KSHV-mediated dysregulation of circRNAs and RNA processing

Katherine Louise Harper

Submitted in accordance with the requirements for the degree of
Doctor of Philosophy

The University of Leeds
Faculty of Biological Sciences
School of Molecular and Cellular Biology

September 2022

Chapters 3 and 4 form work published in Harper KL, Mottram TJ, Anene CA, Foster B, Patterson MR, McDonnell E, Macdonald A, Westhead D, Whitehouse A. Dysregulation of the miR-30c/DLL4 axis by circHIPK3 is essential for KSHV lytic replication. EMBO Rep. 2022 May 4;23(5):e54117. doi: 10.15252/embr.202154117.

The candidate confirms that the work submitted is her own and that appropriate credit has been given where reference has been made to the work of others.

Acknowledgements

I would like to thank my supervisor Professor Ade Whitehouse for all his help and support throughout this PhD. Your level of guidance has only been equalled by your ability to ruthlessly mock the many scientific failures along the way.

I would also like to thank all the members of the Whitehouse lab throughout the years, you have made this PhD much more enjoyable with many laughs along the way. In particular: Becky (RIP Professor Becky Fisher), James (the eternal optimist), Ollie (all of this is meaningless anyway), Freddy (the in-lab pyromaniac), Dr Timothy James Mottram (who requested his full title) and finally toad who instead of writing her thesis managed a beautiful dedication: *'I would like to dedicate this thesis, and my entire life, to Elena May Harrington. She truly has shown me what it is like to live, laugh, AND love. She is my entire support system and saved my thesis by finally providing some meaningful research for it. Thank you Ellie. I owe you my life. Never stop shining bright, and I will always follow the light you leave.'*

Many thanks also go to all my friends outside of the Whitehouse lab, both at home and in Leeds who have all contributed in keeping me sane, particularly Molly who although she has not contributed to my sanity, has always had my back and who shares a single solitary brain cell with me.

Finally I would acknowledge my family, especially my parents and brother, who have supported and helped me from long before I started this PhD.

Abstract

The majority of the human genome is actively transcribed into RNAs with no obvious coding capacity, termed non-coding RNAs (ncRNAs). It is now widely accepted that ncRNAs have crucial roles in many aspects of cell regulation and as such their dysregulation has been implicated in the development of a wide variety of diseases. Non-coding RNAs include long non-coding RNAs, microRNAs (miRNAs) and circular RNAs (circRNAs). Originally it was thought that each ncRNA species worked independently, however there is emerging evidence that different ncRNAs can interact with each other through RNA:RNA interactions, affecting their function, in what is now known as a 'ncRNA regulatory network'. circRNAs are formed through a unique backsplice mechanism and have crucial roles in these ncRNA regulatory networks through their miRNA sponging function. Though research is still in its infancy, evidence suggests circRNA levels are tightly regulated in the cell, with dysregulated circRNA levels being implicated in a range of diseases including viral infection.

We have investigated the role of cellular circRNAs in Kaposi's Sarcoma-Associated Herpesvirus (KSHV) infection. Like all herpesviruses, it has two distinct phases of its life cycle: lytic replication and latency. Chapters 3 and 4 highlight that the cellular circRNA, circHIPK3, is dysregulated during KSHV lytic replication and this dysregulation is important for successful viral replication. Furthermore, a novel non-coding RNA regulatory network is identified, composed of circHIPK3:miR-30c:DLL4, with all aspects of this network contributing to successful KSHV lytic replication. Collectively this data shows how the virus-mediated circRNA dysregulation allows critical regulation of downstream targets essential for viral replication.

Following these findings, Chapter 5 aimed to elucidate how cellular circRNAs are dysregulated during KSHV lytic replication. This led to the analysis and characterisation of the host cell RNA binding protein SFPQ, during KSHV infection. Results demonstrate

that SFPQ forms novel viral-induced condensates, reminiscent of modified paraspeckles during KSHV lytic replication. Further analysis suggest these condensates are likely to have essential roles in RNA processing, including circRNA biogenesis.

Contents

List of Figures

List of Tables:

Abbreviations

A	Adenine
ADAR	Adenosine deaminases acting on RNA
Ago2	Argonaute 2
AIDS	Acquired Immunodeficiency Syndrome
ALS	Amyotrophic lateral sclerosis
ARS2	Serrate RNA effector molecule homolog
BAF	BRG1-associated factor
BCA	Bicinchoninic acid assay
BCL2	BCL2 Apoptosis Regulator
BLV	Bovine leukaemia virus
BSA	Bovine Serum Albumin
BTG1	BTG Anti-proliferation factor 1
C	Cytosine
CARM	Coactivator Associated Arginine Methyltransferase 1
CCL5	C-C Motif Chemokine Ligand 5
CCNB1	Cyclin B1
CCNE1	Cyclin E1
CDC6	Cell division cycle 6
CDK9	Cyclin Dependent Kinase 9
ceRNA	competing endogenous RNA
C/EBPβ	CCAAT enhancer binding protein β
CFIm	Cleavage factor I
circRNA	circular RNA
cNHEJ	Classical non-homologous end-joining
CPSF6	Cleavage and polyadenylation specificity factor subunit 6
CTD	C-terminal domain
C3PO	Component 3 promoter of RISC
DAPI	4’,6-diamidino-2-phenylindole
DBHS	Drosophila behaviour/human splicing
DCC	Deleted in colorectal carcinoma

DC-SIGN	Dendritic Cell-specific Intercellular adhesion molecule-3-grabbing non-integrin
DDX17	DEAD-Box Helicase 17
DDX21	DExD-Box Helicase 21
DGCR8	DiGeorge Syndrome critical region 8
DHX9	DExH-Box Helicase 9
DLL4	Delta Like Canonical Notch Ligand 4
DMEM	Dulbecco's modified eagle medium
DNA	deoxyribose nucleic acid
Ds	double stranded
DTT	Dithiothreitol
E	Early
EBV	Epstein Barr Virus
Ecircs	Exonic circRNAs
E. coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid
EIcircs	Exonic-intronic circRNAs
eIF2α	Eukaryotic Initiation Factor 2α
eIF4a	Eukaryotic Initiation Factor 4a
EJC	Exon junction complex
EPHA2	Ephrin type-A receptor 2 precursor
ER	Endoplasmic reticulum
EZH2	Enhancer of zeste homolog 2
FAK	Focal adhesion kinase
FBS	Foetal Bovine Serum
FISH	Fluorescent *In situ* Hybridisation
FTD	Frontotemporal dementia
FUS	FUS RNA binding protein
G	Guanine
GADPG	Glyceraldehyde-3-phosphate dehydrogenase
gB	glycoprotein B
gC	glycoprotein c
gD	glycoprotein D

GFP	Green fluorescent protein
gH	glycoprotein H
gL	glycoprotein L
GTP	Guanosine triphosphate
HAART	Highly active antiretroviral therapy
HBV	Hepatitis B Virus
HCV	Hepatitis C Virus
HEXIM1	HEXIM P-TEFb Complex Subunit 1
HIF	Hypoxia-inducible factors
HIPK3	Homeodomain interacting protein kinase 3
hnRNP A2B1	Heterogeneous Nuclear Ribonucleoprotein A2/B1
hnRNP C	Heterogeneous Nuclear Ribonucleoprotein C
hnRNP M	Heterogeneous Nuclear Ribonucleoprotein M
hnRNP U	Heterogeneous Nuclear Ribonucleoprotein U
HRE	Hypoxia Response Element
HSC	Haematopoietic stem cell
HSV-1	Herpes Simplex Virus-1
HSV-2	Herpes Simplex Virus-2
HCMV	Human Cytomegalovirus
HHV-1	Human herpesvirus-1
HHV-2	Human herpesvirus-2
HHV-3	Human herpesvirus-3
HHV-4	Human herpesvirus-4
HHV-5	Human herpesvirus-5
HHV-6	Human herpesvirus-6
HHV-7	Human herpesvirus-7
HHV-8	Human herpesvirus-8
HIV	Human immunodeficiency Virus
HPV	Human papillomavirus
HSURs	Herpesvirus saimiri U RNAs
hTREX	Human Transcription-Export complex
HVS	Herpesvirus Saimiri

ICTV	International Committee on Taxonomy of Viruses
IE	Immediate early
IF	Immunofluorescence
IgG	Immunoglobin G
IL-6	Interleukin 6
IL-8	Interleukin 8
IP	Immunoprecipitation
ITCH	Itchy E3 ubiquitin protein ligase
JAK-STAT	Janus kinase signal transducer and activator of transcription
Kb	kilobase
KCl	Potassium chloride
KD	Knockdown
KICS	Kaposi's Sarcoma Inflammatory Cytokine Syndrome
KS	Kaposi's sarcoma
KSHV	Kaposi's sarcoma-associated herpesvirus
LANA	Latency-associated nuclear antigen
LATs	Latency Associated Transcripts
L	Late
LNA	Locked nucleic acid
lncRNA	long non-coding RNA
MALAT1	metastasis associated lung adenocarcinoma
MCD	Multicentric Castleman's disease
MCM	Minichromosome maintenance protein complex
MCMV	Murine cytomegalovirus
McHV-1	Macacine alphaherpesvirus 1
$MgCl_2$	Magnesium Chloride
miRDE	miRNA decay elements
miRNA	microRNA
miR-Seq	microRNA Sequencing
MRE	microRNA response element
mRNA	messengerRNA

MTA	mRNA transcript accumulation
MTE	MTA-response elements
NaCl	Sodium chloride
NaOH	Sodium hydroxide
ncRNA	non-coding RNA
NEAT1	Nuclear Enriched Abundant Transcript 1
NGS	Next generation sequencing
nm	Nanometre
nM	Nanomolar
NONO	Non-POU domain-containing octamer-binding protein
NOPs	NONA/ParaSpeckle
NS4b	Non-structural protein 4b
NTD	N-terminal domain
NUDT21	Nudix Hydrolase 21
ORC	Origin recognition complex
OncomiR	Oncogenic miRNA
ORF	Open Reading Frame
Ori-Lyt	Origins of lytic DNA replication
Ori-P	Origin of latent plasmid replication
Pan	Polyadenylated nuclear RNA
PARP1	Poly (ADP-ribose) polymerase 1
PEL	Primary Effusion Lymphoma
PKR	Protein kinase R
PLD	Prion like domain
Pre-miRNA	Pre microRNA
Pri-miRNA	Primary microRNA
PRMT1	Protein Arginine Methyltransferase 1
PS	Paraspeckle
PSP	Paraspeckle associated protein
PSPC1	Paraspeckle component 1
P/S	Penicillin/Streptomycin
PVT1	Plasmacytoma variant translocation 1

Rac1	Ras-related C3 botulinum toxin substrate 1
RAD51	DNA repair protein RAD51 homolog 1
RBM14	RNA binding motif protein 14
RFP	Red fluorescent protein
RhoA	Ras homolog family member A
RIG-I	Retinoic acid-inducible gene I
RIP	RNA immunoprecipitation
RISC	RNA-induced silencing complex
RNA	Ribose nucleic acid
RNA pol II	RNA polymerase II
RNS	Reactive nitrogen species
ROS	Reactive oxygen species
rRNA	ribosomal RNA
RTA	Replication and transcription activator
RT	Reverse transcription
RUNX	Runt-related transcription factors
SARS-CoV	Severe acute respiratory syndrome-related coronavirus
Scr	Scrambled
SDS	sodium dodecyl sulphate
SDS-PAGE	sodium dodecyl sulphate polyacrylamide gel electrophoresis
SFPQ	Splicing Factor Proline and Glutamine Rich
sfRNA	subgenomic flavivirus RNA
shRNA	short hairpin RNA
SINE	Short interspersed nuclear elements
Sin3A	SIN3 Transcription Regulator Family Member A
SMARCA4	SWI/SNF Related Matrix Associated Actin Dependent Regulator of Chromatin Subfamily A Member 4
snRNPs	small nuclear ribonucleoproteins
SPECC1	sperm antigen with calponin homology and coiled-coil domains 1
SRC	Proto-oncogene tyrosine-protein kinase Src

SRSF2	serine and arginine rich splicing factor 2
SRSF3	serine and arginine rich splicing factor 3
SRSF10	serine and arginine rich splicing factor 10
Ss	single stranded
SWI/SNF	SWItch/Sucrose Non-Fermentable
T	Thymine
TDMD	Target RNA miRNA degradation
TDP-43	TAR DNA binding protein 43
TLR	Toll-like receptor
TRBP	Transactivation response element RNA-binding protein
Tris	tris(hydroxymethyl)aminomethane
tRNA	transfer RNA
U	Uracil
UAP56	ATP-dependent RNA helicase UAP56
UTR	Untranslated region
UV	Ultraviolet
v-FLIP	Viral FADD-like interleukin-1-beta-converting enzyme [FLICE/caspase 8] inhibitory protein
v-IL6	Viral interleukin 6
v-IRF3	Viral interferon regulatory factor 3
v-IRF4	Viral interferon regulatory factor 4
VSR	Viral suppressors of RNA silencing
VZV	Varicella Zoster Virus
XRN2	5′-3′ Exoribonuclease 2
YTDHC1	YTH domain-containing protein 1

Chapter 1

~

Introduction

1.0 Introduction

1.1 Herpesviridae

1.1.1 Classification

The *Herpesviridae*, *Alloherpesviridae* and *Malacoherpesviridae* families comprise the order *Herpesvirales*, itself placed within class I of the Baltimore classification system [1]. *Herpesviridae* include viruses that infect mammalian, reptile and bird species, *Alloherpesviridae* include fish and frog viruses, while members of *Malacoherpesviridae* infect molluscs [2]. *Herpesvirales* is a highly divergent order, with very few conserved proteins amongst all members. Of the few proteins conserved across the order, none are unique to herpesviruses with homologues found in other viruses [2] [3]. It is estimated *Herpesvirales* members diverged around 400 million years ago allowing this large diversification [4].

As the largest and most widely studied family, *Herpesviridae* is further divided into 3 subfamilies: *Alphaherpesvirinae*, *Betaherpesvirinae* and *Gammaherpesvirinae*. The family is estimated to have diverged into its subfamilies 170-209 million years ago leading to over 40 conserved genes within the family, with a range of functions including DNA assembly, virion structure and nucleotide enzymes [5] [6] [4]. Herpesviruses are grouped into subfamilies using a combination of factors, including morphological structure, biological criteria such as host species and type of infection, genome organisation and most recently through genome sequence [7]. Across these 3 subfamilies within *Herpesviridae*, originally 8 human herpesviruses were classified, named Human Herpesvirus 1-8 (HHV) according to the ICTV, with 6 of these also known by another name. However, since the initial classification, 2 variants of HHV-6 have been identified, named as HHV-6a and HHV-6b [7]. Therefore the human herpesviruses are the alphaherpesviruses: HSV-1, HSV-2 and VSV; the betaherpesviruses: HCMV, HHV-6a, HHV-6b and HHV-7, and the gammaherpesviruses EBV and KSHV (Figure 1).

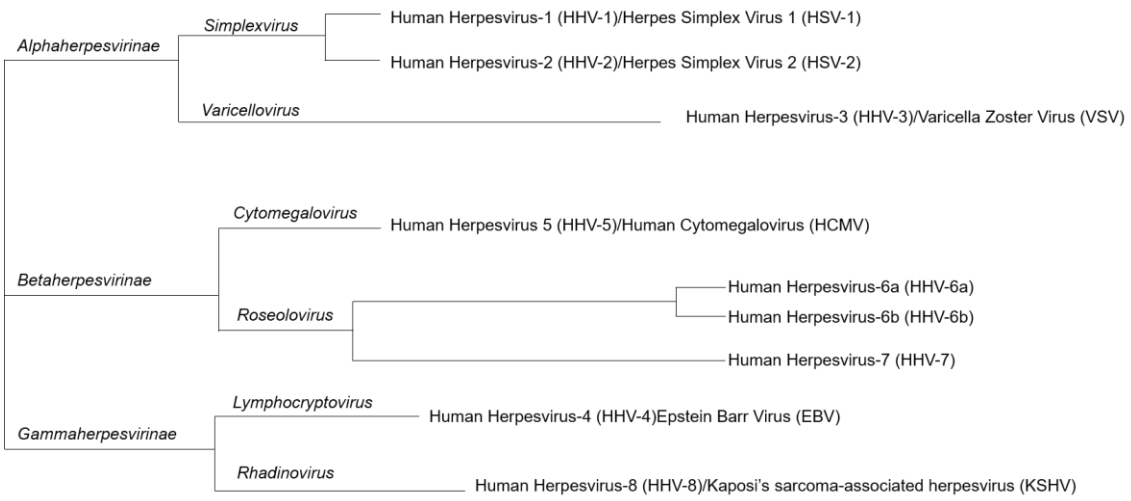

Figure 1: Herpesviridae classification: The phylogenetic tree of the 9 human herpesviruses showing the evolutionary relationship between members of *Herpesviridae* based on amino acid similarity [8].

1.1.1.1 Alphaherpesvirinae

Alphaherpesvirinae contains the genera: *Simplexvirus*, *Varicellovirus*, *Mardivirus*, *Scutavirus* and *Iltovirus*, of which *Simplexvirus and Varicellovirus* include human herpesviruses. Alphaherpesviruses are characterised by having rapid replication cycles, a wide host range, lysing infected cells and maintaining latency in neuronal cells, particularly sensory nerve ganglia [9]. They often infect epithelia cells during lytic replication leading to efficient transmission [10]. The human pathogens include HSV-1, HSV-2 (*Simplexvirus)* and VSV (*Varicellovirus*). Additionally, the macaque virus McHV-1 is also an alphaherpesvirus in the genus *Simplexvirus*, is capable of causing severe neuronal disease in humans with a fatality rate of 80% [7].

1.1.1.2 Betaherpesvirinae

Betaherpesvirinae contain the genera *Cytomegalovirus, Muromegalovirus, Proboscivirus* and *Roseolovirus*, with HCMV belonging to *Cytomegalovirus* and HHV-6 and HHV-7 belonging to *Roseolovirus.* In contrast to *Alphaherpesvirinae, Betaherpesvirinae* viruses are characterised by a slow replication cycle and are far less cytopathic. Infection often leads to the formation of multi-nucleated cells named

cytomegalia. They maintain latent infection in kidney, secretory and lymphoreticular cells [9].

1.1.1.3. Gammaherpesvirinae

Gammaherpesvirinae include the genera *Lymphocryptovirus and Rhadinovirus,* of which EBV and KSHV are respective members. The other genera are *Macavirus* and *Percavirus.* Gammaherpesvirus are characterised by latent infection of lymphoid cells including B or T cells, certain members of this subfamily can also undergo lytic replication in endothelial cells. Both human pathogens in this subfamily are classed as oncogenic [11]. Originally it was thought gammaherpesviruses had a narrow host range, however, on the contrary it has been shown several gammaherpesviruses can infect a range of cell types across different members of an animal order [12].

1.1.2 Genome

Herpesviridae genomes comprise of linear dsDNA ranging from 125-295 kb in length, with the GC content varying from 32-75%. Similarly, the number of ORFs range from 70 to over 200, although there are many conserved genes involved in DNA assembly, virion structure and nucleotide enzymes [2]. Apart from these 40 or so conserved genes across *herpesviridae*, the family has a remarkable number of virus-specific genes, with estimates approximating that around 20% of total ORFs in *herpesviridae* are unique to a single herpesvirus type. For instance, up to 70% of HCMV ORFs are thought to be virus-specific [13]. *Herpesviridae* genomes are classified in 6 different ways depending on their organisation, interestingly, this organisation is not dependent on the virus subfamily, with different genome classifications within the same subfamily (Figure 2) [2] [7]. These classes are shown in Figure 2.

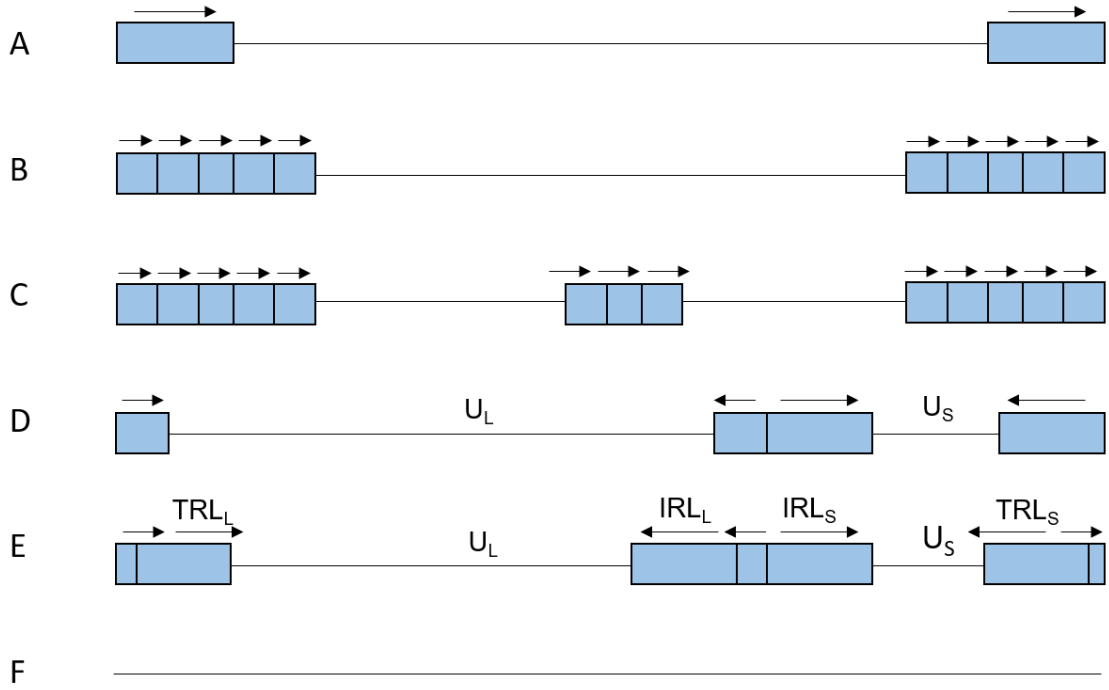

Figure 2: The 6 classifications of Herpesviridae genome organisation. Class A genomes comprise a unique coding region flanked by a single direct repeat, e.g. HHV-7. Class B comprise a unique coding region flanked by a variable number of direct repeats, e.g. KSHV. Class C genomes are similar to class B but contain an internal region of direct repeats, e.g. EBV. Class D have 2 unique coding regions, each of which is flanked by inverted repeats. The larger unique coding region is termed U_L while the smaller is U_S, e.g. VSV. Class E have 2 unique coding regions flanked by inverted repeats and have terminally redundant regions that are located adjacent to the inverted repeats, e.g. HCMV and HSV-1.Class F comprises a unique region with no regions of repeats, only a single identified *Betaherpesvirinae*, Tupaia herpesvirus (THV), is known to contain this structure. Adapted from [2] [7] [14].

1.2 KSHV

1.2.1 Discovery and epidemiology

KSHV was discovered in 1993 by Chang and Moore. DNA fragments isolated from KS tissues from AIDS patients were identified to be gene homologues to the tegument and capsid of known gammaherpesviruses: EBV and herpesvirus saimiri. [15]. Further sequencing characterised a novel gammaherpesvirus now known as HHV-8 or KSHV [16] [17].

Due to its recent discovery, less is known about KSHV than other herpesviruses, not least its worldwide distribution. However, serological investigations suggest it is not as ubiquitous in the population as EBV, with antibodies against KSHV in the USA and Western Europe found in 2-5% of the population. However, uniquely KSHV has an

uneven world distribution, with 10-25% of Mediterranean populations carrying antibodies against KSHV and over 30% of the Sub-Saharan population [18] (Figure 3). Transmission of KSHV is also unknown but is hypothesised to be a combination of childhood transmission, shedding through saliva and mucosal tissues and sexual transmission [19] [20] [21] [22].

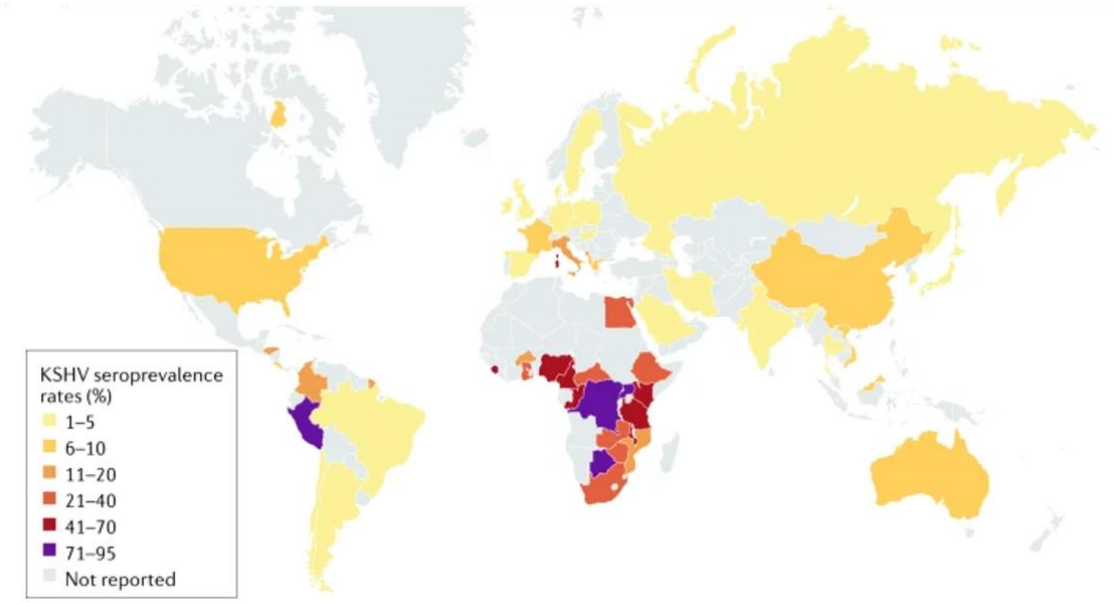

Figure 3: Known seroprevalence distribution of KSHV. Adapted from [23].

1.2.2 KSHV associated diseases

1.2.2.1 Kaposi's Sarcoma

KS is a highly vascularised heterogeneous endothelial cancer and is the most common malignancy associated with KSHV infection [23]. There are 4 types of KS: classic, endemic, iatrogenic and AIDS-related, all of which have varying survival outcomes. Classic KS is mostly found in elderly Mediterranean men and is generally mild with localised lesions and little metastasis to lymph nodes or internal organs [24]. Endemic KS is found mainly in African children and is much more aggressive, spreading to the lymph nodes and having a high fatality rate. Iatrogenic KS occurs after immunosuppression post-transplantation, with KS incidence being around 200-fold higher than in the general population [23]. AIDS-associated KS is the most aggressive

and common form of KS and is rapidly fatal with frequent lymph node invasion. During the HIV/AIDS epidemic of the 1980s there was a surge of KS cases in groups previously not at risk, such as white homosexual men, that lead to its discovery in 1994. Incidence of KS in AIDS patients was estimated to be 20,000-fold higher than in the general population, with the rate being 1/100,000 in the general population, 1/20,000 in AIDS patients and 1/3 for HIV/AIDS+ homosexual men [25] [23]. Since the introduction of HAART therapy against HIV/AIDS the incidence of KS has decreased, although it is still the most common tumour in men in Sub-Saharan Africa and the second most common for women [23].

1.2.2.2 Primary effusion lymphoma

PEL is a B cell non-Hodgkin lymphoma, derived from a clonal expansion of malignant B cells which accumulate in body cavities [24]. PEL is highly aggressive with a mean survival of 2-6 months, similarly to KS it is most common in AIDS patients and males. Interestingly, a large percentage of PEL are co-infected with EBV, with estimates varying between 60 and 90%, however, there are EBV-negative PEL cases particularly in HIV/AIDS+ patients [26].

1.2.2.3 Multicentric Castleman's disease

MCD is a polyclonal expansion of B cells and therefore is not categorised as a true cancer but as a lymphoproliferative disorder. It is highly aggressive and rapidly progresses with a high fatality rate, characterised by systemic inflammatory symptoms. There are 2 classifications of MCD: plasmablastic and hyaline; with KSHV associated with the plasmablastic variant. KSHV is found in nearly all HIV/AIDS+ MCD cases and around 50% of HIV/AIDS- cases [24].

1.2.2.4 Kaposi's Sarcoma Inflammatory Cytokine Syndrome

KICS is the most recently described KSHV-associated disease, first recognised in 2010 with a fatality rate of around 60% [27]. KICS occurs in patients who are both HIV and KSHV + and is characterised as a systemic inflammatory disease, with patients often

exhibiting high levels of IL-6 (both human and viral) and IL-10 [28]. Due to its rarity, there is little research on the condition to date.

1.2.2.5 KSHV-mediated tumourgenesis

How KSHV drives cancer formation is not fully elucidated with many unanswered questions. KS tumours are characterised by being highly vascularised with inflammatory cell invasion and being composed of spindle cells. However, the progenitor of these spindle cells is not known, with tumours exhibiting heterogeneous cell markers. Haematopoietic stem cells (HSCs) are one of the more likely candidates, with the capacity to differentiate into the endothelial cells seen in KS. Furthermore, KS cells are often recruited to sites of injury through inflammatory cytokines with elevated inflammatory cytokine levels capable of inducing spindle cell formation, indicating once again HSCs as the progenitor cells. Notably these cytokines are often found elevated in HIV/AIDS patients [29].

Both the KSHV latent and lytic lifecycles contribute to oncogenesis. During latency, several genes have oncogenic potential including LANA and a range of virally encoded miRNAs, however, although they can lead to some hallmarks of cancer, they are insufficient for complete cellular transformation. Therefore, lytic replication also contributes to the transformation process, where several lytic proteins including K1, vGPCR and vIRF-1 (table 1) are highly pro-inflammatory, pro-angiogenetic and immunosuppressive. As such, a combination of both lifecycles are needed, with lytic replication preventing loss of the KSHV episome over time in the rapidly dividing transforming cell population. It is also hypothesised KSHV lytic replication can promote oncogenesis through the lytic oncoproteins stimulating KSHV latently infected cells in a paracrine manner [30]. Abortive lytic replication is also a contributing factor, here the virus enters lytic replication and expresses its early lytic genes, but then aborts the cascade and re-enters latency. This allows the expression of certain oncogenic proteins without virion productive and lysis of the cell [31]. In HIV+ and immunosuppressed

patients, the repressed immune system also leads to reduced cell death during lytic replication, aiding in the increased likelihood of developing a KSHV-associated malignancy. Some of these oncogenic factors are summarised in the table below.

Name	Lifecycle	Mechanism
K1	Lytic	Aids angiogenesis through triggering secretion of VEGF and MMP-9. Mediates EMT
K3	Lytic	Ubiquitin ligase that downregulates immune response
K5	Lytic	Ubiquitin ligase that downregulates immune response
LANA	Latent	Inactivate p53 and Rb pathways. Antagonises Wnt. Stabilises Notch, promotes VEGF factors.
miR-K1	Latent	Targets IκBα leading to NF-κB activation
miR-K3	Latent	Targets C/EBPβ. Also targets GRK2 leading to increased CXCR2 which induces angiogenesis and cell migration
miR-K6-3p	Latent	Activates STAT3 leading to increased cell proliferation
miR-K11	Latent	Targets C/EBPβ a negative regulator of IL-6, leading to B cell proliferation. Degrades THBS1- a negative regulator of angiogenesis
v-cyclin	Latent	Promotes cell proliferation through CDK4/6
v-FLIP	Latent	Activates NF-κB, leads to production of pro-angiogenic IL-6
vGPCR	Lytic	Induces Erk, PI3K/AKT/mTOR cascade. Activates secretion of VEGF and IL-6. Activates ROS and angiogenesis. Mediates EMT
vIRF-1	Lytic	Inhibits p53 activation

Table 1: The known oncogenic KSHV encoded genes [29] [32] [33] [34] [35] [36].

1.2.3 KSHV Genome

KSHV has a double-stranded DNA genome comprising 165-170kb in length, of which, around 140kb is the long unique region and the GC-rich terminal repeats then vary in number from 16-75 [24]. Encoded within its long unique region are at least 87 ORFs and 25 miRNAs [25]. KSHV was the first human member of the *Rhadinovirus* genus discovered and is therefore closely related to herpesvirus saimiri and thus they share many gene homologues [37]. However, KSHV also encodes multiple KSHV-unique genes, designated K1-K15. These genes have been used to classify KSHV into distinct subtypes, with 5 subtypes currently identified: A, B, C, D and E, each with a unique geographical profile. Furthermore, KSHV has a high number of genes that have no known homologues in other human herpesviruses, nearly 25 out of the 95 genes fulfilling these criteria [38]. Many of these are cellular orthologues, with at least 14 being captured or diverged from eukaryotic cells [25].

1.2.4 Structure

Similar, to all herpesviruses, KSHV has its DNA genome encapsulated into an icosahedral capsid. This is surrounded by a tegument which is further enclosed in a lipid bilayer [39]. KSHV mature virions are 145 nm in diameter with the capsid measuring 110 nm and the core 75 nm [40]. The capsid itself comprises of 162 capsomers and is around 100-110 nm in diameter with a T=16 symmetry [39] [41]. Of the 162 capsomers, 150 are hexamers and 11 pentamers. This capsid also contains a portal at one of the twelve pentameric vertices to enable DNA packaging [42] [43].

The portal has been characterised to different extents across human herpesviruses, with initial studies suggesting it is a conserved structural component [44]. The capsid itself is comprised of 6 proteins, including the major and minor capsid proteins: ORF25 and ORF65 respectively [45]. Within the KSHV virion there are also several cellular proteins including E2-F, tubulin, ezrin, moesin and Hsc-70 [46].

The tegument is a proteinaceous amorphous layer unique to herpesviruses thought to have roles in virion envelopment, virion trafficking, transcriptional/translational regulation, immune evasion and egress [47] [48] [49]. The composition of the tegument is dependent on the herpesvirus, although, there are also highly conserved tegument proteins across all herpesviruses [50]. The composition of the KSHV tegument has not been fully elucidated, however, it is thought to contain at least 9 of the conserved proteins [51].

This tegument is surrounded by the lipid bilayer envelope which contains a set of 3 conserved glycoproteins: gB (ORFF8), gH (ORF22) and gL (ORF47), with gH-gL forming a heterodimer [52]. These glycoproteins are essential for fusion during virus entry with further additional glycoproteins ORF68, K8.1, gN (ORF53) and gM (ORF39) [45].

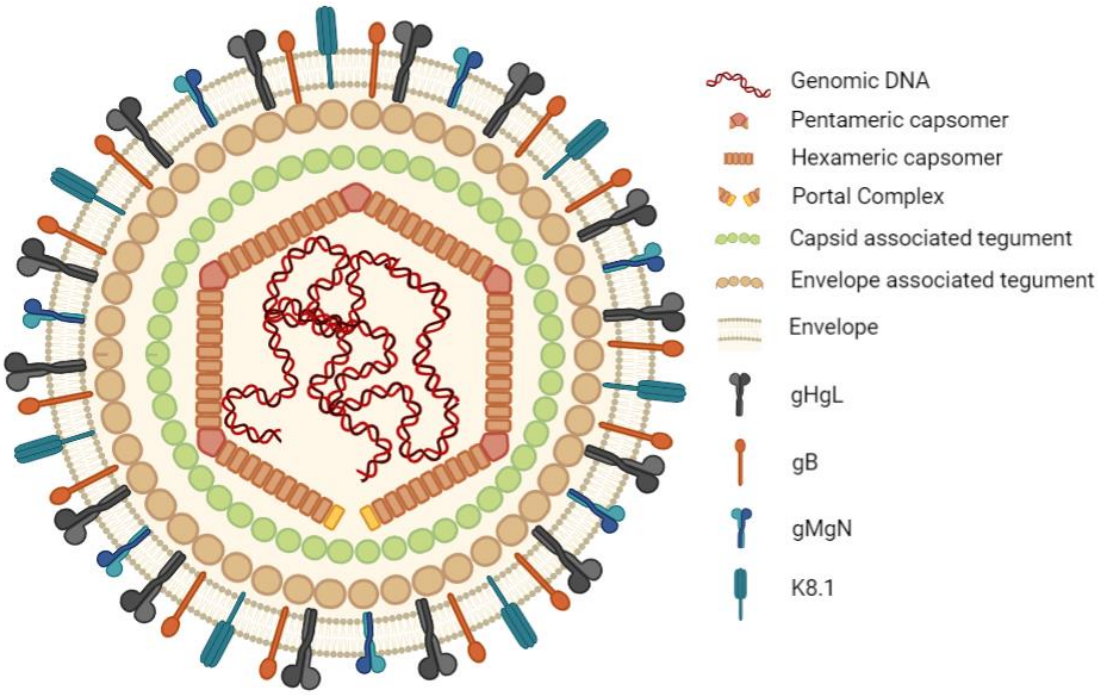

Figure 4: The structure of the KSHV virion. Provided by Dr James Murphy (internal communications)

1.2.5 Virus lifecycle

1.2.5.1 Entry and trafficking

KSHV is capable of infecting many cell types including B cells, monocytes, dendrites and endothelial cells, with the cellular receptor varying depending on cell type. For example, in monocytes, B cells and dendrites KSHV utilises DC-SIGN as the main receptor for cellular entry [53]. Other receptors include heparin sulphate, xCT transporter, tyrosine kinase receptor EPHA2 and the integrins α3β1, αVβ5, and αVβ3 [23]. KSHV has 9 glycoproteins on its surface, with 4 of these having glycoprotein homologues in other herpesviruses [54] [46]. These glycoproteins are ORF4, ORF45, K8.1, ORF28 and ORF68, with gB, gHgL, glycoprotein L and glycoprotein N having homologues. Entry is mediated through KSHV glycoproteins binding to cell surface receptors. The first interactions between glycoproteins and receptors are non-specific, with the glycoproteins, particularly K8.1 binding to heparin sulphate proteoglycans on the cell surface [55]. This concentrates the virions close to the other receptors, allowing specific

stronger interactions to occur, such as binding to ephrins and DC-SIGN, this then leads to induction of cell signalling aiding in virion entry through endocytosis [56]. Heparin sulphate proteoglycans are seen as important but not essential in KSHV viral entry, especially as they are not expressed in B cells. B cell entry mechanisms are still not fully elucidated, however, K8.1 is thought to play an important role in B cell entry even though expression of its only known binder is lacking in B cells [57].

Binding of KSHV to surface receptors leads to induction of multiple signalling pathways including the essential activation of FAK and Src kinases. This results in the induction of PI3-K which in turn activates Rho-GTPases [58]. Rho proteins including RhoA, Rac1 and cCbl drive the formation of filopodia, stress fibres and lamellipodia leading to morphological changes within the membrane, thought to aid in endocytosis [59] [60]. Post endocytosis, the virion uncoats and the capsid is transported to the nucleus via the microtubule network. This cytoskeletal manipulation occurs quickly, with the KSHV genome able to traffic to the nucleus as early as 15 minutes post infection, and peaking by 90 minutes [59]. The microtubules are stabilised through acetylation triggered by the previous activated Rho-GTPases, with the KSHV genome trafficked via dynein motors [59] [61]. Once at the nuclear membrane, viral DNA enters through nuclear pores and circularises, before entering an initial phase of lytic replication. Here low levels of RTA are expressed to allow addition of epigenetic modifications allowing binding of cellular proteins to induce LANA. This leads to the establishment of the latent phase [62].

1.2.5.2 Latency

Latency is characterised by the long term maintenance of the viral genome without the production of infectious virions. Here very few genes are expressed, with the major KSHV latency reservoir thought to be in B cells [63].

Upon nuclear entry and the establishment of a latent state, the viral genome forms circular episomes attached to cellular chromosomes via the KSHV-encoded protein, LANA. DNA replication during latency is initiated at multiple sites within the episome

termed origins of replication (ori-P), which are located in the terminal repeats. DNA replication is performed using the host cell polymerase and allows genome copy number maintenance into daughter cells in line with mitosis [45]. During DNA replication, LANA recruits a number of cellular proteins forming a pre-replication complex, to ori-P sites including cdc6, PARP1, MCMs and the ORC to aid in DNA replication [64].

During latency only a few KSHV ORFs are expressed, including LANA, v-cyclin, v-FLIP, K12, vIL-6, v-IRF3 with their roles including inhibition of apoptosis, modulation of signalling pathways, promoting cell proliferation and maintenance of latency [65] [45] [66].

Interestingly, KSHV encodes its own non-coding RNAs that are expressed during latency. 12 pre-miRNAs are expressed and processed into 25 mature miRNAs. 10 of these pre-miRNAs are found in a cluster between Kaposin and v-FLIP. Another is found within the Kaposin ORF and the final pre-miRNA ORF is found between Kaposin and ORF69 [67]. The location and putative roles of these miRNAs are summarised in Figure 5. Furthermore, several of these miRNAs are classified as oncogenic, in particular miR-K12-11 which shares 100% seed sequence homology with the known human oncomiR miR-155. Both miRNAs target the IL-6 negative regulator C/EBPβ leading to B cell proliferation [67].

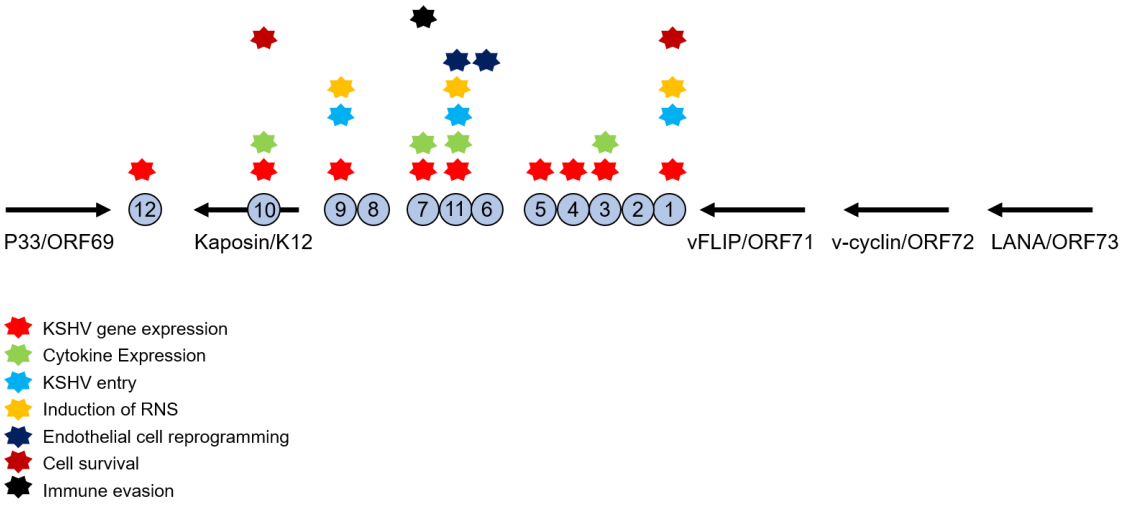

Figure 5: KSHV encoded miRNAs. Schematic showing genome localisation of the 12 KSHV encoded miRNA genes with their proposed functions highlighted [68] [67] [69].

1.2.5.3 Induction of lytic replication

Lytic replication is important for the production of new infectious virions and is crucial to transmit the infection to new cells. This balance between latent and lytic replication is controlled by the competitive interaction between LANA and the latent-lytic switch protein known as RTA. LANA generally suppresses RTA function, however, under certain conditions, RTA can override LANA function. Here, RTA, when at sufficient levels can out-compete LANA for binding to the key cellular protein RBP-Jκ, allowing RTA to bind to over 100 RTA binding sites within the KSHV genome. This leads to transcription of at least 34 KSHV lytic genes [70]. RTA also has a positive feedback loop, binding to its own promoter, increasing its own levels, further promoting lytic replication. RTA also utilises its own E3 ubiquitin ligase activity to target several repressive cellular proteins for proteasomal degradation, as well as several KSHV latent proteins including vFLIP [71] [72].

Many factors lead to the induction of lytic replication, although their roles are not all fully elucidated. Known induction factors include proteins which alter the epigenetic state of the genome, co-infection, hypoxia, cellular stress and oxidative stress activating the MAPK3 signalling cascades [73] [74] [75].

Hypoxia contributes to lytic replication through HIF-1 and HIF-2, which are transcribed in response to low oxygen levels [76]. Numerous hypoxia response elements (HREs) are found throughout the KSHV genome, enabling HIFs binding and promotion of transcription, including RTA [77] [74].

Induction of lytic replication is also linked to the epigenetic state of the KSHV genome. During latency, activating histone modifications are enriched on the latent genes, whereas repressive modifications are seen on lytic genes. However, during lytic replication, EZH2, a member of the polycomb repressive complex 2 which maintains these modifications, dissociates from the complex. This leads to activating histone modifications on lytic genes aiding their expression [74].

Finally infection with other pathogens can lead to induction of lytic replication, for instance HPV, HIV, HCMV and HSV-1 co-infection have been implicated. It is hypothesised viral infection leads to activation of inflammatory cytokines which promote KSHV reactivation [75]. Bacterial pathogens have also been shown to contribute to lytic reactivation, for instance *S. aureus* secretes PAMPs which activate TLR signalling promoting lytic replication [78]. However, the majority of studies investigating co-infection have focused on EBV and HIV.

EBV is often found in up to 70% of PEL tumours and nearly all KSHV-MCD tumours [79], with tumours co-infected with EBV having increased tumourigenicity [80]. However, studies have found that EBV and KSHV repress each other's reactivation [81]. The EBV equivalent of RTA, Zta, can interact with RTA, inhibiting both proteins' function and thereby preventing reactivation [81]. Furthermore the EBV latent ncRNA LMP1 suppresses both EBV and KSHV lytic replication, while KSHV through RTA and the cellular protein RBP-Jκ binds to EBV latency promoters, once again suppressing EBV lytic replication [82]. However, as well as direct interactions, both viruses lead to changes within the cell that can be beneficial or inhibitory to the other virus. For instance, EBV

miRNAs aid in KSHV-mediated immuno-evasion through suppressing immune signalling [83].

KS is an AIDs-defining illness, people with HIV are 20,000 times more likely to develop KS than a HIV-negative individual. Originally, it was thought that KS is so common in HIV/AIDs patients due to the indirect effect of HIV mediated immune-suppression, however, KS is still 300 times more prevalent in HIV+ immunosuppressed individuals than HIV- immunosuppressed people, implicating additional factors at work [74]. Interestingly, this enhancement of KS cases in HIV+ patients has only been observed with HIV-1, not HIV-2, suggesting differences in viral-host cell biology [84].

The HIV encoded proteins Tat and Nef are thought to aid in both KSHV replication and in induction of KS. Tat through stimulation of cytokines, can promote a pro-angiogenic environment. Furthermore when Tat is co-expressed with vIL-6 it leads to enhancement of angiogenesis and tumourgenesis [85]. Additionally it has also been suggested that Tat through IL-6 and the JAK-STAT pathway activation aids in lytic replication induction [86]. Nef, in contrast acts synergistically with K1 to promote cell proliferation and angiogenesis [87]. KSHV also enhances HIV replication, with LANA capable of binding to Tat and improving its transcriptional abilities while several KSHV proteins, including ORF57, ORF45, vFLIP, RTA and LANA, are thought to bind to the HIV LTR which is critical for viral replication [74] [88].

1.2.5.4 Lytic replication

During lytic replication, lytic proteins are expressed in a temporal fashion starting with immediate early (IE genes), followed by early (E) and finally late (L). In the lytic cascade, the first protein expressed is the lytic/latent switch RTA/ORF50 protein, which is necessary and sufficient for lytic induction [74]. RTA acts as a transcription factor leading to the expression of other KSHV genes, thus starting the lytic cascade. Interestingly RTA activates viral promoters using 3 distinct mechanisms. Firstly it can bind directly to promoters through type I RTA-response elements (RREs) such as the ones present in

the PAN promoter [89]. Secondly it can also bind to type II RREs, found in the ORF57 and K8 promoters, however, these also require cellular factors for transcription initiation, such as RBP-Jκ [90]. Finally RTA has E3 ubiquitin ligase activity leading to proteasomal degradation of several lytic suppressing proteins, both KSHV and cellular, such as LANA and Hey1 [74].

Another key lytic protein is encoded by ORF57, which is expressed early in the cascade and has key roles in RNA export, RNA splicing and stability [91] [92] [93]. Many KSHV genes lack introns and therefore are unable to recruit the splicing cellular hTREX complex in a splicing-dependent manner. Therefore ORF57 mimics splicing by binding the viral intronless RNAs and then recruiting hTREX. This interaction stabilises the viral mRNAs as well as enabling the further recruitment of the TAP/p15 nuclear export heterodimer which allows translocation of the viral mRNAs through the nuclear pore [94].

The majority of KSHV-encoded miRNAs are not expressed during lytic replication, however, it does express a long non-coding RNA, termed PAN. PAN is a 1.1kb polyadenylated transcript that accumulates in the nucleus, totalling up to 80% of all KSHV transcript expression. PAN is essential for late viral gene expression and can regulate chromatin structures both negatively and positively, suggesting PAN is a master regulator of KSHV lytic replication [95].

A further important aspect of KSHV lytic replication is host-cell shut off, resulting in the active degradation of around 80% of host cell mRNAs [96]. This is thought to reduce competition for the host cell translation machinery, and occurs through the action of the ORF37 protein (SOX) which can act as an endo and exonuclease [97].

A key aspect of the lytic lifecycle is DNA replication, initially this occurs as part of a theta mechanism before proceeding as a rolling circle mechanism which initiates from 2 origins of replication termed *ori-Lyt(L)* and *ori-Lyt(R)*. DNA replication occurs after the IE and E genes have been expressed as many of these gene products are involved in DNA

replication. Ori-Lyt contain binding sites for multiple replication factors and also the RTA protein, which is essential for replication [98]. RTA binding recruits K8 upstream which leads to unwinding of the episome allowing access of the replication complex. The DNA replication complex comprises of the KSHV proteins ORF9 (polymerase), ORF59 (DNA processivity factor), ORF6, ORF44 (helicase) and ORF56 (primase), as well as cellular proteins including hnRNP U, MSH2, PARP-1 and topoisomerase [99] [100] [74]. DNA replication, along with transcription of viral transcript occurs in viral replication centres (VRCs), these are membraneless nuclear structures induced by KSHV to provide an environment where beneficial host factors are enriched, such as RNA pol II.

Late genes are expressed after viral DNA replication has begun. The exact mechanism controlling the delay in the expression of these genes is unknown, however, it is hypothesised that DNA replication leads to viral DNA remodelling, allowing polymerases and transcription factors access to late gene promoters [74].

The DNA is then packaged into the capsid in the VRCs in the nucleus, capsids then exit by budding through the double nuclear membrane briefly gaining their first envelope before fusing with the second nuclear membrane. The viral tegument is acquired in the cytoplasm and it is proposed the final envelopment occurs by budding into Golgi vesicles. These vesicles then merge with the cellular plasma membrane releasing the mature virion, although details of egress and assembly for KSHV are not well elucidated [74].

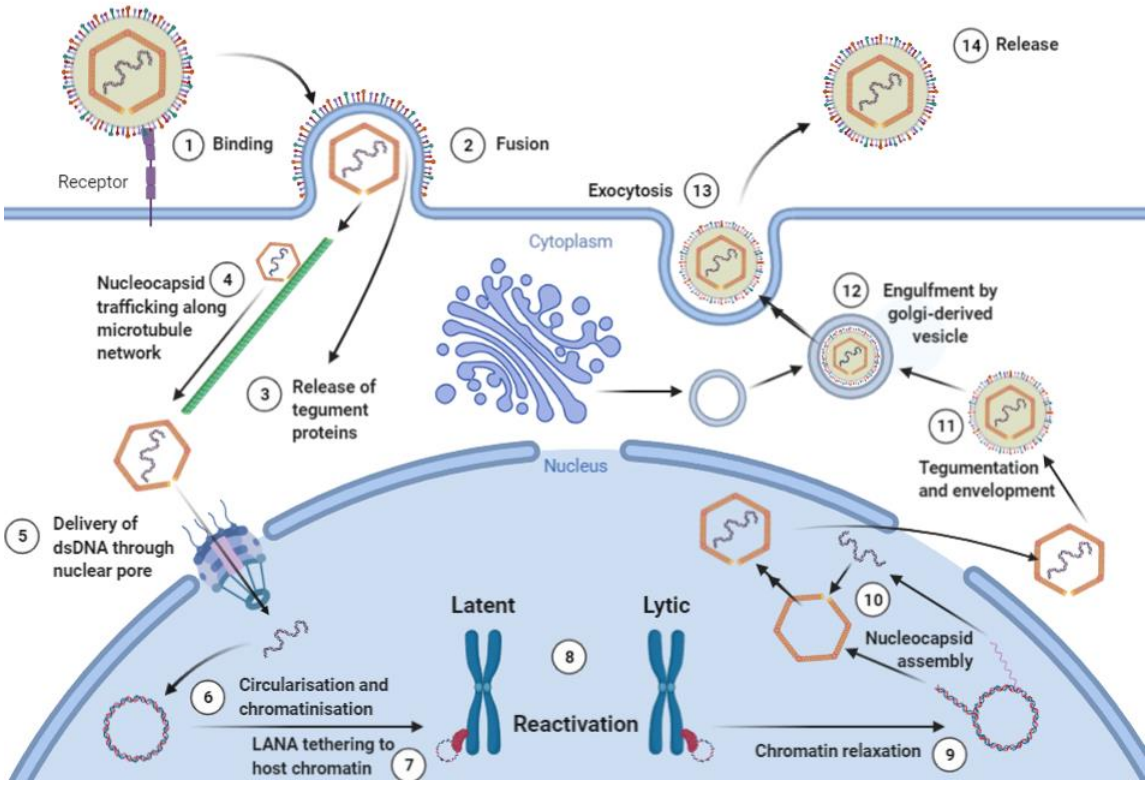

Figure 6: The KSHV lifecycle. Provided by Dr Oliver Manners (personal communication).

1.2.6 KSHV cell culture models

The primary cell culture models for studying KSHV lytic replication and reactivation are B lymphocytes latently infected with KSHV that can be chemically induced to active replication. For instance TREx-BCBL1-RTA are B cell line derived from a PEL, that have been modified to have a doxycycline inducible RTA. This allows consistent KSHV reactivation, essential for reproducible and reliable results. Additionally, there are no KSHV animal model systems, preventing the study of the direct relationship between the latently infected B cell reservoir and secondary sites of infection. However, there are several disadvantages of the system: namely the difficulty of transfecting B cells, poor production of infectious virions and the physiological relevance of simultaneous reactivation of a large proportion of the cell population.

1.3 microRNAs

microRNAs (miRNAs) are a species of small non-coding RNA involved in post-transcriptional regulation of gene expression.

1.3.1 Biogenesis

miRNA genes can be found in both intergenic and intragenic regions of the human genome in around a 50/50 ratio [101]. Of the intergenic miRNA genes, around 1/3 have their own promoters to enable transcription [102]. The vast majority of these miRNAs are transcribed by RNA pol II, although there is also evidence that RNA pol III is involved in the transcription of certain miRNAs [103] [104]. miRNAs are transcribed as primary-miRNAs (pri-miRNAs) that are around 1 kb in length, pri-miRNAs must have a stem loop region as well as single stranded overhangs at each end of the transcript to allow further processing [105]. Of note intragenic miRNA genes are often transcribed as part of the pre-mRNA and therefore must be further processed. There are 2 predominant mechanism for miRNA excision from these mRNAs; either excision mediated through spliceosomal components resulting in a separated pri or pre-miRNA that can enter the miRNA biogenesis machinery, alternatively the pre-miRNA is processed directly from the pre-mRNA by the microprocessor complex [105].

Pri-miRNAs must be further processed to pre-miRNAs, these are shorter at around 60-70 bp and form a hairpin loop. The processing is carried out by the microprocessor complex, which is composed of the RNAse III enzyme Drosha, along with the molecular ruler DGCR8. DGCR8 can recognise the ss overhangs and directs Drosha to cleave 11 bp from the pri-miRNA stem loop [106]. The pre-miRNA is then exported from the nucleus by the Exportin 5 Ran-GTP complex [107].

The final maturation step occurs in the cytoplasm through the activity of the RNAse III enzyme Dicer, which binds to a 2-nucleotide overhang at the 3′ end and cleaves the pre-miRNA into a miRNA: miRNA* duplex, this is assisted by TRBP [108]. This duplex is composed of the more active guide strand and the passenger strand which are designated the miRNA and the miRNA* respectively, usually the passenger strand is degraded. The guide strand is generally the strand with a lower 5′ stability or a U at the 5′ end [109]. This duplex is loaded onto the RISC complex, in humans there are 8

classifications of RISC dependent on their composition, containing Argonaute proteins 1-4, with the majority centred around the protein Argonaute2 (Ago2) [105].

The processing of the duplex into a single stranded mature miRNA is still fully not elucidated, however, there are multiple hypotheses to explain its generation. These include cleavage of the duplex by Dicer, cleavage by Ago2 in tandem with the protein C3PO, mismatches between the duplex promoting unwinding or cleavage by Ago2 in the RISC complex aided by the RNA helicase DHX9 which forms a complex with Dicer, TRBP and Ago2 [105]. This leaves behind the mature miRNA of between 19-22 bp. Within the RISC complex it can then bind to its target mRNA, usually binding to a complementary sequence within the 3′ UTR leading to its translational repression or degradation. This biogenesis is summarised in Figure 7.

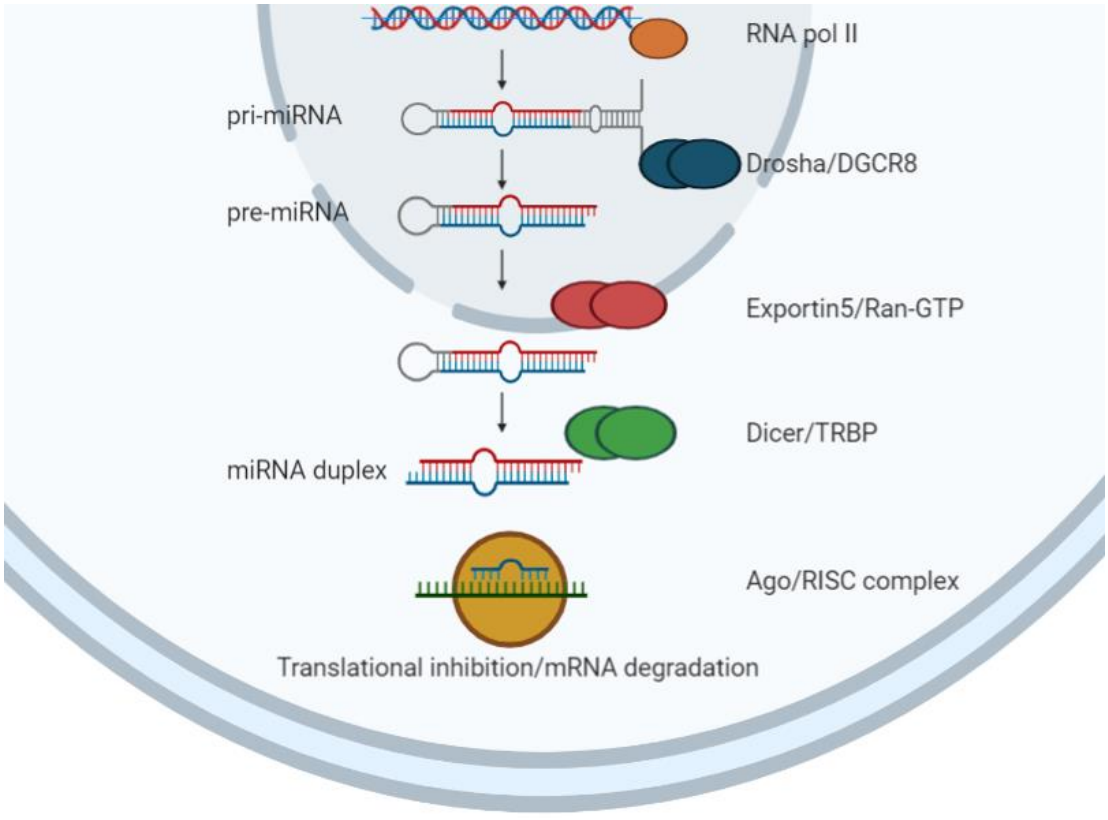

Figure 7: miRNA biogenesis. miRNAs are transcribed by RNA pol II as primary miRNAs. These are processed by the microprocessor complex comprising of DGCR8 which recognises and binds to the overhangs, and Drosha, which cleaves the pri-miRNA into a shorter pre-miRNA. This pre-miRNA is exported from the nucleus utilising Exportin5/Ran-GTP before final cytoplasmic maturation. Dicer, assisted by TRBP cleaves the pre-miRNA into the mature miRNA duplex. The active guide strand of this duplex can be loaded onto the RISC complex, where along with Ago2 it can lead to repression or degradation of target mRNAs through sequence complementarity.

1.3.2 Function

miRNAs are post-transcriptional regulators, with up to 60% of human mRNA transcripts thought to be potentially targeted by miRNAs [110]. They bind to a complementary sequence in its target mRNA, usually in the 3′ UTR leading to either degradation of the mRNA or its translational repression. If the miRNA exhibits high complementarity to its target mRNA it leads to degradation of the mRNA through Ago slicing, followed by deadenylation, decapping and exosomal degradation [105]. Lower complementarity leads to bulges between the miRNA and mRNA mismatches, preventing the action of Ago. Instead, the mRNA undergoes translational repression through a variety of proposed mechanisms. These include inhibition of mRNA circularisation, thereby reducing translational efficiency or competing with eIF4a for binding at the 5′ cap again reducing translation efficiency [111]. Additionally, miRNA binding may interfere with ribosomal loading, either through preventing the recruitment of the 40S subunit or the RISC complex recruiting eIF6 which prevents 80S assembly [112] [113]. mRNA that undergoes translational repression is thought to be stored in cytoplasmic foci called P bodies, forming a pool of mRNAs that can later re-enter the translation cycle if required. Therefore, unlike miRNA mediated degradation, miRNA translational repression is reversible, although the exact role of P bodies and miRNAs is unknown [114].

1.3.3 Regulation of miRNAs

miRNAs are regulated through a variety of mechanisms, some of these mechanisms are briefly summarised below. One such mechanism is target RNA miRNA degradation (TDMD), here when a miRNA binds its target RNA, instead of target RNA degradation, the miRNA is degraded [115] [116]. TDMD appears to be highly specific to select miRNAs, with the mechanism reliant on high complementarity at the 5′ and 3′ end of the miRNA: mRNA duplex with a mismatch in the middle. Originally discovered in viral infections, in recent years cellular examples of TDMD have also been found; although the exact mechanisms and regulatory controls are unknown, however, the degradation

itself centres around the addition of adenines to the miRNA allowing the action of 3′-5′ exonucleases [117].

miRNAs can also be regulated through epigenetic control, such as histone modifications. DNA hypermethylation of CpG islands leads to reduced transcription of the pri-miRNA and as such results in reduced levels of mature miRNAs [118]. RNA editing through enzymes such as ADAR are likewise crucial, with their editing ability leading to base changes which can promote or inhibit the Watson-Crick interactions needed for miRNA: mRNA interactions. It is estimated that up to 16% of all pri-miRNAs undergo some form of RNA editing [119].

However, one of the most important mechanisms of miRNA regulation is through sequestration by other non-coding RNAs species, so-called competing endogenous RNA (ceRNA). These ceRNAs include long non-coding RNAs (lncRNAs), pseudogenes and circular RNAs (circRNA) [117]. ceRNAs work through competitive binding of the miRNA, acting as a sponge, and thereby preventing it binding to its target RNA. These RNA regulatory networks, where RNA species undergo crosstalk through microRNA response elements (MRE) to regulate the cell, independently of protein interactions, were first proposed in 2011 and since then have become an accepted dogma to regulate gene expression [120].

1.3.4 Virus-encoded miRNAs

miRNAs are key post-transcriptional regulators of many cellular processes, therefore it is not surprising that viruses also encode miRNAs. miRNAs are well suited for viruses requiring limited coding capacity and are non-immunogenic [121]. There are over 250 identified virus-encoded miRNAs, with the vast majority in the *herpesviridae* family [122], however, functionally most are not elucidated. Not only do *herpesviridae* collectively encode the majority of miRNAs, individually their genomes contain many more miRNAs than other viruses, for instance, EBV and Rhesus Lymphocryptovirus encoding at least 44 and 50 mature miRNAs respectively [123].

Originally it was thought that only DNA viruses could encode miRNAs, with miRNAs found in *herpesviridae, polyomaviridae, papillomaviridae, adenoviridae* and *baculoviridae* [123], before being found encoded in one retrovirus: Bovine leukaemia virus (BLV). However, these miRNAs have a short hairpin structure allowing bypass of Drosha and immediate cleavage by Dicer. Additionally, there is however some controversy over whether another retrovirus HIV encodes its own miRNAs [122].

Due to the complexity of the genome processing required for miRNA biogenesis by RNase III enzymes it was initially suggested that RNA viruses could not encode for miRNAs, however, recent evidence has contradicted this, with several groups discovering Ebola virus encoded miRNAs [124] [125]. Furthermore, proof-of-principle experiments with artificial miRNA sequences inserted into RNA viruses have been performed in Influenza and tick-borne encephalitis virus [126] [127].

1.3.4 Viral dysregulation of host miRNAs

Viruses are capable of dysregulating nearly every aspect of host cell processes, therefore there are numerous examples of viruses manipulating host cell miRNAs to aid their replication. The best studied is Hepatitis C Virus (HCV) and miR-122. This miRNA is essential for successful HCV replication, with cells lacking miR-122 not capable of sustained infection, conversely, overexpression of miR-122 increased cell permissibility to the virus [128]. Interestingly, miR-122 binds to two sites within the 5′ UTR of the virus, and in contrast to leading to its degradation, it disguises the virus from detection from immune surveillance by RIG-I and prevents degradation by 5′-3′ exonucleases. It is further hypothesised that miR-122 bound to HCV can recruit Ago2, which in turn, blocks the activity of 5′ exonucleases leading to increased viral genome stability [129] [130].

Dengue Virus also dysregulates cellular miRNAs, with infection leading to an overall downregulation of the miRNA machinery. The virally-encoded protein NS4b binds to dicer, inhibiting its function, preventing the maturation of miRNAs [131]. From 151 randomly selected miRNAs 98% were downregulated. Dengue virus also encodes an

ncRNA in its 3′UTR, this subgenomic flavivirus RNA (sfRNA) localises to P bodies and acts as a molecular sponge, competing with host cell miRNAs for Ago binding, thereby dysregulating host cell miRNA machinery by an additional mechanism [132].

Dysregulation of host cell miRNAs has been observed in many more viruses. Briefly these include: West Nile Virus encoding an sfRNA to inhibit dicer [133], Hepatitis B Virus (HBV) downregulating miR-122 [117] as well as several viruses encoding viral suppressors of RNA silencing (VSRs). Viruses known to encode VSRs include Ebola virus, Severe Acute Respiratory Syndrome-related coronavirus (SARS CoV), Influenza A, HCV and HIV [131].

Finally, the herpesviruses are known to alter cellular miRNAs, with several utilising TDMD. Herpesvirus Saimiri (HVS) expresses 7 small nuclear RNAs, termed HSURs, which bind to complementary target miRNAs, including miR-27a and miR-27b, sponging them and leading to their degradation. Murine cytomegalovirus (MCMV) encodes transcript m169 which has multiple binding sites for miR-27, again leading to its degradation via TDMD. Whilst HCMV encodes a bicistronic RNA UL144-145 which contains miRNA decay elements (miRDE) triggering TDMD for its complementary miRNAs, namely miR-17 and miR-20 [134]. Microarray studies have also shown KSHV dysregulates between 170 and 185 cellular miRNAs, although the mechanism behind this dysregulation is not currently known [135]. However, with the majority downregulated, it is likely KSHV utilises some form of miRNA sponge, in a similar fashion to other herpesviruses.

1.4 Circular RNAs

Circular RNAs (circRNAs) are a type of non-coding RNA formed through a unique backsplicing mechanism. Originally thought to be formed from errors in splicing, recent research has shown they may have a range of roles in regulating cellular processes.

1.4.1 Discovery and evolution

The first evidence of RNA forming circles was found in 1976 when 4 plant viroids were visualised under electron microscopy and their genomes were seen to form irregular circular structures [136]. 3 years later in 1979 electron microscopy was used to identify more single stranded circular RNA structures in the cytoplasm of HeLa cells although the source and function was unknown [137]. The first identified circRNA found in human cells followed in 1986, where using electron microscopy and electrophoresis migration the hepatitis δ virus was shown to form a circular RNA genome [138]. The first circular RNA identified as originating in humans was in 1991, here a circular isoform of the gene *DCC* with out of order exons was discovered, although it was initially dismissed as an error in splicing [139]. Further circRNAs were discovered in 1993 in both humans and mice, before in 1995, the importance of inverted flanking repeats was recognised in circRNA biogenesis [140]. With the use of RNA-Seq gaining popularity in the 21st century, an increasing number of circRNAs were identified, although they were still seen as low abundance and originating from errors in splicing. The use of RNase R in 2006 to degrade linear transcripts before RNA-Seq, allowed increased read depth of circRNAs [141]. However, it was advances in NextGen RNA-Seq that really expanded the known number of circRNAs, these advances allowed depth coverage of RNA to be increased through rRNA depletion rather than poly(A) selection. Poly(A) although useful for identification of mRNAs fails to enrich circRNAs due to their lack of a poly(A) tail [142]. Interest in circRNAs was truly peaked in 2013, when a functional role was suggested as a miRNA sponge [143].

These improvements in sequencing have shown that circRNAs can be highly expressed, with up to 10% of circRNAs expressed to higher levels than the gene they are derived from. The variety of circRNAs is also vast, with one study alone finding twice as many circRNAs as there are total identified lncRNAs, although, their functionality is not confirmed. It must be noted that circRNA expression levels appear to be tissue specific,

with up to 64% of leukocyte expressed circRNAs being exclusive to that specific cell type [144].

The recent interest in circRNAs has led to their discovery across the eukaryotic tree of life, with circRNAs found in species from humans and zebrafish to fungi and protists. circRNAs have now been found in eukaryotic species that diverged over 1 billion years ago, highlighting that circRNAs are widely conserved or there have been multiple separate events leading to the formation of circRNAs with both theories advocating the importance of circRNAs [145]. Of note, is the discovery of circRNAs in species such as *Plasmodium falciparum* and *Schizosaccharomyces pombe*, in these species canonical alternative splicing is a rare event, with relatively little exon skipping and other mechanisms proposed to be involved in circRNA biogenesis. This once again highlights that circRNAs are not just a by-product from splicing and suggests that other biogenesis mechanisms have yet to be elucidated. Additionally, the presence of circRNAs across genetically complex and simple species highlights their importance for cellular function.

The conservation across the tree of life has also challenged our knowledge on the requirements for circRNA biogenesis. In humans, circRNAs are often formed from genes with longer than average introns, furthermore the average intron length in humans overall is 5.4 kb long. Therefore, it was assumed long flanking introns was a requirement for circRNA formation, however, the discovery of circRNAs in species where introns average less than 200 nucleotides contradicts this, in fact a circRNA has been found derived from the *MATA1* gene in S. cerevisiae with flanking introns of only 60 nucleotides [145].

Finally, analysis of exons that form circRNAs has shown they are often more conserved in homologues between mice and humans, than exons that only form linear RNA. This would be expected if circRNAs have important non-coding functions [146].

1.4.2 Biogenesis and regulation

The exact mechanisms behind circRNA biogenesis are not fully elucidated, however, all circRNAs undergo some form of backsplicing. In canonical splicing, small nuclear ribonucleoproteins (snRNPs) catalyse the removal of introns and the joining of exons, forming a linear transcript in the 5′-3′ direction, whereas, in circRNA formation, the downstream 3′ splice donor site is joined to an upstream 5′ acceptor site [147] (Figure 8). circRNAs can be formed from either exons or introns or both and as such are classified as exonic (Ecircs), intronic (ciRNAs) or exonic-intronic (EIcircs) [148].

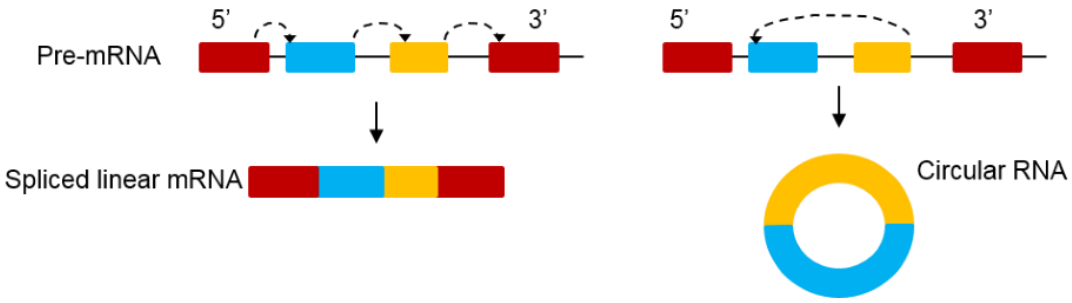

Figure 8: Backsplicing. Canonical splicing vs backsplicing for formation of linear mRNA vs circular RNAs.

Exonic circRNAs form the majority of identified circRNAs, with two proposed mechanisms for their formation: lariat driven or intron pairing (Figure 9). Lariat-driven formation is triggered through exon skipping, where exons are spliced from the main linear transcript forming an intermediate lariat. This lariat can then undergo further splicing to form an exonic circRNA, although the factors involved are not known. In intron pairing, circularisation is driven by the flanking introns having highly complementary regions leading to base pairing between them bringing the splice sites in proximity allowing *trans* factors to complete the formation through direct backsplicing. These *trans* factors are commonly RBPs that are capable of dimerisation; through binding to the distal ends, the RBPs bring the splice sites in closer proximity during dimerisation. Particularly common in the flanking introns of exonic circRNAs are ALU repeats which demonstrate high complementarity, promoting base pairing of reverse complementary matches (RCM), which is critical in promoting circularisation through bringing the distal ends in

close proximity [149]. It must also be noted that canonical splice sites appear to be essential for circRNA formation in the majority of cases, with over 99% of human circRNAs flanked by the canonical splice sequence [150].

Intronic circRNAs are less common than exonic circRNAs and are formed by an alternative mechanism, however, similarly to exonic circRNAs *cis* and *trans* factors are involved in this process. *Cis* factors include consensus motifs at either end of the target introns. At the 5′ end is a 7 nucleotide GU rich element, while at the downstream branch site there is an 11 nucleotide C rich region. These motifs are thought to bring splice sites into close proximity allowing formation of an intermediate intronic lariat. This lariat is formed through a unique nucleophilic substitution leading to the formation of a 2′-5′ phosphodiester bond, followed further processing that removes any exons or overhangs to a mature ciRNA [151]. How the intermediate lariat escapes degradation is not yet known.

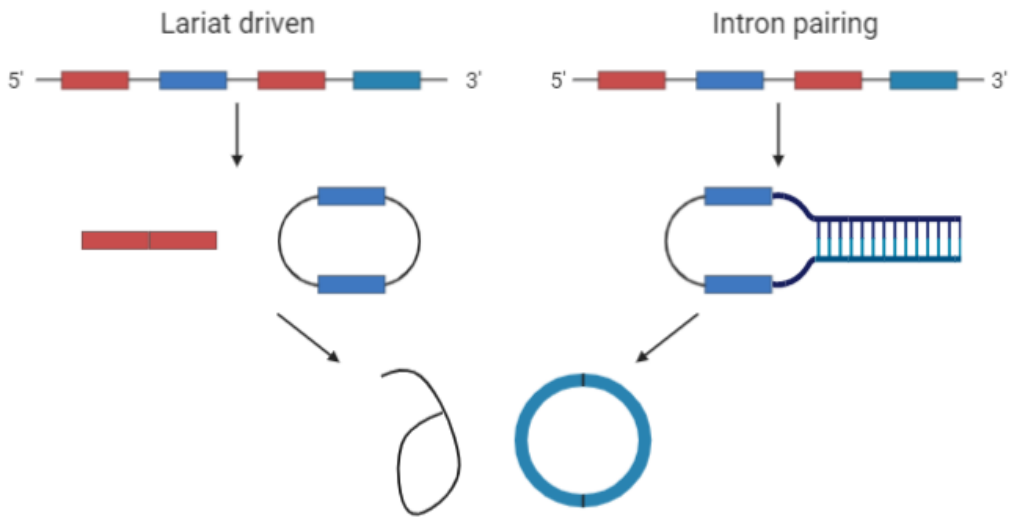

Figure 9: Mechanisms of circRNA biogenesis. The 2 main mechanisms are lariat driven circRNA biogenesis where circRNAs are formed from a lariat formed from spliced exons or intron pairing where base pairing brings the backsplice sites in close proximity.

To date it is unknown whether circRNAs are formed co- or post-transcriptionally. Supporting the co-transcriptional mode, circRNAs have been found in a nascent chromatin-bound form. Further evidence supporting this theory include a mutant RNA

pol II which has a slower elongation rate favours the formation of linear RNA over circRNAs; suggesting canonical splicing and backsplicing usually compete in the formation of circRNAs [152]. Abolition of splice sites flanking circularised exons prevents circRNA formation suggesting the role of the spliceosome in circRNA formation. However, there is evidence for circRNA biogenesis occurring post-transcriptionally, with some introns spliced out post-transcriptionally and biogenesis of circRNAs through lariat formation [148]. It is possible that circRNAs are produced both co- and post-transcriptionally, with one favoured more than the other depending on the cellular environment. Although differences in function between co and post-transcriptional circRNAs have not been established.

Therefore there is still much to be elucidated about circRNA biogenesis, additionally although they are expressed to lower levels, their stability allows circRNA levels to be maintained even with lower transcription rates, with the half-life of mRNA and circRNAs estimated to be 10 hours and over 48 hours respectively [153]. Although the vast majority of circRNAs are often found at lower expression levels than their linear transcripts, a subset have been found that are expressed at higher levels. Furthermore, levels between linear and circular transcripts are not always coupled or inversely correlated, indicating many more factors involved in biogenesis [153].

The regulation of circRNA levels is not surprisingly varied, complex and once more not fully elucidated. One mechanism of circRNA regulation is through promotion of circularisation during biogenesis, for instance by RBPs, such as Mbl. Mbl can bind to the flanking introns of its own *Mbl* gene, and can act as bridge bringing the circRNA splice sites in close proximity and promoting its own circularisation. This mechanism appears to act as a feedback loop, switching the ratio of linear: circular transcripts to favour circularisation when there is an excess of protein [154]. Other RBPs shown to promote circRNA biogenesis include QKI and Fus. Similarly, to Mbl, both bind to the flanking

introns of exons enhancing circularisation, by acting as bridge and bringing the splice sites in closer proximity [155] [156].

RBPs can also act as negative regulators of circRNAs. The RNA-editing enzyme ADAR-1 through its A→I activity disrupts base pairing of flanking introns leading to reduced biogenesis of circRNAs [157]. Whilst the enzyme DHX9 can bind to ALU regions and disrupt their secondary structure, as these regions are crucial for the formation of many circRNAs, DHX9 is a negative regulator of circRNA production. Other ALU-interacting proteins also appear to be capable of regulating circRNA levels, although the exact mechanism is not known, for instance hnRNP C and hnRNP L [158].

circRNA formation can also be regulated by the gene it is spliced from, circRNAs are more commonly formed from genes with longer than average introns, with longer introns having increased efficiency at circularisation [159]. Furthermore, circRNA formation positively correlates with an increased RNA pol II elongation speed [160]. It is hypothesised, that a faster rate of transcription increases the chance that the 5′ end is still available by the time the 3′ end is transcribed. Specific circRNAs have been shown to recruit RNA pol II and U1 snRNP to enhance RNA transcription acting as a positive feedback loop [148].

Another way circRNAs can be regulated is through their degradation and turnover, in recent years several different mechanisms of targeted cellular degradation of circRNAs have been discovered. Due to their closed circular structure, circRNAs cannot be targeted by exonucleases, however, endonuclease-mediated degradation of circRNAs appears to be highly regulated and reliant of specific co-factors. For instance in several known cases of circRNA degradation, their linear counterparts which share a nearly identical sequence are not targeted [161] [162].

RNA secondary structure may also play a role in regulating circRNA turnover. Highly structured circRNAs are targeted for degradation through structure-mediated RNA decay

(SRD). In SRD, the RBPs G3BP1 and UPF1 bind to highly structured circRNAs leading to their degradation, however, the exact mechanism is unknown. Interestingly, both G3BP1 and UPF1 localise to stress granules preventing their targeting of circRNAs during cell stress, this suggests that regulation of circRNAs is dynamic and can respond to cellular stimuli [163].

Stress granules are not the only subcellular foci used to regulate circRNA levels, with evidence suggesting that P bodies are also involved. Depletion of specific P body components in both Drosophila and human cells has led to increased cytoplasmic circRNAs, indicating they may play a role in the negative regulation of circRNA levels, potentially through aiding nuclease functionality [163].

Conditions within the cell can also lead to targeted degradation of circRNAs, for instance during viral infection. In a non-infected cell, circRNAs negatively regulate the innate immune system by forming ~16 bp imperfect duplexes of RNA which interact with PKR preventing its activation and therefore precluding the innate immune response. It is unknown why these dsRNA duplexes themselves do not activate the innate immune response, however, it is hypothesised that the secondary structure from the mismatches in the imperfect duplexes may be responsible. However, during viral infection, the enzyme RNase L is activated, which degrades circRNAs allowing PKR-mediated activation of the immune response [164]. A second mechanism of circRNA regulation by the immune response is through NF90/NF110. Normally, NF90/NF110 can bind to flanking introns of circRNAs and aid their formation, however, during infection, the NF90/NF110 complex is exported from the nucleus to enhance IL-2 mRNA stability, as a consequence this decreases NF90/NF110-mediated circRNA biogenesis [165].

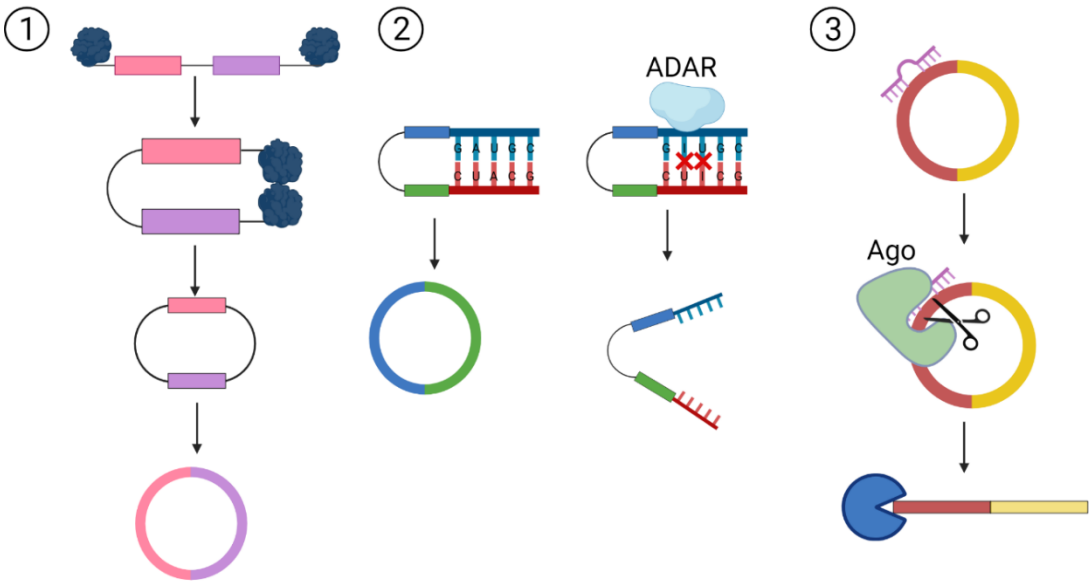

Figure 10: Regulation of circRNAs. 1. RNA binding proteins can bind to the flanking introns of circRNAs. Dimerisation of the RBPs brings the splice sites in closer proximity promoting backsplicing and enhanced biogenesis of the circRNA. 2. ADAR can negatively regulate biogenesis. Through ADAR's catalytic activity converting adenosines to inosines it can disrupt base pairing of flanking introns. These mismatch bulges push apart the splice sites reducing the efficiency of backsplicing. 3. Certain miRNAs once bound to complementary circRNAs lead to recruitment of Ago2 which leads to cutting of the circRNA. The now linear circRNA can then be targeted for degradation by exonucleases.

1.4.3 Function

circRNAs have many potential functions, with function varying on their cellular localisation and other unknown factors, some of these functions are shown in Figure 11.

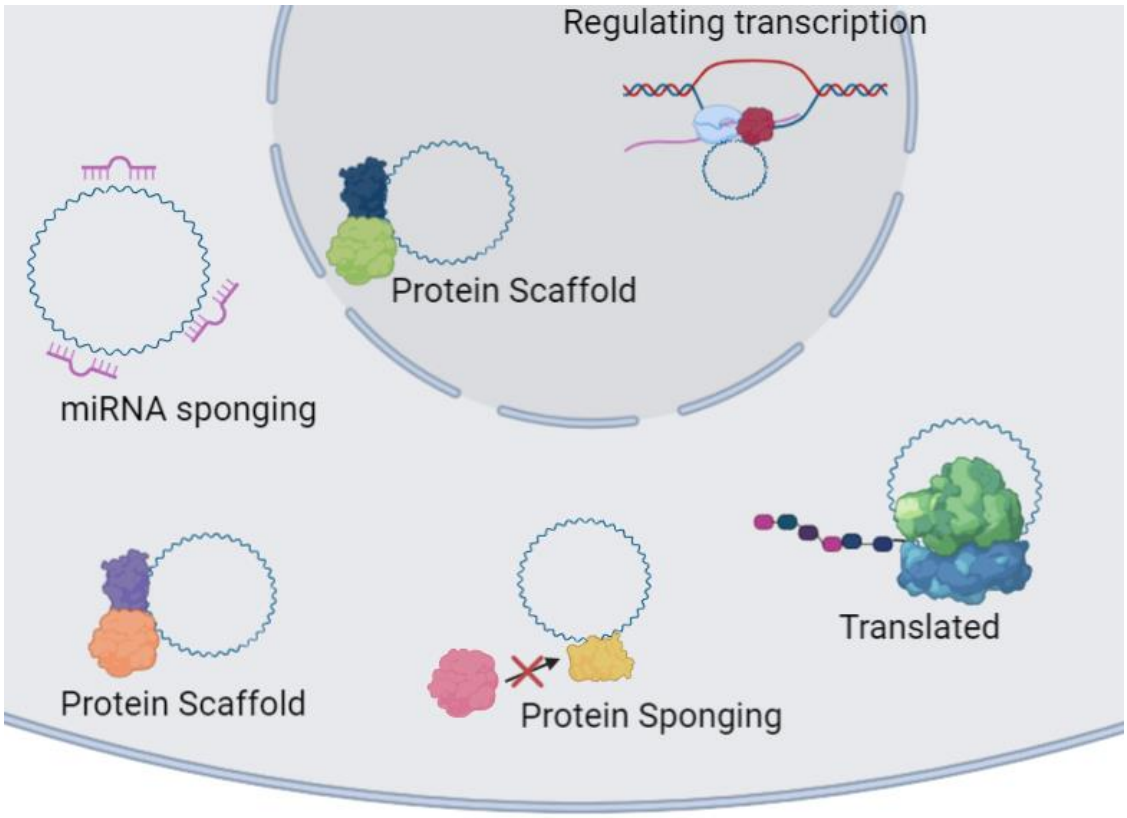

Figure 11: **The different roles of circRNAs.** circRNAs have roles in miRNA sponging, protein sponging or acting as a protein scaffold, a transcriptional regulator or potentially be translated.

1.4.3.1 miRNA sponging

A major role of circRNAs is thought to be as part of a non-coding RNA regulatory network where they function as ceRNA through sponging of miRNAs. Through this sponging, circRNAs could play a primary role in the regulation of gene expression, as such there is increasing evidence that circRNAs are dysregulated during disease. In contrast, to other ceRNAs, circRNAs may be more abundant and highly stable with many potential binding sites for miRNAs allowing efficient regulation [143]. The first circRNA identified to have miRNA sponging capabilities was in 2013, where ciRS-7 (CDR1as) was shown to be capable of sponging miR-7. Using sequence alignment, extensive base pairing particularly at the 3′ end was found between miR-7 and numerous sites along ciRS-7. Furthermore within this base pairing there were mismatches in the central region preventing miRNA-mediated degradation of the transcript [143]. Since then over 90% of ciRS-7 has been found to be composed of hexamer repeats consisting of over 70 binding sites for miR-7 [166]. Through sponging of miR-7, ciRS-7 prevents the degradation of

miR-7 mRNA targets. This mechanism was initially identified in the neuronal cells of zebrafish embryos where dysregulation of ciRS-7 leads to developmental defects through the mTOR pathway. Notably, circRS-7 is also conserved across many eukaryotes underlining its potential importance [146].

Another well characterised miRNA sponging circRNA is circSRY which contains 16 sites for miR-138 [143]. circSRY is encoded on the Y chromosome and as such appears to be testis specific, with studies failing to locate its presence in the brain, liver and kidneys. Interestingly, unlike ciRS-7, circSRY appears to be more species specific, with 16 sites on the murine circRNA but only 1 identified in human circSRY [143].

Unusually some circRNAs can regulate their own linear expression, through miRNA sponging. circITCH contains binding sites for 3 miRNAs: miR-7, miR-17 and miR-214, with all 3 miRNAs complementary to sequences in the 3′ UTR of the parental ITCH transcript [166]. Therefore, when circITCH is upregulated, it also leads to increased levels of the ITCH mRNA and consequently protein levels. ITCH is an ubiquitin ligase which targets several proteins for degradation including Dvl2, this in turn leads to a downregulation of the Wnt/β-catenin pathway. As such circITCH through its regulation of *ITCH* levels exhibits tumour suppressive properties [167].

Many of the miRNA sponging circRNAs have been identified using Ago IPs, with miRNA interacting circRNAs precipitating strongly with Ago. Additionally, overexpression of target miRNAs leads to enrichment of the circRNA within the Ago IP, therefore the target miRNA and circRNA may form a complex interacting with Ago allowing sponging of the miRNAs and a high occupancy of Ago sites [168].

However, the number of circRNAs capable of acting as ceRNAs is hotly debated, with one study only finding 12% of surveyed circRNAs contained any miRNA binding sites or an Ago footprinting site. Furthermore, most circRNAs with miRNA binding sites only

contain 1-2 compared to the 70+ of ciRS-7, and therefore whether they are functionally active is unknown [169].

1.4.3.2 Regulation of gene expression

In contrast to exon-containing circRNAs, intronic circRNAs localise to the nucleus, as such they are thought to have roles in regulating transcription and gene expression, although potential mechanisms are poorly understood. One possible mechanism postulated is by the regulation of parental transcript expression in *cis.* This is supported by depletion of 2 EIcircRNAs: circEIF3J and circPAIP2 which led to decreases in expression of their corresponding linear transcripts, however, depletion of the linear transcript had no effect on transcription of itself or the circular transcript. Additionally, RNA-DNA FISH showed that the circRNAs localise to their parental gene loci in around 50% of cases which was confirmed by CHIP analysis, although this suggests they may have other *trans* regulatory roles. Further experiments using pulldowns showed that these EIcircRNAs associate with members of the transcriptional machinery which may explain their specific role, namely RNA pol II, U1A, U1C and U2 snRNA. Therefore, one mechanism proposed is EIcircRNAs interact with U1 snRNP around 300 base pairs upstream of the transcriptional start site, within the promoter regions to aid in transcription through RNA pol II [148]. This positive in *cis* regulation has also been shown for other intronic circRNAs, for instance circANKRD52 associates with positive elongation factors and therefore positively regulates RNA pol II [170].

Alternatively, certain circRNAs may negatively regulate expression of their linear transcript through a mechanism called mRNA trap. Here it is hypothesised that the backsplicing and formation of circRNAs competes with the formation of their linear transcript, thereby increasing circRNA formation leading to truncated or non-functional linear transcripts. This occurs when the linear translational start codon is in the circRNA, preventing normal translation [171] [172]. How widespread this phenomenon is, is not known, although it could potentially regulate many genes [173].

1.4.3.3 Interactions with proteins

Another function circRNAs may perform is acting as either a protein scaffold or a protein sponge/decoy. Here through binding to the relevant protein they can aid its function by bringing in cofactors or can inhibit it through preventing protein-protein interactions. This can occur in both the nucleus and the cytoplasm and as such can involve both intronic and exonic circRNAs.

One of the best studied circRNAs with numerous characterised interactions with proteins is circFOXO3, and due to these protein interactions it is often classed as a tumour suppressor. One such interaction is with the protein complex formed between FOXO3 and p53. Initially, circFOXO3 appears to be anti-apoptotic, it binds to p53 and induces its ubiquitination and subsequent degradation by MDM2. However, this is balanced by its specific interaction with FOXO3. Here it binds to FOXO3 and blocks the ubiquitination by MDM2, this increase in FOXO3 levels leads to an increase in the downstream PUMA protein. PUMA can then interact with the anti-apoptotic Bcl-2 family members, inhibiting their function and allowing the release of Bax and other downstream targets which induce apoptosis [174]. Further examples of circRNA-proteins are shown in Table 2.

Name	Function	Reference
circPABPN1	Binds to HuR protein preventing its interaction and stabilisation of PABPN1 mRNA. Decreased PABPN1 lead to reduced poly(A) formation.	[175]
circANRIL	Binds to PES1 preventing its role in ribosomal biogenesis	[176]
circDNMT1	Nuclear translocation of p53 and AUF1, leading to degradation of p53. Nuclear AUF1 leads to increased DNMT1 mRNA stability and therefore increased DNMT1 levels	[177]
circMTO1	Binds to TRAF4 and acts as a sponge preventing it from leading to increased kinesin 5 levels	[178]
circAMOT11	Binds to AKT1, leading to its phosphorylation and nuclear translocation, leading to activation of PI3K/AKT pathway	[179]

Table 2: Examples of known circRNA functions.

1.4.3.4 Translation

Although circRNAs are classified as ncRNAs, there is some evidence that under certain conditions specific circRNAs can be translated. The first evidence that circular transcripts

can be translated was early in the study of circRNAs, with the identification that the circular genome of HDV and other plant viroids could be translated [138]. Additional evidence of this proof-of-principle came with artificial cellular circRNA constructs with internal ribosomal entry sites (IRES) in the 1990s, demonstrating that cellular ribosomes do not require a free 5′ end or cap to initiate translation [180]. More recent research has shown that backsplicing allows the efficient translation of a split GFP gene in an artificial construct [181]. Interestingly there is evidence in both human and *E. coli* artificial systems that circular transcripts may also be capable of continuous protein translation through the rolling circle translation [182]. However, it must be noted that much of the evidence for the capability of circRNAs to be translated are provided by artificial systems, with few *in vivo* studies, in part due to the difficulty in distinguishing between linear and circular transcripts. Most experiments use ribosomal profiling to identify whether a circular transcript associated with the ribosomes, although this still fails to distinguish if a transcript is actively translated and relies on the ribosome stalling on the unique splice site.

However, even with these difficulties, there are several studies that have suggested *in vivo* translatable circRNAs. For instance, the *Mbl* locus produces several abundant circRNAs which contain a putative start codon and UTR regulatory sequences, focusing on circMbl3, one group found a 37 kDa protein that could have only been derived from the circular form with its unique backsplice [183]. Another circRNA proposed to be capable of translation is derived from *ZNF690*, once again the circRNA contained an in-frame start codon, and was found associated with polysomes which resulted in a unique protein product [184].

Furthermore, other studies using TrIP-Seq have shown that over 1/5 circRNAs examined were associated with polysomal fractions, suggesting the possibility of translation, although none of corresponding proteins were identified [169].

However, the translation of circRNAs is still a contentious area with more research needed. Most research focuses on proof-of-principle in artificial systems with relatively few *in vivo* studies and no functionality analysis of any protein products. Finally, the only examples of circRNAs being translated are in brain and muscle cells, these cells express exonic circRNAs to high levels, suggesting high abundance of circRNAs may lead to some translation events. Overall, the vast majority of circRNAs appear not to be translated.

1.4.5 Virus-encoded circRNAs

It has recently been discovered that several viruses encode circRNAs, similarly to virally-encoded miRNAs, the vast majority appear in DNA viruses, with the gammaherpesviruses sub-family particularly adept at utilising and encoding their own ncRNA species. Numerous studies have identified multiple KSHV-encoded circRNAs. In one study, KSHV-encoded circRNA expression was induced by lytic replication, which contrasts with KSHV-encoded miRNA expression which is predominately in the latent phase. To date these KSHV-circRNAs are not functionally characterised, however, ectopic expression of kcirc55 and kcirc97 both lead to increased cell growth in infected B and endothelial cells. Furthermore, sequence analysis to identify potential miRNA binding sites within these KSHV-encoded circRNAs, showed enrichment for miRNAs involved in regulating the cell cycle. Additionally, analysis of RBP sites within the KSHV-encoded circRNAs also identified regulators of the ncRNA machinery, namely hnRNP A1, QKI and FUS, suggesting these viral circRNAs may have a role in regulating many aspects of ncRNA biogenesis and function. Finally, KSHV-encoded circRNAs may also regulate the virus life-cycle directly, with miRNA binding sites found complementary to viral encoded-miRNAs and ectopic expression of kcirc55 and kcirc97 repressing RTA expression [185].

In contrast to kcirc55 and kcirc97, a circRNA encoded from the KSHV vIRF4 locus is highly expressed in latently infected KSHV cells, as well as in some KS tumour samples,

although, once again the function is not yet known. Unexpectedly, and in contrast to cellular circRNAs, the intron-containing circVIRF4 actually localises to the cytoplasm. Furthermore, numerous potential circRNAs have been found derived from other regions of the KSHV genome including PAN, miRNA locus, K4, K12 and ORF71, although whether they are functionally relevant has yet to be ascertained. In particular, PAN is highly expressed and therefore any circRNAs may be artefacts from processing or the high copy number [186]. Notably, there is also evidence that KSHV-encoded circRNAs may be incorporated into mature virions, suggesting they may have a role in preparing the cellular environment prior to virus replication [187].

KSHV is therefore thought to encode many low expressed circRNAs and one abundant circRNA, circvIRF4. Interestingly the closely related gammaherpesvirus MHV68 also encodes a similar pattern of circRNAs, with the MHV68 ori-Lyt circRNA highly abundant. In regards to both the KSHV and MHV68 low expressed circRNAs, more evidence is required to determine whether or not they are splicing artefacts, as the majority are clustered around lytic origins of replication, suggesting the complex transcriptional and processing mechanisms around these areas could lead to these circRNAs being generated as by-products [186] [185].

EBV also encodes its own group of circRNAs, with several originating from the BART locus. These circRNAs are expressed in EBV-associated tumours, transformed cells as well as type I and type III latently infected cells, suggesting they may have a role in either maintenance of the viral genomes or in viral-mediated oncogenesis [186]. The closely related rLCV has also been shown to encode circRNAs. Similarly, to KSHV and MHV68, both rLCV and EBV have a cluster of lowly expressed circRNAs around their ori-Lyt, once more raising the possibility that these are general artefacts.

However, there is evidence for several of these herpesvirus-encoded circRNAs having functional significance. The more abundant circRNAs for EBV, KSHV, rLCV and MHV68 all have novel splice acceptor sites that are only used for the formation of the backsplice

and not for any other form of linear splicing. These novel backsplice sites are often within 8 nucleotides of the linear splice sites, however, there is no other function ascertained to them. Unless, the circRNAs formed from these sites are functionally relevant, it would seem to be biologically wasteful to contain such sequences [186] [185].

Additionally, there is arbitrary evidence for a circRNA derived from the HPV E7 oncogene circE7, detected via qPCR and northern blotting. This circRNA is lowly abundant and functionally not elucidated, although knockdown of the circE7 did disrupt HPV-mediated transformation [188]. However, other studies have since stated this circRNA is expressed at too low levels to be functional and any observations are from off-target effects [189].

Recently, surveys have been performed for large-scale identification of virus-encoded circRNAs, one group found nearly 12,000 circRNAs across 23 viruses. Although, the majority, almost 70%, were found in herpesviruses, circRNAs were also identified encoded by RNA viruses including Influenza virus, Zika virus, HIV and Ebola virus. The majority of these circRNAs either contained or were flanked by highly complementary regions or repeating sequences. Once again, there is little functional analysis for any of these circRNAs, with the potential to be artefacts, missplices or false positives. Further analysis and validation is needed [190].

1.4.6 Viral dysregulation of cellular circRNAs

Although not as well studied as dysregulation of miRNAs, several examples of cellular circRNA dysregulation by viruses have been discovered. One of the first viruses shown to dysregulate circRNAs during infection was SV40 polyomavirus, with studies showing at least 134 dysregulated circRNAs during infection, of which 103 were upregulated. Many of these circRNAs are thought to be involved in the regulation of crucial pathways including p53, Wnt and innate immunity regulation through the TLR and JAK pathways [191]. Avian leukosis virus subgroup J (ALV-J) was found to upregulate circVav3. This was initially found due to the downregulation of gga-miR-373 and resulting upregulation of its mRNA target *YAP1* in liver tumour cells. gga-miR-375 downregulation is due to the

sponging activity of circVav3, confirmed by FISH, RIPs and biotinylated pulldowns [192]. Additionally, other studies have found differences in the global circRNA profiles between ALV-J resistant and susceptible chickens, with evidence suggesting these differences may lead to an oncogenic or tumour repressive environment [193].

Dysregulated cellular ncRNA networks have been identified to be caused by a range of viruses, not just oncogenic DNA viruses, including HIV (500 circRNAs, 21 miRNAs and 900 mRNAs), Ebola virus (125 circRNAs, 317 miRNAs and 1911 mRNAs) and HSV-1 (500 circRNAs and 207 miRNAs) [194] [195]. The roles of these dysregulated networks have not been fully elucidated to date, however, initial studies suggest important roles during infection, for instance, of the 1300 dysregulated circRNAs in early HIV infection, through gene ontology, 67 were found to potentially have roles in anti-viral immunity and the host cell response. Of these 67, 3 circRNAs are involved in a complex ncRNA regulatory network. Firstly, a circRNA/miR-524 and miR-101/CDKN1A network regulates the levels of the anti-viral protein CDKN1A. Secondly a circRNA/miR-27b/CCKN network, again targets CCKN due to its negative regulatory activity against HIV, through its inhibition of the Nef protein. Finally, HIV infection led to a decrease in 18 circRNAs that target miR-27b [196].

Another global investigation into dysregulated circRNAs during virus infection was performed in Hantaan orthohantavirus (HTNV)-infected cells, this once again showed that dysregulation of circRNAs by viruses to regulate ncRNA networks appears to be a conserved mechanism with 70 dysregulated circRNAs, 66 miRNAs and 788 mRNAs. Gene ontology analysis showed many of the mRNAs clustered in innate immunity pathways, including IFN-related and cytokine signalling pathways [197]. Therefore, although research into virus-mediated dysregulation of cellular circRNAs is still in its infancy, clearly it is an important mechanism for virus propagation and survival.

1.5 Paraspeckles

There is still much unknown about the regulation of circRNA biogenesis and thus viral dysregulation of circRNA biogenesis, however, there is emerging evidence that nuclear organelles and structures can play a role in their formation and regulation [198] [199]. The nucleus is composed of distinct structures, including the nuclear membrane, chromatin and nucleolus, with each having a distinct role and composition. There are also many membraneless nuclear organelles, including cajal bodies, PML bodies, nuclear speckles, PIKA bodies and paraspeckles. These can have roles in snRNA processing, DNA damage response, splicing, DNA replication and RNA processing respectively [200].

1.5.1 Discovery and formation

Paraspeckles (PS) are membrane-less nuclear bodies, primarily composed around the lncRNA NEAT1 which associates with core paraspeckle proteins including but not limited to SFPQ, NONO, PSPC1 and FUS. NEAT1 has 2 separate transcripts: a shorter polyadenylated NEAT1_1 and the longer NEAT1_2 which forms the scaffold for paraspeckles. Paraspeckles were discovered relatively recently in 2002 during a nuclear mass spectrometry screen which identified several novel proteins, including the now known paraspeckle protein PSPC1, [201]. This screen was followed up with a more in-depth analysis showing PSPC1 is localised diffusely throughout the nucleus but can also form 5-20 denser foci. Within these foci, PSPC1 was observed to associate with two more RNA binding proteins: SFPQ and NONO. In addition, these foci were shown to be distinct, not localising with any other known sub-nuclear compartment including cajal bodies, PML bodies and the nucleolus, and as such these new foci were named paraspeckles [202].

Paraspeckle biogenesis is broadly split into 2 stages, with the first stage occurring co-transcriptionally. Without association of paraspeckle proteins, NEAT1 is fairly unstable with a half-life of approximately 1 hour and is soon degraded, as such paraspeckle

formation is closely tied to active transcription [203]. To prioritise the longer NEAT1_2 transcript that is the essential for paraspeckle formation, the 3′ processing for NEAT1_1 must be inhibited. This occurs through the essential paraspeckle-associated protein (PSP) hnRNP K [204]. hnRNP K binds to NUDT21, a member of the mRNA 3′ cleavage and polyadenylation complex and competes with its binding to CPSF6. This prevents the formation of the CFIm complex, which usually aids in 3′ end cleavage leading to transcription of the longer NEAT1_2 transcript [205]. As NEAT1 is transcribed SFPQ, NONO and RBM14 bind along its length, leading to its stabilisation. SFPQ and NONO through coiled coil domains can then oligomerise forming bunched chain bundles which comprise the basis for paraspeckles [206] [207]. RBM14, SFPQ and NONO are therefore considered the essential proteins for the first stage of paraspeckle biogenesis [206]. This formation is thought to be driven by liquid-liquid phase separation. Here proteins, often with low complexity domains, start to form weak interactions leading to the formation of a highly dynamic membraneless compartment [208]. NEAT1 acts as a scaffold, enabling PS proteins to assemble in this highly concentrated manner.

The latter phase of paraspeckle biogenesis is believed to occur through the PSPs already bound to NEAT1, being able to recruit other PSPs when in close proximity. These interactions can again occur through low complexity regions, with many of the PSPs having prion-like-domains (PLDs) which help in aggregation [209]. The SWI/SNF chromatin remodelling complex is essential for this second phase, by driving the association between core and newly recruited PSPs. SMARCA4, a key member of this complex, is essential for paraspeckle formation, as knockdown inhibits paraspeckle formation. SMARCA4, along with its co-factors BAF47, BAF57, BAF155 and BAF170 also localise to paraspeckles. Interestingly, this localisation and interaction occur even in NEAT1 KD cells, implying phase 2 interactions are independent of phase 1 [210]. This is summarised in Figure 12.

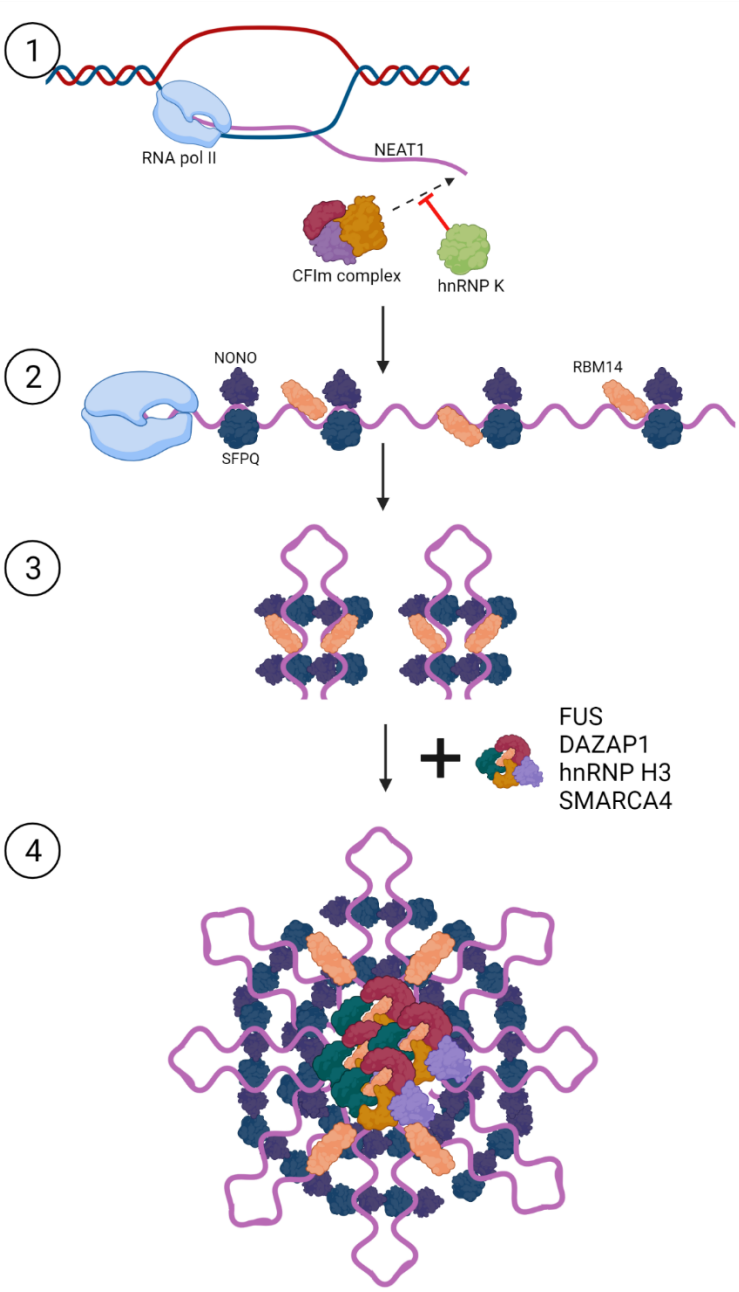

Figure 12: The biogenesis of paraspeckles*:* (1) NEAT1 is transcribed by RNA pol II. hnRNP K binds to the CFIm complex preventing its polyadenylation of NEAT1_1, allowing transcription of the full length NEAT1_2. (2) SFPQ, NONO and RBM14 bind to actively transcribed NEAT1_2 stabilising it. SFPQ and NONO can bind to each other. (3) Through dimerisation and oligomerisation of SFPQ and NONO through coiled coil domains across different NEAT transcripts, NEAT1 starts to bundle and aggregate, leading to liquid: liquid phase interactions (4) Here it is thought that bound proteins to NEAT interact and recruit the final components of paraspeckles including FUS, DAZAP1, hnRNP H3 and SMARCA4 leading to formation of the mature paraspeckle.

1.5.2 Components

Paraspeckle components vary dependent on cell type and condition, with 40-60 different PSPs identified [211] [212], however, the core group of paraspeckle proteins number around 17 proteins [206]. Knockout studies have shown that 8 of these proteins: SFPQ,

NONO, FUS, DAZAP1, hnRNP H3, hnRNP K, RMB14 and SMARCA4, are essential for paraspeckle formation, along with the ncRNA NEAT1 [205].

Advancements in microscopy and EM have shown paraspeckles are not solid homogenous masses but have distinct architecture and sub-organisation. The central core of a paraspeckle comprises of SFPQ, FUS, NONO, PSPC1 and the central region of NEAT1, whilst the shell contains both the 5′ and 3′ ends of NEAT1. The shell and the central core are then interconnected via the essential paraspeckle proteins RBM14 and SMARCA4 [206], summarised in Figure 13.

Additionally, all essential PSPs except hnRNP K contain a prion like domain (PLD), which aid in aggregation of proteins.

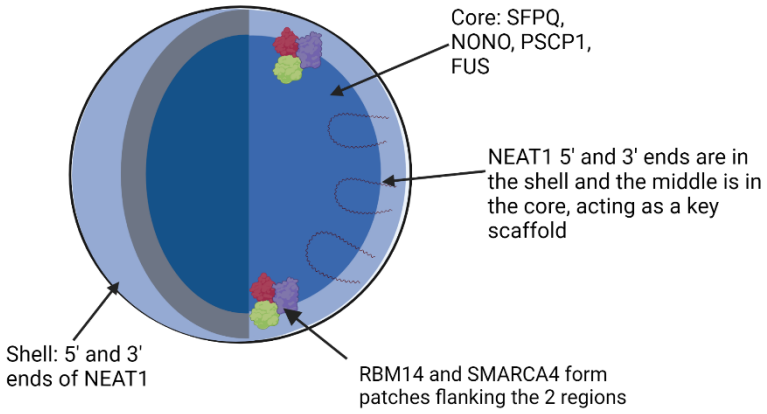

Figure 13: Structure of a paraspeckle. Proposed structure of a paraspeckle highlighting the core and shell.

1.5.3 Function and role

The function of paraspeckles is still not well elucidated, and appears variable dependent on cell conditions and also the composition of the paraspeckles. Another layer of complexity, is many of the PSPs carry out additional roles in the cell, outside of the paraspeckles.

1.5.3.1 RNA retention and processing

One attributed role is nuclear retention of RNAs under specific conditions. In particular, numerous studies have found mRNAs with long inverted repeats in the 3′ UTR (IR-mRNAs) seem to be retained by paraspeckles [213] [214]. Other types of RNA found include mRNAs with purine rich intron sequences and the U1 RNA [215]. Interestingly protein modifications of paraspeckle components have also been shown to play a role in mRNA retention. For instance methylation of NONO by CARM1 reduces its affinity for ALU containing RNA, in contrast, methylation of SFPQ by PRMT1 increases its affinity for ALU containing RNA [216].

1.5.3.2 Transcriptional regulation

SFPQ can act as a transcriptional repressor while present at promoters; when SFPQ Is recruited into paraspeckles the gene can then be transcribed, allowing dynamic regulation of gene expression. For instance SFPQ can bind to the IL-8 promoter, repressing its transcription, but when recruited to paraspeckles, this repression ends and IL-8 is transcribed [217]. Interestingly viral infection, including HSV and influenza, can induce NEAT1 transcription as part of the immune response. Once transcribed, through sequestering of SFPQ, it leads to increases IL-8 transcription and therefore an immune response [218].

SFPQ can also act as a co-repressor for members of the type II nuclear hormone receptor family, which bind to their own promoters leading to transcription. However, SFPQ along with NONO can bind and then recruit Sin3A, leading to recruitment of histone deacetylases resulting in chromatin condensation and loss of transcription [219] [220].

Additionally SFPQ is pro-transcriptional for several genes including ADARB2. SFPQ and NONO form a bridge between RNA pol II and promoters increasing ADARB2 transcription [221][222]. Therefore when SFPQ is sequestered into paraspeckles by NEAT1, ADARB2 transcription is reduced [223].

SFPQ also promotes the transcription of longer genes. It binds to longer introns and through recruitment of CDK9 enhances the interaction with RNA pol II, aiding in long gene transcription [224]. Moreover, SFPQ has been proposed to act like a histone for RNA, with its oligomerisation aiding in CDK9 recruitment for phosphorylation of RNA pol II and therefore elongation, as well as aiding pre-mRNA stability and processing [224] [225].

1.5.3.3 circRNAs and paraspeckles

Currently there have been no specific studies investigating circRNAs involvement or localisation with paraspeckles, however, due to the complex role of paraspeckles in RNA processing and splicing, there is the potential for them to play a role in circRNA biogenesis and regulation. For instance, included in the group of IR-mRNAs that are often retained in paraspeckles, are mRNAs with ALU repeats. Evidence suggests they are poorly exported to the cytoplasm and often further processed in the nucleus [216], this is particularly interesting due to ALU-containing RNAs often being processed into circRNAs. Over 80% of all known circRNAs have ALU repeats in their flanking introns which are thought to drive circularisation through complementary base pairing [226]. Additionally several circRNA biogenesis factors are known PSPs, including the core protein FUS, whose depletion negatively effects the levels of several specific circRNAs [227]. FUS is thought to preferentially bind to RNAs with long flanking introns to drive their circularisation, incidentally these are the RNAs that are often retained by paraspeckles; as such paraspeckles may be the perfect environment for circRNA biogenesis due to high levels of RBPs and enrichment of the necessary RNAs. Finally, SFPQ and NONO have recently been shown to be enriched around circRNA loci, specifically circRNAs with long flanking introns containing ALU repeats [198]. In particular, SFPQ is essential in maintaining correct splicing of long introns, with depletion leading to increases in mis-splicing and decreases in circRNA levels that have long flanking introns [198].

1.5.3.4 Splicing

The exact role of SFPQ in splicing is not currently known, however, it is likely to be involved in alternative splicing. Canonical splicing is co-transcriptional and involves recruitment of U1 and U2 units to the 5′ splice site and the branch site respectively [228]. Splicing requires quick recycling of these spliceosomal subunits to maintain a constant reservoir of factors. Whereas, alternative splicing is regulated through *cis* elements including exonic splicing enhancers (ESEs), intronic splicing enhancers (ISEs) or silencing elements. These elements interact with *trans* factors, particularly the SR and the hnRNP family of proteins, with hnRNPs leading to either exon inclusion or exclusion [229] . SFPQ has been found to promote both exon inclusion and exclusion, dependent on its co-factors and interactions. For instance, SFPQ interacts with the alternative splicing factor, FOX3, recruiting it to a motif that promotes inclusion of exon 30 of NMHC) II-B [230]. In contrast, SFPQ can also lead to exon inclusion, it has been shown to interact with an exon splicing silencer (ESS) on the *CD45* gene. Through this interaction it negates the activity of ESS-bound hnRNP L, modifying the exons skipped on *CD45* from a resting T cell to an activated T cell phenotype [231].

1.5.3.5 Other roles

The SFPQ/NONO dimer has also been implicated in the DNA damage response, stimulating double stranded DNA break repair. The dimer can bind to RAD51 and directs RAD51 to repair double stranded breaks. Moreover SFPQ/NONO promote non-homologous end-joining (NHEJ) through binding to Matrin 3, though the mechanism is not fully known [232] [233] [234].

Several PS components have also been suggested to play a role in regulating miRNA processing through enhancement of pri-miRNA slicing. The SFPQ/NONO dimer binds to the miRNA machinery proteins Drosha and DGCR8. NEAT1 has a pseudo miRNA at its 3′ end which is also thought to aid in the recruitment of the machinery [235]. It has been hypothesised that spatial arrangements also play a role in pri-miRNA processing, with

80% of pri-miRNA genes located in long introns which are often enriched in paraspeckles, aiding in processing [236]. Studies have found that KD of NONO, NEAT1 or SFPQ led to loss of between 64-80% of 500 cellular miRNAs monitored [235].

Finally, the SFPQ/NONO dimer may also recruit the exonuclease XRN2 to promote RNA 3′ processing and termination, facilitating pre-mRNA 3′ processing [237].

1.5.4 Dysregulation in disease

1.5.4.1 Dysregulation in viral infection

Paraspeckle components and SFPQ in particular have been shown to play a role in the lifecycles of multiple viruses; and dependent on the virus can be either inhibitory or beneficial for the virus.

NEAT1 is involved in the immune response to viral infection, with it first being identified in mouse brains infected with either rabies virus or Japanese encephalitis virus [238]. Other studies have shown that in response to viral infection or dsRNA NEAT1 sequesters SFPQ away from the promoters it usually represses. This enhances the transcription of anti-viral genes including IL-8 and CLL5 [218]. Additionally, NEAT1 has been shown to form a ribonuclear complex with several paraspeckle proteins, DNA PK proteins and the and the RBP protein HEXIM1 in response to viral infection. This complex can interact with cGAS, aiding its activation of STING and IRF3 leading to immune activation [239].

During Hepatitis δ Virus (HDV) infection, SFPQ, NONO and PSCP1 co-localise to sites of HDV genome replication and are essential for successful replication. Additionally, PSCP1 partly re-localises to cytoplasmic foci, hypothesised to be part of a cellular stress response and NEAT1 levels were upregulated during HDV infection leading to larger nuclear foci [240]. The exact role of these changes is as of yet unknown, however, the authors hypothesise the upregulation of NEAT1 could be linked to HDV-associated oncogenesis, due to its previously elucidated oncogenic functions or as part of a cellular immune response as previously described.

In HSV-1 infection, speckles form comprising of SRSF2, NEAT1, PSPC1 and NONO, simultaneously NEAT1 can act as a scaffold for SRSF2 in regulating histone modifications near viral genes [241]. Additionally HSV-1 leads to increases in NEAT1 expression through STAT3 activity which can bind to the NEAT1 promoter, moreover STAT3 also localises to paraspeckles. Through binding to NEAT1, PSCP1 and NONO, viral genes are retained in paraspeckles where STAT3 can promote their transcription. Interestingly SFPQ however, is inhibitory to HSV-1, through competing for STAT3 binding [242].

Whereas, in influenza infection SFPQ is pro-viral and required for efficient viral replication. Its role is not however involved in viral splicing but in regulating transcription of viral genes, specifically aiding in polyadenylation [243].

Virus infection can also utilise SFPQ through novel mechanisms, for instance Encephalomyocarditis virus infection results in translocation of SFPQ from the nucleus to the cytoplasm. Once in the cytoplasm SFPQ associates with eukaryotic initiation factors and ribosomal proteins promoting viral IRES mediated translation and as such is essential for viral replication [244].

1.5.4.2 Dysregulation in other diseases

SFPQ dysregulation can also be linked to diseases such as cancer, through its role as a transcription factor for several pro-apoptotic genes such as Bax and Bcl-2. For instance in CML cases, survival can correlate with downregulated NEAT1 expression, this prevents sequestering of SFPQ, leading to transcription of its target genes and therefore increased apoptosis in cancer cells [245]. Overall however, there is evidence for SFPQ, NEAT1 and other paraspeckle components acting as both tumour suppressors or as oncogenes. NEAT1 is upregulated in many cancers and can be associated with a poor prognosis, with both RUNX and HIF proteins shown to induce its expression [246]. In contrast, NEAT1 is a known regulator of the p53 pathway, with a p53 binding motif within NEAT1, and p53 activating NEAT1 expression. It is thought NEAT1 aids p53 in its

function, with p53 effectiveness reduced in NEAT1 KD cells [247]. Overall NEAT1 and SPFQ functionality in cancer is likely to be variable and cell-type dependent.

Paraspeckle components have also been investigated in neurodegenerative diseases. In Amyotrophic lateral sclerosis (ALS) or frontotemporal lobar degeneration (FLD), of the 25 proteins often associated, 8 are paraspeckle proteins, including SFPQ and FUS, with their mutations often driving aggregation [248] [249] [250]. Additionally, in ALS neurones, enhanced paraspeckle formation has been observed. Interestingly this is not linked to an increase in NEAT1 transcription but rather a loss of the negative regulator TDP-43. These hyper paraspeckles are hypothesised to be abnormally stable and contribute to stress signalling and attenuated cytoprotective responses [251].

1.6 Thesis Aims

Non-coding RNAs are emerging as critical regulators of gene expression and as such are often dysregulated in viral replication. This thesis focuses on an understudied species of ncRNA, namely circular RNAs (circRNAs), which are formed through a unique backsplice mechanism.

The initial aim of this project was to investigate the role of cellular circRNAs during KSHV infection. KSHV had previously been found to dysregulate to cellular miRNAs, with the majority downregulated, however, any mechanisms were unknown. It was hypothesised that one potential mechanism could be through circRNAs acting as miRNA sponges. circRNAs, as a rapidly emerging field of molecular biology, had been demonstrated to be capable of miRNA sponging, however, at the beginning of this thesis, there had been no characterisation of cellular circRNA dysregulation by viruses. Thus, initially this thesis aimed to identify if KSHV lytic replication led to circRNA dysregulation, as observed with other ncRNA species. Secondly, whether these dysregulated circRNAs could form part of a ncRNA regulatory network through sponging of miRNAs.

Like all herpesviruses, KSHV has two distinct phases of its life cycle: lytic replication and latency. During the lytic phase, KSHV dysregulates host cell ncRNA species, such as miRNAs, to modulate host cell gene expression thereby enhancing infectious virion production. In chapters 3 and 4, results demonstrate that circHIPK3 is dysregulated during lytic replication and that this dysregulation is important for viral replication. Furthermore, studies highlight that this dysregulated circRNA acts as ceRNA, through sponging of miR-30c forming a non-coding RNA axis allowing KSHV to specifically affect the levels of the target mRNA DLL4. Notably, all aspects of this network are shown to contribute to successful KSHV lytic replication.

Following this, chapter 5 aimed to elucidate mechanisms by which KSHV could dysregulate circRNA biogenesis, leading to the characterisation of the paraspeckle protein SFPQ. Results showed that SFPQ regulates specific circRNAs during lytic

replication, however, results expand the functionality of SFPQ, suggesting it is hijacked by KSHV for further roles in viral and cellular RNA processing. Additionally, studies show that SFPQ forms novel condensates adjacent to VRCs during lytic replication. Notably, these foci are far larger than canonical paraspeckles and also contain the uncharacterised KSHV-encoded ORF11 protein.

Chapter 2

~

Materials and methods

2.0 Materials and methods

2.1 Oligonucleotides

Oligonucleotides for qPCR primers, shRNA sequences and site directed mutagenesis primers were ordered from IDT. Sequences are shown in the tables below.

Gene Name	Sequence
BTG1	F- AGCTGCTGGCAGAACATTAT
	R- CAACGGTAACCCGATCCCTT
CCNB1	F- CATGGTGCACTTTCCTCCTT
	R- AGGTAATGTTGTAGAGTTGGTGTCC
CCNE1	F- CCACACCTGACAAAGAAGATCATCAC
	R- GAGCCTCTGGATGGTGCAATA
CDYL	F- CGAGGAGCTGTACGAGGTTG
	R- ACGAGTGCCTAAGGAGAGGT
circCYDL	F- ACCCACTAGTGCCTCAGGTG
	R- CTCGCTGTCATGCCTTTCC
circHIPK3	F- TATGTTGGTGGATCCTGTTCGGCA
	R- TGGTGGGTAGACCAAGAGTGGTGA
circITCH	F- GCAGAGGCCACCACTGGA
	R- TCCTTGAAGCTACTACGCTGAG
circPVT1	F- TTCAGCACTCTGGACGGACTT
	R- TATGGCATGGGCAGGGTAG
circSPECC1	F- GAGAGCTGCGAAGTTCAAGA
	R- GCCTGTCCGTTTAGTTGTTGT
DLL4	F- CCCTGGCAATGTACTTGTGAT
	R- TGGTGGGTGCAGTAGTTG
GAPDH	F- TGTCAGTGGTGGACCTGA
	R- GTGGTCCTTGAGGGCAATG
HIPK3	F- GGCATCAAGAGTGGAATGGA
	R- TGATGAATGGTTGGGGATGG
K8	F- AGGACCACACATTTCGCAAC
	R- ACCCCTTGTCAGTTCTTC
K8.1	F- GTTCCACACAGATTCGCACA
	R- AGTTCATCCTGCCTAGCCAG
kcircPAN	F- ACCAGACGGCAAGGTTTTA
	R- TCGTTAGTCAACCTAGCAAAACA
kcircvIRF4	F- CTCCGTGTGGATACCAGTGA
	R- TGGTCCCACGCAACAGTCT
MALAT1	F- GAATTGCGTCATTTAAAGCCTAGT
	R- GTTTCATCCTACCACTCCCAATTAAT
Myosin HC Exon 30	F- GCAGTCCCAGTTGGCTGATA
	R- TGTTGACCTCCAGACGAAGC
NEAT1	F- AGTACCCTGAGAGCCAGTATTGGT
	R- GGCAGCTGAGTCAATCTCCTTT
NONO	F- TGATGAAGAGGGACTTCCAGA
	R- AGCGCATGGCATATTCATACT
ORF4 V1	F- CGATTTGTGCACGGAAGA
	R- GAG GGAGTGTTGGTTCTCGC

ORF4 V2	F- CGTTTGCATCCAACACCCAAT
	R- TGTTGGTTCTCGCGGTCTC
ORF4 V3	F- TCCTGC CAACATCCGAAGG
	R- GAGGGAGTGTTGGTTCTCGC
ORF45	F- TACCAGAGGAGGCGGTAGAC
	R- TCCGAACGTGAACAATGAGG
ORF45	F- TACCAGAGGAGGCGGTAGAC
	R- TCCGAACGTGAACAATGAGG
ORF57	F- GCCATAATCAAGCGTACTGG
	R- GCAGAGAAATATTGCGGTGT
ORF59	F- GGTCCGGATATGCTCCTAGTT
	R- CAGCATGCTCACGAGGAATA
18S	F- GATGGTAGTCGCCGTGCC
	R- GCC TGCTGCCTTCTTTGG
ORF65	F- AAGGTGAGAGACCCCGTGAT
	R- TCCAGGGTATTCATGCGAGC
PAN	F- ATAGGCGACAAAGTGAGGTGGCAT
	R- TAACATTGAAAGAGCGCTCCCAGC
Pre/pri-miR-30c	F- GTGTAAACATCCTACACTCTCAGC
	R- TGGCAGAAGGAGTAAACAACCC
Pre/pri-miR-29b	F- TTGCCACTTGAGTCTGTT
	R- CTTCTCTACTGTCACCTCTC
pre-BTG1	F- GTCACCGGCACAATTAACAG
	R- TGCACACAATGGAGTTGATG
pri-miR-29b	F- TGAAAGCAACAGCAGGATGG
	R- ACCCAAGACAACCTAGAAAGGA
Pri-miR-30c	F- CTGATCAACCCTGGACCCTG
	R- GGTCGCATTCTGTCCGATCT
SFPQ	F- ACAGGGAAAGGCATTGTTGA
	R- TCATCTAGTTGTTCAAGTGGTTCC
SPECC1	F- GGAAGATCAGGTGGAACAGCA
	R- CTGCTTGGTCAGAGTCTTCAGC
vIRF4	F- CCCAACAGGCCAGCTACATAA
	R- CTTCGTGGAACTCTGAGACGC
EYA1	F- GAGCTGATGGCTCCGAGTTT
	R- GCTATGCGGGCTGGTTAGAT
RTA/ORF50	F- CGCAATGCGTTACGTTGTTG
	R- GCCCGGACTGTTGAATCG
circEYA1	F- TTGCTTACTGGGTCCTACGC
	R- TACTGCTCCCAATTGCTGAA

Table 3: qPCR primer sequences

Name	Sequence
NEAT1 GapmeR	TAAGCACTTTGGAAAG
DLL4 3′ UTR	F-AAAAAACTCGAGGAGCCTACCTGGACAT
	R-AAAAAAGCGGCCGC TTGGAAACACGGATGCCA
W292A ORF57	F- CACAACCGTGGCCGCCGTCATTTGTTCCTC
	R-GGAGGAACAAATGACGGCGGCCCAGACGGTTGTG
NONO KD 1	F- CCGGAACGTCGCCGATACTAATAAGCTCGAG CTTATTAGTATCGGCGACGTTTTTTTG
	R- AATTCAAAAAAACGTCGCCGATACTAATAAGCTCGAGCTTATTAGTATCGGCGACGTT
NONO KD 2	F-CCGGAAGTCAATTCTGTGTGGTATACTCGAG TATACCACACAGAATTGACTT TTTTTG
	R-AATTCAAAAAAAGTCAATTCTGTGTGGTATACTCGAGTATACCACACAGAATTGACTT
SFPQ KD 1	F- CCGGAATGATTGATGTTGGCTGATACTCGAGTATCAGCCAACATCAATCATTTTTTTG
	R-AATTCAAAAAAATGATTGATGTTGGCTGATACTCGAGTATCAGCCAACATCAATCATT
SFPQ KD 2	F- CCGGAATTAAGACGCATGGAAGAACCTCGAGGTTCTTCCATGCGTCTTAATTTTTTTG
	R- AATTCAAAAAAATTAAGACGCATGGAAGAACCTCGAGGTTCTTCCATGCGTCTTAATT
circHIPK 3 KD	F- CCGGCTACAGGTATGGCCTCACACTCGAGTGTGAGGCCATACCTGTAGTTTTTG
	R- AATTCAAAAACTAXAGGTATGGCCTXAXAXTCGAGTGTGAGGCCATACCTGTAG
DLL4 KD 1	F- CCGGGCAAGAAGCGCAATGACCACTCTCGAGAGTGGTCATTGCGCTTCTTGCTTTTT
	R- AATTCAAAAAGCAAGAAGCGCAATGACCACTCTCGAGAGTGGTCATTGCGCTTCTTG
DLL4 KD 2	F- CCGGGCAGGGAAGCCATGAACAACTCTCGAGAGTTGTTCATGGCTTCCCTGCTTTTTG
	R- AATTCAAAAAGCAGGGAAGCCATGAACAACTCTCGAGAGTTGTTCATGGCTTCCCTG

Table 4: siRNA/shRNA/cloning sequences

2.2 Antibodies

Name	**Brand**	**Species**	**Dilution and use**
GAPDH	Proteintech 60004-1-Ig	Mouse	WB (1/5000)
FLAG	Sigma F7425	Rabbit	IP (1/50) RIP (1/50)
GFP	Proteintech 66002-1-Ig	Mouse	WB (1/5000)
ORF57	Santa Cruz sc-135747	Mouse	WB (1/1000), IF (1/100)
DLL4	Proteintech 21584-1-AP	Rabbit	WB (1/200)
ORF50/RTA	Gift David Lukac, (Rutgers)	Rabbit	WB (1/1000) IF (1/100)
FLAG	Sigma F1804	Mouse	IF (1/50)
ORF65	CRB crb2005224	Rabbit	WB (1/100)
SFPQ	Proteintech 15585-1-AP	Rabbit	WB (1/500), IP (1/50), IF (1/50), RIP (1/50)
NONO	Proteintech 11058-1-AP	Rabbit	IF (1/100), WB (1/1000)
PSPCP1	Proteintech 16714-1-AP	Rabbit	IF (1/50), WB (1/3000)
hnRNP U	Proteintech 14599-1-AP	Rabbit	IF (1/20)
Ago2	abcam ab186733	Rabbit	IP (1/100), RIP (1/100)
DDX17	Proteintech 19910-1-AP	Rabbit	IF (1/10)
DDX21	Proteintech 66925-1-Ig	Mouse	IF (1/50)
hnRNP M	Proteintech 26897-1-AP	Rabbit	IF (1/50)
RNA pol II	Sigma-Aldrich 05-623	Mouse	IF (1/50)
SRSF2	Novus Bio NB100-1774SS	Mouse	IF (1/250)
DLL4	Proteintech 21584-1-AP	Rabbit	WB (1/200), IF (1/50)
hnRNP C	Proteintech 11760-1-AP	Rabbit	IF (1/50)

hnRNP A2B1	Proteintech 14813-1-AP	Rabbit	IF (1/50)
YTHDC1	Proteintech 14392-1-AP	Rabbit	IF (1/50), WB (1/500)
DHX9	Proteintech 67153-1-Ig	Mouse	IF (1/50)
IgG	Sigma Aldrich 12-370	Mouse	IP (1/50), RIP (1/50)
Lamin B	Sigma Aldrich	Rabbit	WB (1/5000)
Histone H3	CST (9715)	Rabbit	WB (1/1000)
SRSF3	Invitrogen (33-4200)	Mouse	WB (1/250), IF (1/50)

Table 5: Antibodies

2.3 Plasmids

Plasmid	**Source**	**Resistance marker**
pVSV.G	Edwin Chen (University of Westminster)	Ampicillin
psPAX2	Edwin Chen (University of Westminster)	Ampicillin
PLKO.1 TRC cloning vector	Gift from David Root; Addgene plasmid #10878	Ampicillin
psiCheck2	James Boyne (Leeds Beckett University)	Ampicillin
GFP	Whitehouse laboratory	Kanamycin
GFP-ORF50	Whitehouse laboratory	Kanamycin
GFP-ORF57	Whitehouse laboratory	Kanamycin
GFP-ORF57 PxxP	Whitehouse laboratory	Kanamycin
GFP-ORF57 RGG1/2	Whitehouse laboratory	Kanamycin
GFP-ORF57-W292A	Cloned	Kanamycin
psiCheck2-DLL4 3′ UTR	Cloned	Ampicillin
PLKO.1 circHIPK3 KD	Cloned	Ampicillin
PLKO.1 DLL4 KDs	Cloned	Ampicillin
PLKO.1 SFPQ KD	Cloned	Ampicillin
PLKO.1 NONO KD	Cloned	Ampicillin
P-lenti CMV GFP	Whitehouse laboratory (Dr Oliver Manners)	Ampicillin
p-lenti CMV GFP-SFPQ	Cloned	Ampicillin

Table 6: Plasmids

2.4 RNA analysis

2.4.1 RNA extraction

Total RNA was extracted either through TRIzol (ThermoFisher Scientific) or Monarch Total RNA Miniprep Kit (NEB) depending on concentration requirements. For Monarch Total RNA Miniprep Kit cell pellets were lysed in 300 µL RNA lysis buffer and transferred to a gDNA removal column. Sample was spun for 30 seconds at 16,000 *x g* before the

flow through was combined with an equal volume of 95% ethanol. Mixture was then transferred to an RNA purification column and spun for 30 seconds at 16,000 *x g* before flow through was discarded. 500 µL RNA wash buffer was added before spinning at 30 seconds for 16,000 *x g* and flow through discarded. For each sample 5 µL DNase I was combined with 75 µL DNase I reaction buffer and added to the column matrix before a 15 minute incubation at room temperature. 500 µL RNA priming buffer was added to column before spinning at 16,000 *x g* for 30 seconds. 2 washes of 500 µL of RNA wash buffer was performed, both spinning at 16,000 *x g*, with the first wash occurring for 30 seconds, and the second for 2 minutes. To elute the RNA the column was placed over a microfuge tube and 30 µL nuclease free water added to the matrix before spinning for 30 seconds at 16,000 *x g.*

For a TRIzol extraction 750 µL of TRIzol was added per 1 x 10^6 TREx-BCBL1-RTA pelleted cells or per well of 6 well plate of pelleted confluent 293Ts. Samples was homogenised before 200 µL chloroform was added and mixed vigorously for 15 seconds. Sample was incubated for 3 minutes at room temperature before centrifugation at 12,000 *x g* for 15 minutes at 4ºC. The clear phase was then transferred to a new RNase free eppendorf and equal volumes of isopropanol was added. 1 µL of Glycoblue (ThermoFisher Scientific) was added and sample incubated for 10 minutes at room temperature. Samples were spun at 12,000 *x g* for 10 minutes at 4ºC and supernatant removed. Pellet was washed in 1 ml 75% ethanol before spun for 5 minutes at 7500 *x g* at 4ºC. Supernatant was removed and pellet re-suspended in 20 µL water.

RNA concentration and purity were analysed via NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies)

2.4.2 gDNA removal

RNA extracted via TRIzol was DNase treated to remove contaminating gDNA. To each sample 1 µL DNase I (Invitrogen) and 0.1 volume DNase Buffer were added before the sample was incubated for 30 minutes at 37ºC. 0.1 volume DNase inactivation reagent

was added and pipetted vigorously before incubated for 2 minutes at room temperature. Sample was spun at 10,000 rpm for 1.5 minutes before supernatant transferred to a new tube.

2.4.3 Reverse Transcription

For each reaction 1000 ng of RNA was used in a total reaction mix of 20 µL with RNase free water and 4 µL of LunaScript RT SuperMix (NEB). Samples were incubated at 25ºC for 2 minutes, 55ºC for 10 minutes and heat inactivated at 95ºC for 1 minute.

2.4.4 miRNA Quantification

Total RNA was extracted via TRIzol. Reverse transcription was performed using the miScript RT II kit (Qiagen). Each reaction consisted of 4 µL 5X miScript HiSpec Buffer, 2 µL 10X miScript Nucleic Acid Mix, 2 µL miScript Reverse Transcriptase Mix, 1000 ng RNA and up to 20 µL RNase free water. Samples were incubated for 60 minutes at 37ºC and 5 minutes at 95ºC. cDNA was then diluted 1/10 in RNase free water. qPCR was performed using miScript PCR kit (Qiagen). Each reaction consisted of 10 µL 2X GoTaq (Promega), 2 µL 10X Universal Primer, 2 µL 10X miScript Primer Assay (Qiagen), 5 µL RNase free water and 1 µL cDNA. qPCR conditions were initial activation at 95ºC for 15 min; followed by 40 cycles of a 3-step consisting of denaturing for 15 seconds at 94ºC, annealing at 55ºC for 30 seconds and extending for 30 seconds at 70ºC.

2.4.5 qPCR

Total cDNA obtained from LunaScript RT was diluted 1/10 to 200 µL total. Primer master mixes were created to final concentration of 10 µM comprising of 20 µL forward primer, 20 µL reverse primer and 160 µL nuclease free water. Each qPCR reaction comprised of 10 µL 2X GoTaq (Promega), 8 µL nuclease free water, 1 µL 10 µM primer master mix and 1 µL cDNA (5 ng). An n + 1 master mix was created for each primer used comprising of GoTaq, nuclease free water and the primer mix before being added to 0.1 ml Rotor-gene style strip tubes (Star lab), with cDNA added separately. qPCR conditions were initial activation at 95ºC for 10 min; followed by 40 cycles of a 3-step consisting of

denaturing for 15 seconds at 95ºC, annealing at 60ºC for 30 seconds and extending for 20 seconds at 72ºC. Runs were analysed using the ΔΔCT method.

2.4.6 GFP-ORF57 RNA Immunoprecipitation

HEK-293T cells were seeded out 24 hours prior to transfection so they would be 90% confluent before being transfected with 2 µg of either GFP or GFP-ORF57 plasmid. Cells were collected 8, 16 or 24 hours after transfection, spun and washed with PBS. Cells were re-suspended in 200 µL ice-cold RIP lysis buffer (150 mM KCl, 25 mM Tris pH7.4, 5 mM EDTA, 0.5% NP-40 v/v, 0.5 mM DTT and 0.25 % Murine RNase inhibitor v/v) and placed on ice for 30 minutes with extensive pipetting every 10 minutes. Tubes were then spun at 18,000 *x g* at 4ºC for 10 minutes and lysate transferred to new pre-cooled tube. 300 µL of Dilution/Wash buffer (10 mM Tris pH 7.5, 150 mM NaCl, 0.5 mM EDTA) was added. 25 µL of GFP-TRAP beads slurry (Chromotek) were used per RIP. 500 µL ice-cold Dilution/Wash buffer was added to beads before centrifuging for 2 minutes at 2500 *x g* at 4ºC and removing supernatant. Wash was repeated twice more. 450 µL cell lysate was added to beads with 50 µL kept as input. Beads were incubated with lysate overnight rotating at 4ºC. Beads were centrifuged for 2 minutes at 2500 *x g* at 4ºC and supernatant removed before washed 5 times in Dilution/Wash buffer. 150 µL Proteinase K buffer (comprising 126 µL RIP wash buffer, 15 µL 10% SDS and 9 µL 20mg/ml Proteinase K) was added to each RIP and inputs. Samples were incubated at 55ºC for 30 minutes before placed on the magnetic separator. The supernatant was transferred to a new tube and 100 µL of RIP wash buffer added to each sample before addition of 750 µL TRIzol LS. RNA was extracted from the samples and cDNA was synthesised before analysis via qRT-PCR.

2.4.7 Antibody RNA Immunoprecipitation

1 million TREx-BCBL1-RTA cells were used for each RIP. Cells were lysed in 110 µL ice cold RIP lysis buffer and left on ice for 5 minutes before frozen at -80ºC. For each RIP 25 µL Magnabeads (Millipore) were used, beads were re-suspended and 500 µL RIP

wash buffer added 20 mM Tris pH7.4, 0.1% NP-40 v/v, 150 mM NaCl). Microfuge tubes were placed on a magnetic separator and supernatant discarded, this was repeated. Beads were then re-suspended in 100 µL RIP wash buffer with addition of either IgG, Ago2 or SFPQ antibody. Beads were incubated for 40 minutes at room temperature with rotation before magnetic separating and the supernatant discarded. Beads were then washed with 500 µL of RIP wash buffer twice. 900 µL of IP buffer (RIP wash buffer, 20 mM EDTA, 0.5% v/v RNase inhibitor) was added to the beads. Cell lysates were thawed and centrifuged at 16,000 *x g* for 10 minutes at 4ºC, with 10 µL taken as input and the remaining 100 µL added to the magnetic beads. Magnetic beads were then incubated rotating overnight at 4ºC while inputs were frozen at -80ºC. Magnetic beads were washed 6 x with 500 µL cold RIP wash buffer. 150 µL Proteinase K buffer was added to each RIP and input. Samples were incubated at 55ºC for 30 minutes before placed on the magnetic separator. The supernatant was transferred to a new tube and 100 µL of RIP wash buffer added to each sample before addition of 750 µL TRIzol LS. RNA was extracted from the samples and cDNA was synthesised before analysis via qRT-PCR.

2.4.8 Biotinylated miR-29b and miR-30c RNA immunoprecipitation

HEK-293T cells were plated into a 12 well plate so they would achieve 90% confluent 24 hours later. 20nM of biotinylated LNA miR-29b or miR-30c mimic (Qiagen) and 2 µL Lipofectamine RNAi Max (ThermoFisher Scientific) were each combined with 100 µL Opti-MEM and incubated for 10 minutes before being combined and incubated for 10 minutes. This was then added to cells and cells were incubated for 24 hours. 25 µL Dynabeads MyOne Streptavidin T1 were used per RIP, beads were re-suspended before being transferred to a low-bind tube. Tube was transferred to magnetic separator and supernatant removed. Beads were washed three times in 100 µL bead wash buffer (10 mM Tris pH 7.5, 0.5 mM EDTA, 1M NaCl). Beads were then re-suspended in RNase freeing solution (0.1M NaOH, 0.05 M NaCl) and incubated at room temperature for 2 minutes. Tubes were placed on a magnetic separator and supernatant removed; this

was performed twice. Beads were re-suspended in 100 µL Solution B (0.1 M NaCl) and incubated for 2 minutes at room temperature before magnetic separation and removal of the supernatant. Beads were re-suspended in 200 µL bead blocking solution (1 µg/µL BSA and 1 µg/µL yeast tRNA in RNase free water) and incubated for 2 hours at room temperature. Cells were collected and lysed in 250 µL lysis buffer (10 mM KCl, 1.5 mM $MgCl_2$, 10 mM Tris pH7.5, 5mM DTT, 0.5% IGEPAL, 60U/ml RNase inhibitor, 1x protease inhibitor). Samples were frozen at -80ºC for full lysis. Lysates were thawed and transferred to a low-bind tube before centrifugation at 4ºC for 2 minutes at 13,000 *x g*. 10 µL of supernatant was kept as input and transferred to a low bind tube and kept on ice. 60 µL 5M NaCl was added to the remaining 240 µL of lysate. Beads in blocking solution was placed on the magnetic separator before supernatant discarded and washed twice using 100 µL wash buffer (10 mM KCl, 1.5 mM $MgCl_2$, 10 mM Tris pH7.5, 5mM DTT, 0.5% IGEPAL, 60U/ml RNase inhibitor, 1x protease inhibitor, 1M NaCl). Beads were then re-suspended in 300 µL wash buffer and 300 µL of cell lysate was added to beads. Tubes were incubated rotating for 30 minutes at room temperature. Tubes were then magnetically separated and supernatant discarded. Beads were washed 3 times with 300 µL wash buffer. After the washes beads were re-suspended in 100 µL RNase free water. RNA purified from the beads using TRIzol extraction as previously described.

2.4.9 RNA Fluorescent *In situ* hybridisation (FISH)

250,000 TREX-BCBL1-RTA cells were seeded out in a 24 well plate on coverslips that had been sterilised in 70% ethanol, incubated in 200 µL poly-L-Lysine (Sigma Aldrich) for 5 minutes and washed with 1 mL PBS. 3 hours after seeding cells were induced. 24 hours after induction, cell media was removed.

RNA FISH was performed using ViewRNA Cell Plus Assay Kit (ThermoFisher Scientific). Fixation/Permeabilisation Solution was prepared by combining equal volumes Fixation/Permeabilisation component A with Fixation/Permeabilisation component B

before 400 µL was added per well and incubated for 30 minutes at room temperature in a humidified staining tray. 1X PBS with RNase inhibitor was prepared by adding RNase inhibitor at 1:100 dilution to 1X PBS. Fixation/Permeabilisation solution was aspirated and cells washed three times with 400 µL 1X PBS with RNase inhibitor. Blocking/Antibody Diluent Solution was prepared by combining Blocking/Antibody Diluent with RNase inhibitor in a 100:1 ratio. Cells were incubated with 400 µL Blocking/Antibody Diluent Solution for 20 minutes at room temperature. Primary antibodies were diluted in Blocking/Antibody Diluent to appropriate dilutions. Blocking/Antibody Diluent solution was aspirate and 200 µL primary antibody solution was added and incubated at room temperature for 1 hour. Primary antibody solution was aspirated, and cells washed 3 times with 400 µL 1X PBS with RNase inhibitor. Secondary antibody was diluted 1 in 500 with Blocking/Antibody Diluent and 200 µL overlaid onto cells and incubated for 1 hour at room temperature. Secondary antibody was aspirated off and washed 3 times with 400 µL 1X PBS with RNase inhibitor. Fixation Solution was prepared by combining 1-part Solution A Fixative with 7 parts Solution B Fixative. Cells were overlaid with 400 µL Fixation Solution for 1 hour at room temperature before it was aspirated and cells washed 3 times with 400 µL 1X PBS with RNase inhibitor.

QG Probe Set for circHIPK3 or NEAT1 were thawed and diluted 1:100 in Probe Set Diluent pre-warmed to 40ºC before 200 µL was added to cells and incubated for 2 hours at 40ºC. Wash Buffer solution was prepared by addition of 3 µL/ml of Wash component 1 and 5 µL/ml of Wash component 2 to ddH_2O. Probe set diluent was aspirated and cells washed in 200 µL Wash Buffer 5 times. Cells were incubated overnight at 4ºc in Wash Buffer. PreAmplifier Solution was prepared by dilution PreAmplifier mix 1:25 to Amplifier Diluent pre-warmed to 40ºC. Wash buffer was aspirated and 200 µL PreAmplifier Solution added for 1 hour at 40ºC. Working Amplifier Solution was prepared by adding AmplifierMix in a 1:25 ratio with Amplifier Diluent pre-warmed to 40ºC. PreAmplifier solution was aspirated and cells washed 5x with 400 µL Wash Buffer. 200 µL Working

Amplifier solution was incubated for 1 hour at 40ºC, before being aspirated and cells washed 5x with 400 µL Wash Buffer. Working Label Probe Solution was prepared by diluting Label Probe Mix 1:25 with Label Probe diluent pre-warmed to 40ºC before 200 µL added to cells and incubated for 1 hour at 40ºC. Working Label Probe Solution was aspirated before cells were washed 5x in 800 µL Wash Buffer, with the cells sitting in the final wash for 10 minutes. Cells were washed with 1X PBS before coverslips were mounted onto slides using Vectashield Antifade mounting medium with DAPI (Vector Labs). Slides were stored at 4ºC before imaging on Zeiss LSM880 with AiryScan Inverted confocal microscope and processing used ZEN 2009 imaging software (Carl Zeiss).

2.5 Mammalian Cell culture

2.5.1 Cell lines and maintenance

TREx-BCBL1-RTA cells are a B lymphoma cell line latently infected with the KSHV genome, engineered to contain a doxycycline inducible myc-RTA (gift Professor JU Jung University of Southern California). BC-3 cells are a primary effusion lymphoma cell line positive for KSHV (ATCC). HEK-293T-rKSHV.219 cells are a HEK-293T cell line infected with the recombinant KSHV.219, constitutively expressing GFP and expressing RFP from the KSHV lytic PAN promoter (gift Dr Jeffery Vieira, University of Washington). BCBL1s are a B lymphoma cell line latently infected with KSHV (gift Dr Andrew Hislop, University of Birmingham) HEK-293Ts are an epithelial embryonic kidney cell line and were purchased from the ATCC. TREx-BCBL1-RTA cells were cultured in RPMI1640 with glutamine (Gibco), supplemented with 1% P/S, 10% Foetal Bovine Serum (FBS) and 1 µg/mL hygromycin B (ThermoFisher Scientific). Overexpression or KD cells were also cultured with 3 µg/mL puromycin. BCBL1s were cultured in RPMI1640 with 1% P/S and 10% FBS. BC-3s were cultured in RPMI 1640 and supplemented with 1% P/S and 20% FBS. HEK-293Ts were cultured in Dulbecco's modified Eagle's medium with glutamine (DMEM) (Gibco) supplemented with 10% FBS and 1% P/S. HEK-293T-rKSHV.219 cells were cultured in DMEN supplemented with 10% FBS, 1% P/S and 3

µg/mL puromycin (Gibco). All cell lines were cultured at 37 °C at 5% CO_2 and tested negative for mycoplasma. Virus lytic replication in TREx-BCBL1-RTA cells was induced via addition of 2 µg/mL doxycycline hyclate (Sigma Aldrich). HEK-293T-rKSHV.219 and BC-3s cells were induced via addition of 20 ng/mL TPA and 4 mM sodium butyrate. BCBL1 cells were induced via addition of 2 mM sodium butyrate.

2.5.2 Lentivirus work

HEK-293Ts were seeded out into a 6 well a day prior to transfections to allow for 60% confluency when transfected. Transfections were performed at 09:00, typically, 650 ng of pVSV.G, 650 ng psPAX2 and 1.2 µg of the shRNA pLKO.1 plasmid DNA were diluted in 250 µL Opti-MEM (ThermoFisher Scientific) and 4 µL Lipofectamine 2000 was combined with µL Opti-MEM. Both were incubated for 10 minutes at room temperature before being combined and incubated for a further 20 minutes at room temperature. This was then added to cells and incubated until the end of the day, where it was removed and replaced with fresh DMEM. Cells were left for 48 hours before supernatant was collected by being spun for 3 minutes at 1300 *x g.* 1 million TREx-BCBL1-RTA cells were seeded out in a 6 well in 1 ml of RPM1 media. The HEK-293T supernatant was passed through a pre-wetted 0.45 µM filter and added to the TREx-BCBL1-RTA cells in a 1:1 ratio. Cells were also incubated with 8 ng/ µL polybrene. After 8 hours media was removed and replaced with fresh RPMI media. 3 days after transductions, RPMI media was replaced with puromycin containing RPMI, with this media replaced every 2-3 days. Puromycin was also added to a non-transduced control. KD was assessed via qPCR and western blot if appropriate 7-9 days post-transduction.

2.5.3 GapmeRs

Antisense LNA-GapmeRs (Qiagen) were utilised against NEAT1 as a nuclear retained RNA to achieve KD. 100 nm of either a scr GapmeR or targeting NEAT1 were added to TREx-BCBL1-RTA cells at the same time as doxycycline induction. KD was assessed via qPCR.

2.5.4 Transfections

HEK-293T cells were seeded out at in a 6 well plate 24 hours prior to transfection so they would achieve 80% confluency at time of transfection. Transfection was carried out using Lipofectamine 2000 (ThermoFisher Scientific) as per manufacturers' instructions. Typically, 2 µg of plasmid DNA and 4 µL Lipofectamine 2000 (ThermoFisher) were each diluted in 250 µL Opti-MEM (ThermoFisher) and incubated for 10 minutes at room temperature. The DNA and Lipofectamine were then combined and incubated for 10 minutes at room temperature before adding to cells.

Luciferase plasmid transfections were carried out in HEK-293T cells at 60% confluency in a 12 well. Transfection was carried out with 100 ng of plasmid DNA, and an appropriate amount of miRNA mimic or antagomiR (detailed in Luciferase methods). Plasmid DNA and the mimic/antagomiR was incubated in 100 µL Opti-MEM while 3 µL Lipofectamine 2000 was incubated in another 100 µL OptiMEM for 10 minutes, before they were combined at incubated for another 10 minutes before adding to cells.

miRNA mimic and antagomiR transfections were performed on 1 million TREx-BCBL1-RTA cells. Briefly 50 nM of mimic or 20-40 nM of antagomiR was combined with 150 µL OptiMEM. Simultaneously 8 µL of RNAi Max (ThermoFisher Scientific) was combined with 150 µL OptiMEM. Tubes were left to incubate at room temperature, before being combined and incubated for a further 10 minutes at room temperature before addition to cells. Cells were reactivated 24 hours post-transfection.

2.5.5 MTS Assay

50,000 TREx-BCBL1-RTA cells were used per well of a 96 well plate in 200 µL media. Drugs were tested across a range of an appropriate concentration, with each concentration performed in triplicate. Additionally untreated cells, media alone and cells + DMSO were used as controls. Post-drug treatment 20 µL CellTiter 96 Aqueous One Solution Cell Proliferation Assay MTS (Promega) was added to the plate and incubate

at 37 ºC for 30 minutes. Values were read by Powerwave XS2 plate reader (BioTek) at wavelength 490 nm.

2.5.6 Harvesting Cells

Cells were collected and spun down at 500 *x g* for 5 minutes at 4ºC before the supernatant was removed. 1X PBS was added to wash the cells before spinning at 500 *x g* for 5 minutes at 4ºC before supernatant was removed. Pellets were stored at -80 ºC. For protein extraction 50-100 µL modified RIPA buffer was added per sample (50 mM Tris HCL pH 7.5, 150 mM NaCl, 1% NP-40, 1X Protease inhibitor (Sigma Aldrich). Samples were then sonicated for 3X 30 seconds with 30 second breaks in between before incubated on ice for 30 minutes. Samples were centrifuged at 12,000 x g for 10 minutes at 4ºC and the supernatant was transferred to a new tube. For RNA extraction the pellet was used in a TRIzol reaction or Monarch Total RNA Miniprep Kit (NEB), see RNA extraction section.

2.6 Virus based Assays

2.6.1 Viral load

500,000 TREx-BCBL1-RTA cells were induced for 72 hours before sample was collected and spun at 500 *x g* at 4°C for 5 minutes. Supernatant was used for the viral reinfection assay while the pellet was washed in PBS and spun again at 500 *x g* at 4°C for 5 minutes. DNA was extracted using DNeasy Blood and Tissue Kit (Qiagen). Cell pellet was re-suspended in 200 µL PBS, before addition of 20 µL Proteinase K and 200 µL Buffer AL. Sample was vortexed before addition of 200 µL 100% ethanol. Sample was pipetted into DNeasy Mini spin column and spin at 6000 *x g* for 1 minute. Flow through was discarded and 500 µL Buffer AW1 was added before repeating spin step. 500 µL Buffer AW2 was added to spin column and spun for 3 minutes at 16,000 *x g.* Spin column was placed in clean 1.5 ml microcentrifuge tube and 100 µL nuclease free water was added to the membrane and incubated for 1 minute. Sample was spun for 1 minute at 6000 *x g* to elute DNA before quantification on the NanoDrop.

2.6.2 Viral reinfection Assay

As described above 500,000 TREx-BCBL1-RTA cells were induced for 72 hours before sample was collected and spun at 500 *x g* at 4°C for 5 minutes. Supernatant was then added to naïve HEK-293T cells in a 1:1 ratio with DMEM. 24 hours afterwards, cells were harvested and RNA extracted using Monarch Total RNA Miniprep Kit as described previously.

2.6.3 K8 splicing assay

50 ng cDNA was obtained at 24 hours post reactivation from TREx-BCBL1-RTA cells. This was diluted 1/1000 and added to 10 µL 2x Q5 Master mix (NEB), x µL nuclease free water and 1 µL of 10 µM Pr75183 forward primer and either 1 µM of 10 µM Pr75828 or 1 µM of 10 µM of Pr75864, to a total volume of 20 µL. Samples were incubated at 98ºC for 30 seconds, followed by 35 cycles of 98ºC for 10 seconds, 55ºC for 20 seconds and 72ºC for 45 seconds, with a 72ºC extension for 2 minutes. Samples were then run on an agarose gel and splicing patterns analysed through visualisation of gel under UV light.

2.7 Molecular techniques

2.7.1 Subcellular Fractionation

1 million TREx-BCBL1-RTA cells were plated out and induced for 24 hours. Cells were pelleted by centrifugation at 1700 rpm for 7 minutes and supernatant discarded. Cells were washed with 1 ml PBS and pelleted at 500 *x g* for 5 minutes at 4ºC before supernatant discarded. Cells were re-suspended in 600 µL PBS with 1% triton and 1 µL/ml RNase inhibitor and lysed for 10 minutes on ice. 250 µL was kept for the whole cell fraction. The remaining 350 µL was centrifuge at 720 *x g* at 4ºC for 5 minutes. The supernatant was transferred to a new eppendorf and kept on ice as the cytoplasmic fraction. The pellet was washed in 1 ml of PBS and centrifuged at 720 *x g* for 5 minutes at 4ºC before the supernatant was removed leaving behind the nuclear fraction. Samples were then either re-suspended in 2x loading dye for analysis via western blot or TRIzol LS (ThermoFisher Scientific) for an RNA extraction.

2.7.2 Nuclear fractionation

1 million TREx-BCBL1-RTA cells were plated out and induced for 24 hours. Cells were collected, spun and washed in PBS as per standard protocol. Cells were resuspended in 50 µl Fractionation Lysis Buffer (10 mM Tris-HCl pH 7.5, 0.05% NP-40, 150 mM NaCl, 1x protease inhibitor) and lysed on ice for 5 minutes. Cell lysate was added to 2.5 volumes of sucrose cushion buffer (Fractionation Lysis buffer containing 24% sucrose) and centrifuged at 16,000 *x g* for 10 minutes at 4 °C before supernatant was discarded. The pellet containing the nucleus was resuspended in 50 µl Glycerol buffer (20 mM Tris-HCl pH 7 7.9, 75 mM NaCl, 0.5 mM EDTA, 0.85 mM DTT, 50% glycerol, 1x protease inhibitor). This was then combined with 50 µl nuclear lysate buffer (10 mM HEPES pH 7.6, 1 mM DTT, 7.5 mM $MgCl_2$, 0.2 mM EDTA, 0.3 M NaCl, 1 M urea, 1% NP-40%, 1x protease inhibitor) and vortexed 2x for 2 seconds. Lysates were chilled on ice for 2 minutes before centrifugation at 10,000 rpm for 2 minutes at 4 °C. Supernatant was collected and kept as nucleoplasmic sample. The pellet was washed in PBS and treated with DNase as previously detailed in TRIzol protocol becoming the chromatin sample. Samples were then resuspended in 2x loading dye for analysis via western blot.

2.7.3 Immunofluorescence

Coverslips were sterilised in 70% ethanol before air drying in 24 well plates. Coverslips were then coated in 200 µL poly-l-lysine (Sigma Aldrich) for 5 minutes before being washed in PBS. 250,000 TREx-BCBL1-RTA cells were added per well before induction via addition of dox 3 hours later. Cells were fixed with 200 µL 4% paraformaldehyde for 15 minutes at room temperature, before 3 x wash steps with PBS. Permeabilisation occurred through addition of 200 µL PBS + 1% triton for 15 minutes followed by 3 x PBS washes. Coverslips were transfected to an incubation chamber and blocked for 1 hour at 37ºC in PBS with 1 % BSA. This was followed by a 1-hour incubation at 37ºC with the appropriate primary antibody diluted in PBS and 1% BSA. Coverslips were washed 5x in PBS before incubation for 1 hour at 37ºC with Alexa-fluor conjugated secondary antibody

(Invitrogen) at 1/500 in PBS and 1% BSA. Coverslips were washed 5x in PBS before mounted using Vectashield Antifade mounting medium with DAPI (Vector Labs). Slides were stored at 4ºC and protected from light before imaging on Zeiss LSM880 with AiryScan Inverted confocal microscope and processing used ZEN 2009 imaging software (Carl Zeiss).

2.7.4 FRAP

A stable GFP-SFPQ overexpression TREx-BCBL1-RTA cell line was created using the lentivirus transduction protocol previously described. Cells were seeded out on coverslips in a 24 well for IF as previously described and reactivated. At 16 and 24 hours post-transfection, coverslips were transferred to 35 mm glass bottomed dish (Ibidi) and imaged using the inverted LSM880 Zeiss microscope that allows live cell imaging. The Zeiss bleaching protocol was used, namely 2 regions were selected, with one bleached and one remaining unbleached. 2 images were taken pre-bleaching to establish an initial fluorescence before bleaching occurred using the 405nm and 488nm lasers at power 100 for 2 seconds, reducing fluorescence to less than 20% of the original. GaAsp was used to protect the microscope. Fluorescent intensity was measured every 5 seconds for 220 seconds.

2.7.5 Immunoprecipitations

5 million of TREx-BCBL1-RTA cells were used per IP condition. 15 µL Dynabeads Protein A or G (ThermoFisher Scientific) were used per IP, with protein A beads used for a rabbit antibody and protein G for a mouse antibody. Beads were washed 3 times with 100 µL RIP lysis buffer, before beads were re-suspended in 100 µL RIP wash buffer and incubated with the antibody of interest at the appropriate concentration. Beads were rotated at 4 ºC for 1 hour before washed 3 times with 100 µL RIP wash buffer. Cells were collected as per normal protocol before lysed in 100 µL RIP lysis buffer and incubated on ice for 20 minutes. Cell lysate was centrifuged at 16, 000 *xg* for 10 minutes at 4 ºC before lysate added to the beads and a 10% input taken. 2x loading dye was added to

the input. Beads were incubated overnight rotating at 4 ºC. Beads were then washed 3 times in RIP wash buffer and re-suspended in 30 µL 2x loading dye and boiled for 10 minutes at 95 ºC. Lysates were analysed via western blot. However, if the primary antibody species during the western blot corresponded to the antibody species used in the pulldown, veriblot (abcam 1/1000) was used instead of secondary to prevent background bands.

2.7.6 GFP-TRAP immunoprecipitations

HEK-293T were plated out in 6cm dishes so they achieved 70% confluency at transfection. Transfections occurred as previously described, scaled so 4 µg of plasmid was used per condition. 24 hours post transfection cells were lysed in 200 µL RIP lysis buffer and incubated on ice for 30 minutes with extensive pipetting every 10 minutes. Cell lysate was spun at 16,000 *x g* at 4 ºC for 10 minutes, and lysate transferred to a new tube. 300 µL Dilution/Wash buffer was added to each sample. 25 µL GFP-TRAP beads were used per IP, slurry was washed twice in dilution buffer as previously described before lysate added to beads and tumbled overnight at 4 ºC. A 10% input was taken. GFP-TRAP beads were washed 3 times in RIP wash buffer before resuspended in 25 µL 2 x loading dye. Samples were boiled for 10 minutes at 95 ºC before analysis via western blot.

2.7.7 Flow cytometry

1 million TREx-BCBL1-RTA cells were used per condition. Cells were harvested at the appropriate time point post-reactivation as previously detailed. Post-PBS wash cell pellet was fixed overnight in 70% ethanol at -20 ºC. Ethanol was removed, and cells resuspended in PBS + 0.5% BSA. Cells were spun at 500 x g for 10 minutes at 4 ºC before PBS + 0.5% BSA removed, this step was repeated twice. Cells were stained in PBS containing 0.5% BSA, 5 µg/ml RNase and 25 µg/ml propidium iodide (Sigma) for 30 minutes at room temperature. Samples were processed on CytoFlex S Benchtop Flow Cytometer (Beckman Coulter).

2.7.8 Cloning PLKO.1 TRC-cloning vector

Forward and Reverse shRNA oligonucleotides against circHIPK3 (IDT) were re-suspended in ddH_2O to 20 µM. The oligonucleotides were annealed by adding the following: 5 µL Forward oligonucleotide, 5 µL Reverse oligonucleotide, 5 µL NEB buffer 2 (NEB) and 35 µL ddH_2O. The mix was incubated at 95 ºC for 4 minutes, 70 ºC for 10 minutes before cooling gradually in a water bath.

PLKO.1 TRC-cloning vector was a gift from David Root (Addgene plasmid # 10878; http://n2t.net/addgene:10878; RRID: Addgene_10878) [252]. 2 µg PLKO.1 TRC-cloning vector was digested by incubating with 1 µL AgeI-HF (NEB), 1 µL EcoRI-HF (NEB), 2 µL CutSmart 10x Buffer (NEB) and up to 20 µL ddH_2O at incubate for 1 hour at 37ºC.

Digested plasmid was combined with 4uL 6x Gel loading dye (NEB) and run on an agarose gel (0.7% agarose, 1x TEA buffer, 0.001% SYBR Safe (ThermoFisher Scientific)) at 90V for 40 minutes. Gel was visualised by UV light and 7kb gel band excised. Gel extraction was performed as described below. Annealed oligonucleotides were ligated with the digested PLKO.1 TRC cloning vector. 2 µL annealed oligonucleotides, 20 ng digested PLKO.1 TRC cloning vector, 2 µL 10x T4 DNA ligase buffer (NEB), 1 µL T4 DNase ligase (Promega) and up to 20 µL total with ddH_2O were incubated overnight at 16ºC. Ligation mix was then transformed as described below.

2.7.9 Cloning psi-chk2 Vector

The insert composed of the 3′ UTR of interest was amplified from genome using Q5. 12.5 µL Q5 master mix (NEB) was combined with 1.25 µL 10 µM forward primer, 1.25 µL 10 µM reverse primer, 0.75 µL DMSO, 1000 ng template DNA and x of water up to 25 µL total volume. Tubes were placed in a thermocycler and a gradient PCR was run as followed: 3 minutes at 98ºC before 25 cycles of 98ºC for 7 seconds, 30 seconds at a gradient of temperatures ranging from 59-71ºC and 15 seconds at 72ºC. 5 µL sample was added to 1 µL 6x DNA loading dye (NEB) and run on an agarose gel (1% agarose, 1x TEA buffer, 0.001% SYBR Safe (ThermoFisher Scientific)) at 90V for 40 minutes to

confirm successful amplification while the remaining DNA underwent a PCR clean-up. Briefly 5x DNA Clean-up binding buffer was added to DNA and loaded into the column. Sample was spun at 16,000 *xg* for 1 minute and flow through discarded. 200 µL DNA wash buffer was added to column and spun at 16,000 *xg* for 1 minute, this step was repeated twice. DNA was eluted through addition of 20 µL of nuclease free water to the column, this was incubated for 1 minute before being spun at 16,000 *xg* for 1 minute.

Psi-chck2 plasmid was double restriction enzyme digested. Briefly at total reaction volume of 20 µL made up of 1 µg of psichk2, x µL of nuclease free water, 2 µL 10x CutSmart buffer (NEB), 1 µL XhoI (NEB) and 1 µL Not1-HF (NEB) was incubated at 37ºC for 1 hour. To confirm linearisation 5 µL of sample was added to was added to 1 µL 6x DNA loading dye (NEB) and run on an agarose gel (0.8% agarose, 1x TEA buffer, 0.001% SYBR Safe (ThermoFisher Scientific)) at 90V for 40 minutes. Linearised plasmid underwent phosphatase treatment, with 1 µL Antarctic Phosphatase (NEB) and 1.7 µL 10x Antarctic Phosphatase buffer (NEB) added to remaining digested plasmid, sample was incubated at 37ºC for 30 minutes and for 2 minutes at 80ºC.

Plasmid and insert were ligated in a 50 µL total ligation reaction comprising of 5 µL of 10x T4 DNA Ligase Buffer (NEB), 1 µL DNA ligase (NEB), 100ng plasmid, insert so in a 1:3 ratio to plasmid and x µL nuclease free water. Samples were incubated at 16 ºC overnight before transformation into competent cells.

2.7.10 Gel extraction

DNA bands were excised from agarose gels using Monarch DNA Gel extraction Kit (NEB). Agarose gels were visualised using UV light and the band cut out using a scalpel before transferring to a pre-weighed micro centrifuge tube. The weight of the gel slice was then calculated before 4 volumes of Monarch Gel dissolving buffer was added to the tube. Sample was incubated for 10 minutes at 55 ºC with occasional agitation. Once the gel was dissolved, sample was transferred to a column and spun for 1 minute at 16,000 *x g.* Flow through was discarded before addition of 200 µL DNA Wash Buffer

followed by another 1 minute spin at 16,000 *x g.* This was step was repeated twice. Column was transferred to a clean microfuge tube and 20 µL of water was incubated on the membrane for 1 minute. The tube was spun for 1 minute at 16,000 *x g* to elute the DNA, purity and concentration was then measured by the NanoDrop.

2.7.11 Cloning plenti-CMV GFP-SFPQ

This plasmid was created using a 3 part Gibson assembly reaction using primers designed by NEB builder. GFP, SFPQ and p-lenti CMV plasmid was all amplified via PCR in 50 µL reaction comprising of 2x Q5 master mix (NEB), 2.5 µL 10 µM primer master mix, x water, 1% DMSO, and either 1ng of plasmid or 50 ng cDNA in the following conditions:

GFP: 98ºC for 30 seconds
35 cycles of: 98ºC for 10 seconds, 68ºC for 30 seconds and 72ºC for 30 seconds
72ºC for 30 seconds

SFPQ: 98ºC for 30 seconds
35 cycles of: 98ºC for 10 seconds, 62ºC for 30 seconds and 72ºC for 30 seconds
72 ºC for 30 seconds

p-lenti CMV plasmid: 98ºC for 30 seconds
35 cycles of: 98ºC for 10 seconds, 66ºC for 30 seconds and 72ºC for 4.5 minutes
72ºC for 30 seconds

Fragments were assembled through a 20 µL Gibson assembly reaction comprising of 50 ng of plasmid with a 3-fold excess of inserts, x nuclease free water, and 10 µL 2X Gibson Master Mix (NEB). Samples were incubated at 50ºC for 15 minutes before being transformed into competent cells.

2.7.12 Site Directed Mutagenesis

W92A GFP-ORF57 plasmid was created using the QuikChange Site-Directed mutagenesis Kit (Agilent Technologies). A 50 µL reaction was prepared comprising of 5 µL 10x reaction buffer, 50 ng DNA template, 2 µL of 100 µM primer master mix, 1 µL dNTP mix and x µL nuclease free water. 1 µL PfuUltra HF DNA polymerase was then

added and a PCR reaction run as follows: 95ºC for 30 seconds followed by 16 cycles of: 95ºC for 30 seconds, 55ºC for 1 minute and 68ºC for 6 minutes. Sample was transformed into competent cells.

2.7.13 Transformations

50 µL competent *E. coli* DH5α cells were thawed on ice and combined with 1 µL plasmid DNA. Samples were incubated for 30 minutes on ice, heat shocked at 42ºC for 30 seconds and cooled on ice for 2 minutes. 950 µL LB broth was added and sample was incubated at 37ºC for 1 hour. 200 µL of cells were plated out onto LB agar plates (1.5% w/v agar) with either 50 µg/mL ampicillin or 25 µg/mL kanamycin. Plates were incubated at 37ºc overnight before a single colony was picked and inoculated in 5ml of LB broth (with either 50 µg/mL ampicillin or 25 µg/mL kanamycin) overnight before plasmid purification. Glycerol stocks were created by combining 500 µL of overnight culture with 500 µL 50% glycerol before storage at -80ºC.

2.7.14 Plasmid Purification

Plasmids were purified using Monarch plasmid Miniprep kit (NEB). Overnight cultures were spun down at 12,000 *x g* for 3 minutes before supernatant removed. Cells were re-suspended in 200 µL Plasmid resuspension buffer B1 and transferred to an eppendorf. 200 µL Plasmid Lysis buffer B2 buffer was added the tube inverted until sample was purple and clear. Sample was incubated for 1 minute, before addition of 400 µL Plasmid neutralisation buffer B3 and inverted until the sample turned yellow. Sample was spun for 5 minutes at 16,000 *x g*, and lysate transferred to a spin column, before the column was spun for 1 minute at 16,000 *x g* and flow-through discarded. 200 µL of Plasmid wash buffer 1 was added before a 1 minute spin at 16,000 *x g* followed by addition of 400 µL of Plasmid wash buffer 2 and centrifuging at 16,000 *x g* for 1 minute. DNA was eluted by transferring the column to a microfuge tube and 50 µL of nuclease free water was added and incubated on the membrane for 1 minute, before spinning at 16,000 *x g* for 1 minute.

DNA purity was measured using NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies) and stored at -20ºC.

2.7.15 Luciferase Reporter Assay

HEK-293Ts were seeded out the day before in a 12 well so they would reach 60% confluency at time of transfection. 100 ng of Psi-Check2 was transfected with either 10, 20 or 50 nM of the miR-30c mimic. Alternatively, 100 ng of Psi-Check2 was transfected with 20 or 40 nM of the miR-30c antagomiR. 1 well was transfected with an appropriate concentration of a scrambled control and 1 well was left un-transfected. 24 hours after transfection, media was removed from the cells and they were prepped for the Dual Luciferase Reporter Assay (Promega). Cells were lysed in plate through addition of 100 µL of 1x Passive Lysis Buffer and rocked for 20 minutes at room temperature.

The plate was then frozen overnight at -20 ºC. The plate was removed and allowed to thaw for 30 minutes prior to reading. 10 µL of sample was added to a well of a white 96 well plate in triplicate. LAR II buffer was diluted 1:4 in water with enough LARII was made for 50 µL per well of the 96 well plate used. 50x Stop and Glo substrate was diluted in Stop and Glo buffer to reach 1 x concentration, before diluted again in a 1:4 ratio with water, with a total volume equal to LAR II buffer. Both buffers were protected from light. 50 µL of LAR II was added to each well before it was read by FLUOStar Optima for the firefly readout. 50 µL of Stop and Glo was then added to each well before being read again for the renilla readout. The un-transfected control value was taken away from each plate to equate for background luminescence.

2.7.16 Cell growth assay

200,000 TREx-BCBL1-RTA KD cells were seeded per well of a 12 well in triplicate in 3 ml of media. 10 µL was taken and combined with equal volume of trypan blue (ThermoFisher Scientific) for counting via haemocytometer at 24, 48 and 72 hours post-plating.

2.8 Western blotting

2.8.1 Cell lysis

Cells were lysed in RIPA buffer (50 mM Tris HCl pH7.6, 150 mM NaCl, 1% NP-40v/v, 1 x protease inhibitor). Lysates sonicated (Machine) for 3 x 30 seconds with 30 second breaks in between followed by incubation on ice for 30 minutes. Samples were clarified by centrifugation at 16,000 *x g* at 4 ºC for 10 minutes. 5 µL of sample was taken for a BCA assay if required and then an equal volume of 2 x loading dye (100mM Tris-HCl pH6.8, 20% glycerol (v/v), 4% SDS, bromophenol blue sodium salt) was then added.

2.8.2 Bicinchoninic acid assay (BCA)

Protein was quantified using Pierce BCA Assay kit (ThermoFisher Scientific). BCA Reagent A and BCA reagent B were mixed in a 50:1 ratio. Standards were made using Bovine Serum Albumin (BSA) in modified RIPA buffer at 2000, 1000, 500, 250, 125 µg/ml. 100 µL of reagent mixed was added to 5 µL of sample in a 96 well plate and incubated at 37 ºC for 30 minutes. Absorption was measured at 562 nm on Powerwave XS2 plate reader (BioTek).

2.8.3 SDS-polyacrylamide gel electrophoresis

Protein samples were separated by molecular weight using a Mini-Protean gel electrophoresis system (Biorad). Polyacrylamide gels comprised of an 8-15% resolving region and a stacking region to allow appropriate separation. The resolving region comprised of 8-15% (v/v) acrylamide/bis-acrylamide 37.5:1, 375 mM Tris HCl pH8.8, 0.1% (w/v) SDS, 0.12% APS (v/v) and 0.012% TEMED (v/v). The stacking gel contained 5% (v/v) acrylamide/bis- acrylamide 37.5:1, 125 mM Tris HCl pH6.8, 0.1% (w/v) SDS, 0.08% APS (v/v) and 0.008% TEMED (v/v).

Samples were heated for 10 minutes at 95 ºC before loaded onto a gel with the Dual Colour Protein Ladder (Biorad). Gels were resolved using running buffer (25 mM Tris, 192 mM glycine, 35 mM SDS) at 180 V until the dye front reached the end of the gel.

2.8.4 Transfer

Proteins were transferred to 0.45 µM nitrocellulose membrane using a semi-dry transfer method and a Trans-Blot Turbo Transfer Rig (Biorad). All components were soaked in transfer buffer (20% v/v methanol, 25 mM Tris and 192 mM glycine). Components were layered on the cathode as follows: blotting paper, nitrocellulose membrane, gel and blotting paper before being transferred at 25 V for 30 minutes.

2.8.5 Immunoblotting

Membranes were incubated at room temperature for 1 hour in 5% (w/v) non-fat milk in 1x TBS-T (150 mM NaCl, 50 mM Tris HCl pH7.5, 1% w/v Tween-20). Membranes were then probed overnight in the appropriate primary antibody diluted in 5% non-fat milk in 1 x TBS-T at 4 ºC. Membranes were washed for 3 x 5 minutes in 1 x TBS-T at room temperature before being incubated for 1 hour at room temperature in secondary antibody. Secondary antibodies were horseradish peroxidase conjugated IgG antibody and were diluted 1/5000 in 5% non-fat milk in 1 x TBS-T. These were then washed for 3 x 5 minutes in 1 x TBS-T at room temperature. Proteins were detected via chemiluminscence using either ECL western blotting substrate (Promega), SuperSignal West Dura Extended Duration Substrate (ThermoFisher Scientific) or SuperSignal West Femto Maximum Sensitivity Substrate (ThermoFisher Scientific). Images were captured using G: Box Chemi XX9 imager (Alpha Metrix). Densitometry analysis was performed using FIJI software and normalised to GAPDH.

2.6 RNA-Seq

2.6.1 mRNA-Seq

mRNA Seq was performed previously by Dr Jim Boyne in the Whitehouse lab. Total RNA was extracted from TREx-BCBL1-RTA cells at 0 and 18 hours post lytic induction. mRNA sequencing libraries were created using an Illumina TruSeq Stranded mRNA Sample Prep Kit (Illumina). cDNA libraries were generated and analysed using an Agilent Technologies 2100 Bioanalyze. Libraries were sequenced using a HiSeq (Illumina) 75

sequencing platform. Reads were aligned to the Homo sapiens genome (Genome Reference Consortium GRCh38) using the Burrows-Wheeler Transform method [253]. 2 biological repeats were performed.

2.6.2 miR-Seq

miR-Seq was performed by Dr Becky Foster. Total RNA was extracted from TREx-BCBL1-RTA cells at 0, 16 and 24 hours post lytic induction. Small RNA libraries were prepared using the TruSeq Small RNA Library Prep Kit (Illumina). Samples were assessed using Agilent High Sensitivity D1000 Screen Tape Station and gel electrophoresis and size based extraction performed. cDNA libraries were analysed via Illumina HiSeq (Illumina) by University of Leeds NGS facility and data deposited at Gene Expression Omnibus (GSE186652). Bioinformatic analysis was performed by Dr Anthony Anene. Raw reads were processed for RNA expression using standard bioinformatics pipeline. Quality filtered ($Q < 20$), and adapter trimmed reads (Trimmomatic v0.39) [254] were aligned to the GRCh38/hg38 assembly of the human genome using Bowtie2 (V 2.4.2) or HISAT2 (V 2.1.0) [255, 256] for miRNA and mRNA expressions, respectively. Then the counts in different genomic features were generated using HTSeq (v0.11.1) [257] on human microRNA annotations (miRbase.org) or GRCh38 annotation (GENCODE Release 32). Expression levels were normalised by “counts per million” (CPM). Differential expression (DE) analyses between two KSHV replication times were performed using limma R package. The DE miRNAs and mRNA were defined at adjusted p-value < 0.05. To reduce rate of false discovery rate in the DE analysis only transcripts with at least 1 CPM in 3 samples were included.

2.6.3 circArray

Total RNA was extracted from TREx-BCBL1-RTA at 0, 16 and 24 hours post lytic induction via a TRIzol extraction and Arraystar processed the samples. Briefly circRNAs were enriched through treatment with RNase R. circRNAs were then amplified and transcribed to fluorescently labelled cDNA and then hybridised onto Arraystar Human

circRNA Arrays V2 (8x15K) and incubated at 65 °C for 17 hours before washing and scanned. Agilent Feature Extraction software was used to collect data. Data was processed by R software limma package for differentially expression analysis. 2 biological repeats were performed.

2.6.4 Statistical analysis

Graphical data shown represent mean ± standard deviation of mean (SD) using three or more biologically independent experiments. Differences between means were analysed by unpaired Student's *t*-test. The t-test was chosen as an appropriate statistical test due to the comparison being between 2 independent groups following normal distributions. Statistics was considered significant at $P < 0.05$, with $*P < 0.05$, $**P < 0.01$ and $***P < 0.001$.

Chapter 3

~

Dysregulation of a novel ncRNA axis

3.0 Dysregulation of a novel ncRNA axis

ncRNAs can be generally categorised based on their size. Small ncRNAs are less than 200 nucleotides in length and include many different RNAs, such as, microRNAs (miRNAs), small-interfering, small nuclear, small nucleolar and PIWI-interacting RNAs. In contrast, long ncRNAs (lncRNAs) vary in length from 200 nucleotides to 100 kilobases and examples include long intergenic, long intronic, sense and antisense RNAs. In addition, circular RNAs (circRNAs) are a novel class of lncRNAs, characterised by a covalently closed loop lacking a 5′ cap or 3′ poly (A) tail.

Tremendous interest in the field has delineated many aspects of ncRNA function, showing they are critical regulators of gene expression involved in many cellular pathways. This is reinforced by ncRNA dysregulation being implicated in the development and progression of many human diseases [258]. However, to date, research has mainly focused on the unidirectional ncRNA-mediated regulation of target protein-coding transcripts, particularly by miRNAs. However, emerging evidence suggests the existence of an unexpected interplay amongst ncRNAs that strongly influences how gene expression is regulated [259]. This hidden cross-talk of ncRNA: ncRNA interactions, results in ncRNAs competing against each other for binding to mRNAs, thus acting as competing endogenous RNAs (ceRNAs) and functioning as a part of an ncRNA regulatory network [260]. The ability of ncRNAs to regulate gene expression offers unique possibilities for viruses to modulate virus gene expression and the host response to infection.

In this chapter results highlight a novel ncRNA network which is dysregulated during KSHV lytic replication, comprising the circHIPK3-miR-30c axis.

circHIPK3 (hsa_circ_0000284) is a circRNA derived from exon 2 of the *Homeodomain Interacting Protein Kinase 3 (HIPK3)* gene located on chromosome 11. HIPK3 is a member of the HIPK family, a group of kinases that act as co-repressors of

homeodomain transcription factors [261]. circHIPK3 is exonic and estimated to be 1099 bp in length. Unusually for a circRNA it is highly expressed, with expression often higher than *HIPK3* itself [262]. Studies have shown it is highly expressed in a number of tissues, with an estimated 60-100 copies per cell, furthermore, it appears to be stable with a half-life of over 24 hours [263].

The flanking introns of exon 2 drive circHIP3 formation, surprisingly these introns are very short for a prolifically expressed circRNA, with only ~1000 nucleotides either side sufficient for circularisation. However, these introns are highly complementary with up to 83% identity for over 250 nucleotides and 28 SINE repeats upstream and 51 downstream. Additionally, several of these SINE regions contain ALU elements, which are particularly effective at promoting circularisation, with 88% of human circRNAs having flanking introns containing these elements [226] [263] [264]. As circHIPK3 is highly expressed, its function has been more widely studied than many other circRNAs, with a proposed pro-proliferative function. It is commonly assigned the role of a miRNA sponge, with over 10 different miRNAs proposed to be sponged [264]. Its functionality is therefore tied to the sponging activity of these miRNAs, often targeting tumour-suppressive or anti-proliferative miRNAs [265].

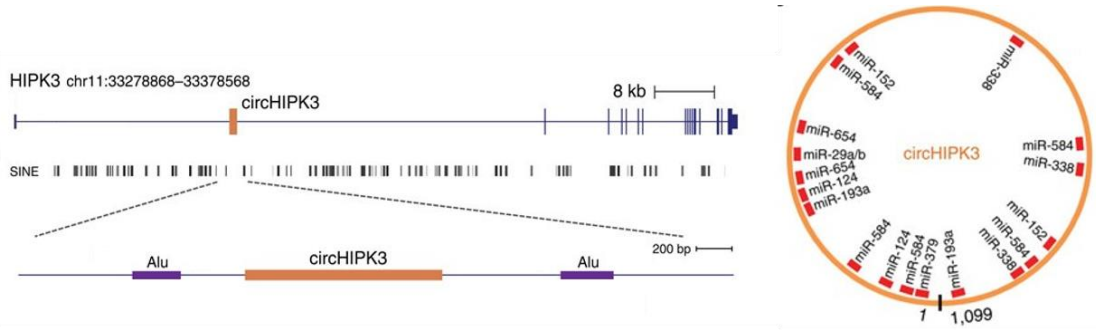

Figure 14: circHIPK3. Schematic of HIPK3 gene with sequence encoding circHIPK3 highlighted with SINE and ALU regions labelled. Several of the proposed miRNA binding sites of circHIPK3 are shown. Modified from [263].

miR-30c is a member of the miR-30 family, which itself contains 5 distinct members encoded on 6 genes distributed between chromosomes 1, 6 and 8 [266] . Members of this family have the same 5′ sequence but a distinct 3′ end and are associated with

regulating a wide range of cellular processes including cell proliferation, apoptosis and differentiation, and as such have been found dysregulated in cancers [267] [268] [269].

Both components of this axis were confirmed to be essential for KSHV lytic replication through a number of assays including late viral protein analysis, viral load and viral reinfection. KSHV-mediated upregulation of circHIPK3 was found to be independent of the parental transcript *HIPK3* and preliminary data suggests that the KSHV ORF57 protein may enhance circHIPK3 backsplicing. Finally, results demonstrate that circHIPK3 functions as a molecular sponge binding miR-30c.

3.1 Identifying dysregulated circRNAs

A preliminary circArray was performed on TREx-BCBL1-RTA cells, a doxycycline promotable KSHV latently infected cell line, at 0 and 24 hours post reactivation to determine if any cellular circRNAs were dysregulated. From this analysis 18 circRNAs were significantly upregulated during lytic replication which were selected for further investigation (Figure 15). The circArray also identified many circRNAs downregulated during lytic replication, however, due to the difficulty in overexpressing circRNAs, it was decided to focus on the upregulated ones.

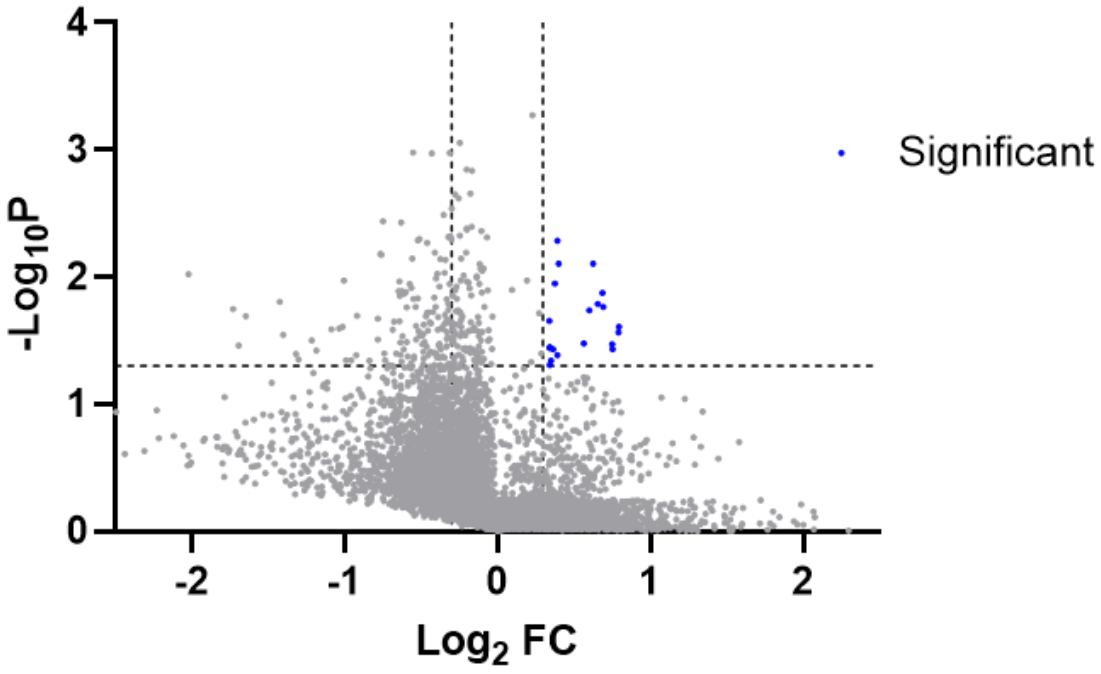

Figure 15: Dysregulated circRNAs. circArray plotting log fold change of circRNAs (log_2 FC) at 0 and 24 hours post KSHV lytic reactivation in TREx-BCBL1-RTA cells against –log of p value ($-log_{10}P$). Horizontal dotted line represents cut off for a p value= 0.05. Vertical horizontal lines represent cut offs for a fold change of 0.6 or 1.4 in circRNA levels between 0 and 24 hours. Significantly upregulated circRNAs above 1.4 fold upregulation were selected for further investigation and are highlighted blue. .N=2.

3.2 Screening circRNAs

To validate the circArray, divergent primers were designed; these primers can only amplify a transcript if a backsplicing event has occurred (Figure 16A). Targets were chosen through a literature search of the more abundantly expressed and characterised circRNAs and levels were analysed via qPCR in TREx-BCBL1-RTA cells, at 0 and 24 hours post lytic induction. circHIPK3, circSPECC1, circPVT1 and circITCH were successful amplified, with both circHIPK3 and circSPECC1 levels significantly upregulated at 24 hours (Figure 16B-D).

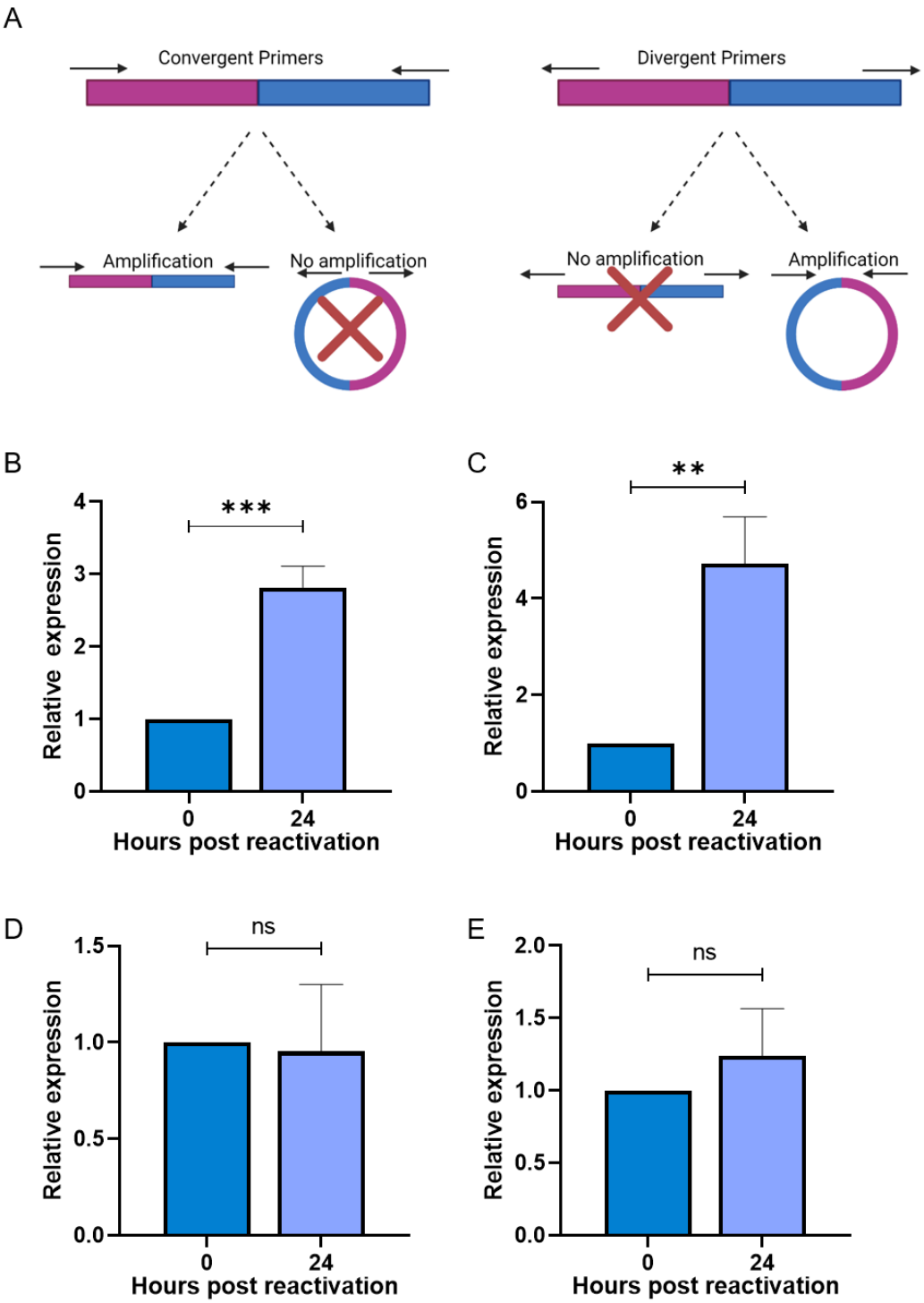

Figure 16: circHIPK3 is upregulated in lytic replication. Divergent primers were designed **(A)** and levels of circHIPK3 **(B).** circSPECC1 **(C).** circPVT1 **(D)** and circITCH **(E).** were analysed during lytic replication via qPCR N=3. Fold change relative to 0hr cells. Error bars represent SD. *P < 0.05, **P < 0.01 and ***P < 0.001 (unpaired Student's t-test

3.3 circRNAs upregulation is specific to the circular transcript

The degree of which circRNA levels are linked to the primary linear transcripts from which they are derived is debated and appears to be highly variable, with some circRNAs

simply being a by-product from regular splicing and thus dependent on the linear levels. However, in other cases the ratio of linear: circular transcripts is not just dependent on the primary linear levels implicating active biogenesis of circRNAs. To identify whether circHIPK3 and circSPECC1 levels were upregulated independently from the linear transcript, the levels of *HIPK3* and *SPECC1* were analysed using qPCR. *HIPK3* levels remained constant while *SPECC1* levels decreased, suggesting that the upregulation of circHIPK3 and circSPECC1 is specific to the circular form and an active change in backsplicing levels (Figure 17A-B).

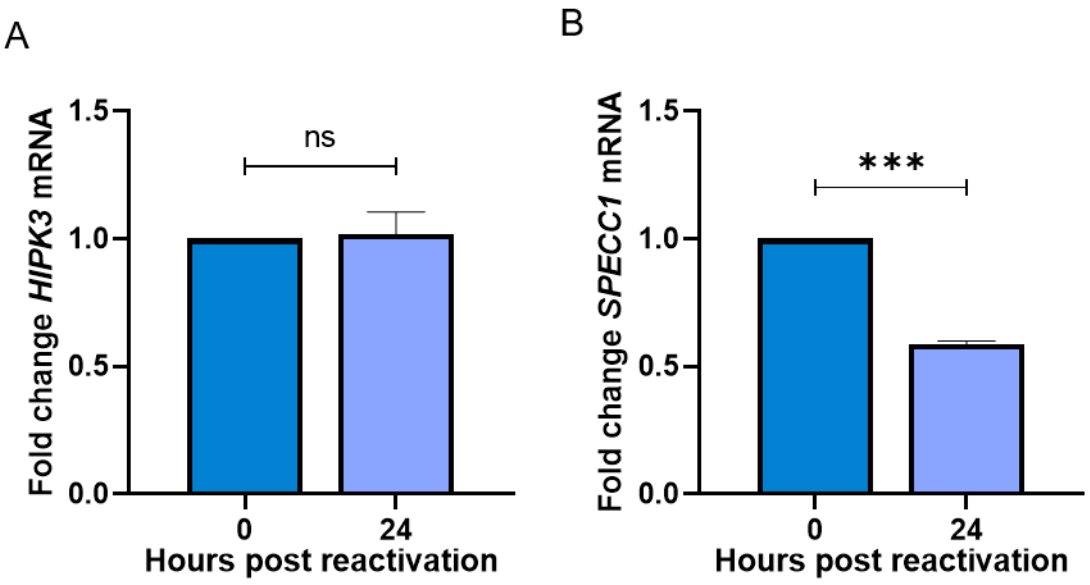

Figure 17: circRNA upregulation is independent of linear levels. Levels of *HIPK3* **(A)** and *SPECC1* **(B)** were analysed via qPCR at 0 and 24 hours post lytic induction. N=3. Fold change relative to 0hr cells. Error bars represent SD. *P < 0.05, **P < 0.01 and ***P < 0.001 (unpaired Student's t-test).

3.4 circHIPK3 and circSPECC1 localise to different areas

circRNA functionality is tied to its localisation, with nuclear intronic circRNAs involved in transcriptional regulation and protein binding, whereas, cytoplasmic circRNAs are primarily involved in miRNA sponging and protein binding. To aid in elucidation of circHIPK3 and circSPECC1 functions during KSHV lytic replication, a rudimentary subcellular localisation assay was therefore designed, with RNA extracted and analysed via qPCR from whole cell (WCF), cytoplasmic and nuclear fractions, although due to the lack of a consistent housekeeper, any nuclear expression of circRNAs was too variable

for error bars. The success of the subcellular fractionation was analysed by western blot using Lamin B and GAPDH-specific antibodies (Figure 18A). Notably circHIPK3 was only detectable in the WCF and cytoplasmic fraction, in contrast, circSPECC1 was mostly amplified in the WCF and nuclear fraction (Figure 18B-C)

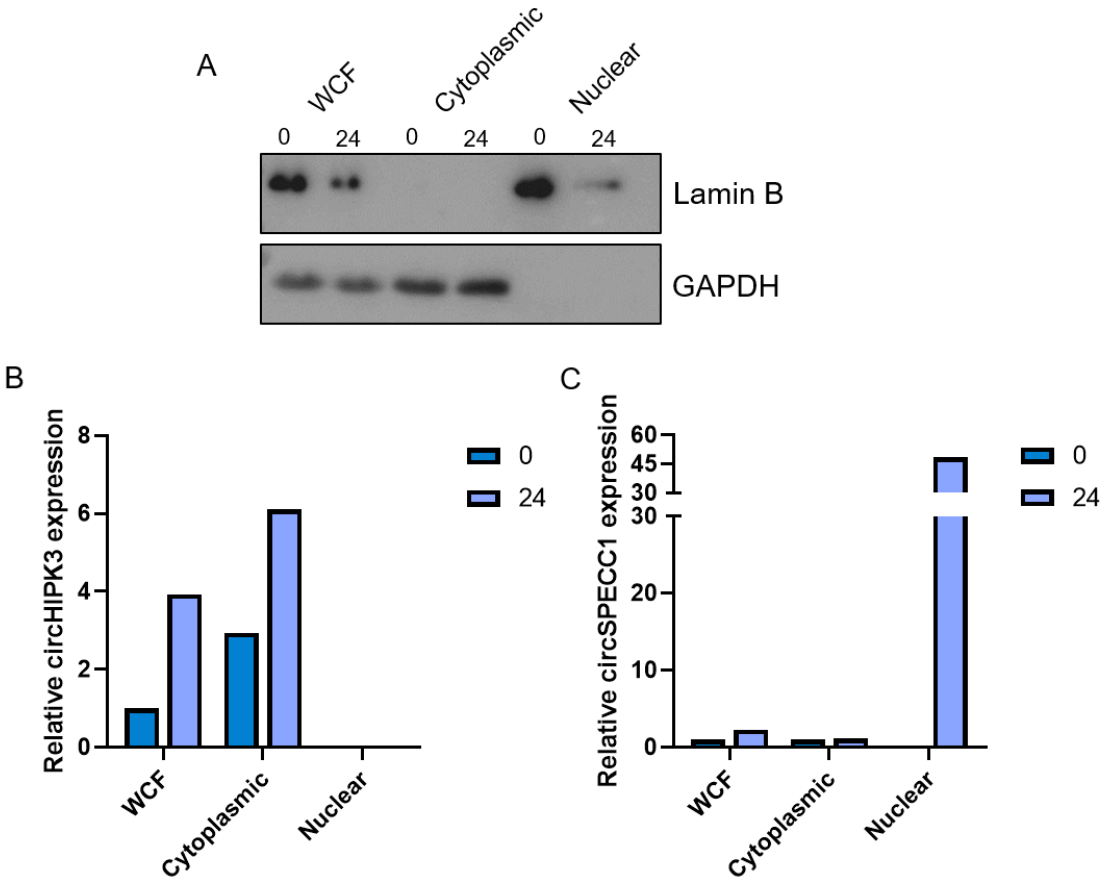

Figure 18*:* circHIPK3 and circSPECC1 localise to distinct areas. Subcellular fractionation of TREx-BCBL1-RTA cells at 0 and 24 hours into whole cell fractions (WCF), cytoplasmic and nuclear fractions. **(A).** Western blot of GAPDH and Lamin B. **(B).** Representative qPCR analysis of circHIPK3 levels. **(C).** qPCR analysis of circSPECC1 levels. Fold change relative to 0hr WCF. N=3.

3.5 circHIPK3 is a cytoplasmic circRNA

Due to our interest in the possibility of KSHV dysregulating host ncRNA regulatory networks through circRNA-miRNA interactions, we therefore selected circHIPK3 for further investigation due to its apparent cytoplasmic localisation. However, due to the lack of a consistent RNA housekeeping gene across all fractions, with no RNA equally expressed in the nucleus and in the cytoplasm for ΔΔCT analysis, we further confirmed circHIPK3 localisation using RNA FISH, with fluorescent probes against the unique

backsplice region. Results confirmed the cytoplasmic localisation of circHIPK3 (Figure 19A-B). As a result of the limited sequence the fluorescent probes could target, signal intensity was weak preventing co-staining with antibodies against cytoplasmic markers. Future work could utilise a GFP-fusion protein to confirm this localisation.

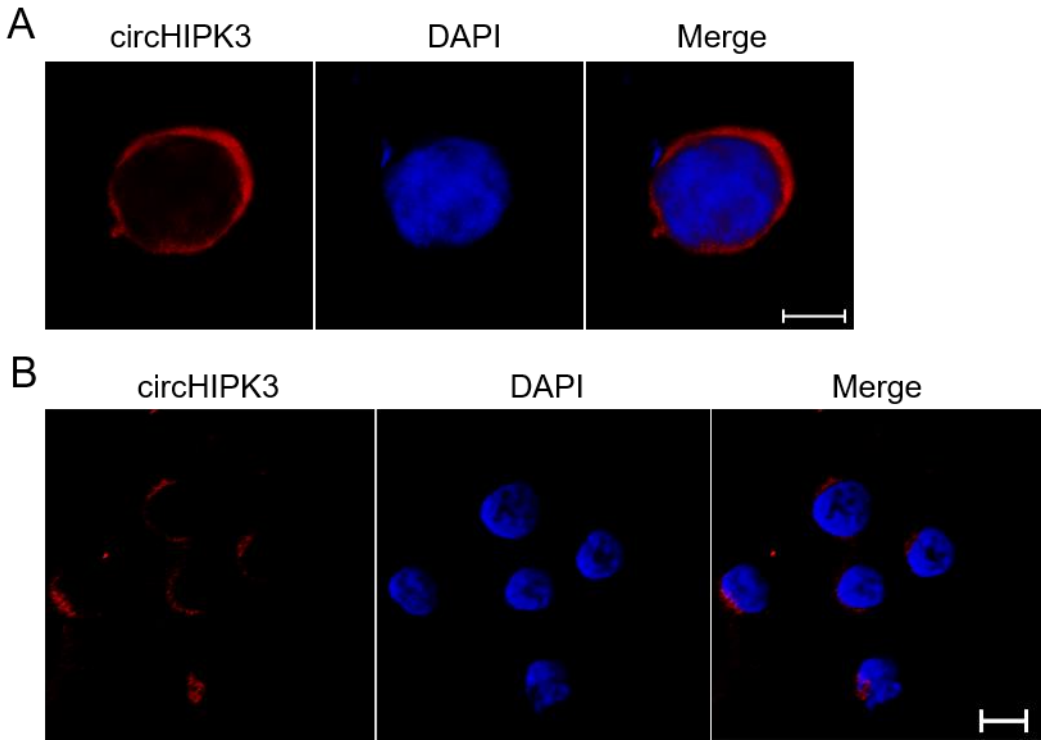

Figure 19: circHIPK3 localises to the cytoplasm. RNA FISH was performed in 16 hour reactivated TREx cells with probes against circHIPK3 (red) and DAPI (blue). **(A)** Scale bars represent 5 μM. **(B).** Scale bars represent 10 μM.

3.6 miRNAs are dysregulated upon lytic replication

To identify potential miRNAs sponged by circHIPK3, we took advantage of a previously produced miR-Seq dataset (Dr Becky Foster, Whitehouse laboratory), which identified dysregulated miRNAs at 16 and 24 hours post KSHV lytic replication compared to the latent phase. This data was analysed using Limma R and a cut off of 1 count per million (CPM) by Dr Anthony Anene (QMUL). Results highlighted 27 of the most dysregulated miRNAs (Figure 20A). Since circHIPK3 is upregulated we focused on downregulated miRNAs, and sequence alignment was performed using RNA binding prediction software (Freiburg RNA tools, IntaRNA-RNA-RNA interactions and University of Bielefeld BiBiSery RNAhybrid program) to identify potential complementary sequences of

downregulated miRNAs. Results highlighted circHIPK3 (hsa_circ_0000284) contained complementary sequences to both miR-29b and miR-30c, with both software identifying 2 and 3 high score binding sites for miR-29b and miR-30c respectively (Figure 20B).

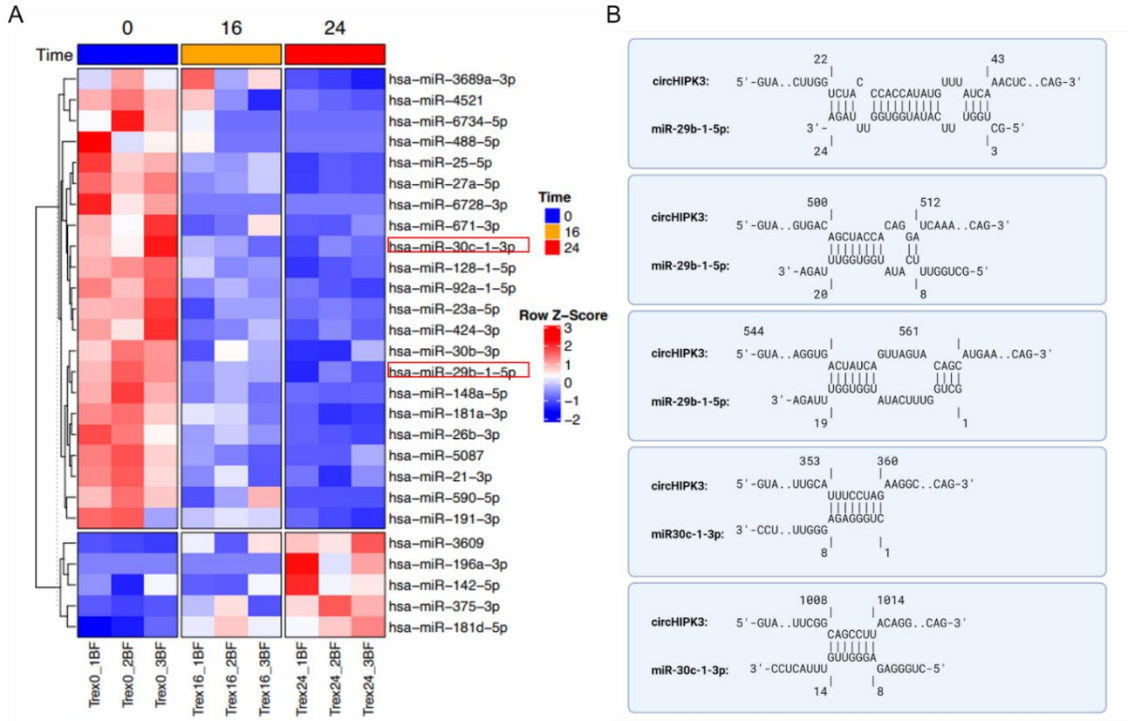

Figure 20: Several miRNA potentially sponged by circHIPK3 are dysregulated during lytic replication. (**A).** Heat map of dysregulated miRNAs identified via miR-Seq at 0, 16 and 24 hours in TREx-BCBL1-RTA cells with miR-29b and miR-30c highlighted. **(B).** circHIPK3 and miR-30c/ miR-29b predicted alignments using Freiburg RNA tools and BiBiSery RNAhybrid program).

3.7 miR-29b and miR-30c are downregulated in KSHV infection

Although miR-Seq identified both miRNAs as significantly downregulated upon lytic induction, we wanted to confirm the validity of this global screen. Firstly a number of publically available microarray GSE datasets looking at miRNA levels were analysed (Figure 21A-D). Both miR-29b and miR-30c were downregulated in KS tumour cells compared to adjacent non-tumourgenic cells. KS tumours contain a mixture of both latent and lytic replicating cells, suggesting these miRNAs are downregulated during KSHV infection. Analysis of a second dataset found miR-30c was also downregulated in KSHV and KSHV/EBV-infected cells, interestingly however, it was upregulated in EBV only infected cells, suggesting miR-30c downregulation is not a pan-gammaherpesvirus mechanism. In this second dataset miR-29b was downregulated in KSHV, KSHV/EBV

and EBV infected cells, however, to a much smaller non-significant level. Therefore to examine the levels of both miRNAs in TREx-BCBL1-RTA cells upon lytic induction, miScript qPCR was utilised at 0 and 24 hours. Results confirmed both miR-29b and miR-30c were significantly downregulated at 24 hours (Figure 21E).

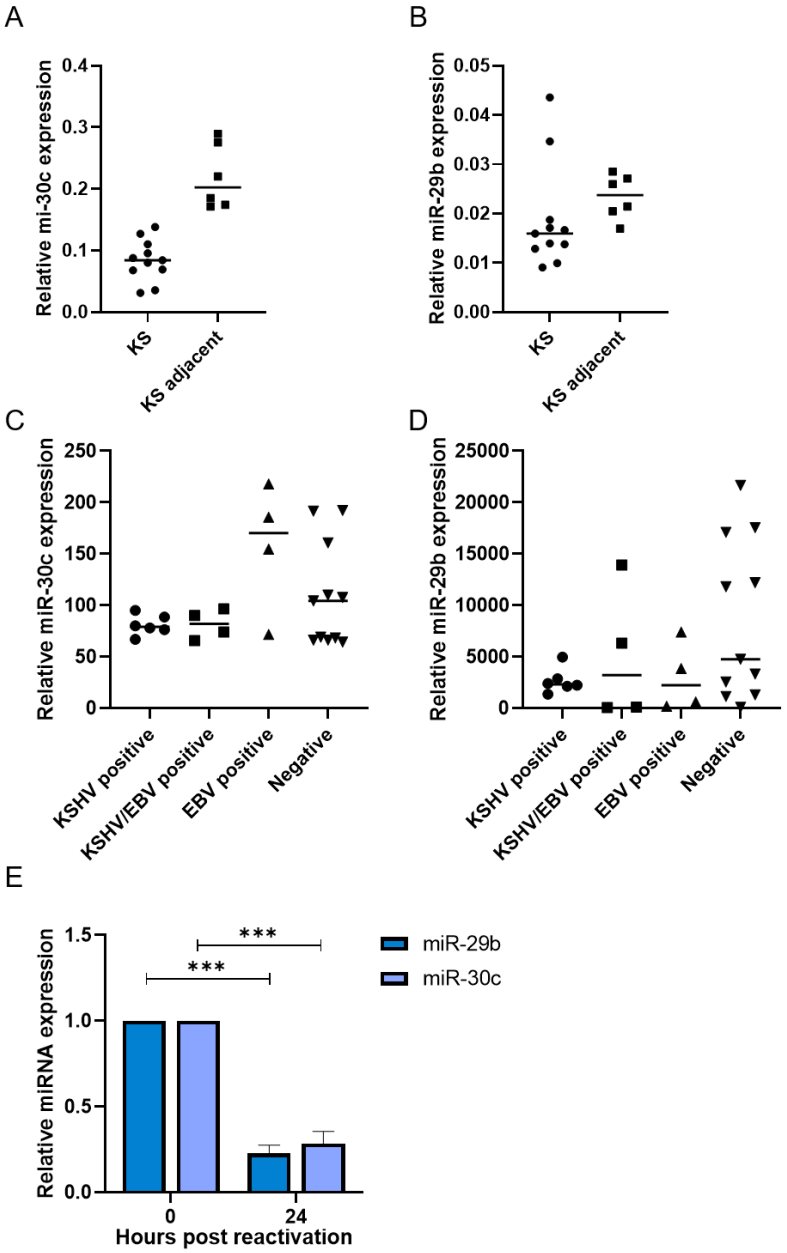

Figure 21: miR-29b and miR-30c are dysregulated during lytic replication. Expression of miR-30c **(A)** and miR-29b **(B)** in KS tissue vs adjacent tissue from GSE 5625. Expression of miR-30c **(C)** and miR-29b **(D)** in KSHV positive, EBV positive, KSHV/EBV positive and negative cells from GSE 18437. **(E).** qPCR analysis of miR-29b and miR-30c in TREx-BCBL1-RTA cells at 0 and 24 hours. SNORD68 was used as a housekeeping gene. (n=3). Fold change relative to 0hr cells. Error bars represent SD. *P < 0.05, **P < 0.01 and ***P < 0.001 (unpaired Student's t-test).

3.8 miR-29b and miR-30c are inhibitory to KSHV lytic replication

As both miR-29b and miR-30c were downregulated we hypothesised they are inhibitory to KSHV lytic replication. To test this hypothesis miRNA downregulation was reversed by transfecting miRNA mimics into TRExBCBL1-RTA cells. TREx-BCBL1-RTA cells have low transfection efficiencies with plasmids due to being a suspension cell line, consequently miRNA mimics were transfected with RNAi Max and successful transfections were confirmed using qPCR (Figure 22A). Once transfection efficiency was optimised, the effect of increasing miR-29b and miR-30c levels upon KSHV lytic replication was analysed by comparing levels of the late ORF65 protein at 24 and 48 hours after viral reactivation via immunoblotting. Transfection of both miRNAs led to a significant decrease in ORF65 levels (Figure 22B-C), suggesting that both miR-29b and miR-30c overexpression are detrimental to KSHV lytic replication. Further assays were then performed to assess the effect of increased miR-30c levels on viral load and infectious virion production. miR-30c was selected for further analysis due to its higher mimic expression level. Briefly for viral load assays, TREx-BCBL1-RTA cells were reactivated for 72 hours and viral DNA collected before analysis via qPCR to confirm viral genome replication. To assess infectious virion production, reactivated TREx-BCBL1-RTA cell supernatants were collected at 72 hours post reactivation and incubated with naïve HEK-293T cells before analysis via qPCR. Both viral load and infectious virion production were significantly reduced upon transfection of the miR-30c mimic over the scrambled control (Figure 22D-E). Taken together these data suggest that KSHV lytic replication is impacted by the overexpression of miR-29b and miR-30c, suggesting they are downregulated during KSHV lytic replication due to their inhibitory effect.

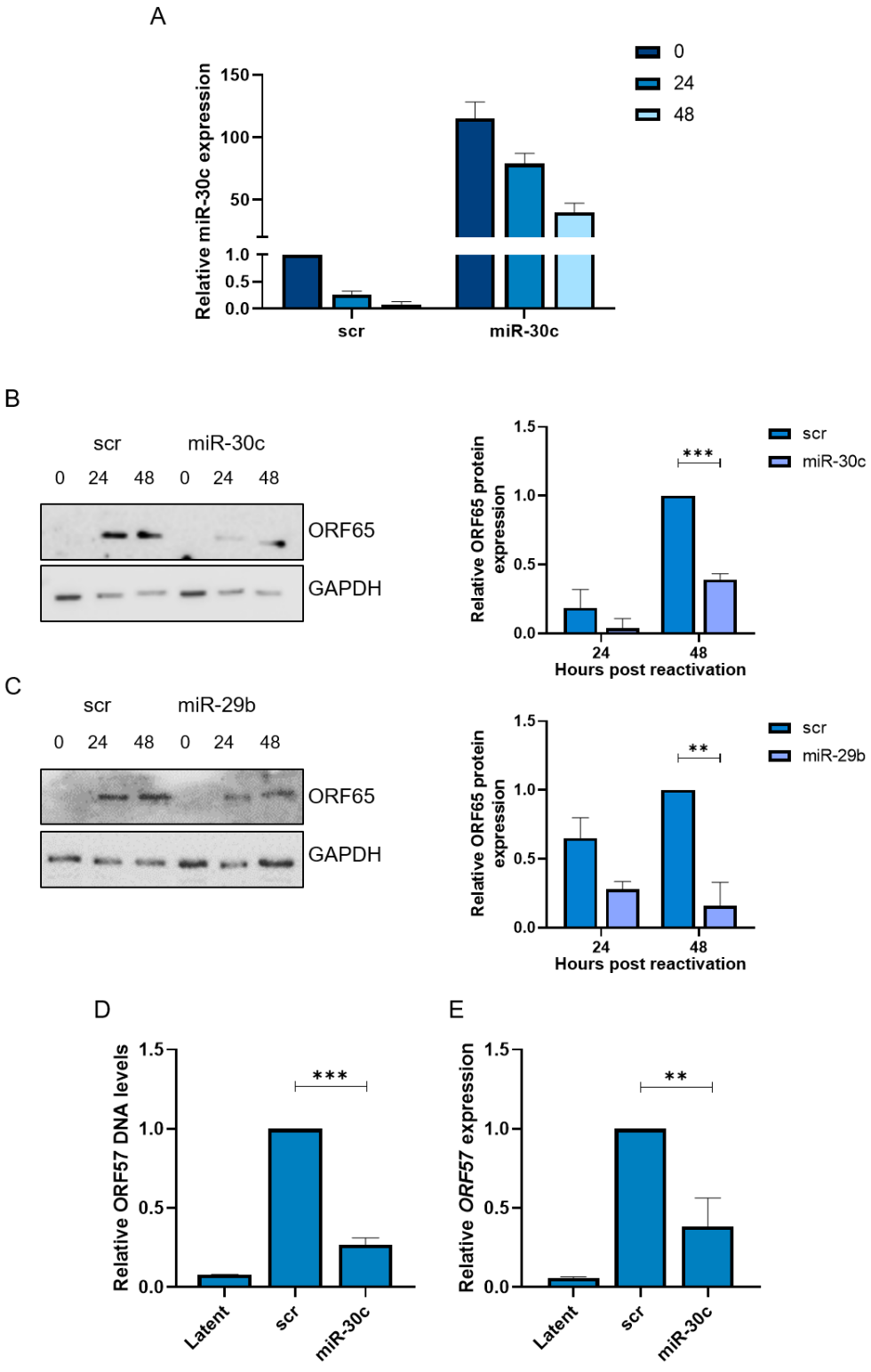

Figure 22 : miR-30c and miR-29b are anti-viral. (A). qPCR analysis of miR-30c levels, TREx-BCBL1-RTA cells were transfected with a scrambled control or a miR-30c mimic before lytic replication induced 24 hours after transfection. miR-30c levels were analysed at 0, 24 and 48 hours post lytic induction. SNORD68 was used as a housekeeper (n=3). **(B).** Representative western blot of ORF65 in miR-30c vs scr transfected TREx-BCBL1-RTA cells with densitometry analysis. **(C).** Representative western blot of ORF65 in miR-29b vs scr transfected TREx-BCBL1-RTA cells with densitometry analysis. **(D).** Viral load of miR-30c vs scr transfected TREx-BCBL1-RTA cells. (E). Viral reinfection of miR-30c vs scr transfected TREx-BCBL1-RTA cells. N=3. Error bars represent SD. *P < 0.05, **P < 0.01 and ***P < 0.001 (unpaired Student's t-test).

3.9 miR-29b and miR-30c are downregulated at the mature level

For circHIPK3 to act as a miRNA sponge, the downregulation of both miRNAs should occur at the mature level, with any occurrence earlier more likely due to transcriptional or processing changes. Therefore to ascertain whether circHIPK3 acts as a miRNA sponge, the primary and pre levels of miR-29b and miR-30c were analysed by qPCR. Results confirmed at 24 hours post reactivation, that the primary and pre levels were not downregulated, implicating that KSHV-mediated regulation of miR-29b and miR-30c was at the mature level (Figure 23A-B).

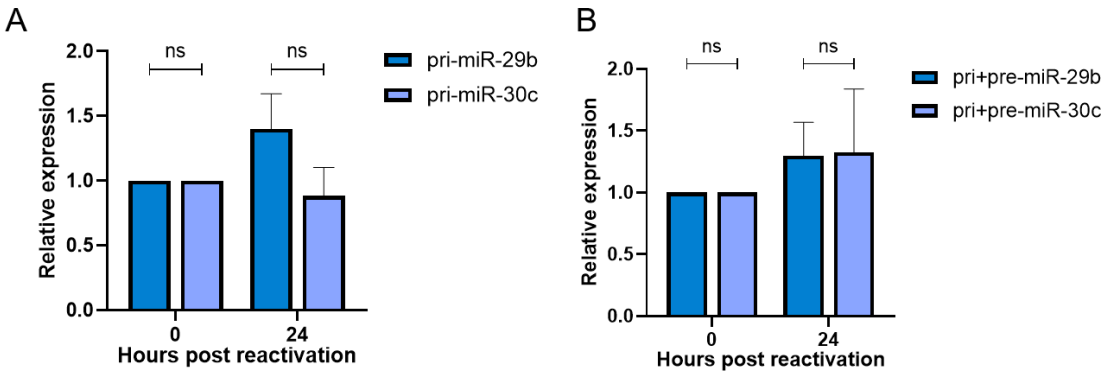

Figure 23*:* miR-29b and miR-30c are downregulated at the mature level. qPCR analysis of primary miR-29b and miR-30c **(A)** and pri + pre miR-30b and miR-30c **(B)** at 0 and 24 hours. GAPDH was used a housekeeper. (n=3). Fold change relative to 0hr cells. Error bars represent SD.

3.10 circHIPK3 associates with the miRNA machinery

To further support a potential sponging activity of circHIPK3, Ago2 RNA immunoprecipitations (RIPs) were performed to determine whether circHIPK3 associates with the miRNA machinery. Latent or 24 hour post reactivation TREx-BCBL1-RTA lysates were incubated with an Ago2 antibody bound to magnetic beads, allowing isolation and identification of bound RNAs. circHIPK3 was found significantly enriched over GAPDH, the negative control (Figure 24A). Unfortunately, a positive control for the Ago2 RIP was not possible, as the only confirmed RNA binders were miRNAs, which require different RNA processing to circRNAs, and due to low yields only one could be performed with each RIP. However, in optimising the RIP, miRNAs were confirmed to bind effectively, validating the RIP. Interestingly transfection of TREx-BCBL1-RTA cells

with either miR-29b or miR-30c 24 hours prior to the RIP significantly increased the association of circHIPK3 with Ago2, with this phenomenon previously observed with other miRNA sponging circRNAs. Whereas transfection of a miR-30c antagomiR, had the opposite effect, decreasing the association of circHIPK3.

To confirm direct binding between circHIPK3 and miR-29b or miR-30c, biotinylated miR-29b and miR-30c mimics were also transfected into HEK-293T cells before a RIP using streptavidin beads was performed. circHIPK3 was found significantly enriched over GAPDH when cells were transfected with miR-29b or miR-30c mimics over scrambled controls, implying a direct interaction (Figure 24B).

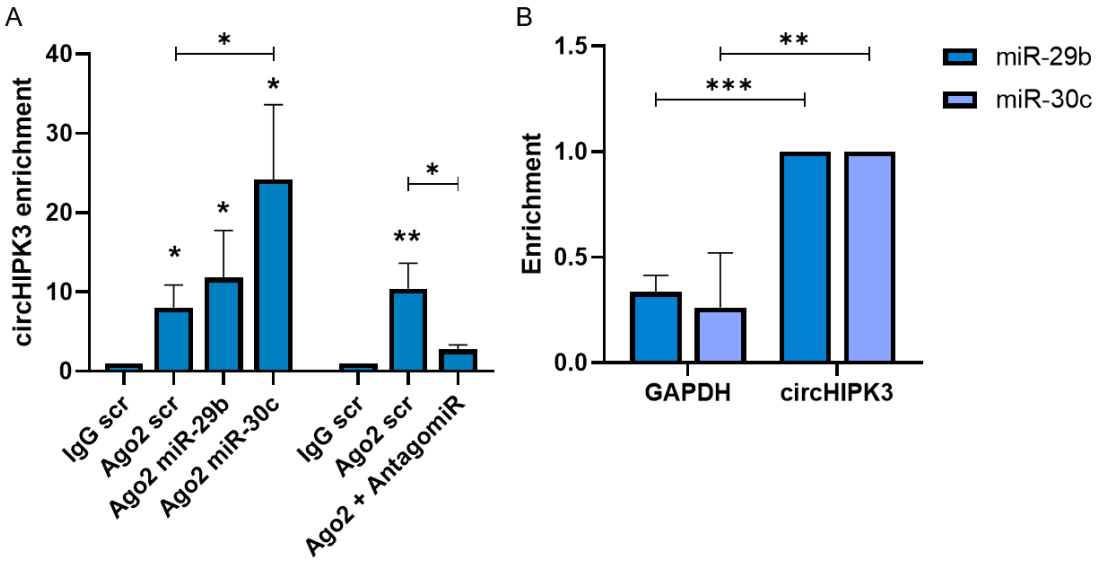

Figure 24: circHIPK3 binds to miRNAs. (A). qPCR analysis of circHIPK3 enrichment over GAPDH in Ago2 RIPs of TREx-BCBL1-RTA cells transfected with scr, miR-29b mimic, miR-30c mimic or a miR-30c antagomiR. Each repeat is normalised to an IgG scr RIP. **(B)** qPCR analysis of streptavidin RIPs of circHIPK3 or GAPDH enrichment in HEK-293T cells transfected with either biotinylated miR-29b or miR-30c over scr. N=3. Error bars represent SD. *P < 0.05, **P < 0.01 and ***P < 0.001 (unpaired Student's t-test).

3.11 miR-29b-3p is not regulated by circHIPK3

Our alignment identified miR-29b-5p as potentially being sponged by circHIPK3, however, miR-29b-3p is the supposedly more highly expressed transcript, with the literature suggesting that over 90% of miR-29b levels are the 3p strand [270]. Therefore, the levels of miR-29b-3p were investigated compared to the 5p strand. miR-Seq showed

that in contrast to miR-29b-5p, no difference in miR-29b-3p levels were observed between 0 and 24 hours post-lytic induction (Figure 25A and C). In addition, miScript qPCR confirmed no change in levels during lytic replication (Figure 25B). Therefore we excluded miR-29b-3p as a potential circHIPK3 target.

A

miRNA name	logFC	AveExpr	P.Value
hsa-miR-29b-1-5p	-3.1992	5.40339	0.00015
hsa-miR-30c-1-3p	-2.2277	4.80493	0.00079
hsa-miR-29b-3p	-0.6202	2.12662	0.43493

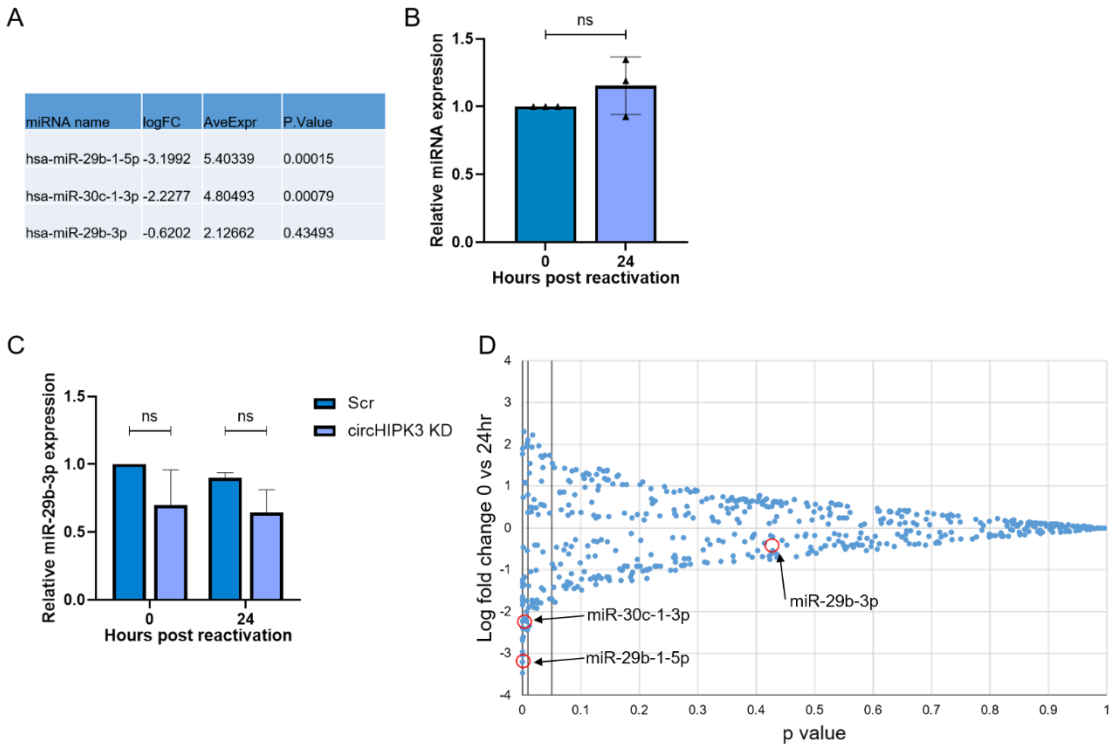

Figure 25: miR-29b-3p is not a target of circHIPK3. (A). Table summarising log fold change, average expression and P value of changes between 0 and 24 hours in the miR-Seq of miR-29b-3p, miR-29b-5p and miR-30c. **(B).** qPCR showing miR-29b-3p levels at 0 and 24 hours. **(C).** Dot plot of miR-Seq log fold changes against p values with lines of p value significances (0.05, 0.01, 0.001). The 3 miRNAs of interest: miR-30c, miR-29b-5p and miR-29b-3p, are marked with red circles. N=3. Error bars represent SD. Fold change relative to 0hr cells.

3.12 circHIPK3 is important for virus replication

To assess the importance of circHIPK3 on KSHV lytic replication, TREx-BCBL1-RTA cells were transduced with lentivirus-based shRNAs targeting the unique backsplice of circHIPK3, these shRNAs were manually designed using standard siRNA guidelines to straddle the backsplice site. qPCR confirmed successful depletion of circHIPK3, with only a small depletion of ~30% in latent cells, however a larger depletion of around 75% at 24 hours. Importantly knockdown had little effect on the parental *HIPK3* transcript (Figure 26A-B), meaning any effects of the knockdown was solely due to circHIPK3

depletion. The effect of circHIPK3 depletion on KSHV lytic replication was then evaluated using immunoblotting to assess if there were any effects on the downstream expression of the late KSHV ORF65 protein. circHIPK3 depletion resulted in a significant reduction in ORF65 protein production compared to the scrambled control (Figure 26C-D). The effect of circHIPK3 knockdown was further evaluated using viral DNA load and reinfection assays, with a significant decrease in both viral load and infectious virion production being observed upon depletion (Figure 26E-F). Together this confirms that KSHV specifically requires the function of circHIPK3 to undergo efficient lytic replication and infectious virion production.

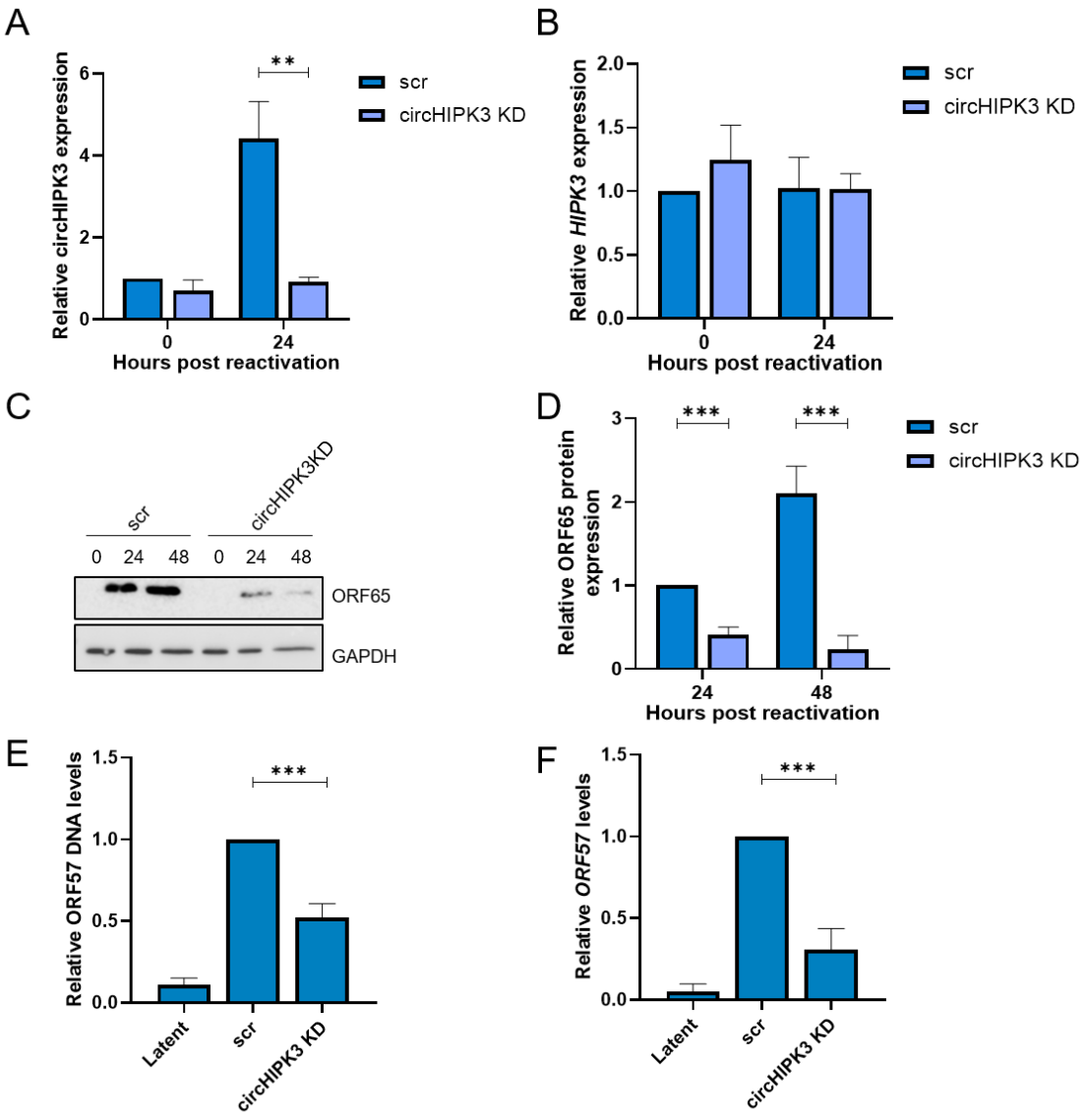

Figure 26: circHIPK3 is important for KSHV. (A). qPCR analysis of circHIPK3 levels in scr and stable circHIPK3 KD TREx-BCBL1-RTA cells. **(B).** qPCR analysis of HIPK3 in scr and stable circHIPK3 KD TREx-BCBL1-RTA cells. **(C).** Representative western blot analysis of ORF65 levels in scr and circHIPK3 KD cells. **(D)** Densitometry analysis of ORF65 protein levels in scr and circHIPK3 KD cells. **(E).** Viral load in scr and circHIPK3 KD cells. **(F).** Viral reinfection in scr vs circHIPK3 KD cells. N=3. Error bars represent SD. *P < 0.05, **P < 0.01 and ***P < 0.001 (unpaired Student's t-test).

3.13 circHIPK3 targets miR-30c

With the importance of circHIPK3 on KSHV lytic replication confirmed, we next assessed what effect circHIPK3 depletion had upon miR-30c levels. Previous results would suggest that if increased levels of circHIPK3 lead to decreases in miR-30c levels due to circHIPk3 sponging activity, miR-30c levels should be partly restored in circHIPK3 KD cells. Therefore the levels of miR-30c were investigated in scr vs circHIPK3 KD cells at 0 and 24 hours post lytic-induction. As previously observed miR-30c was downregulated

in scrambled control cells at 24 hours post reactivation, in contrast no significant downregulation was observed in miR-30c levels upon circHIPK3 depletion. To confirm this effect was specific to the circHIPK3-miR-30c axis and not a result of reduced viral replication, levels of another KSHV-mediated downregulated miRNA: miR-27a were examined. Unlike miR-30c there was no significant change in miR-27a levels between scr and KD cells (Figure 27). Together this shows that KSHV-mediated circHIPK3 dysregulation specifically affected miR-30c levels.

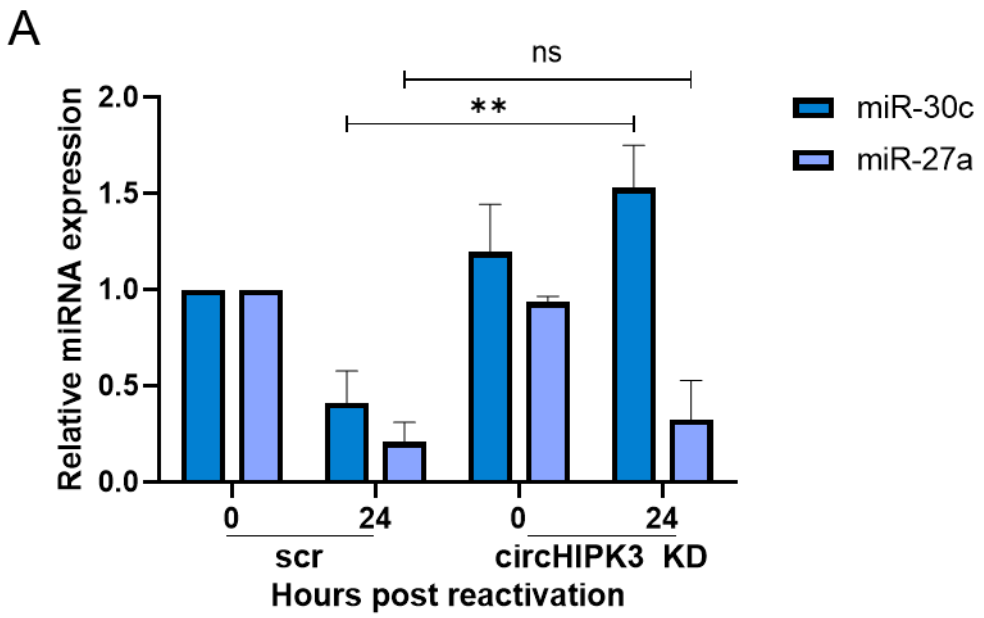

Figure 27: miR-30c is a target of circHIPK3. (A). qPCR analysis of miR-27a and miR-30c levels in scr and circHIPK3 KD TREx-BCBL1-RTA cells at 0 and 24 hours. N=3. Error bars represent SD. *P < 0.05, **P < 0.01 and ***P < 0.001 (unpaired Student's t-test). Fold change relative to 0hr cells.

3.14 circHIPK3 upregulation occurs early in lytic replication

KSHV lytic replication occurs in a temporal cascade, with proteins expressed at immediate-early, early and late time points. To determine if any KSHV-encoded proteins are involved in dysregulating circHIPK3 levels, circHIPK3 levels were firstly assessed at different time points post reactivation. Total RNA was isolated from TREx-BCBL1-Rta cells at 0, 8, 16, 24, 48 and 72 hours post reactivation. qPCR showed a clear upregulation between 16 and 24 hours (Figure 28A), suggestive that an immediate early or early protein is behind the dysregulation. Further small increases were also observed from 24

hours onwards and these are likely due to the inherent stability of circRNAs. Notably linear *HIPK3* levels showed no significant change up to 24 hours followed by a downregulation suggesting that the increase in circHIPK3 is independent of *HIPK3* levels (Figure 28B).

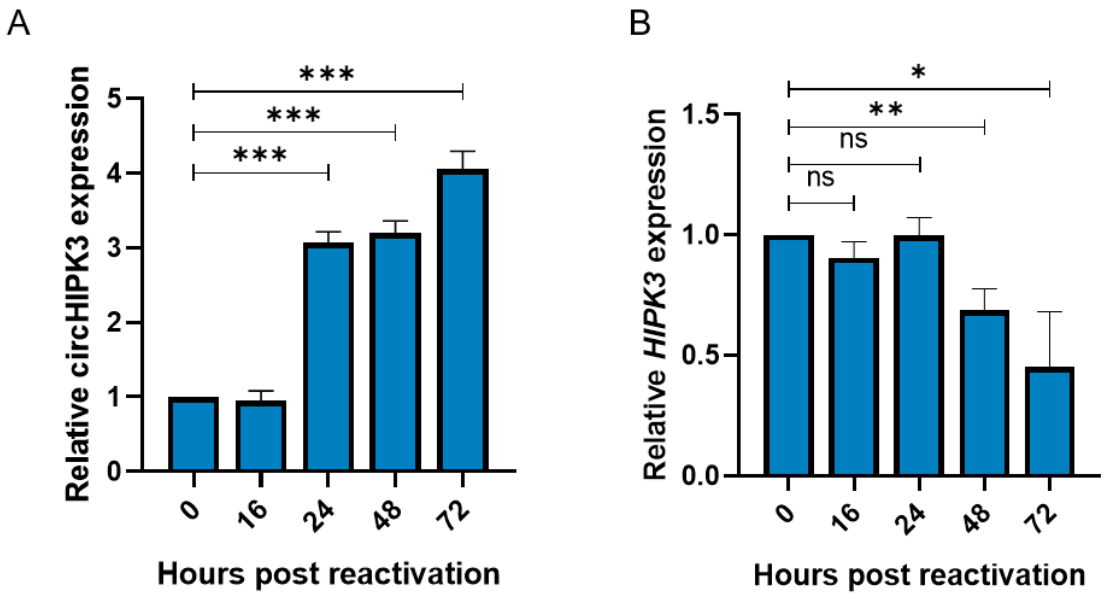

Figure 28: circHIPK3 is upregulated between 16 and 24 hours. qPCR analysis at 0, 16, 24 and 48 hours post lytic induction of circHIPK3 **(A)** and HIPK3 **(B).** N=3. Error bars represent SD. *P < 0.05, **P < 0.01 and ***P < 0.001 (unpaired Student's t-test). Fold change relative to 0hr cells.

3.15 ORF57 leads to increases in circHIPK3

After establishing circHIPK3 upregulation occurred early in KSHV lytic replication, two key early KSHV proteins were tested to determine whether they were sufficient to induce circHIPK3 levels independently. HEK-293T cells were transfected with control GFP, KSHV RTA/ORF50-GFP or ORF57-GFP expression constructs [271] [272] and circHIPK3 levels assessed by qPCR at 24 hours post transfection. Western blotting confirmed successful transfections (Figure 29A). Notably expression of ORF57-GFP alone was sufficient to increase circHIPK3 levels (Figure 29B) and this upregulation was dose-dependent (Figure 29C-D).

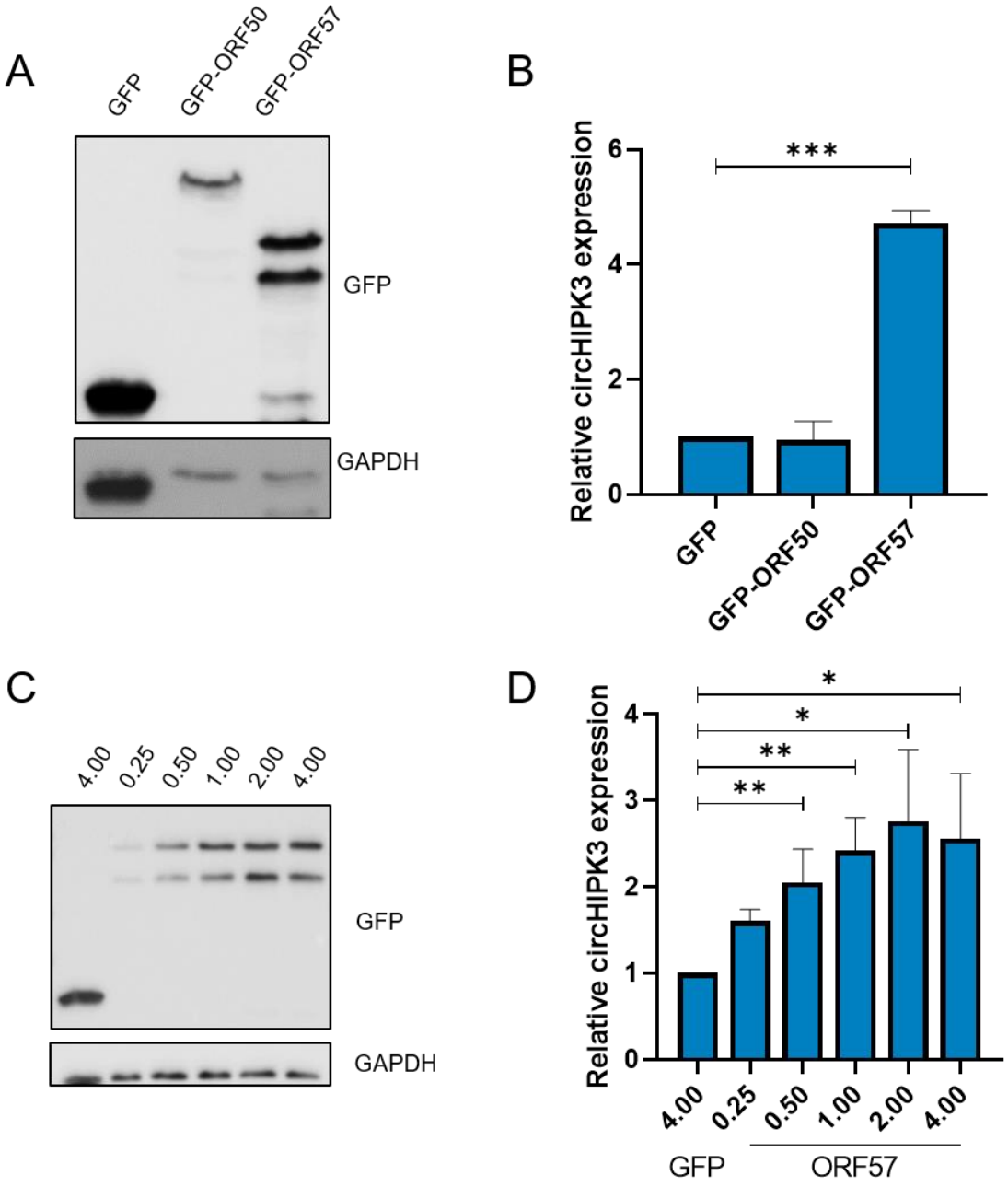

Figure 29: ORF57 upregulates circHIPK3 levels. (A). Western blot for GFP in GFP, GFP-ORF50 and GFP-ORF57 transfected HEK-293Ts. **(B).** qPCR analysis of circHIPK3 levels in in GFP, GFP-ORF50 and GFP-ORF57 transfected HEK-293Ts. **(C).** Western blot GFP in dose-dependent GFP-ORF57 transfected HEK-293Ts. **(D).** qPCR analysis of circHIPK3 levels in dose-dependent GFP-ORF57 transfected HEK-293Ts. N=3. Error bars represent SD. *P < 0.05, **P < 0.01 and ***P < 0.001 (unpaired Student's t-test). Fold change relative to GFP transfected cells.

3.16 ORF57 binds directly to circHIPK3

KSHV ORF57 is a multifunctional RNA binding protein, involved in multiple stages of viral and cellular RNA processing. Therefore, RNA immunoprecipitations were

performed in control or ORF57-GFP transfected HEK-293T cells utilising GFP-TRAP beads, to assess whether ORF57 interacted with circHIPK3.

RIPs were performed at 8, 16 and 24 hours post-transfection in HEK-293Ts. Levels of circHIPK3 and *HIPK3* bound were determined via qPCR at each of the time points, along with GAPDH as a negative control. Pre-BTG1 was used as a positive control for ORF57 binding, with previous literature identifying strong ORF57 binding to the pre-BTG1 intron [273]. At 8 hours ORF57 was only bound to the positive control, pre-BTG1 (Figure 30A). However, by 16 hours strong binding was observed for both *HIPK3* and circHIPK3 (Figure 30B), equal to the binding of pre-BTG1. In contrast, at 24 hours ORF57 showed stronger association with circHIPK3 compared to the linear form (Figure 30C). This suggests that ORF57 binds both the linear and circular forms of HIPK3, with the ratio of binding changing over time.

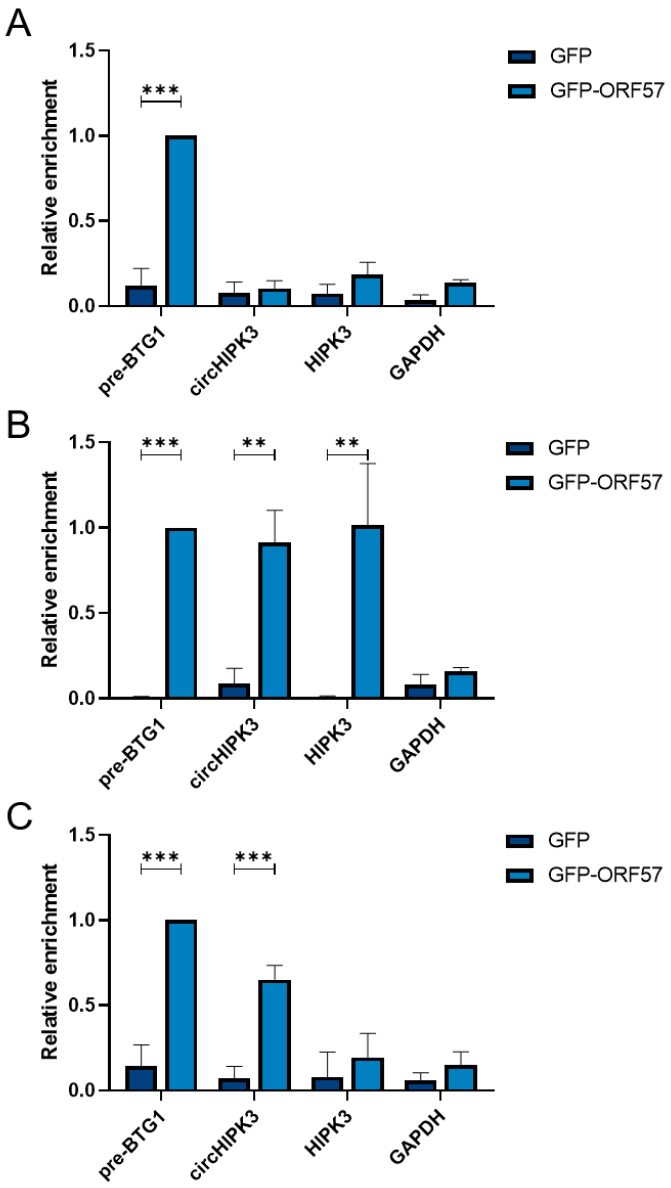

Figure 30: ORF57 binds directly to circHIPK3. qPCR analysis of RNA in GFP RIPs in GFP or ORF57-GFP transfected HEK-293Ts at 8 **(A)** 16 **(B)** and 24 hours **(C)** post-transfection, respectively. Enrichment is normalised to pre-BTG1 in GFP-ORF57 transfected cells. N=3. Error bars represent SD. *P < 0.05, **P < 0.01 and ***P < 0.001 (unpaired Student's t-test). Fold change relative to pre-BTG1 enrichment in GFP-ORF57 transfected cells.

3.17 Inhibition of ORF57 leads to reduction in circHIPK3 levels

Previous work in the Whitehouse group identified CTT018159 as a small molecular inhibitor of ORF57 activity, specifically it prevented the interaction of ORF57 with host cell factors, including the hTREX complex which is required for ORF57-mediated stability and nuclear export of KSHV mRNAs [272]. An MTS assay was performed to assess cytotoxicity of CTT018159, with 1.0 and 2.5 uM the highest concentrations incubated

with over 80% survival, thus they were selected for further experiments work (Figure 31A). To assess the effect of CTT018159 on virus replication and circHIPK3 biogenesis TREx-BCBL1-RTA cells were incubated with non-cytotoxic concentrations of the small molecule and circHIPK3 levels assessed by qPCR 24 hours post treatment. Treatment with CTT018159 led to a reduction in circHIPK3 levels confirming a potential role of ORF57 in the upregulation of circHIPK3 (Figure 31B). The inhibitory activity of CTT018159 on KSHV lytic replication was confirmed by reduced ORF65 protein levels (Figure 31C). These results suggest that ORF57 may recruit host cell factors to enhance circRNA biogenesis or stability.

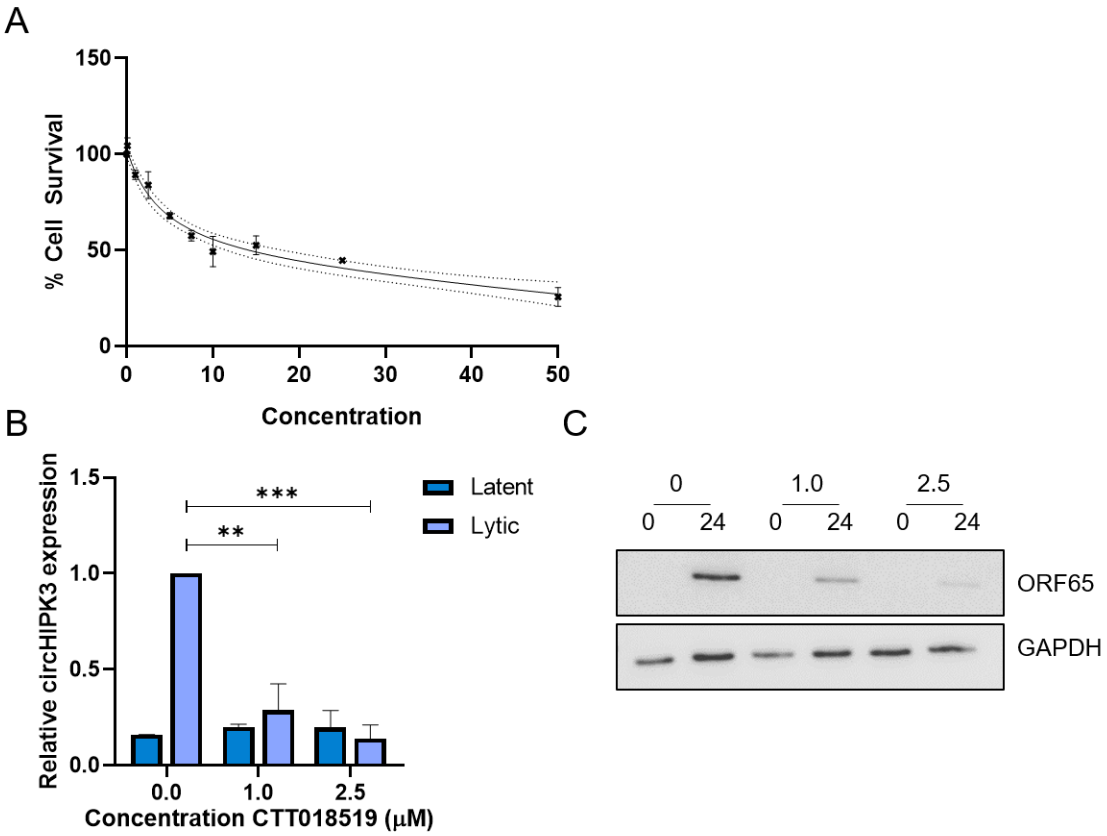

Figure 31: Inhibition of ORF57 prevents an upregulation of circHIPK3. (A) Analysis of % cells survival using MTS assay. Cells were incubated for 24 hours with a range of concentrations of CTT018519 from 0.1–50 µM. **(B).** qPCR analysis of circHIPk3 levels in 0 and 24 hour TREx-BCBL1-RTA cells treated with CTT018519. **(C).** Western blot for ORF65 in CTT018519 treated TREx-BCBL1-RTA cells at 0 and 24 hours. N=3. Error bars represent SD. *P < 0.05, **P < 0.01 and ***P < 0.001 (unpaired Student's t-test).

3.18 RNA binding ORF57 mutants lead to reduced circHIPK3 upregulation

The role of ORF57 was further investigated with previously characterised RNA binding and mRNA export negative mutants, termed RGG1/2 and PxxP, respectively [274] [94]. Dimerisation of RBPs has been shown to play a role in circRNA biogenesis, by promoting close proximity of circRNA splice sites. Interestingly ORF57 is known to dimerise, therefore site-directed mutagenesis was performed to create a dimerisation-negative mutant W292A [275]. These mutants along with wild type ORF57 were transfected into HEK-293Ts cells and levels of circHIPK3 and *HIPK3* were analysed via qPCR after 24 hours. Successful transfection and expression of the respective mutants was assessed by immunoblotting (Figure 32A), however, unfortunately, the W292A mutant was too unstable to utilise in further studies. While the PxxP ORF57 mutant had little if any effect on circHIPK3 levels, transfection of the RGG1/2 ORF57 mutant led to a significant drop in circHIPK3 levels compared to wild-type ORF57 protein (Figure 32B-C). Interestingly transfection with GFP-ORF57 led to a small decrease in linear *HIPK3* levels (Figure 32D), suggesting it could be driving circularisation of *HIPK3*. In contrast, the observed drop in linear *HIPK3* levels upon GFP-ORF57 transfection did not occur with the RGG1/2 mutant. Furthermore, GFP-TRAP RIPs using GFP-ORF57-RGG1/2 also failed to precipitate (Fig 3.18E). Together these results indicate that ORF57 RNA binding domains play a role in the observed increased levels of circHIPK3, although the exact mechanism by which ORF57 increases circHIPK3 levels is yet to be fully elucidated.

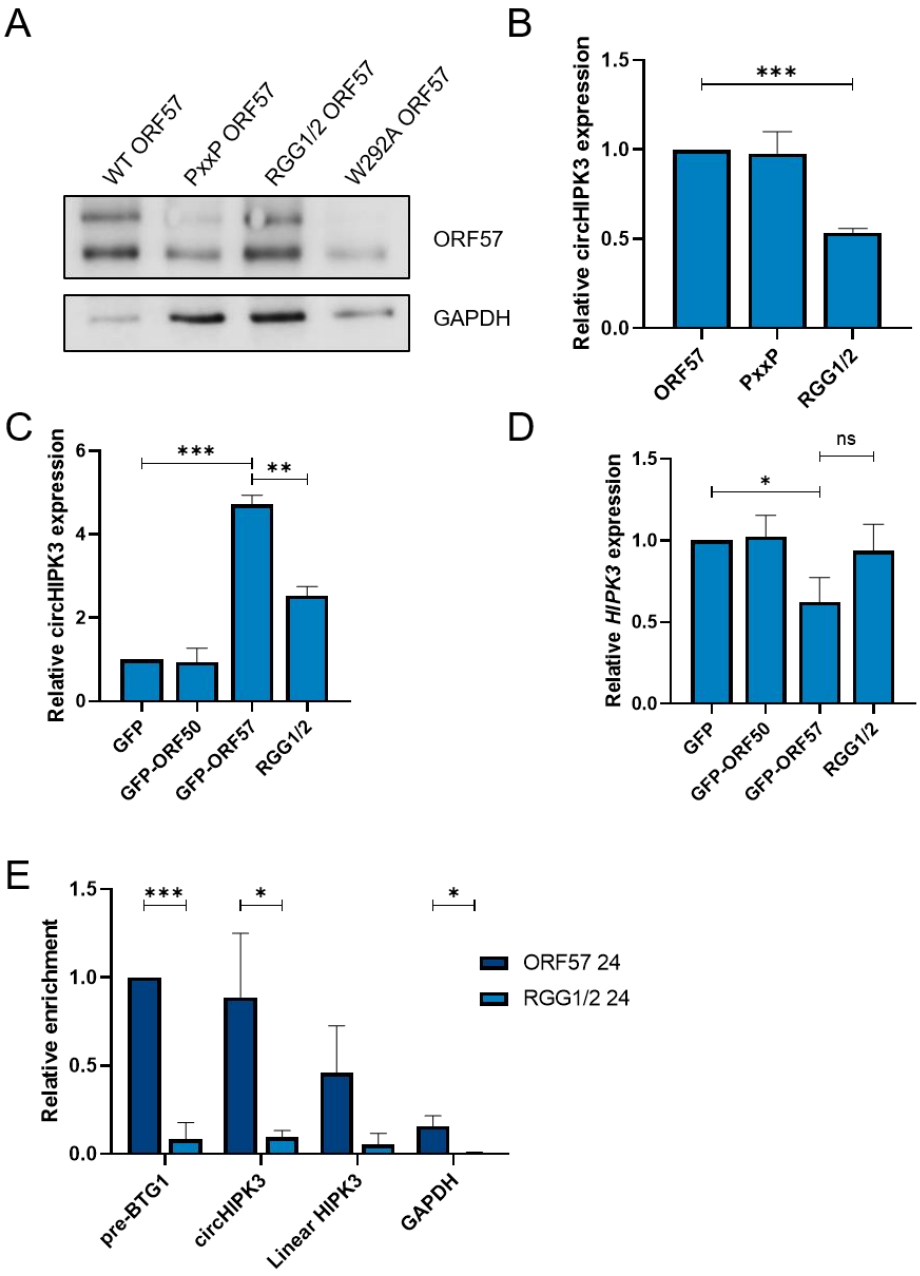

Figure 32: **ORF57's RNA binding domains are important for its role in circHIPK3 upregulation. (A).** Western blot for ORF57 in HEK-293T cells transfected with WT ORF57 and the mutants PxxP, RGG1/2 and W292A. **(B).** qPCR of circHIPK3 levels in HEK-293Ts transfected with ORF57 and the mutants PxxP and RGG1/2. qPCR analysis of circHIPK3 **(C)** and HIPK3 **(D)** in GFP, ORF50, ORF57 and the RGG1/2 mutant. **(E).** GFP RIPs in ORF57-GFP or RGG1/2 transfected HEK-293Ts at 24 hours post-transfection, respectively. Enrichment is normalised to pre-BTG1 in GFP-ORF57 transfected cells. N=3. Fold change relative to GFP-transfected cells (E) or GFP transfected cells. Error bars represent SD. *P < 0.05, **P < 0.01 and ***P < 0.001 (unpaired Student's t-test).

3.19 Discussion

Results in this chapter highlight the identification of a novel ncRNA regulatory axis dysregulated during KSHV lytic replication, with the cornerstone of this network being the virus-mediated upregulation of circHIPK3. circHIPK3 is consistently upregulated

during KSHV lytic replication, in contrast to its primary linear transcript HIPK3, and crucially, reversion of circHIPK3 upregulation leads to inhibition of viral replication and infectious virion production, highlighting the importance to KSHV replication.

Although most circRNAs have no function assigned to them, previous research has identified circHIPK3 is capable of sponging miRNA activity [263]. However, the role of circHIPK3 has not been elucidated in viral infection. As circRNA functionality is tied to their localisation, qPCR and RNA-FISH confirmed its cytoplasmic localisation during lytic replication, supporting its miRNA sponging potential. Further analysis has shown that circHIPK3 dysregulation is essential for the downregulation of antiviral miRNAs, miR-29b and miR-30c. However, the exact mechanism of circHIPK3-mediated dysregulation of the miRNAs is currently not elucidated, with circRNA sponging alone not sufficient to induce degradation and therefore reduce levels of the miRNAs. One potential mechanism for circRNA binding leading to active degradation is through target-RNA-directed microRNA degradation (TDMD), where binding to highly complementary RNAs induces miRNA degradation [134]. Interestingly, of the few miRNAs that have been shown to undergo TDMD, both miR-29b and miR-30c are among them [276] [277]. The exact requirements and factors in TDMD are still not fully known and appear to vary dependent on mechanism, however, a central mismatch bulge and high 3′ complementarity appear to be required although [116], binding site predictions for circHIPK3 and the miRNAs do not fully follow this basis.

Other potential models for degradation of miRNAs through RNA binding include binding leading to recruitment of degradation factors, re-localisation of the miRNA leading to its degradation or most similarly to TDMD: binding could lead to conformational changes that expose the miRNA to degradation machinery [277]. Of note, herpesviruses are particularly adept at regulating and inducing degradation of small RNAs through binding with lncRNAs, with HCMV, MCMV and the close relation of KSHV, herpesvirus saimiri (HVS), all having been shown capable of small RNA degradation [278] [279] [280] [134].

However, at present the mechanism and factors involved in this axis require further investigation.

Finally the mechanism behind KSHV-mediated circHIPK3 upregulation was briefly investigated, with the immediate-early ORF57 protein implicated. At present, the actual mechanism of ORF57-mediated upregulation requires further investigation with cellular factors still to be identified. ORF57 is a promiscuous RBP, it is likely therefore that ORF57 recruits and utilises cellular proteins to aid this dysregulation, however, as of yet these factors are unknown. Furthermore, we show its RNA binding properties are likely important in this role, with the ORF57 RNA binding negative mutant RGG1/2 unable to bind to circHIPK3. Interestingly ORF57 is known to form a dimer [281]; many of the other known circRNA biogenesis promoting proteins act through dimerisation, for instance QKI and Mbl [155] [154]. Here the protein monomers bind to the upstream and downstream splice sites, and as they dimerise they bring the splice sites in close contact promoting circularisation, it is possible that ORF57 promotes circHIPK3 biogenesis by this mechanism. Unfortunately the dimerisation negative ORF57 mutant utilised herein was too unstable for further investigation. Alternatively circHIPK3 dysregulation may be through ORF57-mediated stability enhancement rather than biogenesis, with previous research identifying ORF57 as a key stability factor in the viral ncRNA transcript PAN [92]. Further research is now needed to elucidate exact role of ORF57 in circRNA biogenesis.

Chapter 4

~

circHIPK3 and miR-30c regulate DLL4

4.0 circHIPK3 and miR-30c regulate DLL4

Chapter 3 highlighted the importance of KSHV-mediated dysregulation of the circHIPK3: miR-30c axis to enhance viral lytic replication. It also suggest that KSHV may modulate this ncRNA axis to regulate expression of miR-30c-specific mRNA targets that would be beneficial for virus replication. It was therefore hypothesised that due to the downregulation of miR-30c during lytic replication, its target mRNA would be significantly upregulated, leading to an increase in expression of this pro-viral protein.

In this chapter DLL4 was identified as a target of miR-30c. DLL4 is transmembrane protein, a member of the delta gene family and is predominantly characterised as a notch ligand. Through its stimulation of notch1 and notch4 it can promote angiogenesis and blood vessel remodelling [282]. It has also been identified as a cell cycle regulator through it negative regulation of several cyclins [283].

In line with the hypothesis that the miR-30c mRNA target would be increased in its expression due to dysregulation of its regulating ncRNA axis, we validated the upregulation of DLL4 at both an RNA and at protein level upon KSHV reactivation. Notably, DLL4 expression was specifically regulated by the circHIPK3:miR30c axis utilising a range of knockdown cell lines and miR0-30c mimics/antagomiRs. Furthermore a direct interaction between miR-30c and DLL4 was demonstrated using both 3′ UTR luciferase assays and RIPs.

Finally the role of DLL4 in KSHV lytic replication was investigated. Stable depletion of DLL4 in TREx-BCBL1-RTA cells showed that DLL4 was essential for KSHV lytic replication and infectious virion production. Moreover, evidence is provided suggesting that KSHV-mediated dysregulation of the circHIPK3: miR-30c: DLL4 network is a novel mechanism for cell cycle regulation in B cells.

4.1 Identifying miR-30c target mRNAs

After demonstrating the dependence of the central axis of circHIPK3: miR-30c to enhance KSHV lytic replication, we next aimed to identify miR-30c-specific mRNA targets which completed this ncRNA network. Online miRNA target prediction software miRDB, targetScan and miRBase [284] [285] [286] were used to identify potential mRNA targets. This putative list was then cross referenced with a previously obtained RNA Seq dataset comparing cellular gene expression during latent and lytic replication cycles produced within the Whitehouse laboratory to prioritise a range of targets that were upregulated during lytic replication. Notably, all 3 miRNA target prediction programmes identified DLL4 as a potential miR-30c binding partner. Importantly DLL4 and miR-30c had a matching sequence of 9 bases, with this found at the 5′ end of the miRNA. Although predictions showed the seed sequence was on the shorter end of miRNA: mRNA interactions, previous research suggests an 8 nucleotide match located at the 5’ end is the most important feature for successful miRNA-mRNA targeting [287]. Additionally, DLL4 was also upregulated during lytic replication in the RNA Seq dataset (Figure 33A-B).

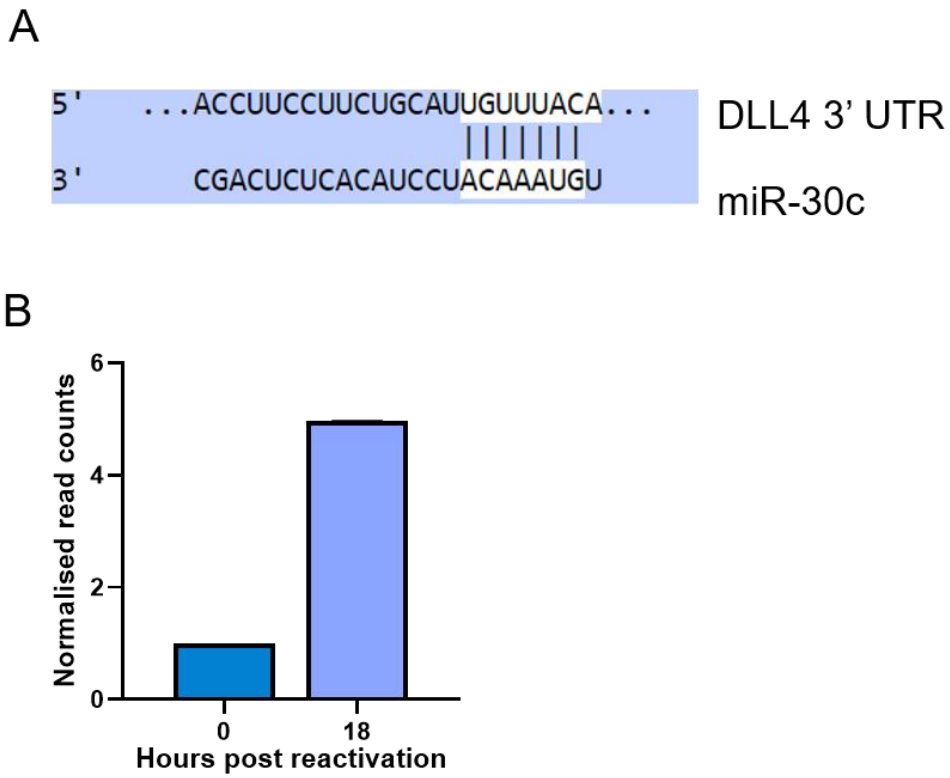

Figure 33: DLL4 is a potential miR-30c target. (A). TargetScan alignment of miR-30c to the DLL4 3′ UTR. **(B).** Normalised read counts of DLL4 in mRNA Seq at 0 and 18 hours in TREx-BCBL1-RTA cells. N=2.

To confirm the RNA Seq dataset and demonstrate that DLL4 was indeed upregulated during KSHV lytic replication qPCR and immunoblotting was performed on RNA and cell lysates from latent and lytically reactivated TREx-BCBL1-RTA cells. Results confirmed that DLL4 is significantly upregulated during lytic replication at both the RNA and protein level (Figure 34A-B).

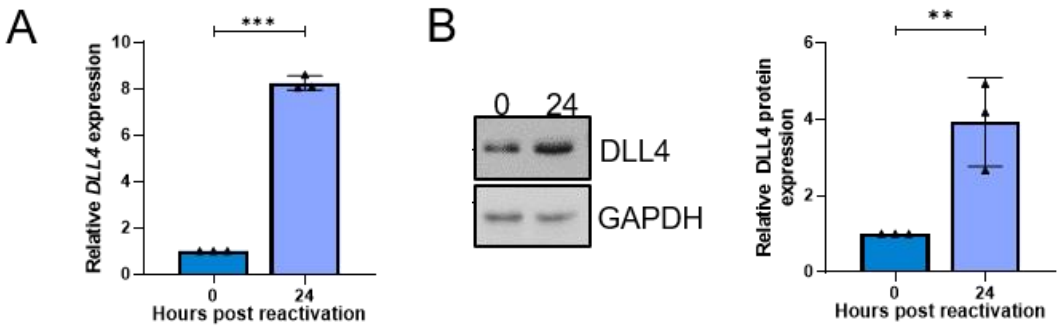

Figure 34: DLL4 is upregulated in lytic replication. (A). qPCR of DLL4 levels at 0 and 24 hours in TREx-BCBL1-RTA cells. **(B).** Representative western blot of DLL4 levels at 0 and 24 hours with densitometry analysis. N=3. Fold change relative to 0hr cells. Error bars represent SD. *P < 0.05, **P < 0.01 and ***P < 0.001 (unpaired Student's t-test).

4.2 circHIPK3 knockdown or use of miR-30c mimics prevents the DLL4 upregulation

In determine whether the circHIPK3: miR-30c axis specifically affected DLL4 expression, the effect of either circHIPK3 depletion utilising the circHIP3 KD stable lentivirus cell lines or overexpression of a miR-30c mimic was assessed on DLL4 expression in TREx-BCBL1-RTA cells at 0 and 24 hours post-reactivation. In both experiments, DLL4 levels were significantly reduced at an RNA and protein level, implying that DLL4 could be a downstream target of this ncRNA axis (Figure 35A-D). Successful depletion of circHIPK3 and overexpression of miR-30c was confirmed for each repeat via qPCR (data not shown)., Ago2 RIPs were performed to determine if DLL4 was associated with the miRNA machinery. Results confirmed that DLL4 was enriched in Ago2 RIPs over IgG, suggesting that it is associating with the miRNA machinery and a potential target of a miRNA (Figure 35E).

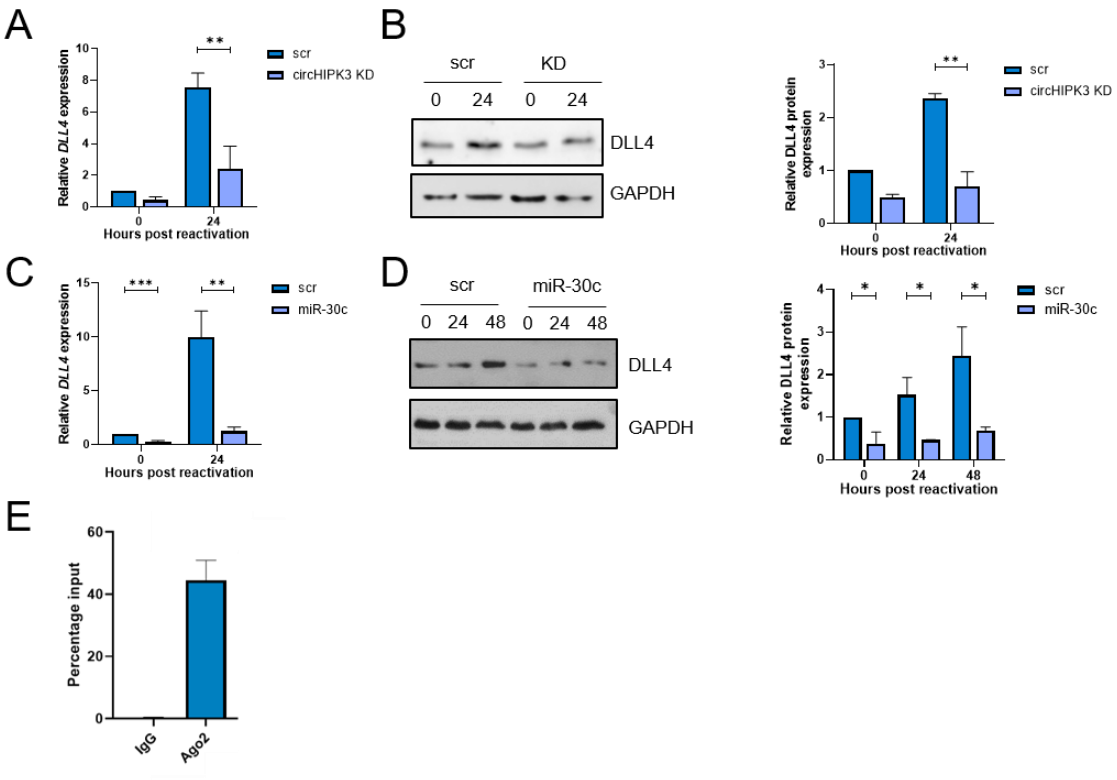

Figure 35: DLL4 is regulated by circHIPK3 and miR-30c. (A). qPCR analysis of DLL4 levels in scr and circHIPK3 KD TREx-BCBL1-RTA cells at 0 and 24 hours. **(B).** Representative western blot of DLL4 levels in scr and circHIPK3 KD TREx-BCBL1-RTA cells at 0 and 24 hours with densitometry analysis. **(C).** qPCR analysis of DLL4 levels in scr and miR-30c transfected TREx-BCBL1-RTA cells at 0 and 24 hours. **(D).** Representative western blot of DLL4 levels in scr and miR-30c transfected TREx-BCBL1-RTA cells at 0, 24 and 48 hours with densitometry analysis**. (E).** qPCR of DLL4 enrichment in Ago2 vs IgG RIP analysed using percentage of input method, with RIP signals divided by inputs to normalise for background levels. N=3. Fold change relative to 0hr cells. Error bars represent SD. *P < 0.05, **P < 0.01 and ***P < 0.001 (unpaired Student's t-test).

4.3 DLL4 is a direct target of miR-30c

To confirm that DLL4 was a definitive target of miR-30c, a series of additional experiments were performed. Firstly, the effect of inhibiting miR-30c with a specific antagomiR was investigated to determine whether it increased DLL4 levels and subsequently had a beneficial effect on virus replication. Results showed effective transfection of the miR-30c antagomiR as measured by a reduction in miR-30c levels (Figure 36A), leading to a significant increase in DLL4 levels seen at both RNA and protein levels (Figure 36B-C). Additionally, further reducing the miR-30c levels observed during lytic replication using the antagomiR had a small but non-significant increase in the late ORF65 protein levels (Figure 36D). The smaller changes observed with the miR-30c antagomiR over the miR-30c mimic during lytic replication were hypothesised due

to already lower levels of miR-30c due to the efficient downregulation of miR-30c during KSHV lytic replication. Moreover, to confirm that inhibition of miR-30c prevented the association of DLL4 with the miRNA machinery, Ago2 RIPs were also performed in the absence and presence of the miR-30c antagomiR. Levels of *DLL4* were precipitated using an Ago2-specific antibody and analysed by qPCR. Results showed a significant reduction in Ago2 associated *DLL4* in the presence of the antagomiR, suggesting that miR-30c targets DLL4 (Figure 36E).

To confirm DLL4 is a direct target of miR-30c, a 3′ UTR luciferase assay was utilised. Here the 3′ UTR of *DLL4* was cloned into a luciferase reporter plasmid and transfected into HEK-293T cells in the absence or presence of a miR-30c mimic. Transfection of miR-30c lead to a reduction in luminescence, confirming miR-30c directly regulates *DLL4* expression (Figure 36F). This was further supported by repeating this experiment in the presence of the miR-30c antagomiR, which lead to small increase in luciferase expression (Figure 36G). Finally, these results were further validated by analysis of *DLL4* levels in miR-30c RIPs, where biotinylated miR-30c or a scrambled control were transfected into HEK-293Ts prior to a RIP being performed using streptavidin beads. *DLL4* showed significant enrichment in the miR-30c RIP over scrambled controls, indicating it directly associates with miR-30c (Figure 36H). Together these data demonstrate that *DLL4* is a direct target of miR-30c.

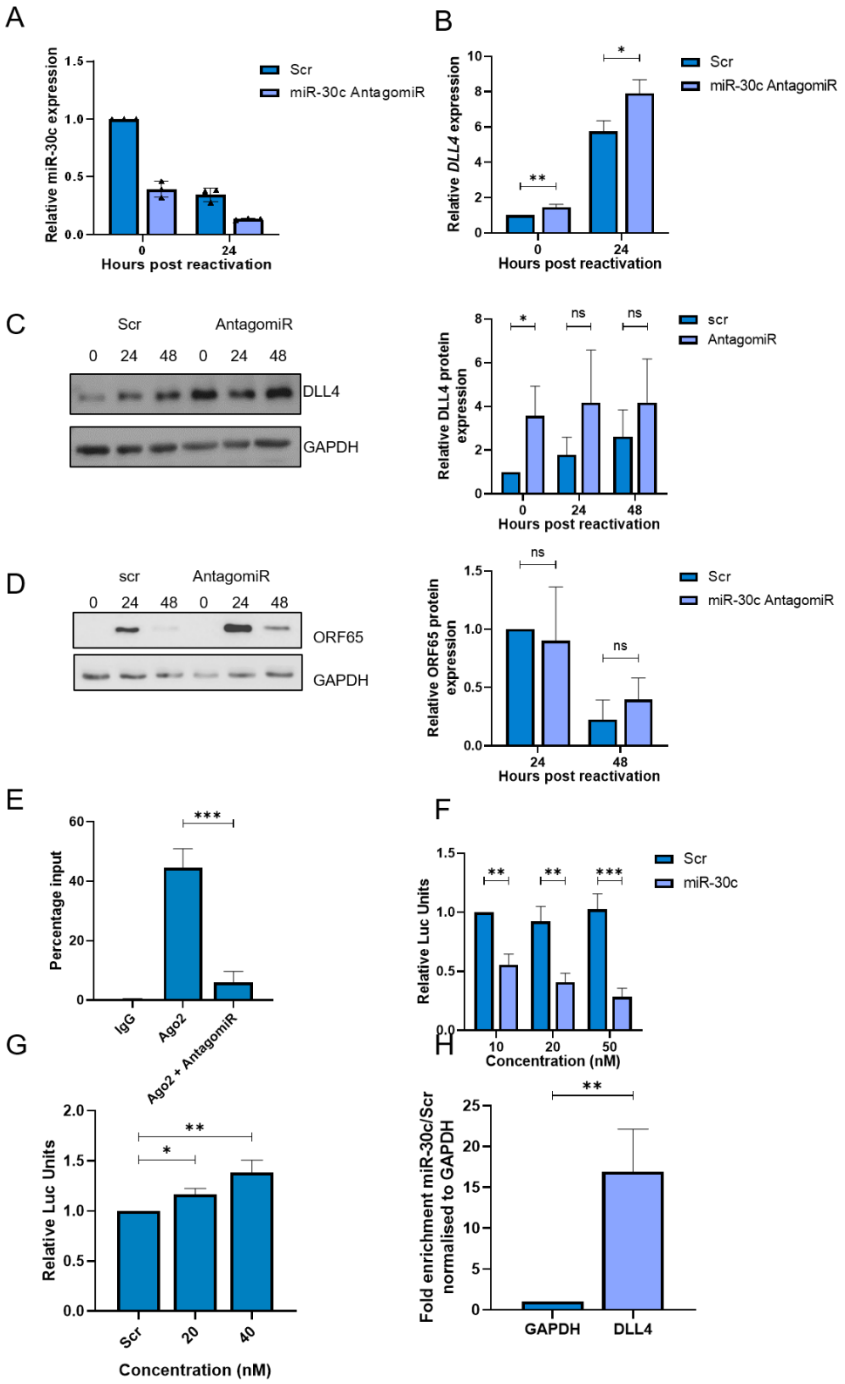

Figure 36: DLL4 is a direct target of miR-30c: A). qPCR analysis of miR-30c levels in scr or miR-30c antagomiR transfected TREx-BCBL1-RTA cells at 0 and 24 hours. **(B).** qPCR analysis of *DLL4* levels in scr or miR-30c antagomiR transfected TREx-BCBL1-RTA cells at 0 and 24 hours. **(C).** Representative western blot of DLL4 in scr and miR-30c antagomiR transfected TREx-BCBL1-RTA cells at 0, 24 and 48 hours with densitometry analysis. **(D).** Representative western blot of ORF65 in scr and miR-30c antagomiR transfected TREx-BCBL1-RTA cells at 0, 24 and 48 hours with densitometry analysis. **(E).** DLL4 levels analysed via qPCR as percentage input in IgG RIPs and Ago2 RIPs (with and without a miR-30c antagomiR). **(F).** Luciferase reporter assay from HEK-293Ts co-transfected with DLL4 3′UTR reporter plasmid and either scr or miR-30c at 10, 20 and 50 nM. Data presented are relative to an internal firefly control. **(G).** Luciferase reporter assay from HEK-293Ts co-transfected with DLL4 3′UTR reporter plasmid and either scr or miR-30c antagomiR with scr at 40 nM and the antagomiR at 20 or 40 nM. Data presented are relative to an internal firefly control. **(H).** Streptavidin RIP of biotinylated miR-30c or scr transfected HEK-293Ts analysing GAPDH or DLL4 enrichment. N=3. Error bars represent SD. *P < 0.05, **P < 0.01 and ***P < 0.001 (unpaired Student's t-test).

4.4 The circHIPK3:miR-30c:DLL4 circuit is dysregulated in multiple KSHV-infected cell lines

To determine whether the dysregulation of the circHIPK3:miR-30c:DLL4 ncRNA network was conserved across different KSHV-infected cell types, the levels of circHIPK3, miR-30c and DLL4 were analysed via qPCR in an additional KSHV-infected B cell line: BC-3s as well as rKSHV.219-HEK-293Ts (clone 9s), a stably infected 293T cell line. In both cell lines, circHIPK3 and DLL4 were upregulated and miR-30c was downregulated upon KSHV lytic replication induction (Figure 37A-F). This demonstrates the importance of this ncRNA network across KSHV biology.

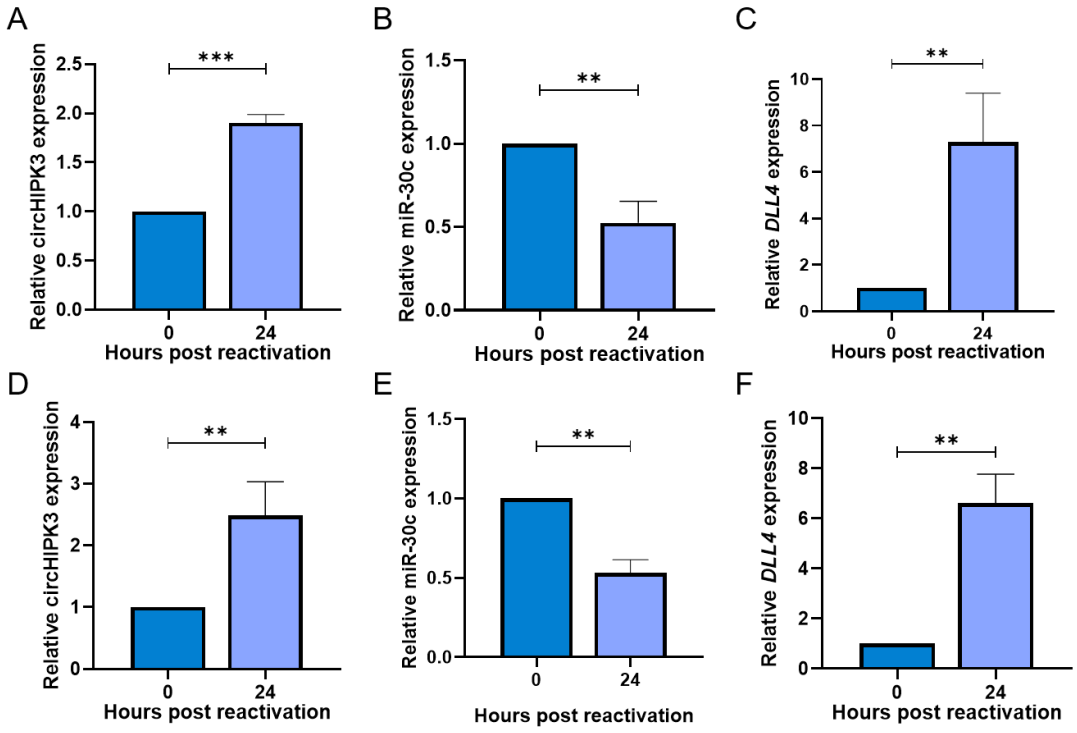

Figure 37: The circHIPK3 ncRNA network is dysregulated across KSHV lytic replication. qPCR analysis of circHIPK3 **(A),** miR-30c **(B)** and DLL4 **(C)** in BC-3s. qPCR analysis of circHIPK3 **(D)**, miR-30c **(E)** and DLL4 **(F)** in rKSHV.219-HEK-293Ts. N=3. Fold change relative to 0hr cells. Error bars represent SD. *P < 0.05, **P < 0.01 and ***P < 0.001 (unpaired Student's t-test).

4.5 DLL4 is essential for KSHV replication

To confirm the importance of DLL4 during KSHV lytic replication, TREx-BCBL1-RTA cells were transduced with lentivirus-based shRNAs, significantly depleting DLL4 RNA and protein expression during lytic replication (Figure 38A-B). The effect of the KD on

the virus was then investigated through assessing levels of the late lytic ORF65 protein. Reactivation assays demonstrated that DLL4 depletion resulted in a significant reduction in ORF65 protein expression compared to the scrambled control (Figure 38C). Similarly, DLL4 depletion resulted in a significant reduction in viral load and infectious virion production, compared to the scrambled control (Figure 38D-E). Together, these data highlight that dysregulation of a circHIPK3:miR-30c:DLL4 regulatory circuit is essential for efficient KSHV lytic replication, with DLL4 having pro-viral activity.

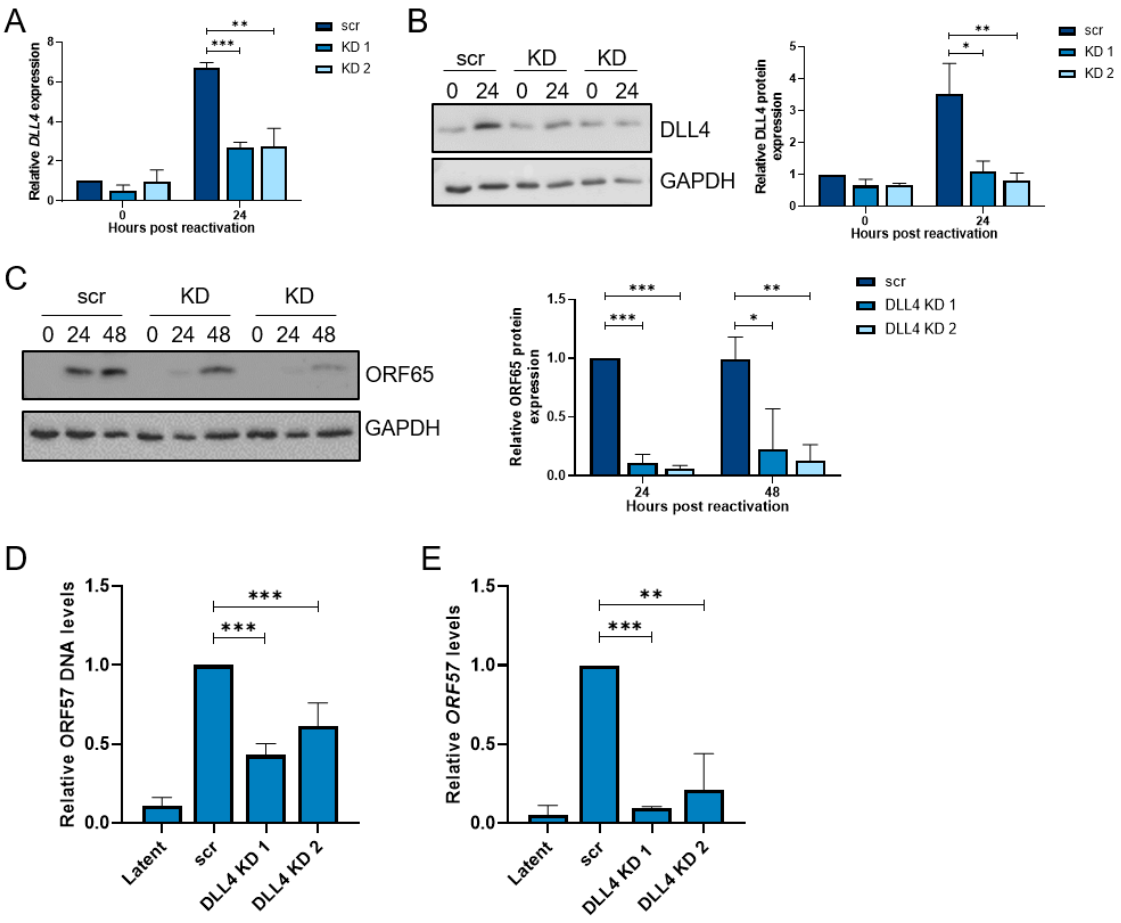

Figure 38: DLL4 is essential in KSHV lytic replication. (A). qPCR of DLL4 levels in scr and DLL4 KD TREx-BCBL1-RTA cell lines at 0 and 24 hours. **(B).** Representative western blot of DLL4 levels at 0 and 24 hours in scr and DLL4 KD cells with densitometry analysis. **(C).** Representative western blot of ORF65 levels at 0, 24 and 48 hours in scr and DLL4 KD cells with densitometry analysis. **(D).** Viral load in scr vs DLL4 KD TREx-BCBL1-RTA cells. **(E).** Viral reinfection in scr vs DLL4 KD TREx-BCBL1-RTA cells. N=3. Fold change relative to Scr 0hr cells. Error bars represent SD. *P < 0.05, **P < 0.01 and ***P < 0.001 (unpaired Student's t-test).

4.6 DLL4 regulates cell cycle progression during lytic replication

DLL4 is a transmembrane protein that acts as a ligand for Notch receptors 1 and 4. As such it has multiple roles, including functioning as a regulator of several cyclins, which are master regulators of the cell cycle. Therefore, it may be the case that KSHV-mediated dysregulation of the circHIPK3:miR-30c:DLL4 circuit was used to modulate the cell cycle during lytic replication [283]. Flow cytometry was therefore performed using propidium iodide to initially analyse changes in the proportion of cells that are in one of the three interphase stages of the cell cycle during KSHV lytic replication in TREx-BCBL1-RTA cells. Results demonstrated a significant drop in the number of cells in the G2/M phase and a maintained number in G1 phase (Figure 39A-B). As expected during KSHV lytic replication and increased number of cells was found to be in G2+ cells due to increased KSHV DNA levels during lytic replication.

Further analysis was then performed to assess the RNA levels of known cyclins regulated by DLL4, namely CCNE1 and CCNB1, during lytic replication. qPCR analysis found *CCNE1* and *CCNB1* were downregulated at 24 hours, which could possibly relate to the changes observed in the cell cycle during KSHV infection (Figure 39C-D). To determine whether the circHIPK3:miR-30c:DLL4 regulatory circuit was directly involved in the observed cell cycle changes, flow cytometry assays were repeated in scrambled control cells versus TREx-BCBL1-RTA cells depleted for either circHIPK3 or DLL4. Results showed similar results in both knockdown cells, with the loss of cells in G2 phase being less pronounced at both 24 and 48 hours post-lytic reactivation (Figure 39A-B). Likewise, the reduction in *CCNE1* and *CCNB1* expression was partly abrogated in both circHIPK3 and DLL4 knockdown cell lines compared to scrambled control (Figure 39C-D).

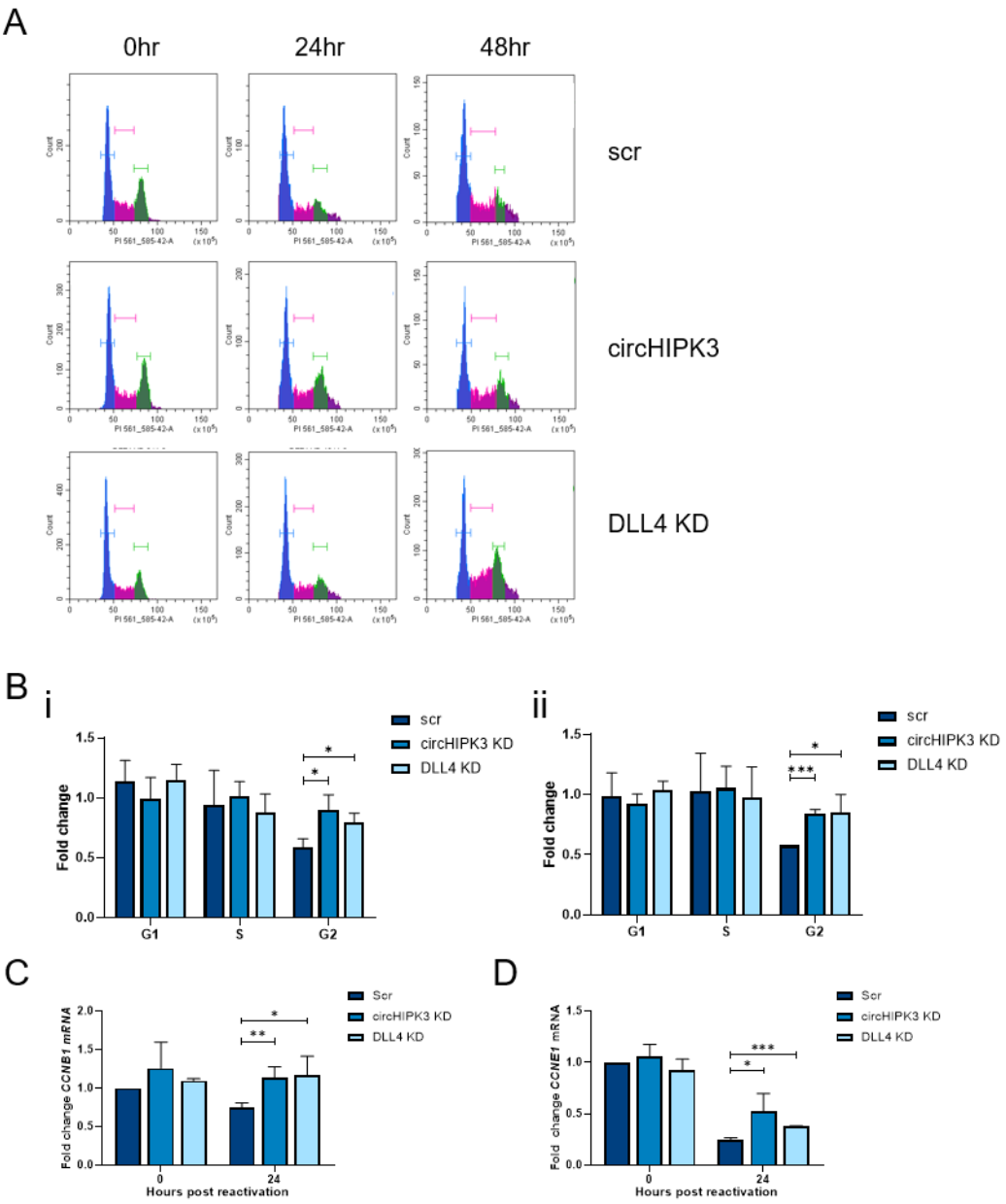

Figure 39: DLL4 regulates the cell cycle. (A). Representative plot of cell cycle distribution at 0, 24 and 48 hours in scr, circHIPK3 and DLL4 KD cells, with G1 phase (blue), S phase (pink) and G2/M phase (green) and G2+ (purple) highlighted**. (B).** Analysis of fold change in number of cells in G1, S and G2/M phases of the cell cycle from 0 to 24 hour **(i)** or 0 to 48 hours **(ii). (C).** qPCR of *CCNB1* levels in scr, circHIPK3 and DLL4 KD TREx-BCBL1-RTA cells at 0 and 24 hours. **(D)**. qPCR of *CCNE1* levels in scr, circHIPK3 KD and DLL4 KD TREx-BCBL1-RTA cells at 0 and 24 hours. Fold change is relative to 0 hour cells (C-D). N=3. Error bars represent SD. *P < 0.05, **P < 0.01 and ***P < 0.001 (unpaired Student's t-test).

4.7 Dysregulation of the cell cycle is important for KSHV replication

The importance of KSHV-mediated cell cycle dysregulation on viral DNA replication was then investigated using a range of cell cycle inhibitors. RO 3306 and nocodazole were used to induce G2 arrest, while thymidine was utilised to induce G1 arrest. The effect of these inhibitors was confirmed via flow cytometry prior to assessment of their effect on

virus replication (Figure 40A-B). As expected, RO 3306 and nocodazole treated cells resulted in an increased proportion of cells in the G2/M phase while thymidine increased cell number in the G1 phase of the cell cycle. TREx-BCBL1-RTA cells were then treated with the inhibitors prior to lytic induction and KSHV DNA measured by qPCR as an indication of viral load. Viral load analysis showed that both RO 3307 and nocodazole led to a significant reduction in viral load, conversely thymidine had little if any effect (Figure 40C). However, interestingly, thymidine treatment of DLL4 depleted cells, which was hypothesised to reverse any DLL4-mediated cell cycle changes, led to rescue of viral DNA replication (Figure 40D). Together these data suggest that the effect of DLL4 on KSHV lytic replication in TREx-BCBL1-RTA cells are directly mediated through its cell cycle activities.

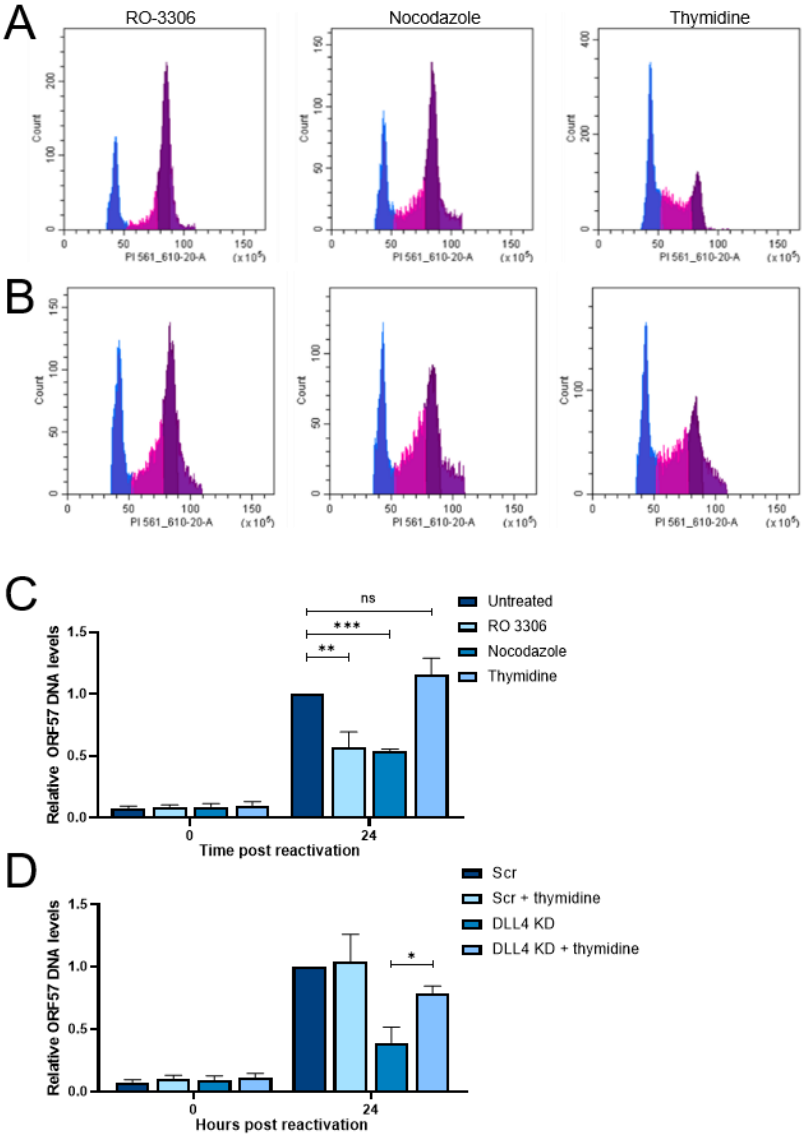

Figure 40: Cell cycle dysregulation is important for KSHV lytic replication. Representative plot of cell cycle distribution at 0 **(A)** and 24 hours **(B)** in TREx-BCBL1-RTA cells treated with RO-3306, nocodazole and thymidine, with G1 phase (blue), S phase (pink) and G2/M phase (purple) highlighted. ***(C)*.** Viral load in untreated, RO 3306, nocodazole and thymidine treated TREx. **(D).** Viral load in scr and DLL4 KD TREx-BCBL1-RTA cells treated with thymidine. N=3. Fold change relative to Scr 24hr cells. Error bars represent SD. *P < 0.05, **P < 0.01 and ***P < 0.001 (unpaired Student's t-test).

4.8 Viral replication centres do not form when the network is disrupted

Dysregulation of the cell cycle is critical for KSHV DNA replication, and as shown above dysregulation of circHIPK3:miR-30c:DLL4 circuit reduces viral load. To further explain this association we hypothesised that due to this inhibition of viral DNA replication, formation of viral replication centres, where viral DNA replication occurs, would be reduced or delayed by inhibition of the ncRNA network. To visualise KSHV replication

centres, TREx-BCBL1-RTA cells were seeded onto poly-L-lysine coated coverslips and after 24 hours induction, fixed, permeabilised and co-stained with an RNA pol II specific antibody and DAPI. Scrambled control cells show the formation of viral replication centres at 24 hours post-reactivation, as indicated by RNA pol II staining. It can be seen that the nuclear architecture undergoes a re-organisation to facilitate viral replication, where RNA pol II is hijacked from a diffuse localisation to form VRCs that will ultimately fill most of the nuclear space compressing the cellular chromatin to the nuclear periphery. However, comparison of replication centre formation in TREx-BCBL1-RTA cells depleted for either circHIPK3 and DLL4 showed that viral replication centre formation was significantly disrupted compared to scrambled control cells, with a diffuse localisation still visible in the KD cells (Figure 41). This further supports the hypothesis that dysregulation of the circHIPK3:miR-30c:DLL4 regulatory circuit is essential for efficient KSHV lytic replication.

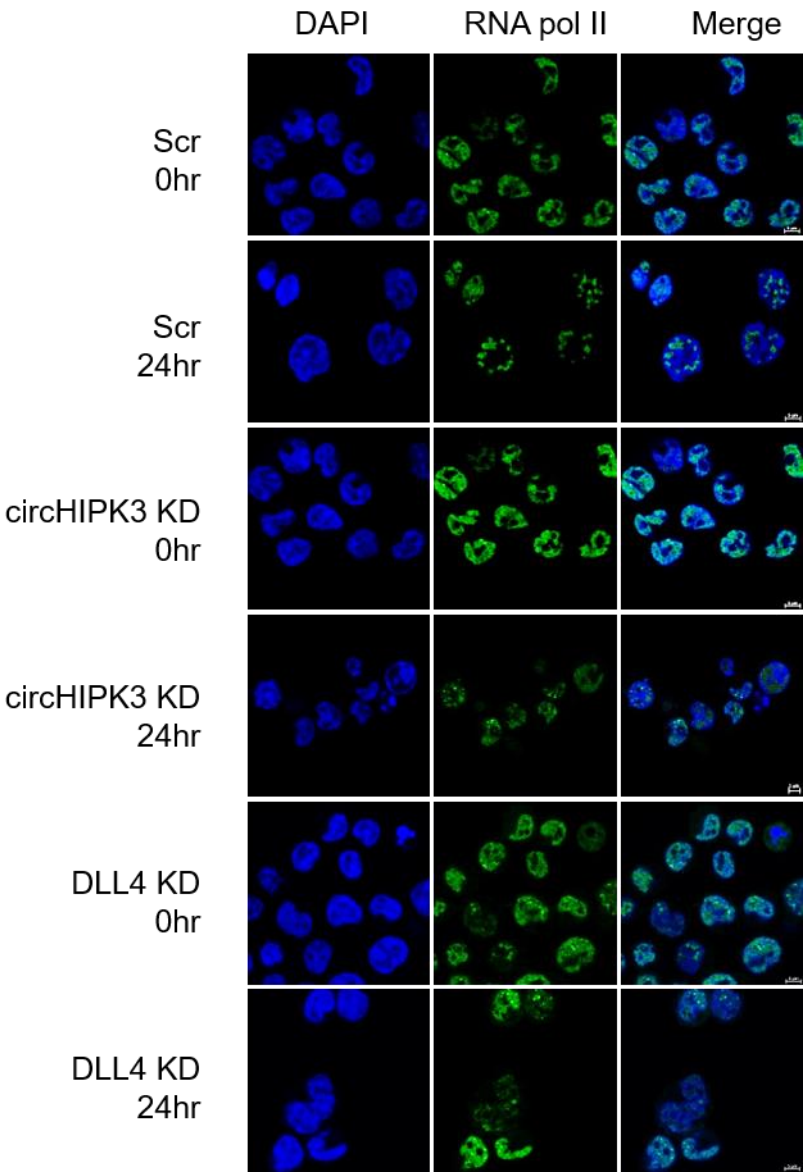

Figure 41: Dysregulation of the ncRNA network inhibits viral replication centres. Scrambled, circHIPK3 knockdown and DLL4 knockdown TREx-BCBL1-RTA cells at 0 and 24 hours post-reactivation stained for DAPI (blue) and RNA pol II (green). Scale bars are 5 μm.

4.9 Discussion

In this chapter we complete the circHIPK3:miR-30c axis into a complete ncRNA regulatory network, by the identification of DLL4 as a direct target of miR-30c. This network is summarised in Figure 42. Supporting this potential network, results show that DLL4 is upregulated following KSHV lytic reactivation in TREx-BCBL1-RTA cells. Importantly, upon circHIPK3 depletion and miR-30c overexpression, we show, DLL4 upregulation is significantly reduced. Further evidence was shown using Ago2 and miR-30c RIPs, confirming *DLL4* association with both the general miRNA machinery and specifically miR-30c, as well as a 3′ UTR luciferase assay confirming DLL4 as a direct target of miR-30c.

Results also highlight an essential role of DLL4 in KSHV lytic replication, as DLL4 depletion has an inhibitory effect on viral lytic replication, affecting late ORF65 protein production, viral load and infectious virion production. This is in agreement with previous reports investigating the importance of DLL4 to KSHV replication in endothelial cells. Here, it was shown that the viral protein vGPCR upregulated DLL4 in an ERK-dependent manner, leading to activation of Notch4 signalling [288] [283]. Due to the importance of DLL4 in viral replication, this cell type redundancy is not surprising. Although it must be noted that ncRNA expression patterns can be cell-type specific, therefore whether circHIPK3 and miR-30c are functionally active in KSHV-infected endothelial cells to regulate DLL4-mediated cell cycle inhibitors needs further investigation

As previously mentioned, DLL4 regulates angiogenesis in endothelial cells, however, its role for KSHV in B cells has not been previously characterised. Therefore, we first examined its potential role in regulating the cell cycle due to its known association with various cyclins [283]. qPCR confirmed depletion of DLL4 led to changes in *CCNE1* and *CCNB1* levels. Additionally flow cytometry analysis confirmed dysregulation of the cell cycle during lytic replication, with maintenance of cell number in G1 phase and a reduction of cells in G2/M phase. Other viruses have also been observed promoting G1

phase, notably EBV, a human gammaherpesvirus, promotes G1 phase during lytic replication [289] [290]. Furthermore, other herpesviruses including HCMV and HSV-1 also require G1 entry for successful lytic replication [291] [292]. Both circHIPK3 and DLL4 depleted cell lines demonstrated the same phenotype using flow cytometry analysis, with a prevention of the G2/M dysregulation seen during lytic replication in scramble control cells, indicative that KSHV-mediated cell cycle dysregulation is being partly blocked. DLL4 has previously been characterised to regulate cyclin and cell cycle factors [283], therefore we hypothesise; that dysregulation of this ncRNA network at any phase prevents KSHV-mediated dysregulation of the cell cycle leading to inhibition of productive viral replication.

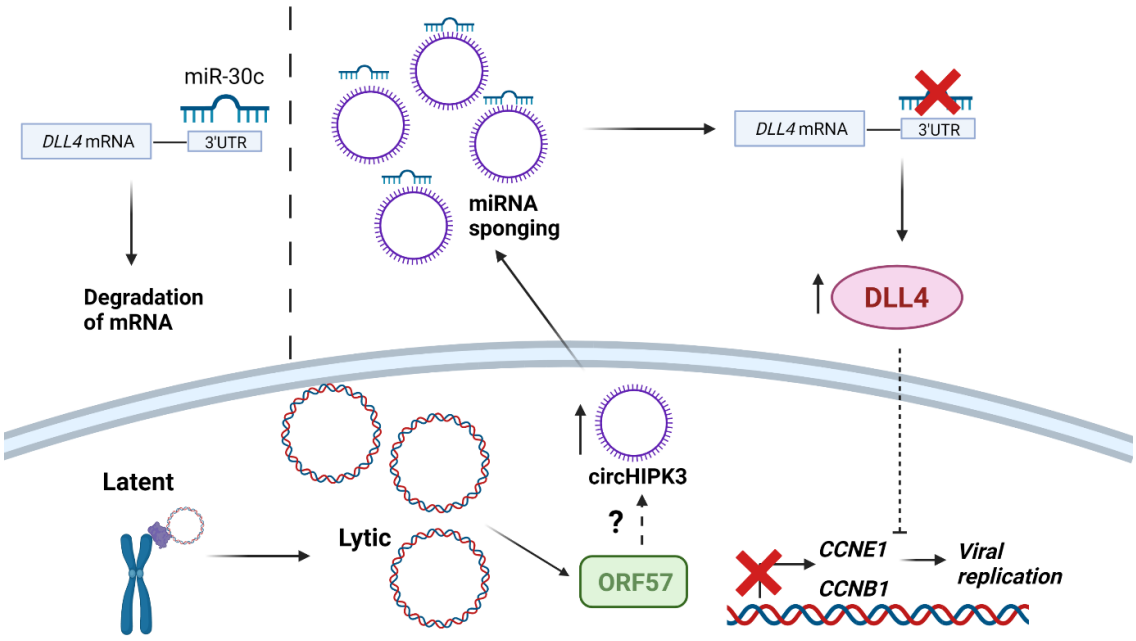

Figure 42: Lytic replication of Kaposi's sarcoma-associated herpesvirus in B cells depends on the dysregulation of a host non-coding RNA regulatory network. The circular RNA circHIPK3 interferes with the miR-30c/DLL4 axis, leading to dysregulation of cyclins and the cell cycle to enhance viral lytic replication.

Chapter 5

~

KSHV forms novel condensates to aid in RNA processing

5.0 KSHV forms novel condensates to aid in RNA processing

The previous results chapters described the KSHV-mediated dysregulation of an ncRNA regulatory network to enhance virus replication. Interestingly, dysregulation of this network is initiated by the modulation of circRNA levels. At present, it is unknown how KSHV manipulates circRNA biogenesis, but preliminary data suggest that the KSHV-encoded ORF57 protein may have a role. However, additional cellular factors may also play an important part in the virus-mediated modulation of circRNA biogenesis pathways. Recent scientific literature has implicated the cellular RBP, SFPQ, as a novel factor in regulating circRNA biogenesis and therefore it was selected for further investigation [198].

SFPQ, along with NONO and PSPC1, is a member of the DBHS (Drosophila Behaviour Human Splicing) family () [293] and is implicated in multiple roles involving transcriptional regulation and RNA splicing. SFPQ has also been identified as an essential core component of paraspeckles. These are nuclear condensates, or membrane-less subnuclear bodies found in the interchromatin space of mammalian cells.

Figure 43: The DBHS family. Schematic showing similar regions of SFPQ, NONO and PSPC1, the members of the DBHS family. Each has RNA binding domains, low complexity regions, a coiled coil and a NONA/ParaSpeckle (NOPs) domain. The NOPs domain aids in binding to other members of the DBHS family whilst the coiled coil plays a role in dimerisation and oligomerisation [294].

Paraspeckles are formed by the interaction between the lncRNA NEAT1 and members of the DBHS family of proteins including NONO and SFPQ [223]. As the core scaffold for

paraspeckle protein, NEAT1 is essential for formation. It has 3 distinct domains each of which is essential for PS formation. Domain A is required for stability of the transcript, potentially through binding to hnRNPs which protect the ends of NEAT1 from exonucleases. The exact role of the B domain is not elucidated, however, it is involved in the regulating the hnRNP K mediated switching of NEAT1_1 to NEAT1_2. Whilst the C domain is involved in binding to the PS proteins, including SFPQ and NONO aiding in PS formation [208], shown in .

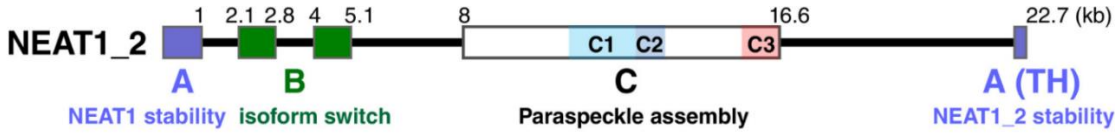

Figure 44: The distinct domains of NEAT1_2. Schematic highlighting the A, B and C domains of NEAT1. Adapted from [208].

Paraspeckles are thought to spatiotemporally control gene regulation through sequestration of diverse proteins and RNAs in cells and through this mechanism regulate gene expression during many cellular processes including differentiation, viral infection and stress responses. Paraspeckles are classified as condensates that are formed through liquid: liquid phase separation (LLPS). Here multivalent interactions between protein: protein, RNA: RNA and protein: RNA interactions drive the formation of membraneless organelles with different physiochemical properties than their surroundings [295], allowing dynamic regulation within the cell.

In this chapter, SFPQ is shown to be essential for KSHV lytic replication. Additionally, we show that SFPQ forms novel foci around VRCs during lytic replication. Notably, these foci are far larger than canonical paraspeckles and also contain the uncharacterised KSHV-encoded ORF11 protein. Overall the exact role of these foci is unknown, however, preliminary analysis suggests a putative role in both viral and cellular RNA processing.

5.1 SFPQ is re-localised into distinct nuclear foci during KSHV lytic replication levels

SFPQ has recently been described as a novel factor in the regulation of circRNAs biogenesis and other complex splicing mechanisms. These include suppression of exon inclusion through binding to 5′ splice sites reducing U1 snRNP access [296]. Therefore, studies were performed to determine if SFPQ had any potential role in the KSHV lifecycle. Initially, the levels of SFPQ and its subcellular localisation were assessed to determine if they were altered during KSHV lytic replication. First qPCR and western blotting were performed to assess SFPQ levels at 0 and 24 hours post-reactivation in TREx-BCBL1-RTA cells. Results showed no significant change during lytic replication at both RNA and protein level observed (Figure 45A-B).

Secondly, as SFPQ can form dynamic sub-nuclear bodies, subcellular fractionation was performed to explore whether its subcellular localisation changed upon lytic induction in TREx-BCBL1-RTA cells. However, once again no change was observed between 0 and 24 hours, with SFPQ showing a strong nuclear localisation (Figure 45C). A further subcellular fractionation was also performed to determine where in the nucleus SFPQ was localised. Fractionation experiments showed it was present in the nucleoplasm at both 0 and 24 hours, however, surprisingly even though SFPQ has known transcriptional regulatory roles, it failed to localise or associate with chromatin, even in latently-infected KSHV cells (Figure 45D).

Finally, immunofluorescence studies were performed to determine if SFPQ was re-distributed in the nucleoplasm during lytic replication. TREx-BCBL1-RTA cells were fixed and permeabilised at 0 or 24 hours post-reactivation, then stained for with SFPQ- and ORF57 specific antibodies (KSHV ORF57 staining was used to confirm successful reactivation). Interestingly, at 0 hours, SFPQ is diffusely localised throughout the nucleus, however, in KSHV lytically replicating cells, SFPQ was re-localised forming 3-10 distinct nuclear foci, averaging around 2 μM in size. (Figure 45E). This result

suggests that KSHV may hijack SFPQ during lytic replication to form these novel sub-nuclear bodies. ORF57 was not found to co-localise with SFPQ, implying another viral protein or mechanism may be behind the observed phenomenon.

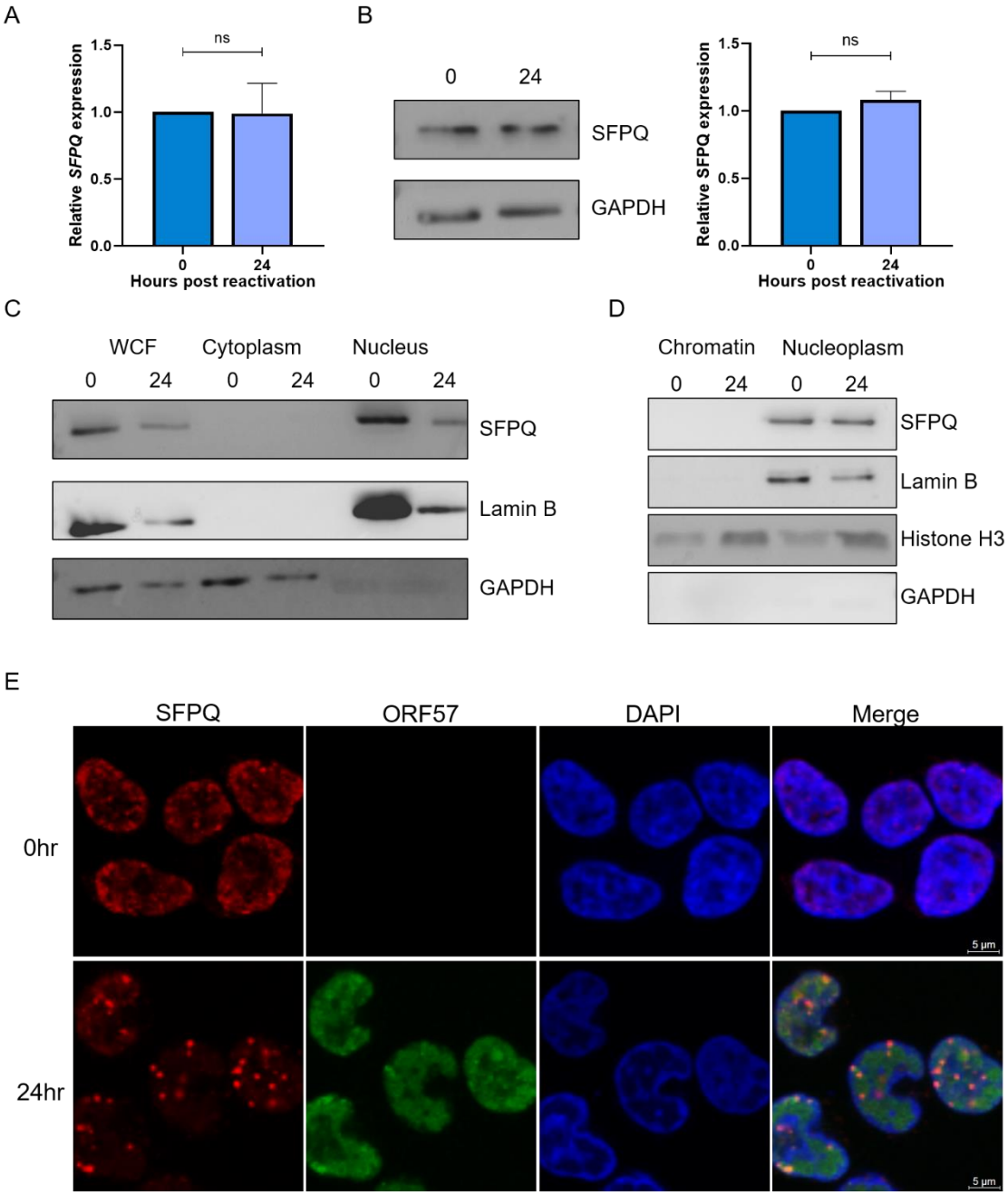

Figure 45: SFPQ levels are stable during lytic replication. (A). qPCR analysis of SFPQ at 0 and 24 hours in TREx-BCBL1-RTA cells. **(B).** Representative western blot of SFPQ levels in TREx-BCBL1-RTA cells at 0 and 24 hours with densitometry analysis. **(C).** Subcellular fractionation of TREx-BCBL1-RTA cells at 0 and 24 hours, probing for SFPQ, Lamin B and GAPDH. **(D).** Subcellular fractionation of TREx-BCBL1-RTA cells at 0 and 24 hours, probing for SFPQ, Lamin B, Histone H3 and GAPDH. N=3. Fold change relative to 0hr cells. Error bars represent SD. *P < 0.05, **P < 0.01 and ***P < 0.001 (unpaired Student's t-test). **(E).** IF performed on TREx-BCBL1-RTA cells at 0 and 24 hours post-reactivation with staining for SFPQ (red), ORF57 (green) and DAPI (blue). Scale bars are 5 µm.

5.2 SFPQ foci are formed in other KSHV-infected cell lines

TREx-BCBL1-RTA cells are derived from a body-cavity-based lymphoma cell line which have been engineered to reactivate upon doxycycline-induced expression of the KSHV RTA gene. To determine whether SFPQ foci formed during KSHV lytic replication in one other widely used model of KSHV infection, we assessed the SFPQ foci in rKSHV.219-HEK-293Ts (Clone 9s), a HEK-293T cell line stably infected with the KSHV genome. KSHV lytic replication was induced by the addition of both TPA and sodium butyrate and then immunofluorescence studies were performed at 24 hours post reactivation using SFPQ-specific antibodies. It is important to note that rKSHV.219-HEK-293Ts cells have been modified to constitutively express GFP, and also express RFP when undergoing lytic replication, therefore SFPQ was visualised using far red. Similar results were seen in rKSHV.219-HEK-293Ts cells as observed previously in TREx-BCBL1-RTA cells. Although, the foci were less numerous and uniform than those observed in TREx-BCBL1-RTA cells (Figure 46). However, it reinforces that SFPQ foci may play a key role in KSHV lytic replication in multiple cell lines and is not B cell specific.

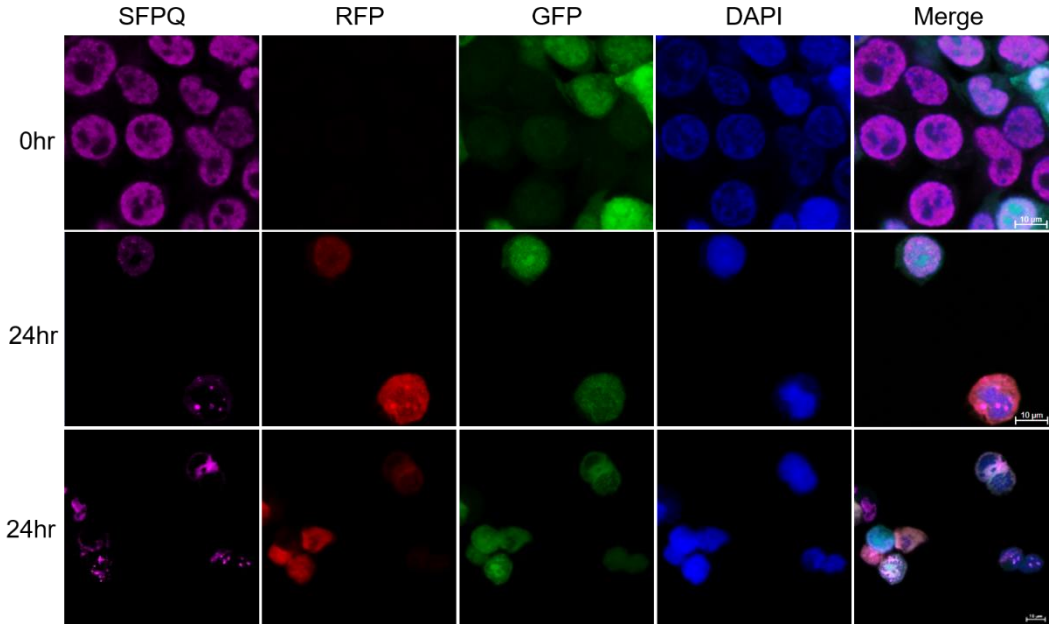

Figure 46: SFPQ foci form in another KSHV+ cell line. Immunofluorescence of rKSHV.219-HEK-293Ts at 0 and 24 hours with channels representing: SFPQ (purple), RFP/lytic marker (red), GFP (green), DAPI (blue). Scale bars represent 10 μm.

5.3 SFPQ foci start forming early in lytic replication

To identify at what point during lytic replication the foci form, IF was performed at 0, 8, 16 and 24 hours post lytic induction, using SFPQ and ORF57-specific antibodies. At 8 hours SFPQ foci appeared to be forming in small clusters in a few cells, although the majority of SFPQ was still mostly diffuse in the nucleus. Whereas, at 16 hours the foci had formed in the majority of cells, and at 24 hours the foci appeared fully mature (Figure 47A).

Consequently we hypothesised these foci were likely to be viral-induced condensates undergoing phase separation through liquid-liquid interactions. To determine whether the foci were indeed condensates FRAP was performed. This required the cloning and stable expression of a GFP-SFPQ overexpression plasmid in TREx-BCBL1-RTA cells. These cells were reactivated for either 16 or 24 hours before FRAP analysis was performed. This used a laser to bleach the condensates and measure fluorescence recovery. Analysis showed that at 16 hours, the foci are more liquid-like, with the foci able to reform after bleaching, reaching about 70% of original fluorescence in 200 seconds. Whereas at 24 hours post-reactivation, the foci failed to recover after bleaching, only gaining around 20% of the initial fluorescent signal, signifying that the condensates have stabilised and there is little dynamic movement of protein (Figure 47B). Overall half-time calculations determine the *thalf* at 16 hours post-reactivation to be 135 seconds post-bleaching, however, due to the lack of recovery, a value could not be calculated for samples at 24 hours post-reactivation. At 16 hours post-reactivation, phase separation is therefore still occurring with multivalent interactions stabilising protein: protein, protein: RNA and RNA: RNA interactions, allowing dynamic movement. However, by 24 hours, these interactions have coalesced into a dense membraneless condensate, with distinct liquid: liquid separation from its surroundings. Results therefore suggest these were novel viral condensates forming through liquid-liquid phase interactions, with this

confirmed through a 1,6-hexanediol treatment performed by Elena Harrington (Whitehouse laboratory).

Finally cells were put under osmotic stress which disrupts liquid-liquid phase interactions, TREx-BCBL1-RTA cells at 0 and 24 hours post-reactivation were incubated in RPMI media diluted 1:5 with water for 30 seconds followed by immunofluorescence using antibodies against ORF57 and SFPQ. Results showed disruption of the foci, once again suggesting they are formed through liquid-liquid phase interactions (Figure 47C). Interestingly the osmotic stress also disrupted the viral replication centres stained by ORF57, with ORF57 and SFPQ showing diffuse co-localisation. Overall, this reinforces that these are novel viral condensates that rely on liquid: liquid phase separation forming early during KSHV lytic replication.

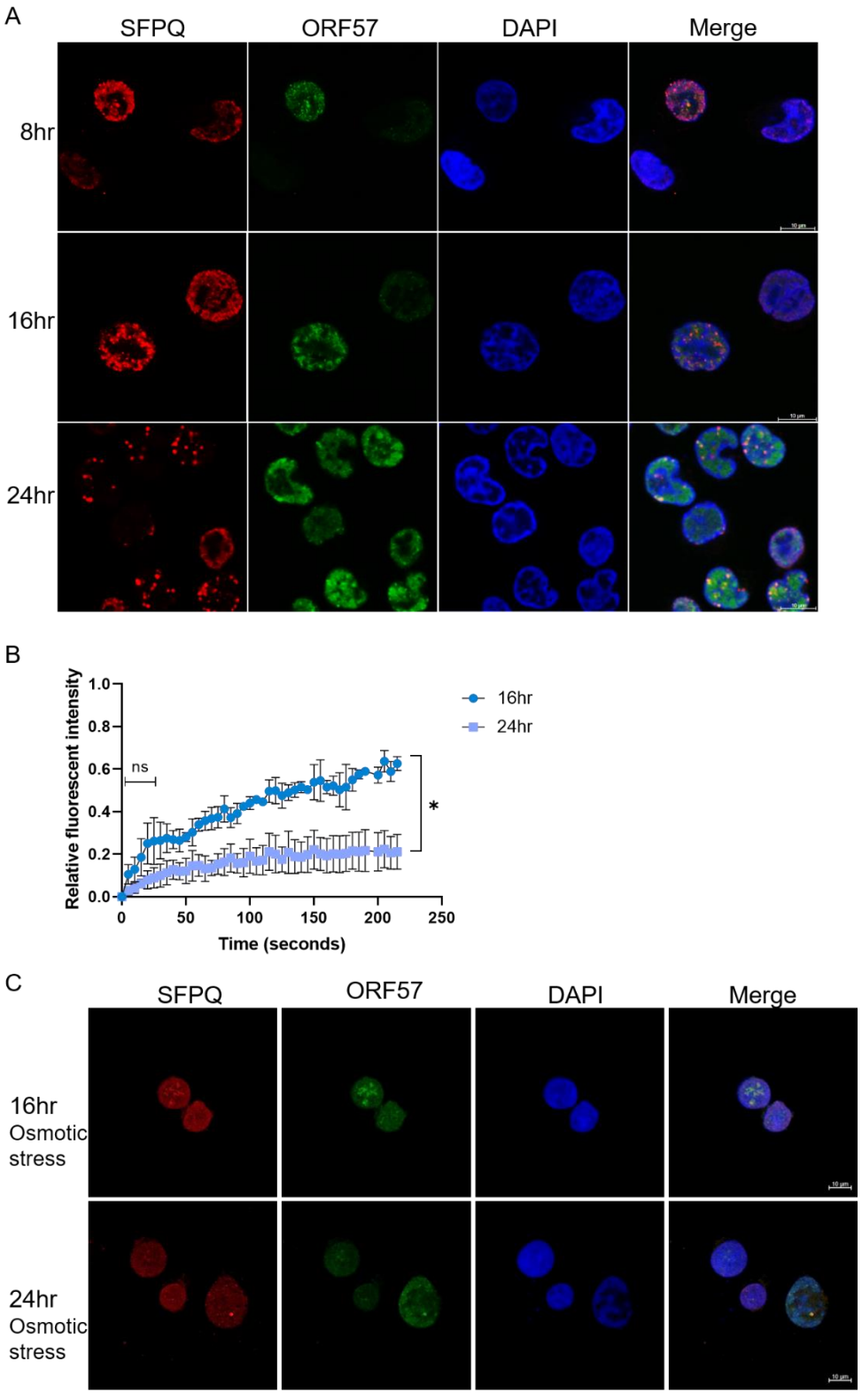

Figure 47: **SFPQ foci are condensates formed through liquid: liquid phase interactions**. **(A).** Immunofluorescence of TREx-BCBL1-RTA cells at 8, 16 and 24 hours with SFPQ (red), ORF57 (green) and DAPI (blue). Scale bars represent 10 μm. **(B)** FRAP analysis performed at 16 hours and 24 hours post-reactivation in GFP-SFPQ-TREx-BCBL1-RTA cells, with fluorescent intensity measured against time (seconds). **(C).** Immunofluorescence at 16 and 24 hours in TREx-BCBL1-RTA cells with SFPQ (red), ORF57 (green) and DAPI (blue) undergoing osmotic stress. Scale bars are 10 μm. N=3. Error bars represent SD. *P < 0.05, **P < 0.01 and ***P < 0.001 (unpaired Student's t-test).

5.4 SFPQ foci localise next to viral replication centres

KSHV forms viral replication centres (VRCs) during lytic replication, where viral transcription, viral DNA replication and assembly of virion particles all occur. To determine whether the SFPQ foci are associated with or co-localise within the VRCs, co-immunofluorescence studies were performed using SFPQ- and RNA pol II-specific antibodies. RNA pol II staining was used to visualise the kidney-shaped VRC structures in the nucleus. Interestingly, SFPQ foci were found to localise adjacent and around the VRCs and not within them. This suggests that SFPQ foci may have a role in viral RNA processing (Figure 48).

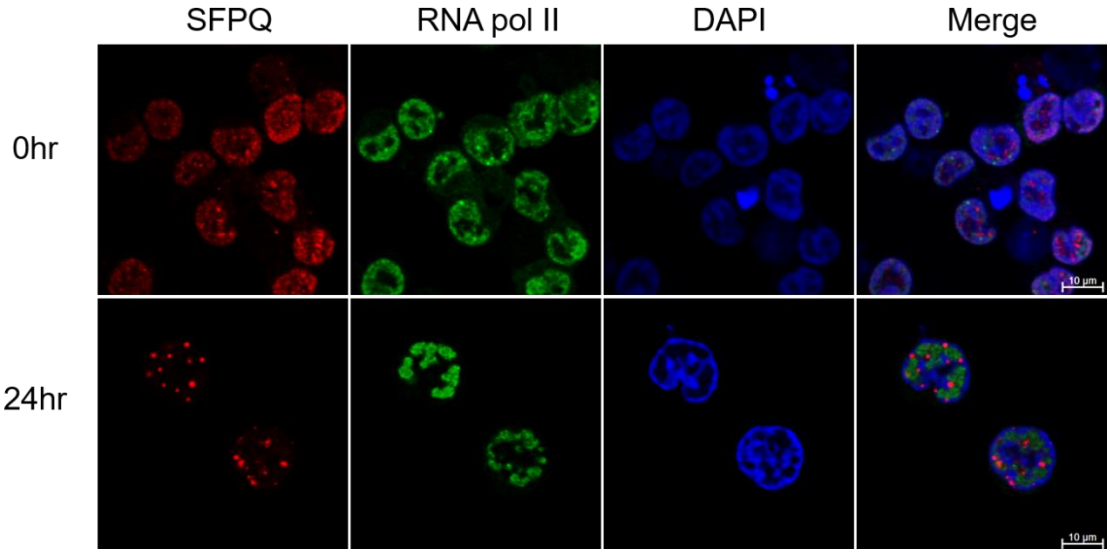

Figure 48: SFPQ foci cluster around VRCs. Immunofluorescence of TREx-BCBL1-RTA cells at 0 and 24 hours probed against SFPQ (red), RNA pol II (green) and DAPI (blue). Scale bars represent 10 µm.

5.5 KSHV-induced SFPQ foci are not due to a cellular stress response

SFPQ is a core component of paraspeckles, which can be induced by a stress response. To determine whether the foci observed during KSHV lytic replication were induced by the virus or were simply a cellular stress response to reactivation, immunofluorescence studies were performed in TREx-BCBL1-RTA cells following stress treatment. Here serum starvation was utilised for 24 hours prior to reactivation for a further 24 hours. Although serum starving can induce changes in the cell cycle and thus have unforeseen

non-specific effects on viral replication and on the cell, previous work on the cell cycle has shown that 24 hours of serum starving does not inhibit viral replication (Figure 40), and due to the foci forming early in lytic replication, any effects on the lytic cascade are likely to be miniminal,TREx-BCBL1-RTA cells were then stained for SFPQ and also the KSHV ORF57 protein to ensure stress treatments alone were not inducing lytic replication. Serum starving for 24 pre lytic induction did not induce SFPQ foci, even in pre-apoptotic cells. Additionally, stress treatments in combination with doxycycline, led to reduced SFPQ foci than usually observed, with the majority of cells demonstrating a diffuse SFPQ staining, while the few foci that were still forming were misshapen, less than 1 μM and numbering fewer than three per cell. This implies a severe stress response can impair their formation (Figure 49).

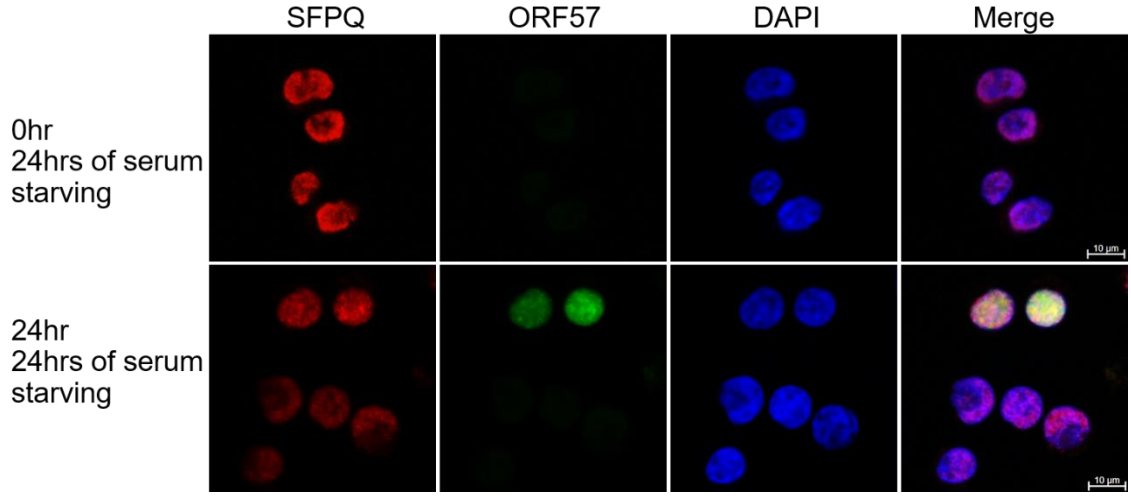

Figure 49: SFPQ foci are not a stress response. TREx-BCBL1-RTA cells were serum starved for either 24 hours or 48 hours prior to 24 hours of reactivation. Immunofluorescence was performed staining for SFPQ (red), ORF57 (green) and DAPI (blue). Scale bars are 10 μm

5.6 SFPQ foci are not nuclear speckles

A variety of sub-nuclear domains have been identified which are utilised by viruses to enhance their replication, for example PML bodies by adenoviruses and nuclear speckles by influenza virus [199] [297] . Paraspeckles are distinct but closely aligned with nuclear speckles, therefore immunofluorescence studies were performed to determine whether the unique SFPQ-containing foci co-localise with a known nuclear speckle marker, SRSF2. Here latent and reactivated TREx-BCBL1-RTA cells were

stained with SRSF2 and SFPQ-specific antibodies to explore whether KSHV-induced SFPQ foci co-localised with nuclear speckles. Nuclear speckles were found to be present in both latent and lytic cells, but in both cases failed to co-localise with the SFPQ foci. Interestingly, SRSF2 labelled nuclear speckles seems to be redistributed into larger foci during KSHV lytic replication, similar to previous observations in other gamma-herpesvirus infections [298] (Figure 50). Moreover, although the SFPQ foci are distinct, it was noted in multiple cases they localise adjacent to the larger SRSF2 foci, which is reminiscent of paraspeckles. This is further evidence that the SFPQ foci are a novel sub-nuclear domain distinct from nuclear speckles.

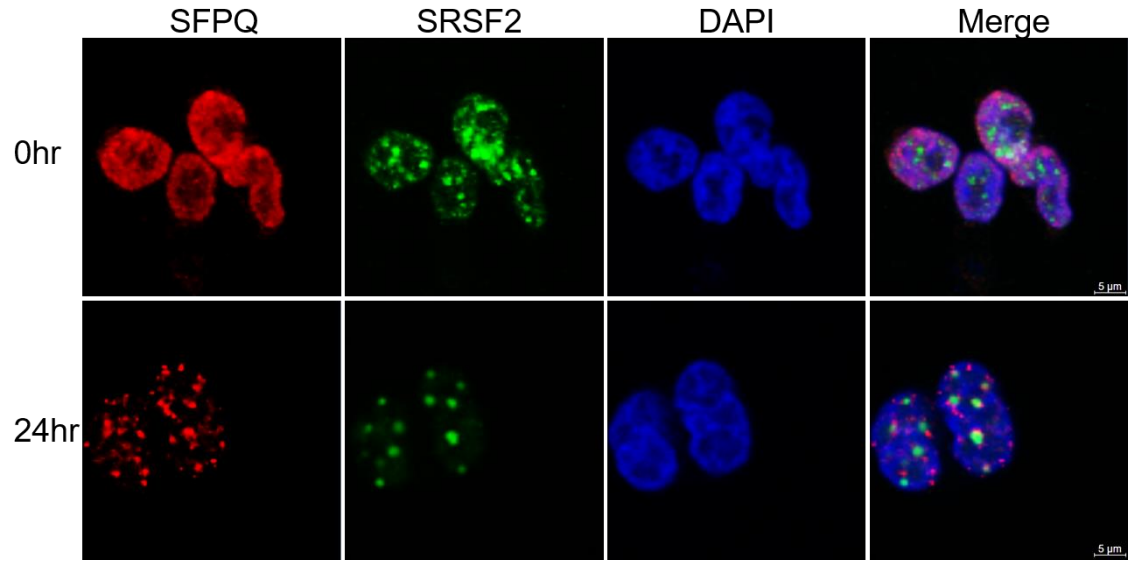

Figure 50: SFPQ foci are not nuclear speckles. TREx-BCBL1-RTA cells were induced for 0 and 24 hours and immunofluorescence performed staining for SFPQ (red), SRSF2 (green) and DAPI (blue). Scale bars are 5 µm.

5.7 SFPQ foci contain NEAT1, suggesting they are virus modified paraspeckles

The lncRNA NEAT1 is the driving force and an essential component for paraspeckle formation. However, qPCR analysis showed that NEAT1 levels do not increase during lytic replication (Figure 51A). This is a rather surprising result, bearing in mind that the KSHV-mediated SFPQ foci are much larger than normal paraspeckles and levels of NEAT1 are meant to correlate with paraspeckle size. Therefore, to examine if NEAT1 is

a component of the KSHV-mediated SFPQ foci RNA FISH was performed to determine NEAT1 localisation during KSHV lytic replication. Intriguingly NEAT1 was localised in small puncta in latent TREx-BCBL1-RTA cells which failed to localise with the diffuse staining of SFPQ. However, at 24 hours post-lytic reactivation, there was a clear co-localisation between NEAT1 and SFPQ in the large foci (Figure 51b). This suggests that the large foci seen in KSHV-infected cells do contain multiple paraspeckle components.

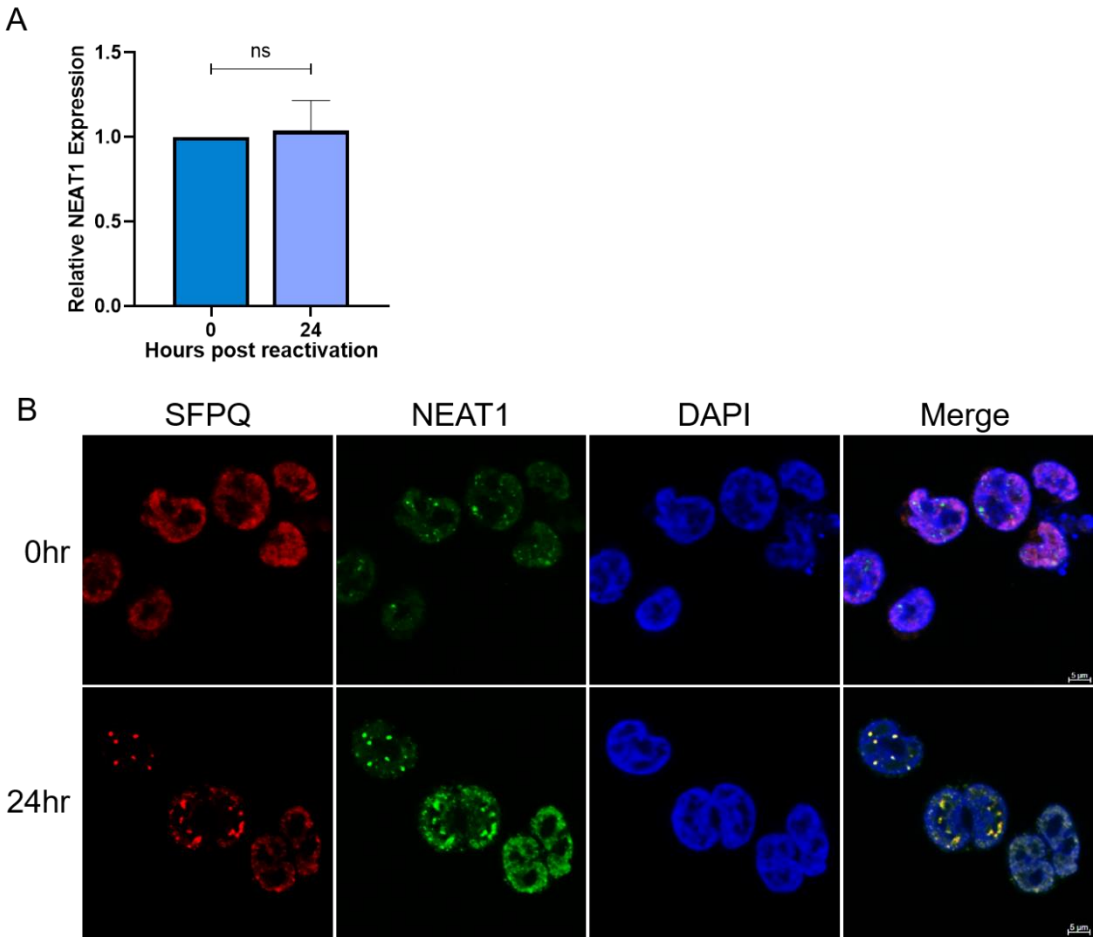

Figure 51: SFPQ foci contain NEAT1. (A). qPCR analysis of NEAT1 levels at 0 and 24 hours post-reactivation in TREx-BCBL1-RTA cells. N=3. Fold change relative to 0 hours. **(B).** Immunofluorescence of TREx-BCBL1-RTA cells at 0 and 24 hours post-reactivation with SFPQ (red), NEAT1 (green) and DAPI (blue). Scale bars represent 5 μm.

5.8 Other core paraspeckle components localise to SFPQ foci

SFPQ is a core paraspeckle protein, therefore to investigate whether the KSHV-induced SFPQ foci contained other known paraspeckle-associated proteins. Immunofluorescence studies were performed at 0 and 24 hours post-lytic reactivation

using an antibody specific to the known core paraspeckle marker, PSPC1. However, due to PSPC1 and SFPQ antibodies being the same species, fluorescent secondary antibodies cannot differentiate between them thus co-localisation with SFPQ could not be performed, instead ORF57 was used as a marker for lytic replication. PSPC1 was indeed localised in foci at 24 hours post-reactivation, suggesting that SFPQ foci are modified paraspeckles (Figure 52).

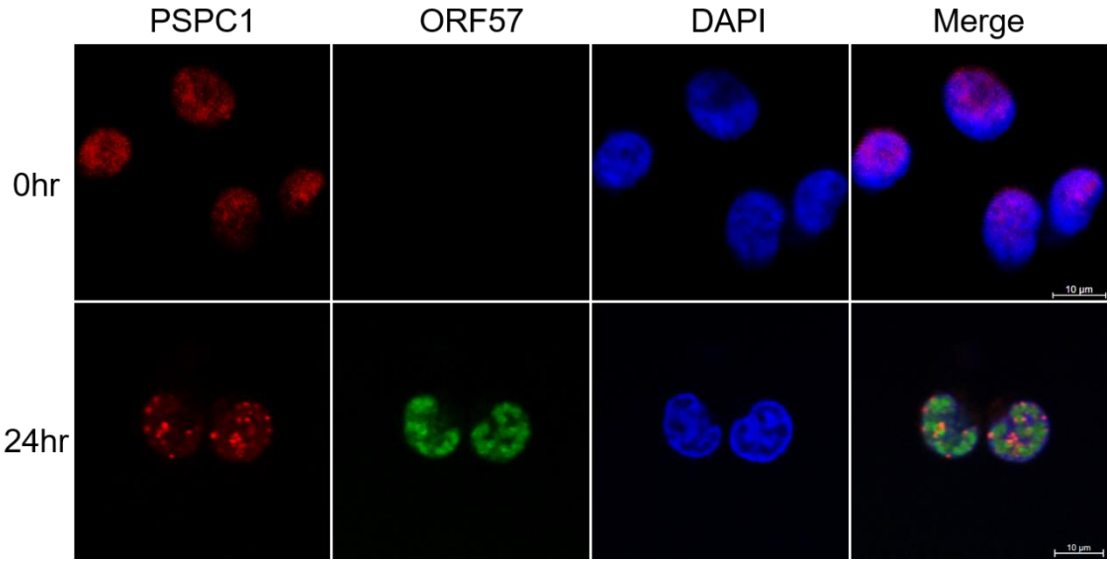

Figure 52: P**SPC1 shows the same pattern during lytic replication**. IF of TREx-BCBL1-RTA cells at 0 and 24 hours post-induction with IF staining against PSPC1 (red), ORF57 (green) and DAPI (blue). Scale bars are 10 µm.

5.9 SFPQ depletion does not affect cell growth of latently-infected TREx-BCBL1-RTA cells

To determine whether SFPQ is important for KSHV lytic replication, a stable knockdown of SFPQ in TREx-BCBL1-RTA cells was established using lentivirus-shRNA transduction. Successful depletion of SFPQ was confirmed at both RNA and a protein level using qPCR and immunoblotting at 0 and 24 hours post-induction, with an approximate 50% and 60% knockdown at RNA and protein levels respectively (Figure 53A-B). As SFPQ disruption has been linked to several disease states, such as ALS, through accumulation of truncated incorrectly spliced RNA transcripts [299], several assays were performed examining whether SFPQ depletion affected cell growth and viability in latently-infected TREx-BCBL1-RTA cells. SFPQ has been shown to be

important for the inclusion of exon 30 of the myosin heavy chain [222], therefore to confirm an efficient inactivation of SFPQ activity, exon 30 levels were examined via qPCR in depleted cells versus scrambled control cells. Levels of exon 30 inclusion were reduced in SFPQ-depleted latently infected TREx-BCBL1-RTA cells compared to scrambled control cell lines, confirming depletion of SFPQ activity (Figure 53C). Furthermore to assess what affect SFPQ depletion had on the growth of latently-infected TREx-BCBL1-RTA cells, cell count assays were performed over a 72 hour time period. No significant change in growth rates between control scramble and SFPQ-depleted cell lines was observed suggesting their viability was not affected by SFPQ knockdown (Figure 53D).

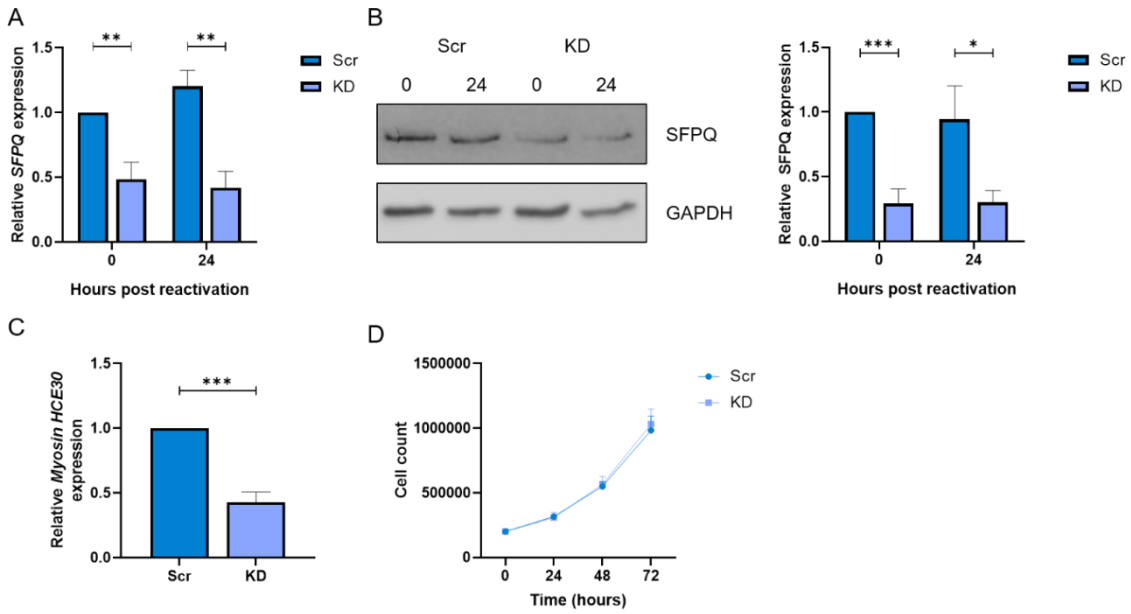

Figure 53: SFPQ depletion does not affect B cell growth. (A). qPCR of SFPQ levels in 0 and 24 hour induced scr and SFPQ KD TREx-BCBL1-RTA cell lines. **(B).** Representative western blot of SFPQ levels in 0 and 24 hours induced scr and SFPQ KD TREx-BCBL1-RTA cells with densitometry analysis. **(C).** qPCR of Myosin Heavy Chain Exon 30 in scr and SFPQ KD TREx-BCBL1-RTA cells at 0 hours. **(D).** Cell counts of scr and SFPQ KD TREx-BCBL1-RTA cells at 0, 24, 48 and 72. N=3. Fold change relative to Scr 0 hours Error bars represent SD. *P < 0.05, **P < 0.01 and ***P < 0.001 (unpaired Student's t-test).

5.10 SFPQ is essential for KSHV lytic replication

To determine what effect SFPQ depletion had on KSHV lytic replication, scrambled and SFPQ stable knockdown TREx-BCBL1-RTA cells were reactivated and lytic replication was assessed by measuring various lytically expressed genes at RNA and protein levels covering the temporal cascade. qPCR analysis showed that SFPQ depletion resulted in

a significant decrease in ORF57, PAN, K8 and ORF65 RNA levels, which was further supported by immunoblotting showing a corresponding reduction in ORF57 and ORF65 levels, compared to scrambled control (Figure 54A-C). This suggests SFPQ is essential in the early stages of KSHV lytic replication. The essential nature of SFPQ on KSHV lytic replication was reinforced using both viral load and virus reinfection assays. Again a significant decrease in both viral genome replication and infectious virion production was observed in SFPQ depleted cells compared to scrambled control (Figure 54D-E).

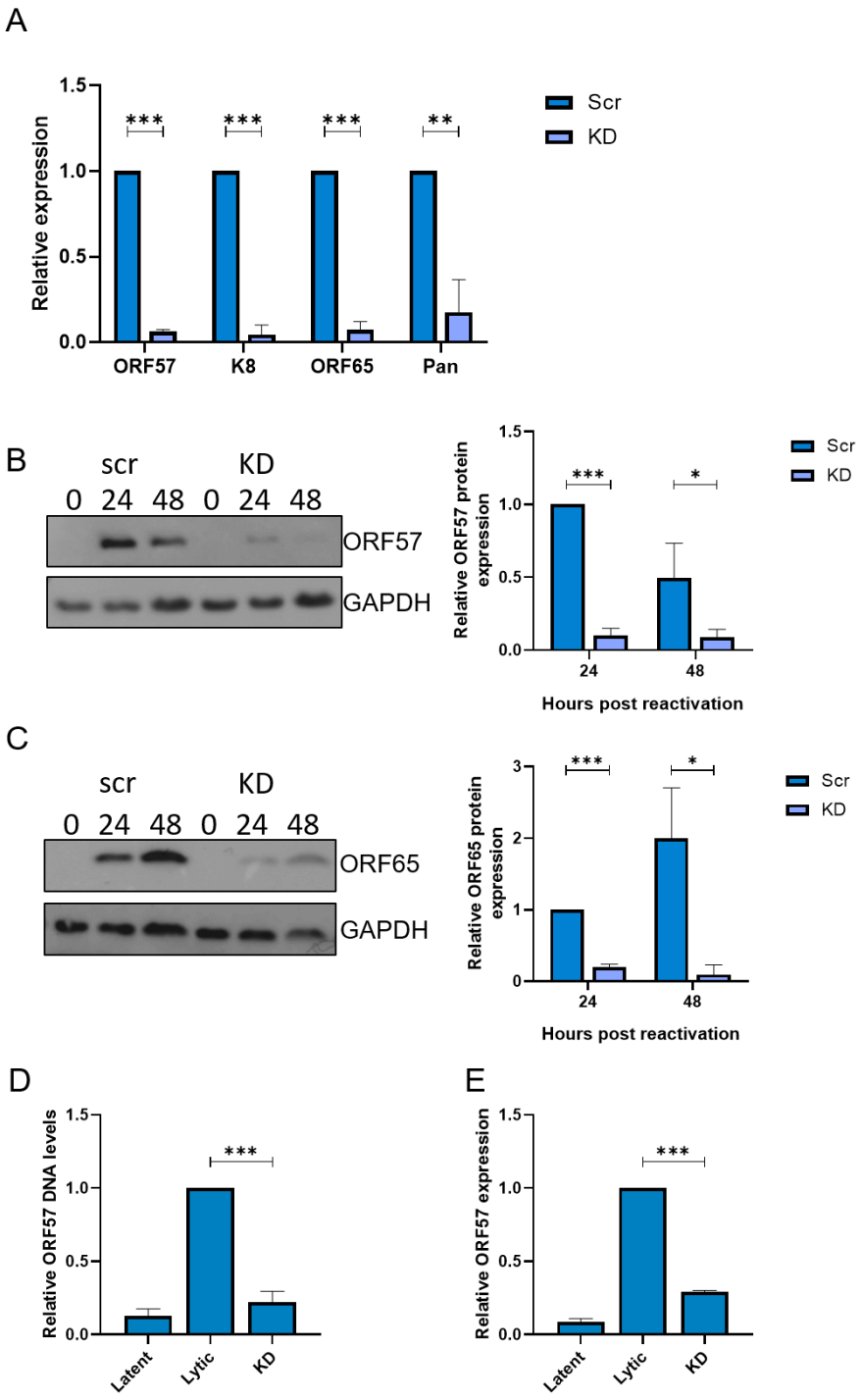

Figure 54: SFPQ is essential for KSHV lytic replication. (A). qPCR of viral genes ORF57, K8, ORF65 and PAN at 24 hours in scr and SFPQ KD TREx-BCBL1-RTA cells. **(B).** Representative western blot of ORF57 at 0, 24 and 48 hours in scr and SFPQ KD TREx-BCBL1-RTA cells with densitometry analysis. **(C).** Representative western blot of ORF57 at 0, 24 and 48 hours in scr and SFPQ KD TREx-BCBL1-RTA cells with densitometry analysis. **(D).** Viral load in scr and SFPQ KD TREx-BCBL1-RTA cell lines. (E). Infectious virion production in scr vs SFPQ KD cells. N=3. Error bars represent SD. *P < 0.05, **P < 0.01 and ***P < 0.001 (unpaired Student's t-test).

5.11 SFPQ KD disrupts foci formation

To determine whether depletion of SFPQ has any effect on the formation of the SFPQ-localised foci, immunofluorescence studies were performed using PSCP1-specific antibodies on reactivated scramble control or SFPQ knockdown TREx-BCBL1-RTA cells. PSCP1 was chosen as a marker for these foci, as a core component of paraspeckles and is known to interact with SFPQ. Results show that SFPQ depletion caused a significant impairment in foci formation with most cells failing to form any foci, or if formed they were much smaller and less regular, instead PSCP1 was diffusely localised throughout the nucleus (Figure 55). In contrast, large foci were again observed upon PSCP1 staining in scrambled control cells. Furthermore as previous observed (Figure 54) ORF57 levels were significantly perturbed in the SFPQ depleted cells at 24 hour post reactivation, supporting the essential nature of SFPQ in KSHV lytic replication. This suggests that SFPQ is essential for virus-mediated foci formation.

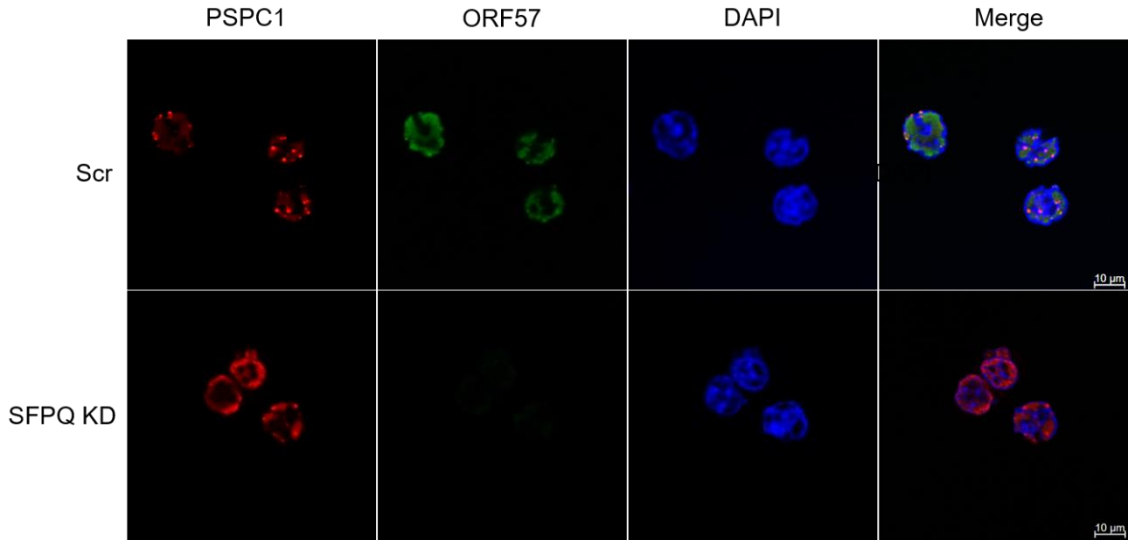

Figure 55: **SFPQ depletion inhibits puncta formation.** Scr and SFPQ knockdown TREx-BCBL1-RTA cells were induced 24 hours and IF stained for PSPC1 (red), ORF57 (green) and DAPI (blue). Scale bars are 10 µm.

5.12 NONO KD affects KSHV lytic replication

Previous experiments have shown that SFPQ is essential for KSHV lytic replication. Therefore, additional experiments were performed to assess whether depletion of other paraspeckle-associated proteins had a similar phenotype. To this end, a stable

knockdown of NONO was performed in TREx-BCBL1-RTA cells using lentivirus-shRNA transduction. Successful depletion of NONO was confirmed at both RNA and protein levels using qPCR and immunoblotting at 0 and 24 hours post-induction (Figure 56A-C).

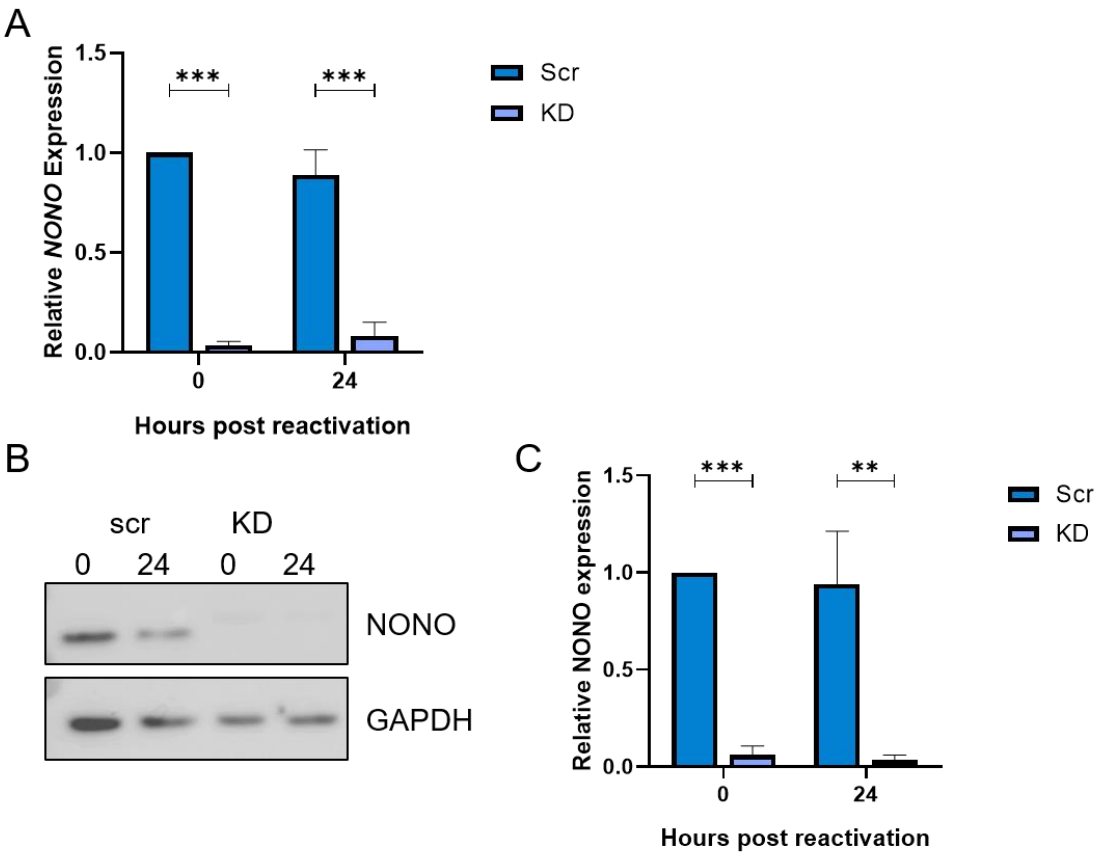

Figure 56: **NONO KD was successful. (A).** NONO levels in scrambled and NONO knockdown TREx-BCBL1-RTA cell lines analysed via qPCR. **(B).** Representative western blot showing NONO protein levels at 0 and 24 hours post lytic replication. GAPDH is a housekeeper with densitometry analysis **(C).** N=3. Fold change relative to Scr 0. Error bars represent SD. *P < 0.05, **P < 0.01 and ***P < 0.001 (unpaired Student's t-test).

The effect of NONO knockdown on KSHV lytic replication was then analysed using qPCR and immunoblotting of a panel of viral genes encompassing the lytic temporal cascade. NONO knockdown led to a significant decrease in viral gene expression, however, the effect was less dramatic than seen with SFPQ depletion (Figure 57A). A corresponding decrease in ORF65 protein levels was also observed via immunoblotting, however, there was no significant effect on the levels of ORF57 (Figure 57B-C). Additionally immunofluorescence was performed in scrambled or NONO depleted TREx-BCBL1-RTA cells at 24 hours post-lytic replication, using ORF57-and SFPQ-specific antibodies

to determine whether NONO is required for the formation of the SFPQ-foci observed during lytic replication. In scrambled TREx-BCBL1-RTA cells, SFPQ formed the previously observed large foci while ORF57 stained to the viral replication centres, whereas in the NONO depleted cells, SFPQ and ORF57 showed diffuse staining throughout the nucleus (Figure 57D). Overall this suggests that NONO plays an essential role in lytic replication and in the formation of the SFPQ foci.

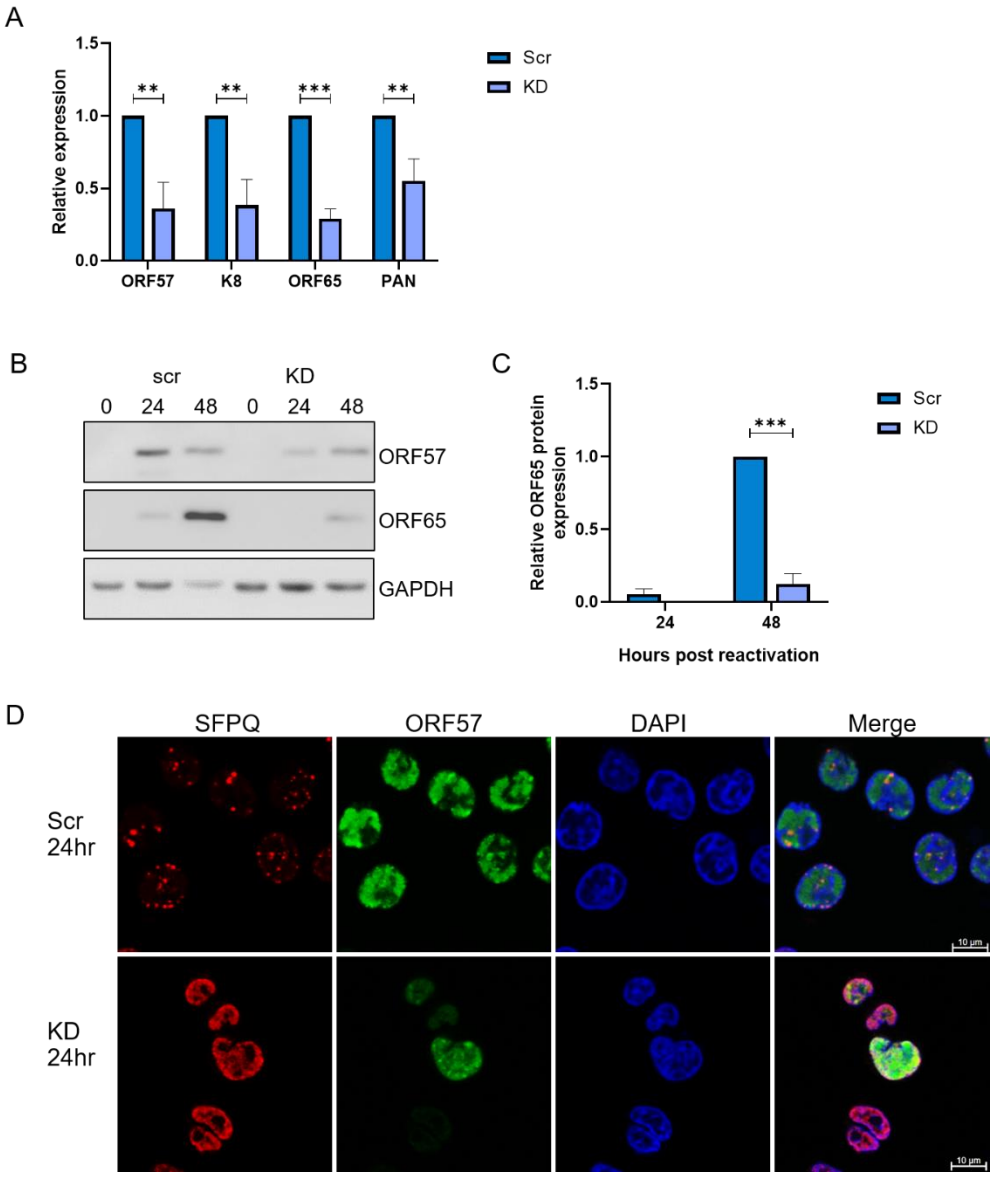

Figure 57: NONO KD reduces viral replication. (A). qPCR of a panel of viral genes at 0 and 24 hours post induction in scrambled and NONO knockdown TREx-BCBL1-RTA cells. **(B).** Representative western blot showing ORF57 and ORF65 levels in scrambled and NONO knockdown TREx-BCBL1-RTA cells at 0, 24 and 48 hours post reactivation, with densitometry analysis **(C)**. **(D).** Immunofluorescence in scrambled and NONO KD TREx-BCBL1-RTA cells at 24 hours post-lytic replication with staining against SFPQ (red), ORF57 (green) and DAPI (blue). Scale bars are 10 µm. N=3. Error bars represent SD. *P < 0.05, **P < 0.01 and ***P < 0.001 (unpaired Student's t-test).

5.13 NEAT1 depletion inhibits KSHV lytic replication

Depletion of both SFPQ and NONO had a detrimental effect on KSHV lytic replication, suggesting that the viral-modified paraspeckles are essential for successful KSHV lytic replication. NEAT1 is the structural scaffold of paraspeckles, as such, depletion of NEAT1 was carried out, and to confirm whether it is the formation of the virally modified paraspeckles that is essential. Depletion was performed using LNA-GapmeRs that rely on enzymatic degradation of the target RNA rather than Ago2 mediated degradation that shRNAs utilise, allowing better depletion of the nuclear retained NEAT1. Successful knockdown was confirmed at the RNA level via qPCR at 0 and 24 hours post-reactivation in TREx-BCBL1-RTA cells (Figure 58A).

The effect on KSHV lytic replication was then analysed via immunoblotting of lytic proteins, no change was detected in early ORF57 protein levels, however, a significant reduction was observed in levels of the late viral protein ORF65 (Figure 58B-C). Moreover, immunofluorescence with SFPQ- and ORF57-specific antibodies was performed to determine if depletion of NEAT1 and the subsequent reduction in viral replication could be connected to a loss of the foci. Depletion of NEAT1 did indeed inhibit the formation of the condensates, with a diffuse staining of SFPQ observed compared to the large foci seen in the scrambled control cell line. Furthermore, similar to NONO depletion, ORF57 failed to localise to viral replication centres in GapmeR treated cells, again showing a diffuse staining (Figure 58D). Overall this suggests that NEAT1 is essential for viral induced paraspeckle formation, with depletion of NEAT1 resulting in a detrimental effect on KSHV lytic replication.

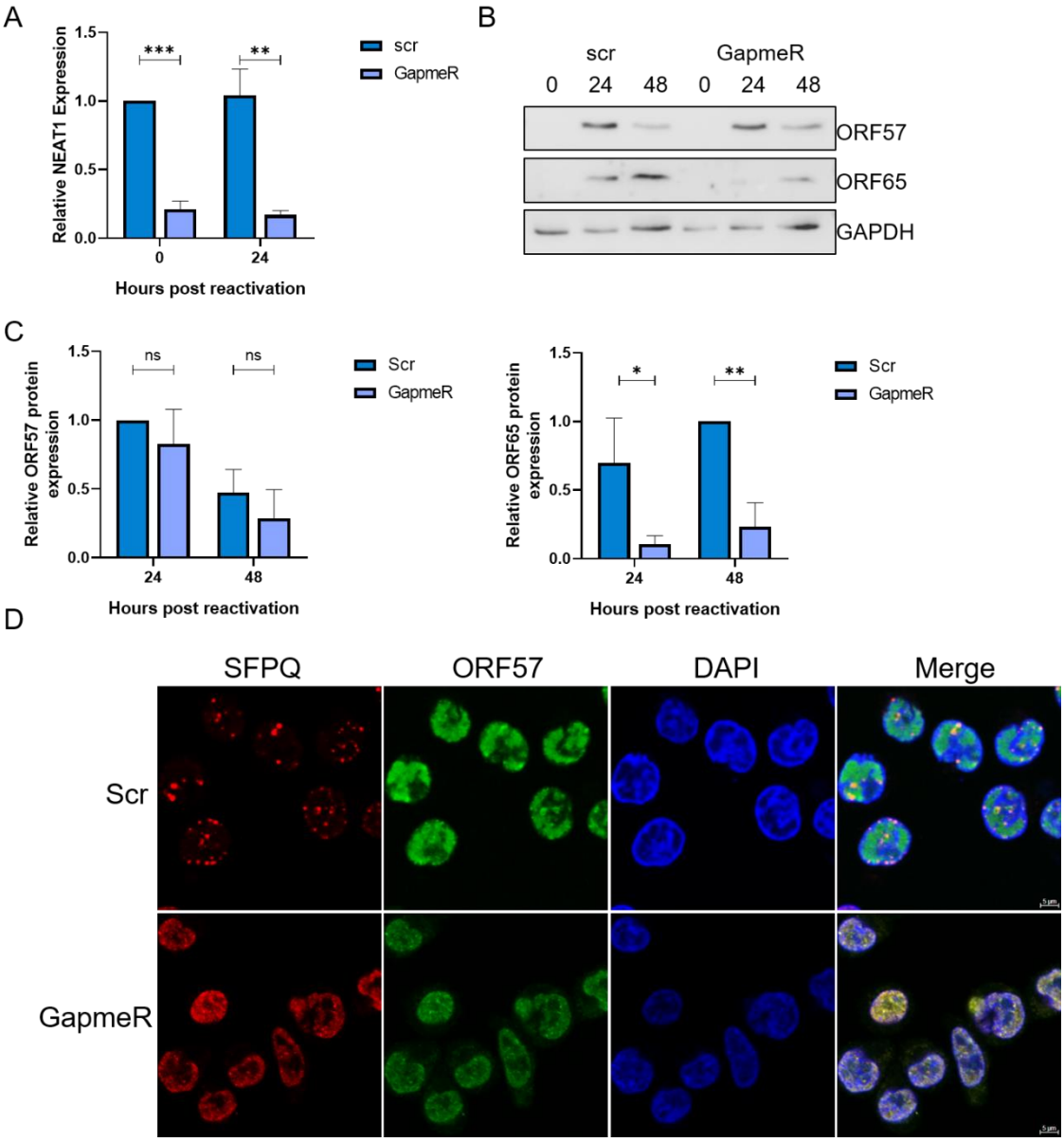

Figure 58: NEAT1 KD reduces viral replication. (A). qPCR of NEAT1 levels in scrambled and GapmeR treated TREx-BCBL1-RTA cells at 0 and 24 hours post. **(B).** Representative western blot showing ORF57 and ORF65 levels in NEAT1 KD at 0, 24 and 48 hours post-reactivation. **(C).** Densitometry analysis of western blotting for ORF57 and ORF65 in scrambled and GapmeR treated TREx-BCBL1-RTA cells at 0, 24 and 48 hours post-reactivation. **(C).** Immunofluorescence in scrambled and NEAT1 KD TREx-BCBL1-RTA cells at 24 hours post-lytic reactivation with staining against SFPQ (red), ORF57 (green) and DAPI (blue). Scale bars are 5 μm. N=3. Fold change relative to 0 hours. Error bars represent SD. *P < 0.05, **P < 0.01 and ***P < 0.001 (unpaired Student's t-test).

5.14 KSHV ORF11 associates with SFPQ foci

Results have shown that the SFPQ foci are formed during the early stages of KSHV lytic replication; between 16 and 24 hours post-reactivation. To identify potential KSHV-encoded proteins which are driving or enhancing the formation of these large virus-induced foci, co-immunoprecipitations were performed to assess potential interactions between SFPQ and a panel of early KSHV proteins, namely RTA/ORF50, ORF57 and

ORF11. Co-immunoprecipitations were performed using a SFPQ-specific antibody in TREx-BCBL1-RTA cells reactivated at 24 hours to assess possible interactions with KSHV RTA/ORF50 or ORF57 proteins, with NONO acting as a positive control. However, due to the lack of a KSHV ORF11 antibody, the putative interaction between SFPQ and ORF11 was assessed in a TREx-BCBL1-RTA cell line stably transduced with a lentivirus expressing ORF11-FLAG (created by Elena Harrington, Whitehouse laboratory). Results showed no association between SFPQ and the ORF50 or ORF57 proteins, however, a potential interaction between SFPQ and ORF11 was observed 24 hours post reactivation (Figure 59A-B). To confirm this association and determine when during lytic replication ORF11 and SFPQ interact, an immunofluorescence time course was performed in TREx-BCBL1-RTA-FLAG-ORF11 cells at 0, 8, 16 and 24 hours post-reactivation using FLAG- and SFPQ-specific antibodies. Intriguingly SFPQ and FLAG-ORF11 showed strong co-localisation early in foci formation at 8 hours, continuing at 16 and 24 hours post-reactivation (Figure 59C). This may suggest ORF11 mediates either a role in the formation of the viral-induced foci or is required for correct formation or functionality.

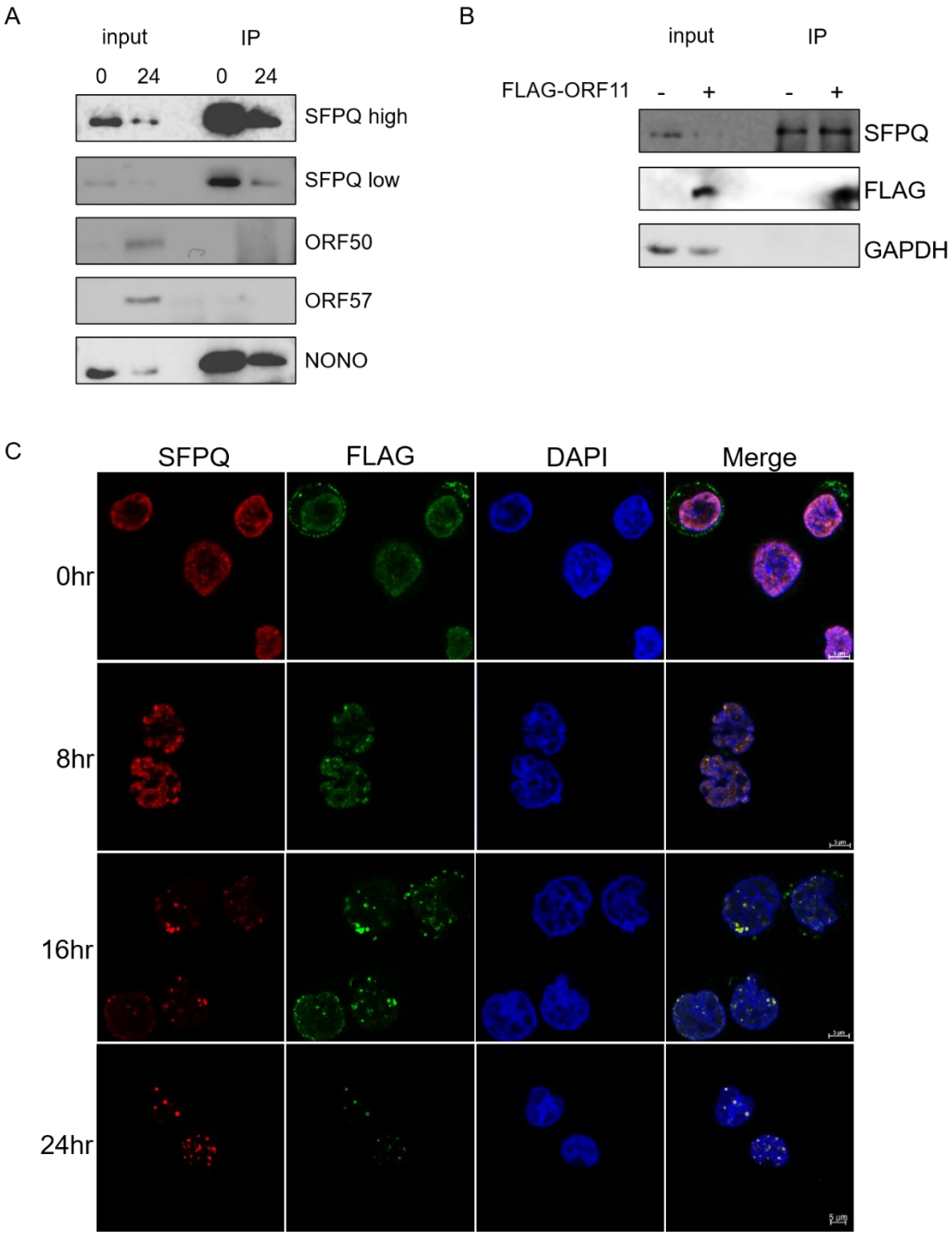

Figure 59: SFPQ and FLAG-ORF11 localise during lytic replication. (A). SFPQ immunoprecipitations were performed in TREx-BCBL1-RTA cells at 0 and 24 hours post-reactivation and probed for SFPQ, ORF50, ORF57 and NONO. Composite blot. **(B).** SFPQ immunoprecipitation in both TREx-BCBL1-RTA and TREx-BCBL1-RTA-FLAG-ORF11 at 24 hours post-lytic reactivation and probed for SFPQ, FLAG and GAPDH. N=3. **(C).** Immunofluorescence in TREx-BCBL1-RTA-FLAG-ORF11 at 0, 8, 16 and 24 hours post-lytic reactivation. Stained for SFPQ (red), FLAG (green) and DAPI (blue). Scale bars represent 5 μm.

5.15 SFPQ and ORF11 co-localise in condensates forming ring structures

To further investigate the co-localisation between ORF11 and SFPQ, AiryScan imaging was performed with a confocal laser scanning microscope equipped with an AiryScan detection unit to visualise the condensates at higher resolution. A clear co-localisation of ORF11 with the SFPQ foci was observed in TREx-BCBL1-RTA-FLAG-ORF11 cells at 24 hours post-reactivation. Surprisingly the higher resolution obtained by AiryScan imaging suggested that the foci are not uniform solid structures, rather, they form rings with both SFPQ and ORF11 localising to the outside of the ring structure (Figure 60A). Z-stack images of the KSHV-induced SFPQ/ORF11 foci were then taken to investigate their structure through different planes of the cell. These images confirmed that the foci were found throughout the cell, confirming that the rings are a subsection of a sphere-shaped focus (Figure 60B)

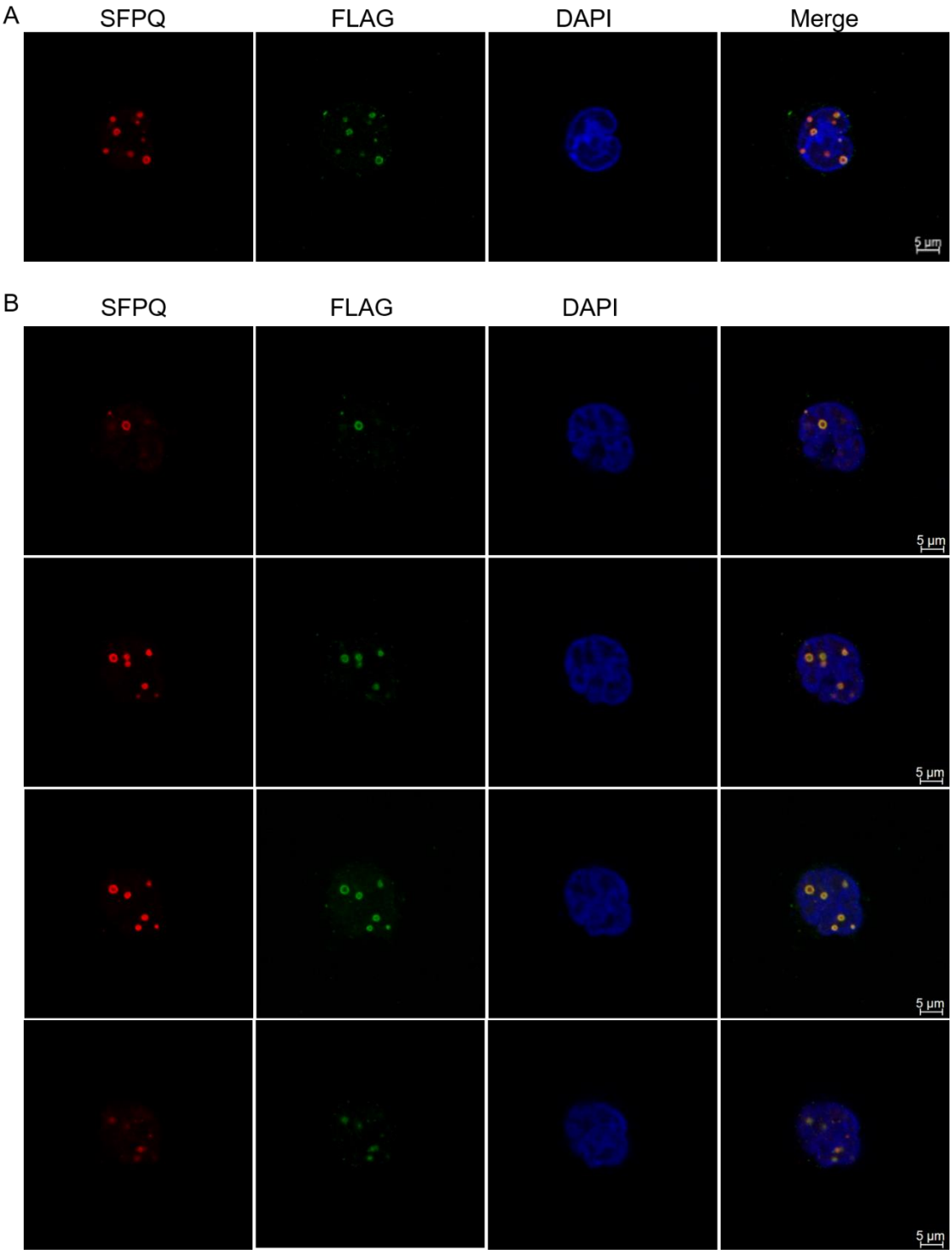

Figure 60: SFPQ and ORF11 co-localise in foci forming ring structures. (A). Immunofluorescence at 24 hours in TREx-BCBL1-RTA-FLAG-ORF11 cells stained for SFPQ (red), green (FLAG), DAPI (blue). Scale bar is 5 µm. **(B).** Z-stack of a TREx-BCBL1-RTA-FLAG-ORF11 cell 24 hours post reactivation at different planes. Cell stained for SFPQ (red), FLAG (green) and DAPI (blue). Scale bars are 5 µm.

FISH was then used in combination with AiryScan imaging to determine whether NEAT1 was also present in the ring structure during lytic replication. Results show that NEAT1

also localised to this ring structure (Figure 61A). Zeiss Zen software was subsequently used in combination with FISH to analyse the size of these foci. Imaging of latently infected TREx-BCBL1-RTA cells suggested the NEAT1 foci were approximately 0.5 µM in diameter, which are reminiscent of canonical paraspeckles which range from 0.1-1 µM. Notably however, the viral-induced foci at 24 hours post-reactivation were much larger, often over 2 µm in diameter, once again implicating that the foci at 24 hours are virally-modified paraspeckles (Figure 61B).

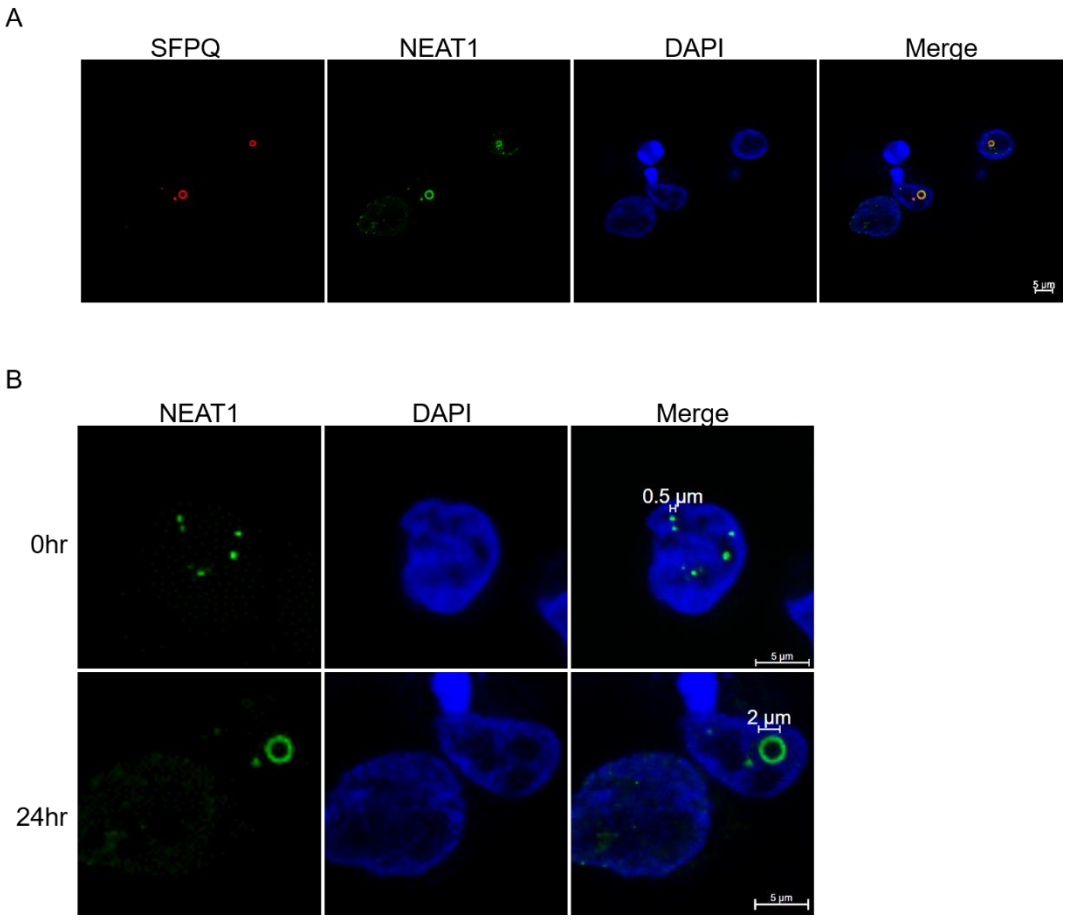

Figure 61: NEAT1 forms a ring structure. (A). FISH in TREx-BCBL1-RTA cells 24 hours post-reactivation probed for SFPQ (red), NEAT1 (green) and DAPI (blue). Scale bars are 5 µm. (B). FISH in TREx-BCBL1-RTA cells at 0 and 24 hours post-reactivation probed for NEAT1 (green) and DAPI (blue). Diameter of foci is measured using Zen blue software.

5.16 ORF11 CRISPR cells have reduced SFPQ foci formation

To determine whether KSHV ORF11 was essential for the formation of the SFPQ foci during KSHV lytic replication, it was first necessary to try and silence ORF11 expression in TREx-BCBL1-RTA cells. However, ORF11 is expressed on a polycistronic transcript,

and as such a standard lentivirus-shRNA approach could not be utilised, instead ORF11 knockout CRISPR cell lines were created by Dr Tim Mottram, Whitehouse laboratory. Immunofluorescence studies were then performed using CRISPR control and ORF11 CRISPR cells comparing latent and lytically replication cells. Cells were stained with SFPQ and ORF57-specific antibodies as a maker of lytically replicating cells. As expected CRISPR control cells showed large SFPQ foci present by 24 hours post-reactivation, which clustered around ORF57 localised in the viral replication centres In contrast, few if any SFPQ foci formed at 24 hours post-reactivation, in the majority of ORF11 CRISPR cells. Due to lower resolution of the images, the rings are not fully visible. The few foci that were identified were drastically reduced in number and size, and also appeared misshapen with more diffuse staining throughout the nuclei (Figure 62).

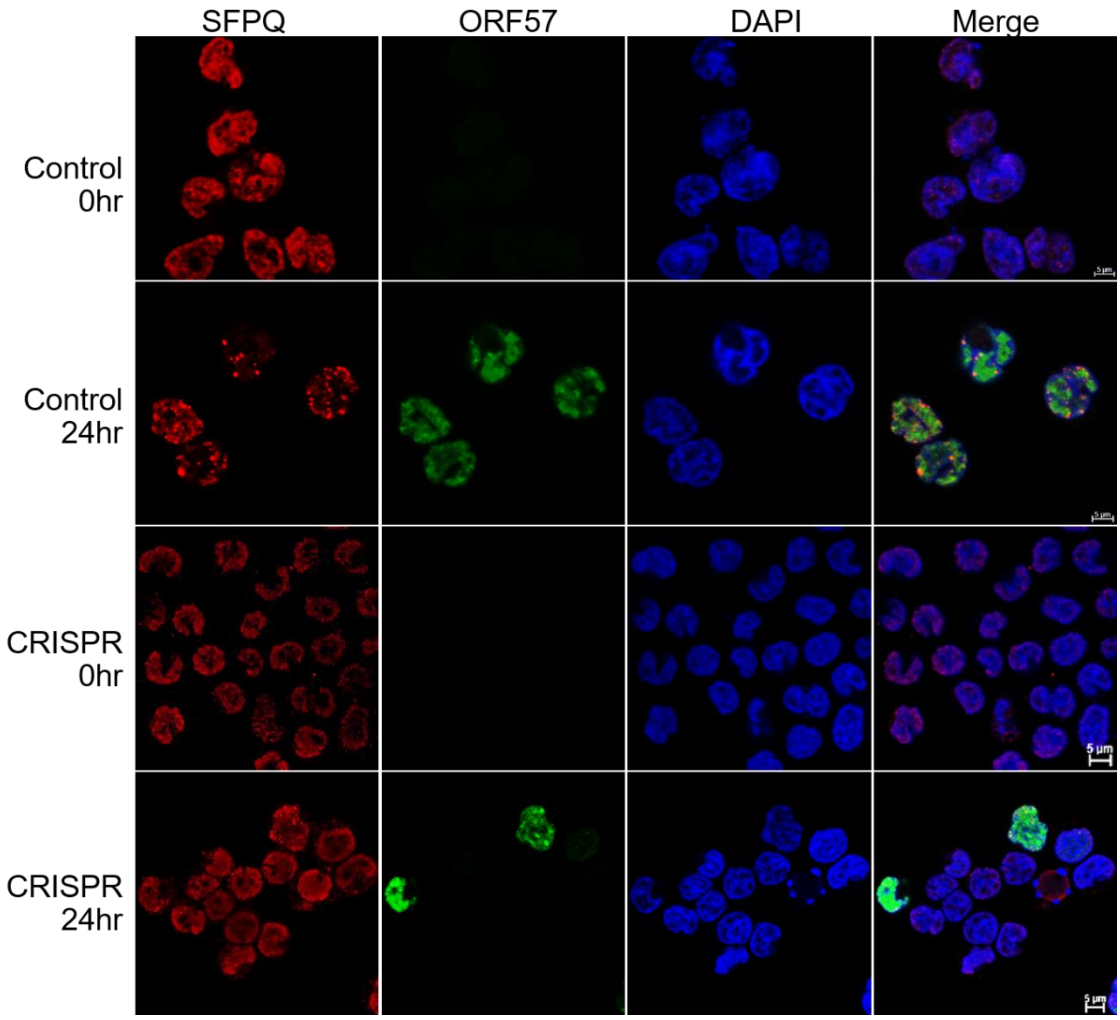

Figure 62: ORF11 CRISPR cells have reduced SFPQ foci formation. Immunofluorescence in either CRISPR control or ORF11 CRISPR TREx-BCBL1-RTA cells at 0 and 24 hours post-reactivation stained for DAPI (blue), SFPQ (red) and ORF57 (green). Scale bars are 5 µm on the 2 upper rows and 5 µm on the lower row.

Additionally, RNA FISH was performed in the CRISPR control and CRISPR ORF11 TREx-BCBL1-RTA cells to analyse the localisation of NEAT1 in the absence and presence of ORF11 during KSHV lytic replication. Similar results were observed to SFPQ localisation, with a clear localisation of NEAT1 in foci at 24 hours in control cells. However, NEAT1 remained diffuse throughout the nucleus in ORF11 CRISPR cell lines (Figure 63). Together this suggests that ORF11 is required for the formation of the large SFPQ foci during KSHV lytic replication, however, it is unknown whether it initiates their production.

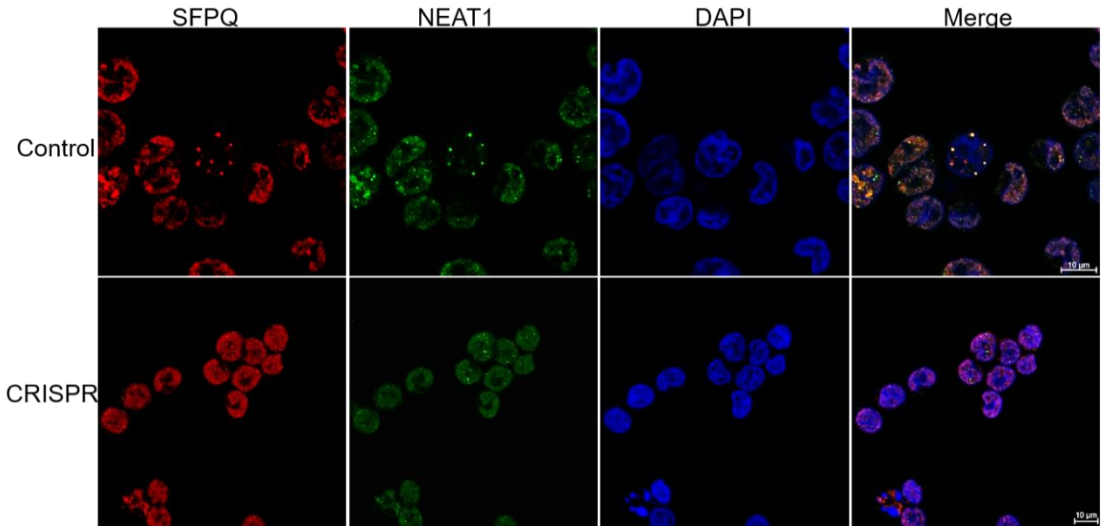

Figure 63: ORF11 CRISPR cells impact NEAT1 foci. Immunofluorescence at 24 hours post-reactivation in control and CRISPR TREx-BCBL1-RTA cells staining for SFPQ (red), NEAT1 (green) and DAPI (blue). Scale bars are 5 μm

5.17 SFPQ foci contain RNA helicases and splicing factors

The previous set of results demonstrate the formation of paraspeckle-like foci during KSHV lytic replication. These foci are essential for KSHV lytic replication, with depletion of their components significantly reducing viral gene expression. Notably, these paraspeckle-like foci are much larger than normal paraspeckle structures, suggesting that KSHV may be modifying paraspeckles to enhance its own replication. To examine the protein content of these KSHV induced foci, LC/LC-MS coupled to TMT quantitative proteomic analysis was performed in conjunction with SFPQ immunoprecipitations in TREx-BCBL1-RTA cells and ORF11-FLAG immunoprecipitations in TREx-BCBL1-RTA-FLAG-ORF11 cells comparing latent and 24 hour reactivated cells (performed by Elena Harrington, Whitehouse laboratory). Results highlighted that multiple RNA helicases and splicing factors associated with SFPQ and ORF11, summarised in tables below. Notably, of the 8 known essential paraspeckle proteins, 5 were identified above the threshold in the proteomic dataset at both 0 and 24 hours. Whereas DAZAP1 and RBM14 had reduced association at 24 hours and SMARCA4/BRG1 was not detected in the immunoprecipitations. Furthermore of the proteins which are classified as important paraspeckle proteins, only 3 were present in both time points. 4 were not above the

threshold at either time point, 2 only at 0 hours and 1 not detectable via MS. This suggests that KSHV-induced foci may have a distinct composition to canonical paraspeckles, however, further studies are required to confirm this hypothesis.

Essential	Important
DAZAP1	CPSF7
FUS	FAM98A
hnRNP H3	FIGN1
hnRNP K	hnRNP A1
NONO	hnRNP R
RBM14	hnRNP UL1
SFPQ	PCED1A
SMARCA4/BRG1	RBM12
	SRSF10
	TAF15

	Above threshold at 0 and 24 hours
	Above threshold at 0 hours
	Below threshold
	Not detected via MS

Table 7: The essential paraspeckles proteins found in the SFPQ proteomic analysis. Paraspeckle proteins that had a greater than 10% abundance in the SFPQ pulldown at 0 or 24 hours post-reactivation in TREx-BCBL1-RTA cells.

Further analysis was performed on the MS dataset to determine if additional proteins were present which may be implicated in the unique roles of the KSHV-modified foci described above. Notably, a large number of associated proteins have been implicated as splicing factors, including members of the hnRNP and dead box RNA helicase families, these proteins are summarised below in table 8. Interestingly, a high proportion of these proteins are not canonically associated with paraspeckles, further supporting the hypothesis that these are virus-modified paraspeckles.

DDX1
DDX17
DDX21
DDX3X
DDX5
DDX6
DHX30
DHX36
DHX9
hnRNP A1
hnRNP A2B1
hnRNP AB
hnRNP C
hnRNP DL
hnRNP H1
hnRNP H3
hnRNP L
hnRNP M
hnRNP R
hnRNP U
hnRNP UL1

	Can associate with PS
	Above threshold at 0 hours
	Above threshold at 24 hours
Bold	Above threshold in ORF11 MS

Table 8: Additional proteins identified in the SFPQ proteomics. Proteins of interest that had a greater than 10% abundance in the SFPQ pulldown at 0 or 24 hours post-reactivation in TREx-BCBL1-RTA cells. Proteins in bold were also above this threshold in the FLAG pulldown performed in FLAG-ORF11-TREx-BCBL1-RTA cells.

To validate the MS dataset, several of the most abundant proteins precipitated in the SFPQ and ORF11 pulldowns were selected for further investigation, these included DHX9, DDX17, DDX21, hnRNP U and hnRNP M. For this validation, depending on species of antibody available, either TREx-BCBL1-RTA cells were stained for SFPQ and the protein of interest or FLAG was stained in FLAG-ORF11 overexpression cells. It was found all 5 proteins of interest co-localised at 24 hours to different extents. hnRNP M and

DDX17 had strong co-localisation, while the others had regions of co-localisation and other distinct staining, suggestive of multiple functions.

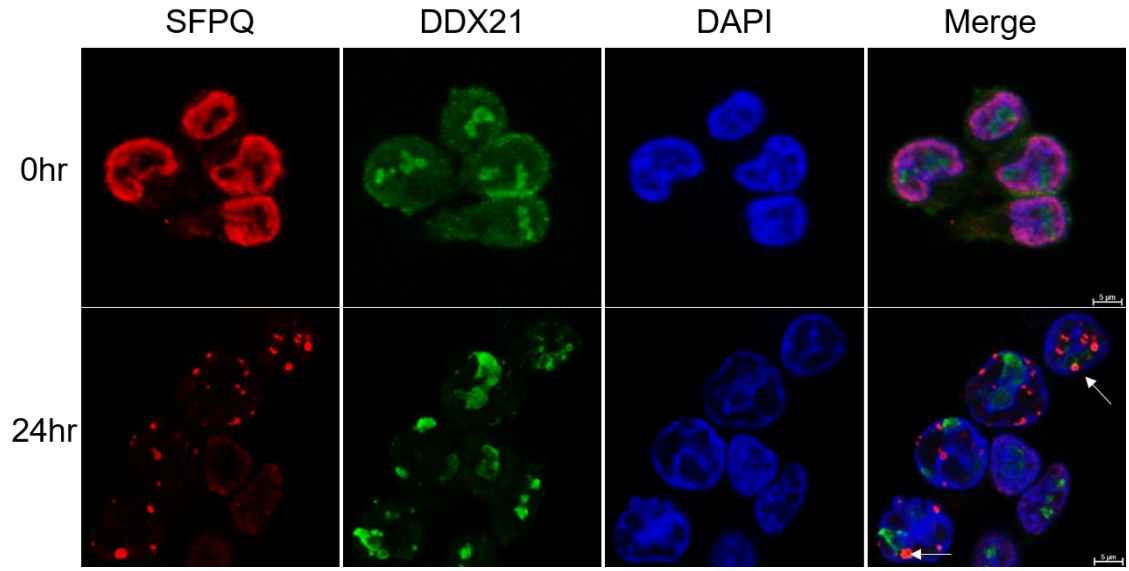

Figure 64: DDX21 staining. IF at 0 and 24 hours post-reactivation in TREx-BCBL1-RTA cells staining for SFPQ (red), DDX21 (green) and DAPI (blue). Arrows highlight co-localisation. Scale bars are 5 μm.

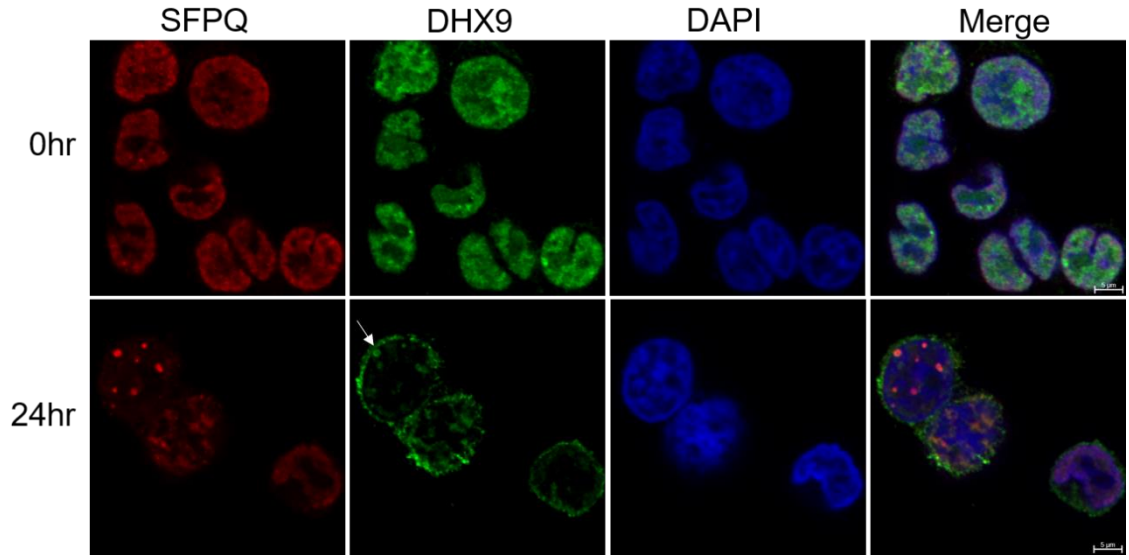

Figure 65: DHX9 staining: IF at 0 and 24 hours post-reactivation in TREx-BCBL1-RTA cells staining for SFPQ (red), DHX9 (green) and DAPI (blue). Arrows highlight co-localisation. Scale bars are 5 μm.

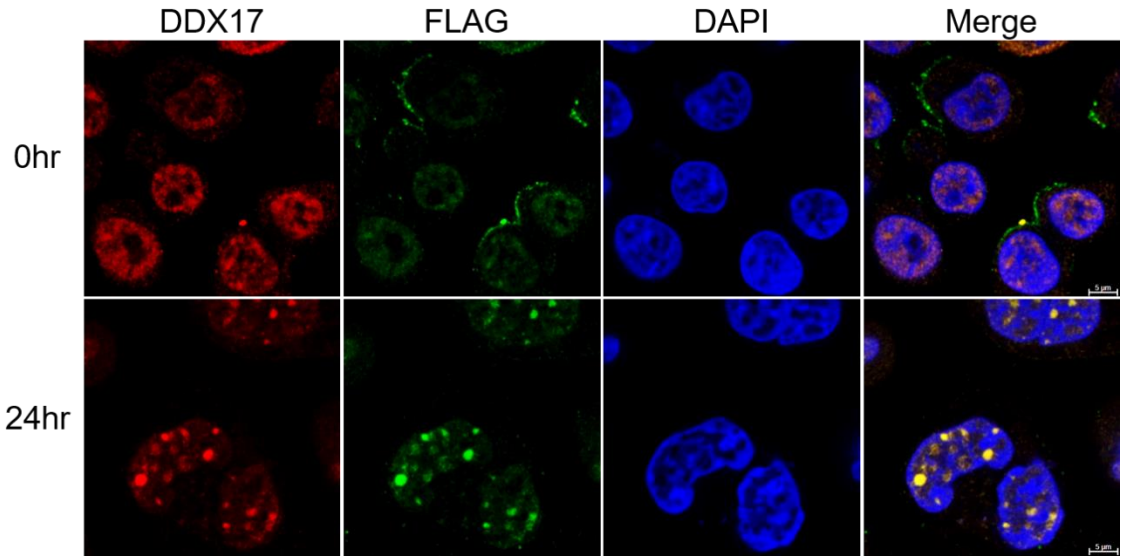

Figure 66: **DDX17 staining.** IF at 0 and 24 hours post-reactivation in FLAG-ORF11 TREx-BCBL1-RTA cells staining for DDX17 (red), FLAG (green) and DAPI (blue). Scale bars are 5 μm.

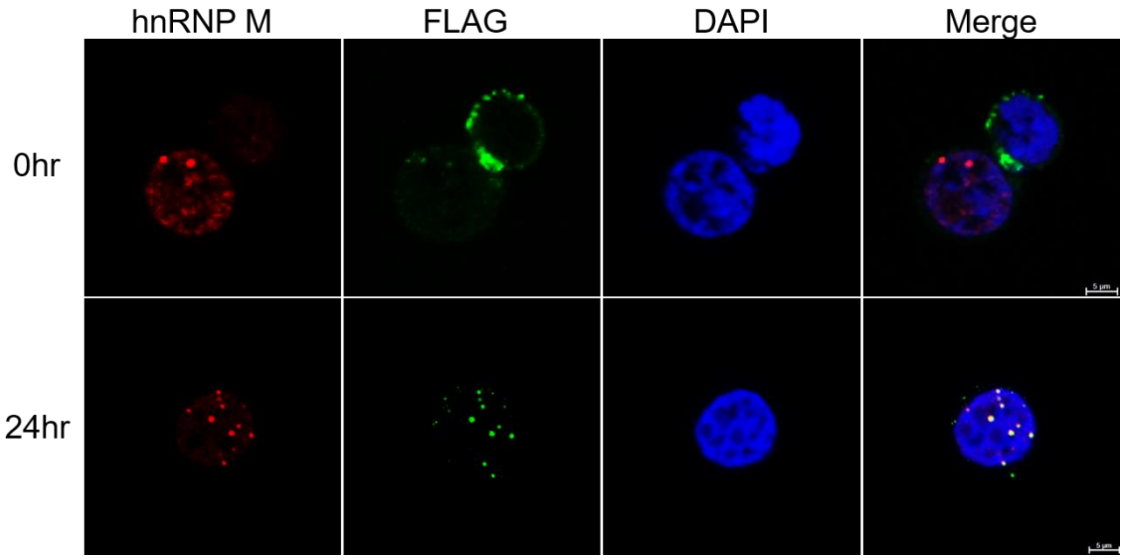

Figure 67: hnRNP M staining. IF at 0 and 24 hours post-reactivation in FLAG-ORF11 TREx-BCBL1-RTA cells staining for hnRNP M (red), FLAG (green) and DAPI (blue). Scale bars are 5 μm.

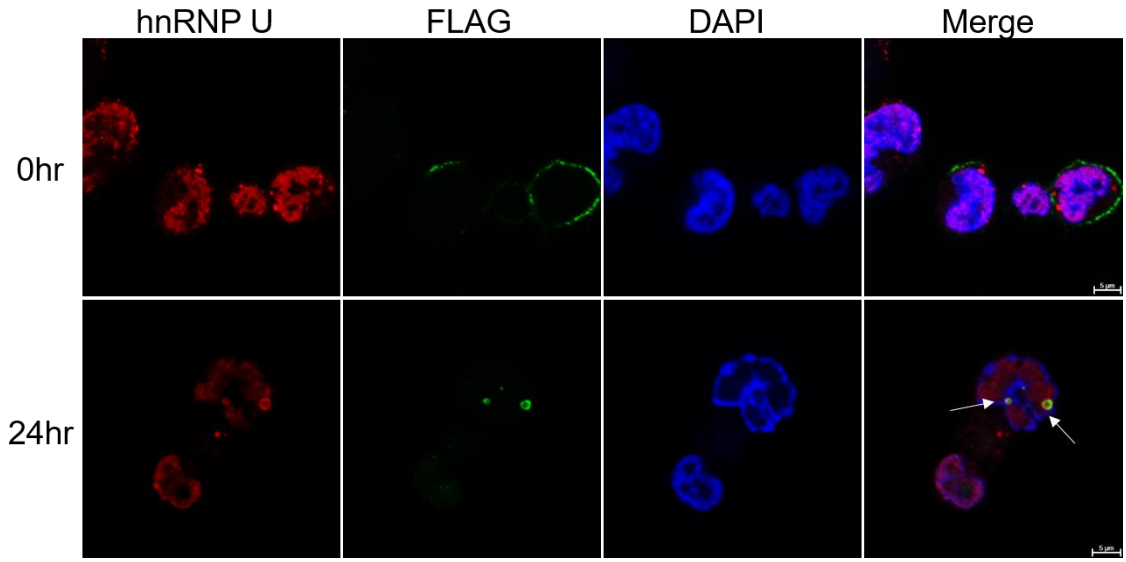

Figure 68: hnRNP U staining. IF at 0 and 24 hours post-reactivation in FLAG-ORF11 TREx-BCBL1-RTA cells staining for hnRNP U (red), FLAG (green) and DAPI (blue). Arrows highlight co-localisation. Scale bars are 5 μm.

5.18 SFPQ associates with viral RNAs

Having shown that SFPQ is re-localised into foci and is essential for KSHV lytic replication, experiments were next performed to try and determine a potential role for SFPQ in the KSHV lifecycle. Due to the known functions of SFPQ in various stages of RNA processing and the enrichment of many RNA splicing factors in the mass spectrometry, SFPQ RIPs were initially performed at 0 and 24 hours post-reactivation in TREx-BCBL1-RTA cells to try and identify viral RNAs which may associate with SFPQ. As a positive control for the SFPQ RIPs, qPCR was also performed on NEAT1. Unsurprisingly NEAT1 was heavily enriched in SFPQ RIPs over IgG controls, however, interestingly, this association decreased during lytic replication. Several viral transcripts also showed a strong association with SFPQ, namely ORF57, K8, ORF4, ORF45 and ORF59, in contrast no association was observed for ORF65, PAN and K8.1 (Figure 69A). Interestingly, a number of these SFPQ-associated viral transcripts undergo splicing which suggests that KSHV-mediated SFPQ-foci formation may be involved in splicing of its own transcripts.

To investigate whether SFPQ foci were involved in the splicing of viral transcripts, the splicing of the K8 transcript was further investigated due to its heavy enrichment in the

SFPQ RIPs. K8 undergoes complex splicing producing different spliced variants of K8 and K8/K8.1. Therefore a rudimentary K8 splicing assay was performed with cDNA isolated from scrambled and SFPQ depleted TREx-BCBL1-RTA cells. PCR was performed with a range of primers designed to amplify the different splicing variants of K8 and K8.1 and the levels of each spliced variant analysed via agarose gel electrophoresis [300]. SFPQ depletion results in an increase in the non-functional β splice variant at the expense of the main functional α variant (Figure 69B-C). This suggests that splicing of K8 is partly regulated by SFPQ, however, it must be noted that overall levels of K8 are reduced in SFPQ depleted cells.

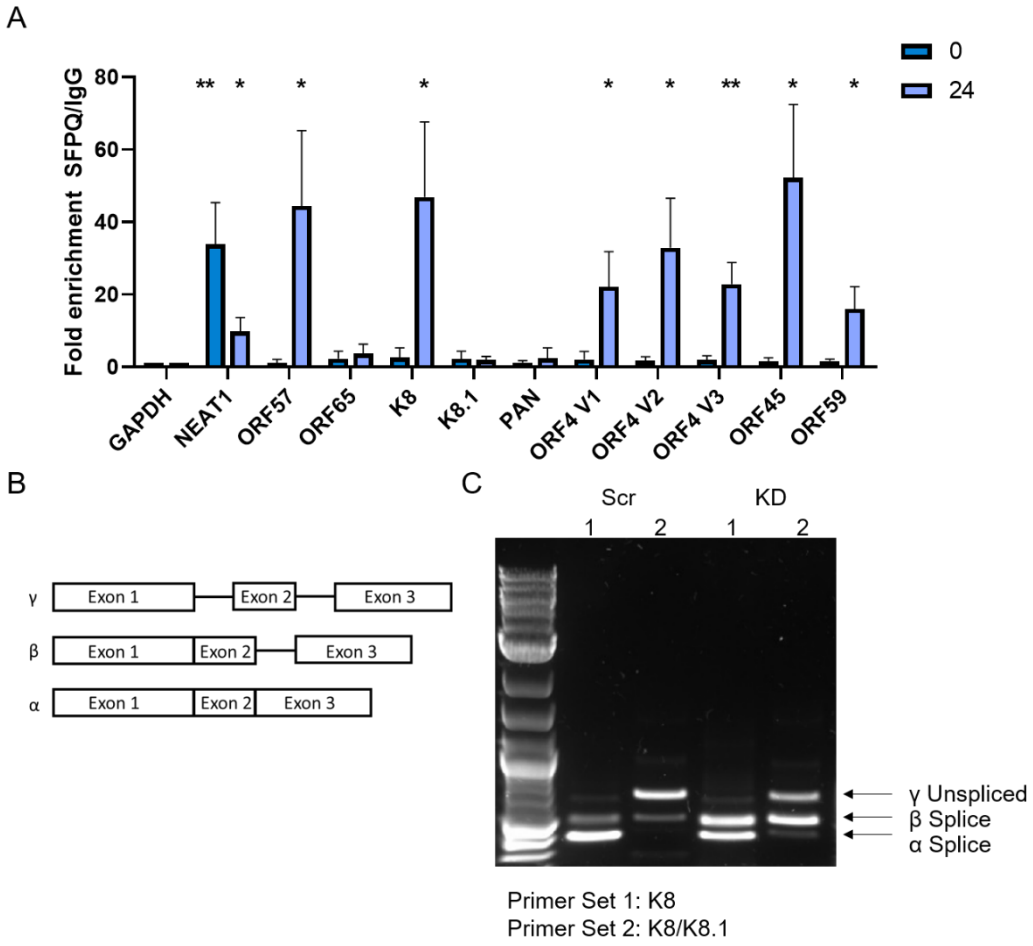

Figure 69: **Viral RNAs associated with SFPQ. (A).** qPCR analysis showing fold enrichment of genes in 0 and 24 hour induced TREx-BCBL1-RTA in SFPQ RIPs over IgG RIP. **(B). Schematic highlighting the 3 K8 variants amplified in the splicing assay**. γ is the unspliced full transcripts, β transcripts retains an intron while α is the full spliced functional variant. **(C). K8 splicing is disrupted by SFPQ KD.** Agarose gel of PCR products from 24 hour reactivated scrambled and SFPQ knockdown TREx-BCBL1-RTA cells amplified using different K8/K8.1 primers. N=3. Error bars represent SD. *P < 0.05, **P < 0.01 and ***P < 0.001 (unpaired Student's t-test).

5.19 SFPQ and cellular circRNAs

SFPQ was initially investigated due to its potential role in modulating circRNA biogenesis [198]. In particular, SFPQ has been implicated to regulate the production of circRNAs with long flanking introns. Therefore circCDYL and circEYA1 were selected due to both possessing these long flanking introns. Interestingly, both of these circRNAs were shown to be upregulated in our previously obtained circArray (Chapter 3, Figure 15). To confirm these specific circRNAs with long flanking introns were upregulated during KSHV lytic replication, qPCR was performed on RNA isolated from latent and reactivated TREx-BCBL1-RTA cells using divergent primers targeting the unique backsplicing site within each circRNA, as well as primers which amplify the linear parental transcripts. Both circRNAs were upregulated during KSHV lytic replication and this upregulation was specific to the circular form, not the linear parental transcripts (Figure 70).

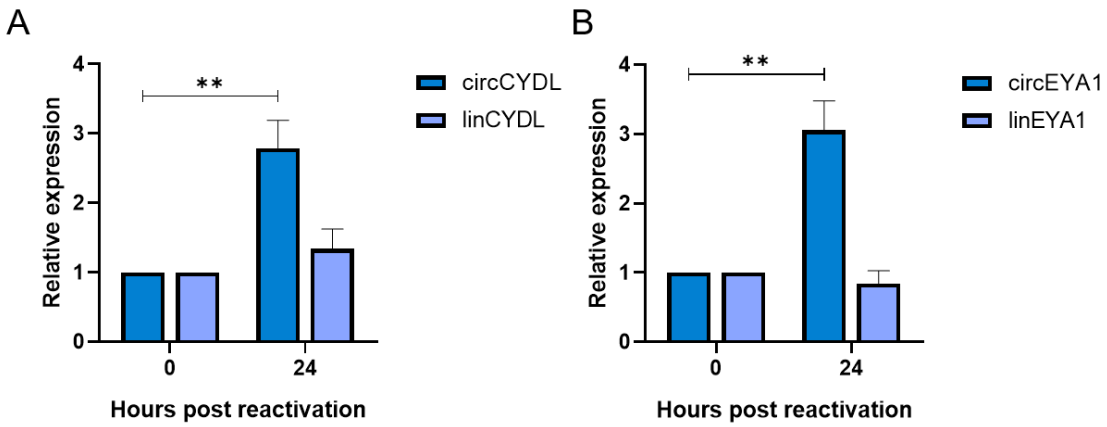

Figure 70: circCDYL and circEYA1 are specifically upregulated. qPCR analysis of the linear and circular forms of CYDL **(A)** and EYA1 **(B)** at 0 and 24 hours post lytic induction. . N=3. Fold change to 0 hour samples. Error bars represent SD. *P < 0.05, **P < 0.01 and ***P < 0.001 (unpaired Student's t-test).

To next assess whether SFPQ was essential in the KSHV-mediated upregulation of circRNAs with long flanking introns, the levels of circCDYL and circEYA1 were compared in scrambled control versus SFPQ depleted TREx-BCBL1-RTA cells. qPCR using divergent primers showed there was a significant reduction in both circRNA levels upon SFPQ depletion at 24 hours post-reactivation (Figure 71A-B).

Finally it was investigated whether SFPQ could associate with host cell circRNAs using SFPQ RIPs in latent or reactivated in TREx-BCBL1-RTA cells. Both circCDYL and circEYA1 were significantly enriched in the SFPQ RIP over IgG in lytic cells. In contrast, circSPECC1 was used as a negative control due to it not containing long flanking regions, failed to be precipitated by SFPQ (Figure 71C). This suggests that the KSHV-mediated upregulation of circCDYL and circEYA1 may be facilitated by SFPQ binding. At present was dictates SFPQ binding to certain circRNAs with long flanking introns is unknown and the exact mechanism of upregulation. Both circCDYL and circEYA1 also showed enhanced binding to SFPQ in latent cells too, however, this enrichment was lower than during KSHV lytic replication. Thus KSHV may be enhancing a regular cellular process during its lytic replication to upregulate pro-viral circRNAs.

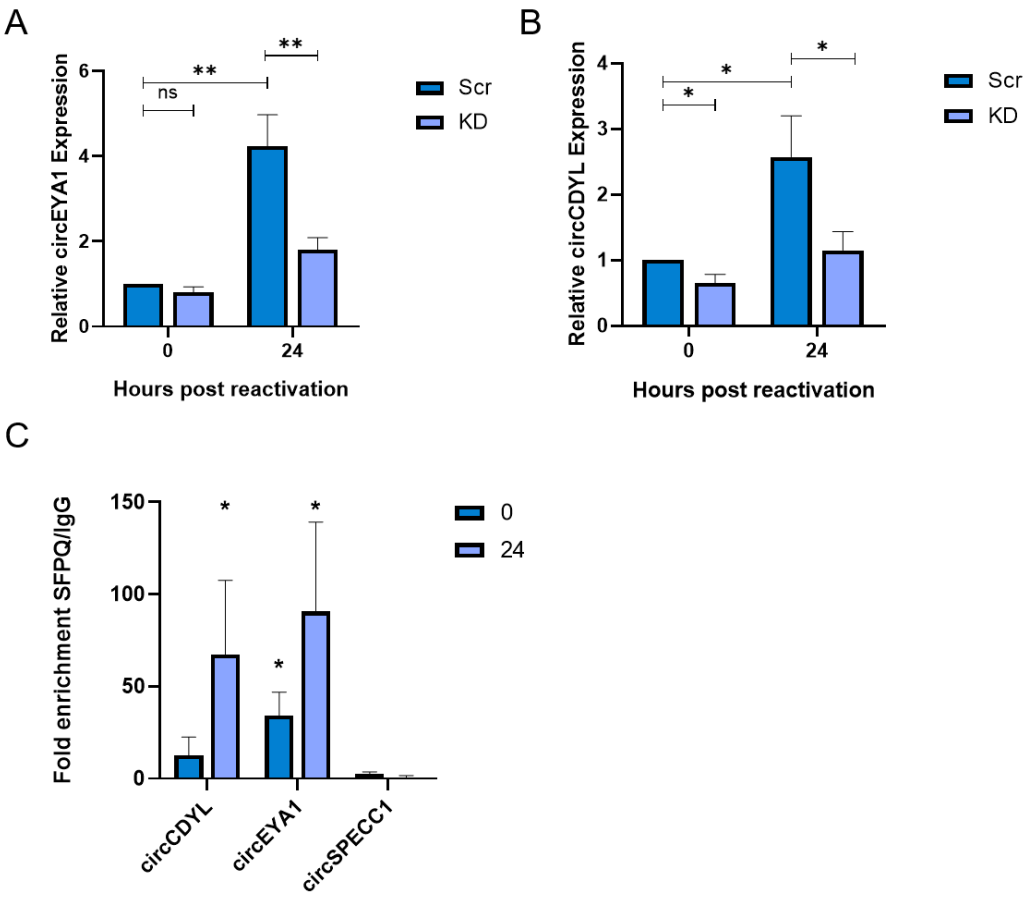

Figure 71: SFPQ dysregulates circRNAs. qPCR of circEYA1 **(A)** and circCDYL **(B)** in scrambled and SFPQ KD TREx-BCBL1-RTA cells at 0 and 24 hours. **(C).** qPCR analysis of fold enrichment of circCDYL, circEYA1 and circSPECC1 in SFPQ and IgG RIPs at 0 and 24 hours post-reactivation in TREx-BCBL1-RTA cells N=3. Error bars represent SD. *P < 0.05, **P < 0.01 and ***P < 0.001 (unpaired Student's t-test).

5.20 SFPQ is involved with KSHV-encoded circRNAs

In addition to KSHV dysregulating host cell circRNAs during infection, the virus encodes its own set of circRNAs (known as kcircRNAs), although some are yet to be confirmed due to their low expression. From a list of putatively-encoded kcircRNAs [187] [301] [186] [185], the 10 most highly expressed were selected for further investigation. Of these, circvIRF4 and circPAN were successfully amplified via qPCR using divergent primers targeting the unique backsplice sequence, with both being upregulated at 24 hours post-reactivation in TREx-BCBL1-RTA cells. circvIRF4 was also expressed at lower levels during latent infection, in agreement with previous reports (Figure 72A) [302].

To assess whether SFPQ was also involved in the biogenesis and upregulation of these kcircRNAs, experiments were repeated in scrambled versus SFPQ-depleted TREx-BCBL1-RTA cells. In addition, linear vIRF4 and PAN parental transcript levels were also assessed. Both kcircRNA levels were significantly affected by SFPQ depletion, it must also be noted that a reduction in the parental viral transcripts was also observed upon SFPQ depletion, however, the effect was less dramatic (Figure 72B-C). Following this observation it was determined whether either kcircRNA was found enriched in SFPQ RIPs at 0 and 24 hours post-reactivation. Results show that SFPQ could associate with both circvIRF4 and circPAN, in contrast, it failed to bind to the parental viral transcripts (Figure 72D). Thus, KSHV encodes its own circRNAs and SFPQ may be involved in their biogenesis or downstream processing.

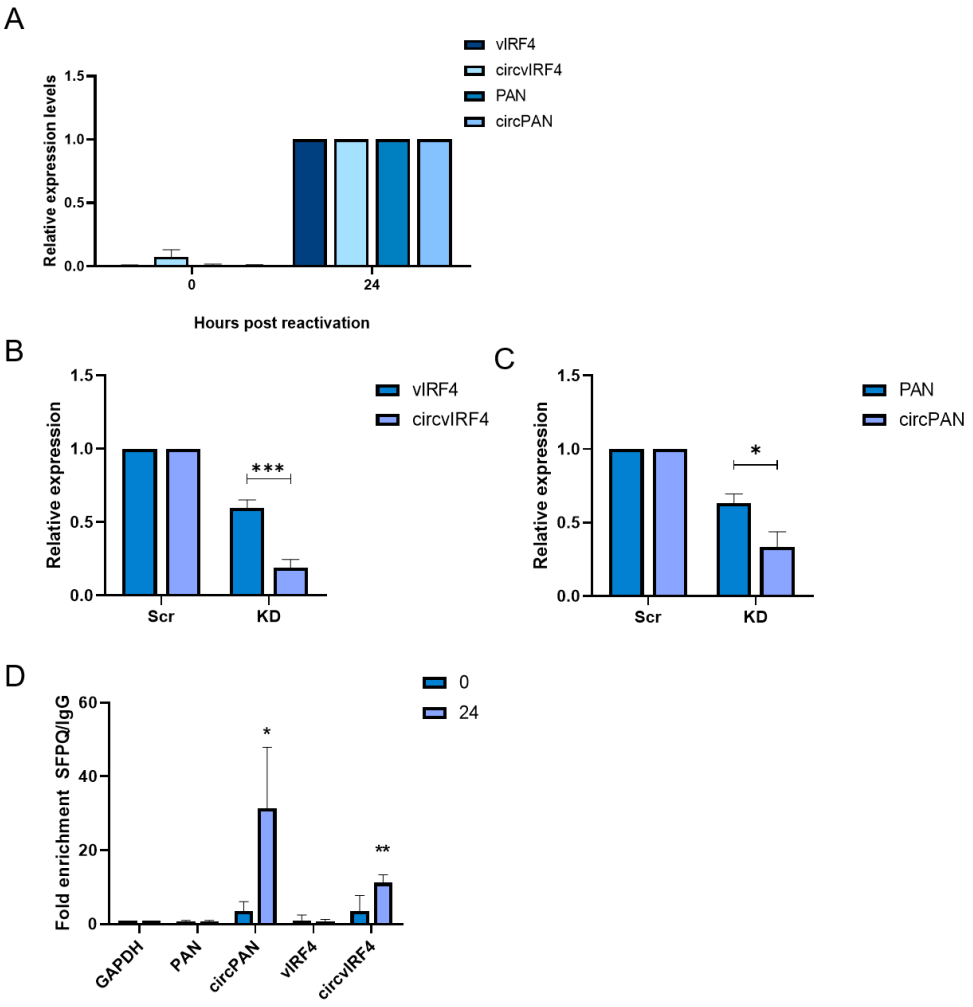

Figure 72: SFPQ regulates kcircRNA biogenesis. (A). qPCR analysis of vIRF4, PAN, circvIRF4 and circPAN at 0 and 24 hours. qPCR analysis of vIRF4 and circvIRF4 **(B)** and PAN and circPAN **(C)** at 24 hours in scr and SFPQ KD cell lines. **(D).** qPCR analysis of circPAN and circvIRF4 at 0 and 24 hours in SFPQ/IgG RIPs. N=3. Error bars represent SD. *P < 0.05, **P < 0.01 and ***P < 0.001 (unpaired Student's t-test).

5.21 Discussion

In this chapter we identify novel nuclear localised condensates that form through liquid-liquid phase separation during KSHV lytic replication. Importantly these condensates are essential for successful viral replication. Analysis suggest these condensates may be virus-modified paraspeckles due to the localisation of SFPQ and NEAT1, as well as other known core paraspeckle proteins, such as NONO and PSCP1. Notably, depletion of SFPQ, a core member of these condensates, results in inhibition of viral replication and infectious virion production. Similar results were also observed upon depletion of other member of the condensates, namely NEAT1 and NONO. Interestingly, SFPQ depletion

affected all stages of KSHV lytic gene expression, notably even early viral transcript were reduced by up to 90% at the RNA level, suggesting that SFPQ and the condensates may play a fundamental role in viral RNA processing. Furthermore, the condensate distribution around but not in the VRCs, suggests that the SFPQ does not regulate KSHV RNA at a transcriptional level, as most KSHV transcription occurs in the VRCs.

Following this, SFPQ RIPs identified multiple viral RNA transcripts bound to SFPQ during lytic replication, including K8, ORF57, ORF45 and ORF59. Interestingly, many of these viral transcripts are known to undergo complex splicing events and further modifications such as m6A methylation [303], which may implicate these condensates as novel sites for viral RNA processing. Proteomic analysis performed by Elena Harrington in the Whitehouse laboratory, identified multiple RNA helicases, hnRNPs and members of the m6A methylation machinery enriched in SFPQ immunoprecipitations, further implicating these condensates are novel hubs of viral RNA processing. It is enticing to speculate that viral transcripts may shuttle between these foci and VRCs, bearing in mind these modified paraspeckles are adjacent to VRCs. Analysis through SFPQ RIPs and SFPQ depleted cell lines, also identified that these foci may be involved in both cellular and KSHV-encoded circRNA biogenesis or processing, although the implication of these findings requires further analysis.

Canonical paraspeckles are classified as sub-nuclear organelles with roles in RNA storage and regulation of gene expression, and while these foci observed during lytic replication contain many canonical paraspeckle components, it must be noted they differ in several aspects. Namely, although FISH and immunofluorescence determined that the viral-induced foci contain the PSPs SFPQ, NONO and PSPC1, as well as the core ncRNA scaffold: NEAT1, proteomic analysis showed that several core essential paraspeckle proteins are missing from the viral induced foci. Previous scientific literature has observed that removal of these missing components should prevent paraspeckle formation [211][215], thus KSHV must be capable of subverting the normal paraspeckle

biogenesis mechanisms to form modified condensates. These condensates were also observed to be up to 10 times larger than canonical paraspeckles, and surprisingly, qPCR showed no corresponding increase in NEAT1 levels during lytic replication. The established dogma is paraspeckle biogenesis occurs co-transcriptionally, with PSPs binding to NEAT1 stabilising the transcript as RNA pol II is actively transcribing, as such, NEAT1 levels correlate with paraspeckle size. This is further evidence suggesting that the virally modified paraspeckles bypass the usually biogenesis mechanisms and are not reliant on active transcription of NEAT1 and the resulting stages of biogenesis.

Although we have not elucidated the alternative biogenesis mechanism for these virally-modified condensates, we implicate the KSHV protein ORF11 to play a key role. ORF11 is classified as an early KSHV encoded protein, expressed from around 8 hours in the lytic cascade, with currently no assigned function. We observed ORF11 co-localises and binds to SFPQ as early as 8 hours post-reactivation. Importantly no virally-modified condensates are observed in ORF11 CRISPR TREx-BCBL1-RTA cells suggesting ORF11 is key in their formation.

Overall, however, further research is needed to fully elucidate both the function and the biogenesis of these novel viral-induced condensates.

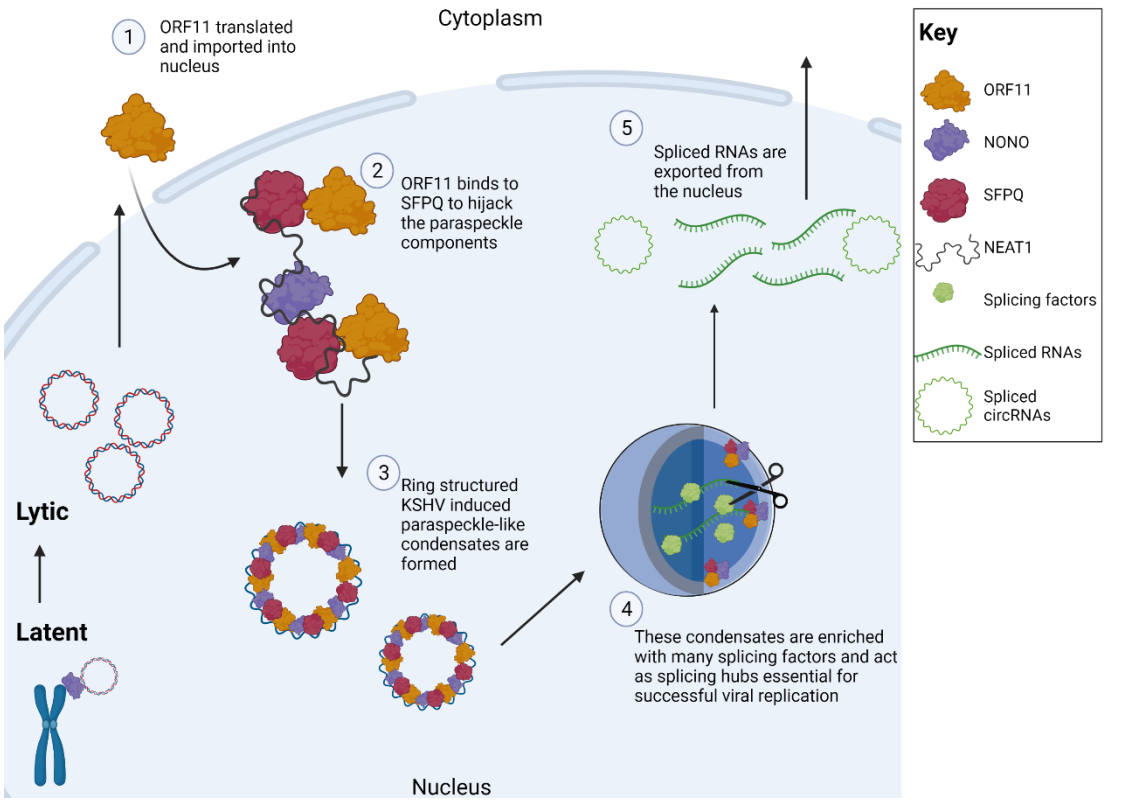

Figure 73: Proposed mechanism for the formation and function of KSHV induced paraspeckle-like condensates. During lytic replication, the viral protein ORF11 is translated and traffics into the nucleus (1). ORF11 interacts with paraspeckle components either directly or indirectly, including SFPQ, NONO and NEAT1 (2). Through these interactions ring structured viral condensates form comprising of ORF11, several core paraspeckle proteins (3). These structures also contain many splicing factors including many hnRNPs, allowing the structures to act as hubs for splicing both cellular and viral RNAs essential for viral replication.

Chapter 6

~

Discussion

6.0 Discussion

Historically, the majority of research has focused on coding RNAs and their protein products, with ncRNAs dismissed as junk RNA or simple by-products of the transcription process. However, with the advancement of global sequencing techniques, it has been established that over half the human transcriptome is non-coding, with multiple different functional roles. ncRNAs comprise many different species including lncRNAs, miRNAs and snRNAs. However, even in the world of ncRNAs, circRNAs have been under-researched, with strong evidence for their functional existence only emerging in 2012 [143]. This is in part due to the prevalence of poly (A) selection in sequencing leading to a bias in their selection, however, since the adaptation of NGS approaches to preferentially enrich circRNAs, they have emerged as more abundant, conserved and stable than previously thought.

Since the first functional analysis of circRNAs as a miRNA sponge, the idea that circRNAs could regulate gene expression through ncRNA regulatory networks has emerged. This would add another layer of complexity and fine-tuning within a cellular environment. This is supported by numerous papers showing circRNAs are dysregulated in disease states, including viral infection [119] [165].

KSHV, as a large double stranded DNA genome-containing herpesvirus, is known to encode its own miRNAs; furthermore, it dysregulates cellular miRNAs to produce a more conducive environment for viral replication [67] [304]. Interestingly, recent work has also identified KSHV can encode for circRNAs, although no functional role of these so called kcircRNAs has yet emerged [186]. We therefore hypothesised that KSHV could also potentially dysregulate cellular circRNAs to form an ncRNA regulatory network to aid its lytic replication.

Another area of research that is currently evolving rapidly is that of condensates; a class of membraneless organelles that are formed through liquid: liquid phase interactions

[305]. It is thought that these condensates can aid in multiple biological activities through either sequestration or enhancement of biochemical reactions by increased concentrations of enzymes and substrates. As circRNA biogenesis is often considered to compete with canonical splicing and formation of the linear transcript, we hypothesised that KSHV could drive formation of novel condensates, through the concentration of factors that promote backsplicing leading to the dysregulation of cellular circRNA biogenesis during lytic replication.

6.1 KSHV dysregulates circRNAs

Since their emergence in the last 10 years as potential regulators of the cellular transcriptome, circRNAs have been implicated in a range of diseases including cardiovascular disease, Alzheimer's and viral infection [148] [306]. KSHV has previously been shown to be prolific at utilising ncRNAs to aid its replication; including encoding its own miRNA and circRNAs [185] [186] and dysregulating cellular miRNAs during both latent and lytic replication. Consequently we investigated whether cellular circRNAs were also dysregulated during viral replication.

Using both the circArray and the literature, divergent primers were designed against the most upregulated, more abundant, or functionally elucidated circRNAs, including circHIPK3, circPVT1, circITCH and circSPECC1. Analysis showed that both circHIPK3 and circSPECC1 were significantly upregulated during lytic replication whereas circPVT1 and circITCH levels showed no change. Importantly, levels of the parental linear genes *SPECC1* and *HIPK3* did not increase, suggesting that circRNA upregulation is not just a by-product formed from an increase in the linear form. Instead these results suggest that the backsplicing process is specifically upregulated with ratio of linear to circular transcripts changing during KSHV lytic replication. This is not an unusual phenomenon, with many functional circRNAs levels not correlating with their parental transcript due to their different functions and regulators [307].

circRNA function is often tied to their localisation, with intronic circRNAs retained in the nucleus and having roles in transcriptional regulation and nuclear protein binding. While exonic circRNAs are proposed to have roles in protein and miRNA binding in the cytoplasm. circSPECC1 has no known function, and has only been investigated as part of a screening while circHIPK3 has proposed to have a role in miRNA sponging [263]. For the initial circRNA function screen, subcellular fractionation was performed, showing circHIPK3 did indeed localise to the cytoplasm, whilst circSPECC1 was predominantly retained in the nucleus. Due to our interest in the virus potentially dysregulating an ncRNA network, circHIPK3 was selected as the most promising candidate, with its cytoplasmic localisation confirmed using RNA FISH.

To identify potential miRNAs downregulated by circHIPK3, we utilised a miR-Seq dataset in the Whitehouse laboratory (produced by Dr Becky Foster) which characterised the most dysregulated miRNAs during lytic replication. Cross-referencing this dataset with potential circHIPK3 binding sites led the selection of miR-29b and miR-30c as candidate sponged miRNAs. qPCR confirmed they were downregulated upon KSHV lytic replication, interestingly qPCR analysis of the precursor primary and pre miRNAs, showed no change, suggesting their downregulation was at the mature stage, implicating a circRNA: miRNA regulatory axis.

The role of both miR-29b and miR-30c during KSHV lytic replication were then investigated through transfection of miRNA mimics into reactivated TREx-BCBL1-RTA cells. Results showed that over-expression of either miRNA had a detrimental role on viral lytic replication and infectious virion production.

To confirm a potential circRNA: miRNA regulatory axis Ago2 RIPs were carried out, showing that circHIPK3 was enriched in Ago2 RIPs over an IgG negative control. Additionally RIPs performed in the presence of miRNA mimics increased circHIPK3 enrichment, whereas an antagomiR reduced association. Although these Ago2 RIPs

strongly indicate circHIPK3 was sponging miR-29b/miR-30c, biotinylated mimics coupled with streptavidin RIPs confirmed a direct interaction.

circHIPK3 as sponge of anti-viral miRNAs, should be pro-viral. To investigate this stable shRNA KD cell lines were designed against the unique backsplice sequence. The success of these KDs was confirmed via qPCR, which also confirmed *HIPK3* was not targeted in the KD, and as such any effects can be attributed to circHIPK3 alone. The role of circHIPK3 KD on the virus was analysed through a range of assays including viral load, viral reinfection and late viral protein production, with all 3 significantly reduced in circHIPK3 KD cell lines. This confirms the essential nature of circHIPK3 to KSHV lytic replication.

If downregulation of miR-29b and miR-30c is essential for efficient virus replication by circHIPK3 sponging, it would follow that depletion of circHIPK3 would be detrimental to the virus. Not surprisingly, a stable shRNA KD designed against the unique backsplice sequence confirmed the essential nature of circHIPK3 to KSHV lytic replication with miR-30c levels increasing. This highlights the restoration of miR-30c is specific to circHIPK3 and is not simply due to inhibition of viral replication.

The exact mechanism of miR-30c degradation via circHIPK3 is currently not elucidated, however, one potential mechanism is Target RNA-Directed MicroRNA Degradation (TDMD). Here when a miRNA binds to its target RNA, it leads to degradation of the miRNA instead of the RNA, however, the molecular basis of the mechanism is not fully understood [277]. It is hypothesised that TDMD requires high complementarity between the RNA and the target miRNA, with a high 3′ complementarity and a central mismatch bulge common features [116]. The mismatched bulged prevents the miRNA: RNA duplex from Ago2 binding and as such cleavage of the RNA cannot occur, instead the miRNA 3′ end is presented for degradation [116]. It is believed there are many different co-factors involved in this process that may be miRNA or cell type specific [134]. Interestingly 2 miRNAs that have been shown to undergo TDMD are miR-29b and miR-30c, suggesting

they are compatible with this method of regulation and can be targeted by even lowly expressed RNAs [277] [276]. However, the classical TDMD characteristics are not seen within the circHIPK3:miR-30c binding site, therefore it may be through an alternative mechanism, such as binding leading to recruitment of degradation factors, re-localisation of the miRNA or most similar to TDMD, binding leads to conformational changes that could expose the miRNA to degradation. It is possible that due to the intrinsic circular nature of circHIPK3 it cannot be loaded onto Ago2 leading to TDMD without the need for the bulge loop seen with canonical mRNA TDMD. Notably, other mechanisms for viral mediated degradation of miRNAs are emerging, with the poxviruses AMEV and VACV utilising their poly(A) polymerase VP55 to add tails to miRNAs leading to their degradation [308].

Emerging evidence also shows that herpesviruses can degrade inhibitory host cell miRNAs by unique mechanisms [279] [309]. For instance HVS, a close relative of KSHV, encodes for small ncRNAs named HSURs, these bind to miR-27 through base pairing, leading to its destabilisation and eventually upregulation of its target *FOXO1* [278]. MCMV also targets miR-27 for degradation through a complementary binding of its transcript the ncRNA m169. Expression of m169 in uninfected cells is sufficient to induce miR-27 degradation [309] [279]. HCMV is also capable of TDMD, with the ncRNAs UL144-145 binding and inducing degradation of miR-17 and miR-20a [134] [280].

We next investigated how KSHV may upregulate circHIPK3 levels, to manipulate the ncRNA axis. Results showed that KSHV ORF57 protein alone could enhance circHIPK3 levels and postulated ORF57 could be binding to *HIPK3* and promoting backsplicing leading to increased levels of the circRNA. Utilising a range of ORF57 mutants we examined whether RNA binding and ORF57 dimerisation were significant. Many circRNA biogenesis factors such as QKI and Fus positively regulate circRNA biogenesis through self-dimerisation [155] [227]. It is thought that monomers of the proteins can bind to the distal splice sites and as they dimerise, they bring the splice sites in close proximity

allowing backsplicing to occur. As ORF57 is a known dimerisation protein [275] we hypothesised that this mechanism could be relevant in KSHV, unfortunately the dimerisation mutant was too unstable to fully characterise this mechanism. However, RNA binding mutants did lead to a reduction in circHIPK3 levels in HEK-293Ts, however, they did not completely ablate circHIPK3 production, indicating that other unknown factors may be involved. One discrepancy between the transfections into HEK-293Ts and KSHV lytic replication, is transfection of ORF57 led to decreases in *HIPK3* levels that are not observed at the equivalent time point in TREx-BCBL1-RTA cells. Other viral factors or differences in cell types could lead to this variation; there is potential that *HIPK3* levels are maintained in TREx-BCBL1-RTA cells through an unknown mechanism to provide a pool of transcript for backsplicing to occur from, although this requires more investigation. Interestingly this small downregulation did not occur with transfection of the ORF57 RNA binding mutants. Alternatively ORF57 could promote circRNA stability leading to increases in circHIPK3 levels, with ORF57 known to have a key role in promoting PAN stability [92] [310], although further research is needed to fully elucidate the role of ORF57 in circHIPK3 upregulation.

6.2 circHIPK3:miR-30c:DLL4 form a novel ncRNA network that is essential for viral replication

To complete this novel ncRNA regulatory network, DLL4 was initially identified as a target of miR-30c through miRNA prediction software and cross referenced with a KSHV replication RNA-Seq time course dataset previously generated in the Whitehouse laboratory. DLL4 was selected due to its high score across a range of software and its upregulation within the RNA-Seq dataset. Further validation confirmed DLL4 as a direct target of miR-30c. Confirmation that the circHIPK3:miR-30c axis was also involved in regulating DLL4 expression, was confirmed by a range of assays including circHIPK3 depletion, use of miRNA antagomiRs and Ago2 RIPs.

DLL4 has previously been shown to be upregulated by the KSHV encoded proteins vGPCR and vIL-6, with vGPCR-mediated upregulation found to be ERK dependent [288], additionally the latent proteins LANA and vFLIP have also been suggested to upregulate DLL4 [283] [288]. However, these studies focus on DLL4 upregulation during latency in endothelial cells focusing on its role as a notch ligand. DLL4 is upregulated by vGPCR and vIL-6 to activate Notch2 and Notch4, promoting cell proliferation, inhibiting apoptosis and promoting angiogenesis and the sprouting of blood vessels [288]. Whereas this study shows for the first time that DLL4 expression is upregulated and essential in KSHV lytic B cells.

Due to the focus of DLL4 and its role in Notch signalling in endothelial cells, we investigated its other known role in regulating the cell cycle in lytic replication [283]. Although cell cycle dysregulation during KSHV infection has been investigated it predominantly focuses on the dysregulation via vCyclin during latency [311] [312], with conflicting results about dysregulation during lytic replication, with either entry into G1 or S phase required [311] [313] .

Results showed that the number of cells in the G1 phase is maintained, with a downregulation of the number of cells in G2 phase at 24 and 48 hours post-reactivation. Notably, although the role of cell cycle in lytic KSHV replication is unknown, other herpesviruses require G1 phase, with both HSV-1 and HCMV requiring G1 entry for successful replication [292] [291]. EBV also promotes cell entry into G1 phase during lytic replication [290] [289]. Interestingly these studies found G1 exit blocked during lytic replication, however, several S phase cellular factors were still promoted for DNA replication, suggesting these conditions allow successful viral DNA replication using the cellular proteins but through blocking of S phase entry, reduced competition from cellular DNA synthesis.

To investigate whether cell cycle dysregulation is beneficial for KSHV lytic replication, cell cycle inhibitors were utilised. G2 arrest significantly reduced DNA viral load whereas

G1/S arrest had no effect, supporting hypothesis that maintenance of cells in G1 phase and loss of cells in G2 phase is essential for KSHV DNA replication.

Together, chapters 3 and 4 identify a novel dysregulated ncRNA network that is essential for efficient KSHV lytic replication.

6.3 KSHV induces novel SFPQ condensates during lytic replication

The previous chapters focused on a single circRNA-associated regulatory network, however, as highlighted in Figure 15, chapter 3, the preliminary circArray data suggests multiple circRNAs may be dysregulated during lytic replication. It is currently unknown how cellular circRNA biogenesis is modulated during lytic replication, initial studies focused on the KSHV ORF57 protein however, it is likely cellular factors are also utilised by the virus to manipulate circRNA biogenesis pathways. Recent literature has implicated the RBP SFPQ as a novel factor in circRNA biogenesis, therefore it was chosen for further investigation.

Initial immunofluorescence studies demonstrated that SFPQ was redistributed from a diffuse nuclear staining to distinct large foci within the nucleus during lytic replication. These foci generally numbered between 3-10 and were around 2 μm in diameter.

Additional immunofluorescence across a time course determined SFPQ foci formation occurred as early as 8 hours post-lytic reactivation, with the majority of the formation occurring between 16 and 24 hours post-reactivation. We hypothesised that these novel SFPQ-containing foci are viral-induced condensates, due to their membraneless nature and maturation process. FRAP was used to confirm the foci were condensates that matured through LLPS. At 16 hours post-reactivation, the foci partly recovered after photobleaching, thus at this point the foci must be dynamic, with any interactions still transient. Whereas at 24 hours post-reactivation, there was little to no recovery after photobleaching, implying they have lost fluidity and have formed fully mature condensates [305]. SFPQ has been previously identified as a core-paraspeckle protein.

However, intriguingly these viral-induced foci are much larger than normal paraspeckles, which suggests they may be virus-modified paraspeckle condensates.

Paraspeckles have previously been characterised as having enhanced formation during times of cellular stress [314], therefore key experiments were performed to determine whether these condensates were just a stress response to the initialisation of lytic replication. Serum starving failed to induce the formation of SFPQ foci, in fact, serum starving seemed to inhibit their formation in lytic cells. This therefore suggests that the SFPQ foci are likely induced directly by KSHV rather than a response to cellular stress.

The lncRNA NEAT1 is the driving force and an essential component for paraspeckle formation, allowing the recruitment of core paraspeckle proteins including SFPQ, NONO and PSCP1 [205]. Surprisingly, qPCR analysis showed NEAT1 levels remained unchanged through lytic replication, this was a particularly intriguing results as canonical paraspeckle formation begins co-transcriptionally, with SFPQ and NONO binding and stabilising NEAT1 as it is being actively transcribed. As such, NEAT1 RNA levels directly correlate with paraspeckle size and abundance. Therefore with the larger than normal SFPQ foci formed during infection we would have expected an increase in NEAT1 levels during lytic replication. This lack of NEAT1 upregulation, reinforces these SFPQ foci are virally modulated and therefore their biogenesis and contents may differ from that of canonical paraspeckles. Following this surprising result RNA FISH was performed to determine if NEAT1 localised in the viral-induced condensates. Results confirmed NEAT1 did co-localise to the foci, suggesting that they are likely to be virally-modified paraspeckles, relying on NEAT1 as the molecular scaffold. However, research is still required to characterise how these modified paraspeckles vary in their content and biogenesis, compared to canonical paraspeckles. Preliminary proteomic analysis coupled with SFPQ pulldowns have suggested that there is a distinct protein composition in these virus-modified paraspeckles and future work will be needed on the functional role of these non-canonical components.

Another question to address is how does KSHV drive the formation of these modified paraspeckle condensates. Findings reported in chapter 5 implicate a previously uncharacterised KSHV-encoded protein, ORF11. ORF11 was initially identified as a binding partner of SFPQ and immunofluorescence studies confirm that ORF11 co-localises with SFPQ from as early as 8 hours post-lytic reactivation, as the foci form. This may suggest ORF11 has a role in the formation of the viral-induced foci or is required for correct functionality. Further support to this theory is the ORF11 CRISPR studies, showing that SFPQ foci fail to form when ORF11 is lacking in reactivated cells. Interestingly, immunofluorescence studies subsequently found that ORF11 and NEAT1 also localise to this outer ring structure of the modified paraspeckles. This leads to the intriguing possibility that the centre of these foci contain cellular and viral RNAs which are being processing during lytic replication. It must be noted that the viral-modified structure is distinct to normal paraspeckles, where both SFPQ and NEAT localise on the other ring, whereas in canonical paraspeckles NEAT1 forms the outer shell and SFPQ is localised in the inner core [206]. This once again implies KSHV is hijacking the paraspeckle components to form distinct virally-modified condensates. Utilising Zen software we also determined these virally-modified condensates are far larger than canonical paraspeckles, reaching sizes over 20 times larger than average paraspeckles. Future work will focus on characterising the virus-modified foci in more detail. AN ideal approach would be to utilise hybridisation proximity (HyPro) labelling technologies coupled with Mass Spectrometry and RNA-Seq [315]. This approach has been used with NEAT1 to characterise canonical paraspeckle components and could be ideal to dissect the differences between viral and canonical paraspeckles.

To investigate the role of SFPQ and implicate the role of the modified foci during KSHV lytic replication, stable SFPQ knockdown cell lines were created and while depletion of SFPQ had little effect on cellular growth, it is highly detrimental to KSHV lytic replication. SFPQ is essential to KSHV lytic replication early in the lytic cascade, as shown through

the reduction of ORF57 protein. This suggest that SFPQ plays a role in RNA processing of early KSHV transcripts, which then leads to reductions in late genes due to interruption of the lytic cascade. It is enticing to suggest that processing of viral transcripts occurs in the condensates due to adjacent localisation to KSHV VRCs allowing easy shuttling and transport of viral RNAs between compartments.

Interestingly, depletion of additional paraspeckle components, namely NONO and NEAT1 also lead to a significant drop in late viral protein expression, however, early gene expression was less disrupted. This may suggest that either SFPQ function during KSHV lytic replication has less redundancy than other paraspeckle components or SFPQ may play a key direct role in viral RNA processing within the foci, while NEAT1 and NONO are just structural components of the condensates.

Finally, attempts have been made to investigate the functionality of these novel SFPQ paraspeckle condensates in virus replication. SFPQ RIPs showed association with many viral RNAs, such as those coding for ORF57, K8 and ORF4. Surprisingly, no association was observed between SFPQ and the highly expressed KSHV PAN RNA. As PAN is highly polyadenylated, this may discount SFPQ mediating polyadenylation of viral transcripts, as has been observed with influenza virus [243]. Notably, ORF57, K8 and ORF4 are all heavily spliced transcripts, with all 3 showing the highest association with SFPQ during lytic replication [303]. As such, SFPQ condensates may be mediating correct processing and splicing of viral transcripts. This was further evidenced through a splicing assay on the viral transcript K8, which showed that the ratio of α, β and γ transcripts was modulated in SFPQ depleted cells. Again, this hypothesis is supported by the proteomic and immunofluorescence analysis which found many RNA helicases and members of the hnRNP splicing protein family associated with SFPQ and localised to the foci. Moreover, several canonical PS proteins were not enriched, suggesting KSHV may modulate paraspeckles, forming hubs that preferentially splice viral transcripts.

In addition, our preliminary findings support a role of SFPQ in the regulation of circRNAs, including the enhanced production of circCDYL and circEYA1. As yet the exact mechanism for SFPQ mediated regulation has not been elucidated, however, several properties of SFPQ and paraspeckles could be contribute to this regulation. Namely, there are several RBPs including FUS localised to paraspeckles that have previously been shown to regulate circRNA biogenesis and RNAs with long flanking introns which are prone to form circRNAs are often enriched in paraspeckles. As such, the LLPS may allow the high concentration of both the biogenesis factors and the putative circRNAs for efficient backsplicing to occur and out-competing the linear splicing machinery. This conducive environment for circRNA formation, may also lead to the promotion of kcircRNAs, with depletion of SFPQ specifically downregulating them.

One of the biggest unanswered questions surrounding these SFPQ condensates however, is how are viral RNAs selected and localised to the foci, and if splicing occurs there, what aids in the regulation of splicing. SFPQ RIP analysis identified many viral spliced RNAs as being present within the condensates, however other viral RNAs including ORF59, which is not heavily spliced were also enriched. Notably, a commonality between these viral RNAs is that, ORF59, along with ORF45, ORF57 and K8 are heavily m^6A methylated early in the lytic cascade [316]. Interestingly, the proteomic analysis identified the m^6A reader, YTHDC1 heavily enriched at 24 hours, which is distinct from canonical paraspeckles. Moreover, its known co-factors hnRNP C and hnRNP A2B1 were also enriched at 24 hours within the SFPQ proteomic dataset. YTHDC1 is involved in the regulation of splicing through recognition and binding to m6A sites, where once bound it inhibits binding of the splicing repressor SRSF10, and promotes binding of the splicing factor SRSF3. This aids in stabilising of U1 snRNP binding to proximal splice sites leading to exon inclusion [317] [318]. Therefore one can theorise that KSHV may be utilising YTHDC1 to aid in correct splicing of its transcripts. Interestingly SRSF10 is considered an important canonical paraspeckle protein,

however, it was below the threshold in the SFPQ proteomic dataset, suggesting that in KSHV lytic replication, the condensates lack the splicing repressor SRSF10 but are enriched for SRSF3 [319].

In summary, we have identified that SFPQ forms novel condensates during KSHV lytic replication, with this formation potentially driven by the viral protein ORF11. These condensates are essential for successful lytic replication, with a putative function in RNA splicing of both circRNAs and viral transcripts.

6.4 Future perspectives

This research has identified a novel ncRNA regulatory network centred on circHIPK3, however, this was just one circRNA identified as dysregulated during KSHV lytic replication. The circArray identified many more significantly dysregulated circRNAs, future research should attempt to functionally elucidate these circRNAs and identify if they play a role in viral replication. Of note, evidence is emerging of circRNA levels correlating with specific disease states, and thus they have potential to be biomarkers, both for identification of disease and prognosis [320] [321]. Given that the circArray identified numerous dysregulated circRNAs, there is the potential for several of these circRNAs to be utilised as novel biomarkers for KSHV infection due to their intrinsic stability and cell-type specificity.

Additionally how these circRNAs are dysregulated is an area that holds much promise, with several avenues to dissect. Assays should focus on determining whether the dysregulation observed is at a transcriptional and promoter level or whether backsplicing itself is affected. KSHV has already been observed to dysregulate numerous RBPs that have already been identified to affect circRNA biogenesis, for instance FUS and QKI [185], with the additional role for the KSHV-encoded RBP ORF57. Interestingly methylation and RNA modifications have also been implicated in the regulation of circRNA biogenesis. For instance in glioblastoma, m^6A peaks were both gained and lost in tumour samples compared to healthy tissues, with m^6A levels positively correlating

with circRNA abundance [322]. Moreover, the Whitehouse lab has previously established key members of the methylation machinery are hijacked by the virus to aid its replication through a variety of mechanisms including increased RNA stability [323]. To corroborate this theory, dysregulated circRNAs could be screened for known m^6A consensus sequences such as DRACH motifs. Advances in CRISPR-Cas9 have allowed removal or inhibition of individual methylation sites for study and although difficulties remain in targeting circRNAs specifically via CRISPR-Cas9 there have been recent breakthroughs in preferential targeting [324].

Furthermore, several key questions remain unanswered about the role of SFPQ and bimolecular condensates during KSHV lytic replication. For instance, canonical paraspeckle formation is tied to NEAT1 transcription, however, no change in NEAT1 levels is observed during KSHV lytic replication. This combined with both the loss of core PS proteins and the gain of non-PS proteins implies these novel condensates are forming through a viral-induced unique mechanism. In particular, the role of ORF11 should undergo further studies to establish if it plays a core role in driving formation such as potentially acting as a scaffold for liquid-phase separation. *In vitro* assays could be utilised to determine whether ORF11 can directly bind and aggregate with SFPQ or other core PS proteins; mutational studies could also be employed to determine key domains for this formation.

Another key question about the condensates is their exact role in viral replication. Immunofluorescence identified SFPQ forming a ring-like structure during lytic replication, leading to the question of what is in the centre of the structures. Given that paraspeckles are known RNA processing hubs and their proximity to VRCs may allow shuttling of viral transcripts, it is likely the centre contains RNA. Initial data from splicing assays, RIPs and MS suggests a role in splicing and RNA processing, however, future experiments should utilise a large-scale screening assay to allow identification of these RNAs. One such experiment is hybridisation-proximity based assays (HyPro). These assays utilise

a digoxigenin probe that hybridises to a target RNA, in this instance against NEAT1. Deployment of the HyPro enzyme binds to the digoxigenin probe leading to labelling of adjacent RNA, allowing identification based on close proximity [315]. Moreover identification of any target RNA would aid in determining how the RNAs are selected or shuttled to the condensates. Whether this be by methylation or another unknown mechanism, sequencing of these RNAs could allow screening for methylation motifs or other common features that may aid in trafficking.

While we have focused on the role of SFPQ within the condensates, the effect of SFPQ and other paraspeckle components being sequestered should be further investigated. During lytic replication, the vast majority of the cell's total amount of SFPQ and other core components are re-localised into the condensates, consequently, their normal cellular role is likely to be abrogated. SFPQ has previously been identified as a transcriptional repressor therefore experiments could be performed to investigate the expression levels of transcripts associated with this function [218]. Additionally preliminary data in the Whitehouse laboratory indicates SFPQ may play a role in the DNA damage response pathway. Analysis could focus on whether the increase in DNA damage seen during KSHV lytic replication may be partly ascribed to the viral hijacking of SFPQ to condensates [325].

In conclusion this thesis identifies a range of ways by which KSHV dysregulates the cellular RNA machinery to aid its replication. It also highlights the complexity and depth of this, with this study elucidating the functionality of a single circRNA forming a ncRNA regulatory network. Additionally, we identify KSHV lytic replication drives the formation of novel viral-induced paraspeckle-like condensates that likely play a role in RNA splicing and processing.

References

1. King, A.M.Q., et al., *Virus Taxonomy-Ninth Report of the International Committee on Taxonomy of Viruses*. 1st ed.: Elsevier Inc.
2. Davison, A.J., et al., *The order Herpesvirales.* Arch Virol, 2009. **154**(1): p. 171-7.
3. Andrade-Martinez, J.S., J.L. Moreno-Gallego, and A. Reyes, *Defining a Core Genome for the Herpesvirales and Exploring their Evolutionary Relationship with the Caudovirales.* Sci Rep, 2019. **9**(1): p. 11342.
4. McGeoch, D.J., et al., *Origin and Evolution of Viruses. Chapter 20- Molecular Evolution of the Herpesvirales*. 2nd ed, ed. E. Domingo, C.R. Parrish, and J.J. Holland. 2008: Academic Press
5. McGeoch, D.J., et al., *Molecular phylogeny and evolutionary timescale for the family of mammalian herpesviruses.* J Mol Biol, 1995. **247**(3): p. 443-58.
6. Mori, I. and Y. Nishiyama, *Herpes simplex virus and varicella-zoster virus: why do these human alphaherpesviruses behave so differently from one another?* Rev Med Virol, 2005. **15**(6): p. 393-406.
7. in *Human Herpesviruses: Biology, Therapy, and Immunoprophylaxis*, A. Arvin, et al., Editors. 2007: Cambridge.
8. McGeoch, D.J., A. Dolan, and A.C. Ralph, *Toward a comprehensive phylogeny for mammalian and avian herpesviruses.* J Virol, 2000. **74**(22): p. 10401-6.
9. Burrell, C.J., C.R. Howard, and F.A. Murphy, *Fenner and White's Medical Virology*. 5th ed. 2016: Elsevier Inc.
10. Devadas, D., et al., *Herpes simplex virus internalization into epithelial cells requires Na+/H+ exchangers and p21-activated kinases but neither clathrin- nor caveolin-mediated endocytosis.* J Virol, 2014. **88**(22): p. 13378-95.
11. Jha, H.C., S. Banerjee, and E.S. Robertson, *The Role of Gammaherpesviruses in Cancer Pathogenesis.* Pathogens, 2016. **5**(1).
12. Gillet, L., et al., *Investigation of the susceptibility of human cell lines to bovine herpesvirus 4 infection: demonstration that human cells can support a nonpermissive persistent infection which protects them against tumor necrosis factor alpha-induced apoptosis.* J Virol, 2004. **78**(5): p. 2336-47.
13. Dunn, W., et al., *Functional profiling of a human cytomegalovirus genome.* Proc Natl Acad Sci U S A, 2003. **100**(24): p. 14223-8.
14. AJ, D., *Comparative analysis of the genomes: Human Herpesviruses: Biology, Therapy, and Immunoprophylaxis*. 2007: Cambridge: Cambridge University Press.
15. Chang, Y., et al., *Identification of herpesvirus-like DNA sequences in AIDS-associated Kaposi's sarcoma.* Science, 1994. **266**(5192): p. 1865-9.
16. Russo, J.J., et al., *Nucleotide sequence of the Kaposi sarcoma-associated herpesvirus (HHV8).* Proc Natl Acad Sci U S A, 1996. **93**(25): p. 14862-7.
17. Cohen, J., *Controversy: is KS really caused by new herpesvirus?* Science, 1995. **268**(5219): p. 1847-8.
18. Belec, L., et al., *High prevalence in Central Africa of blood donors who are potentially infectious for human herpesvirus 8.* Transfusion, 1998. **38**(8): p. 771-5.
19. Pauk, J., et al., *Mucosal shedding of human herpesvirus 8 in men.* N Engl J Med, 2000. **343**(19): p. 1369-77.
20. Koelle, D.M., et al., *Frequent detection of Kaposi's sarcoma-associated herpesvirus (human herpesvirus 8) DNA in saliva of human immunodeficiency virus-infected men: clinical and immunologic correlates.* J Infect Dis, 1997. **176**(1): p. 94-102.

21. Plancoulaine, S., et al., *Human herpesvirus 8 transmission from mother to child and between siblings in an endemic population.* Lancet, 2000. **356**(9235): p. 1062-5.
22. Sosa, C., et al., *Human herpesvirus 8 as a potential sexually transmitted agent in Honduras.* J Infect Dis, 1998. **178**(2): p. 547-51.
23. Cesarman, E., et al., *Kaposi sarcoma.* Nat Rev Dis Primers, 2019. **5**(1): p. 9.
24. Wen, K.W. and B. Damania, *Kaposi sarcoma-associated herpesvirus (KSHV): molecular biology and oncogenesis.* Cancer Lett, 2010. **289**(2): p. 140-50.
25. Mesri, E.A., E. Cesarman, and C. Boshoff, *Kaposi's sarcoma and its associated herpesvirus.* Nat Rev Cancer, 2010. **10**(10): p. 707-19.
26. Fan, W., et al., *Distinct subsets of primary effusion lymphoma can be identified based on their cellular gene expression profile and viral association.* J Virol, 2005. **79**(2): p. 1244-51.
27. Uldrick, T.S., et al., *An interleukin-6-related systemic inflammatory syndrome in patients co-infected with Kaposi sarcoma-associated herpesvirus and HIV but without Multicentric Castleman disease.* Clin Infect Dis, 2010. **51**(3): p. 350-8.
28. Polizzotto, M.N., et al., *Clinical Manifestations of Kaposi Sarcoma Herpesvirus Lytic Activation: Multicentric Castleman Disease (KSHV-MCD) and the KSHV Inflammatory Cytokine Syndrome.* Front Microbiol, 2012. **3**: p. 73.
29. Cavallin, L.E., P. Goldschmidt-Clermont, and E.A. Mesri, *Molecular and cellular mechanisms of KSHV oncogenesis of Kaposi's sarcoma associated with HIV/AIDS.* PLoS Pathog, 2014. **10**(7): p. e1004154.
30. Dupin, N., et al., *Distribution of human herpesvirus-8 latently infected cells in Kaposi's sarcoma, multicentric Castleman's disease, and primary effusion lymphoma.* Proc Natl Acad Sci U S A, 1999. **96**(8): p. 4546-51.
31. Liu, D., Y. Wang, and Y. Yuan, *Kaposi's Sarcoma-Associated Herpesvirus K8 Is an RNA Binding Protein That Regulates Viral DNA Replication in Coordination with a Noncoding RNA.* J Virol, 2018. **92**(7).
32. Radkov, S.A., P. Kellam, and C. Boshoff, *The latent nuclear antigen of Kaposi sarcoma-associated herpesvirus targets the retinoblastoma-E2F pathway and with the oncogene Hras transforms primary rat cells.* Nat Med, 2000. **6**(10): p. 1121-7.
33. Swanton, C., et al., *Herpes viral cyclin/Cdk6 complexes evade inhibition by CDK inhibitor proteins.* Nature, 1997. **390**(6656): p. 184-7.
34. Ye, F.C., et al., *Kaposi's sarcoma-associated herpesvirus latent gene vFLIP inhibits viral lytic replication through NF-kappaB-mediated suppression of the AP-1 pathway: a novel mechanism of virus control of latency.* J Virol, 2008. **82**(9): p. 4235-49.
35. Bais, C., et al., *G-protein-coupled receptor of Kaposi's sarcoma-associated herpesvirus is a viral oncogene and angiogenesis activator.* Nature, 1998. **391**(6662): p. 86-9.
36. Wang, L., et al., *Immortalization of primary endothelial cells by the K1 protein of Kaposi's sarcoma-associated herpesvirus.* Cancer Res, 2006. **66**(7): p. 3658-66.
37. Moore, P.S., et al., *Primary characterization of a herpesvirus agent associated with Kaposi's sarcomae.* J Virol, 1996. **70**(1): p. 549-58.
38. Ye, X., Y. Zhao, and J. Karijolich, *The landscape of transcription initiation across latent and lytic KSHV genomes.* PLoS Pathog, 2019. **15**(6): p. e1007852.
39. Whitley, R., *Medical Microbiology*. 4th ed, ed. S. Baron. 1996: University of Texas Medical Branch
40. Dai, X., et al., *Structure and mutagenesis reveal essential capsid protein interactions for KSHV replication.* Nature, 2018. **553**(7689): p. 521-525.

41. Newcomb, W.W., et al., *Structure of the herpes simplex virus capsid. Molecular composition of the pentons and the triplexes.* J Mol Biol, 1993. **232**(2): p. 499-511.
42. Rochat, R.H., et al., *Seeing the portal in herpes simplex virus type 1 B capsids.* J Virol, 2011. **85**(4): p. 1871-4.
43. Gong, D., et al., *DNA-Packing Portal and Capsid-Associated Tegument Complexes in the Tumor Herpesvirus KSHV.* Cell, 2019. **178**(6): p. 1329-1343 e12.
44. Liu, Y.T., et al., *Cryo-EM structures of herpes simplex virus type 1 portal vertex and packaged genome.* Nature, 2019. **570**(7760): p. 257-261.
45. Yan, L., et al., *Towards Better Understanding of KSHV Life Cycle: from Transcription and Posttranscriptional Regulations to Pathogenesis.* Virol Sin, 2019. **34**(2): p. 135-161.
46. Bechtel, J.T., R.C. Winant, and D. Ganem, *Host and viral proteins in the virion of Kaposi's sarcoma-associated herpesvirus.* J Virol, 2005. **79**(8): p. 4952-64.
47. Bortz, E., et al., *Murine gammaherpesvirus 68 ORF52 encodes a tegument protein required for virion morphogenesis in the cytoplasm.* J Virol, 2007. **81**(18): p. 10137-50.
48. Chevillotte, M., et al., *Major tegument protein pp65 of human cytomegalovirus is required for the incorporation of pUL69 and pUL97 into the virus particle and for viral growth in macrophages.* J Virol, 2009. **83**(6): p. 2480-90.
49. Xu, G., et al., *Viral tegument proteins restrict cGAS-DNA phase separation to mediate immune evasion.* Mol Cell, 2021. **81**(13): p. 2823-2837 e9.
50. Meckes, D.G., Jr. and J.W. Wills, *Dynamic interactions of the UL16 tegument protein with the capsid of herpes simplex virus.* J Virol, 2007. **81**(23): p. 13028-36.
51. Sathish, N., X. Wang, and Y. Yuan, *Tegument Proteins of Kaposi's Sarcoma-Associated Herpesvirus and Related Gamma-Herpesviruses.* Front Microbiol, 2012. **3**: p. 98.
52. Leoni, V., et al., *Herpes simplex virus glycoproteins gH/gL and gB bind Toll-like receptor 2, and soluble gH/gL is sufficient to activate NF-kappaB.* J Virol, 2012. **86**(12): p. 6555-62.
53. Rappocciolo, G., et al., *Human herpesvirus 8 infects and replicates in primary cultures of activated B lymphocytes through DC-SIGN.* J Virol, 2008. **82**(10): p. 4793-806.
54. Zhu, F.X., et al., *Virion proteins of Kaposi's sarcoma-associated herpesvirus.* J Virol, 2005. **79**(2): p. 800-11.
55. Birkmann, A., et al., *Cell surface heparan sulfate is a receptor for human herpesvirus 8 and interacts with envelope glycoprotein K8.1.* J Virol, 2001. **75**(23): p. 11583-93.
56. Dollery, S.J., *Towards Understanding KSHV Fusion and Entry.* Viruses, 2019. **11**(11).
57. Dollery, S.J., et al., *Glycoprotein K8.1A of Kaposi's Sarcoma-Associated Herpesvirus Is a Critical B Cell Tropism Determinant Independent of Its Heparan Sulfate Binding Activity.* J Virol, 2019. **93**(6).
58. Veettil, M.V., et al., *RhoA-GTPase facilitates entry of Kaposi's sarcoma-associated herpesvirus into adherent target cells in a Src-dependent manner.* J Virol, 2006. **80**(23): p. 11432-46.
59. Naranatt, P.P., et al., *Kaposi's sarcoma-associated herpesvirus modulates microtubule dynamics via RhoA-GTP-diaphanous 2 signaling and utilizes the dynein motors to deliver its DNA to the nucleus.* J Virol, 2005. **79**(2): p. 1191-206.
60. Naranatt, P.P., et al., *Host gene induction and transcriptional reprogramming in Kaposi's sarcoma-associated herpesvirus (KSHV/HHV-8)-infected endothelial,*

fibroblast, and B cells: insights into modulation events early during infection. Cancer Res, 2004. **64**(1): p. 72-84.
61. Lyman, M.G. and L.W. Enquist, *Herpesvirus interactions with the host cytoskeleton.* J Virol, 2009. **83**(5): p. 2058-66.
62. Pasdeloup, D., et al., *Herpesvirus capsid association with the nuclear pore complex and viral DNA release involve the nucleoporin CAN/Nup214 and the capsid protein pUL25.* J Virol, 2009. **83**(13): p. 6610-23.
63. Dittmer, D., et al., *Experimental transmission of Kaposi's sarcoma-associated herpesvirus (KSHV/HHV-8) to SCID-hu Thy/Liv mice.* J Exp Med, 1999. **190**(12): p. 1857-68.
64. Lim, C., et al., *Functional dissection of latency-associated nuclear antigen 1 of Kaposi's sarcoma-associated herpesvirus involved in latent DNA replication and transcription of terminal repeats of the viral genome.* J Virol, 2002. **76**(20): p. 10320-31.
65. Toth, Z., et al., *Epigenetic analysis of KSHV latent and lytic genomes.* PLoS Pathog, 2010. **6**(7): p. e1001013.
66. Kang, J.G., et al., *Kaposi's sarcoma-associated herpesviral IL-6 and human IL-6 open reading frames contain miRNA binding sites and are subject to cellular miRNA regulation.* J Pathol, 2011. **225**(3): p. 378-89.
67. Qin, J., et al., *KSHV microRNAs: Tricks of the Devil.* Trends Microbiol, 2017. **25**(8): p. 648-661.
68. Ziegelbauer, J.M., *Functions of Kaposi's sarcoma-associated herpesvirus microRNAs.* Biochim Biophys Acta, 2011. **1809**(11-12): p. 623-30.
69. Cai, X., et al., *Kaposi's sarcoma-associated herpesvirus expresses an array of viral microRNAs in latently infected cells.* Proc Natl Acad Sci U S A, 2005. **102**(15): p. 5570-5.
70. Ellison, T.J., et al., *A comprehensive analysis of recruitment and transactivation potential of K-Rta and K-bZIP during reactivation of Kaposi's sarcoma-associated herpesvirus.* Virology, 2009. **387**(1): p. 76-88.
71. Yang, Z., Z. Yan, and C. Wood, *Kaposi's sarcoma-associated herpesvirus transactivator RTA promotes degradation of the repressors to regulate viral lytic replication.* J Virol, 2008. **82**(7): p. 3590-603.
72. Ehrlich, E.S., et al., *KSHV RTA abolishes NFkappaB responsive gene expression during lytic reactivation by targeting vFLIP for degradation via the proteasome.* PLoS One, 2014. **9**(3): p. e91359.
73. Dittmer, D.P. and B. Damania, *Kaposi sarcoma-associated herpesvirus: immunobiology, oncogenesis, and therapy.* J Clin Invest, 2016. **126**(9): p. 3165-75.
74. Aneja, K.K. and Y. Yuan, *Reactivation and Lytic Replication of Kaposi's Sarcoma-Associated Herpesvirus: An Update.* Front Microbiol, 2017. **8**: p. 613.
75. Li, X., J. Feng, and R. Sun, *Oxidative stress induces reactivation of Kaposi's sarcoma-associated herpesvirus and death of primary effusion lymphoma cells.* J Virol, 2011. **85**(2): p. 715-24.
76. Loboda, A., A. Jozkowicz, and J. Dulak, *HIF-1 and HIF-2 transcription factors--similar but not identical.* Mol Cells, 2010. **29**(5): p. 435-42.
77. Davis, D.A., et al., *Hypoxia induces lytic replication of Kaposi sarcoma-associated herpesvirus.* Blood, 2001. **97**(10): p. 3244-50.
78. Dai, L., et al., *Kaposi Sarcoma-Associated Herpesvirus and Staphylococcus aureus Coinfection in Oral Cavities of HIV-Positive Patients: A Unique Niche for Oncogenic Virus Lytic Reactivation.* J Infect Dis, 2020. **221**(8): p. 1331-1341.
79. Drexler, H.G., et al., *Lymphoma cell lines: in vitro models for the study of HHV-8+ primary effusion lymphomas (body cavity-based lymphomas).* Leukemia, 1998. **12**(10): p. 1507-17.
80. Trivedi, P., et al., *Infection of HHV-8+ primary effusion lymphoma cells with a recombinant Epstein-Barr virus leads to restricted EBV latency, altered*

phenotype, and increased tumorigenicity without affecting TCL1 expression. Blood, 2004. **103**(1): p. 313-6.
81. Jiang, Y., et al., *Mutual inhibition between Kaposi's sarcoma-associated herpesvirus and Epstein-Barr virus lytic replication initiators in dually-infected primary effusion lymphoma.* PLoS One, 2008. **3**(2): p. e1569.
82. Spadavecchia, S., et al., *Convergence of Kaposi's sarcoma-associated herpesvirus reactivation with Epstein-Barr virus latency and cellular growth mediated by the notch signaling pathway in coinfected cells.* J Virol, 2010. **84**(20): p. 10488-500.
83. Gottwein, E., et al., *Viral microRNA targetome of KSHV-infected primary effusion lymphoma cell lines.* Cell Host Microbe, 2011. **10**(5): p. 515-26.
84. Ariyoshi, K., et al., *Kaposi's sarcoma in the Gambia, West Africa is less frequent in human immunodeficiency virus type 2 than in human immunodeficiency virus type 1 infection despite a high prevalence of human herpesvirus 8.* J Hum Virol, 1998. **1**(3): p. 193-9.
85. Zhou, F., et al., *HIV-1 Tat promotes Kaposi's sarcoma-associated herpesvirus (KSHV) vIL-6-induced angiogenesis and tumorigenesis by regulating PI3K/PTEN/AKT/GSK-3beta signaling pathway.* PLoS One, 2013. **8**(1): p. e53145.
86. Zeng, Y., et al., *Intracellular Tat of human immunodeficiency virus type 1 activates lytic cycle replication of Kaposi's sarcoma-associated herpesvirus: role of JAK/STAT signaling.* J Virol, 2007. **81**(5): p. 2401-17.
87. Xue, M., et al., *HIV-1 Nef and KSHV oncogene K1 synergistically promote angiogenesis by inducing cellular miR-718 to regulate the PTEN/AKT/mTOR signaling pathway.* Nucleic Acids Res, 2014. **42**(15): p. 9862-79.
88. da Silva, S.R. and D.E. de Oliveira, *HIV, EBV and KSHV: viral cooperation in the pathogenesis of human malignancies.* Cancer Lett, 2011. **305**(2): p. 175-85.
89. Song, M.J., et al., *Characterization of interactions between RTA and the promoter of polyadenylated nuclear RNA in Kaposi's sarcoma-associated herpesvirus/human herpesvirus 8.* J Virol, 2002. **76**(10): p. 5000-13.
90. Liang, Y., et al., *The lytic switch protein of KSHV activates gene expression via functional interaction with RBP-Jkappa (CSL), the target of the Notch signaling pathway.* Genes Dev, 2002. **16**(15): p. 1977-89.
91. Schumann, S., B. Baquero-Perez, and A. Whitehouse, *Interactions between KSHV ORF57 and the novel human TREX proteins, CHTOP and CIP29.* J Gen Virol, 2016. **97**(8): p. 1904-1910.
92. Massimelli, M.J., et al., *Stability of a long noncoding viral RNA depends on a 9-nt core element at the RNA 5' end to interact with viral ORF57 and cellular PABPC1.* Int J Biol Sci, 2011. **7**(8): p. 1145-60.
93. Majerciak, V., et al., *Kaposi's sarcoma-associated herpesvirus ORF57 functions as a viral splicing factor and promotes expression of intron-containing viral lytic genes in spliceosome-mediated RNA splicing.* J Virol, 2008. **82**(6): p. 2792-801.
94. Jackson, B.R., et al., *An interaction between KSHV ORF57 and UIF provides mRNA-adaptor redundancy in herpesvirus intronless mRNA export.* PLoS Pathog, 2011. **7**(7): p. e1002138.
95. Borah, S., et al., *A viral nuclear noncoding RNA binds re-localized poly(A) binding protein and is required for late KSHV gene expression.* PLoS Pathog, 2011. **7**(10): p. e1002300.
96. Lee, H., et al., *KSHV SOX mediated host shutoff: the molecular mechanism underlying mRNA transcript processing.* Nucleic Acids Res, 2017. **45**(8): p. 4756-4767.
97. Glaunsinger, B. and D. Ganem, *Lytic KSHV infection inhibits host gene expression by accelerating global mRNA turnover.* Mol Cell, 2004. **13**(5): p. 713-23.

98. Wang, Y., et al., *Kaposi's sarcoma-associated herpesvirus ori-Lyt-dependent DNA replication: dual role of replication and transcription activator.* J Virol, 2006. **80**(24): p. 12171-86.
99. Kaul, R., et al., *KSHV lytic proteins K-RTA and K8 bind to cellular and viral chromatin to modulate gene expression.* PLoS One, 2019. **14**(4): p. e0215394.
100. Lin, C.L., et al., *Kaposi's sarcoma-associated herpesvirus lytic origin (ori-Lyt)-dependent DNA replication: identification of the ori-Lyt and association of K8 bZip protein with the origin.* J Virol, 2003. **77**(10): p. 5578-88.
101. de Rie, D., et al., *An integrated expression atlas of miRNAs and their promoters in human and mouse.* Nat Biotechnol, 2017. **35**(9): p. 872-878.
102. Ozsolak, F., et al., *Chromatin structure analyses identify miRNA promoters.* Genes Dev, 2008. **22**(22): p. 3172-83.
103. Lee, Y., et al., *MicroRNA genes are transcribed by RNA polymerase II.* EMBO J, 2004. **23**(20): p. 4051-60.
104. Borchert, G.M., W. Lanier, and B.L. Davidson, *RNA polymerase III transcribes human microRNAs.* Nat Struct Mol Biol, 2006. **13**(12): p. 1097-101.
105. Macfarlane, L.A. and P.R. Murphy, *MicroRNA: Biogenesis, Function and Role in Cancer.* Curr Genomics, 2010. **11**(7): p. 537-61.
106. Han, J., et al., *Molecular basis for the recognition of primary microRNAs by the Drosha-DGCR8 complex.* Cell, 2006. **125**(5): p. 887-901.
107. Yi, R., et al., *Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs.* Genes Dev, 2003. **17**(24): p. 3011-6.
108. Fareh, M., et al., *TRBP ensures efficient Dicer processing of precursor microRNA in RNA-crowded environments.* Nat Commun, 2016. **7**: p. 13694.
109. Khvorova, A., A. Reynolds, and S.D. Jayasena, *Functional siRNAs and miRNAs exhibit strand bias.* Cell, 2003. **115**(2): p. 209-16.
110. Singaravelu, R., et al., *MicroRNAs regulate the immunometabolic response to viral infection in the liver.* Nat Chem Biol, 2015. **11**(12): p. 988-93.
111. Fukaya, T., H.O. Iwakawa, and Y. Tomari, *MicroRNAs block assembly of eIF4F translation initiation complex in Drosophila.* Mol Cell, 2014. **56**(1): p. 67-78.
112. Walsh, D. and I. Mohr, *Coupling 40S ribosome recruitment to modification of a cap-binding initiation factor by eIF3 subunit e.* Genes Dev, 2014. **28**(8): p. 835-40.
113. Chendrimada, T.P., et al., *MicroRNA silencing through RISC recruitment of eIF6.* Nature, 2007. **447**(7146): p. 823-8.
114. Gibbings, D.J., et al., *Multivesicular bodies associate with components of miRNA effector complexes and modulate miRNA activity.* Nat Cell Biol, 2009. **11**(9): p. 1143-9.
115. Kingston, E.R., L.W. Blodgett, and D.P. Bartel, *Endogenous transcripts direct microRNA degradation in Drosophila, and this targeted degradation is required for proper embryonic development.* Mol Cell, 2022. **82**(20): p. 3872-3884 e9.
116. Sheu-Gruttadauria, J., et al., *Structural Basis for Target-Directed MicroRNA Degradation.* Mol Cell, 2019. **75**(6): p. 1243-1255 e7.
117. Gebert, L.F.R. and I.J. MacRae, *Regulation of microRNA function in animals.* Nat Rev Mol Cell Biol, 2019. **20**(1): p. 21-37.
118. Wang, S., W. Wu, and F.X. Claret, *Mutual regulation of microRNAs and DNA methylation in human cancers.* Epigenetics, 2017. **12**(3): p. 187-197.
119. Kawahara, Y., et al., *Frequency and fate of microRNA editing in human brain.* Nucleic Acids Res, 2008. **36**(16): p. 5270-80.
120. Salmena, L., et al., *A ceRNA hypothesis: the Rosetta Stone of a hidden RNA language?* Cell, 2011. **146**(3): p. 353-8.
121. Skalsky, R.L. and B.R. Cullen, *Viruses, microRNAs, and host interactions.* Annu Rev Microbiol, 2010. **64**: p. 123-41.
122. Kincaid, R.P. and C.S. Sullivan, *Virus-encoded microRNAs: an overview and a look to the future.* PLoS Pathog, 2012. **8**(12): p. e1003018.

123. Grundhoff, A. and C.S. Sullivan, *Virus-encoded microRNAs.* Virology, 2011. **411**(2): p. 325-43.
124. Liu, Y., et al., *Ebola virus encodes a miR-155 analog to regulate importin-alpha5 expression.* Cell Mol Life Sci, 2016. **73**(19): p. 3733-44.
125. Duy, J., et al., *Virus-encoded miRNAs in Ebola virus disease.* Sci Rep, 2018. **8**(1): p. 6480.
126. Varble, A., et al., *Engineered RNA viral synthesis of microRNAs.* Proc Natl Acad Sci U S A, 2010. **107**(25): p. 11519-24.
127. Rouha, H., C. Thurner, and C.W. Mandl, *Functional microRNA generated from a cytoplasmic RNA virus.* Nucleic Acids Res, 2010. **38**(22): p. 8328-37.
128. Fukuhara, T., et al., *Expression of microRNA miR-122 facilitates an efficient replication in nonhepatic cells upon infection with hepatitis C virus.* J Virol, 2012. **86**(15): p. 7918-33.
129. Machlin, E.S., P. Sarnow, and S.M. Sagan, *Masking the 5' terminal nucleotides of the hepatitis C virus genome by an unconventional microRNA-target RNA complex.* Proc Natl Acad Sci U S A, 2011. **108**(8): p. 3193-8.
130. Sedano, C.D. and P. Sarnow, *Hepatitis C virus subverts liver-specific miR-122 to protect the viral genome from exoribonuclease Xrn2.* Cell Host Microbe, 2014. **16**(2): p. 257-264.
131. Kakumani, P.K., et al., *Role of RNA interference (RNAi) in dengue virus replication and identification of NS4B as an RNAi suppressor.* J Virol, 2013. **87**(16): p. 8870-83.
132. Clarke, B.D., et al., *Functional non-coding RNAs derived from the flavivirus 3' untranslated region.* Virus Res, 2015. **206**: p. 53-61.
133. Schnettler, E., et al., *Noncoding flavivirus RNA displays RNA interference suppressor activity in insect and Mammalian cells.* J Virol, 2012. **86**(24): p. 13486-500.
134. Fuchs Wightman, F., et al., *Target RNAs Strike Back on MicroRNAs.* Front Genet, 2018. **9**: p. 435.
135. Qin, Z., et al., *Role of host microRNAs in Kaposi's sarcoma-associated herpesvirus pathogenesis.* Viruses, 2014. **6**(11): p. 4571-80.
136. Sanger, H.L., et al., *Viroids are single-stranded covalently closed circular RNA molecules existing as highly base-paired rod-like structures.* Proc Natl Acad Sci U S A, 1976. **73**(11): p. 3852-6.
137. Hsu, M.T. and M. Coca-Prados, *Electron microscopic evidence for the circular form of RNA in the cytoplasm of eukaryotic cells.* Nature, 1979. **280**(5720): p. 339-40.
138. Kos, A., et al., *The hepatitis delta (delta) virus possesses a circular RNA.* Nature, 1986. **323**(6088): p. 558-60.
139. Nigro, J.M., et al., *Scrambled exons.* Cell, 1991. **64**(3): p. 607-13.
140. Dubin, R.A., M.A. Kazmi, and H. Ostrer, *Inverted repeats are necessary for circularization of the mouse testis Sry transcript.* Gene, 1995. **167**(1-2): p. 245-8.
141. Suzuki, H., et al., *Characterization of RNase R-digested cellular RNA source that consists of lariat and circular RNAs from pre-mRNA splicing.* Nucleic Acids Res, 2006. **34**(8): p. e63.
142. Panda, A.C., et al., *High-purity circular RNA isolation method (RPAD) reveals vast collection of intronic circRNAs.* Nucleic Acids Res, 2017. **45**(12): p. e116.
143. Hansen, T.B., et al., *Natural RNA circles function as efficient microRNA sponges.* Nature, 2013. **495**(7441): p. 384-8.
144. Ji, P., et al., *Expanded Expression Landscape and Prioritization of Circular RNAs in Mammals.* Cell Rep, 2019. **26**(12): p. 3444-3460 e5.
145. Wang, P.L., et al., *Circular RNA is expressed across the eukaryotic tree of life.* PLoS One, 2014. **9**(6): p. e90859.

146. Memczak, S., et al., *Circular RNAs are a large class of animal RNAs with regulatory potency.* Nature, 2013. **495**(7441): p. 333-8.
147. Lasda, E. and R. Parker, *Circular RNAs: diversity of form and function.* RNA, 2014. **20**(12): p. 1829-42.
148. Li, Z., et al., *Exon-intron circular RNAs regulate transcription in the nucleus.* Nat Struct Mol Biol, 2015. **22**(3): p. 256-64.
149. Harper, K.L., E. McDonnell, and A. Whitehouse, *CircRNAs: From anonymity to novel regulators of gene expression in cancer (Review).* Int J Oncol, 2019. **55**(6): p. 1183-1193.
150. Xiao, M.S., Y. Ai, and J.E. Wilusz, *Biogenesis and Functions of Circular RNAs Come into Focus.* Trends Cell Biol, 2020. **30**(3): p. 226-240.
151. Chen, L.L., *The expanding regulatory mechanisms and cellular functions of circular RNAs.* Nat Rev Mol Cell Biol, 2020. **21**(8): p. 475-490.
152. Huang, C. and G. Shan, *What happens at or after transcription: Insights into circRNA biogenesis and function.* Transcription, 2015. **6**(4): p. 61-4.
153. Huang, A., et al., *Circular RNA-protein interactions: functions, mechanisms, and identification.* Theranostics, 2020. **10**(8): p. 3503-3517.
154. Ashwal-Fluss, R., et al., *circRNA biogenesis competes with pre-mRNA splicing.* Mol Cell, 2014. **56**(1): p. 55-66.
155. Conn, S.J., et al., *The RNA binding protein quaking regulates formation of circRNAs.* Cell, 2015. **160**(6): p. 1125-34.
156. Verheijen, B.M. and R.J. Pasterkamp, *Commentary: FUS affects circular RNA expression in murine embryonic stem cell-derived motor neurons.* Front Mol Neurosci, 2017. **10**: p. 412.
157. Yang, L., J. Fu, and Y. Zhou, *Circular RNAs and Their Emerging Roles in Immune Regulation.* Front Immunol, 2018. **9**: p. 2977.
158. Stagsted, L.V., et al., *Noncoding AUG circRNAs constitute an abundant and conserved subclass of circles.* Life Sci Alliance, 2019. **2**(3).
159. Wang, M., et al., *Long and Repeat-Rich Intronic Sequences Favor Circular RNA Formation under Conditions of Reduced Spliceosome Activity.* iScience, 2019. **20**: p. 237-247.
160. Zhang, Y., et al., *The Biogenesis of Nascent Circular RNAs.* Cell Rep, 2016. **15**(3): p. 611-624.
161. Hansen, T.B., et al., *miRNA-dependent gene silencing involving Ago2-mediated cleavage of a circular antisense RNA.* EMBO J, 2011. **30**(21): p. 4414-22.
162. Park, O.H., et al., *Endoribonucleolytic Cleavage of m(6)A-Containing RNAs by RNase P/MRP Complex.* Mol Cell, 2019. **74**(3): p. 494-507 e8.
163. Fischer, J.W., et al., *Structure-Mediated RNA Decay by UPF1 and G3BP1.* Mol Cell, 2020. **78**(1): p. 70-84 e6.
164. Liu, C.X., et al., *Structure and Degradation of Circular RNAs Regulate PKR Activation in Innate Immunity.* Cell, 2019. **177**(4): p. 865-880 e21.
165. Li, X., et al., *Coordinated circRNA Biogenesis and Function with NF90/NF110 in Viral Infection.* Mol Cell, 2017. **67**(2): p. 214-227 e7.
166. Kulcheski, F.R., A.P. Christoff, and R. Margis, *Circular RNAs are miRNA sponges and can be used as a new class of biomarker.* J Biotechnol, 2016. **238**: p. 42-51.
167. Wan, L., et al., *Circular RNA-ITCH Suppresses Lung Cancer Proliferation via Inhibiting the Wnt/beta-Catenin Pathway.* Biomed Res Int, 2016. **2016**: p. 1579490.
168. Kluiver, J., et al., *Rapid generation of microRNA sponges for microRNA inhibition.* PLoS One, 2012. **7**(1): p. e29275.
169. Ragan, C., et al., *Insights into the biogenesis and potential functions of exonic circular RNA.* Sci Rep, 2019. **9**(1): p. 2048.
170. Zhang, Y., et al., *Circular intronic long noncoding RNAs.* Mol Cell, 2013. **51**(6): p. 792-806.

171. Chao, C.W., et al., *The mouse formin (Fmn) gene: abundant circular RNA transcripts and gene-targeted deletion analysis.* Mol Med, 1998. **4**(9): p. 614-28.
172. Shen, T., et al., *An intriguing RNA species--perspectives of circularized RNA.* Protein Cell, 2015. **6**(12): p. 871-80.
173. Jeck, W.R. and N.E. Sharpless, *Detecting and characterizing circular RNAs.* Nat Biotechnol, 2014. **32**(5): p. 453-61.
174. Du, W.W., et al., *Induction of tumor apoptosis through a circular RNA enhancing Foxo3 activity.* Cell Death Differ, 2017. **24**(2): p. 357-370.
175. Abdelmohsen, K., et al., *Identification of HuR target circular RNAs uncovers suppression of PABPN1 translation by CircPABPN1.* RNA Biol, 2017. **14**(3): p. 361-369.
176. Holdt, L.M., et al., *Circular non-coding RNA ANRIL modulates ribosomal RNA maturation and atherosclerosis in humans.* Nat Commun, 2016. **7**: p. 12429.
177. Du, W.W., et al., *A circular RNA circ-DNMT1 enhances breast cancer progression by activating autophagy.* Oncogene, 2018. **37**(44): p. 5829-5842.
178. Liu, Y., et al., *Circular RNAMTO1 suppresses breast cancer cell viability and reverses monastrol resistance through regulating the TRAF4/Eg5 axis.* Int J Oncol, 2018. **53**(4): p. 1752-1762.
179. Zeng, Y., et al., *A Circular RNA Binds To and Activates AKT Phosphorylation and Nuclear Localization Reducing Apoptosis and Enhancing Cardiac Repair.* Theranostics, 2017. **7**(16): p. 3842-3855.
180. Chen, C.Y. and P. Sarnow, *Initiation of protein synthesis by the eukaryotic translational apparatus on circular RNAs.* Science, 1995. **268**(5209): p. 415-7.
181. Wang, Y. and Z. Wang, *Efficient backsplicing produces translatable circular mRNAs.* RNA, 2015. **21**(2): p. 172-9.
182. Granados-Riveron, J.T. and G. Aquino-Jarquin, *The complexity of the translation ability of circRNAs.* Biochim Biophys Acta, 2016. **1859**(10): p. 1245-51.
183. Pamudurti, N.R., et al., *Translation of CircRNAs.* Mol Cell, 2017. **66**(1): p. 9-21 e7.
184. Legnini, I., et al., *Circ-ZNF609 Is a Circular RNA that Can Be Translated and Functions in Myogenesis.* Mol Cell, 2017. **66**(1): p. 22-37 e9.
185. Tagawa, T., et al., *Discovery of Kaposi's sarcoma herpesvirus-encoded circular RNAs and a human antiviral circular RNA.* Proc Natl Acad Sci U S A, 2018.
186. Toptan, T., et al., *Circular DNA tumor viruses make circular RNAs.* Proc Natl Acad Sci U S A, 2018. **115**(37): p. E8737-E8745.
187. Abere, B., et al., *Kaposi's Sarcoma-Associated Herpesvirus-Encoded circRNAs Are Expressed in Infected Tumor Tissues and Are Incorporated into Virions.* mBio, 2020. **11**(1).
188. Zhao, J., et al., *Transforming activity of an oncoprotein-encoding circular RNA from human papillomavirus.* Nat Commun, 2019. **10**(1): p. 2300.
189. Yu, L. and Z.M. Zheng, *Human Papillomavirus Type 16 Circular RNA Is Barely Detectable for the Claimed Biological Activity.* mBio, 2022. **13**(1): p. e0359421.
190. Cai, Z., et al., *VirusCircBase: a database of virus circular RNAs.* Brief Bioinform, 2021. **22**(2): p. 2182-2190.
191. Shi, J., et al., *Unique expression signatures of circular RNAs in response to DNA tumor virus SV40 infection.* Oncotarget, 2017. **8**(58): p. 98609-98622.
192. Zhang, X., et al., *Circular RNA Vav3 sponges gga-miR-375 to promote epithelial-mesenchymal transition.* RNA Biol, 2019. **16**(1): p. 118-132.
193. Zhang, X., et al., *Circular RNA alterations are involved in resistance to avian leukosis virus subgroup-J-induced tumor formation in chickens.* Oncotarget, 2017. **8**(21): p. 34961-34970.
194. Wang, Z.Y., et al., *Genome-Wide Search for Competing Endogenous RNAs Responsible for the Effects Induced by Ebola Virus Replication and Transcription Using a trVLP System.* Front Cell Infect Microbiol, 2017. **7**: p. 479.

195. Shi, J., et al., *Deep RNA Sequencing Reveals a Repertoire of Human Fibroblast Circular RNAs Associated with Cellular Responses to Herpes Simplex Virus 1 Infection.* Cell Physiol Biochem, 2018. **47**(5): p. 2031-2045.
196. Zhang, Y., et al., *Crosstalk in competing endogenous RNA networks reveals new circular RNAs involved in the pathogenesis of early HIV infection.* J Transl Med, 2018. **16**(1): p. 332.
197. Lu, S., et al., *RNA-Seq Revealed a Circular RNA-microRNA-mRNA Regulatory Network in Hantaan Virus Infection.* Front Cell Infect Microbiol, 2020. **10**: p. 97.
198. Stagsted, L.V.W., et al., *The RNA-binding protein SFPQ preserves long-intron splicing and regulates circRNA biogenesis in mammals.* Elife, 2021. **10**.
199. Doucas, V., et al., *Adenovirus replication is coupled with the dynamic properties of the PML nuclear structure.* Genes Dev, 1996. **10**(2): p. 196-207.
200. Mao, Y.S., B. Zhang, and D.L. Spector, *Biogenesis and function of nuclear bodies.* Trends Genet, 2011. **27**(8): p. 295-306.
201. Andersen, J.S., et al., *Directed proteomic analysis of the human nucleolus.* Curr Biol, 2002. **12**(1): p. 1-11.
202. Fox, A.H., et al., *Paraspeckles: a novel nuclear domain.* Curr Biol, 2002. **12**(1): p. 13-25.
203. Clark, M.B., et al., *Genome-wide analysis of long noncoding RNA stability.* Genome Res, 2012. **22**(5): p. 885-98.
204. Taiana, E., et al., *LncRNA NEAT1 in Paraspeckles: A Structural Scaffold for Cellular DNA Damage Response Systems?* Noncoding RNA, 2020. **6**(3).
205. Naganuma, T., et al., *Alternative 3'-end processing of long noncoding RNA initiates construction of nuclear paraspeckles.* EMBO J, 2012. **31**(20): p. 4020-34.
206. Fox, A.H., et al., *Paraspeckles: Where Long Noncoding RNA Meets Phase Separation.* Trends Biochem Sci, 2018. **43**(2): p. 124-135.
207. West, J.A., et al., *Structural, super-resolution microscopy analysis of paraspeckle nuclear body organization.* J Cell Biol, 2016. **214**(7): p. 817-30.
208. Hirose, T., T. Yamazaki, and S. Nakagawa, *Molecular anatomy of the architectural NEAT1 noncoding RNA: The domains, interactors, and biogenesis pathway required to build phase-separated nuclear paraspeckles.* Wiley Interdiscip Rev RNA, 2019. **10**(6): p. e1545.
209. Hennig, S., et al., *Prion-like domains in RNA binding proteins are essential for building subnuclear paraspeckles.* J Cell Biol, 2015. **210**(4): p. 529-39.
210. Kawaguchi, T., et al., *SWI/SNF chromatin-remodeling complexes function in noncoding RNA-dependent assembly of nuclear bodies.* Proc Natl Acad Sci U S A, 2015. **112**(14): p. 4304-9.
211. Fong, K.W., et al., *Whole-genome screening identifies proteins localized to distinct nuclear bodies.* J Cell Biol, 2013. **203**(1): p. 149-64.
212. Major, A.T., et al., *Dynamic paraspeckle component localisation during spermatogenesis.* Reproduction, 2019. **158**(3): p. 267-280.
213. Chen, L.L. and G.G. Carmichael, *Altered nuclear retention of mRNAs containing inverted repeats in human embryonic stem cells: functional role of a nuclear noncoding RNA.* Mol Cell, 2009. **35**(4): p. 467-78.
214. Prasanth, K.V., et al., *Regulating gene expression through RNA nuclear retention.* Cell, 2005. **123**(2): p. 249-63.
215. Nakagawa, S., T. Yamazaki, and T. Hirose, *Molecular dissection of nuclear paraspeckles: towards understanding the emerging world of the RNP milieu.* Open Biol, 2018. **8**(10).
216. Hu, S.B., et al., *Protein arginine methyltransferase CARM1 attenuates the paraspeckle-mediated nuclear retention of mRNAs containing IRAlus.* Genes Dev, 2015. **29**(6): p. 630-45.
217. Visa, N., et al., *Intranuclear distribution of U1 and U2 snRNAs visualized by high resolution in situ hybridization: revelation of a novel compartment*

containing U1 but not U2 snRNA in HeLa cells. Eur J Cell Biol, 1993. **60**(2): p. 308-21.
218. Imamura, K., et al., *Long noncoding RNA NEAT1-dependent SFPQ relocation from promoter region to paraspeckle mediates IL8 expression upon immune stimuli.* Mol Cell, 2014. **53**(3): p. 393-406.
219. Dong, X., et al., *Transcriptional activity of androgen receptor is modulated by two RNA splicing factors, PSF and p54nrb.* Mol Cell Biol, 2007. **27**(13): p. 4863-75.
220. Mathur, M., P.W. Tucker, and H.H. Samuels, *PSF is a novel corepressor that mediates its effect through Sin3A and the DNA binding domain of nuclear hormone receptors.* Mol Cell Biol, 2001. **21**(7): p. 2298-311.
221. Rosonina, E., et al., *Role for PSF in mediating transcriptional activator-dependent stimulation of pre-mRNA processing in vivo.* Mol Cell Biol, 2005. **25**(15): p. 6734-46.
222. Lim, Y.W., et al., *The Emerging Role of the RNA-Binding Protein SFPQ in Neuronal Function and Neurodegeneration.* Int J Mol Sci, 2020. **21**(19).
223. Hirose, T., et al., *NEAT1 long noncoding RNA regulates transcription via protein sequestration within subnuclear bodies.* Mol Biol Cell, 2014. **25**(1): p. 169-83.
224. Takeuchi, A., et al., *Loss of Sfpq Causes Long-Gene Transcriptopathy in the Brain.* Cell Rep, 2018. **23**(5): p. 1326-1341.
225. Hosokawa, M., et al., *Loss of RNA-Binding Protein Sfpq Causes Long-Gene Transcriptopathy in Skeletal Muscle and Severe Muscle Mass Reduction with Metabolic Myopathy.* iScience, 2019. **13**: p. 229-242.
226. Ivanov, A., et al., *Analysis of intron sequences reveals hallmarks of circular RNA biogenesis in animals.* Cell Rep, 2015. **10**(2): p. 170-7.
227. Errichelli, L., et al., *FUS affects circular RNA expression in murine embryonic stem cell-derived motor neurons.* Nat Commun, 2017. **8**: p. 14741.
228. Wilkinson, M.E., C. Charenton, and K. Nagai, *RNA Splicing by the Spliceosome.* Annu Rev Biochem, 2020. **89**: p. 359-388.
229. Wang, Y., et al., *Mechanism of alternative splicing and its regulation.* Biomed Rep, 2015. **3**(2): p. 152-158.
230. Damianov, A., et al., *Rbfox Proteins Regulate Splicing as Part of a Large Multiprotein Complex LASR.* Cell, 2016. **165**(3): p. 606-19.
231. Melton, A.A., et al., *Combinatorial control of signal-induced exon repression by hnRNP L and PSF.* Mol Cell Biol, 2007. **27**(19): p. 6972-84.
232. Ha, K., Y. Takeda, and W.S. Dynan, *Sequences in PSF/SFPQ mediate radioresistance and recruitment of PSF/SFPQ-containing complexes to DNA damage sites in human cells.* DNA Repair (Amst), 2011. **10**(3): p. 252-9.
233. Rajesh, C., et al., *The splicing-factor related protein SFPQ/PSF interacts with RAD51D and is necessary for homology-directed repair and sister chromatid cohesion.* Nucleic Acids Res, 2011. **39**(1): p. 132-45.
234. Jaafar, L., et al., *SFPQ*NONO and XLF function separately and together to promote DNA double-strand break repair via canonical nonhomologous end joining.* Nucleic Acids Res, 2017. **45**(4): p. 1848-1859.
235. Jiang, L., et al., *NEAT1 scaffolds RNA-binding proteins and the Microprocessor to globally enhance pri-miRNA processing.* Nat Struct Mol Biol, 2017. **24**(10): p. 816-824.
236. Kim, Y.K. and V.N. Kim, *Processing of intronic microRNAs.* EMBO J, 2007. **26**(3): p. 775-83.
237. Kaneko, S., et al., *The multifunctional protein p54nrb/PSF recruits the exonuclease XRN2 to facilitate pre-mRNA 3' processing and transcription termination.* Genes Dev, 2007. **21**(14): p. 1779-89.
238. Saha, S., S. Murthy, and P.N. Rangarajan, *Identification and characterization of a virus-inducible non-coding RNA in mouse brain.* J Gen Virol, 2006. **87**(Pt 7): p. 1991-1995.

239. Morchikh, M., et al., *HEXIM1 and NEAT1 Long Non-coding RNA Form a Multi-subunit Complex that Regulates DNA-Mediated Innate Immune Response.* Mol Cell, 2017. **67**(3): p. 387-399 e5.
240. Beeharry, Y., et al., *The Hepatitis Delta Virus accumulation requires paraspeckle components and affects NEAT1 level and PSP1 localization.* Sci Rep, 2018. **8**(1): p. 6031.
241. Li, K. and Z. Wang, *Speckles and paraspeckles coordinate to regulate HSV-1 genes transcription.* Commun Biol, 2021. **4**(1): p. 1207.
242. Wang, Z., et al., *NEAT1 modulates herpes simplex virus-1 replication by regulating viral gene transcription.* Cell Mol Life Sci, 2017. **74**(6): p. 1117-1131.
243. Landeras-Bueno, S., et al., *The splicing factor proline-glutamine rich (SFPQ/PSF) is involved in influenza virus transcription.* PLoS Pathog, 2011. **7**(11): p. e1002397.
244. Zhou, B., et al., *Exploitation of nuclear protein SFPQ by the encephalomyocarditis virus to facilitate its replication.* Biochem Biophys Res Commun, 2019. **510**(1): p. 65-71.
245. Zeng, C., et al., *The c-Myc-regulated lncRNA NEAT1 and paraspeckles modulate imatinib-induced apoptosis in CML cells.* Mol Cancer, 2018. **17**(1): p. 130.
246. Choudhry, H., et al., *Tumor hypoxia induces nuclear paraspeckle formation through HIF-2alpha dependent transcriptional activation of NEAT1 leading to cancer cell survival.* Oncogene, 2015. **34**(34): p. 4482-90.
247. Idogawa, M., et al., *Long non-coding RNA NEAT1 is a transcriptional target of p53 and modulates p53-induced transactivation and tumor-suppressor function.* Int J Cancer, 2017. **140**(12): p. 2785-2791.
248. Nishimoto, Y., et al., *The long non-coding RNA nuclear-enriched abundant transcript 1_2 induces paraspeckle formation in the motor neuron during the early phase of amyotrophic lateral sclerosis.* Mol Brain, 2013. **6**: p. 31.
249. Thomas-Jinu, S., et al., *Non-nuclear Pool of Splicing Factor SFPQ Regulates Axonal Transcripts Required for Normal Motor Development.* Neuron, 2017. **94**(4): p. 931.
250. Deng, H.X., et al., *FUS-immunoreactive inclusions are a common feature in sporadic and non-SOD1 familial amyotrophic lateral sclerosis.* Ann Neurol, 2010. **67**(6): p. 739-48.
251. Tollervey, J.R., et al., *Characterizing the RNA targets and position-dependent splicing regulation by TDP-43.* Nat Neurosci, 2011. **14**(4): p. 452-8.
252. Moffat, J., et al., *A lentiviral RNAi library for human and mouse genes applied to an arrayed viral high-content screen.* Cell, 2006. **124**(6): p. 1283-98.
253. Kimura, K. and A. Koike, *Analysis of genomic rearrangements by using the Burrows-Wheeler transform of short-read data.* BMC Bioinformatics, 2015. **16 Suppl 18**(Suppl 18): p. S5.
254. Bolger, A.M., M. Lohse, and B. Usadel, *Trimmomatic: a flexible trimmer for Illumina sequence data.* Bioinformatics, 2014. **30**(15): p. 2114-20.
255. Pertea, M., et al., *Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown.* Nat Protoc, 2016. **11**(9): p. 1650-67.
256. Langmead, B. and S.L. Salzberg, *Fast gapped-read alignment with Bowtie 2.* Nat Methods, 2012. **9**(4): p. 357-9.
257. Anders, S., P.T. Pyl, and W. Huber, *HTSeq--a Python framework to work with high-throughput sequencing data.* Bioinformatics, 2015. **31**(2): p. 166-9.
258. Morris, K.V. and J.S. Mattick, *The rise of regulatory RNA.* Nat Rev Genet, 2014. **15**(6): p. 423-37.
259. Yamamura, S., et al., *Interaction and cross-talk between non-coding RNAs.* Cell Mol Life Sci, 2018. **75**(3): p. 467-484.
260. Tay, Y., J. Rinn, and P.P. Pandolfi, *The multilayered complexity of ceRNA crosstalk and competition.* Nature, 2014. **505**(7483): p. 344-52.

261. Kim, Y.H., et al., *Homeodomain-interacting protein kinases, a novel family of co-repressors for homeodomain transcription factors.* J Biol Chem, 1998. **273**(40): p. 25875-9.
262. Dudekula, D.B., et al., *CircInteractome: A web tool for exploring circular RNAs and their interacting proteins and microRNAs.* RNA Biol, 2016. **13**(1): p. 34-42.
263. Zheng, Q., et al., *Circular RNA profiling reveals an abundant circHIPK3 that regulates cell growth by sponging multiple miRNAs.* Nat Commun, 2016. **7**: p. 11215.
264. Xie, Y., et al., *The circular RNA HIPK3 (circHIPK3) and its regulation in cancer progression: Review.* Life Sci, 2020. **254**: p. 117252.
265. Zhou, J., et al., *CircHIPK3: Key Player in Pathophysiology and Potential Diagnostic and Therapeutic Tool.* Front Med (Lausanne), 2021. **8**: p. 615417.
266. Croset, M., et al., *miRNA-30 Family Members Inhibit Breast Cancer Invasion, Osteomimicry, and Bone Destruction by Directly Targeting Multiple Bone Metastasis-Associated Genes.* Cancer Res, 2018. **78**(18): p. 5259-5273.
267. Zhang, Q., et al., *Role of microRNA-30c targeting ADAM19 in colorectal cancer.* PLoS One, 2015. **10**(3): p. e0120698.
268. Ling, X.H., et al., *MicroRNA-30c serves as an independent biochemical recurrence predictor and potential tumor suppressor for prostate cancer.* Mol Biol Rep, 2014. **41**(5): p. 2779-88.
269. Lai, Y.H., et al., *Collagen triple helix repeat containing-1 negatively regulated by microRNA-30c promotes cell proliferation and metastasis and indicates poor prognosis in breast cancer.* J Exp Clin Cancer Res, 2017. **36**(1): p. 92.
270. Koh, W., et al., *Analysis of deep sequencing microRNA expression profile from human embryonic stem cells derived mesenchymal stem cells reveals possible role of let-7 microRNA family in downstream targeting of hepatic nuclear factor 4 alpha.* BMC Genomics, 2010. **11 Suppl 1**(Suppl 1): p. S6.
271. Gould, F., et al., *Kaposi's sarcoma-associated herpesvirus RTA promotes degradation of the Hey1 repressor protein through the ubiquitin proteasome pathway.* J Virol, 2009. **83**(13): p. 6727-38.
272. Schumann, S., et al., *Targeting the ATP-dependent formation of herpesvirus ribonucleoprotein particle assembly as an antiviral approach.* Nat Microbiol, 2016. **2**: p. 16201.
273. Sei, E., et al., *HITS-CLIP analysis uncovers a link between the Kaposi's sarcoma-associated herpesvirus ORF57 protein and host pre-mRNA metabolism.* PLoS Pathog, 2015. **11**(2): p. e1004652.
274. Taylor, A., et al., *Mutation of a C-terminal motif affects Kaposi's sarcoma-associated herpesvirus ORF57 RNA binding, nuclear trafficking, and multimerization.* J Virol, 2011. **85**(15): p. 7881-91.
275. Majerciak, V., et al., *Stability of structured Kaposi's sarcoma-associated herpesvirus ORF57 protein is regulated by protein phosphorylation and homodimerization.* J Virol, 2015. **89**(6): p. 3256-74.
276. Bitetti, A., et al., *MicroRNA degradation by a conserved target RNA regulates animal behavior.* Nat Struct Mol Biol, 2018. **25**(3): p. 244-251.
277. Ghini, F., et al., *Endogenous transcripts control miRNA levels and activity in mammalian cells by target-directed miRNA degradation.* Nat Commun, 2018. **9**(1): p. 3119.
278. Cazalla, D., T. Yario, and J.A. Steitz, *Down-regulation of a host microRNA by a Herpesvirus saimiri noncoding RNA.* Science, 2010. **328**(5985): p. 1563-6.
279. Libri, V., et al., *Murine cytomegalovirus encodes a miR-27 inhibitor disguised as a target.* Proc Natl Acad Sci U S A, 2012. **109**(1): p. 279-84.
280. Lee, S., et al., *Selective degradation of host MicroRNAs by an intergenic HCMV noncoding RNA accelerates virus production.* Cell Host Microbe, 2013. **13**(6): p. 678-90.

281. Yuan, F., et al., *The crystal structure of KSHV ORF57 reveals dimeric active sites important for protein stability and function.* PLoS Pathog, 2018. **14**(8): p. e1007232.
282. Pitulescu, M.E., et al., *Dll4 and Notch signalling couples sprouting angiogenesis and artery formation.* Nat Cell Biol, 2017. **19**(8): p. 915-927.
283. Emuss, V., et al., *KSHV manipulates Notch signaling by DLL4 and JAG1 to alter cell cycle genes in lymphatic endothelia.* PLoS Pathog, 2009. **5**(10): p. e1000616.
284. Agarwal, V., et al., *Predicting effective microRNA target sites in mammalian mRNAs.* Elife, 2015. **4**.
285. Kozomara, A., M. Birgaoanu, and S. Griffiths-Jones, *miRBase: from microRNA sequences to function.* Nucleic Acids Res, 2019. **47**(D1): p. D155-D162.
286. Wong, N. and X. Wang, *miRDB: an online resource for microRNA target prediction and functional annotations.* Nucleic Acids Res, 2015. **43**(Database issue): p. D146-52.
287. Ellwanger, D.C., et al., *The sufficient minimal set of miRNA seed types.* Bioinformatics, 2011. **27**(10): p. 1346-50.
288. Liu, R., et al., *KSHV-induced notch components render endothelial and mural cell characteristics and cell survival.* Blood, 2010. **115**(4): p. 887-95.
289. Rodriguez, A., E.J. Jung, and E.K. Flemington, *Cell cycle analysis of Epstein-Barr virus-infected cells following treatment with lytic cycle-inducing agents.* J Virol, 2001. **75**(10): p. 4482-9.
290. Kudoh, A., et al., *Reactivation of lytic replication from B cells latently infected with Epstein-Barr virus occurs with high S-phase cyclin-dependent kinase activity while inhibiting cellular DNA replication.* J Virol, 2003. **77**(2): p. 851-61.
291. Salvant, B.S., E.A. Fortunato, and D.H. Spector, *Cell cycle dysregulation by human cytomegalovirus: influence of the cell cycle phase at the time of infection and effects on cyclin transcription.* J Virol, 1998. **72**(5): p. 3729-41.
292. Jordan, R., L. Schang, and P.A. Schaffer, *Transactivation of herpes simplex virus type 1 immediate-early gene expression by virion-associated factors is blocked by an inhibitor of cyclin-dependent protein kinases.* J Virol, 1999. **73**(10): p. 8843-7.
293. Yarosh, C.A., et al., *PSF: nuclear busy-body or nuclear facilitator?* Wiley Interdiscip Rev RNA, 2015. **6**(4): p. 351-67.
294. Knott, G.J., C.S. Bond, and A.H. Fox, *The DBHS proteins SFPQ, NONO and PSPC1: a multipurpose molecular scaffold.* Nucleic Acids Res, 2016. **44**(9): p. 3989-4004.
295. Wang, B., et al., *Liquid-liquid phase separation in human health and diseases.* Signal Transduct Target Ther, 2021. **6**(1): p. 290.
296. Ray, P., et al., *PSF suppresses tau exon 10 inclusion by interacting with a stem-loop structure downstream of exon 10.* J Mol Neurosci, 2011. **45**(3): p. 453-66.
297. Mor, A., et al., *Influenza virus mRNA trafficking through host nuclear speckles.* Nat Microbiol, 2016. **1**(7): p. 16069.
298. Bello, L.J., et al., *The human herpesvirus-8 ORF 57 gene and its properties.* J Gen Virol, 1999. **80 (Pt 12)**: p. 3207-3215.
299. Gordon, P.M., et al., *A conserved role for the ALS-linked splicing factor SFPQ in repression of pathogenic cryptic last exons.* Nat Commun, 2021. **12**(1): p. 1918.
300. Yamanegi, K., S. Tang, and Z.M. Zheng, *Kaposi's sarcoma-associated herpesvirus K8beta is derived from a spliced intermediate of K8 pre-mRNA and antagonizes K8alpha (K-bZIP) to induce p21 and p53 and blocks K8alpha-CDK2 interaction.* J Virol, 2005. **79**(22): p. 14207-21.

301. Ungerleider, N.A., et al., *Comparative Analysis of Gammaherpesvirus Circular RNA Repertoires: Conserved and Unique Viral Circular RNAs.* J Virol, 2019. **93**(6).
302. Ungerleider, N., et al., *Comparative analysis of gammaherpesvirus circRNA repertoires: conserved and unique viral circRNAs.* J Virol, 2018.
303. Arias, C., et al., *KSHV 2.0: a comprehensive annotation of the Kaposi's sarcoma-associated herpesvirus genome using next-generation sequencing reveals novel genomic and functional features.* PLoS Pathog, 2014. **10**(1): p. e1003847.
304. Catrina Ene, A.M., et al., *MicroRNA expression profiles in Kaposi's sarcoma.* Pathol Oncol Res, 2014. **20**(1): p. 153-9.
305. Geiger, F., et al., *Liquid-liquid phase separation underpins the formation of replication factories in rotaviruses.* EMBO J, 2021. **40**(21): p. e107711.
306. Li, X., L. Yang, and L.L. Chen, *The Biogenesis, Functions, and Challenges of Circular RNAs.* Mol Cell, 2018. **71**(3): p. 428-442.
307. Darbani, B., S. Noeparvar, and S. Borg, *Identification of Circular RNAs from the Parental Genes Involved in Multiple Aspects of Cellular Metabolism in Barley.* Front Plant Sci, 2016. **7**: p. 776.
308. Backes, S., et al., *Degradation of host microRNAs by poxvirus poly(A) polymerase reveals terminal RNA methylation as a protective antiviral mechanism.* Cell Host Microbe, 2012. **12**(2): p. 200-10.
309. Marcinowski, L., et al., *Degradation of cellular mir-27 by a novel, highly abundant viral transcript is important for efficient virus replication in vivo.* PLoS Pathog, 2012. **8**(2): p. e1002510.
310. Ruiz, J.C., O.V. Hunter, and N.K. Conrad, *Kaposi's sarcoma-associated herpesvirus ORF57 protein protects viral transcripts from specific nuclear RNA decay pathways by preventing hMTR4 recruitment.* PLoS Pathog, 2019. **15**(2): p. e1007596.
311. Jones, T., et al., *Viral cyclin promotes KSHV-induced cellular transformation and tumorigenesis by overriding contact inhibition.* Cell Cycle, 2014. **13**(5): p. 845-58.
312. Moore, P.S., *KSHV manipulation of the cell cycle and apoptosis*, in *Human Herpesviruses: Biology, Therapy, and Immunoprophylaxis*, A. Arvin, et al., Editors. 2007: Cambridge.
313. Wu, F.Y., et al., *Cell cycle arrest by Kaposi's sarcoma-associated herpesvirus replication-associated protein is mediated at both the transcriptional and posttranslational levels by binding to CCAAT/enhancer-binding protein alpha and p21(CIP-1).* J Virol, 2003. **77**(16): p. 8893-914.
314. An, H., J.T. Tan, and T.A. Shelkovnikova, *Stress granules regulate stress-induced paraspeckle assembly.* J Cell Biol, 2019. **218**(12): p. 4127-4140.
315. Yap, K., T.H. Chung, and E.V. Makeyev, *Hybridization-proximity labeling reveals spatially ordered interactions of nuclear RNA compartments.* Mol Cell, 2022. **82**(2): p. 463-478 e11.
316. Ye, F., E.R. Chen, and T.W. Nilsen, *Kaposi's Sarcoma-Associated Herpesvirus Utilizes and Manipulates RNA N(6)-Adenosine Methylation To Promote Lytic Replication.* J Virol, 2017. **91**(16).
317. Blijlevens, M., J. Li, and V.W. van Beusechem, *Biology of the mRNA Splicing Machinery and Its Dysregulation in Cancer Providing Therapeutic Opportunities.* Int J Mol Sci, 2021. **22**(10).
318. Shkreta, L., et al., *SRSF10: an atypical splicing regulator with critical roles in stress response, organ development, and viral replication.* RNA, 2021. **27**(11): p. 1302-1317.
319. Roundtree, I.A., et al., *YTHDC1 mediates nuclear export of N(6)-methyladenosine methylated mRNAs.* Elife, 2017. **6**.

320. Li, Z., et al., *Tumor-released exosomal circular RNA PDE8A promotes invasive growth via the miR-338/MACC1/MET pathway in pancreatic cancer.* Cancer Lett, 2018. **432**: p. 237-250.
321. Huang, Z., et al., *Plasma Circular RNAs hsa_circ_0001953 and hsa_circ_0009024 as Diagnostic Biomarkers for Active Tuberculosis.* Front Microbiol, 2018. **9**: p. 2010.
322. Zhou, M., et al., *CircSKA3 Downregulates miR-1 Through Methylation in Glioblastoma to Promote Cancer Cell Proliferation.* Cancer Manag Res, 2021. **13**: p. 509-514.
323. Baquero-Perez, B., et al., *The Tudor SND1 protein is an m(6)A RNA reader essential for replication of Kaposi's sarcoma-associated herpesvirus.* Elife, 2019. **8**.
324. Li, S., et al., *Screening for functional circular RNAs using the CRISPR-Cas13 system.* Nat Methods, 2021. **18**(1): p. 51-59.
325. Jackson, B.R., M. Noerenberg, and A. Whitehouse, *A novel mechanism inducing genome instability in Kaposi's sarcoma-associated herpesvirus infected cells.* PLoS Pathog, 2014. **10**(5): p. e1004098.