 and HCF to trigger the transcription of IE genes and the onward lytic gene expression cascade (section 1.3.2). It has been suggested that there is a delay in the viral tegument being transported to the neuronal nucleus, compared to the genome. This asynchronous delivery of VP16 to the cell nucleus (which is required for the lytic replication cycle), results in a failure to express IE genes, allows the virus genome to become silenced by heterochromatin, and results in the establishment of latency (Hafezi et al., 2012; Koyuncu et al., 2018; Nicoll et al., 2012a; Roizman & Sears, 1987). Another factor about neurons that may contribute to the establishment of HSV-1 latency over lytic infection is the host factors involved in the VP16-mediated transactivation of IE genes. It has been shown that HCF is detected in the cytoplasm of neurons, which would limit its ability to form a complex with VP16 and Oct-1 for the onwards transactivation of key viral lytic genes in the nucleus (reviewed in (Nicoll et al., 2012a)). Similarly, Lakin *et al.* demonstrated that Oct-1 is downregulated in non-dividing mature neurons (Lakin et al., 1995). Therefore, there is less Oct-1 available to form a complex with VP16 and HCF.

Evidently, the absence of certain host factors in neurons can mediate the establishment of latency. However, the presence of specific host factors can also contribute to the establishment and maintenance of latency. For example, host cell restriction factors suppress lytic genes in a manner that seems to encourage latency. Promyelocytic leukemia (PML) nuclear bodies (NBs) (or ND10) are such host cell restriction factors. The protein components of the PML NBs include PML, hDaxx and Sp100 and together these proteins function to repress lytic gene expression and tightly control the establishment of latency for multiple herpesviruses (Tavalai & Stamminger, 2009). PML NBs associate with the latent viral genome to form viral DNA-containing PML NBs (vDCP

NBs), resulting in the silencing of the HSV-1 genome (Cohen et al., 2018). Cohen *et al.* demonstrated that PML NBs mediate their repression of the HSV-1 lytic genes through chromatin regulation. Immune-FISH was used to show that PML NBs present in the trigeminal ganglia and associated with latent viral genomes are essential for the chromatinization of the latent genome with histone H3.3 trimethylation, a repressive heterochromatin marker, which as described above are important in the chromatin control of gene expression and for promoting a latent infection (Cohen et al., 2018).

The virus has developed countermeasures to this PML NBs-mediated silencing. ICP0 interacts with PML NBs to target PML for degradation through its E3 ubiquitin ligase domain (Alandijany et al., 2018; Gu & Zheng, 2016). However, such countermeasures may be prevented in neurons via the aforementioned failure of the VP16-Oct1-HCF complex to initiate expression of the ICP0 gene during the earliest stages of infection. This could explain why PML NBs block of viral gene expression and therefore the establishment of latency is favoured in neurons. Ultimately, it is a balance of host defences and viral countermeasures that help determine whether a latent or lytic infection will be established. Usually this equilibrium is considered in an *in vivo* context. Nonetheless, there are *in vitro* models that highlight the importance of the balance between host and viral factors, such as Pourchet *et al*.'s *in vitro* latency model using human embryonic stem-cell derived neurons infected with wildtype HSV-1 (in the presence of viral inhibitors). In this model, when infecting at a low multiplicity of infection (MOI) of 0.001 to 0.01, latency was maintained following removal of viral inhibitors. However, when a higher MOI of 0.1 was used for infection, meaning more virus was added, latency could no longer be maintained and reactivation occurred (Pourchet et al., 2017). This higher MOI seemingly tipped the balance in favour of the viral countermeasures that enable lytic replication over the establishment of latency.

Once latency is established using the various mechanisms described above to largely silence the genome, it is maintained in this state, sometimes for the lifetime of the host, with periodic reactivation events where the genome is derepressed.

Maintenance

As with establishment of latency, it is a balance of host-cell components that contributes to the maintenance of latency. One such viral factor is LAT, which as described above (section 1.4.1), has been shown to contribute to the maintenance of latency. (Nicoll et al., 2012b).

As with establishment of latency, cell-type is important for maintaining latency whereby neurons are senescent. Therefore, the HSV-1 genome does not have to contend with cellular division to maintain viral DNA.

A host-factor shown to be important for the maintenance of latency is the phosphatidylinositol 3-kinase (PI3-K) pathway, which is triggered by nerve growth factor (NGF)-binding to the TrkA receptor tyrosine kinase (RTK) (Camarena et al., 2010). Camarena *et al.* demonstrated by use of a PI3-K inhibitor, that PI3-K signalling is essential for the maintenance of HSV-1 latency, as robust reactivation was observed following use of the PI3-K inhibitor, as demonstrated by increased GFP-positive neurons following establishment of latency with a GFP-expressing HSV-1 reporter virus (Camarena et al., 2010).

Another host factor that is present in neurons that mediates the maintenance of latency is the mammalian target of rapamycin (MTOR) protein and the axonal signalling it mediates, again highlighting the importance of the polarised nature of neurons with respect to latency. The cellular MTOR-containing complex (MTORC1) kinase is a key target of the PI3-K pathway, which as described above is essential for the maintenance of latency (Kobayashi, Wilson, et al., 2012). Furthermore, MTOR signalling regulates cell metabolism, growth, proliferation and survival (Laplante & Sabatini, 2009). Kobayashi *et al*. used qRT-PCR and an EGFP reporter HSV-1 to demonstrate that inhibition of MTOR signalling, specifically targeting MTORC1, causes reactivation in latently infected cultured rat sympathetic neurons (Kobayashi, Wilson, et al., 2012). This suggests that the presence of functional MTORC1 is required for the maintenance of latency.

Given the contribution of these different factors, maintenance of latency continues until reactivation is triggered.

Reactivation

There are many triggers for reactivation both *in vivo* and *in vitro*, that each cause neuronal stress, for example fever, emotional stress, hyperthermic stress/heat shock, PI3-K inhibitors, histone deacetylase (HDAC) inhibitors and NGF withdrawal. As described above PI3-K signalling (which is triggered by NGF) is essential for maintenance of HSV-1 latency, therefore it follows that use of reagents that inhibit these pathways would lead to reactivation from latency (Camarena et al., 2010; Kobayashi, Wilson et al., 2012). Similarly, HDAC inhibitors contribute to a change in chromatin configuration, which as described above is also crucial in determining the outcome of HSV-1 infection (Neumann et al., 2007).

There are broader host-cell changes that ensue following neuronal stress that contribute to a reactivation phenotype, for example, HCF is relocalised from the cytoplasm of neurons to the nucleus, where it can form a complex with Oct-1 and VP16 to activate the IE genes (Kim et al., 2012). Thompson *et al.* were able to show that HSV-1 mutants lacking functional VP16 either completely fail to reactivate or are significantly impaired for reactivation *in vivo* in latently infected mice following hyperthermic stress. Therefore, this work demonstrated that VP16 is crucial for reactivation to occur (Thompson et al., 2009). This is despite data demonstrating that following explant of murine TG latently infected with a VP16 mutant (containing a 12 base-pair insertion, blocking the transactivating function of VP16), reactivation does occur (Steiner et al., 1990; Thompson et al., 2009). This highlights the importance of the model of latency used. Thompson and colleagues have demonstrated physiological differences, such as cell morphology and protein expression, between explant and *in vivo* reactivation that may explain this disparity in the need for VP16 (Sawtell & Thompson, 2004). Thompson *et al.* propose that the stress induced on the neuron during reactivation *in vivo* induces de novo VP16 production, which – when produced in high enough quantities – drives activation of IE genes transcription and in turn lytic replication (Sawtell & Thompson, 2020; Thompson et al., 2009).

Cliffe *et al.* were able to uncover the molecular mechanisms involved in neuronal stress stimuli causing reactivation. They identified a neuronal pathway, common to many stress responses, that involves activation of the c-Jun N-terminal kinase (JNK) induces HSV-1 reactivation. Cliffe *et al.* used both *in vitro* and *in vivo* mouse models of HSV-1 latency to demonstrate that the use of JNK inhibitors blocked reactivation in either dorsal root ganglia (DRG) or superior cervical ganglia (SCG) mouse neurons, as indicated by a significant decrease in the number of GFP-positive cells when using a GFP-expressing reporter virus or by a significant reduction in viral yield (Cliffe et al., 2015). It was shown that activating JNK triggers a histone phosphorylation on a neighbouring serine known as a histone methyl/phospho switch, on HSV-1 ICP27 and ICP8 lytic promoters (Cliffe et al., 2015). This methyl/phospho switch is a means by which gene expression can be initiated despite repressive lysine methylation, as is present during latency (as described above). This explains how JNK mediates the reactivation observed following use of any of the neuronal stress triggers mentioned above.

The initial burst of transcription that follows reactivation does not follow the usual lytic cascade of gene expression but instead there is low-level gene expression from all kinetic classes (Kim et al., 2012; Linderman et al., 2017). Kim *et al.* used a rat embryo SCG model of HSV-1 latency to demonstrate that reactivation is a 2-step process with

different gene expression profiles. Initially, within the first 20 hours post reactivation, in what is deemed “phase I” of reactivation, viral proteins are not required and transient, widespread de-repression of the genome is observed by qRT-PCR to a cross-section of genes from all 3 kinetic classes (Kim et al., 2012). This allows accumulation of VP16, which as described above is crucial for reactivation *in vivo*. Kim *et al.* compared infection with an HSV-1 mutant with an insertion in VP16 to a revertant strain to demonstrate that during phase II but not phase I there are differences in the lytic gene expression observed with these two different viruses following reactivation. Demonstrating that VP16 is required in phase II of reactivation but not phase I. This effect was confirmed by using short-hairpin RNA knockdown of VP16, which during phase I has no impact on the levels of lytic gene transcription observed by qRT-PCR, however, during phase II there is an approximately 5-fold decrease in lytic gene transcription (Kim et al., 2012). During phase II, VP16 is able to bind with its host factors Oct-1 and HCF as described previously, to promote transactivation of viral lytic genes and the onset of the established gene expression kinetics alongside genome replication (Kim et al., 2012)(section 1.3.2).

1.4.3 Models of latency

Our current understanding of HSV-1 latency and the roles of the latency-associated ncRNAs, as described above, have relied on the existence of effective models of HSV-1 latency. The type of model used can also have an impact on the results achieved as indicated with the differences observed between *ex vivo* and *in vivo* reactivation models, as described above (Sawtell & Thompson, 2004).

HSV-1 is an ancient virus that has co-evolved with its host over millions of years. Therefore, studying the intricacies of this relationship would be best done in the human host. However, it is difficult to study HSV-1 latency in the natural human host as this would be too invasive. The only direct study of human neurons is *ex vivo* work using latently infected post-mortem sensory ganglia.

Ex vivo work in post-mortem human neurons has been used to reveal information about the immune-response to latent HSV-1 infection of the human TG. Although the lack of lytic replication during HSV-1 latency is already a form of immune evasion, as there are no replicating DNA and RNA intermediates to be detected by immune sensors, there is still an immune response apparent during latency. CD4+ and CD8+ T-cells recovered from latently infected post-mortem human TG neurons, were found to recognise an array of HSV-1 viral proteins, across different kinetic classes (van Velzen et al., 2013). This provides useful information about the host response to infection during latency and supports the idea that there can be lytic gene expression during latent infection, as

described above, and moreover this gene expression may result in protein production. However, it is possible that the production of protein is from a subset of reactivating neurons or by abortive reactivation episodes that only transiently express lytic genes (Singh & Tscharke, 2019). Nonetheless, human cadaver neurons can provide useful information about the host-virus relationship. However, human cadaver neurons are a limited resource and therefore most of the work described has utilised animal and *in vitro* models of HSV-1 latency.

The success of a model depends on its ability to recapitulate the key characteristics of HSV-1 latency. These characteristics include neuronal cells as the site of latency, repression of lytic gene expression, expression of the latency associated ncRNAs and the potential to reactivate from latency.

1.4.3.1 *In vivo* animal models

Despite the co-evolution of HSV-1 with the human host, this virus has a broad experimental host-range and therefore there are animals available that are permissive to HSV-1 infection that can be used to model HSV-1 latency.

The predominant animal models for HSV-1 latency are murine and rabbit models. There are advantages and disadvantages to each, in terms of how well they recapitulate HSV-1 infection of humans.

Murine models

The mouse is the most widely used mammalian system for modelling human disease as their genetics are well characterised. Therefore, there are many useful mouse models that have been developed such as transgenic knockout mice that allow testing of specific phenotypes and virus response to, or interaction with, specific host gene homologues (Kollias et al., 2015; Webre et al., 2012). Additionally, mouse models are widely available and relatively low in cost, owing to them being small animals (Webre et al., 2012). As such, mouse models have been successfully used to study HSV-1 latency.

Mouse models have revealed some of the key aspects of latency known to-date. One such discovery was that of the structure of the latent genome. Rock and Fraser used southern blot hybridization of viral DNA from latently infected mouse ganglia and brains to demonstrate firstly that HSV-1 genome is present in the CNS tissue and secondly that the latent viral genome exists in a non-linear, “endless” form (Rock & Fraser, 1983). This was confirmed by Efstathiou *et al.* who also used southern blot hybridisation of infected mice neural tissue to demonstrate that during acute phase infection both junction and terminal HSV genomic fragments were detected, suggestive of a linear genome

structure. However, during latent infection only the viral junction fragment could be detected – indicative of an “endless” structure such as a circle (Efstathiou et al., 1986). Efstathiou *et al.* confirmed their findings from latently infected mice in post-mortem human trigeminal ganglia of seropositive individuals. This highlights the important finding gleaned from mice models of latency and demonstrate that those findings effectively recapitulate what is occurring during human latent infection.

Another useful mouse model, that has provided single-cell analysis of latently infected neurons and as previously mentioned provided information about the functions of LAT is the ROSA26R mouse model. This model was used to demonstrate that HSV-1 LAT has a role in the establishment and maintenance of latency (Nicoll et al., 2012b). This model, as briefly described above, expresses a lacZ gene that is activated by Cre recombinase. Therefore, recombinant HSV-1 virus strains that express Cre recombinase can be utilised to biochemically label each latently infected mouse neuron. The infected neurons express lacZ and therefore produce β-galactosidase, which can be visualised using β-galactosidase substrates (see methods 6.2.4.1 for more detail of β-galactosidase staining). This provides a useful single-cell readout for latent infection, a state of infection that is otherwise extremely difficult to quantify due to the lack of viral antigen production from silenced viral genomes.

There are various means of infecting mouse models, through the footpad, flank, ear, snout or corneal route (Wagner & Bloom, 1997). Regardless the route of infection, following primary infection at any of these sites, HSV-1 can travel via retrograde transport up the axons of the neurons of the innervating sensory ganglia to establish a latent infection.

However, mouse models do not fully recapitulate the pathophysiology of human HSV-1 infection. Although mice are permissive to HSV-1 infection, there have been instances where it is acutely fatal to the mice (Wang et al., 2013). However, this is also HSV-1 strain and animal dependent, with variation in mortality observed between different strains or even different genders of mice (Webre et al., 2012). In terms of the HSV-1 strain it has been demonstrated that the HSV-1 McKrae strains showing the most virulence, followed by 17syn+ and finally KOS displaying the least virulence (Thompson & Sawtell, 2001). It is important that mice survive beyond acute infection to be able to study latency. However, in most mouse models it seems that neuronal latency is more tightly restricted than a human infection whereby there is little or no spontaneous reactivation and no replication of virus in the peripheral tissue following reactivation (Dasgupta & BenMohamed, 2011; van Velzen et al., 2013). Clearly the mouse model,

although useful, does not fully recapitulate latency in the natural host. Therefore, to study reactivation in mouse models, external stimulators must be used to trigger reactivation, such as hyperthermic stress (HS). HS was established by Sawtell and Thompson and involves raising the core body temperature of the mice to 43°C. This HS technique was shown to reactivate HSV-1 so that infectious virus were detectable as early as 14 hours following HS with peak reactivation, including observation of HSV-1 antigens at 24 hours posttreatment (Sawtell & Thompson, 1992b). Another means of inducing reactivation from latently infected mice is by explant of mouse neurons whereby the latently infected ganglia are dissected and cultured at 37°C in tissue culture media and then onto a permissive cell-line to view any cytopathic effect from productive virus (Wagner & Bloom, 1997). This provides an inexpensive and amenable model to study reactivation from the mouse trigeminal ganglia despite a lack of spontaneous reactivation.

Rabbit models

Rabbit models more closely model latency in terms of being capable of spontaneous reactivation and accordingly are often utilised when reactivation phenotypes are being tested (Wagner & Bloom, 1997). The usual route of infection when using rabbits as a model for HSV-1 latency is via the cornea. One of the earliest models of reactivation, was a rabbit corneal model of HSV-1 latency utilised by Nesburn and colleagues (Nesburn et al., 1976). To induce reactivation in the Rabbit models, it was discovered by Hill and colleagues that iontophoresis of epinephrine can induce reactivation of HSV-1 in the TG of latently infected rabbits (Bloom et al., 1994; Wagner & Bloom, 1997). Unlike the mouse model, reactivation leads to shedding of infectious virus in the peripheral tissues (Dasgupta & BenMohamed, 2011). In this way rabbit models of latency closely recapitulate human HSV-1 infections, with lifelong latency and recurrent reactivation. The disadvantage of using rabbit models for HSV-1 latency is that they aren't as well characterised as mice, with fewer transgenic models available as a result. Additionally, rabbit models are more expensive than mice models.

These rabbit models have also provided useful information about HSV-1 latency and reactivation and the role of LAT. For example, Perng *et al.* used their rabbit model of HSV-1 corneal infection to show that LAT is protective of neuronal apoptosis in the infected rabbit (TG), as described above (Perng et al., 2000).

Clearly animal models have been key in building our understanding of HSV-1 latency and the role of the latency-associated ncRNAs. For example, the discovery of LAT was first made in mouse and rabbit models. Stevens *et al.* used a combination of *in situ* hybridisation and northern blotting of latently infected mice ganglia to identify the

presence of a small RNA antisense to ICP0, that localises in the nucleus during latency (Stevens et al., 1987). Rock *et al.* used a combination of *in situ* hybridisation and northern blot analysis, this time of latently infected rabbit TG, to confirm the presence of a nuclear-localising latency-related RNA and to show that during latency this is the only abundantly transcribed gene (Rock et al., 1987).

Evidently, small animal models have been useful in elucidating the molecular details of HSV-1 latency. Additionally, a benefit of *in vivo* models is that they can inform a more complete view of the host response to HSV-1, including immune responses (Phelan et al., 2017; Wagner & Bloom, 1997).

1.4.3.2 *In vitro* models

In vitro models are easier to control than *in vivo* models. Additionally, it can be difficult to generate enough latently infected neurons for large-scale molecular analyses from animal models. Accordingly, *in vitro* culture systems that utilise animal cells for models of latency have been developed to study HSV-1 latency.

Initially *in vitro* models for HSV-1 latency proved difficult to achieve as, in dissociated neurons, HSV-1 tends to undergo productive infection instead of establishing a latent infection. Wilcox and colleagues were among the first to develop an HSV-1 latency model, using either SCG or DRG neurons from embryonic rats. The rat neurons were cultured in nerve growth factor (NGF) and infected with HSV-1 in the presence of an HSV-1 inhibitor – acyclovir (ACV), which was maintained for 1 week following inoculation (Wilcox & Johnson, 1988; Wilcox et al., 1990). With this set up, latently infected neurons could be maintained for at least 5 weeks without production of infectious virus. However, following removal of NGF, HSV-1 antigen could be detected within 24 hours, indicating the potential for reactivation from latency (Wilcox et al., 1990).

This model continued to prove useful with Mohr and colleagues adapting it for use with GFP reporter viruses to visualise reactivating HSV-1. Mohr *et al.* were able to explore the mechanism by which NGF maintains latency in neurons. NGF binds to its receptor tyrosine kinase, TrkA, which activates PI3-K, which in turn contributes to the maintenance of HSV-1 latency (Camarena et al., 2010). This *in vitro* model successfully recapitulates all stages of HSV-1 latency and reactivation and provides a largely homogenous cell-type for studying neuron-host interactions.

Another primary neuron *in vitro* model for HSV-1 latency was developed using adult mouse trigeminal ganglia neurons, although in this instance there was no need for ACV. Bertke *et al.* demonstrated that HSV-1 could enter the majority of sensory neurons and

while productive infection occurs in a subset of these neurons, in the A5+ neuronal subtype, which are NGF-responsive, productive infection is blocked and a quiescent infection is established (Bertke et al., 2011).

The benefit of these systems over *in vivo* models is that they provide a pure neuronal population for analysis, with a higher abundance of latently infected cells, which can be more easily manipulated in culture for molecular analyses (Bloom, 2016).

Given the difficulty of studying HSV-1 latency in the natural host, *in vitro* methods have also provided a useful alternative to study the relationship between HSV-1 and human cells.

1.4.3.3 Human *in vitro* models

Although the animal models both *in vivo* and *in vitro* have proved useful, to fully understand the relationship between HSV-1 latency and the natural host, human *in vitro* models must be used. The importance of studying in the natural host is highlighted by instances where virus-host interactions that occur in humans are not fully comparable in rodents, for example, the HSV-1 VP16 protein associates with murine Oct-1 less efficiently than human Oct-1 protein, due to amino acid differences in the proteins (Cleary et al., 1993; Kim et al., 2012).

Preston *et al.* established some of the earliest human models for HSV-1 latency, using human foetal lung fibroblasts. To establish latency in these cells, Preston *et al.* generated replication defective HSV-1 mutant in1814 that contains deletion that abolishes the trans-activation of IE genes by the VP16 (also known as Vmw65). Therefore, the virus cannot progress through the lytic cycle of infection and so is stopped in a quiescent state (Harris & Preston, 1991). Reactivation could be triggered through superinfection with viruses expressing ICP0 such as wildtype HSV. This observation suggested that ICP0 (also known as Vmw110) is important for reactivation. Although this model is an effective means of establishing latency in human cells, the fibroblasts cannot fully model the host-virus relationship during HSV-1 latency as they are non-neuronal. Consequently, the genome of this mutant virus was so tightly repressed that LAT was not expressed in this model (Harris & Preston, 1991). Furthermore, Terry-Allison *et al.* demonstrated that transgene expression was more readily observed in mouse TG neurons than in non-neuronal cells following infection with an IE mutant virus, again suggesting that the genome of a replication-defective virus is more repressed in non-neuronal than neuronal cells (Terry-Allison et al., 2007). Clearly it is important to study latency in its natural site – neurons.

In recent years, alongside this thesis, there has been an increase in the number of human *in vitro* neuronal systems developed with the potential for modelling HSV-1 latency. These include Lund human mesencephalic (LUHMES) cells, immortalized HD10.6 cells and induced pluripotent stem cells (IPSCs) and SH-SY5Y cells (Edwards & Bloom, 2019; Shipley et al., 2017; Thellman et al., 2017). As with the animal models there are positives and negatives of each model.

HD10.6 cells are an immortalized cell line derived from human DRG. The cells proliferate due to a transduced tetracycline-regulated *v-myc* oncogene. Following addition of doxycycline, the HD10.6 cells mature to exhibit neuronal morphology and express neuronal markers and display a sensory neuron-associated phenotype (SNAP) (Thellman et al., 2017). Thellman *et al.* used these neuronally matured HD10.6 or "SNAP" cells to establish a model for HSV-1 latency, by infecting at a low MOI, with the addition of ACV 1hpi. After 3-4dpi the ACV is removed. With these conditions Thellman *et al.* were able to observe HSV-1 genomes in the SNAP cells, although only in a small proportion of cells. Lytic gene expression is observed but restricted, with less than 0.2 ICP0 or TK transcripts per viral genome as the HSV-1 latency model is maintained (following removal of ACV at 4dpi) up to 13dpi. This is in keeping with the expected phenotype during latency. However, LAT levels are similarly low following inhibitor removal, with less than 0.2 transcripts per viral genome when tested at 8 and 13dpi (Thellman et al., 2017). This is less indicative of latency. There was some spontaneous reactivation observed, as indicated by GFP expression when infecting with a GFP reporter virus, suggesting that the virus maintains its capacity to reactivate. However, many of the known inducers of reactivation that have been shown to be effective in other models, for example HS, axonal disruption, UV irradiation, removal of NGF or the use of PI3-kinase inhibitors, did not induce reactivation in this model above the levels of spontaneous reactivation. However, the combination of NGF removal and superinfection with UV irradiated virus did induce reactivation in this model. Therefore, this model can be used to study HSV-1 latency and reactivation. The benefit of this model is that it provides a homogenous population of human neuron-like cells that can be easily manipulated to study the effect of the quiescent HSV-1 infection established. The major disadvantage of this system is the low number of cells that were latently infected in this model, with low genome copy numbers as well as low LAT expression (Thellman et al., 2017).

LUHMES cells are human embryonic precursor cells that similarly to the HD10.6 cells are maintained as proliferating cells due to the expression of a tetracycline-regulatable (Tet-off) *v-myc* transgene. These cells can be differentiated into cells that seem to be

neuronal both in terms of morphology and positive staining for neuronal markers by the addition of tetracycline, glial cell-derived neurotrophic factor (GDNF) and dibutyryl cAMP (Edwards & Bloom, 2019). Edwards and Bloom have developed a model for HSV-1 latency using these neuronally differentiated LUHMES cells, by infecting with wildtype HSV-1 strain 17syn+ at an MOI of 3, in the presence of ACV for the first 48hpi. Under these conditions HSV-1 genome copy number is maintained over time and lytic gene expression decreases with time which is consistent with the establishment of latency. Also consistent with the establishment of latency is the expression of LAT. However, although LAT is initially expressed in the model, the abundance of LAT, as assessed by qRT-PCR, diminished with time. Reactivation seemed to be successfully induced by the addition of a PI3-kinase inhibitor, as RNA-scope demonstrated a large increase in the expression of lytic ICP4 transcripts, although the number of cells that harboured latent virus and then reactivated was not quantified. This model of HSV-1 latency could also be used to study reactivation. There are advantages of the LUHMES model such as it is scalable and produces a homogenous neuronal population that can be used to establish latency. Also, the LUHMES model of latency does not require prolonged chemical treatment to maintain latency although it does require inhibitors to establish latency. However, in this model LAT expression decreases following inhibitor removal to relatively low amounts with approximately 10 LAT transcripts (normalised to GAPDH) observed from 8dpi to 15dpi (Edwards & Bloom, 2019). This decrease in LAT expression is surprising during latency and means this model may not be ideal for testing the functions of the latency-associated ncRNAs – as is the objective of this thesis.

Both these models described above provide human neuronal systems that can be used to model HSV-1 latency and rely on transformed cell-lines differentiated into cells resembling post-mitotic neurons, which although useful does have the risk of the introduction of genotypic modifications occurring during immortalization.

An alternative approach, that most closely resembles natural HSV-1 latency is the use of human stem cell-derived neurons. Pourchet *et al.* have developed an HSV-1 latency model using embryonic stem cell-derived neurons. For this model, the embryonic stem cells are differentiated into neurons over a 12-week process. Once differentiated the stem cell-derived neurons are infected at a low MOI (0.001 – 0.01) in the presence of ACV and high-dose interferon α (IFNα) for the first 6dpi. At the higher doses of IFNα tested, after 6dpi the inhibitors could be removed, and lytic gene expression remained suppressed. In addition to this, LAT expression was shown to increase over time, as shown by qRT-PCR, indicating the establishment of latency. To induce reactivation in this system a histone deacetylase inhibitor was added to the cultures. Reactivation was

monitored by the expression of GFP when using GFP-Us11 reporter virus, whereby GFP expression indicates production of true late protein Us11 and therefore progression through the lytic cycle (Pourchet et al., 2017). The advantages of this system are that it successfully mimicked the hallmarks of latency with repression of lytic gene expression and the increase of LAT expression with establishment of latency, which seemed elusive in the previous two human neuronal models mentioned. They also managed to generate a relatively homogenous culture of neuron-like cells. The disadvantages of this system were that the cells were highly permissible to HSV-1 infection, meaning that high inhibitor doses were required to establish latency (100μm of ACV and up to 500IU/l of IFNα), above levels of inhibitor previously shown to be effective (eg in rat SCG models described above). Additionally, it is a very lengthy and complicated process to differentiate embryonic stem cells into neurons, requires ethical approval and the process is expensive. These could limit the data that can be collected using this model.

Induced-pluripotent stem-cells (IPSCs) have also started to be used to model HSV-1 latency. D'Aiuto *et al.* established HSV-1 latency in human IPSCs (hiPSCs) neurons by infecting at MOI 0.3 in the presence of antivirals IFNα and (E)-5-(2-bromovinyl)-2'-deoxyuridine for 7 days. These conditions successfully recapitulated HSV-1 latency whereby lytic gene expression was downregulated while LAT expression remained (D'Aiuto et al., 2019). Additionally, the virus remained capable of reactivation, triggered by the addition of histone deacetylase inhibitor, sodium butyrate. Reactivation was assessed by the production of infectious virus, as observed by plaque assay. Clearly the hiPSC-derived neurons are effective cells for modelling HSV-1 latency. Although D'Aiuto *et al.* have put work into making the process more scalable, there are disadvantages to this model, for example it takes 60 days to differentiate the hiPSCs into neurons and it is a delicate process (D'Aiuto et al., 2014). Additionally, this differentiation protocol exhibits some heterogeneity in the neuronal cell population (D'Aiuto et al., 2019, 2014). Although this heterogeneity reflects the presence of different sub-populations of neurons within the human host, for experimental purposes this can complicate the interpretation of results found when using this model with bulk cell analyses such as RNA-Seq.

Summary of *in vitro* models

All the *in vitro* models described above, including animal and human models rely on at least brief inhibition of viral replication to establish a quiescent infection. This is often a criticism of *in vitro* models of latency over *in vivo* models as this establishment of latency could be seen to be somewhat artificial and although it still can be used to study the various complexities of HSV-1 latency, the ideal model would eliminate the need for this

artificial block of replication. Accordingly, there has been the development of the use of chamber models that has the potential to allow each of these *in vitro* systems to be used without need for inhibitors or replication defective virus strains. Hafezi *et al.* used Campenot-like 2 compartment chambers to separate the cell bodies of chicken embryonic TG explanted neurons from their distal axon termini. Separation of the neuronal cell body and nerve termini allowed specific infection of the distal axons. Hafezi *et al.* were the first to show that infecting TG neurons at the distal axons with HSV-1 establishes a quiescent infection, without the need for inhibitors (Hafezi et al., 2012). This system has been adapted for use with Campenot chambers consisting of 3 compartments, known as “tri-chambers”, enabling the separation of neuronal cell bodies from axons and finally nerve termini. Koyuncu *et al.* made use of tri-chambers to infect superior cervical rat embryo neurons with pseudorabies virus (PRV) – a swine alphaherpesvirus – at the nerve termini of these neurons. Koyuncu *et al.* showed that the establishment of a quiescent infection with alphaherpesviruses when infecting at the nerve termini is dependent on the use of a low inoculum dose (Koyuncu et al., 2015). This chamber system could be tested with the human neuronal models described above that can establish HSV-1 latency by use of viral inhibitors and this might eliminate the need for this artificial block on replication.

The advantage of this chamber system is that it most closely resembles natural HSV-1 latency, whereby HSV-1 enters the nerve termini of sensory neurons to establish latency.

Nonetheless, there are disadvantages to this model, it is expensive, complicated and time consuming to grow neurons in chambers and there are limitations in scalability and the types of follow-on experiments that can be performed. The use of high throughput analyses for example would be difficult. Therefore, this system might be best used to follow up any phenotypes observed from high-throughput analysis performed in a more scalable model of latency.

Ultimately, it is difficult to fully recreate the complex relationship between HSV-1 latency and a neuronal cell, especially given the complex nature of neurons with various sub-types, delicate axonal processes, senescence and intricate differentiation protocols (D’Aiuto et al., 2014). *In vitro* models, especially when examining a homogenous culture of a particular neuronal sub-type are useful experimentally as you can attribute observations specifically to those cells and it provides easy manipulation that allows the examination of interactions at a molecular level. However, there are complexities that could be missed in using a homogenous *in vitro* model. For example, there may be different responses when using a different neuronal subtype. Cabrera and colleagues

demonstrated that neuronal sub-types differ in LAT promoter activity with neurofilament heavy-positive neurons having higher LAT promotor activity than neurofilament heavy-negative neurons as demonstrated by the use of a reporter virus expressing LAT promotor-driven β-galactosidase (Cabrera et al., 2018). Additionally, during natural HSV-1 latency in the human host there are non-neuronal cells surrounding neurons that may have also an impact on infection. For example, in the CNS there are astrocytes, oligodendrocytes and microglia (Yang & Zhou, 2019). These cells could be pertinent to host response to HSV, especially as it has been demonstrated that microglia and astrocytes are permissive to HSV-1 and mount an immune response to it (Bansode et al., 2019; Lokensgard et al., 2001). This also highlights the final caveat to using *in vitro* models to study HSV-1 latency – doing so limits the study of the immune response to infection. Nonetheless, scalable homogenous *in vitro* methods are useful tools to examine the effect of HSV-1 latency in human neurons, especially when performing high-throughput experiments such as RNA-Seq, where introducing variables such as multiple cell types would complicate the interpretation of any results found.

Evidently, there are advantages and disadvantages to the various types of *in vivo* and *in vitro* animal and human models of HSV-1 latency and information has been gleaned from each. Moving forward, with what is already known about HSV-1 latency, the most relevant route to study the host-virus molecular interactions that occur during HSV-1 latency is by using human neurons as they are the site of natural latency. As described above, there are several approaches developed in recent years to culture human cells to have neuronal properties and establish a latent HSV-1 infection, and benefits and caveats to each. This study makes use of a cell-type that is well-established in the neurology field – differentiated SH-SY5Y cells (Kovalevich & Langford, 2013).

1.4.3.4 SH-SY5Y cell system

The SH-SY5Y neuroblastoma cell line is a thrice cloned sub-line of SK-N-SH cells, originally derived from a bone marrow tumour biopsy (Kovalevich & Langford, 2013). Undifferentiated SH-SY5Y cells contain 2 morphological distinct cell phenotypes: neuroblast-like (N) cells and epithelial-like (E) cells. There have been several methods developed to differentiate SH-SY5Y cells into a more mature neuron-like phenotype, as indicated by morphological changes and the presence of neuronal markers. Encinas *et al.* developed a method to effectively differentiate SH-SY5Y cells into a population of homogenous catecholaminergic neurons by the sequential addition of retinoic acid (RA) and brain-derived neurotrophic factor (BDNF) and the gradual removal of serum. This encourages the differentiation of N cells and eliminates the E cells. Encinas *et al.* showed

that differentiated SH-SY5Y cells were dependent on BDNF for their cell survival and that following RA and BDNF treatment of SH-SY5Y cells, approximately 90% of the cells were arrested in the G1 phase of the cell cycle, which is reflective of senescent neurons (Encinas et al., 2000).

Shipley *et al.* adapted this protocol so that instead of adding BDNF and RA sequentially, the BDNF was added with RA. Accordingly, it has been demonstrated that BDNF enhances the differentiating effects observed with RA alone (Encinas et al., 2000; Shipley et al., 2016).

Using this protocol, Shipley *et al*. showed that neuronally differentiated SH-SY5Y cells are permissive for productive HSV-1 infection (Shipley et al., 2017). Shipley *et al.* performed single-step growth curves by infected the neuronally differentiated SH-SY5Y cells with wild-type HSV-1 McKrae strain and harvested cell-free and cell-associated virus at 6 time-points to demonstrate the viral kinetics of productive HSV-1 infection in these neuron-like cells and show that they are permissive to HSV-1 infection. Morphological changes that demonstrate CPE, such as neurite retraction, were observed over the course of the 24-hour infection, as early as 6hpi. Additionally, the levels of caspase-mediated apoptosis were tested, by measuring levels of inactive pro-caspase 3, which is cleaved when activated to mediate apoptosis. Levels of pro-caspase 3 were reduced from 6hpi during HSV-1 infection of differentiated SH-SY5Y cells, suggesting the onset of caspase-mediated apoptosis, that could be the cause of the CPE observed in this system (Shipley et al., 2017). Shipley *et al.* have demonstrated that this model can be a useful for assessing the neurotropism of lytic HSV-1 infection. However, this model has not been used in exploring HSV-1 latency, arguably the most relevant aspect to use these neuronally differentiated cells for, as human neurons are the site for HSV-1 latency *in vivo*. During this thesis, the neuronally differentiated SH-SY5Y cells will be used to establish a model for HSV-1 latency.

1.4.4 HSV-1 latency: impact on the neuron

Despite sensory neurons being the site of HSV-1 latency, it is not fully understood what the effect of this potentially life-long infection is on the neuronal host cell. As previously described, latency is a more dynamic process than first appreciated and it seems that the idea of a completely restricted virus genome is not ubiquitous across latently infected neurons, instead there is some low-level lytic transcription during HSV-1 latency of the human ganglia (Singh & Tscharke, 2019). The effects of this low-level transcription on the latently infected sensory neurons must be considered. As well as the impact of harbouring this latent HSV-1, reactivation also occurs in neurons, before the virus travels

via retrograde axonal transport to the peripheral tissue. Therefore, the effect of both latency and lytic replication that accompanies reactivation must be considered within human neurons.

There has been data accumulating that support the idea that long-term HSV-1 infection of neurons leads to an increased risk of developing certain neurodegenerative diseases. The most well documented link is with Alzheimer's disease (De Chiara et al., 2019; Mangold & Szpara, 2019). An association was first suggested by Ball, who showed that the same regions of the brain were impacted by Alzheimer's disease and HSV-1 encephalitis (Ball, 1982). Further epidemiological, immunological, genetic and molecular data have emerged since then, adding to the evidence that there is a link between HSV-1 infection or seropositivity and the development of Alzheimer's disease (Linard et al., 2020; Lövheim, Gilthorpe, Johansson et al., 2015; Lövheim, Gilthorpe, Adolfsson et al., 2015; Piacentini et al., 2014; Wozniak et al., 2009). For example, Lövheim *et al.* performed a nested case-control study and showed that in those where plasma was tested at least 6.6 years before the diagnosis of Alzheimer's disease, there was a significant association with the presence of IgG antibodies against HSV-1 and the development of Alzheimer's disease (Lövheim, Gilthorpe, Johansson, et al., 2015). While the exact detailed molecular mechanisms of how HSV-1 contributes to the development of Alzheimer's disease remains elusive, it has been the focus of several recent studies, and there is mounting evidence for why this link may exist. Accordingly, Wozniak *et al.* used a combination of *in situ* PCR and immunohistochemistry to demonstrate that HSV-1 DNA localises to amyloid plaques – a known marker of Alzheimer's disease – in post mortem human brain tissue (Wozniak et al., 2009). Shipley *et al.* infected SH-SY5Y cells to demonstrate some of the mechanisms involved in the interaction between HSV-1 and amyloid beta that may impact on the development of Alzheimer's disease. HSV-1 was shown to reduce full length amyloid precursor protein (APP) levels, as demonstrated by western blot, but increases the amount of novel C-terminal fragment of APP, suggesting that HSV-1 infection alters the processing of APP (Shipley et al., 2005). Furthermore, Chiara *et al.* recently demonstrated that repeat viral reactivation of HSV-1 in mice led to the accumulation of molecular biomarkers of Alzheimer's disease such as amyloid-β protein, tau hyperphosphorylation and neuroinflammation in the brain (De Chiara et al., 2019). This work confirms that recurrent HSV-1 infection is a risk factor for developing Alzheimer's disease.

Another neurodegenerative disease that has been linked to HSV-1 is Multiple Sclerosis (MS)(Duarte et al., 2019). Ferrante *et al.* suggested a role for HSV-1 in triggering MS relapses. Accordingly, Ferrante *et al.* demonstrated that there was HSV-1 DNA found in

the peripheral blood mononuclear cells (PBMCs) of significantly more acute MS patients than stable MS patients or healthy controls (Ferrante et al., 2000). A later case control study confirmed this data, where again the prevalence of HSV-1 DNA was significantly higher in PBMCs of relapsing MS patients compared to PBMCs from healthy control subjects (Najafi et al., 2016). However, the impact, if any, on stable MS patients and within neurons is less clear. Accordingly, Koros *et al.* demonstrated that HSV-1 DNA was only present in the cerebral spinal fluid (CSF) of 4.7% of MS patients tested (Koros et al., 2014).

Although the direct molecular interactions remain elusive, the contribution of HSV-1 to neurodegenerative disorders suggests that HSV-1 has a detrimental effect on neurons over time. In keeping with the disruption of HSV-1 to neuron biology, when HSV-1 infects neurons of the CNS, as mentioned briefly, it can cause herpes simplex encephalitis (HSE). This occurs when there is active replication of HSV-1 in neurons of the brain. This leads to neuronal cell death by apoptosis or necrosis (Duarte et al., 2019).

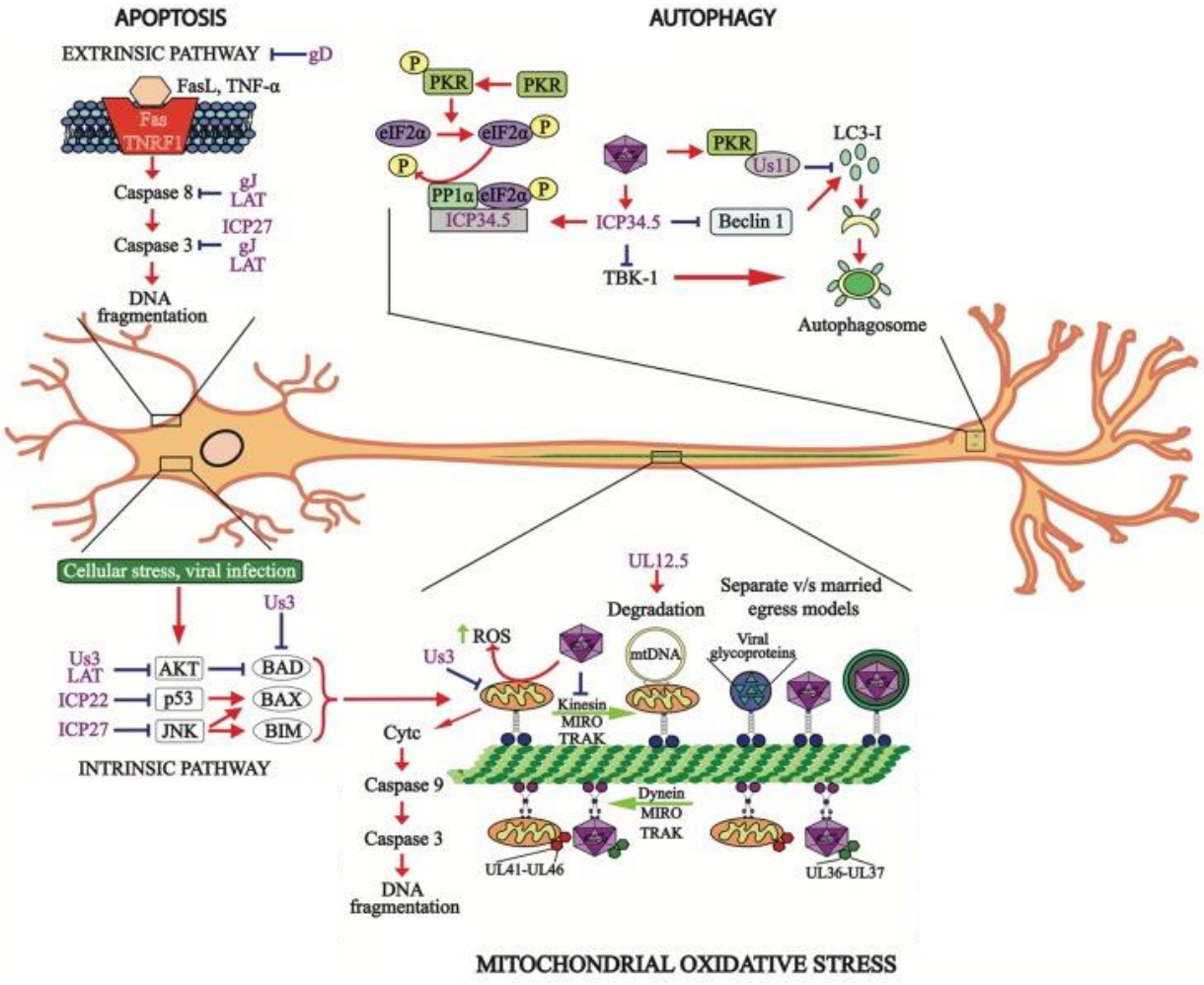
Nonetheless, HSV-1 has been shown not only to induce neuronal cell death but also inhibit it. As described above, the first 1.5kb of LAT has been shown to promote cell survival, through an anti-apoptotic effect (Carpenter et al., 2008; Peng et al., 2004; Perng et al., 2000) (section 1.4.1). Therefore, prolonged neuronal survival could be another effect of long-term latent HSV-1 infection due to the expression of LAT. Lytic HSV-1 proteins are also responsible for regulating apoptosis. Both the extrinsic and intrinsic apoptosis pathways are modulated by an array of HSV-1 proteins, such as IE proteins ICP22, ICP27 and US3, L_1 viral proteins gD and gJ, at different stages of signalling cascades during apoptosis and at different time-points throughout infection (Duarte et al., 2019) (figure 1.7).

Apoptosis is not the only cellular function that HSV-1 has been shown to be affected by HSV-1 in neurons. HSV-1 also modulates autophagy. Viral protein US11 interacts with RNA-activated protein kinase (PKR), causing the inhibition of the conversion of protein LC3-I into LC3-II, which is necessary for proper autophagosome function (Duarte et al., 2019; Runwal et al., 2019). HSV-1 protein ICP34.5 also similarly inhibits LC3-I conversion into LC3-II by blocking beclin-1 and therefore preventing proper autophagosome formation (Duarte et al., 2019).

Mitochondria oxidative stress is also affected by HSV-1 infection in neurons. Both, mitochondria and HSV-1 transport are mediated by retrograde and anterograde processes involving microtubules. Additionally, HSV-1 UL12.5 produces mitochondria DNA degradation and the viral protein US3 protein blocks the electron transport chain

within mitochondria (Duarte et al., 2019). A detailed diagram outlining these interactions between HSV-1 proteins and cellular components during these three cellular processes within neurons is shown in figure 1.7.

Figure 1.7 Schematic diagram showing how HSV-1 modulates cellular processes in neurons, adapted from (Duarte et al., 2019)



The schematic shows how HSV-1 modulates neuronal cellular processes whereby HSV-1 proteins are indicated in purple and the stage in which they inhibit apoptosis, autophagy or mitochondrial oxidative stress is indicated by perpendicular blue lines. The key cellular components and stages for each process are also shown above. Activation is indicated by red arrows. The green arrows indicate direction of travel.

Evidently, there are many host-virus interactions that occur within human neurons. The above focuses on HSV-1 regulation of cellular processes but there are also host cell responses to infection to consider within the neuron.

Originally the immune response was thought to be relatively muted in neurons as a means of avoiding any immune-activated cell death, given that neurons are senescent and therefore largely irreplaceable. However, neurons do still mount an immune response to viruses, and there is now a greater understanding of what this entails.

Toll-like receptors (TLRs) are part of the first line of defence against viruses. TLRs are pattern recognition receptors that initiate the innate immune response by recognising pathogen-associated molecular patterns (PAMPs) on the invading pathogen, in this case HSV-1 (Rosato & Leib, 2016). For example, TLR2, which is expressed on cell surfaces, detects viral glycoproteins. Alternatively, TLRs 3, 7 and 9, which are found in endosomal compartments, detect viral nucleic acids. TLR3 is a dsRNA sensor and has been shown to be important in the neuronal immune response to HSV-1 (Sato et al., 2018). Zhang *et al.* demonstrated that children with a dominant-negative TL3 allele, making them TLR3 deficient, developed HSE (Zhang et al., 2007). This data suggests that TLR3 helps control HSV-1 infection in neurons of the CNS to prevent the development of HSE. Additionally, Zimmer *et al.* demonstrated that trigeminal neurons derived from IPSCs obtained from patients with the TLR3 genetic deficiency are more permissive to HSV-1 (Zimmer et al., 2018). TLR3 activation triggers the production of IFNα and IFNβ, which in turn stimulate signalling cascades that lead to the production of interferon-stimulated genes that ultimately inhibit HSV-1 replication (De Regge et al., 2010; Samuel, 2001). Activation of TLR3 also causes the production of other inflammatory mediators such as nitric oxide, tumour necrosis factor α (TNFα), interleukin 6 (IL-6), IL-12 and chemokines CXCL10 and CCL5, that encourage the clearance of virus-infected cells through recruitment of immune cells such as macrophages and neutrophils (Duarte et al., 2019; Lee et al., 2007). Interestingly, the neuroinflammation caused by HSV-1 occurs during both productive and quiescent infection (Duarte et al., 2019; Lokensgard et al., 2001).

Evidently, the neuronal host cell mounts an antiviral response to HSV-1. However, HSV-1, having co-evolved with its host for millions of years, has also developed immune evasion techniques. For example Xing and colleagues used co-transfection with an HSV-1 VP16 expressing plasmid and an IFNβ promoter reporter plasmid to demonstrate that VP16 inhibits IFNβ promoter activity and does so in a dose dependent manner (Xing et al., 2013).

Together these data highlight that HSV-1 infection has an impact on human neurons and there are a variety of dynamic interactions between the neuronal host cell and HSV-1. However, more work is needed to fully understand this relationship and any further implications of long term latent HSV-1 infection of human neuronal cells.

1.5 Aim of the thesis

In this thesis, the effect of HSV-1 latency on human neurons will be examined further. The aim of this project is to investigate the functions of the HSV-1 latency-associated ncRNAs in the context of HSV-1 latency in human neurons. The focus of this work will primarily be on the role of the major LAT intron as this is the most abundant and stable product during HSV-1 latency, as well as the function of 5 of the most abundantly produced miRNAs during latency. Accordingly, this work aims to distinguish the effects of the major LAT intron or the latency associated miRNAs from the effect of the entire LAT transcript. To achieve these aims, an *in vitro* model was established either by utilising a replication-defective viral mutant or by engineering lentiviral vectors to express the HSV-1 latency-associated ncRNAs in human neurons. The human neurons used in this model were differentiated from SH-SY5Y cells. During this thesis, this model was used to examine the impact of the latency associated ncRNAs on the neuronal transcriptome by using RNA-Seq.

The hypothesis and aims of this thesis are summarised below:

Hypothesis:

- The latency associated ncRNAs will alter the neuronal transcriptome

Aims:

- To establish an *in vitro* model to express the latency-associated ncRNAs in human neurons
- To be able to distinguish the effects of the major LAT intron or the latency associated miRNAs from the effect of the entire LAT transcript
- To use this model to test the effect of the latency associated ncRNAs on the neuronal transcriptome

Chapter 2: Characterising an SH-SY5Y neuronal model system of HSV-1 latency

2.1: Introduction: Modelling HSV-1 latency

Herpes simplex virus 1 is a ubiquitous human virus that can establish lifelong latency in the sensory neurons of its host. Currently, an understanding of the impact of prolonged latency on human neurons harbouring latent HSV-1 genomes, is limited. As discussed in chapter 1.4.4, there has been a correlation suggested between HSV-1 infection and the development of neurodegenerative Alzheimer's disease from epidemiological studies, immunological data and work in explanted human trigeminal ganglia (Ball, 1982; Linard et al., 2020; Lövheim, Gilthorpe, Johansson et al., 2015; Piacentini et al., 2014) but more work is needed to fully understand this relationship and any other implications of HSV-1 long term infection of human neuronal cells.

It is therefore important to develop models of HSV-1 latency that enable us to study the interactions of latent HSV-1 and neuronal cells. As already described in chapter 1, a significant amount of research has been dedicated to developing HSV-1 models of latency. For these systems to successfully reflect the key physiological features of latency, as seen *in vivo* in humans, they must express the latency-associated ncRNAs – as indicated by the presence of LAT – and lytic gene expression must be transcriptionally repressed (Wagner and Bloom, 1997; Margolis et al. 1992). Much of the published literature on HSV latency has utilised various *in vivo* animal models. These have facilitated examination of several aspects of the host response to HSV infection, including the immune response and elucidating several functions of LAT, including a role in the establishment and maintenance of latency and an anti-apoptotic effect, as discussed in chapter 1.4.1 (Nicoll et al., 2012b; Sawtell & Thompson, 1992a; Thompson & Sawtell, 1997; Branco & Fraser, 2005; Perng et al., 2000). Furthermore, a combination of *in vitro* LAT transfected 293T cells and *in vivo* latently infected mouse trigeminal ganglia were used to show that several HSV-1 microRNAs (miRNAs) are processed from the LAT region and that these miRNAs function repress HSV-1 lytic gene expression and protein production (Umbach et al., 2008). Although these examples demonstrate that animal models are indeed useful in elucidating aspects of HSV-1 latency and the potential roles for latency-associated ncRNAs, as discussed, many caveats of the use of animal models remain. This is because these models cannot fully recreate the relationship of HSV infection in human neurons. For example, the mouse model is limited in its capacity for spontaneous reactivation of HSV. Understandably, the best model of any viral-host interaction is one that most closely resembles the natural host cell, in this case, human neurons.

As such, it is useful to consider human neuronal *in vitro* models of HSV-1 latency that better recapitulate the natural site of latency. This enables scrutiny of the host-virus relationship in a way that might not be possible in *in vivo* animal models, for example because the animal model may not express all the same genes that may react to the virus in the natural human host. This is especially likely with HSV-1, as it has co-evolved with its host for millions of years (Davison, 2002; McGeoch & Cook, 1994; McGeoch et al., 1995). There has also been a general shift in research in recent years, alongside this project, to establish tractable *in vitro* models of HSV-1 latency; the most relevant kind being human and neuronal to mimic the site for HSV-1 latency *in vivo*. As described in the introduction (chapter 1.4.3.3) of this thesis, recent human *in vitro* neuronal cells that have been developed to model HSV-1 latency include LUHMES, immortalized HD10.6 cells and stem cell-derived neurons (Edwards & Bloom, 2019; Pourchet et al., 2017; Thellman et al., 2017). As with the animal models there are positives and negatives of each human *in vitro* model, as outlined in section 1.4.3.3.

The aim of this study was to examine the role of the latency-associated ncRNAs on human neurons, particularly concentrating on the effect of these ncRNAs on the neuronal transcriptome. As such, a human neuronal model of HSV-1 latency was needed. In this chapter an *in vitro* model for HSV-1 latency is tested using SH-SY5Y cells. As discussed in chapter 1.4.3.4, these are a thrice cloned subline of the neuroblastoma cell line SK-N-SH, that originated from a metastatic bone tumour biopsy (Encinas et al., 2000). This cell line has been shown to be successfully differentiated into cholinergic or dopaminergic neurons and has been used extensively in the Parkinson's field (Christensen et al., 2011; Kovalevich & Langford, 2013; Shipley et al., 2016). These cells, once differentiated into neurons, have also been shown to be capable of supporting a productive HSV-1 infection and infection with other herpes viruses such as Varicella Zoster virus (VZV) (Christensen et al., 2011; Shipley et al., 2017). It follows that it may be a useful model for HSV-1 latency; this is what this chapter explores.

2.2 Results

2.2.1 SH-SY5Y change morphology and stain positively for neuronal markers following differentiation

SH-SY5Y cells were differentiated in a two-week, two-step process similar to previously published methods by Encinas *et al.* and Shipley *et al.* (Encinas et al., 2000; Shipley et al., 2016). All three protocols work by the gradual addition of neurotrophic factors – retinoic acid (RA) and brain derived neurotrophic factor (BDNF), with simultaneous removal of serum to restrict growth of epithelial-like cells in the SH-SY5Y population.

The combination of the addition of neurotrophic factors and reduction of serum encourages neuronal differentiation and restricts growth of epithelial-like cells. The cells are also plated onto an extracellular matrix or collagen coating to facilitate their adherence and differentiation. However, in the protocol set up by Encinas *et al.* the neurotrophic factors are added sequentially whereas in this study and the work by Shipley *et al.,* the RA and BDNF are added in combination as it has been shown that BDNF enhances the differentiating effects of RA (Encinas et al., 2000; Shipley et al., 2016). The protocol used in this thesis was adapted to incorporate the use of neurobasal media in the differentiation media, which is in keeping with the protocol demonstrated by Shipley and colleagues. Using neurobasal media seemed to enable the cells to be cultured for longer. Whilst optimizing this system, it was observed that full differentiation of the SH-SY5Y cells into neuron-like cells could be achieved by 2 weeks, therefore, the protocol used in this thesis was shortened compared to that published by Shipley *et al*. The differentiation protocol developed in this study is described in more detail in Materials and Methods (chapter 6.2.1.2) and is outlined schematically in figure 2.1A.

This differentiation process results in a morphological change. The SH-SY5Y cells transform from clusters of a mixed population of epithelial-like and neuronal-like cells (figure 2.1B) to just those neuron-like cells (figure 2.1C) and finally to a homogenous population of cells with a neuronal morphology, including the development of neurites (fig 2.1D).

The SH-SY5Y cells appear to be neuronal in terms of morphology with the development of cellular processes, however, to verify this the cells were tested for expression of selected neuronal markers. The neuronal markers; neurofilament medium (NEFM), neurofilament heavy (NF-H), microtubule associated protein 2 (MAP2), phosphorylated neurofilament H (pNF-H) and synaptophysin (SYP) were all tested by immunofluorescent staining using the relevant antibodies at the dilutions indicated in Materials and Methods (table 6.1.2) alongside DAPI nuclear staining. Cells were either fixed prior to differentiation (figures 2.1E, G, I and K) or immediately following differentiation (14 days after the initiation of differentiation) (figures 2.1F, H, J and L) before immunofluorescent staining and imaging.

Figure 2.1 SH-SY5Y cells change morphology and stain positively for neuronal markers following differentiation

A) Schematic representation of SH-SY5Y neuronal differentiation protocol

B-D) Photomicrographs of the morphological changes of SH-SY5Y cells during the differentiation outlined in (A). From undifferentiated SH-SY5Y (B) to differentiating SH-SY5Y (C) and finally fully differentiated SH-SY5Y cells (C). Scale bars represent 100μm.

E) Fluorescent photomicrograph showing immunostaining of neurofilament medium (NF-M) (green) alongside DAPI nuclear staining (blue) in SH-SY5Y cells fixed prior to differentiation, scale bar represents 100μm.

F) Fluorescent photomicrograph showing immunostaining of neurofilament medium (NF-M) (green) alongside DAPI nuclear staining (blue) in differentiated SH-SY5Y cells, grown in differentiation media 2 (methods 6.1.7) and fixed 14 days following the initiation of differentiation, scale bar represents 100μm.

G) Fluorescent photomicrograph showing immunostaining of microtubule associated protein 2 (MAP2) (green) alongside DAPI nuclear staining (blue) in SH-SY5Y cells fixed prior to differentiation, 100 times magnified, scale bar represents 100μm.

H) Fluorescent photomicrograph showing immunostaining of microtubule associated protein 2 (MAP2) (green) alongside DAPI nuclear staining (blue) in differentiated SH-SY5Y cells, grown in differentiation media 1 (methods 6.1.7) and fixed 14 days following the initiation of differentiation, 100x magnified, scale bar represents 100μm.

I) Fluorescent photomicrograph showing immunostaining of phosphorylated neurofilament heavy (pNF-H) (red) alongside DAPI nuclear staining (blue) in SH-SY5Y cells fixed prior to differentiation, 200x magnified, scale bar represents 50μm.

J) Fluorescent photomicrograph showing immunostaining of phosphorylated neurofilament heavy (pNF-H) (red) alongside DAPI nuclear staining (blue) in differentiated SH-SY5Y cells, grown in differentiation media 1 (methods 6.1.7) and fixed 14 days following the initiation of differentiation, 200x magnified, scale bar represents 50μm.

K) Fluorescent photomicrograph showing immunostaining of synaptophysin (SYP) (green) alongside DAPI nuclear staining (blue) in SH-SY5Y cells fixed prior to differentiation, 400x magnified, scale bar represents 25μm.

L) Fluorescent photomicrograph showing immunostaining of synaptophysin (SYP)

(green) alongside DAPI nuclear staining (blue) in differentiated SH-SY5Y cells, grown in differentiation media 1 (methods 6.1.7) and fixed 14 days following the initiation of differentiation, 400x magnified, scale bar represents 25μm.

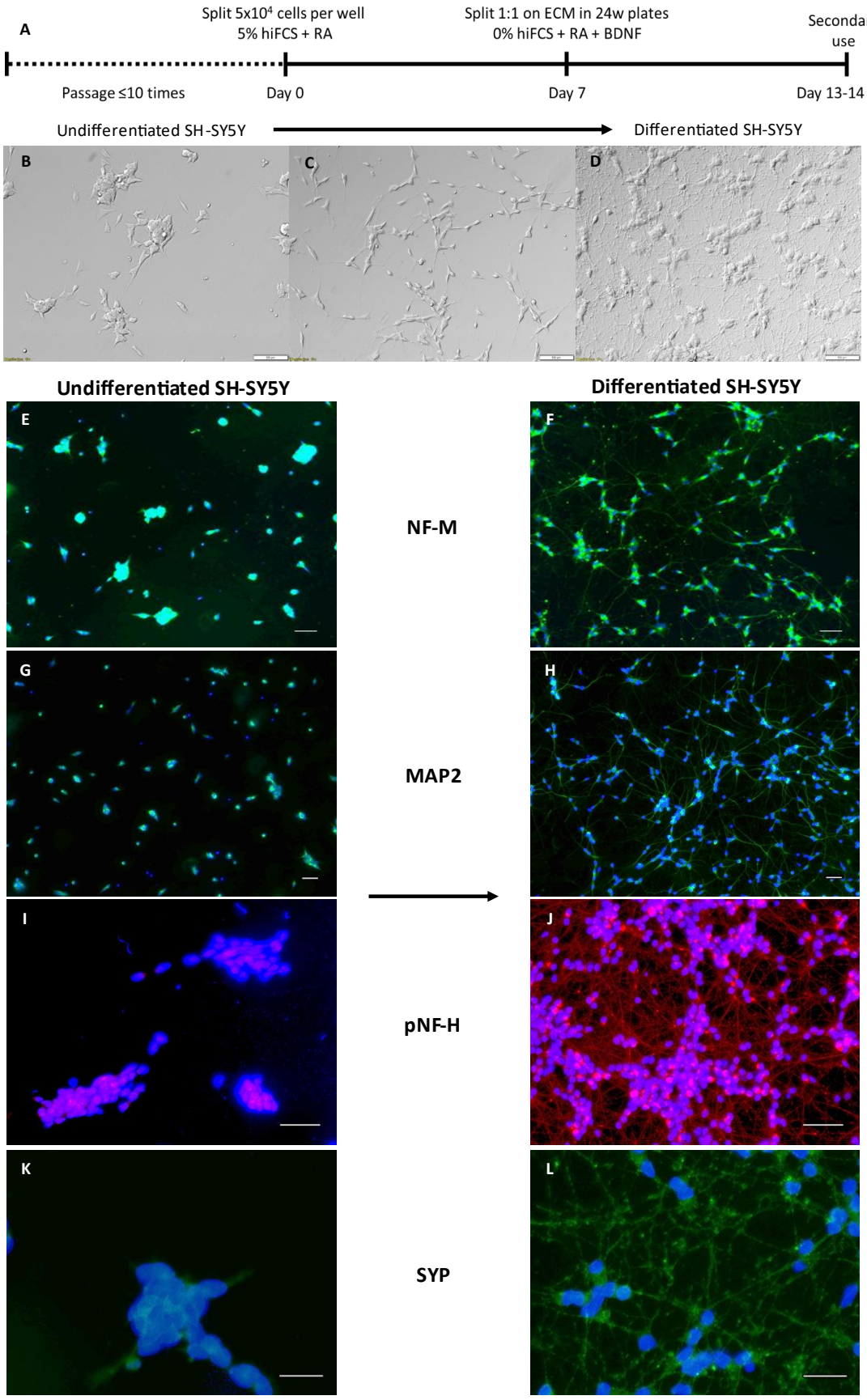


Figure 2.1 Following differentiation SH-SY5Y cells change morphology and stain positively for neuronal markers

Microtubule associated protein 2 is a member of the microtubule associated protein family involved in microtubule assembly, which is in turn essential for neuritogenesis (Ramkumar et al., 2018). Neurofilaments are neuronal structural proteins that are major components of the neuronal cytoskeleton and are required for the development of neurons and their processes, therefore the presence of either NEFM or NF-H could be indicative of axonal development (Park et al., 2016). SYP is a membrane glycoprotein of presynaptic vesicles in neurons, required for neurotransmitter release (Kokotos et al., 2019). Therefore, the presence of these markers would suggest that these cells are neuronal.

The differentiated SH-SY5Y cells stain positively for all three of these neuronal markers (figure 2.1 E-L). There is some positive staining prior to differentiation, but this could be owing to the cells having neuronal lineage at this stage. Furthermore, both pNF-H and SYP are observed in higher intensities following differentiation. In addition, the immunostaining for all four proteins at least partially localise to the cellular projections following differentiation, which are not present in the undifferentiated SH-SY5Y cells, suggesting axonal development occurs after differentiation, which is indicative of neuronal cell growth.

From these results and other studies (Encinas et al. 2000; Shipley et al., 2016), it is clear that the SH-SY5Y cells differentiate appropriately into neuronal-like cells in a controllable manner. Additionally, they are a human cell-line. Consequently, SH-SY5Y cells are a suitable model for studies of HSV-1 infection and latency.

2.2.2 Differentiated SH-SY5Y cells are permissive to HSV-1 infection

Once it was validated that the SH-SY5Y could effectively be differentiated into neuronal type cells, as has been shown in the literature (Encinas et al., 2000; Shipley et al., 2016), the next stage was to test the permissivity of these cells following wildtype (WT) HSV-1 infection.

African green monkey kidney epithelial cells (Vero cells) are frequently used for lytic HSV-1 *in vitro* studies and are permissive to HSV-1 infection (Cui et al., 2006; Milne et al., 2005). As such, Vero cells act as a positive control for virus production. Vero cells and differentiated SH-SY5Y cells were infected in parallel with wildtype (WT) strain 17syn+ HSV-1 at an MOI of 3. The infected cells and supernatant were frozen at -80 degrees Celsius at the time of virus input and 1, 5, 7, 9, 11, 13 and 25 hours post infection before thawing, scraping and diluting onto Vero cell monolayers to quantify progeny virus by plaque assay. Plaque forming units were then counted and normalised to the amount of virus input for either cell type. The results are displayed in figure 2.2A. Virus production

initially occurred more slowly in SH-SY5Y cells than Vero cells, with viral titres relative to input virus load on average 7-fold lower in SH-SY5Y infected cells compared to Vero cells between 7 and 13hpi. However, by 25hpi, equivalent titres of virus were reached in SH-SY5Y cells and Vero cells. From this result, it is evident that differentiated SH-SY5Y cells can support an HSV-1 infection, but HSV-1 production is delayed compared to infection in Vero cells.

These growth curve results indicate HSV-1 can replicate in differentiated SH-SY5Y cells, this was also observed at a single-cell level by use of an enhanced green fluorescent protein (eGFP) reporter virus SC16 HSV-1 strain; C12 (Arthur et al., 2001).This recombinant reporter virus expresses eGFP driven by the human cytomegalovirus (HCMV) immediate early (IE) 1 promotor inserted at the US5 locus in WT HSV-1 strain SC16 virus, and is replication-competent. The eGFP activity of this virus allows visualisation of cells infected with transcriptionally active HSV-1. This not only confirms whether the SH-SY5Y cells can be infected but also how many get infected at a given MOI.

SH-SY5Y cells were infected at a high MOI of 5, which should result in most of the cells being infected. The cells were imaged by fluorescence microscopy at 1, 3 and 7 dpi as indicated in figure 2.2B. The images show strong eGFP signal in most of the cells as early as 1dpi and throughout the 7 days of infection. This data supports the growth curve results showing that SH-SY5Y cells are permissive to productive HSV-1 infection. It also appears that at MOI 5 most, if not all, of the cells are infected, by 1dpi.

Figure 2.2 Differentiated SH-SY5Y cells are permissive to HSV-1 infection

A) Growth curves were performed on Vero cell monolayers of supernatant and cell lysate from MOI 3 infected differentiated SH-SY5Y cells or Vero cells at each timepoint indicated. The titres shown are normalised to the input viral titre. The average titre ± standard deviation was taken across 2 experiments each performed in duplicate.

B) Differentiated SH-SY5Y cells were infected with HSV-1 C12 eGFP reporter virus at MOI 5. Brightfield (BF) and fluorescent (eGFP) photomicrograph images were taken of the live infected cells at 1,3 and 7 days post infection as indicated. Scale bars represent 100μm.

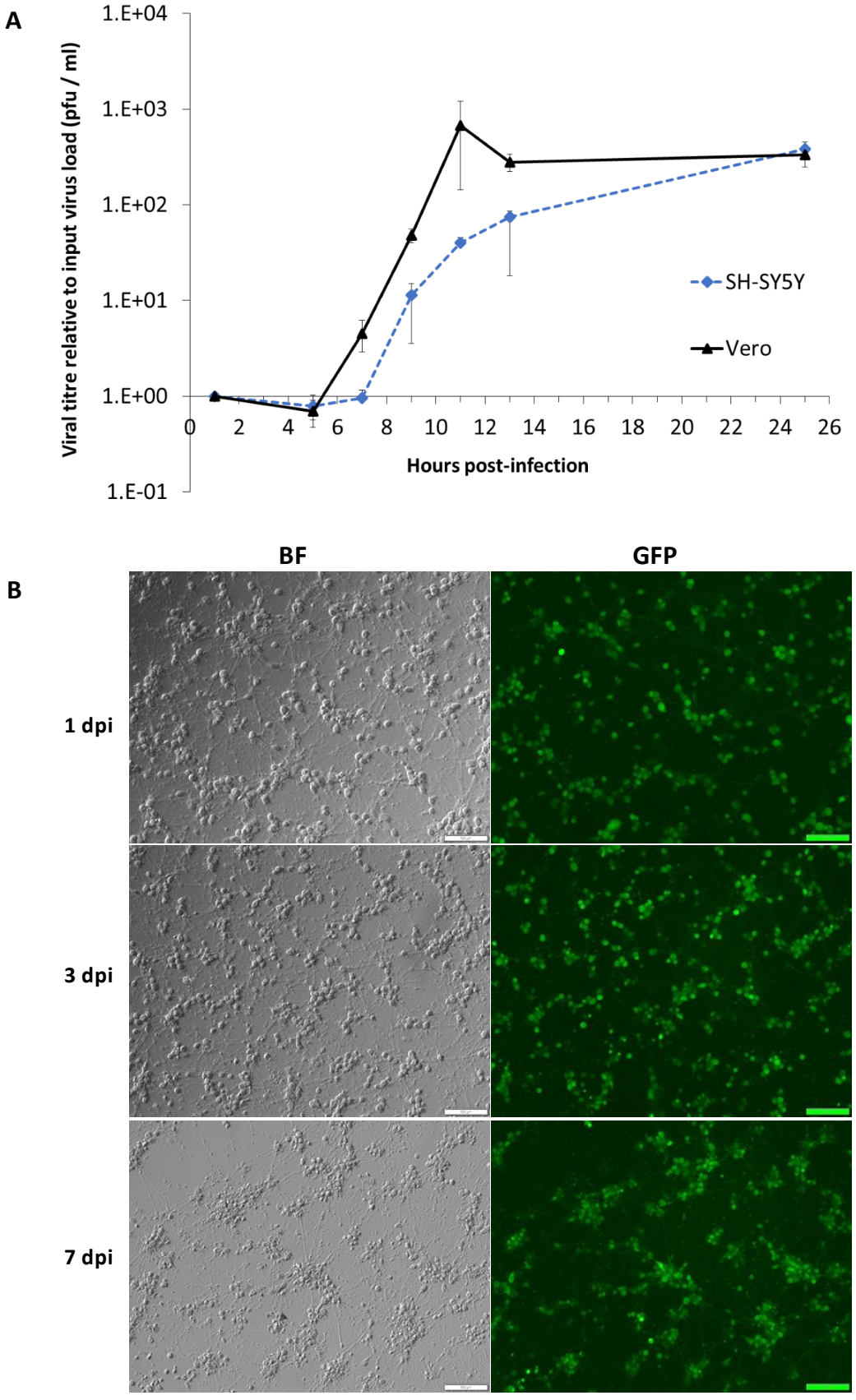


Figure 2.2 Differentiated SH-SY5Y cells are permissive to HSV-1 infection

2.2.3 Pre-treatment of differentiated SH-SY5Y cells with interferon alpha and acyclovir enables the establishment of a latent HSV-1 infection

Once it was established that SH-SY5Y cells could be differentiated into neuronal cells that support lytic HSV-1 infection, the next step was to determine whether a quiescent HSV-1 infection could be established in these cells for use as a model for HSV-1 latency.

The combination of a low MOI infection and the use of viral inhibitors has previously been successful in establishing a latent-like infection in human neurons as well as primary animal neurons (De Regge et al., 2010; Pourchet et al., 2017). This method was tested to try to establish latency in the differentiated SH-SY5Y cells.

Acyclovir blocks viral replication by acting as a deoxyguanosine triphosphate nucleoside analog that competitively inhibits viral DNA polymerase, causing DNA chain termination (Ibáñez et al., 2018; King & Madera, 1982). Acyclovir (ACV) specifically inhibits herpes virus replication as it is activated by a virally encoded protein – thymidine kinase. The type I interferons such as interferon alpha (IFNα) are inducible cytokine proteins that exert an antiviral effect on infected or neighbouring cells by an array of downstream signalling cascades (Samuel, 2001). Together, these inhibitors limit productive HSV-1 infection.

To test the use of the inhibitor pre-treatment as a means of establishing latency, the neuronally differentiated SH-SY5Y cells were pre-treated with acyclovir (ACV) (100μM) and interferon-alpha (IFNα) (500U/ml) which are concentrations previously shown to be effective at establishing a latent HSV-1 infection (De Regge et al., 2010; Kobayashi, Kim et al., 2012; Pourchet et al., 2017). One day following inhibitor pre-treatment, both the inhibitor-treated and control mock-treated differentiated SH-SY5Y cells were infected at an MOI of 0.1, using C12 eGFP reporter virus. This reporter virus allows visualisation of how many cells are infected and whether the virus is transcriptionally active. The infected SH-SY5Y cells were imaged by fluorescence microscopy, and eGFP fluorescence was analysed at 6hpi, 1dpi, 3dpi and 7dpi (Figure 2.3). The inhibitor pre-treated cells were infected in the presence of ACV and IFNα to maintain the block to infection and their concentrations were maintained throughout the experiment.

In the mock-treated, control infected cells there was an increase in the number of cells expressing eGFP over the first 3 days of infection and by 3dpi most, if not all, of the cells were expressing eGFP, this continued until the 7dpi time point. By comparison, the ACV and IFNα pre-treated (and infected) cells showed minimal to no eGFP activity throughout

the 7 days of infection (figure 2.3). Additionally, by 7dpi the untreated infected SH-SY5Y cells exhibited cytopathic effect, with rounding of the cell bodies and retraction of axons as observed by brightfield microscopy whereas the inhibitor treated cells showed no such cytopathic effect (figure 2.3). The ACV and IFNα therefore had a strong restrictive effect on viral replication in this system, in comparison to mock-treated infected cells.

This data suggests that the neuronal SH-SY5Y cells can support a latent HSV-1 infection in the presence of viral inhibitors; ACV and IFNα, at a low MOI infection with replication competent HSV-1.

Figure 2.3 Establishing a latent HSV-1 infection using replication competent HSV-1 and viral inhibitors

Differentiated SH-SY5Y cells were infected with HSV-1 C12 eGFP reporter virus at an MOI of 0.1, with or without 1 day pre-treatment with acyclovir (100μM) and interferon alpha (500U/ml) – as indicated. Brightfield (BF) and fluorescent (eGFP) photomicrographs were taken at 6hpi (scale bar shows 50μm) and at 1, 3 and 7dpi - Scale bars represent 100μm.

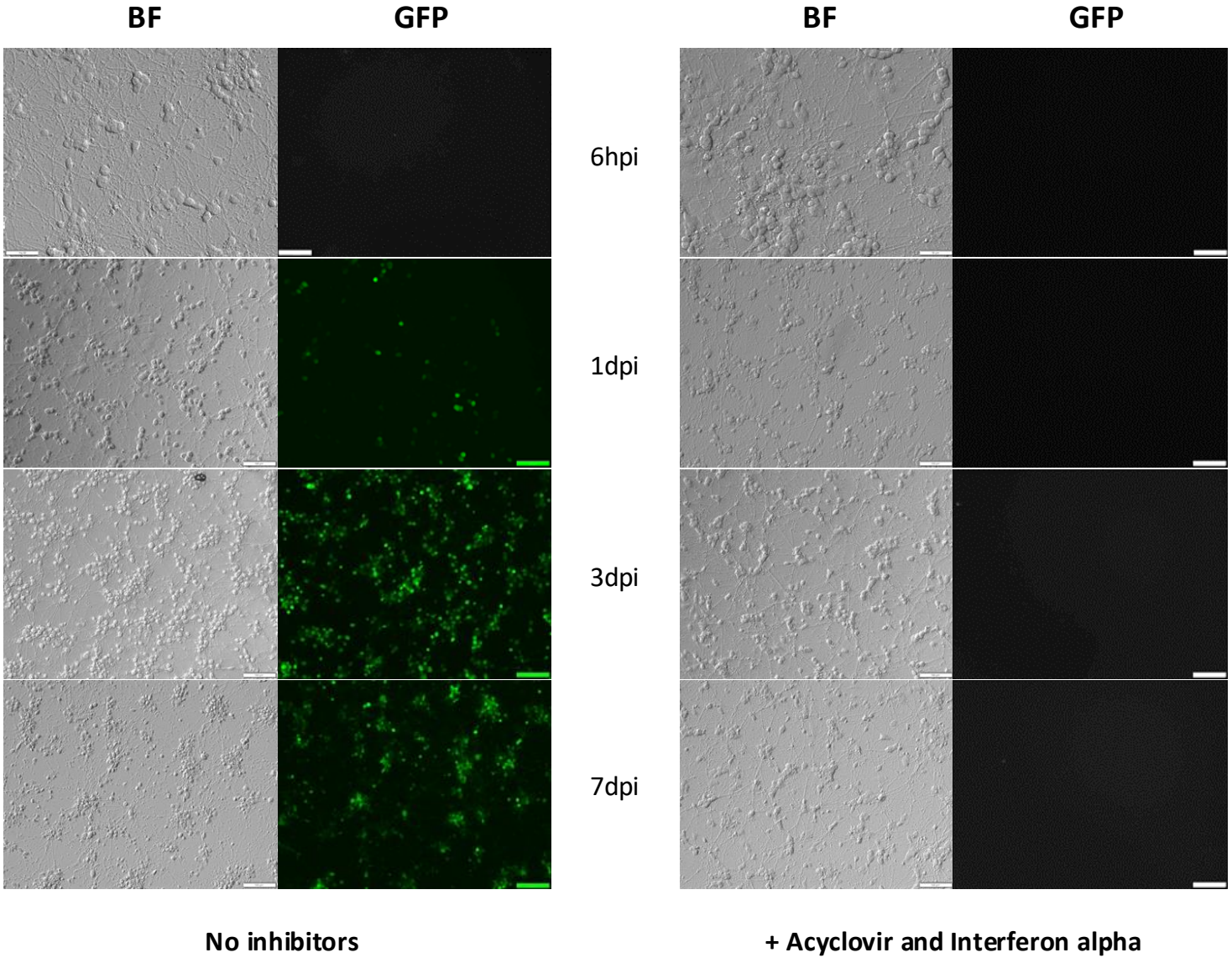


Figure 2.3 Establishing a latent HSV-1 infection using WT HSV-1 and inhibitors

2.2.4 Latent HSV-1 infection established using replication competent virus and viral inhibitors spontaneously reactivates following removal of viral inhibitors

The data in the previous section demonstrates that a latent HSV-1 infection can be established in the neuronal SH-SY5Y cells by using viral inhibitors, which is similar to data seen in the literature using other neuronal models (Kobayashi, Kim et al., 2012; Pourchet et al., 2017). However, the use of viral inhibitors could skew phenotypes that this thesis aims to test. For example, the effect that the HSV-1 latency associated ncRNAs has on the neuronal transcriptome may be masked if these inhibitors also influence the neuronal transcriptome, such as external IFNα inducing interferon-stimulated genes. Therefore, it was important to test whether HSV-1 latency could be stably maintained in differentiated SH-SY5Y cells once the viral inhibitors are removed.

To test this, the latently infected, inhibitor pre-treated cells shown in figure 2.3 were re-fed with neurobasal media without ACV or IFNα at 8dpi. The cells were then imaged at 1, 2, 5, 7, 9, 12 and 16 days post inhibitor removal. Up to 7 days post inhibitor removal, the virus remained mostly quiescent, with rare cells producing eGFP signal. However, from 9 days onwards the number of green cells observed, along with the intensity of eGFP signal, increased with time (figure 2.4). Therefore, in at least a proportion of the infected cells, the previously silenced viral IE CMV promotor that drives eGFP expression in the reporter virus, had become derepressed. This suggested that the latent HSV-1 genomes in these cells had become transcriptionally active, indicative of reactivation. This further indicates that following ACV and IFNα removal from infected but transcriptionally repressed quiescently infected SH-SY5Y cells, the WT HSV-1 reporter virus (C12) is capable of spontaneous reactivation. This spontaneous reactivation was observed to occur in few cells until 7dpi, following this there was either viral spread from reactivating cells or further reactivation as indicated by considerable eGFP expression. Spontaneous reactivation suggests a capability to reactivate that mimics biologically relevant latency. However, it also makes this model of latency difficult to predict and control. Therefore, it was necessary to explore other options that could more reliably maintain latency and that did not require the continual presence of inhibitors of viral replication.

Figure 2.4 Latent HSV-1 infection established using replication competent reporter virus (C12) and viral inhibitors spontaneously reactivates

The differentiated SH-SY5Y cells infected with HSV-1 C12 eGFP reporter virus at MOI 0.1 and pre-treated with acyclovir and interferon alpha, shown in figure 2.3, were refed with fresh media lacking viral inhibitors 8dpi to remove these viral inhibitors. Brightfield (BF) and fluorescent (eGFP) live images were taken at 1-16 days post inhibitor removal as indicated. Scale bars represent 100μm.

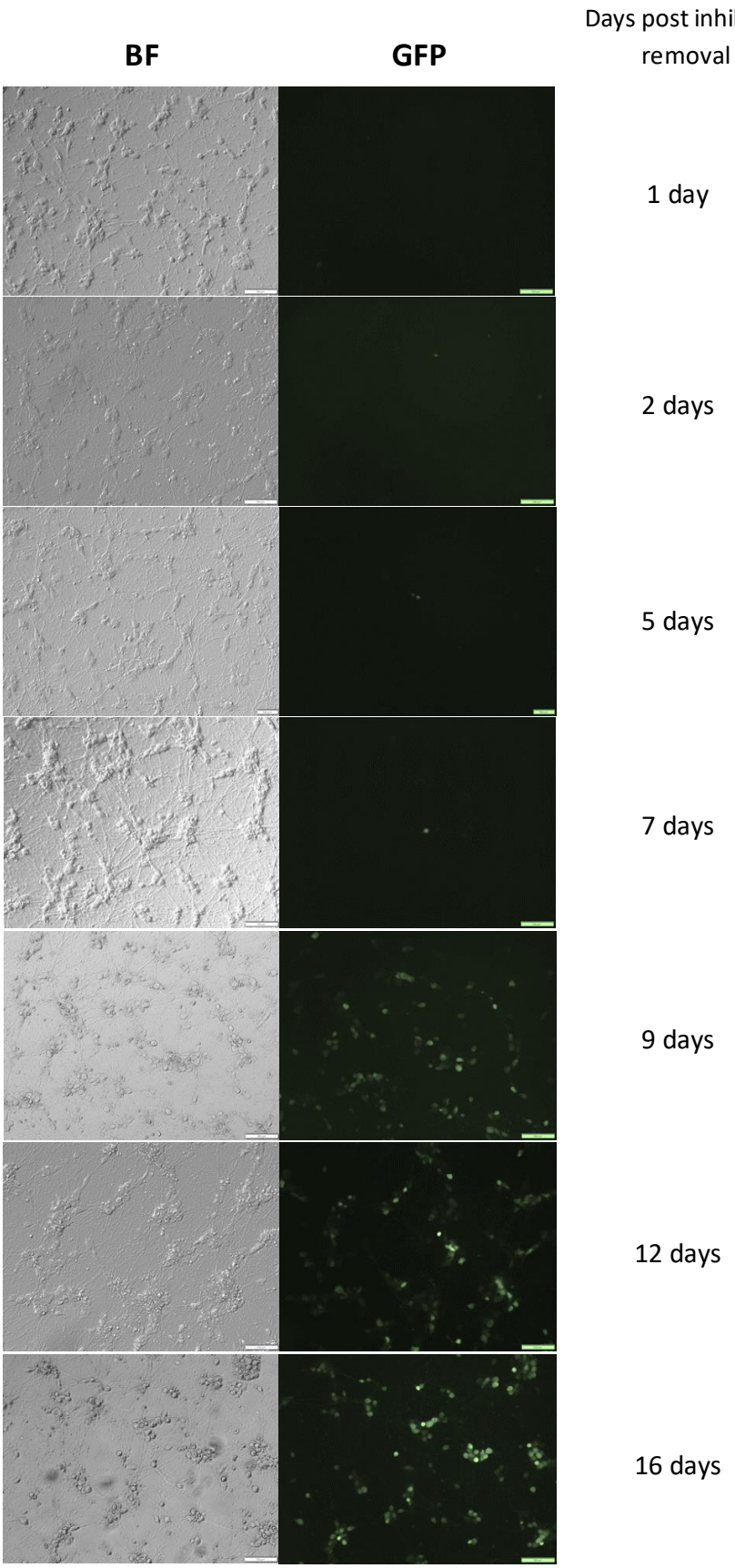


Figure 2.4 Latent HSV-1 infection using WT HSV-1 and inhibitors spontaneously reactivates

2.2.5 Establishing a quiescent HSV-1 infection by utilising replication-defective HSV-1 mutant, in1382

Addition of viral inhibitors to WT HSV-1 infection of differentiated SH-SY5Y cells allows for the establishment of a quiescent infection. However, the spontaneous reactivation observed following inhibitor removal – although perhaps reflective of *in vivo* human HSV-1 infection capable of reactivation – could be problematic when testing various latent phenotypes over a prolonged period. The use of a replication-defective virus removes the need for viral inhibitors to establish latency and blocks the potential for spontaneous reactivation as these viruses lack the key viral transactivators needed to initiate lytic replication and hence viral reactivation. In1382 is a previously characterised replication-defective HSV-1 mutant generated by Preston *et al.* (Preston et al., 1997; Preston & McFarlane, 1998) that contains mutations in 3 of the key viral transactivators.

In1382 has a RING domain deletion in ICP0 that as a result blocks ICP0 activation of gene expression, viral growth and its E3 ubiquitin ligase function responsible for counteracting host intrinsic defence mechanisms such as nuclear domain 10 (Boutell et al., 2002; Everett et al., 1995; Lanfranca et al., 2014; Smith et al., 2011; Preston & McFarlane, 1998). Additionally, in1382 has a 12bp insertion in VP16 that limits its interaction with the cellular proteins Oct-1 and HCF that in turn enables VP16 binding to DNA and activation of immediate early gene transcription (Lai & Herr, 1997; Preston et al., 1997; Preston & McFarlane, 1998). Finally, in1382 also contains a temperature sensitive mutation in ICP4 that means when infecting at a temperature above 38 degrees Celsius (^{o}C) the virus fails to produce functional ICP4 which is essential for early and late gene expression (figure 2.5A) (Preston et al., 1997; Preston & McFarlane, 1998). This renders the virus unable to undergo full lytic replication. Therefore, the virus is stopped in a quiescent state (Preston et al., 1997).

Figure 2.5 Establishing a latent HSV-1 infection using replication defective HSV-1 (in1382)

A) Schematic diagram of the replication defective in1382 virus genome.

B) Differentiated SH-SY5Y cells were infected with HSV-1 in1382 at MOI 1 at 39^{o}C (to block functional ICP4 production) in duplicate, RNA was collected at the timepoints shown, and qRT-PCR performed on HSV genes and normalised to beta actin to examine their transcription over time, until 7dpi.

C) Differentiated SH-SY5Y cells were infected with HSV-1 in1382 at MOI 5, RNA was collected at the timepoints shown, and qRT-PCR performed on HSV genes normalised to beta actin to examine their transcription over time, until 14dpi. The average ± the standard deviation has been taken across 3 technical repeats and 2-3 biological repeats.

D) Differentiated SH-SY5Y cells were infected with HSV-1 in1382 at MOI 5, DNA was collected, and qPCR performed to viral ICP0 DNA load normalised to cellular GAPDH DNA. DNA was extracted at 1hpi, 3dpi, 5dpi, 11dpi and 14dpi.

The average viral DNA copy number per GAPDH copy number across 5, 3, 2, 3 and 3 replicates for the 1hpi to 14dpi time points respectively are shown with standard deviation represented in error bars for those time points with 3 or more biological repeats. A linear trendline is plotted and represented by the dotted blue line.

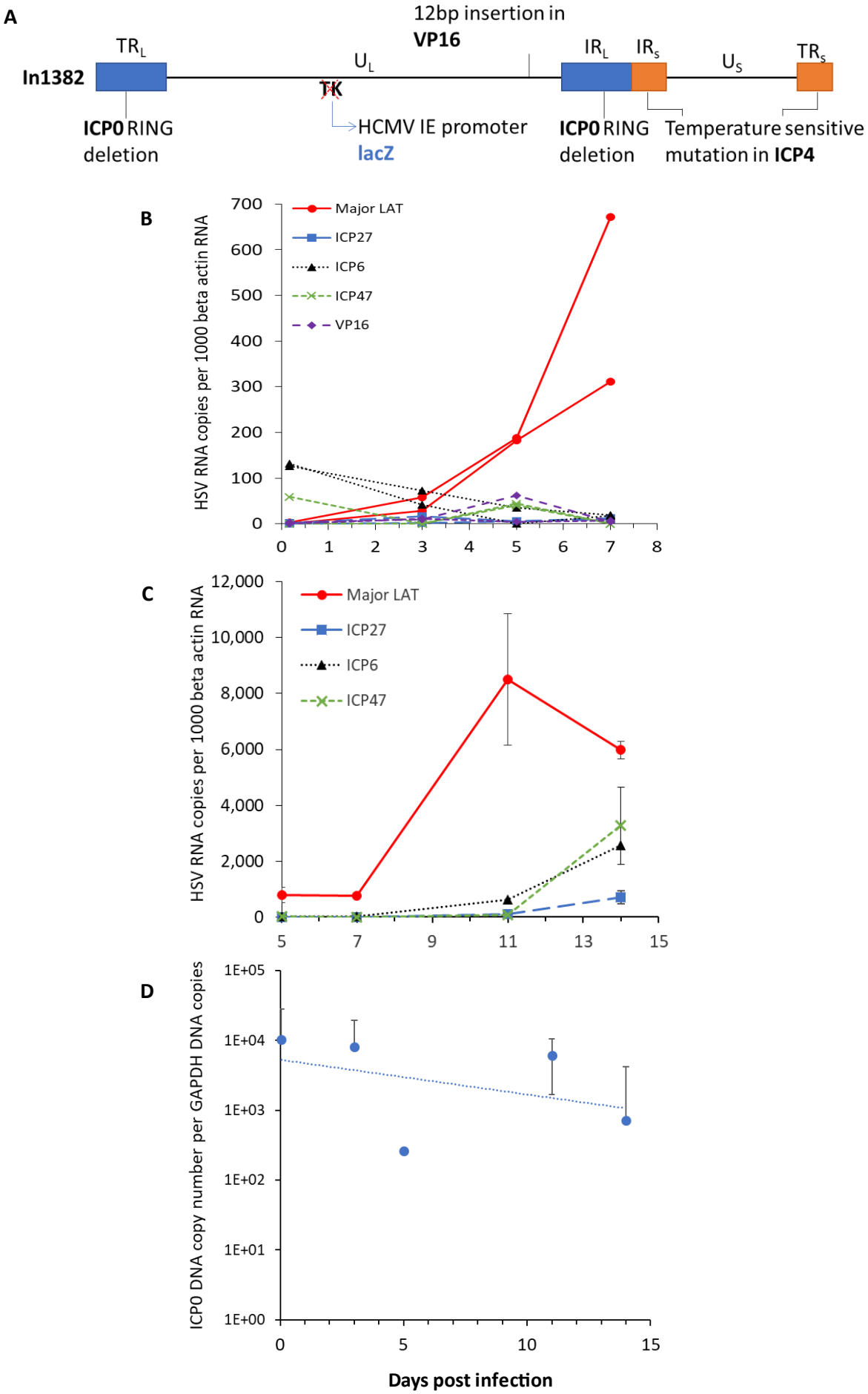


Figure 2.5 Establishing a latent HSV-1 infection using replication defective HSV-1 (in1382)

In an initial pilot experiment SH-SY5Y cells were differentiated as described previously and infected at an MOI of 1 with in1382 virus. RNA was then harvested in duplicate wells at 4hpi and 3, 5 and 7dpi, and qRT-PCRs were performed for major LAT and well as lytic genes; ICP27, ICP6, ICP47 and VP16, and each was normalised to a cellular gene – Beta actin (figure 2.5B). ICP27 and ICP47 are both IE genes, ICP6 is an E gene while VP16 is a virion protein that promotes onwards IE expression (see chapter 1.3.2 for more detail). At 4hpi, there was virtually no LAT transcription and low-level transcription of lytic genes, with 131 and 126 copies per 1000 beta actin of ICP6 and up to 59 copies per 1000 beta actin of ICP47, while ICP27 and VP16 show little or no transcription (less than 3 copies per 1000 beta actin copies). Over the 7 days of infection transcription of these lytic genes decreased to or remained at minimally detectable levels, with the highest lytic gene expression at 7dpi being ICP6 with just 18 and 19 copies per 1000 beta actin. By contrast, LAT expression increases over the course of infection, reaching a peak of either 672 or 311 copies per 1000 beta actin copies (figure 2.5B).

The results from this initial experiment reflect the expected transcription profile of HSV-1 latency, with LAT expression increasing during the establishment of latency while lytic expression decreases from already minimal levels.

Based on the results of this pilot experiment, a second experiment was performed, however, in this instance differentiated SH-SY5Y were infected with in1382 at an MOI of 5. This was to enable all the cells to be infected to try to reach higher LAT levels and, as it appeared that lytic transcription was tightly restricted from the previous experiment, it was hypothesized that these would remain low regardless of the higher MOI. In addition, the experiment examined later time points, up to 14dpi, to test whether LAT expression continued increasing beyond 7dpi. Neuronal SH-SY5Y cells were infected with in1382 at MOI 5 and RNA was harvested at 5, 7, 11 and 14dpi for qRT-PCRs for major LAT and lytic genes; ICP27, ICP6, ICP47, all normalised to beta actin gene expression (figure 2.5C). The results show that with a higher MOI 5, there is a higher level of LAT expression at 7dpi, with an average of 777 copies per 1000 beta actin copies. Furthermore, taking this MOI 5 infection out to 14dpi, an even higher level of LAT expression is reached (figure 2.5C). Major LAT copy number reaches its peak at 11dpi with an average of 8,498 major LAT copies per 1000 Beta actin, compared to an average of just 672 and 311 copies per 1000 Beta actin copies at 7dpi and a lower MOI of 1 in the previous experiment (figure 2.5B). However, at 14dpi LAT expression decreases, with an average of 5984 copies per 1000 Beta actin copies.

Lytic gene expression starts off low and remains less than LAT expression throughout the infection time course. However, there is an increase of lytic gene expression by 14dpi that coincides with a decrease in LAT expression. Nevertheless, even the most highly expressed lytic gene at 14dpi, ICP47, is still approximately half as abundant as major LAT, with 3270 copies per 1000 beta actin copies.

This data suggests that latency can be established by infecting differentiated SH-SY5Y cells with replication defective HSV-1 strain in1382. LAT is strongly expressed, and lytic gene expression remains low by comparison. At the later time points, 11dpi onwards, increased transcription of lytic genes can be observed – indicating that there may be some instability in the system at late time points that result in derepression of a proportion of the latent genomes.

The decrease in LAT expression and slight increase in lytic gene expression at the later time points could also suggest there was some cytotoxicity to cells, and on observation the cells exhibited cytopathic effect at 14dpi. Therefore, to determine to what extent latent genomes were retained over time in this system, differentiated SH-SY5Y cells were infected with in1382 at an MOI of 5 before harvesting the DNA at 1hpi, 3dpi, 5dpi, 11dpi and 14dpi. PCRs were performed to ICP0 and cellular GAPDH DNA. ICP0 copies are normalised to GAPDH copy number. The average viral DNA copy number per GAPDH and linear trendline are displayed in figure 2.5D. While there are some fluctuations in DNA levels throughout the 14-day infection, with an outlier at 5dpi, overall there is a slight decrease of DNA levels and therefore virus copy numbers over time, as indicated by decrease observed with the linear trendline. Viral DNA copy numbers per GAPDH decrease from 10316 to 719 over the course of infection (figure 2.5D). This indicates that there is some loss of DNA over the course of this 14-day infection, most of this loss occurs between 11 and 14 dpi where DNA copies decreased from 6081 to 719 viral copies per 1000 beta actin copies. This loss of viral DNA copies coincides with the cytotoxicity observed at 14 dpi alongside the drop in LAT RNA expression and increase in lytic gene expression in the previous qRT-PCR experiment (figure 2.5C).

2.2.6 Examining the infection kinetics of quiescent infection by single-cell analysis

In1382 is not only a replication-defective mutant, but also a reporter virus. It expresses lacZ driven by the HCMV major immediate early promotor (MIEP). Therefore, infected cells produce beta galactosidase and are seen as blue following staining (methods 6.2.4.1)(figure 2.6A). Cells were infected as described previously with in1382 at an MOI of 5 and fixed in 4% PFA for lacZ assay (methods chapter 6.2.4.1) at 1, 7, 11 and 14 days post infection. At 1dpi, at MOI of 5, most, if not all the cells are blue and

therefore infected with transcriptionally active virus (figure 2.6A). Over the course of the infection, the number of blue cells and intensity of blue qualitatively reduced, consistent with repression of MIEP activity during the establishment of latency. This adds to the evidence that in1382 infection of differentiated SH-SY5Y cells results in a quiescent infection and makes this system a potentially useful model for HSV-1 latency.

Although the reduction in beta galactosidase observed (figure 2.6A) could be due to repression of lytic gene expression that accompanies the establishment of latency, it is possible that the slight loss of viral DNA over time (figure 2.5D) could have contributed to this observation. Therefore, to further characterise this system and verify its use as a latency model, LAT expression was examined at a single cell level. In parallel to the analysis of lac Z expression over time, differentiated SH-SY5Y cells were infected with in1382 virus at MOI 5 and then fixed in 4% PFA at 1, 6, 11 and 14dpi. The *in situ* hybridisation technique of RNAscope (methods chapter 6.2.4.3) that allows monitoring of RNA localisation at a single-cell level was then performed on these infected cells using probes for the LAT intron, shown in green (figure 2.6B+D).

LacZ results revealed that at MOI 5 all the cells can be infected and that there was a slight reduction in CMV MIEP promoter activity as the infection progressed. RNAscope, performed in parallel, demonstrated how many of these infected cells expressed major LAT and how this LAT expression changed over time. The intensity and frequency of green fluorescence, and therefore LAT intron-producing cells qualitatively increased from 1dpi up until 11dpi, where it reached a peak and then decreased slightly by 14dpi (figure 2.6B).

The percentage of LAT-positive SH-SY5Y cells were then quantified using ImageJ by both automatically and manually counting any green-containing cells (methods 6.2.7.3), as a percentage of the total number of cells per image as indicated by DAPI staining of nuclei. Quantification of the LAT positive cells confirmed the trend observed by microscopy whereby the peak of LAT expression occurs at 11dpi (2.6 B and C). Here, the average number of LAT expressing SH-SY5Y cells increases from 13.2% at 7dpi to 27.8% at 11dpi, after which there is a slight dip at 14dpi to 25.57% of cells expressing LAT. Although this is only a small reduction of 2.23%, which given the reduction of DNA copy number over time could be as a result of loss of infected cells, this data is also consistent with the qRT-PCR data showing LAT expression peaks at 11dpi and then decreases slightly by 14dpi (figure 2.5C).

Figure 2.6 Characterising the in1382 replication defective HSV-1 model of latency by single-cell analysis

A) Photomicrographs showing beta-galactosidase assay of differentiated SH-SY5Y cells infected with HSV-1 in1382 at MOI 5, at 39ºC. The cells were fixed for staining at 1, 7, 11 and 14dpi. Scale bars represent 50µm.

B) Fluorescent photomicrographs of differentiated SH-SY5Y cells infected with HSV-1 in1382 at MOI 5, infected in parallel to the cells shown in (A). Cells were fixed at the same time points as in (A) as indicated and RNAscope was performed using a probe to the LAT intron (green) and nuclear DAPI stained (blue). Scale bars represent 50µm. Cells in white dotted outlined boxes are enlarged below the primary image, into boxes with black borders.

C) Quantification of LAT intron positive cells of MOI 5 in1382 infected differentiated SH-SY5Y cells, where each circle represents a different frame imaged and quantified for green LAT-positive cells in imageJ and the lines shows the mean percentage at each time point from a given experiment. The blue circles represent quantification of the infected cells imaged in (B), the green triangles show an independent experiment fixed at 7 and 11 days post infection (MOI 5), performed in parallel to the infection in (2.5C).

D) A high magnification (600x) example image of MOI 5 in1382 infected differentiated SH-SY5Y cells RNAscope probed for LAT intron (green) and DAPI stained (blue) 7 days post infection. Scale bar represents 10µm.

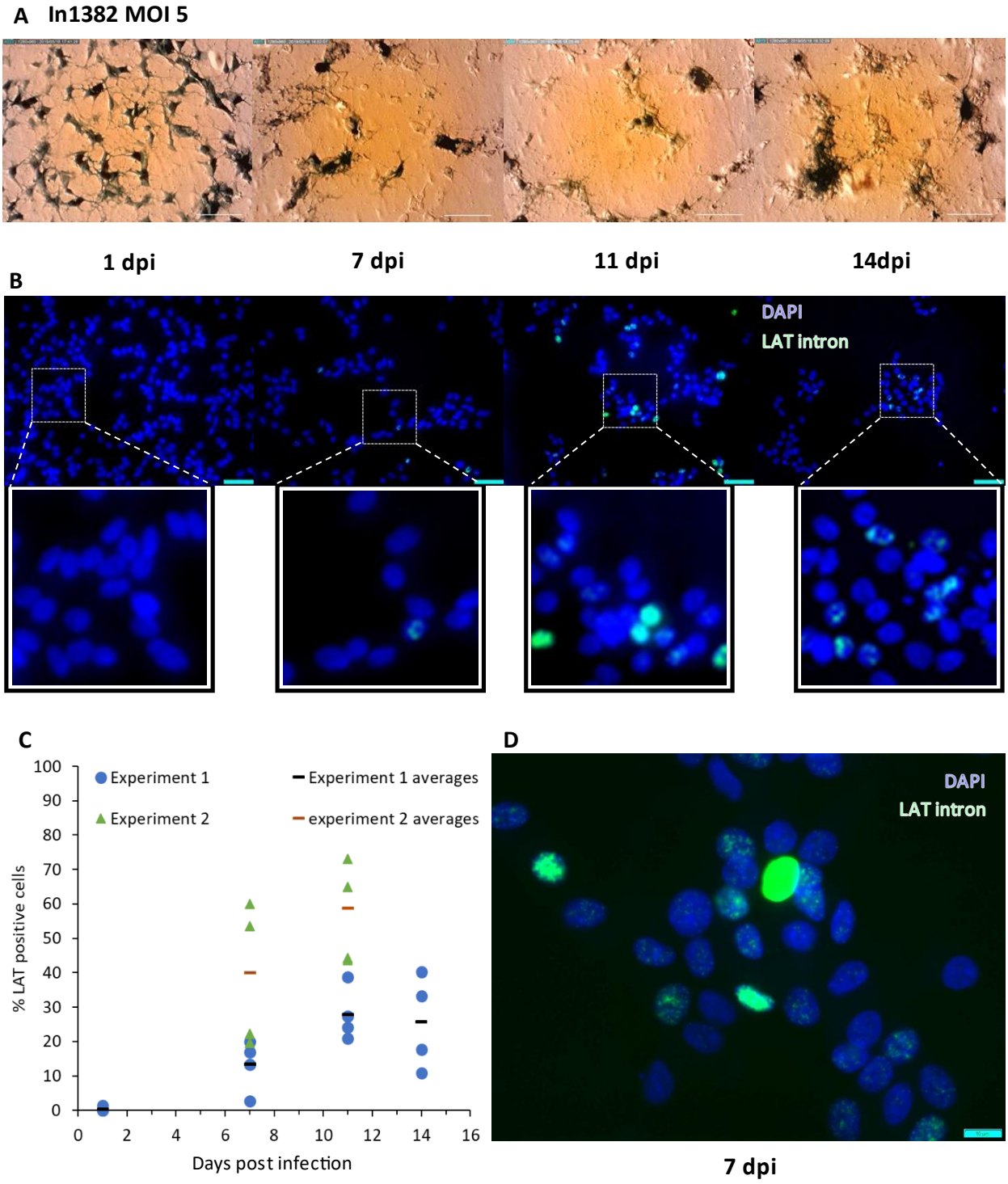


Figure 2.6 Characterising the in1382 replication defective HSV-1 model of latency using single-cell analysis

There is some variation in the system, as shown by a repeat experiment only performed at day 7 and 11dpi where the percentage of LAT expressing cells was higher, with an average of 39.9% and 58.7% SH-SY5Y cells LAT positive at 7 and 11dpi, respectively (figure 2.6C). Despite this variation, 11dpi consistently has the highest level of LAT expression. This second RNAscope experiment was performed in parallel to the qRT-PCR shown in figure 2.5C, which also indicates major LAT expression increased from 7 to 11dpi, where it reached its peak. Therefore, the RNAscope and qRT-PCR expression data support one other.

All the data characterising this replication-defective HSV-1 model of latency provides evidence that the system can be used to establish a quiescent HSV-1 infection, during which LAT is produced, with a peak in production at 11dpi, with an MOI of 5.

Fluorescence RNAscope is highly sensitive and specific and provides additional information aside from which cells are infected, such as the spatial organisation of LAT once expressed. To make use of this, in1382 infected cells probed for LAT by RNAscope were imaged at 600X magnification. Shown here is an example image of an MOI 5 in1382 infection at 7dpi, imaged at 600X magnification (figure 2.6D). This magnified image shows that LAT, either as the LAT intron or the primary transcript prior to splicing, is organised in punctae in the nucleus and that there are varying amounts of LAT produced in the different SH-SY5Y cells. Whether all the LAT positive cells eventually express the same amount of LAT is unclear, but it is important to note that there is heterogeneity in how much LAT is expressed among LAT-positive cells in this system at a given time.

In conclusion, the work described in this chapter shows that latency can be established using a replication-defective HSV-1 in differentiated SH-SY5Y cells. qRT-PCR showed that LAT expression increases until 11dpi while lytic gene expression is repressed. Concurrently, viral DNA was shown by PCR to be relatively stable until 11dpi and then decrease until 14dpi, which correlated with the appearance of CPE and limited lytic gene expression. Single-cell analysis of the system by RNAscope confirmed that LAT is produced in a proportion of cells, peaking at 11dpi with up to an average of 58.7% of in1382-infected differentiated SH-SY5Y cells positive for LAT expression. These techniques provide a useful way of monitoring the efficiency of latency establishment following in1382 infection of SH-SY5Y cells. Having various methods of monitoring the system, as described, offsets the variability that exists in the system, making it a useful model for assessing the effects of the latency associated ncRNAs in human neuronal cells.

2.3 Discussion

In this chapter, differentiated SH-SY5Y cells were verified for their use as a neuronal cell model of HSV-1 latency.

SH-SY5Y cells as a neuronal model

A differentiation protocol was developed that successfully differentiated SH-SY5Y cells into neuronal-like cells both in terms of morphology and presence of neuronal markers; neurofilament medium, microtubule associated protein 2, phosphorylated neurofilament heavy and synaptophysin. This agrees with data previously published showing that the SH-SY5Y cells successfully differentiate into cells that stain positively for neuronal markers, as well as a number of studies that show their use as a model for the investigation of neuronal degenerative diseases such as Parkinsons (Christensen et al., 2011; Encinas et al., 2000; Kovalevich & Langford, 2013; Shipley et al., 2016). The SH-SY5Y cells did show some expression of the candidate neuronal markers prior to differentiation, which could indicate that these markers do not necessarily reflect the development into fully differentiated neurons. Nonetheless, the strong positive staining observed post-differentiation in combination with morphological changes, showed neuronal marker expression along the cellular projections. This result is indicative of the development of axons, which are a characteristic feature of fully differentiated neurons. This, in agreement with the literature (Christensen et al., 2011; Kovalevich & Langford, 2013), suggests that the differentiated SH-SY5Y cells can be used as a neuronal cell model.

Establishing latency using replication-competent virus

From the results in this chapter it is evident that the differentiated SH-SY5Y cells are fully permissive to lytic HSV-1 infection and that although infection is delayed compared to a non-neuronal Vero cell line, the number of virus progeny produced from each cell type is similar by 25 hours following a high multiplicity infection. This observation is consistent with data from Shipley *et al*. who have demonstrated that differentiated SH-SY5Y cells support productive HSV-1 infection (Shipley et al., 2017). This suggested that the use of viral inhibitors would be required in order to establish a latent HSV-1 infection in this cell system.

Accordingly, the first approach to establish latency in neuronally differentiated SH-SY5Y cells utilised viral inhibitors ACV and IFNα with a low MOI WT infection, which has previously proven as an effective method to establish latency in neuronal cell-types such as human embryonic stem cell-derived neurons (Pourchet et al., 2017).

In this study, an eGFP reporter WT SC16 virus was utilised where eGFP is driven by the HCMV major immediate early promotor (MIEP) therefore enabling the visualisation of cells infected with a replication competent virus.

Pre-treating SH-SY5Y cells with ACV and IFNα at concentrations previously shown to induce latency during HSV-1 infection (De Regge et al., 2010; Pourchet et al., 2017) with an inoculum dose also tested previously in models of latency (Pourchet et al., 2017) led to a considerable restriction in eGFP activity. This suggested transcriptional repression, which is consistent with the establishment of latency (figure 2.3). However, lack of eGFP in isolation does not rule out the possibility of a lack of viral genomes altogether rather than the presence of transcriptionally silent genomes. An absence of viral genomes may occur if either no virus was added, or if infection resulted in cell death followed by removal of detached cells during media changes. However, the eGFP expression observed in the control C12 infection of differentiated SH-SY5Y cells that were not pre-treated with inhibitors, plus the lack of cytopathic effect (CPE) observed by brightfield imaging during this experiment suggests that virus was added to cells and that these cells remained viable throughout infection.

Additionally, 9 days post inhibitor removal from these cells, eGFP became visible and therefore genomes became transcriptionally active, confirming that viral genomes had been present but silenced prior to inhibitor removal. This appearance of eGFP and therefore de-repression of latent virus genomes suggests that there was spontaneous reactivation in this culture. However, from this data in isolation it is difficult to know whether this reactivation is as a result of many genomes reactivating throughout the culture or a small proportion of genomes escaping latency and reactivating and in turn enabling virus spread throughout the culture. Regardless, this spontaneous reactivation suggests that this system is unable to maintain latency beyond 7 days without the extended use of viral inhibitors.

Consequently, although the potential to reactivate is one of the indicators of biologically relevant latency and this system could be useful for testing various phenotypes, it is a limitation for this study. The aim of this project is to analyse the effect of the ncRNAs present during HSV-1 latency on the neuronal transcriptome and the prolonged use of viral inhibitors, to maintain latency and prevent reactivation, could also impact on the cellular transcriptome. Alternatively, without the extended use of viral inhibitors, uncontrolled spontaneous reactivation could occur, and any effects of the latency associated ncRNAs would likely be masked by lytic gene expression.

It is possible that a lower MOI of 0.01 might have been more effective at maintaining latency following inhibitor removal, as was the case in Pourchet *et al.*'s work in embryonic stem cells (Pourchet et al., 2017). Embryonic derived neuronal cells infected at an MOI of 0.1 with recombinant GFP-Us11 HSV-1 strain were shown to quickly become GFP positive, and therefore undergoing productive replication, following inhibitor removal, whereas at MOI 0.01 cells remained GFP negative (Pourchet et al., 2017). Nevertheless, such a low MOI might limit the likelihood of seeing an effect of the latency-associated ncRNAs when assessing changes in the host transcriptome by RNA-Seq.

Evidently this system could have been further scrutinised for its use as a latency model, either by testing different MOIs or by performing PCRs to verify LAT expression and restriction of lytic genes that would be expected during latency. However, due to the unpredictable nature of this system with its potential to fully reactivate alongside the intended use of this model (to examine the effects on neuronal transcriptome), an alternative approach was considered that should more reliably maintain latency.

Establishing latency using replication-defective virus

The second approach tested to establish latency in the neuronally differentiated SH-SY5Y cells was by infecting with a replication-defective virus, in1382. This virus lacks fully functional lytic viral transactivator genes (ICP0, ICP4 and VP16) and therefore is stopped in a quiescent state of infection.

Although major LAT has been shown not to be essential for latency, it has been proven to have an important role in the establishment and maintenance of latency and is the most abundant product during HSV-1 latency (Nicoll et al., 2012a, 2012b). As a result, LAT is often used as a surrogate marker for latency, in conjunction with transcriptional silencing of lytic genes. Therefore, to test for the establishment of latency, qRT-PCR was performed to test for the transcription of major LAT and 3-4 lytic transcripts over the course of an in1382 infection in neuronal SH-SY5Y cells.

From the initial 7-day experiment, which involved infection of differentiated SH-SY5Y cells at an MOI of 1 (figure 2.5B), major LAT expression increased over time while lytic gene transcription either decreased with time or remained low throughout infection. This pattern of gene expression closely reflects HSV-1 latency as it has been shown *in vivo,* whereby the number of LAT expressing cells increases with the establishment of latency (Margolis et al., 1992). This suggested potential for this system to be used as a model for HSV-1 latency.

The repeat experiment at a higher MOI (5 pfu/cell) and continued until 14dpi, showed that LAT expression increased beyond 7dpi, until 11dpi. The increased MOI in conjunction with this later time point, led to elevated abundance of LAT. This again seemed conducive with the establishment of latency as seen in the previous experiment and in the literature (Kobayashi, Kim et al., 2012; Pourchet et al., 2017). However, by 14dpi, a reduction in LAT expression was observed that coincided with an increase in lytic gene expression. This reduced LAT expression at 14dpi, is not in keeping the pattern of gene transcription expected with latency that have been apparent up until this point. However the LAT expression observed at 14dpi is still considerably higher than at 7dpi or than any of lytic genes tested.

In effect, this increase in lytic gene expression and decrease in LAT expression could represent a signal consistent with reactivation from latency. However, as this virus lacks functional VP16, ICP0 and ICP4, it is not capable of full reactivation. It is possible that the appearance of lytic transcripts at 14dpi in this experiment could be evidence of the initiation of the first phase of reactivation that is VP16-independent, but the infection would not be able to progress further than this to the second phase of reactivation, as this relies on functional VP16 (Kim et al., 2012). Nevertheless, viral copy number was assessed to examine whether latent genomes were retained over time in this system an whether there was considerably cytotoxicity in this system and there was some loss of viral DNA per cell over time observed. This loss in viral DNA coincides with this decrease in LAT expression and increase in lytic gene transcription. This might suggest that there is some cytotoxicity occurring as a result of this first phase reactivation despite de-repressed genomes being unable to progress to complete reactivation. It is possible that multiple media changes could have triggered neuronal stress and led to both the cell loss and the initiation of first-phase reactivation observed. Regardless, it is important to examine the system for lytic and latent transcription and consider the timepoints for the establishment of latency carefully.

Transcription kinetics of quiescent in1382 infection at a single-cell level

RNAscope *in situ* hybridisation was used to analyse major LAT expression over time at a single-cell level. Additionally, the lacZ reporter function of in1382 was utilised to observe how viral transcription changed over time.

LacZ expression driven by CMV major IE promotor acted as a surrogate marker for transcriptionally active HSV-1 infection and showed that roughly all the cells could be infected in SH-SY5Y culture. Additionally, the results suggest that transcriptional activity of the virus decreases after 1dpi. This agrees with data from the first PCR of lytic gene

expression showing a reduction in lytic gene activity. This decline in lytic viral transcript abundance happens as LAT expression increases, this might be expected given that the latency associated ncRNAs have been shown to repress lytic gene expression (Mador et al., 1998; Umbach et al., 2008)(chapter 1.4.1, section 6.2.4.3).

RNAscope was performed alongside both the LacZ experiment (figures 2.6A and 2.6B) and for the later time points of the qRT-PCR experiments (figures 2.5C and 2.6C). Both experiments revealed 11dpi to show the highest frequency of LAT expressing neurons with 28 and 59% LAT-positive cells respectively, however the percentage of cells that express LAT clearly varies between experiments (figure 2.6C). Nonetheless, both experiments agree with the qRT-PCR data showing 11dpi to have the peak LAT expression (figure 2.5C). In the parallel qRT-PCR and RNAscope experiment (figure 2.5C), the trend of LAT expression increasing from 7 to 11dpi is consistent, however the difference observed between the 2 time points is greater in the qRT-PCR experiment for LAT expression where LAT transcription increases 11-fold from 777 to 8498 copies per 1000 beta actin copies than the quantification of LAT positive cells as assessed by RNAscope which increases 1.5-fold from 40% to 59% LAT-positive cells (figure 2.6C). This highlights the importance and potential differences in how much LAT is expressed in each cell versus how many LAT-positive cells there are. This variation is highlighted when assessing the RNAscope data at high magnification. From these data, it's clear that there are different levels of LAT expression from cell to cell, and although expression in each cell may change with time, it is important to note this heterogeneity when utilising this system to assess the role of LAT. Additionally, from the RNAscope experiments and parallel lacZ experiment, it is worth noting that the percentage of LAT expressing cells, as seen by RNAscope, at most reaches an average of 59%, yet from the lacZ data it appears as though all the cells are infected at this MOI. Taken together, these results indicate that not all the infected cells are expressing LAT. This reinforces the evidence for heterogeneity within the system in terms of LAT expression. Although this data could alternatively indicate that not all the cells are latently infected, it is consistent with previous evidence demonstrating that during HSV-1 latency, only a proportion of infected neurons express LAT at any given time (Chen et al., 2002; Edwards & Bloom, 2019; Mehta et al., 1995).

The advantages and disadvantages of the in1382 latency model

Despite some biological variation within the system and the potential to initiate the earliest stages of reactivation by 14dpi, quiescent infection of neuronal SH-SY5Y

cells with replication-defective HSV-1 is a more amenable system to study the effect of HSV ncRNAs upon neurons compared to ACV and IFNα inhibition of WT HSV-1. The aim of this study moving forward is to test the effect of the latency associated ncRNAs on the neuronal transcriptome. The advantage of using a replication-defective virus is that the neuronal transcriptome won't be affected by use of inhibitors. In addition, the system has been well characterised by a variety of methods to show that until 11dpi major LAT expression and the number of cells expressing LAT increases while lytic gene expression is comparatively low.

In summary, and accepting the caveats discussed above, with methods in place to monitor variation, this system should prove a useful model to investigate aspects of HSV-1 latency. For example, the timing and the MOI used moving forward would need to be carefully considered; while 11dpi was clearly the peak in terms of LAT expression, it was considered pertinent to culture past this point, to give enough time to allow for any changes to the neuronal transcriptome as a result of LAT to manifest, following turnover of existing cellular mRNA. In doing this, there is the risk of lytic genes eliciting a response in the neurons too, so this would have to be considered and monitored. Similarly, while the higher MOI in1382 infection of the SH-SY5Y cells produced a greater quantity of major LAT, and therefore also likely the latency-associated miRNAs processed from the primary LAT transcript, some lytic transcripts production was observed at this MOI.

Notwithstanding a potential utility of this in1382 model of latency, considering some of the qualifications, a more direct reductionist approach was explored next to examine the effects of the latency-associated ncRNAs. In addition, as discussed in the next chapter, the approach taken should also allow discrimination of the effects of LAT and the effects of the miRNAs.

Chapter 3: Characterising vector delivery of the HSV-1 latency-associated ncRNAs to differentiated SH-SY5Y cells

3.1 Introduction

In the previous chapter of this thesis, a model of HSV-1 latency was established. It was found that by differentiating SH-SY5Y cells into neuron-like cells and infecting with replication-defective HSV-1 (in1382), a quiescent infection is established whereby LAT is expressed. This will enable the examination of potential roles of the latency-associated ncRNAs in neurons during latent viral infection. Nevertheless, it would also be useful to explore the impact of major LAT or the latency-associated miRNAs separately so that the effects observed with replication-defective virus can be directly attributed to either major LAT or the latency-associated miRNAs.

As has already been described in chapter 1.4.1, gene expression is greatly restricted during HSV-1 latency and the only abundantly transcribed RNA is the latency-associated transcript (LAT). This primary 8.3 kb primary LAT transcript gets processed into the major LAT introns, which are the most abundant viral product during latency. The 2.0 kb major LAT intron is spliced from the primary LAT transcript during acute and latent HSV-1 infection in neuronal and non-neuronal cells, while the 1.5 kb major LAT intron is a twintron entirely encoded within the 2kb intron and is only produced in neuronal cells during latency (Brinkman et al., 2013). The remaining LAT exon gets further processed into 6 microRNAs (miRNAs) (Umbach et al. 2009). Several of these miRNAs, along with mir-H6 (encoded in the LAT promotor region) are found to be expressed during latency (Umbach et al. 2009). The 5 HSV-1 miRNAs; miR-H2, miR-H3, miR-H4, miR-H6 and miR-H7 were shown to be among the most abundant in human trigeminal ganglia tissue during HSV-1 latency (Umbach et al. 2009).

The role of LAT has been widely examined, and certain functions have been revealed such as its role in establishing and maintaining latency, or the first 1.5kb of LAT aiding protection of cells against apoptosis as described in chapter 1.4.1 (Perng et al., 2000). However, the function of the major LAT introns specifically, which are unusual in terms of their stability and abundance during latency, remain unclear. There has been some work to suggest that the 2.0kb major LAT intron induces heat shock proteins during induced cold-shock, and therefore maintains cell survival (Atanasiu et al., 2006) but the mechanism by which this occurs or any role outside of counteracting cold-shock, needs further elucidation.

The discovery that LAT also encodes miRNAs adds to the complexity of the LAT region and increases the potential for further roles for these ncRNAs during HSV-1 latency. As previously described (chapter 1.4.1), miRNAs are small non-coding RNAs around 22 nucleotides long and are processed from stem loop structures by two cellular

proteins; Drosha and Dicer, which are RNAse III endonucleases (Gosline et al., 2016; O'Brien et al., 2018). MiRNAs have been shown to exert post-transcriptional downregulation of specific mRNA/protein targets. Some of the HSV-1 latency-associated ncRNAs have been shown to function by this mechanism to down-regulate viral lytic genes. For example, miR-H2-3p is transcribed antisense to the key viral immediate early gene – ICP0, reducing the ICP0 protein levels (Umbach et al. 2008). Although several of the HSV-1 miRNA gene targets have been identified (chapter 1.4.1), their effect on human neurons, along with LAT, are still yet to be fully characterised. These miRNAs, together with major LAT, comprise the latency-associated ncRNAs. This chapter explores a reductionist approach of delivering either major LAT or 5 HSV-1 latency-associated miRNAs to differentiated SH-SY5Y cells to study the individual effects of these latency-associated ncRNAs on human neurons.

To deliver the latency-associated ncRNAs, the use of viral vectors was examined. For the viral vectors to be able to deliver the latency-associated ncRNAs to differentiated SH-SY5Y cells, the vectors need to be able to transduce non-dividing neuronal cells. Adeno-associated viruses (AAVs) and Lentiviruses both have uses as delivery vectors into a range of cells, including fully differentiated neurons (Haery et al., 2019; Naso et al., 2017).

AAVs are 20nm, non-pathogenic, non-enveloped single-stranded DNA viruses that can be engineered to deliver DNA to target cells and as such have had many uses as vectors (Hammond et al., 2017; Naso et al., 2017). There are more than 100 different serotypes of AAV with 12 that have been extensively studied and that have known cellular tropisms (Hammond et al., 2017). Half of these have been shown to effectively transduce neurons; AAV1, 2, 5, 8, 9 and 10 (Haery et al., 2019). In this chapter serotype 2 AAV (AAV2) was tested, which was one of the first serotypes to be identified and the first serotype to be cloned into a bacterial plasmid and as such it is one of the most well characterised AAV serotypes. AAV2 has been shown to efficiently transduce cells of the nervous system including neurons (Haery et al., 2019; Hammond et al., 2017). For example, Kaplitt *et al.* showed detection of tyrosine hydroxylase (TH) in striatal neurons of rats following delivery by an AAV2-based vector (Kaplitt et al., 1994). The effective transduction of neurons makes AAV2 a relevant choice to transduce neuron-like cells such as differentiated SH-SY5Y cells.

Lentiviruses are positive single-strand RNA viruses and part of the Retroviridae family. As such they contain a reverse transcriptase, which allows their RNA genome to copied into DNA, which is then integrated into the host cell genome, making lentiviruses a useful vector system. Another feature of lentiviruses, as a complex retrovirus, that

makes it a useful viral vector, is that lentiviruses can transduce nondividing cells in addition to dividing cells. Unlike simple retroviruses that require breakdown of the nuclear membrane – as occurs during mitosis – to enter the nucleus, lentiviruses are transported into the nucleus as part of a pre-integration complex (O'Keefe, 2013). This function of lentiviruses makes them useful for the transduction of neurons. Accordingly, lentiviruses have been shown to successfully transduce neurons previously. Blömer *et al*. demonstrated that lentiviruses efficiently deliver beta-galactosidase to cells of the rat brain that stain positively for neuron specific marker; NeuN (neuronal nuclei) (Blömer et al. 1997). Therefore, lentiviruses could be useful for this study where delivery of transgenes to non-dividing, neuron-like cells is also a requirement.

Clearly, both AAVs and lentiviruses have been successfully used to deliver DNA to neuronal cells previously (Blömer et al. 1997; Haery et al. 2019; Hammond et al. 2017), this chapter tests their effectiveness at encoding and delivering the HSV-1 latency-associated ncRNAs to neuronally differentiated SH-SY5Y cells with a view to testing the effects of these abundant latency products – the 1.5 and 2kb major LAT introns and the 5 most abundant latency-associated miRNAs – in human neuronal cells.

3.2 Results

3.2.1 Lentivirus vectors transduce differentiated SH-SY5Y cells more efficiently than Adeno-associated virus (type 2) vectors

The first test in examining how best to deliver the latency-associated ncRNAs to differentiated SH-SY5Y cells was to compare the efficiency with which AAV or lentivirus vectors transduce differentiated SH-SY5Y cells.

To compare the efficiency of transduction of the AAV and lentivirus vectors, enhanced green fluorescent protein (eGFP) producing viral vectors were utilised; a serotype 2 adeno-associated viral vector expressing eGFP (AAV2-GFP), generated by Vector biolabs (figure 3.1A) and an eGFP lentivirus produced from Addgene plasmid 12252 (pRRLSIN.cPPT.PGK-GFP.WPRE)(figure 3.1B) (see material and methods chapter 6.2.2.5 for lentivirus production). The main genomic features of each viral vector are highlighted schematically in figures 3.1A and 3.1B. There are key differences, such as eGFP is driven by the CMV promoter in the AAV2 vector whereas the human phosphoglycerate kinase 1 (hPGK) promoter drives eGFP expression in the lentivirus vector. Utilising eGFP-expressing viral vectors such as these enables visualisation of transduced cells by expression of the transgene.

To be able to develop and assess the efficiency of the different viral vectors, it was first necessary to be able accurately quantitate and titrate the numbers of transducing viral particles. AAV2-GFP was supplied titrated by qRT-PCR (to the CMV promoter), with genome copy number per ml, by Vector biolabs. The eGFP lentivirus vector was similarly titrated by qRT-PCR on the virus stock. A 10-fold serial dilution was performed on the eGFP lentivirus stock and then qRT-PCR performed using primers designed to hybridise to the woodchuck hepatitis virus post transcriptional regulatory element (WPRE) sequence in the lentivirus vector (figure 3.1B). The virus copy number in the original dilution was calculated by taking an average across the WPRE copy number found by qRT-PCR multiplied by each dilution factor (figure 3.1C). The volume of lentivirus used in the qRT-PCRs was then used to calculate a genome copy number per µl and then per ml, which in turn, with the number of differentiated SH-SY5Y cells per well, was used to calculate genome copies per cell (gpc).

Differentiated SH-SY5Y cells were transduced with eGFP expressing lentiviruses or AAV2 at equivalent genome copy numbers of between 20 and 1000. The transduced cells were observed for eGFP expression up to 18 days post transduction (dpt) and imaged at 2, 4 and 18 dpt. From the results it was clear that the eGFP lentivirus vector was more effective at transducing the differentiated SH-SY5Y cells than the AAV2 eGFP vector (figure 3.2A). Differentiated SH-SY5Y cells transduced with AAV2-eGFP at 500 gpc showed no eGFP expression until 4 dpt and even by 18dpi there was only minimal eGFP expression both in terms of the number of cells and intensity of the eGFP signal. The 1000 gpc AAV2-eGFP infection similarly showed no eGFP expression at either 2 or 4 dpt. However, by 18 dpt there was a higher frequency of cells expressing eGFP than at 500 gpc but this was still only a proportion of the cells and those that were expressing were doing so with very low intensity eGFP signal (figure 3.2A). Conversely, Lentivirus-eGFP transduction of differentiated SH-SY5Y cells showed eGFP expression from 2dpi, faintly at 20 gpc and highly expressed at 1000 gpc. By 4 dpt most if not all the SH-SY5Y cells transduced with lentivirus-eGFP at 1000 gpc virus were expressing eGFP, the levels of which continued to increase by 18 dpt (figures 3.2A and 3.2B). By 18 dpt of the eGFP lentivirus at 1000 gpc, most of the SH-SY5Y cells were expressing eGFP and with strong signal.

The lentiviral vector with hPGK promoter was more efficient than the AAV-2 with CMV promoter at transducing the differentiated SH-SY5Y cells. As already mentioned, it appears as though the vast majority and potentially all the cells expressed eGFP when transduced with 1000 gpc of lentivirus-eGFP (figure 3.2B). This provides valuable information about how much transduction and expression of the transgene should be

Figure 3.1 The genomic structures of the eGFP AAV and eGFP lentivirus plasmids and lentivirus titration method

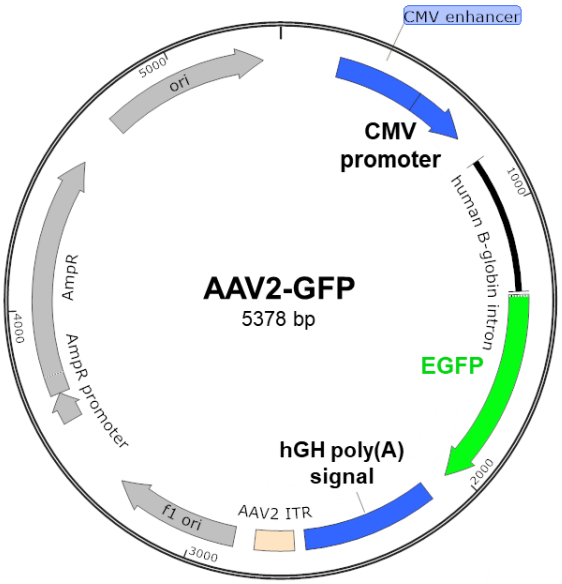
(A) A schematic diagram of the eGFP expressing adeno-associated virus (AAV) plasmid; AAV2-GFP (vector biolabs)

(B) A schematic diagram of the eGFP expressing lentiviral plasmid; pRRLSIN.cPPT.PGK-GFP.WPRE (Addgene)

(C) qRT-PCR was performed in duplicate on 10-fold serially diluted eGFP lentivirus stock, using primers directed to WPRE, the average WPRE across the duplicates for each dilution are displayed on the graph.

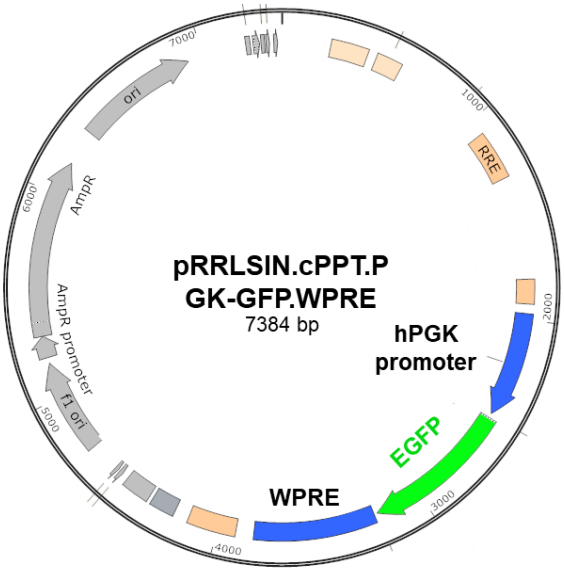
A

Adeno-associated virus (AAV)



B

Lentivirus (LV)



C

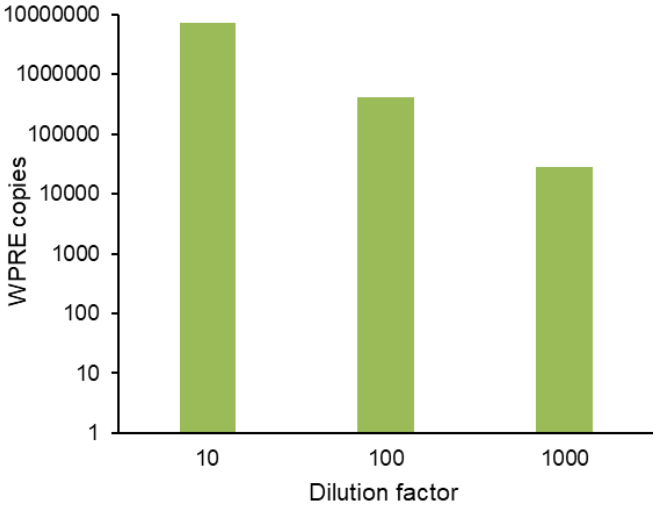


Figure 3.1 The broad genomic structures of the eGFP-AAV2 and eGFP lentivirus plasmids and genome per cell (gpc) titration method of the lentiviral vectors.

Figure 3.2 Lentivirus plasmids are more effective at transducing differentiated SH-SY5Y cells than Adeno-associated virus serotype 2 plasmids

A) Differentiated SH-SY5Y cells were transduced with either eGFP expressing Adeno-associated virus (AAV) or eGFP lentivirus (LV) at 20, 500 or 1000 genomes per cell (gpc) in duplicate. The cells were then imaged by fluorescence microscopy at 2, 4 and 18 days post transduction (dpt) as indicated. Scale bars represent 50μm.

B) Expanded image of the differentiated SH-SY5Y cells transduced with eGFP lentivirus at 1000 gpc taken at 4 dpt, shown in figure A. Scale bars represent 50μm.

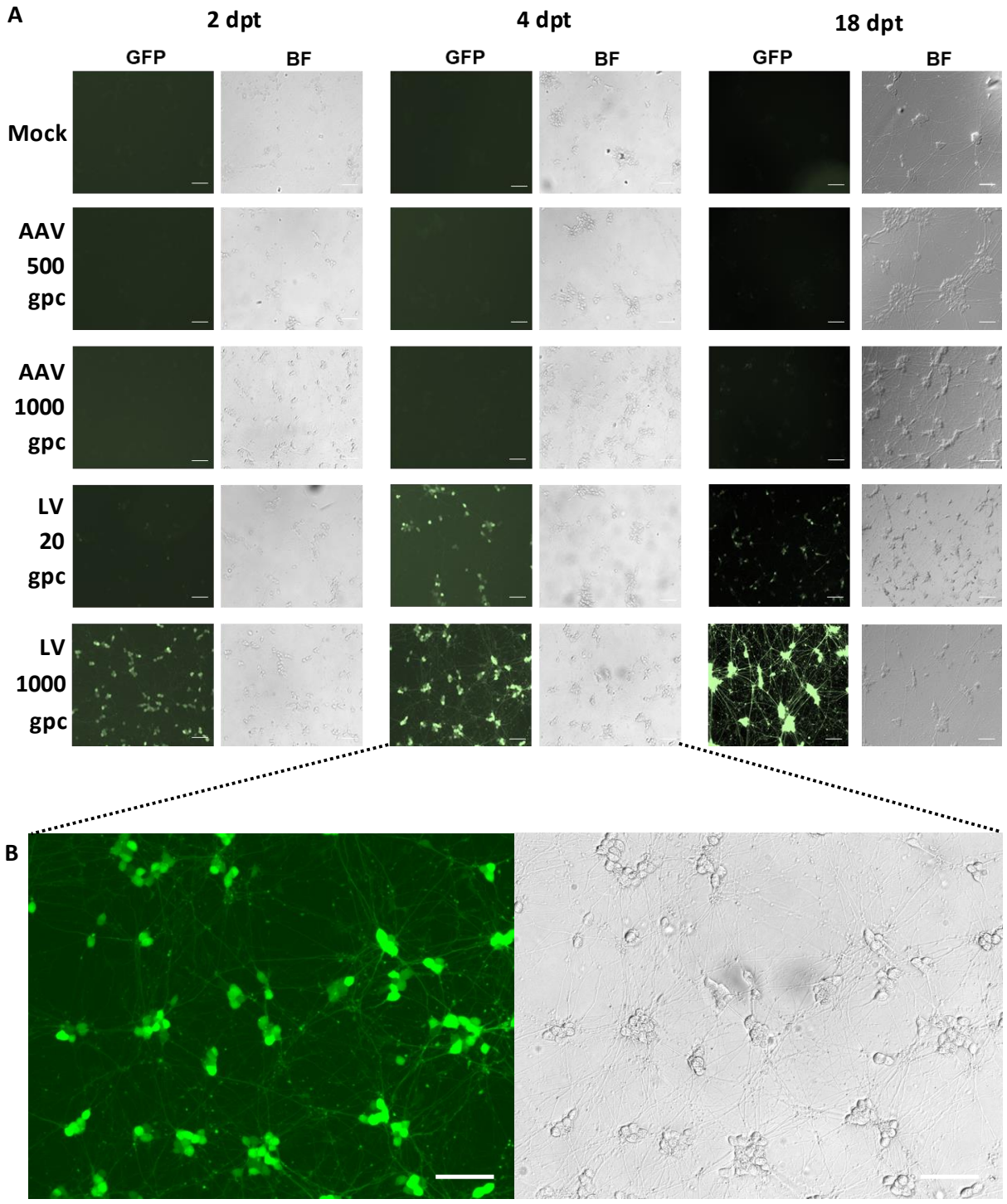


Figure 3.2 Lentivirus plasmids are more effective at transducing differentiated SH-SY5Y cells than Adeno-associated virus serotype 2 plasmids and do so with a high efficiency

expected at the gpcs used of this lentivirus, whereby 20 gpc showed a significant proportion of cells expressing eGFP and 1000 gpc led to strong eGFP expression in most of the cells.

3.2.2 Constructing lentiviral vectors encoding the latency-associated ncRNAs

Having developed an accurate method to compare delivery mechanisms, it was demonstrated that the lentivirus vectors were more effective than AAV-2 based vectors for efficient transduction of differentiated SH-SY5Y cells. Given this, the lentiviral vector system was chosen for the delivery of the latency-associated ncRNAs to differentiated SH-SY5Y cells.

The eGFP lentivirus plasmid (Addgene plasmid 12252) was the lentivirus backbone plasmid used to generate all the lentiviruses in this project. This plasmid is a 3rd generation lentiviral vector, which means that it has a truncated 5'LTR as a safety element that restricts the possibility of recombining to generate replication-competent lentivirus (Dull et al., 1998; Morgan, 2014). The eGFP lentivirus plasmid encodes eGFP and a 3' WPRE driven by the hPGK promoter (figures 3.1B and 3.3A). The WPRE enhances mRNA stability, while the hPGK promoter appears to be constitutively expressed in the differentiated SH-SY5Y cells. This is evident by most of the differentiated SH-SY5Y cells expressing eGFP from this promoter when transduced with the eGFP lentivirus (figure 3.2). There is not only a high frequency of cells expressing eGFP when driven by the hPGK promoter, but this expression is sustained over time (figure 3.2A).

To generate lentiviruses encoding the latency-associated ncRNAs, restriction digests and ligations were utilised to replace eGFP with DNA encoding each ncRNA, as shown schematically (figure 3.3A). Dr Michael Nicoll, division of Virology at NIBSC, constructed 2 lentiviruses using this method: the LAT intron lentivirus expressing the first 3069 bp (~3.1 kb) of the LAT transcript, which includes the potential for splicing of major LAT and the non-splicing LAT (NSLAT) lentivirus, that also expressed the first 3.1 kb of the LAT transcript except with ten nucleotide deletions that remove the splice donor and acceptor sites required for splicing of the major LAT intron (Alvira et al., 1999)(figure 3.3A). The lentiviruses were designed to express the first 3.1kb of the LAT as this contains the entire major LAT intron and the surrounding splice acceptor and donor sites, therefore major LAT should be spliced in the LAT intron lentivirus and the role of major LAT can be addressed. The NSLAT lentivirus acts as a negative control, whereby the major LAT sequence is present, but the intron is not spliced. Therefore, the effects of each lentivirus can be compared to assess the role of the major LAT intron. Additionally, including a

Figure 3.3 Constructing the LAT intron and NSLAT lentivirus plasmids

(A) A schematic diagram of the eGFP lentivirus (addgene plasmid 12252) and its use as the lentivirus backbone plasmid to generate each of the LAT lentiviral plasmids. The key elements that will enable expression of the transgene (for example the hPGK promoter and WPRE) and construction of the LAT lentiviruses (such as the BamHI and NotI restriction sites) are highlighted. The region of the LAT transcript to be encoded following construction of the LAT lentivirus plasmids is shown in red, with splice acceptor and donor deletions annotated for the NSLAT lentivirus.

(B) A schematic diagram of the final LAT lentivirus plasmid, following construction as indicated in (A)

(C) A schematic diagram of the final NSLAT lentivirus plasmid, following construction as indicated in (A)

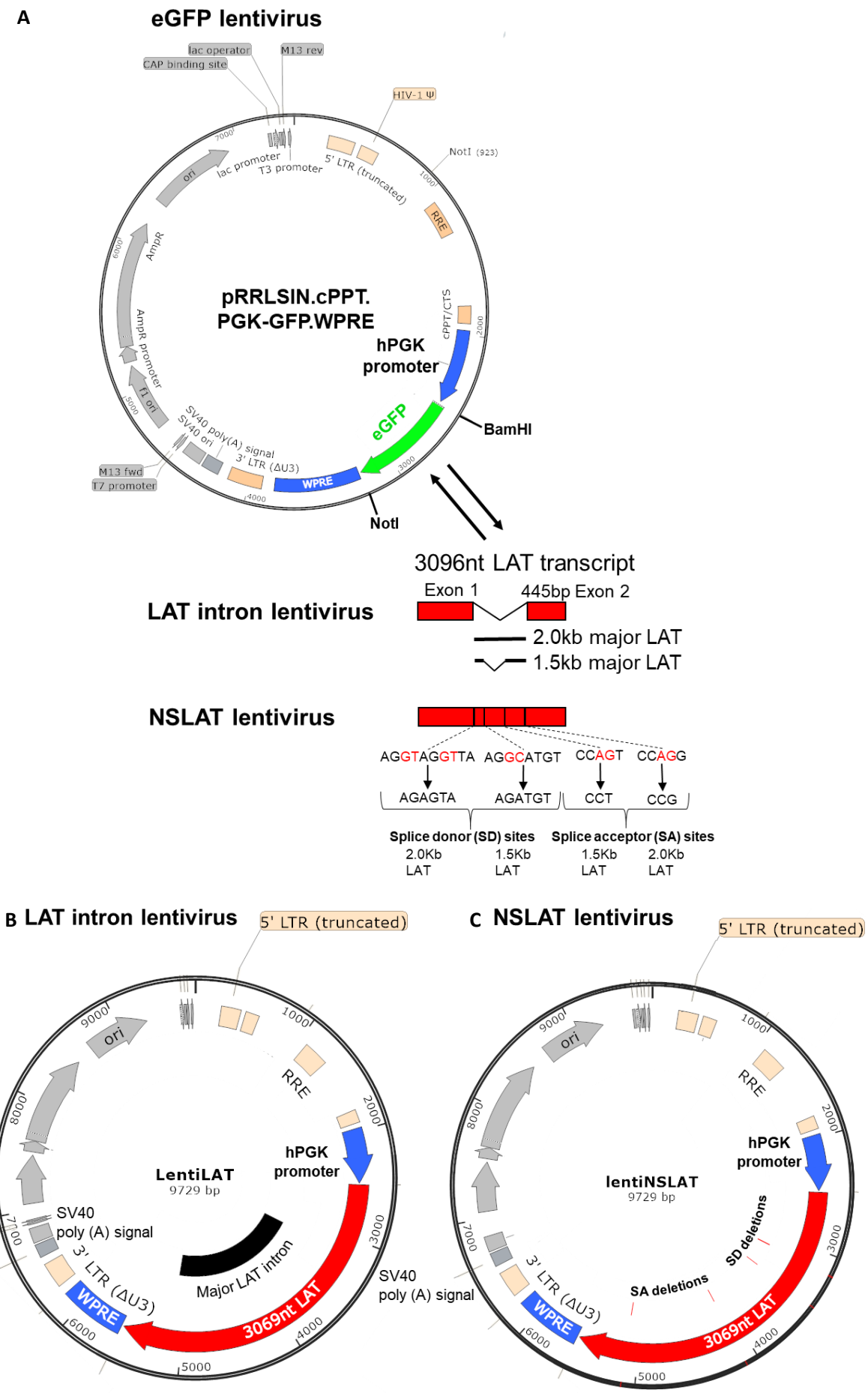


Figure 3.3 Constructing the LAT intron and NSLAT lentivirus plasmids

larger insert size, for example if the entire LAT transcript was encoded, has been shown to lower the viral titres (Kumar et al., 2001). Therefore, the first 3.1kb is sufficient to assess the role of major LAT whilst not compromising on viral titre. Details on the construction of the LAT lentiviruses is provided in materials and methods 6.2.3.2.

The miRNA lentivirus was produced by similar methods except this digestion and ligation was performed in 3 parts. Two synthetic constructs comprised of HSV-1 miRNA stem-loops and flanking HSV-1 sequence were designed and produced (GENEWIZ); HSV-1-miR-H2 (nt 121777-121957); HSV-1-miR-H3 and HSV-1-miR-H4 (nt 125667-126018) (miR-H2-4) as well as HSV-1-miR-H6 (nt 118463-118252); and HSV-1-miR-H7 (nt 123352-123553) (miR-H6-7). These 2 constructs – miR-H2-4 and miR-H6-7, were digested alongside the eGFP lentivirus plasmid (pRRLSIN.cPPT.PGK-GFP.WPRE) and ligated together to generate the miRNA lentivirus (methods 6.2.3.2)(figures 3.4A, 3.4B).

These five miRNAs were chosen as they have been shown to be abundantly produced during HSV-1 latency (Umbach et al. 2009). The tertiary structure of miRNAs are important in their function and processing (Contrant et al., 2014; Cui et al., 2006), which is why the miRNA lentivirus was designed to include the miRNA stem-loops and flanking regions so that their tertiary structure wouldn't be effected. Additionally, including linker regions between some of the miRNAs also means that should further investigation of individual miRNAs become useful, these could be targeted to separate the miRNAs and generate lentiviruses to express the individual miRNAs. The miRNA sequences encoded in the miRNA lentivirus are outlined in methods chapter 6.1.5.

Table 3.1 highlights the sequences that encode each miRNA, as well as the sequence of the combined miRNA constructs and flanking HSV-1 sequence encoded by the miRNA lentivirus following digestion and ligation. As part of the cloning process (see methods 6.2.3 for details) this miRNA plasmid was transformed into bacteria and 7 bacterial colonies were picked. The miRNA lentivirus plasmid grown in these colonies was then tested by diagnostic restriction digest with BamHI enzyme to check that the product was as expected in terms of number of fragments produced and their size. Given the BamHI sites on the miRNA lentivirus (figure 3.4B), digesting with this restriction enzyme should produce 2 DNA fragments; 6748 base pairs (bp) and 940 bp long. From the DNA agarose gel, it is evident that following digestion of the plasmid in each of the 7 selected bacterial colonies, the bands align with the predicted sizes for the miRNA BamH1 digest and there are 2 bands, where as there is only 1 band seen in the undigested columns of the gel (figure 3.4C), indicating successful construction of the miRNA lentiviral plasmid.

Figure 3.4 Constructing the miRNA lentivirus plasmid

(A) A schematic diagram of the eGFP lentivirus (addgene plasmid 12252)(viewed using SnapGene® software (GSL Biotech; available at snapgene.com)) and its use as the lentivirus backbone plasmid to generate the miRNA lentiviral plasmid. The key elements that will enable expression of the transgene and construction of the miRNA lentivirus are highlighted along with the miRNAs to be encoded following construction of the miRNA lentivirus plasmid.

(B) A schematic diagram of the final MiRNA lentivirus plasmid, following construction as indicated in (A)

(C) A diagnostic DNA electrophoresis gel following digest of the constructed miRNA lentivirus plasmid with BamH1 restriction enzyme following cloning into and isolation from bacteria (see methods 6.2.3 for further details). Expected band sizes following BamHI digest are 6748 bp and 940 bp. U lanes show undigested samples while D lanes represent digested samples. A DNA ladder is used and shown with the values in base pairs (bp) indicated.

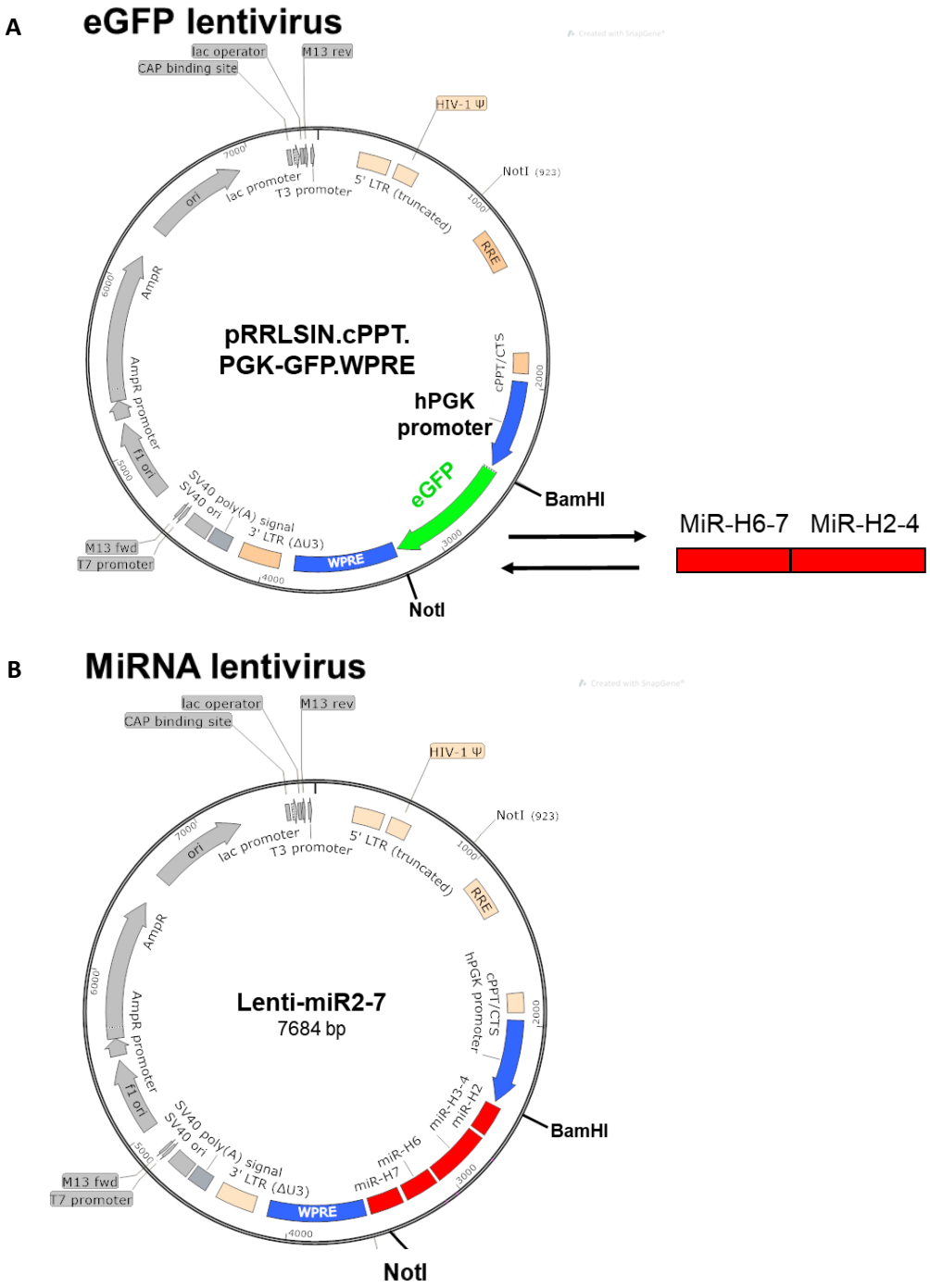


Figure 3.4 Constructing the miRNA lentivirus plasmid

Table 3.1 Table of the miRNA sequences

The DNA coding sequences of the stem-loop of each of the 5 latency-associated miRNAs encoded by the miRNA lentivirus are outlined individually and highlighted in red in the context of the surrounding DNA as they are encoded in the miRNA lentivirus.

MicroRNA	DNA coding sequence
MiR-H2	GCCACCGTCGCACGCGCCCGGCACAGACTCTGTTCTT-GGTTCGCGGCCTGAGCCAGGGACGAGTGCGACTGGGGC
MiR-H3	CCGCGGGCGCGCTCCTGACCGCGGGTTCCGAGTTGGGCGTGGAGGTTACCTGG-GACTGTGCGGTTGGGACGGCGCCCGTGG
MiR-H4	GCCGGGGTGGTAGAGTTTGACAGGCAAGCATGTGCGTGCAGAGGCGAGTAGTGCTT-GCCTGTCTAACTCGCTAGTCTCGGC
MiR-H6	CGGGGGGCCGGAGGGTGGAAGGCAGGGGGGTGTAGGATGGGTATCAGGACTTCCAC-TTCCCGTCCTTCCATCCCCCGTTCCCCTCG
MiR-H7	GAAGAGGGGGGAGAAAGGGGTCTGCAACCAAAGGTGGTCTGGGTCCGTCCTTT-GGATCCCGACCCCTCTTCTTCCCTCTTCTC
MiRNA lentivirus MicroRNA coding sequence	CGCGGGAGGGCCCGTGCCCACCCTCCACGCCCGGCCCCCCCGAGCCGCGCGCCAC-CGTCGCACGCGCCCGGCACAGACTCTGTTCTTGGTTCGCGGCCTGAGCCAGGGACGAGTGCGACTGGGGCACACGGCGCGCGTCCGCGGGGCGGGCGGCCGGCTCCGCCCCGGGGGCCGGGGCGCGGGGGCTTAATTAAGTTTAAACCTGGCGCAGCAGCAGCGAACAAGAAGGCGGGGGCCCACCGGCGGGGGGCGGCGGCGGGGCAGCCGCGGGCGCGCTCCTGACCGCGGGTTCCGAGTTGGGCGTGGAGGTTACCTGGGACTGTGCGGTTGGGACGGCGCCCGTGGGCCCGGGCGGCCGGGGGCGGCGGGGGCCGCGATGGCGGCGGCGGCGGGCCATGGAGACAGAGAGCGTGCCGGGGGTGGTAGAGTTTGACAGGCAAGCATGTGCGTGCAGAGGCGAGTAGTGCTTGCCTGTCTAACTCGCTAGTCTCGGCCGCGGGGGGCCCGGGCTGCCCGCCGCCACCGCTTTAAAGGGCCGCGCGCGACCCCCGGGGGGTTTAAACGTCGACGCGATCGCTGCGGCGGCTGCGGCAGCCGCCGTGTTCGGTCTCGGTAGCCGGCCGGCGGGTGGACTCGCGGGGGGCCGGAGGGTGGAAGGCAGGGGGGTGTAGGATGGGTATCAGGACTTCCACTTCCCGTCCTTCCATCCCCCGTTCCCCTCGGTTGTTCCTCGCCTCCCCCAACACCCCGCCGCTTTCCGTTGGGGTTGTTATTGTGTTGTCGGGAGCGATCGCATTTAAATCACGAGTTCCCCCTCTCCCCCCCCCGCCTCACAAAGTCCTGCCCCCCTGCTGGCCTCGGAAGAGGGGGGAGAAAGGGGTCTGCAACCAAAGGTGGTCTGGGTCCGTCCTTTGGATCCCGACCCCTCTTCTTCCCTCTTCTCCCGCCCTCCAGACGCACCGGAGTCGGGGGTCCCACGGCGTCCCCCAAATATGGCGGGCGGC

Table 3.1: MicroRNA sequences for MiR lentivirus

The resulting 4 plasmids should produce either eGFP, first 3.1kb of LAT with splicing of the major LAT intron, the first 3.1kb of LAT without splicing of the major LAT intron or 5 of the HSV-1 latency-associated miRNAs: miR-H2, miR-H3, miR-H4, miR-H6 and miR-H7 (3.1B, 3.3B and 3.3C and 3.4B).

Once constructed, the lentivirus plasmids were sequenced by Dr Martin Fritzsche, analytical and biological sciences (ABS), NIBSC and the bioinformatics analysis and graphical representation of this sequencing (figure 3.5) was carried out by Dr Mark Preston, ABS, NIBSC. Short reads were aligned to the references sequence (Addgene plasmid 12252 for the eGFP lentivirus or this reference with the eGFP sequence exchanged for the relevant HSV-1 sequence taken from the HSV-1 strain 17 syn+ reference genome NC_001806.2 for the three ncRNA lentiviruses). 100% coverage was achieved for all virus plasmids (figure 3.5A). However, there was a reduction in read depth in the HSV-1 latency-associated ncRNA coding regions, especially in the LAT region of the LAT intron and NSLAT lentiviral plasmids where coverage at its lowest dropped to approximately 56 reads (log10(1.75)) but this is perhaps unsurprising given the high GC content of this region of the HSV genome, which may make it difficult to sequence. A threshold of 100 reads was applied to the data (figure 3.5A) and the number of nucleotides with a read-depth below this counted. The LAT intron lentivirus was found to have the most low-coverage nucleotides with 78 falling below 100 reads. This accounts for 0.8% of the sequence. Therefore, it seems that most of the lentivirus nucleotides aligned to their respective reference sequences.

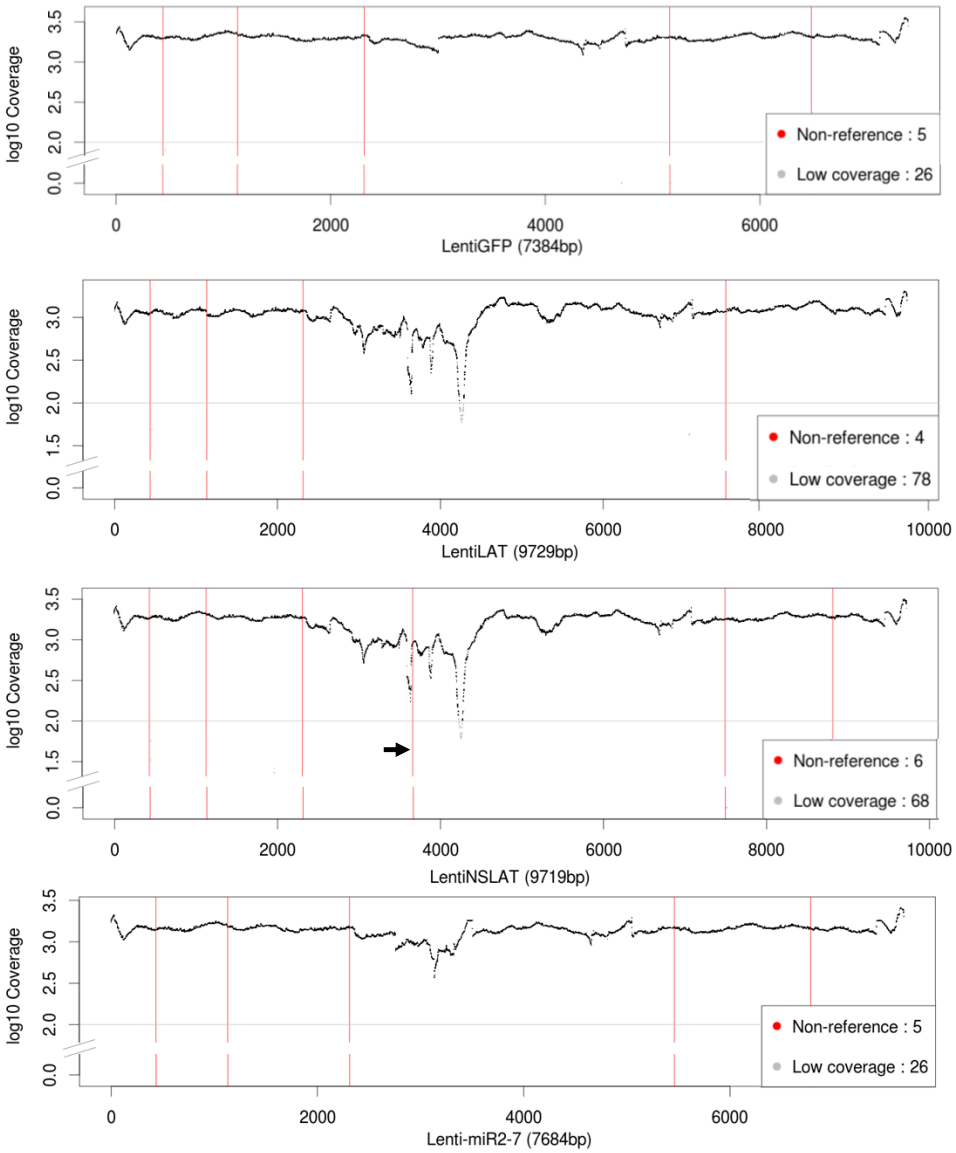
In total there were 6 different single nucleotide polymorphisms (SNPs) found in the lentiviral plasmid sequences, as compared to the reference sequence (addgene plasmid 12252). However, 5 out of 6 of these were conserved in the eGFP lentivirus that acted as the backbone for the construction of these plasmids, therefore, it doesn't appear as though considerable mutations have been introduced when constructing these plasmids (figures 3.5A and 3.5B). The only additional SNP not found in the eGFP lentivirus plasmid was the 4th SNP in the NSLAT lentivirus plasmid (C to A at position 3666)(figures 3.5A and 3.5B) which occurs in the LAT region, which doesn't encode a protein. It occurs in the second exon of the LAT transcript beyond any splice acceptor and donor sites; hence, this SNP should have no impact on the block to major LAT splicing in this virus. Although the SNP found in NSLAT and the low-read depth regions were noted and registered, given the SNP position and generally successful coverage shown, the sequencing data suggested that it was unnecessary to remake the plasmid. Accordingly, the next section describes analysis of the production of ncRNAs from these plasmids.

Figure 3.5 Sequencing the lentivirus plasmids

(A) The graphs (made by Dr Mark Preston, ABS, NIBSC using R software (The R Core Team, 2017)) plot the sequencing coverage across the length of each lentiviral plasmid compared to the relevant reference sequence (addgene plasmid 12252 with the appropriate fragments of the HSV-1 genomes for the ncRNA lentiviruses from the HSV-1 strain 17 syn+ reference genome NC_001806.2, co-ordinates outlined in materials 6.1.5)**.** The log transformed number of reads are shown on the y-axis with the sequence position along the x-axis. Single nucleotide polymorphisms (SNPs)(non-reference)(methods 6.2.7.1) are indicated by red lines. The black arrow indicates the unique SNP found in the NSLAT (mutation 4 in B). The boxes in the bottom right indicate the number of SNPs compared to the reference (non-reference) and the number of nucleotides that fall below a coverage threshold of 100 reads (log10 coverage 2.0) with this threshold indicated by a grey line and any reads that fall below this are represented by grey dots.

(B) The table outlines the details of the SNPs found in each lentivirus plasmid (shown in red in A). Indicated in each column from left to right is the number of the SNP identified in each lentivirus, the lentivirus sequenced, the position along the lentivirus sequence where the SNP lies, the nucleotide found in the reference sequence, the coverage at this point (in number of reads) and finally the relative frequency of each possible nucleotide (Adenine (A), Thymine (T), Cytosine (C) and Guanine (G)) at this position in the lentivirus is shown as a percentage of total number of reads. In other words where ref is C and pT is 99.90, the SNP found is a C to T mutation at the position stated, where, the nucleotide at this position was found to be a T in 99.9% of the reads at this position during sequencing. The unique SNP found in NSLAT is highlighted in the red box (as indicated by a black arrow in A).

A



B

	Lentivirus	pos	ref	cov	pA	pC	pG	pT
1	LentiGFP	436	C	1980	0	0.10	0.00	99.90
2	LentiGFP	1132	C	2158	0	0.00	100.00	0.00
3	LentiGFP	2313	G	2099	100	0.00	0.00	0.00
4	LentiGFP	5161	T	2042	0	100.00	0.00	0.00
5	LentiGFP	6482	C	2034	0	0.20	0.00	99.80
1	LentiLAT	436	C	1117	0	0.00	0.00	100.00
2	LentiLAT	1132	C	1056	0	0.00	100.00	0.00
3	LentiLAT	2313	G	1178	100	0.00	0.00	0.00
4	LentiLAT	7506	T	1161	0	99.70	0.00	0.30
1	LentiNSLAT	436	C	1845	0.00	0	0.00	100
2	LentiNSLAT	1132	C	2053	0.00	0	100.00	0
3	LentiNSLAT	2313	G	1848	99.90	0	0.10	0
4	LentiNSLAT	3666	C	893	100.00	0	0.00	0
5	LentiNSLAT	7496	T	1791	0.00	100	0.00	0
6	LentiNSLAT	8817	C	1857	0.00	0	0.00	100
1	Lenti-miR2-7	436	C	1390	0.00	0.00	0.00	100.00
2	Lenti-miR2-7	1132	C	1514	0.10	0.00	99.90	0.00
3	Lenti-miR2-7	2313	G	1457	99.90	0.00	0.00	0.10
4	Lenti-miR2-7	5461	T	1464	0.00	100.00	0.00	0.00
5	Lenti-miR2-7	6782	C	1408	0.00	0.10	0.00	99.90

Figure 3.5 Sequencing of the lentivirus plasmids

3.2.3 Characterising and titrating the LAT intron and NSLAT lentiviruses

The first step in the characterisation of the LAT lentiviruses was to test whether the LAT intron lentivirus splices the major LAT intron, while the NSLAT lentivirus does not. A northern blot using a probe directed to LAT was chosen as a means of addressing this question because it separates RNA according to size whilst specific RNA sequences can be targeted. The 1.5/2.0kb spliced major LAT intron should differ in size to the non-spliced major LAT sequence, still part of the original 3.1kb of the primary transcript.

In the differentiation protocol of the SH-SY5Y cells there are steps that limit how many cells can be grown and therefore how much RNA can be extracted (for example, the need to grow the cells onto extracellular matrix). Consequently, HEK293T cells were used instead as these do not have such growth limiting steps and therefore more RNA can be successfully extracted to use in a northern blot. HEK293T cells are also the cell line used to grow the lentivirus stocks in, therefore it was known that they could be transfected successfully with the lentivirus plasmids. However, the efficiency of transduction with the lentiviral vectors in the HEK293T cells, was unknown. Therefore, before examining the production and splicing of LAT in the lentiviruses by northern blot, the transduction efficiency of the eGFP lentivirus in HEK293T cells was tested, by using eGFP as a readout of transgene expression (figures 3.6B and 3.6C). Differentiated SH-SY5Y cells have been shown to be efficiently transduced by eGFP lentivirus (figures 3.2A and 3.2B) and so were used as a positive control for transduction efficiency.

Both HEK293T cells and differentiated SH-SY5Y cells were transduced with 2000 gpc of eGFP lentivirus and imaged by fluorescence microscopy 5 days post-transduction. The results show that the HEK293T cells do transduce with the eGFP lentivirus to express the transgene, as indicated by positive eGFP signal (figure 3.6B). However, in most cells this signal is considerably weaker than the eGFP signal observed in the transduced differentiated SH-SY5Y cells (figure 3.6A) and in a smaller proportion of cells (3.6A and 3.6B). Nonetheless, the HEK293T cells were known to successfully be transfected with the plasmids used to produce these lentiviruses (methods 6.2.2.5), and their ability to splice the LAT intron or not should be conserved in the lentiviral plasmid and the lentivirus. Therefore, HEK293T cells were transfected with the lentiviral plasmids (rather than transducing with the lentiviruses produced from these plasmids) for the northern blot testing of major LAT intron splicing.

HEK293T cells were transfected with 2-2.4μg of either the LAT intron, NSLAT or eGFP lentivirus plasmids. Four days post transfection, total RNA was extracted from the transfected cells. 1μg of this extracted total RNA was run on a formaldehyde gel,

alongside a molecular weight marker and then a northern blot was performed using a probe directed to the major LAT sequence.

The results show a band at 5kb (figure 3.6C), occurring from all 3 lentivirus transfections, which therefore cannot be due to LAT. It is also too large for the 3.1kb of the LAT transcript. Given its size and the abundance of ribosomal RNA (rRNA), this 5kb band is likely as a result of cross-hybridisation of the LAT probe to 28S rRNA. Nevertheless, there is a band at 1.9-2kb that only is produced in the LAT intron lentiviral plasmid transfected HEK293T cells. This approximately aligns with the size of the most abundant 2.0kb major LAT intron. The 1.5kb major LAT intron is not observed, but this is likely due to this intron being specifically only produced in neuronal cells and not in the non-neuronal HEK293T cells used here, in contrast to the 2.0kb major LAT (Brinkman et al., 2013). This result indicates that major LAT is produced and spliced following transfection of the LAT intron lentivirus plasmid and not the NSLAT lentivirus plasmid or the negative control eGFP lentivirus plasmid transfection. This is in keeping with the expected result given the designs and sequences of the lentiviruses and the previously published data showing that the mutations included in the NSLAT lentivirus block splicing of the Major LAT intron (Alvira et al., 1999).

The LAT lentivirus plasmids seem to follow the expected splicing phenotype. However, to be a useful delivery vector of LAT, the lentiviruses must drive expression of LAT in the differentiated SH-SY5Y cells. To test this, differentiated SH-SY5Y cells were transduced with lentiviruses produced from the previously described LAT and NSLAT plasmids (see chapter 6.2.2.5 for lentivirus production) and tested for LAT transcription by qRT-PCR, using primers to the major LAT sequence (figure 3.6D). Both the LAT intron and NSLAT lentiviruses encode the first 3.1kb of the primary LAT sequence. Therefore, both lentiviruses should express LAT and it should be detectable with these primers. The difference is that major LAT should be spliced in the LAT intron lentivirus transduction, whilst remaining as part of the primary transcript following NSLAT lentivirus transduction, as indicated in the northern blot.

As previously described, the eGFP lentivirus transduction (figure 3.2) gave information on the number of genomes per cell required to transduce all the SH-SY5Y cells. As all the lentiviruses developed in this chapter share the same backbone plasmid, it was expected that a genome copy number that allows efficient transduction in one of these lentiviruses should be similar for the others. The eGFP lentivirus showed significant eGFP expression with a 20 gpc transduction and strong eGFP expression in most of the cells with a 1000 gpc transduction (figure 3.2A). Therefore, a gpc between

Figure 3.6 Characterising the LAT intron and NSLAT lentiviruses

(A) Fluorescent photomicrograph of differentiated SH-SY5Y cells transduced with 2000 genomes per cell (gpc) of eGFP (green) lentivirus, alongside associated brightfield photomicrograph of the same cells. Scale bars represent 100μm.

(B) Fluorescent photomicrograph of HEK293T cells transduced with 2000 gpc of eGFP (green) lentivirus, alongside brightfield photomicrograph of the same cells. Scale bars represent 100μm.

(C) HEK293T cells were transfected with 2-2.4μg of LAT intron lentivirus plasmid, NSLAT lentivirus plasmid or eGFP lentivirus plasmid per well on a 6-well plate (2 wells per lentivirus). 4 days post transfection total RNA was extracted from 2 wells combined per lentivirus and then 1μg per lentivirus was run in a formaldehyde gel and a northern blot performed on this RNA with a probe directed to major LAT. The sizing of the bands in kilobases (kb) is taken from the sizes given by an RNA molecular weight marker (millennium markers) run on the formaldehyde gel (see methods 6.2.9 for details on the northern blot protocol). The size of the human 28S (5.0kb) and 18S (1.9kb) rRNA are indicated by white arrow heads.

(D) A schematic diagram indicating the position of the LAT qRT-PCR primers and Taqman probe along the LAT transcript, within the LAT intron lentivirus (position along LAT can also be applied to LAT sequence within NSLAT lentivirus or HSV-1). The original LAT lentivirus genome schematics were generated on SnapGene® Viewer (GSL Biotech; available at snapgene.com).

(E) Differentiated SH-SY5Y cells were transduced with LAT intron lentivirus or NSLAT lentivirus at 500 gpc or 1000 gpc across 2 experiments. 14 dpt in each experiment RNA was harvested and qRT-PCR was performed to LAT, using the primers and probe indicated in (D) and normalised to Beta actin. The results show the average from 2-3 biological repeats and 3 technical repeats ± the standard deviation where appropriate (standard deviation has been excluded for the 1000 gpc LAT intron transduction where only 2 biological repeats were used).

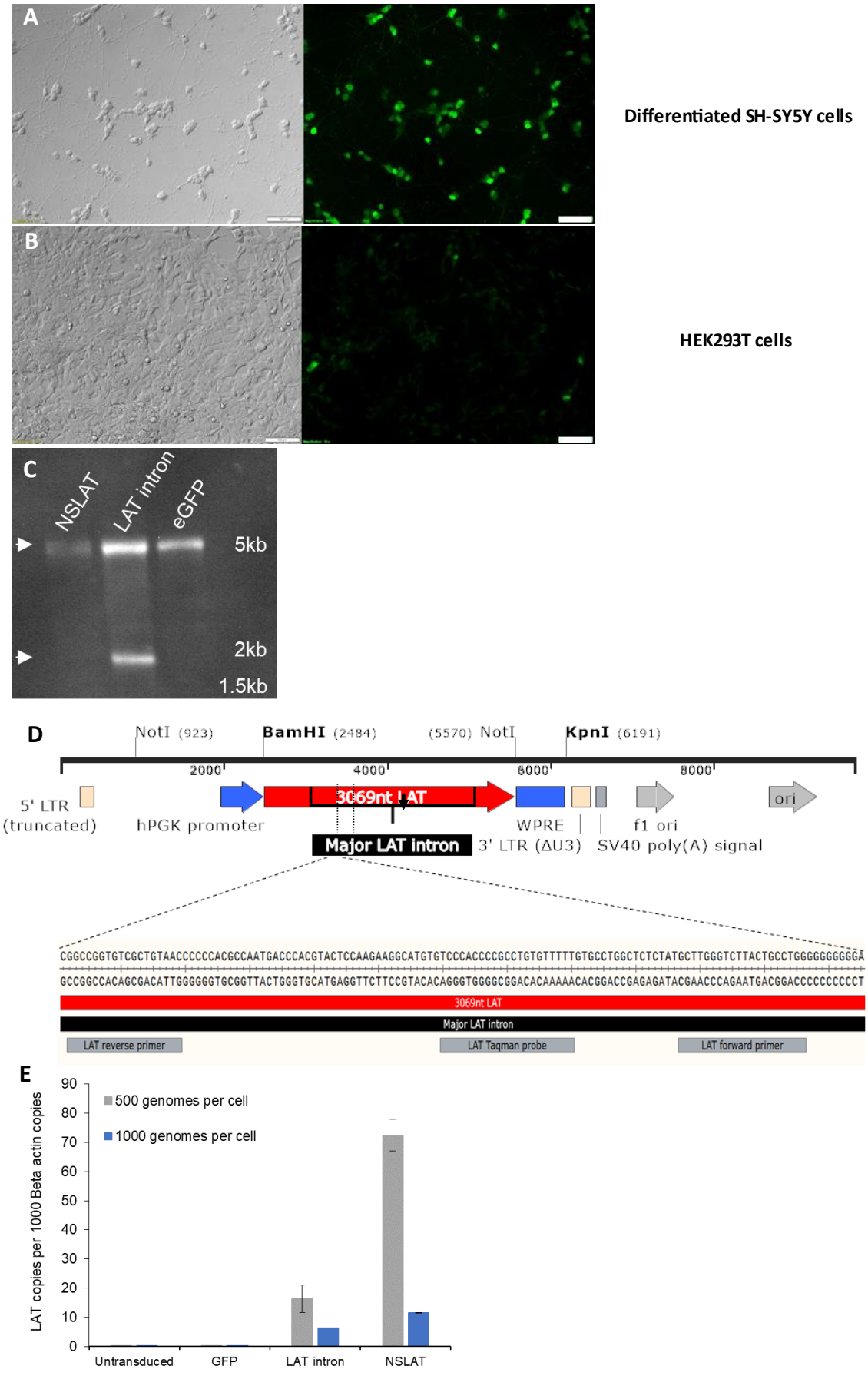


Figure 3.6 characterising the LAT intron and NSLAT lentiviruses

20 and 1000 was utilised (500gpc) to test the expression of LAT during transduction with the LAT intron and NSLAT lentiviruses.

Differentiated SH-SY5Y cells were transduced with eGFP lentivirus, LAT intron lentivirus and NSLAT lentivirus at 500 gpc or mock transduced. 14 dpt total RNA was harvested for extraction and qRT-PCR performed to LAT using the primers described (figures 3.6D and 3.6E). At 500 gpc, both the LAT intron and NSLAT lentivirus transduced SH-SY5Y express LAT, whereas the eGFP lentivirus transduced SH-SY5Y cells or control untransduced SH-SY5Y cells did not (figure 3.6E). This verified that the LAT lentiviruses transcribe LAT. However, the expression of LAT from the LAT intron lentivirus was low with only 16 LAT copies per 1000 beta actin copies, less than the observed LAT expression in the NSLAT transduced cells, which had 72 copies per 1000 beta actin copies (figure 3.6E). This was surprising as the major LAT intron spliced by the LAT intron lentivirus is highly abundant and stable during HSV-1 latency and persists longer than the primary LAT transcript (Thomas et al., 2002). Therefore, LAT might be expected to be detected in higher abundance following LAT intron lentivirus transduction than following NSLAT lentivirus transduction, especially if major LAT is being expressed in every cell, as was the case with the eGFP during eGFP-lentivirus transduction.

This experiment was repeated with 1000 gpc to test whether LAT expression could be improved upon. This gpc was shown to give strong eGFP expression in all SH-SY5Y cells during the earlier eGFP lentivirus transduction (figure 3.2). This again resulted in LAT expression only in the LAT intron or NSLAT transduced SH-SY5Y cells and not in the eGFP lentivirus transduced or untransduced differentiated SH-SY5Y cells, confirming that the LAT lentiviruses produce LAT (figure 3.6E). However, in terms of quantity of LAT produced, this higher transduction resulted in even lower LAT expression with low expression in both LAT intron and NSLAT lentivirus transduced cells with just 6 and 12 copies per 1000 beta actin copies respectively (figure 3.6E).

The low expression levels and variation in the system could be due to a proportion of defective particles in the lentivirus stocks. The titration method used in these experiments to give gpc is achieved by performing qRT-PCR on RNA extracted directly from the virus stocks to give the number of genomes present as previously described. However, this does not account for non-infecting lentiviruses that contain genomes but may not function to enter the SH-SY5Y cells or fail to express the ncRNA in question. This means that calculating genome copy number of virus stocks may not be the most accurate method of titrating the lentiviruses.

An alternative method of titration was tested as a way of measuring functional

lentivirus. Differentiated SH-SY5Y cells were transduced with lentivirus at 1/5, 1/25 and 1/125 dilutions before extracting DNA up to 10 days later and performing qPCR to WPRE, and normalising this to constitutively expressed cellular DNA – GAPDH. This provides the number of lentiviruses that successfully entered the SH-SY5Y cells and were reverse transcribed to produce provirus DNA, while the GAPDH quantity allows the functional lentivirus number to be calculated per cell. At this point, the DNA is ready to be integrated into the host cell DNA. Regardless of whether the genome does integrate or not, the detection of this lentivirus DNA indicates the presence of functional lentivirus capable of expressing the transgene it encodes.

To assess this new titration method LAT intron and NSLAT lentivirus stocks were 5-fold serially diluted until 1 in 125 onto SH-SY5Y cells and the total RNA and DNA harvested 10 dpt. LAT qRT-PCR was performed (using the primers shown in figure 3.6D) on the RNA and normalised to Beta actin, while WPRE qPCR was performed on the DNA and normalised to GAPDH. This allowed for assessment of the relationship between LAT expression and reverse transcribed (and therefore functional) lentivirus, as well as how each of these changed across the lentivirus dilutions (figure 3.7).

Linear regression analysis was performed and as expected, the more virus added, the higher the LAT expression and the more reverse transcribed or “functional” lentivirus as indicated by WPRE per GAPDH (figures 3.7A and 3.7B). The relationship between LAT expression and functional lentivirus was also compared directly and again there was a very strong positive correlation, with R^2 values of 0.9808 and 0.9949 for LAT and NSLAT respectively (figures 3.7C and 3.7D). This highlights the usefulness of titrating lentivirus according to its ability to be reverse transcribed, as the more functional lentivirus there is, the more LAT expression there is. Therefore, functional lentivirus is a more reliable predictor of LAT expression than titration of lentiviral genomes in stock aliquots.

This information was then used to calculate a titration value for infectious lentivirus units, referred to as functional units (FU) in this thesis. The copy number of WPRE DNA per GAPDH produced at each dilution (figure 3.7B) were input into the equation shown in figure 3.7E. This equation takes the WPRE per GAPDH PCR value and multiplies it by 2 to give the number of potentially integrating or ‘functional’ lentivirus copies per cell, as there should be 2 copies of GAPDH per cell. This is then multiplied by the dilution factor of the virus used to transduce the SH-SY5Y cells and the number of cells per well, before dividing by the virus volume used. This gives the FU per ul for each viral stock. Once calculated, FU can then be used to transduce SH-SY5Y cells by

Figure 3.7 Titrating the LAT intron and NSLAT lentiviruses

Differentiated SH-SY5Y cells were transduced by diluting lentivirus stocks in a 5-fold dilution series to titrate the stocks. At 10 days post transduction, DNA and RNA was harvested and PCR performed to WPRE and normalised to GAPDH and qRT-PCR performed to major LAT and normalised to beta actin respectively.

(A) The average of duplicate technical repeats are shown for each biological repeat of which there are 2. Linear regression analysis has been performed between log transformed values of virus dilution values and LAT copies per Beta actin copies and plotted along wtih the linear trendline for each lentivirus. They show a strong positive correlation between virus dilution and LAT expression (as shown by LAT/Beta actin), as indicated by the R^2 values.

(B) The average of duplicate technical repeats are shown for each biological repeat of which there are 2. Linear regression analysis has been performed between log transformed values of virus dilution values and WPRE copies per GAPDH copies and plotted along with the linear trendline for each lentivirus. They show a strong positive correlation between virus dilution and integrating lentivirus (as shown by WPRE/GAPDH), as indicated by the R^2 value.

(C) The average across 2 technical repeats and 2 biological repeats is plotted. Linear regression analysis has been performed between log transformed values of WPRE/GAPDH and LAT/Beta Actin and plotted along with the linear trendline for each lentivirus. They show a strong positive correlation between lentivirus integration (WPRE/GAPDH) and major LAT expression (LAT/Beta actin), as indicated by the R^2 value.

(D) An alternative display of the data from D but instead of a linear regression on log transformed data, the data is plotted as LAT per 1000 beta actin copies, against WPRE/GAPDH. The average across 2 technical repeats and 2 biological repeats is plotted along with a line of best fit.

(E) Schematic representation of the equation used to calculate functional units (FU) (pw = per well) from the information given in graph (B) by performing PCRs on lentiviruses diluted onto differentiated SH-SY5Y cells.

(F) Differentiated SH-SY5Y cells were either untransduced or transduced with 56 functional units (FU) of eGFP lentivirus, LAT intron lentivirus or NSLAT lentivirus. Total RNA was harvested 14 dpt and qRT-PCR was performed to LAT (using the primers shown in figure 3.6D) and normalised to Beta actin. The results show the average from 5 biological repeats and 3 technical repeats ± the standard deviation.

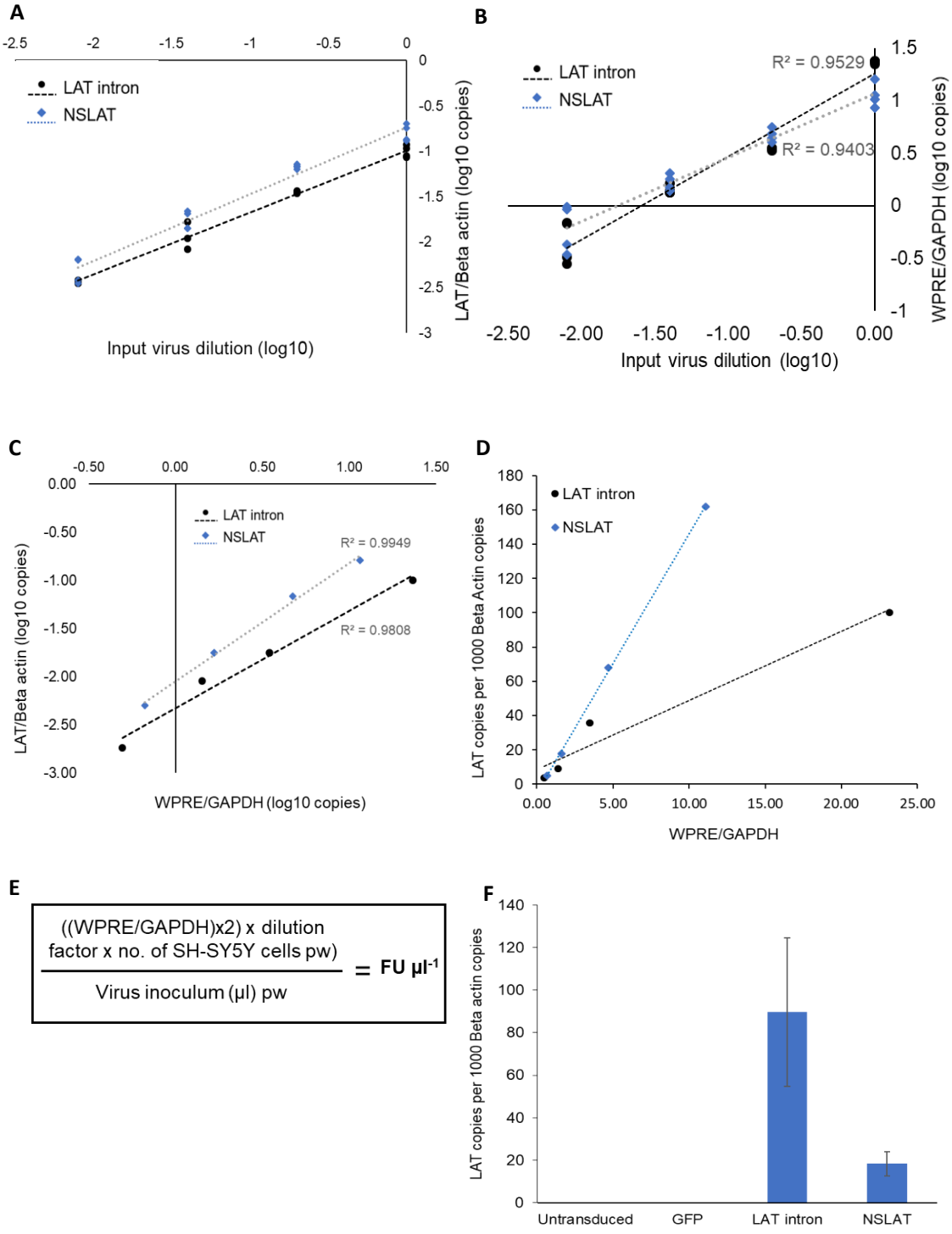


Figure 3.7 Titrating the LAT intron and NSLAT lentiviruses based on functional units (FUs)

functional lentivirus units per cell or “Functional units” (FU) per cell (pc) rather than gpc.

The dilutions used here seemed to produce a higher level of LAT transcription than achieved when transducing with 1000 gpc. When the graph comparing LAT expression and functional lentivirus (figure 3.7C) is displayed according to LAT copies per 1000 beta actin copies (figure 3.7D), LAT transcription reaches 100 and 162 LAT copies per 1000 beta actin copies at the highest dilution for the LAT intron and NSLAT lentiviruses respectively (figure 3.7D). As discussed, the calculation shown in 3.7E expresses the answer per µl, to titrate each lentivirus stock and then FU pc can be calculated from this for each onwards experiment, depending on the number of cells and volume used. However, to work backwards and calculate what the equivalent FU pc at each of the dilutions used for these stocks would have been to achieve the LAT expression shown, a simplified version of this calculation can be used, eliminating the need for volume or cell numbers to be included as the answer is given at a per cell basis. Therefore, the simplified calculation is [((WPRE/GAPDH)x2) x dilution factor = FU pc]. These highest LAT abundance of 100 and 162 copies per 1000 beta actin copies were found with highest dilution factor, which was 1, where the stocks were used neat. Therefore, inputting the average WPRE/GAPDH found from graph 3.7D for each lentivirus at these dilutions into the simplified calculation gives [(23.2)x2) x 1 = 46.4 FU pc] for the LAT lentivirus and [(11.1)x2) x 1 = 22.2 FU pc] for the NSLAT lentivirus. In summary, an average of 46.4 and 22.2 FU pc was used to give 100 and 162 copies LAT and NSLAT respectively with these lentivirus stocks. This information is useful for choosing what FU pc to transduce the differentiated SH-SY5Y cells at to produce substantial levels of LAT in experiments moving forward.

To verify that functional units are a more effective means of titrating and transducing with the lentiviruses, differentiated SH-SY5Y cells were transduced with 56 FU per cell of the LAT intron and NSLAT lentiviruses. Total RNA was extracted 14 days post transduction and qRT-PCR performed to LAT using the primers previously described (3.6B). When transducing in this manner, a higher transcription level of LAT was seen in the LAT intron transduced cells than in the previous genome per cell experiment (figure 3.7F), with an average of 90 LAT copies per 1000 beta actin compared to only 16 and 6. However, NSLAT lentivirus transduced cells showed lower LAT expression than LAT intron-transduced cells, with 18 copies per 1000 beta actin, when also transduced with 56 FU pc. Nonetheless, it is expected that NSLAT lentivirus transduction might give lower abundance than the LAT intron lentivirus transduced cells, given it does not produce the stable LAT intron.

56 FU per cell was chosen by taking into consideration the FU found to reach 100 LAT copies or more per 1000 beta actin as described above (46.4 and 22.2 FU pc). Additionally the ratio of genomes to functional units can be calculated on stocks that were titrated using both gpc and FU methods and although this varies between the lentiviral stocks, it was known and tested for one miRNA lentivirus stock (6,000,000 genomes μl^{-1} and 24,000 FU μl^{-1}) and one GFP lentivirus stock (60,000,000 genomes μl^{-1} and 1,600,000 FU μl^{-1}) and so it was found that there were 250 and 37.5 genomes per FU respectively. Therefore, the 56 FU per cells titre chosen to transduce SH-SY5Y cells in onwards experiments could represent between 2100 and 14000 genomes per cell according to these examples of known ratios of genomes per FU. In both instances, this is higher than the 500 or 1000 gpc used previously to efficiently transduce differentiated SH-SY5Y cells with eGFP lentivirus (figure 3.2). Moreover, the subsequent RNAscope experiment (outlined in section 3.2.4) shows that 56 FU pc of lentivirus leads to LAT expression in between 50 and 80% of the differentiated SH-SY5Y cells (figure 3.8E).

Taken together, this data suggests that the LAT intron and NSLAT lentiviruses successfully transcribe LAT and that assaying for functional lentivirus is the most effective method of titrating and transducing the differentiated SH-SY5Y cells. Using the PCR described provided a measure of functional lentivirus, which was shown to strongly correlate with LAT expression. This was used as the means of titrating the lentivirus stocks and transducing differentiated SH-SY5Y cells for the onwards experiments.

3.2.4 The spatiotemporal dynamics of LAT expression following LAT intron or NSLAT lentivirus transduction of differentiated SH-SY5Y cells

Once constructed and an effective means of titrating the viruses was established, the spatial and temporal aspects of LAT transcription following transduction of differentiated SH-SY5Y cells could be examined. To validate its use as a model for LAT RNA expression, it is of interest to examine whether the major LAT intron accumulates in the nucleus as has been observed to *in vivo* (Arthur et al., 1993). Additionally, to use this model for further experiments, it is crucial to know when LAT expression occurs in this system and how that changes over time.

It has already been shown that there is LAT transcription following transduction with either the LAT intron or NSLAT lentivirus, next it was examined how this LAT expression changes over time following LAT intron lentivirus transduction. Differentiated SH-SY5Y cells were transduced with LAT intron lentivirus at 56 FU pc and total RNA harvested at 5, 7, 11 and 14 dpt. qRT-PCR was carried out for LAT (using the previously described primers) and normalised to beta actin copies in these samples. The results

reveal that there is a peak of LAT expression at 11 dpt with 1203 and 1400 copies per 1000 beta actin, increasing from 7 dpt where LAT has 183 and 242 copies per 1000 beta actin copies (figure 3.8A). There is a decline from 11 dpt to 14 dpt where LAT expression reaches 496 and 578 copies per 1000 beta actin copies.

To examine what proportion of the differentiated SH-SY5Y cells were expressing LAT following transduction with either of the LAT lentiviruses, as well as the localisation of LAT RNA, the *in situ* hybridisation technique of RNAscope was used. Differentiated SH-SY5Y cells were mock transduced or transduced with either the LAT intron lentivirus or the NSLAT lentivirus at 56 FU pc. The cells were fixed 14 dpt and RNAscope performed using a probe directed to the major LAT sequence before imaging by fluorescence microscopy.

From the fluorescent photomicrographs it was apparent that little to no signal was detected in the untransduced negative control differentiated SH-SY5Y cells (figure 3.8B). In contrast, the LAT intron and NSLAT lentivirus-transduced differentiated SH-SY5Y cells both produced LAT RNA (figures 3.8C and 3.8D). However, there are apparent differences in LAT RNA transcription and localisation between the 2 transductions. The LAT intron lentivirus transduction showed more intense signal compared to the NSLAT lentivirus transduced cells. LAT RNA seems to be expressed in puncta in both transductions but in the NSLAT lentivirus transduction, the LAT RNA is more dispersed, with some non-nuclear localisation in addition to nuclear localisation. LAT intron lentivirus transduced SH-SY5Y cells seems to show more concentrated nuclear localisation of the LAT (figures 3.8B and 3.8C).

The number of cells expressing LAT were then quantified by manually counting green cells in imageJ (Schindelin et al., 2012)(methods 6.2.7.3) and this revealed that although the LAT intron transduction showed more LAT expression in terms of intensity of signal, the number of cells that were positive for LAT expression were on average less than in the NSLAT transduced cells with 56% and 73% respectively (figure 3.8E). This helps characterise the system in terms the number of cells that express LAT at a given time when transduced with 56 FU pc but also highlights again that there is a difference between the LAT intron and NSLAT lentivirus transduced cells. This supports the previous northern blot data suggesting there is a difference in the ability of the LAT intron or NSLAT lentivirus to splice the major LAT intron. The accumulation of LAT in the nucleus of LAT intron lentivirus transduced cells (figure 3.8C) is indicative of the accumulation of major LAT (Arthur et al., 1993), suggesting splicing has occurred, this is not observed to the same extent in the NSLAT transduced cells (figure 3.8D).

Figure 3.8 Characterisation of the spatial and temporal aspects of LAT expression following transduction with the LAT intron and NSLAT lentiviruses

(A) Differentiated SH-SY5Y cells were transduced with 56 FU per cell of LAT intron lentivirus in duplicate and total RNA extracted at 5, 7, 11 and 14 dpt then qRT-PCR for major LAT performed and normalised to Beta actin.

(B) Untransduced differentiated SH-SY5Y cells were fixed 11 dpt and RNAscope performed with a probe to major LAT (green) and nuclear DAPI stained (blue). As a negative control for C and D. Scale bar represents 50μm.

(C) Differentiated SH-SY5Y cells were transduced with LAT intron lentivirus, at 56 FU pc. Cells were fixed 11 dpt and RNAscope performed with a probe to major LAT (green) with nuclear DAPI staining (blue). The white boxes have been enlarged below the original image. Scale bar represents 50μm.

(D) Differentiated SH-SY5Y cells were transduced with NSLAT lentivirus, at 56 FU pc. Cells were fixed 11 dpt and RNAscope performed with a probe to major LAT (green) and nuclear DAPI stained (blue). The White boxes have been enlarged below the original image. Scale bar represents 50μm.

(E) The number of green and therefore major LAT positive cells from LAT intron or NSLAT lentivirus transduced differentiated SH-SY5Y cells in C and D) were manually quantified in Image J (as described in methods 6.2.7.3). The blue circles represent individual images counted and the black line shows the mean percentage of LAT-positive cells for either LAT intron or NSLAT transduced cells (5 images each).

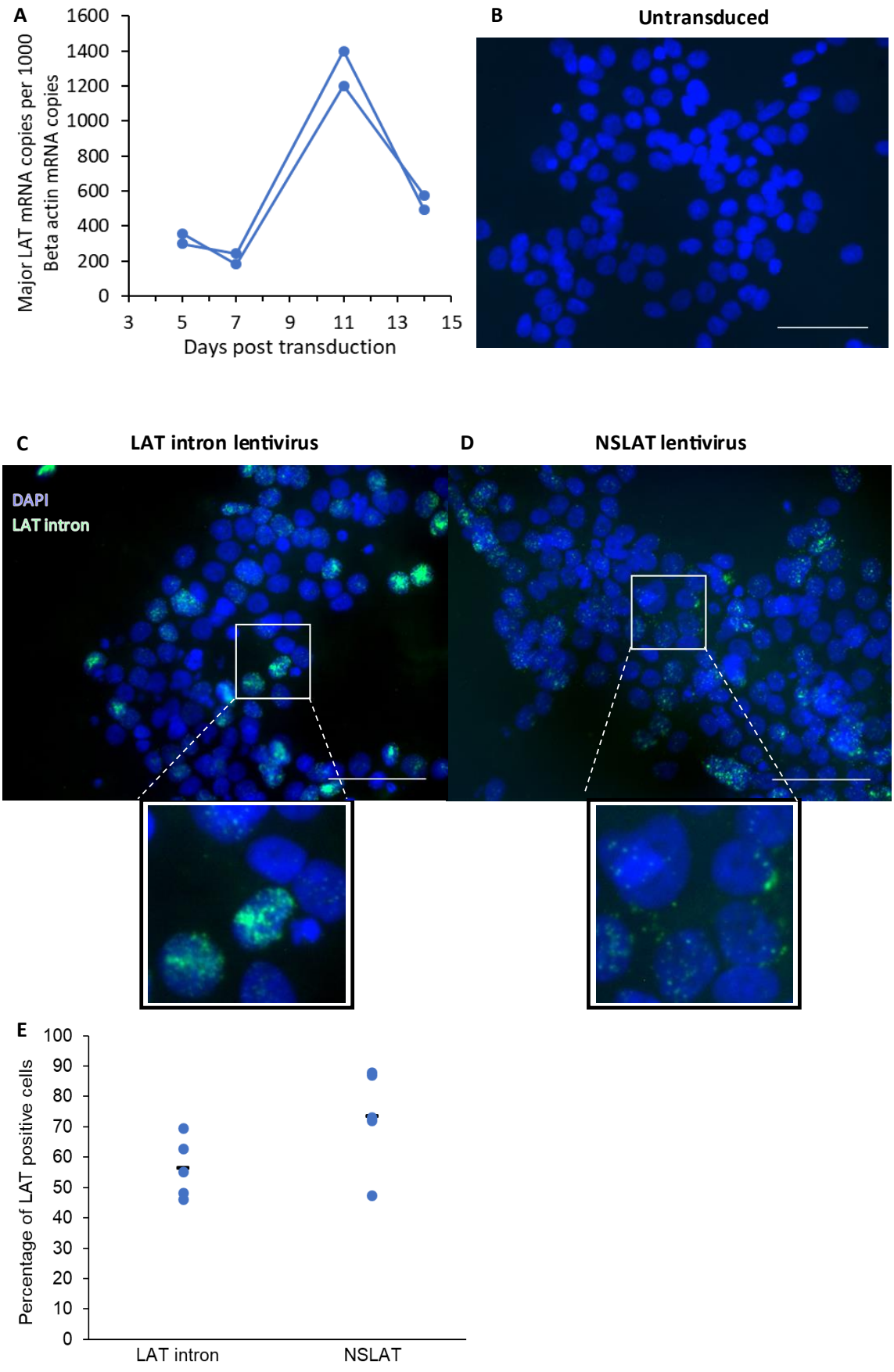


Figure 3.8 Characterisation of the spatial and temporal aspects of LAT expression following transduction with the LAT intron and NSLAT lentiviruses

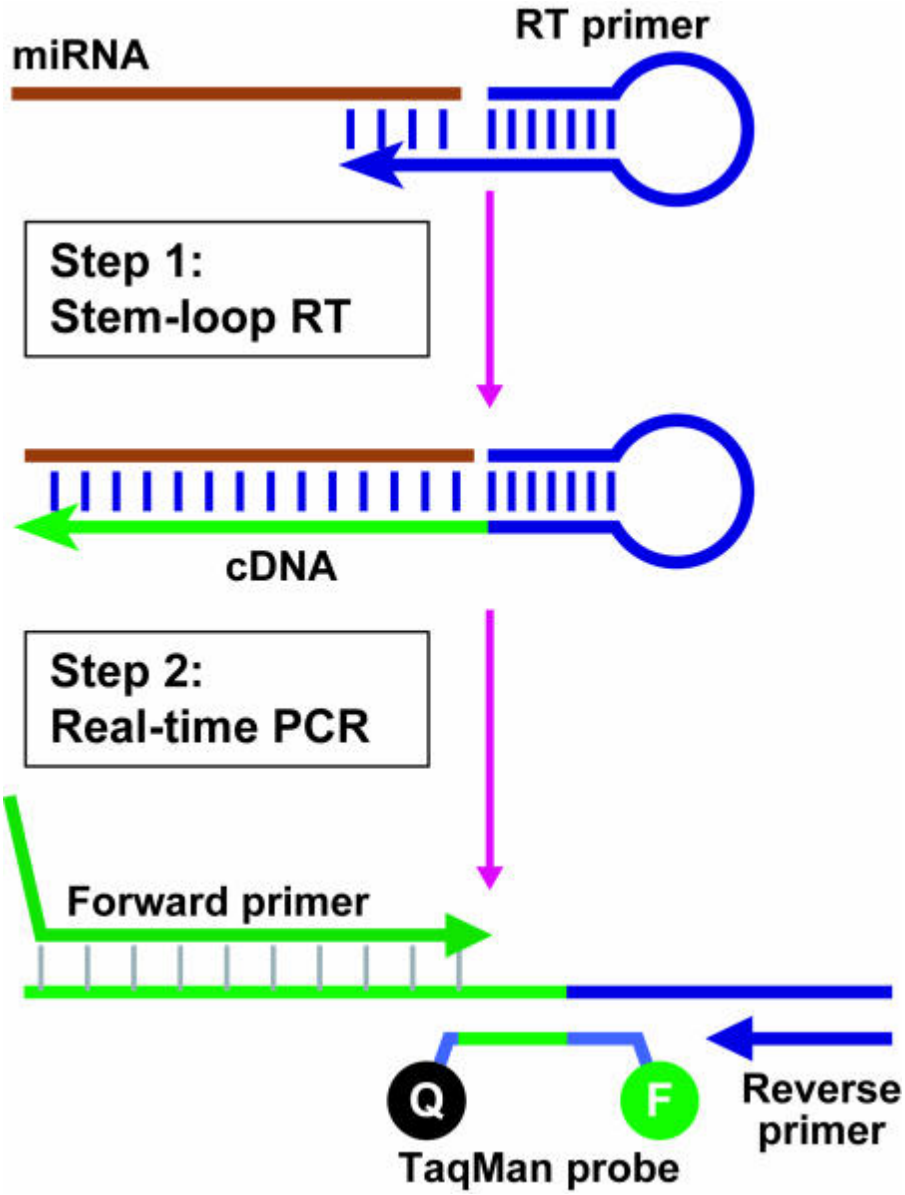
Taken together, this data strongly supports earlier observations indicating that LAT is expressed after transduction with the LAT intron and NSLAT lentiviruses as well as providing information about the spatial and temporal organisation of LAT when expressed by the LAT lentiviruses. In addition, there are differences observed between the LAT expressing lentiviruses, suggesting that as anticipated, only the LAT intron lentivirus splices the LAT transcript to produce the major LAT intron. As such these lentiviruses successfully produce and deliver LAT to differentiated SH-SY5Y cells.

3.2.5 MicroRNAs are produced by the MicroRNA lentivirus

In the previous section, the LAT intron and NSLAT lentiviruses were shown to successfully transcribe LAT in differentiated SH-SY5Y cells, with differences in expression suggesting that the major LAT intron is produced in the LAT intron lentivirus transduced cells only. In this section, a characterisation of the lentivirus construct encoding the latency-associated miRNAs is described.

Stem-loop qRT-PCR was utilised to test whether the lentivirus construct encoding 5 HSV-1 miRNAs can express these miRNAs, following transduction of differentiated SH-SY5Y cells. Stem loop qRT-PCR works by use of a reverse transcription (RT) primer that contains a stem-loop and a 6-nucleotide extension complementary to the miRNA of interest at the 3' end. The stem-loop extends the roughly 22 nucleotide long miRNA to approximately 60 nucleotides or longer during first strand cDNA synthesis so that traditional PCR can be performed subsequently by allowing annealing of a primer pair (figure 3.9) (Chen et al. 2005). This stem-loop qRT-PCR specifically detects mature miRNAs rather than input virus cDNA/RNA.

Figure 3.9 Diagram describing the steps of stem-loop qPCR



Schematic diagram of stem-loop q-RT-PCR, involving 2 steps; stem-loop RT and real-time PCR. The RT stem-loop primer binds the 3' end of the miRNA and the product of this reverse transcription is quantified using traditional TaqMan qPCR as indicated. Taken from Chen et al. 2005.

Differentiated SH-SY5Y cells were either mock transduced or transduced with miRNA lentivirus or eGFP lentivirus at 500 gpc or 56 FU pc for experiment 1 and 2 respectively. RNA was harvested 14 dpt and stem-loop qRT-PCR performed over the course of the 2 experiments; miR-H2, 4 and 6 in experiment 1 and mir-H2, 3 and 7 in experiment 2. Stem loop qRT-PCR was also targeted to a human miRNA, let-7a, as a positive cellular miRNA control.

The results are displayed as the relevant cycle threshold (Ct) values, meaning the number of cycles taken for each sample to reach detectable levels (therefore the lower the Ct value, the more miRNAs initially present in that sample). The negative controls of eGFP transduced or untransduced differentiated SH-SY5Y cells provides baseline Ct values, whereby no HSV miRNAs are produced. The results show that the miRNA lentivirus transduced SH-SY5Y cells produce each miRNA tested, with Ct values lower than the baseline set by GFP-transduced and untransduced cells (which is perhaps indicative of minimal cross-reactivity to cellular miRNAs).

In experiment 2, the baseline Ct value set for mir-H2, 3 and 7 in the control eGFP transduced or mock differentiated SH-SY5Y cells was 33 or above, this was also the case for mir-H6 in experiment 1 (figures 3.10B, 3.10C, 3.10E and 3.10F). In the miRNA lentivirus transduced cells, each miRNA was produced in a much higher abundance than this threshold, reaching levels similar to the cellular let-7a with Ct values of 25 or lower. In experiment 1, a baseline Ct value wasn't required as miR-H2 and 4 were not detected at all during the 40 PCR cycles in the eGFP lentivirus or untransduced SH-SY5Y cells (figures 3.10A and 3.10D). By contrast, in the miRNA lentivirus transduced differentiated SH-SY5Y cells, miR-H2 and 4 were detected at 29 and 32 cycles respectively.

This data from both experiments suggests that the miRNA lentivirus effectively produces miRNAs miR-H2-3p, miR-H3-3p, miR-H4-5p, miR-H6-3p and miR-H7-5p in differentiated SH-SY5Y cells.

Figure 3.10 The microRNA lentivirus transcribes the 5 latency-associated HSV-1 microRNAs in differentiated SH-SY5Y cells

(A) Differentiated SH-SY5Y cells were transduced with 500 gpc of the miRNA lentivirus, eGFP lentivirus or mock transduced before harvesting the RNA 14 dpt for stem-loop qRT-PCRs to miR-H2-3p (blue circles) and let-7a (black squares). Shown are the averages from 3 biological repeats and 3 technical repeats. Error bars show the standard deviation from the mean.

(B) Differentiated SH-SY5Y cells were transduced with 56 FU pc of miRNA lentivirus, eGFP lentivirus or mock transduced before harvesting the RNA 14 dpt for stem-loop qRT-PCRs to miR-H2-3p (red circles) and let-7a (black crosses). Shown are the averages from 6 biological repeats and 2 technical repeats. Error bars show the standard deviation from the mean.

(C) Differentiated SH-SY5Y cells were transduced with 56 FU pc miRNA lentivirus, eGFP lentivirus or mock transduced before harvesting the RNA 14 dpt for stem-loop qRT-PCRs to miR-H3-3p (red squares) and let-7a (black crosses). Shown are the averages from 6 biological repeats and 2 technical repeats. Error bars show the standard deviation from the mean.

(D) Differentiated SH-SY5Y cells were transduced with 500 gpc of miRNA lentivirus, eGFP lentivirus or mock transduced before harvesting the RNA 14dpt for stem-loop qRT-PCRs to miR-H4-5p (blue diamonds) and let-7a (black squares). Shown are the averages from 3 biological repeats and 3 technical repeats. Error bars show the standard deviation from the mean.

(E) Differentiated SH-SY5Y cells were transduced with 500 gpc of miRNA lentivirus, eGFP lentivirus or mock transduced before harvesting the RNA 14 dpt for stem-loop qRT-PCRs to miR-H6-3p (blue triangles) and let-7a (black squares). Shown are the averages from 3 biological repeats and 3 technical repeats. Error bars show the standard deviation from the mean.

(F) Differentiated SH-SY5Y cells were transduced with 56 FU pc miRNA lentivirus, eGFP lentivirus or mock transduced before harvesting the RNA 14 dpt for stem-loop qRT-PCRs to miR-H7-5p (red diamonds) and let-7a (black crosses). Shown are the averages from 6 biological repeats and 2 technical repeats. Error bars show the standard deviation from the mean.

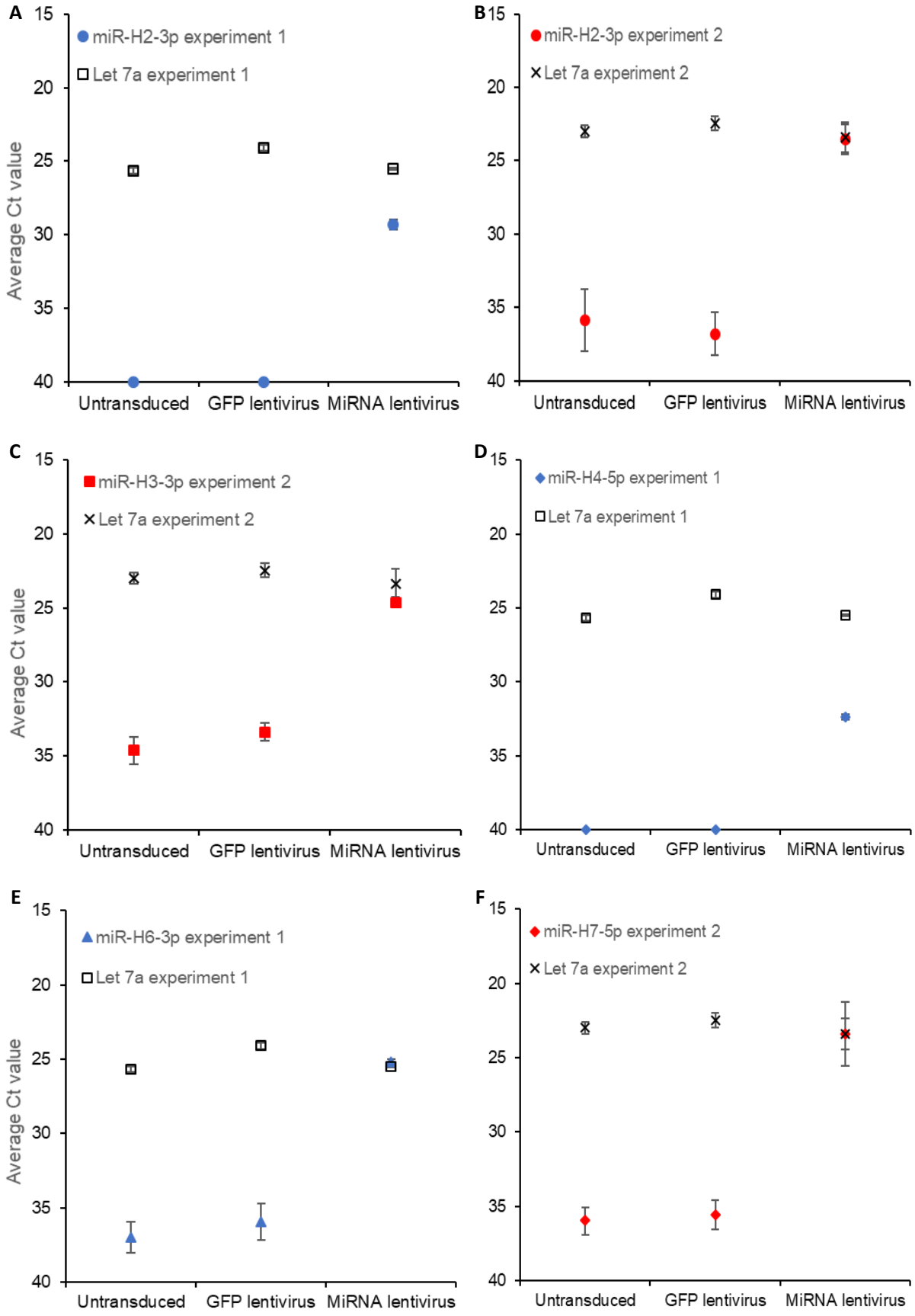


Figure 3.10 The microRNA lentivirus transcribes the 5 latency-associated HSV-1 microRNAs in differentiated SH-SY5Y cells

3.3 Discussion

Summary

In this chapter, methods and efficiency of delivering latency-associated ncRNAs directly to neuronal SH-SY5Y cells were investigated. Lentivirus vectors were found to be an effective means of transducing differentiated SH-SY5Y cells, with a better efficiency on a per genome copy number basis than an AAV2 vector. The lentiviruses were designed to deliver either the first 3.1kb of the LAT transcript – with or without splicing – or 5 latency-associated miRNAs. Two methods were compared for titrating the lentiviruses. Titrating the lentiviruses by FU and using this to transduce the differentiated SH-SY5Y cells was shown to be an effective way to deliver the latency-associated ncRNAs to the cells. This, in conjunction with the replication defective HSV-1 latency model described earlier, could be useful in determining the roles of these ncRNAs during latency in neurons. This system could be used for assessing the effect on the neuronal transcriptome, which is one of the aims of this project.

Comparing transduction efficiency

The first test, comparing transduction efficiencies of AAV and lentivirus vectors in differentiated SH-SY5Y cells demonstrated that the AAV vector was not a suitable one in this cell type (figures 3.1A and 3.1B). When transducing cells with 1000 gpc the AAV2 eGFP vector showed only limited GFP expression even 18 dpt compared to the eGFP lentivirus vector transduction with 1000 gpc showing strong eGFP expression in most of the cells even by 4 dpt (figure 3.2). Although other AAV serotypes have been shown to transduce neuronal cells, Hammond *et al.* demonstrated using GFP expressing AAVs that AAV serotype 2/1 was the most effective at transducing primary murine neurons, therefore it seemed like the most appropriate to test, but it is possible that other serotypes might have been effective where AAV-2 was not (Haery et al., 2019; Hammond et al., 2017). Nonetheless, it was unnecessary to test any further vectors as the lentivirus vector was very effective in entering the differentiated SH-SY5Y cells and expressing eGFP. Besides the type of viral vector used, there are additional factors that may have influenced the efficiency of expression, such as the AAV vector containing a CMV promoter to drive eGFP expression while the lentivirus used the hPGK promoter. Nevertheless, since the results demonstrated that the lentivirus was very effective both in successfully transducing the differentiated SH-SY5Y cells and expressing eGFP marker, it was not necessary to carry out further investigation. Consequently, the lentivirus with hPGK promoter was selected for further expression and analysis of the latency-associated ncRNAs.

Construction and characterisation of the latency-associated ncRNA expressing lentiviruses

Next, the lentiviruses were designed to deliver the latency-associated ncRNAs to the differentiated SH-SY5Y cells. The LAT intron lentivirus was designed to contain the first 3.1kb of LAT (figure 3.3). This not only includes major LAT but also contains key splice acceptor and donor sites which allows for major LAT intron to be spliced. Lentiviruses have a maximum insert size of 9kb, which although could contain the entire 8.3kb LAT transcript, including other elements of the lentivirus such as WPRE, which is required to stabilise the ncRNAs encoded, limits the maximum possible size of LAT to 8.05kb. In addition to this, it has been shown that the larger vector length the lower the viral titres (Kumar et al., 2001). As a key focus of this study is the effect of major LAT, the first 3.1kb is a sufficient in length to focus on the role of the major LAT intron, and not compromise on viral titre.

The NSLAT lentivirus was designed to be a control for the LAT intron lentivirus, where the first 3.1kb of the transcript is present as with the LAT intron lentivirus, but there are deletions in the key splice donor and acceptor sites so that the major LAT intron cannot be spliced (Alvira et al., 1999)(figure 3.3). This allows elucidation of the functions that are specifically as a result of the major LAT intron over those that only require the presence of the major LAT sequence.

The miRNA lentivirus was designed to contain the 5 of the most abundant HSV-1 miRNAs during latency; miR-H2-4 and miR-H6-7 (table 3.1), as were shown by Umbach *et al.* in latently infected human trigeminal ganglia samples (Umbach et al. 2009).

Initial characterisation steps performed directly on these plasmids confirmed that the construction of the lentiviruses matched the design intention. Firstly, sequencing of the lentiviruses showed that there were very few deviations from the reference sequences and at least 99% of each reference sequence had coverage of above 100 reads. Although there were some reductions in read-depths and a few potential mutations in the parental construct used to make the lentiviruses, from the data observed, this had no impact on the transduction efficiency or expression of transgenes in the differentiated SH-SY5Y cells.

The next test performed on the lentivirus plasmids was a northern blot that aimed to address whether the splicing phenotypes of the LAT lentivirus plasmids are as they

have been constructed to be. For this experiment 293T cells were transfected using the lentivirus plasmids and then total RNA extracted. Total RNA was extracted because this northern blot required 1µg of RNA and poly A selection might have been a limiting factor in achieving this, plus the major LAT intron would not have been included if poly A selection had been performed. 293T cells were chosen over differentiated SH-SY5Y cells to obtain enough RNA. As the 293T cells don't require growth on extracellular matrix or any neurotropic factors, harvesting 1-2µg of total RNA was easier to achieve. As the lentiviruses are produced by transfection of 293T cells along with helper plasmids, it was known that these cells can be transfected with these lentivirus plasmids. Considering that the transduction efficiency with the eGFP lentivirus produced from the plasmid was relatively low both in terms of strength of eGFP signal and proportion of cells expressing eGFP (as compared to the differentiated SH-SY5Y cells)(figure 3.6A and B), these cells were transfected with the lentiviral plasmids instead, to compare the LAT lentiviruses splicing phenotypes. Regardless, the splicing of LAT from the plasmids should reflect whether LAT is spliced from the lentiviruses.

In the northern blot analysis, the probe was directed specifically against the LAT RNA, nevertheless, a 5kb RNA was detected in eGFP lentivirus transfected cells as well as the LAT lentivirus transfections (figure 3.6C). Given the sequencing of the lentiviruses and the qRT-PCR experiment testing for LAT expression showing that the eGFP lentivirus does not encode or transcribe LAT (figure 3.6E), this band cannot represent LAT RNA. Due to its size and abundance, this 5 kb band is likely due to cross hybridisation of the LAT specific probe to 28S ribosomal RNA (rRNA) (4.7kb). Cross-hybridisation is plausible here given the huge abundance of rRNA within a cell and as no poly A selection was used here. Additionally, the LAT region is highly GC rich, meaning the probe although directed to LAT might be able to bind GC regions of the rRNA. The 3.1kb of the primary LAT transcript was not detected in either the LAT intron lentivirus or NSLAT lentivirus transfected cells. In the LAT intron transfection this could be due to the primary LAT transcript being present in very low abundance due to efficient splicing of the major LAT intron. In the NSLAT lentivirus the 3.1kb primary LAT might not be observed either due to low abundance or instability of the unspliced LAT transcript. Regardless, there is a band at approximately 2 kb (which is the size of the major LAT intron), only in the LAT intron transfected cells and not the NSLAT lentivirus or eGFP lentivirus transfections. This suggests that this band is likely due to the 2.0kb LAT intron. In support of this, the subsequent RNAscope data showed differences in abundance and localisation of LAT RNA between the LAT intron and NSLAT lentivirus transduced cells. The cytoplasmic localisation of LAT RNA following NSLAT lentivirus transduction (figure

3.8D) indicates that the primary unspliced 3.1kb LAT transcript is exported to the cytoplasm, as can occur with unspliced RNAs (Stewart, 2019). In contrast, the nuclear-only signal observed following LAT intron lentivirus transduction (figure 3.8C), indicates that the major LAT intron is spliced and thus retained in the nucleus (Stewart, 2019). Altogether, the data suggests that during transduction with the LAT intron lentivirus, the primary LAT transcript is efficiently spliced to produce major LAT whereas it is not during the NSLAT lentivirus transduction. There is no band at 1.5kb to represent the alternatively spliced major LAT intron. This is expected given that this sized intron is only observed in neurons, and non-neuronal 293T cells are used here (Brinkman et al., 2013; Spivack and Fraser 1987). Further northern blots would need to be performed to appreciate whether, as expected, there is production of the 1.5kb and 2.0kb major LAT introns in the neuronally differentiated SH-SY5Y cells. Nevertheless, this northern blot distinguished the splicing phenotypes between the two LAT lentiviruses and further analyses performed with qRT-PCR (figures 3.6E and 3.7F) and RNAscope (figures 3.8C and 3.8D) provided evidence of major LAT expression following lentivirus delivery to the differentiated SH-SY5Y cells.

Comparison of titration methods

The data demonstrated that transducing the differentiated SH-SY5Y cells with LAT lentiviruses titrated by genome copy number, resulted in expression of LAT. However, genome copy number titration of the lentivirus stocks led to varied and unexpected results in terms of the amount of LAT transcribed in cells transduced with each of the LAT lentiviruses. For example, a 1000 gpc transduction yielded less LAT expression than a 500 gpc transduction (figure 3.6E). Additionally, the copy numbers on average were low, with LAT intron lentivirus only transcribing between 6 and 16 copies per 1000 beta actin copies. Another unexpected result was that the NSLAT lentivirus transduced cells seemed to transcribe more LAT than the LAT intron lentivirus, with up to 72 LAT copies per 1000 beta actin. This was unexpected as the major LAT intron is highly stable and therefore the expectation would be that the LAT intron lentivirus transduction would result in a greater abundance of 2.0kb major LAT. Accordingly, to test whether expression levels could be improved upon, an alternative method of titration was examined, by measuring functional lentivirus units (FU). However, this higher abundance of LAT RNA following NSLAT lentivirus transduction compared to LAT intron lentivirus transduction is also observed during FU titration assays by qRT-PCR and PCR (figure 3.7A, C and D).

Once titrated using functional lentivirus units (FU) rather than genome copy number, both the LAT intron and NSLAT lentiviruses showed efficient LAT production (figures 3.7 and 3.8). This shows that this means of titration is effective. FU only includes functional lentivirus because the virus needs to have entered the differentiated SH-SY5Y cells and had its genome reverse transcribed in order to produce WPRE DNA, which is indicated by the WPRE per GAPDH PCR. There is still variation in LAT expression between experiments when using FU pc (such as the low LAT expression observed following 56 FU NSLAT lentivirus transduction of differentiated SH-SY5Y cells in figure 3.7F). Nevertheless, with both the LAT intron and NSLAT lentiviruses a strong positive correlation was observed between the FU used and LAT expression (figure 3.7C), confirming this is a useful method as a means of titrating the lentiviruses. An FU of 14 per cell was selected for onwards experiments as this was even higher than the highest virus dilution utilised in the titration experiment that seemed to yield high LAT expression (figure 3.7D). In addition, between 50 and 80% of differentiated SH-SY5Y cells expressed LAT when transduced at 56 FU with either LAT lentivirus (figure 3.8E). Given that this percentage of LAT positive cells either exceeds or is equal to the percentage of LAT positive cells seen with a high MOI quiescent in1382 infection of SH-SY5Y cells, this nominal FU titre seemed appropriate to use for any onwards experiments requiring expression of the latency-associated ncRNAs.

The ratio of genomes per functional lentivirus (gpc vs FU), varied between different lentiviral stocks, which highlights the need to titrate via functional lentiviruses as genome copy number is not a true reflection of the amount of functional lentivirus present. Additionally, this variation between the different lentiviruses suggests that the transgene being expressed has an impact on the success rate of virus production. Altogether, the FU method was shown to be the most relevant and effective means of titrating the lentiviruses.

Single-cell analysis of LAT expression

From the initial 1000 gpc eGFP lentivirus transduction (figure 3.2), it seemed that all the SH-SY5Y cells are transduced and express eGFP. Despite this, when the LAT intron or NSLAT lentiviruses (constructed from the same eGFP lentiviral plasmid) are added in a higher quantity, at 56 FU, not all the cells express LAT by RNAscope. RNAscope provides a means of measuring of LAT expression. However, it is not clear by RNAscope whether all the cells are transduced, as expected from the eGFP lentivirus, and only a proportion of these transduced cells express LAT or whether transduction efficiency is reduced in the LAT lentiviruses. There is precedent for LAT only being

expressed in a proportion of infected cells from what was seen in the previously described in1382 HSV-1 latency model (chapter 2.2.5) and also with that seen in previously published work both in animal and *in vitro* models (Chen et al. 2002; Edwards and Bloom 2019; Mehta et al. 1995). This could suggest that there is something inherent to HSV-1 latency that limits LAT expression occurring in all infected neurons. In *in vivo* models, this variation in LAT expression could potentially be explained by the presence of different neuronal sub-types and regulation by transcription factors that differ between neuronal sub-types (Bertke et al., 2011). However, in this and other *in vitro* models (Edwards & Bloom, 2019), the cell-type is homogenous, which suggests that there could be something intrinsic to LAT (especially as it occurs following lentivirus delivery) that prevents expression in all cells. LAT might self-regulate its own expression in adjacent cells by a state of quiescence being signalled between cells. This could be one reason why examining whether LAT has an impact on the neuronal transcriptome could be of relevance. Nonetheless, it seems likely that most of the cells are transduced but only a proportion of these express LAT given this precedent for limited LAT expression and because the sequences are so similar between the LAT lentiviruses and the eGFP lentivirus, it is expected that the transduction efficiency would be the same. Regardless of the reason why not all the LAT intron or NSLAT lentivirus transduced cells express LAT, it is something to be aware of in this system when using these lentiviruses to assess the roles of the HSV-1 latency-associated ncRNAs. Nevertheless, LAT is successfully delivered by the LAT lentiviruses and on average, LAT is expressed in more cells by LAT lentivirus delivery (56-73% LAT positive cells), especially by the NSLAT lentivirus (73% average) (figure 3.8E), than during in1382 quiescent infection of differentiated SH-SY5Y cells (28-59% LAT positive cells)(figure 2.6C), although there was some variation between separate in1382 experiments as discussed in the previous chapter. The LAT expression being limited to a subset of the in1382-infected differentiated SH-SY5Y cells again might be a sign that there is an unknown method of regulation intrinsic to LAT.

This difference between LAT expression and eGFP expression highlights an issue in using the eGFP lentivirus as a control for transgene expression because although the lentiviruses may be transducing the cells with the same efficiency, it seems that there are less cells producing LAT in equivalent transductions and as miRNAs are too small to be tested by RNAscope, there is no means to test the number of miRNA producing cells in the miRNA lentivirus transduced differentiated SH-SY5Y cells. Taking all this into consideration, in reproducing these lentiviruses it would likely be useful to include a reporter transgene such as eGFP in each lentivirus genome to enable

visualisation of expression in the transduced cells, although this addition in size could have negatively impacted viral titres (Kumar et al., 2001).

Spatial and temporal analysis of LAT expression

The LAT intron lentivirus was also characterised for the kinetics of the production of LAT using time points similar to those used in the previous chapter for the in1382 latency model as these two systems are intended to be compared and even used in parallel. LAT expression in LAT intron lentivirus transduced differentiated SH-SY5Y cells over time reflected what was observed previously in the in1382 HSV-1 latency model, with a peak of LAT expression at 11 dpt. The combination of transducing the cells with 56 FU and harvesting total RNA at the 11 dpt time point resulted in the highest LAT expression that had been observed with these lentiviruses, with over 1000 LAT copies per 1000 beta actin copies (figure 3.8A), validating the use of 56 FU per cell. Beyond 11 dpt there is a reduction of LAT expression but still over 400 LAT copies per 1000 beta actin.

Considering the peak of LAT expression was shown to be at 11dpt by qRT-PCR, RNAscope was performed to LAT RNA on LAT intron or NSLAT lentivirus transduced differentiated SH-SY5Y cells at this time point. By qualitive observation of the LAT and NSLAT lentivirus transduced cells it appears as though LAT expression is higher in the LAT intron lentivirus transduced cells, compared to the NSLAT lentivirus transduced cells. This could be explained by the LAT intron lentivirus being capable of splicing and producing the major LAT intron which is stable and abundant. However, this is not mirrored in the percentage of LAT positive cells. When quantifying the number of LAT positive (green) cells by image J (as described in method 6.2.7.3) there were on average 17% more LAT positive cells in the NSLAT lentivirus transduced cells than in the LAT intron lentivirus transduced cells with 73% and 56% respectively. This single-cell analysis revealed a discrepancy between the lentivirus that produced a higher LAT abundance per cell and the lentivirus with the greatest number of cells expressing LAT. This discrepancy could explain the variation observed in terms of which lentivirus showed a greater abundance of LAT when assessed by qRT-PCR analysis of RNA taken from many cells. Regardless, there is LAT expression in a majority of cells following transduction of either LAT lentiviruses.

RNAscope also showed a difference in the localisation of LAT in each LAT lentivirus transduction. LAT intron lentivirus transduction of differentiated SH-SY5Y cells shows strong nuclear localisation of LAT in puncta while NSLAT lentivirus shows LAT more weakly expressed in both nuclear and cytoplasmic puncta. This difference supports

the idea that the LAT intron lentivirus splices major LAT while the NSLAT does not because the major LAT intron has previously been shown to accumulate in the nucleus (Arthur et al., 1993). In addition, an unspliced LAT transcript, as is expected with the NSLAT lentivirus transductions, would likely be exported to the cytoplasm given that the lentiviruses contain WPRE in their genomes (figures 3.3 and 3.4B), which has been shown to increase the cytoplasmic accumulation of unspliced RNAs (Donello et al., 1998). Additionally, the lentiviruses each contain a poly A signal (figures 3.3, 3.4A and 3.4B) and therefore, if the major LAT intron is not spliced out, the LAT transcript would be polyadenylated, leading to nuclear export (Stewart, 2019). Therefore, these differences in localisation and expression support the idea that there is splicing of major LAT during transduction with the LAT intron lentivirus and not the NSLAT lentivirus, which was the desired outcome when designing these viruses. This agrees with the previous results seen by northern blot.

Analysis of the miRNA lentivirus

Finally, as miRNAs are approximately 22 nucleotides long ncRNAs they are too small to be analysed using RNAscope probes, northern blot probes and traditional qRT-PCR primers, therefore, stem-loop qRT-PCR was utilised instead (figure 3.9). This revealed that differentiated SH-SY5Y cells transduced with the miRNA lentivirus produces the 5 miRNAs that it is designed to (figure 3.10). There are some differences between experiment 1 and 2 in terms of the quantity of miRNAs produced, especially when comparing the same miR-H2-3p in both experiments. However, this can be explained by experiment 1 being transduced with 500 gpc while experiment 2 used 56 FU pc, which could equate to between 4 (as shown for a miRNA stock) and 28 times as many genomes given the examples of genome to FU ratios known. Regardless, there is clearly more lentivirus added in experiment 2 to achieve this FU pc. This explains why higher quantities of the miRNAs were seen in experiment 2. This again highlights the benefit of using FU rather than gpc when titrating and transducing with the lentiviruses.

Conclusion

In conclusion, the lentiviruses are an effective way to deliver the HSV-1 latency-associated ncRNAs to the differentiated SH-SY5Y cells. This will be a useful reductionist approach that alongside the in1382 latency model will help to assess the effect of these latency-associated ncRNAs in neurons.

Chapter 4: The role of the HSV-1 latency-associated ncRNAs on the human neuronal transcriptome

4.1 Introduction

In the previous chapters of this thesis an *in vitro* model for expressing the HSV-1 latency-associated ncRNAs in neuronally differentiated SH-SY5Y cells was developed. The ncRNAs were expressed either by using a replication-defective HSV-1 (in1382) that establishes a quiescent infection or by transducing the cells with lentiviruses engineered to deliver major LAT or 5 HSV-1 miRNAs. This chapter now explores whether these systems can be utilised to reveal any novel functions of major LAT or the HSV-1 latency-associated miRNAs in human neurons, by examining the effect that these latency-associated ncRNAs have on the human neuronal transcriptome.

To test the effect of the latency-associated ncRNAs on the neuronal transcriptome, the technique of RNA-Sequencing (RNA-Seq) was utilised. RNA-Seq is a transcriptome profiling technology that utilizes next-generation sequencing. For this project, total RNA was extracted from samples, quality control measured and filtered for 3' polyadenylation (poly A selected). The RNA transcripts are then reverse transcribed into cDNA and adapters ligated to each end of the cDNA. Next sequencing was performed and aligned to the relevant reference transcriptome. This method provides a genome-wide expression profile (Kadakkuzha & Puthanveettil, 2013). There are several advantages of using RNA-Seq over traditional microarray methods. For example, RNA-Seq can detect novel transcripts as it does not require species or transcript-specific probes. Additionally, and more relevant to this work, RNA-Seq has higher specificity and sensitivity, whereby a larger percentage of differentially expressed genes (DEGs) can be detected, especially those with low expression (Illumina, 2019; Kadakkuzha & Puthanveettil, 2013).

In relation to quiescent HSV-1 infection in neurons, RNA-Seq has been used previously to compare the HSV-1 gene expression profile during productive and quiescent infection in neuronal and non-neuronal cells (Harkness et al., 2014). RNA-Seq has also been used to identify novel neuronal sub-types (Usoskin et al., 2015), which aided the discovery that LAT promoter activity differs depending on the neuronal sub-type HSV-1 infects (Cabrera et al., 2018). However, RNA-Seq has never been used to reveal any effect that the latency-associated ncRNAs expressed during HSV-1 latency might have on the human neuronal transcriptome.

In this chapter, RNA-Seq will be used to analyse whether the expression of HSV-1 latency-associated ncRNAs in differentiated SH-SY5Y cells leads to significant transcriptome changes in these cells. This work aims to reveal any novel functions of the latency-associated ncRNAs and more specifically to observe whether they elicit any

transcriptional changes on human neurons. The hypothesis of this study is that the HSV-1 latency-associated ncRNAs will affect the human neuronal transcriptome.

LAT has previously been shown to restrict lytic HSV-1 gene expression, as discussed in chapter 1.4.1 (Mador et al., 1998; Nicoll et al., 2016). Mador *et al.* used northern blot analysis after infection with a LAT-negative HSV-1 mutant to show that ICP0, ICP4 and ICP27 mRNA levels were greatly reduced following infection of a mouse neuroblastoma cell-line stably expressing LAT compared to a regular neuroblastoma cell-line (Mador et al., 1998). This work suggested a role for LAT in the suppression of HSV-1 immediate early gene mRNA levels. Nicoll *et al.* used reporter viruses expressing either firefly luciferase or beta-galactosidase driven by the CMV promoter to show that the expression of these reporter genes was significantly upregulated in LAT-negative recombinant viruses compared to those with LAT expression, during neuronal latency *in vivo* (Nicoll et al., 2016). This shows that LAT-mediated gene repression is not limited to the HSV-1 immediate-early genes shown by Mador *et al.* or even to HSV-1 genes, suggesting LAT may mediate a global gene silencing effect. It has been suggested that LAT mediates this repression by promoting the formation of facultative heterochromatin, associated with gene silencing (Cliffe et al., 2009). While the work to date focuses on LAT mediating its effects on the HSV-1 genome, it is possible that these functions of LAT extend to host genes too.

Additionally, much of the published data focuses on the role of LAT either in its entirety or focuses on specific regions of the primary LAT transcript (Carpenter et al., 2015). However, as described in the introduction (chapter 1 section 1.4.1), LAT encodes several ncRNAs. The primary transcript is 8.3kb long and is spliced into the 1.5 or 2kb major LAT intron. This major LAT intron is highly abundant and stable during HSV-1 latency. It differs in its structure from the primary transcript as it is a lariat rather than linear transcript. The implications of the abundance of this stable lariat ncRNA during HSV-1 is not fully understood and, as such, is a focus of this study. To address this question, the lentiviruses were designed to either encode the major LAT intron (LAT intron lentivirus) or include mutations that limit splicing of major LAT (NSLAT lentivirus) (chapter 3, figure 3.3).

The remaining exon is further processed into 6 HSV-1 microRNAs. A subset of these, along with miR-H6 which is transcribed in the LAT promotor region, are abundantly produced during HSV-1 latency (chapter 1, figure 1.2). As such, a further focus of this thesis is whether these HSV-1 miRNAs have any novel effects on cellular gene expression during latency. To assess the role of the latency-associated miRNAs, the

miR lentivirus, that encodes 5 HSV-1 miRNAs found to be abundant during HSV-1 latency (Umbach et al., 2009), will be utilised.

MiRNAs have been shown to regulate HSV-1 lytic gene expression. Usually this regulation occurs post-transcriptionally, where HSV-1 miRNAs with sequence complementary to lytic genes guide the RISC to target these genes for degradation (chapter 1.4.1). For example, HSV-1 miR-H2-3p has complementarity to ICP0 and mediates a reduction in ICP0 protein levels but no significant bearing on the transcript levels, suggesting the targeting of ICP0 occurs at the translational level (Umbach et al., 2008). Nevertheless, it has recently been established that miRNAs can mediate their regulatory effects at several stages of mRNA processing, including transcriptional, post-transcriptional and epigenetic modifications, as discussed in chapter 1.4.1 (Gosline et al., 2016).

Clearly the latency-associated ncRNAs have an impact on viral gene expression, with LAT able to mediate epigenetic modifications (Cliffe et al., 2009) while miRNAs can elicit transcriptional and post-transcriptional changes (Gosline et al., 2016; Umbach et al., 2008). It follows that host-gene transcription might be affected by the HSV-1 latency-associated ncRNAs. Additionally, it has been shown that ncRNAs expressed by other members of the Herpesviridae family can induce transcriptional changes in host cells. Marquitz *et al.* used RNA-Seq to show that the Epstein-Barr virus (EBV) BamH1 A rightward transcripts (BARTs), which encode nuclear ncRNAs and 44 miRNAs, regulates host transcription in latent EBV-infected human gastric carcinoma cells (Marquitz et al., 2015). It would be interesting to see if this is a conserved feature of the herpesviruses and, as such, whether the same applies for HSV-1 latency-associated ncRNAs.

Latency allows HSV-1 to persist in human neurons for the lifetime of the host, therefore it is important to understand what effect the virus may have on the physiology of latently infected neurons. This is especially important considering that there have been links suggested between latent and recurrent HSV-1 infection of neurons, especially those of the central nervous system (CNS), and the development of neurodegenerative disorders such as multiple sclerosis (MS) and Alzheimer's disease (Duarte et al., 2019). As discussed in chapter 1.4.4, numerous recent publications have shown a link between long-term HSV-1 infection and an increased risk of developing Alzheimer's disease. The link was first suggested in 1982 in a paper by Ball, showing that the same regions of the brain affected during HSV-1 encephalitis were also impacted by Alzheimer's disease (Ball, 1982). Numerous studies have emerged since then, adding to the data showing a link between HSV-1 infection or seropositivity and the development of Alzheimer's

disease (Lövheim, Gilthorpe, Johansson et al., 2015; Lövheim, Gilthorpe, Adolfsson et al., 2015; Mangold & Szpara, 2019; Piacentini et al., 2014; Wozniak et al., 2009). This highlights the importance of studying the interaction between HSV-1 and human neurons.

The aim of this chapter is to examine whether and how the human neuronal transcriptome changes in response to expression of the HSV-1 latency-associated ncRNAs.

4.2 Results

4.2.1 Lentivirus expression of HSV-1 latency-associated ncRNAs in human neurons for RNA-Seq analysis

To examine the effects of the latency-associated ncRNAs on the neuronal transcriptome, the lentiviruses described in chapter 3, designed to express the HSV-1 latency-associated ncRNAs were utilised. The differentiated SH-SY5Y cells were transduced with 56 functional units (FU) of each previously described lentivirus (section 3.2.2). This includes the eGFP expressing lentivirus; the LAT intron lentivirus which transcribes the first 3.1kb of LAT and produces the major LAT intron; the NSLAT lentivirus that transcribes the first 3.1kb of LAT but has mutations in splice sites to block production of the major LAT intron; and the miRNA lentivirus that encodes 5 HSV-1 miRNAs (miR-H2,3,4,6 and 7) abundantly produced during latency (Umbach et al., 2009). RNA was extracted from duplicate wells 14 days post transduction (dpt), with 5 biological repeats per lentivirus as highlighted schematically (figure 4.1A).

RNA was extracted from cells 14 days post transduction (dpt) because peak expression of LAT was found to be 11dpt for the LAT intron lentivirus transduction (chapter 3, figure 3.8A) and 14dpt allows LAT expression to reach this peak but also give time for the ncRNA to have any effect on the neuronal transcriptome up to a timepoint where LAT is still known to be expressed.

Before performing RNA-Seq on the extracted RNA, the samples were tested for LAT transcription. qRT-PCR was performed on the RNA with primers directed to the major LAT sequence (chapter 3, figure 3.6D) and beta actin as a positive control.

The results show that the average LAT copy numbers were 90 and 18 per 1000 beta actin copies across the 5 biological repeats for LAT intron lentivirus and NSLAT lentivirus transduced cells respectively (figure 4.1B). This confirms that LAT was transcribed in the LAT intron and NSLAT lentivirus transduced cells and not in untransduced or eGFP

lentivirus transduced differentiated SH-SY5Y cells, as expected (figure 4.1B). However, the LAT copy numbers were relatively low. Regardless, LAT is produced following transduction with the LAT lentiviruses. Additionally, LAT abundance is higher in the LAT intron lentivirus transduced cells than NSLAT lentivirus transduced cells, which is expected given that the LAT intron lentivirus produces the major LAT intron, which is highly stable and abundant (Farrell et al., 1991) whereas the NSLAT lentivirus should not produce this intron (Alvira et al., 1999)(chapter 3.2.3). Although confirmatory stem-loop PCR was not performed for the miRNAs on these samples, the HSV-1 miRNAs have been shown to be produced in this system previously (chapter 3, figure 3.10). Additionally, the LAT PCR on these samples confirms that LAT is expressed by the LAT lentiviruses, which are transduced by the same standardised protocol and in parallel to the miR lentivirus (methods 6.2.2.4 and 6.2.2.5 and chapter 3.2.3).

Accordingly, RNA-Seq was performed on these samples to examine the effects of the latency-associated ncRNAs on the human neuronal transcriptome in these lentivirus-transduced samples that express the latency-associated ncRNAs, albeit in slightly low abundance.

Figure 4.1 Lentivirus transduction of differentiated SH-SY5Y cells in preparation for RNA-Seq

(A) Schematic representation of the set-up for the lentivirus RNA-Seq experiment whereby differentiated SH-SY5Y cells were either refed (untransduced) or transduced with 56 functional units (FU) of eGFP lentivirus, LAT intron lentivirus, miR lentivirus or NSLAT lentivirus. Total RNA was harvested (and poly A selected before RNA-Seq was performed) 14dpt with 5 replicates per lentivirus, with RNA being pooled from 2 wells per replicate.

(B) qRT-PCR was performed to LAT sequence and beta actin on total RNA extracted from the cells set-up as described in **(A)** (except the miRNA lentivirus transduced cells). The graph shows the average LAT copies per 1000 beta actin copies from 5 biological repeats and 3 technical repeats ± the standard deviation, as shown previously in figure **3.7F**.

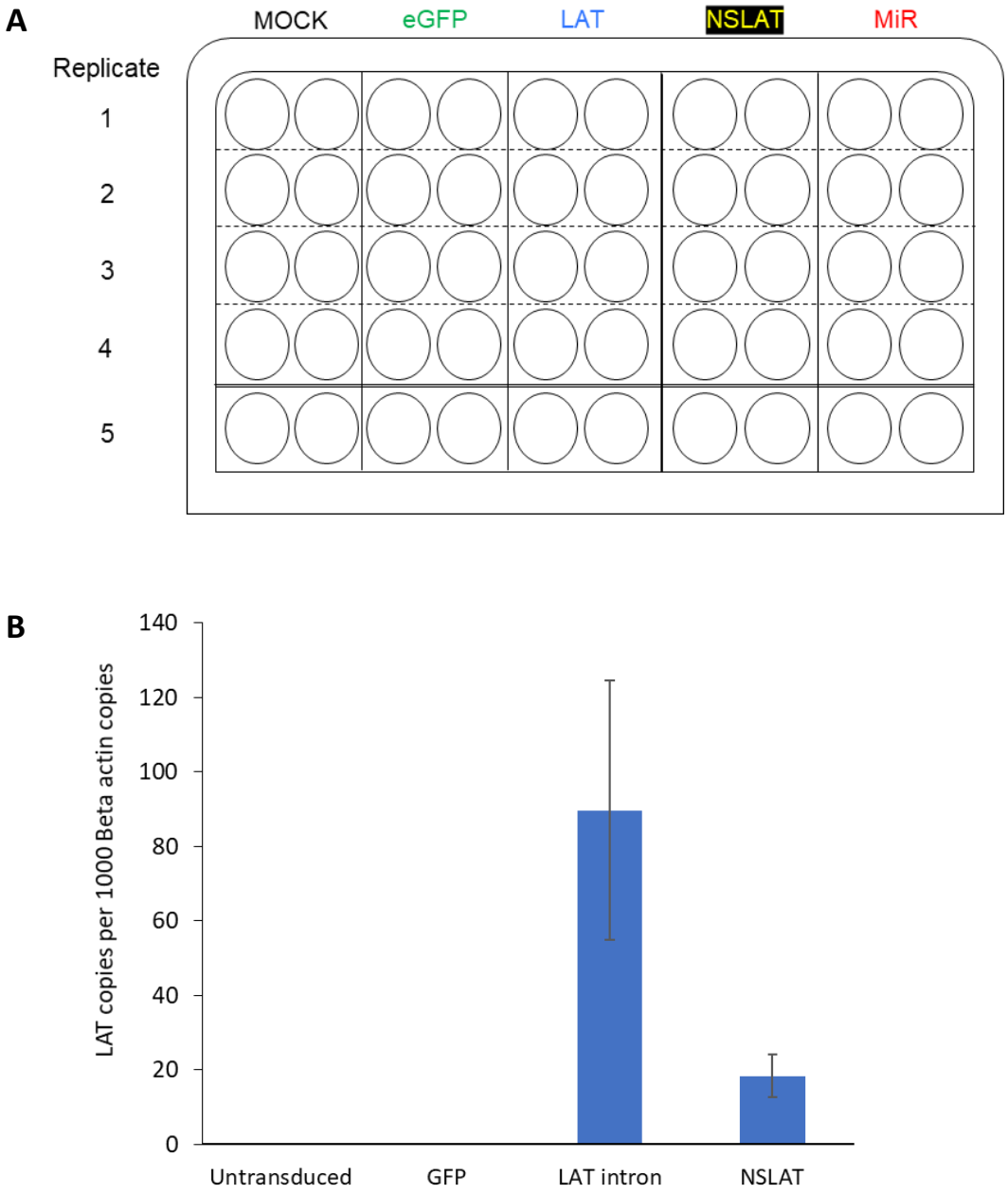


Figure 4.1 Lentivirus transduction of differentiated SH-SY5Y cells in preparation for RNA-Seq

4.2.2 Lentivirus expression of HSV-1 latency associated ncRNAs in human neurons leads to significant host-gene transcriptional changes

Illumina TruSeq stranded mRNA library preparation was carried out to produce adapter-bound cDNA from total RNA harvested from the lentivirus-transduced samples represented schematically in 4.1A, to be able to perform strand-specific RNA-Seq on these samples. The steps to achieve this, including poly-A selection of the total RNA harvested, are described in figure 4.2A and the methods section (chapter 6.2.6) and were carried out by Dr Martin Fritzsche and Mr Ryan Mate, Analytical and Biological Sciences (ABS), NIBSC.

Following RNA-Seq of the samples on the Illumina NextSeq 500 sequencer, bioinformatics analysis was carried out on the output data by Dr Thomas Bleazard, ABS, NIBSC. The bioinformatics steps taken are described in figure 4.2B and the methods section (chapter 6.2.7.1). The bioinformatics analysis provided a list of all the differentially expressed genes, calculated by DESEQ2 analysis using R software (The R Core Team, 2017). These genes could then be filtered to only include those significantly differentially expressed, where p-value < 0.05. These DEGs were then further narrowed down to only include significant genes with a false discovery rate (FDR) of 5% or lower, as calculated by the Benjamini-Hochberg adjustment. This provides a new statistical value, known as the p adjusted value (padj) which reports significance at ≤ 0.05 whilst taking into account the large number of genes being analysed (Benjamini et al., 2009; Benjamini & Hochberg, 1995). All genes that fit these criteria were accepted as differentially expressed genes (DEGs).

The bioinformatics analysis (figure 4.2B) revealed 7301 (3885 upregulated, 3416 downregulated) DEGs in response to transduction with any of the lentiviruses, when each are compared to mock transduced differentiated SH-SY5Y cells (figures 4.1A, 4.3 and 4.5A). However, this included the genes differentially expressed in cells transduced with the control eGFP lentivirus (Figure 4.3A and B). The Venn diagrams produced show the number of DEGs attributable to each lentivirus transduction and how many of these overlap between the different conditions. Although some of the DEGs do overlap between the control eGFP lentivirus and the LAT lentiviruses and miR lentivirus, most do not (figure 4.3). The differences in gene changes observed between the LAT intron lentivirus or MiR lentivirus and the control eGFP lentivirus are highlighted by the pairwise heat maps and dendrogram produced from this data, by Thomas Bleazard, ABS, NIBSC (figure 4.4). The heat map compares the similarity in gene expression across all the genes for the 2 conditions and represents this as a colour on the scale, where red is

Figure 4.2 Flow charts showing the steps involved in performing and analysing the RNA-Sequencing experiments

(A) Flow diagram outlining the experimental steps taken from RNA preparation to RNA-Seq with steps in blue boxes and details of each step explained in bullet points next to each blue box.

(B) Flow diagram outlining the bioinformatics analysis steps taken post RNA-Seq with steps in blue boxes and details of each step explained in bullet points next to each blue box. The information in square brackets applies to the in1382 RNA-Seq experiment but not the lentivirus RNA-Seq experiment.

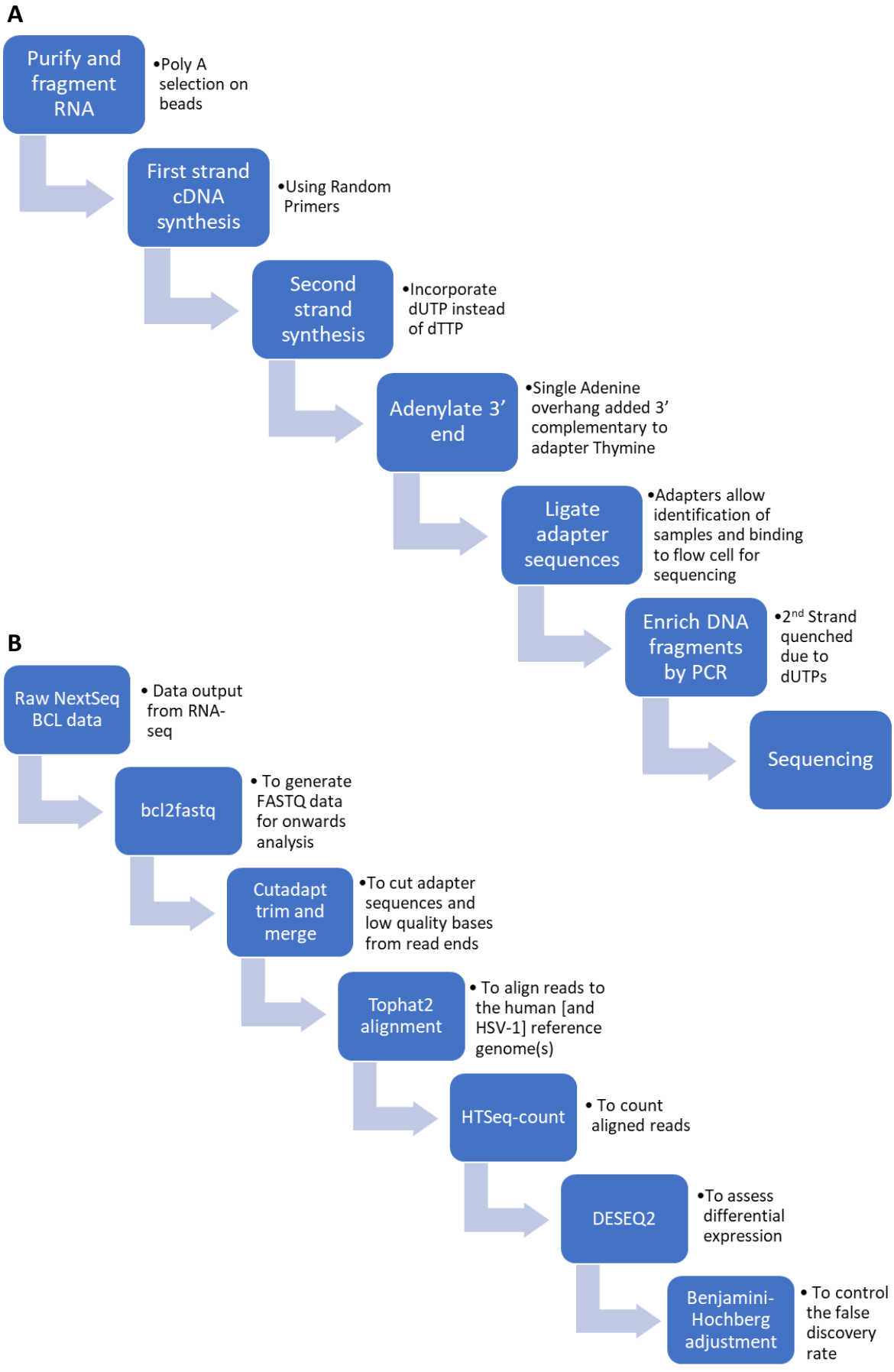


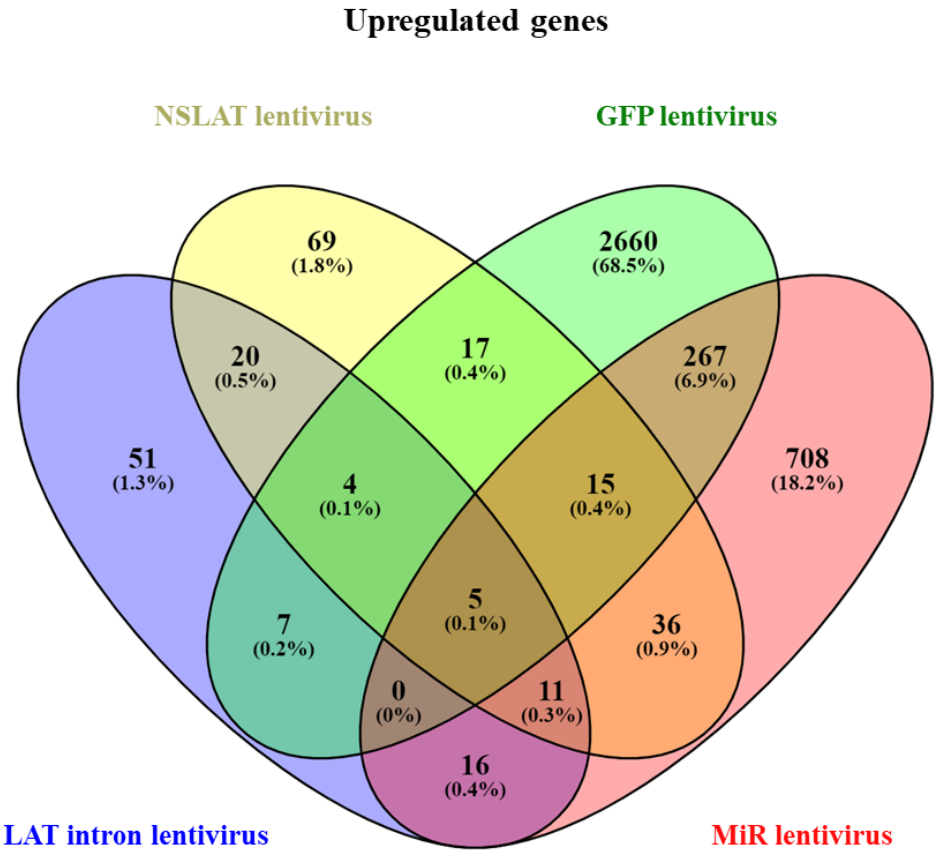
Figure 4.2 Flow charts showing the steps involved in performing and analysing the RNA-sequencing experiments

Figure 4.3 The number of significantly differentially expressed genes attributable to each lentivirus transduction following RNA-Seq

(A) Venn diagram to show the number of significantly upregulated genes following transduction with each lentivirus in differentiated SH-SY5Y cells as set up in **4.1A** and where gene changes overlap between the different transductions, diagram generated using Venny 2.1 (Oliveros, 2015).

(B) Venn diagram to show the number of significantly downregulated genes following transduction with each lentivirus in differentiated SH-SY5Y cells as set up in **4.1A** and where gene changes overlap between the different transductions, diagram generated using Venny 2.1 (Oliveros, 2015).

A



B

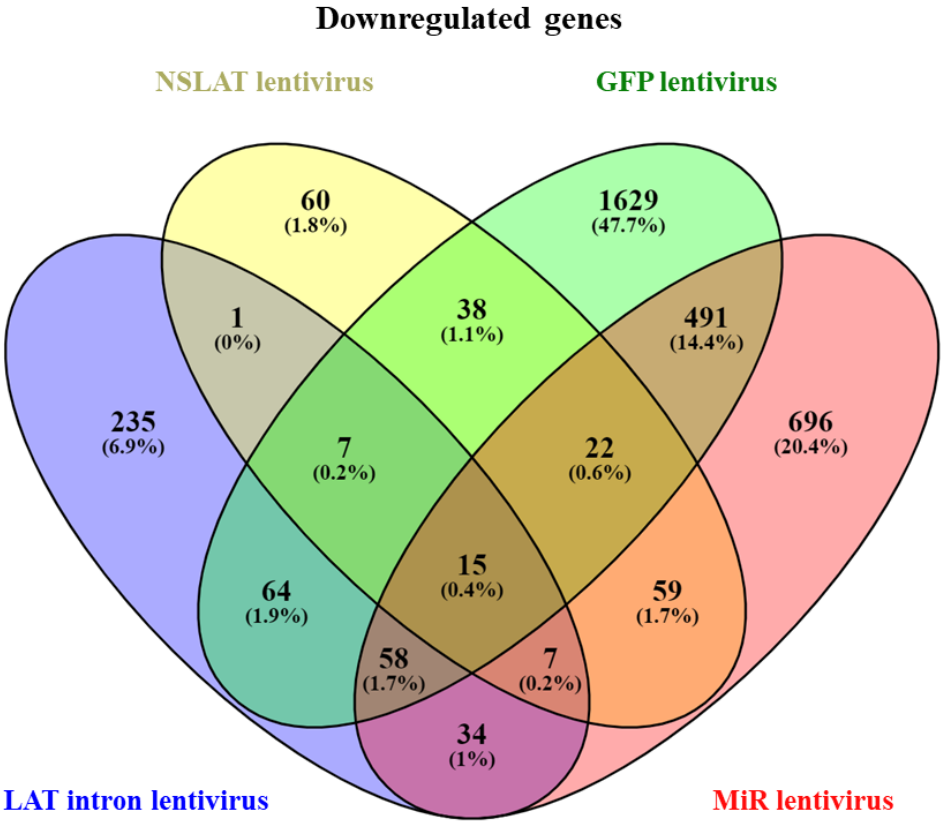


Figure 4.3 The number of significantly differentially expressed genes attributable to each lentivirus transduction following RNA-Seq

Figure 4.4 Heat maps comparing differential gene expression in eGFP lentivirus transduced cells to ncRNA expressing lentivirus transduced cells

(A) Heat map comparing differential expression between eGFP-lentivirus transduced differentiated SH-SY5Y cells and LAT intron lentivirus transduced differentiated SH-SY5Y cells. Whereby the colour represents the Euclidean distance comparing the normalised read count across the entire transcriptome for each of the sample pairs shown. The colour key shows how the Euclidean distance value (representing how different 2 samples are, where the bigger the Euclidean distance, the larger the difference) changes from red, where Euclidean distance is low and the two samples are most alike to yellow, where the Euclidean distance is high, and the samples are most different in terms of gene expression to one another. The histogram (light green line) uses the same x-axis of Euclidean distance, with “count” or number of sample comparisons with each specific Euclidean distance on the y-axis.

(B) Heat map comparing differential expression between eGFP-lentivirus transduced SH-SY5Y cells and MiR lentivirus transduced SH-SY5Y cells. Whereby the colour represents the Euclidean distance between the 2 samples (how much they differ) when comparing the normalised read count across the entire transcriptome for each of the sample pairs shown. The colour key shows how the Euclidean distance value changes with colour, from red, where Euclidean distance is low and the two samples are most alike to yellow, where the Euclidean distance is high, and the samples are most different in terms of gene expression to one another. The histogram (light green line) uses the same x-axis of Euclidean distance, with “count” or number of sample comparisons with each specific Euclidean distance on the y-axis.

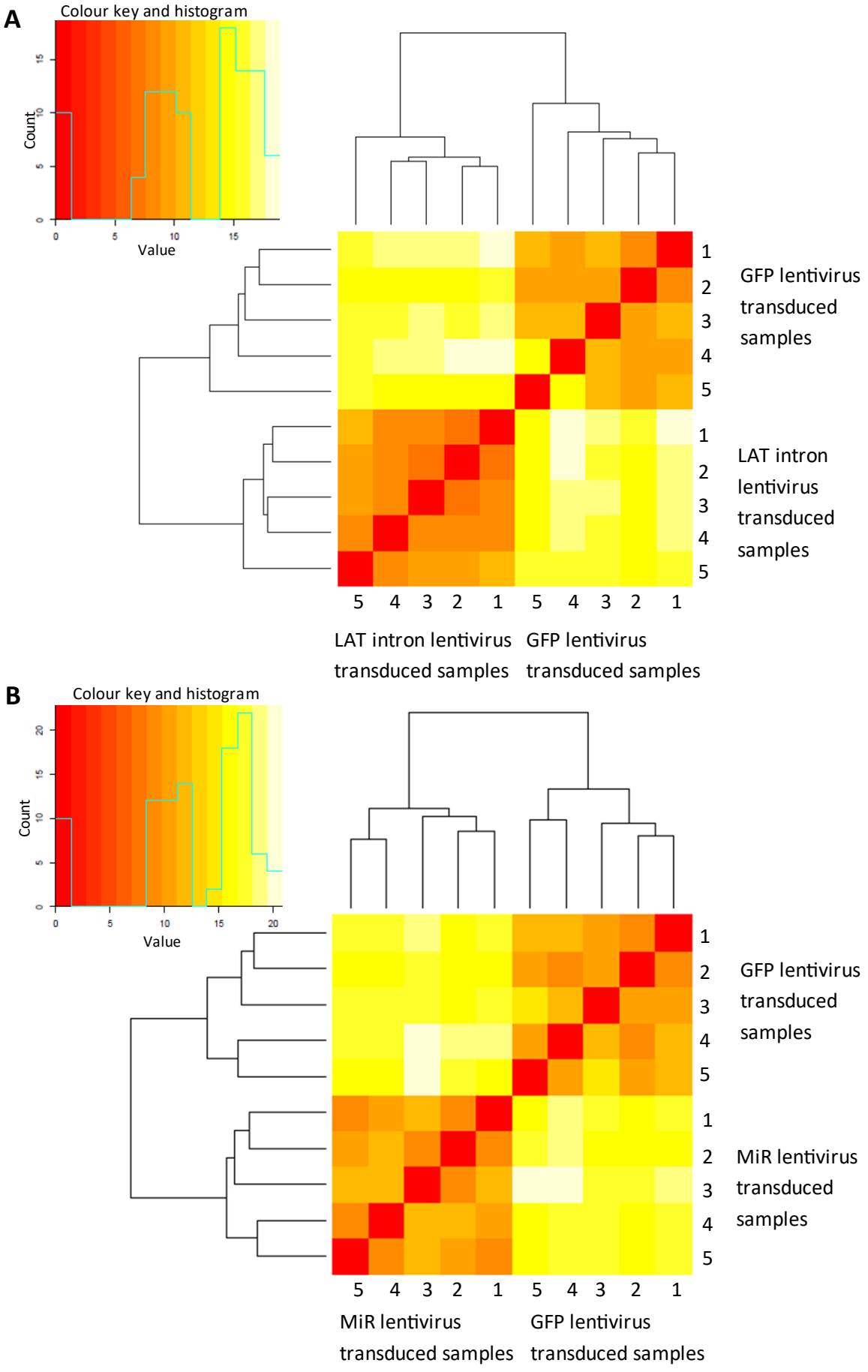


Figure 4.4 Heat maps comparing differential gene expression in eGFP lentivirus transduced cells to ncRNA expressing lentivirus transduced cells

similar and yellow is distant. The accompanying dendrogram shows how the samples cluster in terms of gene expression (figure 4.4).

It is clear from the results of the heat map and dendrogram (figure 4.4A) that the gene expression following transduction with the LAT intron lentivirus differs greatly from gene expression following transduction with the control eGFP lentivirus, as expected. This difference in gene expression is indicated by the yellow squares comparing samples from LAT lentivirus transduced differentiated SH-SY5Y cells and eGFP lentivirus transduced differentiated SH-SY5Y cells (figure 4.4A). Similarly, the gene expression following transduction with miR lentivirus differs greatly from gene expression following transduction with the control eGFP lentivirus (figure 4.4B). There is also clear clustering among samples with the same lentivirus condition, as is expected, as indicated by the dendrogram branching and orange and red squares.

There is a noticeable difference between the gene changes observed in the control eGFP lentivirus and those expressing the HSV-1 latency associated ncRNAs. Therefore, in order to identify the genes significantly differentially expressed directly as a result of the miRNAs and LAT, the gene changes that occur during eGFP lentivirus transduction need to be removed. In this way the eGFP lentivirus acts as a control for non-specific lentivirus-induced transcriptional changes, as it is designed to. The Venn diagrams produced in Venny 2.1 (Oliveros, 2015) as shown in figure 4.3 allows for this elimination of eGFP lentivirus changed genes. Having removed these genes that change as a result of the eGFP lentivirus transduction – whether or not these gene changes occur elsewhere – there are 2003 DEGs (911 upregulated, 1092 downregulated) as a result of the HSV-1 latency-associated ncRNAs delivered by the lentiviruses, including the first 3.1kb of LAT (with or without splicing capabilities) and the latency-associated miRNAs (miR-H2-4 and miR-H6-7) (figures 4.3 and 4.5B). Using the lentivirus constructs it is possible to focus even further and only consider the LAT changes that result from production of the major LAT intron and therefore exclude any gene changes resulting from transduction with the NSLAT lentivirus expressing the non-splicing LAT RNA construct and lacking major LAT intron production. In this way, the NSLAT lentivirus acts as a negative control.

Excluding the genes that were significantly differentially expressed during either eGFP or NSLAT lentivirus transduction of differentiated SH-SY5Y revealed 1740 DEGs (775 upregulated, 965 downregulated) in response to LAT intron lentivirus delivery of LAT, with major LAT splicing phenotype intact, or miR lentivirus transduction or those that overlap in both transductions (figures 4.3 and 4.5C).

Figure 4.5 The number of significantly differentially expressed genes attributable to each lentivirus transduction following RNA-Seq

(A) Flow chart to represent the number of gene changes that occur following transduction with each of the engineered lentiviruses (eGFP, LAT intron, NSLAT and miR lentivirus) compared to mock differentiated SH-SY5Y cells, including how many genes are upregulated and how many are downregulated.

(B) Flow chart to represent the number of gene changes that occur following transduction with the LAT intron lentivirus, NSLAT lentivirus or miR lentivirus compared to mock differentiated SH-SY5Y cells, including how many genes are upregulated and how many are downregulated.

(C) Flow chart to represent the number of gene changes that occur following transduction with only the LAT intron lentivirus or miR lentivirus compared to mock differentiated SH-SY5Y cells, including how many genes are upregulated and how many are downregulated.

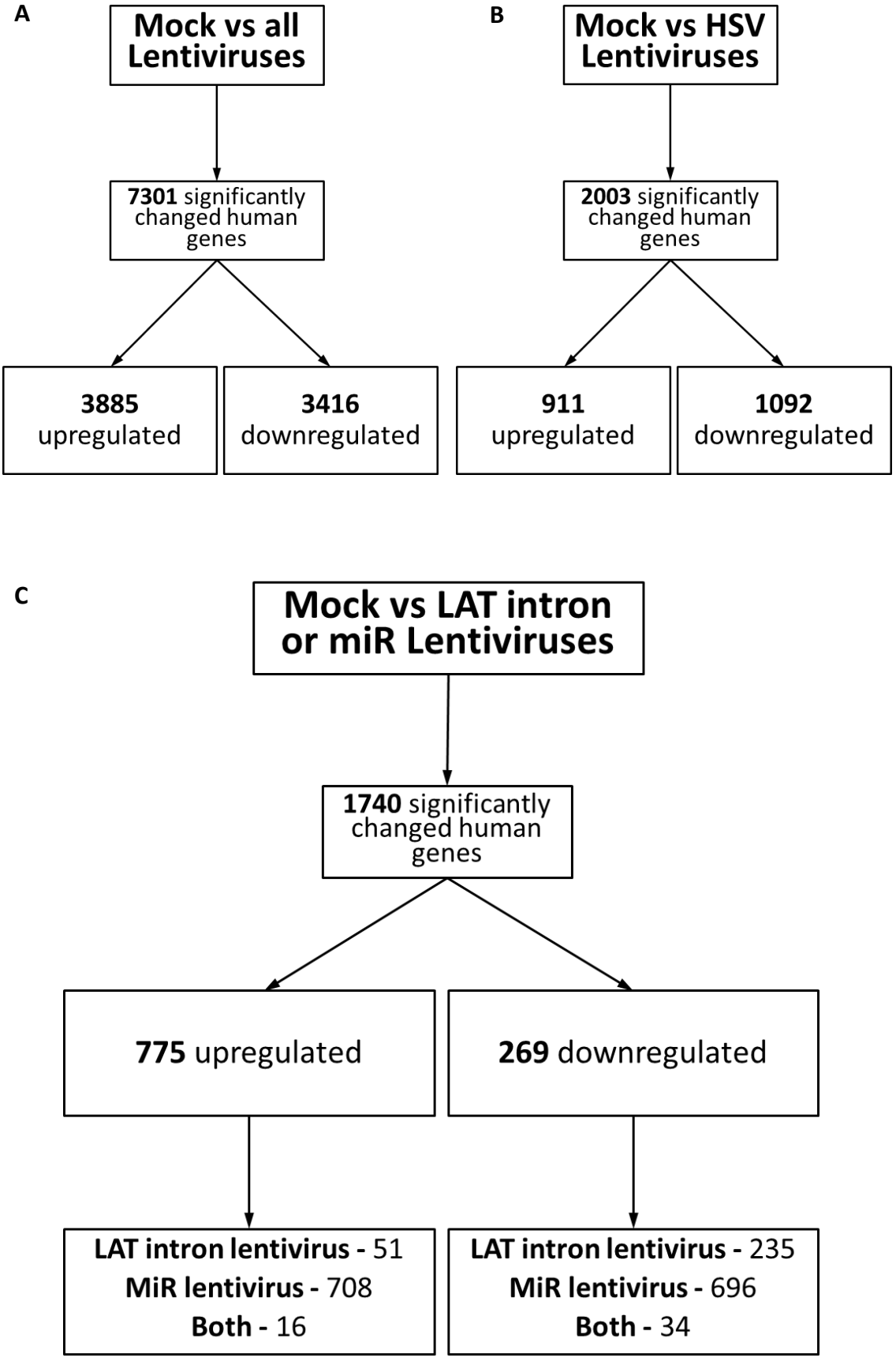


Figure 4.5 The number of significantly differentially expressed genes attributable to each lentivirus transduction following RNA-Seq

From the above data it is evident that the latency-associated ncRNAs affect the neuronal transcriptome, but further analysis is needed to determine the physiological context of these changes.

4.2.3 Functions and fold-changes of lentivirus DEGs

To start to consider the physiological impact that these 1500 plus significant gene changes have within the neuron, it is important to examine the functions of the genes that are differentially expressed and any biological pathways they may be associated with. To achieve this, further analysis was performed on the gene lists provided by the bioinformatics analysis described in figure 4.2B, as highlighted in figure 4.6.

4.2.3.1 The LAT intron and NSLAT lentiviruses DEGs

Initially, the genes that were significantly changed following the LAT intron lentivirus transduction and no other lentivirus transduction were assessed. This focuses on the effect of the major LAT intron on the neuronal transcriptome, as it does not include the genes also changed following transduction with the NSLAT lentivirus where the major LAT sequence is encoded but without the ability for the major LAT intron to be spliced.

The 51 and 235 DEGs following LAT intron lentivirus transduction (figure 4.3) were filtered for significance and the top 5 upregulated and downregulated DEGs according to fold-change were searched for function in NCBI gene search (figure 4.6 and tables 4.1 and 4.2).

The fold-changes following LAT intron lentivirus delivery of major LAT to the SH-SY5Y cells were low with 1.31 upregulation and 1.51 downregulation being the largest fold-changes observed, for the ATP binding cassette subfamily A member 5 (ABCA5) and hepatocyte growth factor activator (HGFAC)/Rhophilin Rho GTPase binding protein 1 (RHPHN1) genes respectively (tables 4.1 and 4.2).

The functions of these top 5 upregulated and downregulated genes are highlighted in tables 4.1 and 4.2. Among these top 5 there are 3 genes upregulated that are involved in transport of molecules across membranes (ABCA5, RAB27B, member RAS oncogene (RAB27B) and solute carrier family 8 member A1 (SLC8A1))(table 4.1).

Figure 4.6 Data analysis performed on RNA-Seq results, following bioinformatics analysis

A flow chart to represent the steps taken to perform further analyses on the RNA-Seq data provided following the bioinformatics analysis steps outlined in figure **4.2B**. The blue boxes show the steps taken with details of the step explained in the adjacent bullet point. The blue dotted brackets identify optional alternate steps depending on the data being analysed.

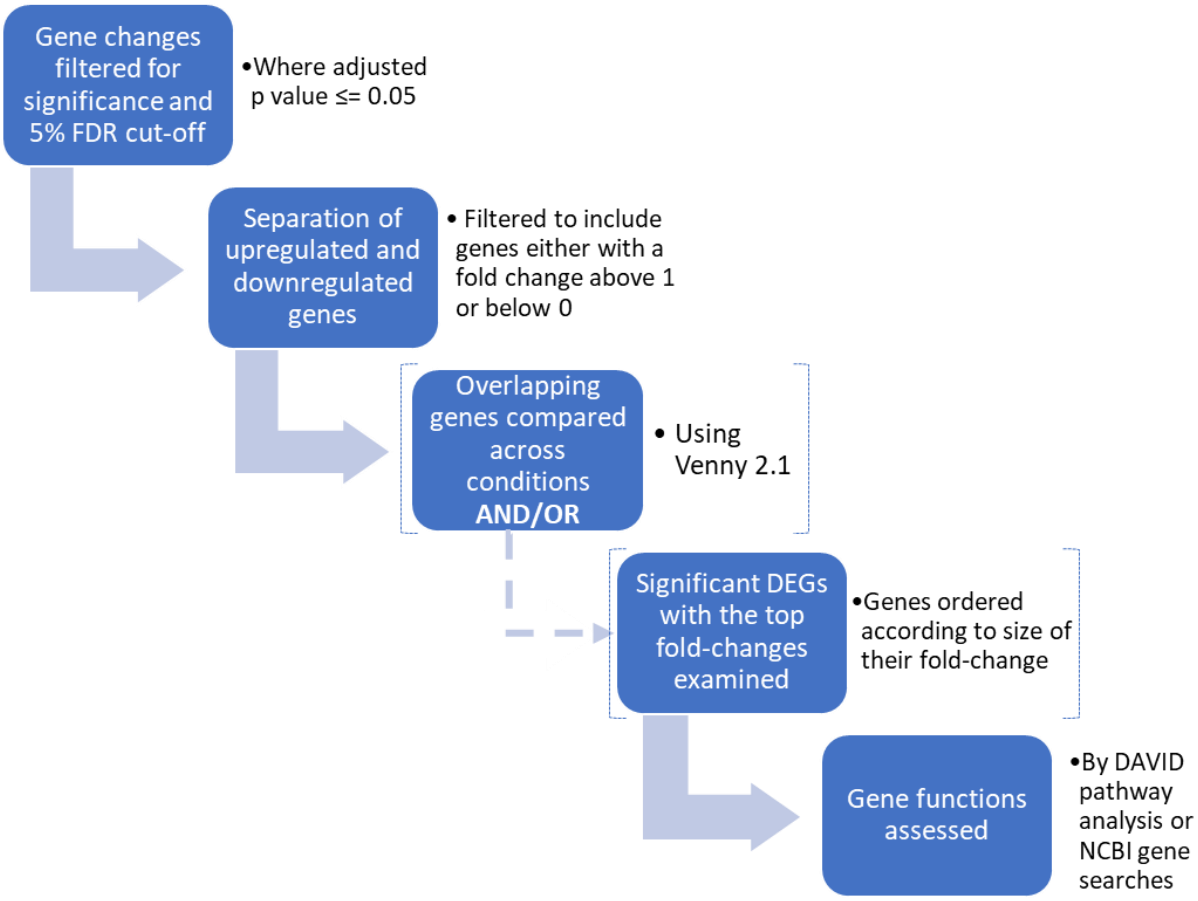


Figure 4.6 Data analysis performed on RNA-seq results, following bioinformatics analysis

Table 4.1 The top 5 significantly upregulated genes exclusively following LAT intron lentivirus transduction in differentiated SH-SY5Y cells

The table includes the top 5 significantly upregulated genes, according to fold change, found by RNA-Seq, exclusively following LAT intron lentivirus transduction of differentiated SH-SY5Y cells compared to untransduced differentiated SH-SY5Y cells, out of the DEGs found following the lentivirus transductions, as indicated in **figure 4.3**. The table shows the fold change, p-adjusted value following Benjamini-Hochberg test and functions of the genes, as found in (NCBI, 2016) unless otherwise indicated.

Table 4.1 The top 5 significantly upregulated genes exclusively following LAT intron lentivirus transduction in differentiated SH-SY5Y cells

Gene	Fold change (2dp)	P adjusted Value (3 sf)	Known protein function
ABCA5	1.31	0.00223	ATP binding cassette subfamily A member 5: • Member of the ATP-binding cassette (ABC) transporters. • Transport various molecules across extra- and intracellular membranes • Promotes cholesterol efflux in the brain (Fu et al. 2015)
RAB27B	1.30	1.26E-08	RAB27B member RAS oncogene family: • Prenylated, membrane-bound proteins involved in vesicular fusion and trafficking
TSHZ2	1.27	0.00411	Teashirt zinc finger homeobox 2: • Transcription factor – transcriptional repressor • Putative role in the development of certain cancers including breast cancer
BMPR1B	1.25	0.00122	Bone morphogenetic protein receptor type 1B: • A member of the bone morphogenetic protein (BMP) receptor family of transmembrane serine/threonine kinases. • Bind BMPs, which are involved in endochondral bone formation
SLC8A1	1.24	9.61E-05	Solute carrier family 8 member A1: • A sodium-calcium exchanger • Aids in returning cardiac myocytes to a resting state

Table 4.2 The top 5 significantly downregulated genes exclusively following LAT intron lentivirus transduction in differentiated SH-SY5Y cells

The table includes the top 5 significantly downregulated genes, according to fold change, found by RNA-Seq, exclusively following LAT intron lentivirus transduction of differentiated SH-SY5Y cells compared to untransduced differentiated SH-SY5Y cells, out of the DEGs found following the lentivirus transductions, as indicated in **figure 4.3**. The table shows the fold change, p-adjusted value following Benjamini-Hochberg test and functions of the genes, as found in (NCBI, 2016) unless otherwise indicated.

Table 4.2 The top 5 significantly downregulated genes exclusively following LAT intron lentivirus transduction in differentiated SH-SY5Y cells

Gene	Fold change (2dp)	P adjusted Value (3 sf)	Known protein function
HGFAC	-1.51	0.0100	Hepatocyte growth factor (HGF) activator: • Endoproteolytic processing activates it into a heterodimeric serine protease that coverts HGF into its active form
RHPN1	-1.51	9.09E-05	Rhophilin Rho GTPase binding protein 1: • No known protein function • Expressed in the brain
COL16A1	-1.48	0.000833	Collagen type XVI alpha1 chain: • A member of the FACIT (fibril-associated collagens with interrupted helices) collagen family S
SCNN1D	-1.42	0.00563	Sodium channel epithelial 1 subunit delta: • A subunit of the epithelial sodium channel (ENaC) that allows the flow of Na+ ions across high resistance epithelia to maintain body salt and water
AMH	-1.39	0.00131	Anti-Mullerian hormone: • A ligand that binds various TGF-beta receptors, which leads to the recruitment and activation of the SMAD family of transcription factors that regulate gene expression • Once proteolytically processed the resulting complex binds to anti-Mullerian hormone receptor type 2 and this causes the regression of Mullerian ducts in the male embryo that would otherwise differentiate into the uterus and fallopian tubes

It would be time-consuming and provide an overwhelming amount of information to search the gene functions of all 286 DEGs following LAT intron lentivirus transduction. Additionally, focusing on the top fold-changed genes may not be particularly relevant here given that the fold-changes are all relatively low. Furthermore, there are many genes that although the fold-change is low they have extremely low p-values and therefore are significantly changed following lentivirus delivery of the major LAT intron (tables 4.1 and 4.2 and appendix 1, sheets 2-7). Instead of looking at each gene function individually, pathway analysis can be performed on all the significantly upregulated or downregulated genes resulting from transduction with the LAT intron lentivirus by using the Database for Annotation, Visualization and Integration Discovery (DAVID) analysis software. This analysis allows testing of whether these DEGs are associated with any common cellular pathways, as defined by gene ontology (GO) biological processes or the Kyoto Encyclopaedia of Genes and Genomes (KEGG). GO is a framework used to organise biological knowledge of genes in a consistent way across genomic resources. Within GO, to be associated with a biological process a gene must either have molecular activity that plays an integral role within said process, regulate the biological process or be part of a separate biological program that is required for the process to occur (Dessimoz & Škunca, 2017). KEGG is a database of gene functions in terms of networks of genes and how they interact, whereby KEGG pathway analysis identifies the biochemical pathways, including metabolic and regulatory pathways, that genes are involved in (Kanehisa & Goto, 2000).

GO analysis and KEGG pathway analysis were performed (methods chapter 6.2.7.2) using DAVID software (Huang et al., 2009a, 2009b) on the 51 genes significantly upregulated following only the LAT intron lentivirus transduction and the 235 genes significantly downregulated following only the LAT intron lentivirus transduction (appendix 1, sheet 8). The processes and pathways found to be associated with the DEGs by DAVID analysis were filtered for significance (p-value < 0.05) whereby a fisher exact statistics method is used to calculate p-values using the DAVID software (Huang et al., 2009a, 2009b). These pathways were then ordered in terms of the number of genes associated with the process or pathway. The graphs indicate the top 10 (or less where there were less than 10 in total) processes or pathways plotted against the -log10(p-value) (figures 4.7 and 4.8). For the full list of genes associated with each process or pathway, see appendix 2, sheets 2-4.

For the upregulated genes associated with major LAT intron expression, there were no known KEGG pathway associations. However, between 11 and 16 of the genes were found to be associated with various GO biological processes. Most of the GO processes

associated with the upregulated genes seem to be linked to the localisation and transport of molecules, for example cellular localisation, protein localisation and protein transport (figure 4.7).

There were certain GO biological processes such as cellular localisation and macromolecule localisation that were represented in both up and downregulated genes following LAT intron expression (figure 4.7 and 4.8A). Nevertheless, there are more downregulated genes associated with these processes than upregulated genes. For example, for cellular localisation there are 44 known gene associations among the 235 downregulated genes and 16 among the 51 upregulated genes. Nevertheless, this represents a larger percentage of the total upregulated genes than total downregulated genes. For a full table of the genes associated with each process and the percentage of the total mapped genes, see appendix 2, sheets 2-4. One GO biological process that was exclusively associated with the downregulated genes, was nervous system development, with 43 of the 235 downregulated DEGs associated with this GO biological process. The downregulated genes also had associations with known KEGG pathways, such as the vascular endothelial growth factor (VEGF) signalling pathway. Nevertheless, none of these pathways had more than 5 genes associated with them.

Figure 4.7 GO biological processes analysis on significantly upregulated genes found by RNA-Seq exclusively following LAT intron lentivirus transduction

Bar graph showing the gene ontology (GO) biological processes following analysis using DAVID software (Huang et al., 2009a, 2009b) on the 51 significantly upregulated genes found by RNA-Seq following LAT intron lentivirus transduction of differentiated SH-SY5Y cells as indicated in figure **4.3A** and listed in **appendix 1**. The graph shows the top 10 significantly associated GO biological processes in terms of number of differentially expressed genes (DEGs), which are indicated in red. The blue bars show the -log10(p-value) following the fisher exact statistics method used by the DAVID software to calculate the p-value, whereby the higher the -log10(p-value), the smaller the p-value, representing significance of association with the DEGs input.

See **appendix 2** for full gene lists associated with each GO biological process.

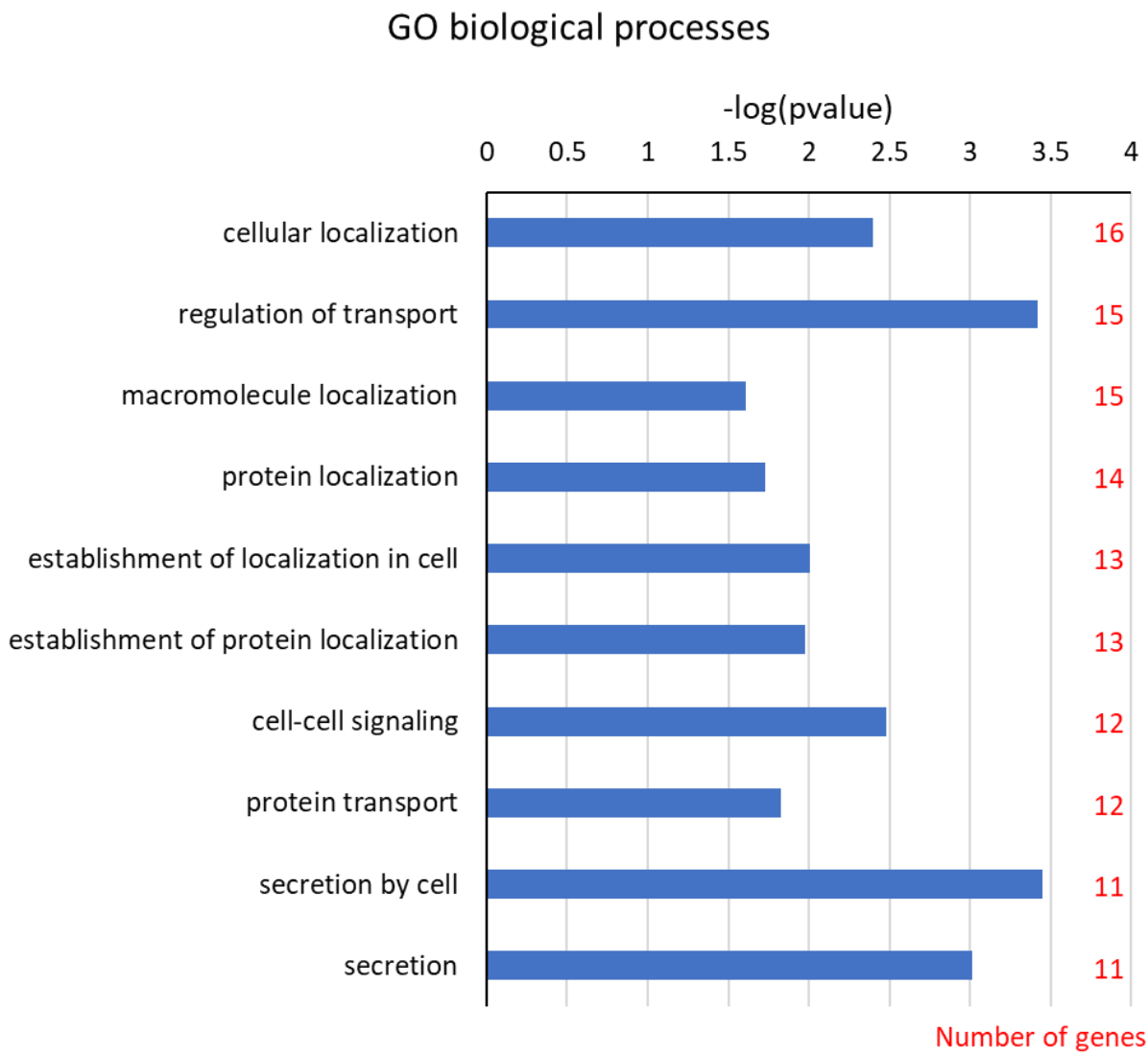


Figure 4.7 GO biological processes analysis on significantly upregulated genes found by RNA-Seq exclusively following LAT intron lentivirus transduction

Figure 4.8 GO biological processes and KEGG pathway analysis on significantly downregulated genes found by RNA-Seq exclusively following LAT intron lentivirus transduction

(A) Bar graph showing the gene ontology (GO) biological processes following analysis using DAVID software (Huang et al., 2009a, 2009b) on the 235 significantly downregulated genes found by RNA-Seq following LAT intron lentivirus transduction of differentiated SH-SY5Y cells as indicated in figure **4.3B** and listed in **appendix 1**. The graph shows the top 10 significantly associated GO biological processes in terms of number of differentially expressed genes (DEGs), which are indicated in red. The blue bars show the -log10(p-value) following the fisher exact statistics method used by the DAVID software to calculate the p-value for each process.

(B) Bar graph to show the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways found by DAVID analysis (Huang et al., 2009a, 2009b) to be significantly associated with the 235 significantly downregulated genes following LAT intron lentivirus transduction of differentiated SH-SY5Y cells as indicated in figure **4.3B** and listed in **appendix 1**. The number of DEGs involved are indicated in red. The blue bars show the -log10(p-value) following the fisher exact statistics used in the DAVID analysis software for each pathway.

See **appendix 2** for full gene lists associated with each GO biological process/KEGG pathway.

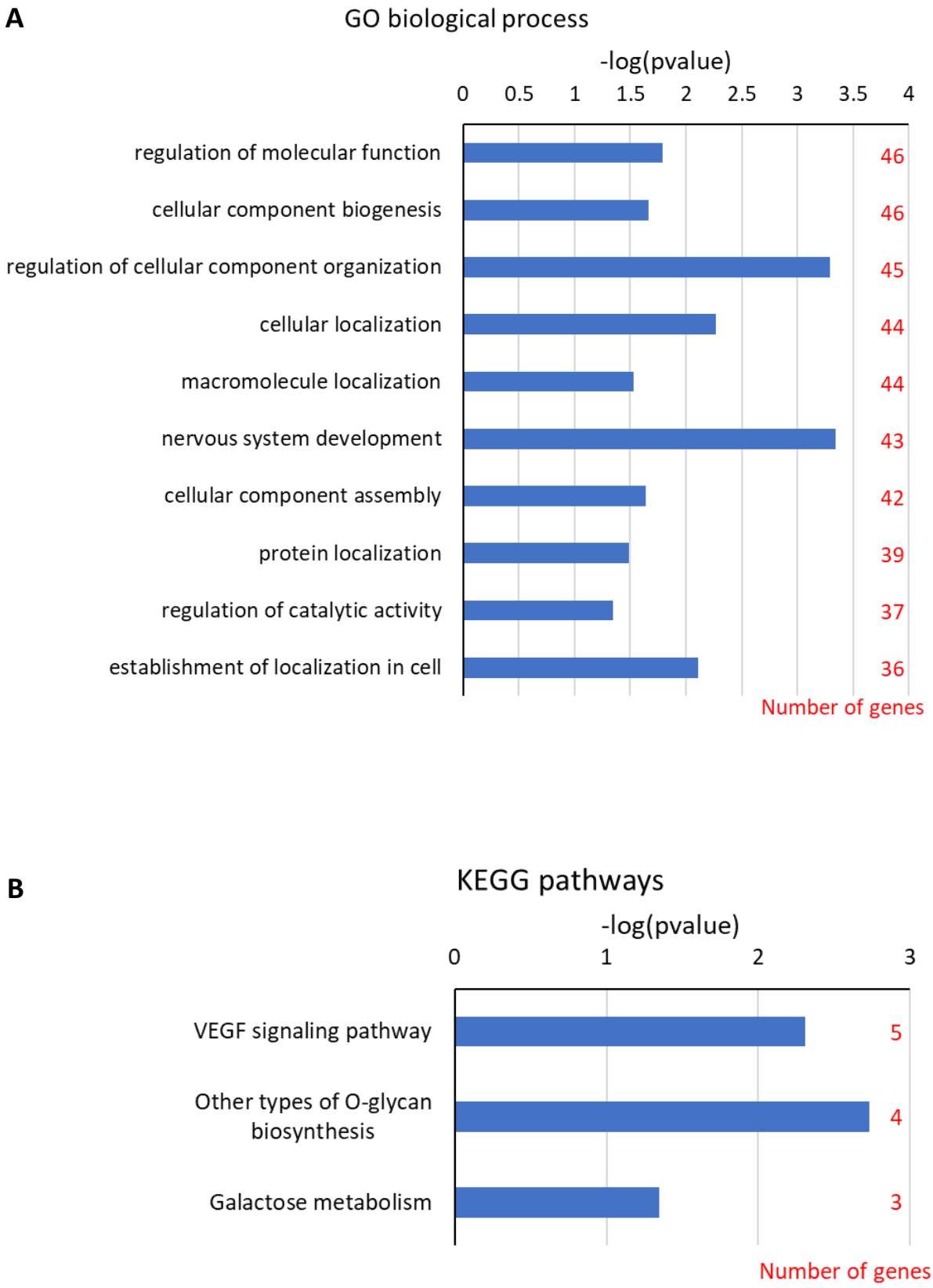


Figure 4.8 GO biological processes and KEGG pathway analysis on significantly downregulated genes found by RNA-seq exclusively following LAT intron lentivirus transduction

4.2.3.2 The MiR lentivirus DEGs

The same analyses were performed for the DEGs observed following transduction with MiRNA lentivirus. Initially the significantly changed genes were ordered according to fold-change as previously described (figure 4.6) and the top 5 upregulated and downregulated gene functions were searched for in NCBI gene search.

The fold-changes observed following miRNA lentivirus transduction are all larger than those observed following transduction with the LAT intron lentivirus with the highest fold-changes being for CD34 (3.11-fold upregulated) and G protein-couples receptor 149 (GPR149) (2.98-fold downregulated) (table 4.3 and 4.4).

There are 2 upregulated DEGs among these top 5 that encode proteins that have an involvement in the development of neurons and the nervous system. This could be relevant when considering the impact of the latency-associated ncRNAs on human neurons. These proteins are the leucine rich repeat transmembrane neuronal 4 protein (LRRTM4) and semaphorin 3A (SEMA3A) protein (table 4.3). In terms of further common functions among the most upregulated/downregulated genes there are also 2 downregulated genes among the top 5 – GPR149 and Neuropeptide Y receptor Y2 (NP2R) (table 4.4) – that encode G protein-coupled receptors.

To further understand the impact of the induced changes, pathway analysis was performed to the 1404 genes that are differentially expressed following lentivirus delivery of 5 latency-associated HSV-1 miRNAs (appendix 1, sheet 9). GO biological process and KEGG pathway analyses were performed using DAVID (Huang et al., 2009b)(methods chapter 6.2.8) and the processes and pathways were ordered according the number of DEGs found to be associated with them (figures 4.9 and 4.10).

Following MiR lentivirus transduction, the GO biological process with the most upregulated genes associated with it was protein modification process with 163 associated genes (figure 4.9A), see appendix 2, sheet 5 for full raw DAVID analysis data with full associated genes list. Among these top 10 GO biological processes there were also metabolic-related processes with phosphate-containing compound metabolic process/phosphorus metabolic process and positive regulation of metabolic process. Additionally, macromolecule localisation came up as being associated with the upregulated genes following miR lentivirus transduction as well as the LAT intron lentivirus transduction.

Unlike the LAT intron upregulated genes, the miR upregulated genes were also found to be associated with known KEGG pathways. The KEGG pathway analysis revealed that

the lysosome KEGG pathway has the most upregulated genes associated with it following miR lentivirus transduction with 23 of the DEGs linked to this pathway (figure 4.9B)(appendix 2, sheet 6).

Following miR lentivirus transduction, the GO biological process associated with the most downregulated genes was (cellular) protein modification process with 150 DEGs (figure 4.10A). However, this process, along with macromolecule localisation, phosphorus metabolic process and phosphate-containing compound metabolic process, are also linked to the upregulated DEGs (figure 4.9A). As previously described, for a gene to be represented as part of a GO biological process, it could directly function within it or regulate it or a crucial related process, therefore the upregulated genes could have different molecular activity relating to each of these processes than the downregulated genes, despite both being associated with the same process (Dessimoz & Škunca, 2017). There were also GO biological processes specifically associated with the downregulated genes following miR lentivirus transduction, several of which are involved in localisation of cellular components (figure 4.10A).

The KEGG pathway analysis on the miR lentivirus downregulated DEGs revealed 4 pathways linked to disease pathogenesis occurring among the top 10 KEGG pathways, with human T-cell leukemia virus type 1 (HTLV-I) infection, viral carcinogenesis, Epstein-Barr virus infection and renal cell carcinoma. HTLV-I infection had the most downregulated genes of the significant KEGG pathways associated with it, with 15 DEGs (figure 4.10B and appendix 2, sheet 8).

Examining the functions of the DEGs following delivery of either the major LAT intron or 5 latency-associated miRNAs or the biological pathways that these DEGs are associated with, starts to give an indication of the physiological impact that these latency-associated ncRNAs may have in neurons during latent HSV-1 infection. However, there are alternative means to consider the most physiologically relevant transcriptome changes during HSV-1 latency.

Table 4.3 The top 5 significantly upregulated genes exclusively following miR lentivirus transduction in differentiated SH-SY5Y cells

The table includes the top 5 significantly upregulated genes, according to fold change, found by RNA-Seq, exclusively following miR lentivirus transduction of differentiated SH-SY5Y cells compared to untransduced differentiated SH-SY5Y cells, out of the DEGs found following the lentivirus transductions, as indicated in **figure 4.3**. The table shows the fold change, p-adjusted value following Benjamini-Hochberg test and functions of the genes, as found in (NCBI, 2016) unless otherwise indicated.

Table 4.3 The top 5 significantly upregulated genes exclusively following miR lentivirus transduction in differentiated SH-SY5Y cells

Gene	Fold change (2dp)	P adjusted Value (3 sf)	Known protein function
CD34	3.11	2.91E-07	CD34 molecule: • Hematopoietic stem/progenitor cell marker • May play a role in the attachment of stem cells to bone marrow extracellular matrix or stromal cells
LGI2	3.00	3.83E-05	Leucine rich repeat LGI family member 2: • Putative association wit partial epilepsy with pericentral spikes (Limviphuvadh et al., 2010)
SEMA3A	2.83	7.80E-18	Semaphorin 3A: • Secreted protein that can be either chemorepulsive, inhibiting axonal out-growth or chemattractive, stimulating the groth of apical dendrites. Either way, it's vital for normal neuronal pattern development. • Increased expression is associated with schitzophrenia and seen in a varie-ty of human tumour cell lines. • Aberrant release of this protein is associated with the progression of Alzhei-mer's disease
CGNL1	2.54	3.50E-26	Cingulin like 1: • Regulates RhoA and Rac1 GTPases to mediate junction assembly
LRRTM4	2.54	9.19E-06	Leucine rich repeat transmembrane neuronal 4: • Putative role in the development and maintenance of the central nervous system (Laurén et al. 2003)

Table 4.4 The top 5 significantly downregulated genes exclusively following miR lentivirus transduction in differentiated SH-SY5Y cells

The table includes the top 5 significantly downregulated genes, according to fold change, found by RNA-Seq, exclusively following miR lentivirus transduction of differentiated SH-SY5Y cells compared to untransduced differentiated SH-SY5Y cells, out of the DEGs found following the lentivirus transductions, as indicated in **figure 4.3**. The table shows the fold change, p-adjusted value following Benjamini-Hochberg test and functions of the genes, as found in (NCBI, 2016) unless otherwise indicated.

Table 4.4 The top 5 significantly downregulated genes exclusively following miR lentivirus transduction in differentiated SH-SY5Y cells

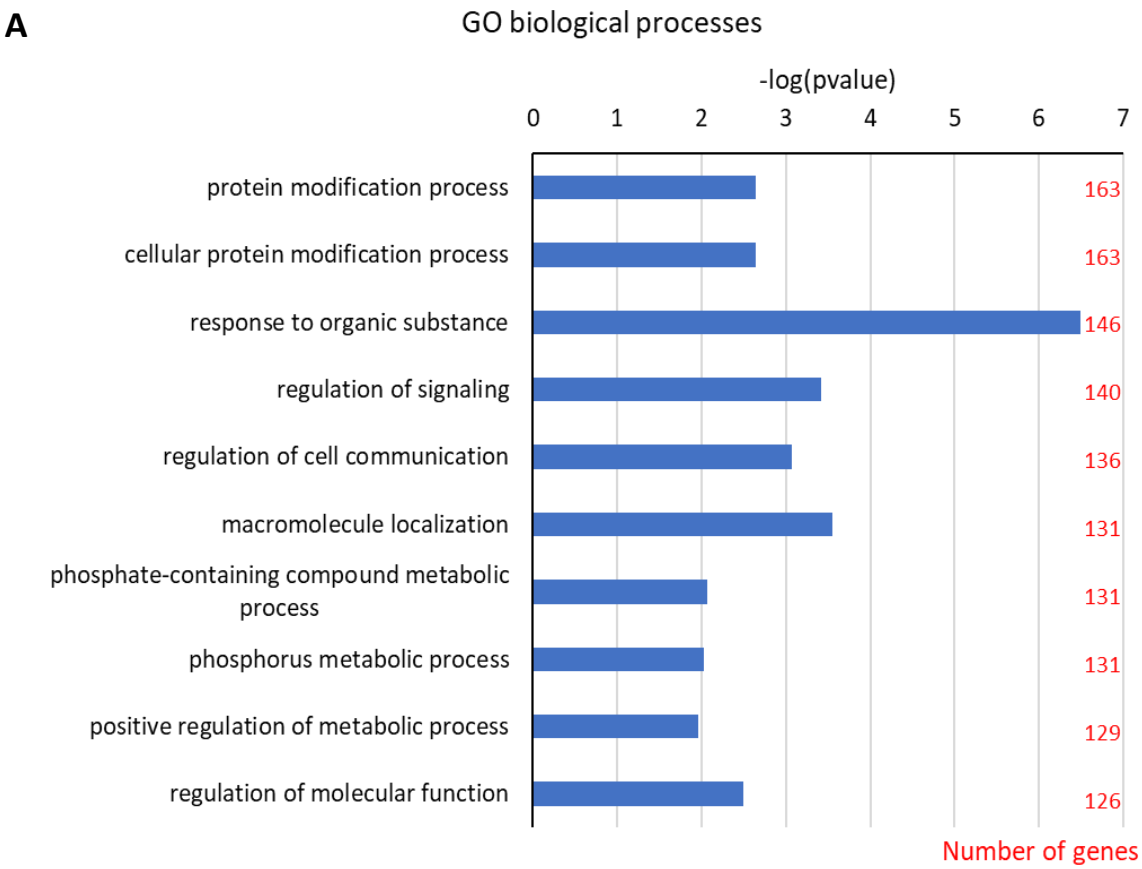
Gene	Fold change (2dp)	P adjusted Value (3 sf)	Known protein function
GPR149	-2.98	1.99E-05	G protein-coupled receptor 149 • G protein-coupled receptor • Important for G protein-coupled signal transduction
ULBP2	-2.72	0.000686	UL16 binding protein 2: • Major histocompatibility complex (MHC) class I-related molecule that binds to the NKG2S receptor on natural killer cells to trigger the release of multiple cytokines and chemokines that contribute to the recruitment and activation of NK cells.
NPY2R	-2.50	1.05E-06	Neuropeptide Y receptor Y2: • G protein-coupled receptor • Binds neuropeptide Y (NPY) (Aerts et al. 2018)
IGFBP7-AS1	-2.44	1.26E-06	Insulin-like growth factor binding protein 7- antisense RNA 1: • Long ncRNA (no known protein-coding function)
CMBL	-2.25	6.39E-46	Carboxymethylenebutenolidase homolog: • Cysteine hydrolase

Figure 4.9 GO biological processes and KEGG pathway analysis on significantly upregulated genes found by RNA-Seq exclusively following miR lentivirus transduction

(A) Bar graph showing the gene ontology (GO) biological processes following analysis using DAVID software (Huang et al., 2009a, 2009b) on the 708 significantly upregulated genes found by RNA-Seq following miR lentivirus transduction of differentiated SH-SY5Y cells as indicated in figure **4.3A**. The graph shows the top 10 significantly associated GO biological processes in terms of number of differentially expressed genes (DEGs), which are indicated in red. The blue bars show the -log10(p-value) following the fisher exact statistics method used by the DAVID software to calculate the p-value for each process.

(B) Bar graph to show the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways found by DAVID analysis (Huang et al., 2009a, 2009b) to be significantly associated with the 708 significantly downregulated genes following miR lentivirus transduction of differentiated SH-SY5Y cells as indicated in figure **4.3B** and listed in **appendix 1**. The number of DEGs involved are indicated in red. The blue bars show the -log10(p-value) following the fisher exact statistics used in the DAVID analysis software for each pathway.

See **appendix 2** for full gene lists associated with each GO biological process/KEGG pathway.



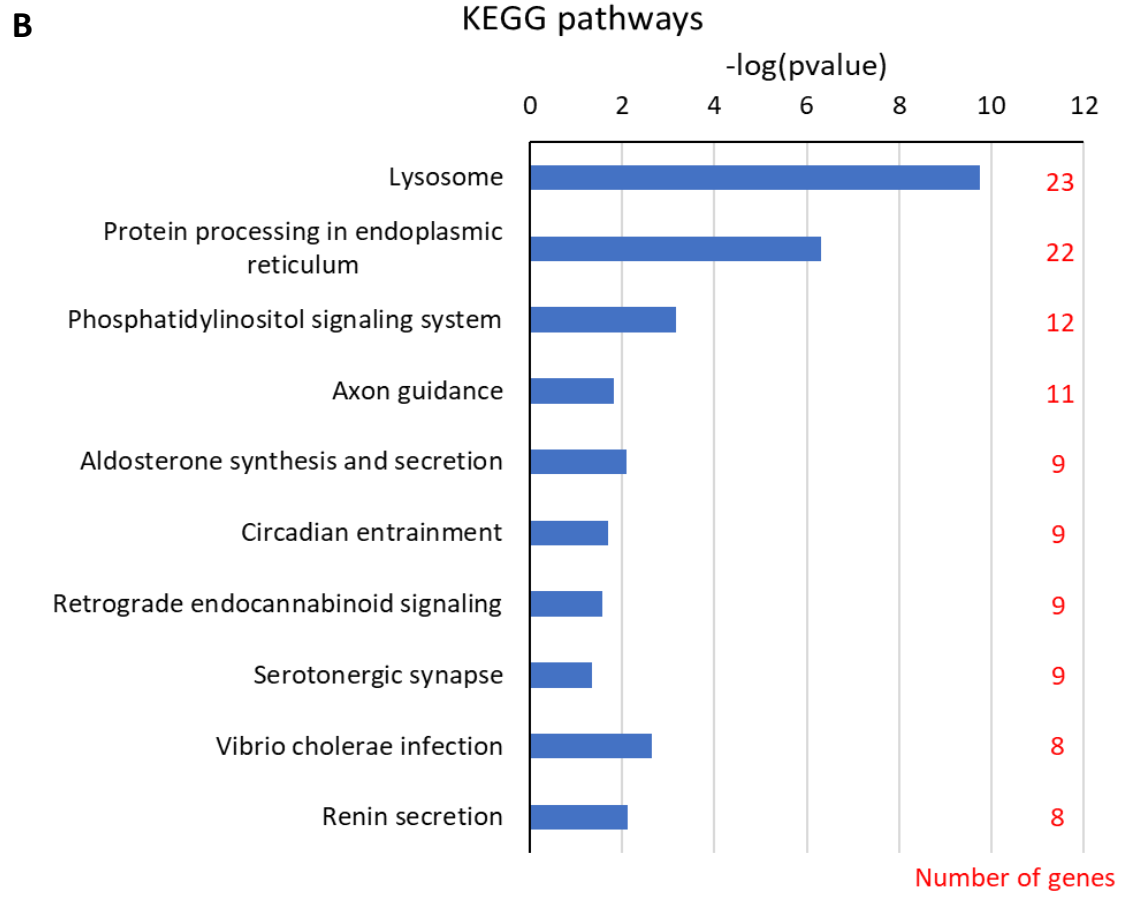


Figure 4.9 GO biological processes and KEGG pathway analysis on significantly upregulated genes found by RNA-seq exclusively following miR lentivirus transduction

Figure 4.10 GO biological processes and KEGG pathway analysis on significantly downregulated genes found by RNA-Seq exclusively following miR lentivirus transduction

(A) Bar graph showing the gene ontology (GO) biological processes following analysis using DAVID software (Huang et al., 2009a, 2009b) on the 696 significantly downregulated genes found by RNA-Seq following miR lentivirus transduction of differentiated SH-SY5Y cells as indicated in figure **4.3B** and listed in **appendix 1**. The graph shows the top 10 significantly associated GO biological processes in terms of number of differentially expressed genes (DEGs), which are indicated in red. The blue bars show the -log10(p-value) following the fisher exact statistics method used by the DAVID software to calculate the p-value for each process.

(B) Bar graph to show the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways found by DAVID analysis (Huang et al., 2009a, 2009b) to be significantly associated with the 696 significantly downregulated genes following miR lentivirus transduction of differentiated SH-SY5Y cells as indicated in figure **4.3B** and listed in **appendix 1**. The number of DEGs involved are indicated in red. The blue bars show the -log10(p-value) following the fisher exact statistics used in the DAVID analysis software for each pathway.

See **appendix 2** for full gene lists associated with each GO biological process/KEGG pathway.

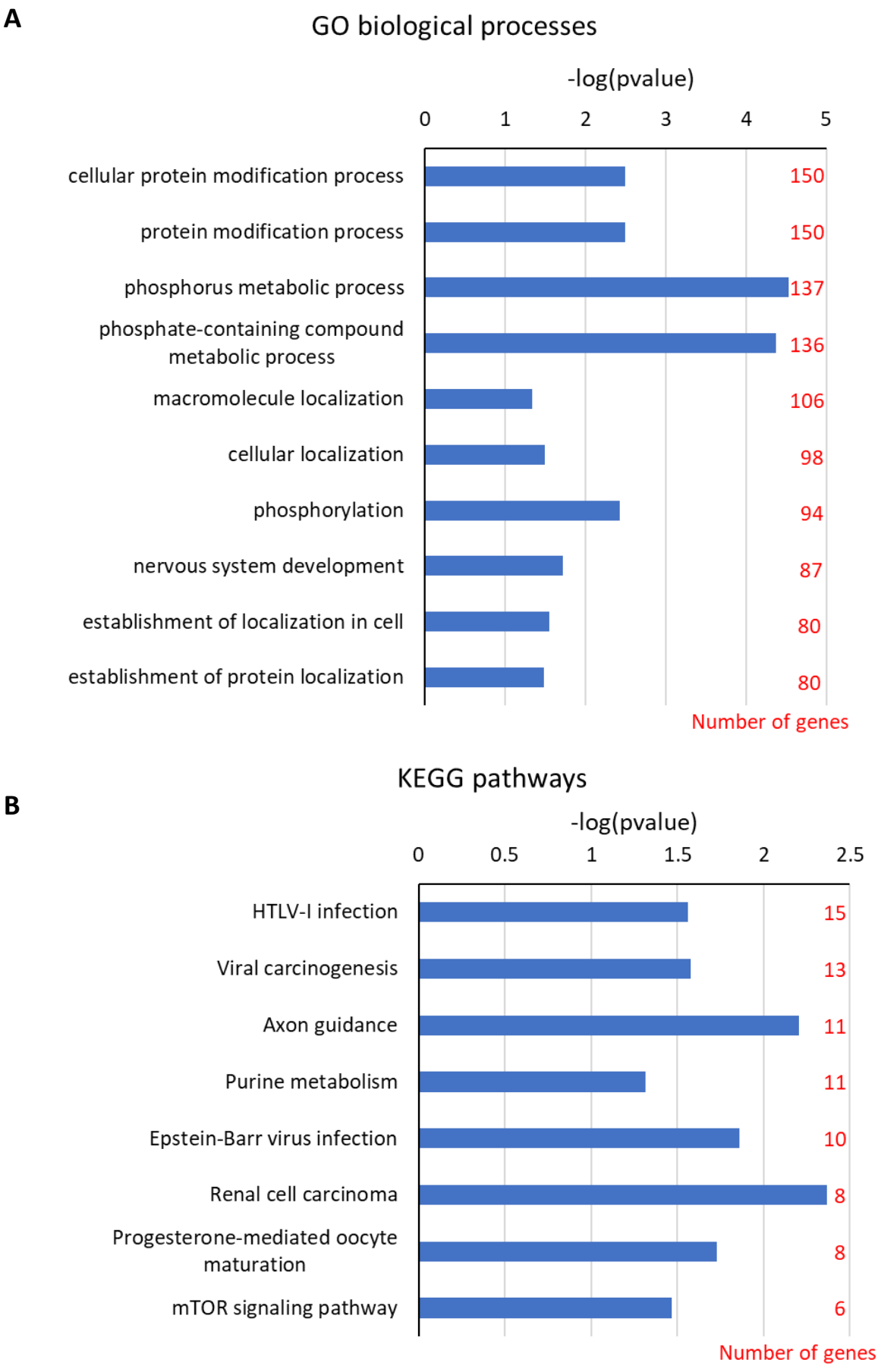


Figure 4.10 GO biological processes and KEGG pathway analysis on significantly downregulated genes found by RNA-seq exclusively following miR lentivirus transduction

4.2.4 Quiescent HSV-1 infection induces over 1500 differentially expressed genes (DEGs)

The lentivirus delivery of the latency associated ncRNAs utilised in the previous experiment has the advantage of separating out transcriptome changes induced by LAT and the 5 latency-associated miRNAs specifically, separate from any effects elicited by the virus genome in its entirety. However, a key test to determining the physiological relevance of these findings was to examine whether they could be recapitulated during latent infection with HSV-1. To this end, cultures of differentiated SH-SY5Y were infected with the HSV-1 recombinant virus, in1382, for onwards RNA-Seq analysis.

Differentiated SH-SY5Y cells were infected with HSV-1 strain in1382 at MOI 5 for 5 biological repeats. At 13 days post infection (dpi), RNA pooled from 2 wells was extracted as indicated in figure 4.11A and analysed by RNA-Seq and qRT-PCR. RNA was harvested at 13dpi because the peak LAT expression was at 11dpi and low-level lytic gene expression was observed from 14dpi (chapter 2, figure 2.5C), which suggested potential reactivation from this time post-infection. Therefore, RNA was extracted at the latest day following the peak of LAT expression, to allow for any effect on the transcriptome, but before reactivation was observed, to try to minimise any effects from lytic transcription.

These samples were analysed by RNA-Seq as described in the previous experiment (sections 4.2.1/4.2.2) and shown schematically in figure 4.2A, by Mr Ryan Mate and Dr Martin Fritzsche, ABS, NIBSC. The data output from the Illumina NextSeq 500 sequencer was then taken through the bioinformatic steps described above and shown in figure 4.2B by Dr Thomas Bleazard, ABS, NIBSC. The quality control measures throughout this process (methods 6.2.7.1) found 2 of the biological samples to fall below the quality threshold required (phred score of 30 or above). The phred score estimates the probability of an error for each nucleobase identified by sequencing, where a score of 30 indicates the probability of an error equals 1 in 1000 (Ewing & Green, 1998). Therefore, bioinformatics analysis was continued on the remaining 3 biological repeats. The bioinformatics analysis revealed 1538 DEGs, where padj ≤ 0.05, in response to in1382 quiescent infection (figure 4.11C).

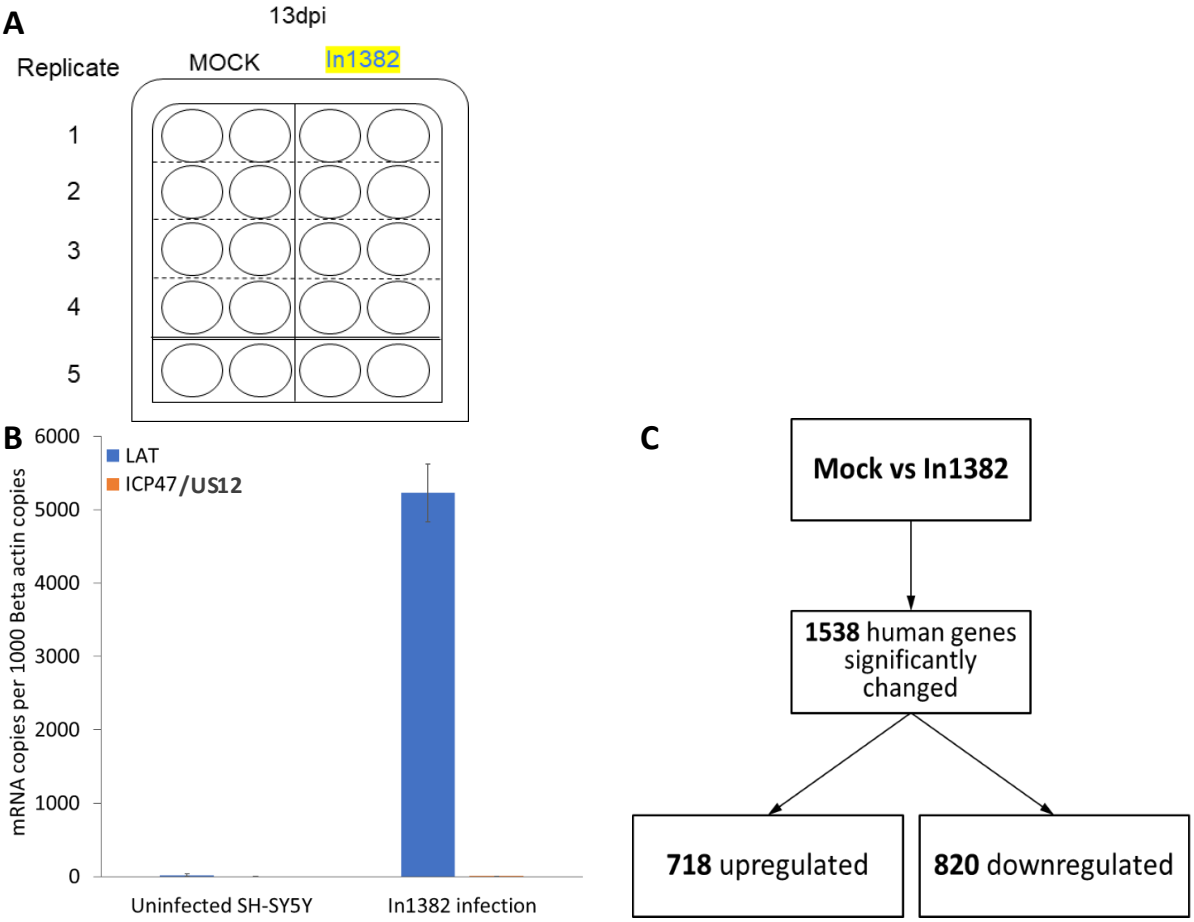
Figure 4.11 RNA-Seq on in1382 infected differentiated SH-SY5Y cells

(A) Schematic representation of the set-up for the in1382 RNA-Seq experiment whereby differentiated SH-SY5Y cells were infected (at 39^{o}C) with HSV-1 strain in1382 at MOI 5 or mock infected. Total RNA was harvested 13 dpi with 5 replicates per condition (3 that were used for onward bioinformatics analysis) and RNA being pooled from 2 wells per replicate.

(B) Differentiated SH-SY5Y cells were infected with HSV-1 strain in1382 or mock infected as described in **(A)** and qRT-PCR performed to LAT, ICP47(/US12) and Beta actin. The graph shows the average copies of LAT (blue) and ICP47 (orange) per 1000 Beta actin copies across 3 biological repeats and 3 technical repeats.

(C) A diagram outlining the number of differentially expressed genes following the RNA-Seq steps outlined in figure **4.2A** on the set-up described in **(A)**.

(D) A table showing all the lytic genes found to be significantly transcribed by in1382 infected differentiated SH-SY5Y cells compared to uninfected differentiated SH-SY5Y cells in the RNA-Seq experiment shown in **(A)**. The table shows gene name, the adjusted p-value (padj) following the Benjamini-Hochberg adjustment and the normalised read counts for each of the mock infected (mock) and the in1382 infected (in1382) samples. The US12(/ICP47) gene that was tested in **(C)** is highlighted in bold with a surrounding box. The reads for HSV-1 LAT found by RNA-Seq was added to the end of the table (also in bold and surrounded by a black box) to compare to the significantly differentially expressed HSV-1 genes.



D

Gene	padj	Mock 1	Mock 2	Mock 3	In1382 1	In1382 2	In1382 3
HSV1_UL18	3.4E-06	0	0	0	30	26	19
HSV1_UL2	8.0E-12	0	0	0	128	145	69
HSV1_UL22	1.9E-06	0	0	0	26	34	21
HSV1_UL24	2.1E-12	0	0	0	141	134	96
HSV1_UL25	8.6E-12	0	0	0	149	94	87
HSV1_UL26	1.4E-10	0	0	0	133	66	61
HSV1_UL26.5	1.3E-08	0	0	0	84	35	39
HSV1_UL27	2.7E-07	0	0	0	33	37	29
HSV1_UL3	2.2E-07	0	0	0	65	37	17
HSV1_UL45	6.0E-14	0	0	0	207	245	114
HSV1_UL46	1.2E-07	0	0	0	68	33	24
HSV1_UL48	1.7E-11	0	0	0	97	122	80
HSV1_US1	1.4E-08	0	0	0	67	55	29
HSV1_US10	3.2E-09	0	0	0	82	60	37
HSV1_US11	2.4E-09	0	0	0	86	64	37
HSV1_US12	**1.8E-09**	**0**	**0**	**0**	**91**	**66**	**38**
HSV1_US5	1.8E-06	0	0	0	43	17	27
HSV1_US6	1.8E-06	0	0	0	43	17	27
HSV1_US7	3.6E-06	0	0	0	40	17	22
HSV1_US8	6.7E-07	0	0	0	34	27	29
HSV1_US9	1.3E-06	0	0	0	32	26	26
HSV1_LAT	**1.64**	**0**	**0**	**0**	**20**	**2**	**3**

Figure 4.11 RNA-seq on in1382 infected differentiated SH-SY5Y cells

To confirm that these changes were due to expression of the latency-associated ncRNAs, qRT-PCR was performed using primers to the major LAT sequence (chapter 3, figure 3.6D) and Beta actin as a positive control, to test that LAT is produced in these samples, as it has previously been shown to be in this model (chapter 2, figure 2.5). As the HSV-1 latency-associated miRNAs are processed from the same part of the HSV-1 genome, LAT abundance also acted as a proxy for HSV-1 miRNA expression here. Following in1382 infection, LAT was highly transcribed with 5230 LAT copies per 1000 beta actin copies in the in1382 infected differentiated SY-SY5Y cells and there was minimal detection in the negative control uninfected differentiated SH-SY5Y cells (figure 4.11B). This shows considerably higher LAT transcription than during lentivirus delivery of LAT (figure 4.1B).

The bioinformatics analysis described in figure 4.2B aligned the reads to both the human genome and the HSV-1 genome, which revealed reads across some HSV-1 lytic transcripts in the in1382 infected samples (figure 4.11D). Among the DEGs, filtered for significance with an FDR cut-off of 5% (padj ≤ 0.05), there were 21 lytic HSV-1 genes aligning to the HSV-1 genome, with a minimum of 17 reads per sample (figure 4.11D). Note that the abundant major LAT species is not detectable from the RNA-Seq analysis due to the poly-A selection that takes place prior to analysis, as it is not polyadenylated. With regards to the low-level of lytic transcription observed, there is a precedent for low-level lytic transcription during HSV-1 latency as discussed in chapter 1.4.2 and demonstrated by Harkness *et al.* by performing RNA-Seq on latently infected murine trigeminal ganglia neurons (Harkness et al., 2014).

In contrast to major LAT, the primary LAT transcript that major LAT is spliced from, along with the minor LAT exon that is the remaining product following splicing of major LAT, are polyadenylated and therefore can be detected within this RNA-Seq experiment, which includes alignment to the HSV-1 genome. As already described (chapter 1.4.1) major LAT is highly stable, largely owing to its lariat structure (Block & Taylor, 1996; Farrell et al., 1991; Wu et al., 1998). The 8.3kb primary LAT and 6.3 kb minor LAT transcripts are relatively unstable by comparison. The read counts for the primary LAT transcript/minor LAT reflect this instability, as they were low and inconsistent across the 3 samples with 20, 2 and 3 reads in each sample (figure 4.11D). As such, LAT was not one of the genes that came up as significantly increased following in1382 infection. However, when qRT-PCR is performed to LAT (prior to poly-A selection) which given this and the position of the primers (figure 3.6D) would include major LAT as well as primary LAT transcript, showed LAT to be highly abundant (figure 4.11B).

In addition to the LAT qRT-PCR described above, qRT-PCR was performed with primers to one of the lytic genes shown to be significantly upregulated during in1382 infection by RNA-Seq analysis – ICP47/US12. This would allow sensitive comparison of the relative transcription of this lytic gene to LAT transcription. RNA-Seq revealed an average of 65 reads aligned to the HSV-1 genome at the position of ICP47/US12. However, compared to the LAT transcription, as assessed by qRT-PCR (described above), this amount was negligible. The qRT-PCR results revealed that despite the reads observed over lytic regions of the HSV-1 genome, LAT is considerably more abundantly transcribed, with 5230 copies per 1000 beta actin, compared to less than 1 copy per 1000 beta actin copies of ICP47 (figure 4.11B). Therefore, despite any potential lytic transcription in the in1382 infection, LAT seems to be the most abundant product, as is expected during a quiescent infection. The discrepancy between the low primary LAT reads and high abundance of major LAT by qRT-PCR also suggests that a considerable proportion of this abundance results from the stability and accumulation of the spliced major LAT intron RNA. Accordingly, it is likely that the major LAT intron is contributing to the neuronal transcriptional changes observed by RNA-Seq. Additionally, it is much more likely that these transcriptional changes are a result of the latency-associated ncRNAs and not lytic transcription.

More than 1500 host genes are significantly differentially expressed as a result of quiescent infection, but to validate these changes most effectively the list of DEGs during HSV-1 latency can be assessed for gene function and narrowed down to look at those potentially most relevant.

4.2.5 Examining the biological functions and pathways affected by the in1382 infection DEGs

As with the lentivirus RNA-Seq experiment, the data was assessed according to highest fold-change as outlined in figure 4.6. Higher fold-changes were observed by RNA-Seq following in1382 infection compared to the ncRNA lentivirus experiment (tables 4.5, 4.6). The highest fold-change observed among the upregulated DEGs was with nescient helix-loop-helix 2 (NHLH2) which was 7-fold upregulated (table 4.5). The gene that was most downregulated was with lipoprotein lipase (LPL), which was 12.4-fold downregulated (table 4.6). The functions of the top 5 upregulated and downregulated genes following the in1382 infection outlined in figure 4.11 are described in tables 4.5 and 4.6, respectively. Of these top DEGs, NHLH2 upregulation may be particularly relevant as it is a neuronal transcription factor, so this upregulation may have downstream effects on the regulation of other genes in these neurons. Additionally, it is

potentially of interest that among the top downregulated genes is CCAAT enhancer binding protein delta (CEBPD), which is a transcription factor involved in regulating genes involved in the immune and inflammatory responses, which could be relevant during a viral infection. Furthermore, IL-6 has been shown to activate CEBPD (Cantwell et al., 1998), while HSV-1 miR-H6 has been shown to reduce production of IL6 (Duan et al., 2012). This could explain the downregulation observed following In1382 infection, as miR-H6 might be mediating this effect through regulating IL-6 which in turn regulates CEPBD.

More than 1500 significant cellular DEGs were found in response to quiescent HSV-1 infection. To consider the potential combined impact of these gene changes initially GO analysis and KEGG pathway analysis was performed in DAVID, to include all significantly changed genes (Huang et al., 2009a, 2009b). Identified pathways were filtered for significance (p-value < 0.05) whereby a fisher exact statistics method is used to calculate p-values in DAVID (Huang et al., 2009a, 2009b). These pathways were then ordered in terms of the number of genes associated with the process or pathway. The graphs (figures 4.12 and 4.13) indicate the top processes or pathways (where p-value < 0.05) – based on the number of DEGs involved – plotted against the -log10(p-value).

GO biological processes analysis of significantly upregulated genes revealed that following in1382 infection of differentiated SH-SY5Y cells the biological process with the most associated upregulated genes was protein modification process with 166 genes (figure 4.12A). Additionally, there were several processes among the top 10 GO processes linked to the upregulated DEGs involved in signalling. These GO processes comprised of regulation of signalling, intracellular signal transduction, regulation of signal transduction and cell surface receptor signalling (figure 4.12A).

The KEGG pathway analysis for the upregulated DEGs revealed some similar associations with pathways involved in signalling, with PI3K-Akt signalling pathway, Rap1 signalling pathway, MAPK signalling pathway, calcium signalling pathway and Ras signalling pathway all among the top 10 (figure 4.12B). The top KEGG pathway associated with the most upregulated DEGs following in1382 infection was proteoglycans in cancer, with 25 upregulated genes associated with it (figure 4.12B).

The same analyses were performed for the significantly downregulated genes. GO biological process analysis revealed that response to organic substance was significantly associated with the most downregulated genes with 167 DEGs (figure 4.13A). For the KEGG pathway analysis the pathway with the most downregulated genes significantly associated with it was the cell cycle with 24 DEGs (figure 4.13B).

There were also multiple KEGG pathways linked to response to viral infection among the top 10, consisting of viral carcinogenesis, herpes simplex infection, influenza A and Epstein-Barr virus infection (figure 4.13B).

Table 4.5 The top 5 significantly upregulated genes following in1382 infection of differentiated SH-SY5Y cells

The table includes the top 5 significantly upregulated genes, according to fold change, found by RNA-Seq, following 13-day in1382 infection of differentiated SH-SY5Y cells compared to uninfected differentiated SH-SY5Y cells, as set up in **figure 4.11A.** The table shows the fold change, p-adjusted value following Benjamini-Hochberg test and functions of the genes, as found in (NCBI, 2016) unless otherwise indicated.

Table 4.5 The top 5 significantly upregulated genes following in1382 infection of differentiated SH-SY5Y cells

Gene	Fold change (2dp)	P adjusted Value (3 sf)	Known protein function
NHLH2	7.02	0.00186	Nescient helix-loop-helix 2: • Neuronal transcription factor • Putative role in controlling physical activity
HSPA6	6.60	0.00683	Heat shock protein family a (Hsp70) member 6: • Suggested protection of differentiated human neuronal cells from cellular stress
CU638689.1	5.95	4.17E-06	Locus position – no known protein coding or functions.
CRLF1	5.57	3.18E-52	Cytokine receptor like factor 1: • Member of the cytokine type I receptor family • Forms a secreted complex with cardiotrophin-like cytokine factor 1 and acts on cells expressing ciliary neurotrophic factor receptors • The complex it forms promotes survival of neurons
GJA5	5.48	0.00372	Gap junction protein alpha 5: • Member of the connexin family • Component of gap junctions, allowing cell-to-cell diffusion

Table 4.6 The top 5 significantly downregulated genes following in1382 infection of differentiated SH-SY5Y cells

The table includes the top 5 significantly downregulated genes, according to fold change, found by RNA-Seq, following 13-day in1382 infection of differentiated SH-SY5Y cells compared to uninfected differentiated SH-SY5Y cells, as set up in **figure 4.11A.** The table shows the fold change, p-adjusted value following Benjamini-Hochberg test and functions of the genes, as found in (NCBI, 2016) unless otherwise indicated.

Table 4.6 The top 5 significantly downregulated genes following in1382 infection of differentiated SH-SY5Y cells

Gene	Fold change (2dp)	P adjusted Value (3 sf)	Known protein function
LPL	-12.43	5.93E-05	Lipoprotein lipase: • Triglyceride hydrolase • Ligand for receptor-mediated lipoprotein uptake
BCAS1	-11.43	4.04E-12	Brain enriched myelin associated protein 1/Breast carcinoma-amplified sequence 1: • Required for myelination • Amplified in a variety of tumours, especially breast cancer
ADRA2A	-8.96	3.37E-12	Adrenoceptor alpha 2A: • Member of the G protein-coupled receptor superfamily • Adrenergic receptor that inhibits adenylate cyclase • Regulates the release of neurotransmitter molecules from sympathetic, adrenergic neurons of the CNS
MMP9	-8.95	1.35E-20	Matrix metallopeptidase 9: • Breakdown of extracellular matrix • Degrades type IV and V collagens
CEBPD	-8.89	0.00515	CCAAT enhancer binding protein delta: • A transcription factor important in the regulation of genes involved in immune and inflammatory responses

Figure 4.12 GO biological process and KEGG pathway analysis on the significantly upregulated genes found by RNA-Seq following in1382 infection

(A) A bar graph to show the gene ontology (GO) biological processes following analysis using DAVID software (Huang et al., 2009a, 2009b) on all significantly upregulated genes found by RNA-Seq following in1382 infection of differentiated SH-SY5Y cells, as listed in **appendix 1**. The graph shows the top 10 significantly associated GO biological processes in terms of number of differentially expressed genes (DEGs), which are indicated in red. The blue bars show the -log10(p-value) following the fisher exact statistics method used by the DAVID software to calculate the p-value for each process. The full gene lists associated with each GO biological process are listed in **appendix 2.**

(B) Bar graph to show the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways found by DAVID analysis (Huang et al., 2009a, 2009b) to be significantly associated with the significantly upregulated genes following in1382 infection of differentiated SH-SY5Y cells, as listed in **appendix 1**. The number of DEGs involved are indicated in red. The blue bars show the -log10(p-value) following the fisher exact statistics used in the DAVID analysis software for each pathway. See **appendix 2** for full gene lists associated with each KEGG pathway.

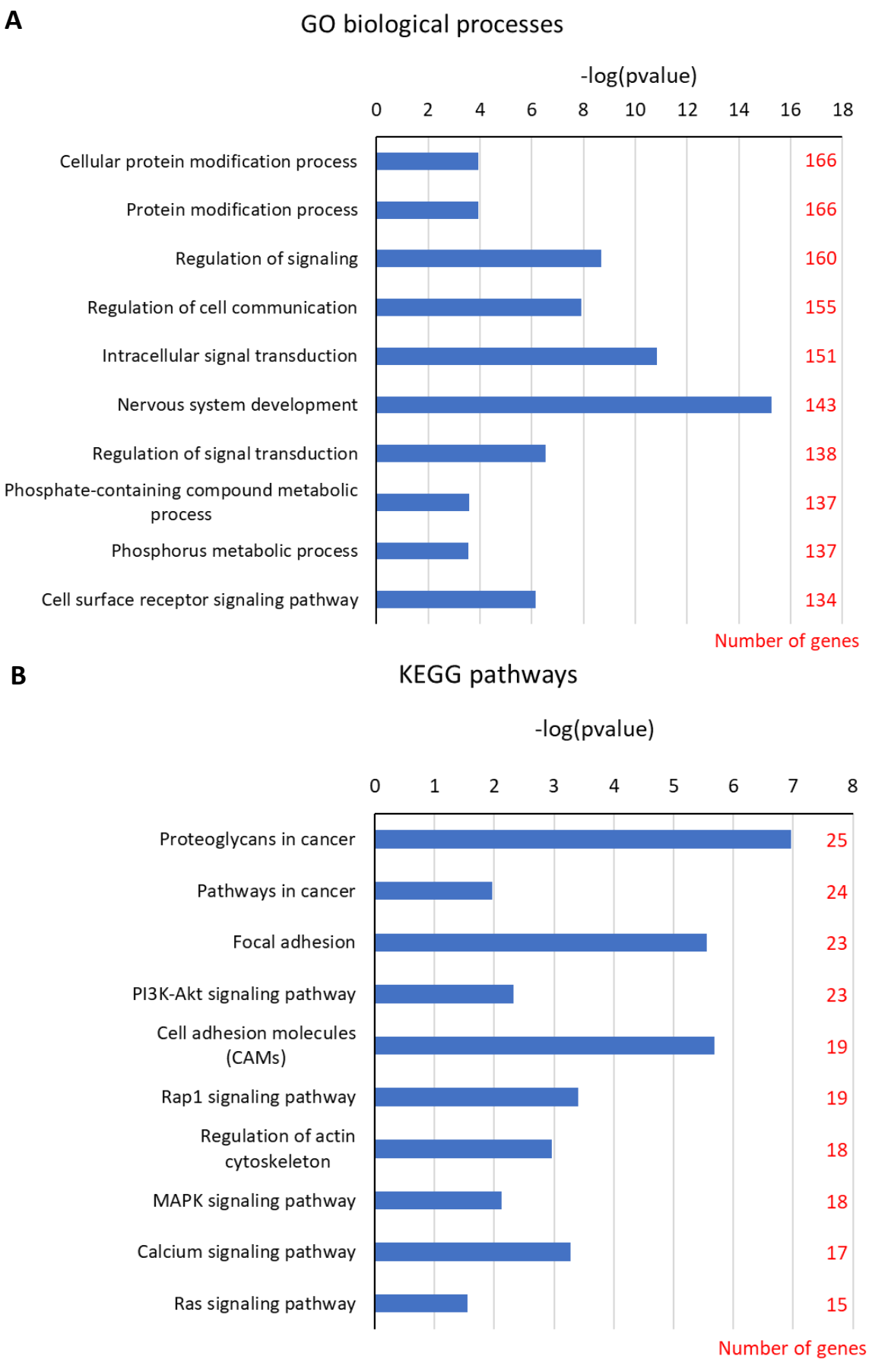


Figure 4.12 GO biological process and KEGG pathway analysis on the significantly upregulated genes found by RNA-seq following in1382 infection

Figure 4.13 GO biological process and KEGG pathway analysis on the significantly downregulated genes found by RNA-Seq following in1382 infection

(A) A bar graph to show the gene ontology (GO) biological processes following analysis using DAVID software (Huang et al., 2009a, 2009b) on all significantly downregulated genes found by RNA-Seq following in1382 infection of differentiated SH-SY5Y cells, as listed in **appendix 1**. The graph shows the top 10 significantly associated GO biological processes in terms of number of differentially expressed genes (DEGs), which are indicated in red. The blue bars show the -log10(p-value) following the fisher exact statistics method used by the DAVID software to calculate the p-value for each process. The full gene lists associated with each GO biological process are listed in **appendix 2.**

(B) Bar graph to show the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways found by DAVID analysis (Huang et al., 2009a, 2009b) to be significantly associated with the significantly downregulated genes following in1382 infection of differentiated SH-SY5Y cells, as listed in **appendix 1**. The number of DEGs involved are indicated in red. The blue bars show the -log10(p-value) following the fisher exact statistics used in the DAVID analysis software for each pathway. See **appendix 2** for full gene lists associated with each KEGG pathway.

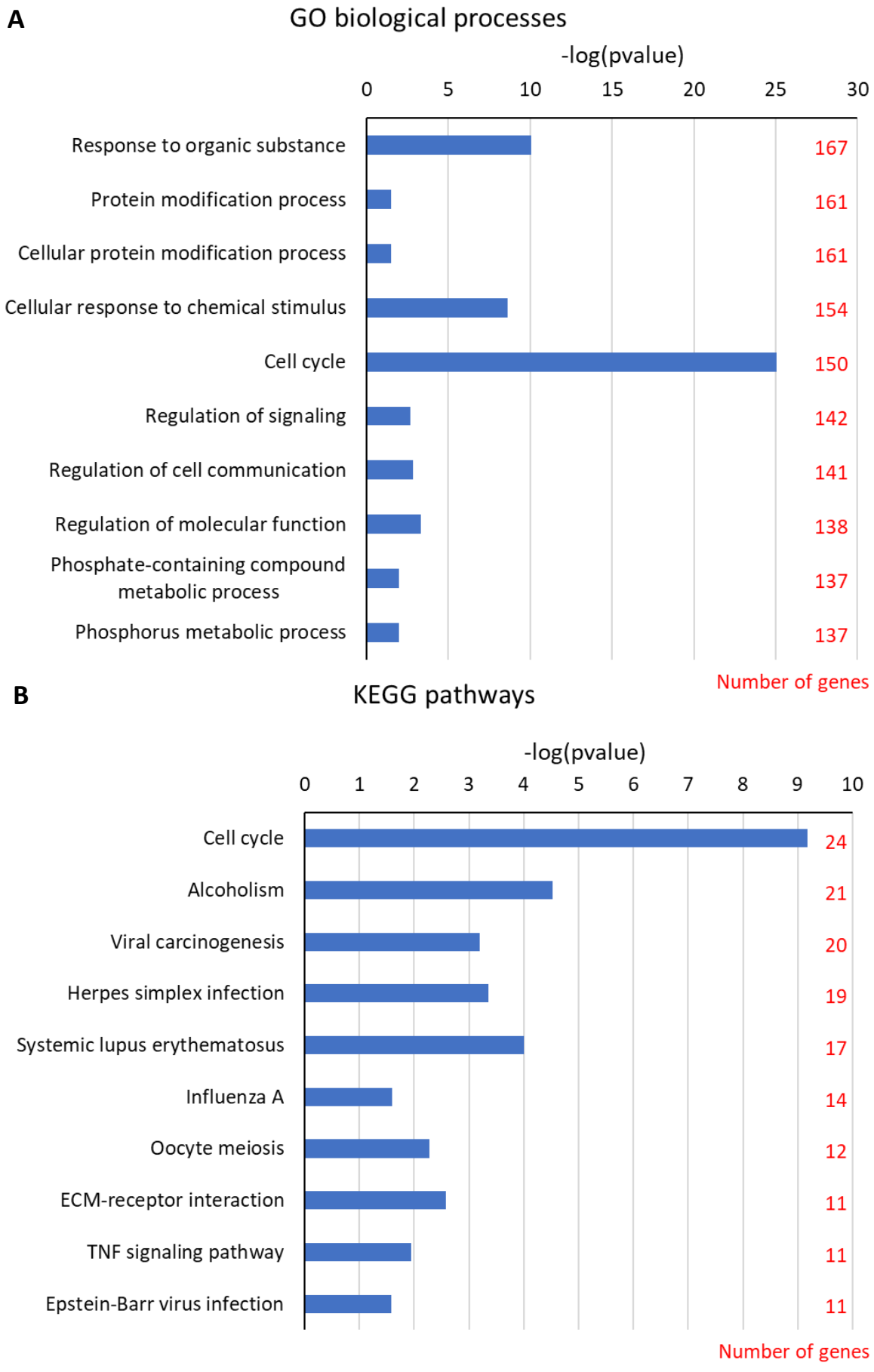


Figure 4.13 GO biological process and KEGG pathway analysis on the significantly downregulated genes found by RNA-seq following in1382 infection

4.2.5.1 Narrowing down the in1382 infection DEGs by fold-change

To focus the analysis, to examine the DEGs with the largest effect, the DEGs were filtered based upon the magnitude of their fold-change. The functions of the genes with the largest fold-changes have already been considered individually (tables 4.5 and 4.6). However, to investigate the cumulative function of the genes exhibiting the greatest change, a threshold of 2-fold change was applied to filter the list of the DEGs. The common functions of these filtered DEGs were then explored via pathway analysis.

The 92 upregulated and 311 downregulated filtered gene lists (fold change >2 for upregulated genes, fold change < -2 for downregulated genes) were analysed in DAVID for GO analysis and KEGG pathways (Huang et al., 2009a, 2009b). Identified pathways were filtered for significance (p-value < 0.05) whereby a fisher exact statistics method is used to calculate p-values in DAVID (Huang et al., 2009a, 2009b). These pathways were then ordered in terms of the number of genes associated with the process or pathway. The graphs (figures 4.14 and 4.15) indicate the top processes or pathways (where p-value < 0.05) – based on the number of DEGs involved – plotted against the -log10(p-value).

GO biological processes analysis of the 92 significantly upregulated genes following in1382 infection of differentiated SH-SY5Y cells revealed enrichment for various processes involved in regulation of signalling, cell-to-cell communication and adhesion (figure 4.14A). The KEGG pathway analysis revealed similar associations with pathways associated with signalling but more specifically associated with neuronal signalling, with enrichment of neuroactive ligand-receptor interaction and calcium signalling pathways (figure 4.14B).

Among the top GO biological processes associated with these upregulated DEGs, there were overlaps to the top GO biological processes observed to be associated with the upregulated genes following transduction with either the LAT intron lentivirus or miR lentivirus. Cell-to-cell signalling, observed to have 17 of the 2-fold or more upregulated genes following in1382 infection, was also one of the top 10 GO pathways associated with the upregulated DEGs following LAT intron lentivirus transduction (figure 4.7). Additionally, regulation of signalling and regulation of cell communication (with 22 and 21 associated upregulated DEGs following in1382 infection - figure 4.14) were also among the top GO biological processes associated with upregulated DEGs following miR lentivirus transduction (figure 4.9A).

GO analysis was also performed on the 311 2-fold or more significantly downregulated genes following in1382 infection of differentiated SH-SY5Y cells (see appendix 1, sheet

10 for full gene list). The top pathways in terms of number of DEGs involved were those to do with cellular responses, to either chemical or organic substances. However, despite their enrichment with upregulated genes, there does also seem to be an association between signalling and cell communication pathways and the downregulated genes as well.

Herpes simplex infection pathway was the joint top KEGG pathway found to be associated with 10 2-fold or more significantly downregulated genes. Associated with the herpes simplex infection pathway are host genes that have been shown to be involved in the response to HSV, for example promyelocytic leukaemia (PML) nuclear body component SP100. Although this is a quiescent infection it seems that responses known to be associated with HSV-1 lytic infection are observed here. For full gene lists associated with each pathway, see appendix 2, sheet 12.

There were two processes and pathways associated with genes downregulated during in1382 quiescent infection (figure 4.15) that were also found to be associated with downregulated DEGs following lentivirus transduction: “regulation of molecular function” (also associated with major LAT intron lentivirus downregulated genes), and “viral carcinogenesis” (also associated with miR lentivirus downregulated genes).(figure 4.10B).

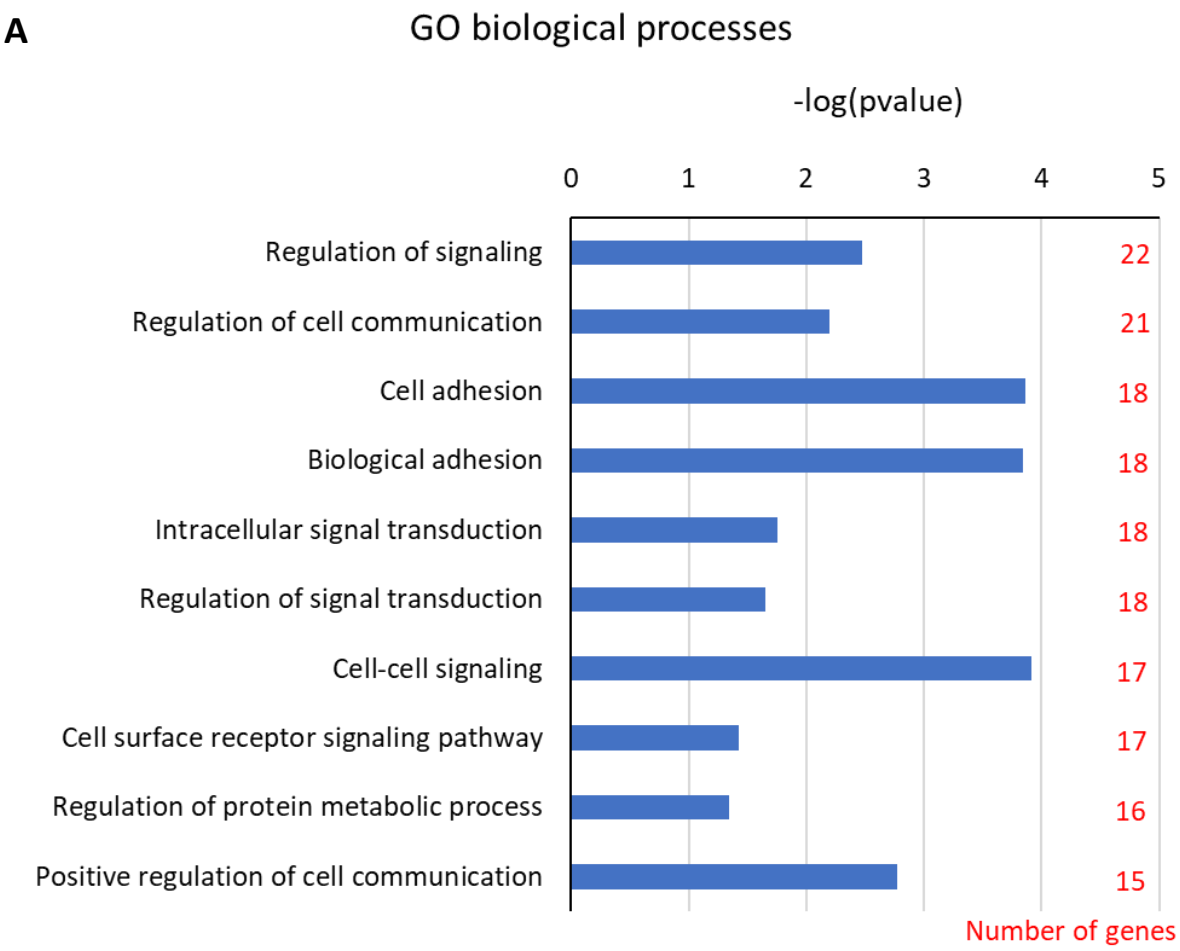
Although cell signalling pathways were associated with both upregulated and downregulated genes, there were also pathways identified to be specifically associated with downregulated DEGs such as cell-cycle, alcoholism and herpes simplex infection In addition, there are processes and pathways that are consistent with the changes observed following lentivirus delivery of latency-associated ncRNAs, providing strong evidence that the HSV-1 latency-associated ncRNAs regulate their function. These overlapping changes will be discussed next.

Figure 4.14 GO biological process and KEGG pathway analysis on the 2-fold or more significantly upregulated genes found by RNA-Seq following in1382 infection

(A) A bar graph to show the top 10 gene ontology (GO) biological processes (based on number of associated genes) following analysis using DAVID software (Huang et al., 2009a, 2009b) on the 2-fold or more significantly upregulated genes found by RNA-Seq following in1382 infection of differentiated SH-SY5Y cells, as listed in **appendix 1**. The graph shows the top 10 significantly associated GO biological processes in terms of number of differentially expressed genes (DEGs), which are indicated in red. The blue bars show the -log10(p-value) following the fisher exact statistics method used by the DAVID software to calculate the p-value for each process.

(B) Bar graph to show the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways found by DAVID analysis (Huang et al., 2009a, 2009b) to be significantly associated with the 2-fold or more significantly upregulated genes following in1382 infection of differentiated SH-SY5Y cells, as listed in **appendix 1**. The number of DEGs involved are indicated in red. The blue bars show the -log10(p-value) following the fisher exact statistics used in the DAVID analysis software for each pathway.

See **appendix 2** for full gene lists associated with each GO biological process/KEGG pathway.



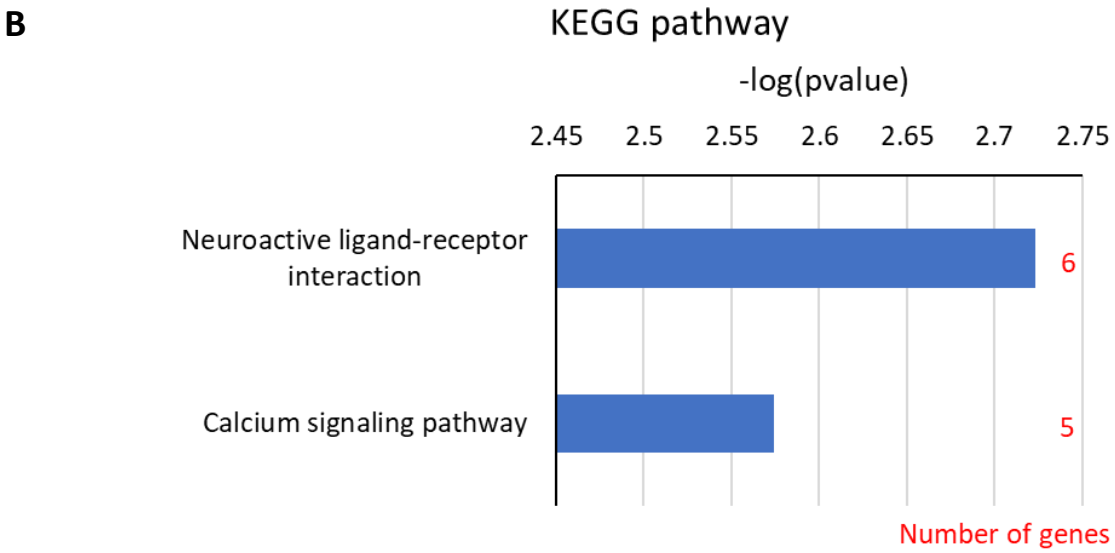


Figure 4.14 GO biological process and KEGG pathway analysis on the 2-fold or more significantly up-regulated genes found by RNA-Seq following in1382 infection

Figure 4.15 GO biological process and KEGG pathway analysis on the 2-fold or more significantly downregulated genes found by RNA-Seq following in1382 infection

(A) A bar graph to show the top 10 gene ontology (GO) biological processes (according to number of genes associated) following analysis using DAVID software (Huang et al., 2009a, 2009b) on the 2-fold or more significantly downregulated genes found by RNA-Seq following in1382 infection of differentiated SH-SY5Y cells, as listed in **appendix 1**. The graph shows the top 10 significantly associated GO biological processes in terms of number of differentially expressed genes (DEGs), which are indicated in red. The blue bars show the -log10(p-values) following the fisher exact statistics method used by the DAVID software to calculate the p-value for each process.

(B) Bar graph to show the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways found by DAVID analysis (Huang et al., 2009a, 2009b) to be significantly associated with the significantly downregulated genes following in1382 infection of differentiated SH-SY5Y cells, as listed in **appendix 1**. The number of DEGs involved are indicated in red. The blue bars show the -log10(p-value) following the fisher exact statistics used in the DAVID analysis software for each pathway.

See **appendix 2** for full gene lists associated with each GO biological process/KEGG pathway.

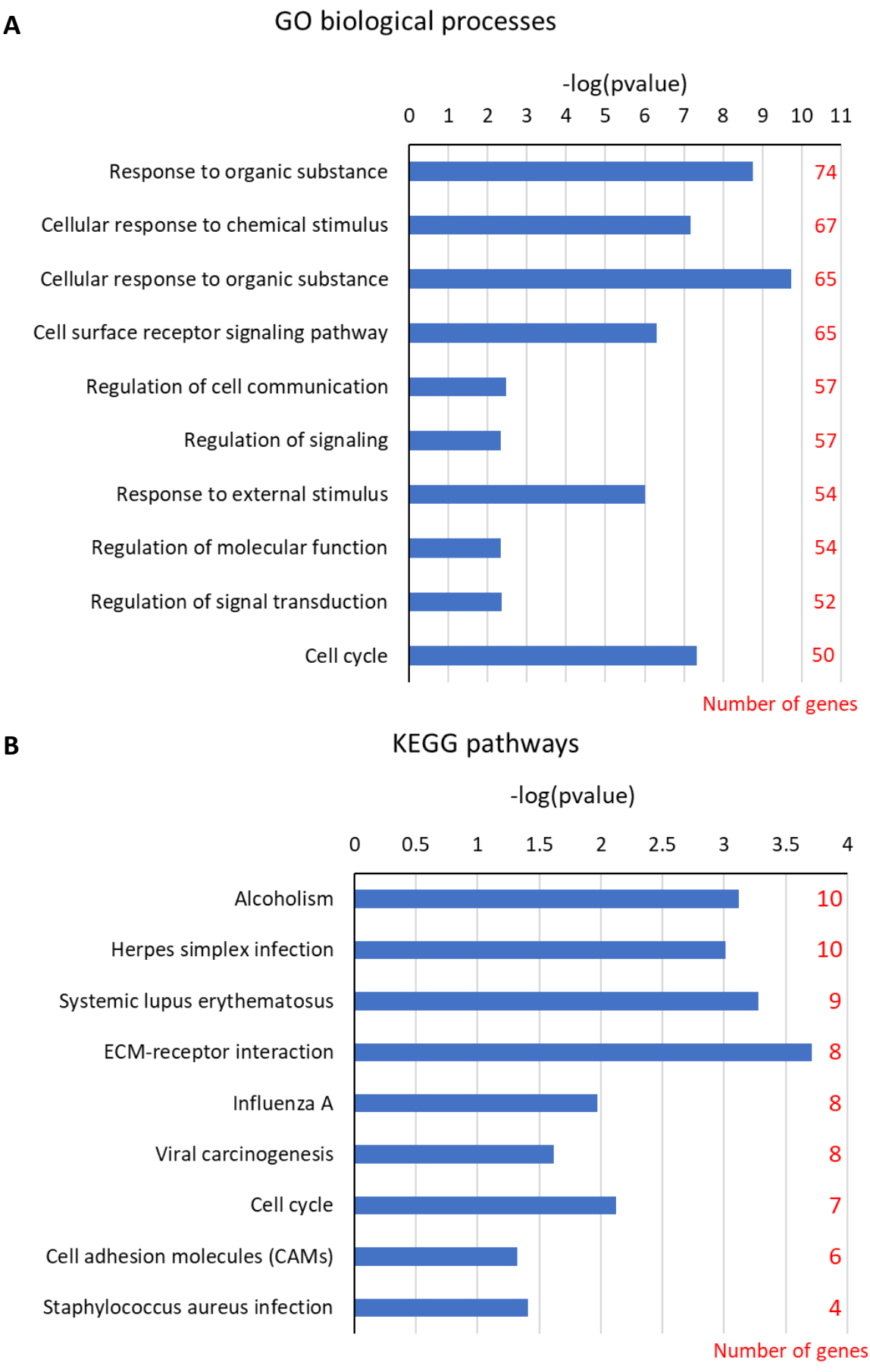


Figure 4.15 GO biological process and KEGG pathway analysis on the 2-fold or more significantly downregulated genes found by RNA-seq following in1382 infection

4.2.6 Analysing the gene changes that overlap during expression of the latency-associated ncRNAs by lentivirus transduction or quiescent HSV-1 infection

Although in1382 quiescence provides more faithful model of natural infection, the lentivirus delivery of the latency-associated ncRNAs allows for their investigation in isolation from the rest of the HSV-1 genome. Each method has provided information about the effects on the neuronal transcriptome (sections 4.2.1 – 4.2.5). However, combining these two methods will stringently assess the genes that change expression both as a direct result of the latency-associated ncRNAs but also during a quiescent latent-like HSV-1 infection. This is an alternative method, as opposed to fold-change, to narrow the focus of this investigation onto the gene changes most likely to be relevant during natural infection.

Therefore, further analysis was performed as shown schematically in figure 4.6, but this time to consider the significantly changed genes that overlapped as a result of expression of LAT and/or miRNAs, as well as during quiescent in1382 infection, and finally examining the physiological functions these genes are involved in. The Venn diagrams produced in Venny 2.1 (Oliveros, 2015) compare the DEGs resulting from latency-associated ncRNA lentivirus transductions and the in1382 infection to reveal the number of overlapping genes (figures 4.16 and 4.17).

The initial aim was to investigate any conserved gene changes elicited by the stable major LAT intron during latency. To do this, the gene changes that occurred during the LAT lentivirus transduction but not the NSLAT lentivirus transduction were compared to the gene changes observed during quiescent in1382 infection. These gene changes only occur in the presence of LAT splicing; therefore, expression of the major LAT intron must be necessary for these gene changes to occur during latency.

The Venny 2.1 comparison revealed that there were 2 significantly upregulated and 3 significantly downregulated overlapping genes as a result of LAT intron transduction and quiescent HSV-1 infection (figure 4.16A and 4.16B respectively). Whilst this is a modest number of genes, narrowing down of the once abundant gene list allows analysis of the function of these potentially key genes as outlined in table 4.7.

Although in terms of LAT the focus of this chapter was to examine the effect of the major LAT intron, as the most abundant and stable product during latency, the NSLAT lentivirus enables examination of the gene changes that occur as a result of LAT without the need for splicing of the major LAT intron. Therefore, the DEGs examined next were those that overlap as a result of the LAT intron lentivirus, NSLAT lentivirus and in1382 infection. These changes represent those that occur following LAT expression but unlike

those overlapping only following the LAT intron lentivirus and in1382 infection (table 4.7) do not rely upon the presence of the major LAT intron, as they also occur in the presence of the non-splicing variant of LAT delivered by the NSLAT lentivirus (figure 4.17C and table 4.8). The results show that there is just 1 gene that overlaps as a result of the LAT intron lentivirus transduction, NSLAT transduction and in1382 infection (figure 4.17C and table 4.8).

There were considerably more significant DEGs that overlapped between miRNA lentivirus transduction and quiescent in1382 infection with 102 upregulated genes and 67 downregulated genes (figure 4.16C and 4.16D respectively). See appendix 1, sheet 11 for full the full list of genes and their fold changes. These Venn diagrams also allow visualisation of the number of genes that do not overlap, demonstrating this is a highly conservative approach to analysing the data. Whilst this approach may potentially dismiss some gene changes that are important during natural infection, there can be far greater confidence in the DEGs conserved in both experimental methods. The DEGs that remain should be those that are altered during quiescent HSV-1 infection and as a direct result of the HSV-1 latency-associated ncRNAs. In this way, although stringent, these DEGs should be the most relevant to those that are altered by the latency-associated ncRNAs during natural HSV-1 latency.

There may be some redundancy between the roles of major LAT and of the miRNAs produced from the LAT region during latency, whereby major LAT and the miRNAs contribute to the same function. To test for this, the overlapping DEGs were also examined following transduction with the LAT intron lentivirus, the miR lentivirus and in1382 infection (figure 4.17A and 4.17B). It was discovered that 3 genes were significantly upregulated, and 1 gene significantly downregulated in response to the HSV-1 latency-associated miRNAs, major LAT and quiescent in1382 infection (figure 4.17A and 4.17B respectively). These 4 genes are named in table 4.9.

There are clearly transcriptional changes that occur to neuronal cells as a result of the HSV-1 latency associated ncRNAs during quiescent infection, with miRNA showing a larger effect (both in terms of the number of significant gene changes but also the magnitude of the fold changes elicited) than LAT. Comparing the changes from the lentivirus transductions and in1382 infection has reduced the number of significant gene changes to a number more manageable in terms of follow-up validation and for examining the functions of these genes in an effort to understand the impact that these gene changes may have on neuronal cell biology.

Figure 4.16 There are overlapping differentially expressed genes in differentiated SH-SY5Y cells transduced with lentiviruses expressing the latency-associated ncRNAs and those quiescently infected with in1382

(A) Venn diagram to show the number of significantly upregulated genes found by RNA-Seq that overlap between just the LAT intron lentivirus transduced differentiated SH-SY5Y cells and in1382 infected differentiated SH-SY5Y cells when compared to uninfected differentiated SH-SY5Y cells. Diagram generated using Venny 2.1 (Oliveros, 2015).

(B) Venn diagram to show the number of significantly downregulated genes found by RNA-Seq that overlap between just the LAT intron lentivirus transduced differentiated SH-SY5Y cells and in1382 infected differentiated SH-SY5Y cells when compared to uninfected differentiated SH-SY5Y cells. Diagram generated using Venny 2.1 (Oliveros, 2015).

(C) Venn diagram to show the number of significantly upregulated genes found by RNA-Seq that overlap between just the miR lentivirus transduced differentiated SH-SY5Y cells and in1382 infected differentiated SH-SY5Y cells when compared to uninfected differentiated SH-SY5Y cells. Diagram generated using Venny 2.1 (Oliveros, 2015).

(D) Venn diagram to show the number of significantly downregulated genes found by RNA-Seq that overlap between just the miR lentivirus transduced differentiated SH-SY5Y cells and quiescent in1382 infected differentiated SH-SY5Y cells when compared to uninfected differentiated SH-SY5Y cells. Diagram generated using Venny 2.1 (Oliveros, 2015).

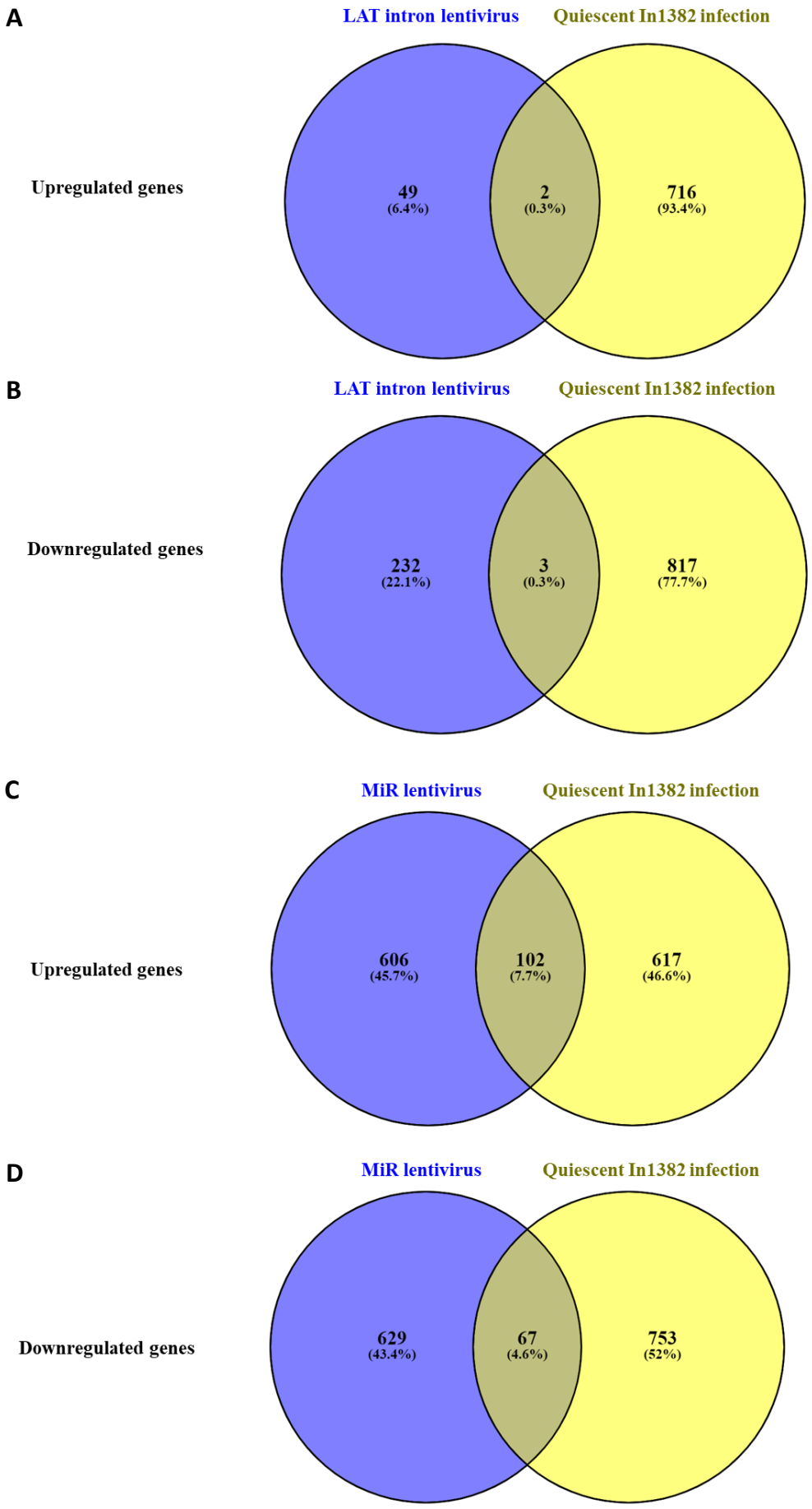


Figure 4.16 There are overlapping differentially expressed genes in differentiated SH-SY5Y cells transduced with lentiviruses expressing the latency-associated ncRNAs and those quiescently infected with in1382

Figure 4.17 Overlapping differentially expressed genes from differentiated SH-SY5Y cells transduced with different ncRNA expressing lentiviruses and those infected with in1382

(A) Venn diagram to show the number of significantly upregulated genes found by RNA-Seq that overlap between the LAT intron lentivirus transduced differentiated SH-SY5Y cells, the miR lentivirus transduced differentiated SH-SY5Y cells and in1382 infected differentiated SH-SY5Y cells when compared to uninfected differentiated SH-SY5Y cells. Diagram generated using Venny 2.1 (Oliveros, 2015).

(B) Venn diagram to show the number of significantly downregulated genes found by RNA-Seq that overlap between the LAT intron lentivirus transduced differentiated SH-SY5Y cells, the miR lentivirus transduced differentiated SH-SY5Y cells and in1382 infected differentiated SH-SY5Y cells when compared to uninfected differentiated SH-SY5Y cells. Diagram generated using Venny 2.1 (Oliveros, 2015).

(C) Venn diagram to show the number of significantly downregulated genes found by RNA-Seq that overlap between the LAT intron lentivirus transduced differentiated SH-SY5Y cells, the NSLAT lentivirus transduced differentiated SH-SY5Y cells and in1382 infected differentiated SH-SY5Y cells when compared to uninfected differentiated SH-SY5Y cells. Diagram generated using Venny 2.1 (Oliveros, 2015).

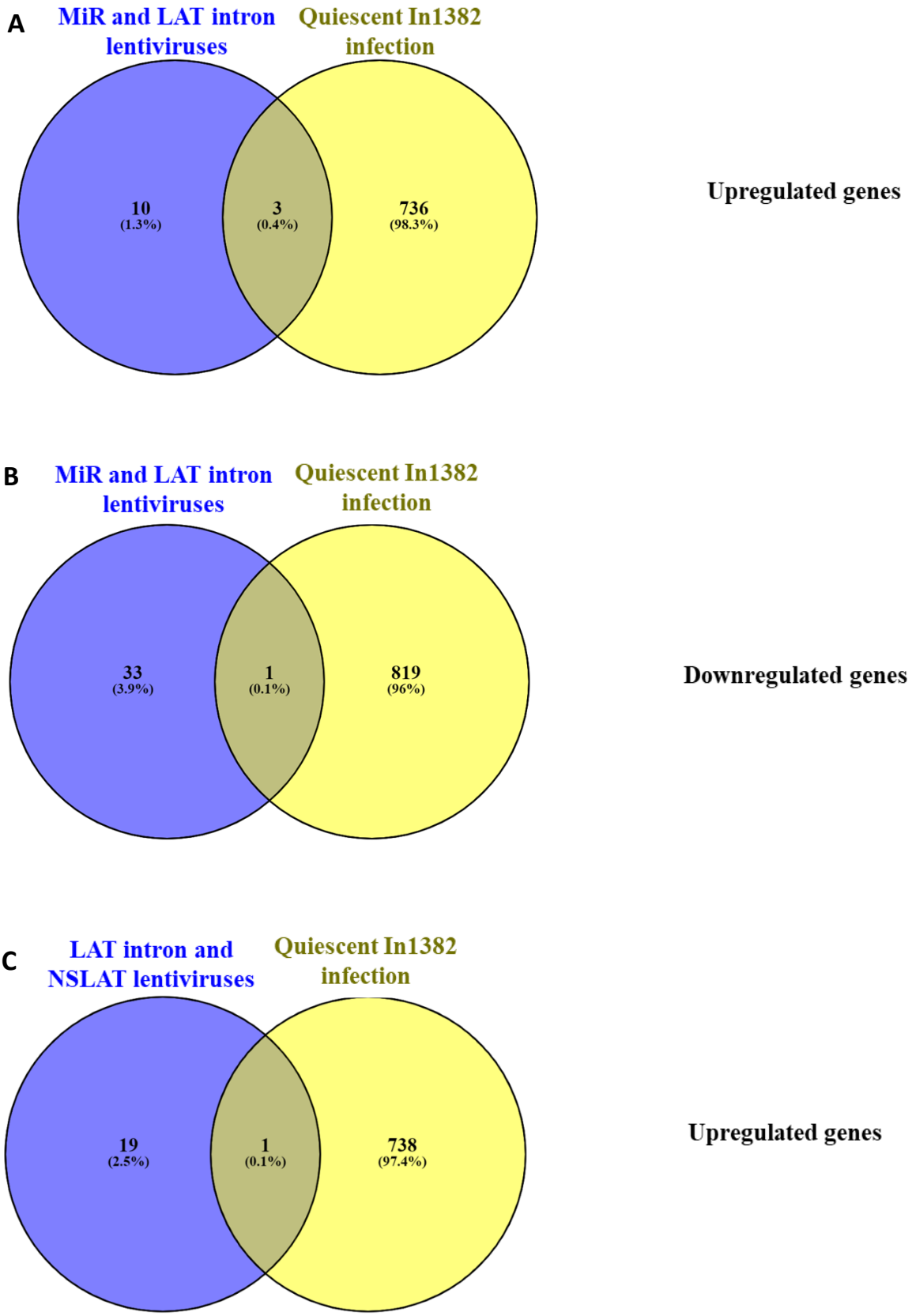


Figure 4.17 Overlapping differentially expressed genes from differentiated SH-SY5Y cells transduced with different ncRNA expressing lentiviruses and those infected with in1382

4.2.6.1 Assessing the functions of the overlapping genes that change during lentivirus delivery of the latency-associated ncRNA and quiescent in1382 infection

The RNA-Seq data has shown that there are neuronal transcriptome changes in response to the latency-associated ncRNAs present during HSV-1 latency. To understand the greater impact these transcriptional changes may have on the neurons and perhaps elucidate novel roles for the latency-associated ncRNAs, it is crucial to understand functions of the genes that are consistently differentially expressed. Filtering down the genes to those most physiological relevant to natural infection, either by fold-change or those that overlap in multiple conditions, provides a focus, whereby the functions of these significant DEGs can be assessed, to understand the further impact of the latency-associated ncRNAs.

To do this, either specific gene functions were checked by searching the NCBI gene database (NCBI, 2016) or pathway analysis was performed in DAVID (Huang et al., 2009a, 2009b) to see the broader biological processes that are affected by these transcriptional changes.

NCBI gene searches were performed for the known functions of the shortlisted DEGs (tables 4.7- 4.11). This includes:

- 5 DEGs overlapping between just the LAT intron lentivirus transduced and in1382 infected differentiated SH-SY5Y cells (figure 4.16A and 4.16B)
- 1 DEG (THBS4) overlapping between the LAT intron lentivirus transduction, NSLAT lentivirus transduction and in1382 infection (figure 4.17C)
- 4 DEGs overlapping between LAT intron lentivirus transduction, miRNA lentivirus transduction and in1382 infection (figure 4.17A and 4.17B)
- The top 5 upregulated DEGs (in terms of fold-change) that overlap between just the miR lentivirus transduction and in1382 infection
- The top 5 downregulated DEGs (in terms of fold-change) that overlap between just miR lentivirus transduction and in1382 infection

The results from these searches are highlighted in table 4.7, 4.8, 4.9, 4.10 and 4.11 respectively, along with the fold-change and adjusted p-value for each DEG (NCBI, 2016).

Table 4.7 The five overlapping DEGs between LAT lentivirus transduced and in1382 infected differentiated SH-SY5Y cells

The table includes the 5 differentially expressed genes (DEGs) found by RNA-Seq to be overlapping following LAT intron lentivirus transduced differentiated SH-SY5Y cells and in1382 infected differentiated SH-SY5Y cells as indicated in figure **4.16A** and **4.16B**. The table shows the fold change, p-adjusted value following Benjamini-Hochberg test and functions of the genes, as found in (NCBI, 2016) unless otherwise indicated.

Table 4.7 The five overlapping DEGs between LAT lentivirus transduced and in1382 infected differentiated SH-SY5Y cells

Gene	change	Fold change (2dp)		P adjusted Value (3 sf)		Known protein function
		In1382	**LAT**	**In1382**	**LAT**	
LPAR1	Up	1.66	1.13	0.00000473	0.0309	Lysophosphatidic acid receptor: • G-protein coupled receptor • Binds LPA inducing cell signalling • Involved in proliferation, smooth muscle contraction and chemotaxis
PCDH17	Up	1.37	1.22	0.00268	0.00549	Protocadherin 17: • Protocadherin • Tumour suppressor • Suggested role in cell-cell connections in the brain
RECQL4	Down	-1.75	-1.24	0.00481	0.00859	RecQ like helicase • DNA helicase
KCNAB3	Down	-1.31	-1.25	0.0241	0.0226	Potassium voltage-gated channel subfamily A regulatory beta subunit 3: • Beta subunit of a voltage-gated potassium channel
DACT1	Down	-1.33	-1.19	0.0196	0.0480	Dishevelled binding antagonist of beta catenin 1: • Dapper protein • Binding antagonist of beta catenin 1

Table 4.8 The overlapping DEG between LAT lentivirus transduced, NSLAT lentivirus transduced and in1382 infected differentiated SH-SY5Y cells

The table includes the differentially expressed gene (DEG) found by RNA-Seq to be overlapping following LAT intron lentivirus transduced, NSLAT lentivirus transduced and in1382 infected differentiated SH-SY5Y cells as indicated in **figure 4.17C**. The table shows the fold change, p-adjusted value following Benjamini-Hochberg test and functions of the genes, as found in (NCBI, 2016) unless otherwise indicated.

Table 4.8 The overlapping DEG between LAT lentivirus transduced, NSLAT lentivirus transduced and in1382 infected differentiated SH-SY5Y cells

Gene	change	Fold change (2dp)			P adjusted Value (3 sf)			Known protein function
		In1382	LAT	NSLAT	In1382	LAT	NSLAT	
THBS4	Up	1.40	1.19	1.24	8.77E-05	0.0201	0.00295	Thrombospondin 4: • Thrombospondin protein • An adhesive glycoprotein that mediates cell-to-cell and cell-to-matrix interactions • Binds heparin and calcium and is involved in local signalling and in the nervous system • Localization in β-amyloid containing plaques

Table 4.9 The four overlapping DEGs between LAT intron lentivirus transduced, miR lentivirus transduced and in1382 infected differentiated SH-SY5Y cells

The table includes the 4 differentially expressed genes (DEGs) found by RNA-Seq to be overlapping following LAT intron lentivirus transduced, miR lentivirus transduced and in1382 infected differentiated SH-SY5Y cells as indicated in **figure 4.17A** and **4.17B**. The table shows the fold change, p-adjusted value following Benjamini-Hochberg test and functions of the genes, as found in (NCBI, 2016) unless otherwise indicated.

Table 4.9 The four overlapping DEGs between LAT intron lentivirus transduced, miR lentivirus transduced and in1382 infected d ifferentiated SH-SY5Y cells

Gene	change	Fold change (2dp)			P adjusted Value (3 sf)			Known protein function
		In1382	LAT	MIR	In1382	LAT	MIR	
PDE4D	Up	1.44	1.19	1.16	0.0447	0.0174	0.0314	Phosphodiesterase 4D: • Has 3',5'-cyclic-AMP phosphodiesterase activity and degrades cAMP
GRIA2	Up	1.42	1.18	1.19	0.00537	0.00910	0.00521	Glutamate ionotropic receptor AMPA type subunit 2: • Encodes a subunit of the AMPA responsive glutamate receptors, which are ligand-activated cation channels and the predominant excitatory neurotransmitter receptor in the brain.
THSD7A	Up	1.36	1.19	1.22	0.0271	0.0105	0.00483	Thrombospondin type 1 domain containing 7A : • Interacts with alpha(V)beta (3) integrin and paxillin to inhibit endothelial cell migration • Putative role in cytoskeletal organisation.
YBX2	Down	-1.44	-1.24	-1.72	0.00113	8.85E-06	5.86E-26	Y-box binding protein 2: • Binds Y-box element in gene promoters to promote transcription

Table 4.10 The top 5 significantly upregulated genes according to fold-change, overlapping between miR lentivirus transduced and in1382 infected differentiated SH-SY5Y cells

The table includes the top 5 significantly upregulated genes, according to fold-change, found by RNA-Seq to be overlapping following miR lentivirus transduced and in1382 infected differentiated SH-SY5Y cells as indicated in **figure 4.16C**. The table shows the fold change, p-adjusted value following Benjamini-Hochberg test and functions of the genes, as found in (NCBI, 2016) unless otherwise indicated.

Table 4.10 The top 5 upregulated DEGs overlapping between miR lentivirus transduced and in1382 infected differentiated SH-SY5Y cells

Gene	Fold change (2dp)		P adjusted Value (3 sf)		Known protein function
	In1382	MiR	In1382	MiR	
PAPPA2	4.25	1.85	4.61E-22	2.46E-41	Pappalysin 2: •metzincin metalloproteinases •Cleaves insulin-like growth factor-binding protein 5 •Regulates insulin-like growth factor (IGF) bioavailability
TFPI2	3.32	2.31	7.31E-32	4.92E-66	Tissue factor pathway inhibitor 2: •Serine protease inhibitor •Tumour suppressor gene
CD34	2.47	3.11	0.000310	2.91E-07	CD34 molecule: •Hematopoietic stem/progenitor cell marker •May play a role in the attachment of stem cells to bone marrow ex-tracellular matrix or stromal cells
GDF15	3.13	1.54	2.46E-14	6.37E-07	Growth differentiation factor 15: •Secreted ligand of TGF-beta proteins •Binds TGF-beta receptors leading to recruitment of SMAD family transcription factors that regulate gene expression •Preproprotein which gets proteolytically processed to generate each subunit of a homodimer involved in the stress response program of cells following injury
GAD1	3.19	1.44	9.59E-14	0.00993	Glutamate decarboxylase 1: •Glutamic acid decarboxylase •Catalyses the production of gamma-aminobutyric acid from L-glutamic acid •Autoantigen in insulin-dependent diabetes

Table 4.11 The top 5 significantly downregulated genes according to fold-change, overlapping between miR lentivirus transduced and in1382 infected differentiated SH-SY5Y cells

The table includes the top 5 significantly downregulated genes, according to fold-change, found by RNA-Seq to be overlapping following miR lentivirus transduced and in1382 infected differentiated SH-SY5Y cells as indicated in **figure 4.16D**. The table shows the fold change, p-adjusted value following Benjamini-Hochberg test and functions of the genes, as found in (NCBI, 2016) unless otherwise indicated.

Table 4.11 The top 5 downregulated DEGs overlapping between miR lentivirus transduced and in1382 infected differentiated SH-SY5Y cells

Gene	Fold change (2dp)		P adjusted Value (3 sf)		Known protein function
	In1382	MiR	In1382	MiR	
BCAS1	-11.43	-1.87	4.04E-12	0.000343	Brain enriched myelin associated protein 1/Breast carcinoma-amplified sequence 1: •Required for myelination •Amplified in a variety of tumour types, especially breast cancer
FGL1	-4.10	-1.83	1.28E-09	8.51E-05	Fibrinogen like 1: •Member of the fibrinogen family •Suggested to have a role in the development of certain cancers sugh as hepatocellular carcinomas
CABP7	-1.90	-3.57	5.66E-05	2.97E-16	Calcium binding protein 7: •Member of the calmodulin superfamily •Interacts with and modulates phosphatidylinositol 4-kinase (McCue *et al., 2012)*
GPR149	-2.38	-2.97	0.00240	1.99E-05	G protein-coupled receptor 149: •G protein-coupled receptor •Important for G protein-coupled signal transduction
PCSK2	-3.64	-1.25	4.62E-17	9.40E-12	Proprotein convertase subtilisin/kexin type 2: •Protease that process protein and peptide precursors trafficking through secretory pathways

To further explore whether any of these genes and their known functions have been linked to HSV-1 previously, an advanced literature search was performed in PubMed for the gene of interest and “HSV” appearing anywhere in any field. None of the genes showed any known links to HSV-1 infection except CD34 and GAD1, for which there were 34 and 2 papers linking them respectively. This means that there could be novel associations to explore between the DEGs and the HSV-1 latency-associated ncRNAs. The purpose of assessing the neuronal transcriptome changes is to examine the impact that the HSV-1 latency-associated ncRNAs have on the neuron. As such, another PubMed search was performed with the term “neuron” and each gene of interest and 16/20 of the genes had papers linked to neurons, suggesting this could be a route worth exploring. The genes in papers linked to neurons were LPAR1, PCDH17, DACT1, KCNAB3, THSD7A, PDE4D, GRIA2, THBS4, PAPPA2, TFPI2, CD34, GDF15, GAD1, BCAS1, CABP7 and PCSK2 (see tables 4.7, 4.8, 4.9, 4.10 and 4.11 for full gene names and functions). Given their known functions and associations with neurons, the potential implications of the expression changes of some of these DEGs during latency are considered in the discussion section.

In terms of fold-changes, most are relatively low, which will have an impact on the size of effect elicited by these gene changes. The LAT intron lentivirus transduction and in1382 infection overlapping DEGs have an average fold-change up or down of 1.35 (table 4.7), the miR and LAT intron lentivirus transductions and in1382 infection overlapping DEGs showed an average fold-change up or down of 1.31 (table 4.9) and the LAT intron and NSLAT lentivirus transductions and in1382 infection overlapping DEG, thrombospondin 4 (THBS4), showed an average fold-change of 1.3 (table 4.8). Although these changes elicited here are small, there is stringency in the approach as they reproducibly occur in the multiple methods of expressing the latency-associated ncRNAs. Nevertheless, among the miR lentivirus and in1382 infection overlapping genes there were larger fold-changes, with 3 genes exhibiting over 2-fold changes in transcription across both conditions (tables 4.10 and 4.11). Therefore, the functions of these 3 genes – Tissue factor pathway inhibitor 2 (TFPI2), CD34 and GPR149 (highlighted in tables 4.10 and 4.11), may be worth looking more closely at as the effect elicited by these genes might be greater.

To examine the cumulative functions of the 169 DEGs that overlapped between the MiR lentivirus transduction and quiescent in1382 infection (figures 4.16C, 4.16D)(appendix 1, sheet 11) DAVID analysis was performed on these genes as described previously. Downregulated genes and upregulated genes were separated as described in figure 4.6 and DAVID software utilised to perform for GO biological processes analysis and KEGG

pathways analysis to see which processes and pathways these genes are involved in. These were filtered according to significance (p-value < 0.05) and then ordered according to the processes or pathways with the highest number of DEGs involved. The graphs show the results in terms of -log10(p-value) and where there were more than 10 processes or pathways found, the top 10 were presented (figures 4.18 and 4.19).

The top GO biological process found to be affected from the 102 genes upregulated during miR lentivirus transduction and quiescent in1382 infection was the protein-modification process, with 32 of the significantly upregulated genes involved. Other than this, the main process enriched for were those related to signal transduction and metabolic processes (figure 4.18A). Lysosome was the top KEGG pathway enriched with the upregulated DEGs overlapping between miR lentivirus transduction and in1382 infection but none of the KEGG pathways enriched had more than 5 of the DEGs involved (figure 4.18B).

Most of the processes associated with the 67 downregulated DEGs overlapping between the miR lentivirus transduction and in1382 infection were related to the cell cycle or cell division (figure 4.19A). The cell cycle was also one of the KEGG pathways that the downregulated genes were shown to be associated with. However, as with the KEGG pathway analysis on the significantly upregulated genes, there were not many downregulated genes associated with each KEGG pathway, none with more than 4 genes associated (figure 4.19B)(see appendix 2 for full pathway analysis gene lists).

There are many different functions and biological pathways associated with the DEGs that change in response to the HSV-1 latency-associated ncRNAs. Understanding these functions is the first step to better understanding the downstream physiological effects that the HSV-1 latency-associated ncRNAs may have on human neurons.

Figure 4.18 GO biological process and KEGG pathway analysis on the significantly upregulated genes overlapping in miR lentivirus transduced differentiated SH-SY5Y cells and in1382 infected differentiated SH-SY5Y cells

(A) A bar graph to show the gene ontology (GO) biological processes following analysis using DAVID software (Huang et al., 2009a, 2009b) on the 102 significantly upregulated genes overlapping in miR lentivirus transduced and in1382 infected differentiated SH-SY5Y cells listed as listed in **appendix 1**. The graph shows the top 10 significantly associated GO biological processes in terms of number of differentially expressed genes (DEGs), which are indicated in red. The blue bars show the -log10(p-values) following the fisher exact statistics method used by the DAVID software to calculate the p-value for each process.

(B) Bar graph to show the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways found by DAVID analysis (Huang et al., 2009a, 2009b) to be significantly associated with the 102 significantly upregulated genes overlapping in miR lentivirus transduced and in1382 infected differentiated SH-SY5Y cells, as listed in **appendix 1**. The number of DEGs involved are indicated in red. The blue bars show the -log10 p-values following the fisher exact statistics used in the DAVID analysis software for each pathway.

See **appendix 2** for full gene lists associated with each GO biological process/KEGG pathway.

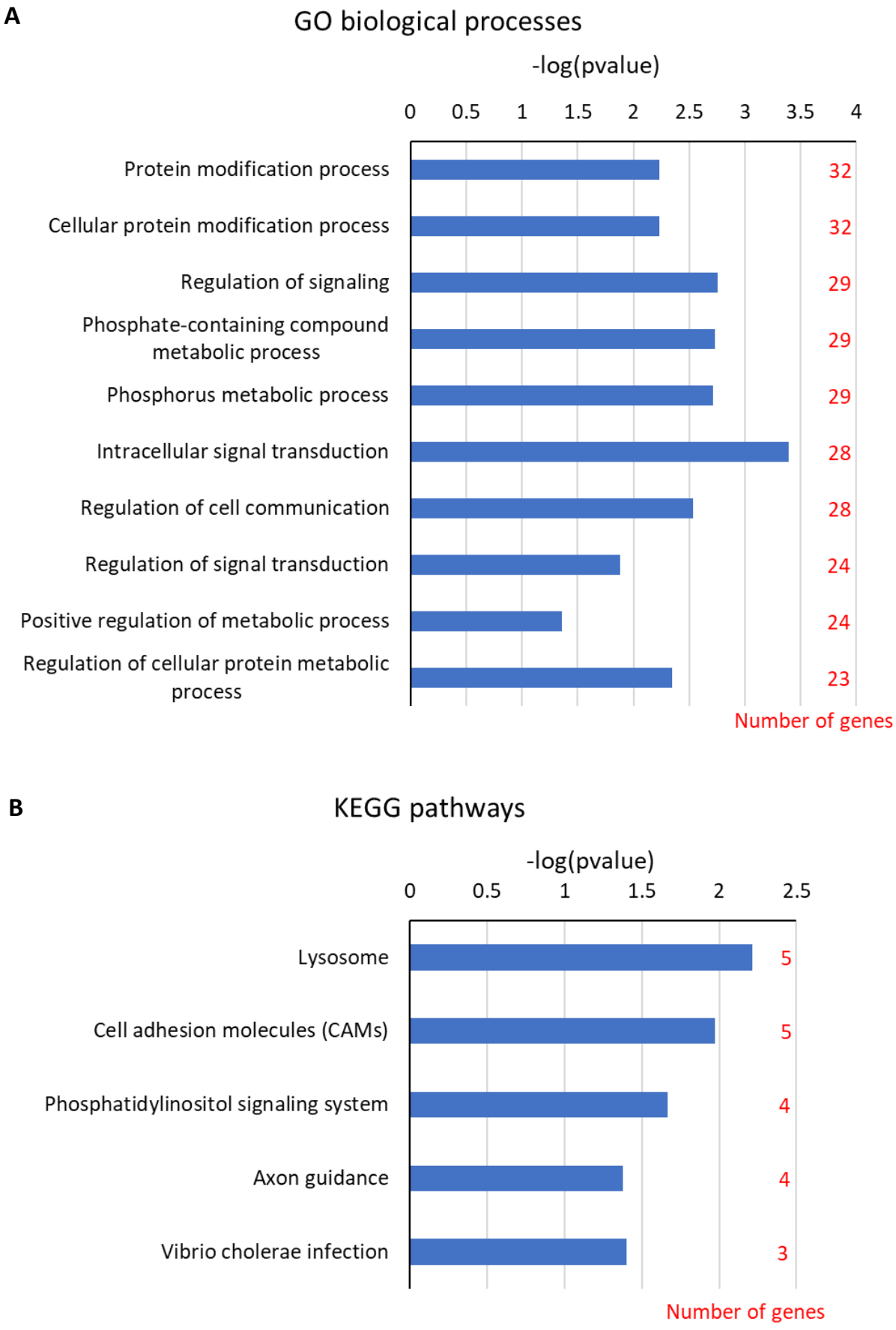


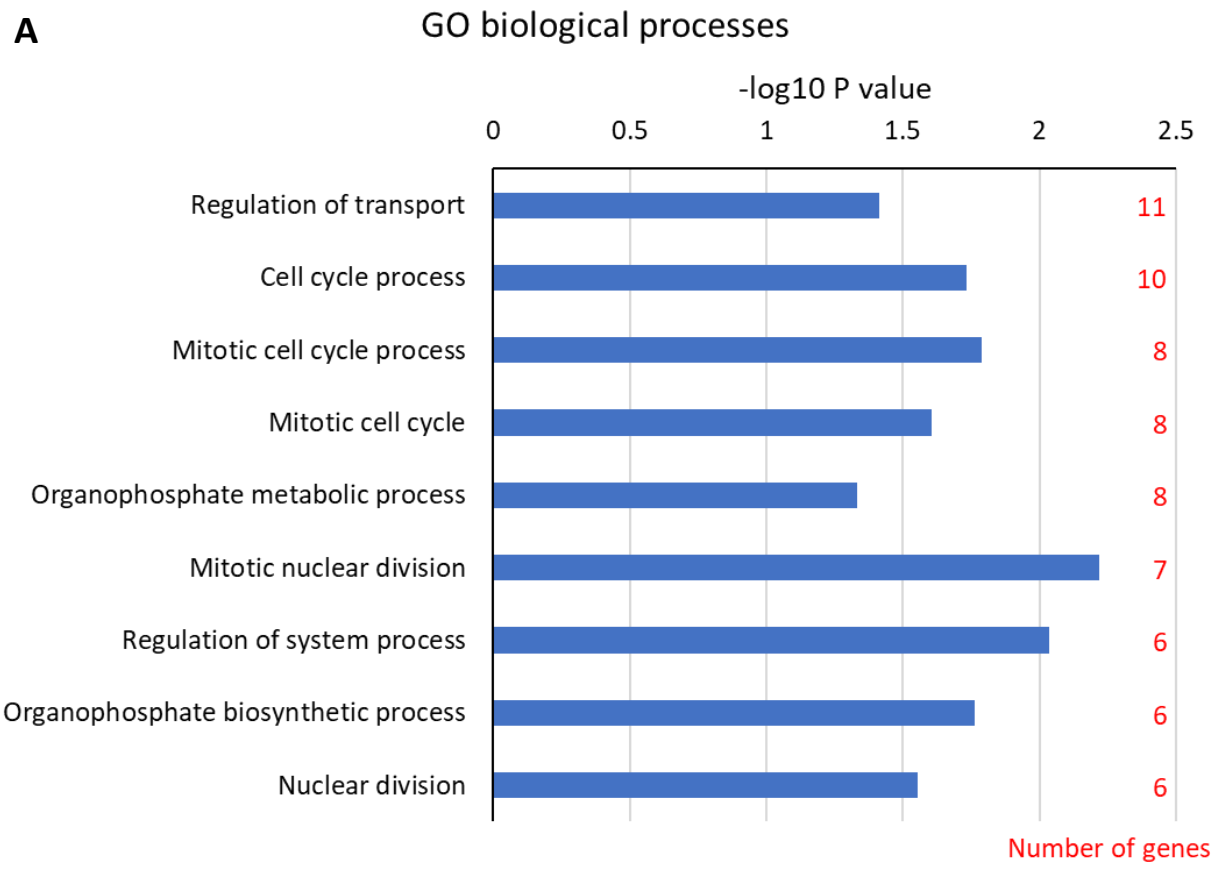
Fig 4.18 GO biological process and KEGG pathway analysis on the significantly upregulated genes overlapping in either miR lentivirus transduced or in1382 infected differentiated SH-SY5Y cells

Figure 4.19 GO biological process and KEGG pathway analysis on the significantly downregulated genes overlapping in miR lentivirus transduced differentiated SH-SY5Y cells and in1382 infected differentiated SH-SY5Y cells

(A) A bar graph to show the gene ontology (GO) biological processes following analysis using DAVID software (Huang et al., 2009a, 2009b) on the 67 significantly downregulated genes overlapping in miR lentivirus transduced and in1382 infected differentiated SH-SY5Y cells, as listed in **appendix 1**. The graph shows the top 10 significantly associated GO biological processes in terms of number of differentially expressed genes (DEGs), wshich are indicated in red. The blue bars show the -log10 p-values following the fisher exact statistics method used by the DAVID software to calculate the p-value for each process.

(B) Bar graph to show the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways found by DAVID analysis (Huang et al., 2009a, 2009b) to be significantly associated with the 67 significantly downregulated genes overlapping in miR lentivirus transduced and in1382 infected differentiated SH-SY5Y cells, as listed in **appendix 1**. The number of DEGs involved are indicated in red. The blue bars show the -log10 p-values following the fisher exact statistics used in the DAVID analysis software for each pathway.

See **appendix 2** for full gene lists associated with each GO biological process/KEGG pathway.



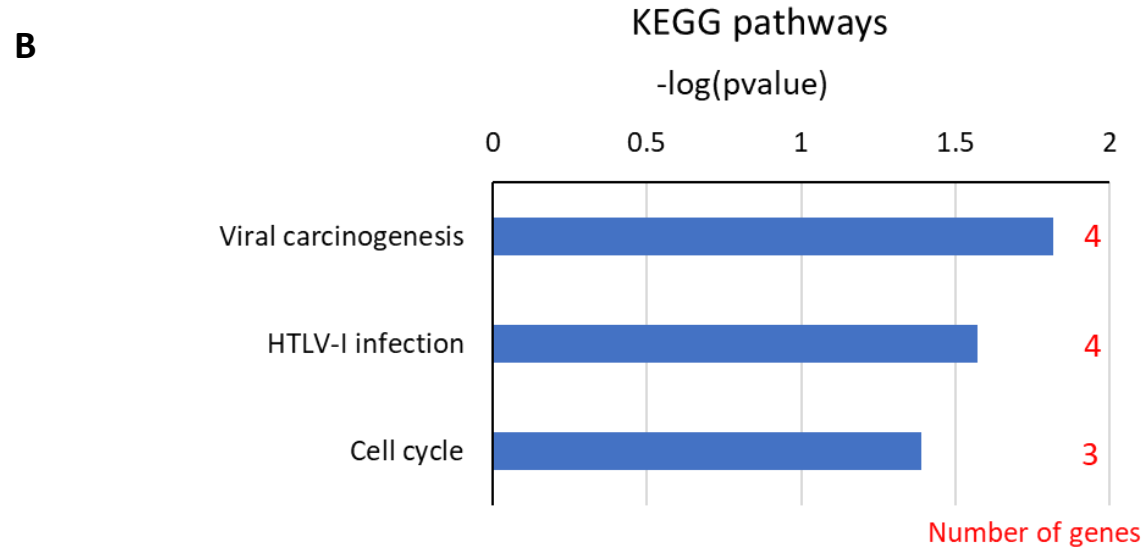


Fig 4.19 GO biological process and KEGG pathway analysis on the significantly downregulated genes overlapping in miR lentivirus transduced and in1382 infected differentiated SH-SY5Y cells

4.2.7 Further examination of how the HSV-1 latency-associated miRNAs mediate their effect on the neuronal transcriptome

In the previous sections it was found that the miRNAs elicit a stronger effect on the neuronal transcriptome than major LAT does. The DAVID software analysis then revealed a variety of biological pathways that are altered as a result of these transcriptional changes. Given the potentially substantial physiological impact of the DEGs resulting from expression of the HSV-1 latency miRNAs, in conjunction with the various ways in which miRNAs are known to regulate gene expression (Catalanotto et al., 2016)(chapter 1.4.1), it is useful to consider how the miRNAs are eliciting their effect on the neuronal transcriptome.

Additional DAVID analyses were run to examine whether any transcription factors were associated with the DEGs overlapping as a result of miR lentivirus transduction or in1382 infection. 172 and 58 transcription factors were found to be associated with the upregulated and downregulated DEGs respectively. The transcription factors analysis found transcription factors associated with 98 of the 102 upregulated DEGs and 58 of the 67 downregulated DEGs (see appendix 2 for complete DAVID analysis table). The top 10 transcription factors significantly associated with the most DEGs are shown in tables 4.12 and 4.13 for the significantly upregulated and downregulated genes respectively. AREB6 (also known as ZEB1 – zinc finger E-box binding homeobox 1) was the top transcription factor for the upregulated DEGs, with 91 significantly associated DEGs (table 4.12). Paired box 4 (PAX4) was the top transcription factor for the downregulated DEGs, with 48 significantly associated DEGs (table 4.13). As a result, it could be that instead of directly mediating transcriptional changes on the 169 DEGs, the HSV-1 miRNAs may alter the levels of these transcription factors that, in turn, mediate transcriptional changes.

Nevertheless, when investigating these transcription factors in the RNA-Seq dataset, the miRNAs do not elicit a significant effect on the transcriptional levels of all but one of these transcription factors (appendix 2, sheet 3). The only transcription factor that was significantly differentially expressed following miR lentivirus transduction, with a false discovery rate of less than 5% was myocyte enhancer factor 2A (MEF2). This transcription factor was found to be associated with both the upregulated and downregulated DEGs (tables 4.12 and 4.13), therefore it is not specific to either. Nevertheless, it is possible for a transcription factor can cause the upregulation of certain genes and the downregulation of others (Latchman, 2001). There were also 2 transcription factors, CDP (also known as CUX1 – cut like homeobox 1) and POU class

3 homeobox 2 (POU3F2), which were specifically associated with the significantly downregulated genes (table 4.13), that although not significantly differentially expressed with a false discovery rate of 5%, their original p-values were < 0.05 (both 0.008) and their adjusted p-values following Benjamini-Hochberg adjustment were just over the threshold set of 0.05 (0.058 and 0.052 respectively). It is possible that the miRNAs are impacting the transcription of these 3 transcription factors, which in turn elicits the effect on the DEGs observed. However, it is also possible that, despite significant association with transcription factors, the miRNAs elicit their effects on the DEGs directly. Alternatively, HSV-1 miRNAs could elicit their effect on the transcription factors or DEGs post-transcriptionally.

Table 4.12 Transcription factors associated with miR lentivirus and in1382 infection overlapping upregulated DEGs

Transcription factor	Count	List Total	PValue
AREB6	91	98	8.92E-07
OCT1	90	98	0.000131
GATA1	87	98	1.04E-05
AML1	85	98	4.64E-06
PAX4	85	98	7.99E-06
PPARG	84	98	0.000357
MEF2	82	98	1.05E-05
EVI1	82	98	0.002695
YY1	80	98	4.3E-06
MYCMAX	79	98	1.85E-08

A table to show the results from transcription factor analysis using DAVID analysis software (Huang et al., 2009a, 2009b). The table shows the top 10 transcription factors significantly associated with the most significantly upregulated genes overlapping in miR lentivirus transduced and in1382 infected differentiated SH-SY5Y cells. The table show the transcription factor, the number of DEGs associated with it (count), the total number of upregulated DEGs that were included in the transcription factor analysis results and the p-value representing the association of the transcription factor with its DEGs. See **appendix 2, sheet 17** for complete list of associated DEGs.

Table 4.13 Transcription factors associated with miR lentivirus and in1382 infection overlapping downregulated DEGs

Transcription factor	Count	List Total	PValue
PAX4	48	58	0.006971
MEF2	46	58	0.008545
CDP	41	58	0.002011
RFX1	39	58	0.014337
AP4	39	58	0.038821
PAX5	37	58	0.012111
POU3F2	37	58	0.029929
EN1	35	58	0.016338
TAXCREB	34	58	0.039588
BRACH	34	58	0.046666

A table to show the results from transcription factor analysis using DAVID analysis software (Huang et al., 2009a, 2009b). The table shows the top 10 transcription factors significantly associated with the most significantly downregulated genes overlapping in miR lentivirus transduced and in1382 infected differentiated SH-SY5Y cells. The table show the transcription factor, the number of DEGs associated with it (count), the total number of downregulated DEGs that were included in the transcription factor analysis results and the p-value representing the association of the transcription factor with its DEGs. See **appendix 2, sheet 18** for complete list of associated DEGs.

4.2.7.1 RNAhybrid analysis of MiRNA targeting of host genes

MiRNAs are known to mediate post-transcriptional degradation or translational inhibition of gene targets that have partially complementary sequence through binding of its 3 prime (3') untranslated region (UTR) (Huntzinger & Izaurralde, 2011; Orang et al., 2014; Selbach et al., 2008). Therefore, the HSV-1 miRNAs may mediate a post-transcriptional effect on the cellular mRNAs, that could affect gene expression and/or protein levels. To assess whether the 5 HSV-1 miRNAs delivered by the miR lentivirus have the capacity to bind to and mediate downregulation of the DEGs post-transcriptionally, RNAhybrid software was utilised to predict putative binding sites for the HSV-1 miRNAs (Krüger & Rehmsmeier, 2006; Rehmsmeier et al., 2004).

To test potential binding capacity or "hits" of the 5 latency-associated miRNAs expressed during quiescent infection and delivered by the miR lentivirus, a subset of the DEGs found to overlap as significantly downregulated following either miR lentivirus transduction or in1382 infection were tested as potential targets for the miRNAs. Additionally, one of the transcription factors, CDP (CUX1), that was revealed through the DAVID analysis described to be associated with 41 downregulated DEGs (table 4.13) was tested as a binding partner for the latency-associated miRNAs. RNAhybrid software was used, to find the minimum free energy hybridization of a long (target) and a short (query) RNA, which in this case is the 3' UTRs of the DEGs or associated transcription factor and each HSV-1 miRNA respectively (Krüger & Rehmsmeier, 2006; Rehmsmeier et al., 2004). Dr Michael Nicoll, division of virology, NIBSC, generated control miRNA sequences by inputting each miRNA sequence into a sequence scrambling tool (GenScript, n.d.).

These new "false" scrambled miRNA sequences were input alongside the 5 HSV-1 miRNA sequences (both 5p and 3p forms) to control for any false positive predicted binding to the sequences on the basis of specific nucleotide frequency, instead of sequence. The scrambled miRNAs are named "false 2-7" according to which miRNA was scrambled to generate it. All the miRNA sequences were input into RNAhybrid alongside the 3' UTRs of a selection of 8 downregulated DEGs that overlapped following either miR lentivirus transduction or in1382 infection, as well as CDP(CUX1). The first test was performed with parameters that requires the miRNA to bind specifically between its short seed sequence (nucleotides 2-7), with no G:U pairing, to each of the 3' UTR(s) for each gene inputted.

The results from this first test are shown in table 4.14 and reveal that the 3' UTR of the transcription factor CDP(CUX1) along with downregulated DEGs MYC target 1

(MYCT1), brain enriched myelin associated protein 1/breast carcinoma-amplified sequence 1 (BCAS1), radial spoke head component 4A (RSPH4A), troponin T1 (TNNT1) and prostaglandin E receptor 3 (PTGER3) all have at least one hit with at least 1 miRNA each. The transcription factor CDP has the most hits with the most miRNAs, showing hits with all the miRNA pairs. However, there were also a total of 6 false positive hits, 4 with CDP(CUX1), found using the scrambled miRNA sequences, suggesting the binding prediction may not be specific to any of the HSV-1 miRNA sequences.

To try to decrease the false-positive binding predictions, the test was re-run with stricter parameters, whereby the miRNA must bind specifically between its full seed sequence (nucleotides 2-8), with no G:U pairing, to the 3' UTR region of each gene.

The results are shown in table 4.15. There was a reduction in the number of hits found with the scrambled "false" miRNAs but also less hits among the DEGs. From this test CDP(CUX1), MYCT1, BCAS1 and RSPH4A showed hits with at least 2 miRNAs but only CDP showed multiple hits with any one miRNA. Both miR-H2-5p and miR-H6-5p had 2 predicted hits with CDP (table 4.15 and figure 4.20). MiR-H7-3p also had 1 hit with the transcription factor, but as the false 7 miRNA sequence also had 1 hit, this and the hits found with BCAS2 and RSPH4A may not predict true binding. These data suggested that at least 2 of the latency-associated miRNAs target the transcription factor, CDP(CUX1), which from the transcription factor DAVID analysis is seen to be associated with 41 of the downregulated DEGs (table 4.13) (appendix 2, sheet 18). This could provide insight into how the miRNAs are eliciting such a large response on the neuronal transcriptome. The predicted binding sites of the hits found between CDP(CUX1) and the HSV-1 miRNAs (including those with just one predicted hit) are shown in figure 4.20.

The latency-associated ncRNAs appear to have a significant effect on the neuronal transcriptome, with an array of potential physiological implications. The miRNAs seem to elicit some of their effect through direct interaction of gene targets, as shown by RNAhybrid analysis. It is possible in this way miRNA regulate transcription factors such as CDP(CUX1) that, in turn, mediate transcriptomic changes. Before embarking on significant follow-up work to test the downstream physiological effects from these interactions and the transcriptome changes, the gene changes observed by RNA-Seq were validated by qRT-PCR.

Table 4.14 RNA-hybrid analysis requiring binding of 6 nucleotides reveals host cellular targets for HSV-1 latency-associated miRNAs

A table indicating the number of hits found by RNAhybrid (Krüger & Rehmsmeier, 2006; Rehmsmeier et al., 2004) corresponding to a predicted binding between nucleotides 2-7 of the HSV-1 latency-associated miRNAs and the 3' UTR(s) of host genes (not including G:U pairing). These host genes are a subset of those found to be significantly downregulated when performing RNA-Seq on either miR lentivirus transduced differentiated SH-SY5Y cells or in1382 infected differentiated SH-SY5Y cells, plus a transcription factor (CDP) found to be associated with a significant number of these downregulated DEGs by DAVID analysis (Huang et al., 2009a, 2009b). False 2,3,4 and 7 refer to scrambled miRNAs of their number HSV-1 miRNA counterparts (i.e. false 2 is scrambled HSV-1 miR-H2) to act as control miRNA sequences for false positive hits. Where there are multiple hit numbers separated by a slash, multiple 3' UTRs were tested for the one gene and hits came up in both where indicated.

Table 4.14 RNA-hybrid analysis requiring binding of 6 nucleotides reveals host cellular targets for HSV-1 latency-associated miRNAs

	CDP (Cux1)	GPR149	MYCT1	KCNS3	BCAS1	RSPH4A	TNNT1	PTTG1	PTGER3
H2-3p	4/1				1	1			
H2-5p	2								
H3-3p					1		1		
H3-5p	4		1/1		2				
H4-3p	2		1/1						
H4-5p			1		1				
H6-3p					1				
H6-5p	4	0	1	0		1		0	1
H7-3p	1		1						
H7-5p	4/1					1			
false 2									
false 3					1				
false 4	3		1			1			
false 7	1								

Table 4.15 RNA-hybrid analysis requiring binding of 7 nucleotides still reveals host cellular targets for HSV-1 latency-associated miRNAs

A table indicating the number of hits found by a repeat RNAhybrid analysis (Krüger & Rehmsmeier, 2006; Rehmsmeier et al., 2004), this time corresponding to a predicted binding between nucleotides 2-8 of the HSV-1 latency-associated miRNAs and the 3' UTR of host genes (not including G:U pairing). These host genes are a subset of those found to be significantly downregulated when performing RNA-Seq on either miR lentivirus transduced differentiated SH-SY5Y cells or in1382 infected differentiated SH-SY5Y cells, plus a transcription factor (CDP) found to be associated with a significant number of these downregulated DEGs by DAVID analysis (Huang et al., 2009a, 2009b). False 2,3,4 and 7 refer to scrambled miRNAs of their number HSV-1 miRNA counterparts (i.e. false 2 is scrambled HSV-1 miR-H2) to act as control miRNA sequences for false positive hits.

Table 4.15 RNA-hybrid analysis requiring binding of 7 nucleotides still reveals host cellular targets for HSV-1 latency-associated miRNAs

	CDP (Cux1)	GPR149	MYCT1	KCNS3	BCAS1	RSPH4A	TNNT1	PTTG1	PTGER3
H2-3p		0		0	1	1	0	0	0
H2-5p	2								
H3-3p					1				
H3-5p					1				
H4-3p									
H4-5p									
H6-3p									
H6-5p	2		1						
H7-3p	1								
H7-5p						1			
false 2									
false 3					1				
false 4			1			1			
false 7	1								

Figure 4.20 The predicted binding sites between CDP (CUX1) and the latency-associated HSV-1 miRNAs

A schematic representation of the predicted binding sites between the host transcription factor CDP(CUX1) and the latency-associated HSV-1 miRNAs as determined by RNAhybrid (Krüger & Rehmsmeier, 2006)

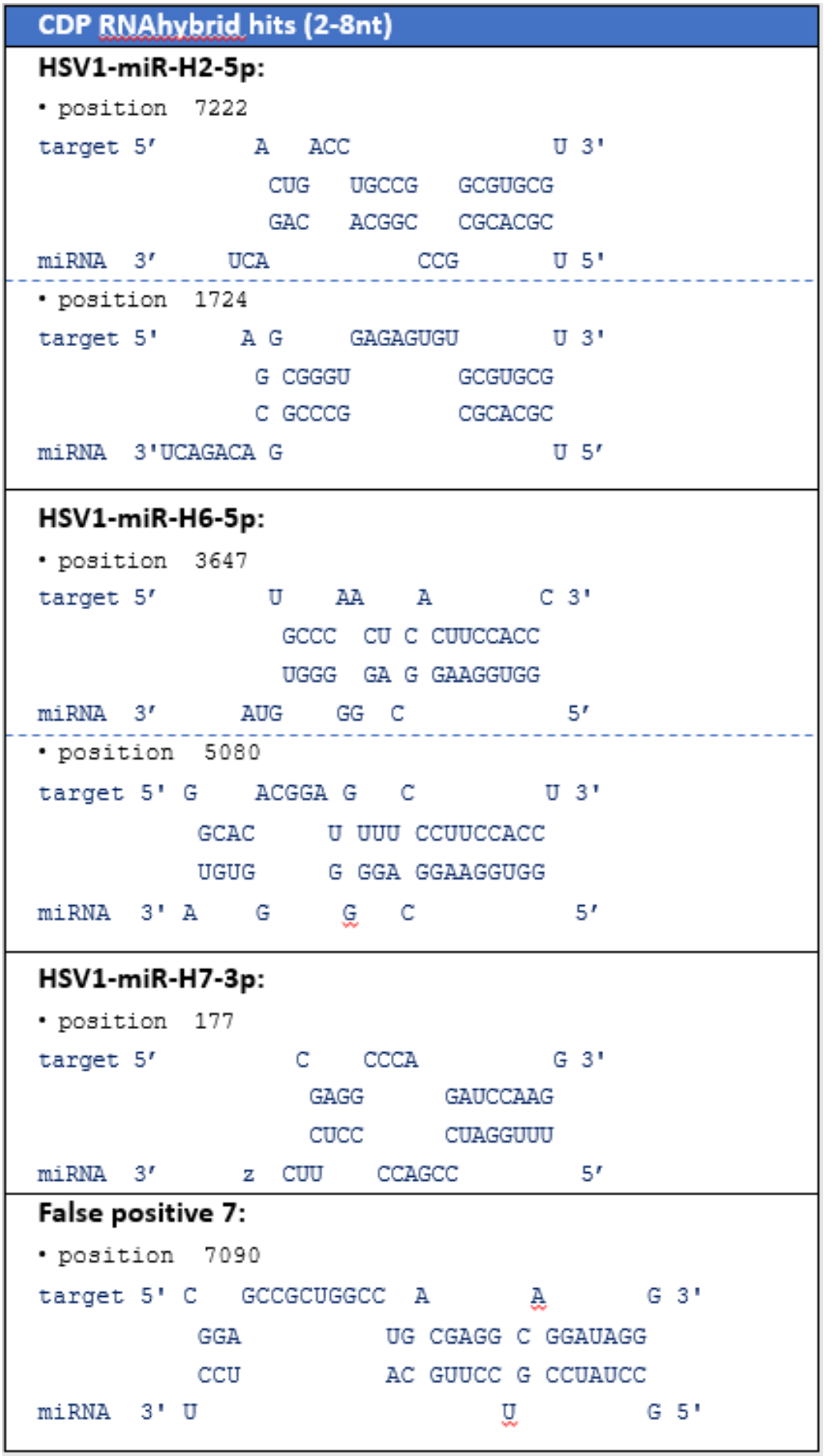


Figure 4.20 The predicted binding sites between CDP (CUX1) and the latency-associated HSV-1 miRNAs

4.2.8 Independent validation of transcriptome changes by qRT-PCR

To validate the RNA-Seq results observed and the reproducibility of the changes observed, qRT-PCR was performed to a subset of the genes found to be significantly differentially expressed in both the lentivirus and in1382 RNA-Seq experiments as outlined above (figures 4.16 and 4.17) .

Initially, qRT-PCR was performed towards a subset of the DEGs on the same RNA samples extracted for the RNA-Seq experiments (figures 4.1A and 4.11A) to validate the RNA-Seq methodology. The ΔΔCt formula ($2^{-\Delta\Delta Ct}$) (methods chapter 6.2.5.3) was used to calculate a fold-change when comparing the transcription in the transduced or infected cells to mock differentiated SH-SY5Y cells. This fold change was then compared to that found using RNA-Seq (table 4.16).

For the overlapping DEGs observed in LAT intron lentivirus and in1382 infections, most of the fold-changes observed by qRT-PCR were similar to that found in the initial RNA-Seq experiments (table 4.16A). For example, protocadherin 17 (PCDH17) was found to be 1.37-fold upregulated in in1382 infected differentiated SH-SY5Y cells by RNA-Seq and 1.32-fold upregulated by q-RT-PCR (table 4.16A). RecQ like helicase (RECQL4) was found to be 1.75-fold downregulated in in1382 infected SH-SY5Y cells by RNA-Seq and 1.82-fold downregulated by qRT-PCR. There was just 1 DEG – Potassium voltage-gated channel subfamily A regulatory beta subunit 3 (KCNAB3), that did not follow the pattern of transcriptional change from RNA-Seq to qRT-PCR testing. KCNAB3 was shown to be 1.3-fold downregulated in LAT-transduced differentiated SH-SY5Y cells by RNA-Seq but 1.5-fold upregulated when examined by qRT-PCR. However, this upregulation seems to be an outlier because the in1382 infected cells have a similar 1.31 and 1.38 downregulation when measured by RNA-Seq or qRT-PCR respectively (table 4.16A).

The table also allows comparison between the fold-changes observed during LAT intron lentivirus transduction and those observed following infection with in1382. The fold-changes are similar regardless of the route of LAT delivery. Generally, there seems to be considerable correlation between not only the pattern of expression in terms of upregulation or downregulation when assessing either LAT intron lentivirus transduced differentiated SH-SY5Y cells or in1382 infected differentiated SH-SY5Y cells by either RNA-Seq or PCR, but also in terms of the magnitude of the change. However, they are all small fold-changes, close to 1 (table 4.16A).

A similar comparison and follow up qRT-PCRs were performed on the miR lentivirus transduced SH-SY5Y RNA samples for a subset of the DEGs that were significantly downregulated overlapping following either miR lentivirus transduction or in1382 infection (Table 4.16B).

Firstly, the fold-changes across the 5 DEGs were consistently downregulated across all the genes tested by RNA-Seq and qRT-PCR on the miR lentivirus transduced RNA. The magnitude of this downregulation was also similar for most of the genes, for example MYCT1 was shown to be 1.92-fold downregulated by RNA-Seq and 1.72-fold downregulated by qRT-PCR. However, there is some variation in the magnitudes of the fold-changes observed by either RNA-Seq or PCR. This is the case for GPR149 where RNA-Seq revealed an almost 3-fold decrease following miR lentivirus transduction of differentiated SH-SY5Y cells, whereas qRT-PCR showed an almost 5-fold decrease. Therefore, although the fold-changes are consistent in terms of direction of change for all the DEGs using RNA-Seq or PCR, there does seem to be variation in the sensitivity of each method, resulting in some variation in the magnitude of change observed for a specific gene.

The magnitude of changes did also show some differences when comparing the fold-change observed in the original RNA-Seq experiment in miR lentivirus transduced cells or in1382 infected cells. In the miR lentivirus transduced differentiated SH-SY5Y cells, BCAS1 was 1.87-fold downregulated whereas in the in1382 infected SH-SY5Y cells there was an 11.43-fold downregulation, both shown by RNA-Seq (table 4.16B). Therefore, it is possible that in1382 has a more significant impact on the neuronal transcriptome. Nevertheless, there were genes that changed by a similar magnitude in both miR lentivirus transduced cells and in1382 infected cells. For example, RNA-Seq revealed that RSPH4A was 1.68-fold downregulated in miR lentivirus transduced cells and 1.61-fold downregulated in in1382 infected cells (table 4.16B).

In conclusion, from both sets of data (table 4.16A and B), it appears that the qRT-PCR data largely validates the data observed in the RNA-Seq experiment, with almost completely consistent directions of change and mostly similar magnitudes of change.

Next, the reproducibility of the significant gene changes observed was assessed. In addition to comparing the RNA-Seq and qRT-PCR consistency among overlapping DEGs in the same RNA samples, qRT-PCR was performed on separately transduced differentiated SH-SY5Y cells for a subset of the overlapping DEGs observed. Differentiated SH-SY5Y cells were transduced with 56 FU pc of either LAT or miR

lentivrus, from which RNA was extracted 14dpt in keeping with the original lentivirus RNA-Seq experiment.

This was tested not only for the DEGs that overlapped in miR lentivirus transduced cells and in1382 infected cells but also for a subset of DEGs that overlapped between miR lentivirus transduction, LAT intron lentivirus transduction and in1382 infection.

Most of the genes tested in a separate transduction had fold-changes in keeping with those found in the RNA-Seq samples, as shown by qRT-PCR (tables 4.16B and 4.16C). For example, qRT-PCR revealed that in the miR lentivirus and LAT lentivirus transduced cells, Y-box binding protein 2 (YBX2) was on average downregulated 1.19-fold in the original RNA-Seq experiment samples assessed by qRT-PCR and 1.15-fold in the separate transduction as assessed by qRT-PCR (table 4.16C). In fact, all the DEGs tested in this separate transduction by PCR were upregulated or downregulated in accordance to the change observed in the original RNA-Seq samples. This demonstrates the reproducibility and reliability of the data. However, for GPR149 although it is consistently and reproducibly downregulated, the magnitude of change is lower in this separate transduction. Nevertheless, this consistency and high fold-changes seen for GPR149 in the original RNA-Seq experiment suggest it might still be worth considering the physiological implications of this gene, despite some experiment-to-experiment variation.

The PCR data largely validated the RNA-Seq data. Regardless, it is also possible to use the PCR results to remove any gene outliers from any further analysis. Therefore, the qRT-PCR experiments not only confirmed the success of the original RNA-Seq experiment but can also aid in narrowing the focus of onwards experiments.

Table 4.16 qRT-PCR was performed to validate the DEGs observed during RNA-Seq

(A) A table comparing the fold-changes (to 2 decimal places (dp)) of the overlapping DEGs observed following RNA-Seq of 56 FU pc LAT intron lentivirus transduced or MOI 5 in1382 infected differentiated SH-SY5Y cells, to the fold-changes observed following qRT-PCR performed on the same samples as set-up in figures **4.1A** and **4.11A**. "-" means there was no PCR performed.

(B) A table comparing the fold-changes (to 2 dp) of a subset of the overlapping DEGs observed following RNA-Seq of 56 FU pc miR lentivirus transduced or MOI 5 in1382 infected differentiated SH-SY5Y cells, to the fold-changes observed following qRT-PCR performed on the same samples as set-up in figures **4.1A** and **4.11A** or from qRT-PCR on samples from a separate 56 FU pc miR lentivirus transduction of differentiated SH-SY5Y cells, with RNA harvested 14dpt for qRT-PCR.

(C) A table comparing the fold-changes (to 2 dp) of a subset of the overlapping DEGs observed following RNA-Seq of 56 FU pc LAT intron lentivirus transducted, miR lentivirus transduced or MOI 5 in1382 infected differentiated SH-SY5Y cells, to the fold-changes observed following qRT-PCR performed on the same samples as set-up in figures **4.1A** and **4.11A** or from qRT-PCR on samples from a separate 56 FU pc LAT intron lentivirus or miR lentivirus transduction of differentiated SH-SY5Y cells, with RNA harvested 14dpt for qRT-PCR. "-" means there was no PCR performed.

A

Gene	conditions	change	Fold change in:			
			RNA-seq		RNA-seq PCR	
			LAT	IN1382	LAT	IN1382
PCDH17	LAT lenti; In1382	upregulated	1.22	1.37	1.41	1.32
LPAR1	LAT lenti; In1382	upregulated	1.13	1.66	1.05	-
RECQL4	LAT lenti; In1382	downregulated	-1.2	-1.75	-1.04	-1.82
KCNAB3	LAT lenti; In1382	downregulated	-1.3	-1.31	1.50	-1.38
DACT1	LAT lenti; In1382	downregulated	-1.2	-1.33	-1.23	-1.21

B

		Fold change in:	RNA-seq		RNA-seq PCR	Separate PCR
Gene	conditions	change	MIR	IN1382	MIR	MIR
GPR149	MiR lenti; In1382	downregulated	-2.97	-2.38	-4.45	-1.41
MYCT1	MiR lenti; In1382	downregulated	-1.97	-1.67	-1.72	-1.96
KCNS3	MiR lenti; In1382	downregulated	-1.92	-2.89	-1.21	-1.06
BCAS1	MiR lenti; In1382	downregulated	-1.87	-11.43	-1.03	-1.20
RSPH4A	MiR lenti; In1382	downregulated	-1.68	-1.61	-1.19	-1.12

C

		Fold change in:	RNA-seq			RNA-seq PCR		Separate PCR	
Gene	conditions	change	LAT	MIR	IN1382	LAT	MIR	LAT	MIR
YBX2	LAT & MiR lenti In1382	downregulated	-1.24	-1.72	-1.44	-1.25	-1.12	-1.15	-1.15
THSD7A	LAT & MiR lenti In1382	upregulated	1.19	1.22	1.36	-1.02	1.28	-	-
GRIA2	LAT & MiR lenti In1382	upregulated	1.18	1.19	1.42	-	-	1.09	1.07

Table 4.16 qRT-PCR was performed to validate the DEGs observed during RNA-seq

4.2.9 Assessing the downstream impact of transcriptional changes on protein levels

The DEGs that come through the rigorous selection criteria employed in this chapter are worth following up to start to examine the downstream effect that these transcriptomic changes may have. One downstream effect is to assess whether there has been an impact on the protein production. A means to test the effect on protein production is by western blot. This has scope for future work but as a proof of principle, a pilot experiment western blot was performed on 3 of the genes significantly changed as a result of LAT intron lentivirus transduction as well as quiescent in1382 infection. These were PCDH17, lysophosphatidic acid receptor (LPAR1) and RECQL4. These 3 genes were chosen for the initial western blot experiment as they came up as significantly differentially expressed in both the LAT intron lentivirus transduction and the in1382 infection and therefore could be key in uncovering a novel function of major LAT. Additionally, they are have 3 distinct sizes so it should be easy to distinguish each protein on a western blot and antibodies were available for each of them.

Differentiated SH-SY5Y cells were transduced at 56 FU pc with LAT intron lentivirus or mock transduced in keeping with the RNA-Seq experiment described previously. Simultaneously, mock HEK293T cells were grown as a western blot positive control. Fourteen days post transduction the SH-SY5Y cells, and the HEK293T cells once confluent, were lysed and cellular proteins harvested. 200µg/µl of the differentiated SH-SY5Y cell proteins alongside 500, 200 and 50µg/µl of the HEK293T cell proteins were run on parallel western blots. The blots were probed with a combination of PCDH17, LPAR1 and RECQL4 antibodies initially and then washed and re-probed using beta actin antibody as a positive control. The expected size of each of the proteins are 20.18, 41.11 and 133.07 kilodaltons (kDa) for PCDH17, LPAR1 and RECQL4 respectively and approximately 42 kDa for Beta actin.

The results show clear and constant bands between 40 and 50kDa for the differentiated SH-SY5Y cells protein samples on the beta actin antibody probed blot (figure 4.21B). This is consistent with the size of beta actin. The HEK293T samples showed progressively less beta actin as the protein concentration decreased, as expected. Together, this suggests consistent loading and a successful western blot protocol.

The results when the blot was probed for the 3 DEGs are less clear. However, there are faint bands just above 20 in the SH-SY5Y samples and the most concentrated HEK293T samples, consistent with the size of PCDH17 (20.18kDa). There are also faint bands in the 500 and 200µg/µl HEK293T samples between 120 and 220 kDa, consistent with the

size of RECQL4 (133.07kDa) but this is not observable in the SH-SY5Y samples (figure 4.21A).

Quantification analysis was performed on Bio-Rad image lab software on these blots to test whether the changes in transcription had a similar impact on protein production. PCDH17 transcription was between 1.2 and 1.4-fold upregulated in the neurons expressing the major LAT intron, either through the LAT intron lentivirus delivery or quiescent in1382 infection, as shown by RNA-Seq and qRT-PCR. Quantitative analysis revealed that on average PCDH17 protein levels (normalised to beta actin levels) were 1.6-fold reduced in cells expressing LAT compared to untransduced, differentiated SH-SY5Y cells (figure 4.21C). Here, the protein levels decreased, which is opposite to how the pattern of transcription changed, as PCDH17 was upregulated following LAT intron lentivirus transduction. Although this seems unlikely, it is not impossible as protein levels and transcript levels do not always correlate (Kumar et al., 2016; Maier et al., 2009; Vogel & Marcotte, 2012).

To confirm these results, this blot would need repeating with the PCDH17 antibody in isolation and re-probed on the same blot for beta actin. It would also be worth testing more of the DEGs, especially those largest fold-changes in transcription, for example LPL (table 4.6). Additionally, it would be useful to perform western blots on the genes shown to have hits for miRNA binding such as CDP (table 4.15), to test whether this putative binding has an impact on protein levels.

Regardless, the western blots and quantitative analysis software successfully indicated protein levels in the LAT intron lentivirus transduced differentiated SH-SY5Y cells compared to mock SH-SY5Y cells. Therefore, western blot and quantitative analysis is a useful approach to examine the protein levels of the DEGs that were consistently significantly differentially expressed in response to the latency-associated ncRNAs.

The experiments to follow-up the RNA-Seq dataset is in its infancy, but it is evident already that there are many potentially interesting routes to follow on from this data. Determining the onwards functions that the transcriptomic changes elicited by the HSV-1 latency-associated ncRNAs has on human neurons could be invaluable to better understanding HSV-1 latency and to neuronal biology research.

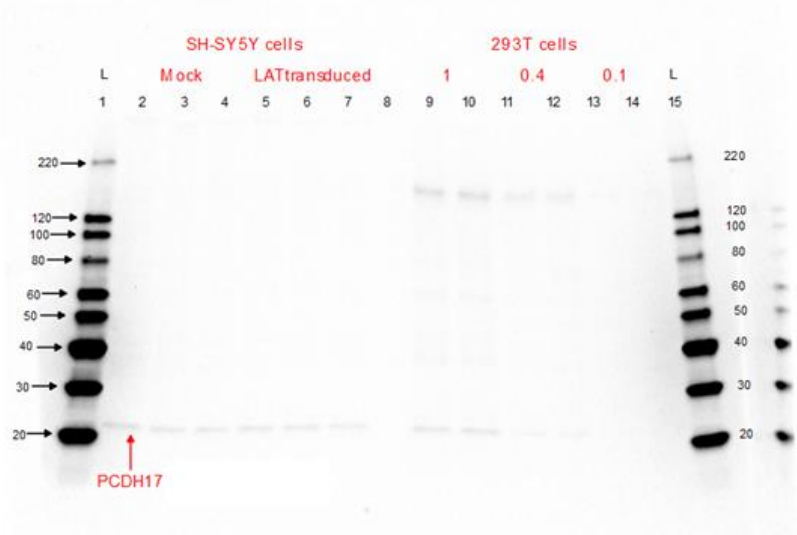
Figure 4.21 Western blot validation of transcriptional changes

(A) Western blot performed on proteins harvested from the LAT transduced and mock transduced differentiated SH-SY5Y cells or LAT or mock transfected HEK293T cells, using antibodies to PCDH17, LPAR1 and RECQL4. Lanes 2-4 show technical repeats of mock differentiated SH-SY5Y cells, lanes 5-7 show technical triplicates of LAT transduced differentiated SH-SY5Y cells and lanes 9-14 show technical duplicates of mock transfected HEK293T cells loaded at the ratios indicated above each lane (1, 0.4 and 0.1 for 500, 200 and 50µg/µl respectively). Bands running at the size of PCDH17 are indicated in red. Imaged on a BIO-RAD ChemiDocTM MP imaging system. L refers to protein molecular weight markers, and the band sizes are indicated.

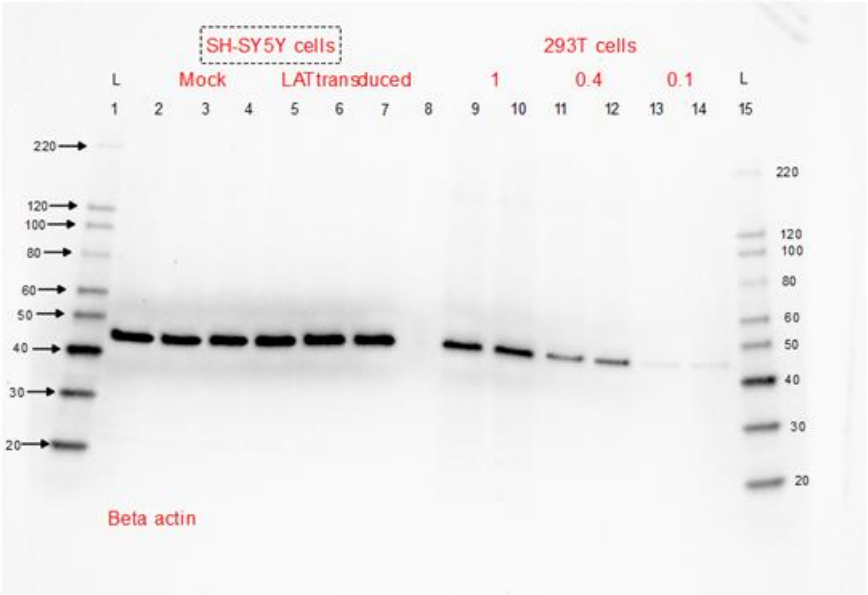
(B) The western blot described in **(A)** on proteins harvested from 56 FU pc LAT lentivirus transduced and mock transduced differentiated SH-SY5Y cells as well as LAT lentivirus or mock transfected HEK293T cells, was washed and re-probed using antibody to beta actin (methods chapter 6.2.8). Lanes 2-4 show technical repeats of mock differentiated SH-SY5Y cells, lanes 5-7 show technical triplicates of LAT transduced differentiated SH-SY5Y cells and lanes 9-14 show technical duplicates of mock transfected HEK293T cells loaded at the ratios indicated above each lane (1, 0.4 and 0.1 for 500, 200 and 50µg/µl respectively). Imaged on a BIO-RAD ChemiDocTM MP imaging system. L refers to protein molecular weight markers and the band sizes are indicated.

(C) A table showing the results from quantitative analysis on bio-rad image lab 6.0 software (2017, Bio-Rad laboratories incorporated) indicating relative protein quantities following capture of western blots shown in **(A)** and **(B)**.

A



B



C

Sample	**Mock** SH-SY5Y cells 200µg/µl			**LAT** transduced SH-SY5Y cells 200µg/µl		
Lane	2	3	4	5	6	7
	Relative PCDH17 quantity normalised to relative Beta actin quantity:					
	1	0.774	0.626	0.598	0.507	0.436
Average	0.8			0.5		
Relative difference	**1.6**					

4.2.10 Examining the transcriptional changes that occur 2hpi and 13dpi with repeat quiescent in1382 infection

There is substantial evidence that at least a subset of the gene changes observed following delivery of the latency-associated ncRNAs are directly as a result of their expression. There were reproducible gene changes despite varying the method of delivery of the latency-associated ncRNAs or the method of analysis, be it RNA-Seq or PCR. However, it is possible that some of the gene changes observed resulted from a persistent cellular response to encountering virus, rather than the HSV-1 latency-associated ncRNAs specifically.

To help address this question, differentiated SH-SY5Y cells were either mock infected or infected with in1382 at an MOI of 5 and total RNA harvested 2hpi and 13dpi (figure 4.22A). This allowed observation of the gene changes that occurred early during HSV-1 infection, to distinguish them from the changes that only occur following expression of the latency-associated ncRNAs.

LAT qRT-PCR was performed alongside a key lytic gene – ICP27 (introduction chapter 1.3.2) – on the samples harvested at 2hpi and 13dpi to ensure that the gene expression patterns were as expected for the establishment of a quiescent infection. From 2hpi to 13dpi, ICP27 expression decreased from 36.6 copies per 1000 beta actin copies to 0.2 copies per 1000 beta actin copies. Simultaneously LAT expression increased, from 4.7 copies per 1000 beta actin copies to 32.8 (figure 4.22B). This decrease in lytic gene expression and increase in LAT expression suggests that – as shown in previous experiments (chapter 2.2.5 and 4.2.3) – a quiescent infection is established with in1382 by 13dpi. However, the abundance of LAT is much lower in this repeat experiment than the previous in1382 infection for RNA-Seq (figure 4.11B), this must be considered when looking at onward gene changes and any differences observed between the two experiments. Nevertheless, RNA-Seq was performed as outlined in figure 4.2A (methods chapter 6.2.6ii) by Dr Martin Fritzche and Mr Ryan Mate. Following this, Dr Thomas Bleazard carried out bioinformatics analysis as outlined in figure 4.2B (methods chapter 6.2.7.1).

The bioinformatics analysis revealed that at 2hpi, 12,448 human genes were significantly differentially expressed following in1382 infection compared to uninfected SH-SY5Y cells with 6029 significantly upregulated genes and 6419 significantly downregulated genes (figure 4.22C).

At 13dpi, there are less transcriptional changes, but still a considerable number of significant gene changes with 2971 significantly differentially expressed human genes resulting from 1560 upregulated genes and 1411 downregulated genes (figure 4.22D).

Figure 4.22 RNA-Seq on 2h and 13d repeat in1382 infected differentiated SH-SY5Y cells

(A) Schematic representation of the set-up for the in1382 RNA-Seq experiment whereby differentiated SH-SY5Y cells were infected (at 39ºC) with HSV-1 strain in1382 at MOI 5 or mock infected. Total RNA was harvested at 2hpi and 13dpi with 5 replicates per condition and RNA being pooled from 2 wells per replicate.

(B) Differentiated SH-SY5Y cells were infected with HSV-1 strain in1382 or mock infected as described in **(A)** and qRT-PCR performed to LAT, ICP27 and Beta actin. The graph shows the average copies of LAT (blue) and ICP27 (orange) per 1000 Beta actin copies across 3 biological repeats and 3 technical repeats with the standard deviation represented by error bars.

(C) A flow chart outlining the number of differentially expressed genes following the RNA-Seq steps outlined in figure **4.2A** on the 2hpi harvested RNA set-up as described in **(A)**.

(D) A flow chart outlining the number of differentially expressed genes following the RNA-Seq steps outlined in figure **4.2A** on the 13dpi harvested RNA set-up as described in **(A)**.

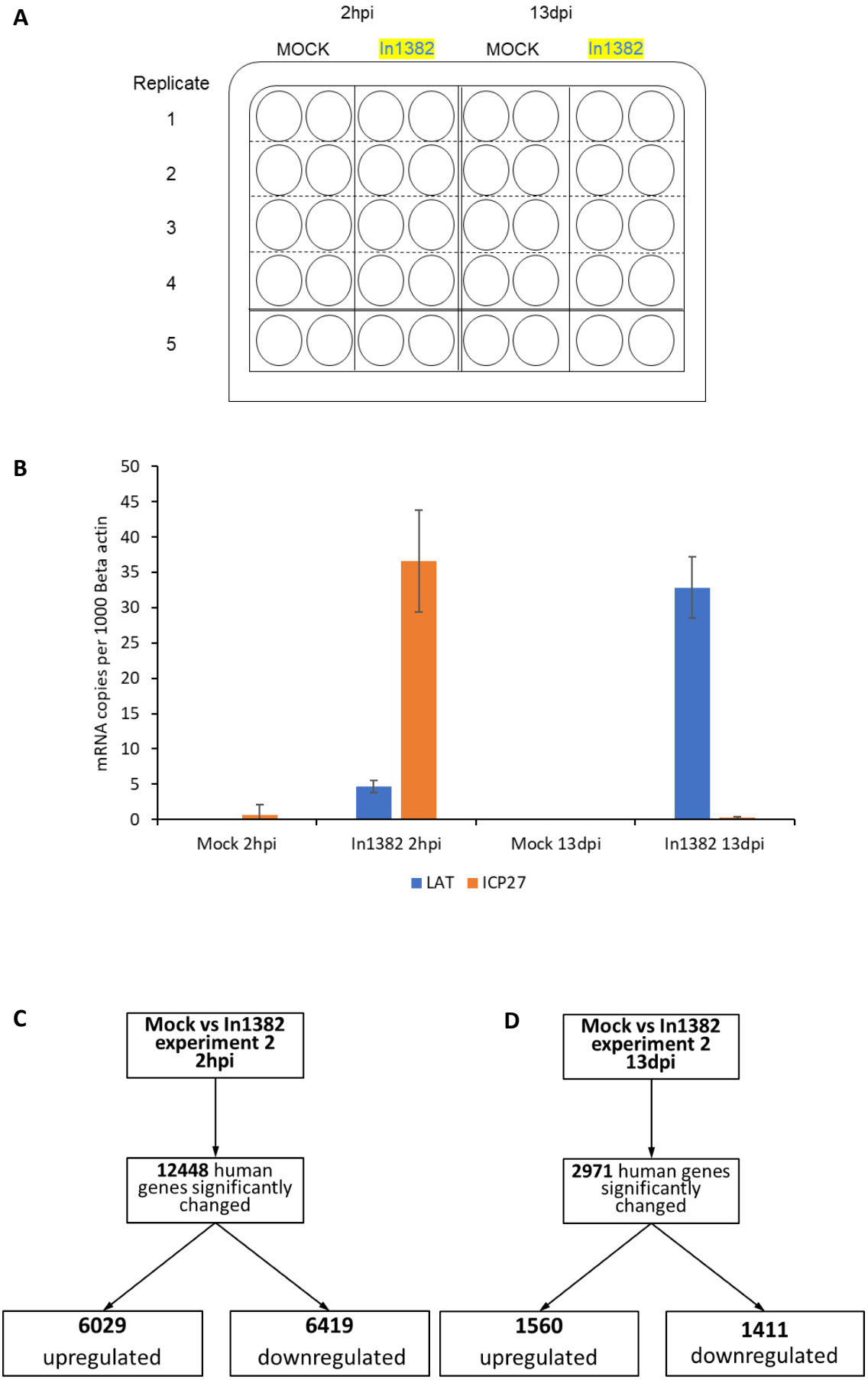


Fig 4.22 RNA-Seq on 2h and 13d in1382 infected differentiated SH-SY5Y cells

4.2.10.1 Comparing the transcriptional changes that occur at 13dpi during in1382 infection 1 and 2

Before examining the specific gene-changes observed at the 2 different time-points, it would be useful to know the reproducibility of the changes observed at 13dpi, compared to the previous 13d quiescent in1382 infection, despite each RNA-Seq experiment being performed in isolation but internally controlled. This is especially pertinent given the observation that the abundance of LAT differs in these 2 experiments (figures 4.11B and 4.22B).

Venny 2.1 was used to examine overlaps between the first and second 13d in1382 infections. There were 1164 (534 upregulated and 630 downregulated) and 1901 (1133 upregulated and 768 downregulated) gene changes specific the first and second 13d in1382 experiments respectively. These differences could be owing to the differences observed in LAT expression (figures 4.11B and 4.22B). Nevertheless, it was also observed that there were 374 genes (184 significantly upregulated genes and 190 significantly downregulated) that overlapped in both quiescent in1382 infections. This equates to 9.9% and 12% of all the significant upregulated and downregulated DEGs observed respectively (figure 4.23). For full gene-lists see appendix 1, sheet 14. These genes that come through in both 13d in1382 experiments, despite experimental variation could be crucial to the latency-associated ncRNA functions. Given that there was some reproducibility between quiescent infections, it was still useful address the question of whether any of the changes observed at 13dpi are also seen at 1dpi and therefore are a general response to infection or whether the changes observed at 13dpi differ and are a specific response to expression of the latency-associated ncRNAs.

Figure 4.23 Overlapping differentially expressed genes following 13d in1382 infection 1 and 2 in differentiated SH-SY5Y cells

(A) Venn diagram to show the number of significantly upregulated genes found by RNA-Seq that overlap following either 13dpi in1382 infection 1 or exclusively at 13dpi in in1382 infection 2 in differentiated SH-SY5Y cells when compared to uninfected differentiated SH-SY5Y cells. Diagram generated using Venny 2.1 (Oliveros, 2015).

(B) Venn diagram to show the number of significantly downregulated genes found by RNA-Seq that overlap following either 13dpi in1382 infection 1 or exclusively at 13dpi in in1382 infection 2 in differentiated SH-SY5Y cells when compared to uninfected differentiated SH-SY5Y cells. Diagram generated using Venny 2.1 (Oliveros, 2015).

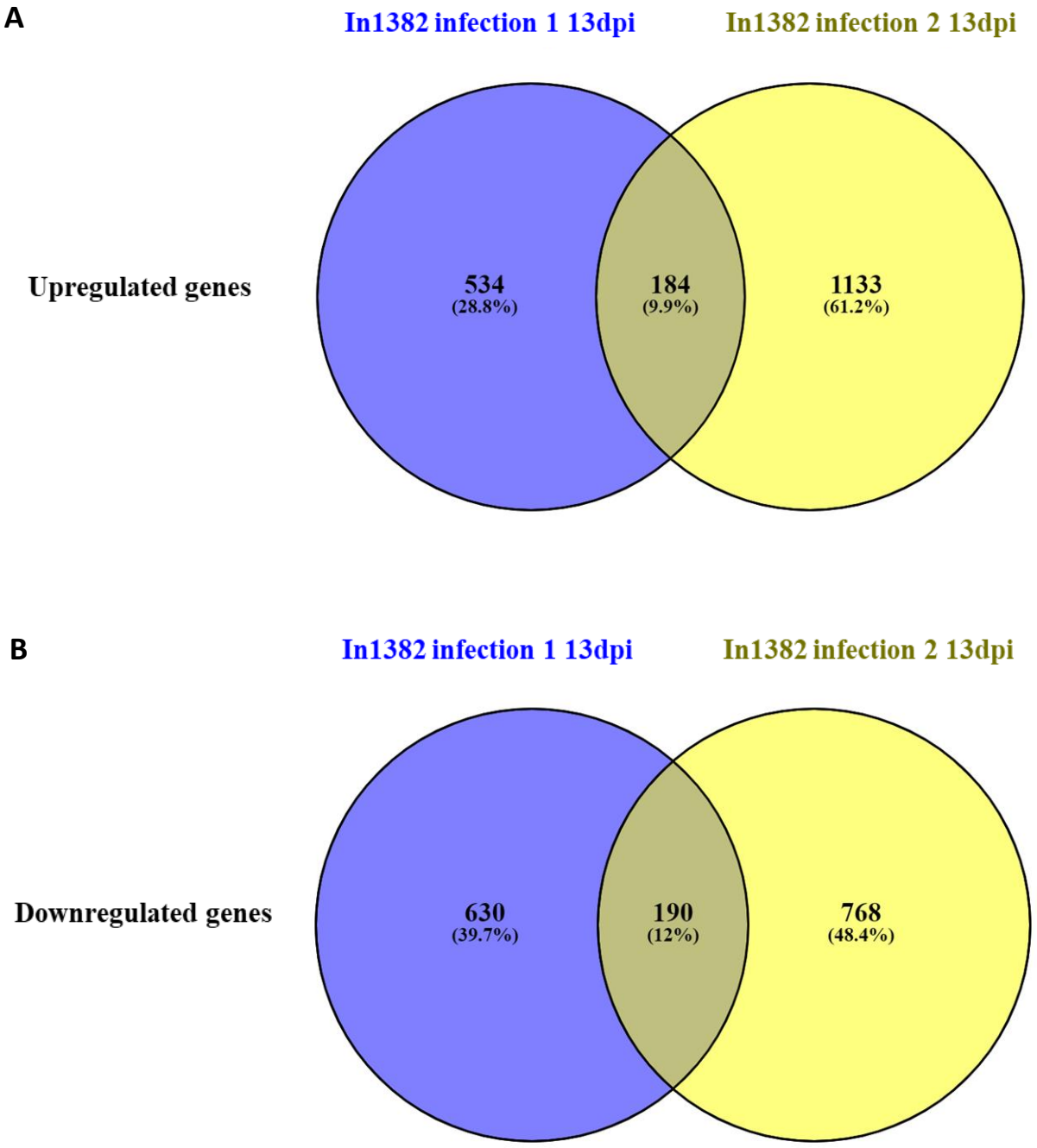


Figure 4.23 Overlapping differentially expressed genes following 13d in1382 infection 1 and 2 in differentiated SH-SY5Y cells

4.2.10.2 Comparing the transcriptional changes that occur at 13dpi with in1382 from those seen at 2hpi

To more closely examine the gene changes observed at each timepoint, with the aim of comparing them, further analysis was performed as described in figure 4.6, to look at the DEGs with the highest fold-change.

As well as there being a vast number of transcriptional changes at 2hpi, in response to the early stages of viral infection, there were also large fold-changes observed (tables 4.17 and 4.18), with the top upregulated gene being AC092573.1 which was 27.45-fold upregulated (table 4.17) and the top downregulated gene being C9orf135, which was 19.84-fold downregulated. Although the protein-coding functions of these genes are currently unknown, there is a large response to initial infection with replication defective HSV-1 mutant, in1382.

At 13dpi with this second in1382 RNA-Seq experiment, although smaller than at 2hpi, there were still sizable fold-changes. The top upregulated gene was AP001024.1 with a 12.98-fold change (table 4.19) and the most downregulated gene was MT-TE which was 7.61-fold downregulated (table 4.20). The size of these fold-changes is similar to the largest fold-changes seen in the previous 13d in1382 RNA-Seq experiment (tables 4.5 and 4.6) despite the lower LAT abundance in this repeat experiment. It is noteworthy, that many of the top downregulated genes following this second 13d in1382 infection were mitochondrial genes (table 4.20).

Table 4.17 The top 5 significantly upregulated genes following 2-hour in1382 infection 2 of differentiated SH-SY5Y cells

The table includes the top 5 significantly upregulated genes, according to fold change, found by RNA-Seq, following 2-hour in1382 infection of differentiated SH-SY5Y cells compared to uninfected differentiated SH-SY5Y cells, as set up in **figure 4.22A.** The table shows the fold change, p-adjusted value following Benjamini-Hochberg test and functions of the genes, as found in (NCBI, 2016; “UniProt,” n.d.) unless otherwise indicated.

Table 4.17 The top 5 significantly upregulated genes following 2-hour in1382 infection 2 of differentiated SH-SY5Y cells

Gene	Fold change (2dp)	P adjusted Value (3 sf)	Known protein function
AC092573.1	27.45	0.000124	No known protein coding function
ACTG2	25.94	0.000454	Actin gamma 2, smooth muscle: • Smooth muscle component • Found in enteric tissues
ID1	25.33	5.57E-85	Inhibitor of DNA binding 1, HLH protein: • Helix-loop-helix (HLH) protein that forms heterodimers with the basic HLH family of transcription factors to inhibit their DNA binding and transcriptional activation ability. • Putative role in cell growth, senescence and differentiation.
NPAS4	16.23	1.44E-25	Neuronal PAS domain protein 4: • NXF is a member of the basic HLH-PER-ARNT-SIM transcriptional regulators • Neuronal transcription factor
NPIPA8	14.86	0.000905	Nuclear pore complex interacting protein family member A8: • Ubiquitous expression in skin epithelial cells • No function information known

Table 4.18 The top 5 significantly downregulated genes following 2-hour in1382 infection 2 of SH-SY5Y cells

The table includes the top 5 significantly downregulated genes, according to fold change, found by RNA-Seq, following 2-hour in1382 infection of differentiated SH-SY5Y cells compared to uninfected differentiated SH-SY5Y cells, as set up in **figure 4.22A.** The table shows the fold change, p-adjusted value following Benjamini-Hochberg test and functions of the genes, as found in (NCBI, 2016) unless otherwise indicated.

Table 4.18 The top 5 significantly downregulated genes following 2-hour in1382 infection 2 of differentiated SH-SY5Y cells

Gene	Fold change (2dp)	P adjusted Value (3 sf)	Known protein function
C9orf135	-19.84	0.000602	Chromosome 9 open reading frame 135: No known protein coding function
AC112721.2	-17.78	0.00165	No known protein coding function
AC253536.2	-16.96	0.00193	No known protein coding function
FAM27D1	-16.37	0.00347	Family with sequence similarity 27 member D1: No known protein functions
AL078645.2	-16.37	0.0112	No known protein coding function

Table 4.19 The top 5 significantly upregulated genes following 13-day in1382 infection 2 of SH-SY5Y cells

The table includes the top 5 significantly upregulated genes, according to fold change, found by RNA-Seq, following 13-day in1382 infection of differentiated SH-SY5Y cells compared to uninfected differentiated SH-SY5Y cells, as set up in **figure 4.22A.** The table shows the fold change, p-adjusted value following Benjamini-Hochberg test and functions of the genes, as found in (NCBI, 2016) unless otherwise indicated.

Table 4.19 The top 5 significantly upregulated genes following 13-day in1382 infection 2 of differentiated SH-SY5Y cells

Gene	Fold change (2dp)	P adjusted Value (3 sf)	Known protein function
AP001024.1	12.98	0.0151	No known protein coding function
RPL13AP25	11.18	0.00498	Ribosomal protein L13a pseudogene 25 – no protein coding function
AC007969.1	5.29	0.00882	No known gene or protein coding function
RPL37P2	5.22	0.0416	ribosomal protein L37 pseudogene 2 – no protein coding function
PARM1	4.03	2.51E-06	Prostate androgen-regulated mucin-like protein 1: • Regulated progesterone metabolism

Table 4.20 The top 5 significantly downregulated genes following 13-day in1382 infection 2 of SH-SY5Y cells

The table includes the top 5 significantly upregulated genes, according to fold change, found by RNA-Seq, following 13-day in1382 infection of differentiated SH-SY5Y cells compared to uninfected differentiated SH-SY5Y cells, as set up in **figure 4.22A.** The table shows the fold change, p-adjusted value following Benjamini-Hochberg test and functions of the genes, as found in (NCBI, 2016) unless otherwise indicated.

Table 4.20 The top 5 significantly downregulated genes following 13-day in1382 infection 2 of differentiated SH-SY5Y cells

Gene	Fold change (2dp)	P adjusted Value (3 sf)	Known protein function
MT-TE	-7.61	4.45E-20	Mitochondrially encoded tRNA glutamic acid: • Contributes to the biosynthesis of mitochondrial proteins
MT-ND6	-7.53	3.95E-35	Mitochondrially encoded NADH dehydrogenase 6: • A component of enzyme complex I in mitochondria that contributes to oxidative phosphroylation
CNTNAP4	-6.75	2.20E-26	Contactin associated protein family member 4: • Member of the neurexin protein family • Cell adhesion molecule/receptor of the nervous system • Putative role in proper neurotransmission in dopaminergic and GABAergic systems
MT-TD	-6.20	3.59E-23	Mitochondrially encoded tRNA aspartic acid: • Contributes to the biosynthesis of mitochondrial proteins
MT-TF	-5.82	5.22E-35	Mitochondrially encoded tRNA phenylalanine: • Contributes to the biosynthesis of mitochondrial proteins

This experiment was designed to examine whether the genes differentially expressed following expression of the latency-associated ncRNAs at 13dpi with in1382 are specific or merely reflecting the initial gene changes that occur 2hpi, in the early stages after encountering virus. To address this question directly, the number of DEGs that overlap at 2hpi and 13dpi with in1382 were assessed.

Venny 2.1 (Oliveros, 2015) was employed to see how many DEGs overlapped in both timepoints. There are gene changes that occur at 2hpi that are also present at 13dpi with 243 upregulated and 453 downregulated overlapping gene changes. Nevertheless, this only equates to 3.3% and 6.1% of all the upregulated and downregulated genes respectively across the two timepoints. Therefore, most of the gene changes observed were specific to each time point. There were 1317 genes that were significantly upregulated at 13hpi and not 2hpi (figure 4.24A). Additionally, 958 genes were significantly downregulated at 13hpi that were not significantly downregulated at 2hpi (figure 4.24B). At 2hpi there were 5786 significantly upregulated and 5966 significantly downregulated genes that were specifically differentially expressed at 2hpi and not at 13dpi (figure 4.24A and 4.24B respectively). This suggests that there is an initial burst of transcriptional changes in the early stages of infection, but the gene changes observed at 13dpi are as a result of the latency-associated ncRNAs and not an ongoing cellular response after encountering virus.

To better appreciate the distinction between the initial transcriptional response to infection to those that occur as a result of the latency-associated ncRNAs, the transcriptional changes observed at either 2hpi or 13dpi were input into DAVID software (Huang et al., 2009b) for pathway analysis.

Given the large data-set (figure 4.24A and 4.24B) and high fold-changes observed (tables 4.17 to 4.20), the pathway analysis was performed on the DEGs with a fold-change of 2 or more as there is a maximum of 3000 genes that can be input as official gene symbols into DAVID (see methods section 6.2.7.2). This also means that the analysis is in keeping with that performed on the previous in1382 infection, also performed on genes filtered by a fold-change of 2 or more (figures 4.14 and 4.15). Filtering for a fold-change of 2-fold or more, narrowed down the gene changes to 1535 upregulated and 2083 downregulated genes at 2hpi with in1382 (figure 4.24C) and 40 upregulated and 108 downregulated genes at 13dpi with in1382 (figure 4.24D).

Figure 4.24 Overlapping differentially expressed genes from 2h and 13d in1382 infected differentiated SH-SY5Y cells

(A) Venn diagram to show the number of significantly upregulated genes found by RNA-Seq that overlap following either 2hpi of in1382 infection 2 in differentiated SH-SY5Y cells or 13dpi of in1382 infection 2 in differentiated SH-SY5Y cells when compared to uninfected differentiated SH-SY5Y cells. Diagram generated using Venny 2.1 (Oliveros, 2015).

(B) Venn diagram to show the number of significantly downregulated genes found by RNA-Seq that overlap following either 2hpi of in1382 infection 2 in differentiated SH-SY5Y cells or 13dpi of in1382 infection 2 in differentiated SH-SY5Y cells when compared to uninfected differentiated SH-SY5Y cells. Diagram generated using Venny 2.1 (Oliveros, 2015).

(C) Flow chart to represent the 2-fold or more significantly DEGs exclusively from 2hpi in1382 infection 2 out of those shown in **(A)** and **(B)**. Full gene lists in **appendix 1**.

(D) Flow chart to represent the 2-fold or more significantly DEGs exclusively from 13dpi in1382 infection 2 out of those shown in **(A)** and **(B)**. Full gene lists in **appendix 1**.

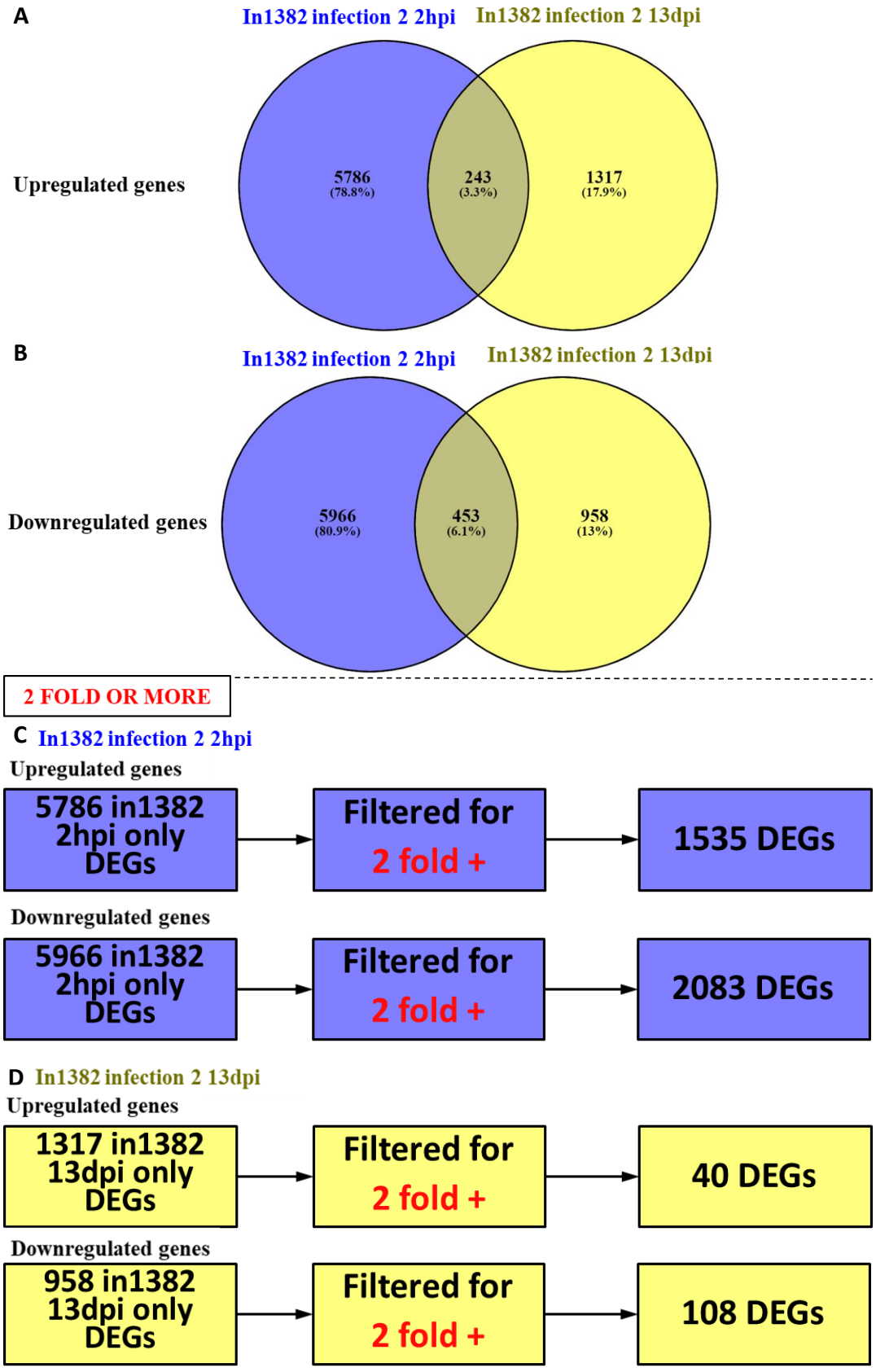


Fig 4.24 Overlapping differentially expressed genes from 2h and 13d in1382 infected differentiated SH-SY5Y cells

Figure 4.25 GO biological process and KEGG pathway analysis on 2-fold or more significantly upregulated genes following 2hpi with in1382 in differentiated SH-SY5Y cells

(A) A bar graph to show the gene ontology (GO) biological processes following analysis using DAVID software (Huang et al., 2009a, 2009b) on the 1535 2-fold or more significantly upregulated genes exclusively found at 2hpi with in1382 infected differentiated SH-SY5Y cells, as indicated in **figure 4.24C** and listed in **appendix 1**. The graph shows the top 10 significantly associated GO biological processes in terms of number of differentially expressed genes (DEGs), which are indicated in red. The blue bars show the -log10(p-values) following the fisher exact statistics method used by the DAVID software to calculate the p-value for each process.

(B) Bar graph to show the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways found by DAVID analysis (Huang et al., 2009a, 2009b) to be significantly associated with the 1535 2-fold or more significantly upregulated genes exclusively found at 2hpi with in1382 infected differentiated SH-SY5Y cells, as indicated in **figure 4.24C** and listed in **appendix 1**. The number of DEGs involved are indicated in red. The blue bars show the -log10 (p-values) following the fisher exact statistics used in the DAVID analysis software for each pathway. See **appendix 2** for full gene lists associated with each GO biological process/KEGG pathway.

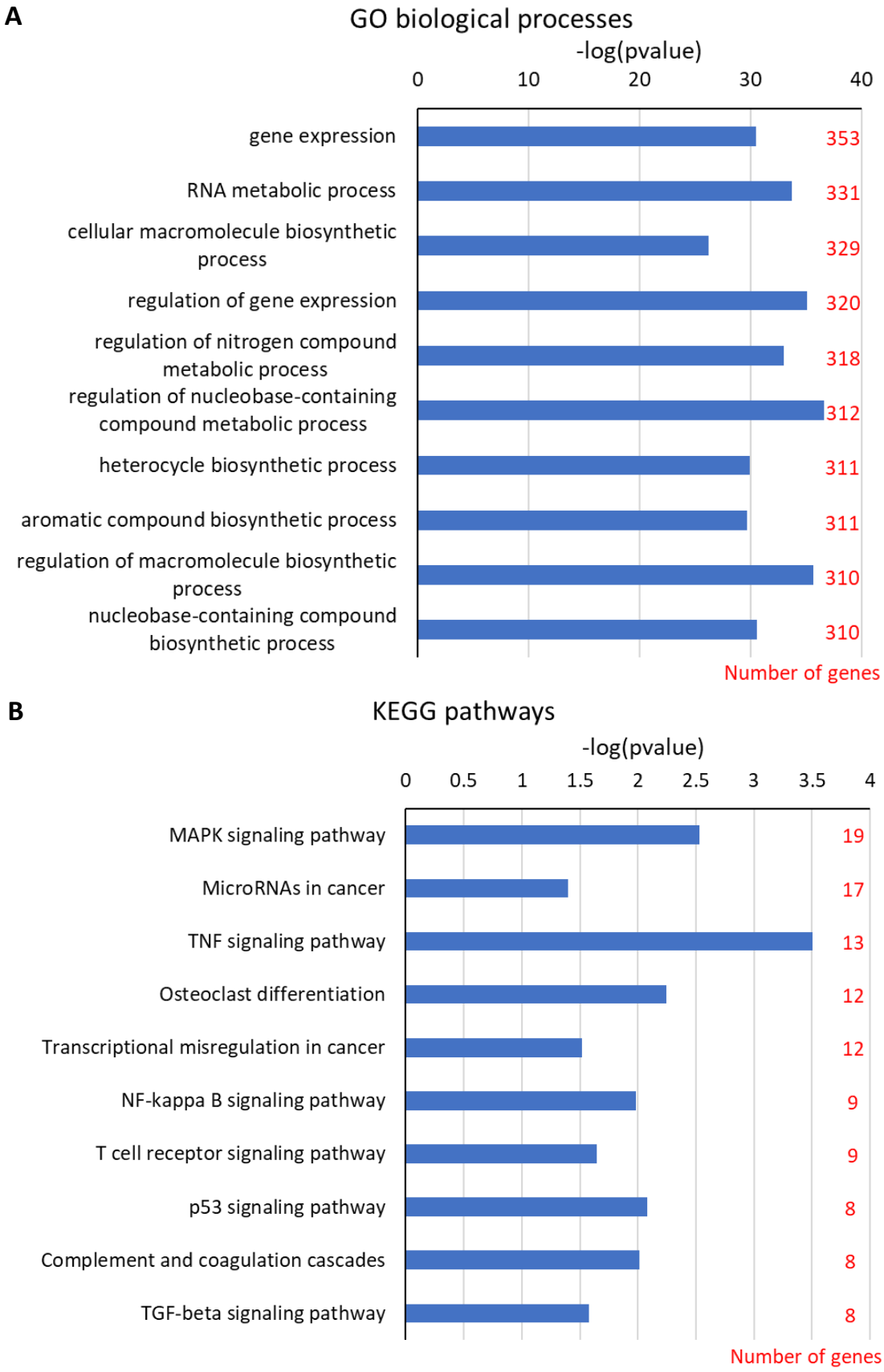


Figure 4.25 GO biological process and KEGG pathway analysis on 2-fold or more significantly upregulated genes following 2hpi with in1382 in differentiated SH-SY5Y cells

Figure 4.26 GO biological process and KEGG pathway analysis on 2-fold or more significantly downregulated genes following 2hpi with in1382 in differentiated SH-SY5Y cells

(A) A bar graph to show the gene ontology (GO) biological processes following analysis using DAVID software (Huang et al., 2009a, 2009b) on the 2083 2-fold or more significantly downregulated genes exclusively found at 2hpi with in1382 infected differentiated SH-SY5Y cells, as indicated in **figure 4.24C** and listed in **appendix 1**. The graph shows the top 10 significantly associated GO biological processes in terms of number of differentially expressed genes (DEGs), which are indicated in red. The blue bars show the -log10(p-values) following the fisher exact statistics method used by the DAVID software to calculate the p-value for each process.

(B) Bar graph to show the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways found by DAVID analysis (Huang et al., 2009a, 2009b) to be significantly associated with the 2083 2-fold or more significantly downregulated genes exclusively found at 2hpi with in1382 infected differentiated SH-SY5Y cells, as indicated in **figure 4.24C** and listed in **appendix 1**. The number of DEGs involved are indicated in red. The blue bars show the -log10(p-values) following the fisher exact statistics used in the DAVID analysis software for each pathway. See **appendix 2** for full gene lists associated with each GO biological process/KEGG pathway.

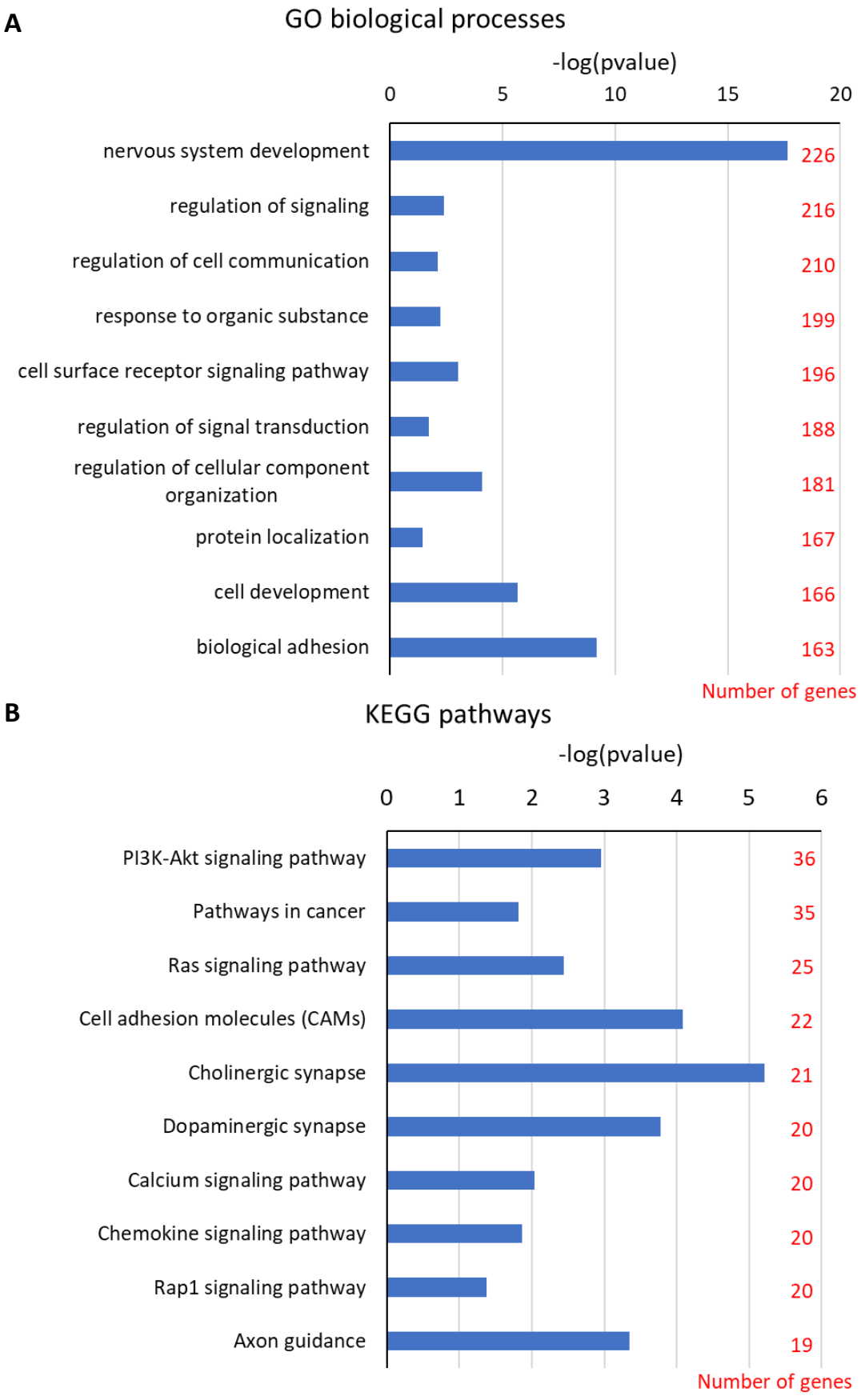


Figure 4.26 GO biological process and KEGG pathway analysis on 2-fold or more significantly downregulated genes following 2hpi with in1382 in differentiated SH-SY5Y cells

GO biological processes and KEGG pathway analyses were first performed on the 1535 2-fold or more significantly upregulated genes at 2hpi with in1382. The top GO biological process associated with the 2-fold or more significantly upregulated genes was gene expression, with 353 associated DEGs (figure 4.25A)(appendix 2, sheet 19). In addition to this, several of the top GO biological pathways associated with these upregulated DEGs were involved in biosynthetic processes. The KEGG pathway analysis indicated that the pathway with the most upregulated DEGs associated with it was the MAPK signalling pathway with 19 DEGs (figure 4.25B)(appendix 2, sheet 20).

Pathway analysis was similarly performed on the 2083 2-fold or more significantly downregulated genes found at 2hpi with in1382. The GO biological process analysis revealed that the GO process with the most DEGs significantly associated with it was nervous system development with 226 associated genes (figure 4.26A)(appendix 2, sheet 21). The KEGG pathway analysis revealed that the PI3K-Akt signalling pathway had the most significant gene associations with 36 associated genes (figure 4.26B) (appendix 2, sheet 22).

To compare the changes observed at 2hpi and 13hpi, DAVID pathway analysis was performed on the 40 genes that were 2-fold or more significantly upregulated following 13d in1382 infection (figure 4.24D)(Appendix 1, sheet 13). This analysis revealed GO biological processes but no KEGG pathways. The GO biological process analysis demonstrated that the process with the most upregulated DEGs associated with it was positive regulation of macromolecule metabolic process with 9 genes associated with it (figure 4.27).

From these GO biological processes observed from upregulated 13dpi genes, there are no overlaps with the top 10 GO biological processes seen following pathway analysis on the 2hpi 2-fold or more upregulated genes (figure 4.25A). This suggests that not only are different gene changes observed but also the downstream physiological effects of these transcriptional changes are different. This suggests that the transcriptional changes observed at 13dpi are specifically observed following expression of the latency-associated ncRNAs.

The same DAVID pathway analyses were performed on the 108 genes that were 2-fold or more downregulated exclusively at 13dpi, and not at 2hpi with in1382 HSV-1 strain. The analyses revealed that the GO biological process with the most significant 2-fold or more downregulated genes associated with it was the cell surface receptor signalling pathway with 18 significantly associated genes (figure 4.28A)(appendix 2, sheet 24). The KEGG pathway analysis revealed that the alcoholism pathway was linked to the most 2-

fold or more downregulated genes with 6 associated genes (figure 4.28B)(appendix 2, sheet 25). This KEGG pathway was also associated with the most 2-fold or more downregulated genes following the first in1382 13d infection (figure 4.15B).

Figure 4.27 GO biological process analysis on 2-fold or more significantly upregulated genes following 13dpi with in1382 infection 2 in differentiated SH-SY5Y cells

(A) A bar graph to show the gene ontology (GO) biological processes following analysis using DAVID software (Huang et al., 2009a, 2009b) on the 40 2-fold or more significantly upregulated genes exclusively found at 13dpi with in1382 infected differentiated SH-SY5Y cells, as indicated in **figure 4.24D** and listed in **appendix 1**. The graph shows the top 10 significantly associated GO biological processes in terms of number of differentially expressed genes (DEGs), which are indicated in red. The blue bars show the -log10(p-values) following the fisher exact statistics method used by the DAVID software to calculate the p-value for each process.

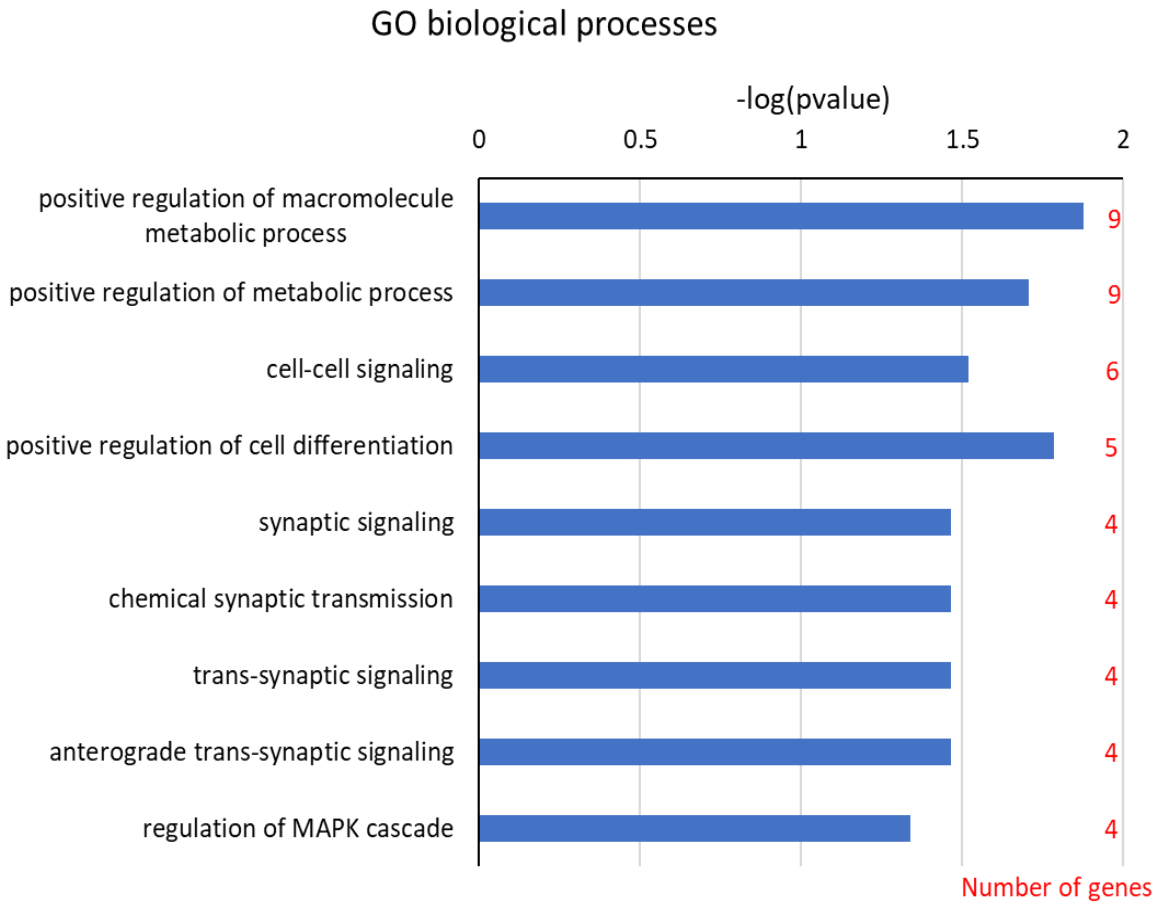


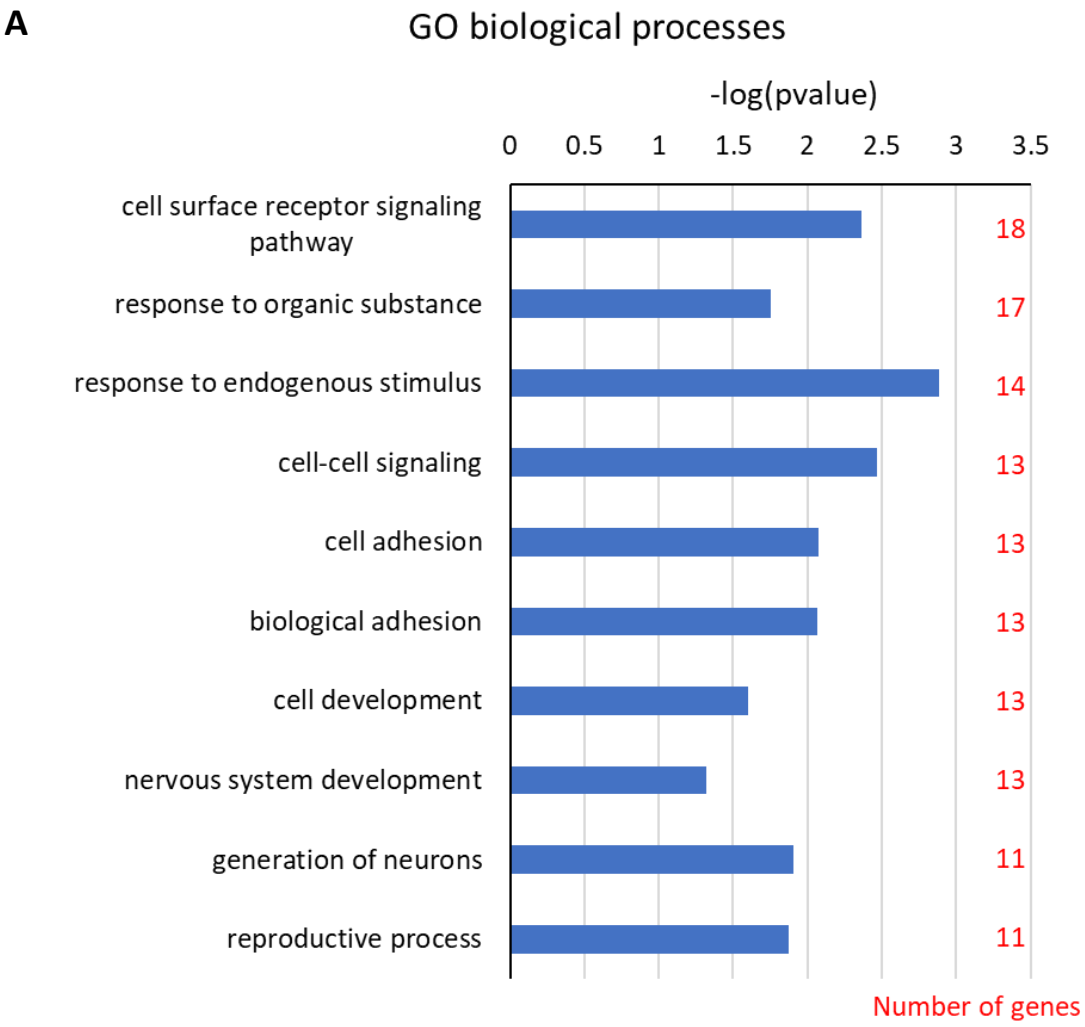
Figure 4.27 GO biological process analysis on 2-fold or more significantly upregulated genes following 13dpi with in1382 in differentiated SH-SY5Y cells

Figure 4.28 GO biological process and KEGG pathway analysis on 2-fold or more significantly downregulated genes following 13dpi with in1382 infection 2 in differentiated SH-SY5Y cells

(A) A bar graph to show the gene ontology (GO) biological processes following analysis using DAVID software (Huang et al., 2009a, 2009b) on the 108 2-fold or more significantly downregulated genes exclusively found at 13dpi with in1382 infected differentiated SH-SY5Y cells, as listed in **appendix 1**. The graph shows the top 10 significantly associated GO biological processes in terms of number of differentially expressed genes (DEGs), which are indicated in red. The blue bars show the -log10(p-values) following the fisher exact statistics method used by the DAVID software to calculate the p-value for each process.

(B) Bar graph to show the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways found by DAVID analysis (Huang et al., 2009a, 2009b) to be significantly associated with the 108 2-fold or more significantly downregulated genes exclusively found at 13dpi with in1382 infected differentiated SH-SY5Y cells, as listed in **appendix 1**. The number of DEGs involved are indicated in red. The blue bars show the -log10 p-values following the fisher exact statistics used in the DAVID analysis software for each pathway.

See **appendix 2** for full gene lists associated with each GO biological process/KEGG pathway.



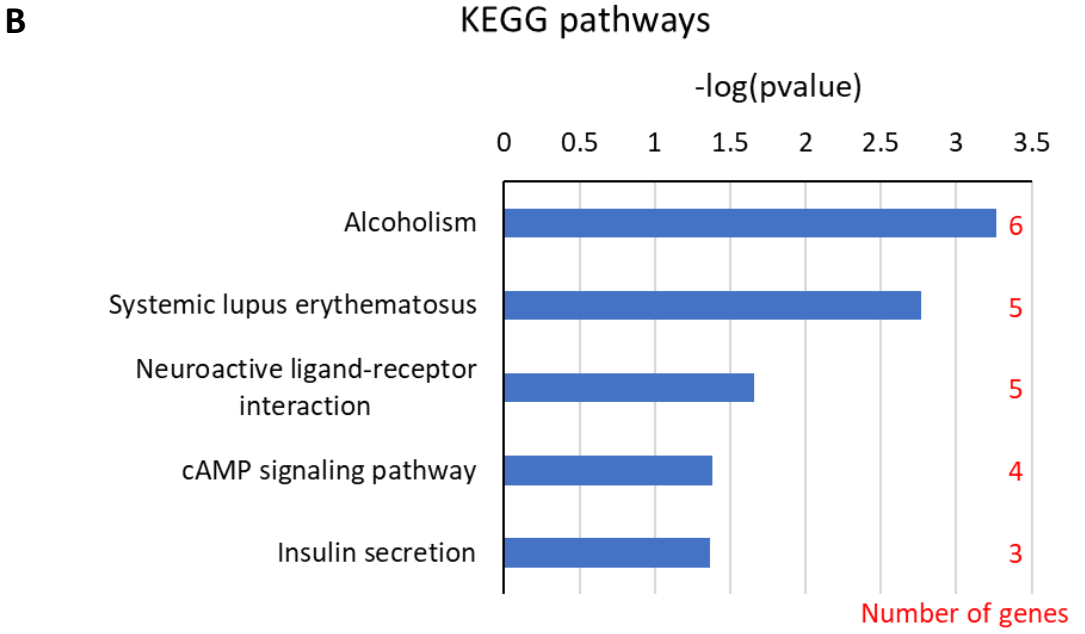


Figure 4.28 GO biological process and KEGG pathway analysis on 2-fold or more significantly downregulated genes following 13dpi with in1382 infection 2 in differentiated SH-SY5Y cells

4.2.11 Overlapping gene changes between latency-associated ncRNAs and both quiescent infections

In support of the changes observed at 13dpi being due to expression of the latency-associated ncRNAs, there were overlaps between with some of the GO pathways seen in previous pathway analyses performed on conditions that also delivered the latency-associated ncRNAs. For example, positive regulation of metabolic process was associated not only with these 13d in1382 infection 2, 2-fold upregulated genes but also the significantly upregulated genes following MiRNA lentivirus transduction (figures 4.9A and 4.27). Additionally, cell-cell signalling was associated with the 2-fold or more upregulated genes following this second 13d in1382 infection as well as those following the first 13d in1382 infection and the upregulated genes following the LAT intron lentivirus transduction (figures 4.14A and 4.27). This suggests that despite differences in the DEGs across the difference approaches to deliver the latency-associated ncRNAs and in repeat experiments, there are still overlaps in the downstream physiological effects from the transcriptional changes that occur in response to the latency-associated ncRNAs.

Again, there are overlaps with the KEGG pathways associated with the downregulated DEGs and those linked to the downregulated genes that occurred in the previous 13d in1382 infection and lentivirus transductions. Aside from the alcoholism pathway, the systemic lupus erythematosus pathway was also associated with the 2-fold or more downregulated genes from both 13d in1382 infections (figures 4.15B and 4.28B). Conversely, there are no overlaps with the KEGG pathways linked to the 2-fold or more downregulated genes that occurred 2hpi with in1382 (figure 4.26B). Additionally, there were GO biological processes that were revealed to be linked to the downregulated genes shown in this 13d in1382 experiment but also the previous 13d in1382 experiment and LAT intron lentivirus experiment. For example, the cell surface receptor signalling pathway is associated with the 2-fold or more downregulated genes found following in1382 13d infection 1 as well as 13d infection 2 (figures 4.15A and 4.28A). However, this process was also found to be associated with the 2-fold or more downregulated genes that changed following 2hpi in1382 infection 2 (figure 4.26A). This means that despite differences in transcriptional changes, some of the onwards pathways effected are the same following initial response to infection and during establishment of a quiescent infection.

Regardless, most of the gene changes observed at 2hpi and 13dpi with in1382 are different. Additionally, overlaps can be observed between the pathways impacted

following 13d in1382 infection 2 and the previous in1382 13d infection as well as lentivirus delivery of the latency-associated ncRNAs. Given this overlap in pathways analysis and the overlap of 374 genes between the two 13d in1382 experiments (figure 4.23) it is interesting to consider whether, despite the differences in the approaches and experiment to experiment variation, there are any transcriptional changes that are reproducible across both 13d in1382 experiments and transductions with either the LAT intron or miR lentiviruses. Especially given that the first in1382 13d experiment did have gene changes that overlapped with the lentiviruses (figures 4.16 and 4.17).

Venny 2.1 was utilised (Oliveros, 2015) to consider overlaps between both the 13d in1382 infections and either ncRNA expressing lentivirus transductions. The Venn diagrams show that although there are 0 overlapping upregulated genes, there are 3 genes that are reproducibly significantly downregulated across the in1382 13d experiments and the LAT intron lentivirus transduction (figure 4.29A and 4.29B). These repeatedly downregulated genes are dishevelled binding antagonist of beta catenin 1 (DACT1), KCNAB3 and RECQL4. The functions of these genes are outlined in table 4.7 along with their fold-changes across 2 of the experiments, see appendix 1, sheet 7 for the fold-changes during the second 13d in1382 infection.

For the miR lentivirus transduction and in1382 13d infections, there are 28 genes that are significantly upregulated in all 3 conditions (figure 4.29C)(appendix 1, sheet 15). There are also 14 downregulated genes that overlap in all 3 conditions (figure 4.29D)(appendix sheet 15).

These consistently reproducible gene changes could be key to uncovering the latency-associated ncRNA functions and would be a useful area to follow up.

Figure 4.29 Overlapping differentially expressed genes following 13d in1382 infection 1 and 2 and transductions with the latency-associated ncRNA expressing lentiviruses in differentiated SH-SY5Y cells

(A) Venn diagram to show the number of significantly upregulated genes found by RNA-Seq that overlap following either 13dpi in1382 infection 1 or exclusively at 13dpi (and not 2hpi) in in1382 infection 2 or exclusively following LAT intron lentivirus transduction (and not any other lentivirus transduction) in differentiated SH-SY5Y cells when compared to uninfected differentiated SH-SY5Y cells. Diagram generated using Venny 2.1 (Oliveros, 2015).

(B) Venn diagram to show the number of significantly downregulated genes found by RNA-Seq that overlap following either 13dpi in1382 infection 1 or exclusively at 13dpi in in1382 infection 2 or exclusively following LAT intron lentivirus transduction in differentiated SH-SY5Y cells when compared to uninfected differentiated SH-SY5Y cells. Diagram generated using Venny 2.1 (Oliveros, 2015).

(C) Venn diagram to show the number of significantly upregulated genes found by RNA-Seq that overlap following either 13dpi in1382 infection 1 or exclusively at 13dpi in in1382 infection 2 or exclusively following miR lentivirus transduction in differentiated SH-SY5Y cells when compared to uninfected differentiated SH-SY5Y cells. Diagram generated using Venny 2.1 (Oliveros, 2015).

(D) Venn diagram to show the number of significantly downregulated genes found by RNA-Seq that overlap following either 13dpi in1382 infection 1 or exclusively at 13dpi in in1382 infection 2 or exclusively following miR lentivirus transduction in differentiated SH-SY5Y cells when compared to uninfected differentiated SH-SY5Y cells. Diagram generated using Venny 2.1 (Oliveros, 2015).

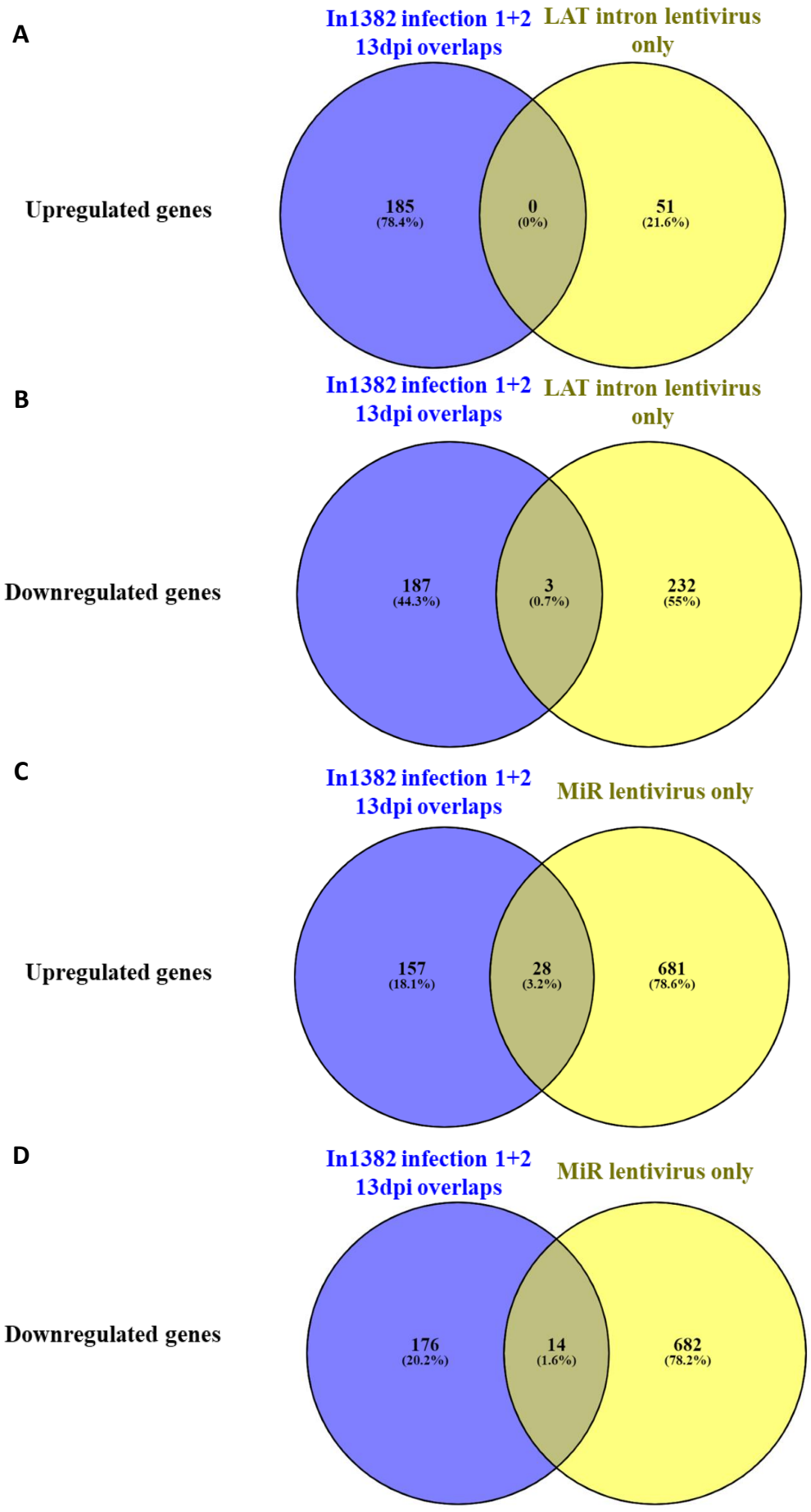


Figure 4.29 Overlapping DEGs following 13d in1382 infection 1 and 2 and transductions with the latency -associated ncRNA lentiviruses in differentiated SH-SY5Y cells

4.3 Discussion

In this chapter, RNA-Seq was used, alongside the HSV-1 latency models set-up in the previous chapters of this thesis (chapter 2 and 3), to show that the HSV-1 latency-associated ncRNAs affect the human neuronal transcriptome.

Initially, the lentiviruses constructed in chapter 3 (figure 3.3 and 3.4) were used to express the HSV-1 latency-associated ncRNAs in differentiated SH-SY5Y cells. RNA-Seq revealed more than 7000 genes were significantly differentially expressed following transduction with each of the lentiviruses (including the eGFP control lentivirus) compared to untransduced differentiated SH-SY5Y cells (figures 4.3 and 4.5A).

The eGFP lentivirus was used here as a control for the gene changes that occur in response to lentiviral transduction rather than the specific HSV-1 latency-associated ncRNAs they express. Therefore, any DEGs that were observed during eGFP lentivirus transduction were taken either to be a general response to lentivirus transduction or a reaction to the ectopic production of a protein (eGFP), which could explain the large number of gene changes that were exclusively observed following transduction with the eGFP lentivirus (figure 4.3). Regardless, these gene changes are not the result of expression of the HSV-1 latency-associated ncRNAs and, as such, were removed from further analysis. This meant that 2003 genes were significantly differentially expressed in response to the LAT intron lentivirus, miR lentivirus or NSLAT lentivirus (figure 4.5B).

The neuronal transcriptome response to the LAT intron lentivirus differs greatly from the response to the NSLAT lentivirus. There were 286 genes that were exclusively differentially expressed following transduction with the LAT intron lentivirus while transduction with the NSLAT lentivirus exclusively led to 129 DEGs (figure 4.3). This is compared to just 21 significant gene changes that overlapped between the two LAT lentiviruses. Therefore, most of the gene changes observed by the LAT lentiviruses are specific to each. Along with sequence analysis and the RNAscope data (chapter 3, figure 3.7B, 3.7C and 3.7D), this is further evidence that these two lentiviruses have different splicing phenotypes (as they are designed to, see chapter 3, figure 3.3), resulting in a different effect on the neuronal transcriptome. Therefore, the NSLAT lentivirus can act as a useful control whereby removing the 21 gene changes that occur in both transductions, leaves the 286 DEGs that occur specifically as a result of the presence of spliced major LAT intron (and not just the major LAT sequence). Accordingly, the number of gene changes were further narrowed down to 1740 genes that significantly changed as a result of the LAT intron lentivirus that produces the major LAT intron and/or the latency-associated miRNAs (figure 4.5C).

LAT intron lentivirus exclusively induced DEGs

LAT lentivirus individual DEG functions

Exploring the specific functions of the top 10 DEGs (in terms of largest fold-changes) following LAT intron lentivirus transduction only, revealed that three of these top DEGs encoded proteins that mediate the transport of molecules (NCBI, 2016). These genes were ABCA5, RAB27B and SLC8A1 (table 4.1). The proteins that these genes encode have roles within neurons that make them interesting targets for upregulation by major LAT.

ABCA5 encodes an ATP binding cassette subfamily A (ABCA) protein (table 4.1), which are known to be transporters that mediate the removal of excess cholesterol from neurons. Fu *et al.* showed that ABCA5 is present in human neurons, and by using SK-N-SH cells (the human neuroblastoma cell-line that SH-SY5Y cells originate from) showed that like other members of the ABCA family, ABCA5 mediates cholesterol efflux from neurons. Additionally, Fu *et al.* showed that ABCA5 can reduce amyloid β peptide production in neurons, this is especially interesting as amyloid β plaques are a hallmark of Alzheimer's disease pathology (Fu et al., 2015). Therefore, ABCA5 may potentially have a protective role in Alzheimer's neuropathology. It is therefore perhaps surprising that ABCA5 is upregulated by HSV-1 LAT, given that persistent HSV-1 infection has been linked to increased risk of developing Alzheimer's disease (Lövheim, Gilthorpe, Johansson et al., 2015; Wozniak et al., 2009; De Chiara et al., 2019)). This perhaps illustrates that there may be a more complex relationship between HSV-1 and Alzheimer's disease than currently understood. Accordingly, the relationship between HSV-1 and ABCA5 may be worth following up.

RAB27B has also been linked to several neurodegenerative disorders (Ginsberg et al., 2011; Underwood et al., 2020). Ginsberg *et al.* revealed that several RAB GTPases including RAB27B were upregulated in post-mortem brain microdissections from individuals with Alzheimer's disease and mild cognitive impairment. While Underwood *et al.* showed that RAB27B protein levels were raised in post-mortem brain lysates from individuals with Parkinson's disease or dementia with Lewy bodies when compared to samples from healthy individuals. Underwood *et al.* used a paracrine α-synuclein model that utilises a doxycycline-inducible neuroblastoma cell line that measures toxicity of α-synuclein to show that knockdown of RAB27B led to an increase in α-synuclein toxicity but a decrease in α-synuclein release. As α-synuclein is the main component of Lewy bodies, which are proteinaceous aggregates that cause synucleinopathies such as

Parkinson's disease, this data suggested that RAB27B has a protective role against Parkinson's disease by regulating the release and toxicity of α-synuclein (Underwood et al., 2020).

This is another example of a gene that encodes a protein with a putative protective role against neurodegenerative diseases being upregulated in response to major LAT expression, despite long-term HSV-1 infection having been linked to an increased risk of neurodegenerative diseases such as Alzheimer's disease. This further suggests that the link between HSV-1 infection and the development of neurodegenerative diseases could be more complex than currently understood and looking at the direct impact during latency might be an important route to consider.

SLC8A1, is a candidate gene for an association with bipolar disorder (Douglas et al., 2016; Le-Niculescu et al., 2009). This is interesting in the light of evidence suggesting that HSV-1 infection of neurons in the temporal lobe has a significant impact on cognitive function in schizophrenia and bipolar patients (Tucker & Bertke, 2019). Perhaps SLC8A1 upregulation is one way in which HSV-1 elicits its effect on schizophrenia.

Described above are just three interesting examples of gene functions that could be useful to follow up. However, as genes function as part of complex pathways and gene networks, the cumulative effect of the gene changes following LAT intron lentivirus transduction were also examined, by DAVID analysis.

Pathway analysis on LAT intron lentivirus changed genes

The pathway analyses revealed a significant association of the LAT intron lentivirus exclusively upregulated genes with several GO biological processes. However, there was a maximum of 16 genes associated with any GO biological process, and no significant KEGG pathways found to be associated with the LAT intron lentivirus exclusively upregulated genes. This is likely due to it being a relatively small sample size of 51 (figure 4.7).

There were a larger number of LAT intron lentivirus exclusively downregulated genes. Therefore, the DAVID analysis performed on these genes yielded significant KEGG pathways as well as more genes associated with each GO biological process (between 36 to 46 associated DEGs)(figure 4.8A). Four of these GO biological processes that were associated with the LAT intron lentivirus downregulated genes were also associated with the LAT intron lentivirus upregulated genes. This could be due to these being very common processes with a vast array of gene associations and therefore LAT may not specifically target them. Alternatively, the upregulated genes could have different

molecular activity relating to each of the GO biological processes than the downregulated genes do, therefore the major LAT could be tightly regulating these processes.

Altogether it seems that major LAT has a significant impact on the neuronal transcriptome despite low fold-changes in expression of the significant DEGs. The low fold-changes in the DEGs following LAT intron lentivirus transduction might be a result of either the low LAT expression observed by qRT-PCR (figure 4.1B) or because only between ~50-70% of cells express LAT at a given time (chapter 3 figure 3.7D).

MiR lentivirus exclusively induced DEGs

The latency-associated miRNAs also had a significant impact on the neuronal transcriptome, with a larger number of DEGs observed than following LAT intron transduction. There were 708 exclusively upregulated DEGs and 696 exclusively downregulated DEGs following miR lentivirus transduction. Additionally, following miR lentivirus transduction there were DEGs with larger fold-changes than with the LAT-lentivirus transduction (tables 4.3 and 4.4) with a top fold-change of 3.11 with CD34.

Although unavoidable circumstances during this PhD project meant that a stem-loop qRT-PCR was not carried out on the RNA-Seq RNA samples, LAT expression acted as a proxy for miR expression for the RNA-Seq experiments. There is evidence to support this approach as Kramer *et al.* demonstrated by use of recombinant viruses that the deletion of LAT significantly reduces expression of the HSV-1 miRNAs during HSV-1 latency. This suggests that LAT is required for optimum miRNA production. Additionally, in previous characterisation steps of the lentivirus system, the miRNA lentivirus was shown to express the HSV-1 miRNAs in the differentiated SH-SY5Y cells (chapter 3, figure 3.8).

MiR lentivirus specific gene functions

Included among the top 5 upregulated significant DEGs following miR lentivirus transduction were two that have been shown to have roles in neurons. LRRTM4 has been suggested to have a role in the development and maintenance of the central nervous system, which includes neurons (Laurén et al., 2003). Similarly, SEMA3A is vital for normal neuronal pattern development (NCBI, 2016). Given these functions, it could be that these genes are upregulated in a neuronal protective capacity. Maintaining neuronal viability would allow for the maintenance of HSV-1 latency within the neurons.

Additionally, increased expression of SEMA3A has been associated with schizophrenia (Eastwood et al., 2003) and aberrant release of SEMA3A protein has been linked to

progression of Alzheimer's disease (Good et al., 2004). These are interesting findings given that expression of HSV-1 miRNAs leads to the upregulation of this gene and as already mentioned HSV-1 has been linked to increased risk of Alzheimer's disease (Lövheim, Gilthorpe, Johansson et al., 2015; Lövheim, Gilthorpe, Adolfsson et al., 2015; De Chiara et al., 2019) and cognitive function in schizophrenia (Tucker & Bertke, 2019). Therefore, this interaction may begin to explain some of the molecular dynamics involved in these associations.

There is only so much that can be learnt from looking at the functions of specific gene changes alone, therefore, the cumulative functions of these gene changes that occurred following miR lentivirus transduction only, as examined by pathway analysis, will be discussed next.

Pathway analysis of miR lentivirus changed genes

Pathway analysis using DAVID software revealed many associations of both the significantly upregulated and downregulated DEGs following miR lentivirus transduction with GO biological processes or KEGG pathways.

For the GO biological processes, up to 150 significantly downregulated genes following miR lentivirus transduction were linked to localisation processes in the cell (macromolecule localisation, cellular localisation, establishment of localisation in the cell and establishment of protein localisation). Three of these localisation-related GO processes were also associated with LAT lentivirus downregulated DEGs (macromolecule localisation, cellular localisation and establishment of localisation)(figures 4.8A and 4.10A). This overlap could indicate that these localisation-related biological processes are particularly crucial to the function of the latency-associated ncRNAs. Macromolecule localisation is linked with both the significantly upregulated and downregulated genes following LAT intron lentivirus and miR lentivirus transduction. This could be due to it being a very broad GO biological process with many gene associations and not specifically targeted by the latency-associated ncRNAs. Alternatively, the latency-associated ncRNAs could be particularly tightly regulating macromolecule localisation by both the major LAT intron and the HSV-1 miRNAs targeting this process by both upregulation and downregulation of genes linked to it. Additionally, to be found to be associated with a GO process, a gene may be involved with the regulation of the process, which could be inhibitory or stimulatory. Therefore, the upregulation of certain genes linked to a GO process and downregulation of others may have the same ultimate impact on the process in question.

For the KEGG pathways that were associated with the downregulated DEGs, there were several related to viral infection, with HTLV-I infection, viral carcinogenesis and Epstein-Barr virus infection (figure 4.10B). The miRNAs targeting genes related to response to viral infection for downregulation could be of relevance to the balance between lytic and latent infection. When examining the specific significantly downregulated genes associated with these KEGG pathways following miRNA transduction, there are some that suggest the miRNAs might be involved in the regulation of reactivation from latency (appendix 1, sheet 8). For example, there were three histone-deacetylases (HDACs) downregulated (HDAC1, HDAC3 and HDAC8). As mentioned in chapter 1.4.2, inhibition of HDACs leads to reactivation from latency. This is because HDACs alter chromosome structure, blocking transcription factor access to DNA, repressing transcription (NCBI, 2016). In the context of latency, lytic gene repression encourages latency, which is why inhibition of this HDAC-mediated repression leads to reactivation. Therefore, downregulation of these 3 HDACs by the latency-associated miRNAs could be a self-regulation method of promoting reactivation. In support of this idea, another KEGG pathway that was significantly associated with the significantly downregulated DEGs was the MTOR signalling pathway (figure 4.10B). As described (introduction chapter 1.4.2), MTOR mediates latency and its components are targets for the PI3-K pathway, which when inhibited also leads to reactivation. Therefore, downregulation of components in this pathway could also mediate reactivation from latency. When further examining the specific downregulated genes that were shown to be linked to this and the viral KEGG pathways, the idea of associations with reactivation is strengthened. For example, phosphoinositide-3-kinase regulatory subunit 3 (PIK3R3) and phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit delta (PIK3CD), which are a subunit of PI3-K and a class I PI3K respectively, are significantly downregulated. PI3-K is important for the maintenance of latency, as inhibition of PI3-K causes reactivation (Camarena et al., 2010). Therefore, downregulation of these PI3-K component genes could also mediate reactivation. Although the role of the latency associated ncRNAs in reactivation has been somewhat disputed in the literature (chapter 1.4.1), with LAT having been suggested both to promote reactivation (Block et al., 1993; Perng et al., 1994) but also suppress reactivation (Nicoll et al., 2016), recent evidence has suggested a role for two latency-associated HSV-1 miRNAs, miR-H1 and miR-H6 in reactivation. Barrozo *et al*. demonstrated by use of a recombinant virus lacking these miRNAs led to an impaired reactivation phenotype following explant of latently-infected mouse DRG (Barrozo et al., 2020). This data supports the idea that the HSV-1 miRNAs may promote reactivation from latency as Barrozo *et al.* suggests.

The lentivirus model for delivering the latency-associated ncRNAs was able to demonstrate that the latency-associated ncRNAs do affect the neuronal transcriptome. From the data it seems that the latency-associated HSV-1 miRNAs have a larger impact on the neuronal transcriptome than the major LAT intron does, both in terms of fold-change and number of genes affected. The physiological functions affected by the transcriptome changes observed vary but seem to include a tight regulation of localisation of molecules in the cell, regulation of reactivation and a potential neuroprotective role for the latency-associate ncRNAs. This would agree with the literature that states LAT has a protective, anti-apoptotic function in neurons (Branco & Fraser, 2005; da Silva & Jones, 2013).

Advantages and disadvantages of the lentivirus model

Evidently the lentivirus model of expressing the HSV-1 latency-associated ncRNAs in differentiated SH-SY5Y cells enabled examination of the impact of major LAT or the HSV-1 miRNAs on the neuronal transcriptome.

Being able to attribute the gene changes to each of the lentiviruses and in turn the relevant latency-associated ncRNA is a major advantage of the lentivirus method of delivering the latency-associated ncRNAs and analysing their effects on the neuronal transcriptome. This enabled exploration of the specific physiological impact of the changes specifically elicited by each latency-associated ncRNA.

The success of this model was highlighted by the clear distinction between the gene changes attributable to the control eGFP-lentivirus and those attributable to the latency-associated ncRNAs, as indicated by the heat maps comparing these conditions (figure 4.4A and 4.4B). Another positive of this lentivirus RNA-Seq experiment is that there were 5 replicates per condition (figure 4.1A). Therefore, although some of the fold-changes of the DEGs were low, they recurred in all the replicates and were found to be significant with a maximum of 5% FDR. Together, this suggested that these changes, at least initially, were worth investigating.

The main disadvantages to this lentivirus delivery system for investigating the effect on the neuronal transcriptome is that it has relatively low-level LAT expression (figure 4.1B) and many of the gene changes observed had low fold-changes, especially following LAT intron lentivirus transduction (tables 4.1 and 4.2). These low-fold changes could be due to the low-level LAT expression observed and because potentially not all the transduced cells expressed LAT at any given time (chapter 3, figure 3.7). Additionally, in part because of this low LAT transcription, the lentivirus transductions do not fully reflect the

delivery of the latency-associated ncRNAs in true latency, whereby the expression of the latency-associated ncRNAs would be in the presence of each other and the rest of the viral genome.

Quiescent in1382 infection DEGs

To account for these caveats of the lentivirus RNA-Seq experiment, an alternative method was employed to express the latency-associated ncRNAs in differentiated SH-SY5Y cells. Differentiated SH-SY5Y cells were quiescently infected with replication defective HSV-1 strain, in1382. This method not only had higher LAT transcription (figure 4.11B) but it is also more reflective of true latency whereby the ncRNAs are produced in the presence of the latent HSV-1 virus genome and each other. RNA-Seq was performed on the in1382 infected differentiated SH-SY5Y cells and revealed more than 1500 significantly differentially expressed genes (DEGs) at a false discovery rate of 5% compared to mock differentiated SH-SY5Y cells (figure 4.11C).

There were larger fold-changes observed following quiescent infection compared to the lentivirus transductions with the top upregulated gene being 7-fold upregulated (table 4.5) and the top downregulated gene 12.4-fold downregulated (table 4.6). The larger fold-changes observed during in1382 infection (compared to LAT lentivirus transduction) could be due to there being a higher abundance of LAT, as shown by qRT-PCR (figures 4.1B and 4.11B).

In1382 DEGs pathway analyses

To get more of an overview of the physiological impact of the transcriptional changes observed following quiescent infection, DAVID analysis software was used to examine the biological pathways affected by these gene changes. Initially all significant DEGs observed following quiescent in1382 HSV-1 infection were input into DAVID. However, given the higher fold-changes observed, to focus on the DEGs that might have the biggest onwards impact, the DEGs were filtered for a fold change of 2-fold and entered into DAVID. This included 403 DEGs with a 2-fold change or more, with 92 significantly upregulated and 311 significantly downregulated compared to mock differentiated SH-SY5Y cells (appendix 1, sheet 10).

Looking at the GO biological processes and KEGG pathway analyses for the 92 2-fold or more upregulated DEGs, the main processes that appeared to be enriched seem to relate to cell communication and signalling (figure 4.14), which was also similar to the processes enriched before a 2-fold threshold was applied to the upregulated genes (figure 4.12). Therefore, these processes seem to be considerably affected by quiescent

in1382 infection. It could be suggested that the upregulation of cell signalling promotes the regular functioning of neuronal cells and in this way the latency-associated ncRNAs are neuroprotective, as had been observed from the lentivirus experiment. However, upon examining the gene ontology biological processes that were associated with the 311 downregulated DEGs using DAVID analysis software, it also revealed processes involved in signalling and cell communication, as well as cell response to stimuli (figure 4.15A). As observed with the lentiviruses, an overlap in the association of biological processes with both upregulated and downregulated genes, in this case related to cell signalling and communication, could be because a careful balance is needed in the neurons and the latency-associated ncRNAs help achieve this. However, it could be that these are highly common pathways with many known gene associations. Therefore, the role of the latency-associated ncRNAs may not directly be linked to cell-signalling or communication. This overlap of processes could also be a symptom of having input too many genes into the DAVID analysis software and narrowing down the DEG list further could yield results more specific to the effect of the HSV-1 latency-associated ncRNAs.

Nonetheless, DAVID analysis software did also identify processes that were specifically associated with the downregulated genes and not upregulated genes (figures 4.13 and 4.15) and this included the cell cycle. Cell cycle was observed to be associated with the downregulated genes both as a GO biological process and as a KEGG pathway both before and after filtering for 2-fold or higher significantly changed genes. There were 150 (50 when 2-fold filtered) DEGs and 24 (7 when 2-fold filtered) DEGs involved with the cell cycle following GO biological process analysis and KEGG pathway analysis respectively.

The focus of the latency-associated ncRNAs to quell the cell-cycle is interesting in the context of neurons, which once differentiated are generally believed to be senescent cells, unable to proliferate. Nevertheless, expression of certain cell-cycle components such as cyclin D1 can be detected in neurons and although these have been suggested to have alternative functions in neurons such as neuronal maturation and migration, it means that it is possible for neurons to re-enter the cell cycle (Frade & Ovejero-Benito, 2015). Cell-cycle re-entry occurs following upregulation of key cell-cycle components such as cyclin D and Cdk4/6 (Frade & Ovejero-Benito, 2015; Liu & Greene, 2001). This re-entry into the cell-cycle in neurons has been shown to lead to apoptosis of the cell (Liu & Greene, 2001). Accordingly, the latency-associated ncRNAs could be restricting cell-cycle components, again as a neuroprotective measure.

Additionally, the latency-associated ncRNAs inhibition of the cell-cycle may be a means of maintaining latency, as viral replication is restricted during latency and therefore limiting the replication components of the cell may aid in preserving this state. In support of this idea, it has been shown that there is an increase in cell cycle components such as cyclin D2 in response to HSV-1 reactivation (reviewed in Traylen et al. 2011). Cyclin D2 contributes to the switching from G1 phase to S phase of the cell-cycle as well as in the sequestering of cell-cycle inhibitor p27 (Traylen et al. 2011). It follows that remaining in G1 phase and limiting the cell-cycle and division may contribute to the maintenance of HSV-1 latency.

Another interesting observation from the KEGG pathway analysis of the downregulated DEGs following quiescent in1382 infection was that there was an association with the herpes simplex infection KEGG pathway. Looking more closely at which of the DEGs are associated with this pathway, it is immune response components. This includes interferon-inducible genes such as interferon induced protein with tetratricopeptide repeats 1 (IFIT1), interferon induced with helicase C domain 1 (IFIH1), 2'-5'-oligoadenylate synthetase (OAS) 1 and 3, genes encoding cytokines such as C-C motif chemokine ligand 2 (CCL2) and major histocompatibility complex (MHC) class 1 protein-encoding genes – HLA-B and HLA-C.

The proteins each gene encodes contributes to an antiviral response in different ways. For example, IFIH1 which encodes the viral RNA sensor MDA5, directly detects viral RNA and acts on an infected cell by inducing production of the type I interferons (Bansode et al., 2019). The production of type I interferons induces a variety of onwards signalling pathways that put the cell in an antiviral state (Samuel, 2001). One onward effect of production of the type I interferons is the induction of interferon stimulated genes such as IFIT1, another downregulated DEG observed. IFIT1 can bind a number of cellular proteins, such as eukaryotic translation initiation factor 3 (eIF3) and viral RNA, which in turn can stall HSV-1 replication (Bansode et al., 2019; Fensterl & Sen, 2015). Conversely, the MHC class I molecules elicit their effect indirectly by presenting viral antigen to recruit CD8+ T lymphocytes, which can then mediate killing of infected cells (Yewdell & Bennink, 2002).

Regardless of the approach, the presence of any of these immune response genes would help restrict HSV-1 lytic replication. Therefore, it is surprising to see downregulation of these genes (all found to be associated with the herpes simplex infection KEGG pathway) following expression of the latency-associated ncRNAs. However, this could exhibit an immune evasion technique during latency. For example,

downregulation of the MHC class I molecules could prevent presentation of HSV-1 antigen to cytotoxic T cells and therefore avoid T-cell mediated neuronal cell death, so that latent infection can continue in these cells. However, also included among these downregulated DEGs associated with HSV infection was PML nuclear body component, SP100. The cellular protein this gene encodes has been shown to co-localise with the incoming HSV-1 genome upon nuclear entry in neurons, to block viral replication and in turn supports HSV-1 latency (Alandijany et al., 2018; Cohen et al., 2018). Therefore, while downregulation of these genes may be an immune evasion technique, again it is somewhat surprising that they are downregulated by the latency-associated ncRNAs. It is possible that the lytic transcription shown to also occur during in1382 infection might be having an impact here as downregulation of these genes seems to favour lytic infection over latency. Although its effect cannot completely be ruled out, lytic transcription was extremely low, with less than 5000 times the transcription of major LAT as shown by qRT-PCR (figure 4.11B). Alternatively, the latency-associated ncRNA could be self-regulating latency and preparing for reactivation by downregulating this pathway. Self-regulation could explain why LAT expression seems to reproducibly only be expressed in a subset of latently infected cells, both in the systems used in this thesis sand in the literature (chapter 2.2.5 and 3.2.4) (Chen et al., 2002; Edwards & Bloom, 2019; Mehta et al., 1995). The idea of the latency-associated ncRNAs having a role in driving reactivation was also suggested by the earlier miRNA data as well as in the literature (Barrozo et al., 2020; Block et al., 1993; Perng et al., 1994). Nevertheless, the effect of LAT in reactivation is complex and in batch cell experiments, such as this, can be masked by the effect on the establishment of latency as already discussed (chapter 1.4.1)(Nicoll et al., 2016; Sawtell & Thompson, 1992a; Thompson & Sawtell, 1997).

Advantages and disadvantages of in1382 quiescent HSV-1 infection model

There are benefits of using in1382 to express the latency-associated ncRNAs, such as high LAT expression and that any effect that occur as a result of a combination of LAT and the miRNAs can be tested. However, there are also caveats to this system. For example, the infections are performed at 39^{o}C, to rule-out the production of functional ICP4, which could affect the neurons and skew the transcriptomic results found. However, this raised temperature might also have an impact on the neurons, which could impact their transcriptome. Nevertheless, the mock uninfected differentiated SH-SY5Y cells are also kept at 39^{o}C, therefore, any differential gene expression observed when comparing in1382 infected and the mock uninfected neuronal SH-SY5Y cells will be due to quiescent in1382 infection and not any effect of temperature. Another potential caveat of using in1382 infection to test the effect of the latency-associated ncRNAs on the

neuronal transcriptome is that the changes observed are not attributable to each specific latency-associated ncRNA. Additionally, lytic transcription was also observed in this system (figure 4.11D) meaning that this lytic transcription could account for some of the host transcriptional changes observed. Nevertheless, it is possible that the observation of lytic transcripts represents read-through transcription rather than lytic gene expression. Additionally, there is precedent for low-level lytic gene expression during HSV-1 latency as this has been observed in other HSV-1 latency models. Harkness *et al.* used a replication defective HSV-1 mutant to infect murine neurons and revealed by RNA-Seq that genes aside from LAT remain active during quiescent infection (Harkness et al., 2014). Therefore, this low-level lytic transcription with higher LAT expression is still reflective of HSV-1 latency. In support of this, as already described, in this in1382 model qRT-PCR revealed that major LAT is transcribed over 5000 times more than the lytic gene ICP47, that had shown transcription by RNA-Seq (figure 4.11B). Therefore, despite lytic transcription it is more likely that the neuronal gene changes observed are attributable to the abundant latency-associated ncRNAs than the lytic genes.

Accordingly, there were overlaps in the processes that were linked to DEGs observed following quiescent in1382 infection and those observed following the lentivirus transductions, which occur specifically as a result of either major LAT or the latency-associated miRNAs. For the upregulated DEGs, DAVID analysis revealed overlaps in GO processes of regulation of signalling and regulation of cell communication following either miR lentivirus transduction or in1382 infection and cell-cell signalling following either LAT intron lentivirus transduction or in1382 infection (figures 4.7, 4.9 and 4.14).

For the significantly downregulated genes, DAVID analysis revealed that both the LAT intron downregulated genes and the in1382 downregulated genes were associated with the GO biological process of regulation of molecular function and both the miR lentivirus downregulated genes and in1382 downregulated genes were associated with the viral carcinogenesis KEGG pathway (figures 4.10 and 4.15). Having this overlap in the processes and pathways effected by lentivirus expression of the latency-associated ncRNAs and in1382 quiescent infection, gives supporting evidence that these processes and pathways are targeted by the latency-associated ncRNAs.

Overlapping DEGs following expression of the latency-associated ncRNAs through either lentivirus transduction or quiescent in1382 infection

The similarities between transcriptomic effects elicited by the lentivirus transductions and the quiescent in1382 infection go further than these overlaps observed in the separate pathway analyses. Examining the specific DEGs that overlap following the lentivirus

transductions and in1382 infection can help reveal which of the changes observed during quiescent infection are attributable to which of latency-associated ncRNAs. Additionally, combining the results in this manner eliminates the caveats described above for either method in isolation. Any potential effect from any lytic expression during quiescent in1382 infection is removed, the gene changes can be attributed to either the miRNAs or major LAT and finally the changes occur during a more physiologically relevant model of HSV-1 latency. Accordingly, the DEGs that occur in both systems were examined (figures 4.16 and 4.17)(appendix 1, sheet 11).

Examining the overlapping gene changes from the ncRNA lentiviruses and in1382 infection is a stringent method for accepting genes as significantly changed due to the HSV-1 latency-associated ncRNAs. This means that the gene must be significantly differentially expressed, with a maximum 5% FDR in multiple repeats across multiple conditions in both experiments. It is possible that being so stringent in accepting an overlapping DEG and eliminating so many other gene changes in the process means that some genes that are truly affected by the HSV-1 latency-associated ncRNAs could be missed, which is why the gene changes from each experiment in isolation have also been considered prior to this overlapping assessment. However, being so stringent does somewhat counteract the fact that many of these overlapping DEGs have low fold-changes, close to 1 (tables 4.7-4.9). Although these low fold changes could be considered a caveat, it also might be expected given that not all the in1382 infected or LAT lentivirus transduced cells express LAT (chapter 2, figure 2.6D and chapter 3, figure 3.8C and D). This means that any effect that the latency-associated ncRNAs are having on the neuronal transcriptome is only in a proportion (between approximately 30 and 80%) of cells and so the results from a bulk cell RNA-Seq experiment like this could be limited. A focus for future studies could be to increase the proportion of cells that express LAT to see whether this affects the DEGs observed, especially in terms of fold-change. Alternatively, utilising single-cell analysis techniques could be a useful next step to examine the effect of the latency-associated ncRNAs, in just the cells expressing the ncRNAs.

There were 169 gene changes overlapping between the miR lentivirus transduction and in1382 infection compared to just 5 between the LAT intron lentivirus transduction and in1382 infection (figure 4.16), plus 4 DEGs overlapping with both lentivirus transductions and quiescent infection (fig 4.17A, 4.17B and table 4.9). There is clearly a larger effect elicited by the latency-associated miRNAs than major LAT in terms of the number of genes affected. In addition to this, the fold-change of the effects of the miRNAs tended to be larger than that elicited by LAT. The DEG with the highest fold-change resulting

from expression of the LAT intron, either from LAT intron lentivirus transduction or in1382 infection, as observed by RNA-Seq was RECQL4 with an average fold change of 1.5. For the latency-associated miRNAs the DEG with the largest fold change exhibited as shown by RNA-Seq was TFPI2 with an average of 2.9-fold change across the miR lentivirus and in1382 experiments. This is perhaps unsurprising given that there are 5 miRNAs in the miR lentivirus construct that could have additive roles plus miRNAs have a known gene expression regulation role. Although miRNA regulation is usually post-transcriptional there is data to suggest that miRNAs can regulate transcription as well (Catalanotto et al., 2016). LAT has also been shown to have a crucial role in the establishment and maintenance of latency (Nicoll et al., 2012b; Sawtell & Thompson, 1992a; Thompson & Sawtell, 1997) so it could be that the role of major LAT lies in viral effect with a more modest effect on the host cell transcriptomics. Following transduction with the NSLAT lentivirus, whereby LAT is expressed but the major LAT intron cannot be spliced, only one of the DEGs was found to overlap with those resulting following LAT intron lentivirus transduction or in1382 infection, this suggests that it is important to have major LAT splicing to elicit the effects that LAT does have on the neuronal transcriptome, even though they are smaller than those seen with the HSV-1 miRNAs.

The importance of the production of major LAT intron and its function within neurons, perhaps in eliciting these transcriptional changes, could be due to its structure as a lariat (Wu et al., 1998). It has been shown that a debranching defect causes the HSV-1 major LAT intron to form a lariat and remain stable in this structure (Block & Taylor, 1996; Wu et al., 1998). In keeping with this, it has been suggested that circular RNAs, are less likely to become degraded (Salzman, 2016), which may explain why, as this lariat structure, major LAT is so stable and the most abundant product during latency. In recent years the importance of circular RNAs has started to be recognised, with the discovery that circular RNAs are ubiquitous across various species of eukaryotes (Salzman, 2016). In addition to this, circular RNAs have recently been shown to be important in Gammaherpesvirus biology whereby Ungerleider *et al.* revealed that both EBV and KSHV express various viral circular RNAs at different points in their life cycles. For example, different circular RNAs are predominant during KSHV lytic infection than those during latency (Ungerleider et al., 2018). Therefore, it might be this structure that allows the major LAT intron to accumulate and elicit the transcriptional changes described in this study.

Utilising the lentivirus and in1382 results in combination allowed the list of genes with significantly altered expression to be narrowed down to less than 200 genes as described above (figure 4.16 and figure 4.17). This enabled the functions of certain

genes and the wider biological processes and pathways that are affected during HSV-1 latency to be considered.

Functions of individual overlapping DEGs following LAT transduction or in1382 infection

PCDH17 is one of the two neuronal genes significantly upregulated in the presence of the major LAT intron either delivered by the LAT intron lentivirus or during quiescent HSV-1 infection. It has been shown previously that DNA viruses, including other members of the herpesvirus family, regulate protocadherin production at the protein level, during lytic infection (Soday et al., 2019; Weekes et al., 2014). Weekes' group developed a technique called quantitative temporal viromics, a large-scale proteomic analysis using tandem-mass-tag-based mass spectrometry, to show how HCMV directs the expression of more than 8000 cellular proteins, including the downregulation of 8 protocadherins, 6 potently so (Weekes et al., 2014). In a related follow-up experiment, Soday *et al.* performed temporal proteomic analysis on vaccinia virus infected cells to quantify the levels of roughly 9000 cellular proteins plus around 80% of viral proteins, at various time points post infection. Protocadherin proteins were among those downregulated again in response to lytic infection (Soday et al., 2019). On observing the downregulation of multiple protocadherins in response to HCMV alongside some known natural killer ligands, Weekes *et al.* hypothesised that the protocadherins may represent unknown natural killer ligands. Weekes *et al.* found a reduction in degranulation following knockdown of certain protocadherins that suggest they are immunomodulator molecules (Weekes et al., 2014). Therefore, the viruses could downregulate these proteins as an immune evasion tactic.

Following either transduction with the LAT intron lentivirus or in1382 infection, the opposite effect can be seen at the transcriptional level with a protocadherin; PCDH17, being upregulated. Therefore, perhaps during lytic infection viruses downregulate these proteins as a means of evading the immune system, however as latency itself is a means of evading the immune response as well (as there is no virion production or DNA replication to be detected by immune components), downregulation may be unnecessary. Of possible relevance to HSV-1 neuronal latency, PCDH17 has also been shown to induce cell cycle arrest in its role as a tumour suppressor gene (He et al., 2019), cell cycle arrest could be useful for maintenance of neuronal latency, as described previously. Additionally, PCDH17 has been linked to neuronal function. Hayashi *et al.* showed that knockdown of PCDH17 in mice led to misorientated axons in the amygdala, suggesting that PCDH17 is required for normal axon extension in the amygdala neurons

(Hayashi et al., 2014). Furthermore, Hoshina *et al*. showed that PCDH17 has a role in the regulation of presynaptic vesicle assembly and that a deficiency in PCDH17 can lead to altered presynaptic function (Hoshina et al., 2013). Therefore, by promoting cell-cycle arrest and effective neuronal cell-to-cell connections, the upregulation of PCDH17 by LAT during HSV-1 latency may support the ongoing maintenance of viral latency whilst sustaining the survival of the host cell. This could be yet another example of a potential neuroprotective role for major LAT.

DACT1 was one of the 3 genes significantly downregulated due to presence of the major LAT intron, through delivery by either the LAT intron lentivirus or quiescent in1382 infection (but not due to LAT without splicing capabilities as delivered by the NSLAT lentivirus). DACT1 is known as dishevelled binding antagonist of β-catenin 1 and as the name indicates DACT1 interacts with dishevelled – a key protein component of the Wnt signalling pathway – to target its degradation. Degradation of dishevelled leads to inhibition of Wnt signalling. A reduction in Wnt signalling causes decreased β-catenin levels, and in turn diminished activation of β-catenin responsive genes (Cheyette et al., 2002; Zhang et al., 2006).

The Wnt/β-catenin pathway is a key signalling pathway, crucial for embryonic development, cell fate and adult tissue homeostasis (MacDonald et al., 2009) and has been shown to promote neuronal differentiation, axonal repair and cell survival, which is important for the maintenance of HSV-1 latency (Harrison et al., 2020; Hirabayashi et al., 2004).

Clinton Jones' group first showed a link between the latency-reactivation cycle of alphaherpesviruses and the Wnt/β-catenin pathway using bovine herpes virus 1 (BoHV-1) (Liu et al., 2016; Workman et al., 2018). Liu *et al.* revealed through microarray and RNA-Seq that Wnt/β-catenin signalling pathway components were differentially expressed in calf neurons during BoHV-1 latency-reactivation cycle (Liu et al., 2016). In keeping with this, Clinton Jones' group also later discovered that ORF2 – the protein encoded by BoHV-1 latency related RNA during latent infection – interacts with components of the Wnt pathway leading to increased β-catenin dependent transcription (Workman et al., 2018; Zhu et al., 2017). Following on from this work, Harrison *et al.* demonstrated that mice trigeminal ganglia neurons latently infected with wild-type HSV-1 expressed more β-catenin relative to trigeminal ganglia of either uninfected mice or those infected with a LAT-negative mutant. This suggests that LAT has a role in mediating the Wnt/β-catenin pathway, causing an upregulation in β-catenin which in turn could promote neuronal viability (Harrison et al., 2020).

In keeping with these previous reports, the data shown here suggests a role for LAT in the downregulation of DACT1 – a β-catenin antagonist. This suggests an alternative route by which LAT regulates the Wnt/β-catenin pathway and therefore promotes neuronal differentiation and survival. It would seem from this data and the literature, also suggesting an anti-apoptotic role for LAT (da Silva & Jones, 2013; Jiang et al., 2011; Jin et al., 2003) that neuronal survival is a key function for major LAT. This is unsurprising given that the neuron is the site of HSV-1 latency, therefore keeping the cell viable would promote the maintenance of latency.

Another gene found to be downregulated following LAT intron lentivirus transduction or in1382 infection was RECQL4, which encodes a DNA helicase. The RECQL4 protein acts to unwind DNA at the initiation of DNA replication, however, during HSV-1 latency and in post-mitotic neurons there should be no DNA replication. It has been shown that in some instances, re-initiation of the cell-cycle in post-mitotic neurons can lead to cell death (reviewed in Copani et al. 2007). This adds to the evidence that LAT has a protective role in neurons and that LAT may contribute to inhibiting replication in the cell and in turn maintaining latency.

As described in chapter 1.4.4, a link between Alzheimer's disease and long-term latent HSV-1 infection has been established. There has been epidemiological, immunological, genetic and molecular data to suggest that persistent and recurrent HSV-1 infection in human neurons leads to an increased risk of developing Alzheimer's disease (Piacentini et al., 2014). In terms of genetic factors for developing Alzheimer's disease, existence of the APOE-ε4 allele of the APOE gene has been shown to be linked to the development of Alzheimer's disease. APOE-ε4 has been shown to have limited ability to bind to amyloid-β protein, therefore its expression leads to the accumulation of amyloid-β, characteristic of Alzheimer's disease pathology (Piacentini et al., 2014). Interestingly, APOE-ε4 has also been implicated during HSV-1 infection, whereby the presence of APOE-ε4 has been shown to cause a greater viral load in the brain (Burgos et al., 2003). Evidently, it is worth exploring genetic factors that may influence the association of persistent HSV-1 infection and Alzheimer's disease. Therefore, it could be useful to assess the list of DEGs found to be directly impacted by the HSV-1 latency-associated ncRNAs in this study, for any impact on the protein misfolding associated with Alzheimer's disease. In accordance with this, the DEG upregulated in response to LAT, regardless of major LAT intron production, following in1382 infection or transduction with either the LAT intron lentivirus or NSLAT lentivirus – THBS4 – has been shown to localize to amyloid-β plaques in human Alzheimer's disease cases (Cáceres et al., 2007). It is possibly that THBS4 could have a role in the accumulation of amyloid-β that

contributes to Alzheimer's disease and therefore its upregulation by LAT could contribute to the association observed between long-term HSV-1 infection and Alzheimer's disease, especially given that HSV-1 DNA has been shown to localise to amyloid plaques (Wozniak et al., 2009).

These are examples of how DEGs consistently found in this study may have physiological relevance during HSV-1 latency. It would be interesting to continue to explore the potential onwards effect of the transcriptional regulation of the functionally relevant DEGs, especially those that reproducibly change in both methods of expressing the latency-associated ncRNAs or those with the largest fold-changes. One such way to continue analysis would be to perform large scale proteomics, such as the quantitative temporal viromic technique described by Weekes *et al.*, using the in1382 infection and lentivirus transductions described above to deliver the HSV-1 latency-associated ncRNAs.

Overlapping pathways of the DEGs found following miR lentivirus transduction and in1382 infection

The biological processes and pathways enriched in association with upregulated genes following either miRNA lentivirus transduction or in1382 infection were mainly associated with metabolic process and signalling (figure 4.18). Initially an association with signalling processes was shown between both the upregulated and downregulated DEGs following in1382 infection. However, when DAVID analysis was performed only on the overlapping DEGs following miR transduction or in1382 infection, the association with signalling pathways is specific to upregulation. In line with this, there has been published data to suggest that herpesvirus miRNAs preferentially target host genes in cell signalling pathways (Gao et al., 2009). Although miRNAs are most often associated with downregulation of gene targets, there are also incidents of miRNA-mediated upregulation (reviewed in Orang et al. 2014), and so this could explain the upregulated DEGs observed here. Moreover, there is evidence to suggest that the same miRNA can mediate upregulation or downregulation depending on the target mRNA (Vasudevan, 2012). For example, it has been suggested that miR-145 upregulates myocardin (MYOCD) protein but represses Kruppel like factor 4 (KLF4) protein in a cell-specific manner, which in turn contributes to determining cell-fate (Cordes et al., 2009). There has been evidence suggesting that upregulation by miRNAs may occur specifically in quiescent cells, such as the differentiated neurons used here (reviewed in Vasudevan 2012). This could also explain the association observed between the downregulated

DEGs and the cell-cycle, as inhibition of the cell cycle might allow for the miRNA upregulation of cell-signalling components observed here.

In further support of the latency-associated ncRNAs inhibiting the cell-cycle, the downregulated genes that overlap following either miR lentivirus transduction or quiescent infection seem largely to be associated with processes and pathways related to the cell cycle and division (figure 4.19). This supports the results from the earlier DAVID software analysis showing downregulated DEGs following in1382 to be associated with the cell-cycle. This data together suggests that despite some potential overlap in this functions of the latency-associated ncRNAs (i.e. with the downregulation of the DNA helicase – RECQL4, by major LAT), the miRNAs are the HSV-1 latency-associated ncRNAs largely responsible for inhibiting the cell-cycle and cell division during in1382 quiescent infection. This highlights the benefit of utilising the lentivirus method of delivery in addition to quiescent infection because it enables the transcriptome changes observed during quiescent infection to be attributed to either LAT, major LAT or the miRNAs. As already described, this restriction of the cell cycle may be a means of maintaining latency as keeping the cell senescent limits the replication components of the cell, which aids in the maintenance of HSV-1 latency (reviewed in Traylen et al. 2011).

Evidently there are advantages to combining the two systems. Firstly, it highlights the gene changes that are reproducibly observed despite the mode of delivering the latency-associated ncRNAs differing. Additionally, combining the models enables the changes observed during quiescent infection to be attributed to the specific latency-associated ncRNA.

The disadvantage in combining the results from the both systems is that these are two stand-alone experiments so there will be experimental variation and differences in each system. For example, the in1382 infections are performed at 39^{o}C and the lentivirus transductions are performed at 37^{o}C which could have an impact on neurons. However, both experiments are internally controlled with their respective mock differentiated SH-SY5Y cells harvested alongside either transduced or infected cells, at the same temperature. Accordingly, a gene that significantly changes across the different experiments, despite the different methods of expressing the HSV-1 latency-associated ncRNAs, would suggest a consistent effect of the latency-associated ncRNAs.

Further analysis of the changes elicited by the HSV-1 latency-associated miRNAs

As miRNAs have been shown to mediate their effects post-transcriptionally (Gosline et al., 2016), a further analysis was performed for these DEGs that were observed following

either quiescent in1382 infection or miR lentivirus transduction. This additional analysis was performed in DAVID (Huang et al., 2009b) to show any transcription factors that the DEGs found are known to be associated with. Several transcription factors were found to be associated with the DEGs tested (tables 4.12 and 4.13) (appendix 1, sheets 17 and 18).

Among these transcription factors observed following either in1382 infection or miR lentivirus transduction, there are some of particular interest. MEF2 has previous known associations with herpesviruses. Lomonte *et al.* showed that HSV-1 ICP0 counteracts HDAC repression of MEF2 (Lomonte et al., 2004). Additionally, MEF2 has been suggested to play a role in EBV reactivation from latency, whereby MEF2 has been shown to bind to the ZI elements of viral BZLF1 gene promotor (Zp), which is an essential component for EBV reactivation (Bryant & Farrell, 2002; Liu et al., 1997; Murata & Tsurumi, 2009). Furthermore, MEF2 has been shown to have a neuroprotective role (Dietrich, 2013). The transcription factor, CDP(CUX1), is also known to have a role within neurons, whereby CDP(CUX1) has been shown to be important in neuronal development, especially of the upper cortex and in dendritogenesis (Cubelos et al., 2010; Weiss & Nieto, 2019). In addition to this, CDP has been shown to have a role in the regulation of the cell-cycle, accelerating entry into the S (synthesis) phase of the cell cycle whereby replication occurs (Sansregret & Nepveu, 2008). Both neuroprotection and cell-cycle inhibition have been mentioned as downstream physiological effects of the transcriptional changes that the latency-associated ncRNAs elicit, therefore, these transcription factors are functionally relevant targets for the HSV-1 latency-associated miRNAs.

Of these transcription factors that were shown to be associated with the downregulated DEGs following in1382 infection or miR lentivirus transduction, CDP (CUX1) was chosen for onward RNAhybrid analysis to see whether it was a target for any of the HSV-1 latency-associated miRNAs. CDP(CUX1) was chosen not only because it was one of the transcription factors with the top association specifically to the downregulated DEGs, both in terms of number of DEGs and low p-value (table 4.13) but also because of its relevant physiological functions in neurons already mentioned, plus it's been shown to be a functional repressor (Sansregret & Nepveu, 2008) and therefore its association with these downregulated DEGs is in keeping with this. This makes CDP a good candidate for regulation by the HSV-1 miRNAs to at least partially mediate their downregulation of the DEGs observed, however, CDP was not found to be a significant DEG by RNA-Seq, which suggests if this transcription factor is targeted by the HSV-1 latency-associated miRNAs it may not be blocked at the transcriptional stage.

Using RNAhybrid, the transcription factor CDP was shown to be a putative target for 2 of the latency-associated miRNAs (table 4.15). This suggests these miRNAs could post-transcriptionally target CDP. This could explain why although DAVID analysis revealed CDP(CUX1) and other transcriptional factors to be associated with the miRNA downregulated DEGs (table 4.13), it did not have significant changes at the transcriptional level. The miRNAs may target transcription factors such as CDP and impact translation of the resulting protein, which in turn affects the levels of the genes that these transcription factors regulate. This could be another way the miRNAs mediate their effect on the neuronal transcriptome observed in this study.

The results showing that some of the downregulated DEGs or transcription factors associated with DEGs were cellular targets for the HSV-1 latency-associated miRNAs describes one method by which the HSV-1 miRNAs may be eliciting their effect on mRNA levels. It has been shown previously that HSV-1 miRNAs target viral mRNAs for degradation (Umbach et al., 2008). Similarly, cellular miRNAs have been shown to downregulate cellular target genes largely by re-directing target mRNA to the mRNA decay pathway. Although there is also evidence suggesting that miRNAs cause translational repression of targets and therefore inhibit protein production rather than mRNA levels (Huntzinger & Izaurralde, 2011; Selbach et al., 2008). This potential post-transcriptional targeting of host genes could also provide a reason why changes in transcript levels did not correlate with the changes seen protein levels for PCDH17 by western blot.

Validation of DEGs found and testing the impact on protein levels

Western blotting was one means of validation that was used for the DEGs to test whether the transcriptomic changes affected the protein changes. The pilot western blots performed on the proteins tested here showed inconclusive results. RECQL4 and LPAR1 were essentially undetectable and PCDH17 protein levels seem to decrease in conditions where transcription had increased. Nevertheless, these initial tests showed that this method could be effective at testing whether gene transcriptional changes correlate with protein level changes or not.

The western blots performed here (figure 4.21) had caveats that might explain the inconclusive results. One such issue was that the multiple antibodies for the 3 DEGs were combined in one blot. Although this is a cost and time efficient approach, there could have unwanted interactions between the antibodies and there was insufficient time to validate the antibodies properly to test this. The results show detection of PCDH17, which suggests that there are not significant spurious interactions with the PCDH17

antibody as it still detected protein. However, RECQL4 and LPAR1 proteins are not or only minimally detected in this blot (figure 4.21A), which could be in part due to cross interactions between these two antibodies, as this was not tested. Similarly, although this blot was washed before re-probing for beta-actin (methods 6.2.8), there is still a risk of bound antibody remaining and interfering with the results. On repeating the western blot protocol to test more of the DEGs observed by RNA-Seq and validated by qRT-PCR, it would be beneficial to strip and re-probe the same blot with each antibody in turn, to avoid such caveats. Nevertheless, western blotting would be a useful next step to determine which of the DEGs results in a change in protein levels.

It is possible that there could be a reduction of PCDH17 protein following LAT transduction, as observed by western blot (4.21C), despite an upregulation of the gene at the transcriptional level, as indicated by RNA-Seq. It has been shown in the literature that protein levels and transcript levels do not always correlate (Kumar et al., 2016; Maier et al., 2009; Vogel & Marcotte, 2012). Furthermore, mRNA can be targeted for post-transcriptional modifications that block translation, for example through binding of complementary miRNA, which clearly from the data the HSV-1 miRNAs appear to have host cellular targets. However, this study focused on host targets from downregulated genes because this is the main known function of miRNAs (Catalanotto et al., 2016). There are also physiological conditions that can effect translation efficiency such as temperature may affect tertiary folding of mRNA which in turn could impact how accessible the ribosomal binding site is or the abundance of proteins that act as translational modulators (Maier et al., 2009). This means that having expression of a gene does not necessarily equate to protein production. In addition, there are experimental considerations, such as RNA-Seq is highly sensitive and can detect even low abundance transcripts whereas western blotting can only detect relatively abundant proteins, therefore the protein levels may be underestimated compared to the transcript levels. Any of the above could explain why LAT seemed to cause an increase in transcription but a slight decrease in the protein levels of PCDH17.

In addition to this, performing western blots on proteins that had relatively small changes at the transcriptional level was always going to be difficult to observe a change at the protein level. Future work might benefit from focusing on performing western blots on those genes that had larger fold-changes – despite not overlapping in the different approaches of delivering the latency-associated ncRNAs. An alternative option, considering the low-fold changes and that LAT is not expressed in every cell, would be to assess protein levels using a single-cell approach such as immunofluorescence, which was attempted using antibodies for some of the DEGs observed. However, there

were some technical difficulties when validating the antibodies and not enough time to troubleshoot them. But this could be a useful future approach to test whether the transcriptional changes affected protein levels or localisation.

If further western blots or immunofluorescence experiments were performed to determine which of the validated DEGs results in a change in protein levels, as was started with PCDH17 here, further functional analysis assays could be designed, considering the specific protein's function. For example, for the DEGs seemingly with a neuroprotective function, an apoptosis assay could be performed, such as those looking as caspase protein levels, to determine whether the changes in the DEG protein levels had an impact on the level of apoptosis observed. This is just one example of the plethora of further avenues that could be tested, following on from the transcriptional changes observed in this study.

The follow-up work to the RNA-Seq experiments that were performed here mainly focused on validating the system through qRT-PCR and largely it seemed that the gene changes observed by RNA-Seq follow-through with qRT-PCR, showing similar fold changes with each method (table 4.16). Performing qRT-PCR on separately transduced SH-SY5Y cells also allowed the reproducibility of each DEG to be tested. Although not all the gene changes were followed up by qRT-PCR, from those that were it seems that the RNA-Seq data was mostly reproducible and accurate.

Comparing the gene changes that occur early after infection with in1382 and those that occur following establishment of latency

A repeat in1382 infection was performed, with an additional 2hpi timepoint along with the original 13dpi timepoint (figure 4.22A). This experiment was designed to determine whether the genes that change expression in response to virus at the early stages of infection, at 2hpi, differed from those that change once the expression of the latency-associated ncRNAs have been established. Seeing a difference in gene expression at these different time points would indicate that those changes observed at 13dpi are not simply a result of lasting transcriptional changes upon encountering virus but more specific to that time point, namely following expression of the latency-associated ncRNAs.

Although this experiment was performed exactly as the previous in1382 infection, aside from the additional timepoint, the LAT expression observed at 13dpi was considerably lower than in the previous experiment (figures 4.22B and figure 4.11B). This suggests variability in the system. Nevertheless, the pattern of expression at 13dpi compared to

2hpi was consistent with the establishment of latency, with LAT expression increasing and lytic gene expression decreasing, as indicated by ICP27 (figure 4.22B).

While there was a massive response early after encountering virus, at 2hpi, with over 12,000 significant gene changes and large fold-changes (figure 4.22C and tables 4.17 and 4.18), most of these changes did not overlap with those that occurred at 13dpi (figure 4.24) with only 696 overlapping DEGs out of the 12444 at 2hpi and out of the 2971 at 13dpi.

Examining the top DEGs that changed in response to the second 13d in1382 infection, the fold-changes are in a similar range to the initial 13dpi infection, despite a difference in the LAT abundance. It was also notable that there were many downregulated mitochondrial genes. This is interesting because it has previously been shown that mitochondrial DNA can trigger an antiviral response and that to counteract this, HSV-1 induces MT-DNA stress (West et al., 2015). Therefore, it seems this downregulation of mitochondrial genes is in keeping with this known immune evasion technique employed by HSV-1. HSV-1 employing immune evasion techniques like this one, despite quiescent infection, may contribute to the success of the virus in remaining latent in neurons for the lifetime of the host.

This difference in the specific gene changes, along with differences observed when pathway analysis was performed on these DEGs that had a 2-fold change or more following this second in1382 infection, suggested that the DEGs observed at 13dpi are largely unique and not a non-specific result of gene changes that remain after encountering virus. However, there were some overlaps in processes impacted at 2hpi and 13dpi with in1382, despite differences in the DEGs observed. There were 5 biological processes that overlapped as associated with downregulated DEGs observed at 2hpi or at 13dpi (figures 4.26 and 4.28). It seems that these processes, such as cell surface receptor signalling, are downregulated in response to HSV-1, no matter at what stage of infection. There were also processes and pathways unique to each time point, for the 2-fold or more upregulated DEGs there were no overlaps in the GO biological processes or KEGG pathways observed to be associated at 2hpi or 13dpi. This suggests that the processes associated with the 2-fold or more upregulated DEGs following 13dpi in1382 infection are specific to a time after the latency associated ncRNAs are produced.

Accordingly, there were some overlaps in the pathways and processes found to be associated with these 2-fold or more DEGs following 13dpi in1382 infection 2 and those following the 13d in1382 infection 1 or the DEGs observed after transduction with the latency-associated ncRNA expressing lentiviruses. For example, the cell-cell signalling

GO process and the alcoholism KEGG pathway were repeatedly associated with the DEGs following expression of the latency-associated ncRNAs (figures 4.7A, 4.14A, 4.15B, 4.27A and 4.28B).

This overlap in the DAVID (Huang et al., 2009b) pathway analysis results, indicated some reproducibility in the effect of the latency-associated ncRNAs, despite variations in the systems used to deliver them and having seperate RNA-Seq experiments. To test this reproducibility further, Venny 2.1 (Oliveros, 2015) was used to observe any overlaps in the DEGs that occurred in response to the second 13dpi in1382 infection and the previous 13d in1382 infection, eliminating any gene changes that also occurred at 2hpi. This revealed 374 overlapping significantly DEGs observed in both 13d in1382 experiments (figure 4.23). This test in reproducibility was taken even further, to observe whether any genes were consistently upregulated or downregulated following either 13d in1382 infection and the latency-associated ncRNA lentivirus transductions. Three genes were found to be consistently downregulated following either 13d in1382 infections and LAT intron lentivirus transduction, while 42 DEGs were observed to be either upregulated or downregulated in response to either 13d in1382 infections or transduction with the miR lentivirus. It might be that these genes that are consistently significantly differentially expressed in response to the latency-associated ncRNAs might be the ones worth performing further analysis on, along with those with the largest fold-changes.

Summary

Altogether the data showed that the latency-associated ncRNAs do impact the neuronal transcriptome with more than 1500 genes significantly differentially expressed during quiescent HSV-1 infection. More of these changes were attributable to the 5 latency-associated miRNAs tested than to major LAT. Regardless, this data suggests putative novel functions for the HSV-1 latency-associated ncRNAs in terms of their effect on the neuronal transcriptome as well as potentially aiding in explaining the mechanisms behind previously described functions, such as a neuroprotective role, an association with neurodegenerative diseases, maintenance of latency via inhibition of cell-cycle elements and regulation of reactivation. Further understanding the physiological impact of these transcriptomic changes may reveal important information about the long-term effect of latent HSV-1 infection on the human neuronal cell.

Although this study revealed many neuronal transcriptional changes elicited by the latency-associated ncRNAs, there is a caveat in testing the effect of the ncRNAs on the neuronal transcriptome in differentiated SH-SY5Y cells. Although once differentiated

these cells stain positively for neuronal markers and take on neuronal morphology (chapter 2), the SH-SY5Y cells are a transformed cell-line, originating from a metastatic bone tumour biopsy. Therefore, some of the observed transcriptome changes may be specific to the SH-SY5Y cells and not found in human neurons during latent HSV-1 infection. For example, PCDH17 and TFPI2 both have tumour suppressor roles. Therefore, the upregulation of these genes and the downregulation of cell-cycle components as indicated by DAVID analysis software, although relevant to maintaining latency as described above, could be a response to maintaining latency in cells derived from a malignant tumour. It could be necessary to dampen the cell-cycle for example, in these cells that prior to differentiation are transformed to divide indefinitely as opposed to post-mitotic human neurons that are senescent. Nevertheless, the DEGs observed in this study could be used as a basis for further study in other neuronal cell-types such as post-mortem explant of human primary neurons or induced pluripotent stem cell derived neurons. The data found here suggests a role for the HSV-1 latency-associated ncRNAs in exhibiting changes on the neuronal transcriptome and provides many pathways to examine further, to explore the interaction between human neurons and HSV-1 latency.

Future work

To fully understand the physiological implications of the DEGs observed during HSV-1 latency, follow-up work is required. However, the list of genes that significantly change in response to expression of major LAT or the latency-associated miRNAs (either by lentiviruses delivery or as part of a quiescent infection) provide a useful starting off point for further studies into the role of the latency-associated ncRNAs in neurons.

In following up this work, single-cell analyses, such as immunofluorescence or single-cell RNA-Seq might be useful to be able to distinguish between the cells expressing the latency-associated ncRNAs and those not. It could also be beneficial to use siRNA to knockdown a validated DEG and observe if there are any changes to latent or lytic HSV-1 infection. To analyse changes to latent HSV-1 infection, qRT-PCR could be performed to major LAT following knockdown of a DEG. To examine any effects on lytic HSV-1 infection viral titre could be examined by plaque assay during WT HSV-1 infection following knockdown of a DEG. Immunoprecipitation assays could be used to see if any of the gene changes observed are elicited through a direct interaction with the latency-associated ncRNAs. Another approach could be to look at the functional downstream effects of any of the changes observed. For example, to determine whether any of the gene changes suggested to be neuroprotective have an impact on neuronal cell death, necroptosis or apoptosis assays could be performed on the transduced or infected SH-

SY5Y cells. It would also be interesting to see whether any of these gene changes could be reproduced in primary neurons, in response to delivery of the latency-associated ncRNAs. For example, if primary mouse neurons were used and revealed the same gene homologues were differentially expressed, then knockout mice for any one of the DEGs could be used, to test how latency is impacted *in vivo*. Alternatively, latently infected post-mortem human trigeminal ganglia could be used to observe any differences in expression of the DEGs identified in this study between LAT positive and LAT negative neurons. Employing in1382 infection or the latency-associated ncRNA expressing lentiviruses in neurons derived from induced pluripotent stem-cells could be another route of exploring how reproducible these DEGs are in different neuronal models. Regardless of the approach taken, there are various exciting avenues of research that this data opens. Elucidating the onwards physiological impact of these transcriptional changes that the HSV-1 latency-associated ncRNAs induce in human neurons could be key to better understanding the balance between lytic and latent infection as well as the long-term impact of HSV-1 latency on human neurons and potentially the links with neurodegenerative diseases.

Chapter 5: General Discussion

Small animal models of HSV-1 latency have revealed crucial information about this host-virus interaction. For example, the discovery of LAT was first made in mouse and rabbit models before confirming in human ganglia, as described in chapter 1.4.3.1 (Rock et al., 1987; Stevens et al., 1987). Nonetheless, the impact of the latency-associated ncRNAs on latently infected human neurons is still not fully understood. The natural host of HSV-1 is humans and this ancient virus has co-evolved with its host for millions of years (McGeoch & Cook, 1994).Therefore, not all human genes and host-virus interactions will be represented in animal models. Hence, there needs to be an effective human model for expressing the latency-associated ncRNAs to fully examine their impact on human neurons.

Accordingly, the aim of this thesis was to establish an *in vitro* model to express the latency-associated ncRNAs in human neurons and use this to examine the effects that major LAT and the latency-associated HSV-1 miRNAs have on the human neuronal transcriptome.

Chapter 2 initially focused on the culture of SH-SY5Y into neurons, optimising a protocol for their differentiation. Once these differentiated SH-SY5Y cells were shown to successfully become neuronal in terms of morphology and positively staining for the candidate neuronal markers tested (NF-M, pNF-H, MAP2 and synaptophysin) the aim of this chapter was to establish latency in these cells using a selection of HSV-1 strains. The first method for establishing latency utilised a fully replication competent HSV-1 with eGFP reporter activity. Pre-treatment of cells with viral inhibitors (IFNα and ACV) prior to infection with this virus did lead to establishment of latency whereby reporter gene expression was silenced. Spontaneous reactivation was observed following removal of inhibitors, which would mean that viral inhibitors were required to maintain latency. As discussed in chapter 2.3, the use of these inhibitors could skew the effects seen on the neuronal transcriptome. Therefore, a replication-defective HSV-1 mutant (in1382) was explored to remove the need for using viral inhibitors. Using the HSV-1 mutant in1382, it was possible to establish a quiescent infection in differentiated SH-SY5Y cells whereby LAT expression increased over time to reach a peak at 11dpi and lytic expression remained low within the same timeframe. In this system LAT was expressed in an average of ~30-60% of cells, despite beta-galactosidase staining suggesting that the virus had infected almost 100% of cells in a parallel experiment at the same MOI (figure 2.6). This data mirrors data showing that only a proportion of latently infected mouse ganglia express LAT *in vivo* (Mehta et al., 1995).

In chapter 3 a more reductionist approach was taken, whereby lentiviruses were designed to express the latency-associated ncRNAs. The lentiviruses were engineered to encode either the first 3.1kb of LAT – including production of the major LAT intron, the first 3.1kb of LAT but with mutations to stop splicing of the major LAT intron, 5 of the latency-associated miRNAs or eGFP. The lentiviruses were shown to transduce differentiated SH-SY5Y cells and express LAT and the 5 latency-associated miRNAs. Peak LAT expression was observed 11dpi with RNAscope revealing that LAT was expressed in an average of 50 to 80% of the differentiated SH-SY5Y cells, at this timepoint. This is despite the data showing that the eGFP lentivirus (that the LAT lentiviruses are developed from) has approximately 100% transduction efficiency when using lower transduction units, as discussed in chapter 3.3. Nevertheless, this is a higher proportion of cells shown to express LAT than during quiescent HSV-1 infection.

In the final results chapter (chapter 4), the systems established in chapter 2 and 3 were utilised to examine the effects of the latency-associated ncRNAs on the neuronal transcriptome by employing RNA-Seq. The RNA-Seq experiments revealed that the latency-associated ncRNAs do impact the neuronal transcriptome, with more than 1500 significantly differentially expressed genes observed in each experiment. While more gene changes were found to be attributable to the latency-associated miRNAs than the major LAT intron, both revealed gene changes that could have interesting physiological implications either in their specific known functions or the biological pathways and processes they contribute to. Many of the changes observed following quiescent infection and/or lentivirus delivery of the latency-associated ncRNAs either had a known function within neurons or could have putative roles within some of the previously established phenotypes attributed to the latency-associated ncRNAs. For example, there were DEGs linked to regulation of reactivation, control of the Wnt pathway, a neuroprotective role and links to neurodegeneration, as discussed in more detail in chapter 4.3. This could aid in uncovering the mechanisms behind roles previously attributed to the latency-associated ncRNAs as well as revealing novel roles in altering the neuronal transcriptome. This work highlighted the benefit of using the models established in this thesis. Although there are advantages and disadvantages associated with each method of expressing the latency-associated ncRNAs, as discussed in chapters 2.3 and 4.3, combining the 2 approaches revealed previously unknown information on the impact that these ncRNAs have on the human neuronal transcriptome.

Throughout the duration of this PhD other human neuron HSV-1 latency models have been established such as LUHMES cells and HD10.6 CELLS, as discussed in chapter

1.4.3.3 (Edwards & Bloom, 2019; Thellman et al., 2017). This highlights the shift in the study of HSV-1 latency to focus on human *in vitro* methods. Both these *in vitro* models for HSV-1 latency provide scalable neuron-like cells that can be infected with HSV-1 to establish an infection whereby genome is present, but no viral replication is observed, consistent with latency. Therefore, either model could be used as alternative human *in vitro* models for latency. However, there are advantages to the *in vitro* approaches of delivering latency-associated ncRNAs to differentiated SH-SY5Y cells, developed in this thesis. In the published models, the LAT transcript levels decrease from 8dpi to very low levels whereas in the model herein LAT expression increases until 11dpi. Even at 14dpi where LAT levels are seen to decrease, they are maintained much higher than the lytic transcript levels, which is reflective of HSV-1 latency, but this is not the case for either published models, where LAT and lytic genes show equivalent levels (Edwards & Bloom, 2019; Thellman et al., 2017). To examine the effects of the latency-associated ncRNAs, their expression must be maintained. Additionally, the models used for RNA-Seq analysis in this thesis did not require antiviral pre-treatment, which both the published models do, which as described could influence the transcriptional changes observed. Furthermore, the combination of approaches discussed in this thesis have the added benefit of attributing effects found during quiescent infection to the specific latency-associated ncRNAs. Accordingly, the HSV-1 latency systems described in this thesis proved useful for the onwards analysis of transcriptome changes.

Despite the benefits of this model, there are aspects that could be improved upon. Given that between 30-80% of the differentiated SH-SY5Y cells were shown to express LAT following transduction with the LAT lentiviruses, it might have been useful to have a positive-selection method for eliminating non-transduced cells designed into each lentivirus. However, the initial data depicting the high transduction efficiency of the eGFP lentivirus in the differentiated SH-SY5Y cells made it seem as though a means of positively selecting only cells expressing the transgene was unnecessary given nearly 100% of cells were found to express eGFP (figure 3.2). Given that all the lentiviruses share the same backbone lentivirus and are produced in the same way as the eGFP lentivirus (chapter 3, figure 3.3A and methods 6.2.2.5 and 6.2.3.2), it was expected that transgene expression would be similar in all the lentiviruses. Accordingly, a positive-selection method was not included in the lentivirus genome. Additionally, the use of a positive regulator could have unknown consequences. For example, if an antibiotic resistance gene was used as a selectable marker (in the lentiviruses genome, it would not be known if the antibiotic itself may have an impact on the neuronal transcriptome. There is scope for antibiotics influencing the host cell transcriptome as Morgun *et al.*

demonstrated by microarray of ileum samples of mice treated with a cocktail of antibiotics. Morgun *et al.* observed over 1500 differentially expressed genes in these samples compared to samples from untreated mice, and 645 of these were still differentially expressed in germ-free mice where the gene expression changes could not be as a result of the antibiotics depleting the microbiota (Morgun et al., 2016). If an antibiotic used to positively select transduced cells did impact the human neuronal transcriptome, it could skew experiments testing the impact of the latency-associated ncRNAs delivered by the lentiviruses on the neuronal transcriptome.

An alternative option to including a selectable marker in the lentivirus could be to have a means of measuring transduction efficiency in all the lentiviruses much like the eGFP lentivirus has eGFP as a read-out of transduction. If all the lentiviruses encoded a reporter molecule, such as a fluorophore, downstream of the ncRNA to be transcribed it could be a visual representation of transduction efficiency as well as expression of the transgene. Nonetheless, the RNAscope *in situ* hybridisation technique utilised did provide an alternative means of measuring of LAT expression that perhaps eliminates the need for eGFP as a read-out for expression. However, the number of LAT-positive cells revealed by RNAscope did not correlate with the predicted frequency of transduced cells according to eGFP expression with the eGFP lentivirus transduction. Therefore, having eGFP in the LAT lentiviruses might have confirmed whether all the cells are actually transduced (as expected from the eGFP lentivirus) and only a proportion of these transduced cells express LAT or whether the transduction efficiency is reduced in the LAT lentiviruses. Thus, including a read-out for transduction efficiency could have been useful, but increasing the size of the lentiviral insert could also have had detrimental effects, such as reducing lentiviral titres (Kumar et al., 2001).

An alternative solution to the observation that only a proportion of the cells express LAT at any given time, especially as this may be an intrinsic characteristic of the LAT promotor (as discussed in chapter 3.3), would be to deploy single-cell RNA-Seq with this system. This would enable the examination of individual transduced or infected cells and compare the neuronal transcriptome in cells expressing LAT to those that are not. This could reveal any correlations between the latency-associated ncRNAs and cellular transcript expression, on a cell-by-cell basis. Therefore, this could be a useful next utility for the models established in this thesis, especially given that single-cell RNA-Seq has proven effective for examining the relationship between lytic HSV-1 infection and host cell genes (Wyler et al., 2019).

Another potential caveat to the model of expressing the latency-associated ncRNAs described here-in, is the use of differentiated SH-SY5Y cells. As discussed in chapters 1.4.3.4 and 4.3, these cells originate from a metastatic bone tumour biopsy. Therefore, although these cells become neuron-like in terms of morphology and positively staining for neuronal markers (chapter 2.2.1), as a transformed line, these cells may not fully recapitulate the molecular events of latency in human sensory neurons (Laemmle et al., 2019). There are *in vitro* neuronal models derived from human stem cells that more closely reflect human adult ganglionic sensory neurons. Both human embryonic stem cell-derived neurons and human induced-pluripotent stem-cells have been shown to be permissive to HSV-1 infection and with the use of viral inhibitors, allow for the establishment of latency (D'Aiuto et al., 2019; Pourchet et al., 2017), as described in chapter 1.4.3.3. Therefore, the results found in this thesis, in differentiated SH-SY5Y cells, might be best followed up in one of these stem-cell derived neuronal models for latency, to verify that the transcriptional changes observed are physiologically relevant.

Despite the potential caveats described, the models used in this thesis revealed novel functions for the HSV-1 latency-associated ncRNAs, in their effect on the neuronal transcriptome. A better understanding of the role of viral ncRNAs in herpesviruses is important given that herpesviruses encode the most ncRNAs of any virus family and it is becoming clear that viral ncRNAs can have a significant effect and diverse roles during infection including impacting the host cell (Tycowski et al., 2015). For example, Naqvi *et al.* used microarray to demonstrate that herpesvirus miRNAs, including HSV-1 miR-H1, have a significant impact on the host cell transcriptome in human oral keratinocytes (Naqvi et al., 2018). They were they able to use the downregulated DEGs observed to show that miR-H1 has predicted binding sites in multiple genes involved in the immune response and autophagy. This is especially interesting given the identification of several human herpesviruses in tissue associated with oral inflammatory diseases and the observation that there is increased expression of viral miRNAs (including HSV miR-H1) in gingival tissue during periodontitis (severe gum disease) (Naqvi et al., 2018). Clearly the ncRNAs produced by the human herpesviruses can have a significant impact on host cells, which in turn could contribute to disease pathogenesis. Additionally, it is particularly important to consider the impact of latent herpesvirus infections on the host cell latency can occur for the lifetime of the host, therefore any impact from the ncRNAs expressed may be long-lasting.

There is further evidence of herpesviruses having a large impact on the host cell, with EBV infection of resting B lymphocytes leading to the transformation of these cells into lymphoblastoid cell lines, that can grow indefinitely, *in vitro*. Wang *et al.* used RNA-Seq

to demonstrate that latent EBV infection induces significant transcriptome changes in human B lymphocytes (Wang et al., 2019). Moreover, these transcriptome changes included 94% of genes essential for lymphoblastoid cell line growth, as identified by a genome-wide clustered regularly interspaced short palindromic repeats (CRISPR) screen. Additionally, cell proliferation was among the pathways enriched. Therefore, it seems that these transcriptome changes induced by EBV encourage the transformation of host cells from resting B lymphocytes into lymphoblastoid cell lines (Wang et al., 2019). These are important findings given that EBV can cause tumorigenic diseases *in vivo* such as EBV-related lymphoproliferative diseases, Burkitt lymphomas, Hodgkin lymphomas, natural killer/T-cell lymphomas and epithelial carcinomas (Kanda, 2018). It also further demonstrates that there is precedent for herpesviruses altering the transcriptome of host cells during latency. This is especially relevant given that EBV also encodes ncRNAs and miRNAs during latency (Kanda, 2018). Therefore, it is possible that as LAT and the HSV-1 latency-associated miRNAs were found to in this thesis, the EBV latency-associated ncRNAs might also contribute to the host transcriptional changes observed. In support of the idea that the ncRNAs encoded by EBV have a significant impact on the host cells, Zhang *et al.* found that EBV microRNAs – miR-BART7 and miR- BART13 – levels are elevated in plasma samples from nasopharyngeal carcinoma patients when compared with control patients, suggesting a putative role for these EBV miRNAs in the tumorigenic nature of EBV, transforming the host cell (Zhang et al., 2015).

Similarly, HCMV has been found to have considerable impact on the state of the host cell. HCMV goes latent in myeloid progenitor cells. Shnayder *et al.* used single-cell RNA-Seq to demonstrate a relationship between latent HCMV infection and the expression of cell-surface molecules in human monocytes. The expression of MHC class II and CD74 (a MHC class II chaperone) inversely correlated with level of viral transcript. This high viral transcripts and low CD74/MHC class II phenotype in turn led to significantly improved reactivation. In addition, the cells with high viral transcript levels exhibited reduced immune-response gene expression (Shnayder et al., 2020). Taken together it seems HCMV transcripts expressed during latency are driving the cell to an anergic-like state with a reduction in the expression of MHC class II and CD74 molecules, which may in turn lead to the increased reactivation phenotype observed. This is especially interesting given that, as discussed in chapter 3 section 4.3, the downregulated genes and associated pathways following lentivirus delivery of the HSV-1 miRNAs suggested a putative role for these HSV-1 latency-associated ncRNAs in promoting reactivation. This observation is also in keeping with published data showing a role for miR-H1 and

miR-H6 in reactivation (chapter 4.3)(Barrozo et al., 2020). Therefore, there may be a conserved role in promoting reactivation for the herpesvirus miRNAs during latency. Furthermore, this is another example of a latent infection with a herpesvirus impacting the host-cell transcriptome.

The long-term impact of herpesvirus infection on host cells is of great significance given life-long latency can occur. This is an important consideration in an ageing population where people are living to ages where even slow effects elicited by these viruses may eventually manifest. This is especially pertinent with alphaherpesviruses, which can latently infect neurons of the CNS. The long-term effects of this have started to be examined with the mounting evidence that long-term HSV-1 infection in neurons is linked to an increased risk of developing Alzheimer's disease as described in chapter 1.4.4 (De Chiara et al., 2019). Better understanding the molecular interactions that occur during latency could uncover more details about the impact of HSV-1 on human neurons as well as provide targets for novel treatments to target latency. Accordingly, the list of differentially expressed genes found in this thesis is a useful starting point to further examine the impact that the latency-associated ncRNAs have on human neurons.

Chapter 6: Materials and Methods

6.1 Materials

6.1.1 General reagents

General-purpose reagents were supplied by Sigma-Aldrich, Promega and Thermofisher Scientific. Plasticware and materials for cell culture were obtained from STARLAB and Corning. Tissue culture media were purchased from Gibco (Thermofisher Scientific). Foetal calf serum (FCS) was supplied by Gibco (for sh-sy5y cells) and PAN-biotech (all other cells).

Probes and primers were purchased from Abcam, ACD (Bio-Techne) and Eurofins. Antibodies were purchased from the companies indicated in table 6.1.2. Molecular biology enzymes were purchased from Promega.

6.1.2 Table of Antibodies

Antibody	Manufacturer – product code	Dilution used
Donkey anti-Mouse IgG (H+L) secondary antibody, Alkaline phosphatase (AP)	Thermofisher – Catalogue number (CN): A16014	1:5000 for western blot (WB)
Goat anti Mouse IgG (H+L) Cross-adsorbed secondary antibody, Alexa fluor 568	Thermofisher – CN: A11004	1:500
Goat anti Rabbit IgG (H+L) Cross-adsorbed secondary antibody, Alexa fluor 488	Thermofisher – CN: A11008	1:500
Goat anti-Rabbit IgG (H+L) superclonal secondary antibody, horseradish peroxidase (HRP)	Thermofisher – CN: A27036	1:5000 for western blot (WB)
Mouse anti-Beta-actin monoclonal antibody	Sigma-Aldrich – CN: A2228	1:10000 for WB
Mouse anti pNF-H (SMI31)	BioLegend - CN: 801601100	1:500
Rabbit anti LPAR1	Abcam – CN: ab23698	IF: 1:50, WB: 1:2000
Rabbit anti MAP2	Sigma-Aldrich – CN: HPA012828	1:32
Rabbit anti NF-M (NEFM)	Sigma-Aldrich – CN: HPA023138	1:50
Rabbit anti PCDH17	Sigma-Aldrich – CN: HPA026817	Immunofluorescence (IF): 1:50-1:200 WB: 1:2000
Rabbit anti RECQL4	Novusbio (Bio-Techne) – CN: 25470002	1:10-1:50 for IF, 1:2000 for WB
Rabbit anti Synaptophysin	Abcam – CN: 14692	1:100

6.1.3 Table of Buffers, inhibitors, solutions or fixatives

Reagent	Composition
2-Mercaptoethanol	2-Mercaptoethanol, CN: M6250, Sigma
3-(N-Morpholino)propanesulfonic acid, 4-Morpholinepropanesulfonic acid (MOPS) buffer	10X MOPS buffer, consisting of: 0.2M MOPS free acid 0.05M sodium acetate 0.01M Na_2EDTA
5-bromo-4-chloro-3-indolyl β-D-galactopyranoside (X-gal) solution	5mM potassium ferrocyanide, 1mg ml^{-1} X-Gal in detergent solution
Acyclovir (ACV)	100µM final concentration ACV
Detergent solution	0.01% sodium deoxycholate, 0.02% Octylphenoxy poly(ethyleneoxy)ethanol (IGEPAL CA-630) in PBS/$MgCl_2$
Extracellular matrix (ECM) gel	ECM gel from Engelbreth-Holm-Swarm murine sarcoma, 8-12mg/ml in Dulbecco's modified eagle medium |(DMEM) with 50mg/l gentamicin sulfate. CN: E1270, Sigma-Aldrich.
Interferon α (IFNα)	Final concentration of 500U/ml IFNα
Laemmli sample buffer	2x Laemmli sample buffer (CN: 1610737, BIO-RAD)
MagicMark™ XP Western Protein Starndard	125mM Tris-HCl (pH 6.8), 10mM DTT, 17.4% Glycerol, 3% SDS, and 0.025% Bromophenol Blue providing protein molecular weight markers at 20, 30, 40, 50, 60, 80, 100, 120 and 220 kDa. (CN: LC5602, Thermofisher)
Methyl Violet	0.5% Methyl Violet

Northern blot washing buffer	0.1 M Maleic acid, 0.15 M NaCl; pH 7.5 (20° C); 0.3% (v/v) Tween 20
paraformaldehyde (PFA)	4% paraformaldehyde in PBS
PBS	138 mM NaCl, 2.7 mM KCl, 8 mM Na2HPO4, 1.5 mM KH2HPO4 pH 7.6
PBS/$MgCl_2$	2mM $MgCl_2$ in PBS
Radioimmunoprecipitation assay *buffer (*RIPA) buffer	50mM Tris HCl (pH 7.4), 150mM NaCl, 0.5% deoxycholate, 1.0% NP40, 1mM EDTA, Complete Protease Inhibitor Cocktail Tablets (PI) with EDTA - Roche
Sodium dodecyl sulfate (SDS)	10% SDS in dH_2O
Sodium Carboxymethyl Cellulose (CMC)	5% Sodium CMC in PBS
Saline Sodium Citrate (SSC)	20X SSC containing 0.2µm filtered 3M NaCl in 0.3M sodium citrate (pH 7.0). CN:S6639, Sigma-Aldrich
TAE	40 mM Tris-Acetate, 1 mM EDTA, pH 7.85
(1X) TBS solution	10% 10x tris buffered saline (TBS)(200mM Tris base and 1.5M NaCl) in distilled water (ddH_2O)
TBST solution	10% 10x TBS and 0.1% TWEEN 20 in ddH_2O
TGS buffer	1x concentrate Tris-Glycine-SDS buffer
Trypsin EDTA solution	0.25% Sterile – filtered Trypsin, BioReagent, 2.5g porcine trypsin and 0.2g EDTA in 4Na per Litre of Hanks' Balanced Salt Solution with phenol red (SIGMA – T4049)

6.1.4 Table of Cell lines

Cell-line	Source	Description	Use
Mammalian			
ATCC-Vero	ATCC	African green monkey kidney epithelial cells	• growing and titrating wildtype HSV-1 virus stocks • Control cell-line
BHK	Cambridge university	Baby hamster kidney fibroblasts	• Growth and titration of HSV-1 strain in1382
HEK293T	NIBSC	human embryonic kidney cells	• lentivirus transfections for growing lentivirus stocks • Control cell-line
SH-SY5Y	ATCC	human neuroblastoma cells	• Differentiation into neuron-like cells for HSV-1 latency models and titrating lentivirus stocks
Bacterial			
HB101	Promega	Single-use competent Escherichia coli cell-line	• Transformation of lentiviral plasmids

6.1.5 Table of Plasmids

HSV-1 genetic coordinates used throughout this project were determined from GenBank accession number NC_001806.2 for HSV-1 strain 17 (Duncan J. McGeoch, 1987).

Plasmid	Origin
pCMVδR8.91 (Zufferey et al., 1997)	Packaging plasmid used for lentivirus production. Mammalian expression of lentiviral genes gag-pol, tat and rev, driven by CMV promoter (Zufferey et al., 1997)
pLentiLAT (figure 3.3B)	Derived from pRRLSIN.cPPT.PGK-GFP.WPRE and pUC57, contains the first 3069 bp of the primary LAT transcript (nt 118800 - 121870) in place of eGFP on pRRLSIN.cPPT.PGK-GFP.WPRE. Generation described in 6.2.3.
pLenti-miR-2-7 (figure 3.4B)	Derived from pRRLSIN.cPPT.PGK-GFP.WPRE, with eGFP replaced by a synthetic construct (GENEWIZ) comprised of HSV-1 microRNA stem-loops (and 60bp of 5' and 3' flanking sequence): HSV-1-miR-H2 (nt 121777-121957); HSV-1-miR-H3 and HSV-1-miR-H4 (nt 125667-126018); HSV-1-miR-H6 (nt 118463-118252); and HSV-1-miR-H7 (nt 123352-123553) encoded in pUC57 as described in 6.2.3
pLentiNSLAT (figure 3.3C)	Derived from pRRLSIN.cPPT.PGK-GFP.WPRE, and is identical to pLentiLAT except for ten nucleotide deletions to remove both pairs of splice donor and acceptor sites (nt 119465-119466; 119469-119470; 119740-119741; 120414-120415; 121419-121420) which were introduced during HSV-1 sequence synthesis into pUC57(GENEWIZ). Generation described in 6.2.3.
pMD2.G	VSV-G-expressing envelope plasmid

(Addgene plasmid no. 12259)	
pRRLSIN.cPPT.PGK-GFP.WPRE (Addgene plasmid no. 12252)	A 3rd generation lentiviral vector engineered to encode an enhanced GFP (eGFP) and 3' woodchuck hepatitis virus posttranscriptional regulatory element (WPRE), expressed from the human phosphoglycerate kinase promoter (hPGKp)
pSLAT4 (Arthur et al., 1998)	HSV-1 sequence 119292-120078 (787bp) consisting of the LAT intron. Used to generate the LAT intron probe for northern blot analysis
pUC57 (GENEWIZ)	Cloning vector with ampicillin resistance marker, designed to encode each of the HSV-1 inserts described within this table above and therefore used to generate the plasmids pLentiLAT, pLentiNSLAT and pLenti-miR-2-7 as described in materials 6.2.2.5 and 6.2.3 below

6.1.6 Table of Primers and probes

Primer	Sequence	Concentration (nm)	USE (Taqman (TM) /SYBR green (SG) /Stem-loop (SL) PCR or RNAScope (RS))
Human			
Beta-Actin forward	AGGCCAACCGCGAGAAGATG	50	TM
Beta-Actin reverse	AGGATCTTCATGAGGTAGTCAGTCAG	900	TM
Beta-Actin probe	ATGCCCTCCCCCATGCCATCCTGCGT	50	TM
Beta-Actin primer mix	Hs_ACTB_2_SG QuantiTect Primer Assay (QT01680476)		SG
GAPDH forward	CGGCTACTAGCGGTTTTACG	300	TM
GAPDH reverse	AAGAAGATGCGGCTGACTGT	900	TM
GAPDH probe	CACGTAGCTCAGGCCTCAAGACCT	50	TM
GAPDH primer mix	Hs_GAPDH_1_SG QuantiTect Primer Assay (QT00079247)	Not Applicable (N/A) – used at 1x concentration	SG
ATF3 forward	AGCCCCTGAAGAAGATGAAAGG	120	SG
ATF3 reverse	GCAGGCACTCCGTCTTCTC	120	SG

BCAS1 forward	TCACTTGGGTCTTTGTTCGC	120	SG
BCAS1 reverse	CGGATGCCAACGGAAAGAAT	120	SG
CD34 forward	CAGGCTGGTACTTCCAAGGG	120	SG
CD34 reverse	ACGGTACTGCTACCCCAGAG	120	SG
CUX1 forward	AGCCAGATCTCACAGCTTGA	120	SG
CUX1 reverse	GCGGTTCTTCTCCAGCAAC	120	SG
DACT1 forward	GAGTTTTGTGGCACCAGCTG	120	SG
DACT1 reverse	GGAGACACTGCCCTGAAGAC	120	SG
GPR149 forward	CGATCTTCTGGAGGCTGTCT	120	SG
GPR149 reverse	TCCTGTCGGTGACCATCTTC	120	SG
GRIA2 forward	TGTCAAAGTGGTGAGAGGCA	120	SG
GRIA2 reverse	TCCACCACGCTTTGTCTGTA	120	SG
KCNAB3 forward	GTGACCCTAACCTCCCAAGC	120	SG
KCNAB3 reverse	GGGGAACTGCAGTCTTCCTC	120	SG
KCNS3 forward	ACTGGGGAAGCTGCTTACTT	120	SG
KCNS3 reverse	TGATGCCCCAGTACTCGATC	120	SG
LPAR1 forward	TTCACGTTGTCGCTCTCCTC	120	SG
LPAR1 reverse	GGGAGAAACTTTGCGCCTTC	120	SG

MYCT1 forward	CCCTTTCTGCAATGTCCACC	120	SG
MYCT1 reverse	GTTGAAAGGCCCACTCGAAG	120	SG
PAPPA2 forward	CACCGACACCATGATCCATG	120	SG
PAPPA2 reverse	AGCCACAGGTGTCACTAGTG	120	SG
PCDH17 forward	CCAGAAAGAGCCAGCCTGAG	120	SG
PCDH17 reverse	TGTCACAGCAGGAGCCTTTG	120	SG
PDE4D forward	CGTGAAGATGTCCACCTTGC	120	SG
PDE4D reverse	ATGTCACATGCCACAACCAC	120	SG
RECQL4 forward	GAGGAGCCTGCTACGGAGTG	120	SG
RECQL4 reverse	GGCAATTACGTACGGCTCAAC	120	SG
RSPH4A forward	TTCTACATAGGCTGGGGTCA	120	SG
RSPH4A reverse	GCTGCGAGTAGAACTGCTTC	120	SG
TFPI2 forward	TTTCTTTGGTGCGCAGAAGC	120	SG
TFPI2 reverse	TGCAAGTGAGTGTGGACGAC	120	SG
THSD7A forward	TGCTGCCATCCAAACTGAAC	120	SG
THSD7A reverse	CAAACTGGTTGCCTGCATCT	120	SG
YBX2 forward	CTGGTTCCTTCTCAGCCTGA	120	SG
YBX2 reverse	TGACCTGTACCCTACCCAGT	120	SG

Lentiviral vectors			
WPRE forward	ACCACCTGTCAGCTCCTTTC	900	TM
WPRE reverse	GAATTGTCAGTGCCCAACAG	300	TM
WPRE probe	CTATTGCCACGGCGGAACT	100	TM
HSV-1			
Major LAT intron probe	514951 – V-HHV1-Lat-intron probe	N/A	RS
Major LAT RNA forward	CCAGGCAGTAAGACCCAAGC	850	TM
Major LAT reverse	GGCCGGTGTCGCTGTAAC	300	TM
Major LAT probe	TCCCACCCCGCCTGTGTTTTT	50	TM
ICP27 forward	CCTTTCTCCAGTGCTACCTGAA	300	TM
ICP27 reverse	CCAGAATGACAAACACGAAGGAT	50	TM
ICP27 probe	TCCTTAATGTCCGCCAGACGCC	25	TM
ICP0 forward	GTGATGCCCCCCGAGTA	300	TM
ICP0 reverse (& RT primer)	GGGCGTCCCTTATTGTTTTC	900	TM
ICP0 probe	CACGGGTGCCGAGACCGC	50	TM

ICP47 forward	CCCACCCAAGGTGCTTAC	900	TM
ICP47 reverse	CCACCCACGAAACACAGG	300	TM
ICP47 probe	GTGGGTTTCTGTCGTCGGAG	100	TM
ICP6 forward	TGCCAGACCTGTTTTTCAAG	50	TM
ICP6 reverse	CTCCTGGATGGGTATCTGCT	300	TM
ICP6 probe	ACGGGGAGGAGTTCGAGAAG	100	TM
HSV-1 miR-H2-3p/miR-H3-3p/miR-h4-5p/miR-H6-3p/miR-H7-5p	TaqMan® MicroRNA Assays (thermofisher, CN:6444088)	N/A – used at 1x concentration	SL, TM

6.1.7 Table of Tissue culture media

Cell line	Media used
ATCC-Vero	Dulbecco's modified eagle medium (DMEM), 10% heat inactivated foetal calf serum (hiFCS), 1% 200mM L-glutamine (L-glut) and 1% penicillin-streptomycin (10 000 units penicillin, 10mg streptomycin) (penstrep)
BHK	DMEM (+ high glucose), 10% hiFCS, 1% L-glut, 1% penstrep

HEK293T	DMEM, 10% hiFCS, 1% 200mM L-glut, 1% penstrep
SH-SY5Y basic culture media	MEM, 15% FCS, 1% 200mM L-glut, 1% penstrep
SH-SY5Y primary differentiation media	DMEM nutrient mixture F12 + L-glut + sodium bicarbonate, 1% penstrep, 5% FCS supplemented with 10μM *all-trans* retinoic acid (ATRA) immediately before use
SH-SY5Y secondary differentiation media 1	Dulbeccos modified eagle medium (DMEM) nutrient mixture F12 + L-glut + sodium bicarbonate, 1% penstrep, 1% B-27, 20mM potassium chloride, 1% glutamax Supplemented with 10μM ATRA, 100ng/ml brain-derived neurotropic factor (BDNF)(Sigma-Aldrich) immediately before use
SH-SY5Y secondary differentiation media 2	Neurobasal medium (Thermofisher), 10μM retinoic acid, 1% B-27 (Thermofisher), 20mM potassium chloride, 1% glutamax (gibco, Thermofisher) Supplemented with 10μM ATRA, 50ng/ml BDNF (sigma-aldrich) immediately before use.

6.1.8 Table of Viruses

Virus strain	Origin
17syn+	HSV-1 wildtype strain 17 (NCBI Reference Sequence: NC_001806.2 (McGeoch, 1987))
AAV2-GFP (Vector Biolabs)	Recombinant AAV (serotype 2) with eGFP expression driven by a ubiquitous CMV promoter. CN: 7004
C12	Replication competent recombinant reporter virus expressing eGFP driven by the human cytomegalovirus (HCMV) major immediate early (IE) 1 promotor (MIEP) inserted at the US5 locus in WT HSV-1 strain SC16 virus (Arthur et al., 2001)
eGFP lentivirus	Generated by co-transfecting HEK293T cells with pRRLSIN.cPPT.PGK-GFP.WPRE, pCMV8.91 and pMD-G plasmids
In1382	Replication defective recombinant HSV-1 mutant with RING domain deletion in ICP0, 12bp deletion from VP16 and temperature sensitive mutation in ICP4 and reporter virus expressing lacz driven by the HCMV MIEP in place of the HSV-1 thymidine kinase gene. (Coleman et al. 2008; Preston et al., 1997)
LAT intron lentivirus	Generated by co-transfecting HEK293T cells with pLentiLAT, pCMV8.91 and pMD-G plasmids
MiR lentivirus	Generated by co-transfecting HEK293T cells with pLenti-miR-2-7 , pCMV8.91 and pMD-G plasmids
NSLAT lentivirus	Generated by co-transfecting HEK293T cells with pLentiNSLAT, pCMV8.91 and pMD-G plasmids

6.2 Methods

6.2.1 Cell culture

6.2.1.1 Cell passage

All cells were incubated at 37°C and 5% CO_2 in a humidified incubator, unless otherwise specified, and genetically verified to confirm cell type (Eurofins, UK). The cell-types, their uses and the associated culture medium are described in tables 6.1.4 and 6.1.7 respectively.

To passage cells, the relevant culture media was removed, and the monolayer washed with PBS and trypsin-EDTA added to the tissue culture flask. SH-SY5Y cells were incubated in trypsin at room temperature (RTemp) for approximately 1 minute and no more than 1 minute 30 seconds, all other cell types were incubated for 5 minutes at 37°C and 5% CO_2. Trypsin was inactivated by the addition of culture media containing FCS and the cells were spun down (1000 rpm on a VWR MEGA STAR 1.6 centrifuge with TX-400 swing-out rotor for 2 minutes for SH-SY5Y cells, 2000 rpm for 3 minutes for all other cell-lines), supernatant removed, and cell pellet resuspended in the appropriate media. An appropriate proportion of cells (either calculated through the splitting ratio e.g. splitting 1:10 or by counting exact number of cells that needed seeding) were added along with the relevant media to a new tissue culture flask and returned to 37°C and 5% CO_2.

6.2.1.2 SH-SY5Y differentiation

SH-SY5Y cells were differentiated using modifications to previously published methods (Christensen et al., 2011; Encinas et al., 2000; Shipley et al., 2017). SH-SY5Y cells were gently lifted from flasks by short incubation with trypsin-EDTA (Sigma-Aldrich), added to 24 well plates ($5x10^4$ cells per well) and maintained in primary differentiation medium (table 6.1.7) for one week, with refeeding on days 3 and 5. On day 7, 24-well plate wells are coated with a 1/20 dilution of 8-12mg/ml ECM gel in DMEM (Sigma-Aldrich)(table 6.1.3). 300µl of ECM gel is added at 4^{o}C to each well (with or without a glass coverslip – depending on the final use of the cells), the plate is then left at 37^{o}C for 1-2 hours which activates polymerisation of the ECM. The differentiating SH-SY5Y cells were briefly incubated in trypsin EDTA and added to ECM-coated 24-well plates. Cells were then maintained in secondary differentiation media (table 6.1.7) for a further week, with media changes on days 9 and 11. Cells are considered differentiated from day 14, where they can be observed to take-on neuronal morphology and stain positively for neuronal markers (figure 2.1). Once differentiated, the cells are maintained for the length of time

required for any given experiment with thrice weekly medium changes in secondary differentiation media. Secondary differentiation media 2 was used unless otherwise stated.

6.2.2 Virus work

6.2.2.1 Virus stocks

Crude HSV virus stocks: ATCC-Vero cells were grown until confluent (~$5x10^7$ cells) in $175cm^2$ tissue culture flasks and infected at 0.01 pfu/cell of virus, in 10ml of ATCC-Vero tissue culture media (table 6.1.7). Cells and virus were incubated for 1 hour at 37°C and 5% CO_2 before adding 40ml more of tissue culture media to each flask. Virus-infected Vero cells were incubated for 2 days at 37°C and 5% CO_2, resulting in approximately 80% cytopathic effect (CPE) of the infected cell monolayer. Cells were harvested by scraping into the cell medium, this then underwent 1 freeze-thaw cycle at -80°C. The harvested cells were then centrifuged for 25 minutes at 2000rpm (VWR MEGA STAR 1.6 centrifuge with TX-400 swing-out rotor), the supernatant from this was harvested for media purifying virus stocks (described below). The cell and associated virus pellets were sonicated in a cup-horn (Q700 QSonica) Sonicator for 20 seconds on and 20 seconds off for a total of 1 minute. These were then resuspended in tissue culture media (either Vero media or secondary media minus supplements – see table 6.1.7 for details) and aliquoted into vials (100µl per vial) for long-term storage at -80°C.

For in1382, the same process was carried out except with measures to complement the replication defect introduced due to the mutations present in this virus (table 6.1.8). The virus was grown on BHK-21 cells in BHK media supplemented with 3mM Hexamethylene bisacetamide (HMBA) at 34°C.

Media purified virus stocks: Stocks were generated as above, except following 2000rpm centrifugation, the resulting supernatant was centrifuged (on Beckman Coulter Optima™/Avanti® ultracentrifuges) on a type 19 rotor at 15 000 rpm for 2 hours to pellet cell free virus. Supernatant was removed and pellets left in residual media at 4°C overnight before resuspending in fresh media and aliquoted into vials (50µl per vial) for long-term storage at -80°C.

Media purified stocks were used in the experiments included in this thesis except for C12 infections, in which a crude stock was used.

6.2.2.2 Plaque assays

Plaque assays were used to titrate viral stocks. Vero cells for wildtype HSV-1 stocks or BHK cells for in1382 virus stocks were plated on 6-well plates until approximately an 80% confluency was established. A 10-fold dilution series was performed on each virus stock using Vero tissue culture media (table 6.1.7) from the neat virus stock up to 10^{-9} dilution on a new stock. Media was aspirated from the Vero cell monolayers and 200µl of each virus dilution was added to the Vero cell monolayer and incubated for 1 hour at 37ºC and 5% CO_2, the plates were agitated at ~15-minute intervals to disperse virus inoculum. Following this incubation time, a pre-warmed 5% CMC (table 6.1.3) and BHK/Vero media (table 6.1.7) mixture (1:4 CMC:Media → 1% CMC in media) was added on top of the infected Vero cells, to completely cover each 6-well plate well (2-3ml). These were incubated for ~2 days at 37ºC and 5% CO_2. Following this incubation time, the CMC-media overlay was removed, the Vero cells rinsed gently with PBS and fixed and stained with methyl-violet (table 6.1.3) for at least 30 minutes. The excess methyl-violet was then rinsed off and plaques counted (on a 10X magnifying Microtec HM-2 stereo microscope). The viral titre could then be calculated considering the titration value, the volume of virus dilution added to the Vero cells and the number of plaques observed. Where multiple titrations yielded countable plaques, the average viral titre generated across these was taken and often across duplicate wells also.

6.2.2.3 In vitro growth curve analysis

Growth curves were performed on Vero and differentiated SH-SY5Y cell cultures at MOI 3. Cells were incubated for one hour and extracellular virus washed away by rinsing three times with PBS. Infected cells were sampled at set time points over a 24-hour period and stored at -70°C prior to plaque assay (section 7.2.2.2).

6.2.2.4 Infections and transductions

Differentiated SH-SY5Y cells were infected with the different HSV-1 strains (table 6.1.8) in the relevant media for the cells used and at the relevant MOI, as indicated in each experiment, calculated according to the 50,000 cells plated per well.

Differentiated SH-SY5Y cells were transduced at the genome per cell (gpc) or functional units (FUs) outlined in the results chapters in secondary differentiation media 2 (table 6.1.7).

Establishment of latency

Latency was established with HSV-1 in the differentiated SH-SY5Y by two different approaches as discussed in the results chapters:

Replication-defective mutant strain, in1382 (figure 2.5A), was used at MOI 1-5 (as described in the results), at 39^{o}C (to block expression of functional ICP4) on differentiated SH-SY5Y cells, the cells were maintained at this temperature for the duration of infection.

Replication-competent reporter virus, C12, was used in conjunction with viral inhibitors. Differentiated SH-SY5Y cells were pre-treated with 100μm ACV and 500U/ml IFNα for 24-hours before infecting at MOI 0.1. The culture was maintained at the concentrations of inhibitor mentioned until removal at 8dpi.

The final alternative method to express the latency-associated ncRNAs in the differentiated SH-SY5Y cells and their effects during latency, instead of using whole virus, was to transduce with lentiviruses expressing the latency-associated ncRNAs. The transductions were carried out as described above.

6.2.2.5 Lentivirus production

Transfections

The third-generation lentivirus vectors, LAT intron lentivirus, NSLAT lentivirus and miR lentivirus were generated by co-transfection of support plasmids pCMVδR8.91 (kindly gifted by Mary Collins) and pMD-G with pRRLSIN.cPPT.PGK-GFP.WPRE, pLentiLAT, pLentiNSLAT, or pLenti-miR-2-7 (table 6.1.5), into HEK293T cells using Fugene 6 transfection reagent, according to the manufacturer's instructions (Promega).

To achieve this, in a T175 flask, HEK293T cells were grown to ~50% confluency. 2.5μg of pMD-G, 2.5μg of pCMVδR8.91 and 3.75μg of the specific transgene-expressing lentivirus plasmid to be grown (pRRLSIN.cPPT.PGK-GFP.WPRE or pLentiLAT or pLentiNSLAT or pLenti-miR-2-7) was mixed in dH_2O to make it up to 55μl per flask transfecting and added to Fugene 6 Transfection reagent in a 1.2:1 ratio. The DNA:Fugene mix was rested at RTemp for 5 mins before adding to 10ml of HEK293T media and adding it to the HEK293T cells. This was incubated at 37^{o}C for 20 minutes before topping up with 20ml fresh HEK293T media and leaving to incubate at 37^{o}C for 48hours until the first supernatant collection for purification of the lentivirus.

Purification of lentiviral vector stocks

Lentivirus vectors were purified from cell supernatants 48- and 72-hours post-transfection. Filtered supernatants (0.45μm) were centrifuged in an SW32Ti rotor (Beckman) at 28,000 rpm for 2 hours at 4°C, and pelleted virions were resuspended in

secondary differentiation medium before aliquoting in 50-100µl stocks for long-term storage at -80ºC.

Titration of lentiviral vector stocks

To quantify functional lentivirus particles (i.e. reverse transcribed DNA provirus genomes), differentiated SH-SY5Y (approximately 50,000 cells per well) in 24-well plates were transduced with a dilution series of each lentivirus. DNA was extracted three days post-transduction (10 days used on initial titration test shown in figure 3.7, but once this method was established the transduction time was shortened to 3 days and still found to give similar titres), and reverse-transcribed provirus genomes were quantified with primer sets hybridising to lentiviral WPRE and cellular GAPDH DNA sequences (table 6.1.6).

6.2.3 Cloning techniques

6.2.3.1 Restriction digests

Restriction enzymes were chosen depending on their sites on the plasmid to be cloned. The restriction digests were performed using 1µg of plasmid DNA. With the relevant buffer and temperature for the restriction enzymes used (buffer E for BamH1 and Not1 digests and multicore buffer where Acc1 was also used). The restriction digests described below were all performed at 37ºC for ~1 hour. An approach was taken that 4 units of each restriction enzyme are required per µg of DNA over 1 hour, however as partial digests were used to encourage only partial cleavage with suboptimal buffer conditions to allow for double digests, the enzymes were working at ~25% capacity, therefore 10 units of each enzyme were used per µg of DNA.

6.2.3.2 Lentivirus construction

The pLentiLAT, pLentiNSLAT and pLenti-miR-2-7 plasmids (described in table 6.1.5 and shown in figures 3.3B, 3.3C and 3.4B respectively) cloning were constructed as follows:

1µg of pRRLSIN.cPPT.PGK-GFP.WPRE plasmid (Addgene plasmid #12252) was digested with BamH1 and Not1 to remove the DNA encoding eGFP, ready for insertion of DNA encoding the relevant HSV-1 ncRNA. A double digest was performed with BamH1 and Not1 (Promega), using a partial digest to account for there being 2 Not1 sites along the pRRLSIN.cPPT.PGK-GFP.WPRE plasmid and encourage only the cutting of the Not1 site surrounding the eGFP DNA.

Restriction digests were similarly performed on the GENEWIZ HSV-1 ncRNA constructs using BamH1 and Not1 on pUC57 synthesized to encode either the first 3069 bp of the

primary LAT transcript (NC_001806 nucleotides (nt) 118800 – 121870) or the first 3069 bp of the primary LAT transcript (nt 118800 – 121870) with ten nucleotide deletions introduced to remove both pairs of splice donor and acceptor sites (nt 119465-119466; 119469-119470; 119740-119741; 120414-120415; 121419-121420) (GENEWIZ).

To produce the miR lentivirus plasmid there was an additional restriction digest step as the HSV-1 sequence was engineered into 2 separate plasmid constructs initially: pUC57 was synthesized to encode either miR-H2-4 (NC_001806 nt 121777-121957 and 125667-126018) or miR-H6-7 (NC_001806 nt 118463-118252 and 123352-123553) stem-loops plus flanking sequence (GENEWIZ). The construct encoding MiR-H2-4 was digested with BamH1 and Acc1 (Promega) while the pUC57 plasmid encoding miR-H6-7 was digested with Not1 and Acc1.

Each of the pUC57 HSV-1 constructs now had BamHI and NotI primers engineered on to the 5' and 3' end of the HSV-1 sequence, ready for ligation with the previously digested pRRLSIN.cPPT.PGK-GFP.WPRE plasmid, in the presence of a T4 ligase (1-3u)(Promega) overnight (a triple ligation was performed for the 2 pUC57 miR constructs and pRRLSIN.cPPT.PGK-GFP.WPRE).

In this way the relevant HSV-1 sequences were synthesised (GENEWIZ) and inserted into the pRRLSIN.cPPT.PGK-GFP.WPRE parent plasmid to generate the plasmids described in table 6.1.5 and outlined in figures 3.3B, 3.3C and 3.4B.

Restriction digests were also performed as a diagnostic tool to confirm the structure of each Lentivirus vector. This involved performing restriction digests on colonies picked from bacteria transformed with the newly ligated plasmids using relevant restriction enzymes with known positions on the ligated plasmid. The known DNA fragment sizes, given the position of restriction sites, were then checked by DNA agarose gel electrophoresis.

6.2.3.3 DNA agarose gel electrophoresis

The products of each restriction digest were run on a DNA agarose gel by electrophoresis for gel DNA extraction. DNA gel electrophoresis was also performed following diagnostic digest of plasmids on transformed DNA colonies (as in 3.4C) to test that the band sizes matched those expected following restriction digestions of the ligated plasmids.

0.8% agarose gels were used when gel extraction would be required or 1% agarose gels for the analysis of diagnostic restriction enzyme digests.

Making the gel

The relevant amount of agarose (Invitrogen) was dissolved in the necessary amount of TAE buffer (table 6.1.3) depending on onwards use and the size of the gel to be used (E.g. 1.2g agarose in 150ml TAE for an 0.8% agarose gel approximately 15cm^2 in size, that would be used for onwards gel extraction). The agarose and TAE solution was gently mixed and microwaved until the agarose had fully dissolved. This was then left to cool slightly, so still liquid and hot to touch but no vapours emitted. At this point 0.01% of ethidium bromide (EtBr) was added to the agarose-TAE gel mix (i.e. 15µl in 150ml mix) and gently mixed. The gel mixture was poured into a gel rig that had been prepared with masking tape around the peripheries and a teeth comb for the appropriate number of wells. The gel was left to solidify in the rig. Once set, the masking tape was removed and the gel and rig were added to a gel electrophoresis tank and TAE poured in, enough to submerge the gel.

Running the gel

The gel was run in TAE (table 6.1.3). 4.5µg of 1kb DNA ladder (promega) was added to the first and last wells of the gel. The samples were then mixed with with loading buffer (Thermofisher, 6x DNA gel loading dye, R0611)(1x final concentration) and loaded to the rest of the wells.

The cables were connected to the gel tank and the power pack and run for an appropriate time and voltage, again varying depending on the gel size and the expected size of the DNA samples being run (e.g. 100V for 1-2 hours), gels were visualised and run further where appropriate.

Visualizing the gel

Gels where no DNA was to be extracted were visualised and imaged on a SYNGENE G:BOX Chemi XX6 gel doc system.

Where DNA was to be extracted from a gel, the gel was visualised on a maestrogen MLB-21 UltraBright UV Transilluminator.

6.2.3.4 Gel extraction of DNA

For purification of DNA fragments for cloning, the band of the expected size was extracted using QIAquick gel extraction kit (Qiagen) following electrophoresis in a 0.8% agarose gel as described above. DNA was visualised on a low-power UV lightbox and the relevant sized DNA fragments excised from the gel using a scalpel. The DNA was

then purified from the gel using the QIAquick gel extraction kit (Qiagen), according to the manufacturer's instructions.

6.2.3.5 Ligation of DNA into plasmids

The extracted DNA was ligated into the lentivirus backbone plasmid (pUC57 HSV ncRNA inserts into the pRRLSIN.cPPT.PGK-GFP.WPRE with GFP removed lentivirus vector backbone). To allow the best chance of successful ligation, the lentiviral backbone plasmid:insert DNA were added in 3 different ratios: 1:2, 1:5 and 1:10 (amounting to between 20-84ng of DNA, depending on the ratio and size of DNA inserts) along with 1-3u T4 ligase in a (1:1) ratio with 1x T4 buffer, and 10x T4 ligase buffer diluted to 1x in enough dH_2O to facilitate this. The mixture was left at RTemp overnight.

6.2.3.6 Transformation

The various ratios of overnight ligated plasmid DNA (20-84ng) were added to thawed competent HB101 Escherichia coli cells (Promega). After half-an-hour on ice the bacteria were heat-shocked at 42^{o}C for 45 seconds and then cold-shocked on ice for at least 45 seconds. Following addition of 500µl of Luria broth (LB), the bacteria were incubated at 37^{o}C for 30 to 45 minutes. The transformed bacteria were spread onto an 100µg/ml ampicillin agar plate and grown overnight. The ampicillin-resistant colonies were picked and grown up in 5ml 100µg/ml ampicillin and LB for 7-8 hours, this was then added to 50ml with 100µg/ml ampicillin LB overnight.

6.2.3.7 Plasmid DNA preparation

The overnight bacterial cultures grown up in 50ml LB were miniprepped using the Wizard plus SV Miniprep Kit (Promega) according to the manufacturer's instructions, to isolate the lentivirus plasmid DNA. Glycerol stocks were also made from these overnight cultures and frozen at -70.

6.2.3.8 Formaldehyde gel electrophoresis

1.2% agarose (Invitrogen) and 1x MOPS buffer (table 6.1.3) were mixed in dH_2O and microwaved until the agarose dissolved (approximately 2 mins). Once this had cooled slightly but not yet solidified, 2% formaldehyde was added and poured into the gel rig to set.

Meanwhile, the 1-2µg of the RNA samples for onwards northern blot analysis were mixed with 1µg loading buffer (Sigma-Aldrich, RNA sample loading buffer, R4268). These were denatured at 65^{O}C for 15mins and then put on ice for 10 minutes. EtBr was added

(0.01%). These samples were then loaded onto the set formaldehyde gel and run alongside 2µg Millennium™ Formamide RNA Markers for 4.5 hours at 45V.

The gel was UV imaged on a SYNGENE G:BOX Chemi XX6 gel doc system alongside a fluorescent ruler for later sizing of the northern blot the RNA from this gel is to be transferred onto.

6.2.4 Staining techniques

6.2.4.1 β-galactosidase assay

Cells were fixed in 4% paraformaldehyde in PBS for 20 minutes and rinsed in PBS. The cells are then rinsed in PBS/MgCl2 twice (table 6.1.3). The cells can be left at 4°C overnight at this point. The cells are covered in X-gal solution (table 6.1.3) and incubated at room temperature overnight. The X-gal solution is removed, and cells rinsed with PBS followed by H_2O. If the cells were plated onto a coverslip and require imaging, the coverslip could then be mounted using Fluoroshield (with DAPI) mounting media (sigma-aldrich) onto a glass slide. These slides were left to dry, sealed with nail varnish then imaged on an Olympia IX71 fluorescent microscope.

6.2.4.2 Immunofluorescence analysis

SH-SY5Y cells were plated and differentiated at 50,000 cells per 24-well plate well and onto an ECM-coated glass coverslip as described in 6.2.1.2. Following differentiation, infection or transduction cells were rinsed in phosphate buffered saline (PBS) and then fixed in 4% PFA for 20 minutes before a further 2 rinses in PBS at room temperature. Cells were permeabilised in 0.1% Triton X for 10 minutes before a further 2 washes in PBS. Cells were then blocked for 30 minutes in 5-20% normal horse serum. Primary antibodies were diluted as indicated (table 6.1.2) and added to the coverslips. These were incubated for 1.5 hour(s) at room temperature before rinsing 3 times with PBS and repeating with the fluorescently labelled secondary antibodies for 45 minutes. Finally, the coverslips were rinsed in PBS then H_2O before mounting onto glass slides in Fluoroshield (with DAPI) mounting media (Sigma-aldrich). These slides were left to dry, sealed with nail varnish then imaged on an Olympia IX71 fluorescent microscope.

6.2.4.3 RNAscope (Fluorescent in-situ hybridisation)

Coverslips each containing approximately 50,000 differentiated SH-SY5Y cells (infected or transduced, as indicated) were fixed in 4% PFA for 30 minutes. The cultured adherent cell sample preparation for the RNAscope® Multiplex fluorescent v2 Assay protocol was followed according to the manufacturer's instructions (Advanced Cell Diagnostics Inc.

(ACD)). Briefly, cells were dehydrated by exposure to increasing concentrations of ethanol (50-100%) stored at -20 for no longer than 6 months. Cells were then rehydrated by exposure to decreasing concentrations of ethanol (100-50%) and submerged in PBS for 10 minutes. Then the rehydrated differentiated SH-SY5Y cells were completely covered in RNAscope® Hydrogen Peroxide (ACD) (1-3 drops per coverslip) for 10 minutes then before rinsing twice in distilled water, followed by RNAscope® Protease III (ACD)(diluted 1 in 15 in PBS) for 10 minutes at room temperature. The RNAscope® Multiplex Fluorescent v2 Assay was then performed as per the manufacturer's instructions (ACD). Briefly, cells were hybridized for 2hrs at 40^{o}C in a HybEZ^{TM} II oven, using probes designed against the LAT intron (Reference 514951 - C1, ACD-Bio-Techne), before undergoing a series of treatments to develop and amplify fluorescent probe signal, washing the samples in a 1x diluted 50x RNAscope® wash buffer (ACD) in dH_2O between each step. Finally, the cells were mounted in ProLong Gold Antifade Mountant, containing DAPI. These slides were left to dry, sealed with nail varnish then imaged on an Olympia IX71 fluorescent microscope.

6.2.5 Molecular biology techniques

6.2.5.1 RNA extraction

RNA was extracted from cells for onwards PCR or RNA-Seq experiments using TRIzol® (Thermofisher) according to the manufacturer's instructions. 200µl TRIzol® reagent (contains Phenol and guanidine isothiocyanate) was added to each well containing ~50,000 differentiated SH-SY5Y cells and incubated at room temperature for 5 minutes. Cell lysates were collected from each well and frozen at -80^{o}C until completing the RNA extraction protocol. For this, 40µl chloroform was added (0.2ml for every 1 ml of trizol originally added) to the samples and mixed vigorously by hand and incubated for 2-3 minutes at room temperature. The samples were centrifuged at 12,000g for 15 minutes at 4^{o}C (on a Heraeus Fresco 21 refrigerated centrifuge, with 2.0mL rotor) to separate the upper aqueous phase, containing the RNA, from an interphase and the lower phenol-chloroform phase (containing the DNA and proteins). This RNA is taken up by careful pipetting and added to an Eppendorf containing 10µg of glycogen as a carrier for this aqueous phase. 500µl of isopropanol was added to the aqueous phase and incubated for 10 minutes at room temperature. These samples were centrifuged at 12,000g for 10 minutes at 4^{o}C. The supernatant was removed, leaving only the RNA pellet, which was washed in 200µl 75% ethanol (the RNA can be stored at -20^{o}C for up to a year at this stage). The samples were vortexed and centrifuged at 7,500g for 5 minutes at 4^{o}C before discarding the wash. The RNA pellets were air dried, for approximately 10 to 15 minutes.

RNA was resuspended in 50µl RNase-free water and incubated at 55-60ºC for 10-15 minutes. The RNA was then taken through a DNA removal procedure to remove any contaminating DNA not removed with the interphase layer, using a TURBO DNA-free™ kit (thermofisher), according to the manufacturer's instructions. Briefly, TURBO DNase buffer and TURBO DNase was added to the RNA and mixed. This combination was incubated at 37ºC for 30min and then the DNase Inactivation Reagent was added and mixed. This is incubated for 5 mins at RTemp (mixing occasionally). Finally, the samples were centrifuged for 1.5min at 10,000g (on a Heraeus Fresco 21 refrigerated centrifuge, with 2.0mL rotor) and the RNA was transferred to a new tube, free from DNA and inactivated DNase.

6.2.5.2 DNA extraction

Nucleic acid was extracted from differentiated SH-SY5Y cells by incubating with 200µl TRIzol (Thermofisher) per wells. DNA was extracted from TRIzol preparations following removal of RNA in the aqueous phase as described in section 7.2.5.1. Following centrifugation, DNA pellets were washed with 70% ethanol and incubated in 0.5% sodium dodecyl sulfate (SDS) and 50µg of proteinase K/ml in TE buffer (10 mM Tris HCl, 1 mM EDTA [pH 8]). DNA was purified using phenol-chloroform extraction and ethanol precipitation.

6.2.5.3 Polymerase chain reaction (PCR)

For reverse transcription PCRs (RT-PCR) on RNA, reverse transcription (RT) was performed using the SuperScript IV First-Strand synthesis system (Thermofisher) according to the manufacturer's instructions, using random primers for all but ICP0 (where a specific, reverse primer was used, see table 6.1.5). The RT was performed alongside samples where the reverse transcriptase was excluded to provide a RT-negative control for the onwards PCRs.

Quantitative PCRs (qRT-PCR/qPCR) were performed in triplicate for all HSV-1 genes and control host genes and were conducted using Hot-Start TaqMan (HotStarTaq) reagents (Qiagen). Reactions were performed in 20µl with 8mM of each dNTP, 0.8U HotStarTaq DNA polymerase, 1x HotStarTaq buffer, 1.5-5mM $MgCl_2$ and 8-16% DMSO with primer and TaqMan quantities as outlined in table 6.1.5. A dilution series of plasmids containing target sequences (10 copies to 100,000,000 copies) were utilised to generate a standard curve of each target sequence and QuantStudio™ Design & Analysis Software v1.5.0 used to quantify the sample sequence copy numbers using from each curve. The PCR conditions for all the samples were 95ºC for 15 minutes followed by 45

cycles of 30 seconds at 95ºC followed by 1 minute at 60ºC and were performed on a QuantStudio 3 system (Applied biosystems, Thermofisher).

For examining the fold-changes of host genes found to be differentially expressed following infections/transductions by RNA-Seq underwent SYBR green qRT-PCR in triplicate, alongside control host genes. Each PCR reaction was performed in 50µl using 1X SYBR green PCR master mix (Thermofisher) and 1X primer mix or as indicated in table 6.1.5. PCR conditions were 95ºC for 10 minutes followed by 40 cycles of 15 seconds at 95ºC followed by 1 minute at 60ºC and were performed on a QuantStudio 3 system (Applied biosystems, Thermofisher). The Ct values achieved during the PCRs, as recorded by the QuantStudio™ Design & Analysis Software were added to the delta-delta-CT (ddCT) formula [2^{-ddCt} = (sample A (eg infected): Ct gene of interest – Ct internal control) – (sample B (eg mock): Ct gene of interest – Ct internal control)]. This formula was used to calculate fold-changes between the different samples (eg mock samples vs infected samples).

HSV-1 miRNAs miR-H2-3p, miR-H3-3p, miR-H4-5p, miR-H6-3p and miR-H7-5p, and human miRNA let-7a were quantified using pre-designed TaqMan stem-loop qRT-PCR miRNA Assays as per the manufacturer's instructions (Thermofisher). These were performed on a Quantstudio3 system (Applied biosystems, Thermofisher). The main principles of stem-loop qRT-PCR are outlined in figure 3.9.

6.2.6 Sequencing

All sequencing was performed by Dr Martin Fritzsche and Mr Ryan Mate, Analytical Biosciences (ABS), NIBSC.

6.2.6i Sequencing of lentivirus plasmids

NGS libraries were constructed using the DNA Prep kit (formerly known as Nextera DNA Flex) and dual-indexed using Nextera DNA CD indexes (both Illumina, USA).

Library preparation is based on tagmentation, a process where a bead-linked transposome complex fragments the DNA and simultaneously tags it with sequencing adapters.

These libraries were pooled in equimolar concentrations and sequenced with 250 bp paired-end reads on a MiSeq v2 kit (Illumina, USA) for 500 cycles.

6.2.6ii RNA sequencing (RNA-Seq)

Following RNA extraction (section 6.2.5.1), RNA-Seq libraries were constructed using the TruSeq Stranded mRNA kit (Illumina, 20020594) following manufacturer's instructions.

In brief, polyadenylated RNA molecules were captured, purified, and fragmented on oligo-dT covered magnetic beads. RNA was then converted into first strand cDNA using random primers followed by second strand synthesis and dUTP incorporation (for second strand amplification quenching as DNA polymerase is used during onwards PCR amplification and so will not incorporate past these dUTPs). Resulting blunt-end cDNA was 3' adenylated and indexing/sequencing adapters ligated via this adenine overhang. The cDNA fragments that have adapters ligated to both ends were enriched by PCR. The concentration and fragment size are measured for each indexed sample library and the molarity calculated, then each sample is diluted to a concentration of 10nM and all the samples are pooled. Finally, these pooled and normalised libraries were sequenced on the NextSeq 500 (Illumina) sequencing platform.

6.2.7 Computational analyses

6.2.7.1 Bioinformatic analysis

The bioinformatics analysis following RNA-seq, described below, was carried out by Dr Thomas Blaezard, ABS, NIBSC.

Following sequencing on the Illumina NextSeq 500 sequencer, the raw NextSeq output was processed with the Illumina tool bcl2fastq v2.17.1.14 software (© 2019 Illumina) to convert binary base call (BCL) files into FASTQ (text-based) files for onwards analysis. Quality control assessment was performed using fastqc v0.11.7 (Andrews, 2010) and multiqc v1 software (Ewels et al., 2016). Raw FASTQ data was trimmed for adapter sequences and quality using Cutadapt 1.16 (Martin, 2011). A phred score threshold of 30 was applied whereby the phred score estimates the probability of an error for each nucleobase identified by sequencing, where a score of 30 indicates that the probability of an error equals 1 in 1000 (Ewing & Green, 1998). Read pairs with a read shorter than 50 bp after trimming were discarded. The human reference genome hg38 (GenBank assembly accession: GCA_000001405.15) (Rosenbloom et al., 2015) and HSV-1 strain 17 reference genome NC_001806.2 (McGeoch, 1987) were combined and indexed as a reference genome. Forward and reverse read pairs for each sample were aligned to the reference using TopHat v2.1.1 software (Kim et al., 2013) using an inner distance of 50 bp, and using the transcriptome index as a guide. Picard Tools v2.18.9 (Institute,

2019) was used to process and index alignments. Read pairs uniquely aligned to the given strand in gene regions were counted using HTSeq-count v0.10.0. These counts were then processed using DESeq2 to assess differential expression. A Benjamini-Hochberg adjustment (Benjamini & Hochberg, 1995) was then applied in R software (The R Core Team, 2017) to control for a false discovery rate (FDR) of 5% or less. This analysis provided data tables of all the significantly differentially expressed genes with a 5% or lower FDR across all conditions. Heat maps (figure 4.4) were also generated from this data. Regularised log transformations were performed to normalise the expression profiles from the raw read counts and a distance matrix generated from these using the DESseq2 package in R. Heatmaps were then generated to compare samples using the R package gplots (Warnes, 2011).

The bioinformatics for the sequencing of the lentivirus plasmids was performed by Dr Mark Preston. In brief, the sequences produced with the MiSeq v2 kit (Illumina) were quality control assessed using fastqc v0.11.7 (Andrews, 2010). Low-quality bases were trimmed using trimmomatic software (Bolger et al., 2014). Once quality assessed and trimmed, the sequencing data was aligned against the reference sequence for each lentivirus plasmid (outlined in table 6.1.5), using bwa software (Li & Durbin, 2009). Next the variants (from regions with a minimum read depth of 100) were identified and this data was compiled using SAMtools software (Li et al., 2009). Finally graphs and tables outlining the SNPs found and the coverage and read-depth along each sequence were produced (figure 3.5) using R software (The R Core Team, 2017).

6.2.7.2 DAVID pathway analyses

DAVID software was accessed at https://david.ncifcrf.gov/tools.jsp (Huang et al., 2009a, 2009b). The gene list to be analysed was copied and pasted into the “paste a list" field, official gene symbol selected (maximum of 3000 genes can be entered), homo sapiens background genome selected and functional analysis tool run. For the gene ontology and KEGG pathway analyses GOTERM_BP_FAT under “Gene_Ontology” and KEGG_pathway under “Pathways” charts were selected (Huang et al., 2009a). For transcription factor analysis, the UCSC_TFBS chart was selected under “Protein_Interactions”. The charts were downloaded and imported into excel. In excel the lists were sorted from those with the most genes associated to least no. of genes and filtered for significance according a p-value < 0.05.

6.2.7.3 Imaging/counting software

(Fiji Is Just) ImageJ software (Schindelin et al., 2012) was used following immunofluorescence microscopy to make composite colour photomicrographs with scale bars and to count cells.

For the proportion of LAT-positive cells following RNAscope *in situ* hybridization (6.2.4.3) a mixture of automatic counting and manual counting were used to confirm automatic results.

For automatic counting, a macro was created in (Fiji Is Just) ImageJ to automatically count the SH-SY5Y cells (selected based on size and roundness) in DAPI, and then a percentage of positive cells based on wherever signal in the required channel overlapped with the DAPI cells.

For manual counting (for verifying automatic counting and where cytoplasmic localisation rendered the automatic counting ineffective), the cell counter plugin tool was utilised to manually click and count every cell in the DAPI field and then the same for the field with signal (FITC for LAT intron expression by RNAscope).

A percentage of positive cells were calculated and plotted for each field of view (and averaged across these) for each timepoint, as indicated in each graph (figures 2.6C and 3.8E).

6.2.8 Western blotting

Cells were lysed using RIPA buffer and spun at 4000rpm (on a VWR MEGA STAR 1.6 centrifuge with TX-400 swing-out rotor), 4^{o}C for 5mins. Following centrifugation, the cell pellet and supernatant are separated and kept on ice with pellet resuspended in PBS (both can be assayed for proteins).

A Pierce™ BCA Protein Assay was performed with Pierce™ BCA Protein Assay Kit and Pierce™ Bovine Serum Albumin (BSA) Standard Pre-Diluted Set as per the manufacturer's (Thermofisher) instructions to measure total protein quantity to determine the protein levels in the cell pellet and supernatant samples. In brief, the BSA pre-diluted standards and the protein samples were each mixed with the BCA Protein Assay Kit components on a 96-well plate, this combination led to chelation of copper with protein in an alkaline environment to form a light blue complex. This allows for a colorimetric readout of protein concentrations when the plate is read by a FLUOstar Omega luminometer. A standard curve can be created of the luminometer readouts of the known BSA standards. The concentration of protein in the samples can then be estimated based on their colorimetric readouts by comparison.

Known amounts of protein (at least 100µg) could then be run on a gel (BIO-RAD mini-PROTEAN® TGX™ Precast gels) in a 1:1 ratio with 2x Laemmli sample buffer and 10% 2-mercaptoethanol, after shaking and boiling at 1400rpm for 10 minutes on a Thermomixer comfort heated shaker (Eppendorf).

The gels were run alongside MagicMark™ XP Western Protein Standard (Invitrogen, Thermofisher) in a mini-PROTEAN® Tetra cell, at 200V for 35 minutes submerged in TGS buffer.

The gels were transferred onto membranes using a BIO-RAD Trans-Blot Turbo Transfer System and trans-blot transfer pack (Trans-Blot® Turbo™ Mini PVDF Transfer Packs).

The membranes were blocked in 5% Marvel dried skimmed milk in TBST (see table 6.1.3) on a shaker for 30mins at RT, washed in TBST for 5mins and then probed using the relevant antibody/antibodies at the appropriate dilution (table 6.1.2) in TBST, shaking at RT for 1 hour. The membrane was further washed in TBST (2x 5mins) before probing with the relevant secondary antibody (table 6.1.2) for 1 hour and washing in TBST again (2x 5mins). The membrane was rinsed with TBS (3x 5mins or until no bubbles appear). For re-probing of the same blot, the initial probed blot was washed in TBST overnight before starting the blocking and probing process again. The membrane was drained and ECL detection reagents (GE Lifescience - Amersham ECL Western Blotting Detection Reagent) were added to the blot for 10 mins before imaging on a BIO-RAD ChemiDoc™ MP imaging system. The BIO-RAD image lab 6.0 software was used for quantification of size and abundance of bands compared to known markers and other bands on the blots.

6.2.9 Northern blotting

HEK293T cells were transfected with 2-2.4µg of LAT intron lentivirus plasmid, NSLAT lentivirus plasmid or eGFP lentivirus plasmid per well on a 6-well plate (2 wells per lentivirus). 4 days post transfection total RNA was extracted from 2 wells combined per lentivirus and then 1µg per lentivirus was run on a formaldehyde gel (see section 6.2.3.8 for the making, running and imaging of the RNA on a formaldehyde gel).

A chemiluminescent northern blot was performed on this RNA according to the manufacturer's instructions (DIG Northern Starter Kit version 10 – Roche) with a probe directed to major LAT. Briefly, the RNA was transferred from the formaldehyde gel to a nylon Membrane through capillary action overnight. This was followed by UV cross-linking (twice at 1200J UV)(UV Stratalinker 2400, Stratagen), hybridisation and immunological detection of the RNA using a LAT-RNA specific probe, with washing steps

in between using 2 × SSC, 0.1% SDS and 0.1 × SSC, 0.1% SDS and northern washing buffer (table 6.1.3).

The LAT intron probe was generated following the manufacturer's instructions (Roche). Briefly, pSLAT4 was linearized by restriction digest with 5µg of pSLAT4 mixed with HindIII (20U) in buffer E (Promega) and dH_2O (2:1 ratio of dH_2O:plasmid/enzyme/buffer miz) and incubated at 37^OC for 1 hour). 1µg of linearized pSLAT4 was incubated (table 6.1.5) in dH_2O was incubated with DIG-labelling mix (Roche), transcription buffer (Roche) and T3/T7 RNA polymerase (roche) for 1h at 42^OC. This was then DNAse treated to remove template DNA by incubating with DNase I (Roche) for 15min at 37^OC. The reaction was stopped by the addition of 0.2M EDTA (Roche).

Following immunological detection, the blot was imaged on a SYNGENE G:BOX Chemi XX6 gel doc system alongside a fluorescent ruler. This image was overlapped with the original image with gel and fluorescent ruler to size the RNA bands found.

REFERENCES

ACE, C. I., DALRYMPLE, M. A., RAMSAY, F. H., PRESTON, V. G., & PRESTON, C. M. (1988). Mutational Analysis of the Herpes Simplex Virus Type 1 Giycoprotein E Promoter. *J. Gen. Virol.*, *196*(69), 2595–2605. https://doi.org/10.1006/viro.1993.1508

Aerts, E., Geets, E., Sorber, L., Beckers, S., Verrijken, A., Massa, G., … Van Hul, W. (2018). Evaluation of a Role for NPY and NPY2R in the Pathogenesis of Obesity by Mutation and Copy Number Variation Analysis in Obese Children and Adolescents. *Annals of Human Genetics*, *82*(1), 1–10. https://doi.org/10.1111/ahg.12211

Alandijany, T., Roberts, A. P. E., Conn, K. L., Loney, C., McFarlane, S., Orr, A., & Boutell, C. (2018). Distinct temporal roles for the promyelocytic leukaemia (PML) protein in the sequential regulation of intracellular host immunity to HSV-1 infection. *PLoS Pathogens*, *14*(1). https://doi.org/10.1371/journal.ppat.1006769

Alvira, M. R., Goins, W. F., Cohen, J. B., & Glorioso, J. C. (1999). Genetic studies exposing the splicing events involved in herpes simplex virus type 1 latency-associated transcript production during lytic and latent infection. *Journal of Virology*, *73*(5), 3866–3876.

Andrews, S. (2010). Babraham Bioinformatics - FastQC A Quality Control tool for High Throughput Sequence Data. Retrieved September 19, 2020, from http://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Arthur, J., Efstathiou, S., & Simmons, A. (1993). Intranuclear foci containing low abundance herpes simplex virus latency-associated transcripts visualized by non-isotopic in situ hybridization. In *Journal of General Virology* (Vol. 74).

Arthur, J L, Scarpini, C. G., Connor, V., Lachmann, R. H., Tolkovsky, A. M., & Efstathiou, S. (2001). Herpes simplex virus type 1 promoter activity during latency establishment, maintenance, and reactivation in primary dorsal root neurons in vitro. *Journal of Virology*, *75*(8), 3885–3895. https://doi.org/10.1128/JVI.75.8.3885-3895.2001

Arthur, Jane L., Everett, R., Brierley, I., & Efstathiou, S. (1998). Disruption of the 5’ and 3’ splice sites flanking the major latency-associated transcripts of herpes simplex virus type 1: Evidence for alternate splicing in lytic and latent infections. *Journal of General Virology*, *79*(1), 107–116. https://doi.org/10.1099/0022-1317-79-1-107

Atanasiu, D., Kent, J. R., Gartner, J. J., & Fraser, N. W. (2006). The stable 2-kb LAT intron of herpes simplex stimulates the expression of heat shock proteins and protects cells from stress. *Virology*, *350*(1), 26–33. https://doi.org/10.1016/j.virol.2006.02.001

Ball, M. J. (1982). “Limbic Predilection in Alzheimer Dementia: Is Reactivated Herpesvirus Involved?” *Canadian Journal of Neurological Sciences / Journal Canadien Des Sciences Neurologiques*, *9*(3), 303–306. https://doi.org/10.1017/S0317167100044115

Bansode, Y. D., Chattopadhyay, D., & Saha, B. (2019). Innate immune response in astrocytes infected with herpes simplex virus 1. *Archives of Virology*, *164*(5), 1433–1439. https://doi.org/10.1007/s00705-019-04197-x

Barrozo, E. R., Nakayamaa, S., Singhc, P., Neumannc, D. M., & Bloom, D. C. (2020). *Deletion of Herpes Simplex Virus 1 microRNAs miR-H1 and miR-H6 Impairs Reactivation.* (February). https://doi.org/10.1128/AAC.02511-19

Benjamini, Y., Heller, R., & Yekutieli, D. (2009). Selective inference in complex research. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *367*(1906), 4255–4271. https://doi.org/10.1098/rsta.2009.0127

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, *57*(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Bertke, A. S., Swanson, S. M., Chen, J., Imai, Y., Kinchington, P. R., & Margolis, T. P. (2011). A5-Positive Primary Sensory Neurons Are Nonpermissive for Productive Infection with Herpes Simplex Virus 1 In Vitro. *Journal of Virology*, *85*(13), 6669–6677. https://doi.org/10.1128/jvi.00204-11

Blo¨mer, U., Blo¨mer, B., Naldini, L., Kafri, T., Trono, D., Verma, I. M., & Gage, F. H. (1997). Highly Efficient and Sustained Gene Transfer in Adult Neurons with a Lentivirus Vector. In *JOURNAL OF VIROLOGY* (Vol. 71).

Block, T. M., Deshmane, S., Masonis, J., Maggioncalda, J., Valyi-Nagi, T., & Fraser, N. W. (1993). An HSV LAT Null Mutant Reactivates Slowly from Latent Infection and Makes Small Plaques on CV-1 Monolayers. *Virology*, *192*(2), 618–630. https://doi.org/10.1006/viro.1993.1078

Block, T. M., & Taylor, J. M. (1996). Evidence that Two Latency-Associated Transcripts of Herpes Simplex Virus Type 1 Are Nonlinear. In *JOURNAL OF VIROLOGY* (Vol. 70).

Bloom, D C, Devi-Rao, G. B., Hill, J. M., Stevens, J. G., & Wagner, E. K. (1994). Molecular analysis of herpes simplex virus type 1 during epinephrine-induced reactivation of latently infected rabbits in vivo. *Journal of Virology*, *68*(3), 1283–1292. https://doi.org/10.1128/jvi.68.3.1283-1292.1994

Bloom, David C. (2016). Chapter Two – Alphaherpesvirus Latency: A Dynamic State of Transcription and Reactivation. In *Advances in Virus Research* (Vol. 94, pp. 53–80). https://doi.org/10.1016/bs.aivir.2015.10.001

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Boutell, C., Sadis, S., & Everett, R. D. (2002). Herpes Simplex Virus Type 1 Immediate-Early Protein ICP0 and Its Isolated RING Finger Domain Act as Ubiquitin E3 Ligases In Vitro. *Journal of Virology*, *76*(2), 841–850. https://doi.org/10.1128/jvi.76.2.841-850.2002

Branco, F. J., & Fraser, N. W. (2005). Herpes simplex virus type 1 latency-associated transcript expression protects trigeminal ganglion neurons from apoptosis. *J Virol*, *79*(14), 9019–9025. https://doi.org/10.1128/JVI.79.14.9019

Brdovčak, M. C., Zubković, A., & Jurak, I. (2018). Herpes simplex virus 1 deregulation of host microRNAs. *Non-Coding RNA*, *4*(4). https://doi.org/10.3390/ncrna4040036

Brinkman, K. K., Mishra, P., & Fraser, N. W. (2013a). The half-life of the HSV-1 1.5-kb LAT intron is similar to the half-life of the 2.0-kb LAT intron. *Journal of NeuroVirology*, *19*(1), 102–108. https://doi.org/10.1007/s13365-012-0146-6

Brinkman, K. K., Mishra, P., & Fraser, N. W. (2013b). The half-life of the HSV-1 1.5-kb LAT intron is similar to the half-life of the 2.0-kb LAT intron. *Journal of NeuroVirology*, *19*(1), 102–108. https://doi.org/10.1007/s13365-012-0146-6

Browne, H., Bell, S., Minson, T., & Wilson, D. W. (1996). An endoplasmic reticulum-retained herpes simplex virus glycoprotein H is absent from secreted virions: evidence for reenvelopment during egress. *Journal of Virology*, *70*(7), 4311–4316. https://doi.org/10.1128/jvi.70.7.4311-4316.1996

Bryant, H., & Farrell, P. J. (2002). Signal Transduction and Transcription Factor Modification during Reactivation of Epstein-Barr Virus from Latency. *Journal of Virology*, *76*(20), 10290–10298. https://doi.org/10.1128/jvi.76.20.10290-10298.2002

Burgos, J. S., Ramirez, C., Sastre, I., Bullido, M. J., & Valdivieso, F. (2003). ApoE4 is more efficient than E3 in brain access by herpes simplex virus type 1. *Neuroreport*, *14*(14), 1825–1827. https://doi.org/10.1097/00001756-200310060-00013

Cabrera, J. R., Charron, A. J., & Leib, D. A. (2018). Neuronal subtype determines HSV-1 Latency-Associated-Transcript (LAT) promoter activity during latency. *Journal of Virology*, (April), JVI.00430-18. https://doi.org/10.1128/JVI.00430-18

Cáceres, M., Suwyn, C., Maddox, M., Thomas, J. W., & Preuss, T. M. (2007). Increased cortical expression of two synaptogenic thrombospondins in human brain evolution. *Cerebral Cortex*, *17*(10), 2312–2321. https://doi.org/10.1093/cercor/bhl140

Camarena, V., Kobayashi, M., Kim, J. Y., Roehm, P., Perez, R., Gardner, J., … Chao, M. V. (2010). Nature and duration of growth factor signaling through receptor tyrosine kinases regulates HSV-1 latency in neurons. *Cell Host and Microbe*, *8*(4), 320–330. https://doi.org/10.1016/j.chom.2010.09.007. Nature

Cantwell, C. A., Sterneck, E., & Johnson, P. F. (1998). Interleukin-6-Specific Activation of the C/EBPδ Gene in Hepatocytes Is Mediated by Stat3 and Sp1. *Molecular and Cellular Biology*, *18*(4), 2108–2117. https://doi.org/10.1128/mcb.18.4.2108

Carpenter, D., Henderson, G., Hsiang, C., Osorio, N., BenMohamed, L., Jones, C., & Wechsler, S. L. (2008). Introducing point mutations into the ATGs of the putative open reading frames of the HSV-1 gene encoding the latency associated transcript (LAT) reduces its anti-apoptosis activity. *Microbial Pathogenesis*, *44*(2), 98–102. https://doi.org/10.1016/j.micpath.2007.07.001

Carpenter, D., Hsiang, C., Jiang, X., Osorio, N., BenMohamed, L., Jones, C., & Wechsler, S. L. (2015). The herpes simplex virus type 1 (HSV-1) latency-associated transcript (LAT) protects cells against cold-shock-induced apoptosis by maintaining phosphorylation of protein kinase B (AKT). *Journal of NeuroVirology*, *21*(5), 568–575. https://doi.org/10.1007/s13365-015-0361-z

Catalanotto, C., Cogoni, C., & Zardo, G. (2016, October 13). MicroRNA in control of

gene expression: An overview of nuclear functions. *International Journal of Molecular Sciences*, Vol. 17. https://doi.org/10.3390/ijms17101712

Chen, C., Ridzon, D. A., Broomer, A. J., Zhou, Z., Lee, D. H., Nguyen, J. T., … Guegler, K. J. (2005). Real-time quantification of microRNAs by stem-loop RT-PCR. *Nucleic Acids Research*, *33*(20). https://doi.org/10.1093/nar/gni178

Chen, X. P., Mata, M., Kelley, M., Glorioso, J. C., & Fink, D. J. (2002). The relationship of herpes simplex virus latency associated transcript expression to genome copy number: A quantitative study using laser capture microdissection. *Journal of NeuroVirology*, *8*(3), 204–210. https://doi.org/10.1080/13550280290049642

Cheyette, B. N. R., Waxman, J. S., Miller, J. R., Takemaru, K. I., Sheldahl, L. C., Khlebtsova, N., … Moon, R. T. (2002). Dapper, a Dishevelled-associated antagonist of β-catenin and JNK signaling, is required for notochord formation. *Developmental Cell*, *2*(4), 449–461. https://doi.org/10.1016/S1534-5807(02)00140-5

Christensen, J., Steain, M., Slobedman, B., & Abendroth, A. (2011). Differentiated neuroblastoma cells provide a highly efficient model for studies of productive varicella-zoster virus infection of neuronal cells. *Journal of Virology*, *85*(16), 8436–8442. https://doi.org/10.1128/JVI.00515-11

Cleary, M. A., Stern, S., Tanaka, M., & Herr, W. (1993). Differential positive control by Oct-1 and Oct-2: Activation of a transcriptionally silent motif through Oct-1 and VP16 corecruitment. *Genes and Development*, *7*(1), 72–83. https://doi.org/10.1101/gad.7.1.72

Cliffe, A. R., Arbuckle, J. H., Vogel, J. L., Geden, M. J., Rothbart, S. B., Cusack, C. L., … Deshmukh, M. (2015). Neuronal Stress Pathway Mediating a Histone Methyl/Phospho Switch Is Required for Herpes Simplex Virus Reactivation. *Cell Host & Microbe*, *18*(6), 649–658. https://doi.org/10.1016/j.chom.2015.11.007

Cliffe, A. R., Garber, D. A., & Knipe, D. M. (2009). Transcription of the herpes simplex virus latency-associated transcript promotes the formation of facultative heterochromatin on lytic promoters. *Journal of Virology*, *83*(16), 8182–8190. https://doi.org/10.1128/JVI.00712-09

Coding, C. R., & O'Hare, P. (1989). Herpes simplex virus Vmw65-octamer binding protein interaction: A paradigm for combinatorial control of transcription. *Virology*, *173*(2), 363–367. https://doi.org/10.1016/0042-6822(89)90548-5

Cohen, C., Corpet, A., Roubille, S., Maroui, M. A., Poccardi, N., Rousseau, A., … Lomonte, P. (2018a). Promyelocytic leukemia (PML) nuclear bodies (NBs) induce latent/quiescent HSV-1 genomes chromatinization through a PML NB/Histone H3.3/H3.3 Chaperone Axis. In *PLoS Pathogens* (Vol. 14). https://doi.org/10.1371/journal.ppat.1007313

Cohen, C., Corpet, A., Roubille, S., Maroui, M. A., Poccardi, N., Rousseau, A., … Lomonte, P. (2018b). Promyelocytic leukemia (PML) nuclear bodies (NBs) induce latent/quiescent HSV-1 genomes chromatinization through a PML NB/Histone H3.3/H3.3 Chaperone Axis. *PLoS Pathogens*, *14*(9). https://doi.org/10.1371/journal.ppat.1007313

Coleman, H. M., Connor, V., Cheng, Z. S. C., Grey, F., Preston, C. M., & Efstathiou, S. (2008). Histone modifications associated with herpes simplex virus type 1 genomes during quiescence and following ICP0-mediated de-repression. *Journal of General Virology*, *89*(1), 68–77. https://doi.org/10.1099/vir.0.83272-0

Contrant, M., Fender, A., Chane-Woon-Ming, B., Randrianjafy, R., Vivet-Boudou, V., Richer, D., & Pfeffer, S. (2014). Importance of the RNA secondary structure for the relative accumulation of clustered viral microRNAs. *Nucleic Acids Research*, *42*(12), 7981–7996. https://doi.org/10.1093/nar/gku424

Copani, A., Caraci, F., Hoozemans, J. J. M., Calafiore, M., Angela Sortino, M., & Nicoletti, F. (2007, April 1). The nature of the cell cycle in neurons: Focus on a “non-canonical” pathway of DNA replication causally related to death. *Biochimica et Biophysica Acta - Molecular Basis of Disease*, Vol. 1772, pp. 409–412. https://doi.org/10.1016/j.bbadis.2006.10.016

Cordes, K. R., Sheehy, N. T., White, M. P., Berry, E. C., Morton, S. U., Muth, A. N., … Srivastava, D. (2009). MiR-145 and miR-143 regulate smooth muscle cell fate and plasticity. *Nature*, *460*(7256), 705–710. https://doi.org/10.1038/nature08195

Cubelos, B., Sebastián-Serrano, A., Beccari, L., Calcagnotto, M. E., Cisneros, E., Kim, S., … Nieto, M. (2010). Cux1 and Cux2 regulate dendritic branching, spine morphology, and synapses of the upper layer neurons of the cortex. *Neuron*, *66*(4), 523–535. https://doi.org/10.1016/j.neuron.2010.04.038

Cui, C., Griffiths, A., Li, G., Silva, L. M., Kramer, M. F., Gaasterland, T., … Coen, D. M. (2006). Prediction and identification of herpes simplex virus 1-encoded microRNAs. *Journal of Virology*, *80*(11), 5499–5508. https://doi.org/10.1128/JVI.00200-06

D'Aiuto, L., Bloom, D. C., Naciri, J. N., Smith, A., Edwards, T. G., Mcclain, L., … Nimgaonkar, L. (2019). *Modeling Herpes Simplex Virus 1 Infections in Human Central Nervous System Neuronal Cells Using Two- and Three- Dimensional Cultures Derived from Induced Pluripotent Stem Cells*. *93*(9), 1–23.

D'Aiuto, L., Zhi, Y., Kumar Das, D., Wilcox, M. R., Johnson, J. W., Mc Clain, L., … Nimgaonka, V. L. (2014). Large-scale generation of human ipsc-derived neural stem cells/early neural progenitor cells and their neuronal differentiation. *Organogenesis*, *10*(4), 365–377. https://doi.org/10.1080/15476278.2015.1011921

da Silva, L. F., & Jones, C. (2013). Small non-coding RNAs encoded within the herpes simplex virus type 1 latency associated transcript (LAT) cooperate with the retinoic acid inducible gene I (RIG-I) to induce beta-interferon promoter activity and promote cell survival. *Virus Research*, *175*(2), 101–109. https://doi.org/10.1016/j.virusres.2013.04.005

Danaher, R. J., Jacob, R. J., Steiner, M. R., Allen, W. R., Hill, J. M., & Miller, C. S. (2005). Histone deacetylase inhibitors induce reactivation of herpes simplex virus type 1 in a latency-associated transcript-independent manner in neuronal cells. *Journal of NeuroVirology*, *11*(3), 306–317. https://doi.org/10.1080/13550280590952817

Dasgupta, G., & BenMohamed, L. (2011). Of mice and not humans: How reliable are animal models for evaluation of herpes CD8 +-T cell-epitopes-based immunotherapeutic vaccine candidates? *Vaccine*, *29*(35), 5824–5836. https://doi.org/10.1016/j.vaccine.2011.06.083

Davison, A. J. (2002). Evolution of the herpesviruses. *Veterinary Microbiology*, *86*(1–2), 69–88. https://doi.org/10.1016/S0378-1135(01)00492-8

Davison, A. J. (2010). Herpesvirus systematics. *Veterinary Microbiology*, *143*(1), 52–69. https://doi.org/10.1016/j.vetmic.2010.02.014

Davison, A. J., Eberle, R., Ehlers, B., Hayward, G. S., McGeoch, D. J., Minson, A. C., … Thiry, E. (2009, January). The order Herpesvirales. *Archives of Virology*, Vol. 154, pp. 171–177. https://doi.org/10.1007/s00705-008-0278-4

De Chiara, G., Marcocci, M. E., Civitelli, L., Argnani, R., Piacentini, R., Ripoli, C., … Palamara, A. T. (2010). APP processing induced by herpes simplex virus type 1 (HSV-1) yields several APP fragments in human and rat neuronal cells. *PloS One*, *5*(11), e13989. https://doi.org/10.1371/journal.pone.0013989

De Chiara, G., Piacentini, R., Fabiani, M., Mastrodonato, A., Marcocci, M. E., Limongi, D., … Palamara, A. T. (2019). Recurrent herpes simplex virus-1 infection induces hallmarks of neurodegeneration and cognitive deficits in mice. *PLoS Pathogens*, *15*(3), 1–30. https://doi.org/10.1371/journal.ppat.1007617

De Regge, N., Van Opdenbosch, N., Nauwynck, H. J., Efstathiou, S., & Favoreel, H. W. (2010a). Interferon alpha induces establishment of alphaherpesvirus latency in sensory neurons in vitro. *PLoS ONE*, *5*(9). https://doi.org/10.1371/journal.pone.0013076

De Regge, N., Van Opdenbosch, N., Nauwynck, H. J., Efstathiou, S., & Favoreel, H. W. (2010b). Interferon alpha induces establishment of alphaherpesvirus latency in sensory neurons in vitro. *PLoS ONE*, *5*(9). https://doi.org/10.1371/journal.pone.0013076

DeLuca, N. A., & Schaffer, P. A. (1985). Activation of immediate-early, early, and late promoters by temperature-sensitive and wild-type forms of herpes simplex virus type 1 protein ICP4. *Molecular and Cellular Biology*, *5*(8), 1997–2008. https://doi.org/10.1128/mcb.5.8.1997

Depledge, D. P., Sadaoka, T., & Ouwendijk, W. J. D. (2018). Molecular aspects of varicella-zoster virus latency. *Viruses*, *10*(7), 1–21. https://doi.org/10.3390/v10070349

Dessimoz, C., & Škunca, N. (2017). The gene ontology handbook. In *springer protocols* (Vol. 1). https://doi.org/10.1038/s41559-017-0340-2

Dietrich, J. B. (2013, May 9). The MEF2 family and the brain: From molecules to memory. *Cell and Tissue Research*, Vol. 352, pp. 179–190. https://doi.org/10.1007/s00441-013-1565-2

Do¨hner, K., Wolfstein, A., Prank, U., Echeverri, C., Dujardin, D., Vallee, R., & Sodeik, B. (2002). Function of Dynein and Dynactin in Herpes Simplex Virus Capsid Transport. *Molecular Biology of the Cell*, *13*(6), 2795–2809. https://doi.org/10.1091/mbc.01

Donello, J. E., Loeb, J. E., & Hope, T. J. (1998). Woodchuck Hepatitis Virus Contains a Tripartite Posttranscriptional Regulatory Element. *Journal of Virology*, *72*(6), 5085–5092. https://doi.org/10.1128/jvi.72.6.5085-5092.1998

Douglas, L. N., McGuire, A. B., Manzardo, A. M., & Butler, M. G. (2016). High-resolution chromosome ideogram representation of recognized genes for bipolar

disorder. *Gene*, *586*(1), 136–147. https://doi.org/10.1016/j.gene.2016.04.011

Duan, F., Liao, J., Huang, Q., Nie, Y., & Wu, K. (2012). HSV-1 miR-H6 inhibits HSV-1 replication and IL-6 expression in human corneal epithelial cells in vitro. *Clinical and Developmental Immunology*, *2012*. https://doi.org/10.1155/2012/192791

Duan, Y., Zeng, J., Fan, S., Liao, Y., Feng, M., & Wang, L. (2019). Herpes Simplex Virus Type 1–Encoded miR-H2-3p Manipulates Cytosolic DNA – Stimulated Antiviral Innate Immune Response by Targeting DDX41. *Viruses*, *11*.

Duarte, L. F., Farías, M. A., Álvarez, D. M., Bueno, S. M., Riedel, C. A., & González, P. A. (2019, January 29). Herpes simplex virus type 1 infection of the central nervous system: Insights into proposed interrelationships with neurodegenerative disorders. *Frontiers in Cellular Neuroscience*, Vol. 13, pp. 1–23. https://doi.org/10.3389/fncel.2019.00046

Dull, T., Zufferey, R., Kelly, M., Mandel, R. J., Nguyen, M., Trono, D., & Naldini, L. (1998). A Third-Generation Lentivirus Vector with a Conditional Packaging System. *Journal of Virology*, *72*(11), 8463–8471. https://doi.org/10.1128/jvi.72.11.8463-8471.1998

Eastwood, S. L., Law, A. J., Everall, I. P., & Harrison, P. J. (2003). The axonal chemorepellant semaphorin 3A is increased in the cerebellum in schizophrenia and may contribute to its synaptic pathology. *Molecular Psychiatry*, *8*(2), 148–155. https://doi.org/10.1038/sj.mp.4001233

Edwards, T. G., & Bloom, D. C. (2019). Lund Human Mesencephalic (LUHMES) Neuronal Cell Line Supports HSV-1 Latency in vitro. *Journalfile:///C:/Users/Ajacobs/Downloads/JVI.02210-18. Bloom LUHMES Paper.Pdf of Virology*, (January). https://doi.org/10.1128/JVI.02210-18

Efstathiou, S., Minson, A. C., Field, H. J., Anderson, J. R., & Wildy, P. (1986). Detection of herpes simplex virus-specific DNA sequences in latently infected mice and in humans. *Journal of Virology*, *57*(2), 446–455. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3003377

Efstathiou, S., & Preston, C. M. (2005). Towards an understanding of the molecular basis of herpes simplex virus latency. *Virus Research*, *111*(2 SPEC. ISS.), 108–119. https://doi.org/10.1016/j.virusres.2005.04.017

Eisenberg, R. J., Atanasiu, D., Cairns, T. M., Gallagher, J. R., Krummenacher, C., & Cohen, G. H. (2012). Herpes virus fusion and entry: A story with many characters.

Viruses, *4*(5), 800–832. https://doi.org/10.3390/v4050800

Encinas, M., Iglesias, M., Liu, Y., Wang, H., Cefia, V., Gallego, C., & Comella, X. (2000). *Sequential Treatment of SH-SY5Y Cells with Retinoic Acid and Brain-Derived Neurotrophic Factor Gives Rise to fully Differentiated, Neurotrophic Factor-Dependent, Human Neuron-Like Cells*. https://doi.org/10.1046/j.1471-4159.2000.0750991.x

Enk, J., Levi, A., Weisblum, Y., Yamin, R., Charpak-Amikam, Y., Wolf, D. G., & Mandelboim, O. (2016). HSV1 MicroRNA Modulation of GPI Anchoring and Downstream Immune Evasion. *Cell Reports*, *17*(4), 949–956. https://doi.org/10.1016/j.celrep.2016.09.077

Everett, R., O'hare, P., O'rourke, D., Barlow, P., & Orr, A. (1995). Point Mutations in the Herpes Simplex Virus Type 1 Vmw110 RING Finger Helix Affect Activation of Gene Expression, Viral Growth, and Interaction with PML-Containing Nuclear Structures. In *JOURNAL OF VIROLOGY* (Vol. 69).

Ewels, P., Magnusson, M., Lundin, S., & Käller, M. (2016). MultiQC: Summarize analysis results for multiple tools and samples in a single report. *Bioinformatics*, *32*(19), 3047–3048. https://doi.org/10.1093/bioinformatics/btw354

Ewing, B., & Green, P. (1998). Phred_Error. *Genome Research*, (8), 1–9. Retrieved from sftp://cerca@192.168.2.5/home/cerca/Desktop/data/laptop_files/info/biologia/dna_sequencing/phred_error.pdf%5Cnpapers2://publication/uuid/3E49BD7A-1745-4E0D-9280-2EB0FE015C56%0Ahttp://bioinformatics.bc.edu/~marth/BI820-2004S/files/Ewing-Phred2-GR-1998.pdf

Farrell, M. J., Dobson, A. T., & Feldman, L. T. (1991). Herpes simplex virus latency-associated transcript is a stable intron. *Proceedings of the National Academy of Sciences of the United States of America*, *88*(3), 790–794. https://doi.org/10.1073/pnas.88.3.790

Fensterl, V., & Sen, G. C. (2015). Interferon-Induced Ifit Proteins: Their Role in Viral Pathogenesis. *Journal of Virology*, *89*(5), 2462–2468. https://doi.org/10.1128/jvi.02744-14

Ferrante, P., Mancuso, R., Pagani, E., Guerini, F. R., Calvo, M. G., Saresella, M., … Caputo, D. (2000). Molecular evidences for a role of HSV-1 in multiple sclerosis clinical acute attack. *Journal of NeuroVirology*, *6*(SUPPL. 2).

Frade, J. M., & Ovejero-Benito, M. C. (2015, March 13). Neuronal cell cycle: The neuron itself and its circumstances. *Cell Cycle*, Vol. 14, pp. 712–720. https://doi.org/10.1080/15384101.2015.1004937

Fu, Y., Hsiao, J. H. T., Paxinos, G., Halliday, G. M., & Kim, W. S. (2015). ABCA5 Regulates amyloid-β peptide production and is associated with Alzheimer's disease neuropathology. *Journal of Alzheimer's Disease*, *43*(3), 857–869. https://doi.org/10.3233/JAD-141320

Gao, G., Li, J. T., Kong, L., Tao, L., & Wei, L. (2009, May). Human herpesvirus miRNAs statistically preferentially target host genes involved in cell signaling and adhesion junction pathways. *Cell Research*, Vol. 19, pp. 665–667. https://doi.org/10.1038/cr.2009.45

Garber, D. A., Schaffer, P. A., & Knipe, D. M. (1997). A LAT-associated function reduces productive-cycle gene expression during acute infection of murine sensory neurons with herpes simplex virus type 1. *Journal of Virology*, *71*(8), 5885–5893. https://doi.org/10.1128/jvi.71.8.5885-5893.1997

GenScript. (n.d.). Sequence Scramble. Retrieved April 23, 2020, from https://www.genscript.com/tools/create-scrambled-sequence

Ginsberg, S. D., Mufson, E. J., Alldred, M. J., Counts, S. E., Wuu, J., Nixon, R. A., & Che, S. (2011). Upregulation of select rab GTPases in cholinergic basal forebrain neurons in mild cognitive impairment and Alzheimer's disease. *Journal of Chemical Neuroanatomy*, *42*(2), 102–110. https://doi.org/10.1016/j.jchemneu.2011.05.012

Good, P. F., Alapat, D., Hsu, A., Chu, C., Perl, D., Wen, X., … Kohtz, D. S. (2004). A role for semaphorin 3A signaling in the degeneration of hippocampal neurons during Alzheimer's disease. *Journal of Neurochemistry*, *91*(3), 716–736. https://doi.org/10.1111/j.1471-4159.2004.02766.x

Gosline, S. J. C., Gurtan, A. M., JnBaptiste, C. K., Bosson, A., Milani, P., Dalin, S., … Fraenkel, E. (2016). Elucidating MicroRNA Regulatory Networks Using Transcriptional, Post-transcriptional, and Histone Modification Measurements. *Cell Reports*, *14*(2), 310–319. https://doi.org/10.1016/j.celrep.2015.12.031

Grewal, S. I. S., & Moazed, D. (2003). *Heterochromatin and Epigenetic Control of Gene Expression*. *301*(August), 798–803.

Gu, H., & Zheng, Y. (2016). Role of ND10 nuclear bodies in the chromatin repression

of HSV-1. *Virology Journal*, *13*(1), 1–8. https://doi.org/10.1186/s12985-016-0516-4

Haery, L., Deverman, B. E., Matho, K. S., Cetin, A., Woodard, K., Cepko, C., … Fan, M. (2019, November 26). Adeno-Associated Virus Technologies and Methods for Targeted Neuronal Manipulation. *Frontiers in Neuroanatomy*, Vol. 13. https://doi.org/10.3389/fnana.2019.00093

Hafezi, W., Lorentzen, E. U., Eing, B. R., Müller, M., King, N. J. C., Klupp, B., … Kühn, J. E. (2012). Entry of herpes simplex virus type 1 (HSV-1) into the distal axons of trigeminal neurons favors the onset of nonproductive, silent infection. *PLoS Pathogens*, *8*(5). https://doi.org/10.1371/journal.ppat.1002679

Hammond, S. L., Leek, A. N., Richman, E. H., & Tjalkens, R. B. (2017). Cellular selectivity of AAV serotypes for gene delivery in neurons and astrocytes by neonatal intracerebroventricular injection. *PLOS ONE*, *12*(12), e0188830. https://doi.org/10.1371/journal.pone.0188830

Harkness, J. M., Kader, M., & DeLuca, N. A. (2014). Transcription of the herpes simplex virus 1 genome during productive and quiescent infection of neuronal and nonneuronal cells. *Journal of Virology*, *88*(12), 6847–6861. https://doi.org/10.1128/JVI.00516-14

Harris, R. A., & Preston, C. M. (1991). Establishment of latency in vitro by the herpes simplex virus type 1 mutant in1814. *Journal of General Virology*, *72*(4), 907–913. https://doi.org/10.1099/0022-1317-72-4-907

Harrison, K. S., Zhu, L., Thunuguntla, P., & Jones, C. (2020). Herpes simplex virus 1 regulates β-catenin expression in TG neurons during the latency-reactivation cycle. *Plos One*, *1*, 1–17. https://doi.org/10.1371/journal.pone.0230870

Hayashi, S., Inoue, Y., Kiyonari, H., Abe, T., Misaki, K., Moriguchi, H., … Takeichi, M. (2014). Protocadherin-17 Mediates Collective Axon Extension by Recruiting Actin Regulator Complexes to Interaxonal Contacts. *Developmental Cell*, *30*(6), 673–687. https://doi.org/10.1016/j.devcel.2014.07.015

Hayward, G. S., Jacob, R. J., Wadsworth, S. C., & Roizman, B. (1975). Anatomy of herpes simplex virus DNA: evidence for four populations of molecules that differ in the relative orientations of their long and short components. *Proceedings of the National Academy of Sciences of the United States of America*, *72*(11), 4243–4247. https://doi.org/10.1073/pnas.72.11.4243

He, Y., Wang, Z., Liu, C., Gong, Z., Li, Y., Lu, T., & Hu, G. (2019). Protocadherin 17 is a tumor suppressor and is frequently methylated in nasopharyngeal carcinoma. *Cancer Management and Research*, *11*, 1601–1613. https://doi.org/10.2147/CMAR.S191102

Hirabayashi, Y., Itoh, Y., Tabata, H., Nakajima, K., Akiyama, T., Masuyama, N., & Gotoh, Y. (2004). The Wnt/β-catenin pathway directs neuronal differentation of cortical neural precursor cells. *Development*, *131*(12), 2791–2801. https://doi.org/10.1242/dev.01165

Hoshina, N., Tanimura, A., Yamasaki, M., Inoue, T., Fukabori, R., Kuroda, T., … Yamamoto, T. (2013). Protocadherin 17 regulates presynaptic assembly in topographic corticobasal ganglia circuits. *Neuron*, *78*(5), 839–854. https://doi.org/10.1016/j.neuron.2013.03.031

Huang, D. W., Sherman, B. T., & Lempicki, R. A. (2009a). Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Retrieved April 21, 2020, from Nucleic acids research website: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2615629/

Huang, D. W., Sherman, B. T., & Lempicki, R. A. (2009b). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols*, *4*(1), 44–57. https://doi.org/10.1038/nprot.2008.211

Huntzinger, E., & Izaurralde, E. (2011, February 18). Gene silencing by microRNAs: Contributions of translational repression and mRNA decay. *Nature Reviews Genetics*, Vol. 12, pp. 99–110. https://doi.org/10.1038/nrg2936

Ibáñez, F. J., Farías, M. A., Gonzalez-Troncoso, M. P., Corrales, N., Duarte, L. F., Retamal-Díaz, A., & González, P. A. (2018, October 11). Experimental dissection of the lytic replication cycles of herpes simplex viruses in vitro. *Frontiers in Microbiology*, Vol. 9. https://doi.org/10.3389/fmicb.2018.02406

Illumina. (2019). RNA-Seq Technology | Comparison vs Microarrays. Retrieved April 23, 2020, from RNA-Seq Technology Overview website: https://www.illumina.com/science/technology/next-generation-sequencing/microarray-rna-seq-comparison.html

Institute, B. (2019). Picard Toolkit. Retrieved from http://broadinstitute.github.io/picard/

Izumi, K. M., McKelvey, A. M., Devi-Rao, G., Wagner, E. K., & Stevens, J. G. (1989). Molecular and biological characterization of a type 1 herpes simplex virus (HSV-1)

specifically deleted for expression of the latency-associated transcript (LAT). *Microbial Pathogenesis*, *7*(2), 121–134. https://doi.org/10.1016/0882-4010(89)90031-4

Jiang, X., Alami Chentoufi, A., Hsiang, C., Carpenter, D., Osorio, N., BenMohamed, L., … Wechsler, S. L. (2011). The Herpes Simplex Virus Type 1 Latency-Associated Transcript Can Protect Neuron-Derived C1300 and Neuro2A Cells from Granzyme B-Induced Apoptosis and CD8 T-Cell Killing. *Journal of Virology*, *85*(5), 2325–2332. https://doi.org/10.1128/JVI.01791-10

Jin, L., Peng, W., Perng, G. C., Brick, D. J., Nesburn, A. B., Jones, C., & Wechsler, S. L. (2003). Identification of herpes simplex virus type 1 latency-associated transcript sequences that both inhibit apoptosis and enhance the spontaneous reactivation phenotype. *J Virol*, *77*(11), 6556–6561. https://doi.org/10.1128/JVI.77.11.6556-6561.2003

Jurak, I., Silverstein, L. B., Sharma, M., & Coen, D. M. (2012). Herpes Simplex Virus Is Equipped with RNA- and Protein-Based Mechanisms To Repress Expression of ATRX, an Effector of Intrinsic Immunity. *Journal of Virology*, *86*(18), 10093–10102. https://doi.org/10.1128/JVI.00930-12

Jurak, Igor, Kramer, M. F., Mellor, J. C., van Lint, A. L., Roth, F. P., Knipe, D. M., & Coen, D. M. (2010). Numerous Conserved and Divergent MicroRNAs Expressed by Herpes Simplex Viruses 1 and 2. *Journal of Virology*, *84*(9), 4659–4672. https://doi.org/10.1128/jvi.02725-09

Kadakkuzha, B. M., & Puthanveettil, S. V. (2013). Genomics and proteomics in solving brain complexity. *Molecular BioSystems*, Vol. 9, pp. 1807–1821. https://doi.org/10.1039/c3mb25391k

Kanda, T. (2018). EBV-encoded latent genes. *Advances in Experimental Medicine and Biology*, *1045*, 377–394. https://doi.org/10.1007/978-981-10-7230-7_17

Kanehisa, M., & Goto, S. (2000). KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Research*, *28*(1), 27–30. https://doi.org/10.1093/nar/28.1.27

Kaplitt, M. G., Leone, P., Samulski, R. J., Xiao\, X., Pfaff, D. W., O'malleys, K. L., & During, M. J. (1994). *Long-term gene expression and phenotypic correction using adena-associated virus vectors in the mammalian brain*. Retrieved from http://www.nature.com/naturegenetics

Kennedy, P. G. E., Rovnak, J., Badani, H., & Cohrs, R. J. (2015). A comparison of

herpes simplex virus type 1 and varicella-zoster virus latency and reactivation. *Journal of General Virology*, *96*(7), 1581–1602. https://doi.org/10.1099/vir.0.000128

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., & Salzberg, S. L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions Daehwan. *Genome Biology*, *14*(R36).

Kim, J. Y., Mandarino, A., Chao, M. V., Mohr, I., & Wilson, A. C. (2012). Transient reversal of episome silencing precedes VP16-dependent transcription during reactivation of latent HSV-1 in neurons. *PLoS Pathogens*, *8*(2). https://doi.org/10.1371/journal.ppat.1002540

King, D. H., & Madera, P. D. C. (1982). *History, pharmacokinetics, of acyclovir and pharmacology*. 176–179.

Knipe, D. M., & Cliffe, A. (2008). Chromatin control of herpes simplex virus lytic and latent infection. *Nature Reviews Microbiology*, *6*(3), 211–221. https://doi.org/10.1038/nrmicro1794

Kobayashi, M., Kim, J. Y., Camarena, V., Roehm, P. C., Chao, M. V., Wilson, A. C., & Mohr, I. (2012). A primary neuron culture system for the study of Herpes simplex virus latency and reactivation. *Journal of Visualized Experiments*, (62), 3823. https://doi.org/10.3791/3823

Kobayashi, M., Wilson, A. C., Chao, M. V., & Mohr, I. (2012). Control of viral latency in neurons by axonal mTOR signaling and the 4E-BP translation repressor. *Genes and Development*, *26*(14), 1527–1532. https://doi.org/10.1101/gad.190157.112

Kokotos, A. C., Harper, C. B., Marland, J. R. K., Smillie, K. J., Cousin, M. A., & Gordon, S. L. (2019). Synaptophysin sustains presynaptic performance by preserving vesicular synaptobrevin-II levels. *Journal of Neurochemistry*, *151*(1), 28–37. https://doi.org/10.1111/jnc.14797

Kollias, C. M., Huneke, R. B., Wigdahl, B., & Jennings, S. R. (2015). Animal models of herpes simplex virus immunity and pathogenesis. *Journal of NeuroVirology*, *21*(1), 8–23. https://doi.org/10.1007/s13365-014-0302-2

Koros, C., Ioannidis, A., Acquaviva, T., Zoga, M., Nikolaou, C., Chatzipanagiotou, S., … Anagnostouli, M. C. (2014). HSV1 and 2 detection in the CSF of multiple sclerosis patients by real-time PCR. *In Vivo*, *28*(6), 1201–1206.

Kovalevich, J., & Langford, D. (2013). Considerations for the use of SH-SY5Y neuroblastoma cells in neurobiology. *Methods in Molecular Biology*. https://doi.org/10.1007/978-1-62703-640-5_2

Koyuncu, O. O., Macgibeny, M. A., & Enquist, L. W. (2018). Latent versus productive infection: The alpha herpesvirus switch. *Future Virology*, *13*(6), 431–443. https://doi.org/10.2217/fvl-2018-0023

Koyuncu, O. O., Song, R., Greco, T. M., Cristea, I. M., & Enquist, L. W. (2015). The number of alphaherpesvirus particles infecting axons and the axonal protein repertoire determines the outcome of neuronal infection. *MBio*, *6*(2), 1–12. https://doi.org/10.1128/mBio.00276-15

Krüger, J., & Rehmsmeier, M. (2006). RNAhybrid: MicroRNA target prediction easy, fast and flexible. https://doi.org/10.1093/nar/gkl243

Kukhanova, M. K., Korovina, A. N., & Kochetkov, S. N. (2014). Human herpes simplex virus: Life cycle and development of inhibitors. *Biochemistry (Moscow)*, *79*(13), 1635–1652. https://doi.org/10.1134/S0006297914130124

Kumar, D., Bansal, G., Narang, A., Basak, T., Abbas, T., & Dash, D. (2016). Integrating transcriptome and proteome profiling: Strategies and applications. *PROTEOMICS*, *16*(19), 2533–2544. https://doi.org/10.1002/pmic.201600140

Kumar, M., Keller, B., Makalou, N., & Sutton, R. E. (2001). Systematic determination of the packaging limit of lentiviral vectors. *Human Gene Therapy*, *12*(15), 1893–1905. https://doi.org/10.1089/104303401753153947

Laemmle, L., Goldstein, R. S., & Kinchington, P. R. (2019). Modeling Varicella Zoster Virus Persistence and Reactivation – Closer to Resolving a Perplexing Persistent State. *Frontiers in Microbiology*, *10*(July), 1–15. https://doi.org/10.3389/fmicb.2019.01634

Lai, J. S., & Herr, W. (1997). Interdigitated residues within a small region of VP16 interact with Oct-1, HCF, and DNA. *Molecular and Cellular Biology*, *17*(7), 3937–3946. https://doi.org/10.1128/mcb.17.7.3937

Lakin, N. D., Palmer, R., Lillycrop, K. A., Howard, M. K., Burke, L. C., Thomas, N. S. B., & Latchman, D. S. (1995). Down regulation of the octamer binding protein Oct-1 during growth arrest and differentiation of a neuronal cell line. *Molecular Brain Research*, *28*(1), 47–54. https://doi.org/10.1016/0169-328X(94)00183-F

Laplante, M., & Sabatini, D. M. (2009). mTOR signaling at a glance. *Journal of Cell Science*, *122*(20), 3589–3594. https://doi.org/10.1242/jcs.051011

Latchman, D. S. (2001, April 1). Transcription factors: Bound to activate or repress. *Trends in Biochemical Sciences*, Vol. 26, pp. 211–213. https://doi.org/10.1016/S0968-0004(01)01812-6

Laurén, J., Airaksinen, M. S., Saarma, M., & Timmusk, T. (2003). A novel gene family encoding leucine-rich repeat transmembrane proteins differentially expressed in the nervous system. *Genomics*, *81*(4), 411–421. https://doi.org/10.1016/S0888-7543(03)00030-2

Le-Niculescu, H., Patel, S. D., Bhat, M., Kuczenski, R., Faraone, S. V., Tsuang, M. T., … Niculescu, A. B. (2009). Convergent functional genomics of genome-wide association data for bipolar disorder: Comprehensive identification of candidate genes, pathways and mechanisms. *American Journal of Medical Genetics, Part B: Neuropsychiatric Genetics*, *150*(2), 155–181. https://doi.org/10.1002/ajmg.b.30887

LEE, H., PARK, C., CHO, I.-H., KIM, H. Y., JO, E.-K., LEE, S., … LEE, S. J. (2007). Double-Stranded RNA Induces iNOS Gene Expression in Schwann Cells, Sensory Neuronal Death, and Peripheral Nerve Demyelination. *GLIA*, *55*, 712–722. https://doi.org/10.1002/glia

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, *25*(14), 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., … Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, *25*(16), 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Limviphuvadh, V., Ling Chua, L., Eisenhaber, F., Adhikari, S., & Maurer-stroh, S. (2010). IS LGI2 THE CANDIDATE GENE FOR PARTIAL EPILEPSY WITH PERICENTRAL SPIKES? *Journal of Bioinformatics and Computational Biology*, *6*(1), 117–127. https://doi.org/10.1142/S0219720010004550

Linard, M., Letenneur, L., Garrigue, I., Doize, A., Dartigues, J.-F., & Helmer, C. (2020). Interaction between *APOE4* and herpes simplex virus type 1 in Alzheimer's disease. *Alzheimer's & Dementia*, *16*(1), 200–208. https://doi.org/10.1002/alz.12008

Linderman, J. A., Kobayashi, M., Rayannavar, V., Fak, J. J., Darnell, R. B., Chao, M. V, … Mohr, I. (2017). Immune Escape via a Transient Gene Expression Program Enables Productive Replication of a Latent Pathogen. *Cell Reports*, *18*(5), 1312–1323. https://doi.org/10.1016/j.celrep.2017.01.017

Liu, D. X., & Greene, L. A. (2001). Neuronal apoptosis at the G1/S cell cycle checkpoint. *Cell and Tissue Research*, *305*(2), 217–228. https://doi.org/10.1007/s004410100396

Liu, S., Liu, P., Borras, A., Chatila, T., & Speck, S. H. (1997). *Cyclosporin A-sensitive induction of the Epstein – Barr virus lytic switch is mediated via a novel pathway involving a MEF2 family member*. *16*(1), 143–153.

Liu, Y., Hancock, M., Workman, A., Doster, A., & Jones, C. (2016). β-Catenin, a Transcription Factor Activated by Canonical Wnt Signaling, Is Expressed in Sensory Neurons of Calves Latently Infected with Bovine Herpesvirus 1. *Journal of Virology*, *90*(6), 3148–3159. https://doi.org/10.1128/jvi.02971-15

Lokensgard, J. R., Hu, S., Sheng, W., vanOijen, M., Cox, D., Cheeran, M. C. J., & Peterson, P. K. (2001). Robust expression of TNF-α, IL-1β, RANTES, and IP-10 by human microglial cells during nonproductive infection with herpes simplex virus. *Journal of NeuroVirology*, *7*(3), 208–219. https://doi.org/10.1080/13550280152403254

Lomonte, P., Thomas, J., Texier, P., Caron, C., Khochbin, S., & Epstein, A. L. (2004). Functional Interaction between Class II Histone Deacetylases and ICP0 of Herpes Simplex Virus Type 1. *Journal of Virology*, *78*(13), 6744–6757. https://doi.org/10.1128/jvi.78.13.6744-6757.2004

Looker, K. J., Magaret, A. S., May, M. T., Turner, K. M. E., Vickerman, P., Gottlieb, S. L., & Newman, L. M. (2015). Global and regional estimates of prevalent and incident herpes simplex virus type 1 infections in 2012. *PLoS ONE*, *10*(10), 1–17. https://doi.org/10.1371/journal.pone.0140765

Lövheim, H., Gilthorpe, J., Adolfsson, R., Nilsson, L. G., & Elgh, F. (2015). Reactivated herpes simplex infection increases the risk of Alzheimer's disease. *Alzheimer's and Dementia*, *11*(6), 593–599. https://doi.org/10.1016/j.jalz.2014.04.522

Lövheim, H., Gilthorpe, J., Johansson, A., Eriksson, S., Hallmans, G., & Elgh, F. (2015). Herpes simplex infection and the risk of Alzheimer's disease: A nested case-control study. *Alzheimer's and Dementia*, *11*(6), 587–592.

https://doi.org/10.1016/j.jalz.2014.07.157

MacDonald, B. T., Tamai, K., & He, X. (2009, July 21). Wnt/β-Catenin Signaling: Components, Mechanisms, and Diseases. *Developmental Cell*, Vol. 17, pp. 9–26. https://doi.org/10.1016/j.devcel.2009.06.016

Mador, N., Goldenberg, D., Cohen, O., Panet, A., & Steiner, I. (1998). Herpes simplex virus type 1 latency-associated transcripts suppress viral replication and reduce immediate-early gene mRNA levels in a neuronal cell line. *Journal of Virology*, *72*(6), 5067–5075. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9573277

Maier, T., Güell, M., & Serrano, L. (2009). Correlation of mRNA and protein in complex biological samples. *FEBS Letters*, *583*(24), 3966–3973. https://doi.org/10.1016/j.febslet.2009.10.036

Mangold, C. A., & Szpara, M. L. (2019). Persistent Infection with Herpes Simplex Virus 1 and Alzheimer's Disease—A Call to Study How Variability in Both Virus and Host may Impact Disease. *Viruses*, *11*(10), 966. https://doi.org/10.3390/v11100966

Margolis, T. P., Sedarati, F., Dobson, A. T., Feldman, L. T., & Stevens', J. G. (1992). Pathways of Viral Gene Expression during Acute Neuronal Infection with HSV-1. In *VIROLOGY* (Vol. 189).

Margolis, T. P., Sedarati, F., Dobson, A. T., Feldman, L. T., & Stevens, J. G. (1992). Pathways of viral gene expression during acute neuronal infection with HSV-1. *Virology*, *189*(1), 150–160. https://doi.org/10.1016/0042-6822(92)90690-Q

Marquitz, A. R., Mathur, A., Edwards, R. H., & Raab-Traub, N. (2015). Host Gene Expression Is Regulated by Two Types of Noncoding RNAs Transcribed from the Epstein-Barr Virus BamHI A Rightward Transcript Region. *Journal of Virology*, *89*(22), 11256–11268. https://doi.org/10.1128/JVI.01492-15

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EBMnet*, *17*(1), 10–12.

MARY JO GROVES, MD. (2016). Genital herpes: a review. *American Family Physician*, *72*(8), 1527–1534. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/16273819

McCue, H. V., Patel, P., Herbert, A. P., Lian, L. Y., Burgoyne, R. D., & Haynes, L. P. (2012). Solution NMR structure of the Ca2+-bound N-terminal domain of CaBP7:

A regulator of golgi trafficking. *Journal of Biological Chemistry*, *287*(45), 38231–38243. https://doi.org/10.1074/jbc.M112.402289

McGeoch, D. J., & Cook, S. (1994). Molecular phylogeny of the alphaher- pesvirinae subfamily and a proposed evolutionary timescale. *J. Mol. Biol.*, (238), 9–22. Retrieved from http://ac.els-cdn.com/S0022283684712642/1-s2.0-S0022283684712642-main.pdf?_tid=976f791a-300b-11e7-8f4e-00000aacb361&acdnat=1493821380_cb1fd6233dcf307c8002c55a5a8641a2

McGeoch, D. J., Dalrymple, M. A., Davison, A. J., Dolan, A., Frame, M. C., McNab, D., … Taylor, P. (1988). The complete DNA sequence of the long unique region in the genome of herpes simplx virus type 1. *Journal of General Virology*, *69*(7), 1531–1574. https://doi.org/10.1099/0022-1317-69-7-1531

McGeoch, Duncan J. (1987). *THE GENOME OF HERPES SIMPLEX VIRUS: STRUCTURE, REPLICATION AND EVOLUTION*. *4*.

McGeoch, Duncan J., Cook, S., Dolan, A., Jamieson, F. E., & Telford, E. A. R. (1995). Molecular phylogeny and evolutionary timescale for the family of mammalian herpesviruses. *Journal of Molecular Biology*, *247*(3), 443–458. https://doi.org/10.1006/jmbi.1995.0152

McLennan, J. L., & Darby, G. (1980). Herpes simplex virus latency: The cellular location of virus in dorsal root ganglia and the fate of the infected cell following virus activation. *Journal of General Virology*, *51*(2), 233–243. https://doi.org/10.1099/0022-1317-51-2-233

Mehta, A., Maggioncalda, J., Bagasra, O., Thikkavarapu, S., Saikumari, P., Valyi-Nagy, T., … Block, T. M. (1995). In situ DNA PCR and RNA hybridization detection of herpes simplex virus sequences in trigeminal ganglia of latently infected mice. *Virology*, *206*(1), 633–640. https://doi.org/10.1016/S0042-6822(95)80080-8

Milne, R. S. B., Nicola, A. V, Whitbeck, J. C., Eisenberg, R. J., & Cohen, G. H. (2005). Glycoprotein D receptor-dependent, low-pH-independent endocytic entry of herpes simplex virus type 1. *Journal of Virology*, *79*(11), 6655–6663. https://doi.org/10.1128/JVI.79.11.6655-6663.2005

Morgan, K. (2014). Addgene. Retrieved August 19, 2020, from https://blog.addgene.org/your-lentiviral-plasmid-faqs-answered

Morgun, A., Dzutsev, A., Dong, X., Greer, R., Sexton, D., Ravel, J., … Shulzhenko, N. (2016). Using Transkingdom Gene Networks. *Gut*, *64*(11), 1732–1743.

https://doi.org/10.1136/gutjnl-2014-308820.Uncovering

Murata, T., & Tsurumi, T. (2009). Switching of EBV cycles between latent andlytic states. *Reviews in Medical Virology*, *19*(1), 57–64. https://doi.org/10.1002/rmv

Najafi, S., Ghane, M., Poortahmasebi, V., Jazayeri, S. M., & Yousefzadeh-Chabok, S. (2016). Prevalence of herpes simplex virus infection in patients with relapsing-remitting multiple sclerosis: A case-control study in the North of Iran. *Archives of Clinical Infectious Diseases*, *11*(3), 1–6. https://doi.org/10.5812/archcid.36576

Naqvi, A. R., Seal, A., Shango, J., Navarette, M. B., Martinez, G., Chapa, G., … Nares, S. (2018). Herpesvirus-encoded microRNAs detected in human gingiva alter host cell transcriptome and regulate viral infection. *Biochim Biophys Acta.*, *1861*(5), 497–508. https://doi.org/10.1016/j.physbeh.2017.03.040

Naso, M. F., Tomkowicz, B., Perry, W. L., & Strohl, W. R. (2017, August 1). Adeno-Associated Virus (AAV) as a Vector for Gene Therapy. *BioDrugs*, Vol. 31, pp. 317–334. https://doi.org/10.1007/s40259-017-0234-5

NCBI. (n.d.). Human herpesvirus 1 strain 17, complete genome. Retrieved from FASTA website: https://www.ncbi.nlm.nih.gov/nuccore/NC_001806.2

NCBI. (2016). Home - Gene - NCBI. Retrieved April 23, 2020, from https://www.ncbi.nlm.nih.gov/gene/

Nesburn, A. B., Dickinson, R., Radnoti, M., & Green, M. J. (1976). *in Rabbits*. *21*(October), 185–190.

Neumann, D. M., Bhattacharjee, P. S., Giordani, N. V., Bloom, D. C., & Hill, J. M. (2007). In Vivo Changes in the Patterns of Chromatin Structure Associated with the Latent Herpes Simplex Virus Type 1 Genome in Mouse Trigeminal Ganglia Can Be Detected at Early Times after Butyrate Treatment. *Journal of Virology*, *81*(23), 13248–13253. https://doi.org/10.1128/jvi.01569-07

Nicoll, M. P., Proença, J. T., Connor, V., & Efstathiou, S. (2012). Influence of Herpes Simplex Virus 1 Latency-Associated Transcripts on the Establishment and Maintenance of Latency in the ROSA26R Reporter Mouse Model. *Journal of Virology*, *86*(16), 8848–8858. https://doi.org/10.1128/JVI.00652-12

Nicoll, Michael P., Hann, W., Shivkumar, M., Harman, L. E. R., Connor, V., Coleman, H. M., … Efstathiou, S. (2016). The HSV-1 Latency-Associated Transcript Functions to Repress Latent Phase Lytic Gene Expression and Suppress Virus

Reactivation from Latently Infected Neurons. *PLoS Pathogens*, *12*(4). https://doi.org/10.1371/journal.ppat.1005539

Nicoll, Michael P., Proença, J. T., & Efstathiou, S. (2012). The molecular basis of herpes simplex virus latency. *FEMS Microbiology Reviews*, *36*(3), 684–705. https://doi.org/10.1111/j.1574-6976.2011.00320.x

Nishiyama, Y. (1996). Herpesvirus genes: molecular basis of viral replication and pathogenicity. *Nagoya Journal of Medical Science*, *59*(3–4), 107–119. https://doi.org/10.18999/nagjms.59.3-4.107

O'Brien, J., Hayder, H., Zayed, Y., & Peng, C. (2018, August 3). Overview of microRNA biogenesis, mechanisms of actions, and circulation. *Frontiers in Endocrinology*, Vol. 9, p. 402. https://doi.org/10.3389/fendo.2018.00402

O'Keefe, E. P. (2013). Nucleic Acid Delivery: Lentiviral and Retroviral Vectors. *Materials and Methods*, *3*. https://doi.org/10.13070/mm.en.3.174

Oliveros, J. C. (2015). Venny 2.1.0. Retrieved April 23, 2020, from An interactive tool for comparing lists with Venn's diagrams. website: http://bioinfogp.cnb.csic.es/tools/venny/

Orang, A. V., Safaralizadeh, R., & Kazemzadeh-Bavili, M. (2014). Mechanisms of miRNA-Mediated Gene Regulation from Common Downregulation to mRNA-Specific Upregulation. *International Journal of Genomics*, *2014*. https://doi.org/10.1155/2014/970607

Park, M. H., Jang, J. H., Song, J. J., Lee, H. S., & Oh, S. H. (2016). Neurofilament heavy chain expression and neuroplasticity in rat auditory cortex after unilateral and bilateral deafness. *Hearing Research*, *339*, 155–160. https://doi.org/10.1016/j.heares.2016.07.010

Peng, W, Jin, L., Henderson, G., Perng, G. C., Brick, D. J., Nesburn, A. B., … Jones, C. (2004). Mapping herpes simplex virus type 1 latency-associated transcript sequences that protect from apoptosis mediated by a plasmid expressing caspase-8. *Journal of Neurovirology*, *10*(4), 260–265. https://doi.org/10.1080/13550280490468690

Peng, Weiping, Vitvitskaia, O., Carpenter, D., Wechsler, S. L., & Jones, C. (2008). Identification of two small RNAs within the first 1.5-kb of the herpes simplex virus type 1-encoded latency-associated transcript. *Journal of NeuroVirology*, *14*(1), 41–52. https://doi.org/10.1080/13550280701793957

Perng, G.-C., Jones, C., Ciacci-Zanella, J., Stone, M., Henderson, G., Yukht, A., … Wechsler, S. L. (2000). Virus-Induced Neuronal Apoptosis Blocked by the Herpes Simplex Virus Latency-Associated Transcript. *Science*, *287*(5457). Retrieved from http://science.sciencemag.org/content/287/5457/1500.full

Perng, G. C., Dunkel, E. C., Geary, P. A., Slanina, S. M., Ghiasi, H., Kaiwar, R., … Wechsler, S. L. (1994). The latency-associated transcript gene of herpes simplex virus type 1 (HSV-1) is required for efficient in vivo spontaneous reactivation of HSV-1 from latency. *Journal of Virology*, *68*(12), 8045–8055. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7966594

Perng, G. C., Ghiasi, H., Slanina, S. M., Nesburn, A. B., & Wechsler, S. L. (1996). The spontaneous reactivation function of the herpes simplex virus type 1 LAT gene resides completely within the first 1.5 kilobases of the 8.3-kilobase primary transcript. *Journal of Virology*, *70*(2), 976–984. https://doi.org/10.1128/jvi.70.2.976-984.1996

Perng, G. C., Slanina, S. M., Yukht, A., Ghiasi, H., Nesburn, A. B., & Wechsler, S. L. (2000). The latency-associated transcript gene enhances establishment of herpes simplex virus type 1 latency in rabbits. *Journal of Virology*, *74*(4), 1885–1891. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10644361

Perusina Lanfranca, M., Mostafa, H., & Davido, D. (2014). HSV-1 ICP0: An E3 Ubiquitin Ligase That Counteracts Host Intrinsic and Innate Immunity. *Cells*, *3*(2), 438–454. https://doi.org/10.3390/cells3020438

Phelan, D., Barrozo, E. R., & Bloom, D. C. (2017). HSV1 latent transcription and non-coding RNA: A critical retrospective. *Journal of Neuroimmunology*. https://doi.org/10.1016/j.jneuroim.2017.03.002

Piacentini, R., Chiara, G. De, Domenica, D. L. P., Ripoli, C., Marcocci, M. E., Garaci, E., … Grassi, C. (2014). HSV-1 and Alzheimer's disease: More than a hypothesis. *Frontiers in Pharmacology*. https://doi.org/10.3389/fphar.2014.00097

Pourchet, A., Modrek, A., Placantonakis, D., Mohr, I., & Wilson, A. (2017a). Modeling HSV-1 Latency in Human Embryonic Stem Cell-Derived Neurons. *Pathogens*, *6*(2), 24. https://doi.org/10.3390/pathogens6020024

Pourchet, A., Modrek, A. S., Placantonakis, D. G., Mohr, I., & Wilson, A. C. (2017b). Modeling HSV-1 latency in human embryonic stem cell-derived neurons. *Pathogens*, *6*(2). https://doi.org/10.3390/pathogens6020024

Preston, C M, Mabbs, R., & Nicholl, M. J. (1997). Construction and characterization of herpes simplex virus type 1 mutants with conditional defects in immediate early gene expression. *Virology*, *229*(1), 228–239. https://doi.org/10.1006/viro.1996.8424

Preston, Chris M., Mabbs, R., & Nicholl, M. J. (1997). Construction and Characterization of Herpes Simplex Virus Type 1 Mutants with Conditional Defects in Immediate Early Gene Expression. *Virology*, *229*(1), 228–239. https://doi.org/10.1006/viro.1996.8424

Preston, Chris M., & McFarlane, M. (1998). Cytodifferentiating Agents Affect the Replication of Herpes Simplex Virus Type 1 in the Absence of Functional VP16. *Virology*, *249*(2), 418–426. https://doi.org/10.1006/viro.1998.9314

Proença, J. T., Coleman, H. M., Connor, V., Winton, D. J., & Efstathiou, S. (2008). A historical analysis of herpes simplex virus promoter activation in vivo reveals distinct populations of latently infected neurones. *Journal of General Virology*, *89*(12), 2965–2974. https://doi.org/10.1099/vir.0.2008/005066-0

Rajč, L. A., & Uchova, R. (2004). Peculiarities of Herpes Simplex Virus (HSV) Transcription: An overview. In *Virus Genes* (Vol. 28).

Ramkumar, A., Jong, B. Y., & Ori-McKenney, K. M. (2018, January 1). ReMAPping the microtubule landscape: How phosphorylation dictates the activities of microtubule-associated proteins. *Developmental Dynamics*, Vol. 247, pp. 138–155. https://doi.org/10.1002/dvdy.24599

Rechenchoski, D. Z., Faccin-Galhardi, L. C., Linhares, R. E. C., & Nozawa, C. (2017, March 1). Herpesvirus: an underestimated virus. *Folia Microbiologica*, Vol. 62, pp. 151–156. https://doi.org/10.1007/s12223-016-0482-7

Rehmsmeier, M., Steffen, P., Höchsmann, M., & Giegerich, R. (2004). Fast and effective prediction of microRNA/target duplexes. *RNA*, *10*(10), 1507–1517. https://doi.org/10.1261/rna.5248604

Rock, D., & Fraser, N. (1983). Detection of HSV-1 genome in central nervous system of latently infected mice. *Nature*, *302*(5908), 523–525.

Rock, D. L., Nesburn, A. B., Ghiasi, H., Ong, J., Lewis, T. L., Lokensgard, J. R., & Wechsler, S. L. (1987). Detection of latency-related viral RNAs in trigeminal ganglia of rabbits latently infected with herpes simplex virus type 1. *Journal of Virology*, *61*(12), 3820–3826. https://doi.org/10.1128/jvi.61.12.3820-3826.1987

Rødahl, E., & Haarr, L. (1997). Analysis of the 2-Kilobase Latency-Associated Transcript Expressed in PC12 Cells Productively Infected with Herpes Simplex Virus Type 1: Evidence for a Stable, Nonlinear Structure. In *JOURNAL OF VIROLOGY* (Vol. 71).

Roizman, B., & Sears, A. (1987). An Inquiry Into The Mechanisms of Herpes simplex virus latency. *Expert Reviews in Molecular Medicine*, *5*(29). https://doi.org/10.1017/S1462399403006975

Roizman, B., & Whitley, R. J. (2013). An inquiry into the molecular basis of HSV latency and reactivation. *Annual Review of Microbiology*, *67*, 355–374. https://doi.org/10.1146/annurev-micro-092412-155654

Roizmann, B., Desrosiers, R. C., Fleckenstein, B., Lopez, C., Minson, A. C., & Studdert, M. J. (1992). The family Herpesviridae: an update. The Herpesvirus Study Group of the International Committee on Taxonomy of Viruses. *Archives of Virology*, *123*(3–4), 425–449. https://doi.org/10.1007/BF01317276

Rosato, P. C., & Leib, D. A. (2016). Neurons versus herpes simplex virus: the innate immune interactions that contribute to a host – pathogen standoff. *Future Virology*, *10*(6), 699–714. https://doi.org/10.2217/fvl.15.45.Neurons

Rosenbloom, K. R., Armstrong, J., Barber, G. P., Casper, J., Clawson, H., Diekhans, M., … Kent, W. J. (2015). The UCSC Genome Browser database: 2015 update. *Nucleic Acids Research*, *43*(D1), D670–D681. https://doi.org/10.1093/nar/gku1177

Runwal, G., Stamatakou, E., Siddiqi, F. H., Puri, C., Zhu, Y., & Rubinsztein, D. C. (2019). LC3-positive structures are prominent in autophagy-deficient cells. *Scientific Reports*, *9*(1), 1–14. https://doi.org/10.1038/s41598-019-46657-z

Russell, T. A., & Tscharke, D. C. (2016). Lytic Promoters Express Protein during Herpes Simplex Virus Latency. *PLoS Pathogens*, *12*(6), 1–20. https://doi.org/10.1371/journal.ppat.1005729

Salzman, J. (2016, May 1). Circular RNA Expression: Its Potential Regulation and Function. *Trends in Genetics*, Vol. 32, pp. 309–316. https://doi.org/10.1016/j.tig.2016.03.002

Samuel, C. E. (2001). Antiviral actions of interferons. *Clinical Microbiology Reviews*, Vol. 14, pp. 778–809. https://doi.org/10.1128/CMR.14.4.778-809.2001

Sandri-Goldin, R. M. (1998). ICP27 mediates HSV RNA export by shuttling through a leucine-rich nuclear export signal and binding vital intronless RNAs through an RGG motif. *Genes and Development*, *12*(6), 868–879. https://doi.org/10.1101/gad.12.6.868

Sansregret, L., & Nepveu, A. (2008). The multiple roles of CUX1: Insights from mouse models and cell-based assays. *Gene*, *412*(1–2), 84–94. https://doi.org/10.1016/j.gene.2008.01.017

Sato, R., Kato, A., Chimura, T., Saitoh, S. I., Shibata, T., Murakami, Y., … Miyake, K. (2018). Combating herpesvirus encephalitis by potentiating a TLR3–mTORC2 axis. *Nature Immunology*, *19*(10), 1071–1082. https://doi.org/10.1038/s41590-018-0203-2

Sawtell, N. M., & Thompson, R. L. (2004). Comparison of Herpes Simplex Virus Reactivation in Ganglia In Vivo and in Explants Demonstrates Quantitative and Qualitative Differences. *Journal of Virology*, *78*(14), 7784–7794. https://doi.org/10.1128/jvi.78.14.7784-7794.2004

Sawtell, N M. (1997). Comprehensive quantification of herpes simplex virus latency at the single-cell level. *Journal of Virology*, *71*(7), 5423–5431. https://doi.org/10.1128/jvi.71.7.5423-5431.1997

Sawtell, N M, & Thompson, R. L. (1992a). Herpes simplex virus type 1 latency-associated transcription unit promotes anatomical site-dependent establishment and reactivation from latency. *Journal of Virology*, *66*(4), 2157–2169. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=289008&tool=pmcentrez&rendertype=abstract

Sawtell, N M, & Thompson, R. L. (1992b). Rapid in vivo reactivation of herpes simplex virus in latently infected murine ganglionic neurons after transient hyperthermia. *Journal of Virology*, *66*(4), 2150–2156. https://doi.org/10.1128/jvi.66.4.2150-2156.1992

Sawtell, Nancy M, & Thompson, R. L. (2020). Alphaherpesvirus Latency and Reactivation with a Focus on Herpes Simplex Virus. *Curr Issues Mol Biol.*, *4*(41), 267–356.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., … Cardona, A. (2012). Fiji: an open-source platform for biological-image analysis.

Nature Methods, *9*(7), 676–682. https://doi.org/10.1038/nmeth.2019

Sedarati, F., Izumi, K. M., Wagner, E. K., & Stevens, J. G. (1989). Herpes simplex virus type 1 latency-associated transcription plays no role in establishment or maintenance of a latent infection in murine sensory neurons. *Journal of Virology*, *63*(10), 4455–4458. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=251069&tool=pmcentrez&rendertype=abstract

Selbach, M., Schwanhäusser, B., Thierfelder, N., Fang, Z., Khanin, R., & Rajewsky, N. (2008). Widespread changes in protein synthesis induced by microRNAs. *Nature*, *455*(7209), 58–63. https://doi.org/10.1038/nature07228

Shipley, M M, Mangold, C. A., Kuny, C. V, & Szpara, M. L. (2017). Differentiated Human SH-SY5Y Cells Provide a Reductionist Model of Herpes Simplex Virus 1 Neurotropism. *Journal of Virology*, *91*(2Shipley, M. M., Mangold, C. A., Kuny, C. V, Szpara, M. L. (2017). Differentiated Human SH-SY5Y Cells Provide a Reductionist Model of Herpes Simplex Virus 1 Neurotropism. Journal of Virology, 91(23), 1. https://doi.org/10.1128/JVI.00958-173), 1. https://doi.org/10.1128/JVI.00958-17

Shipley, Mackenzie M., Mangold, C. A., Kuny, C. V., & Szpara, M. L. (2017). Differentiated Human SH-SY5Y Cells Provide a Reductionist Model of Herpes Simplex Virus 1 Neurotropism. *Journal of Virology*, *91*(23). https://doi.org/10.1128/jvi.00958-17

Shipley, Mackenzie M., Mangold, C. A., & Szpara, M. L. (2016). Differentiation of the SH-SY5Y human neuroblastoma cell line. *Journal of Visualized Experiments*, *2016*(108). https://doi.org/10.3791/53193

Shipley, S. J., Parkin, E. T., Itzhaki, R. F., & Dobson, C. B. (2005). Herpes simplex virus interferes with amyloid precursor protein processing. *BMC Microbiology*, *5*(1), 48. https://doi.org/10.1186/1471-2180-5-48

Shnayder, M., Nachshon, A., Rozman, B., Bernstein, B., Lavi, M., Fein, N., … Schwartz, M. (2020). Single cell analysis reveals human cytomegalovirus drives latently infected cells towards an anergic-like monocyte state. *ELife*, *9*, 1–24. https://doi.org/10.7554/eLife.52168

Singh, N., & Tscharke, D. C. (2019). Herpes Simplex Virus Latency Is Noisier the Closer We Look. *Journal of Virology*, *94*(4), 1–7. https://doi.org/10.1128/jvi.01701-

Smith, M. C., Boutell, C., & Davido, D. J. (2011). HSV-1 ICP0: Paving the way for viral replication. *Future Virology*, *6*(4), 421–429. https://doi.org/10.2217/fvl.11.24

Soday, L., Lu, Y., Albarnaz, J. D., Davies, C. T. R., Antrobus, R., Smith, G. L., & Weekes, M. P. (2019). Quantitative Temporal Proteomic Analysis of Vaccinia Virus Infection Reveals Regulation of Histone Deacetylases by an Interferon Antagonist. *Cell Reports*, *27*(6), 1920-1933.e7. https://doi.org/10.1016/j.celrep.2019.04.042

Spivack, J. G., & Fraser, N. W. (1987). Detection of herpes simplex virus type 1 transcripts during latent infection in mice. *Journal of Virology*, *61*(12), 3841–3847.

Steiner, I., Spivack, J. G., Deshmane, S. L., Ace, C. I., Preston, C. M., & Fraser, N. W. (1990). A herpes simplex virus type 1 mutant containing a nontransinducing Vmw65 protein establishes latent infection in vivo in the absence of viral replication and reactivates efficiently from explanted trigeminal ganglia. *Journal of Virology*, *64*(4), 1630–1638. https://doi.org/10.1128/jvi.64.4.1630-1638.1990

Steiner, I., Spivack, J. G., Lirette, R. P., Brown, S. M., MacLean, A. R., Subak-Sharpe, J. H., & Fraser, N. W. (1989). Herpes simplex virus type 1 latency-associated transcripts are evidently not essential for latent infection. *The EMBO Journal*, *8*(2), 505–511. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=400833&tool=pmcentrez&rendertype=abstract

Stevens, J G, Wagner, E. K., Devi-Rao, G. B., Cook, M. L., & Feldman, L. T. (1987). RNA complementary to a herpesvirus alpha gene mRNA is prominent in latently infected neurons. *Science*, *235*(1979), 1056–1059. https://doi.org/10.1126/science.2434993

Stevens, Jack G., & Cook, M. L. (1971). Latent herpes simplex virus in spinal ganglia of mice. *Science*, *173*(3999), 843–845. https://doi.org/10.1126/science.173.3999.843

Stewart, M. (2019). Polyadenylation and nuclear export of mRNAs. *Journal of Biological Chemistry*, *294*(9), 2977–2987. https://doi.org/10.1074/jbc.REV118.005594

Strahl, B. D., & Allis, C. D. (2000). *The language of covalent histone modifications*. *403*(January), 41–45.

Strang, B. L., & Stow, N. D. (2005). Circularization of the Herpes Simplex Virus Type 1 Genome upon Lytic Infection. *Journal of Virology*, *79*(19), 12487–12494. https://doi.org/10.1128/jvi.79.19.12487-12494.2005

Tang, S., Patel, A., & Krause, P. R. (2009). Novel less-abundant viral microRNAs encoded by herpes simplex virus 2 latency-associated transcript and their roles in regulating ICP34.5 and ICP0 mRNAs. *Journal of Virology*, *83*(3), 1433–1442. https://doi.org/10.1128/JVI.01723-08

Tavalai, N., & Stamminger, T. (2009). Interplay between herpesvirus infection and host defense by PML nuclear bodies. *Viruses*, *1*(3), 1240–1264. https://doi.org/10.3390/v1031240

Terry-Allison, T., Smith, C. A., & DeLuca, N. A. (2007). Relaxed Repression of Herpes Simplex Virus Type 1 Genomes in Murine Trigeminal Neurons. *Journal of Virology*, *81*(22), 12394–12405. https://doi.org/10.1128/jvi.01068-07

The R Core Team. (2017). R: A Language and Environment for Statistical Computing. Retrieved April 22, 2020, from Cran website: https://cran.r-project.org/doc/manuals/r-release/fullrefman.pdf

Thellman, N. M., Botting, C., Madaj, Z., & Triezenberg, S. J. (2017). An Immortalized Human Dorsal Root Ganglia Cell Line Provides a Novel Context to Study Herpes Simplex Virus Type-1 Latency and Reactivation. *Journal of Virology*, JVI.00080-17. https://doi.org/10.1128/JVI.00080-17

Thomas, D. L., Lock, M., Zabolotny, J. M., Mohan, B. R., & Fraser, N. W. (2002). The 2-Kilobase Intron of the Herpes Simplex Virus Type 1 Latency-Associated Transcript Has a Half-Life of Approximately 24 Hours in SY5Y and COS-1 Cells. *Journal of Virology*, *76*(2), 532–540. https://doi.org/10.1128/jvi.76.2.532-540.2002

Thompson, R. L., & Sawtell, N. M. (2001). Herpes Simplex Virus Type 1 Latency-Associated Transcript Gene Promotes Neuronal Survival. *Journal of Virology*, *75*(14), 6660–6675. https://doi.org/10.1128/jvi.75.14.6660-6675.2001

Thompson, R L, & Sawtell, N. M. (1997). The herpes simplex virus type 1 latency-associated transcript gene regulates the establishment of latency. *Journal of Virology*, *71*(7), 5432–5440. https://doi.org/10.1128/JVI.75.14.6660-6675.2001

Thompson, Richard L., Preston, C. M., & Sawtell, N. M. (2009). De novo synthesis of VP16 coordinates the exit from HSV latency in vivo. *PLoS Pathogens*, *5*(3). https://doi.org/10.1371/journal.ppat.1000352

Traylen, C. M., Patel, H. R., Fondaw, W., Mahatme, S., Williams, J. F., Walker, L. R., … Akula, S. M. (2011, April). Virus reactivation: A panoramic view in human infections. *Future Virology*, Vol. 6, pp. 451–463. https://doi.org/10.2217/fvl.11.21

Tucker, J. D., & Bertke, A. S. (2019, July 1). Assessment of cognitive impairment in HSV-1 positive schizophrenia and bipolar patients: Systematic review and meta-analysis. *Schizophrenia Research*, Vol. 209, pp. 40–47. https://doi.org/10.1016/j.schres.2019.01.001

Tycowski, K. T., Guo, Y. E., Lee, N., Moss, W. N., Vallery, T. K., Xie, M., & Steitz, J. A. (2015). Viral noncoding RNAs: More surprises. *Genes and Development*, *29*(6), 567–584. https://doi.org/10.1101/gad.259077.115

Umbach, J. L., Nagel, M. A., Cohrs, R. J., Gilden, D. H., & Cullen, B. R. (2009). Analysis of Human Alphaherpesvirus MicroRNA Expression in Latently Infected Human Trigeminal Ganglia. *Journal of Virology*, *83*(20), 10677–10683. https://doi.org/10.1128/jvi.01185-09

Umbach, Jennifer Lin, Kramer, M. F., Jurak, I., Karnowski, H. W., Coen, D. M., & Cullen, B. R. (2008). MicroRNAs expressed by herpes simplex virus 1 during latent infection regulate viral mRNAs. *Nature*, *454*(7205), 780–783. https://doi.org/10.1038/nature07103

Underwood, R., Wang, B., Carico, C., Whitaker, R. H., Placzek, W. J., & Yacoubian, T. A. (2020). The GTPase Rab27b regulates the release, autophagic clearance, and toxicity of α-synuclein. *The Journal of Biological Chemistry*, *295*(23), 8005–8016. https://doi.org/10.1074/jbc.RA120.013337

Ungerleider, N. A., Jain, V., Wang, Y., Maness, N. J., Blair, R. V., Alvarez, X., … Flemington, E. K. (2018). Comparative Analysis of Gammaherpesvirus Circular RNA Repertoires: Conserved and Unique Viral Circular RNAs. *Journal of Virology*, *93*(6), 1952–1970. https://doi.org/10.1128/jvi.01952-18

UniProt. (n.d.). Retrieved from https://www.uniprot.org/

Usoskin, D., Furlan, A., Islam, S., Abdo, H., Lönnerberg, P., Lou, D., … Ernfors, P. (2015). Unbiased classification of sensory neuron types by large-scale single-cell RNA sequencing. *Nature Neuroscience*, *18*(1), 145–153. https://doi.org/10.1038/nn.3881

van Velzen, M., Jing, L., Osterhaus, A. D. M. E., Sette, A., Koelle, D. M., & Verjans, G. M. G. M. (2013). Local CD4 and CD8 T-Cell Reactivity to HSV-1 Antigens

Documents Broad Viral Protein Expression and Immune Competence in Latently Infected Human Trigeminal Ganglia. *PLoS Pathogens*, *9*(8), 12–15. https://doi.org/10.1371/journal.ppat.1003547

Vasudevan, S. (2012). Posttranscriptional Upregulation by MicroRNAs. *Wiley Interdisciplinary Reviews: RNA*, *3*(3), 311–330. https://doi.org/10.1002/wrna.121

Vogel, C., & Marcotte, E. M. (2012). Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. *Nature Reviews Genetics*, *13*(4), 227–232. https://doi.org/10.1038/nrg3185

Wagner, E K, & Bloom, D. C. (1997). Experimental investigation of herpes simplex virus latency. *Clinical Microbiology Reviews*, *10*(3), 419–443. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9227860

Wagner, Edward K., & Bloom, D. C. (1997). *Experimental Investigation of Herpes Simplex Virus Latency*. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC172928/pdf/100419.pdf

Wang, C., Li, D., Zhang, L., Jiang, S., Liang, J., Narita, Y., … Zhong, Q. (2019). *RNA Sequencing Analyses of Gene Expression during Epstein-Barr Virus Infection of Primary B Lymphocytes*. (February), 1–15.

Wang, H., Davidob, D. J., & Morrison, L. A. (2013). HSV-1 strain McKrae is more neuroinvasive than HSV-1 KOS after corneal or vaginal inoculation in mice Hong. *Virus Research*, *173*(2), 436–440. https://doi.org/10.1038/jid.2014.371

Warnes, G. R. (2011). gplots: Various R Programming Tools for Plotting Data.

Webre, J. M., Hill, J. M., Nolan, N. M., Clement, C., McFerrin, H. E., Bhattacharjee, P. S., … Thompson, H. W. (2012). Rabbit and mouse models of HSV-1 latency, reactivation, and recurrent eye diseases. *Journal of Biomedicine and Biotechnology*, *2012*. https://doi.org/10.1155/2012/612316

Weekes, M. P., Tomasec, P., Huttlin, E. L., Fielding, C. A., Nusinow, D., Stanton, R. J., … Gygi, S. P. (2014). Quantitative temporal viromics: An approach to investigate host-pathogen interaction. *Cell*, *157*(6), 1460–1472. https://doi.org/10.1016/j.cell.2014.04.028

Weiss, L. A., & Nieto, M. (2019). The crux of Cux genes in neuronal function and plasticity. *Brain Research*, *1705*(2019), 32–42. https://doi.org/10.1016/j.brainres.2018.02.044

West, A. P., Khoury-Hanold, W., Staron, M., Tal, M. C., Pineda, C. M., Lang, S. M., … Shadel, G. S. (2015). Mitochondrial DNA stress primes the antiviral innate immune response. *Nature*, *520*(7548), 553–557. https://doi.org/10.1038/nature14156

Whisnant, A. W., Jürges, C. S., Hennig, T., Wyler, E., Prusty, B., Rutkowski, A. J., … Dölken, L. (2020). Integrative functional genomics decodes herpes simplex virus 1. *Nature Communications*, *11*(1), 1–14. https://doi.org/10.1038/s41467-020-15992-5

Wilcox, C. L., & Johnson, E. M. (1988). Characterization of nerve growth factor-dependent herpes simplex virus latency in neurons in vitro. *Journal of Virology*, *62*(2), 393–399. https://doi.org/10.1128/jvi.62.2.393-399.1988

Wilcox, C. L., Smith, R. L., Freed, C. R., & Johnson, E. M. (1990). Nerve growth factor-dependence of herpes simplex virus latency in peripheral sympathetic and sensory neurons in vitro. *Journal of Neuroscience*, *10*(4), 1268–1275. https://doi.org/10.1523/jneurosci.10-04-01268.1990

Workman, A., Zhu, L., Keel, B. N., Smith, T. P. L., & Jones, C. (2018). The Wnt Signaling Pathway Is Differentially Expressed during the Bovine Herpesvirus 1 Latency-Reactivation Cycle: Evidence That Two Protein Kinases Associated with Neuronal Survival, Akt3 and BMPR2, Are Expressed at Higher Levels during Latency. *Journal of Virology*, *92*(7). https://doi.org/10.1128/jvi.01937-17

Wozniak, M., Mee, A., & Itzhaki, R. (2009). Herpes simplex virus type 1 DNA is located within Alzheimer's disease amyloid plaques. *The Journal of Pathology*, *217*(1), 131–138. https://doi.org/10.1002/path.2449

Wu, T. T., Su, Y. H., Block, T. M., & Taylor, J. M. (1998). Atypical splicing of the latency-associated transcripts of herpes simplex type 1. *Virology*, *243*(1), 140–149. https://doi.org/10.1006/viro.1998.9036

Wyler, E., Franke, V., Menegatti, J., Kocks, C., Boltengagen, A., Praktiknjo, S., … Landthaler, M. (2019). Single-cell RNA-sequencing of herpes simplex virus 1-infected cells connects NRF2 activation to an antiviral program. *Nature Communications*, 1–14. https://doi.org/10.1038/s41467-019-12894-z

Wysocka, J., & Herr, W. (2003). The herpes simplex virus VP16-induced complex: The makings of a regulatory switch. *Trends in Biochemical Sciences*, *28*(6), 294–304. https://doi.org/10.1016/S0968-0004(03)00088-4

Xing, J., Ni, L., Wang, S., Wang, K., Lin, R., & Zheng, C. (2013). Herpes Simplex Virus 1-Encoded Tegument Protein VP16 Abrogates the Production of Beta Interferon (IFN) by Inhibiting NF- B Activation and Blocking IFN Regulatory Factor 3 To Recruit Its Coactivator CBP. *Journal of Virology*, *87*(17), 9788–9801. https://doi.org/10.1128/jvi.01440-13

Yang, Q. qiao, & Zhou, J. wei. (2019). Neuroinflammation in the central nervous system: Symphony of glial cells. *Glia*, *67*(6), 1017–1035. https://doi.org/10.1002/glia.23571

Yewdell, J. W., & Bennink, J. R. (2002). *Mechanisms of Viral Interference with MHC Class I Antigen Processing and Presentation*. Retrieved from https://www.ncbi.nlm.nih.gov/books/NBK2226/

You, Y., Cheng, A. C., Wang, M. S., Jia, R. Y., Sun, K. F., Yang, Q., … Chen, X. Y. (2017). The suppression of apoptosis by α-herpesvirus. *Cell Death and Disease*, *8*(4), 1–12. https://doi.org/10.1038/cddis.2017.139

Yousuf, W., Ibrahim, H., Harfouche, M., Abu Hijleh, F., & Abu-Raddad, L. (2020). Herpes simplex virus type 1 in Europe: Systematic review, meta-analyses and meta-regressions. *BMJ Global Health*, *5*(7), 1–15. https://doi.org/10.1136/bmjgh-2020-002388

Zhang, G., Zong, J., Lin, S., Verhoeven, R. J. A., Tong, S., Chen, Y., … Chen, H. (2015). Circulating Epstein-Barr virus microRNAs miR-BART7 and miR-BART13 as biomarkers for nasopharyngeal carcinoma diagnosis and treatment. *International Journal of Cancer*, *136*(5), E301–E312. https://doi.org/10.1002/ijc.29206

Zhang, L., Gao, X., Wen, J., Ning, Y., & Chen, Y. G. (2006). Dapper 1 antagonizes Wnt signaling by promoting dishevelled degradation. *Journal of Biological Chemistry*, *281*(13), 8607–8612. https://doi.org/10.1074/jbc.M600274200

Zhang, S. Y., Jouanguy, E., Ugolini, S., Smahi, A., Elain, G., Romero, P., … Casanova, J. L. (2007). TLR3 deficiency in patients with herpes simplex encephalitis. *Science*, *317*(5844), 1522–1527. https://doi.org/10.1126/science.1139522

Zhao, H., Zhang, C., Hou, G., & Song, J. (2015). MicroRNA-H4-5p encoded by HSV-1 latency-associated transcript promotes cell proliferation, invasion and cell cycle progression via p16-mediated PI3K-Akt signaling pathway in SHSY5Y cells.

International Journal of Clinical and Experimental Medicine, *8*(5), 7526–7534. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/26221296

Zhu, L., Workman, A., & Jones, C. (2017). Potential Role for a β-Catenin Coactivator (High-Mobility Group AT–Hook 1 Protein) during the Latency-Reactivation Cycle of Bovine Herpesvirus 1. *Journal of Virology*, *91*(5). https://doi.org/10.1128/jvi.02132-16

Zimmer, B., Ewaleifoh, O., Harschnitz, O., Lee, Y., & Peneau, C. (2018). *Human iPSC-derived trigeminal neurons lack constitutive TLR3-dependent immunity that protects cortical neurons from HSV-1 infection*. *115*(37). https://doi.org/10.1073/pnas.1809853115

Zufferey, R., Nagy, D., Mandel, R. J., Naldini, L., & Trono, D. (1997). *Multiply attenuated lentiviral vector achieves efficient gene delivery in vivo Romain*. *3*(11), 3–8.

APPENDICES

Appendix 1

Please see Excel file: Appendix 1 – Gene lists.xlsx

Appendix 1 lists the significantly differentially expressed genes (DEGs) found by each RNA-Seq experiment, organised into relevant sub-sections. Where possible, the genes are listed along with their fold change and/or adjusted p-value. Where fold-changes are included the genes are ordered according to these (from largest to smallest), otherwise the gene lists are arranged alphabetically.

Appendix 2

Please see Excel file: Appendix 2 DAVID analysis tables.xlsx

Appendix 2 shows the results of the DAVID analyses performed (methods 6.2.7.2) with the full gene lists associated with each pathway.