Nup98-mediated Regulation of Gene Expression and RNA Metabolism

by

Juliana Silva Capitanio

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Cell Biology
University of Alberta

Abstract

The nucleus is a hallmark of eukaryotic cells. It segregates the genetic material from the cytoplasm. The nucleus is encapsulated by the nuclear envelope which is perforated with nuclear pore complexes (NPCs), macromolecular structures formed by proteins termed nucleoporins (Nups). NPCs and nuclear transport factors regulate nucleocytoplasmic transport. Many Nups also function away from NPCs, within the nucleoplasm, where they associate with chromatin and other nuclear factors to regulate gene expression and RNA metabolism.

One such Nup, Nup98, has been shown to interact with chromatin to regulate gene expression in metazoan cells. To gain further insight into how Nup98 contributes to gene expression regulation in human cells, we began by identifying novel Nup98 binding partners and focusing on those with known roles in gene expression. Here, we report the identification of DHX9, a DExH/D-box helicase, as an intranuclear Nup98 binding partner. Our results show that the N-terminal domain of Nup98, containing FG/GLFG amino-acid repeats, binds to DHX9 in an RNA-facilitated manner. DHX9 interacts with Nup98 via its N- and C-terminal domains, which bind double-stranded and single-stranded RNAs, respectively. Importantly, the RNA-dependent ATPase activity of DHX9 is stimulated by its interaction with Nup98, resulting in increased DHX9-stimulated transcription of a reporter gene. Consistent with these observations, Nup98 and DHX9 bind interdependently to similar gene loci and their transcripts, regulating gene expression and RNA splicing. Based on our results, we propose that Nup98 affects gene expression and RNA metabolism by functioning as a cofactor that regulates DHX9 and, potentially, other RNA helicases in the cell.

Several viruses exploit Nups and RNA helicases during infection. Nup98 and DHX9 are among these host factors hijacked by viruses to participate in their life cycle. Thus, we explored how the Nup98-DHX9 complex, characterized as a regulator of gene expression and RNA metabolism in the host cell, can be exploited by positive-strand RNA viruses of the *Flaviviridae* family (i.e. the Hepatitis C virus and Zika virus). We also examined the importance of other host RNA-binding nuclear factors and nuclear transport factors (NTFs) in the life cycle of these cytoplasmic RNA viruses. Despite the cytoplasmic replication cycle of the Hepatitis C virus (HCV) and Zika virus (ZIKV), their positive-strand RNA genomes (+vRNA) are present in the nuclei of infected cells. NTFs dynamically transport this nuclear pool of +vRNA, and altering nucleocytoplasmic transport can affect the nuclear localization of the viral genome in infected cells. Disrupting the nuclear localization of +vRNA changes its interaction with several RNA-binding host nuclear factors (including Nup98 and DHX9), consequently affecting the viral life cycle. These data point towards a previously unknown role for the nuclear environment in the life cycle of viruses that undergo replication in the cytoplasm of the host cells.

Preface

Chapter 3 of this thesis has been previously published in: **Capitanio, J. S., Montpetit, B., and Wozniak, R. W. (2017). Human Nup98 regulates the localization and activity of DExH/D-box helicase DHX9.** ***eLife*****, 6**. I was responsible for performing the described experiments and data analysis. I composed an initial draft of the manuscript. Dr. Richard W. Wozniak was involved with concept formation, manuscript composition, and data analysis. Editing of the final manuscript was performed by all three authors.

The research described in chapter 4 was also performed under Dr. Wozniak's supervision. I was responsible for the described experiments and their analysis. Dr. Zhongjing Su and Brett Roughead assisted with experiments and analysis, when indicated in the corresponding figures. Dr. Wozniak and I conceived of the project based on preliminary observations made by Dr. Christopher J. Neufeldt. Conceptual input was provided by our collaborators Dr. Ben Montpetit (UC Davis, previously at University of Alberta) and Dr. Lorne J. Tyrrell (University of Alberta).

"The important thing is to not stop questioning. Curiosity has its own reason for existence. One cannot help but be in awe when he contemplates the mysteries of eternity, of life, of the marvelous structure of reality. It is enough if one tries merely to comprehend a little of this mystery each day."

Albert Einstein

In: "Old Man's Advice to Youth: 'Never Lose a Holy Curiosity.'" LIFE Magazine (2 May 1955) p. 64

For my husband, family, and friends.

Acknowledgements

I am grateful to many who have played indispensable roles through my graduate studies. There have been a number of individuals that have provided the guidance and support allowing the completion of this thesis. First and foremost, I am thankful to Professor Richard W. Wozniak for giving me the opportunity to be a part of his research group and study under his supervision. Dr. Wozniak's enthusiasm for science and drive for excellence in research are an inspiration for all in his laboratory. I am grateful for all the knowledge and support he has given me throughout my PhD. I can truly say Dr. Wozniak has had a large impact in my life. Being his student has changed the way I think about science and research. He has helped me through many milestones during my graduate program and molded the student I am today. I hope his example of perseverance, attention to detail, ethics, and success will always remain a positive influence in my future. I would also like to thank all of the members of the Wozniak Lab; their personal and professional support has greatly enriched my education and experience as a graduate student. I owe any successes during my studies to the thorough critique, feedback, and insights on my work provided by these invaluable colleagues. Especially, I would like to thank Dr. Zhongjing Su, Dr. Christopher J. Neufeldt, and Brett Roughead for their direct collaborations to part of the work presented here. It has been a pleasure working with all of them.

Part of the work presented here was done in collaboration with Professor Ben Montpetit, whose expertise, perspicacity, and criticism were very valuable. I am very thankful for the opportunity to have worked with Dr. Montpetit and members of his lab. This work was also influenced by the contributions of Professor Lorne J. Tyrrell and his research group, whose support with insights and materials facilitated the development of this research.

I am also thankful to the members of my graduate committee, Professors Paul Melançon and Mark Glover. Their feedback and encouragement during these years have been much appreciated, and a great addition to my education. I'm grateful to have had their contribution to my development as a graduate student.

I appreciatively acknowledge the funding received during my graduate studies from the Department of Cell Biology, the Faculty of Medicine and Dentistry, the Faculty of Graduate Studies and Research and Alberta Innovates Health Solutions. I'm thankful to have been deemed worthy of receiving scholarships and awards from

these institutions during my PhD program. This support, along with the support provided by Dr. Wozniak, has made this work possible.

Finally, the unwavering support of my family and their constant encouragement have been essential during this endeavor. My greatest admiration goes to my parents and sister. I am extremely fortunate to have such extraordinary people in my life. I'm thankful to my parents, whose love and care have given me all I have in life and made me who I am, as well as to my sister, whose constant friendship has always been most valuable. I have no word to thank them for everything they have done for me. Nothing is enough to express the importance they have in my life. I would also like thank my husband. I am very fortunate to have met someone so supportive, caring and loving to share my life. He has my greatest love and admiration as a scientist and as a person. Without him by my side, I would not be who I am. His constant care and love through all these years continue to inspire me to grow and improve.

Contents

List of Tables

List of Figures

List of Symbols

3D Three dimentional

aa or AA amino acid residue

ACE1 Ancestral Coatomer Element 1

Ago2 Argonaute 2

APC anaphase promoting complex

ApoB apolipoprotein B

ApoE apolipoprotein E

Arg arginine

ATCC American Type Culture Collection

ATM Ataxia Telangiectasia Mutated

ATR Ataxia Telangiectasia And Rad3-Related Protein

CBP cAMP response element-binding (CREB) binding protein

CDC Centers for Disease Control and Prevention

Cdk1 Cyclin-dependent kinase 1

cNLS classical NLS

CRM1 chromosomal maintenance 1

Ct Threshold cycle

CTE constitutive transport element

DAPI 4',6-diamidino-2-phenylindole

DMEM Dubelcco's modified Eagle's media

DMSO Dimethyl sulfoxide

DMVs double membrane vesicles

DNA-PK DNA-dependent protein kinase

dsRBDs double stranded RNA-binding domains

dsRNA Doubled stranded RNA

EGFR EGF receptor

ELAVL1 ELAV (Embryonic Lethal, Abnormal Vision, Drosophila)-Like 1

EMCV encephalomyocarditis virus

ER endoplasmic reticulum

Erk extracellular signal regulated kinase

ESCRT endosomal sorting complexes required for transport

FBS Fetal bovine serum

FG phenylalanine-glycine

FISH Fluorescence in situ hybridization

FMDV foot-and-mouth disease virus

FOXO1 Forkhead box O1

GDP Guanosine diphosphate

GFP Green fluorescent protein

GLFG glycine-leucine-phenylalanine-glycine

GST Glutathione S-transferase

GTP Guanosine triphosphate

HCV Hepatitis C virus

HEAT Huntingtin elongation factor 3 protein phosphatase 2A target of rapamycin 1

HIV LTR HIV long terminal repeat

HIV TAR HIV trans-activation response element

HLA-DRA major histocompatibility complex, class II, DR alpha

hnRNP U Heterogeneous Nuclear Ribonucleoprotein U

HRP Horseradish peroxidase

HuR Hu Antigen R

IFN-γ interferon gamma

IgG Immunoglobulin G

INM inner nuclear membrane

IPZ Importazole

IRES internal ribosome entry site

Iver or IVM Ivermectin

Kapβ1 Karyopherin (importin) beta 1

Kapβ3 Karyopherin (importin) beta 3

Kaps karyopherins

KASH Klarsicht, ANC-1, Syne Homology

KTs kinetochores

LARP6 La ribonucleoprotein domain family member 6

LB Luria Broth

LINC linker of nucleoskeleton and cytoskeleton

LMB Leptomycin B

LMX1B LIM homeobox transcription factor 1 beta

Lys lysine

m6A N6-adenosine methylation

MBD2a methyl-CpG binding domain protein 2

MBNL1 muscleblind 1

MCAK mitotic centromere-associated kinesin

MDR1 multidrug resistance gene 1

MEF1 MDR1 promoter-enhancing factor

mESCs mouse embryonic stem cells

miRNA microRNA

MLE Maleless

MMVs multi membrane vesicles

MOI Multiplicity of infection

MPMV Mason-Pfizer monkey virus

mRNPs messenger ribonucleoprotein particles

MSL Male Specific Lethal

MTAD minimal transactivation domain

MTP microsomal triglyceride transfer protein

NA numerical aperture

NAs nucleic acids

NDHII Nuclear DNA Helicase II

NE nuclear envelope

Neks NimA related kinases

NES nuclear export signal

NF110 nuclear factor 110

NLS nuclear localization signal

NPCs nuclear pore complexes

NS3 Non-structural protein 3

NSL nonspecific lethal

NTF2 Nuclear transport factor 2

NTFs nuclear transport factors

Nups nucleoporins

NXF1 nuclear RNA export factor 1

NXT1 NTF2-related export protein 1

ONM outer nuclear membrane

p.i. post infection

PBS Phosphate buffered saline

PCE post-transcriptional control element

PCK1 phosphoenolpyruvate carboxykinase 1

PCR Polymerase chain reaction

PIC pre-integration complex

Pom pore membrane domain

PRC1 polycomb-repressive complex 1

PRKDC Protein Kinase, DNA-Activated, Catalytic Polypeptide

Ran RAs-related Nuclear protein

RanGAP1 Ran GTPase-activating protein 1

RanGEF Ran guanine nucleotide exchange factor

RBD Rae1-binding domain

RCC1 Regulator Of Chromosome Condensation 1

RFs replication factories

RHA RNA helicase A

RISC RNA-induced silencing complex

RNA Pol II RNA polymerase II

RNP ribonucleoprotein

RPMI Roswell Park Memorial Institute 1640

RRE Rev response element

shRNA short hairpin RNA

SMN survival motor neuron

snRNPs small nuclear ribonucleoproteins

SSB Sjogren Syndrome Antigen B (Autoantigen La)

SSC Saline sodium citrate

SUN Sad1p, UNC-84

TAM Tyro3, Axl, and Mer

TAP tip-associated protein

TCP80 translation control protein 80

TIM-1 T cell immunoglobulin and mucin domain 1

TonEBP TonE-binding protein

TPR tetratricopeptide repeat

TRBP transactivation responsive RNA binding protein

VLDL very low density lipoprotein

-vRNA negative strand viral RNA

VSV Vesicular Stomatitis virus

XPO1 exportin 1

ZIKV Zika virus

CHAPTER 1

Introduction

1.1 Preface

In eukaryotic cells, cellular processes are compartmentalized. This compartmentalization allows for local concentration of factors while decreasing interference between cellular activities, consequently increasing the efficiency of individual cellular processes. Cellular compartmentalization also increases organizational complexity, facilitating the formation of multicellular organisms.

The nucleus is a hallmark organelle of eukaryotic cells. It is encapsulated by the nuclear envelope (NE), which segregates the genetic material from the cytoplasm of the cell. The NE is perforated with nuclear pore complexes (NPCs), which are selectively permeable channels between the nucleus and cytoplasm. NPCs, coupled with nuclear transport factors (NTFs), function as gateways regulating macromolecule transport across the NE. Situated at the interface between the nucleus and the cytoplasm, NPCs are positioned within the cell to participate in a vast number of cellular processes, including the regulation of mitosis, organization of the genome, gene expression and RNA metabolism (reviewed in Ptak and Wozniak, 2016; Wozniak et al., 2010; Ibarra and Hetzer, 2015; Burns and Wente, 2014). In higher eukaryotes, proteins known as nucleoporins (the building blocks of the NPC) also take leave of the NE and function within the nucleoplasm, in association with chromatin and other nuclear factors, to regulate gene expression and RNA metabolism.

Given the wide range of cellular functions performed by Nups, NPCs, and NTFs, it is not surprising that defects in this machinery have been implicated in a broad variety of human diseases, including cancer and viral infections. Understanding how nucleoporins (Nups) and NTFs interface with components of other cellular pathways will undoubtedly shed light on the mechanisms of numerous cellular processes while providing insight into disease states influenced by NPC-associated processes (reviewed in Nofrini et al., 2016; Dickmanns et al., 2015). An improved understanding of the communication between NPCs and the cellular processes they influence has the potential to transform the approach to diverse areas of biomedical research, including cancer, immunity, viral infections, aging and regenerative medicine.

1.2 The nuclear envelope

An important feature of eukaryotic cells is their complex intracellular organization and compartmentalization. This compartmentalization is achieved through membrane-bound organelles that allow for the spatial separation of cellular processes, while simultaneously leading to the local concentration of factors involved in a given cellular activity. The nucleus is the most prominent of such organelles

in eukaryotic cells. It encapsulates chromatin in a double-phospholipid bilayer with selective permeability termed nuclear envelope (NE) (Hetzer, 2010).

The NE forms the interface between the nucleoplasm and the cytoplasm of the cell. At the cytoplasmic face of the NE is the outer nuclear membrane ONM. This membrane is continuous with the endoplasmic reticulum (ER) and it is functionally and biochemically similar to this organelle.

The nucleoplasmic face of the NE is formed by the inner nuclear membrane INM. The INM contains several unique integral membrane proteins that allow it to interact with the underlying chromatin and form structural links to the cellular cytoskeleton (Hetzer et al., 2005).

In some eukaryotes, including human cells, the nucleoplasmic face of the INM is also lined by a meshwork of filamentous proteins (lamins) that form the nuclear lamina. The nuclear lamina can interact with INM proteins (such as the lamin B receptor and lamina-associated polypeptides 1 and 2) and with the underlying chromatin. The interaction of the nuclear lamina with chromatin is involved in a myriad of nuclear processes, like gene expression, DNA damage repair, development, cell differentiation, cell proliferation, and cellular aging (Burke and Stewart, 2013). The INM proteins interact with chromatin not only through their association with the lamina. INM proteins containing LEM domains (such as Lap2, Emerin, and MAN1) can also bind DNA directly or interact with other chromatin-associated factors (Hetzer, 2010).

In addition to binding chromatin, INM proteins are linked to the cytoskeleton by the linker of nucleoskeleton and cytoskeleton complexes (LINC). These complexes are formed by the interaction of SUN and KASH domain-containing proteins. SUN proteins are INM localized. They interact with the underlying chromatin and lamina and traverse the perinuclear space to interact with ONM proteins containing KASH domains. KASH domain proteins at the ONM interact with the cellular cytoskeleton. This molecular bridge between the cytoskeleton and the nuclear interior maintains NE structure and provides a unique link between signaling pathways in the cytoskeleton and gene regulation (Rothballer and Kutay, 2013).

The ability of the nuclear envelope to separate the nucleoplasmic and cytoplasmic compartments, while maintaining selective transport between the two, has allowed eukaryotic cells to evolve a highly complex system of gene expression regulation not available to prokaryotes. Transport through the NE occurs throughout its surface, where the ONM and the INM fuse to form aqueous channels that harbor macromolecular protein structures called nuclear pore complexes. These channels allow selective transport between the cytoplasm and nucleoplasm (Knockenhauer and Schwartz, 2016).

1.3 The nuclear pore complex

The nuclear pore complex is composed of approximately thirty different proteins termed nucleoporins (Nups) arranged in octagonal symmetry. These Nups form three concentric rings: central, nucleoplasmic and cytoplasmic (von Appen and Beck, 2016). Specific nucleoporins, containing phenylalanine-glycine repeats, line the central channel of the NPC creating a selective barrier across the pore. Small water-soluble molecules can passively diffuse through the NPC. However, most cellular proteins are too large for diffusion and must be actively transported through the NPC with nucleocytoplasmic transport factors (Musser and Grünwald, 2016).

While the function of Nups in NPC structure and nuclear transport has been well established, several observations indicate that nucleoporins also function outside of NPCs in the cytoplasm and nucleoplasm (Raices and D'Angelo, 2012). NPCs not only control nucleocytoplasmic transport, but also influence gene expression, chromatin organization, chromosome inheritance, establishment of epigenetic maintenance, DNA repair, and cell cycle regulation (Ibarra and Hetzer, 2015).

Deciphering the structure of the nuclear pore complex has been a significant achievement in the field of structural biology. However, recent evidence indicates that despite high conservation in its overall structure, not all NPCs contain the same protein composition. NPC composition can vary in different cell types and tissues, increasing the diversity of their functions and properties (Raices and D'Angelo, 2012). Therefore, the description of the NPC provided below paints a general picture of a metazoan NPC, and thereby provides a framework for comprehending the overall principles governing the structure-function relationship of NPCs and nuclear trafficking.

1.3.1 Nuclear pore complex structure

The eukaryotic NPC is a massive cylindrical structure that spans the inner and outer membranes of the NE. It ranges from a molecular mass of approximately 125 MDa in vertebrates to 66 MDa in yeast (Reichelt, 1990 and Rout, 1993). Longitudinally, eight interconnected, symmetrical spokes surround a central channel (Gall, 1967). As mentioned, the organization of the NPC can be subdivided into three concentric regions. The outermost region of the NPC faces the cytoplasm of the cell and it consists of the cytoplasmic ring and protruding, flexible, cytoplasmic filaments. The central spoke region of the NPC harbors the central transport channel and allows the anchoring of the NPC to the NE membrane. The innermost portion of the NPC faces the nucleoplasm of the cell. This portion is composed of a nucleoplasmic ring

and rigid nucleoplasmic filaments that form the nuclear basket. The human NPC is approximately 85 nm in length, excluding cytoplasmic filaments (Figure 1.1). It presents an hourglass morphology, with the inner diameter of the central channel being around 50 nm and expanding to circa 90 nm at the nucleoplasmic and cytoplasmic rings (Maimon et al., 2012).

The NPC is composed of approximately 30 unique nucleoporins. Each Nup is usually present in two or four copies per NPC spoke, and therefore a typical NPC will have 400 to 1000 proteins (Cronshaw et al., 2002; Alber et al., 2007). Nups are organized in subunits within the NPC, based on location and function (Knockenhauer and Schwartz, 2016).

The first group of Nups contains the integral membrane proteins that facilitate anchoring of the NPC to the NE membrane, generally referred to as pore membrane domain (Pom) proteins. The Poms participate in pore assembly and connect the NE to the NPC core scaffold proteins. The core scaffold proteins are the next group of Nups, forming and maintaining the membrane curvature of the NPC and its structural scaffold. Core scaffold nucleoporins are among the most stable and long-lived components of the NPC and often contain domains involved in curving membranes (α-solenoid and β-propeller folds), homologous to those found in vesicle coat proteins (Brohawn et al., 2008; Field and Dacks, 2009). The central channel of the NPC is lined with Nups containing unstructured phenylalanine-glycine (FG) repeat domains that create the selective permeability barrier between the cytosol and nucleoplasm (Figure 1.1) (Li et al., 2016a). The FG-Nups are highly dynamic, and they mediate interactions between the NPC and transport proteins to facilitate the active transport of large cargoes across the NE, they also serve as platforms for numerous NPC-associated proteins (Wente and Rout, 2010).

In metazoan cells the function and localization of nucleoporins can also be altered by post-translational modifications. For example, O-linked glycosylation and phosphorylation of FG-Nups by Erk can affect nucleocytoplasmic transport (Labokha et al., 2013; Kosako et al., 2009). Moreover, Cyclin-dependent kinase 1 (Cdk1) and NimA related kinases (Neks) phosphorylate several Nups triggering the global dismantling of NPCs and the NE at the onset of mitosis (Laurell et al., 2011; Glavy et al., 2007). Furthermore, Nups can be mono- and polyubiquitinated in a cell cycle-specific manner; Nup ubiquitination leads to changes in their protein level, which can regulate cell cycle progression and phase-specific gene expression (Chakraborty et al., 2008). It is highly likely that much remains to be discovered about the post-translational regulation of Nups and NPCs.

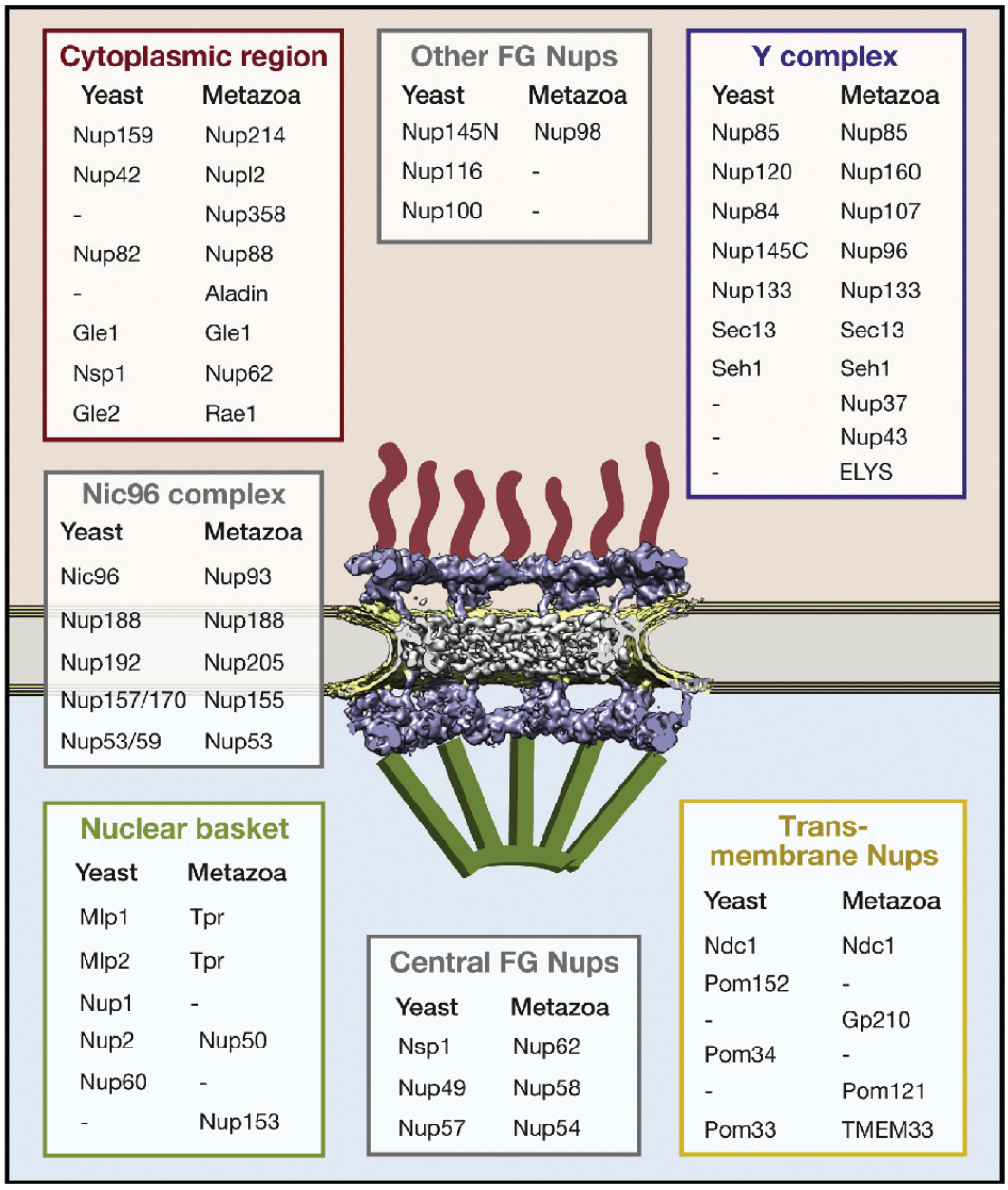


Figure 1.1: Diagram of Nup positions in the NPC.

Diagram of the human NPC shown in a cut-away view displaying half of an NPC embedded in the NE (shown in yellow). Nups from yeast and metazoa are listed and color-matched according to their approximate positions within the NPC. Adapted from (Knockenhauer and Schwartz, 2016). Reprinted with permission from Cell Press, original publication DOI: http://dx.doi.org/10.1016/j.cell.2016.01.034

1.3.2 Pore Membrane nucleoporins

The Pom proteins facilitate anchoring of the NPC to the NE and, also contribute to NPC assembly and nucleocytoplasmic transport. The Poms identified in mammalian cells include gp210, NDC1, and Pom121 (Gerace et al., 1982; Hallberg et al., 1993; Mansfeld et al., 2006; Stavru et al., 2006).

The Pom gp210 has a single transmembrane domain containing a short N-terminal region that extends into the pore and can interact with the core scaffold nucleoporins (Greber et al., 1990; Wozniak and Blobel, 1992; Wozniak et al., 1989). Pom121 also has a single transmembrane portion, along with a larger, domain that extends into the interior of the NPC (Hallberg et al., 1993; Söderqvist and Hallberg, 1994). The third vertebrate Pom, NDC1, has six predicted transmembrane domains containing three luminal loops and a NPC associated C-terminal domain (Mansfeld et al., 2006; Stavru et al., 2006) (Figure 1.2). The different Pom proteins appear to have significant functional redundancy, and they play a major role in NPC assembly. NDC1 has been shown to participate in NPC assembly and maintenance of NPC structure. Similarly, Pom121 has also been linked to NPC biogenesis (Mansfeld et al., 2006; Funakoshi et al., 2011; Mitchell et al., 2010).

In mammalian cells, NPC assembly can occur during NE reformation after mitosis. NPC assembly after mitosis occurs via recruitment of ER membranes and disassembled NPC components to the chromatin, allowing simultaneous NE reformation and NPC assembly (Schooley et al., 2012). This post-mitotic NPC assembly involves a stepwise association of Nups that is initiated by recruitment of the Nup107-160 complex and formation of a pre-pore (Dultz et al., 2008; Harel et al., 2003; Walther et al., 2003). Assembly of this pre-pore is followed by the recruitment of Pom121 and NDC1 that leads to the binding of the linker complexes and the successive association of the remaining FG-repeat and peripheral Nups (Antonin et al., 2008, 2005; Hawryluk-Gara et al., 2008; Mitchell et al., 2010).

NPC assembly can also occur *de novo* during interphase. NPC assembly during interphase begins with the recruitment of Poms (Doucet et al., 2010; Dultz and Ellenberg, 2010; Funakoshi et al., 2011; Talamas and Hetzer, 2011). Recruitment of Poms is followed by the rapid, synchronous association of the core scaffold Nup complexes, leading to the recruitment of FG-repeat and peripheral Nups (Dultz and Ellenberg, 2010).

The two main theories for NPC biogenesis postulate that NPCs are either assembled directly into regions of the NE (devoid of NPCs) or into lipid membranes of the ER that subsequently fuse to the NE. There is evidence to support both hypotheses, so cells might utilize multiple mechanisms to generate new NPCs. The NE has been shown to expand during interphase, maintaining regions devoid of NPCs

(Maeshima et al., 2006). NPC assembly occurs in these pore-free islands, from both sides of the NE, presumably from pre-assembled Nup subcomplexes (D'Angelo et al., 2006). This process is dependent on RanGTP and Cdk1/Cdk2 activity, and it is negatively regulated by importin-β (D'Angelo et al., 2006; Harel et al., 2003; Maeshima et al., 2010; Walther et al., 2003). This pathway of NPC assembly is surveilled by ESCRT-III/Vps4, which clears defective NPCs, ensuring the fidelity of global nucleocytoplasmic transport (Webster et al., 2014). NPC-containing arrays of stacked ER-derived membranes, termed annulate lamellae, also occur within the cytoplasm of many cell types, especially embryonic and transformed cells (Kessel, 1992). It is possible that annulate lamellae represent NPC assembly factories or storage centers for improperly assembled or defective NPCs.

1.3.3 Core Scaffold nucleoporins

The scaffold nucleoporins form the characteristic three-ring architecture of the NPC. These Nups are enriched for coiled-coil, α-helical repeats, and β-propeller domains, allowing for the formation of extensive protein-protein interactions (Figure 1.2).

A significant portion of the NPC scaffold is built from β-propeller domains that function in the structural strengthening of the NPC (Figure 1.2). These β-propeller domains are 4 to 8 bladed sheets arranged around a central axis. Each β-sheet has anti-parallel strands that twist to allow interaction between the hydrophobic faces. These interactions create rigid structures that strengthen the final protein structure, and consequently the structure of the NPC (Paoli, 2001; Dickmanns et al., 2015) .

The α-helical domains comprise over half of the NPC scaffold mass. Many α-helical domains are found organized in α-helical solenoids, hairpin-like structures with two or three α-helical repeats stacked to form a superhelical domain (Figure 1.2). The α-solenoids are often found in combination with β-propellers in large protein assemblies. The Nups α-helical domains can also be found in the Ancestral Coatomer Element 1 (ACE1), a structural manifestation only common to Nups and components of COPII vesicle coat, indicating that the NPC and coated vesicles may share a common evolutionary origin (Devos et al., 2004; Dickmanns et al., 2015).

Scaffold Nups interact extensively with each other creating lattices that form the basic structural framework of the NPC (Figure 1.1). They are essential for NPC assembly and anchoring at the NE (Theerthagiri et al., 2010). The central scaffold of the NPC is composed of two major sub-complexes. The first, Nup107-Nup160 sub-complex, also known as the Y-complex, is formed by 9-10 proteins, Nup160, 133, 107, 96, 85, 43, and 37, Seh1, Sec13, and ELYS. The second, Nup93 sub-complex, contains five nucleoporins, with three large structural nucleoporins

(Nup155, Nup188 and Nup205), and two linker nucleoporins (Nup93 and Nup53), whose short linear motifs provide a bridge between the large structural Nups (Dickmanns et al., 2015) (Figure 1.2).

1.3.4 Cytoplasmic filaments and nuclear basket

The cytoplasmic face of the NPC is formed by the Nup214 complex. Nup358 is part of this complex, and the largest nucleoporin in the mammalian NPC. It can interact with single-stranded RNA through three highly positively charged TPR motifs. Nup358 also has four RanGTP binding domains, eight consecutive zinc finger motifs, an E3 ligase domain, and a C-terminal domain with homology to cyclophilin A. These domains are inter-spaced by non-structured regions containing FG-repeats (Figure 1.2). Nup358 is thought to localize near Nup214 in the mammalian NPC (Dickmanns et al., 2015). Nup214 is anchored to the NPC by its N-terminal propeller and a central domain with predicted coiled-coils (Figure 1.2). These regions interact with the Y-complex in the cytoplasmic ring structure (Paulillo et al., 2005; Bui et al., 2013). The coiled-coil region of Nup214 binds to Nup88. Nup88 not only interacts with Nup214, but also can bind lamin A and Nup98 (Fornerod et al., 1997; Lussi et al., 2011; Griffis et al., 2003). Nup98 will be discussed in more detail in upcoming sections. Its N-terminal FG-repeat region contains a Rae1 interaction domain that can dock mRNA export factors NXF1 and Rae1 to the NPC (Blevins et al., 2003).

The nucleoplasmic face of the NPC harbors the nuclear basket. The nuclear basket is formed by eight filaments that emanate from the nuclear ring of the NPC and converge into a distal ring (Figure 1.1). In mammalian NPCs, the nuclear basket is formed by Nup50, Nup153, and Tpr. Tpr contains a large coiled-coil domain and an acidic globular C-terminal (Figure 1.2). It can form homodimers, thought to make up the filaments of the nuclear basket. Nup153 contains three domains, and its N-terminal targets it to the NPC, anchoring it to the nuclear ring. The central domain of Nup153 has four zinc-fingers that allow its interaction with the distal ring of the nuclear basket. Its C-terminal contains FG-repeats. This type of domain will be further discussed in the next section (Figure 1.2). At the nuclear basket, Nup153 interacts with Tpr and Nup50. Nup50 is a mobile nucleoporin that associates with the nuclear basket through its direct binding to Nup153, showing no interaction with Tpr (Duheron et al., 2014).

1.3.5 FG-repeat Nups

The FG-repeat Nups are characterized by the presence of 4-48 FG, GLFG, and/or FXFG-repeats dispersed in long unfolded domains (Zeitler and Weis, 2004; Denning et al., 2003; Radu et al., 1995) (Figure 1.2). The FG-repeats of these nucleoporins form a network of natively unfolded domains that fills the central channel of the NPC. This FG-repeats network also protrudes into the cytoplasm and nucleoplasm, and it appears thinner towards the center of the NPC (Alber et al., 2007; Lemke, 2016).

FG-repeat Nups lining the central channel of the NPC maintain the selective permeability of the nuclear pore by passively forming a permeability barrier that can block access to macromolecules, unless they are associated with the soluble nuclear transport factors (NTFs) that can interact with FG-repeats to facilitate the movement of cargoes (Wente and Rout, 2010; Lemke, 2016).

Though the components that make the FG-repeat Nups are well defined, the exact organization of these domains and the way they interact with NTFs to facilitate transport are still controversial. The FG-repeat domains can also change their localization within the NPC during NTF-cargo translocation across the central channel (Paulillo et al., 2005). There are many models of how FG-Nups can mediate cargo translocation through the NPC, and these models will be discussed in more detail in subsection 1.4.3.

The FG-Nups can be found asymmetrically distributed in the previously discussed cytoplasmic filaments and nuclear basket of the NPC, where they provide docking sites for specific transport and regulatory proteins associated with the nuclear pore. FG-Nups are also part of the Nup62 subcomplex, composed of symmetrically distributed FG-repeat nucleoporins (Nup54, Nup58, Nup62, and Nup98) that line the central NPC channel (Figure 1.1). Recent models for NPC transport suggest that these symmetrical FG-Nups (including Nup54, Nup58 and Nup62), along with the linker and scaffold Nups, form a symmetrical core that can facilitate reversible dilation of the transport channel (Melcák et al., 2007; Solmaz et al., 2013, 2011). This reversible alteration in NPC structure could allow the NPC to accommodate transport of large macromolecules, while preventing concomitant bidirectional leakage of smaller proteins (Solmaz et al., 2013).

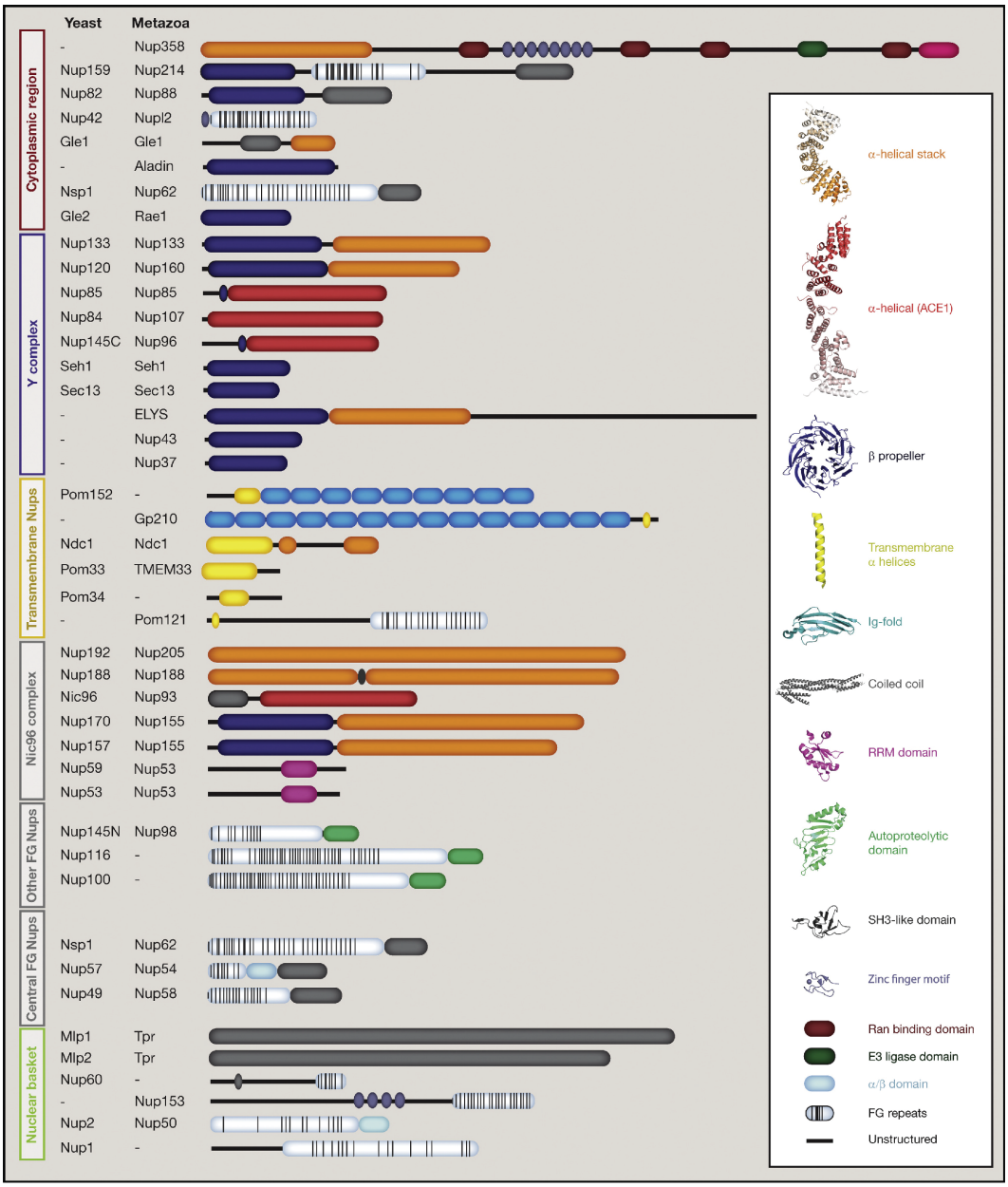


Figure 1.2: Domain Architecture of Nucleoporins.

Metazoan nucleoporins with their yeast homologs are listed. The domain architecture is derived primarily from X-ray crystallographic data, combined with structure prediction whenever experimental data are not yet available. The vast majority of nucleoporins are built from a limited set of structural modules, which characteristically occur in multiple proteins. Adapted from Knockenhauer and Schwartz, 2016. Reprinted with permission from Cell Press, original publication DOI: http://dx.doi.org/10.1016/j.cell.2016.01.034

1.4 Nucleo-cytoplasmic transport

The NPC can maintain selective permeability between the nucleus and cytoplasm by being simultaneously a barrier and an interactor for cargoes. Early studies using dextran beads of varying diameters have shown that the NPC has a size exclusion barrier, blocking passive diffusion of molecules with approximately 40 kDa, or 5 nm diameter (Feldherr and Akin, 1997; Keminer and Peters, 1999). Molecules above this size threshold have to overcome this transport barrier by interacting, either directly with the NPC or with soluble NTFs (Fried and Kutay, 2003; Güttler and Görlich, 2011; Pemberton and Paschal, 2005).

The NPC allows active transport of an assortment of cargoes through a variety of mechanisms, including the transport of numerous RNA species and proteins, that can reach up to a diameter of 39 nm (Feldherr et al., 1984; Panté and Kann, 2002; Wente and Rout, 2010). The NPC can translocate around 1000 macro-molecules per second, accommodating a mass flow of nearly 100 MDa/second (Ribbeck and Görlich, 2001).

The nuclear transport machinery consists of roughly 80 distinct proteins, which can be separated into NPC elements (Nups), soluble NTFs, and the RanGTP/RanGDP system. The Ran system provides energy and directionality to the transport process (Güttler and Görlich, 2011), see Figure 1.3. Though considerable progress has been made in identifying the components necessary for NPC-mediated transport, there is still no unifying model for the process of transport.

1.4.1 Nuclear transport factors and their cargoes

Cargoes destined for transport across the NPC contain nuclear localization signals (NLS) or nuclear export signals (NES) that specify the directionality of transport as either nuclear import or export, respectively. These NLS or NES sequences are sufficient for transport of proteins across the NE, as demonstrated by increased nuclear localization of soluble cytoplasmic proteins after addition of specific NLS sequences (Goldfarb et al., 1986).

In general, NLS sequences show stretches of basic residues, with variable amino acid residue (aa) composition in different cargo proteins. Some cargo proteins may contain a classical NLS (cNLS) sequence, of five residues (Lys-Lys-Lys-Arg-Lys), which is sufficient for transport. Other cargoes require a bipartite NLS sequence, with two regions of basic residues separated by ten amino acid residues (Dingwall et al., 1988; Goldfarb et al., 1986; Kalderon et al., 1984). In contrast to import cargoes, nuclear export cargo NES sequences generally contain short stretches rich

in leucine amino acid residues (Wen et al., 1995).

Most nuclear transport factors are homologous proteins known as importins and exportins, collectively identified as karyopherins (Kaps). Each Kap recognizes a specific NLS or NES, or interacts with its cargo indirectly via adapter molecules. There are at least 22 known Kaps in humans (Matsuura, 2016; Soniat and Chook, 2015).

Structurally, NTFs fall into three classes. The most prominent one is the karyopherin β superfamily. It is commonly made up of HEAT repeats, composed of several helix-turn-helix motifs packed side by side, forming elongated molecules with a superhelical twist. In Kaps, the HEAT repeats stacked arrangement results in an overall superhelical shape, with the A-helices located at the outer convex surface of the molecule and the B-helices lining the inner concave side (Dickmanns et al., 2015).

Crystal structures of Kapβ cargo complexes demonstrate marked differences in the way Kaps recognize substrates destined for import (Cingolani et al., 1999, 2002; Lee et al., 2003, 2005). This flexibility in cargo recognition allows for the import of a broad range of different cargoes (Lee et al., 2005). Kapβ1 can also interact with the adaptor protein Kapα (Cingolani et al., 1999). The domain of Kapα that mediates its interaction with Kapβ contains a short sequence of basic amino acid residues resembling a cNLS (Görlich et al., 1996a). Alpha importins (Kapα) also recognize NLS-containing cargoes, but they require Kapβ to mediate their interaction with the FG-repeats lining the NPC. The use of Kapα as an adaptor in cargo recognition, interacting with Kapβ, defines the classical nuclear import system (Lange et al., 2007). Other importins, including Kapβ3, do not require adaptor proteins, but rather they interact with both cargo proteins and FG-Nups directly. For nuclear export, the majority of export cargoes associate with the exportin CRM1 (also known as XPO1), although at least eight different exportins have been identified in metazoans. Despite the variety of nuclear transport factors present in mammalian cells, the nuclear transport pathway displays much redundancy. Several cargoes can be shuttled by multiple importins and exportins. Additionally, CRM1 and Kapα/β complexes can recognize a variety of unrelated cargoes (Wente and Rout, 2010).

The second family of NTFs is involved in mRNA export. The mRNA export receptor is formed by a heterodimer of two proteins, NXF1 (also known as TAP) and NXT1. NXF1 is a multidomain protein. Its N-terminal contains an NLS, an RNA-binding domain, and a leucine-rich repeat domain. The C-terminal half of NXF1 harbors two domains, one with structural similarity to NTF2 and NXT1, and the other predicted to have a ubiquitin-associated fold. The NTF2-like domains of NXF1 and NXT1 are composed of a six-stranded, highly twisted β-sheet, shielded on one side by two of its helices and sandwiched on the other side to the β-sheet of the

other subunit of the hetero-dimer (Dickmanns et al., 2015).

NTF2, the nuclear import receptor for RanGDP, is the characteristic member of the third family of NTFs and it has distinct binding sites for Ran and FG-Nups. NTF2 shows striking overall structural similarity to NXF1, however, NTF2 forms homo-dimers (Dickmanns et al., 2015).

Since multiple hydrophobic patches on NTFs can interact with FG-repeat domains within the NPC channel, it is assumed that NTFs travel through the NPC by transient low-affinity contacts to the numerous FG-repeats present in the transport channel (Tetenbaum-Novatt et al., 2012). On the other hand, these multiple strong avidity interactions between NTFs and FG-Nups, which ensure NPC transport selectivity by enhancing stability and specificity, may also contradict the rapid (2.5 to 5 ms) dwell times of nucleocytoplasmic transport cargoes.

A recent idea proposes that the NPC transport selectivity and speed are regulated by the occupancy of Kaps in the pore. By analogy, we can imagine the NPC as a "dirty velcro", where Kap adhesion to the FG-repeats network is reduced as more soluble Kap molecules occupy it. This "Kap-occupancy" can determine transport rates by impacting the diffusion volume inside the NPC, and the binding and transport kinetics of subsequent Kaps within the channel. Therefore, populating the NPC with strongly bound Kaps may reinforce barrier functionality, while providing for a finely tuned micro-environment that facilitates transport selectivity and speed. Changes in the expression level or availability of Kaps, may serve to regulate nucleocytoplasmic transport by fine-tuning the NPC microenvironment according to the functional needs of the cell (Lim et al., 2015).

1.4.2 Transport directionality

Nuclear transport can be referred to as facilitated, rather than active transport, since import of cargo through the NPC is thought to be a passive process, with energy expenditure required only for the dissociation of NTF-cargo complexes in the nucleus (Stewart et al., 2007; Güttler and Görlich, 2011). In the classical model of nuclear transport, the vectorial movement of Kapα/β cargo complexes through the NPC is driven by the affinity of Kapβ for the nucleotide-bound state of Ran, which cycles between GDP- and GTP-bound forms (Figure 1.3). During nucleocytoplasmic transport, the small GTPase Ran regulates binding between NTFs and their cargoes.

In the nucleus, the Ran guanine nucleotide exchange factor (RanGEF), RCC1, is physically tethered to chromatin through interactions with histones H2A and H2B, restricting the localization of RanGTP to the nuclear interior (Ohtsubo et al., 1989; Bischoff and Ponstingl, 1991; Nemergut et al., 2001). Conversely, Ran GTPase-

activating protein 1 (RanGAP1) is physically tethered to Nup358 on the NPC cytoplasmic filaments (Matunis et al., 1996; Mahajan et al., 1997; Saitoh et al., 1997), restricting hydrolysis of Ran-bound GTP to GDP to the cytoplasm. By physically separating the activities of RanGEF and RanGAP, a gradient of RanGTP is established that is 100 to 1000-fold higher in the nuclei than the cytoplasm of cells (Görlich et al., 1996b; Izaurralde et al., 1997).

The export of NES-containing cargoes from the nucleus depends on their interaction with exportins and RanGTP. This complex moves through the pore to the cytoplasm. In the cytoplasm, RanGTP is hydrolyzed into RanGDP by RanGAP1 and the complex dissociates, releasing the cargo and RanGDP (Figure 1.3).

During nuclear import, importins bind their NLS-containing cargoes in the cytoplasm. This complex interacts with FG-repeats at the nuclear pore complex moving through the channel. Importins have a much higher affinity for Ran in its GTP versus GDP-bound state. Thus inside the nucleus, interaction with RanGTP causes importins to release their cargoes (Rexach and Blobel, 1995; Görlich et al., 1996b; Bischoff and Görlich, 1997; Gilchrist et al., 2002; Timney et al., 2006).

Inside the nucleus, the importin-RanGTP complex is recycled to the cytoplasm, where RanGAP hydrolyzes RanGTP to RanGDP, allowing separation of RanGTP from the importin. RanGDP is returned to the nucleoplasm by the nuclear transport factor NTF2. In the nucleoplasm, RanGDP interacts with the RanGEF RCC1 for exchange into RanGTP, restarting the transport cycle (Figure 1.3) (Wente and Rout, 2010).

While protein import, protein export, and some RNA export processes (notably those of tRNA, rRNA, snRNA and some mRNAs) depend on binding to Kaps and the RanGTP gradient across the NE (Arts et al., 1998; Zemp and Kutay, 2007), the bulk of mRNA molecules does not require a Ran gradient for their export. NXF1-NXT1 are the major transport factors for these mRNAs. These transport factors interact with mRNA transcripts and other adaptor proteins in the nucleus and translocate through the pore by interacting with the FG-repeats in the channel. The directionality of transport is provided by an ATP-dependent DEAD-box helicase (DDX19) on the cytoplasmic face of the NPC. DDX19 interacts with Nup214 at the cytoplasmic face of the pore and it remodels the NXF1-NXT1-mRNP complex, releasing the transport factors from the transcript and ensuring directionality of transport (Montpetit et al., 2011; Napetschnig et al., 2009). In addition to the described nuclear export factors, RNA export also requires several adaptor proteins, further increasing pathway complexity (Sloan et al., 2015).

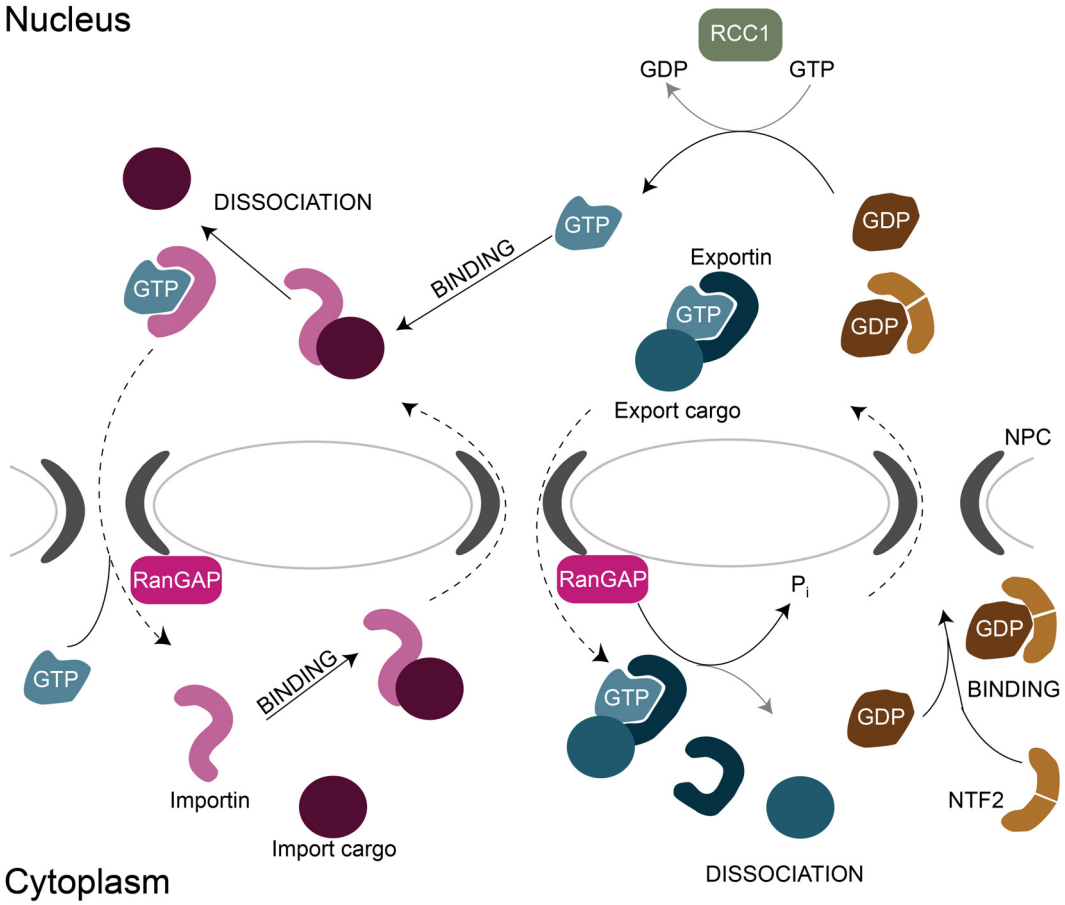


Figure 1.3: NPC-mediated transport cycle

Cargo proteins that undergo nuclear import (purple) contain an NLS sequence that is recognized by an importin (light purple) on the cytoplasmic side of the NPC. This complex interacts with Nups and translocates through the NPC. On the nucleoplasmic side of the NPC, the high concentration of RanGTP (light blue), maintained by RCC1, releases the cargo protein into the nucleus. Importin-RanGTP is exported to the cytoplasm, RanGTP is hydrolyzed to RanGDP and the complex dissociated. Conversely, cargo proteins undergoing nuclear export (blue) contain an NES sequence, which facilitates the formation of a trimeric complex with an exportin (dark blue) and RanGTP (light blue). The export complex translocates through the NPC where RanGAP (pink) hydrolyzes RanGTP to RanGDP (light brown), releasing the export cargo into the cytoplasm. RanGDP interacts with NTF2 and gets re-imported into the nucleoplasm, where it gets converted to RanGTP by RCC1. Adapted from Atkinson, 2012. Reprinted with permission from Dr. Claire Atkinson and Digital Commons @ Rockefeller University, original publication: http://hdl.handle.net/10209/509

1.4.3 Transport models

Although much information about the players involved in nucleocytoplasmic transport has been elucidated from experimentation, there is still considerable debate on how exactly cargoes, bound by their appropriate NTFs, actually move through the NPC. Several models have been proposed to explain translocation through the pore. All of these models build on the well-established binding between nuclear transport factors and the FG-repeat domain of Nups, see Figure 1.4.

The virtual gate model, also known as the entropic bristle model, proposes that the nuclear pore complex selectivity is based on changes in entropy (Rout et al., 2000, 2003). For a cargo to move from a freely diffusing state (in the cytoplasm or nucleoplasm) to a much less mobile state within the central channel of the nuclear pore complex, there would be a rapid and substantial drop in its entropy. However, when a molecule is bound to a NTF, its interaction with the FG-repeats in the channel lowers the activation energy of translocation across the NE, allowing NTF-cargo complexes to enter the NPC. Additionally, in this model the cytoplasmic filaments and the nuclear asymmetric FG-Nups are thought to act as a polymer brush. The movement of these entropic bristles serves to exclude non-transporting proteins from entering the pore, increasing the efficiency of transport (Lim et al., 2006, 2007).

Atomic force microscopy of purified FG-domains attached to a surface suggested that the polymer brush-like structure they form collapses down to the surface upon addition of cargo. This observation lead to the development of the "collapse" model, in which the interaction between NTF-cargo complexes and the FG-repeats causes the FG-repeats to retract, pulling the transport complex into the NPC (Figure 1.4) (Lim et al., 2006, 2007). However, these experiments only tested, Nup153, a component of the nuclear basket that may not be representative of the central and cytoplasmic Nups. Also, the geometry and density of these domains, when tethered to gold nanodots on a flat surface, may have an effect on the observed behavior, which may not be recapitulated in the NPC *in vivo*. A further modification of this model suggests that the presence of NTF-cargo within the pore also helps prevent non-specific molecules from entering the channel (Timney et al., 2006; Zilman et al., 2007, 2010). This possibility was tested using an artificial nanopore lined with the FG-repeats from Nsp1. In these experiments, non-specific cargo was prevented from transiting the nanopore only in the presence of transported cargo, which was selectively favored (Jovanovic-Talisman et al., 2009).

Alternative models suggest that selectivity of the pore originates from the physical barrier created by interactions between FG-repeat domains. One such model, nicknamed the "oily-spaghetti" model, advocates that the many copies of flexible

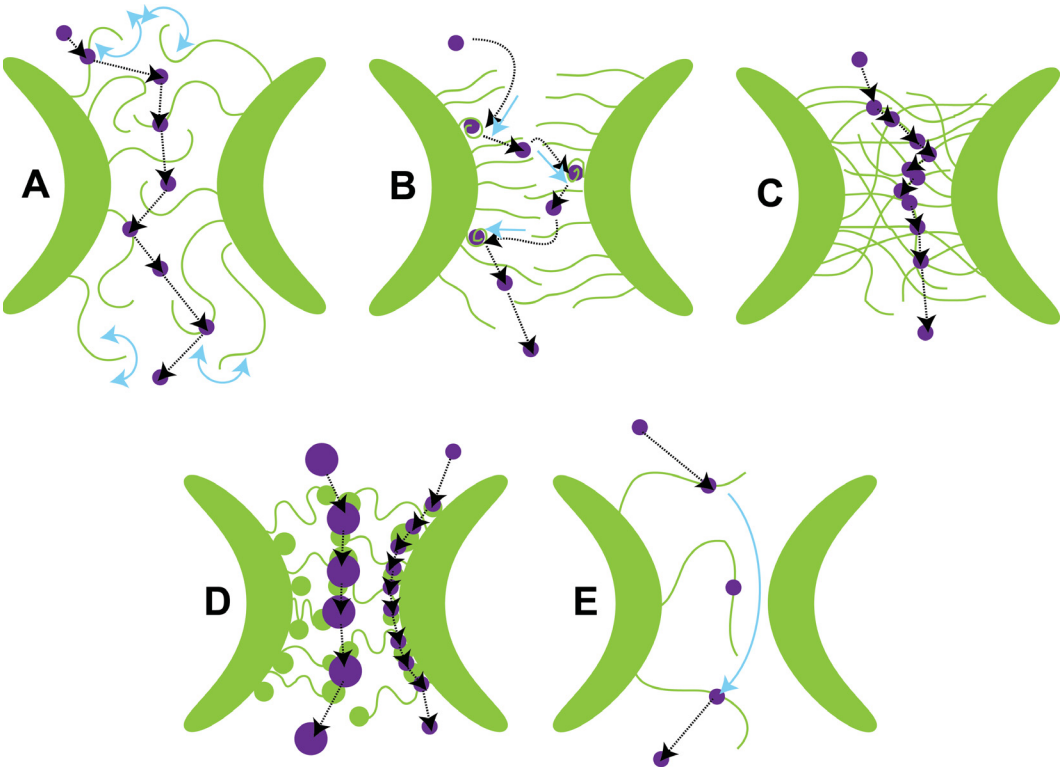


Figure 1.4: Models for NTF translocation through the NPC central channel.

A) Virtual gate model. B) Collapse model. C) FG-Hydrogel model. D) Forest model. E) Brownian ratchet model. FG-Nups are shown in green; cargo is shown in purple; cargo binding and unbinding are shown with black dashed arrows; movement of the FG-repeats is shown with blue solid arrows. Adapted from (Atkinson, 2012). Reprinted with permission from Dr. Claire Atkinson and Digital Commons @ Rockefeller University, original publication: http://hdl.handle.net/10209/509

FG-repeats form a network that fills the central channel of the pore (Macara, 2001). NTF-cargo complexes navigate through the pore by transiently binding to the meshwork. The selective phase model (also termed the FG-hydrogel model) builds on the "oily-spaghetti" model by suggesting that the FG-repeats interact with each other through hydrophobic interactions in the central channel, generating a connected meshwork or phase (FG-hydrogel) (Ribbeck and Görlich, 2001). Nuclear transport factors locally dissolve the meshwork by competing for and disrupting the FG-FG interactions, which allows for passage of cargo through the pore. Molecules that do not interact with the FG-repeats would be excluded, as they would be unable to disrupt the FG-hydrogel structure (Figure 1.4). In support of this model, it has been shown that a single FG-Nup (Nup98), with highly cohesive FG-repeats, can form hydrogels of NPC-like permeability. Nuclei with engineered NPCs reveal that the NPC barrier, and NTF-cargo transport across the pore, require multivalent cohesion between FG-repeats, and cannot rely solely on the interaction between FG-repeats and NTFs (Hülsmann et al., 2012).

The dual gate model attempts to reconcile the FG-hydrogel and the virtual gate models. It is based on data showing that, in a bead binding assay, only a subset of the FG-Nups was able to bind each other via FG-FG interactions (Patel et al., 2007). In this model, the FG-Nups in the center of the pore would form a cohesive FG-hydrogel, while those at the periphery would form an entropic gate.

The forest model proposes that some cohesive FG-Nups exhibit intermolecular FG-FG interactions and adopt a collapsed coil structure. This coiled structure can either line the wall of the pore, as a 'shrub', or can be at the end of relaxed or extended coil structures, filling the center of the pore, as a 'tree'. The collapsed-coil domains would form a transport zone in the middle of the NPC, and the extended-coil domains would create a peripheral transport zone. Small molecules, nuclear transport factors, and complexes pass through one of the two formed zones, depending on their physicochemical properties (Figure 1.4) (Yamada et al., 2010).

The reduction of dimensionality model proposes that the FG-repeats of Nups coat the inner surface of the NPC channel, forming a hydrophobic lining while leaving a narrow central channel that would allow passive diffusion of small molecules through the pore. Continuous interactions between the FG-repeats at the wall and NTFs or NTF-cargo complexes would cause two-dimensional movement along the inner wall of the channel, while molecules too large to diffuse through the central channel, and unable to bind the FG-repeats, would be excluded from the pore (Peters, 2005).

The Brownian ratchet model places importance on the NTF-FG-repeats interaction. This model, based on computational exploration of the NPC parameter space, suggests that once an NTF interacts with an FG-repeat, it remains bound, traversing

the pore in a complex with that FG-domain until its dissociation by RanGTP in the nucleus (Figure 1.4) (Mincer and Simon, 2011).

Clearly, FG-repeats play a vital role in nucleocytoplasmic transport. They have an inherent ability to create an entropy barrier, which can be overcome by transport factors. The different models have divergent views on how the FG-Nups interact with each other, and on the effect the interaction of NTFs with FG-repeats has on these FG-Nups.

1.5 Non-transport functions of nucleoporins

In addition to controlling nuclear transport, Nups can affect several other cellular processes. Many NPC components have been shown to influence diverse genome functions. Nups have been shown to influence gene expression (Capelson et al., 2010; Kalverda et al., 2010; Ptak and Wozniak, 2016), chromatin organization (Liang and Hetzer, 2011; Kalverda and Fornerod, 2010; Ptak et al., 2014), chromosome segregation (Jeganathan et al., 2005, 2006), cell cycle regulation (Kalverda and Fornerod, 2010; Laurell et al., 2011; Lusk et al., 2007; Wozniak et al., 2010; Wozniak and Goldfarb, 2008) and the cellular immune response (Satterly et al., 2007; Enninga et al., 2002).

1.5.1 Roles for the disassembled NPC during mitosis

During mitosis, the NPC is disassembled and some of its components have mitotic roles. In the early stages of mitosis, centrosomes migrate to opposite sides of the nucleus using the NE and NPCs as a sliding platform. Nup358 helps centrosome anchoring by its association with the molecular motors dynein and kinesin through the adaptor protein Bicaudal D2 (Cai et al., 2001; Payne et al., 2003; Splinter et al., 2010). The Nup107-160 complex also contributes to centrosome tethering, by a mechanism distinct from the Nup358-Bicaudal D2-mediated pathway (Bolhy et al., 2011).

In the next mitotic stage, during nuclear envelope breakdown, NPCs are triggered for disassembly by hyperphosphorylation of nucleoporins (Laurell et al., 2011; Glavy et al., 2007). Following nuclear envelope breakdown, microtubules emanate from the centrosomes and bind kinetochores to align the chromosomes along the equatorial plate of the cell. At this stage, the transmembrane Nups are redistributed into the ER, and several soluble Nups become diffusely localized in the mitotic cytoplasm. The Nup107-Nup160 subcomplex, however, relocates to the kinetochore

(Loïodice et al., 2004). The Nup107-Nup160 components and Elys recruit the microtubule nucleator γ-tubulin ring complex to unattached kinetochores (KTs), driving microtubule nucleation (Mishra et al., 2010; Yokoyama et al., 2014). Additionally, Seh1, of the Nup107 subcomplex, is required for proper localization of Aurora B at centromeres (Platani et al., 2009), and attachment of the Nup107 subcomplex to KTs is related to the attachment of Nup358 (Zuccolo et al., 2007), which regulates the stability of kinetochore-microtubule interactions (Joseph and Dasso, 2008). Nup358 is also required in progression into anaphase, as it promotes sumoylation of TOPOIIa, enabling its recruitment to centromeres, where it functions in chromatid separation (Dawlaty et al., 2008). Rae1 and Nup98, redistribute throughout the cell in mitosis. Rae1 is necessary for normal mitotic progression and stabilization of microtubules at the kinetochore (Babu et al., 2003; Blower et al., 2005). Rae1 interacts with the checkpoint kinase Bub1, the spindle assembly factor NuMA, and microtubules (Wang et al., 2001; Babu et al., 2003; Wong et al., 2006). Nup98 also facilitates bipolar spindle formation, interacting with microtubules and antagonizing MCAK (Cross and Powers, 2011). Nup98 and Rae1 also regulate APC/Cdh1 activity during prometaphase/metaphase, preventing premature sister chromatid separation (Jeganathan et al., 2005).

By late anaphase, Nups start to be recruited to the chromosomes, driving NPC and NE reassembly (Doucet et al., 2010; Mackay and Ullman, 2011). Nups might also affect mitotic exit. Nup153 and Nup50 interact with the protein phosphatase PP1 and its targeting subunit, Repo-Man (Moorhead et al., 2008; Vagnarelli et al., 2011). This interaction could influence proper PP1 localization/activity affecting the dephosphorylation cascades of mitotic kinase targets required to recycle the nuclear components after mitosis (Ibarra and Hetzer, 2015).

1.5.2 Nups participate in DNA replication and repair

Proper DNA replication during S-phase is essential for genomic integrity maintenance. The link between NPCs and DNA replication was initially identified when the NPC component Elys was reported to interact with the Mcm2-7 replicative helicase (Gillespie et al., 2007). This observation led to the hypothesis that NPC assembly and replication licensing are coordinated within the cell cycle. In accordance, zebrafish harboring a mutation in the Elys gene showed reduced interaction between Elys and Mcm2 and sensitization to replication stress and DNA damage (Davuluri et al., 2008). Several other Nups have also been reported to play a role in DNA damage repair in yeast (Bennett et al., 2001). Interestingly, DNA double strand breaks and collapsed replication forks were also found to be spatially relocated to

yeast NPCs (Nagai et al., 2008; Palancade et al., 2007). It was proposed that Mec1 and Tel1 (the yeast homologs of ATR and ATM) recruit damaged DNA to NPCs via Nup84 phosphorylation at sites of interaction between the damaged DNA and the NPC-bound SUMO E3 ligase Slx5/Slx8 (Nagai et al., 2008; Palancade et al., 2007). It is possible a similar mechanism exists in higher eukaryotes, as both Nup107 and Nup88 are known ATM/ATR substrates following γ-irradation in human cells (Matsuoka et al., 2007).

1.5.3 Regulation of gene expression by NPCs and Nups

Originally, a relationship between NPCs and actively expressed genes was implied by high-resolution images of mammalian nuclei showing a distinct non-random association of heterochromatin with regions of the NE devoid of NPCs, while regions of the NE containing NPCs showed a possible association with euchromatin (Raices and D'Angelo, 2012; Ptak et al., 2014; Ibarra and Hetzer, 2015). In line with these observations, the gene gating hypothesis (Blobel, 1985) proposed that nuclear pores interact with actively expressed genes to promote co-regulation of transcription and mRNA export.

Consistent with the gene gating hypothesis, NPCs have been shown to associate with transcriptional co-activators that promote chromatin relaxation and gene accessibility. Nup98, Nup93, and Kapα can interact with the histone acetyltransferase CBP/p300 (Kasper et al., 1999; Ryan et al., 2006). Interactions with epigenetic modifiers also enable NPCs to separate transcriptionally active and inactive regions, functioning as barrier insulators to block heterochromatin spreading (Ishii et al., 2002; Krull et al., 2010).

The interaction between NPCs and chromatin can occur in cis or trans, affecting the 3D nuclear landscape. In cis, NPCs mediate chromatin looping between promoters and enhancers over large genomic distances to boost gene expression (Tan-Wong et al., 2009). Trans-chromatin interactions of NPCs can enhance coordinated expression of gene networks, for example, yeast genes on different chromosomes that share similar gene recognition sequences (DNA zip codes) can cluster at NPCs to promote synchronized expression (Brickner et al., 2012).

While in yeast most interactions between chromatin and nucleoporins occur at the NE, in metazoans Nups are not restricted to the NPC but can interact with chromatin within the nucleoplasm. The mobility of nucleoplasmic Nups (e.g. Nup153, Nup98, Nup50) is dependent on active RNA Pol II transcription, indicating that this nucleoporin pool might actively participate in transcriptional regulation (Griffis et al., 2002; Buchwalter et al., 2014).

The involvement of nucleoplasmic Nups in gene expression regulation was first described in Drosophila cells. In these cells, several Nups (Nup98, Nup50, Sec13, Nup62, Nup153 and Mtor/Tpr) have been shown to interact with actively transcribing genes in the nucleoplasm (Capelson et al., 2010; Kalverda et al., 2010). Consistently, the alteration of the expression level of these Nups deregulates the expression of hundreds of genes in these cells (Capelson et al., 2010; Kalverda et al., 2010; Vaquerizas et al., 2010). Many of these Nups were shown to co-occupy the same genomic regions in Drosophila chromosomes, suggesting potential interactions among Nups to regulate gene expression (Kalverda et al., 2010; Vaquerizas et al., 2010). Nup-regulated target genes in Drosophila cells are enriched for developmental regulators and developmentally induced genes (Capelson et al., 2010; Kalverda et al., 2010), indicating Nups may play a direct role in controlling developmental transcriptional programs (e.g. Nup98 is necessary for tissue-specific Hox gene expression in developing flies) (Pascual-Garcia et al., 2014).

Recent publications indicate that the function of intranuclear Nups as regulators of gene expression function observed in Drosophila is conserved in mammalian cells. Attempts to map Nup-genome interactions in differentiated human cells support a context-dependent association of Nups (Nup50, Nup210, Nup98, Nup153, sPom121) with transcriptionally active and repressed regions (Liang et al., 2013; Buchwalter et al., 2014; Franks et al., 2016; Brown et al., 2008; D'Angelo et al., 2012; Jacinto et al., 2015). Additionally, mammalian Nups (Nup50, Nup210, Nup98, Nup153) have been implicated in cell differentiation and cell identity maintenance (D'Angelo et al., 2012; Buchwalter et al., 2014; Jacinto et al., 2015; Liang et al., 2013). For example, in the human cell lines IMR90 and U2OS, Nup153, and Nup93, can interact with, and regulate the expression of, super enhancer-associated genes, which are critical regulators of cell type-specific transcriptional programs (Ibarra et al., 2016). Similarly, mouse embryonic stem cells (mESCs) require Nup153 to maintain their pluripotency by regulating the transcription of several lineage-specific genes (Jacinto et al., 2015). Nup50 and gp210 have also been shown to regulate gene expression during the differentiation of mouse myoblasts to myotubes (D'Angelo et al., 2012; Buchwalter et al., 2014). Consistent with the ideas above, Nup98 binds to distinct genomic regions in different human cell types and Nup98 target genes in human cells change during differentiation, likely influencing specific transcriptional programs as cell identity evolves (Liang et al., 2013).

Nups appear to play a fundamental role in the regulation of cell type-specific gene expression, and they likely have a pivotal role in differentiation and development. However, the molecular mechanisms underlying Nup-mediated gene regulation are not as well understood. For example, in mESCs, Nup153 can recruit the polycomb-repressive complex 1 (PRC1) to the transcription start site of several

lineage-specific genes. Polycomb group proteins promote predominantly gene silencing, but can also activate genes (Aranda et al., 2015; Jacinto et al., 2015). In addition, Nup153, Nup93, and Nup98 have also been reported to interact with the histone acetyltransferase CBP/p300, responsible for H3K27 acetylation and transcriptional activation (Kasper et al., 1999; Vahedi et al., 2015).

The role of Nup98 in gene expression regulation and possible molecular mechanisms will be further discussed in section 1.6. It is possible that Nups might impact gene expression by controlling the chromatin structure and accessibility of transcription factors. Nups could also serve as functional hubs formed in specific chromatin loci to locally enrich the machinery required for the multi-step process of gene expression.

1.6 Nup98

Nup98 is a 920 amino acid protein, and it is one of the peripheral FG-Nups. Nup98 is distinct from the other FG nucleoporins in that it is the only metazoan nucleoporin to contain a substantial number of a characteristic FG-repeat signature of multiple non-tandem glycine-leucine-phenylalanine-glycine (GLFG) repeats. The N-terminal half of Nup98 contains not only its characteristic GLFG but also an FG and two FXFG (where X is any amino acid) nucleoporin repeat motifs. This N-terminal repeats domain is bisected by a small α-helical binding site for Rae1 (Pritchard et al., 1999), and together the Rae1-Nup98 complex is capable of binding single-stranded RNA, see Figure 1.5. Nup98 is post-translationally modified by O-linked N-acetylglucosamine at two or more sites (Finlay et al., 1987; Holt et al., 1987). Nup98 is also phosphorylated during interphase and hyperphosphorylated during mitosis (Macaulay et al., 1995).

The C-terminus of Nup98 (aa 711-870) forms a unique structural fold that is required for autoproteolysis following amino acid 863 (Figure 1.5). The C-terminal domain of Nup98 targets it to the NPC via interactions with the symmetric nucleoporin Nup96 or the cytoplasmic nucleoporin Nup88 (Hodel et al., 2002; Vasu et al., 2001; Griffis et al., 2003). Thus, Nup98 localizes to the cytoplasmic and the nuclear faces of the central channel of the NPC. At nuclear pores the FG/GLFG-repeats of Nup98 are thought to function as docking sites for karyopherins during trafficking of molecules through the NPC. Accordingly, the FG/GLFG domain of Nup98 has been shown to interact with the importin-β family (Allen et al., 2001), the exportin CRM1 (Oka et al., 2010) and the mRNA export factor NXF1 (Blevins et al., 2003; Powers et al., 1997).

In addition to the NPC, Nup98 is present in the cytoplasm and nucleoplasm,

and it is capable of rapidly moving between all these locations (Griffis et al., 2002). Inside the nucleus, Nup98 can be found dispersed in the nucleoplasm and in intranuclear foci termed GLFG bodies, because the GLFG domain of Nup98 is required for targeting to this structure. GLFG bodies can be nucleoplasmic or nucleolar, and examples are shown in Figure 1.5. In HeLa cells, an increased level of Nup98 is associated with an increase in the number of GLFG bodies present. Nups from the Nup107-160 complex, together with Elys, have been shown to colocalize with Nup98 in GLFG bodies. Unlike their behavior at NPCs, the Nups from the Nup107-160 complex can dynamically shuttle into and out of GLFG bodies (Morchoisne-Bolhy et al., 2015).

Nup98 can move between the nucleoplasm, cytoplasm, GLFG bodies and NPCs. However, its mobility within the nucleus is dependent on ongoing transcription in the cell (Griffis et al., 2002). Inhibition of all cellular polymerases by actinomycin D drastically reduces the mobility of all nuclear Nup98 pools (NPC, GLFG bodies and nucleoplasm) and decreases the exchange of nuclear and cytoplasmic Nup98, without affecting the mobility of cytoplasmic Nup98. Specific inhibitors of RNA Pol II (DRB) also reduce the mobility of Nup98 present at NPCs and nucleoplasmic GLFG bodies. However, these inhibitors do not affect the mobility of Nup98 in the nucleoplasm or nucleolar GLFG bodies (Griffis et al., 2004).

These observations link the mobility of intranuclear pools of Nup98 to ongoing transcription. As previously mentioned, studies in Drosophila cells revealed the association of nucleoplasmic Nup98 with actively transcribed genes, especially those involved in developmental regulation and the cell cycle. Modulation of Nup98 expression (overexpression or knock-down) alters the transcription level of these genes (Capelson et al., 2010; Kalverda et al., 2010).

In human cells, the role of Nup98 in transcriptional regulation is still conserved. Nup98 has been shown to associate with chromatin and regulate gene expression during the differentiation of embryonic stem cells into neural progenitor cells. In these cells, Nup98 preferentially associates with developmentally regulated genes, and changes in the level of Nup98 can alter gene expression (Liang et al., 2013).

Several recent observations have provided further insight into the role of Nup98 in transcription. Pascual-Garcia and colleagues showed the binding of Nup98 to the promoter regions of certain developmental genes, such as the Hox family of genes, and a requirement for Nup98 in their transcription. Nup98 binding to these genes was dependent on TRX and MBD-R2, a component of the NSL (nonspecific lethal) complex that directs histone H4K16 acetylation. However, the loss of Nup98 did not change H4K16 acetylation or Trx-mediated H3K4 trimethylation patterns, both of which are required for active transcription and memory. Thus, the function of Nup98 in transcription of these loci remains unclear (Pascual-Garcia et al., 2014).

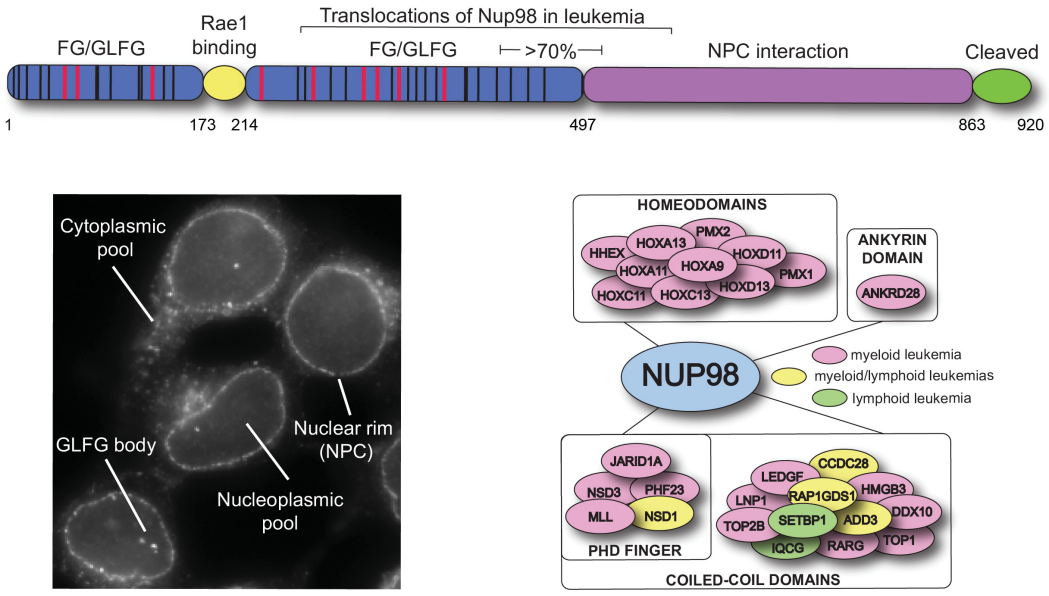


Figure 1.5: Nucleoporin 98

Top: Schematic representation of Nup98 protein. Amino-acid residues are indicated on the bottom, FG/GLFG repeats are marked by lines and domains are represented by ellipses. The mapping region of most known *Nup98* gene translocations in leukemia is indicated in brackets (see Appendix A for further information). **Bottom left**: The cellular distribution of Nup98, in HEK293T cells, was examined by indirect immunofluorescence using antibodies directed against Nup98. The diverse cellular pools of Nup98 are indicated. **Bottom right**: Diagram exemplifying known *Nup98* translocation partner genes in leukemia patients, and the partner gene domains retained in the oncogenic fusion protein (see Appendix A for further information).

Like its counterpart in yeast, mammalian Nup98 also appears to play a role in transcriptional memory, i.e. the rapid induction of a gene with a recent history of being activated. Light and colleagues showed that Nup98 binds to recently expressed *HLA-DRA*, an interferon-induced gene exhibiting transcriptional memory. This gene promoter containing Nup98 accumulates poised RNA Pol II along with histone H3K4 dimethylation. Depletion of Nup98 leads to the loss of this histone mark and absence of poised RNA polymerase, causing slower re-induction of gene expression upon a secondary stimulus with IFN-γ (Light et al., 2013).

Similar results by Panda and collaborators suggest the role of Nup98 in transcriptional memory might also be conserved in Drosophila cells. In these cells, Nup98 primes virus-stimulated genes by regulating the occupancy of active RNA Pol II at their promoters, poising them for rapid induction, and coordinating a robust and complex antiviral response (Panda et al., 2014). Nup98's access to these specific genes and its regulation of their expression is dependent on the transcription factor FoxK; depletion of this TF significantly reduces Nup98-dependent induction of antiviral genes (Panda et al., 2015).

Nup98 has also been shown to participate in Forkhead box O1 (FOXO1)-mediated transcriptional regulation in the human hepatoma cell line HepG2, where Nup98 can interact with phosphorylated FOXO1 and CBP inside the nuclei. This interaction can increase H3K9/14 acetylation in the promoter region of a FOXO1 target gene (PCK1), leading to increased gene expression. The presence of insulin, known to represses FOXO1-dependent transcription, promotes the relocalization of the FOXO1-PCK1 gene complex to the nuclear periphery, in a CRM1-dependent manner. At the nuclear envelope, the FOXO1-PCK1 gene complex associates with NPC-bound Nup98 and EHMT2, an H3K9 dimethylase that, through histone post-translational modifications, causes inhibition of PCK1 transcription. FOXO1 seems to use Nup98 as a transcriptional regulator through histone modifications. FOXO1 interaction with Nup98 in the nucleoplasm allows recruitment of CBP, a histone acetyltransferase, which activates PCK1 transcription. On the other hand, FOXO1 interaction with NPC-bound Nup98 leads to histone dimethylation and transcriptional repression. Thus, FOXO1 may use nucleoplasmic or NPC-bound Nup98 for either transcriptional activation or repression, respectively (Arai et al., 2015).

Nup98 also affects gene expression at the post-transcriptional level. In response to genotoxic stimuli, p53 upregulates genes, such as p21, to regulate cell cycle and DNA repair. In normal cells, Nup98 binds to the 3'UTR of these mRNAs (i.e., p21 and other p53 target genes) to protect them from degradation. In the absence of Nup98, these exposed mRNAs are targeted for exosome-mediated degradation. Therefore, it is not surprising that Nup98 expression is downregulated in some types of cancer (such as hepatocellular carcinomas), which may lead to destabilization of selected

p53 target mRNAs, contributing to tumorigenesis (Singer et al., 2012).

More evidence for the role of Nup98 in gene expression regulation also comes from studies of hematopoietic malignancies. More than twenty-eight different chromosomal rearrangements involving the Nup98 gene have been identified (Figure 1.5). The resulting fusion proteins have been shown to alter transcription through fusing the N-terminal domain of Nup98 (Bai et al., 2006; Kasper et al., 1999) to a C-terminal domain that usually contains a chromatin/DNA interacting region (Figure 1.5) (Capitanio and Wozniak, 2012). The oncogenicity of several Nup98 fusions has been demonstrated in mouse models, where Nup98 fusions lead to acute myeloid leukemia recapitulating the human disease phenotype (Gough et al., 2011; Moore et al., 2007). For more information on gene expression deregulation by Nup98-fusions in mouse models of acute myeloid leukemias see Appendix A.

All the evidence discussed above indicates that Nup98 and other NPC components are important regulators of gene expression, influencing several different steps in this process. Nups have been shown to alter epigenetic marks in chromatin, modulate gene transcription, increase mRNA stability, and control mRNA export. It is therefore not surprising that deregulation of nucleoporins can be observed in diverse pathologies, such as neoplasias and viral infections (see section 1.9). Despite our incomplete knowledge on the mechanism by which Nup98 and other Nups affect these processes, their relevance in gene expression regulation is well established. In this dissertation, we will further discuss a possible mechanism by which nucleoporins can exert their role in gene expression by acting in complex with, and regulating the activity of, cellular helicases (chapter 3).

1.7 Helicases

RNA and DNA helicases are enzymes that can catalyze the energy dependent separation of double-stranded nucleic acids (NAs) (Cordin et al., 2006). Helicases are highly conserved on the structural level. However, different helicases perform diverse functions that range from unwinding thousands of DNA base pairs during replication to destabilizing short RNA helices during processes such as pre-mRNA splicing and ribosome biogenesis. The ubiquitous presence of helicases in prokaryotes, eukaryotes, and viruses reflects their fundamental importance in DNA and RNA metabolism, such as replication, recombination, DNA repair, transcription, translation, splicing, RNA export, RNA turnover, and processing of microRNAs (Jarmoskaite and Russell, 2014; Bourgeois et al., 2016).

Helicases can be classified into five superfamilies (SF1 to SF5) depending on the occurrence and characteristics of conserved motifs in the primary protein sequence

(Singleton et al., 2007). All helicases possess the Walker A and B motifs, these nucleotide triphosphate (NTP) binding motifs are also found in many NTPases (Walker et al., 1982), allowing them to bind and hydrolyze NTPs.

SF1 and SF2 contain a large number of DNA and RNA helicases from archaea, eubacteria, eukaryotes, and viruses. They can unwind duplexes in a 3' to 5' direction as well as 5' to 3' direction. SF3 includes small putative helicase domains of 100 amino acid residues that are found in DNA and RNA viruses. This family has only three conserved motifs, including the two classical ATP-binding motifs described above. SF4 consists of helicases that are related in sequence to the *E. coli* DnaB protein. They have five conserved motifs, unwind DNA in the 5' to 3' direction, and generally form hexameric ring structures. The last superfamily of helicases, SF5, is recognized as a family with sequence similarity to the β-subunit of proton-translocating ATPases, as exemplified by the transcription termination factor Rho (Singleton et al., 2007).

Helicases can also be subdivided by mechanistic differences. Enzymes with specificity for DNA, RNA, or DNA-RNA hybrids have all been identified. SF1 and SF2 enzymes are considered to be primarily monomeric or dimeric, while the other superfamilies (SF3, SF4, and SF5) arrange in hexameric rings formed from six individual RecA folds. The three-dimensional fold of the ATP-binding domains, however, is conserved in the helicases of different superfamilies (Singleton et al., 2007, 2000). To date, most of the known eukaryotic RNA helicases belong to SF2 and only a few belong to SF1. Our focus here will be on RNA helicases from superfamily 2.

The central helicase core of RNA helicases is highly conserved and is surrounded by variable N- and C-terminal domains. In most RNA helicases, the flanking domains are larger in size than the helicase domain and are not well conserved within or between helicase families. Since RNA helicases are usually found within large RNP complexes, a current hypothesis is that the specificity of binding to target RNAs is conferred by the interactions of the flanking domains of the helicase with other proteins on the target RNA (Tanner et al., 2003). Thus, the activity of individual RNA helicases is highly context dependent and influenced by the availability of interacting partners (Fuller-Pace, 2013a).

Most RNA molecules exhibit significant inter- or intra-molecular interactions, which serve as negative regulators of RNA function. Thus, a prerequisite for most biological processes is a rearrangement of these interactions into alternative RNA-RNA or RNA-protein interactions. RNA helicases have three broad functions: catalyzing progressive rearrangements between RNA, DNA, and/or protein molecules; functioning as chaperones to remodel ribonucleoprotein (RNP) complexes; and scaffolding for protein interactions that are required for a given step in RNA metabolism. Via these capabilities RNA helicases can affect from gene transcription to every step

in post-transcriptional gene expression (see Figure 1.6 for examples) (Bourgeois et al., 2016).

A high degree of structural similarity is observed in the catalytic cores of all known RNA helicases. The helicase core is composed of two distinct but highly similar domains arranged in tandem. These two domains share nine conserved sequence motifs: motifs Q, I (Walker A), Ia, Ib, II (Walker B), and III in domain 1 and motifs IV, V and VI in domain 2 (motif Q is only present in DEAD-box helicases) (Bleichert and Baserga, 2007; Tanner et al., 2003). All SF2 RNA helicases contain Walker A and Walker B motifs and at least five to seven of the other conserved motifs. SF2 helicases are subdivided into five distinct families based on the consensus sequence in conserved motifs and similar structural and functional properties (Fairman-Williams et al., 2010). The five families of the SF2 of RNA helicases include the following: DEAD-box helicases, DEAH/RHA helicases, RIG-I-like proteins, Ski2-like proteins, and the NS3/NPH-II subfamily (comprised only of viral proteins). Our focus here will be on the second RNA helicase family mentioned (DEAH/RHA), although DEAD-box and DEAH/RHA helicases present several similarities and can be collectively identified as DExD/H RNA helicases.

Given the role played by RNA helicases in various steps of RNA metabolism (Figure 1.6), it is not surprising that multiple developmental disorders, neurodegenerative disorders, and cancers have been linked to deregulated expression or loss of functionality of RNA helicases (Abdelhaleem, 2004; Hanada and Hickson, 2007). RNA helicases are also commonly hijacked by viruses during their infection life cycle (Kwong et al., 2005).

1.7.1 DExD/H box helicases

The DEAD/H-box family is a large group of multifunctional helicases that can unwind RNA and DNA. Members of this helicase family share eight highly conserved amino acid motifs located in two different domains (motifs I, Ia, Ib, II and III - domain 1, motifs IV, V and VI - domain 2), including the four amino acid residues in Motif II, Asp-Glu-Ala-Asp/His, for which the family is named.

Motifs I and II are NTP-binding motifs, also known as the Walker A and B motifs. Motif I forms a loop structure (P-loop), that creates a pocket that can bind the phosphates of NTP. Mutations of the first alanine residue, the conserved lysine, or the last threonine of Motif I abolish ATPase activity (Caruthers and McKay, 2002; Cordin et al., 2004). Motif II forms interactions with β and γ-phosphates of NTP through a coordinated $Mg_2{}^+$ (Fry et al., 1986). Mutations within Motif II decrease or abolish ATPase and helicase activities without altering RNA binding by the helicase domain

(Pause and Sonenberg, 1992; Iost et al., 1999). Motif III was proposed to participate in linking ATPase and helicase activities. Mutations in Motif III can cause a loss of helicase activity while having minimal effects on NTP hydrolysis and RNA binding to the helicase domain (Pause and Sonenberg, 1992; Schwer and Meszaros, 2000). Motifs Ia, Ib, and IV have been poorly studied in DExD/H proteins. Motifs Ia and Ib are part of domain 1, whereas Motif IV is found at the bottom of domain 2. Motif V is proposed to be an RNA-binding motif in association with Motifs Ia, Ib, and IV (Cordin et al., 2006). Nevertheless, a role for these motifs in ATPase activity and in coupling the ATPase and helicase activities cannot be excluded (Caruthers et al., 2000). Motif V is part of a loop at the interface between domains 1 and 2 that points towards the RNA-binding region. Motif VI has been shown to be essential for ATPase activity and RNA binding (Pause et al., 1993). Mutations in Motif VI (from the basic residues histidine or arginine, to uncharged glutamine) abolish RNA binding and reduce ATP hydrolysis, resulting in reduced helicase activity (Pause and Sonenberg, 1992).

The main activity of a helicase is to couple ATP binding and hydrolysis to conformational changes that result in nucleic acid base pairs separation and/or translocation along a nucleic acid substrate. The detailed molecular mechanism of dsDNA or dsRNA unwinding by helicases is still not completely known, but there are certain features of unwinding and translocation that are common to all helicases.

Recent structural information on DEAH/RHA and Ski2-like proteins indicates that they may be able to unwind RNA helices by a mechanism that is fundamentally similar to SF1 and viral SF2 helicases (Jarmoskaite and Russell, 2014). Conventional helicase mechanisms include translocation of the protein along one of the nucleic acid strands. This movement is directly linked to unwinding since movement along one strand results in displacement of the complementary strand. When bound to a NA that includes a 3' extension, the helicase encircles this single-strand region by domain insertions in domain 1 (D1) and domain 2 (D2). Translocation occurs as ATP binding induces closure of the two core domains (D1 and D2), resulting in movement of D1 while D2 remains stationary on the NA. ATP hydrolysis and product release allow a transient loosening of D2 from contacts with both single-stranded nucleic acids and D1, resulting in domain opening and movement of D2 by one nucleotide in the direction of translocation. Reformation of contacts by D2 resets the core for ATP binding and another turn of the cycle. The mechanistic features of DEAH/RHA and Ski2-like helicases suggest a general requirement for a 3' single-stranded extension to allow for helicase loading, which most likely generates important constraints on their physiological substrates and specific functions (Jarmoskaite and Russell, 2014).

DExD/H-box helicases are ubiquitous and essential for most aspects of the cel-

lular RNA metabolism (Figure 1.6). The processing and fate of mRNAs rely on their packaging into messenger ribonucleoprotein particles (mRNPs). RNA helicases, with their capacity to remodel or lock the composition of mRNP complexes, have pleiotropic functions at different steps of gene expression (Figure 1.6). RNA helicases bridge different gene expression steps, coordinating gene expression programs. As RNA helicases control a subset of mRNAs along different steps of the gene expression process, they guide these subsets of mRNAs towards dedicated molecular factories, monitoring the specific direction of the flow of genetic information (Figure 1.6) (Bourgeois et al., 2016). In this dissertation, we will focus on one such multifunctional RNA helicase, DHX9, also termed RNA helicase A (RHA) and use it as an example to characterize the broad range of functions an RNA helicase performs in the cell.

1.8 DHX9

DHX9 (also known as Nuclear DNA Helicase II (NDHII) and RNA Helicase A (RHA)) is a 142 kDa (1270 amino acid residues) helicase, a member of the DExH-box family, and capable of unwinding both RNA and DNA (Zhang and Grosse, 1994), as well as aberrant polynucleotide structures (Jain et al., 2010) in an NTP-dependent manner. RNA-containing duplexes are unwound more efficiently than dsDNA (Chakraborty and Grosse, 2010), and as previously described for other DExH-box helicases, DHX9 also shows a preference for substrates with a short single-stranded 3' overhang. DHX9 translocates in the 3' to 5' direction and can utilize all dNTPs and rNTPs for its unwinding activity. DHX9 is a multidomain, multifunctional RNA helicase, with regulatory roles in DNA replication, transcription, translation, RNA processing and transport, microRNA processing, and maintenance of genomic stability. DHX9 homologs have been characterized in humans, bovines, mice, Drosophila, *C. elegans*, and Arabidopsis (Lee and Pelletier, 2016).

DHX9 was originally identified as a critical factor involved in the development of male fruit flies, a protein known as Maleless (MLE), necessary for X-chromosome dosage compensation (Belote and Lucchesi, 1980). Mammalian DHX9 was originally purified from the nuclear fraction of calf thymus and designated NDHII due to its DNA unwinding activity (Lee and Hurwitz, 1992). The human homolog was isolated shortly afterward from nuclear extracts of HeLa cells and termed RHA due to its dsRNA unwinding activity (Lee and Hurwitz, 1993). Today, these proteins are recognized as homologous and DHX9 from diverse species are known to unwind both DNA and RNA in an NTP-dependent manner (Lee and Pelletier, 2016). We will focus mainly on human DHX9 below.

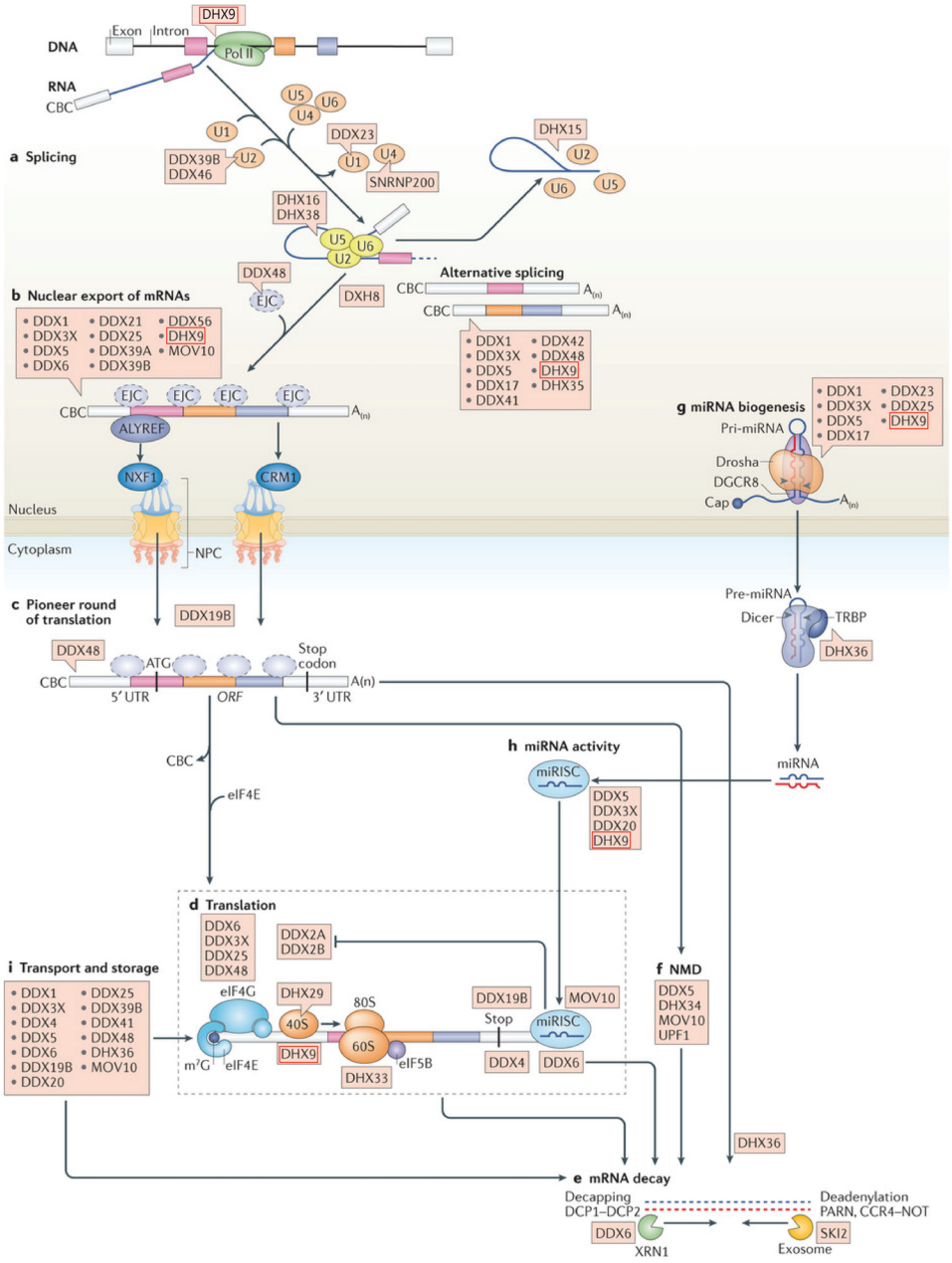


Figure 1.6: RNA helicases in the maturation and fate of an mRNA.

RNA helicases can interact with transcriptional regulators and chromatin to regulate transcription. They also interact with the produced mRNA co-transcriptionally and participate in all subsequent steps of an mRNAs life cycle. **a.** Following binding of the cap-binding complex, the pre-mRNA undergoes splicing. DHX16 promotes the formation of the spliceosome. DHX8 and DHX15 help to disassemble the spliceosome. RNA helicases also contribute to alternative splicing. **b.** Upon splicing, the exon junction complex assembles upstream of exon-exon junctions and contributes to mRNA export from the nucleus. This requires recruitment of Aly/REF, NXF1 or CRM1, and several RNA helicases. **c.** The mRNP is remodeled by DDX19B in the cytoplasm and undergoes CBC-dependent translation with the help of DDX48. **d.** eIF4E replaces CBC, and the eIF4F complex initiates translation through recruitment of the 40S and 60S ribosomal subunits. **e.** RNA degradation can occur from 5' to 3' or 3' to 5' via different mechanisms. **f.** Nonsense-mediated mRNA decay relies on interactions between the pioneering translation machinery, the EJC and the UPF1 RNA helicase and its cofactors. **g** In the nucleus, RNA helicases modulate cleavage of the pri-miRNA transcript into hairpin intermediates. In the cytoplasm, the pre-miRNA is cleaved by Dicer-TRBP. **h.** Several helicases contribute to the formation of the miRNA-induced silencing complex, and its binding to mRNA targets. **i** Translationally stalled mRNAs in transport or storage granules can be redirected to translation or mRNA decay. DHX9 is highlighted by red squares. Adapted from Bourgeois et al., 2016. Reprinted with permission from Springer Nature, original publication DOI: https://dx.doi.org/10.1038/nrm.2016.50

1.8.1 Structure and domain organization

DHX9 contains a central helicase core domain (aa 380 to 830) consisting of eight motifs (see subsection 1.7.1 for detailed descriptions), Motifs I-III are in domain 1 and Motifs IV-VI in domain 2. In addition to the helicase domain, DHX9 contains two double-stranded RNA-binding domains (dsRBDs) at its N-terminus (aa 1-91 and 165-264). The minimal transactivation domain (MTAD) is another N-terminal domain of DHX9 (aa 331-380). It is the site of RNA polymerase II (RNA Pol II) interaction, and it is adjacent to the start of the helicase domain. Adjacent to the end of the helicase domain is the helicase-associated domain 2 (HA2) (aa 830-958). The C-terminal portion of DHX9 contains three other domains, an oligonucleotide/oligosaccharide-binding fold (OB-fold) (aa 958-1074), overlapping NLS and NES sequences (aa 1155-1173), and a glycine-rich RGG-box that can interact with single-stranded nucleic acids (aa 1173-1270) (Lee and Pelletier, 2016).

The structure of the helicase core of DHX9 is conserved with what has been described above for the DExD/H helicases. The MTAD domain consists of two short β-strands that lie in a hydrophobic groove on the surface of the helicase core (Schütz et al., 2010). The two dsRBD domains are arranged into a core α-β-β-β-α fold, with the two α-helices lying on one surface of the three-stranded antiparallel β-sheets (Nagata et al., 2012). The remaining domains of DHX9 do not have resolved 3D structures to date.

Mutational analysis indicated that the two dsRBD domains, the OB-fold and RGG-box of DHX9 are dispensable for its NTPase and helicase activities, suggesting that aa 313-1160 comprise a minimal functional helicase (Zhang and Grosse, 1997; Xing et al., 2014). A point mutation in Motif I (GCGKT to GCGRT, identified as K417R) of the helicase core of DHX9 abrogates ATP binding and ATPase activity, supporting this to be the site of NTP binding (Aratani et al., 2001). The two dsRBDs show optimal binding to dsRNA, and they enhance DHX9's catalytic activity by promoting binding of DHX9 to substrate RNA (Xing et al., 2014). The RGG-box, on the other hand, binds specifically to single strand nucleic acids (Zhang and Grosse, 1997). The absence of the RGG-box or the two dsRBD diminishes the nucleic acid-stimulated ATPase activity of DHX9. While the structure of full-length DHX9 has not been elucidated, current models indicate that the N-terminal, helicase core, and C-terminal domains may be in close spatial proximity, allowing the dsRBDs and RGG-box domains to help regulate or modulate the activity of the helicase domain (Zhang and Grosse, 1997). For example, the dsRBDs and RGG-box domains may initiate binding to nucleic acids, and may cooperatively recognize single-stranded/double-stranded junctions. This binding may effect an allosteric change to activate the NTPase/helicase activity of DHX9, as has been previously ob-

served for other dsRBD-containing helicases (Figure 1.7) (Lai et al., 1995; Patel and Sen, 1992). The kinetic and molecular mechanism involved in DHX9's RNA helicase activity were characterized by single-molecule Forster Resonance Energy Transfer (smFRET), a technique that enables detection of unwinding by a single DHX9 on a duplex RNA molecule. The smFRET of wild type and dsRBD-deleted DHX9 proteins showed that the dsRBDs increase the binding affinity and contribute to the stability of DHX9 binding to dsRNA (Koh et al., 2014).

1.8.2 Cellular localization

In human cells, DHX9 is typically localized to the nucleoplasm and excluded from the nucleolus (Zhang et al., 1999a; Fuchsová and Hozák, 2002). DHX9 is phosphorylated, in a RNA-dependent manner, by the DNA-dependent protein kinase (DNA-PK) and this modification determines its subnuclear localization, with exclusion from the nucleoli. However, under conditions of RNA Pol II mediated transcriptional inhibition, growth arrest, or stress (viral replication or low temperature), DHX9 is translocated into the nucleolus (Zhang et al., 1999a; Fuchsová and Hozák, 2002; Liu et al., 2007). Transport into the nucleolus is dependent on DHX9's NTPase and helicase activity and is mediated by the dsRBDII and C-terminal nuclear transport domains (Liu et al., 2007).

Despite being a predominantly a nuclear protein, DHX9 also shuttles to the cytoplasm to carry out some of its functions in translational regulation and miRNA processing (Zhang et al., 1995). DHX9 can also be seen in the cytoplasm as a consequence of transcriptional inhibition and during mitosis (Zhang et al., 1999b). During mitosis, DHX9 is released into the cytoplasm in prophase, during chromosomal condensation and breakdown of the nuclear envelope, and it reenters the nucleus during telophase, when the nuclear envelope reforms (Zhang et al., 1999b; Pfaller et al., 1991).

DHX9's nucleocytoplasmic shuttling depends on its NLS and NES, both located in the C-terminal region. Nuclear import is mediated by the classical Kapα/β-dependent pathway (Aratani et al., 2006; Tang et al., 1999) and requires methylation of arginine residues in the NLS by the protein arginine methyltransferase PRMT1 (Smith et al., 2004). The nuclear export pathway utilized by DHX9 remains to be elucidated, but it has been determined to be insensitive to leptomycin B, a drug that specifically blocks the CRM1-dependent nuclear export pathway (Tang et al., 1999; Kudo et al., 1999; Wolff et al., 1997). Aside from the NLS and NES region, subcellular localization may also depend on other functional domains and the interactions they mediate with other proteins and nucleic acids (Fujita et al.,

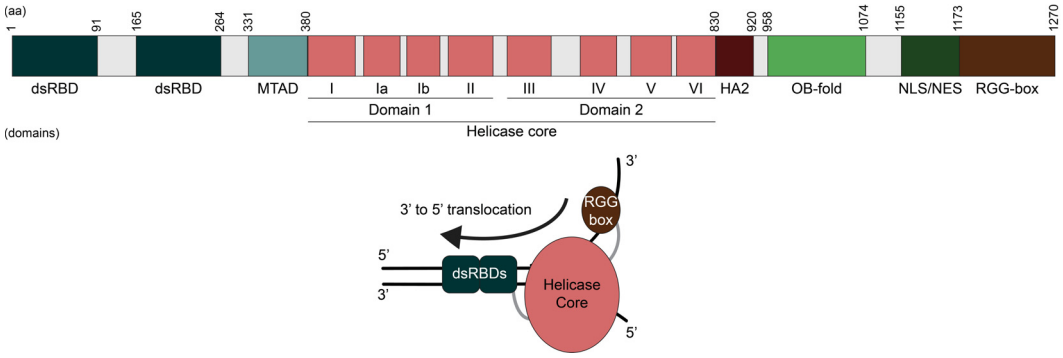


Figure 1.7: DHX9 domain organization and putative positioning on a nucleic acid substrate

Top: Schematic representation of DHX9 functional domains: Numbers indicate amino acid positions in human DHX9. dsRBD, double-stranded RNA binding domain; MTAD, minimal transactivation domain; HA2, helicase-associated domain 2; OB-fold, oligonucleotide/oligosaccharide-binding fold; NLS, nuclear localization signal; NES, nuclear export signal. Adapted from Lee and Pelletier, 2016. **Bottom:** Hypothetical diagram of DHX9 interaction with a forked substrate. The dsRBDs and RGG-box domains may initiate binding to nucleic acids, and may cooperatively recognize single-stranded/double-stranded junctions. This binding may effect an allosteric change to activate the NTPase/helicase activity of DHX9. Adapted from Gibson and Thompson, 1994.

2005).

1.8.3 The role of DHX9 in gene expression and RNA metabolism

In vivo and *in vitro* data indicate that DHX9 performs a number of functions spanning transcription, RNA processing, translation, loading of RISC complex, and innate immune sensing (Nakajima et al., 1997; Hartman et al., 2006; Anderson et al., 1998; Robb and Rana, 2007; Kim et al., 2010; Tang et al., 1997). The versatility of DHX9 is attributable to its multidomain architecture that facilitates numerous protein-protein and protein-nucleic acid interactions.

Transcriptional regulation and splicing

DHX9 was initially characterized as an intermediate factor bridging RNA Pol II to the transcriptional co-factor CBP (Nakajima et al., 1997). The interaction of DHX9 with RNA Pol II is mediated by aromatic residues in the minimal transactivation domain (MTAD, aa 331-380), and is necessary for transcriptional activation. The interaction of DHX9 with CBP is mediated by DHX9's N-terminal dsRBD (aa 1-88) (Aratani et al., 2001). The interaction of coactivators (CBP) and RNA Pol II with different domains of DHX9 indicates it could act as a hinge, transmitting regulatory signals from CBP (or other binding factors) to RNA Pol II. Of note, the interaction between CBP and DHX9 might be relevant in oncogenesis. CBP is implicated in the regulation of the expression of genes involved in malignant transformation, such as *MYC, JUN, FOS*, transforming viral proteins, like E1A, E6 and large T antigen, and tumor-suppressor proteins, such as p53, E2F, Rb, Smads, RUNX and BRCA1 (Iyer et al., 2004). Mutations in the CBP-binding region of DHX9 occur in several human tumors and were reported to strongly affect regulation of gene expression (Chen et al., 2014).

DHX9 also activates transcription by interacting directly with BRCA1 (via DHX9 aa 230-325) and bridging its association with RNA Pol II. Of note, breast cancer-related BRCA1 mutants display low ability to bind DHX9, thus reducing BRCA1 tumor suppressor activity and promoting cancer growth (Anderson et al., 1998). Nuclear β-actin is another component of the transcription pre-initiation complex (Hofmann et al., 2004), and DHX9 serves as an adaptor to link it with RNA Pol II. This interaction enhances transcription from the actin-dependent CSF-1 promoter. Contrary to what was observed for the aforementioned transcriptional regulators, DHX9 interacts with β-actin via its C-terminal RGG-box, and its NTPase activity is not required for transcriptional activation (Tang et al., 2009).

DHX9's N-terminal domain (aa 1-649) can also interact with the proto-oncoprotein p65. As described for CBP and BRCA1, the interaction of p65 with DHX9 could contribute to RNA Pol II recruitment for the formation of a transactivation complex. DHX9 binding activates NF-$\varkappa$B-mediated transcription, and DHX9 depletion reduces the expression of these genes (Tetsuka et al., 2004). Moreover, the ATPase activity of DHX9 is required for the transcriptional activation mediated by NF-$\varkappa$B. Since p65 is known to utilize CBP/p300 as a coactivator (Gerritsen et al., 1997), it is possible that DHX9, p65, CBP/p300 and RNA Pol II may all be part of the same transactivation complex. Activation of CREB/CBP/p300-mediated transcription by the methyl-CpG binding domain protein 2 (MBD2a) (Fujita et al., 2003), and the E2-like enzyme UBC9 (Argasinska et al., 2004) are also dependent on interaction with DHX9.

Other known DHX9-binding transcriptional activators include the osteoblast-specific transcriptional factor osterix (Amorim et al., 2007), nuclear factor 110 (NF110) (Reichman et al., 2003), the Zic2 zinc finger protein (Ishiguro et al., 2007), topoisomerase IIα (Zhou et al., 2003), and LMX1B (Hoekstra et al., 2013). In addition to activating transcription, DHX9 can also repress it. Association of DHX9 with the transcriptional activator TonE-binding protein (TonEBP) inhibits TonEBP activity (Colla et al., 2006). Although the mechanism is not clear, it is possible that DHX9 may recruit other proteins that directly inhibit TonEBP.

DHX9 also binds directly to promoters in a sequence-specific manner. It enhances transcription of the tumor suppressor *CDKN2A* by binding specifically to the sequence 5'-CGGACCGCGTGCGC-3' within its promoter (Myöhänen and Baylin, 2001). Another example of selective transcriptional regulation is that of the multidrug resistance gene 1 (MDR1). DHX9 is a component of the MDR1 promoter-enhancing factor (MEF1) complex and binds to the CAAT-like cis-acting element in the MDR1 promoter (Zheng, 2004). As well, DHX9 participates in EGF receptor (EGFR)-mediated transcriptional activation. DHX9 mediates the interactions between EGFR and an AT-rich sequence in the promoter of target genes (Lin et al., 2001).

Splicing is an RNA processing step that usually occurs closely coupled to transcription. Therefore it is not surprising that DHX9 may also participate in splicing regulation. DHX9 has been identified in human pre-spliceosomes (Hartmuth et al., 2002). It binds to both mRNA and pre-mRNA (Zhang et al., 1999b), and it interacts with the splice regulator muscleblind 1 (MBNL1) (Paul et al., 2011), as well as the survival motor neuron (SMN) protein, a component of small nuclear ribonucleoproteins (snRNPs) involved in pre-mRNA splicing (Pellizzoni et al., 2001b; Terns and Terns, 2001; Pellizzoni et al., 2001a). The role of DHX9 in mRNA splicing is further supported by evidence that DHX9, in concert with ADAR2 (an adenosine deaminase

that acts on RNA) coordinates the editing and splicing of the glutamate receptor subunit B pre-mRNA. mRNA editing and splicing are competing events - ADAR2 editing requires a stable stem-loop, which may sequester the 5' splice site. It is thought that DHX9 helps overcome this splicing inhibition by resolving the stem-loop (Bratt and Ohman, 2003). An example of coordinated editing and splicing can also be found in Drosophila. In a process distinct from its role in dosage compensation, MLE links editing and splicing of the para sodium channel pre-mRNA. A mutation near the NTP-binding site of MLE results in aberrant splicing and exon skipping, again suggesting that the Drosophila DHX9 homolog may act to resolve secondary structures concealing splice sites (Reenan et al., 2000; Kernan et al., 1991).

Post-transcriptional regulation: translation and RNA interference

Through its ability to recognize and bind complex secondary RNA structures, DHX9 can participate in other post-transcriptional steps of RNA metabolism, such as regulating the translation of select mRNAs or inhibiting gene expression, by neutralizing targeted mRNA molecules through RNA interference.

One such example is the DHX9-dependent translation of *JUND* mRNA, which contains a highly structured post-transcriptional control element (PCE). Found in the 5'UTRs of retroviruses and some cellular mRNA transcripts, PCEs contain two redundant stem-loop structures that must be rearranged by DHX9 for efficient ribosome loading and eIF4A function (Hartman et al., 2006). Similarly, DHX9 is required for the effective translation of LIN28 target mRNAs, which possess highly structured 5'UTRs. LIN28, a protein that is highly expressed in stem cells, actively recruits DHX9 to the polyribosomes for the translation of these select mRNAs (Jin et al., 2011). Similarly, DHX9 is necessary for sufficient levels of collagen synthesis. LARP6 recruits and binds DHX9 to the start codon, within the 5' stem-loop structure of collagen mRNA, to facilitate unwinding and translation initiation (Manojlovic and Stefanovic, 2012), leading to enhanced collagen translation. DHX9 also helps regulate IRES-mediated translation. Exposure to DNA damaging agents increases p53 translation (Grover et al., 2009; Giaccia and Kastan, 1998; Halaby and Yang, 2007) through an IRES in the 5'UTR of p53 mRNAs (Ray et al., 2006; Yang et al., 2006). DHX9 was found to bind simultaneously to the p53 IRES and to translation control protein 80 (TCP80) to stimulate p53 IRES-mediated translation. It is predicted that DHX9 likely helps unwind the p53 5'IRES, thereby promoting efficient translation (Halaby et al., 2015a,b).

Another function of DHX9 in post-transcriptional gene expression is its role in microRNA (miRNA) biogenesis and processing. The BRCA1-DHX9 complex participates in miRNA maturation. BRCA1 interacts with the DROSHA microprocessor

complex and regulates the processing of a small set of precursor and mature miRNAs. RNA immunoprecipitation showed that both BRCA1 and DHX9 associate with similar miRNAs, and DHX9 depletion suppressed the processing of these pri-miRNAs (Kawai and Amano, 2012). Further downstream in the miRNA processing cascade, DHX9 can also directly interact with the main components of the RNA-induced silencing complex (RISC), including siRNA, TRBP, Dicer, and Ago2. DHX9 participates in RISC loading and unloading of siRNA. In cells depleted of DHX9, gene silencing of another endogenous mRNA target, such as *CDK2*, is significantly impaired - indicating that DHX9 is necessary for the formation of the active RISC (Robb and Rana, 2007).

The interactions of DHX9 with an extensive and varied array of nuclear and cytoplasmic protein and nucleic acid partners indicates that it may be a central regulator of gene expression. Therefore, it is not surprising that DHX9 has been implicated in several pathologies, especially neoplasias and viral infections (see section 1.9), as well as the regulation of immune response (Fullam and Schröder, 2013).

1.9 NPCs and helicases in health and disease

The cellular distribution of proteins and their regulation are crucial for normal cellular function and human health. Deregulation of nucleoporins and nucleocytoplasmic transport, allowing mislocalization of nuclear and cytoplasmic proteins, is therefore increasingly recognized as a basis for human diseases. Structural changes in Nups and NTFs or defects in transport pathways leading to nuclear or cytoplasmic accumulation of molecules are correlated with a number of diseases, such as cancer, immune system disorders, nervous system diseases and viral infections. Similarly, DExD/H-box helicases are involved with most steps of RNA metabolism, acting as master regulators of gene expression in the cell. Thus, RNA helicases play critical roles in cellular metabolism and in many cases have been implicated in disease development, such as participating in viral life cycles, sensing viral infections and triggering an immune response, participating in neurological disorders, neoplastic transformation, and aging. In this dissertation, we will describe the role of helicases, nucleoporins, and nuclear transport factors in viral infections, with particular emphasis on their role in the life cycle of viruses from the *Flaviviridae* family.

1.9.1 The NPC in viral infections

The NPC and nucleocytoplasmic transport pathways of eukaryotic cells can be hijacked during several viral infections. Many viruses have a nuclear replication phase and must interact with the NPC during their life cycle to gain access to the nuclear compartment. Viruses such as Adenovirus, HSV-1, Influenza A, HIV-1, and HBV, with a nuclear stage in their replication, require NPCs or components of the nucleocytoplasmic transport machinery to remove the capsid coat from the viral genome allowing import into the nucleus.

The *Retroviridae* family of viruses (especially HIV-1) represent a well-characterized example of viruses with a nuclear stage in their life cycle. These viruses must, therefore, interact with NPCs and exploit the nucleocytoplasmic transport pathway during infection. Retroviruses have an RNA genome that is reverse transcribed into cDNA in the cytoplasm before import into the nucleus and integration into the host cell genome. During reverse transcription, viral and cellular proteins bind the cDNA to form the pre-integration complex (PIC) (Jayappa et al., 2012). Many viral proteins incorporated into the PIC contain NLS domains that contribute to nuclear import of the viral genome, via a still undefined mechanism. Interestingly, HIV-1 infection can alter NPC morphology, significantly decreasing the levels of 18 of the 30 Nups (Monette et al., 2011). Several siRNA screens have reported that nine different Nups (Nup85, Nup98, Nup107, Nup133, Nup153, Nup155, Nup160, Nup214, and Nup358) and two transport proteins (CRM1 and transportin 3) are host factors required for HIV infection (Brass et al., 2008; König et al., 2008; Zhou et al., 2008). The specific roles of these Nups in the viral life cycle have been partially defined. Nup358 interacts with the viral capsid, promoting docking of the PIC at the NPC and regulating karyopherin dependent import (Di Nunzio et al., 2012; Lin et al., 2013; Schaller et al., 2011). Nup153 aids PIC exit from the NPC and release into the nucleoplasm, through a still undefined mechanism (Di Nunzio et al., 2013; Lee et al., 2010; Matreyek and Engelman, 2011). Additionally, HIV infection increases the intranuclear localization of Nup62, Nup98, and Nup153, where they promote viral genome integration (Ao et al., 2012; Di Nunzio et al., 2012). Finally, export of the transcribed viral RNA also requires the nuclear transport machinery, and it is mediated by interactions between the viral Rev protein and CRM1 (Bogerd et al., 1998).

The nuclear transport of influenza virus proteins has also been characterized. In influenza A virus infected cells, PB1 and PA form a dimer and enter the nucleus interacting with Kapβ3. PB2 separately enters the nucleus where it interacts with the PA-PB1 dimer. Additionally, two NLSs on NP allow transport of RNPs into the nucleus. Nuclear import of incoming vRNPs and newly distributed NPs are mediated

by their exposed NLS. Once inside the nucleus, NPs assemble into RNP for nuclear export, and these exported RNPs are no longer able to enter the nucleus. Exported RNPs have a hidden NLS that hides either in the NPs before RNP assembly or after the NPs are assembled to form the RNP. Therefore, selective exposure of NLSs plays a significant role in the regulation of the directionality in nuclear transport of their genome (Jamali et al., 2011).

Other viral families also exploit the nucleocytoplasmic transport pathway for their biology. For example, active CRM1-mediated export is required for the targeting of the adenovirus nucleocapsid to the nucleus, and capsid uncoating occurs through its association with Nup214 on the cytoplasmic side of the NPC (Strunze et al., 2005; Trotman et al., 2001). One study suggests that the adenovirus nucleocapsid tracks along microtubules on kinesin motors until it reaches the NPC (where it associates with Nup214) (Strunze et al., 2011). Nup358 then activates the kinesin heavy chain, which exerts a pulling force on the capsid and NPC facilitating capsid disassembly. The pulling force also disrupts NPC structure, promoting viral DNA entry into the nucleus. Similarly, Herpes virus capsid uncoating and genome import require docking with the NPC, facilitated by interactions with Nup214, Nup358, and Kapβ (Copeland et al., 2009; Ojala et al., 2000; Pasdeloup et al., 2009). In other viruses, such as Hepatitis B virus, the intact viral capsid appears to be imported to the nuclear face of the NPC, where interactions with Nup153 facilitate capsid uncoating and genome release into the nucleus (Schmitz et al., 2010). These observations demonstrate the importance of the NPC and, more specifically, Nup153, Nup214, and Nup358 in nucleocapsid uncoating and genome localization for several viral infections.

Some viruses also manipulate the nuclear transport of host cargoes to alter the cellular environment favorably for viral proliferation. The blocking of host mRNA export to limit host cell protein production, and consequently immune responses to viral infection, is an example of such a viral strategy (Yarbrough et al., 2014). Several picornaviruses, for instance, disrupt NPC transport by targeting specific Nups for degradation. Virus-derived 2A proteases produced by either poliovirus or human rhinovirus specifically cleave several Nups, including Nup62, Nup98, and Nup153, and a second rhinovirus protease (3C protease) also cleaves Nup214 and Nup358 (Belov et al., 2004; Castelló et al., 2009; Ghildyal et al., 2009; Gustin and Sarnow, 2001, 2002; Park et al., 2008, 2010; Watters and Palmenberg, 2011). Nup98, in particular, is a very early Poliovirus target and unlike other nucleoporins, Nup98 degradation does not require viral replication (Park et al., 2008). Thus, Nup98 may be targeted by a protease delivered with the infecting viral particle. Proteolysis of these Nups disrupts NPC permeability, inhibiting nuclear transport.

The encephalomyocarditis virus (EMCV) alters nuclear trafficking by modify-

ing specific Nups. The ECMV leader protein hyperphosphorylates Nup62, Nup153, and Nup214 and suppresses the activity of RanGTPase, altering the Ran gradient, and consequently changing the transport capacity of the nuclear transport system (Porter et al., 2006; Porter and Palmenberg, 2009). Other viruses can disrupt NPC-mediated transport without physically altering the NPC structure. The SARS-CoV and Ebola virus inhibit specific import pathways by competing for binding sites on certain Kaps or by causing the mislocalization of Kaps (Frieman et al., 2007; Mateo et al., 2010; Reid et al., 2006). Consequently, both these viruses inhibit the nuclear localization of STAT1, an important transcription factor for the induction of interferon stimulated genes that participate in immune response (Figure 1.8) (Frieman et al., 2007; Kopecky-Bromberg et al., 2007; Mateo et al., 2010; Reid et al., 2006).

Interfering with host mRNA export is a common strategy for several viruses since inhibiting host cell protein production can limit immune responses to viral infection. Although the Vesicular Stomatitis virus (VSV) replicates in the cytoplasm of infected cells, it inhibits mRNA export in part through the viral Matrix (M) protein, which binds to and inhibits a component of the mRNA export pathway. The VSV M protein is targeted to the nucleus by an internal NLS sequence. Inside the nuclei, M protein interacts with Rae1/Gle2, a Nup98 interacting protein (Enninga et al., 2002; Faria et al., 2005; von Kobbe C et al., 2000; Quan et al., 2014).

Influenza viruses are negative strand segmented RNA viruses that replicate in the nucleus of infected cells. During influenza A virus infection, the viral protein NS1 interacts with several components of the mRNA export machinery, including Nup98, Rae1, NXF1, NXT1 and E1B-AP5, causing degradation of Nup98 and a block in the mRNA export pathway (Satterly et al., 2007). Intranuclear viral RNA is exported through the CRM1 export pathway rather than the NXF1 dependent pathway utilized by most host mRNAs for export (Elton et al., 2001; Watanabe et al., 2001). The Influenza protein NS1 also disrupts mRNA processing, cleaving several polyadenylation factors (Satterly et al., 2007). Influenza polymerase functions to polyadenylate viral RNA itself, thus avoiding deleterious effects on viral RNA production and export. Therefore, the inhibition of host cell mRNA processing and export pathways by the influenza virus acts to both prevent the expression of immune effectors and promote viral replication (Figure 1.8).

Interestingly, cellular immune responses (increased IFN-γ) can combat viral interference with cellular mRNA export by increasing levels of components of the mRNA export pathway, including Nup98 and Rae1 proteins. The promoters of both the *NUP98* and *RAE1* genes contain IFN-γ response elements. Presumably, increased IFN-γ produced in response to a viral infection leads to upregulation of both these proteins to counteract viral pathogenesis. Interferon-stimulated increased expression of Nup98 and Rae1, or exogenous overexpression of Nup98, Rae1, or NXF1,

can reverse the mRNA export block induced by influenza or VSV infection (Enninga et al., 2002; Satterly et al., 2007). Also of note is the fact that Nup98 itself has an essential role in immune response, especially in the IFN-γ pathway. This cytokine increases the expression and intranuclear pool of Nup98 (Enninga et al., 2002) and Nup98, in turn, can affect the epigenetic state of interferon response genes allowing for faster reinduction of their expression (Light et al., 2013).

The NPC, Nups, and NTFs can also play a role in the life cycle of viruses whose genome replication occurs in the cytoplasm, such as positive-strand RNA viruses. Several Hepatitis C virus (HCV) non-structural proteins have been shown to interact with Kaps, leading to disruption of specific nuclear transport pathways. The HCV NS5A protein interacts with Kapβ3 and disrupts its nuclear import pathway (Chung et al., 2000). Proteomics analysis of HCV-infected cells revealed that viral NS3 interacts with two other importins and two exportins (Germain et al., 2014). Cells infected with HCV relocate several nucleoporins and NTFs into viral sites of replication and assembly (membranous web). HCV infection also causes a slight increase in the levels of some Nups, and their depletion decreases the production of infectious virus. Depletion of different Nups shows effects at various stages of the HCV life cycle. Nup98 and Nup153 for example, possibly participate in replication of the viral genome, while Nup155 more likely has a role in virion assembly (Neufeldt et al., 2013). Various HCV proteins also contain NLS and NES sequences, and they can interact with NTFs such as Kapα5, Kapβ3, and CRM1. Interactions between HCV proteins and Kapα5 play a role during the replication phase of the HCV life cycle, while the interaction between HCV proteins and Kapβ3 have roles during early replication and early assembly of HCV virions (Levin et al., 2014b).

It is possible that the interaction between viral molecules and host nucleocytoplasmic transport components is not restricted to the cytoplasm of HCV-infected cells. Interestingly, four of the ten HCV proteins, core, NS2, NS3, and NS5A, contain putative NLS sequences, and can enter the nucleus when mutated or produced outside of the context of viral infection (de Chassey et al., 2008; Ide et al., 1996; Isoyama et al., 2002; Kim and Nikodem, 1999; Levin et al., 2014b; Suzuki et al., 2005). However, only core has been suggested to enter the nucleus of HCV-infected hepatocytes (Cerutti et al., 2011). Viral RNA molecules can also be seen in the nuclei of HCV-infected cells (as observable in images and movies from Fiches et al., 2016; Shulla and Randall, 2015). Moreover, several nuclear host proteins are hijacked by cytoplasmic positive-strand RNA viruses to participate in their life cycle, with many remaining nuclear upon infection. Most of these nuclear factors are RNA-binding proteins that aid vRNA translation, replication, and virion assembly (Lloyd, 2015), such as the DEAH/RHA helicases, discussed below.

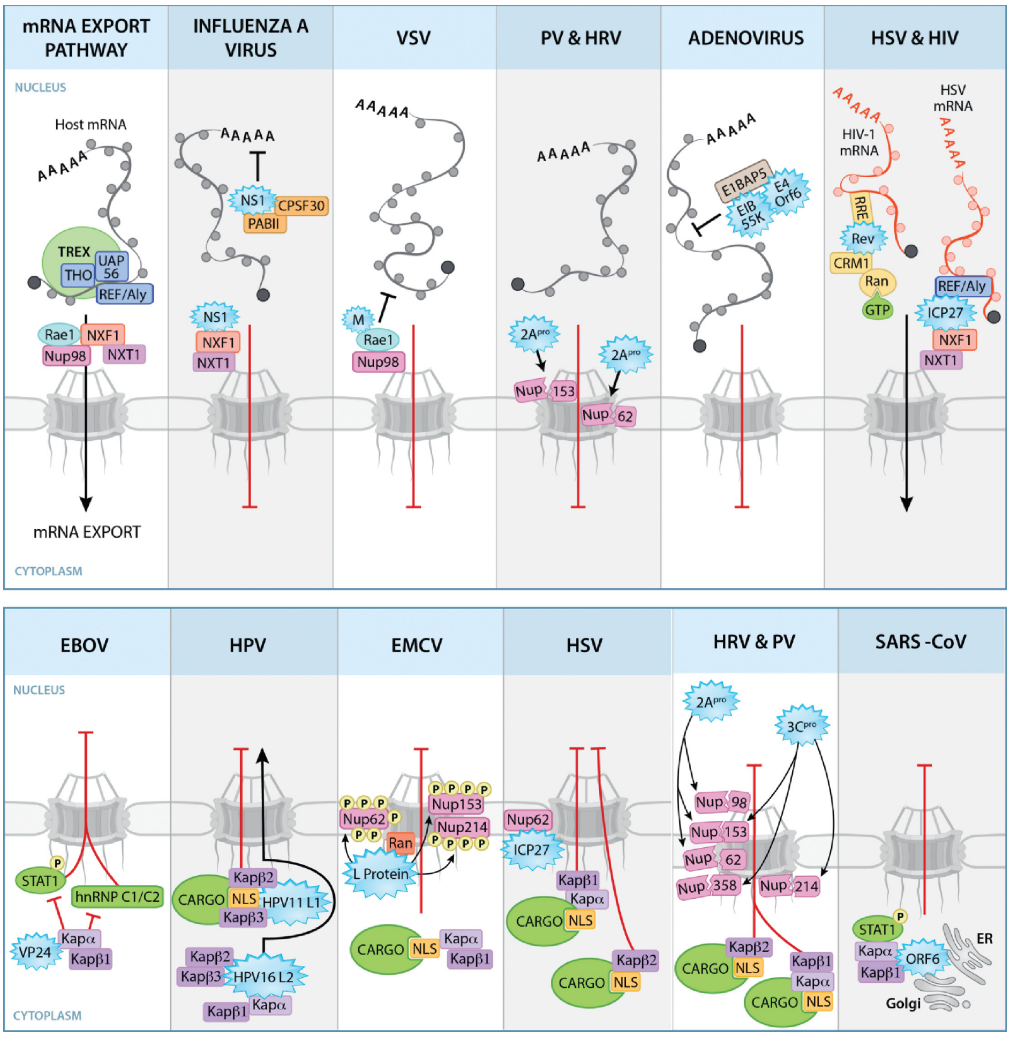


Figure 1.8: Viral strategies to disrupt nucleocytoplasmic trafficking

Top: From left to right, diagram of host mRNA export pathway. Viral proteins (blue starbursts) disrupt mRNA nuclear export. IAV NS1 disrupts mRNA processing and export. VSV M protein causes mRNA nuclear export block by binding Rae1 and Nup98. 2Apro of PV and HRV cleaves Nups disrupting NPC architecture. AdV E1B-55K and E4orf6 proteins disrupt NXF1-mediated mRNA export. Other viruses (e.g. HSV and HIV), utilize nucleocytoplasmic transport to promote viral RNA export. The HSV protein ICP27 interacts with REF/Aly and NXF1 to promote export of vRNAs. HIV-1 Rev protein exports unspliced or partially spliced vRNAs using CRM1 and RanGTP. **Bottom:** Viral proteins disrupt nuclear protein transport. From left to right, EBOV-VP24 binds Kapαcausing cytoplasmic accumulation of STAT1 and hnRNP C1/C2. In HPV, HPV11 L1 binds Kapβ2/β3, to disrupt import, while HPV16 L2 enters the nucleus via Kapβ2, Kapβ3 and Kapα/β1. L protein of EMCV hyperphosphorylates Nups and binds Ran to inhibit import. HSV ICP27 interacts with Nup62 and blocks nuclear import via Kapα/β1 and Kapβ2. 2Apro and 3Cpro of HRV and PV degrade Nups and block nuclear import via Kapα/β1 and Kapβ2. SARS-CoV ORF6 protein tethers STAT1-Kapα/β complex to ER/Golgi disrupting nuclear import of phosphorylated STAT1. Adapted from Yarbrough et al., 2014. Reprinted with permission from Wiley Online Library (John Wiley and Sons), original publication DOI: https://dx.doi.org/10.1111/tra.12137

1.9.2 DEAH/RHA helicases and viral infections

Most viruses require RNA helicases to facilitate their replication. VIruses from a few families (e.g. *Flaviviridae, Poxviridae* and *Potyviridae*) are known to encode their RNA helicases (Ranji and Boris-Lawrie, 2010a; Gross and Shuman, 1996, 1998; Kwong et al., 2005). Helicase activity of these RNA helicases is an essential function for the cognate virus. These viral RNA helicases exhibit the same modular structure as cellular RNA helicases and provide genetically separable catalytic and scaffold functions (Ranji and Boris-Lawrie, 2010a). Other viral families rely on hijacking the host cellular machinery, including RNA helicases, to facilitate their replication (Ranji and Boris-Lawrie, 2010a; Jeang and Yedavalli, 2006). Usually, viruses that synthesize their genomes within the nucleus of the host cell utilize host cellular RNA helicases. However, many cytoplasmic viruses have also been described to hijack host helicases during their life cycle, including many nuclear RNA helicases (Lloyd, 2015).

A number of host RNA helicase genes show altered expression upon viral infection. Host RNA helicases are involved in most stages of viral life cycle, suach as virus entry by receptor-mediated endocytosis, reverse transcription, integration into the host genome, transcription, RNA processing, export from the nucleus, polysomal translation, and viral assembly (Ranji and Boris-Lawrie, 2010a). The role of RNA helicases in the replication cycles of different viruses has not been systematically investigated. However, siRNA screens continue to identify DExD/H-box helicases as essential cofactors for the replication of specific viruses. Proteomics screens with viral proteins as baits also often identify interacting DExD/H-box helicases (Fullam and Schröder, 2013).

RNA helicases have a dual role, participating in viral infection as well as the host immune response. In some viral infections, RNA helicase activity benefits the virus by promoting viral gene expression and squelching the antiviral response. In others, RNA helicase activity benefits the host by sensing viral nucleic acid and triggering antiviral responses (Sharma and Boris-Lawrie, 2012). A large number of helicases participate in immune response and various viruses hijack this class of proteins during their life cycles. Here we will focus on the role of DHX9 in viral infection (Ranji and Boris-Lawrie, 2010a; Fullam and Schröder, 2013).

DHX9 is required for the replication of several different viruses. Interestingly, the structure of DHX9 is quite similar to that of the Flavivirus NS3 helicase (Schütz et al., 2010). A role for DHX9 in the life cycle of two viruses from the *Flaviviridae* family has been described. DHX9 and other NF/NFAR proteins can bind the 5' and 3' UTR of the HCV RNA, and they are required for HCV replication. The currently proposed model hypothesizes that DHX9 is involved in the creation of a circular loop

structure between the 3' and 5' UTR of the HCV RNA aiding in the coordination of translation and replication (Isken et al., 2007). Consistent with this idea, partial depletion of DHX9 causes a gradual reduction of HCV RNA and protein (He et al., 2008), suggesting DHX9 is a necessary host factor for HCV infection. DHX9 can also interact with the 3' and 5' UTR of the RNA from another *Flavivirus*, bovine viral diarrhea virus. As described for HCV, DHX9 interacts with this viral RNA along with other NF/NFAR proteins and this protein complex is thought to act as a scaffold to promote a configuration of the vRNA that aids in its replication (Isken et al., 2003; Ranji and Boris-Lawrie, 2010a; Fullam and Schröder, 2013).

DHX9 plays a role in the replication of picornaviruses. DHX9 depletion decreases viral replication of Foot-and-mouth disease virus (FMDV). FMDV infection also changes the subcellular localization of DHX9 from prominently nuclear to prominently cytoplasmic. In the cytoplasm of infected cells, DHX9 localizes to sites of viral replication, where it associates with the 5' UTR of the FMDV genome and with viral replication proteins, 2C and 3A. DHX9 is speculated to connect the 5' and 3' UTRs of the vRNA to switch between replication and protein synthesis, which may occur on a circular polysome (Lawrence and Rieder, 2009; Ranji and Boris-Lawrie, 2010a; Fullam and Schröder, 2013).

Influenza virus replication can exploit DHX9 (Lin et al., 2012). DHX9 interacts with viral NS1 in an RNA-dependent manner to enhance viral replication and transcription. Interestingly, Influenza virus-induced IFN-α production is decreased in cells upon DHX9 deletion (Zhang et al., 2011), indicating a possible dual function for DHX9, mediating IFN-induction in response to a virus that also actively recruits it to facilitate viral replication (Ranji and Boris-Lawrie, 2010a; Fullam and Schröder, 2013).

Many steps of the HIV life cycle make use of DHX9 (Li et al., 1999). DHX9 can interact with the double-stranded stem of the HIV trans-activation response element (HIV TAR) leader RNA to increase HIV long terminal repeat (HIV LTR) promoter transcription (Fujii et al., 2001). The interaction between DHX9 and HIV TAR is decreased by phosphorylation of DHX9's dsRBD by PKR (Sadler et al., 2009). DHX9 can also release incompletely spliced HIV RNAs from spliceosomes for export out of the nucleus, increasing the levels of unspliced HIV CTE (constitutive transport element) and RRE (Rev response element) containing mRNAs (Li et al., 1999). DHX9 also facilitates the export of these RRE- and CTE-containing viral RNAs by a nuclear export pathway mediated by Sam68 and NXF1 (Reddy et al., 2000). In the cytoplasm, DHX9 interacts with the 5' UTR of HIV RNAs to facilitate translation (Bolinger et al., 2010). Through its interaction with HIV gag and HIV RNA, DHX9 facilitates the annealing of $tRNA^{Lys3}$ to HIV gag mRNA and this entire complex is then incorporated into assembling virions (Xing et al., 2011). HIV virions that do

not contain DHX9 are less infective and have a defective reverse transcription (Roy et al., 2006). It is possible that the DHX9-HIV gag complex remodels the viral RNA to promote a switch from translation to incorporation into assembling virions (Xing et al., 2011). Thus, DHX9 appears to be involved in most steps regulating the expression of HIV genes and assembly of new virions (Ranji and Boris-Lawrie, 2010a; Fullam and Schröder, 2013).

DHX9 also participates in translation initiation in several other lymphotropic retroviruses; it interacts with the posttranscriptional control element (PCE), a stem-loop structure located at the 5' UTR of retroviral (and some cellular) mRNAs. DHX9 is likely to recognize common structural features of the PCE rather than a specific RNA sequence and to promote translation by facilitating ribosome access (Bolinger et al., 2007). DHX9 interaction with the 5' UTR requires the ATP-binding activity of the helicase domain, and it increases polysome association (Hartman et al., 2006; Bolinger et al., 2010). DHX9's helicase activity rearranges the RNP to facilitate ribosome scanning and translation initiation. This rearrangement may also facilitate protein-protein interactions that secure a circular polysome for efficient translation reinitiation (Hartman et al., 2006). In the case of the simian Mason-Pfizer monkey virus (MPMV), DHX9 interacts not only with the 5' UTR to facilitate translation (Hull and Boris-Lawrie, 2002), but also with the 3' UTR in a region necessary for nuclear export of the unspliced viral transcript (Boris-Lawrie et al., 2001). This interaction of DHX9 with the CTE in the 3' UTR of the vRNA (Tang et al., 1997) facilitates nuclear export by recruiting NXF1 (Grüter et al., 1998). It is possible that DHX9 binds to the MPMV vRNA in the nucleus and facilitates consecutive remodeling of the viral RNP for nuclear export and translation in the cytoplasm (Hull and Boris-Lawrie, 2003; Ranji and Boris-Lawrie, 2010a; Fullam and Schröder, 2013).

In summary, DHX9 interacts with the terminal regions of many viral RNAs, participating in different steps of these viral life cycles. In this dissertation we focus on the role Nups, NTFs, and DHX9 play in the life cycle and vRNA metabolism of two viruses from the *Flaviviridae* family, the Hepatitis C virus and the Zika virus.

1.10 The *Flaviviridae* family of viruses

The *Flaviviridae* family of positive-strand RNA viruses consists of three genera: *Flavivirus, Pestivirus*, and *Hepacivirus*. A fourth genus, *Pegivirus*, has also recently been proposed. The viruses from this family have diverse biological properties but share similarities in virion morphology, genome organization, and replication strategy (Knipe and Howley, 2013). The family *Flaviviridae* includes many viruses of medical importance, such as dengue, yellow fever, West Nile, Japanese encephalitis, Zika

and Hepatitis C viruses (Burrell et al., 2017).

These enveloped virions contain a lipid bilayer with envelope (E) glycoproteins surrounding a nucleocapsid composed of a single-stranded, positive-sense RNA genome bound by multiple capsid proteins. Viral entry into host cells involves receptor-mediated endocytosis. Once in the endosome, the low pH in this organelle induces fusion of the virion envelope with cellular membranes, and nucleocapsid uncoating to release the RNA genome into the cytoplasm. RNA replication is cytoplasmic, occurring in close association with intracellular membranes. Progeny virions are assembled by budding into an intracellular membrane compartment, most likely the ER, and then transiting through the secretory pathway for release at the cell surface (Figure 1.9) (Knipe and Howley, 2013).

The viral genome has three discrete roles during the viral life cycle: mRNA for translation of viral proteins, a template for RNA replication, and genetic material packaged in new viral particles. All *Flaviviridae* genera share similar genome organization. Viral proteins are translated as a single polyprotein that is cleaved by a combination of host and viral proteases. The N-terminal portion of the polyprotein contains the structural proteins, with the nonstructural proteins in the remainder. Among the nonstructural proteins, a serine protease, RNA helicase, and an RNA-dependent RNA polymerase are found in similar locations in the polyproteins of all genera of *Flaviviridae* (Figure 1.9) (Knipe and Howley, 2013).

In this dissertation, we focus on two viruses from the *Flaviviridae* family, the Hepatitis C virus, from the genus *Hepacivirus*, and the Zika virus, from the genus *Flavivirus*. The genus *Hepacivirus* contains only the human pathogens hepatitis C virus and related viruses of primates. The genus *Flavivirus* contains more than 75 viruses, with approximately 30 viruses of medical importance, such as yellow fever virus, the five dengue viruses, West Nile virus, Japanese encephalitis virus, Zika virus, and several tick-borne encephalitis viruses. *Flaviviridae* viruses rank among the most important human viral pathogens on a global scale (Burrell et al., 2017).

1.10.1 Hepatitis C virus

The Hepatitis C virus was identified in 1989 as the primary cause of non-A, non-B hepatitis infections (Choo et al., 1989). Currently, it is estimated that the worldwide burden of HCV sits at 170 million, roughly 3% of the global population. HCV remains a leading cause of chronic liver disease, cirrhosis, and hepatocellular carcinoma. In the Western hemisphere, HCV infection is the leading cause of liver transplantation (Mohd Hanafiah et al., 2013; Tsoulfas et al., 2009). The global distribution of HCV varies widely, but the estimated rate of HCV infection in Europe

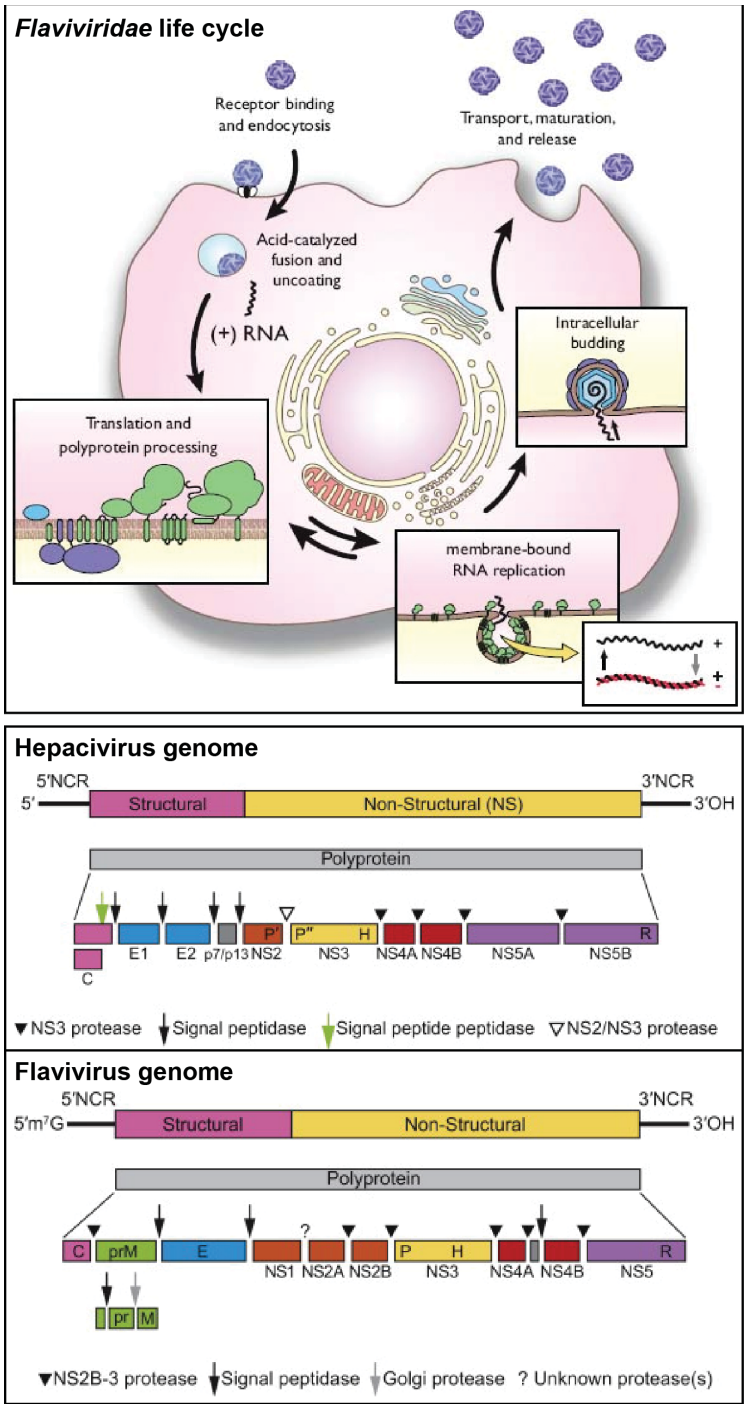


Figure 1.9: The *Flaviviridae* family of viruses

Top: Life cycle of *Flaviviridae* family viruses. Reproduced with permission from Knipe, DM, Howley, PM, 2013. Fields virology, 6th ed. Wolters Kluwer/Lippincott Williams and Wilkins Health, original publication ISBN: 9781451105636 **Bottom:** *Hepacivirus* and *Flavivirus* genome organization and polyprotein processing with generated viral proteins indicate. Reproduced with permission from King, AMQ et al., 2012. Virus Taxonomy. Ninth Report of the International Committee on Taxonomy of Viruses. Academic Press, original publication DOI: 10.1016/B978-0-12-384684-6.00086-0.

is around 0.5% of the population, 0.7% of the total Canadian population and 1.3-1.5% in the USA and Australia. The prevalence is higher in several countries in the north of the African continent, the Middle East, and South Asia (CDC, 2016). Based on its evolutionary history, HCV is classified into seven genotypes, each further divided into numerous subtypes. HCV genotypes differ by more than 30% at the nucleotide level, and they vary in worldwide distribution, disease progression, and susceptibility to treatment (Knipe and Howley, 2013).

Once the infection is established, HCV is rarely naturally cleared. More often, it establishes a persistent chronic infection (Micallef et al., 2006). Yet the life cycle of HCV lacks a DNA intermediate, which precludes it from ever taking a latent form that can be camouflaged from immune detection (Chung et al., 2014). Instead, it must continuously replicate, mutate sufficiently to avoid a strong humoral response, and rely on the rapid production of virus and continuous cell-to-cell spread to survive (Chen and Morgan, 2006).

HCV selectively infects hepatocytes by using CD81 and SCARE 1 as its primary receptors. Upon clathrin-mediated endocytosis, the virus enters the cell and the pH decrease in the endosomal compartment allows viral uncoating and release of its positive-strand RNA genome into the cytoplasm (Fénéant et al., 2015; Bianchi et al., 2011; Lavillette et al., 2007).

The HCV genome is 9.6 kb long and encodes a single, long open reading frame. An internal ribosomal entry site (IRES) recruits ribosomes to the 5' UTR for the synthesis of a 3011 amino acid polyprotein at the ER. The polyprotein is then cleaved into three structural (core, El, and E2) and seven nonstructural proteins (p7, NS2, NS3, NS4A, NS4B, NS5A, and NS5B). The junctions between core/El, E1/E2, E2/p7, and p7/NS2 are cleaved by the host ER signal peptidase (Knipe and Howley, 2013). The remaining junctions are cleaved by viral encoded proteases, NS2/3 cysteine protease and NS3-4A serine protease (Knipe and Howley, 2013).

The cleaved viral proteins associated with the ER membrane induce morphological changes to create the membranous web. Replication occurs in double-membrane folds within this membranous web (termed replication complexes), where vRNA and proteins can be relatively isolated from the rest of the cytoplasm. At replication complexes, the positive-strand genome is copied into a complementary negative-strand RNA by the RNA-dependent RNA polymerase NS5B. The negative-strand RNA serves as a template for additional rounds of replication (also by NS5B). Approximately ten positive-strand RNA genomes are made per negative-strand RNA template. Newly synthesized positive sense viral genomes can then be transcribed, translated, or packaged into nascent virions.

HCV genomes destined for packaging are coated with core protein and bud into the ER lumen, resulting in the envelopment of HCV particles with ER-derived

membranes studded with HCV glycoproteins. Assembly and secretion of infectious HCV particles are tightly linked to the very low density lipoprotein (VLDL) biogenesis pathway. Infectious HCV secretion requires the microsomal triglyceride transfer protein (MTP), which is critical for VLDL formation, as well as apolipoprotein B (ApoB) (Huang et al., 2007; Gastaminza et al., 2008) and/or apolipoprotein E (ApoE) (Jiang and Luo, 2009).

Virions are transported through the cellular secretory pathway, where the glycoproteins are further matured by glycosylation in the Golgi and are released from cells (Gastaminza et al., 2008). Cell-free infectious HCV virions have a low buoyant density and circulate in complex with VLDLs as lipoviral particles (Thomssen et al., 1992). Such complexes may be composed of virions and individual lipoprotein. The HCV life cycle is depicted in Figure 1.10.

Of note, recent work by Neufeldt and collaborators has uncovered an intricate interaction network between HCV viral proteins and the NPC, identifying a novel function for the NPC and nuclear transport machinery in positive-strand virus replication and immune evasion. These novel interactions between the nuclear transport machinery and viral proteins were shown to support the formation of the membranous web and limit host cell immune activation. Multiple components of the nuclear transport pathway were found to interact with HCV proteins and accumulate in the membranous web. Additionally, several Nups and NTFs were identified as host factors necessary to support HCV infection. These observations support a role for the nuclear transport machinery in the formation of distinct viral compartments that maintain a selective barrier with the surrounding cytosol. This selective barrier, created at the membranous web by Nups and NTFs, limits access to proteins that negatively impact viral replication, such as RIG-I-like receptors, while allowing traffic of proteins containing an NLS sequence, such as viral proteins (Neufeldt, 2014; Neufeldt et al., 2013; Levin et al., 2014b).

Despite the cytoplasmic life cycle of HCV, several viral proteins contain nuclear localization signals and nuclear export signals (Levin et al., 2014b). These NLSs and NESs allow trafficking of viral proteins in and out of the membranous web, as described above. However, HCV core, NS3, and NS5A proteins can also be seen in the nuclei of infected cells. So far, no known nuclear function of NS3 has been described in HCV-infected cells. Nuclear core and NS5A have been shown to modulate host gene expression to promote cell survival, immune evasion, and increased ribosome biogenesis (Bonamassa et al., 2015). Several nuclear host proteins are also known to participate in HCV vRNA replication and translation. Most of these nuclear factors interact with the viral RNA, but only some show relocation to the cytoplasm of infected cells (Lloyd, 2015). Since core, NS3, and NS5A are all RNA binding proteins that interact with the HCV vRNA, they may possibly carry some vRNA into the

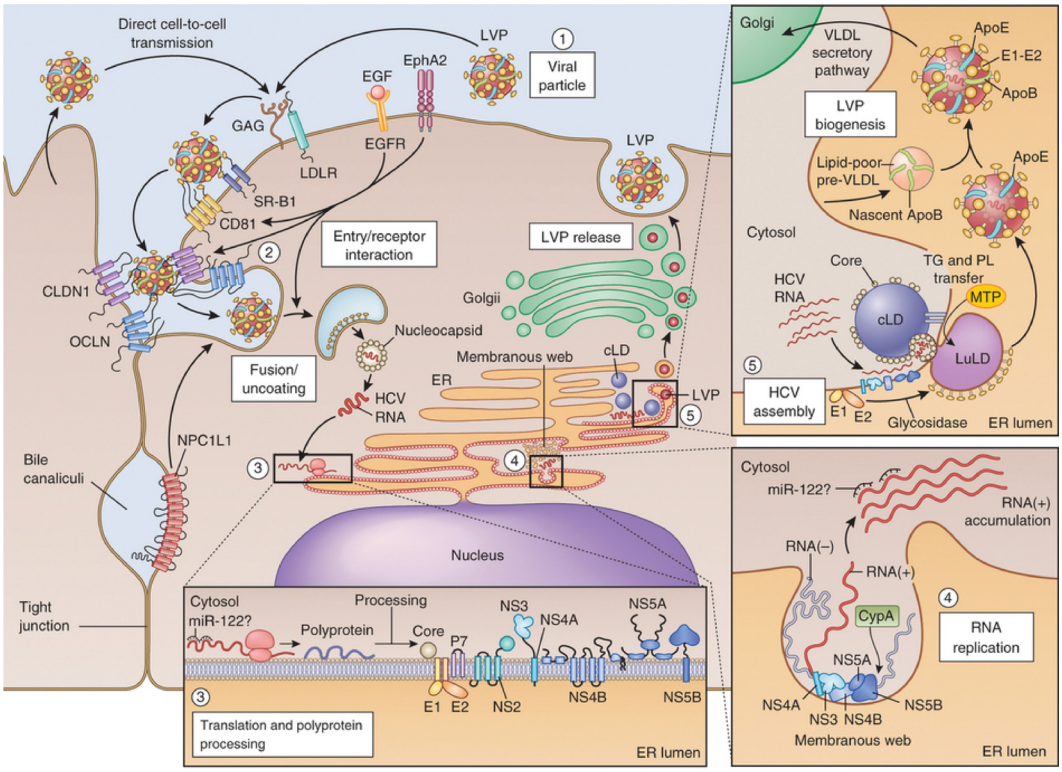


Figure 1.10: The Hepatitis C virus life cycle

Interaction of extracellular HCV LVPs (1) with cellular surface receptors initiates the entry process (2), which can also occur from direct cell-to-cell transmission. After pH-dependent fusion and uncoating, the incoming HCV genome is translated and the resulting polyprotein processed (bottom inset and (3)). Replication takes place in ER-derived membrane spherules (membranous web, bottom right inset, (4)). In the assembly and release process (top right inset, (5)), core protein is transferred from cytoplasmic lipid droplets (cLDs) to form nucleocapsids that, assisted by NS5A, are loaded with RNA. The p7, NS2, and NS3-NS4A proteins are also involved in coordination of assembly. HCV virion morphogenesis is coupled to the VLDL pathway, and particles are produced as lipoviroparticles (LVPs). Particles released from cell culture have less ApoB association. EphA2, ephrin receptor type A2; GAG, glycosaminoglycans; PL, phospholipids; TG, triglycerides; LuLD, luminal lipid droplets; MTP, microsomal transfer protein; CypA, cyclophilin A. Reproduced with permission from Scheel and Rice, 2013 Nature Publishing Group, original publication DOI: http://dx.doi.org/10.1038/nm.3248.

nuclei of infected cells, where it could interact with the nuclear factors necessary for its translation and replication.

1.10.2 Zika virus

The Zika virus (ZIKV) is a mosquito-borne *Flavivirus* discovered from the *Rhesus macaque* in the Zika forest of Uganda in April of 1947 (Hayes, 2009). ZIKV is an enveloped, icosahedral virus with a non-segmented, single-stranded positive-strand RNA genome. Like other *Flaviviruses*, ZIKV is transmitted to humans and non-human primates via arthropod vectors, namely mosquitoes that bite vertebrate animals. The virus is primarily transmitted by the *Aedes aegypti* and *Aedes albopictus* mosquitoes (Musso et al., 2014), which are found abundantly throughout warm and humid territories (CDC, 2016). Unlike any other known *Flaviviruses*, ZIKV can also be sexually transmitted in humans, passed via laboratory exposure and blood transfusion, or from infected mothers to their fetuses through vertical transmission, (Musso et al., 2015; D'Ortenzio et al., 2016).

ZIKV is hard to target since around 80% of the people who are infected are asymptomatic, while those who are symptomatic show an acute onset of fever, arthralgia, maculopapular rash, and nonpurulent conjunctivitis. In Brazil, the Dominican Republic, Puerto Rico, and other american countries ZIKV RNA has been seen in the tissues of many infants with microcephaly (Pan American Health Organization and World Health Organization, 2017). These infants show abnormal brain development. As a result, they have poor motor skills, speech, hearing, vision, abnormal facial features, seizures, and are intellectually disabled (Butler, 2016; Zanluca et al., 2015). The Zika virus has also been seen in mothers who lost their fetus during pregnancy (Zammarchi et al., 2015). The Brazil Ministry of Health has shown that there has been an increase in the number of cases of microcephaly in 2015 from approximately 0.5 to 20 cases per 10,000 live births (CDC, 2016) (Butler, 2016).

The ZIKV genome ($\approx$ 11 kb) is a positive single-stranded RNA that encodes a polyprotein flanked by 5' UTR (capped) and 3' UTR with secondary structures. The ZIKV vRNA also contains N6-adenosine methylation (m6A), which may regulate genome stability (Gokhale et al., 2016; Lichinchi et al., 2016). The vRNA from the incoming virion can be used directly for protein translation, producing viral proteins required for viral RNA replication. The translated polyprotein is then cleaved by viral and host proteases to produce ten individual viral proteins (Knipe and Howley, 2013). Polyprotein processing occurs co- and post-translationally on the ER membrane. There are three structural proteins (capsid, prM, and E), which are found in

the viral particles, and seven nonstructural proteins (NS1, NS2A, NS2B, NS3, NS4A, NS4B, and NS5), which participate in viral replication (Knipe and Howley, 2013).

As is common among *Flaviviruses*, replication of the ZIKV genome occurs via a negative-strand replicative intermediate (minus-strand RNA) (Garcia-Blanco et al., 2016). The negative-strand RNA is produced by the RNA-dependent RNA polymerase NS5, as are the newly synthesized positive-strand viral genomes. The helicase activity of NS3 also participates in RNA replication. Other NS proteins contribute to the assembly and maintenance of replication complexes on intracellular membranes (Welsch et al., 2009).

The formation of membrane-associated replication complexes, which may serve to increase the local concentration of nonstructural proteins, is a common feature of positive-strand RNA replication. The capsid protein binds to the viral genome to form the nucleocapsid (Jones et al., 2003), which is further enveloped by a lipid bilayer membrane derived from the host cell. The viral envelope is embedded with prM/M and E, two transmembrane viral glycoproteins that mediate viral entry into the host cell (Rey et al., 1995). E is the receptor-binding and fusion protein, while prM functions as a chaperone of E protein, preventing premature fusion (Guirakhoo et al., 1992; Kuhn et al., 2002). In the Golgi, as the virion exits the cell via the secretory pathway, furin cleaves prM to generate M protein on the surface of mature virions (Stadler et al., 1997; Kuhn et al., 2002).

ZIKV entrance to target cells (such as human fibroblasts) is mediated by several cell-surface adhesion factors, including TAM receptor proteins (AXL and Tyro3) and T cell immunoglobulin and mucin domain 1 (TIM-1) (Hamel et al., 2015). Notably, AXL is highly expressed in radial glia cells, orthotopic retinal ganglion cells, astrocytes, and microglia in the developing human brain (Nowakowski et al., 2016) as well as in trophoblast progenitors in placenta (Tabata et al., 2016), partially explaining the high efficiency of ZIKV infection in these cell types. However, deletion of AXL does not affect ZIKV infection in human neural progenitors in monolayer or organoid cultures (Wells et al., 2016), underscoring the lack of definitive identification of bona fide cellular receptor(s) for ZIKV. Once inside the host cell, ZIKV hijacks cellular pathways for its replication and assembly, interfering with host cell proliferation and survival of neural progenitors.

One of the host cell pathways activated by ZIKV infection is autophagy (Ming et al., 2016), as exemplified by the formation of autophagosomes in ZIKV infected human skin fibroblasts (Hamel et al., 2015). This phenotype has been previously described for other positive-strand RNA viruses. For example, dengue virus and HCV-mediated modulation of cellular autophagy benefits their replication (Hamel et al., 2015; Heaton and Randall, 2010; Sir et al., 2012), while autophagy inhibition attenuates their replication in host cells (Hamel et al., 2015; Liang et al., 2016).

Transcriptional profiling of ZIKV-infected human neural progenitors has revealed dysregulation of autophagy-related genes (ATGs), including upregulation of ATG2A, ATG4A, ATG16L1, STK38L, RAB7A, and ULK1 and downregulation of LAMP2A and CASP2 (Tang et al., 2016). ZIKV infection of human fetal neural progenitors causes inhibition of the Akt-mTOR pathway, leading to aberrant activation of autophagy and defective neurogenesis (Liang et al., 2016).

Cell death, via caspase-3 activation and apoptosis, is also observed in ZIKV infected neural progenitors cells (in culture), in animal models, and in clinical samples of ZIKV-infected fetuses (Ming et al., 2016). ZIKV infection increases p53 levels in the host cell, p53 nuclear accumulation, and Ser15 phosphorylation (Ghouzzi et al., 2016). Consequently, p53 inhibitors can block the apoptosis induced by ZIKV in human neural progenitor cells (Zhang et al., 2016). These data indicate that ZIKV infection can interfere with the main survival and homeostasis pathways in the host cell.

Another striking consequence of ZIKV infection in neural progenitors is deregulated gene expression. Transcriptome profiling of ZIKV-infected human neural progenitors, brain organoids, and mouse cortical tissues revealed many genes with deregulated expression. These genes with altered expression are related to cell cycle, transcription, metabolism, cell death, DNA replication and repair, and viral responses (Ming et al., 2016). Notably, human cortical neural progenitors cells (in culture and in embryonic mouse cortex) infected with ZIKV show downregulation of many currently known microcephaly-associated genes (ASPM/MCPH5, CASC5/MCPH4, CENPF, Microcephalin/MCPH1, RBBP8, STIL/MCPH5, and TBR2) (Li et al., 2016b; Tang et al., 2016).

ZIKV infection can also affect epigenetic and epitranscriptomic regulation of the host cells, processes with critical roles in stem cells regulation and neurogenesis (Yao et al., 2016). In human cortical neural progenitors, ZIKV infection upregulates DNA oxidases (TET1,2,3), and downregulates DNA methyltransferases (DNMT1 and DNMT3A) (Tang et al., 2016), possibly allowing genome-wide demethylation. Among possible mRNA modifications, m6A is the most abundant, affecting RNA structure and function, such as mRNA decay, microRNA production, and translational control (Yue et al., 2015). ZIKV RNA is modified at m6A and 2'-O-methylated nucleosides, and depletion of m6A methyltransferases or m6A demethylases, respectively, increases or decreases infectious production of ZIKV and HCV in infected cells (Gokhale et al., 2016; Lichinchi et al., 2016). ZIKV infection in human HEK293 cells also alters host m6A mRNA profiles (Lichinchi et al., 2016). Future studies will be necessary to determine whether similar epitranscriptomic deregulation occurs in ZIKV infected neural progenitors, as well as if other viruses of the *Flaviviridae* family have similar effects in the epitranscriptome of their host cells.

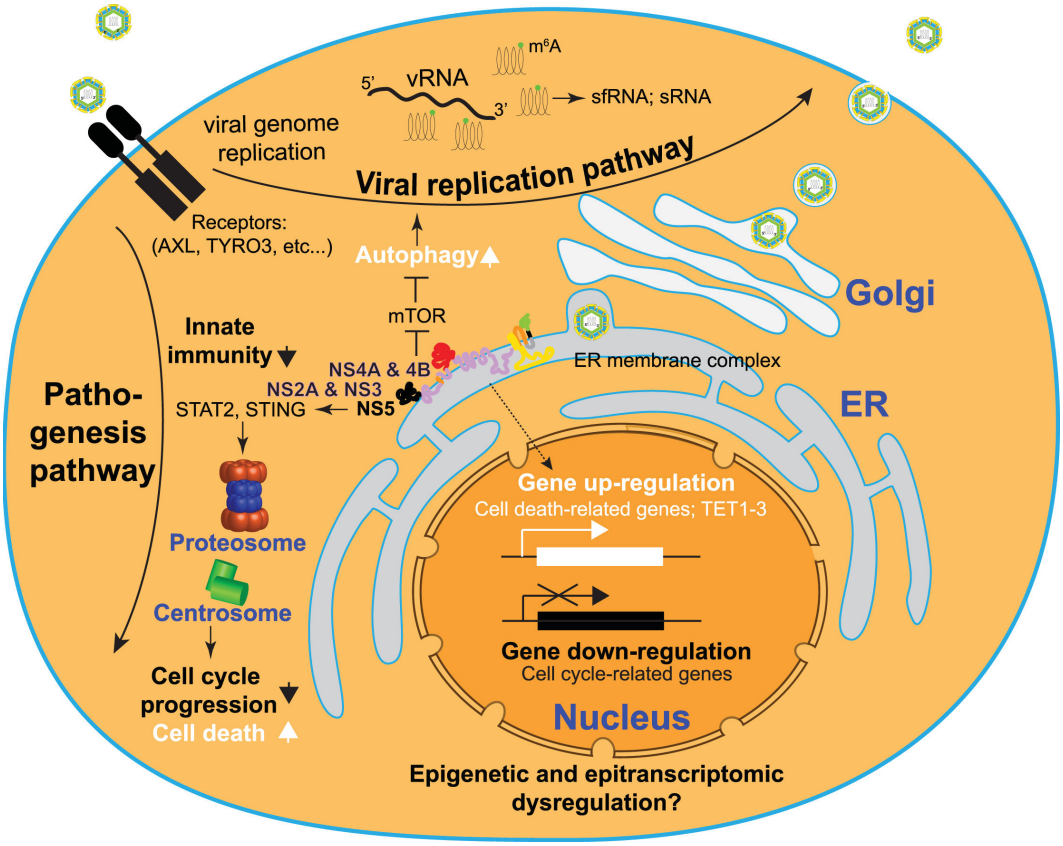


Figure 1.11: The Zika virus life cycle

ZIKV enters the host cell through receptor-mediated endocytosis (AXL and others). Viral RNA is then released and used for synthesis of a polyprotein at the endoplasmic reticulum. ZIKV proteins enhance viral replication by blocking two pathways that inhibit viral replication, interferon and mTOR signaling. NS5 protein promotes degradation of STAT2, preventing transcription of interferon-regulated genes. NS4A and NS4B inhibit mTOR signaling from receptor tyrosine kinases (RTK). ZIKV proteins modulation of host pathways also inhibits neurogenesis and promotes cell death. sfRNA, subgenomic flavivirus RNA; sRNA, viral small RNA. Reproduced with permission from (Ming et al., 2016) Cell Press, original publication DOI: http://dx.doi.org/10.1016/j.stem.2016.11.014.

1.11 Dissertation focus

While the roles of Nups in NPC structure and nuclear transport have been well established, numerous observations indicate that Nups also function outside of NPCs in the cytoplasm and nucleoplasm (Chatel and Fahrenkrog, 2012; Hou and Corces, 2010; Ptak and Wozniak, 2016; Raices and D'Angelo, 2012). For example, various FG-Nups have been detected in the nucleoplasm, and shown to move between intranuclear sites and NPCs (Rabut et al., 2004). In addition to contributing to nuclear transport (Sakiyama et al., 2016; Zahn et al., 2016), these intranuclear Nups have been reported to regulate gene expression through binding transcription sites (Capelson et al., 2010; Kalverda et al., 2010; Ptak et al., 2014), including immune response genes (Enninga et al., 2002; Faria et al., 2006; Light et al., 2013; Panda et al., 2014; Satterly et al., 2007), and influencing chromatin organization (Kalverda and Fornerod, 2010; Liang and Hetzer, 2011; Ptak and Wozniak, 2016).

Among the most studied Nups exhibiting intranuclear localization is Nup98 (Griffis et al., 2002; Iwamoto et al., 2010; Radu et al., 1995). Despite the growing evidence linking Nup98 to the regulation of chromatin structure and gene expression, little is known about the mechanism by which Nup98 affects these processes. In Chapter 3 of this dissertation, we have focused on identifying novel Nup98 binding partners and assembling a Nup98 interaction network. Of the Nup98 interactors, one of the strongest binding partners was the RNA helicase DHX9. We demonstrate that the FG/GLFG region of Nup98 binds to N- and C-terminal regions of DHX9 in an RNA facilitated manner. Nup98 binds DHX9 in the nucleoplasm, regulates DHX9 localization, and influences DHX9 RNA-binding and ATPase activity, which ultimately influences gene expression *in vivo* at the level of DHX9-mediated transcription and splicing. Consistent with these observations, our analysis revealed that Nup98 and DHX9 bind interdependently to similar gene loci and their transcripts. Based on our results, we propose that Nup98 functions as a co-factor that regulates DHX9 and, potentially, other RNA helicases. These data provide evidence for a novel mechanism by which the nucleoporin Nup98 can regulate gene expression away from NPCs.

The nucleocytoplasmic transport pathways of eukaryotic cells are also hijacked during several viral infections; Nup98 has been identified as a host factor affecting the life cycle of many viruses. Moreover, several nuclear host proteins are also appropriated by viruses to participate in their replication, one such example is DHX9. In Chapter 4, we explore how the Nup98-DHX9 complex, characterized in Chapter 3 as a regulator of RNA metabolism, may be exploited by positive-strand RNA viruses of the *Flaviviridae* family. We also further explore the importance of different nuclear proteins and nuclear transport factors in the life cycle of these positive-strand

RNA viruses (HCV and ZIKV). Despite their cytoplasmic replication cycle, we show here that the viral RNA of ZIKV and HCV can be found in the nuclei of infected cells. Altering nucleocytoplasmic transport can bias the localization of these +vRNAs, change the interaction between +vRNA and host nuclear RNA binding proteins, and affect the viral life cycle, pointing towards a previously unknown role for the nuclear and perinuclear environment in the metabolism of the vRNA from viruses of the *Flaviviridae* family.

CHAPTER 2

Materials and Methods

2.1 Cell culture

2.1.1 Cell lines

The human cell lines HEK293T (ATCC), HeLa S3 (ATCC), A549 (ATCC), Huh7.5 (Blight et al., 2002), Huh7 (ATCC) and African Green Monkey Vero cells (ATCC) were grown in Dubelcco's modified Eagle's media (DMEM) (Thermo Fisher Scientific Cat No. 11965-092) supplemented with 10% fetal bovine serum (FBS) (Thermo Fisher Scientific Cat No. 12483-020). Huh7 cells expressing the HCV JFH-1 subgenomic replicon (Lohmann et al., 1999) were grown in DMEM media supplemented with 10% FBS and 1 mg/ml of Geneticin (Thermo Fisher Scientific Cat No. 10131035). All adherent cell lines were kept at 25% to 80% confluence. U937 cells were cultured in Roswell Park Memorial Institute (RPMI) 1640 media (GIBCO Cat No. 22400-105) supplemented with 10% FBS (GIBCO Cat No. 12483-020) and kept at a cell density of $1-2\times10^6$ cells/ml. All cell lines were maintained at 37°C with 5% CO_2. Trypsin-EDTA 0.05% solution (Thermo Fisher Scientific Cat No. 25300062) was used for detachment of adherent cell lines for sub-culturing when necessary. Cell lines were preserved by freezing in Recovery Cell Culture Freezing Medium (Thermo Fisher Scientific Cat No. 12648010) at -80°C.

2.1.2 Cell viability assays

When freezing or sub-culturing cells viability was determined using a 0.4% Trypan Blue solution (Thermo Fisher Scientific Cat No. 15250061). In this dye exclusion test, viable cells do not take up the impermeable dye Trypan Blue, while dead cells are permeable and take up the dye. Viability is quantified by examining cells in a hemocytometer under an inverted light microscope.

Experiments requiring more precise determination of cell cytotoxicity under different conditions were performed with the Cell Cytotoxicity Assay Kit (Abcam, Cat No. ab112119). Cells were seeded in 96 well plates (10^4 cells/well) and submitted to the appropriate treatments, described in sections 2.4, 2.5 and 2.6. Assay solution was added to cells ($1/5$ volume of growth media) and incubated for 1 hour at 37°C. This assay measures mitochondrial dehydrogenases activity based on the reduction of oxidized non-fluorescent blue resazurin present in the assay solution to red fluorescent resorufin by acceptance of an electron from mitochondrial respiratory chain in live cells. Resorufin produced is directly proportional to the number of living cell. Fluorometric quantification of resorufin was done in a CLARIOStar microplate reader (BMG Labtech).

2.2 Mammalian expression constructs

2.2.1 Plasmid construction

Nup98 cDNA inserts encoding the indicates amino-acid residues (aa) (see Table 2.1) were generated from HEK293T cell RNA through reverse transcription and amplification using the SuperScript III One-Step RT-PCR System with Platinum Taq DNA Polymerase (Thermo Fisher Scientific, Cat No. 12574018). Polymerase chain reaction (PCR) products were purified with the QIAquick PCR Purification kit (QIAGEN, Cat No. 28104) and digested with appropriate restriction endonuclease (see Table 2.1)(New England BioLabs), as per manufacturer's protocol. Digested DNA was purified using the QIAquick Gel Extraction kit (QIAGEN, Cat No. 28704), and ligated into the indicated vectors (see Table 2.1) using the Quick Ligation kit (New England BioLabs, Cat No. M2200S). Ligated plasmids were transformed into Sub-cloning Efficiency DH5α Competent Cells (Thermo Fisher Scientific, Cat No. 18265017). Transformed bacteria were platted into Luria Broth (LB) agar plates containing the appropriate selection antibiotic. Successful ligations were determined by PCR using Taq DNA Polymerase (New Englad BioLabs, Cat No. M0273S). Plasmids were isolated from overnight cultures of DH5α cells using the EndoFree Plasmid Maxi Kit (QIAGEN, Cat No. 12362). The presence and orientation of cDNAs was confirmed by sequence analysis (TAGC applied genomics core). Nup98, Nup98$^{1\text{-}504}$, and GFP cloned into MSCV-DamID-Gateway plasmid, along with the empty vector and the Ampho plasmid were kindly provided by Dr. Tobias M. Franks and Professor Martin W. Hetzer (Salk Institute for Biological Studies) and have been described (Franks et al., 2016). To produce Dam-DHX9, the DHX9 ORF in a pShuttle vector (GeneCopoeia, Cat. No. GC-H1793) was recombined into the MSCV-DamID-Gateway plasmid using the Gateway LR Clonase II Enzyme Mix (ThermoFisher Scientific, Cat. No. 11791020) per the manufacturers protocol. The pcDNA3-HA DHX9 constructs (I347A, W339A, K417R) were a kind gift from Professor Toshihiro Nakajima (Tokyo Medical University) and have been previously described (Aratani et al., 2001). The NLS, NES and their respective controls cloned into pEGFP-N1 vector were a generously provided by Dr. Avid Levin and Professor Lorne Tyrrell (University of Alberta), and have been previously described (Levin et al., 2014b). The pGL4 plasmids are commercially available (Promega, pGL4.75 - E6931, pGL4.29 - E8471).

Table 2.1: Mammalian expression plasmids

Insert cDNA (AA)	Cloning sites	Plasmid	Reference
none	none	pEGFP-C1	Mitchell et al., 2010
Nup98 (1-920)	EcoRI / BamHI	pEGFP-C1	Mitchell et al., 2010
Nup98 (1-497)	EcoRI / BamHI	pEGFP-C1	Mitchell et al., 2010
Nup98 (498-920)	EcoRI / BamHI	pEGFP-C1	Mitchell et al., 2010
Rev NLS (34-50)	NheI / AflII	pEGFP-N1	Levin et al., 2014b
Rev SLN	NheI / AflII	pEGFP-N1	Levin et al., 2014b
SV40 NLS (125-132)	NheI / AflII	pEGFP-N1	Levin et al., 2014b
SV40 NLS mutant	NheI / AflII	pEGFP-N1	Levin et al., 2014b
Rev NES (73-84) SV40 NLS	XbaI / ApaI	pEGFP-N1	Levin et al., 2014b
Rev SEN SV40 NLS	XbaI / ApaI	pEGFP-N1	Levin et al., 2014b
none	none	pEGFP-N1	Levin et al., 2014b
GFP	Gateway RFC	pMSCV-DAMID	Franks et al., 2016
DHX9 (1-1270)	Gateway RFC	pMSCV-DAMID	Franks et al., 2016
Nup98 (1-504)	Gateway RFC	pMSCV-DAMID	Franks et al., 2016
Nup98 (1-863)	Gateway RFC	pMSCV-DAMID	Franks et al., 2016
DHX9 (1-1270)	not applicable	pcDNA3-HA	Aratani et al., 2001
DHX9 (1-1270) I347A	not applicable	pcDNA3-HA	Aratani et al., 2001
DHX9 (1-1270) W339A	not applicable	pcDNA3-HA	Aratani et al., 2001
DHX9 (1-1270) K417R	not applicable	pcDNA3-HA	Aratani et al., 2001
hRluc/CMV	not applicable	pGL4.75	AY738231.1
luc2P/CRE/Hygro	not applicable	pGL4.29	DQ904461

2.2.2 Transfection of plasmids into mammalian cells

Direct transfection: cells were seeded into appropriate plates at a concentration of 1.5×10^5 cells/ml 16 hours before transfection. Plasmids were transfected with FuGENE 6 (Promega, Cat No. E2691), Lipofectamine 2000 (Thermo Fisher Scientific, Cat No. 11668019), Lipofectamine 3000 (Thermo Fisher Scientific, Cat. No. L3000001) or TransIT LT1 (Mirus Bio LCC Cat No. MIR 2300) transfection reagents, according to the manufacturer's instructions. Briefly, plasmids were mixed with transfection reagents in serum free media and incubated at room temperature for 5-20 minutes before being added to cells at 80% confluence. Cells were incubated with transfection complexes at 37°C, 5%CO_2 for 6-12 hours before media replacement.

Reverse transfection: transfection complexes were prepared by mixing transfection reagent with plasmids, as per manufacture's protocol. Transfection complexes were added to appropriate plates first, and newly passaged cells were added to the transfection complex containing plates at a concentration of 4×10^5 cells/ml.

Transfected cells were collected for appropriate assays 24 hours after transfection, unless otherwise indicated.

2.3 Luciferase assay

HEK293T cells (10^5 cells per well in 24 well plate) were transfected with 455 ng of pGL4.29 [luc2P/CRE/Hygro] (firefly luciferase gene under control of a cAMP-response element) plasmid (Promega Cat. No. E8471) and 45 ng of pGL4.75 [hRluc/CMV] (renilla luciferase gene under control of the CMV promoter) vector (Promega Cat. No. E6931). Cells were simultaneously co-transfected with 500 ng of pEGFP-C1 or pEGFP-NUP98$^{1\text{-}920}$ and 500 ng of pcDNA3 or pcDNA3-HA-DHX9 constructs (WT, I347A, K417R). Transfections were performed with Lipofectamine 3000. When indicated 5 μM Forskolin (Sigma-Aldrich Cat. No. F3917) diluted in DMSO (Sigma-Aldrich Cat. No. D2650), or DMSO alone was added to transfected cells 5 hours before sample collection. Samples were collected 24 hours after transfection in Passive Lysis Buffer (Promega Cat. No. E1941), according to manufacturers instructions and stored at -80°C. Luciferase activity was quantified in a BioTek Synergy 4 microplate reader using the Dual-Luciferase Reporter Assay System (Promega Cat. No. E1910), and data analyzed as recommended by the manufacturer (Schagat et al., 2007). Briefly, the activity of firefly luciferase was normalized to the activity of renilla luciferase in the same sample. The normalized firefly luciferase activity of lysates from HEK293T cells, transfected with pGL4.29

and pGL4.75, were normalized to 1 and used to calculate the relative firefly luciferase activity for HEK293T cells co-transfected with pGL4.29, pGL4.75, and the other constructs described above.

2.4 Production of lentiviruses and shRNA mediated knockdown

2.4.1 Producing lentiviruses in HEK293T cells

Lentiviral pseudo-particles were produced in HEK293T cells (2.5×10^6 cells) in tissue culture treated plates (100 mm diameter) as previously described (Neufeldt et al., 2013; Schoggins et al., 2011). These cells were transfected using FuGENE 6 reagent (Promega Cat No. PRE2691) and the necessary plasmids for lentiviral production. The following previously described (Schoggins et al., 2011) plasmids encoding: HIV gag-pol (11.4 μg), vesicular stomatitis virus glycoprotein (4.5 μg), and the appropriate pLKO.1puro shRNA (15 μg, Sigma-Aldrich Mission shRNA), see Table 2.2, were used. For production of lentiviruses indicated as control, a pLKO.1puro plasmid encoding an shRNA sequence targeting a non-mammalian transcript was used. Cells were transfected in 5 ml of DMEM media containing 3% FBS. After 6 hours, the media was removed and 10 ml of fresh DMEM containing 3% FBS was added to each dish.

The supernatant containing viral particles was harvested 72 hours following transfection and cellular debris was removed by cetrifugation at $500g$ for 10 minutes. Lentiviral pseudo-particles present in the collected supernatants were concentrated using Lenti-X Concentrator (Clontech, Cat No. 631231), aliquoted and stored at -80°C. Viral titers were determined by infecting HEK293T cells with serially diluted lentiviral pseudo-particles and selecting for transduced cells with 1 μg/ml puromycin dihydrochloride (Thermo Fisher Scientific Cat No. A1113803) for 6 days. Colonies of surviving cells in each well were counted and the number of infectious lentiviral pseudo-particle units per milliliter calculated.

2.4.2 Lentiviral pseudo-particle transduction

Cells were seeded 16 hours before transduction with lentiviral pseudo-particles (10^4 cells/ml). Cells were transduced with a multiplicity of infection (MOI) of approximately 2 in DMEM media with 3% FBS and 8 μg/ml of hexadimethrine bromide (Sigma-Aldrich, Cat. No: H9268). Cells were incubated with viral particles at 37°C,

Table 2.2: Lentiviral shRNA clones

Target mRNA	Clone ID
NUP98	NM_005387.3-2874s1c1
DHX9	NM_001357.3-2863s21c1
HNRNPU	NM_031844.2-1136s21c1
XPO1	NM_003400.2-174s1c1
NXF1	NM_006362.3-1149s1c1
KAPNB1	NM_002265.4-1595s1c1
IPO5	NM_002271.3-1953s1c1
Control	SHC002

All clones were purchased from Sigma-Aldrich

5% CO_2 for 24 hours before the media was replaced with fresh DMEM with 10% FBS. Cells were sub-cultured as necessary and collected 72 to 96 hours after transduction for use in the appropriate assays. When necessary, puromycin (1 μg/ml) was added to the cell culture media 24 hours after transduction to select for cells actively expressing shRNAs.

2.5 Viral infections

2.5.1 HCV Virus production

Huh7.5 cells were seeded at a density of 10^5 cells/ml in T75 flasks 16 hours before infection. Cells were infected with 3 RNA genome equivalents, or an MOI of 1, of a serially passaged JFH-1 strain of HCV (provided by Takaji Wakita through Dr. Lorne Tyrrell)(Wakita et al., 2005). Virus containing supernatants were collected four days and six days after infection, and cleared of cell debris by centrifugation at $500g$ for 10 minutes. HCV containing supernatants were supplemented with 20% FBS, aliquoted and stored at -80°C.

2.5.2 Zika virus production

Vero cells were seeded at a density of 10^5 cells/ml in T75 flasks 16 hours before infection. Cells were infected with 3 RNA genome equivalents, or an MOI of 1, of a serially passaged ZIKV (provided by Dr. Tom Hobman, University of Alberta). Virus containing supernatants were collected two and three days after infection, and cleared of cell debris by centrifugation at $500g$ for 10 minutes. Zika virus containing supernatants were aliquoted and stored at -80°C.

2.5.3 Quantification of viral titers

HCV titers were determined by infecting Huh7.5 cells with serially diluted virus for 48 hours. Viral focus-forming units were determined by indirect immunofluorescence microscopy (described in section 2.10) using antibodies specific to HCV Core protein (Thermo Fisher Scientific, Cat No. MA1-080).
ZIKV titers were determined by infecting A549 cells with serially diluted virus for 24 hours. Viral focus-forming units were determined by indirect immunofluorescence microscopy (described in section 2.10) using antibodies against *Flavivirus* envelope

glycoproteins (Millipore, Cat No. MAB10216).
The number of viral focus-forming units per milliliter of viral containing supernatant was used as the virus titer for further experiments.
Images were obtained using the Operetta high content imaging system (PerkinElmer Inc.) and processed and analyzed in MATLAB (subsection C.2.2) and ImageJ using custom macros, available in subsection C.1.3.

Quantification of viral titers in NTF inhibitor treated samples

In the experiments quantifying viral titers in NTF inhibitor treated samples, the collected supernatant containing produced viruses also contained the NTF inhibitors used in the treatment of infected cells. One of these inhibitors, IVM, has been shown to decrease ZIKV infection at a concentration of 10 μM, but not 1 μM or less in different cell lines when added 1 hour prior to virus infection (Barrows et al., 2016). Therefore, the samples of serially diluted virus with the highest concentrations of NTF inhibitors were excluded from analysis.

2.5.4 Infection of Huh7.5 cells with HCV

Non-synchronized infection: Huh7.5 cells were seeded at a density of 2×10^5 cells/ml in appropriate plates. The cells were infected with serially passaged JFH-1 HCV at an MOI of 1 and incubated at 37°C in a 5% CO_2 for 6 hours. After incubation, cell media containing virus was replaced with DMEM containing 10% FBS. Infected cells were collected at 72 hours post-infection, unless otherwise described.

Synchronized infection: Huh7.5 cells were seeded at a density of 2×10^5 cells/ml 16 hours before infection. Cells were infected with serially passaged JFH-1 HCV virus at an MOI of 2 on ice. Cells then were incubated at 4°C for 4 hours. Following incubation with virus cells were washed three times in PBS and incubated with DMEM 10% FBS at 37°C in 5% CO_2 for 72 hours, unless otherwise stated.

2.5.5 Infection of A549 cells with Zika virus

Non-synchronized infection: A549 cells were seeded at a density of 2×10^5 cells/ml in appropriate plates. Cells were infected with serially passaged ZIKV at an MOI of 1 and incubated at 37°C in a 5% CO_2 for 6 hours. After incubation, cell media containing virus was replaced with DMEM containing 10% FBS. Infected cells were collected at 24 hours post-infection, unless otherwise described.

Synchronized infection: A549 cells were seeded at a density of 2×10^5 cells/ml 16 hours before infection. Cells were infected with serially passaged ZIKV at an MOI of 2 on ice. Cells then were incubated at 4°C for 4 hours. Following incubation with virus cells were washed three times in PBS and incubated with DMEM 10% FBS at 37°C in 5% CO_2 for 24 hours, unless otherwise stated.

2.6 Treatment of cell lines with nucleocytoplasmic transport inhibitors

Cells were plated 16 hours before beginning experimental assays. Additional assay dependent steps were carried out previous to treatment of cells with nucleocytoplasmic inhibitors (e.g. infection of cells with HCV or ZIKV or transfection of cells with mammalian expression plasmids).
Three nucleocytoplasmic transport inhibitors were used, Leptomycin B (LMB) (Sigma-Aldrich, Cat No. L2913 or Enzo Life Sciences, Cat No. ALX-380-100-C100), Importazole (IPZ) (Sigma-Aldrich, Cat No. SML0341) and Ivermectin (IVM) (Sigma-Aldrich, Cat No.I8898). Leptomycin B was diluted in ethanol at a stock concentration of 18.5 μM. Importazole and Ivermectin were diluted in DMSO at stock concentrations of 40 mM and 50 mM, respectively.
For treatment of cells, nucleocytoplasmic inhibitors were diluted in DMEM media containing 10% FBS and added to cells for 3 hours, unless otherwise stated. Final concentration of inhibitors during treatment was 18.5 nM Leptomycin B, 10 μM Ivermectin and 40 μM Importazole.

2.7 RNA purification and reverse transcription

Total cellular RNA: Total RNA was purified from appropriate cells using Trizol Reagent (Thermo Fisher Scientific, Cat No. 15596026) per manufacturers specifications. RNA (1μg RNA per reaction) was treated with DNase I (ThermoFisher Scientific 18068015) and reverse transcribed into cDNA using random primers (Thermo Fisher Scientific Cat No. 48190011) and Superscript II reverse transcriptase (Thermo Fisher Scientific Cat No. 18064014) per manufacturer's protocol.

Intracellular viral RNA: Total RNA was purified from appropriate cells using Trizol Reagent according to the manufacturers specifications. RNA (1μg RNA per reaction) was treated with DNase I (ThermoFisher Scientific 18068015) and reverse transcribed into cDNA using primers specific to the Hepatitis C or Zika viral RNA and Superscript III One-Step RT-PCR (Thermo Fisher Scientific Cat No. 18064014).

Extracellular viral RNA: Viral RNA present in the media supernatant of infected cells was purified with a High Pure Viral RNA Kit (Roche Life Science, Cat No. 11858882001) according to the manufacturer's instructions. Briefly, viruses were lysed by a detergent and Proteinase K to release total viral nucleic acids. In the presence of guanidine-HCl and appropriate buffers, viral nucleic acids bind selectively to a spin filter column, while contaminating substances do not. The purified RNA is then eluded in a small volume of low-salt buffer. RNA was reverse transcribed into cDNA using primers specific to the Hepatitis C or Zika viral RNA and Superscript III One-Step RT-PCR.

2.8 Antibodies

Anti-Nup53 antibodies (Hawryluk-Gara et al., 2008) and anti-GFP rabbit polyclonal antibodies (Makhnevych et al., 2003) were previously described. Nup98 specific antibodies were also previously described (Mitchell et al., 2010). Of note, the Nup98 antibody may show non-specific staining of a 'Golgi-like' structure adjacent to the NE on immunofluorescence microscopy experiments. An example of this non-specific staining can be observed in Figure 3.5 (right panel), as it remains visible upon Nup98 depletion. Commercial antibodies include mAB414, against Nup62, Nup153, Nup214, and Nup358 (Abcam Cat No. ab24609), anti-α-tubulin (Sigma-Aldrich Cat No. T6074), anti-DHX9 (Abcam Cat No. ab54593), hnRNP U (Abcam Cat No. ab10297), PRKDC (Thermo Scientific Cat No. MS-423-P1), GFP (Sigma-Aldrich Cat No. 11814460001), GST (GE Healthcare Life Sciences Cat No. 27-4577-01), NXF1 (Abcam, Cat No. ab50609), ELAVL1 (Abcam, Cat No. ab110081), SSB (Abcam, Cat No. ab75927), XPO1 (Millipore, Cat No. ABS1626), Kapβ3 (Santa Cruz Biotechnology, Cat No. sc-84578), Kapβ1 (Abcam, Cat No. ab2811), lamin B (Abcam, Cat No. ab16048), HCV core (Thermo Fisher Scientific, Cat No. MA1-080 or Abcam, Cat No. ab58713), HCV NS3 (Millipore, Cat No. MAB8691), dsRNA (Scicons, Cat No. J2) and ZIKV envelope glycoproteins (Millipore, Cat No. MAB10216). Goat antirabbit IgG-HRP (Bio-Rad Cat No. 170-6515), goat antimouse IgG-HRP (Bio-Rad Cat No. 170-6516), Alexa Fluor 750 goat anti-rabbit IgG (Thermo Fisher Scientific Cat No. A21039), and Alexa Fluor 680 goat anti-mouse IgG (Thermo Fisher Scientific Cat No. A21057) were used for Western blotting. Alexa Fluor 488 donkey anti-rabbit (Thermo Fisher Scientific Cat No. A21206), Alexa Fluor 488 donkey anti-mouse (Thermo Fisher Scientific Cat No. A21202), Alexa Fluor 594 donkey anti-mouse (Thermo Fisher Scientific Cat No. A21203), Alexa Fluor 594 goat anti-rabbit (Thermo Fisher Scientific Cat No. A11012), Alexa Fluor 647 goat anti-mouse (Thermo Fisher Scientific, Cat No. A21237) and Alexa 647 goat anti-rabbit (Thermo

Fisher Scientific, Cat No. A21246) were used for immunofluorescence microscopy.

2.9 SDS-PAGE and Western blot

Proteins resolved by SDS-PAGE were either stained with BioSafe Coomassie Stain (Bio-Rad, Cat No. 161-0786) or silver nitrate (Sigma-Aldrich Cat No. S6506) to detect proteins or transferred to nitrocellulose membranes (0.2 μm, Bio-Rad Cat No. 9004-70-0) for western blotting. These membranes were blocked in 5% skim milk in PBS-T (PBS containing 0.1% Tween 20) and incubated overnight at 4°C with the appropriate primary antibodies. Secondary antibodies conjugated to HRP or fluorescent dyes were used to visualize primary antibody binding.

2.10 Immunofluorescence microscopy

Cells were grown on a surface adequate for imaging and submitted to experimental conditions of interest. Cells were washed twice with PBS, fixed for 10 minutes at room temperature with PBS containing 3.6% formaldehyde (Sigma-Aldrich Cat No. F8775) and permeabilized for 2 to 4 minutes at room temperature with PBS containing 0.2 to 0.4% Triton X-100 (Thermo Fisher Scientific Cat No. BP151-500). Following two washes with PBS, samples were blocked in 2.5% skim milk in PBS-T for 2 hours at 4°C, probed with primary antibodies diluted in 2.5% skim milk in PBS-T overnight at 4°C, washed 3 times for 10 minutes with PBS-T, probed with secondary antibodies diluted in 2.5% skim milk in PBS-T for 2 hours at 4°C, and then washed 3 times for 10 minutes in PBS-T. Cell nuclei were marked with nucleic acid stains (DAPI or Hoechst 33342) before mounting with appropriate mounting media. All immunofluorescence microscopy images shown were acquired with an Axio Observer Z1 epifluorescence microscope, 63x/1.40 NA Oil UPlanS-Apochromat objective lens (Carl Zeiss, Inc.) as Z-stacks with a 0.24 μm distance between slices. Z-stack images were deconvolved under conservative settings using an iterative algorithm using Axiovision software (Carl Zeiss, Inc.). The 3 center-most slices of the deconvolved Z-stacks were submitted to Image J (Schneider et al., 2012) average intensity Z projection.

2.11 Single molecule RNA fluorescent *in situ* hybridization

We utilized a fluorescence *in situ* hybridization (FISH) protocol that allows the visualization of single RNA molecules in the cell. This is achieved due to the strong signal to noise ratio obtained when between 45-50 probes of 20 nucleotides, labeled with one flour each, are hybridized sequentially on a single RNA molecule. Since several probes are used to detect a single RNA a sufficient number of probes will interact even with partial RNA sequences (possibly due to degradation, partial replication, obstruction by RNA-binding proteins) reducing the possibility of false negatives. At the same time off target probe-RNA interactions will generate only a weak diffuse fluorescence that is not detectable, since we need tens of probes in a single RNA for a strong enough signal, also decreasing the probability of false positives.

For our experiments we designed 7 sets of probes (see Appendix B). The probe sets 1 and 2 hybridize with the GAPDH mRNA and are thus identified in figures. They contain probes with the same sequences, tagged with Quasar 570 or Quasar 670 fluors (LGC Biosearch Tachnologies, Cat No. VSMF-2150-5 and VSMF-2151-5, proprietary sequences).

Probe sets 3, 4 and 5 hybridize with the HCV JFH-1 vRNA (GenBank Accession: AB047639.1). Sets 3 and 4 were designed to hybridize with the positive-strand vRNA. Set 3 is labeled with Quasar 570 fluors, identified as +vRNA (5' probes) in figures, and it maps to nucleotides 700-3000 (Table B.1). Set 4 is labeled with Quasar 670 fluors, identified as +vRNA (3' probes) in figures, and it maps to nucleotides 3700-9500 (Table B.2). Probe set 5, identified as -vRNA in figures, was designed to hybridize with the HCV JFH-1 negative strand vRNA, mapping to nucleotides 400-3100 and labeled with Quasar 570 fluors (Table B.3).

Probe sets 6 and 7 hybridize with the positive-strand Zika virus vRNA (GenBank Accession: KF993678.1). Probe set 6, identified as +vRNA (5' probes) in figures, maps to nucleotides 1500-3500 and it is labeled with Quasar 570 fluors (Table B.4). Probe set 7, identified as +vRNA (3' probes) in figures, hybridizes to nucleotides 6500-8500 and is labeled with Quasar 670 fluors (Table B.5).

2.11.1 Sample preparation

After appropriate experimental procedures (e.g. HCV or Zika virus infection, treatment with nucleocytoplasmic transport inhibitors, etc.) cells were plated in 6 channel μ-Slide VI$^{0.4}$ (Ibidi, Cat No. 80606) for 16 to 24 hours. Cells were washed with PBS and incubated at room temperature for 10 minutes in fixation buffer (3.7%

formaldehyde (Sigma-Aldrich, Cat No. F8775) in PBS). Cells were washed twice with PBS and permeabilized with 70% ethanol for one hour at 4°C. Cells were incubated with a wash buffer containing 10% formamide (Sigma-Aldrich, Cat No. 221198) in 2X saline sodium citrate buffer (SSC)(Sigma-Aldrich, Cat No. S6639) for 5 minutes at room temperature. Probe hybridization was carried out in a humidified chamber at 37°C for 16 hours in a hybridization buffer (125 nM of probes - Appendix B, 100 mg/ml dextran sulfate (Millipore, Cat No. S4030) and 10% formamide in 2X SSC). Following hybridization cells were incubated with wash buffer at 37°C for 30 minutes. Nuclei were stained by incubating cells with wash buffer containing DAPI (1 μg/ml) at 37°C for 30 minutes. Cells were washed once with 2x SCC buffer for 5 minutes at room temperature before mounting with Vectashield anti-fade mounting media (VectorLabs, Cat No. H-1000).

2.11.2 Image acquisition and Processing

Cells hybridized to probes were imaged using the DeltaVision Elite imaging system (GE Healthcare Life Sciences, Cat No. 29065728) with SoftWoRx software version 6.5.2 (GE Healthcare Life Sciences). Images were acquired at 60 times magnification with the 1.42 NA oil, Plan Apo N objective (Olympus, ref: 1-U2B933). All images were acquired as Z-stacks spanning a region of approximately 12 μm across the Z dimension of a cell. For all images a pixel is equivalent to 0.1084 μm in the X and Y directions and 0.2 μm in the Z direction. Images were deconvolved with SoftWoRx software deconvolution module, under conservative setting for 15 cycles. Deconvolved images were processed with custom ImageJ macros (subsection C.1.1) before being imported into MATLAB 2015a (MathWorks) for analysis.

2.11.3 Computational modeling of cellular structures

Due to the high number of RNA foci present in each cell, automated focus counting and localization in relation to the nuclei is required for these images.
Before mapping of nuclei and foci, images were processed with the adequate filters to highlight features of interest in each channel of the images. The channel containing the DAPI stained nuclei was submitted to a band pass filter with a Gaussian kernel, and a gray scale threshold was used to create initial masks grossly marking the nuclei edges. These masks along with the gray scale images were submitted to an active contour algorithm to create the nuclei contour masks. These nuclei contour masks were used to eliminate any background signal in the DAPI channel images, before these images were used to create isosurfaces delineating the nuclei

of cells in the 3D plane. This work-flow was inspired by some previously described functions from (Zhao, 2016).
Images from channels containing the RNA foci were processed, as previously described (Wu and Rifkin, 2015). Briefly, RNA spots are recognized and classified by measuring several features of local intensity maxima in the images and submitting these features to a user trained supervised random forest classifier.
Custom MATLAB scripts (subsection C.2.1) and functions (subsection C.2.3) were used for analysis, these are adapted from previously mentioned published methods for foci identification (Wu and Rifkin, 2015) and nuclei modeling in microscopy images (Zhao, 2016). These scripts create a 3D representation of a cell containing surfaces representing the nuclei (DAPI stain bound to DNA used for mapping) and points indicating RNA foci (mapped from FISH signal). For +vRNA foci, only spots mapped for both viral sets of probes within 250 nm of each other were considered. We also wanted to map a region equivalent to the nuclear envelope in these cells, a region of $\pm$ 1 μm from the DAPI surface was identified as likely to contain the NE of cells.

2.11.4 Data aggregation and analysis

Once the 3D model of the cells were created, information on the localization of RNA foci in relation to nuclei was collected. This allowed the quantification of foci localizing to the nucleoplasm, nuclear envelope or cytoplasm of cells. We also used this information to estimate the density of RNA foci per unit area of nuclear envelope. All data described pertains to 3 biological replicates containing a minimum of 10 cells for each experimental condition. For statistical tests, the values of all cells in a single biological replicate were averaged. The mean foci numbers, in each biological compartment, for each biological replicate, were used in analysis of variance tests, followed by Tukey's honest significant difference post-tests. Statistical significant was assigned to testing results presenting p-values smaller than 0.05.

2.12 Protein immunoprecipitation

Protein G Dynabeads (ThermoFisher Scientific 10004D) were conjugated to antibodies according to the manufacturers instruction. Briefly, 200 μl of beads were conjugated to 10 μg of commercial anti-GFP, anti-DHX9, anti-hnRNP U mouse monoclonal antibodies or 10 μg of anti-Nup98 (Mitchell et al., 2010), anti-GFP (Makhnevych et al., 2003) rabbit polyclonal antibodies prepared in our laboratory. The mixture of

beads and antibodies were prepared in 0.8 ml of PBS containing 0.02% Tween-20 (Sigma-Aldrich P9416-50ML) and incubated at room temperature for 10 minutes with rotation. Beads conjugated to antibodies were resuspended with 200 μl of PBS with 0.02% Tween-20.

HEK293T cells or HEK293T cells transfected with GFP, GFP-NUP98$^{1\text{-}920}$, GFP-NUP98$^{1\text{-}497}$, or GFP-NUP98$^{498\text{-}920}$ were detached from plates with trypsin (GIBCO, 25300-062) and washed twice with PBS (137 mM NaCl, 2.7 mM KCl, 4.3 mM Na_2HPO_4, 1.4 mM KH_2PO_4, pH 7.4). Cells were lysed with NP-40 cell lysis buffer (50 mM Tris, pH 7.4, 250 mM NaCl, 5 mM EDTA, 50 mM NaF, 1 mM Na_3VO_4, 1% Igepal CA-630, 0.02% NaN_3, and protease inhibitor cocktail (Sigma-Aldrich 11873580001)) on ice for 30 minutes, using 400 μl of buffer per 100 mm diameter plate (approximately 6×10^6 cells/plate at time of collection). Lysates were further disrupted by centrifugation through a QIAshredder column (QIAgen 79654) and samples were clarified by centrifugation at $14,000g$ for 20 minutes at 4 °C. Cell lysate supernatant fractions were combined with antibody conjugated beads (50 μl of bead solution per 0.5 ml of cell lysate supernatant derived from 6×10^6 cells) and incubated at 4°C for 1 hour with rotation. Protein complexes bound to the beads were washed five times with 0.5 ml of NP-40 cell lysis buffer for 10 minutes at room temperature. Protein complexes were eluted from beads by heating to 100°C for 3 minutes in 25 μl Laemmli sample buffer with DTT (2% SDS, 10% glycerol, 0.01% bromophenol blue, 0.2 M of DTT and 0.06 M Tris-HCl, pH 6.8) per 50 μl of bead solution. Eluted proteins were analyzed by SDS-PAGE and western blotting.

Immunoprecipitation reactions from HEK293T cells transfected with GFP-NUP98$^{1\text{-}920}$ that were submitted to LC-MS/MS were prepared according to the protocol described above, but were scaled up by a factor of ≈ 4-fold (200 μl of bead solution per 2 ml of cell lysate supernatant derived from 3×10^7 cells).

Immunoprecipitation of Nup98 from U937 cells that were submitted to LC-MS/MS followed the same protocol described above with the following modifications. U937 cells were grown in suspension to 2×10^6 cells/ml in 100 mm culture plates. Harvested cells (1×10^8) were lysed in 4 ml of NP-40 cell lysis buffer and samples were clarified by centrifugation. Cell lysate supernatant fractions derived from 5×10^7 cells (2 ml) were combined with 200 μl of antibody conjugated beads and incubated at 4°C for 1 hour. Bound protein complexes were washed, eluted and processed for SDS-PAGE.

When performing RNase treatment of immunoprecipitated complexes, HEK293T cells (1.2×10^7 cells) were lysed in 800 μl of NP-40 cell lysis buffer containing 40 μl of RNase OUT RNase Inhibitor (ThermoFisher Scientific, 10777-019) on ice for 30 minutes. Lysates were further disrupted by centrifugation through a QIAshredder column, and samples were clarified by centrifugation at $14,000g$ for 20 minutes

at 4°C. Cell lysate supernatant fractions were combined with antibody conjugated beads (100 μl of bead solution and 800 μl of cell lysate supernatant) and incubated at 4°C for 1 hour. Protein complexes bound to the beads were washed three times with 1 ml of NP-40 cell lysis buffer without RNase OUT RNase Inhibitor. Samples were then split and half of the beads were resuspended in NP-40 cell lysis buffer with RNase OUT RNase Inhibitor (100 μl) and the other half was incubated with 100 μl of NP-40 cell lysis buffer supplemented with 1 μl of the RNase A (10 mg/ml) (ThermoFisher Scientific EN0531). Both samples were incubated at 37°C for 30 minutes and then washed three times with 0.5 ml of NP-40 cell lysis buffer for 10 minutes at room temperature. Protein complexes were eluted from beads by addition of 25 μl Laemmli sample buffer with DTT and heating to 100°C for 3 minutes.

2.13 Immunoprecipitation from HeLa cell nuclear envelope and nucleoplasm

Nuclei from HeLa cells were kindly provided by Dr. Paul Melanon (University of Alberta), and isolated from 10^7 cells according to a previously published protocol (Balch et al., 1984). Pelleted nuclei were resuspended by drop-wise addition of 250 μl of ice-cold solution A (0.1 mM $MgCl_2$ solution, protease inhibitor cocktail), and constant slow speed vortexing. Nuclei were then immediately diluted by addition of 1 ml of ice-cold solution B (10% sucrose, 20 mM triethanolamine (pH 8.5), 0.1 mM $MgCl_2$, 1 mM DTT and protease inhibitor cocktail) containing 1 μg/ml DNAse I (Sigma-Aldrich Catalogue No. D5025) and incubated at room temperature for 15minutes. A 10 μl sample of total nuclei was removed for western blotting and diluted with 10 μl of PBS. The nuclei suspension was centrifuged at $4,100g$ for 15minutes at 4°C to separate the nuclear envelope (pellet) and nucleoplasm (supernatant) fractions. A 20 μl sample of the nucleoplasmic fraction was removed for western blot and the remaining stored for immunoprecipitation.The nuclear envelope (pellet) was resuspended in 1.24 ml of ice-cold solution C (10% sucrose, 20 mM triethanolamine, pH 7.5, 0.1 mM $MgCl_2$, 1 mM DTT and protease inhibitor cocktail). A 20 μl sample was taken for western blot and the remainder used for immunoprecipitation.

One tenth volume of NP-40 cell lysis buffer stock solution (250 mM Tris-HCl, pH 7.5, 1.25 M NaCl, 25 mM EDTA, 5% NP-40, 5 mM VRC and protease inhibitor cocktail) was added to the nuclear envelope and nucleoplasm fractions and samples were divided into 3 equal volumes. Each sample received 10 μg of anti-Nup98, anti-DHX9, or anti-GFP antibody. Anti-GFP antibodies were used in negative control immunoprecipitation reactions and are identified as control IgG in figures. Samples

were incubated at 4°C with rotation for 1 hour. 100 μl of Protein G Dynabeads was then added and samples were incubated with rotation at 4°C for an addition 30 minutes. Beads were washed five times with 400 μl of NP-40 cell lysis buffer (25 mM Tris-HCl pH 7.5, 125 mM NaCl, 2.5 mM EDTA, 0.5% NP-40, 0.5 mM VRC and protease inhibitor cocktail). Samples were eluted into 40 μl of Laemmli sample buffer by heating to 100°C for 3 minutes and analyzed by western blotting.

2.14 Mass spectrometry

Proteins present in immunoprecipitates of GFP-Nup98$^{1\text{-}920}$ from HEK293T cell lysates and endogenous Nup98 from U937 cell lysates were used for mass spectrometry analysis, as previously described (Mitchell et al., 2010). Briefly, proteins were resolved by SDS-PAGE, stained with Bio-Safe Coomassie Stain, bands excised from gel lanes, and subjected to in-gel trypsin digestion followed by LC-MS/MS using a mass spectrometer (Q-TOF Premier; Waters Corp.). Protein identification was performed by peptide mass fingerprinting using PEAKS mass spectrometry software (Bioinformatics Solutions, Inc.). MS/MS data exported from PEAKS software is available as an excel file (MSdata.xls), deposited online at http://github.com/jucapitanio/thesis or in attached CD-ROM.

2.14.1 Creating and analyzing protein-protein interaction networks

Curated protein-protein interactions (PPI) among identified Nup98 binding partners were extracted using the Search Tool for the Retrieval of Interacting Genes/Proteins (Szklarczyk et al., 2011). Only PPIs from curated databases or curated published experiments were included in the PPI retrieval, and a minimum integrated confidence score of 0.5 was required for each interaction (for details see von Mering et al., 2003). Identified interactions were visualized using Cytoscape (Smoot et al., 2011). The PPI network edge thickness represents the integrated confidence score for the interaction (ranging for medium confidence score of 0.5 to high confidence score of 1). Node colour in gray scale from light to dark indicates increasing abundance of the interactor in the GFP-Nup98 immunoprecipitation, based on number of unique peptides present in LC-MS/MS data (ranging from 5 to 30 unique peptides). Clustering of the resulting PPI network with the Cytoscape plugin MCODE (Bader and Hogue, 2003) identified highly interconnected proteins, which are likely to represent protein complexes in PPI network. We used the BinGO Cytoscape plugin (Bader and Hogue, 2003; Maere et al., 2005) to perform GO annotation enrichment

analysis on the protein clusters identified by the MCODE plugin to infer biological processes for protein complexes. The node clusters identified by MCODE, representing putative protein complexes, along with the biological processes identified as enriched for each complex are indicated on the network as a coloured Venn diagram.

Network and node level statistics were extracted from the resulting network using Cytoscape. Specifically, network level statistics refers to statistically significant (p-value < 0.001) protein-protein interaction enrichment, comparing the number of interactions observed in the network to the expected number of interactions from a random graph with the same number of nodes. Node level statistics refers to node degree, i.e. the number of edges connected to each node in the network. Node degree was used as a selection criterion for which identified Nup98 interactors would be further investigated.

2.15 RNA immunoprecipitation

The RNA immunoprecipitation protocol described below is based on previously described assays (Conrad, 2008; Jensen and Darnell, 2008; Licatalosi et al., 2008). Briefly, HEK293T cells were seeded in 150 mm diameter tissue culture plates, grown for 48 hours to 75% of confluency ($\approx 10^7$ cells), washed once with PBS, and cross-linked with 0.5% formaldehyde in PBS for 10 minutes at room temperature under slow shaking (approximately 70 rpm). Cross-linking was quenched with 220 mM glycine pH 7.0 for 5 minutes at room temperature, and cells were harvested by scraping and centrifugation ($700g$ for 3 minutes at 4°C). Cells were washed four times with ice-cold PBS before lysis in 500 µl of lysis buffer (1.06 mM KH_2PO_4, 155 mM NaCl, 2.97 mM Na_2HPO_4, 0.1% SDS, 0.5% sodium deoxycholate, 0.5% Igepal CA-630, protease inhibitor cocktail and 2 units/µl of RNaseOUT Recombinant Ribonuclease Inhibitor). Cell lysates were spun through a QIAshredder spin column twice and the insoluble fraction was cleared by centrifugation at $16,000g$ for 10 minutes at 4°C. Samples of the input (10%) were removed for quantitation of total RNA and proteins present in the cleared cell lysates, and the remaining 90% of the samples were used for immunoprecipitation. For each immunoprecipitation reaction, 100 µl of Protein G Dynabeads was conjugated to 5 µg of anti-DHX9, anti-Nup98, or anti-GFP antibodies. Anti-GFP rabbit polyclonal antibodies were used as negative control in immunoprecipitation reactions and are identified as control IgG in figures. Cleared cell lysates were added to antibody conjugated beads and incubated with rotation at 4°C for 2 hours. Beads were washed at 4°C with 1 ml of PBS cell lysis buffer, twice with 1 ml of high-salt buffer (5.3 mM KH_2PO_4, 775 mM NaCl, 14.18 mM Na_2HPO_4, 0.1% SDS, 0.5% sodium deoxycholate, 0.5% Igepal

CA-630, protease inhibitor cocktail, and 2 units/μl of RNaseOUT Recombinant Ribonuclease Inhibitor), and then three times with 1 ml of PBS cell lysis buffer each wash for 10 minutes at 4°C. Sample cross-linking was reversed by incubating beads with 140 μl of reverse buffer (10 mM TrisHCl, pH 6.8, 5 mM EDTA, 10 mM DTT and 2 units/μL of RNaseOUT Recombinant Ribonuclease Inhibitor) for 45 minutes at 70°C. Input samples were also incubated with 120 μl of reverse buffer for 45 minutes at 70°C. Ten percent of each sample was removed for SDS-PAGE and western blotting. The remaining sample was treated with an equal volume of 2 x Proteinase K solution (0.2 mg/ml proteinase K, 40 mM TrisHCl pH 7.5, 5 mM EDTA, 33.4 ng/ml GlycoBlue (ThermoFisher Scientific AM9515), 0.2 mg/ml total yeast RNA (Sigma-Aldrich R6625)) for 30 minutes at 37°C to digest protein bound to the beads and release RNA. RNA was subsequently purified using TRIzol LS Reagent (ThermoFisher Scientific 10296-010) and treated with DNase I (ThermoFisher Scientific 18068015) before quantification. Reverse transcription reactions were performed on 1 μg of purified RNA using random primers (ThermoFisher Scientific 48190011) and Superscript II reverse transcriptase kit reagents (ThermoFisher Scientific 18064014) in a total volume of 20 μl. cDNA from the reverse transcription reaction (5 μl) was used as template in a 25 μl PCR reactions using Phusion High-Fidelity DNA Polymerase (New England BioLabs, M0530S). PCR amplification products were resolved in 2% agarose gels in TBE buffer (100 mM Tris, 90 mM boric acid, and 1 mM EDTA) and visualized with SYBR Safe DNA Gel Stain (ThermoFisher Scientific, S33102) in a Safe Imager 2.0 Blue Light Transilluminator (ThermoFisher Scientific, G6600).

2.16 DamID assays in mammalian cell cultures

DamID assays were performed as previously described (Franks et al., 2016; van Steensel et al., 2001; Vogel et al., 2007). Nup98, Nup98$^{1\text{-}504}$, and GFP cloned into MSCV-DamID-Gateway plasmid and the pCL-Ampho plasmid were kindly provided by Drs. Tobias Franks and Martin Hetzer (Salk Institute for Biological Studies, CA, USA) and have been described (Franks et al., 2016). To produce Dam-DHX9, the DHX9 ORF in a pShuttle vector (GeneCopoeia, Cat. No. GC-H1793) was recombined into the MSCV-DamID-Gateway plasmid using the Gateway LR Clonase II Enzyme Mix (ThermoFisher Scientific) per the manufacturers protocol.

Retroviruses encoding the Dam constructs described above were generated by co-transfection of pCL-Ampho plasmid and the MSCV-DamID vector of choice (5 μg of each plasmid) into HEK293T cells, using Lipofectamine 3000 (Thermo Fisher Scientific) with media replacement 6 hours after transfection. Two days after transfection media containing retroviruses was collected from HEK293T cells, retroviruses

were added to naive HEK293T cells in 6 well plates (6×10^5 cells/well) and incubated for 6 hours in the presence of hexadimethrine bromide (Sigma-Aldrich). To select cell lines stably expressing the Dam constructs, two days after transduction, cells were grown in medium containing 1.5 μg/ml puromycin (Thermo Fisher Scientific A1113803) for a minimum of 10 days.

Stable cell lines expressing Dam constructs were transduced with lentiviral pseudoparticles encoding shRNAs targeting Nup98, DHX9 or a control sequence as described above. Protein depletion was allowed to proceed for 6 days and depletion was verified by western blotting. DNA was harvested from 2×10^6 cells using the Qiagen DNAeasy blood and tissue kit as described by the manufacturer (QIAgen 69504). Purified DNA (2.5 μg) was digested in a 10 μl reaction containing 0.5 μl of DpnI restriction enzyme (New England Biolabs R0176S) and CutSmart Buffer overnight. DpnI-digested DNA was ligated with a DamID adapter primer duplex (supplemental file 1D) in a 20 μl ligation reaction with T4 DNA ligase (New England Biolabs M0202S) for 4 hours at 16°C. The ligation reaction was digested with DpnII (New England Biolabs R0543S) in a 50 μl reaction for 1 hour. Ten microliters of the DpnII digested ligation sample was amplified by PCR with Expand High Fidelity PCR System as described by the manufacturer (Sigma-Aldrich 11732641001). The resulting amplified DNA was purified with QIAquick PCR Purification Kit (QIAgen 28104) and used as template for real-time PCRs, as described above. Data from DamID qPCR was normalized to background amplification of genomic gene desert regions and is shown as Dam-Nup98, Dam-Nup98$^{1\text{-}504}$ or Dam-DHX9 enrichment over Dam-GFP control.

2.17 Quantitative real time PCRs

All quantitative PCR reactions were performed with SYBR green super mix (Quanta, Cat No. 95055100), per the manufacturers protocol, in a Mx3000P QPCR System (Agilent Technologies 401403). All qPCR primers (subsection 2.17.5) were designed using Primer3Plus software (Untergasser et al., 2007), and qPCR results were analyzed with MxPro QPCR Software (Agilent Technologies), Microsoft Excel (Microsoft, Office 2013) or R statistical computing language (Team, 2016).

2.17.1 qPCR of RNA immunoprecipitations

RNA immunoprecipitations were performed as described in section 2.15, with the following additions. HEK293T cells were seeded into 100 mm diameter tissue cul-

ture plates (3×10^6 cells/plate) and transduced with lentiviral pseudoparticles encoding shRNAs targeting Nup98, DHX9, or control mRNAs. Cells were cultured 60 hours after transduction, seeded into 150 mm diameter plates (5×10^6 cells/plate) and ninety-six hours after transduction cells were cross-linked, RNA immunoprecipitations performed, and cDNA made. Real-time PCR (qPCR) was performed with the resulting cDNA. Real time PCR results were analyzed as described (Hellemans et al., 2007), using reactions containing cDNA from input samples for normalization of relative quantities. Fold changes in the abundance of mRNA species bound to Nup98 upon DHX9 depletion or bound to DHX9 upon Nup98 depletion were calculated relative to mRNA amounts present in immunoprecipitated complexes from cells transduced with control shRNAs (i.e. mock depleted cells).

2.17.2 qPCR of alternative splicing

HEK293T cells were seeded into 24 well plates (5×10^4 cells/well) 16 hours before transduction with lentiviral pseudoparticles. Cells were cultured approximately 60 hours after transduction, seeded into 12 well plates (10^5 cells/well), and ninety-six hours after transduction total RNA was purified from cells using Trizol Reagent (ThermoFisher Scientific 15596026). Purified RNA samples were treated with DNAse I before quantification and quality check using a Nanodrop 2000. Reverse transcription reactions contained 1 μg of purified RNA, random primers and Superscript II reverse transcriptase kit reagents. The produced cDNA was used as a template in qPCR reactions. Real time PCR results were analyzed as described previously (Hellemans et al., 2007), with 4 reference genes (HPRT, GAPDH, ACTB, TUBA1A) used for normalization of splice isoforms. The relative quantity of each E1A splice isoform was normalized to the total E1A transcripts present in the same sample (determined using an E1A primer amplifying all splice isoforms and pre-mRNA). Fold changes in the abundance of different E1A splice isoforms upon Nup98 or DHX9 depletion was calculated relative to isoform abundance in cells transduced with control shRNA (mock depleted cells).

2.17.3 Quantification of nuclear and cytoplasmic transcripts

HEK293T cells were transduced with lentiviral pseudoparticles encoding shRNAs targeting Nup98, DHX9 or a control sequence. Four days after transduction, 5×10^6 cells were collected and processed using the PARIS Kit (Thermo Fisher Scientific AM1921). The manufacturers protocol for protein and RNA purification from cultured cells was used and followed with parallel samples for purification of total

cellular protein/RNA, nuclear protein/RNA, and cytoplasmic protein/RNA. Appropriate nuclear/cytoplasmic fractionation was verified by western blotting. RNA samples were treated with DNase I and then reverse transcribed, before being used as template in real-time PCR reaction. Change in the level of each specific mRNA in the nuclear or cytoplasmic fraction is represented as a fold-change, between depleted and mock-depleted cells, in the ratio of fraction:total mRNA of each transcript examined (i.e. transcript amount present in nuclear or cytoplasmic fraction / total transcript amount present in cell lysates). Thus the fold changes in the ratios of fraction:total amount of any given transcript account for changes in the levels of that transcript in the depleted cells.

2.17.4 qPCR of viral RNAs

RNA purified from cells (see section 2.7) was reverse transcribed into cDNA using primers specific to the Hepatitis C or Zika viral RNA and Superscript III One-Step RT-PCR. Real-time PCR reactions were performed with SYBR green super mix (Quanta Cat No. 95055-100). Concomitantly with qPCR of samples of interest, a standard curve qPCR was performed. For the standard curve a sample containing a known copy number of viral cDNA template was used in a 10-fold serial dilution to produce a range of template concentrations. The serial dilution of known quantities of cDNA was used in qPCRs and each standard curve was defined as the regression line of the logarithm of copy number (in serial dilution samples) versus threshold cycle or Ct (the cycle where fluorescence first exceeds baseline in qPCR reactions). Using the standard curves the number of viral genome copies present can be inferred from their Ct values in qPCR reactions.

2.17.5 List of PCR primers

Table 2.3: Primers list

Gene	Primer	Primer sequence
NUP98	Forward	ACCACCCAGAACACTGGCTT
NUP98	Reverse	GGCTGTGAGGCTTGGGTTAC
DHX9	Forward	CTCCACATCTGGCTCTCAAA
DHX9	Reverse	TTTTCCAAGGTCCAGTTTCC
ZFY mRNA	Forward	TCAGTGTGAGTACTGTGAGTATAGC
ZFY mRNA	Reverse	TTAGGGCAGACCAACTTCTTTATGG

Continued on next page

Table 2.3 – Primers

Gene	Primer	Primer sequence
ZFY exon 8	Forward	AAACCTAGTACCATCCAAAACC
ZFY exon 8	Reverse	GGACAGTAAAAATCAGGTAGGG
HOXA2 mRNA	Forward	GAGCTGGCCTAAACAATGACAGTCC
HOXA2 mRNA	Reverse	AAGTCGATTGTGGTGAGTGTGTCTG
HOXA2 intron 1	Forward	AGCTATTGTGCTGCCTTTCC
HOXA2 intron 1	Reverse	AAAGTTTGCTCCCGGATG
GADD45A	Forward	GAGCTCCTGCTCTTGGAGAC
GADD45A	Reverse	TTCCCGGCAAAAACAAATAA
NHLH2	Forward	GGAGAGTAGCTTCTTTGTGTGTG
NHLH2	Reverse	TCACAGCAGAGAACATGAAACA
IER3	Forward	CAGCCGCAGGGTTCTCTAC
IER3	Reverse	GTTAGGGGCGTCCTCTGG
MYC 5' TSS	Forward	TCTCCACTTGCCCCTTTTAG
MYC 5' TSS	Reverse	CGGAGTTCCCAATTTCTCAG
MYC mRNA	Forward	GCTGCTTAGACGCTGGATTT
MYC mRNA	Reverse	CACCGAGTCGTAGTCGAGGT
JUND 5' TSS	Forward	TCGCTCATTTGCATGGAG
JUND 5' TSS	Reverse	CTGTTGTGGCGTTTACCG
JUND exon 1	Forward	CTGGCGTAACGAGACTTTACTG
JUND exon 1	Reverse	CGAGTCCACATTCCTGTTTG
JUND mRNA	Forward	GCCCATCGACATGGACAC
JUND mRNA	Reverse	TGGCTGAGGACTTTCTGCTT
TNF alpha	Forward	GCAGTCAGATCATCTTCTCG
TNF alpha	Reverse	ACCACCAGCTGGTTATCTCT
HEXIM1	Forward	GCATGGAGGACGAGAACAAC
HEXIM1	Reverse	GGTCAGCAGCTGGAGGTTC
FOXP2 mRNA	Forward	GCAACTCTCATAAGGCAGGC
FOXP2 mRNA	Reverse	TTTGTGACCTTCGCTTCTGG
FOXP2 intron1	Forward	AAGTGCTATGCCCCAAGATG
FOXP2 intron1	Reverse	GAGTATAACGCCTTGTTGTTCC
E1A pre-mRNA	Forward	TTATCACCGGAGGAATACGG
E1A pre-mRNA	Reverse	CTGCCCATAATTTTCACTTACTG
E1A	Forward	TTATCACCGGAGGAATACGG
E1A	Reverse	CTGCCCATAATTTTCACTTACTG
13S E1A	Forward	GATCGAAGAGGTACTGGCTG
13S E1A	Reverse	ACTCCTCACCCTCTTCATCC
12S E1A	Forward	GATCGAAGAGGTACTGGCTG
12S E1A	Reverse	GACACAGGACCCTCTTCATC
11S E1A	Forward	GATCGAAGAGCCCGAGCAGC
11S E1A	Reverse	CAAGACCTGCAACCGTGCCC
10S E1A	Forward	GATCGAAGAGCCCGAGCAGC
10S E1A	Reverse	GACACAGGACCCTCTTCATC
9S E1A	Forward	GATCGAAGAGGTCCTGTGTC

Continued on next page

Table 2.3 – Primers

Gene	Primer	Primer sequence
9S E1A	Reverse	TCAGGATAGCAGGCGCCATT
HPRT	Forward	CCTGGCGTCGTGATTAGTG
HPRT	Reverse	ACACCCTTTCCAAATCCTCAG
ACTIN	Forward	CTGTGGCATCCACGAAACTA
ACTIN	Reverse	AGCACTGTGTTGGCGTACAG
TUBULIN	Forward	GGAACCCACAGTCATTGATGAA
TUBULIN	Reverse	GCCCTCGGGCATAGTTATTG
GAPDH	Forward	GCACCGTCAAGGCTGAGAAC
GAPDH	Reverse	TGGTGAAGACGCCAGTGGA
HCV	Forward	TCTGCGGAACCGGTGAGTA
HCV	Reverse	GTGTTTCTTTTGGTTTTTCTTTGAGGT
Zika	Forward	CCTTGGATTCTTGAACGAGGA
Zika	Reverse	AGAGCTTCATTCTCCAGATCAA
Chr4	Forward	AAGTGTGGCAAACACCTTCC
Chr4	Reverse	AATGGTTAGCTCCGTTGTGC
Chr12	Forward	TCCCTCCATTTGTCCGTAAG
Chr12	Reverse	ATCCCTGTAAAACCCCAACC
adapter primer	amplification	GGTCGCGGCCGAGGATC
adapter primer	duplex bottom	TCCTCGGCCG
adapter primer	duplex top	CTAATACGACTCACTATAGGG CAGCGTGGTCGCGGCCGAGGA

2.18 Bacterial expression of recombinant proteins

2.18.1 Plasmid manipulation and purification

Plasmids for expression of recombinant GST tagged Nup98 were created by cloning Nup98's coding sequence (NM_016320.4) into EcoRI and NotI sites of the pGEX-6P-1 plasmid (GE Life Sciences Cat No. 28-9546-48). Nup98 full-length, N-terminal and C-terminal cDNA sequence are the same as previously described for mammalian plasmids (Table 2.1). Plasmids for the expression of recombinant GST tagged DHX9 were created by cloning the DHX9 coding sequence (NM_001357.4) into the EcoRI and XhoI cloning sites of the pGEX-6P-1 plasmid. The full-length DHX9 plasmids contained the cDNA sequence encoding DHX9s amino acids 1 to 1270, DHX9s N-terminal plasmids comprised amino acids 1 to 380, DHX9s helicase domain plasmids amino acids 381 to 820 and DHX9s C-terminal plasmids amino acids 821 to 1270. Plasmids for expression of recombinant N-terminally GST-tagged and C-terminally eGFP tagged DHX9 were created by adding the eGFP sequence from the pEGFP-C1 plasmid to the XhoI and NotI sites of the pGEX-6P-1 plasmids described.

Table 2.4: Bacterial expression plasmids

Insert		Cloning sites	Plasmid	
cDNA (AA)	GenBank		Backbone	GenBank
none	not applicable	none	pGEX-6P-1	U78872.1
Nup98(1-920)	NM_016320.4	EcoRI/NotI	pGEX-6P-1	U78872.1
Nup98(1-497)	NM_016320.4	EcoRI/NotI	pGEX-6P-1	U78872.1
Nup98(498-920)	NM_016320.4	EcoRI/NotI	pGEX-6P-1	U78872.1
DHX9(1-1270)	NM_001357.4	EcoRI/XhoI	pGEX-6P-1	U78872.1
DHX9(1-380)	NM_001357.4	EcoRI/XhoI	pGEX-6P-1	U78872.1
DHX9(381-820)	NM_001357.4	EcoRI/XhoI	pGEX-6P-1	U78872.1
DHX9(821-1270)	NM_001357.4	EcoRI/XhoI	pGEX-6P-1	U78872.1
DHX9(1-380)eGFP	NM_001357.4-U55763.1	XhoI/NotI	pGEX-6P-1	U78872.1
DHX9(381-820)eGFP	NM_001357.4-U55763.1	XhoI/NotI	pGEX-6P-1	U78872.1
DHX9(821-1270)eGFP	NM_001357.4-U55763.1	XhoI/NotI	pGEX-6P-1	U78872.1

2.18.2 Recombinant protein expression and purification

Expression and purification of recombinant proteins was performed as previously described (Mitchell et al., 2010). Briefly, *E. coli* BL21-CodonPlus(DE3) cells (Agilent Technologies 230245) were transformed with the pGEX-6P-1 based plasmids, grown to an $O.D._{600}$ of 0.6, and protein expression induced with 1 mM IPTG (ThermoFisher Scientific, BP175510) for 2 hours at 37°C (Nup98) or overnight at 16°C (DHX9). After collection by centrifugation, bacterial cells were resuspended in lysis buffer (50 mM Tris, pH 7.5, 300 mM NaCl, 150 mM KOAc, 2 mM MgOAc, 10% glycerol, 0.1% Igepal CA-630, 1 mM DTT, and protease inhibitor cocktail) and sonicated. The soluble fractions of the lysates were cleared by centrifugation at $27,000g$ for 20 minutes. Purification of recombinant GST fusion proteins using glutathione-Sepharose 4B Media (GE Healthcare Life Sciences 17-0756-01) was performed as previously described (Mitchell et al., 2010). When appropriate, the GST tag was cleaved from the recombinant proteins using PreScission Protease as described by the manufacturer (GE Healthcare Life Sciences 27-0843-01).

2.19 *In vitro* assays using recombinant proteins

2.19.1 *In vitro* pull down assay

Protein G Dynabeads (300 μl) were conjugated to 30 μg of anti-DHX9 antibody as described by the manufacturer (ThermoFisher Scientific 10004D) and incubated with approximately 3.6 nmoles of GST-tagged DHX9 in 1.2 ml of PBS-T at room temperature for 1 hour with rotation. After washing to remove unbound protein, beads were resuspended in a total volume of 1.2 ml of PBS-T and 400 μl aliquots were incubated for 10 minutes at room temperature with either RNase A (final concentration 100 μg/ml), poly I:C RNA (Sigma-Aldrich P1530) (final concentration 100 μg/ml), or buffer alone. In parallel, GST-tagged Nup98 (1.2 nmoles in 200 μl of PBS-T) and purified GST (1.2 nmoles in 200 μl of PBS-T) were similarly treated with RNase A, poly I:C RNA, or buffer alone. Each of the three bead bound samples of GST-tagged DHX9 were then divided into 2 equal parts and combined with similarly treated GST-tagged Nup98 (0.6 nmoles in 200 μl of PBS-T per sample) or GST alone (6 nmoles in 200 μl of PBS-T per sample) and incubated at 4°C with rotation for 30 minutes. The protein complexes were washed five times with 500 μl of PBS-T, eluted from beads with 15 μl of Laemmli sample buffer with DTT, and analyzed by SDS-PAGE and western blotting. The same procedure described above

was also performed with DHX9 and Nup98 after GST tag removal by cleavage with PreScission Protease.

2.19.2 Bead halo assay

The bead-halo assay was performed as previously described (Patel and Rexach, 2008; Zhou et al., 2013), with some modifications. To prepare bait samples, 15 μl of protein G Dynabeads were conjugated to 1.5 μg of anti-DHX9 antibody, divided into 2 equal samples and approximately 150 pmoles of purified recombinant DHX9 in 15 μl of PBS-T or buffer alone was added to each. These samples were further divided into 3 equal parts for the addition of RNase A (1 μg), poly I:C RNA (1 μg), or PBS-T alone and all six bait samples were then incubated at 4°C for 30 minutes. Prey samples were prepared by mixing 300 pmoles of purified recombinant Nup98 with 2 μg of anti-Nup98 and 2 μg fluorescently tagged Alexa Fluor 488 donkey anti-rabbit antibody in a final volume of 45 μl in PBS-T. This sample was incubated for 10 minutes at room temperature and then divided into 3 equal parts (15 μl per sample) for addition of RNase A (2 μg), poly I:C RNA (2 μg) or PBS-T alone, followed by incubation at 4°C for 30 minutes. Bait and prey samples were mixed per their additives (RNase A, poly I:C RNA or buffer alone) and incubated together for 10 minutes at 4°C. To define the domains of DHX9 that mediate Nup98 binding, GST or GST-Nup98$^{1\text{-}863}$ was immobilized on Glutathione High Capacity Magnetic Agarose beads (Sigma-Aldrich G0924) as the bait. DHX9 domains (1-380, 381-820, and 821-1270) tagged with eGFP at the C-terminus of DHX9 acted as prey, with both bait and prey samples being treated and combined for binding as above. All bead samples were washed three times with 60 μl of PBS-T before acquisition of epifluorescence images as described for immunofluorescence.

Image analysis was performed using ImageJ (Schneider et al., 2012) with custom macros. Data processing was done in R (R Core Team, 2016). Briefly, images were opened in ImageJ and processed for background subtraction. Masks were created to identify beads and fluorescence intensities were measured for each masked bead. Bead fluorescence intensity measurement files were imported into R and aggregate averages were calculated for different experimental conditions and biological replicates. The fluorescence intensity of negative control samples was subtracted from corresponding experimental conditions. Mean and standard deviation of arbitrary fluoresce units of biological replicates were calculated and plotted in bar graphs.

2.19.3 ATPase assay

ATPase reactions were carried out in 96 or 384 well plates at 37°C using a previously described enzyme-coupled assay (Panaretou et al., 1998). Each 50 or 100 μL reaction contained 25 mM HEPES, 1 mM phosphoenolpyruvate (Sigma-Aldrich P7127), 3 mM $MgCl_2$, 1 mM DTT, 2.5 μl of Pyruvate Kinase/Lactic Dehydrogenase enzymes (Sigma-Aldrich P0294), 0.5 mM NADH (Sigma-Aldrich N8129), 2 mM ATP (Sigma-Aldrich L510327) and 30-40 nM purified recombinant DHX9. Where specified, the reaction also contained 100 μg/ml poly I:C RNA (indicated as RNA on figures) and/or purified recombinant Nup98, GST-Nup98$^{1\text{-}920}$, GST-Nup98$^{1\text{-}497}$, GST-Nup98$^{498\text{-}920}$, or GST at an amount equimolar to DHX9 or as indicated on the figure. Control reactions for each condition contained the same reagents and recombinant proteins except for DHX9. ATP hydrolysis was monitored indirectly using absorbance of NADH at 340 nm, which was measured each minute for 120 minutes using a BioTek Synergy 4 microplate reader. The decrease of NADH absorbance over time was subsequently converted to micromoles of ATP consumed as previously described (Montpetit et al., 2012). The specific activity of DHX9 was calculated by subtracting the ATP consumption rate of control reactions (i.e. no DHX9) and normalizing the corrected rate to the concentration of DHX9 present, resulting in the ATP hydrolysis rate of DHX9 per second.

2.20 Comparative analysis of large scale DNA/RNA-seq data

Genome-wide Nup98 interaction with chromatin was assessed through available DamID-seq data, by comparing enriched DNA sequences from Dam-Nup98 or Dam-Nup98ΔCTD (also termed Dam-Nup98$^{1\text{-}504}$) expressing HeLa-C cells to those of Dam-GFP expressing cells (GSE83692). Data analysis was performed as described in the corresponding dataset and its publication (Franks et al., 2016).

Transcriptome-wide interaction of Nup98 with mRNA molecules was determined from available sequencing data for Nup98 RNA immunoprecipitations from K562 cells (GSE67963) (Hendrickson et al., 2016). DHX9 interaction with RNA was determined from sequencing data for DHX9 RNA immunoprecipitation from TC32 cells, kindly provided by Drs. Hayriye Erkizan and Jeffrey Toretsky (Georgetown University, USA) (Erkizan et al., 2015). Data analysis was performed as described in the corresponding datasets and their indicated publication, transcripts were considered as interacting with target proteins if showing a fold enrichment above 1.5 and adjusted p-value < 0.05.

Transcriptome-wide changes in transcript or splicing isoform abundance were

determined from RNA-sequencing data for HepG2 or IMR90 cells upon Nup98 depletion (GSE83551) (Franks et al., 2016). Transcriptome changes in NB1 cells upon DHX9 depletion were determined from available RNA-sequencing data (GSE44585) (Chen et al., 2014). Transcriptome sequencing data was analyzed as previously described (Wolfien et al., 2016), using Galaxy (Afgan et al., 2016), R (Team, 2016) and Bioconductor (Huber et al., 2015). An adjusted p-value < 0.05 was used to identify transcripts/isoforms whose abundance was significantly altered upon target protein depletion. All datasets were aligned to human reference sequence GRCh37/hg19 and annotated with corresponding UCSC genes and Ensembl genes (Huang et al., 2013; Rosenbloom et al., 2015; Yates et al., 2016; Yu et al., 2015). Statistically significant overlap between gene sets were calculated using the Fishers exact test based on the hypergeometric distribution through the R package GeneOverlap (Shen, 2013).

A script for reproduction of the analysis described in this section is provided in the appendix section C.3.

CHAPTER 3

Nup98 regulates the localization and activity of DExH/D-box helicase DHX9

A version of this chapter has been previously published in: Capitanio, J. S., Montpetit, B., and Wozniak, R. W. (2017). Human Nup98 regulates the localization and activity of DExH/D-box helicase DHX9. *eLife*, 6.

3.1 Summary

Beyond their role at nuclear pore complexes, some nucleoporins function in the nucleoplasm. One such nucleoporin, Nup98, binds chromatin and regulates gene expression. To gain insight into how Nup98 contributes to this process, we focused on identifying novel binding partners and understanding the significance of these interactions. Here we report on the identification of the DExH/D-box helicase DHX9 as an intranuclear Nup98 binding partner. Various results, including *in vitro* assays, show that the FG/GLFG region of Nup98 binds to N- and C-terminal regions of DHX9 in an RNA facilitated manner. Importantly, binding of Nup98 stimulates the ATPase activity of DHX9, and a transcriptional reporter assay suggests Nup98 supports DHX9-stimulated transcription. Consistent with these observations, our analysis revealed that Nup98 and DHX9 bind interdependently to similar gene loci and their transcripts. Based on our results, we propose that Nup98 functions as a co-factor that regulates DHX9 and, potentially, other RNA helicases.

3.2 Introduction

As described in chapter 1, while the roles of Nups in NPC structure and nuclear transport have been well established, numerous observations indicate that Nups also function outside of NPCs in the cytoplasm and nucleoplasm (Chatel and Fahrenkrog, 2012; Hou and Corces, 2010; Ptak and Wozniak, 2016; Raices and D'Angelo, 2012). For example, various FG-Nups have been detected in the nucleoplasm, which have been shown to move between intranuclear sites and NPCs (Rabut et al., 2004). In addition to contributing to nuclear transport (Sakiyama et al., 2016; Zahn et al., 2016), these intranuclear Nups have been reported to regulate gene expression through interactions with transcription sites (Capelson et al., 2010; Kalverda et al., 2010; Ptak et al., 2014), (Enninga et al., 2002; Faria et al., 2006; Light et al., 2013; Panda et al., 2014; Satterly et al., 2007), and influencing chromatin organization (Kalverda and Fornerod, 2010; Liang and Hetzer, 2011; Ptak and Wozniak, 2016).

Among the most studied Nups exhibiting intranuclear localization is Nup98 (Griffis et al., 2002; Iwamoto et al., 2010; Radu et al., 1995). Nup98 binds to the mRNA export factors Rae1 and NXF1 and has been shown to mediate mRNA export (Bachi et al., 2000; Blevins et al., 2003; Powers et al., 1997). Nup98 also participates in nuclear import and export of proteins though its interaction with importin-βfamily members (Allen et al., 2001) and the exportin CRM1 (Oka et al., 2010). Several distinct domains are present in Nup98, including an N-terminal region containing FG/GLFG repeats, a putative RNA-interacting domain, a binding site for Rae1 (RBD), and a C-terminal region that interacts with other Nups (Chatel and Fahrenkrog, 2012; Sun and Guo, 2008).

Immunofluorescence analysis revealed that Nup98 is visible throughout the nucleoplasm, but can accumulate at intranuclear structures termed GLFG bodies (Griffis et al., 2002). Further, the mobility of NPC-associated and nucleoplasmic Nup98 is dependent on ongoing transcription in the cell (Griffis et al., 2002). Studies in Drosophila revealed the association of Nup98 with actively transcribed genes, especially those involved in development and cell cycle regulation, with modulation of cellular Nup98 levels (over-expression or knock-down) altering transcription of these genes (Capelson et al., 2010; Kalverda et al., 2010). Similarly, in mammalian cells, Nup98 has been shown to associate with chromatin and regulate gene expression during the differentiation of embryonic stem cells into neural progenitor cells (Liang et al., 2013). In these cells, Nup98 preferentially associates with the promoter regions of developmentally regulated genes, and changes in the levels of Nup98 are again found to alter gene expression.

Several recent observations have provided further insight into the role of Nup98 in transcription. Light and colleagues showed that mammalian Nup98 binds to

HLA-DRA (Light et al., 2013), an interferon-γinduced gene exhibiting transcriptional memory (i.e. a gene that displays rapid induction given a recent history of being activated). In the absence of Nup98, transcriptional memory was lost and the binding of RNA polymerase II at promoters poised for reactivation was reduced, which matched similar findings in yeast (Light et al., 2010, 2013). In Drosophila, Pascual-Garcia and colleagues showed binding of Nup98 to the promoter regions of certain active genes and a requirement for Nup98 in their transcription (Pascual-Garcia et al., 2014). Nup98 binding to these genes was dependent on TRX and MBD-R2, a component of the NSL (nonspecific lethal) complex that directs histone H4K16 acetylation.

More evidence for the role of Nup98 in gene expression regulation comes from studies of hematopoietic malignancies, especially acute myeloid leukemias, where more than twenty-eight different chromosomal rearrangements involving the *NUP98* gene have been identified. The resulting fusion proteins have been shown to alter transcription through fusing the N-terminal domain of Nup98 (Bai et al., 2006; Kasper et al., 1999) to a C-terminal domain that usually contains a chromatin/DNA interacting region (Capitanio and Wozniak, 2012). The oncogenicity of several NUP98 fusions has been demonstrated in mouse models where *NUP98* fusions lead to acute myeloid leukemia recapitulating the human disease phenotype (Gough et al., 2011; Moore et al., 2007). For more information, see Appendix A.

Despite the growing evidence linking Nup98 to the regulation of chromatin structure and gene expression, little is known about the mechanism by which Nup98 affects these processes. In this study, we have focused on identifying novel Nup98 binding partners and assembling a Nup98 interaction network. Of the Nup98 interactors, one of the strongest binding partners was the DEAH/RHA helicase DHX9. We demonstrate that Nup98 binds DHX9 in the nucleoplasm of cells, where it regulates DHX9 localization. Nup98 also influences DHX9 RNA-binding and ATPase activity, which ultimately affects DHX9-mediated transcription and splicing *in vivo*. These data provide evidence for a novel mechanism by which the nucleoporin Nup98 can regulate gene expression away from NPCs, in partnership with a DEAH-box helicase.

3.3 Results

3.3.1 Identification of Nup98 interacting partners

Nup98 is a component of NPCs, but it has also been shown to reside in the cytoplasm and nucleoplasm (Griffis et al., 2002). The presence of this Nup in different locations likely reflects the participation of Nup98 in distinct cellular processes. To further understand these putative non-NPC functions, we focused on identifying Nup98 binding partners. GFP-NUP98 or GFP alone was expressed in HEK293T cells and immunoprecipitated (IP) using antibodies directed against GFP. Mass spectrometry (MS) analysis of purified protein complexes (Figure 3.1) identified previously characterized Nup98 interactors, such as Nup88 (Griffis et al., 2003), Rae1 (Pritchard et al., 1999), NXF1 (Bachi et al., 2000), and CRM1 (Oka et al., 2010), as well as several other proteins.

Nup98-interacting proteins were prioritized for analysis based on the number of unique peptides mapped to the protein, the percent coverage of the protein sequence (Liu et al., 2004), and absence of the protein in a database of common contaminants identified by IP-MS (Mellacheruvu et al., 2013). Nup98 interactors were further annotated with curated protein-protein interactions (PPI) to create a PPI network (Figure 3.2a). Network clustering identified highly interconnected nodes within the network, possibly reflecting protein complexes that may interact with GFP-Nup98 within the context of distinct cellular processes. Enrichment analysis based on Gene Ontology (GO) annotations indeed showed that these sets of GFP-Nup98 interacting proteins function in specific mRNA metabolism events including mRNA processing, splicing, stabilization, and transport (Figure 3.2a).

We speculated that proteins of the highest abundance (i.e. highest number of unique peptides detected by LC-MS/MS) are likely the nearest neighbors of Nup98 (Mazloom et al., 2011), and proteins with the highest number of PPIs in the network (i.e. hubs) may represent key components that interact with Nup98 within the context of these processes (He and Zhang, 2006). Therefore, we selected Nup98 interactors for further study on the basis of abundance in the GFP-Nup98 immunoprecipitation (IP) and the number of PPIs occurring with the other proteins within the network (Figure 3.2). Based on these features, DHX9 and hnRNP U were selected for further analysis. Importantly, the association of both DHX9 and hnRNP U with Nup98 was confirmed by reciprocal immunoprecipitations of endogenous Nup98, DHX9, and hnRNP U (Figure 3.3).

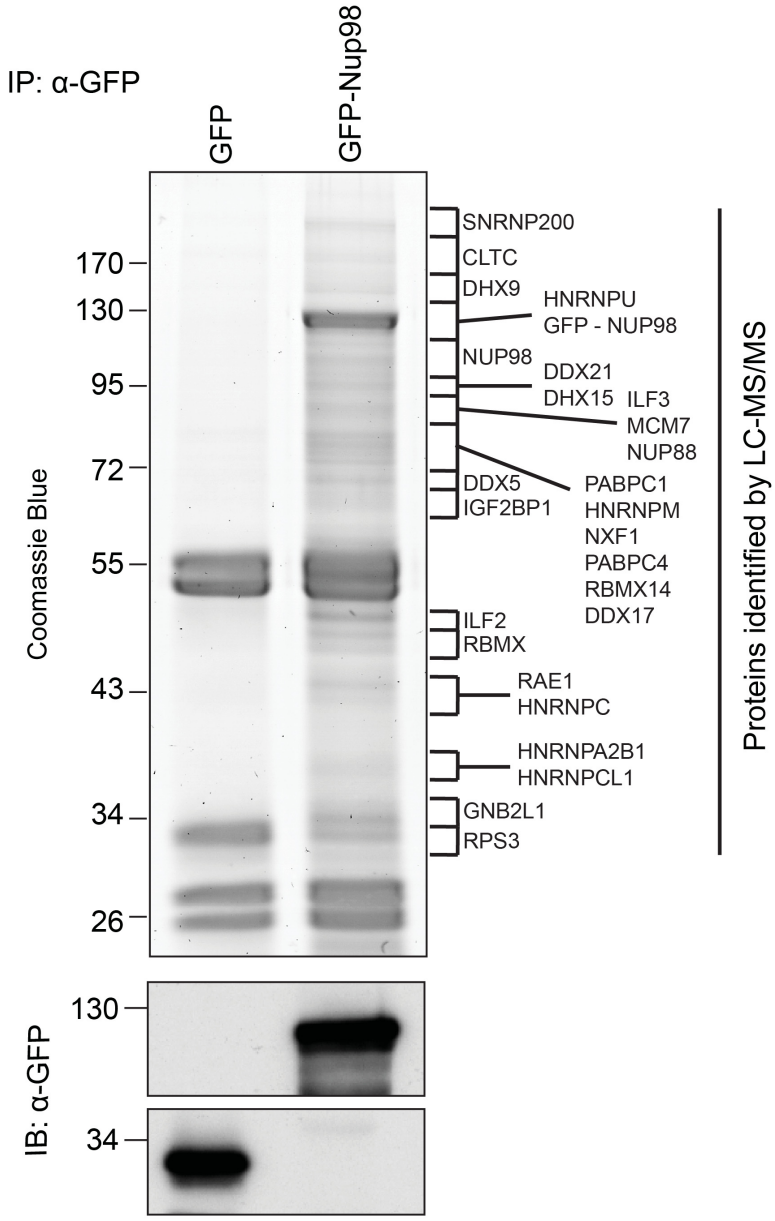
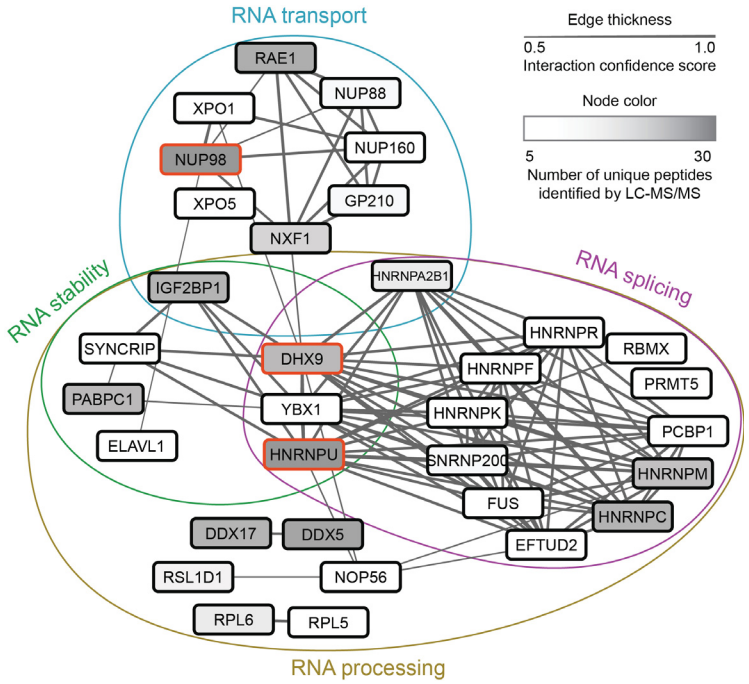
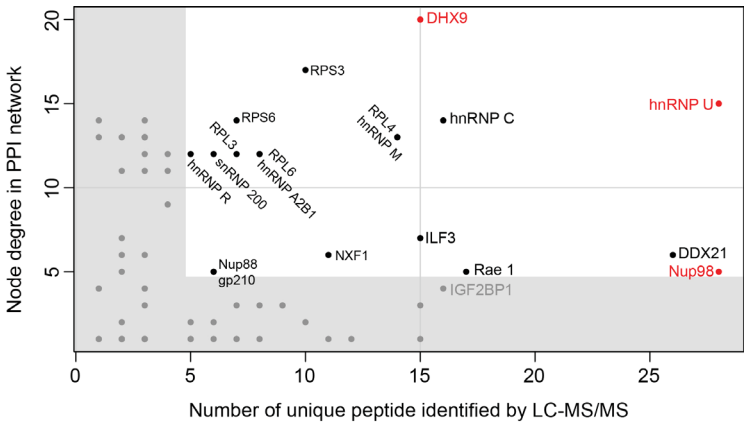


Figure 3.1: Identification of Nup98-interacting proteins.

Plasmids encoding GFP-Nup98 or GFP alone were transfected into HEK293T cells. These proteins were immunoprecipitated from whole cell lysates using an antibody directed against GFP. Co-immunoprecipitated proteins were analyzed by SDS-PAGE and gel pieces containing regions of interest were analyzed by LC-MS/MS to identify proteins co-immunoprecipitated with GFP-Nup98. Western blotting of these fractions using anti-GFP antibodies confirmed the presence of GFP and GFP-Nup98 (bottom panel). Protein species indicated on the right of the gel represent those producing peptides most frequently identified by LC-MS/MS in the GFP-Nup98 immunoprecipitated fractions. The positions of molecular mass markers (shown in kDa) are indicated on the left.



(a) Nup98 PPI network identified by IP-MS.



(b) Prioritizing Nup98 interactors.

Figure 3.2: DHX9 and hnRNP U stand out among Nup98 interactors.

3.2a Curated PPI among identified Nup98 binding partners are represented in a network. Edge thickness indicates the confidence score for the interaction and node color indicates abundance of the interactor in the GFP-Nup98 immunoprecipitation. Biological functions of identified protein complexes are indicated in the colored Venn diagram superimposed on the network. Nup98, DHX9 and hnRNP U are indicated by a red border. **3.2b** Nup98 interactors were prioritized based on their degree of interconnection and the number of unique peptides identified by MS. In the scatterplot, node degree in the PPI network (y-axis) identifies hubs in the GFP-Nup98 PPI network, while number of unique peptides (x-axis) reflects the abundance of the indicated protein in the purified GFP-Nup98 protein complex. The Nup98 interactors DHX9 and hnRNP U are shown in red.

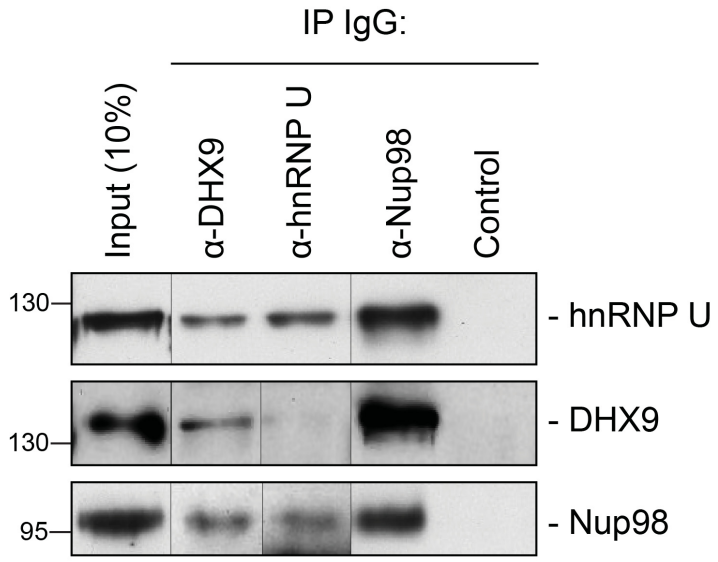


Figure 3.3: Immunoprecipitation of endogenous Nup98 with DHX9 and hnRNP U.

Endogenous DHX9, hnRNP U, and Nup98 were immunoprecipitated from whole cell lysates (Input) using specific antibodies as indicated above the panels. Anti-GFP antibodies were used to assess non specific binding (Control). Immuno-purified proteins were analyzed by western blotting using antibodies against the target proteins as indicated on the right. The positions of molecular mass markers (shown in kDa) are indicated on the left. N = 3 biological replicates, representative images of western blots are shown.

3.3.2 Nup98 influences the intranuclear distribution of DHX9

Both DHX9 and hnRNP U are RNA-binding proteins that reside in the nucleoplasm (Dreyfuss et al., 1984; Uhlén et al., 2015; Weidensdorfer et al., 2009; Zhang and Grosse, 1991). We compared the localization of these proteins and Nup98 in HEK293T cells using immunofluorescence microscopy (Figure 3.4). Similar to previous reports (Uhlén et al., 2015; Uhlen et al., 2010), DHX9 and hnRNP U were broadly distributed within the nucleoplasm in a punctate pattern, but appeared excluded from nucleoli. Neither protein appeared concentrated at the NE (Figure 3.4, inset). Nup98 was also detected within the nucleoplasm (Figure 3.4); however, the broad intranuclear distributions of Nup98, DHX9, and hnRNP U made it difficult to judge the significance of any signal overlap.

As an alternative approach to assess the physical relationship between these proteins, we examined the consequences of depleting or overproducing Nup98 on the nuclear distribution of DHX9 and hnRNP U. Depletion of Nup98 resulted in no detectable changes in the distribution pattern of hnRNP U (Figure 3.5). However, the loss of Nup98 caused the appearance of bright DHX9 foci in the nucleoplasm. Moreover, the exclusion of DHX9 from nucleoli observed in mock-treated cells was less pronounced in Nup98-depleted cells, suggesting that DHX9 has greater access to the nucleolus in the absence of Nup98. In contrast, depletion of DHX9 did not noticeably alter Nup98 localization (Figure 3.6).

Combined with the protein-interaction data, these results are consistent with a model in which Nup98 contributes to the steady-state localization of DHX9 within the nucleoplasm. To further test this idea, we increased cellular levels of Nup98 and examined the distribution of DHX9 (Figure 3.8). Elevated levels of Nup98 accumulate in intranuclear foci termed GLFG-bodies (Griffis et al., 2002), and we observed the formation of these foci in cells producing GFP-Nup98 (Figure 3.8). Importantly, DHX9 was recruited to the GFP-Nup98 foci. By contrast, no visible impact on hnRNP U distribution was observed in GFP-Nup98 producing cells (Figure 3.9). The change in DHX9 distribution was not accompanied by alterations in the cellular levels of DHX9 (Figure 3.10), thus the DHX9 associated with GFP-Nup98 foci was likely recruited from other locations.

Our observation that DHX9 is localized to intranuclear Nup98-containing foci and did not appear to accumulate at NPCs with Nup98 (Figure 3.4) suggests that these proteins interact in the nucleoplasm. To further test this model, HeLa cell nuclei were fractionated to make nucleoplasmic and NE enriched fractions, which could be used to further assess the location of the Nup98-DHX9 interaction. Consistent with localization data, DHX9 was primarily present in a nucleoplasmic fraction, while Nup98 was detected in both the nucleoplasmic and NE fractions, which

were further validated with antibodies against other NE and nucleoplasmic proteins (Figure 3.11a). Immunoprecipitations from these fractions showed that DHX9 was only detected in association with nucleoplasmic Nup98, although similar amounts of Nup98 were purified from both fractions, and that nucleoplasmic DHX9 was able to immunopurify Nup98 (Figure 3.11b).

We also examined the interactions of DHX9 with an N-terminal region of Nup98 (GFP-Nup98$^{1\text{-}497}$). When expressed in cells, GFP-Nup98$^{1\text{-}497}$ can enter the nucleoplasm and induce the formation of GLFG bodies, however it does not associate with NPCs (Griffis et al., 2002; Kalverda et al., 2010). As observed with the full-length GFP-Nup98, DHX9 was recruited to GLFG-bodies formed by GFP-Nup98$^{1\text{-}497}$ truncation (Figure 3.12a). Consistent with this result, DHX9 was detected in association with immunopurification of GFP-Nup98$^{1\text{-}497}$, but not a C-terminal fragment, GFP-Nup98$^{498\text{-}920}$ (Figure 3.12b). Cumulatively, these results strongly argue that the Nup98-DHX9 complex is primarily present in the nucleoplasm, and that the N-terminal domain of Nup98 (containing FG/GLFG repeats and the RBD) interacts with DHX9.

3.3.3 Binding of Nup98 to DHX9 is enhanced by RNA

Since DHX9 and Nup98 both interact with RNA (Fuller-Pace, 2006; Ren et al., 2010), we also investigated the importance of RNA in their association. In Nup98 immunoprecipitates from HEK293T cell lysates we detected DHX9, PRKDC, and several Nups bound to Nup98 (Figure 3.13), consistent with the results presented in Figure 3.1. However, when parallel samples of bead-bound complexes were incubated with RNase A in amounts sufficient to degrade all detectable RNA (Höck et al., 2007; Moore et al., 2014; Ule et al., 2005; Zhang et al., 2008), levels of Nup98-associated DHX9 were reduced, while PRKDC and associated Nups were unaffected. These results imply that RNA, directly or indirectly, contributes to interaction of Nup98 with DHX9.

To extend our characterization of the *in vivo* interactions between Nup98, DHX9, and RNA, we examined whether recombinant Nup98 and DHX9 could interact using *in vitro* binding assays. Magnetic beads coupled to anti-DHX9 antibodies were incubated with GST-DHX9. GST alone or GST-Nup98 was then added to bead-bound GST-DHX9. Only the GST-Nup98 protein bound to the beads, suggesting a direct interaction between DHX9 and Nup98 (Figure 3.14a). Similar results were obtained using DHX9 and Nup98 lacking the GST tags (Figure 3.14b). As it was possible that RNA present within the *E. coli* extracts could contribute to the *in vitro* DHX9-Nup98 interaction, we also conducted binding reactions after pre-treating the recombinant

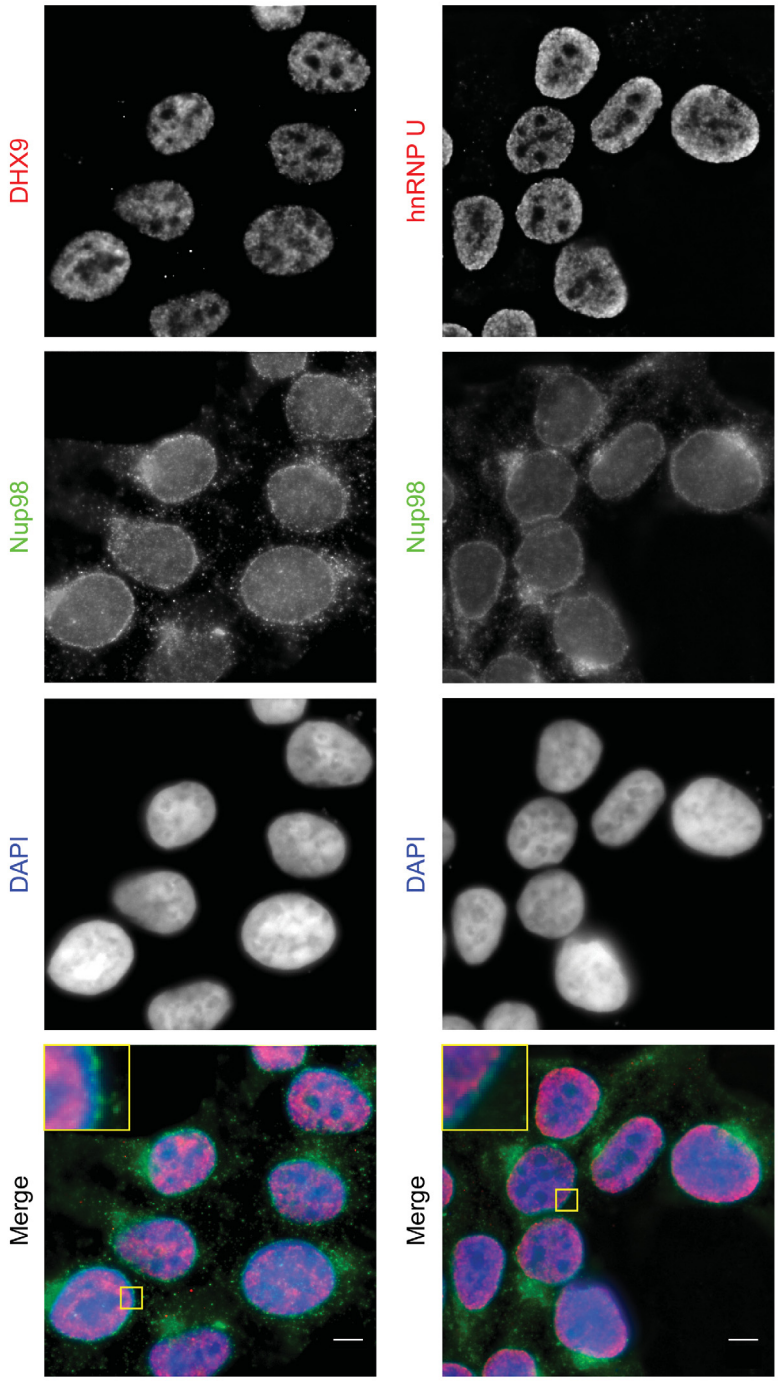


Figure 3.4: Localization of Nup98 with DHX9 and hnRNP U.

The cellular distribution of Nup98, DHX9, and hnRNP U in HEK293T cells was examined by indirect immunofluorescence using antibodies directed against each protein as indicated. The positions of nuclei were determined using the DNA stain DAPI. Merged images showing DHX9 or hnRNP U (red), Nup98 (green), and DAPI-stained DNA (blue) are shown. Note, DHX9 and hnRNP U are partially excluded from the nucleoli, which exhibits reduced DAPI staining. Scale bars, 5 μm. N = 3 biological replicates, representative immunofluorescence microscopy images are shown.

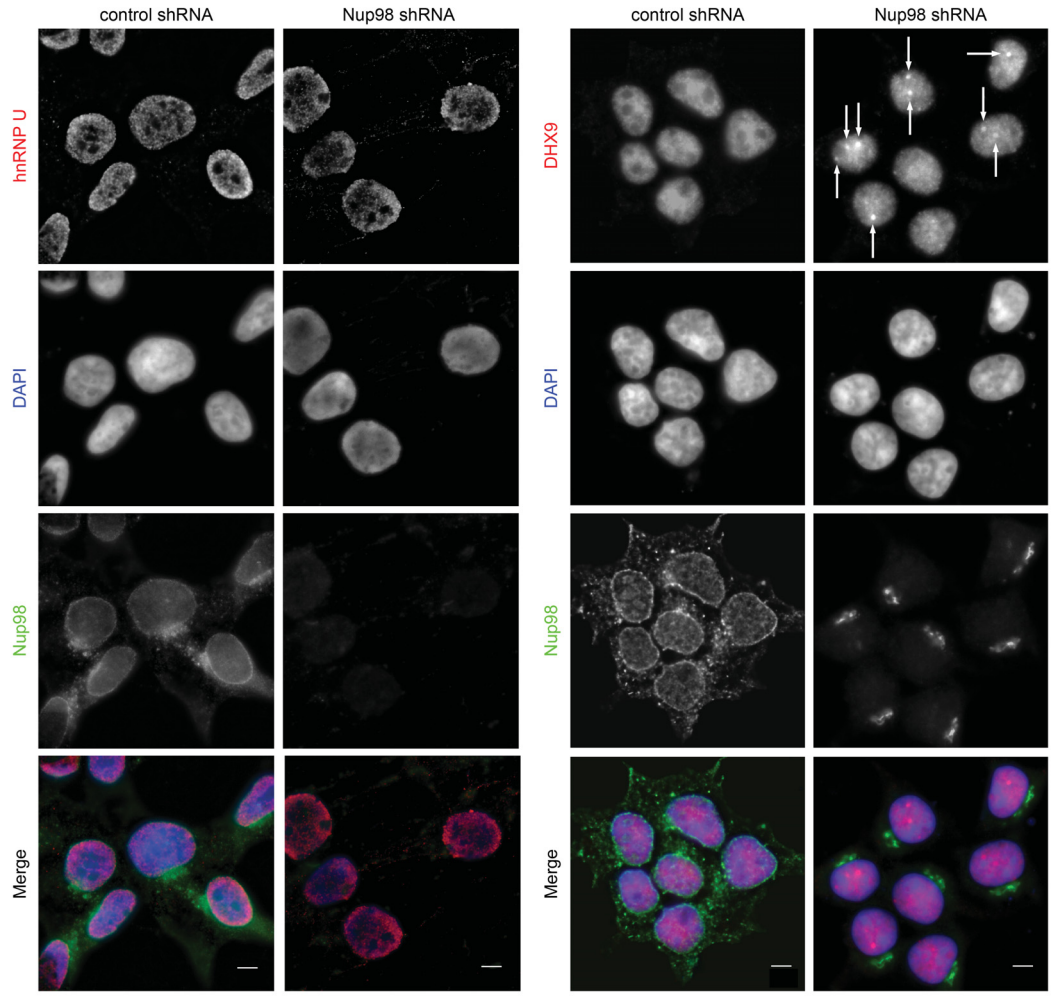


Figure 3.5: Nup98 depletion alters the intranuclear distribution of DHX9, but not hnRNP U.

HEK293T cells were transduced with lentiviruses encoding an shRNA targeting Nup98 or a control shRNA. Four days later the cellular distributions of Nup98 and either DHX9 or hnRNP U were examined by indirect immunofluorescence. Cells depleted of Nup98 show partial relocation of DHX9 into intranuclear foci (white arrows). Merged images show DHX9 or hnRNP U (red), Nup98 (green), and DAPI-stained DNA (blue). Scale bars, 5 μm. N = 3 biological replicates, representative immunofluorescence microscopy images are shown. The cellular localization of Nup98 is not affected by depletion of DHX9 (Figure 3.6). Protein depletion in these experiments was confirmed by immunoblot analysis (Figure 3.7).

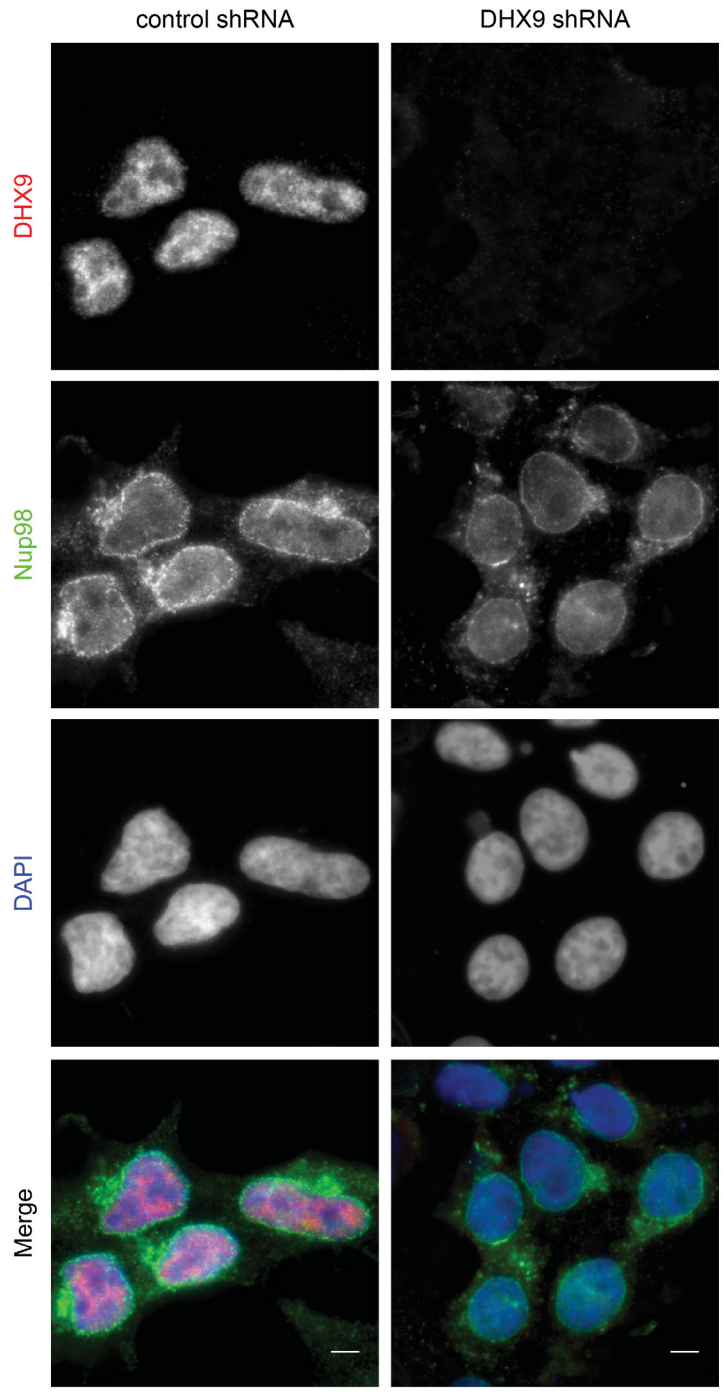


Figure 3.6: DHX9 depletion does not alter Nup98 localization in the cell.

HEK293T cells were transduced with lentiviruses encoding a control shRNA or an shRNA targeting DHX9. Four days later the cellular distribution of Nup98 and DHX9 were examined by indirect immunofluorescence. Merged images show DHX9 (red), Nup98 (green), and DAPI stained DNA (blue). Scale bars, 5 μm. N = 3 biological replicates, representative immunofluorescence microscopy images are shown. Protein depletion in these experiments was confirmed by immunoblot analysis (Figure 3.7).

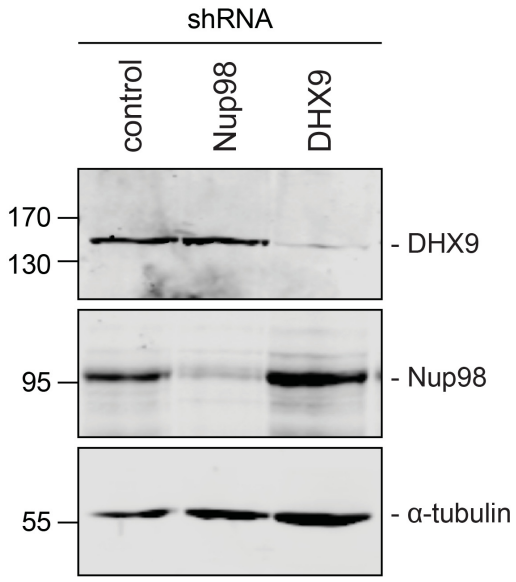


Figure 3.7: Immunoblotting of cell extracts following shRNA-mediated protein depletion.

Four days after transduction with lentiviruses encoding indicated shRNAs, lysates from cells depleted of the indicated protein (listed above top panel) were analyzed by western blotting using antibodies directed against Nup98, DHX9, and α-tubulin (load control) as indicated to the right of the panels. The positions of molecular mass markers (shown in kDa) are indicated on the left. N = 3 biological replicates, representative western blot images are shown. Corresponding immunofluorescence images are shown in Figure 3.5 and Figure 3.6.

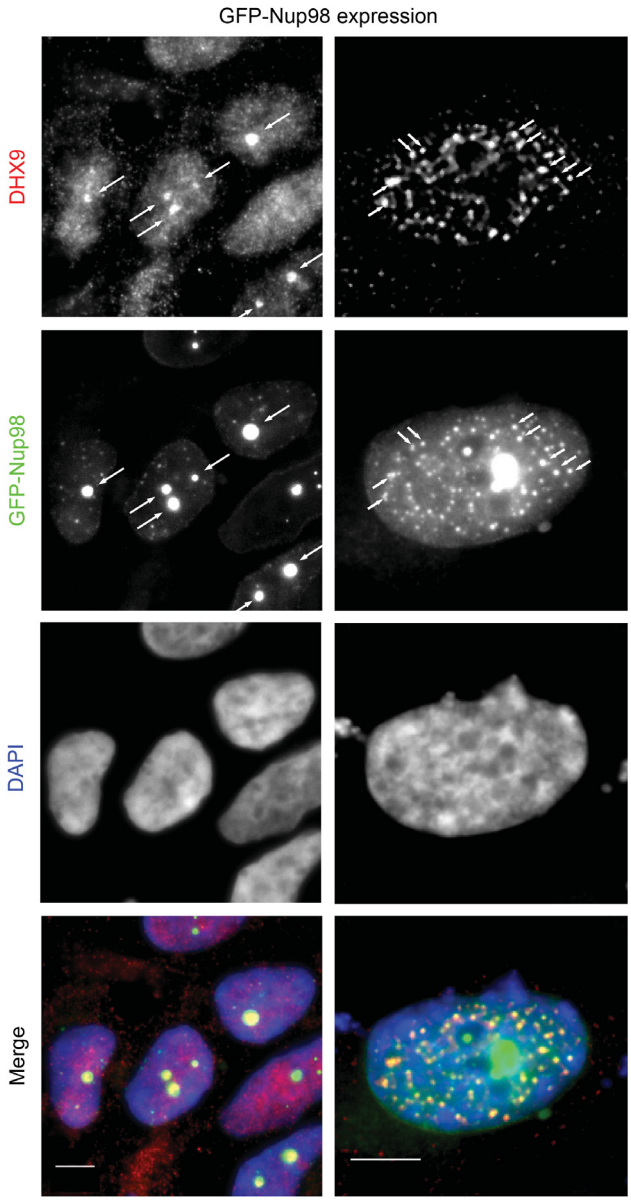


Figure 3.8: Intranuclear colocalization of GFP-Nup98 and DHX9

HEK293T cells expressing GFP-NUP98 were used to compare DHX9 and GFP-Nup98 localization by immunofluorescence microscopy. Two magnifications are shown, each showing that upon GFP-NUP98 expression intranuclear GFP-Nup98-containing foci form that contain DHX9. Examples of GFP-Nup98 colocalization with DHX9 are marked by arrows. Cells expressing higher levels of GFP-NUP98 (right column) contain greater numbers of GFP-Nup98 foci and display even more pronounced DHX9 colocalization. Merged images show DHX9 (red), GFP-Nup98 (green), and DAPI-stained DNA (blue). Scale bars, 5 μm. N = 3 biological replicates, representative immunofluorescence microscopy images are shown. In contrast, expression of GFP had no effect on DHX9 localization, and expressing GFP or GFP-NUP98 had no impact on hnRNP U localization (Figure 3.9). The presence of GFP-Nup98 was confirmed by immunoblot (Figure 3.10).

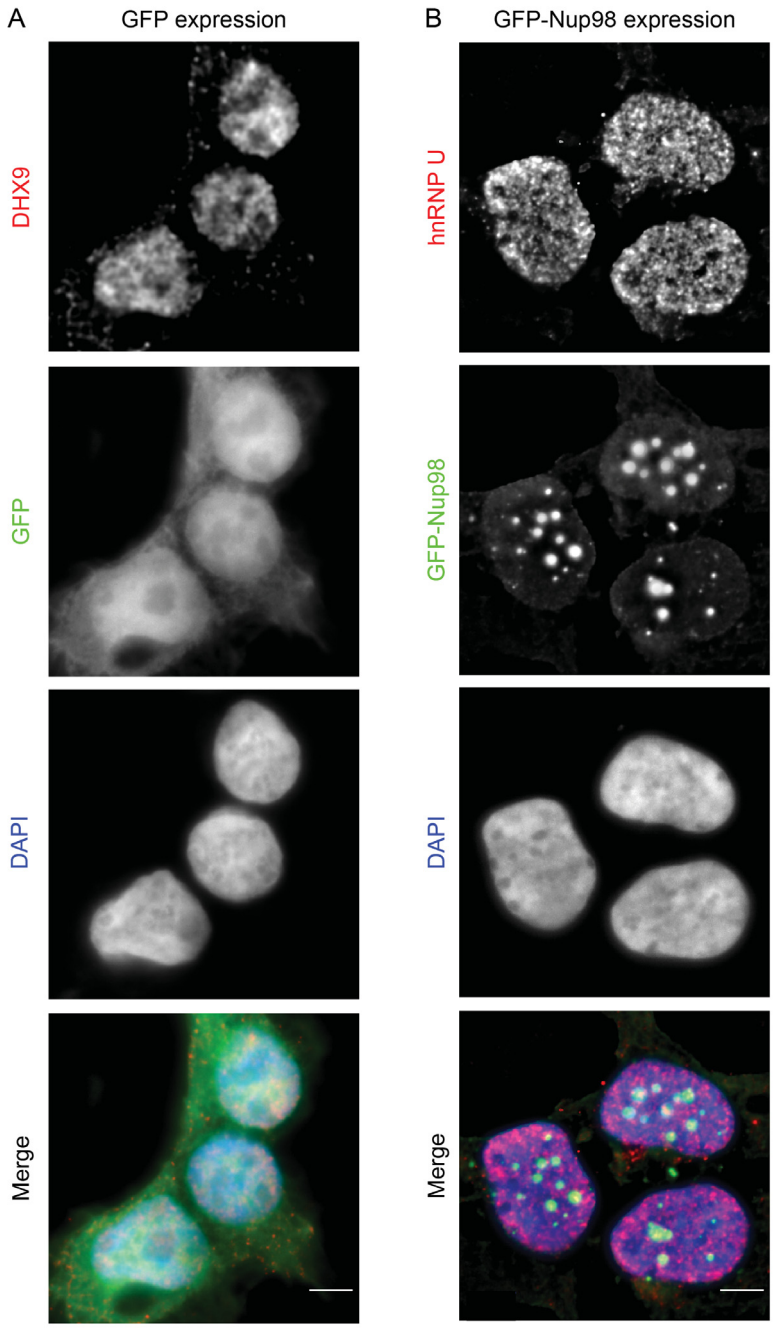


Figure 3.9: GFP expression does not alter the localization of DHX9, nor does GFP-Nup98 alter hnRNP U localization.

Figure 3.9 HEK293T cells expressing GFP or GFP-NUP98 were used to compare their localization to DHX9 or hnRNP U detected by indirect immunofluorescence microscopy using specific antibodies. Merged images show DHX9 or hnRNP U (red), GFP or GFP-Nup98 (green), and DAPI-stained DNA (blue). Scale bars, 5 μm. N = 3 biological replicates, representative immunofluorescence microscopy images are shown.

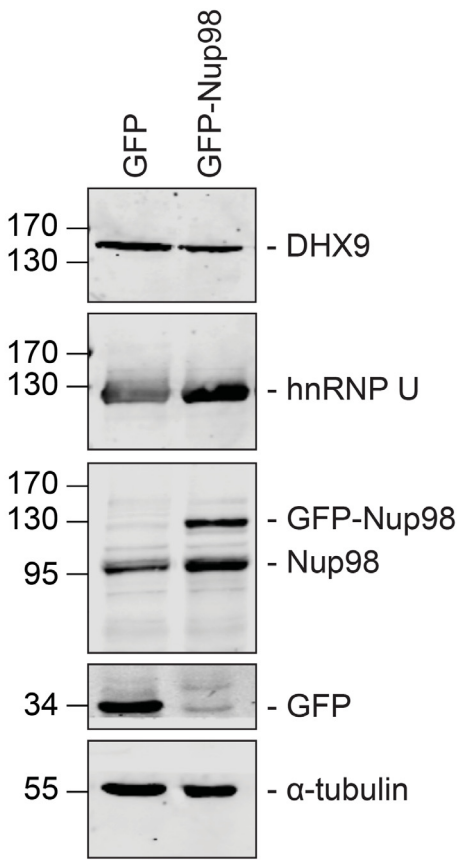
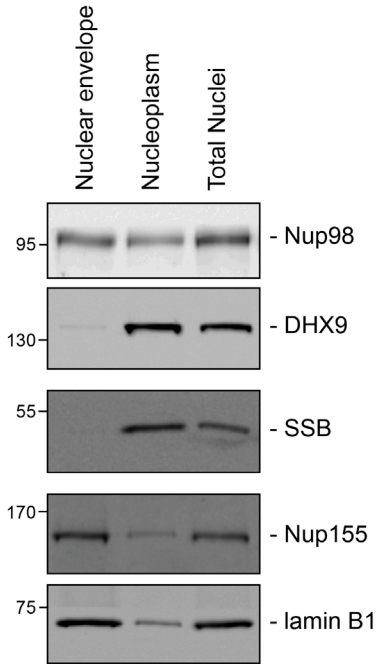
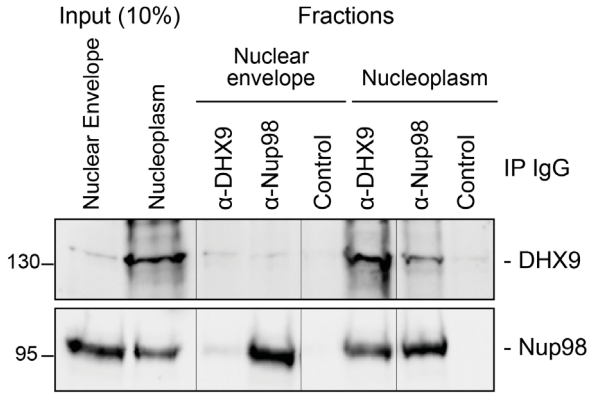


Figure 3.10: GFP or GFP-Nup98 expression does not alter cellular levels of DHX9 or hnRNP U.

Figure 3.10 Western blots of proteins derived from HEK293T cells lysates expressing GFP or GFP-Nup98 are shown. Antibodies were used for immunoblotting to detect the proteins indicated to the right of the panels. The positions of molecular mass markers (shown in kDa) are indicated on the left. Tubulin was used as loading control. N = 3 biological replicates, representative western blot images are shown. Corresponding immunofluorescence images are shown in Figure 3.8 and Figure 3.9.



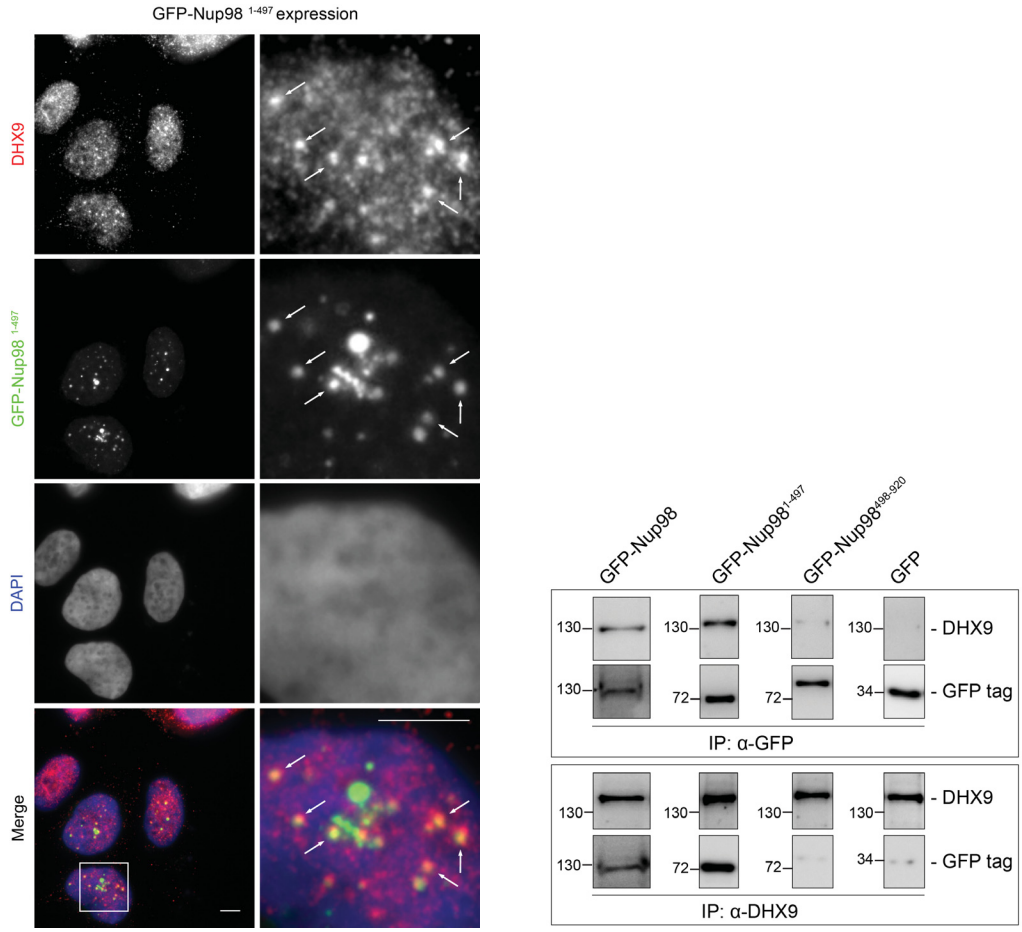
(a) Nuclei fractionation into nuclear envelope and nucleoplasm.



(b) DHX9 and Nup98 interact in the nucleoplasm.

Figure 3.11: DHX9 interacts with intranuclear Nup98.

3.11a HeLa cell nuclei were fractionated to produce nucleoplasmic and nuclear envelope fractions. Fractions were analyzed by western blotting using antibodies directed against the indicated proteins (right). The positions of molecular mass markers (shown in kDa) are indicated on the left. The fractionation procedure was evaluated by western blotting using antibodies directed against NE (lamin B and Nup155) and nucleoplasmic (SSB) proteins. N = 3 biological replicates, representative western blot images are shown. **3.11b** Nup98 or DHX9 were immunoprecipitated from nucleoplasmic and NE fractions derived from HeLa cell nuclei. Co-purifying proteins from the samples indicated above the panels were separated by SDS-PAGE and analyzed by immunoblotting to detect DHX9 and Nup98 as specified to the right of the panels. The positions of molecular mass markers (shown in kDa) are indicated on the left. N = 3 biological replicates, representative western blot images are shown.



(a) DHX9 is recruited to intranuclear GFP-Nup98$^{1\text{-}497}$ foci.

(b) DHX9 interacts with the N-terminal domain of Nup98.

Figure 3.12: DHX9 binds to GFP-Nup98$^{1\text{-}497}$ in intranuclear foci.

3.12a HEK293T cells expressing GFP-NUP98$^{1\text{-}497}$ were used to assay the localization of DHX9 detected by immunofluorescence microscopy. Two magnifications are shown and examples of GFP-Nup98$^{1\text{-}497}$ colocalization with DHX9 are highlighted with arrows. Merged images show DHX9 (red), GFP-Nup98$^{1\text{-}497}$ (green), and DAPI-stained DNA (blue). Scale bars, 5 μm. N = 3 biological replicates, representative immunofluorescence microscopy images are shown. **3.12b** HEK293T cells were transfected with GFP-Nup98 constructs (full length: 1-920, N-terminal domain: 1-497, C-terminal: 498-920) for 24 hours and lysed. Anti-DHX9 or anti-GFP IgG was used to affinity purify protein complexes from cell lysates. Co-purifying proteins were separated by SDS-PAGE and analyzed by immunoblotting using anti-DHX9 or anti-GFP antibodies. Antibodies were used for immunoblotting to detect the proteins indicated to the right of the panels. The positions of molecular mass markers (shown in kDa) are indicated on the left. N = 3 biological replicates, representative western blot images are shown.

proteins with RNase A (to remove any residual RNA) or adding excess RNA (poly I:C). Binding reactions conducted under these conditions reveal that the interaction of untagged or GST-tagged DHX9 with Nup98 was partially reduced by the addition of RNase A, while the addition of RNA (poly I:C) did not appear to alter the binding of DHX9 to Nup98 (Figure 3.14a and Figure 3.14b).

A second *in vitro* assay previously employed to assess Nup-Nup interactions (Patel and Rexach, 2008; Zhou et al., 2013) was also used to evaluate the interaction of Nup98 with DHX9. Termed the bead halo assay, protein (DHX9) is bound to beads and then incubated with a potential binding partner (Nup98). Binding of Nup98 to bead-associated DHX9 is detected using anti-Nup98 antibodies and fluorescently labeled secondary antibodies. Interactions between the proteins are visualized by a fluorescent signal on the surface of the beads (Figure 3.15b). The level of bead-associated fluorescence signal provides a relative measure of the strength of the interaction (Patel and Rexach, 2008; Zhou et al., 2013). Using this assay, we detected and quantified the binding of recombinant Nup98 to DHX9 (Figure 3.15a and 3.15b), and again the addition of RNA did not significantly alter the relative strength of this interaction, but the inclusion of RNase A reduced the level of DHX9 binding to Nup98 (Figure 3.15a and 3.15b).

The bead halo assay was also used to identify regions of DHX9 that interact with Nup98. GST-Nup98 bound to beads was incubated with three consecutive, non overlapping domains of DHX9-GFP (Figure 3.16a). We observed that an N-terminal region (aa 1-380), containing two dsRNA binding motifs (DRBM1 and DRBM2) followed by the MTAD, and a C-terminal segment (aa 821-1270), containing an (OB)-binding fold and a ssRNA-binding RGG-box, bound to Nup98. Conversely, DHX9s central region (aa 381-820), containing its helicase domain (Zhang and Grosse, 1997), did not bind GST-Nup98 under these conditions (Figure 3.16a and 3.16b). The interactions of the N- and C-terminal domains of DHX9-GFP with GST-Nup98 appeared to be facilitated by the presence of RNA, as these interactions were sensitive to RNase A. Furthermore, the addition of RNA prior to mixing of the two proteins stimulated binding of GST-Nup98 and the N-terminal domain of DHX9-GFP (Figure 3.16a). Cumulatively, these data lead us to conclude that DHX9 can directly bind Nup98 and that their association is augmented by RNA. The DHX9-Nup98 interaction is likely mediated by the N- and C-terminal domains of DHX9.

3.3.4 Nup98 stimulates the ATPase activity of DHX9

Like other RNA helicases, DHX9 can bind and hydrolyze ATP, which can promote unwinding of duplex RNA and remodelling of RNA-protein complexes (Fullam and

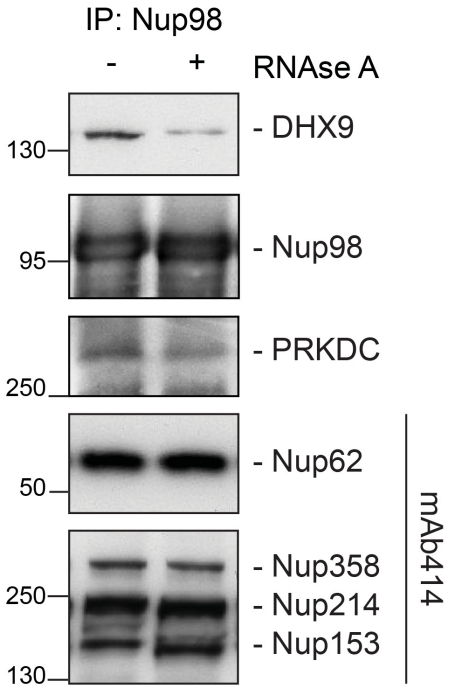
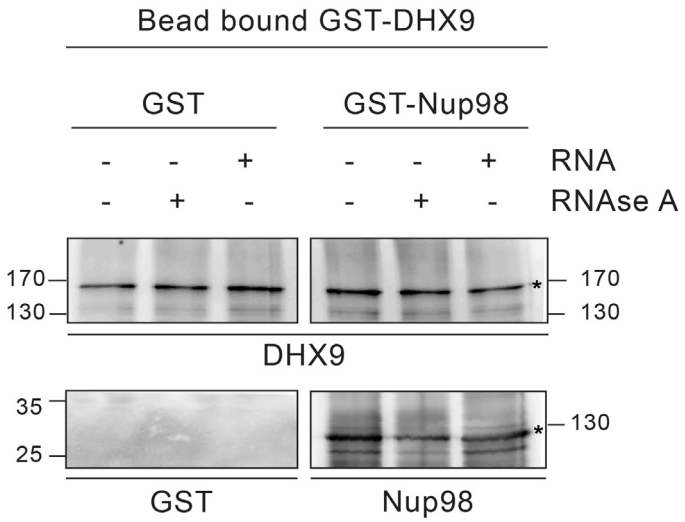
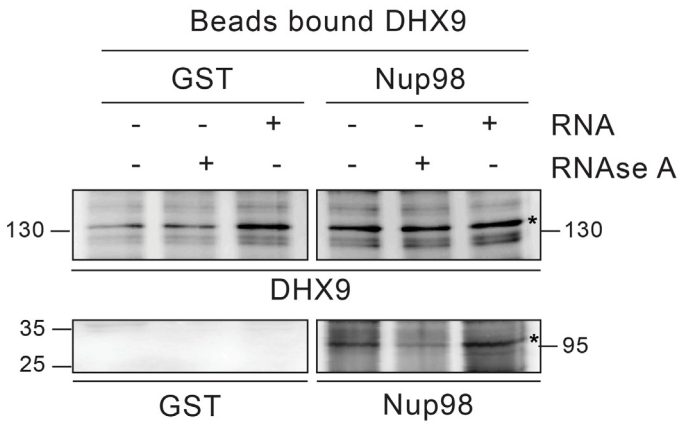


Figure 3.13: RNAse A decreases the interaction of Nup98 with DHX9 *in vivo*

Nup98 was affinity purified from HEK293T cell lysates. Bead-bound protein complexes were then incubated with or without RNase A, and proteins remaining bound to Nup98 were analyzed by western blotting using antibodies directed against the indicated proteins (right). The positions of molecular mass markers (shown in kDa) are indicated on the left. N = 3 biological replicates, representative western blot images are shown.



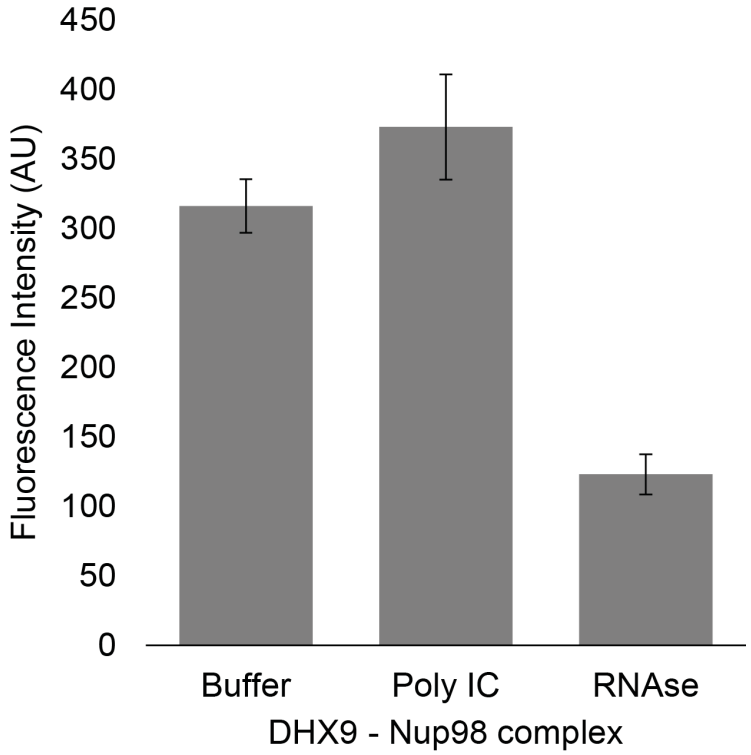
(a)



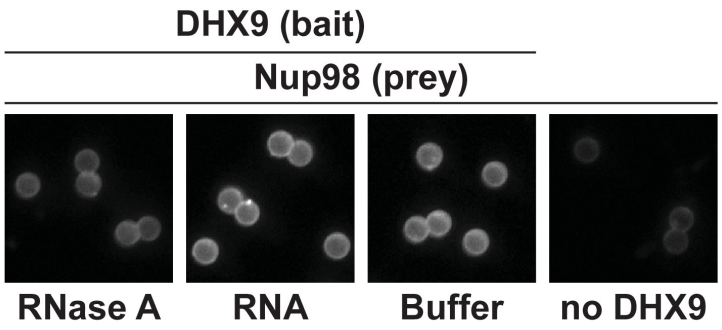
(b)

Figure 3.14: Nup98 binds directly to DHX9 *in vitro*.

3.14a Anti-DHX9 antibodies coupled to beads were used to immobilize GST-DHX9. Bead-bound GST-DHX9 was incubated with GST-Nup98 or GST alone in the presence of RNA (poly I:C), RNase A, or buffer alone. Bound proteins were analyzed by western blotting using the indicated antibodies (below each panel). The top row of images shows the GST-DHX9 bait bound to beads. The bottom row of images shows GST and GST-Nup98 that bound to GST-DHX9 under the indicated conditions. Asterisks denote positions of GST-DHX9 and GST-Nup98. The positions of molecular mass markers (shown in kDa) are indicated. **3.14b** The assay described above was reproduced with untagged recombinant proteins. A similar interaction between untagged recombinant DHX9 and Nup98 was also detected. N = 3 biological replicates, representative western blot images are shown.



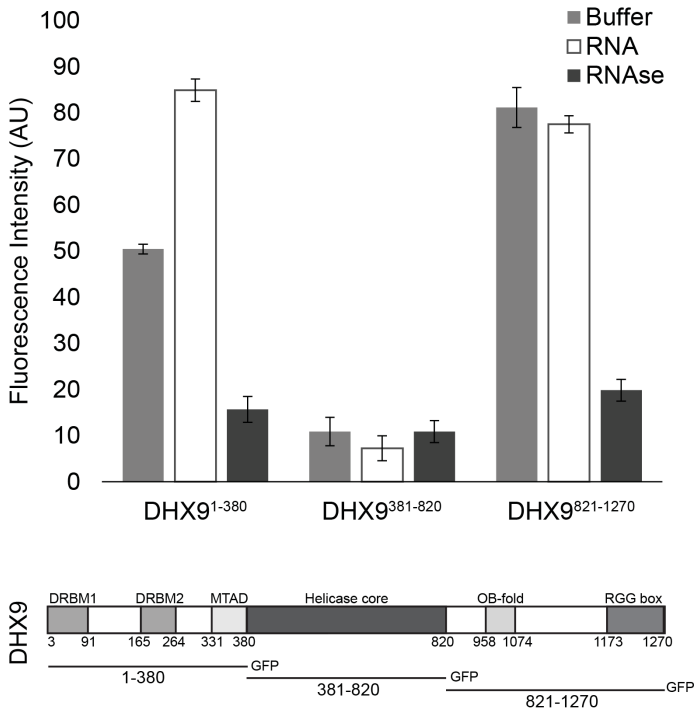
(a) Quantification of Nup98 and DHX9 interaction *in vitro*.



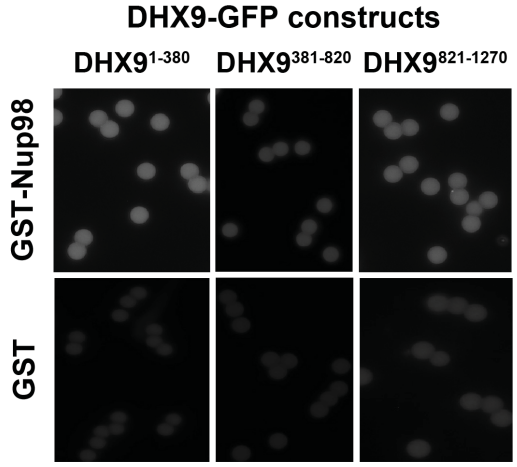
(b) Examples of images of bead-bound complexes

Figure 3.15: RNAse A decreases the Nup98-DHX9 interaction *in vitro*.

Bead halo assays were performed with purified recombinant proteins, using DHX9 immobilized on beads with an anti-DHX9 antibody as bait and Nup98 as prey. Prior to the binding step, both proteins were incubated with RNase A, RNA (poly I:C), or buffer alone. Interactions of Nup98 with bead-bound DHX9 were detected by fluorescence microscopy with rabbit anti-Nup98 antibodies and Alexa Fluor 488 donkey anti-rabbit antibodies. Plots show average fluorescence intensity values of beads (arbitrary units) corrected against negative control assays. Error bars indicate standard deviation between biological replicates (N = 3 biological replicates).



(a) Quantification of Nup98 and DHX9 domains interaction *in vitro*.



(b) Examples of images of bead-bound complexes

Figure 3.16: RNAse A decreases the interaction of Nup98 with the N and C-terminal domains of DHX9

Bead halo assays were performed with purified recombinant proteins, using bead-bound GST-Nup98 (bait) and different domains of DHX9-GFP (prey; see bottom diagram). Proteins were incubated with RNA (poly I:C), RNase A, or buffer alone before binding. The interaction of bead-bound GST-Nup98 with DHX9-GFP domains was detected by fluorescence microscopy. Plots show average fluorescence intensity values of beads (arbitrary units) corrected against negative control assays. Error bars indicate standard deviation between biological replicates (N = 3 biological replicates).

Schröder, 2013; Fuller-Pace, 2006; Zhang and Grosse, 1994).

Binding partners of RNA helicases have been shown to regulate helicase function by inhibiting or stimulating their ATPase activity (Bourgeois et al., 2016) and we hypothesized that Nup98 could play a similar role with DHX9. To test this, ATPase activity of recombinant GST-DHX9 was examined in the presence and absence of GST-Nup98. We observed a basal ATPase rate for recombinant DHX9 that was stimulated by the addition of RNA to levels comparable to that previously reported for DHX9 (Figure 3.17) (Schütz et al., 2010; Zhang and Grosse, 1991).

In the presence of excess RNA (poly I:C), the addition of GST-Nup98 induced a dose-dependent increase in GST-DHX9 ATPase activity reaching levels approximately five-fold higher than GST-DHX9 and RNA alone at the highest GST-Nup98 concentration tested (Figure 3.18a). A similar level of stimulation was also observed using untagged versions of DHX9 and Nup98 (Figure 3.18b).

DHX9 ATPase activity was also increased upon addition of the N-terminal domain of Nup98 (containing FG/GLFG repeats and a RBD). By contrast, neither the C-terminal domain of Nup98 or GST alone caused significant changes in DHX9 ATPase activity, nor could GST-Nup98 stimulate GST-DHX9 in the absence of RNA (Figure 3.19). These data indicate that Nup98 functions as a positive regulator of DHX9 ATPase activity in the context of RNA.

3.3.5 Nup98 and DHX9 interact with a shared subset of mRNAs and gene loci

Given that Nup98 and DHX9 exists in a complex *in vivo*, and that Nup98 stimulates DHX9 activity in the presence of RNA, we would expect that Nup98 and DHX9 interact with a shared set of mRNAs. To assess this, we compared recently published mRNA binding datasets for Nup98 (Hendrickson et al., 2016) and DHX9 (Erkizan et al., 2015). We find a statistically significant overlap in these datasets with 37% of the Nup98-interacting mRNAs also detected bound to DHX9 and 40% of the DHX9 bound transcripts interacting with Nup98 ($p = 2.5 \times 10^{-93}$). To directly test whether these proteins bind similar mRNAs, we immunoprecipitated DHX9 and Nup98 from cell lysates following crosslinking. By using stringent conditions that disrupt the DHX9-Nup98 interaction (Figure 3.20b), we could assess the ability of each protein to bind RNA independent of one another and determine whether they interact with similar RNA species. RT-PCR was used to test whether specific mRNA species were associated with the immunopurified proteins. We tested for the presence of several potential interacting mRNAs, encoding JunD, Myc, FoxP2, HoxA2, and ZFY (Erkizan et al., 2015; Hendrickson et al., 2016; Hartman et al., 2006; Ranji et al., 2011; Wei-

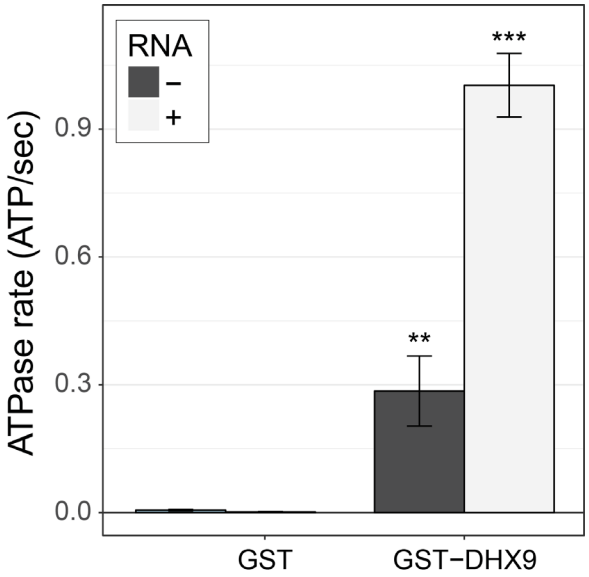


Figure 3.17: ATPase activity of recombinant GST-DHX9.

The average ATP hydrolysis rate (ATP/sec) of purified recombinant GST-DHX9 or GST alone in the presence or absence of RNA was examined. Error bars indicate standard deviation (N = 3 biological replicates). Results from 3 biological replicates were submitted to ANOVA followed by Tukey HSD tests (*** indicates adjusted p-values < 0.001 for Tukey HSD in pairwise comparison between reactions containing GST-DHX9 and reactions containing GST alone in the presence or absence of RNA).

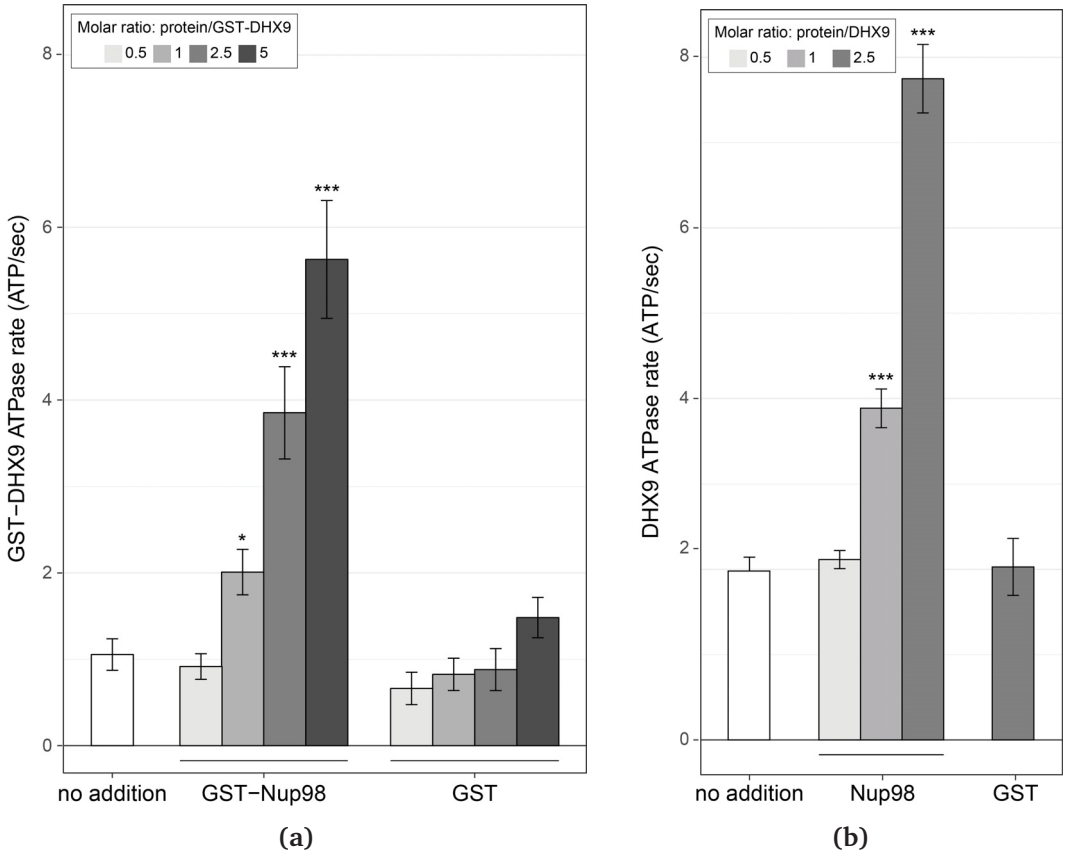


Figure 3.18: GST-Nup98 stimulates GST-DHX9 ATPase activity.

The average ATPase activity (ATP hydrolysis rate) of purified recombinant GST-DHX9 (a) or untagged DHX9 (b) in the presence of RNA alone (no addition) or following the addition of increasing concentrations of GST-Nup98 (a), untagged Nup98 (b) or GST (a and b) (show as the molar ratio of the added protein to that of DHX9) is shown on the y-axis. Error bars indicate standard deviation (N = 3 biological replicates). Results from 3 biological replicates were submitted to ANOVA followed by Tukey HSD tests (*** indicates adjusted p-values < 0.001 and * < 0.05 for Tukey HSD in pairwise comparison between a reaction containing Nup98 and a reaction containing GST in similar molar amounts).

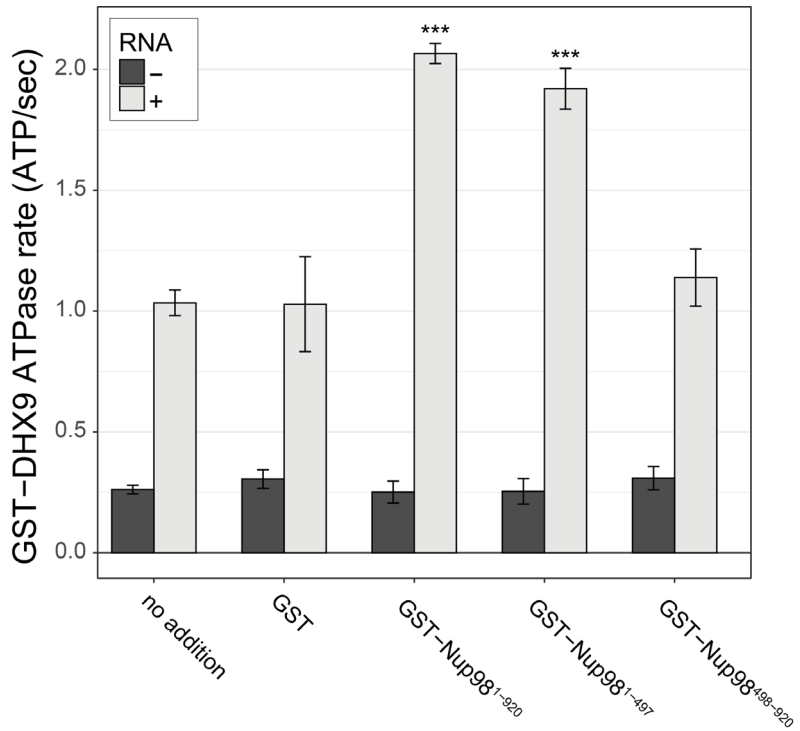


Figure 3.19: Nup98 stimulates DHX9 ATPase activity via its FG/GLFG domain

The average ATP hydrolysis rate (ATP/sec) of purified recombinant GST-DHX9 in the presence or absence of RNA (poly I:C), and either alone (no addition) or following the addition of equal molar amounts of GST or GST-Nup98 constructs containing full length Nup98 (1-920), the N-terminal region (1-497), or C-terminal region (498-920). Error bars indicate standard deviation (N = 3 biological replicates). Results from 3 biological replicates were submitted to ANOVA followed by Tukey HSD tests (*** indicates adjusted p-values < 0.001 for Tukey HSD in pairwise comparison between a reaction containing a Nup98 construct and a reaction containing GST).

densdorfer et al., 2009; Yugami et al., 2007), and two predicted negative controls NHLH2, and HEXIM1. As anticipated, JunD, Myc, FoxP2, HoxA2, and ZFY encoding mRNAs were detected bound to both DHX9 and Nup98 (Figure 3.20a). By contrast, both NHLH2 and HEXIM1, showed no interaction with either DHX9 or Nup98. Both Nup98 and DHX9 also interacted with the Adenovirus early region 1A (E1A) encoding RNA, a well known splicing reporter whose metabolism is regulated by several hnRNPs and RNA helicases (Zheng, 2010). These results suggest that Nup98 and DHX9 interact with, and potentially regulate, a shared set of mRNAs *in vivo*.

To investigate the interdependencies of mRNA-binding, we depleted either Nup98 or DHX9 and evaluated mRNA binding by the other factor. Note that depletion of Nup98 or DHX9 did not alter cellular levels of the other protein or the efficiency of immunoprecipitation (Figure 3.22). As shown in Figure 3.21, upon depletion of DHX9, five of six mRNAs (*E1A, FOXP2, HOXA2, MYC* and *ZFY*) showed a significant decrease in Nup98 association relative to the input as compared to mock-depleted cells. By contrast, depletion of Nup98 led to a significant increase in the amount of each of the six mRNAs bound to DHX9.

The changes in the association of Nup98 or DHX9 with these mRNAs does not appear to be due to a change in the nuclear export status of the tested mRNAs (Figure 3.23a). These results show that Nup98 and DHX9 influence each others association with mRNA, and are consistent with a model in which DHX9 promotes the association of Nup98 with specific mRNAs, and Nup98 facilitates the release of these mRNAs from DHX9.

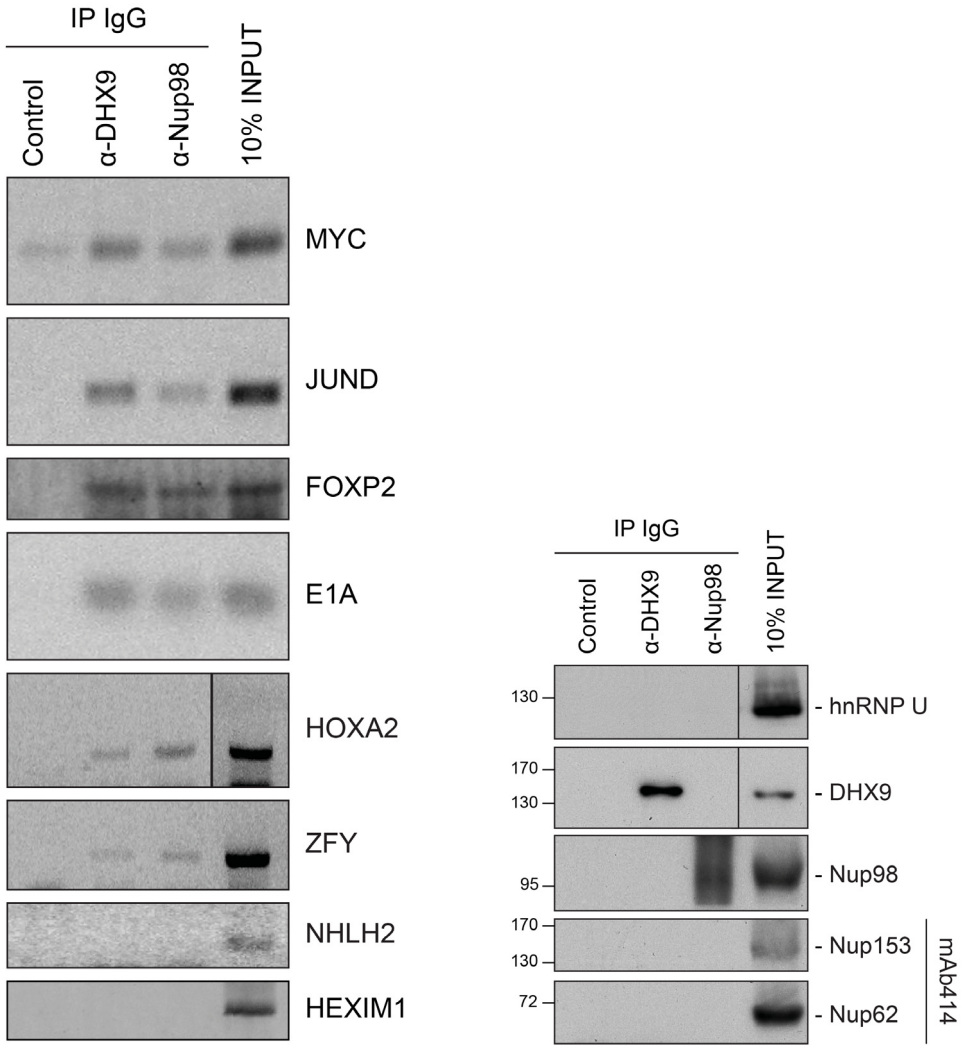
To further evaluate the nature of the shared binding of Nup98 and DHX9 to this set of mRNAs, we used the DamID assay (Franks et al., 2016) to determine whether the Nup98-DHX9 complex interacted with the gene loci encoding these mRNAs. For this analysis, genes encoding Nup98 or DHX9 fused to *E.coli* DNA methyltransferase (Dam) were introduced in to HEK293T cells. Modified DNA was then amplified, purified, and used in qPCR reactions to assess whether specific regions of the genome were bound to Nup98 and DHX9. As shown in Figure 3.24, both Nup98 and DHX9 mapped to the six gene loci whose transcripts were bound to Nup98 and DHX9. For *JUND* and *MYC*, two regions were examined: a 5' region containing the promoter and a region within the 3' half of the ORF. For both Nup98 and DHX9, robust binding was detected to the 5' promoter of these genes, while binding was greatly reduced or absent from regions within their ORFs. Similarly, no detectable binding was observed to the NHLH2 and HEXIM1 genes, consistent with our observation that their transcripts were not detected in association with Nup98 and DHX9 (Figure 3.24). Of note, DamID experiments performed with a fusion (Dam-Nup98$^{1\text{-}504}$) containing only the N-terminal region of Nup98 shown to be sufficient for DHX9 binding (Figure 3.12b) displayed a similar chromatin-binding profile (Figure 3.24). The binding

of Nup98 and DHX9 to the gene loci tested was also interdependent. Depletion of Nup98 or DHX9 significantly reduced the interactions of its binding partner with the target gene (Figure 3.24). We therefore suggest that the Nup98-DHX9 complex binds to specific genes and their transcripts.

3.3.6 Nup98 stimulates DHX9-mediated transcription.

DHX9 and Nup98 have been linked to various steps in mRNA metabolism, including transcription (Franks et al., 2016; Lee and Pelletier, 2016; Liang et al., 2013; Light et al., 2013; Pascual-Garcia et al., 2014; Fidaleo et al., 2016). Consistent with these data, the specific genes we detected bound to Nup98 and DHX9 exhibited altered transcript levels (Figure 3.25). Furthermore, analysis of RNA-Seq data (Chen et al., 2014; Franks et al., 2016) revealed shared sets of genes with altered transcription upon depletion of these proteins. A comparison of these data sets shows significant overlap in the identity of gene products affected by the depletion of either protein (287 genes with altered expression upon DHX9 or Nup98 depletion, $p = 3.24 \times 10^{-36}$), consistent with the idea that these proteins form a functional complex. Interestingly, a significant number (p-value 2.38×10^{-4}) of those genes exhibiting altered expression upon Nup98 depletion contain a putative cAMP-response element (CRE) (Zhang et al., 2005), a regulatory element whose transcriptional activity can be regulated by DHX9 (Aratani et al., 2001; Fidaleo et al., 2016; Lee and Pelletier, 2016).

To more directly assess the functional role of Nup98 in DHX9-mediated transcription, we used a CRE-luciferase reporter assay. This assay has been used to evaluate the role of DHX9 in transcription, including defining the contributions of its ATPase activity to its role in transcription (Aratani et al., 2001). Similar to previous reports, expression of exogenous DHX9 in cells containing the CRE-luciferase reporter increased production of luciferase (Figure 3.26a). Point mutants in DHX9 that reduce ($DHX9^{I347A}$) or eliminate ($DHX9^{K417R}$) ATPase activity show reduced stimulation of reporter expression (Aratani et al., 2001). Since our $^{\text{in vitro}}$ assays showed that Nup98 could stimulate the ATPase activity of DHX9, we tested whether overexpression of Nup98 could stimulate the DHX9-mediated expression of CRE-luciferase. In the absence of exogenous DHX9, expression of Nup98 had no significant affect on the expression of luciferase. However, Nup98 expression stimulated luciferase production in the presence of DHX9 or the ATPase compromised $DHX9^{I347A}$ mutant, while having no significant stimulatory impact on luciferase expression in the presence of the ATPase dead mutant ($DHX9^{K417R}$). These results are consistent with our *in vitro* observations showing Nup98 can stimulate the ATPase



(a) RT-PCR **(b) Western blot**

Figure 3.20: Nup98 and DHX9 directly interact with target mRNA molecules.

3.20a Following crosslinking of HEK293T cells to preserve protein/RNA complexes, cell lysates were incubated with beads coupled to a control IgG (anti-GFP) or beads coupled to Nup98 or DHX9 specific antibodies. RNA present in immunoprecipitated complexes and total cellular RNA (10% input) was used as template in RT-PCR reactions containing primers specific to regions of several cDNAs whose genes are denoted on the right. **3.20b** Western blot analysis of protein immunoprecipitation fractions from RNA-IPs used to detect associated RNA. The indicated IP samples (IP IgG) were analyzed by western blotting using antibodies directed against the proteins indicated on the right. The positions of molecular mass markers (shown in kDa) are indicated on the left. N = 3 biological replicates, representative images of western blots and agarose gels are shown.

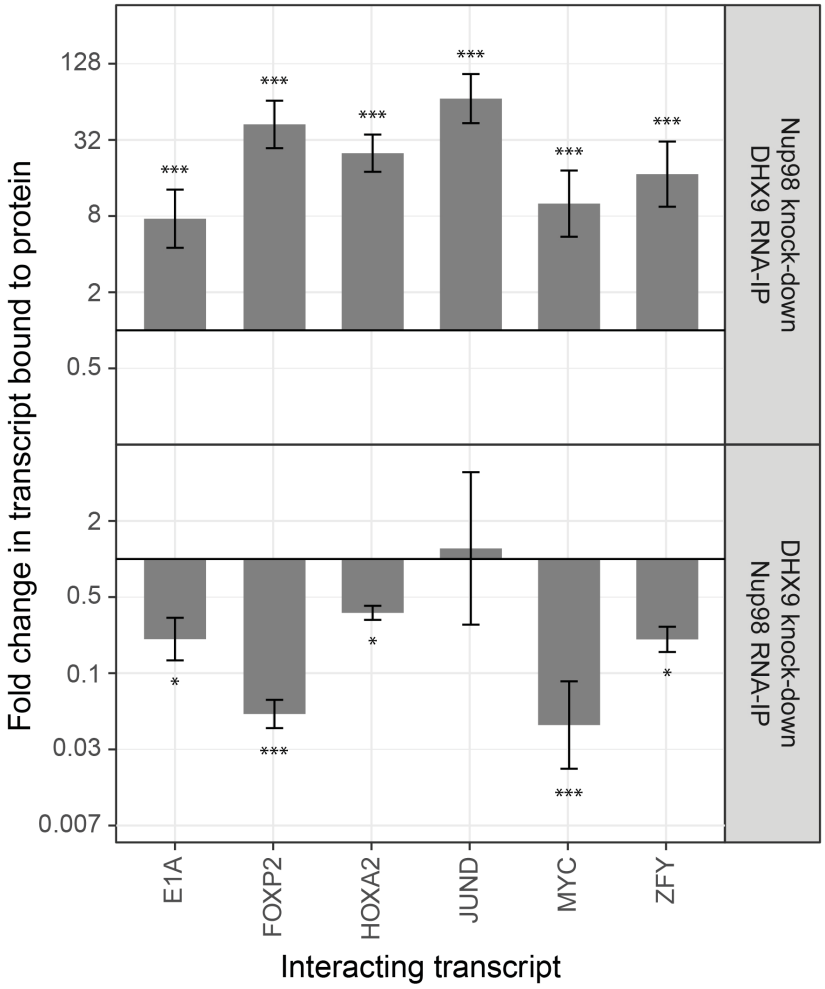


Figure 3.21: The association of Nup98 or DHX9 with specific mRNAs is altered by depletion of its binding partner.

HEK293T cells were transduced with lentiviruses encoding a control shRNA or an shRNA targeting Nup98 or DHX9. RNA immunopurified with Nup98 or DHX9 was reverse transcribed and used in qPCR reactions to assess the levels of indicated transcripts (x-axis). The average ratio of bound mRNA relative to input was determined for each transcript listed. The fold-change in this ratio, relative to that determined from mock-depleted cells, is shown on the y-axis. Top panel: mRNA bound to DHX9 upon Nup98 depletion; bottom panel: mRNA bound to Nup98 upon DHX9 depletion. Error bars indicate standard deviation (N = 3 biological replicates). Results from 3 biological replicates were submitted to ANOVA followed by Tukey HSD tests. The *** indicates adjusted p-values < 0.001 and * < 0.05 for Tukey HSD in pairwise comparisons between depleted and mock depleted samples.

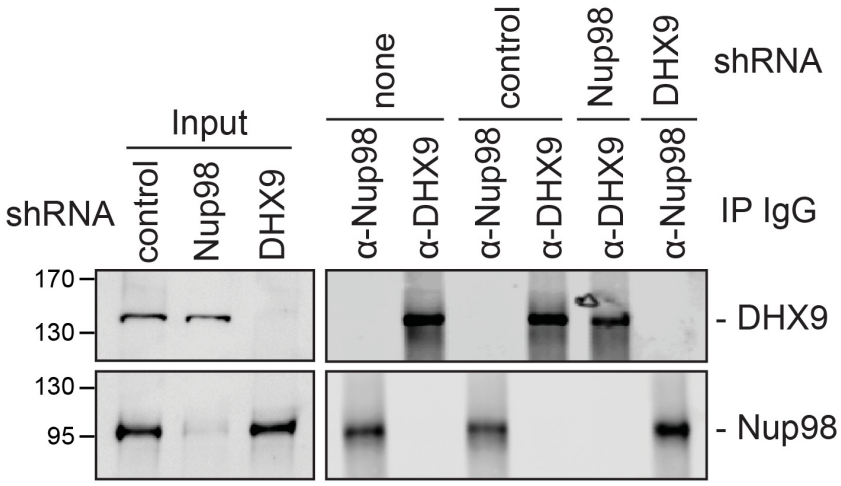
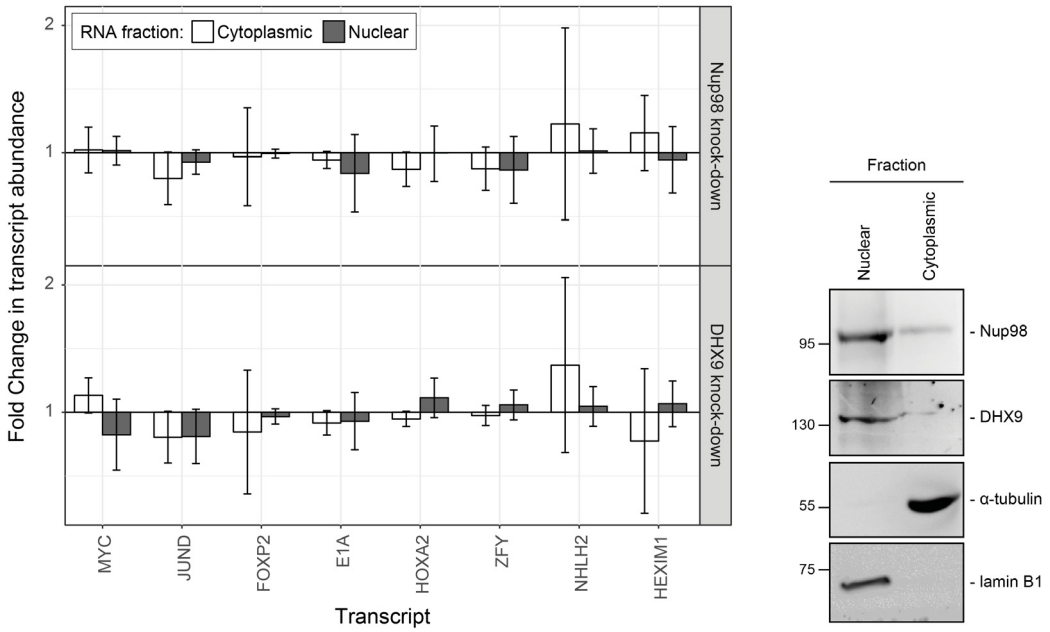


Figure 3.22: Nup98 or DHX9 RNAA-IP upon depletion of its binding partner.

HEK293T cells were transduced with lentiviruses encoding a control shRNA or an shRNA targeting Nup98 or DHX9. Western blot analysis of cell lysates and protein immunoprecipitation fractions are shown. The indicated cell lysates and IP samples (list above the panels) were probed with antibodies directed against DHX9 (top row) and Nup98 (bottom row). The positions of molecular mass markers (shown in kDa) are indicated on the left. N = 3 biological replicates, representative images of western blots are shown.



(a) Real time PCR **(b) Western blot**

Figure 3.23: Nup98 or DHX9 depletion has no significant impact on the nuclear or cytoplasmic abundance of target mRNAs

HEK293T cells were transduced with lentiviruses encoding a control shRNA or an shRNA targeting Nup98 or DHX9. Cells were fractionated into nuclear and cytoplasmic samples. **3.23a** The levels of the indicated gene transcripts were quantified by qPCR. Average nuclear and cytoplasmic transcript abundance was normalized to average total transcript abundance in the cell. Fold-changes (y-axis) in nuclear or cytoplasmic RNA levels upon Nup98 (top) or DHX9 (bottom) depletion relative to mock-depleted cells are shown. Error bars indicate standard deviation (N = 3 biological replicates). Results from 3 biological replicates were submitted to ANOVA tests and show no statistically significant changes. **3.23b** Fractions were evaluated by western blotting using antibodies directed against the Nup98, DHX9, the nuclear protein lamin B1, and cytoplasmic protein α-tubulin. The positions of molecular mass markers (shown in kDa) are indicated on the left. N = 3 biological replicates, representative images of western blots are shown.

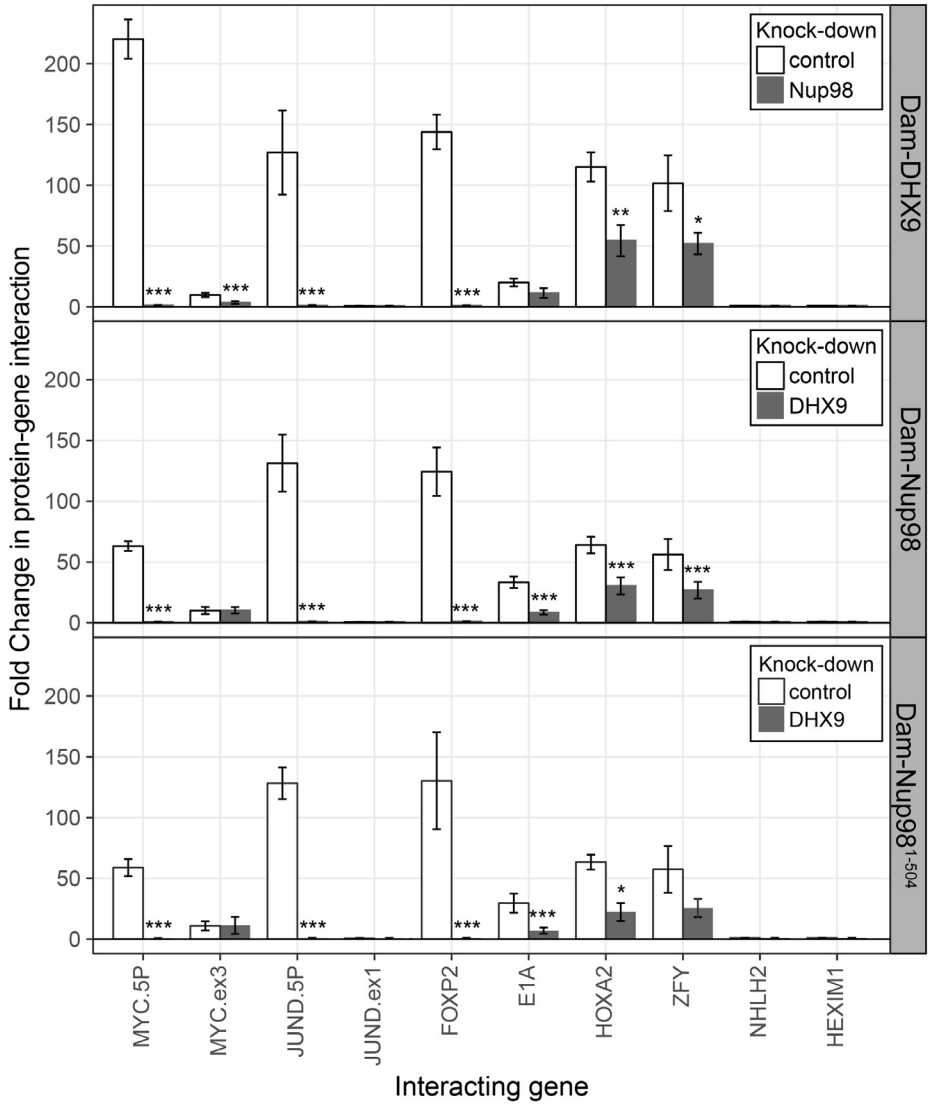


Figure 3.24: Nup98 and DHX9 associate with similar gene loci and their binding is interdependent.

HEK293T cells stably expressing Dam-GFP, Dam-Nup98, Dam-Nup98$^{1\text{-}504}$ or Dam-DHX9 were transduced with lentiviruses encoding a control shRNA (white) or an shRNA targeting Nup98 or DHX9 (gray) and Dam-ID analysis was performed. The average association of Dam-Nup98 and Dam-DHX9 to the indicated gene loci is represented as the fold-change (x-axis) relative to a Dam-GFP control. Error bars indicate standard deviation (N = 3 biological replicates). For the top and bottom graphs, results from 3 biological replicates were submitted to two-way ANOVA followed by Tukey HSD tests. Adjusted p-values are indicated as *** $<$ 0.001 $<$ ** $<$ 0.01 $<$ * $<$ 0.05 for Tukey HSD in pairwise comparisons between mock and Nup98 or DHX9 depleted cells for each gene tested.

activity of DHX9, and they suggest that the stimulatory effect of Nup98 binding to DHX9 supports its role in transcription.

Steps in mRNA metabolism are often tightly coupled, including transcription and mRNA splicing (Alpert et al., 2016; Saldi et al., 2016). Among the curated DHX9 protein-protein interactions (Figure 3.2a), factors functioning in mRNA splicing are among the most abundant. In addition, DHX9 has been implicated in splicing regulation (Bratt and Ohman, 2003; Selvanathan et al., 2015), raising the possibility that the interactions of Nup98 and DHX9 may also play a role in this process. Data sets from RNA-Seq analysis of Nup98 and DHX9 depleted cells reveal a significant overlap in gene products exhibiting altered splicing upon depletion of each protein. DHX9 depletion altered the splicing of 866 genes, of these 217 genes also show altered splicing upon Nup98 depletion ($p = 2.03 \times 10^{-43}$). Based on this information, we examined splicing isoforms of the well-characterized E1A mRNA, which interacted with DHX9 and Nup98 (Figure 3.20a). Different E1A splicing intermediates (13S, 12S, 11S, 10S and 9S) have been characterized (Stephens and Harlow, 1987), the abundance of which could be quantified following depletion of Nup98 or DHX9 (Figure 3.27). Depletion of Nup98 or DHX9 resulted in a 1.9 or 1.8-fold increase in pre-spliced isoform of the transcript (Figure 3.25). Furthermore, we observed differential effects on levels of the various splicing isoforms. Depletion of Nup98 led to significantly increased levels of the 12S, 11S and 10S isoforms. A similar increase in 12S and 11S isoform was detected in cells depleted of DHX9. DHX9 depletion also caused significant decreases in 9S and 13S abundance. These results suggest that the interactions of Nup98 and DHX9 with specific mRNAs, such as E1A, regulates their splicing.

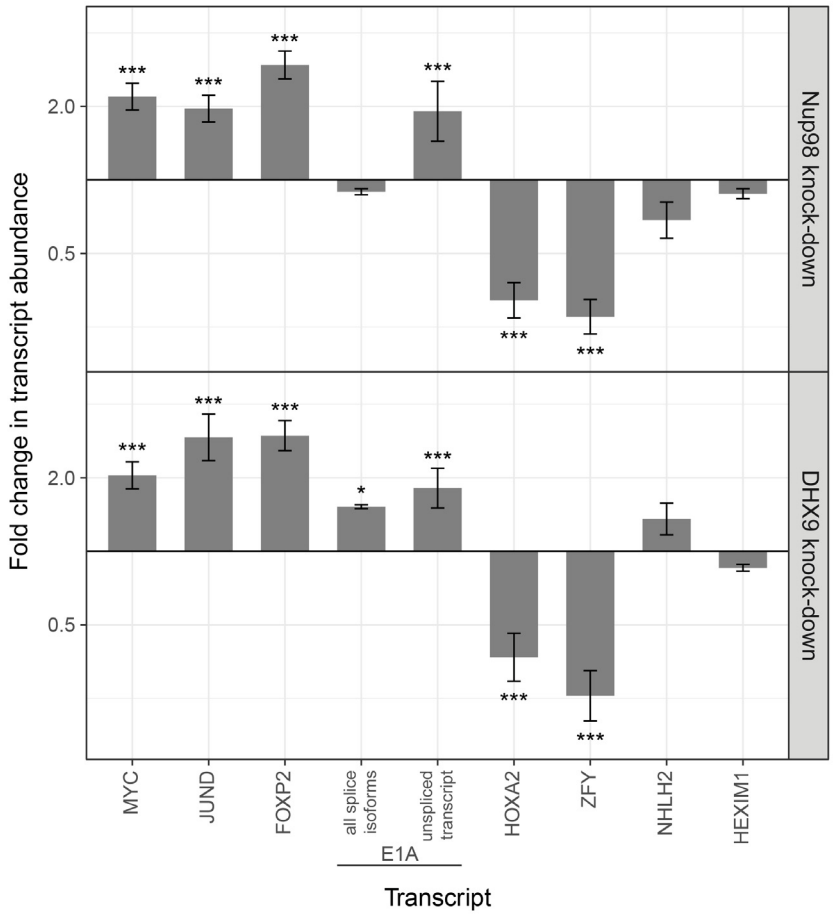
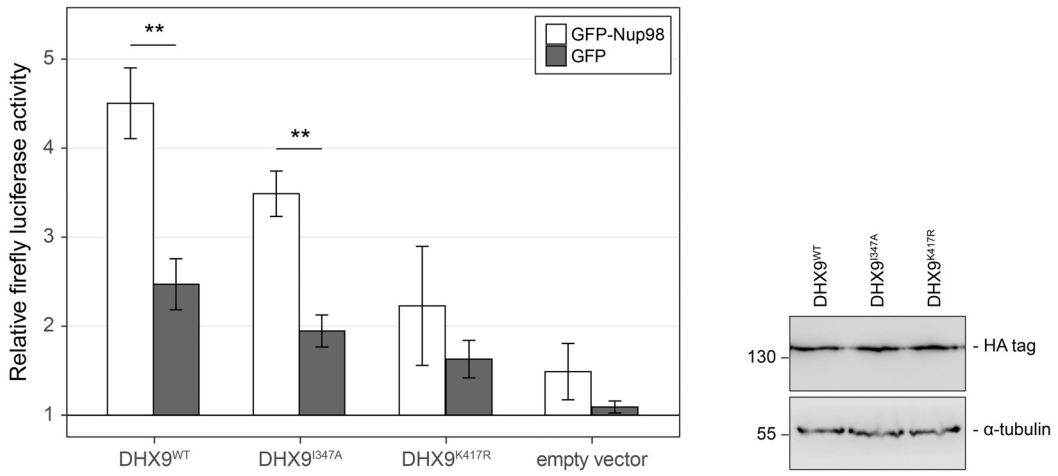


Figure 3.25: Nup98 or DHX9 depletion alters the abundance of target mRNAs

HEK293T cells were transduced with a control shRNA or an shRNA targeting Nup98 or DHX9. RNA was purified and transcript levels from the indicated genes (x-axis) were reverse transcribed and quantified by qPCR. Fold-changes (y-axis) in the average abundance of different transcripts upon Nup98 (top) or DHX9 (bottom) depletion relative to average transcript abundance in mock depleted cells are shown. Error bars indicate standard deviation (N = 3 biological replicates). Results from 3 biological replicates were submitted to ANOVA tests followed by Tukey HSD tests. p-values are indicated as *** < 0.001 and * $<$ 0.05 for Tukey HSD in pairwise comparisons between mRNA levels from depleted and mock depleted cells.



(a) Luciferase assay (cAMP response element) **(b) Western blot**

Figure 3.26: Nup98 stimulates the transcriptional activity of DHX9.

3.26a HEK293T cells transfected with the luciferase gene under control of a cAMP response element (CRE) were co-transfected with two plasmids, one containing GFP-NUP98 or GFP and another containing either $DHX9^{WT}$, the point mutant $DHX9^{I347A}$, the point mutant $DHX9^{K417R}$ or an empty plasmid. Luciferase activity is shown on the y-axis. Luciferase activity from cells transfected with luciferase plasmid alone was designated 1. Each value of relative luciferase activity represents the mean standard deviation (N = 3). (** indicate p-value < 0.01 in T-test comparing normalized luciferase activity in cells transfected with GFP-Nup98 versus GFP). **3.26b** DHX9 point mutant constructs are expressed at levels similar to WT. Western blots of proteins derived from HEK293T cells lysates expressing the indicated DHX9 constructs were performed. These constructs contain a C-terminal HA-tag, allowing detection and comparison of protein levels. The positions of molecular mass markers (shown in kDa) are indicated on the left. α-tubulin was used as loading control. N = 3 biological replicates, representative images of western blots are shown.

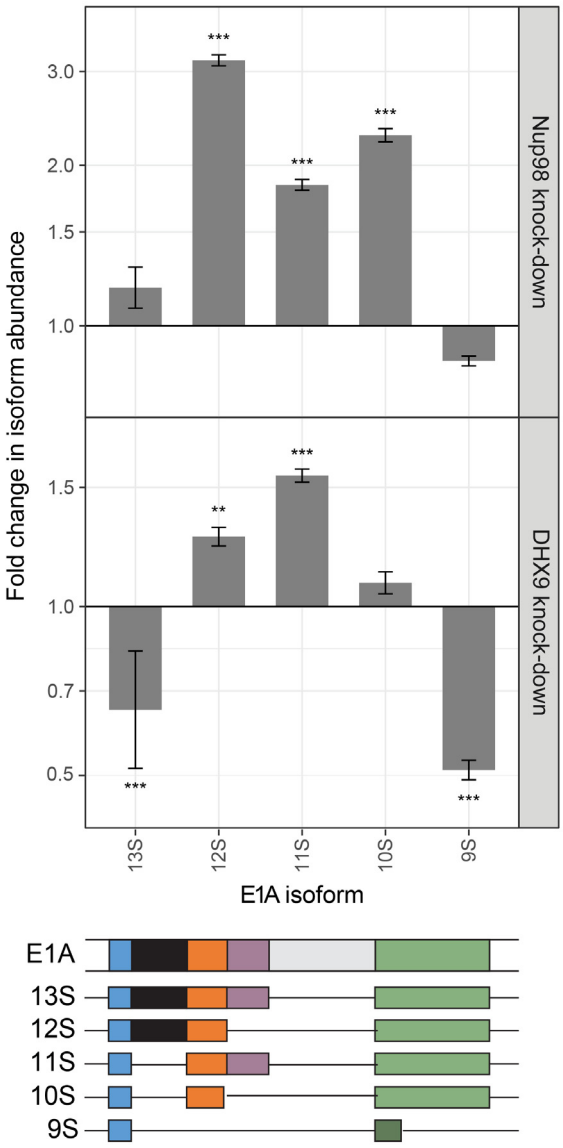


Figure 3.27: Nup98 or DHX9 depletion affects alternative splicing of E1A mRNA.

HEK293T cells were transfected with a control shRNA or an shRNA targeting Nup98 or DHX9. RNA from these cells was reverse transcribed and cDNAs used as template in qPCR reactions containing primers specific to different splice isoforms of the E1A mRNA (see bottom diagram). The average abundance of each E1A splice isoform was normalized to the average total E1A transcript present in the same sample. Fold-change in the normalized abundance of different E1A splice isoforms between the knock-down and control samples are shown in the y-axis. Error bars indicate standard deviation (N = 3 biological replicates). Results from 3 biological replicates were submitted to ANOVA followed by Tukey HSD tests. Adjusted p-values are indicated as *** $<$ 0.001 $<$ ** $<$ 0.01 for Tukey HSD in pairwise comparisons between E1A splice isoforms mRNA amounts from cells depleted of DHX9 or Nup98 and mock-depleted cells.

3.4 Discussion

Several previous publications have established the importance of Nup98 in regulating gene expression (Kalverda et al., 2010; Liang et al., 2013; Light et al., 2013; Pascual-Garcia et al., 2014; Singer et al., 2012). However, our understanding of the mechanistic role of Nup98 in this process has lagged due to our limited knowledge of Nup98 binding partners, most notably in the nucleoplasm, and the consequences of these interactions on the functions of the interacting partners. In this work, we focused on a possible mechanism by which Nup98 can alter gene expression through its interaction with, and regulation of, the RNA helicase DHX9. We have shown using a combination of *in vitro* and *in vivo* assays that Nup98 directly binds DHX9 in the nucleoplasm and this interaction is facilitated by RNA. Importantly, binding of Nup98 to DHX9 can stimulate the ATPase activity of DHX9 and support the role of this DExH/D-box protein in the transcription and splicing of a subset of genes. Consistent with these observations, our analysis revealed that Nup98 and DHX9 bind to similar gene loci and mRNAs, and these interactions are interdependent upon one another. In aggregate, our observations lead us to conclude that intranuclear Nup98 functions as a regulator of DHX9.

3.4.1 The interaction of Nup98 with DHX9

Immunopurified Nup98 revealed associated proteins with known roles in mRNA metabolism (Figure 3.1 and 3.2), mainly hnRNP proteins and RNA helicases (Han et al., 2010), suggesting a functional link between Nup98 and mRNA metabolism. We envisage that many of these proteins are components of Nup98-interacting protein complexes, but most are unlikely to bind directly to Nup98. As others have concluded from MS data (Cox and Mann, 2011; Liu et al., 2004; Mazloom et al., 2011), we speculated that proteins directly interacting with Nup98 were more likely to be among those species most highly represented by unique peptides in our MS analysis, which led us to focus on DHX9. This seemed reasonable as previous studies have described interactions between RNA helicases and nucleoporins, including the interactions of yeast Dbp5 with Nup159 and human DDX19 with Nup214, and the role of these interactions in modulating ATP-dependent helicase activity and mRNA export (Montpetit et al., 2011; Napetschnig et al., 2009; Noble et al., 2011; von Moeller et al., 2009; Weirich et al., 2004).

DHX9, also termed RNA helicase A (RHA), is a member of the helicase superfamily 2. Most of the proteins known to function as RNA chaperones or RNA-protein complex remodelers in this superfamily are found within the DEAD-box (DDX) and

the DEAH/RHA (DHX) families (Jarmoskaite and Russell, 2014). Both DHX and DDX helicases contain a highly conserved helicase domain that mediates nucleotide binding and hydrolysis and is linked to binding of nucleic acid (Jarmoskaite and Russell, 2014; Stevens, 2010). Flanking the helicase domain, DHX family members possess variable N- and C-terminal domains that, while often containing shared sequence motifs, contribute to the diverse functions of family members (Jankowsky and Fairman, 2007; Linder and Jankowsky, 2011). DHX9 contains two double-stranded RNA binding domains (dsRBD) followed by a minimal transactivation domain (MTAD), the site of RNA polymerase II interaction, within the N-terminal third of the protein (Zhang and Grosse, 1997). The C-terminal third of DHX9 contains a oligonucleotide/oligosaccharide (OB)-binding fold and an RGG-box, a domain that characteristically binds single-stranded nucleic acids. Several proteins have been shown to bind the N-terminus, C-terminus, or both regions of DHX9 (Lee and Pelletier, 2016).

The N- and C-terminal regions of DHX9 containing the dsRBDs and the RGG-box are thought to be spatially positioned in close proximity and contribute to the nucleic acid binding properties of DHX9 (Zhang and Grosse, 1997). As mentioned above, these regions also contribute binding surfaces for interacting proteins, and in this study we show that both N- and C-terminal regions of DHX9 bind the N-terminal region of Nup98 (containing FG/GLFG repeats and a RBD) (Figure 3.12b and 3.16a). While the nature of these interactions remains to be examined in greater detail, studies of another DHX member, yeast Prp43, and its binding partner Ntr1 offer possible insights into the nature of the interactions between DHX9 and Nup98. Like DHX9, Prp43 contains an OB-binding fold, which binds to an intrinsically unstructured, N-terminal region of Ntr1 (Christian et al., 2014). Similarly, the C-terminal region of DHX9 contains an OB-binding fold that binds the unstructured N-terminal region of Nup98 (Figure 3.12b and 3.16a). Intriguingly, the unstructured regions of Ntr1 contain a G-patch motif rich in glycines and bulky, hydrophobic residues (Aravind and Koonin, 1999), a compositional property shared with the FG/GLFG repeats present on the N-terminal of Nup98. When Ntr1 binds to Prp43, the conformation imparted on the G-patch motif facilitates its binding to RNA (Christian et al., 2014). We speculate that the binding of Nup98 to DHX9 may also impart structural features on the FG/GLFG repeat regions of Nup98 that facilitates binding to RNA. This idea is consistent with our observation that the binding of certain mRNAs to Nup98 is facilitated by DHX9 (Figure 3.21; see below).

Reciprocally, the binding of Nup98 to N- and C-terminal regions of DHX9 increases the ATPase activity of DHX9 (Figure 3.18a, 3.18b and 3.19). Other factors also function similarly to regulate DHX9. For example, Werner Syndrome helicase interacts with both N- and C-terminal regions of DHX9 to inhibit DHX9 activity

(Friedemann et al., 2005). In another case, the catalytic subunit of DNA-dependent protein kinase (PRKDC) interacts with DHX9 and increases ATPase activity (Mischo et al., 2005). Mechanistically, the binding of Ntr1 to Prp43 again offers a precedent for how these binding events could regulate DHX9 activity (Christian et al., 2014; Tanaka et al., 2007).

3.4.2 Nup98 and DHX9 interact in the nucleoplasm

Previous analyses of the subcellular distribution of DHX9 suggest it is largely restricted to the nucleoplasm (Zhang et al., 1995), but it is excluded from nucleoli and shows no obvious accumulation at NPCs (Figure 3.4) (Lee and Pelletier, 2016; Uhlén et al., 2015; Zhang et al., 1995). Several observations lead us to conclude that the Nup98-DHX9 complex also resides in the nucleoplasm. First, either depletion or overexpression of Nup98 altered the intranuclear distribution of DHX9, including the recruitment of DHX9 to intranuclear Nup98 foci formed upon Nup98 overexpression (Figure 3.8). Second, nuclear fractionation revealed that Nup98-DHX9 complexes are present primarily in the nucleoplasm (Figure 3.11b). Finally, DamID analysis established interdependent binding of Nup98 and DHX9 to specific gene loci (Figure 3.24). Together these results strongly support the existence of an intranuclear Nup98-DHX9 complex. Given that DHX9 appears to be a more abundant protein than Nup98 (Montague et al., 2015; Schaab et al., 2012; Wang et al., 2012; Wilhelm et al., 2014), and that other DHX9 interacting partners appear to bind similar regions of DHX9 (Anderson et al., 1998; Erkizan et al., 2015; Jin et al., 2011; Nakajima et al., 1997; Pellizzoni et al., 2001b; Robb and Rana, 2007; Sadler et al., 2009; Smith et al., 2004; Tetsuka et al., 2004), we assume that Nup98 bind and regulates a subpopulation of DHX9. Overall, the competition between various binding partners (e.g. Nup98, Werner Syndrome helicase, and PRKDC) for DHX9 is likely key to determining the localization of DHX9, the specific mRNAs it binds, and its overall involvement in gene expression (Anderson et al., 1998; Erkizan et al., 2015; Jin et al., 2011; Nakajima et al., 1997; Pellizzoni et al., 2001b; Robb and Rana, 2007; Sadler et al., 2009; Smith et al., 2004; Tetsuka et al., 2004).

3.4.3 The Nup98-DHX9 complex regulates transcription and mRNA processing

Nup98 and DHX9 have each independently been shown to play a role in regulating gene transcription (Capelson et al., 2010; Kalverda et al., 2010; Liang et al., 2013; Lee and Pelletier, 2016). We propose that at least some of the regulatory functions

ascribed to these proteins are performed by the Nup98-DHX9 complex. Several results support this conclusion including our findings that both proteins associate with similar gene loci and their RNA products. For example, our examination of previously published data sets (Chen et al., 2014; Erkizan et al., 2015; Franks et al., 2016; Hendrickson et al., 2016) revealed a strong correlation between gene loci bound to Nup98 and the association of Nup98 and DHX9 with the mRNA products of these genes. Specifically, of the gene loci that interact with nucleoplasmic Nup98, 70% produce transcripts that are also bound to Nup98 ($p = 6.14 \times 10^{-215}$) and 27% produce transcripts bound to both Nup98 and DHX9 ($p = 3.35 \times 10^{-58}$). Furthermore, our analysis of several putative DHX9-interacting gene loci revealed the interdependent binding of Nup98 and DHX9 to these genes and their transcripts (Figure 3.21 and 3.24), as well as a role for these proteins in the expression of these genes (Figure 3.25).

Of note, many of the genes showing altered expression following depletion of Nup98 or DHX9 contain a putative cAMP-response element (CRE). CRE regulated genes represent 50% of the Nup98 interacting gene loci detected in Nup98-Dam-ID studies (Franks et al., 2016), and of these genes 72% have their transcripts bound by Nup98 ($p = 4.2 \times 10^{-205}$) and 36% bound by DHX9 ($p = 2.3 \times 10^{-5}$). Consistent with these observations, both Nup98 and DHX9 have been reported to bind to the CREB-binding protein (CBP)/p300 (Aratani et al., 2001; Kasper et al., 1999; Nakajima et al., 1997), a transcriptional co-activator (Vo and Goodman, 2001).

On the basis of their physical and functional links to CRE regulated genes, we used a CRE-luciferase reporter assay to assess the role of Nup98 in DHX9-mediated transcription. Aratani and colleagues previously used this assay to characterize two modes of DHX9-stimulated reporter expression, one dependent on, and the other independent of, DHX9 ATPase activity (Aratani et al., 2001). Using DHX9 point mutants that either reduce or eliminate its ATPase activity (Aratani et al., 2001), we assessed the ability of Nup98 to stimulate the ATPase activity of DHX9 and modulate its transcriptional activity. Importantly, the expression of Nup98, while itself unable to stimulate reporter expression, could suppress the transcriptional defects of the DHX9 point mutant with reduced ATPase activity ($DHX9^{I347A}$), but not an ATPase dead mutant (Figure 3.26a). These results are consistent with our *in vitro* analysis showing Nup98 can stimulate the ATPase activity of DHX9 and supports the hypothesis that Nup98 functions as a cofactor to regulate the ATPase-dependent transcriptional functions of DHX9.

Nup98, by virtue of its ability to bind DHX9 and increase its RNA-dependent ATPase activity, is predicted to stimulate the cellular activities of DHX9 including facilitating efficient processing and release of mRNAs (Jarmoskaite and Russell, 2014; Jankowsky, 2011). This idea is supported by our data showing that depletion of

Nup98 causes an increase in the binding of RNA to DHX9 (Figure 3.21). Given the RNA binding properties of the Nup98-DHX9 complex, and that transcription and mRNA splicing are often coupled, we envisage that defects associated with disruption of this complex would alter splicing. This proved to be the case as the analysis of cells depleted of Nup98 or DHX9 revealed shared splicing defects at the level of E1A reporter (Figure 3.27) and throughout the transcriptome. Our observations are the first to suggest a role for Nup98 in mRNA splicing, and are consistent with a previously proposed role for DHX9 in splicing regulation in mammals (Hartmuth et al., 2002) and Drosophila (Pellizzoni et al., 2001b; Reenan et al., 2000).

3.4.4 Nup98 and RNA helicases beyond DHX9

Finally, we must make note that the role for Nup98 in regulating DHX9 may extend to other DExH/D-box proteins, since five other RNA helicases were identified as Nup98 interactors in our study, including DDX21, DDX5, DDX17, DHX15, and DDX3. DDX5 and DDX17 are highly similar proteins that act as corepressors and coactivators through their interaction and modulation of transcription factors (Fuller-Pace, 2013b). Like Nup98 and DHX9, DDX21, DDX3, and DHX15 have also been implicated in the antiviral immune response by sensing viral dsRNA and contributing to the regulation of expression of interferon and interferon-stimulated genes (Fullam and Schröder, 2013; Lu et al., 2014; Wang et al., 2015). Consequently, it will be important to determine whether Nup98 regulates the localization and activity of these other RNA helicases, which may reflect a more general role for Nup98 in the regulation of RNA processing through an association with multiple members of the DExH/D-box family of proteins.

CHAPTER 4

A role for the nuclear environment in the life cycle of viruses from the *Flaviviridae* family

4.1 Summary

The Hepatitis C virus (HCV) is a positive-strand RNA virus from the *Flaviviridae* family. HCV replication and assembly occurs in cytoplasmic virus-induced rearrangements of host cell membranes termed the membranous web (MW). Nups and nuclear transport factors (NTFs) are also present in the MW of cells infected with HCV and other positive-strand RNA viruses. These Nups and NTFs regulate access of proteins and vRNAs into the interior of the MW, contributing to the establishment of an environment conducive to viral infection (Neufeldt et al., 2013, 2016). These same Nups and NTFs regulate nucleocytoplasmic transport in the cell. Thus, these RNA viruses can also exploit nucleocytoplasmic trafficking to gain access to the nuclei of infected cells. Many recent studies show that, despite their cytoplasmic replication, viral proteins from positive-strand RNA viruses enter the nuclei during infection (Levin et al., 2014b,a; Bonamassa et al., 2015; Flather and Semler, 2015; Lopez-Denman and Mackenzie, 2017). Due to the limited coding capacity of their genome, HCV and other viruses face challenges in supporting the functions required for their propagation. Therefore, many steps in vRNA metabolism require host RNA-binding proteins, several of which are nuclear and normally function in host gene expression (Lloyd, 2015). Here, we show that nuclear host RNA-binding proteins (Nup98, DHX9, hnRNP U, SSB and ELAVL1) can interact with the HCV vRNA and participate in different steps of the HCV life cycle, despite maintaining their nuclear localization during infection. This observation led to the characterization of a dynamic nuclear pool of HCV vRNAs in infected cells. These nuclear HCV vRNAs are imported into the nuclei via Kapβ3, and to a lesser extent, Kapβ1. HCV vRNA retention in the nuclei depends on the RNA helicase DHX9. HCV vRNA nuclear export is CRM1-dependent and it requires Nup98 and hnRNP U. Finally, we show that the Zika virus also has a dynamically transported nuclear pool of vRNAs, pointing towards a role for the nuclear environment in the life cycle of other viruses from the *Flaviviridae* family.

4.2 Introduction

The nuclei is an organelle characteristic of eukaryotic cells. It is delimited by the nuclear envelope and perforated by several nuclear pore complexes. NPCs are formed by approximately 30 different Nups arranged in concentric rings with octagonal symmetry. The center of the pore is lined with Nups containing FG repeats, creating a selective barrier to transport across the channel. Small cargoes can cross through NPCs via passive diffusion. However larger molecules must be actively transported by NTFs (Knockenhauer and Schwartz, 2016).

Active transport through the NPC relies on nuclear transport factors, also known as karyopherins, that can interact with the FG repeats lining the pore. These NTFs interact with NLS or NES sequences in cargo proteins, mediating their transport across the NPC. The Kapβ family of proteins has at least 20 members in metazoans. In vertebrates, Kapβ1 is the predominant transport factor for cargo import into the nucleus. Kapβ1 can also interact with alpha karyopherins (Kapα1 to Kapα6). Kapα can recognize NLS containing cargoes, but it requires Kapβ1 to mediate its interaction with the FG repeats lining the NPC. (Matsuura, 2016; Soniat and Chook, 2015). For nuclear export, the majority of export cargoes associates with the exportin CRM1. At least eight exportins have been identified in metazoans so far (Wente and Rout, 2010).

During nucleocytoplasmic transport the small GTPase Ran regulates binding between NTFs and their cargoes. In the nucleus, due to the presence of a RanGEF, RanGTP is dominant. In the cytoplasm, a RanGAP ensures RanGDP is prevalent. The export of NES-containing cargoes from the nucleus depends on their interaction with exportins and RanGTP. This complex moves through the pore to the cytoplasm where it dissociates, releasing the cargo and RanGTP. In the cytoplasm, RanGTP gets hydrolyzed into RanGDP by RanGAP. During nuclear import, importins bind their NLS-containing cargoes in the cytoplasm and this complex moves through the NPC. Inside the nucleus, interaction with RanGTP causes importins to dissociate from their cargoes. The importin-RanGTP complex is recycled to the cytoplasm, where RanBP separates RanGTP from the importin, allowing RanGAP to hydrolyze RanGTP to RanGDP. RanGDP is returned to the nucleoplasm by the nuclear transport factor NUTF2. In the nucleoplasm, RanGDP interacts with RanGEF for exchange into RanGTP, restarting the transport cycle (Wente and Rout, 2010).

Some cargoes, such as large mRNA molecules, do not require a Ran gradient for their export. NXF1 and NXT1 are the major transport factors for mRNAs. These proteins interact with mRNA transcripts in the nucleus and translocate through the pore by interacting with the FG repeats in the channel. Directionality of transport is provided by an ATP-dependent DEAD box helicase (termed DDX19 in humans) on

the cytoplasmic face of the NPC. DDX19 interacts with Nup214 at the cytoplasmic face of the pore and it remodels the NXF1-NXT1-mRNA complex releasing the transport factors from the transcript (Montpetit et al., 2011; Napetschnig et al., 2009). Other types of RNA, such as rRNA, snRNA, and a subset of mRNAs are also dependent on CRM1 and the Ran gradient for their nuclear export. In addition to the described nuclear export factors, the export of RNA also requires several adaptor proteins, increasing pathway complexity (Sloan et al., 2015).

The nucleocytoplasmic transport pathways of eukaryotic cells are also hijacked during several viral infections. Viruses such as Adenovirus, HSV-1, Influenza A, HIV-1, and HBV can interact with the NPC for uncoating and import of their genomes into the nucleus. A number of viruses also manipulate the nuclear transport of host cargoes to alter the cellular environment favorably for viral proliferation. The blocking of host mRNA export to limit host cell protein production and immune responses is an example of such viral strategy (Yarbrough et al., 2014).

The NPC, Nups and NTFs can also play a role in the life cycle of viruses whose genome replication occurs in the cytoplasm, such as positive-strand RNA viruses. Cells infected with HCV relocate several Nups and NTFs into viral sites of replication and assembly (membranous web). HCV infection also causes a slight increase in the levels of some Nups, and their depletion decreases production of infectious virus. Depletion of different Nups affects different stages of the HCV life cycle. Nup98 and Nup153, for example, participate in viral genome replication, while Nup155 plays a role in virion assembly (Neufeldt et al., 2013). Various HCV proteins contain NLS and NES sequences that mediate their interaction with NTFs such as Kapα1, Kapβ3, and CRM1. Interactions between HCV proteins and Kapα1 play a role during the replication phase of HCV life cycle, while interactions between HCV proteins and Kapβ3 have dual functions during replication and assembly of HCV virions (Levin et al., 2014b).

Despite the cytoplasmic replication of HCV, it is possible that interactions between viral proteins and host nuclear transport pathway components are not restricted to the cytoplasm of infected cells. Several viral proteins containing NLS and NES sequences (Levin et al., 2014b) can be seen in the nuclei of HCV-infected cells (core, NS3 and NS5A) (Bonamassa et al., 2015). Nuclear core and NS5A have been proposed to modulate host gene expression to promote cell survival, immune evasion and increased ribosome biogenesis (Bonamassa et al., 2015). So far, no known nuclear function of NS3 has been described.

Moreover, several nuclear host proteins are hijacked by positive-strand RNA viruses to participate in their replication. Most of these nuclear factors are RNA-binding proteins that interact with the viral RNA and aid translation, replication and virion assembly (Lloyd, 2015). Some of these nuclear factors show relocation

to the cytoplasm of infected cells, but others remain nuclear (Lloyd, 2015). Since core, NS3, and NS5A interact with the HCV +vRNA, it is possible these proteins may carry the vRNA when they enter the nuclei of infected cells, allowing vRNA interaction with host nuclear factors. In fact, viral RNA molecules can be seen in the nuclei of HCV-infected cells in previously published images and movies, although the presence of this nuclear vRNA pool was not described or further investigated in these publications (Fiches et al., 2016; Shulla and Randall, 2015). Despite the possible presence of HCV vRNA in the nuclei of infected cells, this phenomena has never been characterized or discussed in the literature, and the possibility and implications of an intranuclear pool of positive-strand viral RNAs have never been addressed.

In this chapter, I begin the characterization of a nuclear pool of viral RNA in cells infected with HCV, exploring the interaction of this vRNA with host NFs that participate in the viral life cycle. I also investigate the role NTFs play in the nucleocytoplasmic shuttling of vRNAs, showing that nucleocytoplasmic transport disruption can affect vRNA interaction with host NFs and alter the HCV life cycle. Finally, we show that another virus, the Zika virus, also has a dynamic pool of nuclear vRNA, hinting towards a possible role for the nuclear environment in the life cycle of other *Flaviviridae* family viruses.

4.3 Results

4.3.1 Nuclear host proteins can interact with the HCV vRNA

Several host nuclear RNA-binding proteins participate in the HCV life cycle. These proteins have roles in virion assembly, vRNA translation, replication, or the switch between these processes (Lloyd, 2015). Nup98, DHX9, and hnRNP U are among these candidate host NFs that participate in HCV vRNA metabolism.

Previous publications have identified Nup98 as an important host factor in HCV infection. Upon HCV infection, Nup98 partially relocates to the membranous web in the cytoplasm of infected cells (Neufeldt et al., 2013). Nup98 depletion causes a two-fold reduction in viral RNA replication and it decreases the production of infectious virus in HCV-infected cells by 80 to 90%. Nup98 has also been shown to interact with the HCV protein core and it has been detected associated with HCV virions (Neufeldt et al., 2013; Lussignol et al., 2016).

DHX9 in an RNA helicase that interacts with, and is regulated by, Nup98 (see chapter 3). DHX9 is also another nuclear host factor previously shown to play a role in HCV infection. DHX9 can interact with the HCV vRNA, its depletion reduces HCV vRNA replication more than two-fold, and it decreases the production of viral proteins by almost 90% (Isken et al., 2007; Li et al., 2014; He et al., 2008).

Another Nup98 binding partner described in chapter 3 is hnRNP U. This protein contains an N-terminal domain that binds to DNA, and a C-terminal RGG domain capable of RNA-binding. In mammalian cells, hnRNP U participates in high order organization of chromatin in the nucleus, gene expression regulation (transcriptional initiation and elongation), and various aspects of RNA metabolism (RNA stability, splicing, and silencing) (Xiao et al., 2012). HnRNP U, in a complex with DHX9, has also been shown to interact with the HCV protein NS3 and with the HCV vRNA in infected cells (Chatel and Fahrenkrog, 2012; Lu et al., 2004; Upadhyay et al., 2013).

We can draw a parallel between the role Nup98, DHX9, and hnRNP U might play in regulating the metabolism of HCV vRNA and our previous results showing the relevance of these proteins in the regulation of host mRNAs (chapter 3). Therefore, we next decided to further explore the role these nuclear proteins may have in HCV infection, especially their interplay with the HCV vRNA.

We began by determining if Nup98, DHX9, and hnRNP U interact with the HCV vRNA in Huh7.5 cells infected with HCV genotype 2a strain JFH-1. We used stringent buffer conditions to immunoprecipitate Nup98, DHX9, and hnRNP U from crosslinked lysates of HCV-infected cells (three days post infection). RT-PCR was used to test whether HCV vRNA was associated with the immunopurified proteins (Figure 4.1). We confirmed that DHX9 and hnRNP U interact with the HCV vRNA,

as previously described (Isken et al., 2007; Li et al., 2014; Chatel and Fahrenkrog, 2012; Lu et al., 2004; Upadhyay et al., 2013). We also detected a novel interaction between Nup98 and HCV vRNA. As expected, no interaction is observed between HCV vRNA and a negative control IgG (α-GFP). As previously described (Beran et al., 2007), a strong interaction is observed between the viral RNA helicase NS3 and HCV vRNA (Figure 4.1).

4.3.2 Nuclear host proteins participate in the HCV life cycle

Once we determined that the host NFs Nup98, DHX9, and hnRNP U interact with the HCV vRNA, we next examined the importance of these proteins during HCV infection. Lentivirus expressing shRNAs were used to deplete Nup98, DHX9, or hnRNP U, and an shRNA targeting a non-mammalian sequence was used as control. Three days after the lentivirus transduction, we saw $>$ 70% decrease in mRNA levels and $>$ 60% protein depletion for the targeted host factors (Figure 4.2a), with negligible effect on cell viability (Figure 4.2c).

Cells were coinfected with lentivirus and HCV to assess the effect of Nup98, DHX9, or hnRNP U depletion on HCV replication. Three days post coinfection (p.i.), changes in intracellular HCV vRNA levels upon NF depletion were determined using qPCR. The intracellular levels of vRNA were significantly decreased upon depletion of Nup98 or DHX9, while reduced levels of hnRNP U significantly increased intracellular vRNA when compared to control shRNA in HCV-infected cells (Figure 4.3a). These data indicate that Nup98 and DHX9 facilitate proper replication of the HCV vRNA, as previously demonstrated (Neufeldt et al., 2013; Li et al., 2014; He et al., 2008; Isken et al., 2007). HnRNP U, on the other hand, has been characterized as an important antiviral protein and a hotspot of viral perturbation strategies (Pichlmair et al., 2012). Therefore, the role of hnRNP U in immune response might restrict HCV replication in infected cells.

Proper translation of the HCV vRNA in HCV-infected and Nup98-, DHX9- or hnRNP U-depleted cells was verified by detection of the HCV core protein by western blot. In accordance with decreased intracellular vRNA levels, cells depleted of Nup98 or DHX9 also show similar decreases in HCV core abundance (Figure 4.3b). HnRNP U-depleted cells show an even greater reduction in HCV core protein levels, despite the increased intracellular vRNA amounts (Figure 4.3b). This possibly indicates that, while hnRNP U is not required for replication of the HCV vRNA, it is necessary for HCV vRNA translation and production of viral proteins.

We also evaluated the consequence of Nup98, DHX9 or hnRNP U depletion on the production of infectious HCV. Cells were coinfected with HCV and lentivirus

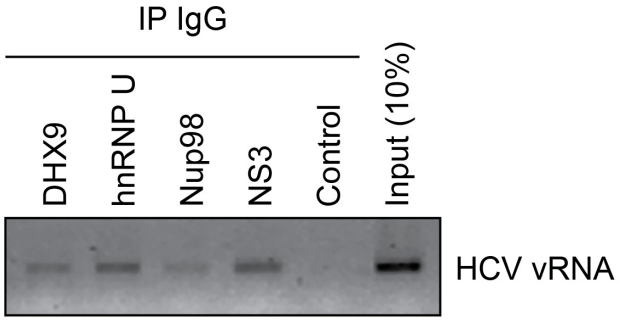
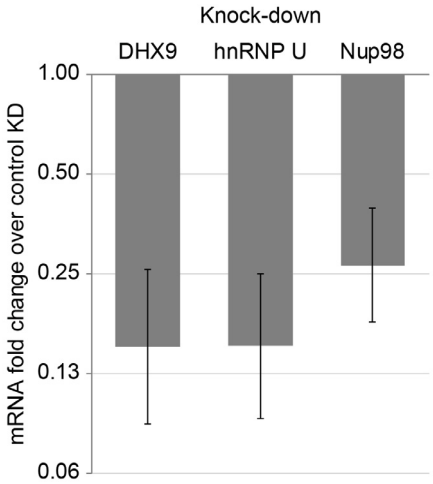
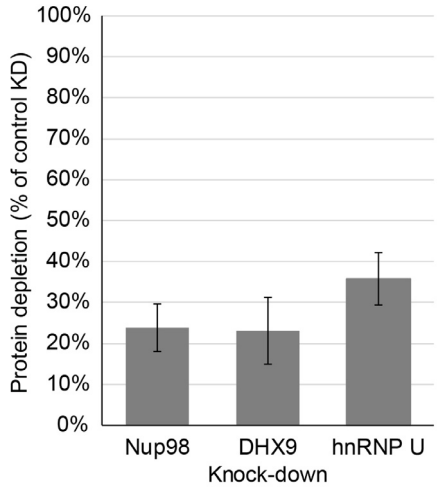


Figure 4.1: Nup98, DHX9, and hnRNP U interact with the HCV vRNA.

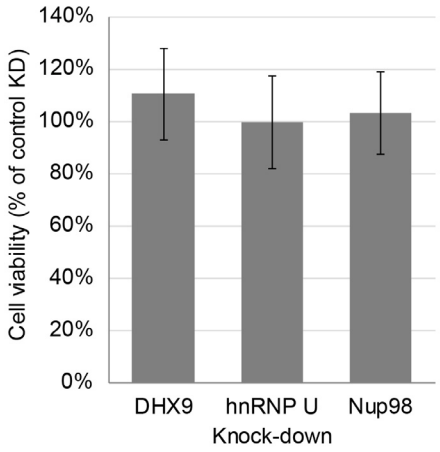
Crosslinked cell lysates from Huh7.5 cells infected with HCV (72h p.i.) were incubated with beads coupled to a negative control IgG (α-GFP), or coupled to antibodies against Nup98, DHX9, hnRNP U, or NS3 (positive control IgG). RNA present in immunoprecipitated protein-RNA complexes and total cellular RNA (10% input) were used as template in RT-PCR reactions containing primers specific to the HCV +vRNA. N = 3 biological replicates, representative agarose gel images are shown. This data was produced in collaboration with Dr. Zhongjing Su.



(a) mRNA depletion (qPCR)



(b) protein depletion (WB)



(c) cell viability

Figure 4.2: Depletion of host NFs has no effect on cell viability.

Huh7.5 cells were coinfected with HCV and lentiviruses encoding shRNAs directed against Nup98, DHX9, hnRNP U, or a non-mammalian control sequence for three days. **4.2a** Average transcript depletion was evaluated by qPCR with primers for Nup98, DHX9, or hnRNP U. Samples were normalized to house-keeping transcripts and fold-change relative to cells expressing the control shRNA were calculated. **4.2b** Protein depletion following shRNA expression was evaluated by western blot with antibodies against Nup98, DHX9, or hnRNP U. Samples were normalized to a load control (α-tubulin) and the average percentage of protein depletion relative to cells expressing the control shRNA was calculated. **4.2c** The viability of HCV-infected cells expressing different shRNAs was evaluated by mitochondrial dehydrogenases activity. Average viability of cells depleted of the indicated proteins is expressed as a percentage of HCV-infected cells expressing the control shRNA. Error bars indicate standard deviation. N $\geq$ 3 biological replicates.

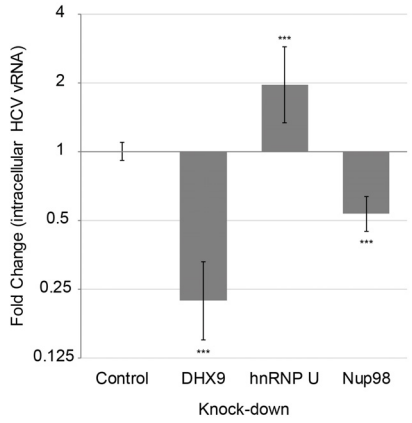
for approximately 16 hours, followed by replacement of infection supernatant with fresh cell culture media. The infectious titer of newly produced HCV released into the cell culture media was determined three days post infection. Depletion of Nup98, DHX9, or hnRNP U caused a decreased of ≈ 70-80% in the titer of infectious HCV produced by infected cells (Figure 4.3c). This decrease is consistent with our previous results showing deceased viral protein levels under these conditions.

The results described above indicate that depletion of Nup98, DHX9, or hnRNP U have deleterious effects in HCV infection, corroborating their role in the HCV life cycle. DHX9 depletion in HCV-infected cells caused a decrease of similar magnitude (≈ 4-fold) in intracellular vRNA, viral protein, and infectious virus titer, indicating that the role of DHX9 in the HCV life cycle is likely linked to replication of the viral RNA. Nup98 depletion in HCV infected cells caused a similar decrease (≈ 2-fold) in intracellular vRNA and viral protein abundance, confirming its previously described role in HCV vRNA replication (Neufeldt et al., 2013). However, Nup98 depletion in HCV-infected cells caused an even more pronounced (≈ 3-fold) decrease in the production of infectious virus, corroborating a further requirement for Nup98 at a post-replication stage of virus assembly or egress (Neufeldt et al., 2013; Lussignol et al., 2016). Depletion of hnRNP U in HCV-infected cells did not impair HCV replication, but resulted in a drastic decrease in viral protein production (nearly 90% reduction). The production of infectious virus was also similarly decreased (≈ 80% reduction), indicating that hnRNP U likely participates at a post-replication stage affecting viral protein production.

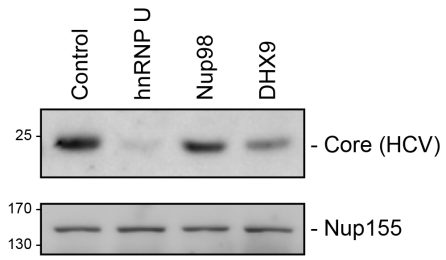
4.3.3 Host NFs with a role in HCV infection do not relocate to the cytoplasm of infected cells

HCV is a positive-strand RNA virus with a cytoplasmic life cycle. In line with this idea, Nup98 has been shown to partially relocalize to the cytoplasm of HCV-infected cells (Neufeldt et al., 2013), as confirmed here (Figure 4.4b). However, this is not the case for several other host NFs that play a role in HCV life cycle. Several nuclear RNA-binding proteins, for instance, have been shown to interact with the HCV vRNA and participate in HCV infection. But not all of them show relocation to the cytoplasm in HCV-infected cells (Lloyd, 2015).

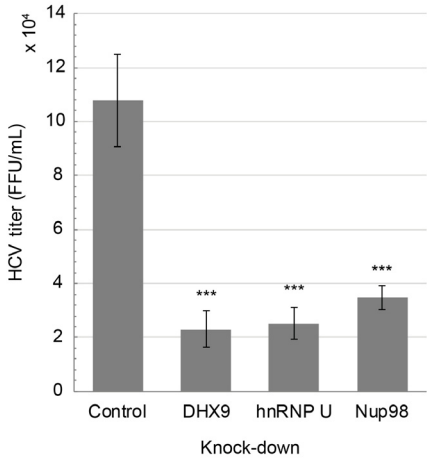
DHX9 and hnRNP U are among such RNA-binding proteins that reside in the nucleoplasm (Dreyfuss et al., 1984; Uhlén et al., 2015; Weidensdorfer et al., 2009; Zhang and Grosse, 1991), interact with the HCV vRNA (Figure 4.1), and participate in the HCV life cycle (Figure 4.3). However, previous reports indicate a lack of cytoplasmic DHX9 and hnRNP U in HCV-infected cells (Chatel-Chaix et al., 2011). We used immunofluorescence microscopy to determine the cellular localization of



(a) intracellular HCV vRNA



(b) HCV core protein



(c) HCV infectious titer

Figure 4.3: Nup98, DHX9, or hnRNP U depletion alters the HCV life cycle.

Huh7.5 cells were coinfected with HCV and lentiviruses encoding shRNAs directed against Nup98, DHX9, hnRNP U, or a control sequence for three days. **4.3a** Average changes in intracellular HCV vRNA were evaluated by qPCR. Samples were normalized to house-keeping transcripts and fold-change relative to cells expressing the control shRNA was calculated. **4.3b** HCV core levels were determined by western blot. Nup155 was used as loading control that is unaffected by HCV infection (Neufeldt et al., 2013). Positions of molecular mass markers (in kDa) are indicated on the left. **4.3c** The infectious titer of HCV produced (focus-forming units per mL of medium) was determined by indirect immunofluorescence microscopy. For all panels, N $\geq$ 3 biological replicates. Plot error bars indicate standard deviation for 3 biological replicates. Results from $\geq$ 3 biological replicates were submitted to ANOVA followed by Tukey HSD tests. Adjusted p-values $<$ 0.001 for Tukey HSD in pairwise comparisons between protein depleted sample and control are indicated as ***. Part of these data were produced in collaboration with Dr. Zhongjing Su.

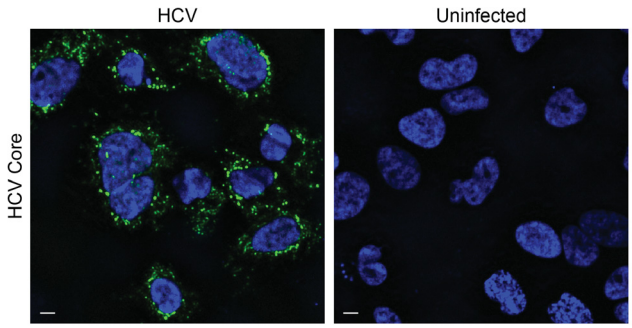
hnRNP U and DHX9 in HCV infected cells. Huh7.5 cells infected with HCV (MOI 1) for three days show nearly 100% of cells positive for HCV antigen (Core protein), as seen in Figure 4.4a (Yi, 2010; Scheel et al., 2008). The same pool of HCV-infected cells shows no change in DHX9 or hnRNP U localization three days post infection (Figure 4.4b).

Given the lack of a cytoplasmic pool of DHX9 and hnRNP U in HCV infected cells, we next decided to assess the cellular localization of two other nuclear RNA-binding proteins with well characterized roles in the HCV life cycle; ELAVL1 (also termed HuR) and SSB (also known as autoantigen La). ELAVL1 interacts with the vRNA, regulates its stability, promotes genome circularization, and promotes vRNA replication and translation. ELAVL1 depletion decreases HCV replication and viral protein abundance by half in HCV infected cells. ELAVL1 can interact with NS3 and NS5B (viral proteins that participate in vRNA replication) and it facilitates the interaction of SSB with the HCV vRNA (Korf et al., 2005; Rivas-Aravena et al., 2009; Shwetha et al., 2015; Spångberg et al., 2000). SSB can interact with the HCV +vRNA IRES and alter its conformation to promote genome circularization for translation or translation-replication switching (Fontanes et al., 2009; Martino et al., 2012; Shirasaki et al., 2010). Similarly to what was observed for DHX9 and hnRNP U, HCV infected cells show no change in SSB localization three days post infection. ELAVL1 also remains mainly nuclear upon HCV infection. However, small cytoplasmic puncta containing this protein are occasionally visible in HCV-infected cells (Figure 4.4b).

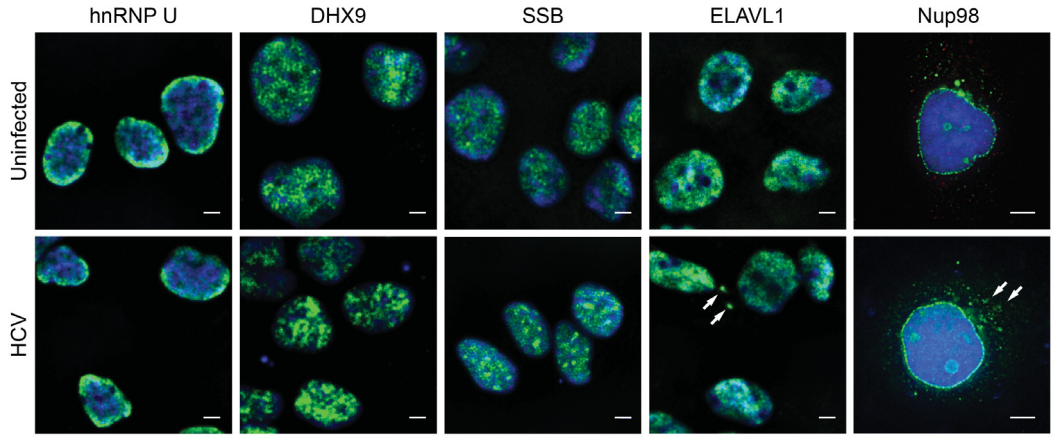
One intriguing possibility is that, despite being a virus with a cytoplasmic life cycle, the HCV vRNA might enter the nuclei of infected cells to interact with nuclear host RNA-binding proteins that can facilitate the different steps in its life cycle, such as DHX9, hnRNP U, ELAVL1, SSB, and possibly Nup98.

4.3.4 HCV vRNA can be detected in the nuclei of infected cells

To determine the localization of the HCV vRNA in infected cells we used fluorescence *in situ* hybridization (FISH) to visualize RNA molecules in the cell. Three sets of FISH probes were used. The first set contains control probes that hybridize with the GAPDH mRNA. The two other sets contain probes that hybridize to the 5' of the HCV +vRNA (sequence encoding from Core to NS2) and probes that bind the 3' of the HCV +vRNA (sequence encoding NS3 to the 3' UTR), identified as 5' probes and 3' probes, respectively. Only fluorescent foci showing hybridization to 5' and 3' probes were used in the analysis. This selection ensured only full length +vRNAs were evaluated, excluding any degradation products or non-specific staining.



(a) Widespread HCV infection in Huh7.5 cells



(b) Host NFs localization

Figure 4.4: Localization of host NFs in HCV-infected cells.

Huh7.5 cells were uninfected or infected with HCV for three days. **4.4a** Monitoring of the virus spread (HCV-positive cells) was evaluated by indirect immunofluorescence microscopy using antibodies specific for HCV core (green). **4.4b** Localization of hnRNP U, DHX9, SSB, ELAVL1 or Nup98 in uninfected or HCV-infected (HCV) cells was evaluated by indirect immunofluorescence microscopy using antibodies specific for the indicated proteins (green). Cytoplasmic relocalization of host NFs is indicated by white arrows. DNA was stained with DAPI (blue). Scale bar, 5 μm. For all panels, N = 3 biological replicates. Representative immunofluorescence microscopy images are shown. This experiment was performed in collaboration with Dr. Zhongjing Su.

Using the protocol above, we can visualize HCV +vRNA molecules present inside the nuclei of infected cells. Uninfected cells show no fluorescence from non-specific binding of the +vRNA probes. We can also observe high colocalization between the HCV +vRNA 5' and 3' probes, while the GAPDH mRNA probes show no significant colocalization with the +vRNA 3' probes. Furthermore, HCV infection shows no effect in the localization of the GAPDH mRNA (Figure 4.5).

Stable Huh7 cell lines containing autonomously replicating subgenomic HCV JFH-1 +vRNAs (missing the genomic regions encoding from Core to NS2) can facilitate the study of viral protein expression and HCV replication, in the absence of other stages of the viral life cycle (i.e., virus entry into the cell, +vRNA uncoating, virion assembly, virus budding, and release). This system is known as HCV replicon and it is extensively used to study specific steps of the HCV life cycle (Joyce and Tyrrell, 2010; Scheel and Rice, 2013; Kato et al., 2003). In HCV replicon cells, the FISH probes mapping to the sequence encoding from Core to NS2 (5' probes) show no fluorescence, while the probes mapping to the sequence encoding NS3 to the 3'UTR (3' probes) are visible, confirming the low non-specific binding of these probes. Interestingly, this HCV replicon system also contains an intranuclear pool of HCV +vRNA (Figure 4.6). The presence of the HCV replicon in Huh7 cells had no effect on the localization of GAPDH mRNA (Figure 4.6).

Since these FISH images have a high number of RNA foci present in each cell, we used automated foci counting and localization in relation to the nuclei for analysis. We used previously published methods for foci identification (Wu and Rifkin, 2015) and nuclei modeling in microscopy images (Zhao et al., 2016) to create a custom analysis workflow (see appendices C.1.1 and C.2.1). This analysis allowed the representation of each image in a three dimensional plane, containing isosurfaces representing the nuclei of cells (DAPI stain bound to DNA used for mapping) and points representing RNA foci (FISH probes used for mapping). For +vRNA foci, only points mapped for 5' and 3' probes within 250 nm of one another were considered. To further investigate the localization of the created DAPI isosurfaces in relation to NPCs in these cells, we used immunofluorescence to detect Nup98 (with specific antibodies) and DNA bound DAPI stain in Huh7.5 cells. In a histogram of the distribution of distances between detected Nup98 foci and the closest DAPI isosurface, we see a peak of most Nup98 foci localizing 0.5 μm outside the DAPI isosurface (on its cytoplasmic side). This distribution likely indicates that the DAPI isosurface is created approximately 0.5 μm on the inside of the NE boundary, since most Nup98 foci were detected at the nuclear rim (NPC-associated Nup98) (Figure 4.7). Of note, Nup98 is also present within the nucleoplasm, and to a lesser extent the cytoplasm, of Huh7.5 cells as visible in the tails of the histogram from Figure 4.7.

Mapping RNA FISH images onto a 3D plane, we can obtain quantitative infor-

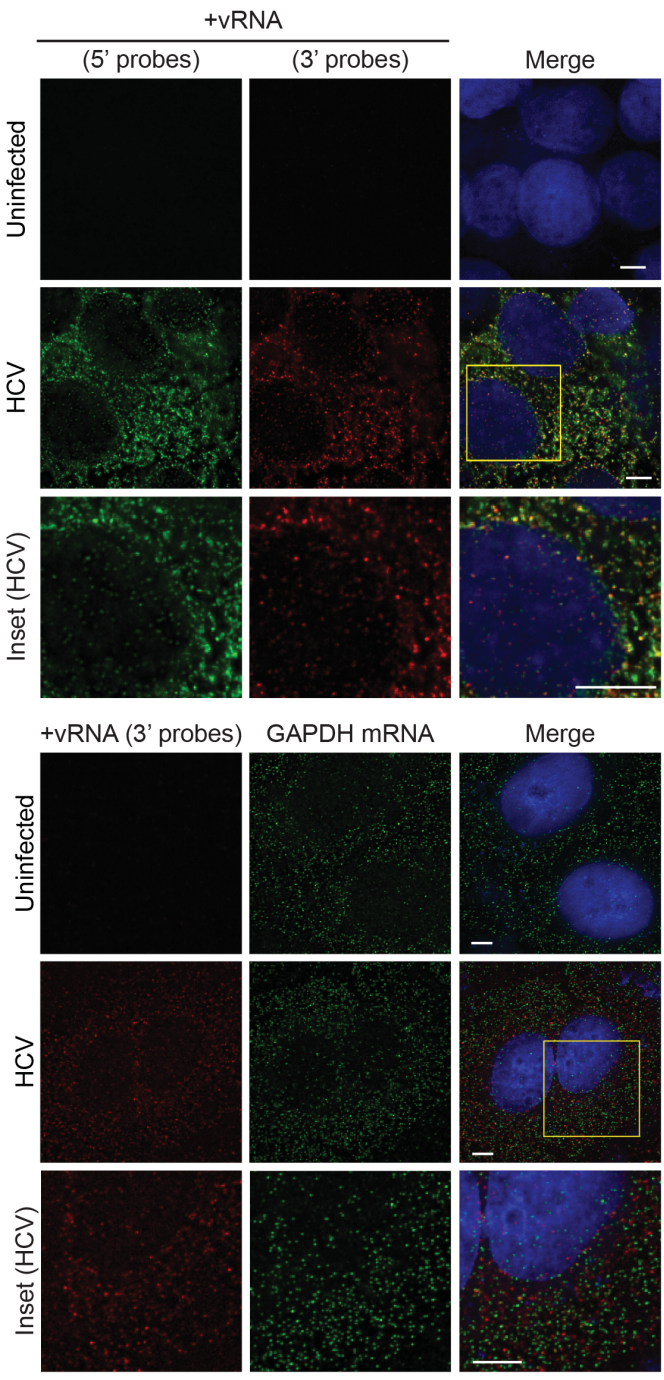


Figure 4.5: HCV-infected cells have an intranuclear pool of +vRNA.

Huh7.5 cells were uninfected or infected with HCV for four days. Localization of +vRNA (3' and 5' probes) or GAPDH mRNA in uninfected or HCV-infected (HCV) cells was evaluated by FISH visualized by epifluorescence microscopy. Insets showing higher magnification of regions marked by a yellow square are shown below each panel. DNA was stained with DAPI (blue). Scale bar, 5 μm. N = 3 biological replicates, representative FISH images are shown.

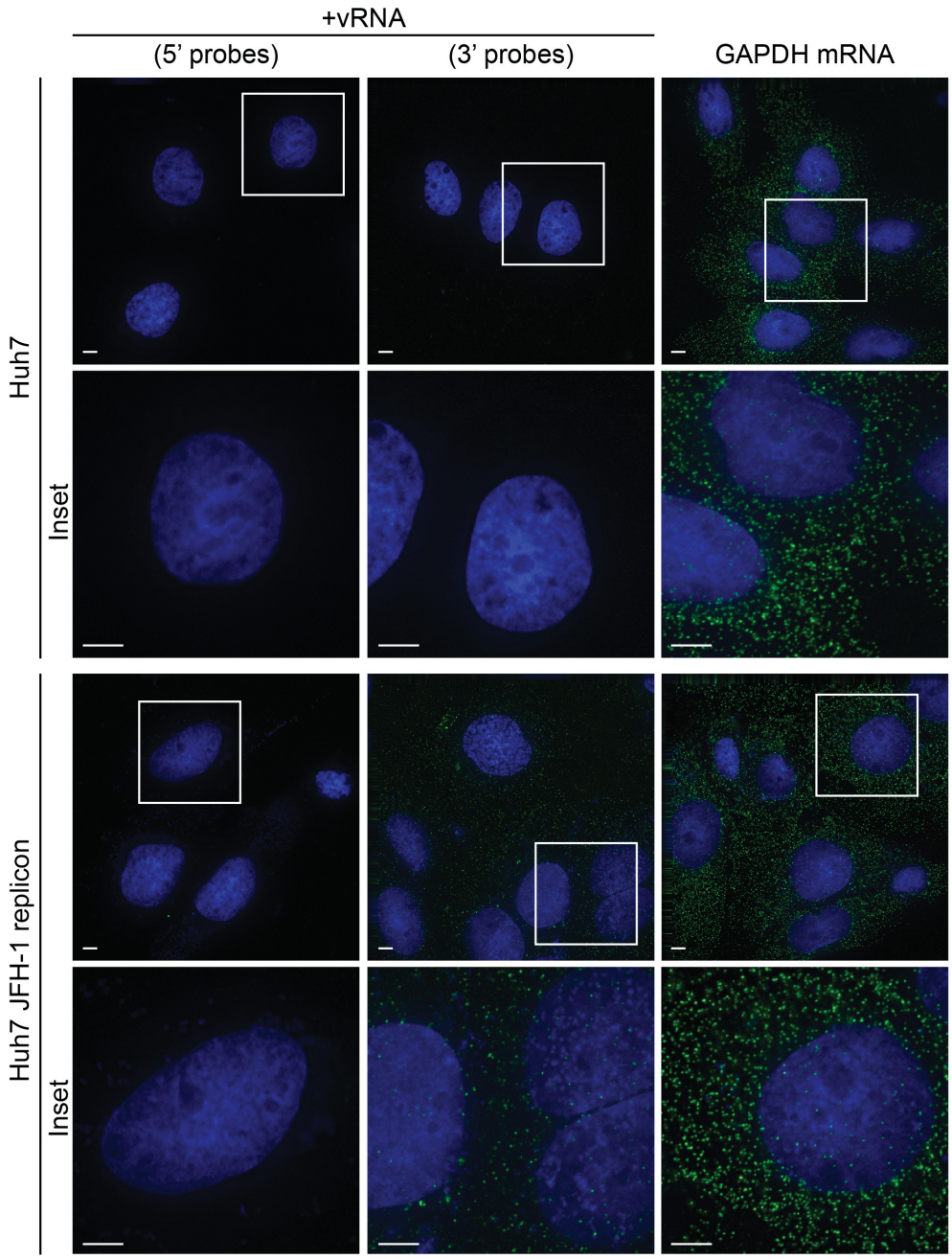


Figure 4.6: HCV replicon cells have an intranuclear pool of +vRNA.

Huh7 cells or Huh7 cells stably transfected with a subgenomic HCV JFH-1 replicon were analyzed by FISH with probes targeting the GAPDH mRNA or probes targeting the HCV +vRNA genome (5' and 3' probes). The sequence targeted by the 5' probes is absent from the subgenomic HCV JFH-1 replicon. GAPDH mRNA and +vRNA localization were evaluated by epifluorescence microscopy and are shown in green. Insets showing higher magnification of regions marked by a white square are shown below each panel. DNA was stained with DAPI (blue). Scale bar, 5 μm. N = 3 biological replicates, representative FISH images are shown.

mation regarding the localization of detected RNA foci in relation to the nuclei of cells. One caveat of this methodology is that the DAPI isosurface, used as a proxy for nuclei boundaries, is likely localized within ≈ 0.5 μm of the NE, in the interior of the nucleus. Thus, we may underestimate the number of nuclear RNA foci by excluding those in close proximity to the NE. Quantifying images of +vRNA FISH in Huh7.5 cells infected with HCV JFH-1 for 4 days (Figure 4.5), we see that approximately 5% of the detected +vRNA foci are nuclear. As expected, the control GAPDH mRNA also shows nuclear localization of around 13% of its mRNA foci (Figure 4.5).

4.3.5 Nuclear +vRNA amounts change during the time course of HCV infection

Once the presence of nuclear +vRNA foci was confirmed in HCV-infected cells, we next examined the localization of these +vRNAs during different time points of HCV infection. Huh7.5 cells were infected with HCV (synchronized infection) for two, three, or four days and analyzed by FISH. The HCV +vRNA is visible in infected cells two days after infection, with almost 6% of the +vRNA foci showing nuclear localization at this time point. At three days post infection, the proportion of nuclear +vRNA increases to $\approx$ 13% of the total HCV +vRNA foci observed. The nuclear pool of +vRNA decreases again at four days post infection, to about 7% of the total HCV +vRNA foci present (Figure 4.8). The control GAPDH mRNA can also be seen inside the nuclei of HCV-infected cells. The proportion of nuclear GAPDH mRNA foci remains constant, at around 13%, on all experimental conditions described (Figure 4.9).

We also examined the localization of the HCV negative strand vRNA (-vRNA) during the same time points of infection. The -vRNA of viruses with a positive-strand RNA genome serves as a template for replication. Therefore, +vRNA colocalized with -vRNA represent an RNA pool that could be involved in active replication (Shulla and Randall, 2015). Using FISH to detect the HCV -vRNA in combination with the +vRNA (5' probes), we see that the +vRNA and -vRNA do not colocalize in the nuclei of HCV-infected cells, and only partially colocalize in cytoplasmic regions (likely at replication complexes). This lack of colocalization indicates that the nuclear pool of +vRNA is not actively replicating (Figure 4.10). Nonetheless, we detect $\approx$ 5% of -vRNA foci localizing to the nuclei of HCV-infected cells at later infection time points (three and four days after infection) (Figure 4.11).

The detectable presence of a nuclear vRNA pool does not necessarily imply that these vRNAs are actively participating in the life cycle of the virus. Given that human cells undergo open mitosis and that HCV-infected cells contain a large number of HCV vRNA copies, it is possible vRNAs could get trapped in the nucleus during

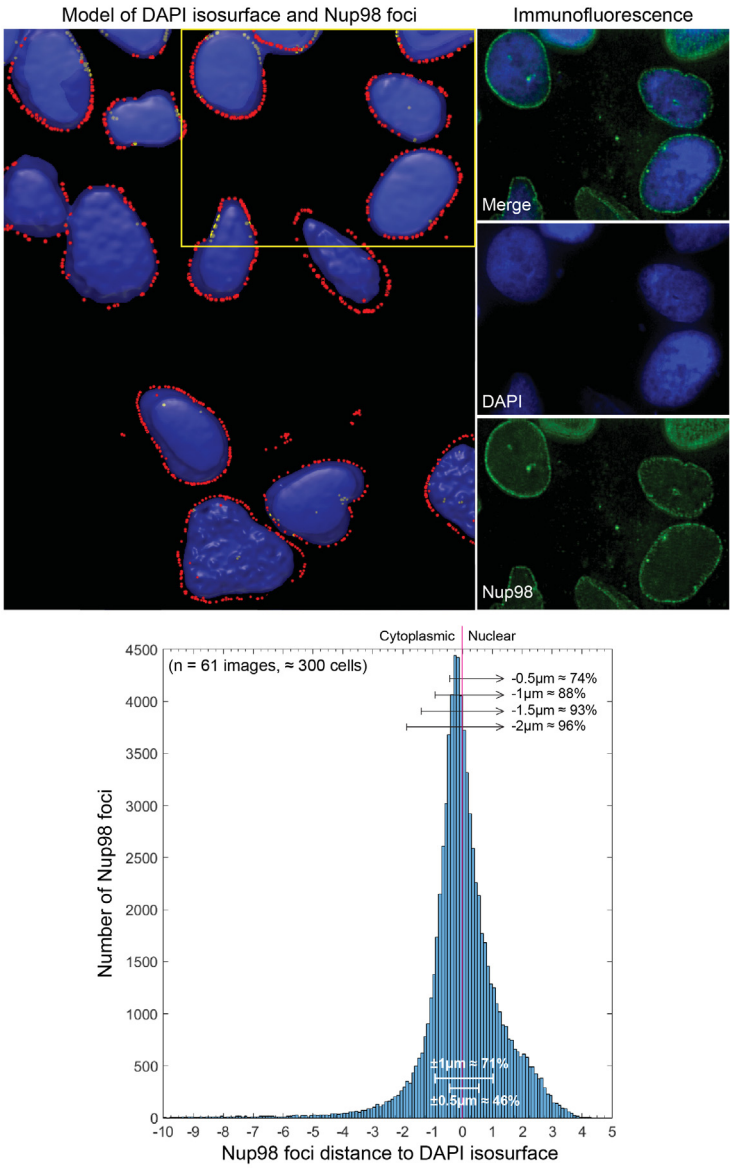


Figure 4.7: Localization of Nup98 foci and DAPI isosurfaces in Huh7.5 cells.

Localization of Nup98 in Huh7.5 cells was determined by indirect immunofluorescence microscopy using antibodies specific for Nup98. DNA was stained with DAPI. Images were mapped onto a 3D plane to create isosurfaces encompassing the DAPI stained intranuclear region, and points indicating Nup98 foci localization. **Top panels:** an example image of the mapped DAPI isosurfaces (blue) and Nup98 foci (red - outside, yellow - inside DAPI isosurface) are shown on the left. A portion of the corresponding immunofluorescence images are shown on the right (Nup98 - green, DNA bound DAPI stain - blue). **Bottom panel:** Histogram showing the distribution of distances between Nup98 foci and the closest DAPI isosurface. N= 61 images, corresponding to approximately 300 cells from 3 biological replicates.

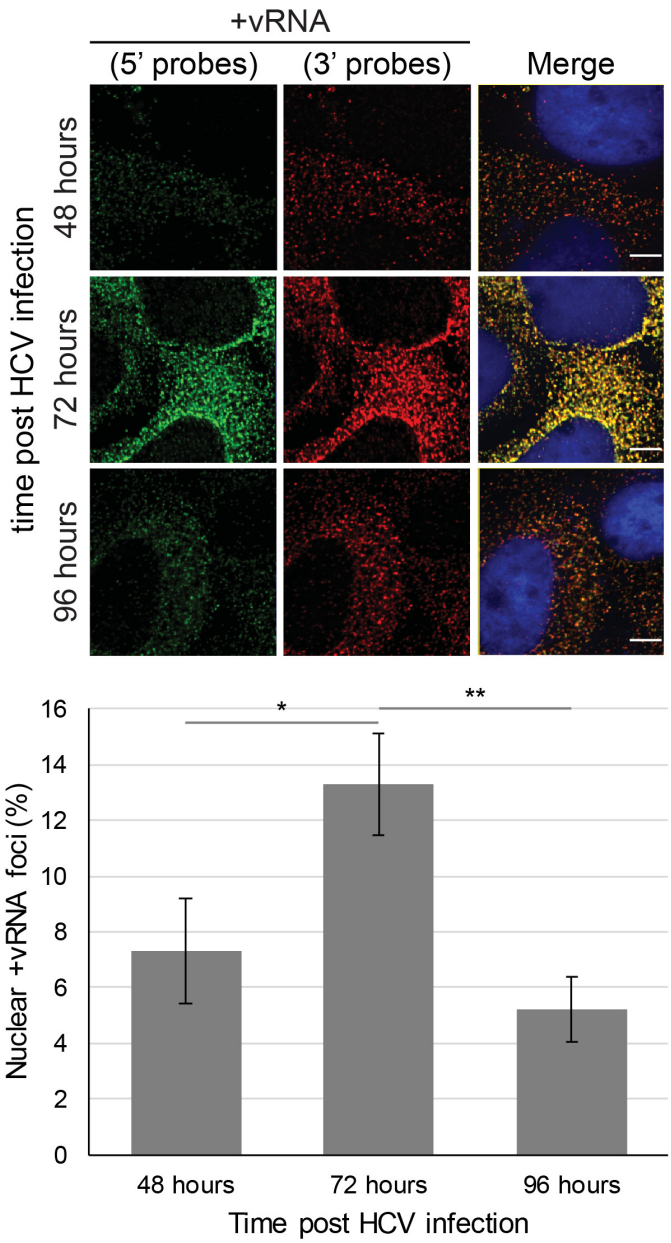


Figure 4.8: The nuclear pool of +vRNA changes during the time course of HCV infection.

Huh7.5 cells were infected with HCV for 48 to 96 hours. Localization of +vRNA (5' and 3' probes) in HCV-infected cells was evaluated by FISH visualized by epifluorescence microscopy. Time point of infection is indicated to the left of the image. DNA was stained with DAPI (blue). Scale bar, 5 μm. Plot shows the average percentage of HCV +vRNA foci with nuclear localization in HCV-infected cells at different times post infection. Error bars indicate standard deviation. N $\geq$ 3 biological replicates, representative FISH images are shown. Results from $\geq$ 3 biological replicates were submitted to ANOVA followed by Tukey HSD tests. Adjusted p-values for Tukey HSD in pairwise comparisons between indicated time points are represented as ** $<$ 0.01 $<$ * $<$ 0.05.

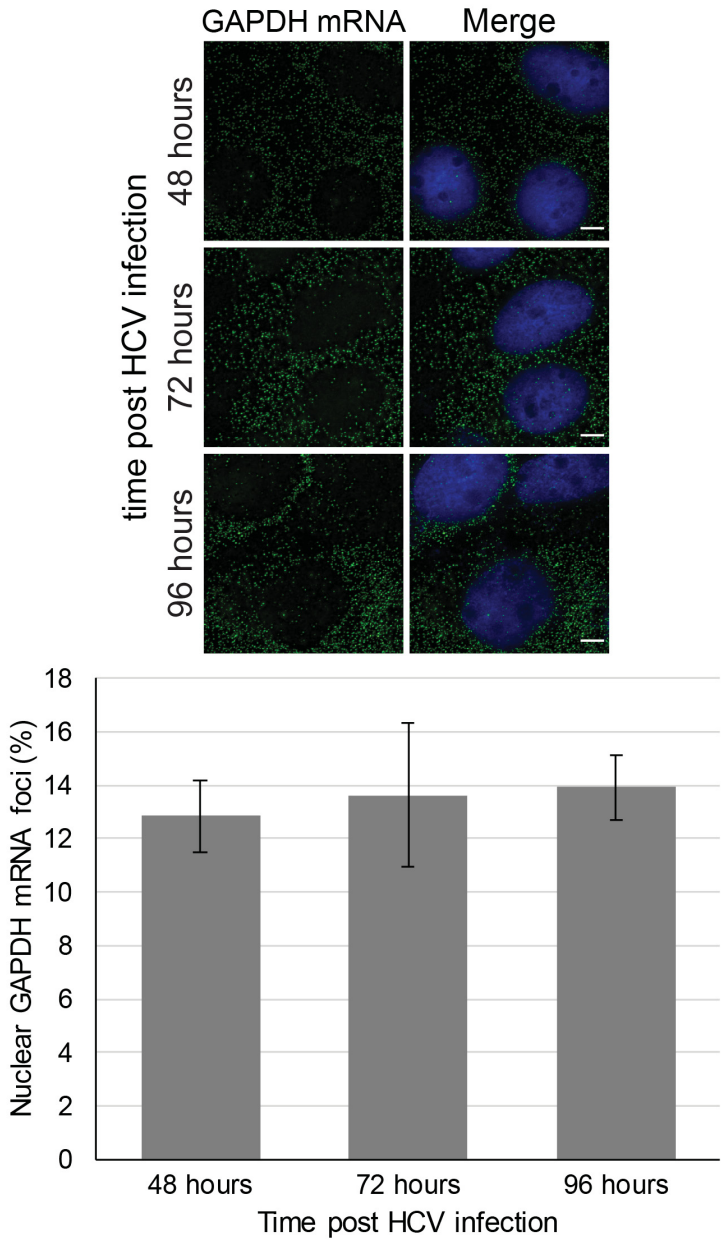


Figure 4.9: The nuclear pool of GAPDH mRNA does not change during the time course of HCV infection.

Huh7.5 cells were infected with HCV for 48 to 96 hours. Localization of GAPDH mRNA in HCV-infected cells was evaluated by FISH visualized by epifluorescence microscopy. Time post infection is indicated on the left of images. DNA was stained with DAPI (blue). Scale bar, 5 μm. Plot shows the average percentage of GAPDH mRNA foci with nuclear localization in HCV-infected cells at different time points of infection. Error bars indicate standard deviation. N $\geq$ 3 biological replicates, representative FISH images are shown. Results from $\geq$ 3 biological replicates were submitted to ANOVA test, with resulting p-value $>$ 0.05.

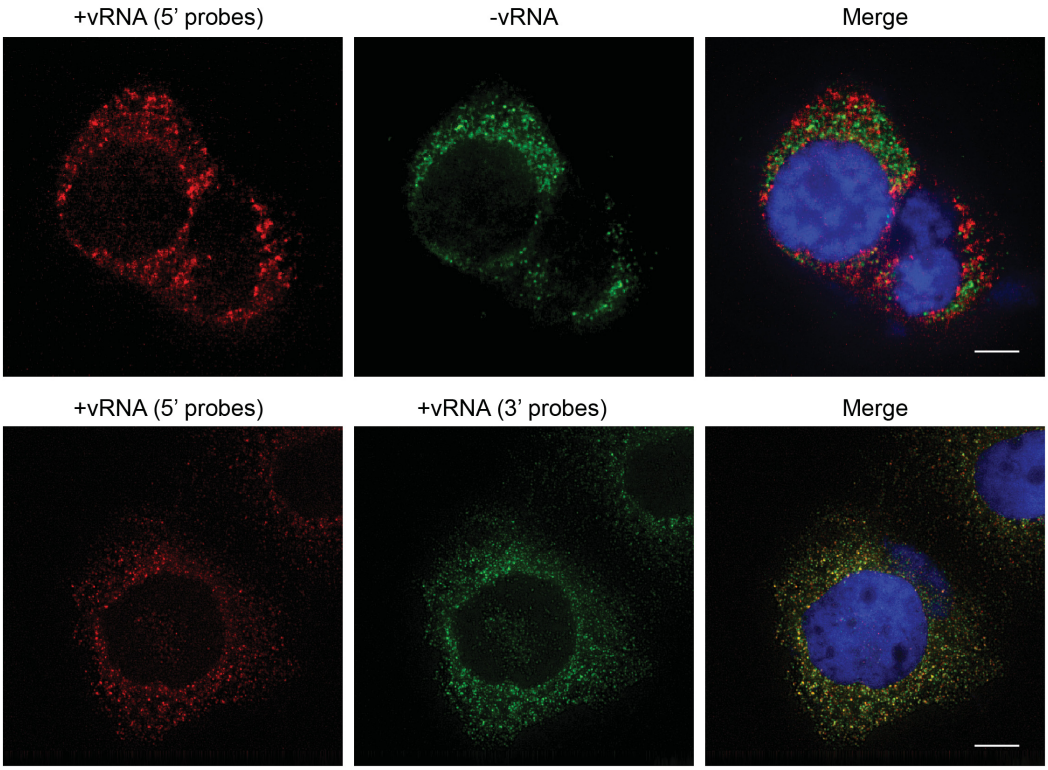


Figure 4.10: HCV -vRNA does not colocalize with +vRNA inside the nuclei of infected cells.

Huh7.5 cells were infected with HCV for three days. Colocalization of HCV -vRNA (red) and +vRNA (green) in HCV-infected cells (shown in top panels) was compared to colocalization of two probe sets hybridizing with +vRNA (5' and 3' probes), shown in bottom panels. Localization of vRNAs was evaluated by FISH and visualized by epifluorescence microscopy. DNA was stained with DAPI (blue). Scale bar, 5 μm. N = 3 biological replicates, representative FISH images are shown.

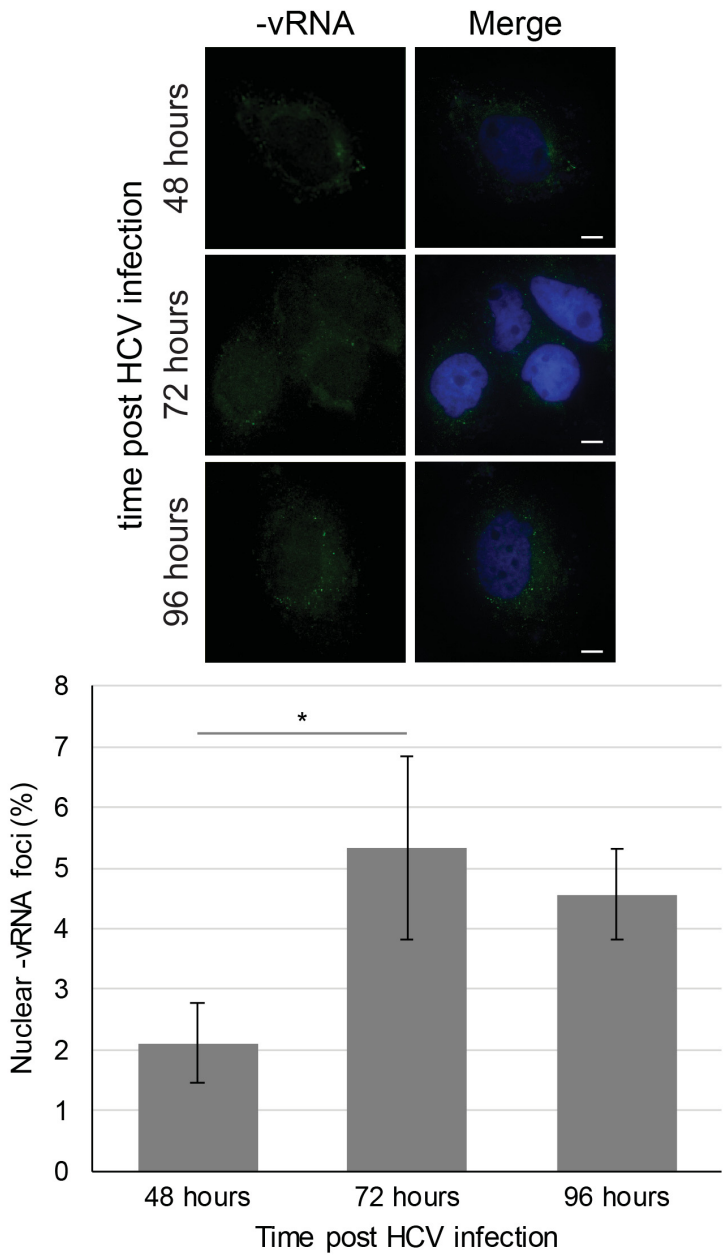


Figure 4.11: The nuclear pool of HCV -vRNA changes during the time course of infection

Huh7.5 cells were infected with HCV for 48 to 96 hours. Localization of HCV -vRNA in HCV-infected cells was evaluated by FISH visualized by epifluorescence microscopy. Time post infection is indicated on the left of images. DNA was stained with DAPI (blue). Scale bar, 5 μm. Plot shows the average percentage of HCV -vRNA foci with nuclear localization in HCV-infected cells at different time points of infection. Error bars indicate standard deviation. N $\geq$ 3 biological replicates, representative FISH images are shown. Results from $\geq$ 3 biological replicates were submitted to ANOVA followed by Tukey HSD tests. Adjusted p-values for Tukey HSD in pairwise comparisons between indicated time points are shown as * $<$ 0.05.

cell division. In this case the proportion of nuclear +vRNA would most likely be static. However, the proportion of nuclear +vRNA changes during a time course of HCV infection (increasing and decreasing, see Figure 4.8), while the total amount of +vRNA in infected cells has been shown to continually increase during the same time course (Neufeldt et al., 2013). This hints at the possibility that this is may not be a static pool of HCV +vRNA trapped in the nuclei of infected cells, but a dynamic pool of actively transported RNA molecules.

4.3.6 NTFs regulate HCV +vRNA nuclear localization

In order to determine if the nuclear pool of HCV +vRNA was dynamically transported in and out of the nuclei we tried to alter the nuclear localization of this +vRNA with nuclear export and nuclear import inhibitors. We attempted to further accumulate +vRNA in the nucleus by altering nuclear export with leptomycin B (LMB), a drug that inhibits CRM1-dependent nuclear export (Sun et al., 2013). HCV-infected cells were treated with LMB for three hours, starting 69 hours after HCV infection. Samples were analyzed by FISH at 72 hours p.i., the time point showing the most nuclear +vRNA foci during HCV infection (Figure 4.8). These HCV-infected cells show nuclear accumulation of HCV +vRNA in the presence of LMB. The percentage of intranuclear +vRNA foci increased to $\approx$ 21% in the presence of LMB, from $\approx$ 13% in cells treated with a drug vehicle (indicated as control) (Figure 4.12). Interestingly, LMB has also been shown to cause accumulation of HCV core protein in the nuclei of infected cells (Cerutti et al., 2011), a viral protein that can interact with the HCV +vRNA.

We also attempted to alter the nuclear pool of +vRNA in HCV-infected cells via nuclear import block, using Importazole (IPZ) or Ivermectin (IVM) treatments. IPZ inhibits nuclear import through the Kapβ1 pathway. It disrupts the RanGTP-Kapβ1 interaction preventing the release of cargoes in the nuclei of cells. IVM is an inhibitor of Kapα/β nuclear import. It disrupts the interaction between Kapα and Kapβ1, preventing nuclear import of Kapα bound cargoes. As described previously, Huh7.5 cells infected with HCV for 69 hours were treated with IPZ or IVM for three hours and then analyzed by FISH assays (at 72 hours p.i.). The presence of IPZ and IVM caused a non statistically significant reduction in the pool of nuclear HCV +vRNA (from $\approx$ 13% in cells treated with drug vehicle to $\approx$ 10% in cells treated with IPZ or IVM) (Figure 4.12).

No change was observed in the nuclear localization of HCV -vRNA or GAPDH mRNA under any of the above mentioned inhibitor treatments (i.e., LMB, IPZ, or IVM) (Figure 4.13). The presence of nucleocytoplasmic inhibitors (3 hours) and

HCV-infection (72 hours) had a negligible effect in the viability of Huh7.5 cells (Figure 4.14).

Since LMB-mediated CRM1-dependent nuclear export block can affect the nuclear localization of HCV +vRNA, we next decided to investigate if disrupting other NTFs could have a similar effect. We used shRNA-mediated knock-down to deplete the import factors Kapβ1 and Kapβ3, and the export factors CRM1 and NXF1. Huh7.5 cells were simultaneously infected with HCV and lentiviruses containing the target shRNAs. Cells were analyzed by FISH three days after infection. Proper target mRNA depletion was quantified by qPCR (Figure 4.15a). Protein depletion was determined by quantitative western blot (Figure 4.15b). None of the experimental conditions described above had a significant effect in cell viability (Figure 4.15c). Of note, we could only deplete 50% of NXF1 with no significant loss of cell viability during this time frame (Figure 4.15). Depletion of these NTFs (Kapβ1, Kapβ3, CRM1 and NXF1) had no statistically significant effect on the nuclear localization of the HCV -vRNA or the host GAPDH mRNA (Figure 4.16). Only a minor (not statistically significant) increase in the nuclear pool of host GAPDH mRNA was observed upon NXF1 depletion, likely due to this protein's function on mRNA export (Okamura et al., 2015).

Depletion of nuclear import factors Kapβ1 or Kapβ3 caused a decrease in the nuclear localization of HCV +vRNA in infected cells. Kapβ1 depletion caused a $\approx$ 2% decrease in nuclear +vRNA foci, similar to what was observed for the inhibitors IPZ and IVM (Figure 4.16a and 4.12). Kapβ3 depletion had a more significant effect, decreasing the amount of nuclear HCV +vRNA foci by half (Figure 4.16a). Interestingly, Kapβ3 can interact with the NLS of HCV proteins that interact with the +vRNA (core, NS3, and NS5A) (Levin et al., 2014b; Neufeldt et al., 2013).

Depletion of the nuclear export factors CRM1 or NXF1 caused an increase in the nuclear localization of HCV +vRNA foci, from $\approx$ 11% in control cells to $\approx$ 15% in NXF1-depleted cells and $\approx$ 17% upon CRM1 depletion (Figure 4.16a). This increase was similar to what was previously observed for the LMB-mediated CRM1-dependent nuclear export block (Figure 4.12). It is possible that the HCV +vRNA can utilize the CRM1 and NXF1-dependent export pathways, depending on the proteins it interacts with, and/or the current function of the +vRNA in the HCV life cycle. Similar exploitation of CRM1 and NXF1-dependent nuclear export has been observed for the influenza A virus (York and Fodor, 2013) and some retroviruses (Fontoura et al., 2005). Of note, CRM1 has also been shown to interact with the NES of two HCV proteins (NS2 and core). Given that the viral proteins core, NS3, and NS5A are capable of interacting with the +vRNA and that these HCV proteins have been previously detected in the nuclei of infected cells (Bonamassa et al., 2015), it is possible that the +vRNA is transported in and out of the nucleus in complex with

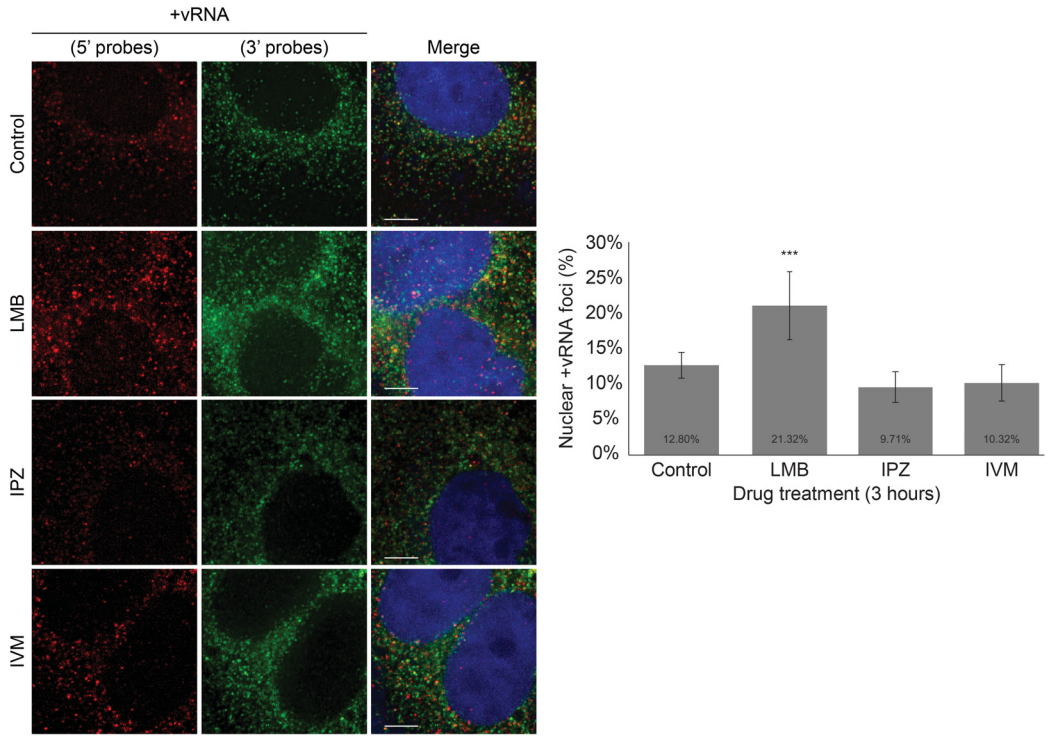


Figure 4.12: NTF inhibitors alter the nuclear localization of HCV +vRNAs.

Huh7.5 cells were infected with HCV for three days and treated with drug vehicle (control), LMB, IPZ or IVM for 3 hours (at 69 hours p.i.). Localization of HCV +vRNA (5' and 3' probes) in HCV-infected cells treated with the indicated inhibitors was determined by FISH visualized by epifluorescence microscopy. DNA was stained with DAPI (blue). Scale bars, 5 μm. Plot shows the average percentage of HCV +vRNA foci with nuclear localization in HCV-infected cells in the presence of indicated NTF inhibitors. Error bars indicate standard deviation. N $\geq$ 3 biological replicates, representative FISH images are shown. Results from $\geq$ 3 biological replicates were submitted to ANOVA followed by Tukey HSD tests. Adjusted p-values $<$ 0.001 for Tukey HSD in pairwise comparisons between drug treated and control cells are indicated as ***.

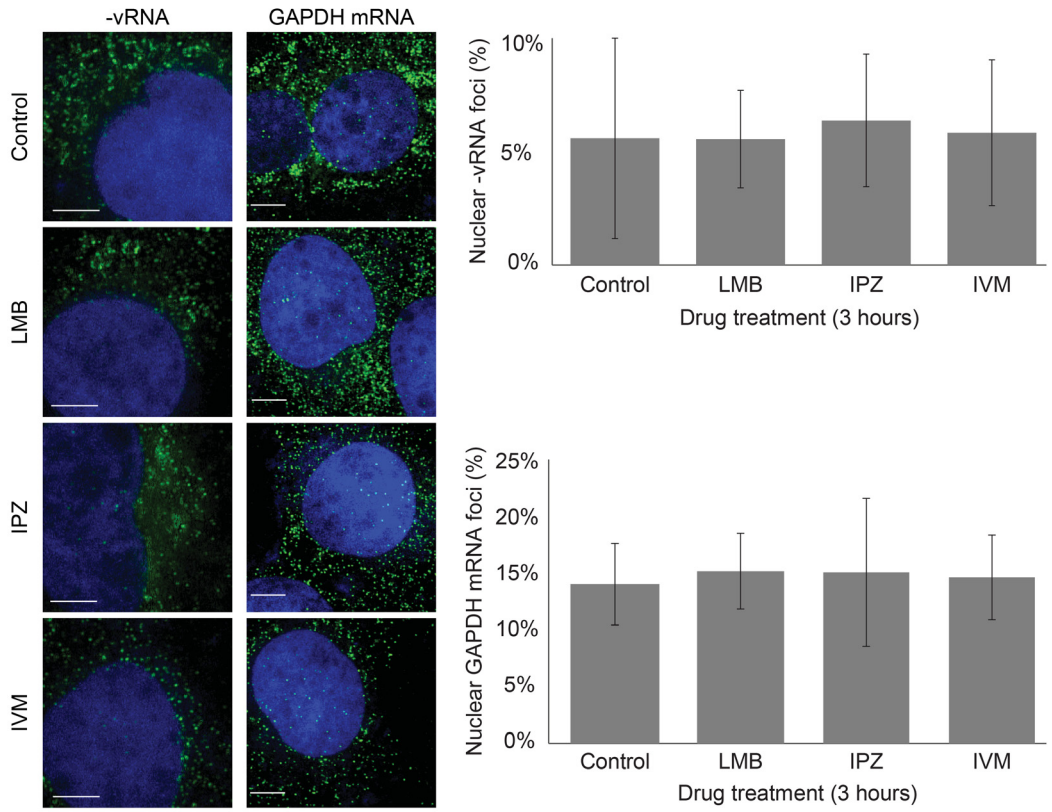


Figure 4.13: NTF inhibitors do not affect the nuclear localization of HCV -vRNA or GAPDH mRNA.

Huh7.5 cells were infected with HCV for three days and treated with drug vehicle (control), LMB, IPZ or IVM for 3 hours (at 69 hours p.i.). Localization of HCV -vRNA or GAPDH mRNA (indicated above images), shown in green, was determined by FISH visualized by epifluorescence microscopy. DNA was stained with DAPI (blue). Scale bars, 5 μm. Plots show the average percentage of HCV -vRNA foci (top) or GAPDH mRNA foci (bottom) with nuclear localization in HCV-infected cell in the presence of indicated NTF inhibitors. Error bars indicate standard deviation. N $\geq$ 3 biological replicates. Results from $\geq$ 3 biological replicates were submitted to ANOVA tests and showed p-value $>$ 0.05.

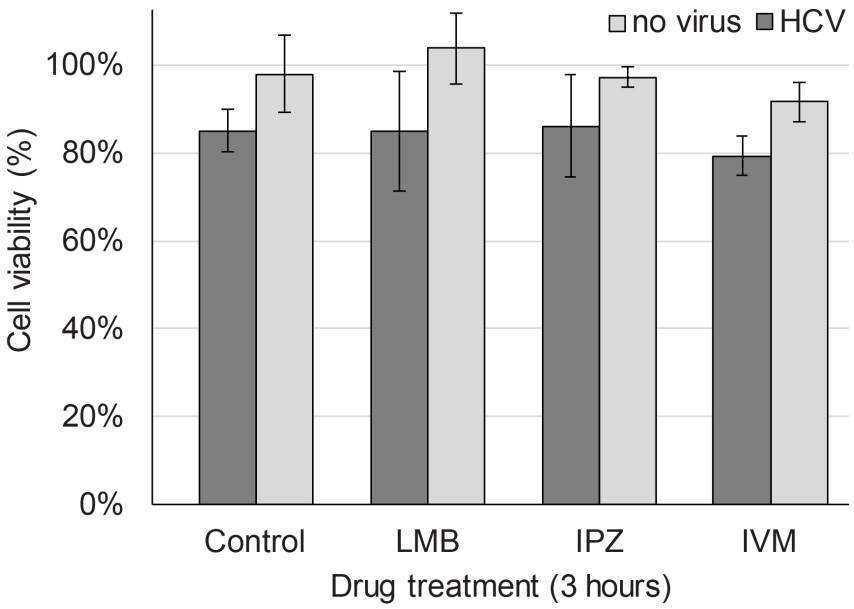
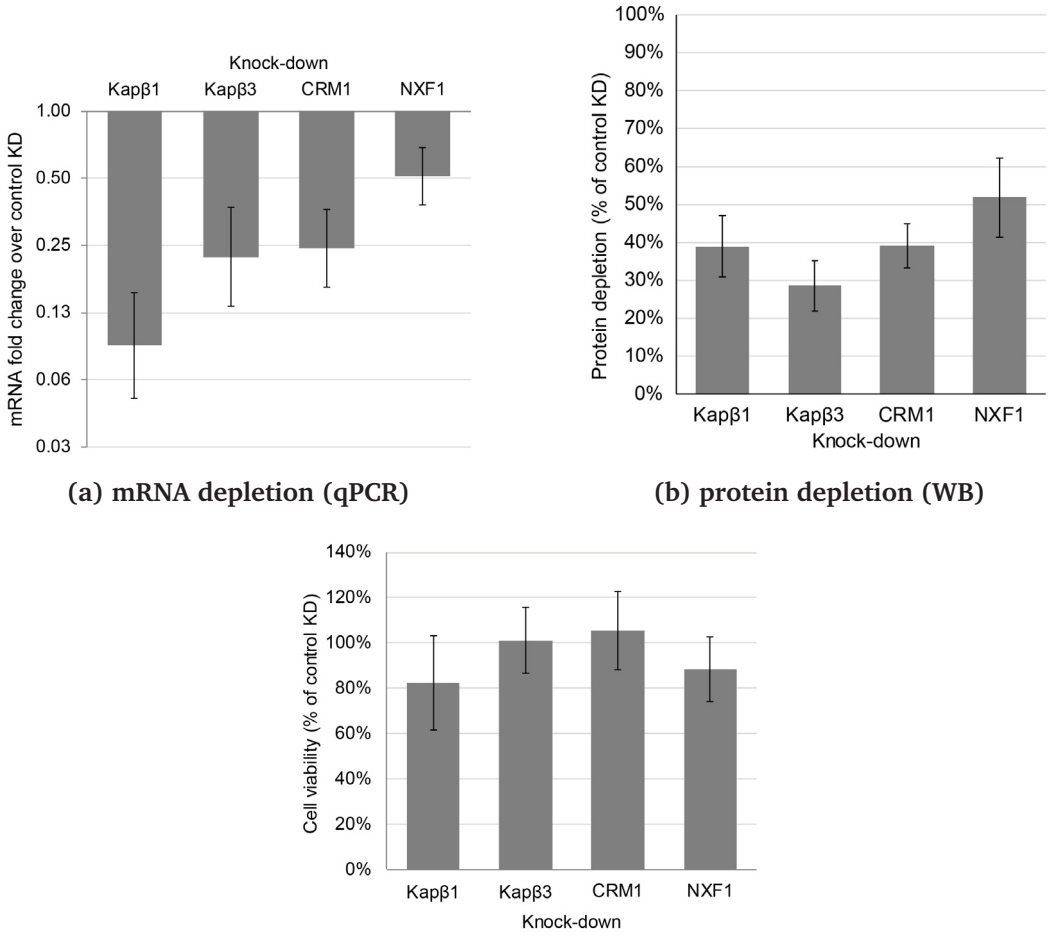


Figure 4.14: NTF inhibitors have no major effect in the viability of Huh7.5 cells infected with HCV.

Huh7.5 cells were infected with HCV for 72 hours and treated with indicated NTF inhibitors for 3 hours. The cytotoxic effects of HCV infection and/or drug treatments on Huh7.5 cells were evaluated by mitochondrial dehydrogenases activity. Average values for cell viability are expressed as a percentage of the viability of Huh7.5 cells (uninfected, in the absence of inhibitors). Error bars indicate standard deviation. N $\geq$ 3 biological replicates. Results from $\geq$ 3 biological replicates were submitted to ANOVA followed by Tukey HSD tests. Adjusted p-values were $>$ 0.05 for Tukey HSD in pairwise comparisons between inhibitor treated and corresponding control cells. This data was produced in collaboration with Brett Roughead.



(a) mRNA depletion (qPCR)

(b) protein depletion (WB)

(c) cell viability

Figure 4.15: NTF depletion in HCV-infected Huh7.5 cells.

Huh7.5 cells were coinfected with HCV and lentivirus encoding shRNAs directed against Kapβ1, Kapβ3, CRM1, NXF1, or a non-mammalian control sequence for three days. **4.15a** Average transcript depletion was evaluated by qPCR with primers for Kapβ1, Kapβ3, CRM1, NXF1. Samples were normalized to house-keeping transcripts and fold-change relative to cells expressing the control shRNA was calculated. **4.15b** Protein depletion following shRNA expression was evaluated by western blot with antibodies against Kapβ1, Kapβ3, CRM1, NXF1. Samples were normalized to a load control and average percentage of protein depletion relative to cells expressing the control shRNA was calculated. **4.15c** The viability of HCV-infected cells expressing shRNAs was evaluated via mitochondrial dehydrogenases activity. Average viability of cells depleted of the indicated proteins is expressed as a percentage of HCV-infected cells expressing the control shRNA. For all panels, N = 3 biological replicates. Part of these data were produced in collaboration with Dr. Zhongjing Su.

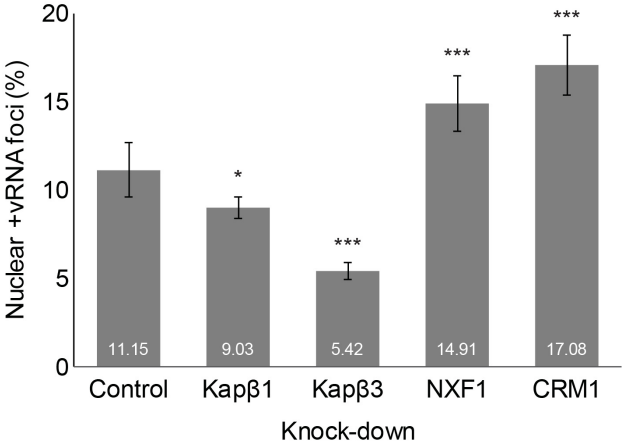
these viral proteins.

4.3.7 Disrupting nucleocytoplasmic transport alters the HCV life cycle

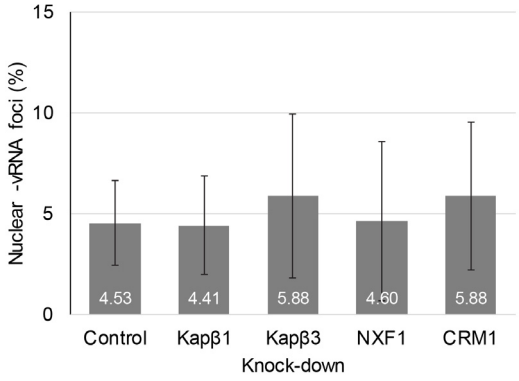
We assessed the consequences of disrupting nucleocytoplasmic transport on the HCV life cycle by focusing on three steps of HCV infection: +vRNA genome replication, +vRNA translation, and production of infectious virus. Viral genome replication was determined by quantifying the intracellular pool of HCV +vRNA by qPCR, +vRNA translation was assessed through quantification of intracellular HCV proteins, and production of infectious virus was examined by quantifying the infectious titer of HCV produced through re-infection assays.

We began by treating HCV-infected Huh7.5 cells (69 hours p.i.) with NTF inhibitors (LMB, IPZ, and IVM) for three hours. The presence of inhibitors had no detectable effect on the levels of intracellular HCV +vRNA (Figure 4.17a) or viral protein (Figure 4.17b). It should be noted, however, that three days after HCV infection, the HCV +vRNA and viral proteins are present in high abundance in infected cells. The estimated half-lifes of HCV +vRNA and HCV NS proteins are 11 to 16 hours (Quinkert et al., 2005; Pause et al., 2003; Pietschmann et al., 2001). Our treatments with LMB, IPZ or IVM lasted only three hours, followed by immediate sample collection. Therefore, given the abundance and stability of HCV +vRNA and protein previously present in the treated cells, it is possible that any effects the inhibitors had in HCV +vRNA replication or translation would only affect a comparably small pool of nascent +vRNA or protein. These factors could cause us to underestimate the effect of NTF inhibitors on HCV +vRNA replication or translation.

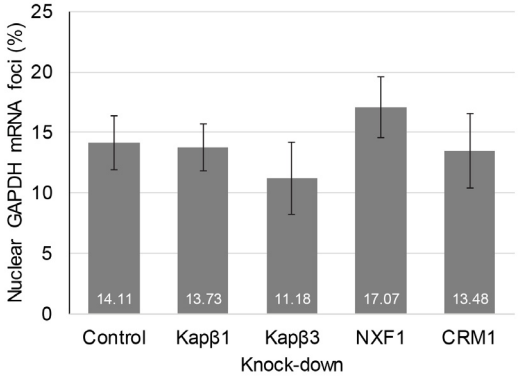
We also evaluated the effects of NTF inhibitors in the production of HCV by infected cells. Huh7.5 cells infected with HCV for 69 hours were repeatedly washed and cell culture media was replaced with media containing nucleocytoplasmic transport inhibitors (LMB, IPZ or IVM), drug vehicle or media alone. Media containing newly produced HCV were collected three hours later and used for re-infection of newly plated Huh7.5 cells, allowing quantification of the titer of infectious virus present. Due to the presence of NTF inhibitors in the media containing the viruses tested in these re-infection assays, we performed parallel control experiments to determine if the presence of these inhibitors in the media could affect HCV infection of Huh7.5 cells. For these control re-infection assays we added the same concentration of inhibitors to the virus present in the supernatant collected from HCV-infected cells in the presence of media alone described above. These re-infection assays measured the infectious HCV produced only during the three hours of drug treatment, and could be more sensitive to the effects of NTF inhibitors than the above mentioned



(a) HCV +vRNA foci localizaton



(b) HCV -vRNA foci localizaton



(c) GAPDH mRNA foci localization

Figure 4.16: Depleting host NTFs alters the nuclear localization of HCV +vRNA.

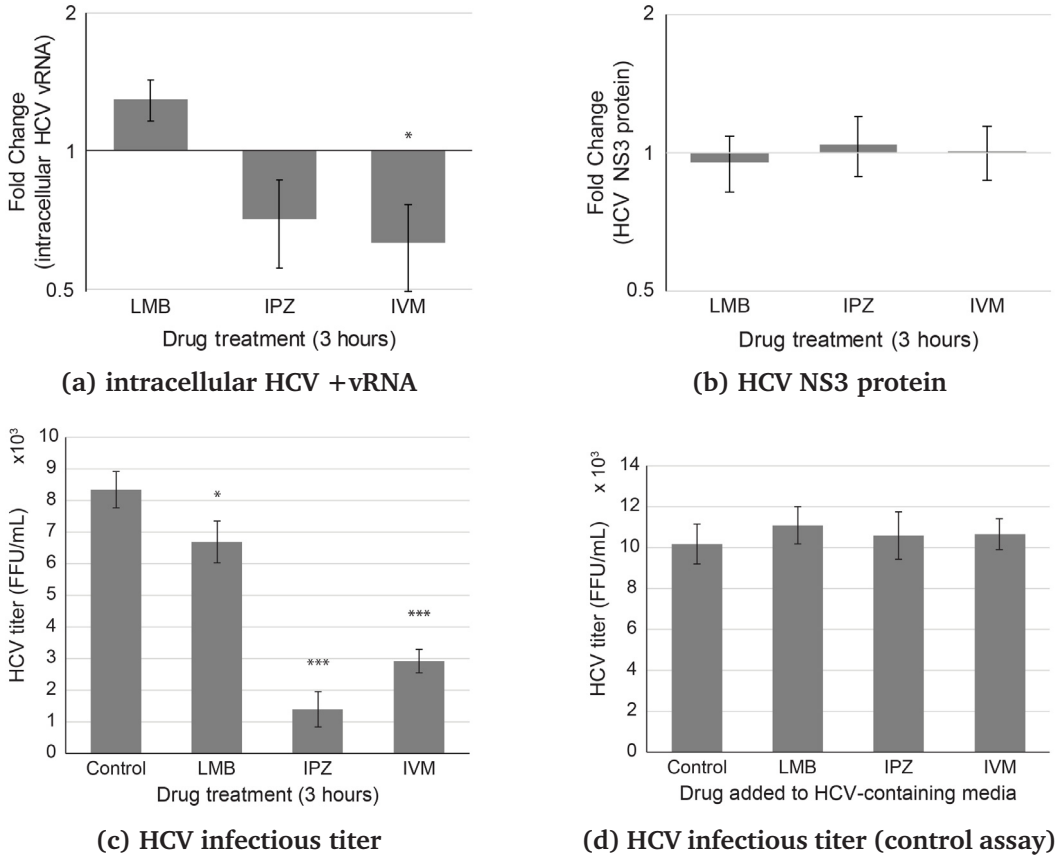
Huh7.5 cells were coinfected with HCV and lentivirus encoding shRNAs directed against Kapβ1, Kapβ3, CRM1, NXF1, or a non-mammalian control sequence for three days. Localization of HCV +vRNAs, -vRNAs or GAPDH mRNAs was determined by FISH visualized by epifluorescence microscopy. Plot shows the average percentage of HCV +vRNA (**4.16a**), -vRNA (**4.16b**) or GAPDH mRNA (**4.16c**) foci with nuclear localization in HCV-infected cells depleted of the indicated proteins (x-axis). Error bars indicate standard deviation (N = 3 biological replicates). Results from 3 biological replicates were submitted to ANOVA followed by Tukey HSD tests. Adjusted p-values for Tukey HSD in pairwise comparisons between target knock-downs and mock depleted cells are shown as *** $< 0.001 <$ * < 0.05.

quantification of intracellular HCV +vRNA or viral proteins.

We observed a decrease in the titers of infectious HCV following treatment of infected cells with all nucleocytoplasmic transport inhibitors. Inhibition of Kapβ1-dependent import decreased virus production by over 80%, while Kapα/β-dependent import inhibition led to over 60% decrease in viral titers. Inhibiting CRM1-dependent nuclear export had a less severe effect on viral production, decreasing it by ≈ 20% (Figure 4.17c). The control re-infection assays showed that addition of NTF inhibitors to the media containing HCV does not affect infection of Huh7.5 cells (Figure 4.17d).

These NTF inhibitors will not only affect transport between the nucleus and the cytoplasm, but also transport at the membranous web (MW) of HCV-infected cells (Neufeldt et al., 2013). A recent publication has shown that HCV-infected cells treated with IPZ or IVM display a significant decrease in the presence of an NLS-cargo at the MW, further indicating that Kapβ1 and Kapα/β regulate the access of proteins to this compartment (Neufeldt et al., 2016). Therefore, we may not be able to differentiate if the changes observed in the HCV life cycle upon NTF inhibition are due to altered transport at the nucleus or the MW.

We also evaluated the consequences of shRNA-mediated depletion of NTFs in the HCV life cycle. Previous publications have shown that depletion of the nuclear import factor Kapβ3 has no effect in the abundance of intracellular HCV +vRNA or in the level of viral core protein in infected cells, although Kapβ3 depletion decreased the level of extracellular HCV +vRNA (Neufeldt et al., 2013). These results would indicate that Kapβ3 participates in the production of secreted HCV, but not in +vRNA replication or translation. Similarly, we observe no change in the abundance of HCV +vRNA or NS3 protein in HCV-infected cells depleted of Kapβ3. However, Kapβ3 depletion decreased the production of infectious virus by more than half (Figure 4.18), corresponding to what was previously observed (Neufeldt et al., 2013). Depletion of another nuclear import factor, Kapβ1, also caused no change in the level of intracellular HCV +vRNA (Figure 4.18a), possibly indicating this NTF also does not play a role in HCV +vRNA replication. Unlike Kapβ3 depletion, Kapβ1 depletion caused a decrease in the level of viral protein NS3 in HCV-infected cells (Figure 4.18b). Therefore, it is possible that Kapβ1-mediated transport plays a role in the translation of HCV +vRNA. Kapβ1 inhibitors (IPZ and IVM) show no effect in the level of viral protein NS3 in HCV-infected cells (Figure 4.17b). However, as previously mentioned, any effect of NTF inhibitors on vRNA replication or translation may not be apparent due to the short time of drug treatment (three hours). Kapβ1 depletion in HCV-infected cells also caused a significant decrease in the production of infectious virus (≈ 80% decrease) (Figure 4.18c), similarly to what was observed in HCV-infected cells treated with Kapβ1 inhibitors (Figure 4.17c). This decrease



(a) intracellular HCV +vRNA

(b) HCV NS3 protein

(c) HCV infectious titer

(d) HCV infectious titer (control assay)

Figure 4.17: Nucleocytoplasmic transport inhibitors decrease the production of HCV by infected cells.

Huh7.5 cells were infected with HCV for 69 hours and treated with the indicated NTF inhibitors for three hours before sample collection (72 hours p.i.). **4.17a** Changes in average intracellular HCV vRNA were evaluated by qPCR. Samples were normalized to housekeeping transcripts and fold-change relative to cells treated with drug vehicle was calculated. **4.17b** HCV NS3 protein average level was determined by western blot. Samples were normalized to a load control and fold-change relative to cells treated with drug vehicle was calculated. These data (4.17a, 4.17b) were produced by Dr. Zhongjing Su. The average infectious titers of HCV virus, in focus-forming units per mL of medium (FFU/mL), produced in the presence of indicated inhibitors (**4.17c**), or produced in the absence of drugs and combined with the inhibitors before reinfection assays (**4.17d**) were determined by indirect immunofluorescence microscopy. For all plots, error bars indicate standard deviation ($N \geq 3$ biological replicates). Results from ≥ 3 biological replicates were submitted to ANOVA followed by Tukey HSD tests. Adjusted p-values for Tukey HSD in pairwise comparisons between cells treated with drug vehicle and cells treated with the indicated inhibitors are shown as *** < 0.001, * < 0.05.

may indicate Kapβ1 also has a role in virus assembly and release.

Depletion of nuclear export factors in HCV-infected cells can also alter the virus life cycle. CRM1 depletion in HCV-infected cells led to a ≈ 50% increase in intracellular HCV +vRNA (Figure 4.18a), slightly more than was observed for infected cells treated with the CRM1 inhibitor LMB for three hours (≈ 30%) (Figure 4.17a). On the other hand, CRM1 depletion caused a decrease in the level of viral protein NS3 and in the titer of infectious HCV produced (Figure 4.18). It is possible that the presence of CRM1 has an inhibitory effect on the HCV +vRNA replication or stability. But CRM1 depletion decreases protein translation and assembly/release of infectious virus in HCV-infected cells. One possibility is that CRM1 depletion may impact the immune response in the cells. The importance of CRM1 activity in immunity has been previously described (Xu et al., 2015; Browne et al., 2006; Aggarwal and Agrawal, 2014; Cheng et al., 2014; Xylourgidis et al., 2006). It is possible that CRM1 depletion or inhibition allows increased viral infection and replication. However, if CRM1 is required for the later stages of the HCV life cycle, such as protein translation and/or virion production, then we would expect a decrease in viral proteins and infectious virus produced upon CRM1 depletion, as was observed (Figure 4.18).

NXF1 is the main mRNA nuclear export factor in mammalian cells, so it is not surprising that NXF1 is an essential gene. Consequently, we were only able to partially deplete its expression (50%) in Huh7.5 cells without significant loss of viability. As described for CRM1, NXF1 is also important for proper immune response. Numerous viruses, including those with cytoplasmic replication such as HCV, are known to restrict export of host mRNAs to the cytoplasm. This bulk viral block of host poly(A) mRNA export is usually mediated by interference with the proteins involved in the NXF1-dependent mRNA export pathway; however, disruption of CRM1-dependent mRNA export has also been described. This mRNA export restriction is key to inhibiting host gene expression, deregulating immune responses, and affecting many other host processes (Yarbrough et al., 2014; Kuss et al., 2013). NXF1 depletion caused an increase in intracellular HCV +vRNA, viral protein NS3, and titer of infectious HCV produced in HCV-infected cells (Figure 4.18). It is possible that this increased HCV replication and viral production is a consequence of the dampened immune response in these cells, due to the reduced levels of NXF1. However, since we were only able to partially deplete NXF1 in HCV-infected cells, we cannot discount the possibility that this protein may have a role in the HCV life cycle that is masked if the remaining NXF1 is sufficient for proper viral replication and production.

As previously mentioned for NTF inhibitors, depleting NTFs will likely not only affect nucleocytoplasmic transport, but also transport at the membranous web of

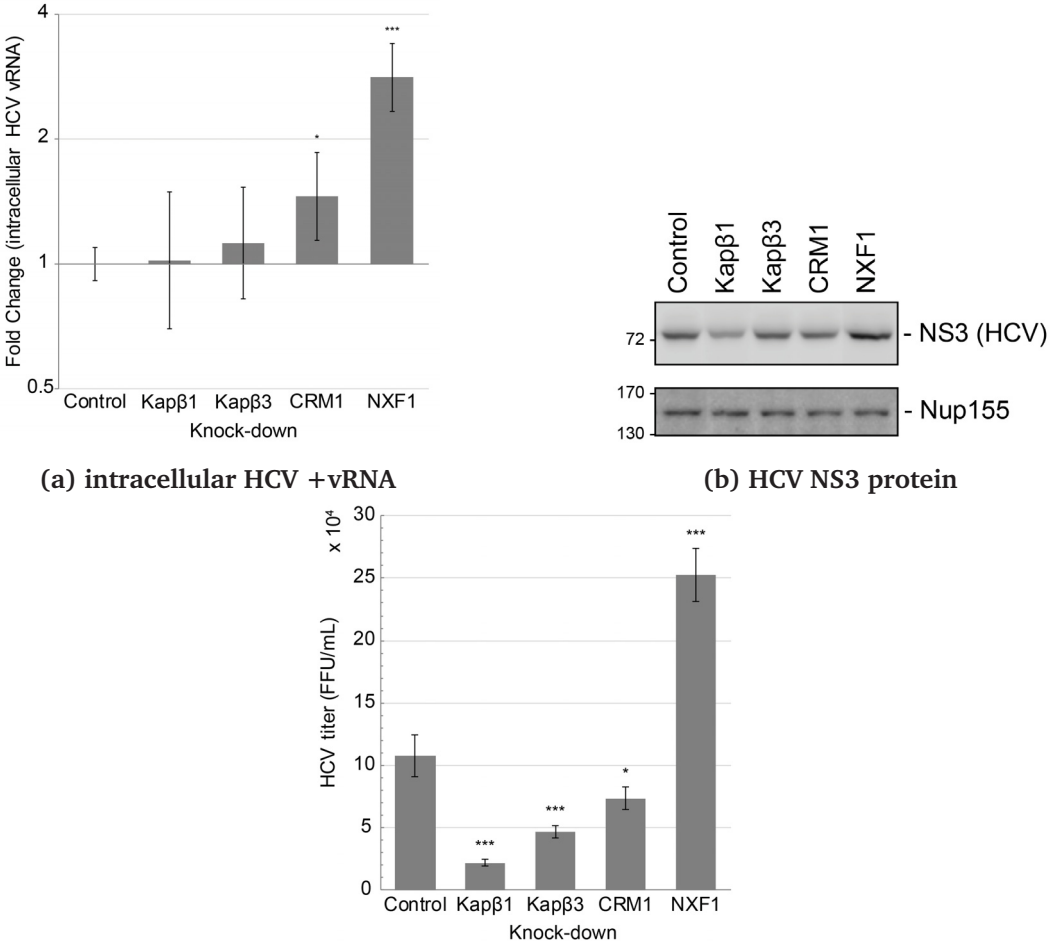
HCV-infected cells (Neufeldt et al., 2013, 2016). Thus, the changes observed in the HCV life cycle upon NTF depletion may be due to altered transport at the nucleus and/or the MW. However, we can hypothesize that if some of the effects of NTFs disruption on the HCV life cycle are due to changes in the nuclear localization of HCV +vRNA, these effects may be mediated by altered interactions of HCV +vRNA with host NFs that participate in the viral life cycle (subsection 4.3.2). Therefore, our next step was to determine if altering nucleocytoplasmic transport in HCV-infected cells could affect the interaction of host NFs exploited during infection with the HCV +vRNA.

4.3.8 Host NFs can bind HCV +vRNA and affect its nuclear localization

In previous subsections (see subsection 4.3.1, 4.3.2, and 4.3.3), we have identified nuclear host proteins that play important roles in the HCV life cycle and interact with the HCV +vRNA despite retaining their nuclear localization in infected cells. We have also detected HCV +vRNA in the nuclei of infected cells and shown that CRM1-export block increases the pool of intranuclear +vRNA (subsection 4.3.6) affecting the HCV life cycle (Figure 4.18). We may, therefore, hypothesize that the increase in nuclear +vRNA caused by LMB treatment of HCV-infected cells might increase the amount of +vRNA bound by nuclear host factors that participate in the HCV life cycle but localize to the nucleus.

In order to test this hypothesis Huh7.5 cells were infected with HCV for three days and treated with LMB or drug vehicle (control) for three hours. Cells were crosslinked and cleared cell lysates were combined with beads coupled to antibodies against the host NFs DHX9, hnRNP U, Nup98, SSB, ELAVL1 or viral protein NS3. Viral RNA bound to precipitated proteins was purified, reverse transcribed and used as template in qPCR reactions with prmers specific to the HCV +vRNA. LMB treatment of HCV-infected cells increased the interaction of hnRNP U, Nup98, DHX9 and SSB with HCV +vRNA, but it did not alter the interaction of ELAVL1 or the viral protein NS3 with +vRNA. Interestingly, ELAVL1 shows cytoplasmic puncta in HCV-infected cells, therefore it is possible that its interaction with the HCV +vRNA is cytoplasmic and not nuclear. All the other host factors retain their nuclear localization in infected cells, thus the increase in nuclear +vRNA promoted by LMB could facilitate the interaction between vRNA and these nuclear factors that aid HCV infection (Figure 4.19).

If we follow on our initial hypothesis that the HCV +vRNA may enter the nuclei of infected cells to interact with host NFs that participate in its life cycle, we could imagine that depletion of these NFs may alter the nuclear localization of +vRNA.



(a) intracellular HCV +vRNA

(b) HCV NS3 protein

(c) HCV infectious titer

Figure 4.18: Impact of NTF depletion on the HCV life cycle.

Huh7.5 cells were coinfected with HCV and lentivirus encoding shRNAs directed against Kapβ1, Kapβ3, CRM1, NXF1, or a control non-mammalian sequence for three days. **4.18a** Changes in average intracellular HCV vRNA were evaluated by qPCR. Samples were normalized to house-keeping transcripts and fold-change in intracellular HCV vRNA in knock-down cells relative to mock depleted cells was calculated. **4.18b** HCV NS3 levels were determined by western blot. Nup155 was used as load control that is unaffected by HCV infection (Neufeldt et al., 2013). Positions of molecular mass markers (in kDa) are indicated on the left. N = 3 biological replicates, representative western blot images are shown. **4.18c** The average infectious titers of HCV, in focus-forming units per mL of medium (FFU/mL), produced upon NTFs depletions were determined by indirect immunofluorescence microscopy. All plot error bars indicate standard deviation ($N \geq 3$ biological replicates). Results from ≥ 3 biological replicates were submitted to ANOVA followed by Tukey HSD tests. Adjusted p-values for Tukey HSD in pairwise comparisons between protein depleted and mock depleted samples are indicated as $*** < 0.001 < * < 0.05$. Part of the data were produced in collaboration with Dr. Zhongjing Su and Brett Roughead.

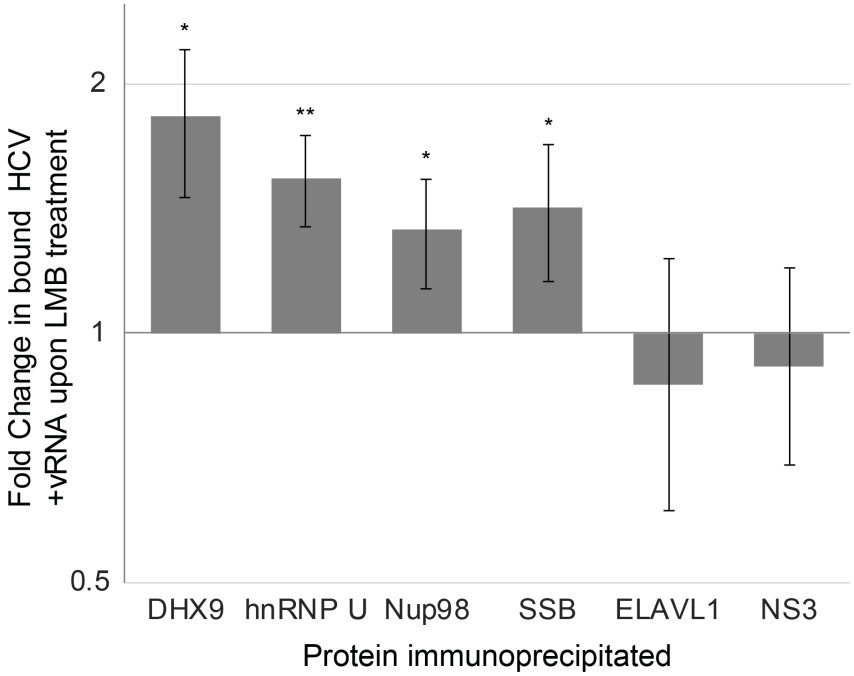


Figure 4.19: Nuclear accumulation of HCV +vRNA increases its interaction with host NFs.

Huh7.5 cells infected with HCV (72h p.i.) and treated with LMB or drug vehicle (3 hours) were crosslinked to preserve protein/RNA complexes. Cell lysates were incubated with beads coupled to a negative control IgG (α-GFP) or beads coupled to Nup98, DHX9, hnRNP U, SSB, ELAVL1 or NS3 (positive control IgG) specific antibodies. RNA present in immunoprecipitated complexes and total cellular RNA (10% input) was reverse transcribed and used as template in qPCR reactions containing primers specific to the HCV +vRNA. HCV +vRNA bound to proteins was normalized to input samples. Fold change of +vRNA bound to different nuclear host factors (indicated in x-axis) in the presence of LMB when compared to drug vehicle is indicated in the y-axis. Error bars indicate standard deviation (N = 3 biological replicates). Results from 3 biological replicates were submitted to ANOVA followed by Tukey HSD tests. Adjusted p-values for Tukey HSD in pairwise comparisons between LMB and drug vehicle treated cells are indicated as ** $<$ 0.01 $<$ * $<$ 0.05. Data produced by Dr. Zhongjing Su.

The +vRNA that enters the nuclei may need to interact with a given NF before it can be exported back to the cytoplasm, becoming trapped in its absence. On the other hand, a +vRNA that enters the nuclei may need to interact with a NF for nuclear retention and get quickly exported back out in the absence of this binding partner.

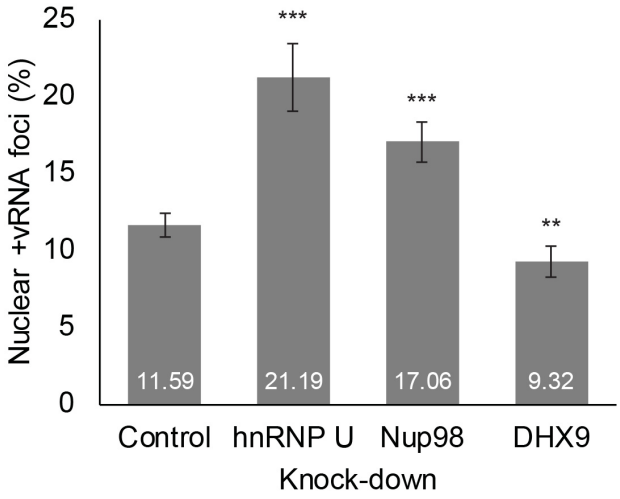
To test this possibility we depleted Nup98, DHX9 or hnRNP U in HCV-infected cells and used FISH to detect the localization of HCV +vRNA, -vRNA and host GAPDH mRNA. Depletion of these host NFs had no effect on the nuclear localization of the HCV -vRNA or host GAPDH mRNA but it altered the nuclear pool of HCV +vRNA (Figure 4.20). Nup98 depletion increased the nuclear localization of HCV +vRNA by $\approx$ 6% (Figure 4.20a). This could be due to the function of Nup98 in RNA export (Powers et al., 1997; Blevins et al., 2003; Ren et al., 2010) or it could reflect a role for nucleoplasmic Nup98 in the metabolism of HCV +vRNA.

Depletion of hnRNP U in HCV-infected cells almost doubled the accumulation of nuclear +vRNA, from $\approx$ 11% to 21% of the total +vRNA in the infected cell (Figure 4.20a). This is in agreement with a previously described role for hnRNP U in the nuclear export of vRNAs from another virus, HIV-1 (Valente and Goff, 2006). HnRNP U interaction with the HIV-1 vRNAs was implicated in vRNA translocation from the nucleus to the cytoplasm and vRNA translatability. Since we observed similar effects upon hnRNP U depletion in HCV-infected cells (Figure 4.3b and 4.20a), it is possible this protein plays a similar role during HCV infection.

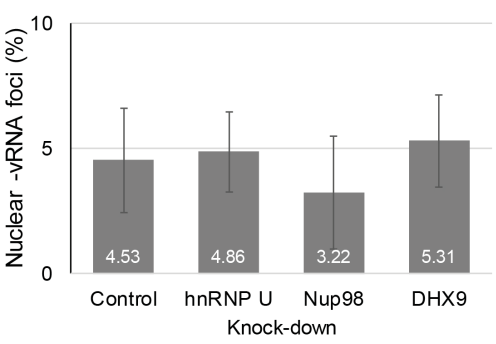
DHX9 depletion, on the hand, causes a slight decrease ($\approx$ 2%) in HCV +vRNA nuclear localization (Figure 4.20a), this is consistent with the role DHX9 plays in the life cycle of another virus, HIV-1. In HIV-1 infected cells DHX9 regulates vRNA replication and translation but not nuclear export (Bolinger et al., 2010; Lorgeoux et al., 2012). It is possible that the HCV +vRNA that enters the nuclei of infected cells is remodeled by DHX9, allowing its nuclear retention. This remodeling could regulate the interaction of other host NFs with the +vRNA. A similar mechanism was previously described in subsection 3.3.5, where the interaction of Nup98 with several cellular mRNAs is decreased upon DHX9 depletion.

4.3.9 Zika virus +vRNA is also present in the nuclei of infected cells.

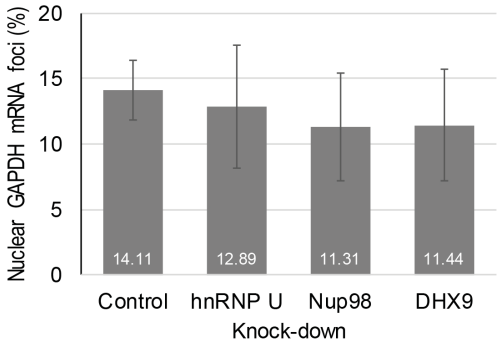
In the results described above we have identified a nuclear pool of +vRNA in HCV-infected cells, determined that this nuclear +vRNA is dynamically transported by NTFs, and that, inside the nuclei of infected cells, the +vRNA can interact with host NFs that participate on the HCV life cycle. Thus, we next decided to explore the possibility that something similar might occur in another virus from the *Flaviviridae* family, such as Zika virus (ZIKV). This virus has been declared a public health



(a) HCV +vRNA foci localizaton



(b) HCV -vRNA foci localizaton



(c) GAPDH mRNA foci localization

Figure 4.20: Depleting host NFs alters the nuclear localization of HCV +vRNA.

Huh7.5 cells were coinfected with HCV and lentivirus encoding shRNAs directed against Nup98, DHX9, hnRNP U, or a non-mammalian control sequence for three days. Localization of HCV +vRNAs, -vRNAs or GAPDH mRNAs was determined by FISH visualized by epifluorescence microscopy. Plot shows the average percentage of HCV +vRNA (**4.20a**), -vRNA (**4.20b**) or GAPDH mRNA (**4.20c**) foci with nuclear localization in the presence of indicated shRNA-mediated NF depletion. Error bars indicate standard deviation (N = 3 biological replicates). Results from 3 biological replicates were submitted to ANOVA followed by Tukey HSD tests. Adjusted p-values for Tukey HSD in pairwise comparisons between mock depleted cells and cells with indicated target knock-down are indicated as *** $<$ 0.001 $<$ ** $<$ 0.01.

emergency by the World Health Organization, but not much is known about its cell biology.

Production of infectious ZIKV in tissue culture cells achieves its best titers when produced in Vero cells for 24 to 48 hours (Kumar et al., 2016). ZIKV can infect numerous human cell lines (Chan et al., 2016), but A549 cells are easily infected by ZIKV with a short time course of infection (Kumar et al., 2016). In ZIKV-infected A549 cells, significant cell death starts to occur between 48 and 72 hours after infection (Figure 4.21). Accordingly, for all experiments described below, ZIKV produced in Vero cells was used to infect A549 cells for no more than 24 hours, unless otherwise stated.

We began by observing the localization of the ZIKV +vRNA during a time course of infection in A549 cells. We used two sets of probes that hybridize with different regions of the ZIKV +vRNA in a FISH assay (5' and 3' probes, see section 2.11). Experimental procedures and image analysis were performed as previously described for Huh7.5 cells infected with HCV (subsection 4.3.4). Probes hybridizing with the GAPDH mRNA were used as control. We observed ZIKV +vRNA and GAPDH mRNA localization in A549 cells at 3, 12, 18, 24 and 48 hours after ZIKV infection (Figure 4.22). GAPDH mRNA localization remained unchanged during the infection time course (Figure 4.22b). ZIKV +vRNA appears dispersed through the cytoplasm of cells during early infection time points ($\approx$ 3 to 18 hours p.i.). Later in infection ($\approx$ 24 to 48 hours p.i.) the ZIKV +vRNA accumulates in a region of the cytoplasm, adjacent to the NE.

During early infection time points (3 and 12 hours p.i.), $\approx$ 20 to 25% of ZIKV +vRNA foci are nuclear. The percentage of nuclear ZIKV +vRNA foci starts to decrease at 18 hour p.i., reaching a minimum of 2.5% at 48 hours p.i. (Figure 4.22a). Therefore, similarly to what we observed with HCV-infected cells, ZIKV-infected cells have a pool of nuclear +vRNA. The proportion of nuclear ZIKV +vRNA also changes during a time course of infection, initially increasing and then decreasing, as seen for HCV infected cells (Figure 4.22a and 4.8). Of note, the ZIKV +vRNA accumulates in the nuclei of infected cells early in infection (as early as three hours p.i.), decreasing its nuclear pool at later infection time points. This would argue against the possibility of +vRNAs getting trapped within the nuclei of cells during mitosis, since we expect very few cells to have divided within such a short time fram (3 hours). this would more likely indicate that ZIKV infected cells may also have a dynamic nuclear pool of actively transported vRNA molecules.

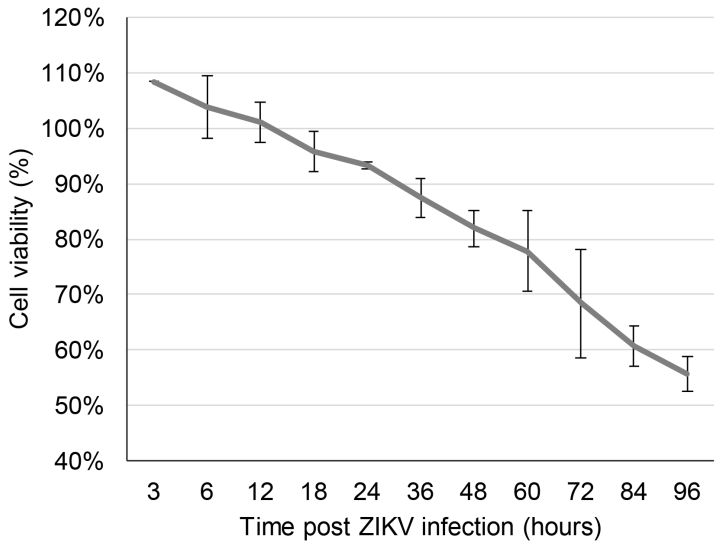
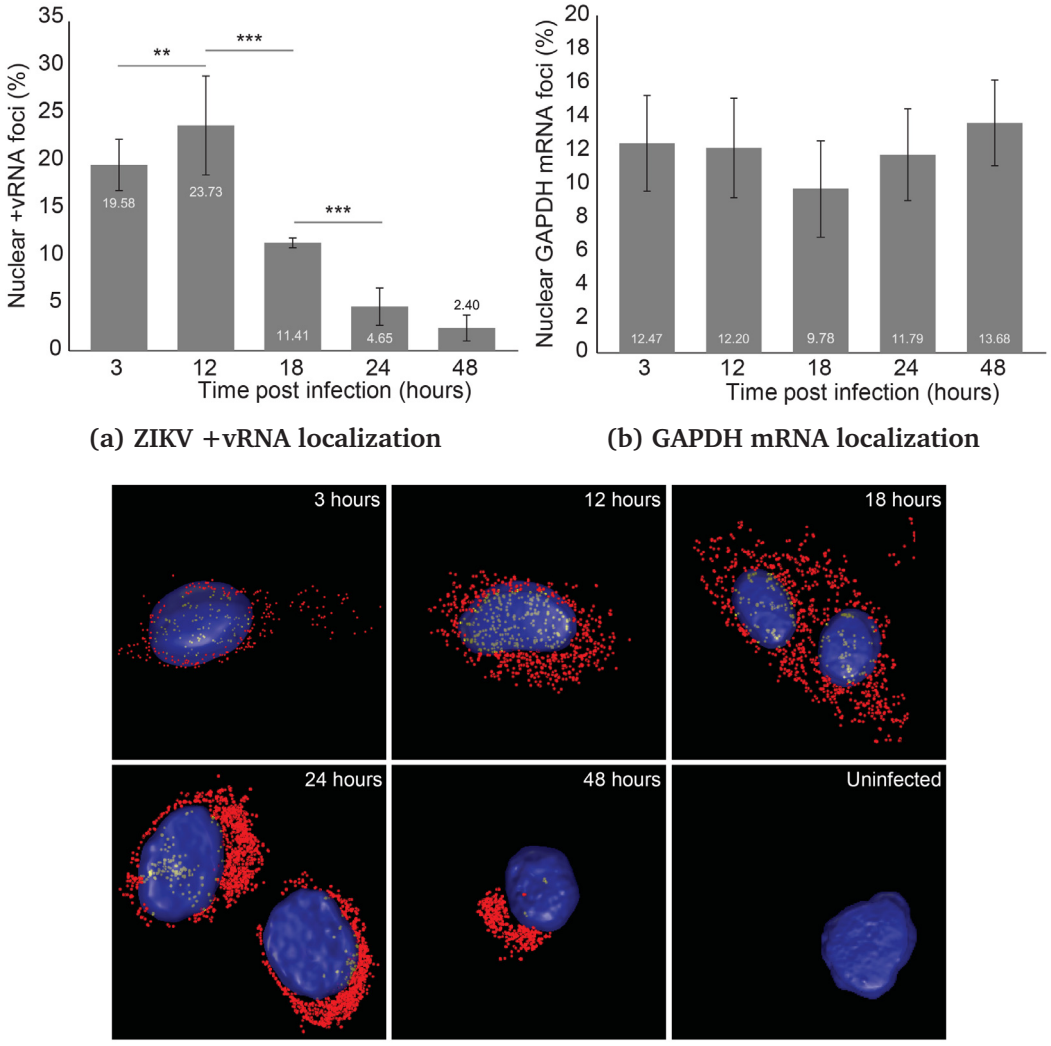


Figure 4.21: ZIKV infection causes a time dependent decrease in cell viability.

A549 cells were infected with ZIKV for 3 to 96 hours. Cell viability cells was evaluated by mitochondrial dehydrogenases activity. Average values for cell viability are expressed as a percentage of the viability of uninfected A549 cells. Error bars indicate standard deviation (N = 3 biological replicates). Data produced in collaboration with Brett Roughead.



(a) ZIKV +vRNA localization **(b) GAPDH mRNA localization**

(c) Example images of ZIKV +vRNA foci localization

Figure 4.22: Nuclear localization of ZIKV +vRNA changes during the time course of infection.

A549 cells infected with ZIKV for 3, 12, 18, 24 or 48 hours were evaluated by FISH visualized by epifluorescence microscopy. Plots show the average percentage of ZIKV +vRNA foci (**4.22a**), or GAPDH mRNA foci (**4.22b**) with nuclear localization in ZIKV-infected cells at the indicated time points post infection. Error bars indicate standard deviation (N = 3 biological replicates). Results from 3 biological replicates were submitted to ANOVA followed by Tukey HSD tests. Adjusted p-values for Tukey HSD in pairwise comparisons between each infection time point and the one directly preceding it are indicated as *** $< 0.001 <$ ** < 0.01. Example representative images of the mapped nuclei (shown as DAPI isosurfaces in blue) and ZIKV +vRNA foci (red foci - outside, yellow foci - inside DAPI isosurface) are shown in **4.22c**.

4.3.10 NTF inhibitors alter the ZIKV +vRNA localization and its life cycle

In order to determine if the ZIKV +vRNA is dynamically transported across the NE by NTFs, we attempted to alter the nuclear localization of ZIKV +vRNA with NTF inhibitors (LMB, IPZ, and IVM). One of these inhibitors (Ivermectin - IVM) has been shown to prevent replication of other Flaviviruses when added up to 14 hours after infection, by inhibiting viral NS3 helicase unwinding activity (Mastrangelo et al., 2012). Therefore, we chose to treat ZIKV-infected cells with NTF inhibitors at 15 hours p.i. to prevent blocking viral replication. LMB, IPZ or IVM were used to block CRM1-dependent nuclear export, Kapβ1 or Kapα/β-dependent nuclear import, respectively. The effect of NTF inhibitor treatments was evaluated three hours later, at 18 hours p.i., a time point in ZIKV infection when $\approx$ 11% of the +vRNA foci are intranuclear. Treatment of ZIKV-infected A549 cells with these inhibitors showed no effect on cell viability (Figure 4.23) or the localization of GAPDH mRNA foci (Figure 4.24c).

CRM1-dependent nuclear export block (LMB) nearly doubled the proportion of nuclear ZIKV +vRNA foci ()from $\approx$ 11% to almost 20%). Nuclear import block via Kapβ1 or Kapα/β inhibitors (IPZ or IVM), on the other hand, decreased the ZIKV +vRNA foci with nuclear localization by more than half (from $\approx$ 11% to $\approx$ 5%) (Figure 4.24b and 4.24a). These results indicate that the nuclear pool of ZIKV +vRNA is dynamically transported in and out of the nuclei of infected cells, at least in part in a CRM1 and Kapβ1-dependent manner.

Since inhibiting NTFs alters the nuclear localization of ZIKV +vRNA foci, we next explored the effects of NTF inhibitors on the ZIKV life cycle. We assessed the ZIKV life cycle by focusing on 2 steps of infection: +vRNA genome replication and production of infectious virus. Unfortunately, +vRNA translation could not be assessed at this time due to a lack of available antibodies specific for ZIKV proteins. Viral genome replication was determined by quantifying intracellular ZIKV +vRNA by qPCR and production of infectious virus was examined by quantifying the infectious titer of ZIKV produced through re-infection assays.

To this end, A549 cells were infected with ZIKV for 15 hours and treated with inhibitors (LMB, IPZ, IVM) for an additional 3 hours. We quantified intracellular ZIKV +vRNA to identify any changes in viral genome replication in the presence of NTF inhibitors. CRM1-dependent nuclear export block by LMB had no effect in the levels of ZIKV +vRNA in infected cells. This could indicate that CRM1-dependent nuclear export is not required during the +vRNA replication step of the ZIKV life cycle. Although it is also possible treatment duration was not sufficient to allow detection of replication changes. IPZ and IVM treatment, blocking nuclear import via Kapβ1 or

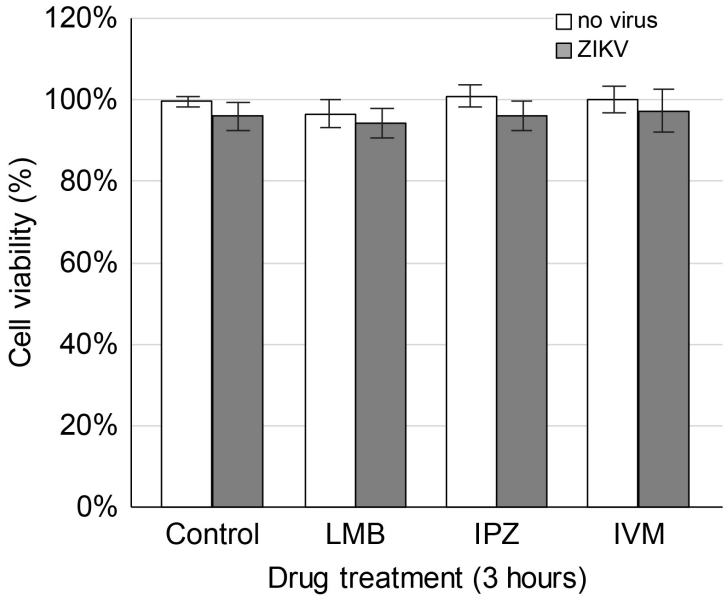
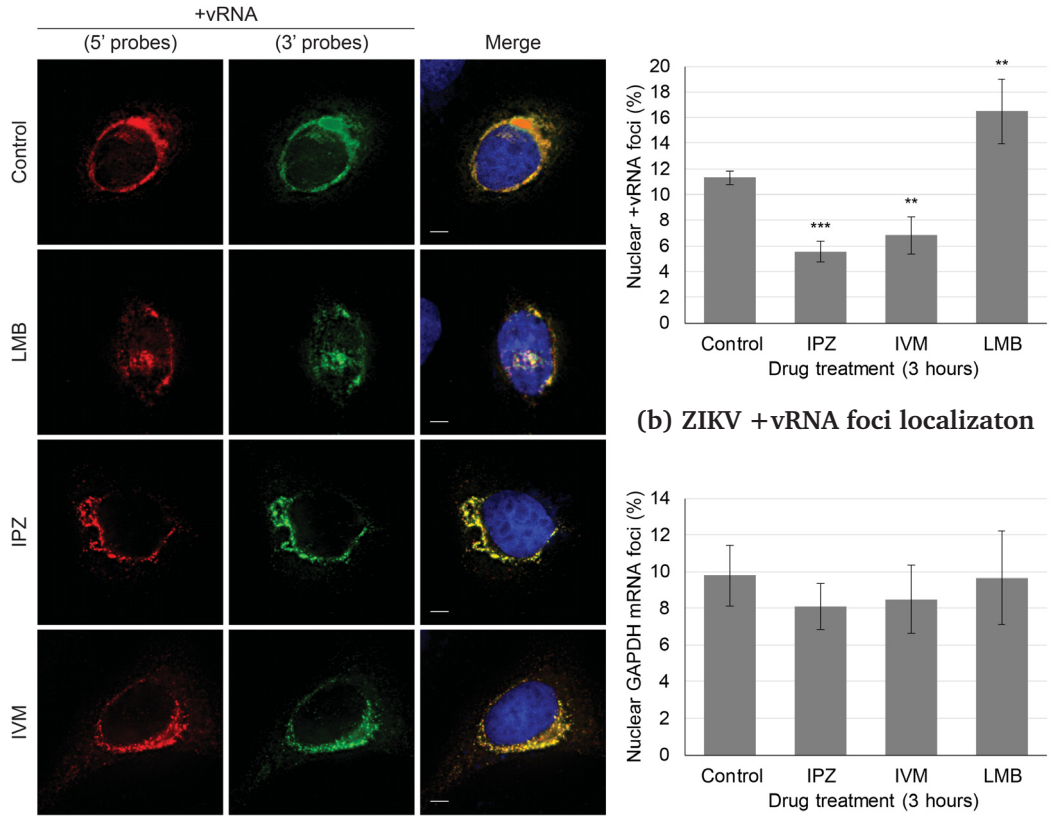


Figure 4.23: NTF inhibitors do not affect the viability of ZIKV-infected cells.

A549 cells were infected with ZIKV for 15 hours and treated with drug vehicle (control), LMB, IPZ or IVM for an additional 3 hours (total 18 hours of infection). Cell viability was evaluated by mitochondrial dehydrogenases activity. Average values for cell viability are expressed as a percentage of the viability of A549 cells (uninfected, in the absence of inhibitors or drug vehicle). Error bars indicate standard deviation (N = 3 biological replicates).



(a) ZIKV +vRNA FISH

(b) ZIKV +vRNA foci localizaton

(c) GAPDH mRNA foci localization

Figure 4.24: NTF inhibitors alter the nuclear localization of ZIKV +vRNA.

A549 cells were infected with ZIKV for a total of 18 hours before sample collection. Infected cells weretreated with LMB, IPZ or IVM for 3 hours at 15 hours p.i.. **4.24a** Localization of ZIKV +vRNA (5' and 3' probes) was evaluated by FISH visualized by epifluorescence microscopy. DNA was stained with DAPI (blue). Scale bars, 5 μm. N $\geq$ 3 biological replicates, representative FISH images are shown. Plots show the average percentage of ZIKV +vRNA foci (**4.24b**), or GAPDH mRNA foci (**4.24c**) with nuclear localization in ZIKV-infected cells in the presence of indicated NTF inhibitors. Error bars indicate standard deviation (N $\geq$ 3 biological replicates). Results from $\geq$ 3 biological replicates were submitted to ANOVA followed by Tukey HSD tests. Adjusted p-values for Tukey HSD in pairwise comparisons between inhibitor and drug vehicle treated (control) cells are indicated as *** $<$ 0.001 $<$ ** $<$ 0.01.

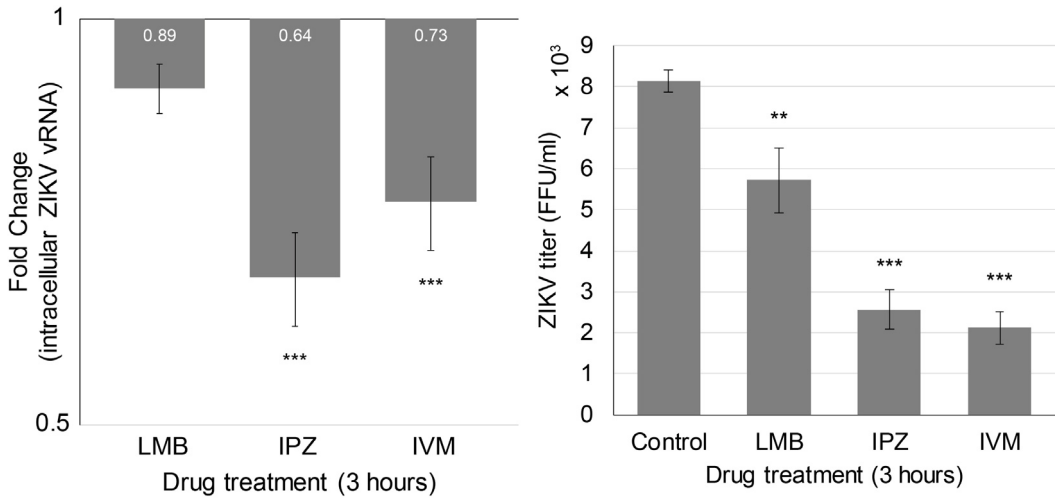
Kapα/β, caused $\approx$ 30% decrease in ZIKV +vRNA levels (Figure 4.25a). This might indicate that Kapα/β or Kapβ1-dependent transport promotes +vRNA replication or stability during ZIKV infection. However, at this time we cannot determine if this effect is due directly to changes in the nuclear pool of ZIKV +vRNA or if it reflects other effects of these NTF inhibitors on host or viral proteins that participate in ZIKV +vRNA replication and stability.

We also evaluated the effect of NTF inhibitors in the production of infectious virus by ZIKV-infected cells. Cells infected with ZIKV for 15 hours were repeatedly washed and culture media was replaced with media containing nucleocytoplasmic transport inhibitors (LMB, IPZ or IVM) or drug vehicle (control). Three hours later newly produced ZIKV in the media was collected for re-infection of naive A549 cells, allowing quantification of the infectious titer of virus produced. We observe a decrease in the production of infectious ZIKV following treatment of infected cells with all NTF inhibitors (LMB, IPZ, and IVM). CRM1-dependent nuclear export inhibition decreased the production of infectious virus by $\approx$ 25% (Figure 4.25b). Inhibition of Kapβ1 or Kapα/β-dependent nuclear import shows a more significant effect, decreasing virus titers by $\approx$ 75%.

The effect of NTF inhibitors on the ZIKV life cycle may be a direct consequence of altering the nuclear localization of the +vRNA. However, we cannot discount the possibility that other effects of these inhibitors may also contribute to these phenotypes. For instance, it is possible that ZIKV proteins may also enter the nuclei of infected cells, as described for HCV (Bonamassa et al., 2015) and other Flaviviruses (Lopez-Denman and Mackenzie, 2017); thus, inhibiting NTFs might disrupt their nuclear transport. Another possibility is that transport at the ZIKV replication factories (Cortese et al., 2017) might also be regulated by NTFs, as has been shown for the HCV MW (Neufeldt et al., 2013, 2016), and NTF inhibitors could disrupt transport at these structures as well.

4.3.11 NFs and NTFs that are exploited by HCV are also required for ZIKV infection.

In the sections above we have described many similarities observed between HCV- and ZIKV-infected cells. Cells infected with either virus show a dynamic pool of nuclear +vRNA that changes during the time course of infection and whose localization can be altered by NTF inhibitors (subsection 4.3.5, 4.3.6, 4.3.9 and 4.3.10). Additionally, NTF inhibitors disrupt the life cycle of both viruses, ultimately decreasing viral production by infected cells (subsection 4.3.7 and 4.3.10). Therefore, we next decided to explore if other parallels could be identified between HCV- and ZIKV-infected cells with regards to their host factor requirements. We focused on



(a) intracellular ZIKV vRNA **(b) ZIKV infectious titer**

Figure 4.25: NTF inhibitors affect the ZIKV life cycle.

A549 cells were infected with ZIKV for 15 hours and treated with indicated inhibitors for an additional three hours before sample collection. **4.25a** Changes in average intracellular ZIKV +vRNA were evaluated by qPCR. Sample were normalized to house-keeping gene transcripts and fold-change in ZIKV +vRNA in cells treated with inhibitors relative to cells treated with drug vehicle was calculated. **4.25b** Average Infectious titers of ZIKV, in focus-forming units per mL of medium (FFU/mL), produced in the presence of inhibitors (LMB, IPZ and IVM) or drug vehicle (Control) were determined by indirect immunofluorescence microscopy. For both panels, error bars indicate standard deviation ($N \geq 3$ biological replicates). Results from ≥ 3 biological replicates were submitted to ANOVA followed by Tukey HSD tests. Adjusted p-values for Tukey HSD in pairwise comparisons between control sample (drug vehicle) and inhibitor treated samples are indicated as *** $< 0.001 <$ ** < 0.01.

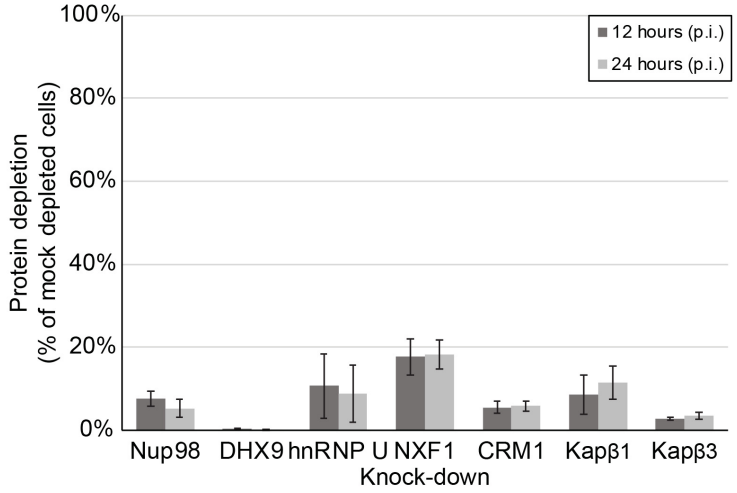
the previously described NTFs (CRM1, NXF1, Kapβ1 and Kapβ3) and NFs (DHX9, hnRNP U and Nup98) shown to affect the HCV life cycle and the nuclear localization of HCV +vRNA in infected cells.

In order to achieve this A549 cells were transduced with lentivirus encoding shRNAs targeting CRM1, NXF1, Kapβ1, Kapβ3, DHX9, hnRNP U or Nup98 three days before ZIKV infection. Since the time course of ZIKV infection showed that the proportion of nuclear +vRNA varied significantly from 12 to 24 hours after infection (Figure 4.22a), we decided to further investigate both time points (12 and 24 hours) of infection in these assays. This infection protocol differs from the one previously described for HCV-infected cells, where cells were simultaneously infected with HCV and shRNA-encoding lentiviruses. In this case, given the infection of cells with shRNA-encoding lentiviruses three days before ZIKV infection, at the time of ZIKV addition target proteins were already depleted in the host cells. Depletion of target mRNAs and proteins was verified by qPCR and western blot (Figure 4.26). No significant change in cell viability was observed under these experimental conditions, except for NXF1 depletion, which caused a decrease of $\approx$ 30% in cell viability (Figure 4.26).

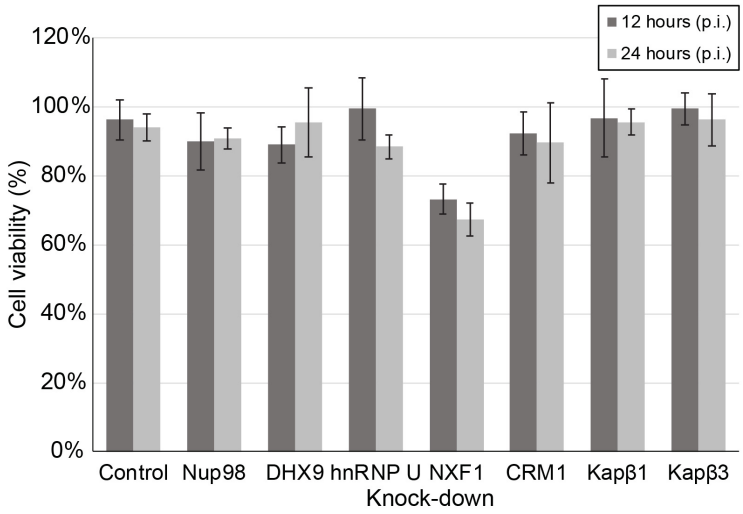
Due to the depletion of host factors that are likely participants in the ZIKV life cycle at the time of infection, we observed a significant reduction in the number of ZIKV-infected cells under these conditions. The shRNA-mediated decrease of most host factors described above (excepting DHX9 and Kapβ3) significantly hindered the capacity of ZIKV to establish an infection in these cells.

Depletion of Nup98, NXF1 and CRM1 showed the most deleterious effects in ZIKV infection. At 24 hours post infection over 70% of cells are infected with ZIKV in the presence of control shRNA. However, the percentage of ZIKV-infected cells remained below 10% for Nup98 or NXF1-depleted cells and 20% for CRM1-depleted cells (24 hours p.i.). Similarly, Kapβ1 and hnRNP U depletion also hindered ZIKV infection, decreasing the percentage of ZIKV-infected cells by more than half (Figure 4.27). This likely indicates these host factors are necessary for the establishment of productive ZIKV infection in the cell (Figure 4.27).

The decrease in ZIKV infection of cells depleted of these host factors raises the possibility that any remaining infected cells might represent a subpopulation where host factor depletion was partial or absent. Thus, any further analysis of these cells would not accurately represent the effect depletion of these host factors may have in the ZIKV life cycle or +vRNA nuclear localization. Nevertheless, drawing a parallel between HCV and ZIKV-infected cells, it seems likely that the host NFs Nup98 and hnRNP U, and the NTFs CRM1 and Kapβ1, are exploited by HCV and ZIKV during infection. However, further evidence will be required to confirm these speculations and determine the role played by these host factors in ZIKV infected cells.



(a) protein depletion



(b) cell viability

Figure 4.26: NTFs and NFs depletion in ZIKV-infected A549 cells.

A549 cells were infected with lentiviruses encoding shRNAs directed against Kapβ1, Kapβ3, CRM1, NXF1, Nup98, DHX9, hnRNP U or a non-mammalian control sequence for three days before being infected with ZIKV for 12 or 24 hours. **4.26a** Protein depletion following shRNA expression was evaluated by western blot with antibodies against indicated proteins. Samples were normalized to a load control and average percentage of protein depletion relative to cells expressing a control shRNA calculated. **4.26b** The viability of ZIKV-infected cells expressing shRNAs was evaluated via mitochondrial dehydrogenases activity. Average viability of cells depleted of the indicated proteins is expressed as a percent of uninfected A549 cells (no lentivirus, no ZIKV infection). For both panels, errors bars indicate standand deviation (N = 3 biological replicates). Part of these data were produced in collaboration with Brett Roughead.

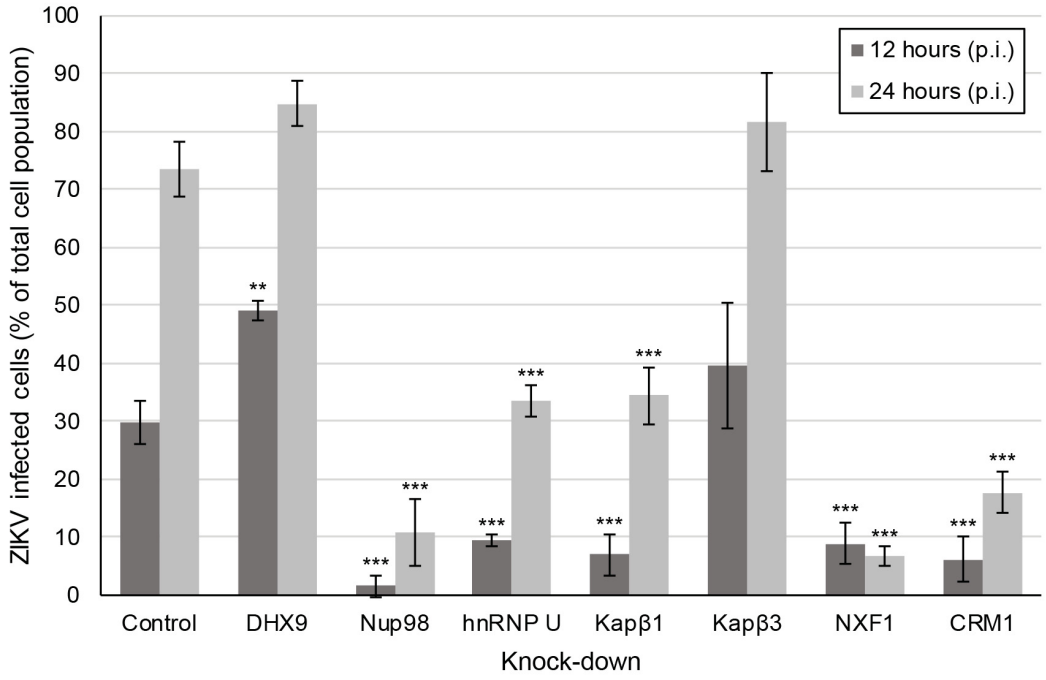


Figure 4.27: Depletion of host NFs or NTFs decreases ZIKV infection.

A549 cells were infected with lentiviruses encoding shRNAs directed against Kapβ1, Kapβ3, CRM1, NXF1, Nup98, DHX9, hnRNP U or a non-mammalian control sequence for three days before being infected with ZIKV for 12 or 24 hours. Monitoring of the virus spread (ZIKV-positive cells) was evaluated by FISH (probes targeting the ZIKV +vRNA) visualized by epifluorescence microscopy. Plot shows the average percentage of A549 cells depleted of the indicated proteins infected by ZIKV 12 or 24 hours after virus addition. Error bars indicate standard deviation (N = 3 biological replicates). Results from 3 biological replicates were submitted to ANOVA followed by Tukey HSD tests. Adjusted p-values for Tukey HSD in pairwise comparisons between target knock-down and mock depleted cells are shown as $*** < 0.001 < * < 0.05$.

4.4 Discussion

In this work we identified a nuclear pool of vRNAs that were dynamically transported by NTFs in and out of the nuclei of HCV-infected cells. Disrupting the nucleocytoplasmic transport of the vRNA affected the HCV life cycle, decreasing production of infectious virus. These nuclear vRNAs interacted with host NFs that participate in the HCV life cycle but maintain their nuclear localization in infected cells. Depleting these host NFs also altered the nuclear pool of vRNA and disrupted the life cycle of the virus. Finally, we showed that another virus from the *Flaviviridae* family, Zika virus, also had a nuclear pool of vRNA, dynamically transported by NTFs. Disruption of nucleocytoplasmic transport in ZIKV-infected cells not only altered the nuclear localization of the vRNA, but also disrupted the viral life cycle, decreasing the production of infectious virus. Depletion of several NTFs and host NFs that participate in HCV infection severely decreased ZIKV infection of host cells, indicating these host factors were also exploited by ZIKV during infection.

Several previous publications have established the importance of numerous host factors in the life cycle of positive-strand RNA viruses (Nagy and Pogany, 2011; Reid et al., 2015; Li and Nagy, 2011). All stages of the life cycle of these viruses are thought to occur in the cytoplasm of infected cells. However, many host factors exploited by these viruses are nuclear host proteins that in many cases do not relocate to the cytoplasm in infected cells. The nuclear proteins hijacked by these viruses often bind RNA and control several features of RNA biology and gene expression (i.e. splicing, transport, stability and translation) in the host cell. RNA viruses employ the functions of numerous host RNA-binding proteins to facilitate various steps in the virus life cycle, especially +vRNA translation and replication (Lloyd, 2015). On the other hand, vRNA and viral proteins, have also been observed in the nuclei of infected cells. This nuclear pool of viral proteins are thought to affect host gene expression, creating a more favorable environment for viral replication (Bonamassa et al., 2015; Lopez-Denman and Mackenzie, 2017).

Since these viruses can utilize both the nuclear and cytoplasmic environments, it is not surprising that they have an intricate relationship with components of the nucleocytoplasmic transport pathway. RNA viruses disrupt host nucleocytoplasmic transport to achieve a variety of objectives, such as: allowing viral proteins and vRNAs to enter the nucleus; relocating nuclear proteins exploited by these viruses to cytoplasmic sites of replication or translation; and interfering with host antiviral defenses and gene expression regulation (Mettenleiter, 2016; Flather and Semler, 2015; Bonamassa et al., 2015; Kuss et al., 2013; Yarbrough et al., 2014).

Recently, a novel mechanism by which cytoplasmic RNA viruses hijack components of the nucleocytoplasmic transport pathway has been described (Levin et al.,

2014b; Neufeldt et al., 2013, 2016). Neufeldt and collaborators have identified multiple components of the nuclear transport machinery that interact with HCV proteins and accumulate in the virus-induced membranous web in the cytoplasm of infected cells (Neufeldt et al., 2013). These Nups (possibly as part of assembled or partially assembled NPCs) create a selective barrier at the MW, controlling the movement of NTF-cargo complexes. This selective transport limits the access of proteins that negatively impact viral replication (such as innate immunity intracellular pattern recognition receptor), while allowing traffic of viral proteins and nuclear host factors exploited by the virus (proteins containing an NLS sequence) (Levin et al., 2014b; Neufeldt et al., 2016). These findings also raise the intriguing possibility that the Nups recruited to the cytoplasm might influence membrane curvature at virus created organelle structures. However, this idea remains to be functionally demonstrated (Neufeldt, 2014). Depletion of Nups and NTFs in HCV-infected cells decreases extracellular HCV vRNA and the titer of infectious virus produced, indicating they play a role in HCV assembly. Some of the Nups and NTFs recruited to the MW in HCV-infected cells could also have non-transport related functions in the viral life cycle. Nup98 and Nup153, two RNA-binding Nups (Ren et al., 2010; Ball et al., 2004), affect not only virus assembly, but also HCV vRNA replication (Neufeldt et al., 2013).

The results presented in this chapter further confirm a role for NTFs in HCV assembly. Depletion of CRM1, Kapβ1, Kapβ3, or treatment of HCV-infected cells with LMB, IPZ, or IVM all had deleterious effects in the production of infectious virus by HCV-infected cells. However, depletion of these NTFs did not decrease HCV replication, as measured by intracellular vRNA levels (subsection 4.3.7). Nup98 depletion on the other hand, affected vRNA replication (decreasing it in half) and HCV assembly ($\approx$ three-fold decrease) (subsection 4.3.2), confirming this Nup might participate in these two stages of the HCV life cycle. We also show that disrupting NTFs with LMB, IPZ, or IVM can affect the life cycle of another virus, ZIKV, decreasing the production of infectious virus by infected cells (Figure 4.25).

Since HCV proteins containing NLSs can interact with, and be transported by, NTFs recruited to the MW (Levin et al., 2014b; Neufeldt et al., 2013, 2016), it is not surprising that they also interact with the nucleocytoplasmic transport machinery at NEs. Accordingly, HCV proteins (core, NS3, NS5A) have been previously observed in the nuclei of infected cells (Bonamassa et al., 2015). We have shown here that nuclear localization of vRNAs can also be observed in HCV-infected cells (chapter 4). The presence and relevance of these nuclear vRNAs has never been characterized or discussed in the literature (Shulla and Randall, 2015; Fiches et al., 2016; Ramanan et al., 2016). Similarly, viral proteins from Flaviviruses (ZIKV, DENV, WNV, JEV, etc.) also contain NLSs and can be observed in the nuclei of infected cells (Tay

et al., 2013; Uchil et al., 2006; Zhu et al., 2016; Li et al., 2017).

The presence of viral proteins and possibly vRNAs in the nuclei of cells infected with positive-strand RNA viruses (Uchil et al., 2006) would hint towards a possible role for the nuclear environment in their life cycle. These observations also indicate that the nuclear environment might be more likely to play a role early in infection, before the cytoplasmic membrane rearrangements induced by these viruses have been fully established.

ZIKV infection, for example, is susceptible to nucleocytoplasmic transport disruption early in infection. Treating cells with IVM an hour before ZIKV addition will prevent infection establishment in various human cell lines (Barrows et al., 2016). IVM has been shown to almost completely inhibit the replication of several other Flaviviruses, especially when present within the first 14 hours of infection, loosing its inhibitory effect on replication after approximately 22 hours of infection (Mastrangelo et al., 2012). We observed something similar in ZIKV-infected cells, when cells were treated with IVM or IPZ 15 hours after infection ZIKV replication decreased 30-40% (Figure 4.25).

Interestingly, recent work characterizing the cytoplasmic membrane rearrangements induced by ZIKV for the formation of membranous replication factories (RFs) shows that these structures appear between 12 and 24 hours after ZIKV infection (Cortese et al., 2017). This coincides with the time frame when we see the proportion of ZIKV +vRNA foci with nuclear localization decrease from over 20% to less than 5% of the vRNA present in the cell. It is tempting to speculate that early in infection, before ZIKV RFs have been established, nuclear import of +vRNA may be necessary to promote its replication. In fact, this idea has been previously proposed for other viruses from the Flavivirus genus (Uchil et al., 2006). This hypothesis is also in agreement with our results showing that addition of IVM or IPZ 15 hours after ZIKV infection will significantly decrease the pool of nuclear +vRNA, decreasing +vRNA replication and the production of infectious virus.

The capsid proteins (also known as C or core) of many flaviviruses contain NLSs that allow their transport into the nuclei of infected cells (DENV, WNV, JEV and Kunjin virus) (Zhu et al., 2016). The ZIKV C protein also contains a NLS that is conserved in different viral strains and allows exogenously expressed ZIKV C protein to localize to the nuclei of yeast cells (Li et al., 2017; Zhu et al., 2016). Therefore, it is possible that ZIKV C could act as an adaptor protein for the nucleocytoplasmic transport of the +vRNA in infected cells. Upon formation of the RFs however, the ZIKV C protein and vRNA would be sequestered to these membranous structures decreasing their transport in and out of the nuclei of infected cells.

It is possible the nuclear environment also plays a role during early HCV infection. Other than the envelope proteins, two other viral proteins that bind the vRNA

are found in the HCV virion: core and NS3 (Lussignol et al., 2016). These viral proteins contain multiple NLSs and can interact with Kapα/β and Kapβ3. Similarly to what we have described for the HCV +vRNA (subsection 4.3.4), these proteins have also been observed in the nuclei of HCV infected cells (Bonamassa et al., 2015). HCV core can be observed in the nuclei of infected cells as early as 20 minutes after viral entry (Cerutti et al., 2011), up to approximately eight hours after HCV infection (Levin et al., 2014a). LMB treatment of HCV infected cells can increase the nuclear pool of core (48 hours p.i.) (Cerutti et al., 2011) and prevent the establishment of infection if added to cells within the first 14 hours of HCV infection (Levin et al., 2014b).

During synchronized HCV infection, virus entry and the initial rounds of translation occur within the first 4 to 10 hours after infection. This is followed by continued vRNA translation and increased replication at 10 to 24 hours post infection. Viral assembly and egress follow at around 30 to 48 hours after infection (Levin et al., 2014b; Shulla and Randall, 2015). Rearrangement of host cytoplasmic membranes for formation of the membranous web includes the creation of double membrane vesicles (DMVs) and multi membrane vesicles (MMVs). Fomation of these vesicles requires the presence of sufficient amounts of viral proteins in the cytoplasm of HCV-infected cells. Thus, DMVs only appear in significant numbers in HCV-infected cells at 24 hours p.i.; while MMVs, formed by the HCV envelope proteins necessary in viral assembly, appear in infected cells at 36 to 48 hours p.i. (Romero-Brey et al., 2012).

A nuclear role for HCV core and CRM1-dependent transport within the first 14 hours of infection will precede the formation of the MW and likely coincide with the stages of vRNA replication and/or translation, similarly to what was described above for a possible role for the nuclear environment early in ZIKV infection. Unfortunately, the FISH assay used here prevented determination of the HCV +vRNA localization in early infection time points. However, since core strongly binds the HCV +vRNA we could speculate that these molecules might enter the nuclei of infected cells together. Also, we were still able to increase the nuclear accumulation of HCV +vRNA treating infected cells with LMB at 48 or 72 hours p.i. (subsection 4.3.6), similarly to what has been shown for LMB-dependent nuclear accumulation of core later in infection (Cerutti et al., 2011).

One possible reason for the +vRNA of viruses to enter the nuclei of infected cells could be to recruit the function of nuclear host factors necessary for its life cycle. In accordance with this idea, several host nuclear RNA-binding proteins have been shown to participate in the life cycle of positive-strand RNA viruses. The interaction of these host NFs with +vRNAs provides the RNA chaperone/remodeling capabilities necessary for +vRNA replication and translation in infected cells (Lloyd,

2015).

Many of these nuclear RNA-binding host factors participate in the HCV life cycle (Li et al., 2015; Bartenschlager, 2013). Here, we have shown that the RNA-binding NFs Nup98, DHX9, and hnRNP U are among the host factors that can interact with the HCV +vRNA in infected cells (Figure 4.1). Depletion of these NFs has deleterious effects in HCV infection, especially decreasing +vRNA replication (Nup98 or DHX9 depletion) or translation (hnRNP U depletion) (subsection 4.3.2). This agrees with the hypothesis that the +vRNA might enter the nuclei of infected cells in order to gain access to the functions of nuclear host factors necessary for its life cycle. Interestingly, while some host NFs are recruited to the cytoplasm of infected cells later in infection (Nup98 and ELAVL1) others retain their nuclear localization (DHX9, hnRNP U and SSB) (subsection 4.3.3).

If the interaction between host RNA-binding NFs and +vRNAs can occur in the nuclei of infected cells, these interactions could also affect the nuclear localization of the HCV +vRNA. An interaction between the HCV +vRNA and a NF may be necessary to either allow its nuclear retention or to promote its nuclear export. This being the case, depleting these NFs should also alter the nuclear pool of HCV +vRNA. As expected, depletion of DHX9, hnRNP U or Nup98 all alter the proportion of HCV +vRNA with nuclear localization in infected cells (Figure 4.20a).

DHX9 depletion caused a decrease in the nuclear pool of HCV vRNA. We can hypothesize that DHX9-mediated remodeling of the imported vRNP in the nucleus might remove proteins bound to the imported vRNA allowing its nuclear retention. This remodeling could also allow nuclear retention of the +vRNA by promoting interactions between the +vRNA and other NFs, such as the DHX9 interactors Nup98 or hnRNP U. We have shown in a previous chapter (chapter 3) that the interaction of Nup98 with host mRNAs is decreased upon DHX9 depletion (subsection 3.3.5). It is possible that the same mechanism used by the Nup98-DHX9 complex to regulate host mRNAs in the nuclei of cells is exploited by the +vRNA in the nuclei of HCV-infected cells. Similarly, protein complexes containing DHX9 and hnRNP U have been previously described as participating in the regulation of the metabolism of several host mRNAs (Weidensdorfer et al., 2009; Lee and Pelletier, 2016; Chu et al., 2012; Francisco-Velilla et al., 2016), and could be exploited by HCV.

An increase in the nuclear pool of +vRNA upon hnRNP U depletion is also in agreement with a previously described role for hnRNP U in the nuclear export of HIV-1 vRNAs from the nuclei of infected cells (Valente and Goff, 2006). HnRNP U interaction with the HIV-1 vRNAs was implicated in their translocation from the nucleus to the cytoplasm, promoting vRNA stability, and translatability. It is possible hnRNP U might play a similar role during HCV infection. Our results showing that hnRNP U depletion leads to nuclear accumulation of the HCV +vRNA (Fig-

ure 4.20a), along with severe decrease in viral protein levels (Figure 4.3), likely indicating decreased translation of the +vRNA, are in accordance with this hypothesis.

Similar to hnRNP U, depletion of Nup98 in HCV-infected cells leads to the accumulation of HCV +vRNA in the nuclei (Figure 4.20a). Again drawing a parallel with the role played by Nup98-DHX9 in the metabolism of host mRNAs (chapter 3), we see that Nup98 binds to DHX9 increasing its RNA-dependent ATPase activity and facilitating efficient processing and release of mRNAs in the cells (Jarmoskaite and Russell, 2014; Jankowsky, 2011). Consequently, depletion of Nup98 leads to an increase in the binding of mRNAs to DHX9 (subsection 3.3.5). Therefore, we can imagine that if the Nup98-DHX9 complex behaves similarly with regards to its interaction with the HCV +vRNA, upon Nup98 depletion the interaction of HCV +vRNAs with DHX9 is increased due to a slower rate of DHX9-dependent +vRNA remodeling; thus, increasing the pool of nuclear +vRNA that is retained in the nuclei of infected cells.

If we assume that the interaction between these host RNA-binding NFs and the HCV +vRNA can occur in the nuclei of infected cells, we may hypothesize that increasing the nuclear pool of HCV +vRNAs by disrupting nucleocytoplasmic transport would increase the interaction of host RNA-binding NFs with the +vRNA. Given that HCV +vRNA nuclear export is at least in part CRM1-dependent, inhibiting CRM1-dependent nuclear export leads to the accumulation of +vRNA in the nuclei of HCV-infected cells (Figure 4.12). As predicted, this increase in the nuclear pool of +vRNA, consequently increases the interaction of host RNA-binding NFs with the HCV +vRNA (Figure 4.19). This modulation in the interaction between NFs and HCV +vRNA by shifting the amount of nuclear +vRNA further indicates a nuclear interaction between +vRNA and NFs in HCV-infected cells.

These nuclear interactions between RNA-binding NFs and the +vRNA could facilitate the recruitment of NFs to replication factories in the cytoplasm of infected cells. One hypothesis is that the HCV +vRNA might enter the nuclei of infected cells early during infection, possibly even accompanied by the HCV protein core. In the nuclei, this vRNP complex is remodeled by DHX9 or other RNA helicases, promoting the interaction between NFs (such as Nup98, ELAV1, and others) and the HCV +vRNP. Given that many of these proteins (Nup98, ELAV1, core, etc.) can interact with CRM1, it is possible that the vRNP complexes can be exported by a CRM1-dependent pathway. If by this point during infection the cytoplasmic membrane rearrangements promoted by the virus have started to form, these NFs could become trapped by their interactions with viral proteins or the +vRNA, now sequestered to these membranous structures. This could represent one of the mechanisms by which Nup98, ELAV1 or other exploited RNA-binding NFs are recruited

to the cytoplasm of HCV-infected cells (Neufeldt et al., 2013; Isken et al., 2007; Li et al., 2014; Lloyd, 2015).

This role played by RNA-binding NFs in HCV replication can be expanded to other Flaviviruses. The 3' UTRs of viral genomes from the *Flavivirus* genus are complex structures that contain cis-acting elements required for translation, circularization and replication. Accordingly, this regulatory RNA region binds several nuclear host factors known to participate in RNA metabolism, including many that also interact with the HCV +vRNA, such as SSB, PTB, YB-1, NF90, DHX9, and NF45. RNA replication in Flaviviruses is dependent on many of these co-opted host NFs. They remodel the +vRNA to an alternate conformation based on complex long-range binding interactions that form a new panhandle structure. This panhandle structure positions 5' and 3' termini of the genome close together to promotes RNA replication. Binding of a complex of dsRNA-binding factors (NF90, DHX9, and NF45) to the 3' UTR occurs in concert with the formation of the panhandle to promote RNA replication. The possibility exists that these factors may also bind the 5' UTR and stabilize the long-range looped structure similarly to their role in HCV +vRNA circularization (Lloyd, 2015).

The ability of HCV to exploit the functions of the Nups and Kaps to create an environment conducive to its replication and assembly may also represent a mechanism widely used by other positive-strand RNA viruses. Increased amounts of cytoplasmic Nup98-containing foci (that may represent NPCs) have also been observed in cells infected with hepatitis A virus and dengue virus (Neufeldt et al., 2013). Consistent with this observation, electron microscopy studies have reported increased levels of annulate lamellae, a cytoplasmic structure containing Nups, in hepatitis A virus infected cells (Marshall et al., 1996), as well as cells infected with Japanese Encephalitis virus (Wang et al., 1997), Rubella virus (Kim and Boatman, 1967) and Rous Sarcoma virus (Courington and Vogt, 1967). Therefore, Nups and NTFs may represent a conserved target of positive-strand RNA viruses.

Since the role of host NTFs and RNA-binding NFs in the life cycle of positive-strand RNA viruses is not restricted to HCV (Lloyd, 2015; Fullam and Schröder, 2013; Neufeldt et al., 2013; Neufeldt, 2014), we also attempted to explore their function in the life cycle of ZIKV. To this end, we depleted the nuclear import factors Kapβ1 and Kapβ3, the nuclear export factors CRM1 and NXF1, or the RNA-binding NFs Nup98, DHX9, and hnRNP U in A549 cells three days before ZIKV infection. Depletion of the RNA-binding NFs Nup98 or hnRNP U significantly reduced the percentage of ZIKV-infected cells in the population (Figure 4.27). Similarly, depletion of the NTFs Kapβ1, CRM1 or NXF1 also prevented the establishment of ZIKV infection (Figure 4.27), as has been observed with the Kapα/β inhibitor IVM (Barrows et al., 2016). Therefore, we were unable to properly assess the effect of depleting these

host factors in the ZIKV +vRNA localization or the consequences of this depletion to the viral life cycle. Nonetheless, it is clear that CRM1, NXF1, Kapβ1, Nup98 and hnRNP U have important roles in the ZIKV life cycle, since ZIKV infection cannot be properly established in the absence of these factors. Since these host factors are either nuclear proteins or proteins that participate in nucleocytoplasmic transport, the dependence of ZIKV on these proteins for infection points towards a possible role for the nuclear environment in the life cycle of this virus as well.

CHAPTER 5

Perspectives

The nucleus is a hallmark of eukaryotic cells. It is encapsulated by the nuclear envelope, which separates the genome from the cytoplasm. Nuclear pore complexes perforate the NE, establishing selectively permeable channels between the nucleus and cytoplasm. NPCs, along with nuclear transport factors, regulate macromolecule transport across the NE. Situated at the interface between the nucleus and the cytoplasm, NPCs are positioned to participate in a vast number of cellular processes, including the regulation of gene expression and RNA metabolism (Ibarra and Hetzer, 2015; Burns and Wente, 2014; Ptak and Wozniak, 2016)). In higher eukaryotes, nucleoporins also take leave of the NE and function within the nucleoplasm (Rabut et al., 2004). Nucleoplasmic Nups associate with chromatin and other nuclear factors, to regulate gene expression and RNA metabolism (Capelson et al., 2010; Kalverda et al., 2010; Ptak et al., 2014), including immune response genes (Enninga et al., 2002; Faria et al., 2006; Light et al., 2013; Panda et al., 2014; Satterly et al., 2007), and influencing chromatin organization (Kalverda and Fornerod, 2010; Ptak and Wozniak, 2016; Liang and Hetzer, 2011).

Nup98 is among the most studied Nups exhibiting intranuclear localization (Griffis et al., 2002; Iwamoto et al., 2010; Radu et al., 1995). Despite the growing evidence linking Nup98 to the regulation of chromatin structure, gene expression, and RNA metabolism, gaps remain in our knowledge of the mechanisms by which Nup98 affects these processes. In Chapter 3 of this dissertation, we identify novel interactions between Nup98 and two classes of well known regulators of gene expression and RNA metabolism, DExH/D-box helicases (Jarmoskaite and Russell, 2014; Bourgeois et al., 2016; Hardwick and Luisi, 2013) and hnRNPs (Han et al., 2010; Piccolo et al., 2014). Of these Nup98 interactors, one of the strongest binding partners is the RNA helicase DHX9. We demonstrate that Nup98 and DHX9 interact in the nucleoplasm in an RNA facilitated manner. The interaction of Nup98 with DHX9 regulates its localization, RNA-binding and ATPase activity, which ultimately influences gene expression *in vivo* via DHX9-regulated transcription and splicing. Nup98 and DHX9 bind interdependently to similar gene loci and their transcripts, where Nup98 functions as a co-factor that regulates DHX9. This function of Nup98 could potentially also regulate the cellular activities of other RNA helicases, indicating a novel mechanism by which Nup98 can regulate gene expression away from NPCs.

Nup98 and other nucleoporins have been previously identified as a host factors affecting the life cycle of many viruses (Lopez-Denman and Mackenzie, 2017; Mettenleiter, 2016; Neufeldt, 2014; Flather and Semler, 2015). The same is true for several other nuclear host proteins, such as DExH/D-box helicases (Fullam and Schröder, 2013; Ahmad and Hur, 2015) and hnRNPs (Li and Nagy, 2011; Lloyd, 2015), which are also appropriated by viruses to participate in their life cycles. Therefore, it is not surprising that these proteins, along with other components of

the nucleocytoplasmic transport pathways of eukaryotic cells are exploited during several viral infections, not only to promote viral replication, but also to alter host gene expression and immune response. In chapter 4, we explore how the Nup98-DHX9 complex, as well as other host nuclear RNA-binding proteins and NTFs, can be exploited by positive-strand RNA viruses of the *Flaviviridae* family. We show that the viral RNA of ZIKV and HCV can be found in the nuclei of infected cells, despite their cytoplasmic replication cycle. Altering nucleocytoplasmic transport can bias the localization of these +vRNAs, change the interaction between +vRNA and host nuclear RNA-binding proteins, and affect the viral life cycle. These data point towards an unexplored role for the nuclear environment during positive-strand RNA virus infection.

In this chapter, I discuss the impact of our observations on the understanding of how Nups, NTFs, DExH/D-box helicases, and other RNA-binding nuclear factors can collaborate to regulate gene expression and RNA metabolism in the cell. We also explore how these host factors get exploited during viral infections, including by viruses with cytoplasmic replication. Finally, we'll explore the dual role these host factors play in infected cells, as host factors that participate in the viral life cycle and as regulators of innate immunity in the host cell.

5.1 Nups and DExH/D-box helicases in gene expression regulation

Many recent publications have made clear the importance of nucleoporins in the regulation of gene expression, from yeast to human cells (recently reviewed in Ptak and Wozniak, 2016). The yeast model system has led the way in understanding the role of NPC and chromatin interactions in the regulation of gene expression. Followed by studies in mammalian cells that have also detected Nups in association with active genes and epigenetic chromatin modifiers such as histone deacetylases and histone acetyltransferases. These interactions have been shown to occur at the NE, and at gene loci positioned inside the nucleus. The interaction of nucleoporins with chromatin has been most studied at sites of active transcription. However, it is still largely unclear what the molecular bases for these interactions are, and why current active genes bind Nups (Ptak and Wozniak, 2016). One intriguing example of Nups regulating the transcription of active genes comes from the X chromosome in male Drosophila cells.

In flies, dosage compensation of the X-linked genes leads to a two-fold upregulation in males in comparison to females. This is mediated by the dosage compensation complex, also known as Male Specific Lethal complex (MSL), due to the

male specific lethality phenotype upon loss-of-function of its major components. The Drosophila MSL is a ribonucleoprotein complex and is composed of two functionally redundant long non-coding RNAs, roX1 and roX2, and at least five proteins, MSL-1 (male-specific lethal 1, scaffolding protein), MSL-2 (male-specific lethal 2, RING finger protein), MSL-3 (male-specific lethal 3, chromodomain protein), MOF (males absent on the first, histone acetyltransferase) and MLE (maleless, RNA helicase), whose human homolog is the helicase DHX9. Similarly to what was described in chapter 3, the ATPase activity of MLE is also required for the transcriptional enhancement promoted by the MSL complex at the male X chromosome (Morra et al., 2008).

The MSL complex decorates the male X chromosome which is also hyperacetylated at histone H4K16 (Georgiev et al., 2011). Interestingly, H4K16 is also hyperacetylated at transcriptionally active, nucleoporin-interacting genes, including those bound by Nup50, Nup62, and Nup98 throughout the genome of Drosophila cells (Kalverda et al., 2010). Nup153 and Mtor have also been shown to bind genome-wide large domains in Drosophila cells, which are heavily enriched for active transcription marks, including H4K16 hyperacetylation. Up to 75% of the male X chromosome is enriched in these Nup associated domains, with the majority of these loci also showing an enrichment for members of the MSL complex (Vaquerizas et al., 2010). Nucleoporins could possibly coordinate the expression of a number of genes genome-wide, including the extreme case of chromosome-wide upregulation of X-linked genes underlying dosage compensation in Drosophila males. In these cells the hyperactivation of the X chromosome seems to correlate with the formation a dosage-compensated chromosomal domain in a Nup and MSL complex dependent manner (Georgiev et al., 2011). Intriguingly, the Drosophila MSL complex member MOF has been shown to co-purify with components of the NPC including Nup153, Mtor, and Nup98. RNAi-mediated knockdown of Nup153 and Mtor led to reduction of the typical MSL binding pattern on the X chromosome, suggesting a role for Nups the dosage compensation process (Mendjan et al., 2006).

We can draw a parallel between the overlapping role played by Nups and MSL, containing MLE (DHX9) and MOF, in dosage compensation in Drosophila cells and our observations on the Nup98-DHX9 complex regulating gene expression in a CBP-dependent manner in human cells (chapter 3). In both cases the Nups and helicases localize to the same gene loci, and depletion of Nups decreases the interaction of the helicase with the target genes (Mendjan et al., 2006 and Figure 3.24). The MLE-dependent transcriptional activation in Drosophila dosage compensation requires its ATPase activity (Morra et al., 2008). Similarly, the CBP-dependent transcriptional activation induced by the Nup98-DHX9 complex also requires DHX9's ATPase activity (Figure 3.26a). Finally, the histone acetyltransferase MOF has been shown

to interact with MLE (Morra et al., 2011), as well as with several Nups (Mendjan et al., 2006) in Drosophila cells. Likewise, the histone acetyltransferase CBP has been shown to interact with DHX9 (Aratani et al., 2001; Nakajima et al., 1997) and Nup98 (Kasper et al., 1999) in human cells. In chapter 3, we have shown that Nup98, via its N-terminal domain (containing FG/GLFG repeats and a RBD), can interact with DHX9, via its N- and C-terminal domains (containing two dsRBDs followed by the MTAD and an OB-fold followed by an RGG-box, respectively) in an RNA facilitated manner. This interaction between Nup98 and DHX9 can regulate DHX9's ATPase activity, and consequently the CBP-dependent DHX9 trancriptional activition. Comparably, in Drosophila cells, MLE requires its N and C-terminal dsRBDs and RGG-box, as well as the interaction with specific RNAs in order to promote its MOF-dependent transcriptional activation, which also requires MLE's ATPase activity (Morra et al., 2011). Some Nups shown to interact with the MSL complex (containing MLE and MOF) contain a FG-repeats domains (Nup153 and Nup98), which may be able to induce MLE's ATPase activity, as was observed for the N-terminal domain of Nup98 (containing FG/GLFG repeats) with DHX9 in human cells, followed by a histone acetyltransferase dependent transcriptional induction.

It is possible that this could represent a conserved mechanism by which nucleoporins can regulate gene expression via their interaction with RNA helicases. The Nup-helicase interaction could modulate the ATPase activity of the helicase, and consequently affect its role in transcriptional regulation, which could be dependent on histone modifiers, such as HATs and HDACs (Fuller-Pace, 2006). Further study is necessary to verify this model, characterizing the effect of Nup98 interaction on the function of other DEAD/DExH-Box RNA helicases known to participate in gene expression regulation (Bourgeois et al., 2016; Fuller-Pace, 2006) and identified as Nup98 interactors (Figure 3.1 and 3.2) could be a starting point of interest. Similarly, since many other Nups shown to regulate gene expression (e.g. Nup50, Nup62, Nup153, sPom121) are FG-Nups (Kalverda and Fornerod, 2010; Capelson et al., 2010; Franks et al., 2016), verifying if these Nups can also bind and regulate DEAD/DExH-Box RNA helicases would be an interesting next step in this project.

Characterizing the interaction of Nup-helicase complexes with epigenetic modifiers (i.e. histone acetylases and deacetylases) and identifying the effect these interactions have on the activity of a given histone modifier would be an important next step. Since the interaction of DHX9 with CBP has been previously characterized (Nakajima et al., 1997; Aratani et al., 2001), and give that we have shown Nup98 can bind DHX9 (subsection 3.3.3), an interesting starting point would be to determine how these three proteins interact with one another in a complex. Secondly, since we have seen that Nup98 can alter the transcriptional activity of the DHX9-CBP complex, we should further explore how these interactions affects the histone

acetyltransferase activity of CBP.

Lastly, we have not yet identified the mechanism by which this complex is targeted to specific genes. DHX9 has been shown to interact with transcription factors that direct its activity to specific target genes (for reviews see Fidaleo et al., 2016; Lee and Pelletier, 2016). DHX9 can also interact with double stranded DNA, and it has been shown to directly bind to the promoter of p16INK4A in a sequence specific manner. However, no genome-wide binding studies of DHX9 have been performed thus far. Comparing the genome-wide chromatin binding profile of Nup98 and DHX9 may provide interesting insights as to how the genes they regulate are targeted, either via specific sequences recognized by DHX9 itself, via DHX9 binding transcription factors, or through epigenetic marks recognized by chromatin modifiers interacting with the Nup98-DHX9 complex.

5.2 The dual roles of Nups and DExD/H-box RNA helicases in innate immunity and viral replication

In chapter 4, we discuss a possible role for the nuclear environment in the life cycle of positive-strand RNA viruses from the *Flaviviridae* family. Viral proteins and vRNAs from HCV and ZIKV are transported to the nuclei of infected cells via NTFs and the NPC. In the nucleus, vRNAs interact with host nuclear RNA-binding proteins that affect the viral life cycle. Many other RNA viruses have been described to exploit the nucleocytoplasmic pathway during infection. Similarly, most of these same viruses also hijack nuclear host RNA-binding proteins to participate in their life cycle. Although in chapter 4 we discuss the viral exploitation of the functions of these host factors in different stages of the virus life cycle, many of these host protein are also mediators of anti-viral innate immunity. The duality of the function of these proteins during viral infection suggests that these host factors are highly contested targets in the ongoing 'arms race' between viruses and the host immune system. In this chapter we will further explore how RNA viruses target these host factors during their life cycle and how that might provided the simultaneous benefit of disrupting the host anti-viral immune responses.

5.2.1 Nups and RNA helicases in antiviral innate immunity

Nucleoporin-dependent gene expression regulation plays an important role in antiviral innate immune response. On the other hand, the same subset of Nups is exploited during infection to support viral propagation. Among these Nups are those

linked to the regulation of gene expression, described above (section 5.1), including Nup98. (Ptak and Wozniak, 2016). Recent work by Pandas and collaborators has shown that Nup98 plays an essential antiviral role in Drosophila cells against human insect-borne viruses. In these cells, Nup98 was upregulated by viral infection and Nup98 depletion significantly enhanced viral infection. Upon viral infection Nup98 was shown to bind the promoter region of a subset of antiviral genes, whose expression was Nup98-dependent and induced by viral infection. This promoter-bound Nup98 primes virus-stimulated genes by regulating the occupancy of active RNA polymerase at these promoters poising them for rapid induction, during a coordinated, robust, and complex antiviral response to restrict human arboviruses in Drosophila (Panda et al., 2014). The interaction of Nup98 with the promoter of these genes is dependent on the transcription factor FoxK, and their expression upon viral infection requires both, Nup98 and FoxK in Drosophila cells (Panda et al., 2015).

The observations described above for Drosophila cells can be extended to the mammalian orthologs, Nup98 and FOXK1, which also have an antiviral function, promoting gene expression from promoters containing a IFN-β-stimulated response element (Panda et al., 2015). Nup98 has also been shown to promote the expression of IFN-β itself during antiviral immune response induced by the RIG-I-like receptor pathway in human cells (van der Lee et al., 2015). Additionally, Nup98 expression is induced in the presence of IFN-γ and IFN-α (Enninga et al., 2002) and IFN-γ-stimulated genes showing transcriptional memory depend on the presence of Nup98 for their transcriptional reactivation. In human cells depleted of Nup98, the marks of transcriptional memory present in the promoter of these genes (i.e. H3K4me2 and poised RNA polymerase II association) are lost, and the rate of transcriptional reactivation is reduced (Light et al., 2013). This interplay between Nup98, IFN, and IFN-stimulated genes during antiviral innate immune response hints towards the induction of a positive feedback loop between these factors under these conditions. Similarly, Nup96 expression is also induced by IFN-γ and IFN-α (Enninga et al., 2002) and Nup96 downregulation leads to impaired expression of IFN-γ and IFN-α stimulated genes (Faria et al., 2006).

Similarly to what has been described above for Nups, DExD/H-box RNA helicases also participate in the regulation of the gene expression program involved in antiviral innate immune response. Several DExD/H-box helicases can also contribute to anti-viral immunity, either by acting as sensors for viral nucleic acids or by facilitating downstream signaling events. In general, pathogens are detected by innate immune cells with the help of pattern recognition receptors (PRRs) that recognize conserved structures present in certain pathogen classes. The main groups of PRRs sensing viral RNA are endosomal Toll-like receptors (TLRs), the RIG-like

helicases (RLHs) and several DExD/H-box helicases (not part of the RIG-I family), including DDX3, DDX60, DDX41, DDX1, DHX9 and DHX36 (Fullam and Schröder, 2013). A general characteristic of anti-viral PRR signaling is the induction of type I interferons (IFN-α and IFN-β), whose potent anti-viral activity is largely mediated by interferon-stimulated genes that encode proteins with direct anti-viral functions (Schoggins et al., 2011).

DHX9 has been previously identified as a sensor for double-stranded RNA in myeloid cells (Zhang et al., 2011), and as a sensor for CpG DNA in plasmacytoid dendritic cells (pDCs). In these cells, DHX9 is required for the induction of IFN-α and TNF-α (Kim et al., 2010). Similarly, in myeloid DCs, type I IFN and pro-inflammatory cytokine production in response to dsRNA (poly I:C) is decreased upon DHX9 depletion (Zhang et al., 2011). On the other hand, DHX9 itself is an interferon-inducible gene (Sadler et al., 2009). The function of DHX9 during antiviral immune response is not restricted to its role as a sensor for double stranded nucleic acids. DHX9 also functions as a transcriptional regulator. It is recruited into PML nuclear bodies after IFN-α-stimulation where it associates with actively transcribing gene loci, enriched for RNA polymerase II and nascent transcripts, suggesting DHX9 has a role in regulating gene expression at these IFN-stimulated gene loci (Fuchsová and Hozák, 2002).

DHX9 and Nup98 share many similarities in the roles they might play during antiviral innate immune response, both show increased expression and relocalization within the nuclei in the presence of IFNs (Enninga et al., 2002; Sadler et al., 2009; Fuchsová and Hozák, 2002). Both are necessary for IFN induction itself (van der Lee et al., 2015; Kim et al., 2010; Zhang et al., 2011), as well as the expression of IFN-stimulated genes (Panda et al., 2015; Light et al., 2013; Fuchsová and Hozák, 2002). A next step of interest would be to investigate if Nup98-DHX9 act as a complex during these steps of antiviral innate immune response and if their roles are interdependent on one another.

5.2.2 Nups and RNA helicases are hijacked by viral infections

In the context of the evidence described above for one or more functions of DHX9 and Nup98 in antiviral immunity, it is very interesting to note that these proteins are actively recruited by many different viruses and required for their replication. This is true for many other Nups, NTFs and RNA helicases that not only have been shown to participate in viral recognition and anti-viral immunity, but they are also actively recruited by viruses to facilitate their replication cycles (Fullam and Schröder, 2013; Lopez-Denman and Mackenzie, 2017; Bonamassa et al., 2015; Yarbrough et al.,

2014; Le Sage and Mouland, 2013). Many viruses rely heavily on host RNA helicases to mediate RNA remodelling events that are part of their replication cycle or required for viral gene expression. Several DExD/H-box helicases have been identified as essential host factors for the replication of different viruses, including the major global health threats HIV and HCV. In many cases, the same RNA helicases that are exploited by viruses during their life cycle are also the ones implicated in anti-viral innate immune responses (Fullam and Schröder, 2013).

The role of DHX9 in the life cycle of HIV has been well characterized. DHX9 is required for HIV infection and appears to be involved in several different steps of the viral life cycle. DHX9 binds the HIV TAR leader RNA, enhancing transcription from the HIV-LTR. DHX9 also increases the unspliced HIV CTE- and RRE-containing mRNAs by releasing incompletely spliced HIV RNAs from spliceosomes, while facilitating their export out of the nucleus through an alternative nuclear export pathway mediated by NXF1 and Sam68. Additionally, DHX9 can facilitate translation of HIV RNAs by binding to their 5' UTR and together with HIV gag promote the switch from translation to vRNA incorporation into the newly assembling virions. Finally, DHX9 is incorporated into HIV virions via interactions with the HIV gag protein and HIV RNA. DHX9 present in virions is thought to participate in reverse transcription. As previously stated, DHX9 appears to be involved in virtually all steps regulating the expression of HIV genes as well as the assembly of new virions (Lorgeoux et al., 2012). Similarly, DHX9 has been shown to participate in the life cycle of many other retroviruses, flaviviruses, picornaviruses and influenza virus. In most of these cases, DHX9 binds to the highly structured 5' or 3' UTRs of the viral RNA genome, remodeling it to facilitate one or several of the steps in the virus life cycle, such as its replication, transcription, translation, or packaging (Ranji and Boris-Lawrie, 2010b).

Similarly, Nups also participate in the life cycle of many viruses (Ptak and Wozniak, 2016). Again using the retrovirus HIV-1 as an example, Nups and NTFs have been shown to participate not only in the nucleocytoplasmic transport of HIV-1 vRNAs and genomes, but also in the integration of the viral genome into the host cell genome. HIV-1 integration has been shown to occur in chromatin regions located at the periphery of the nucleus in close proximity to NPCs. This region contains host genes characterized by the presence of active transcription chromatin marks before viral infection, which are preferentially targeted by HIV-1 for integration (Marini et al., 2015). Several nucleoporins that participate in this process have been identified, such as Nup98, Nup358, Nup153 and Tpr (Di Nunzio et al., 2012; Marini et al., 2015; Lelek et al., 2015). Nup153 and Nup358 contain FG-repeats domains, both Nups are required for the nuclear import of HIV-1 in infected cells. Consequently, Nup153 or Nup358 depletion prevents HIV-1 import and integration into the host

genome (Di Nunzio et al., 2012; König et al., 2008; Lelek et al., 2015). Nup98 depletion, on the other hand, has no effect on the nuclear import of HIV-1, but it decreases viral integration by half (Di Nunzio et al., 2012; König et al., 2008). At NPCs, Nup153 anchors Tpr at the nuclear basket. The nuclear basket participates in chromatin organization, maintaining transcriptionally active chromatin adjacent to the NPCs (Krull et al., 2010). Tpr does not play a role in HIV-1 nuclear import, or integration into the host genome. However, depletion of Tpr reduces the pool of actively transcribed chromatin adjacent ot the NPC. Therefore, Tpr-depleted cells exhibited less HIV-1 integrations mapping to actively transcribed genes (Marini et al., 2015; Lelek et al., 2015). It is likely that through its ability to maintain chromatin in an active state adjacent to NPCs, Tpr creates an environment favorable for HIV integration at sites of active transcription. Once integrated at active genes close to the NPC, the HIV-1 DNA interacts with several Nups, including Tpr, Nup153, Nup98, Nup62, and other FG-Nups (recognized by the mAB414). The interaction of Nups with the HIV-1 DNA participates in the transcriptional regulation of the viral genome, and it coincides with the association of RNA Polymerase II and the TFs USF1 and p65/RelA to the viral DNA (Marini et al., 2015).

Along with HIV-1, many other viruses also manipulate or utilize Nups and the NPC to facilitate viral capsid uncoating and trafficking of their genome into the nucleus. Several viruses also exploit the NPC to decrease the efficiency of transport for immune signaling molecules or host mRNA transcripts. Adenovirus, HSV-1 and influenza virus all have been shown to both utilize and inhibit nuclear transport pathways during infection. These viruses replicate in the nucleus and interact with specific Nups for uncoating and nuclear import of their genome. Additionally, proteins encoded by these viruses bind components of the mRNA export machinery to decrease host mRNA export and increase viral RNA export (reviewed in Le Sage and Mouland, 2013; Yarbrough et al., 2014). Positive strand RNA viruses, that replicate in the cytoplasm, have also been shown to utilize and disrupt nucleocytoplasmic transport (reviewed in Lloyd, 2015; Lopez-Denman and Mackenzie, 2017). By targeting specific components of the nuclear transport machinery, these viruses can alter the host cell environment, facilitating viral propagation. Since several of the Nups targeted by viruses not only participate in nucleocytoplasmic transport, but also in gene expression regulation, including regulation of the gene expression programs involved in antiviral innate immunity, it is very likely that viral disruption of these Nups goes beyond altering nucleocytoplasmic transport, to directly affect host gene expression regulation.

Bibliography

Abdelhaleem, M. (2004). Do human RNA helicases have a role in cancer? *Biochimica et biophysica acta*, 1704(1):37–46.

Afgan, E., Baker, D., van den Beek, M., Blankenberg, D., Bouvier, D., Čech, M., Chilton, J., Clements, D., Coraor, N., Eberhard, C., Grüning, B., Guerler, A., Hillman-Jackson, J., Von Kuster, G., Rasche, E., Soranzo, N., Turaga, N., Taylor, J., Nekrutenko, A., and Goecks, J. (2016). The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. *Nucleic acids research*, 44(W1):W3–W10.

Aggarwal, A. and Agrawal, D. K. (2014). Importins and exportins regulating allergic immune responses. *Mediators of inflammation*, 2014:476357.

Ahmad, S. and Hur, S. (2015). Helicases in Antiviral Immunity: Dual Properties as Sensors and Effectors. *Trends in biochemical sciences*, 40(10):576–85.

Alber, F., Dokudovskaya, S., Veenhoff, L. M., Zhang, W., Kipper, J., Devos, D., Suprapto, A., Karni-Schmidt, O., Williams, R., Chait, B. T., Sali, A., and Rout, M. P. (2007). The molecular architecture of the nuclear pore complex. *Nature*, 450(7170):695–701.

Allen, N. P., Huang, L., Burlingame, A., and Rexach, M. (2001). Proteomic analysis of nucleoporin interacting proteins. *J Biol Chem*, 276(31):29268–29274.

Alpert, T., Herzel, L., and Neugebauer, K. M. (2016). Perfect timing: splicing and transcription rates in living cells. *Wiley Interdisciplinary Reviews: RNA*.

Amorim, B. R., Okamura, H., Yoshida, K., Qiu, L., Morimoto, H., and Haneji, T. (2007). The transcriptional factor Osterix directly interacts with RNA helicase A. *Biochemical and biophysical research communications*, 355(2):347–51.

Anderson, S. F., Schlegel, B. P., Nakajima, T., Wolpin, E. S., and Parvin, J. D. (1998). BRCA1 protein is linked to the RNA polymerase II holoenzyme complex via RNA helicase A. *Nat Genet*, 19(3):254–256.

Antonin, W., Ellenberg, J., and Dultz, E. (2008). Nuclear pore complex assembly through the cell cycle: Regulation and membrane organization. *FEBS Letters*, 582(14):2004–2016.

Antonin, W., Franz, C., Haselmann, U., Antony, C., and Mattaj, I. W. (2005). The Integral Membrane Nucleoporin pom121 Functionally Links Nuclear Pore Complex Assembly and Nuclear Envelope Formation. *Molecular Cell*, 17(1):83–92.

Ao, Z., Jayappa, K. D., Wang, B., Zheng, Y., Wang, X., Peng, J., and Yao, X. (2012). Contribution of host nucleoporin 62 in HIV-1 integrase chromatin association and viral DNA integration. *The Journal of biological chemistry*, 287(13):10544–55.

Arai, T., Kano, F., and Murata, M. (2015). Translocation of forkhead box O1 to the nuclear periphery induces histone modifications that regulate transcriptional repression of ¡i¿PCK1¡/i¿ in HepG2 cells. *Genes to Cells*, 20(4):340–357.

Aranda, S., Mas, G., and Di Croce, L. (2015). Regulation of gene transcription by Polycomb proteins. *Science advances*, 1(11):e1500737.

Aratani, S., Fujii, R., Oishi, T., Fujita, H., Amano, T., Ohshima, T., Hagiwara, M., Fukamizu, A., and Nakajima, T. (2001). Dual Roles of RNA Helicase A in CREB-Dependent Transcription. *Molecular and Cellular Biology*, 21(14):4460–4469.

Aratani, S., Oishi, T., Fujita, H., Nakazawa, M., Fujii, R., Imamoto, N., Yoneda, Y., Fukamizu, A., and Nakajima, T. (2006). The nuclear import of RNA helicase A is mediated by importin-alpha3. *Biochem Biophys Res Commun*, 340(1):125–133.

Aravind, L. and Koonin, E. V. (1999). G-patch: a new conserved domain in eukaryotic RNA-processing proteins and type D retroviral polyproteins. *Trends in biochemical sciences*, 24(9):342–4.

Argasinska, J., Zhou, K., Donnelly, R. J., Hay, R. T., and Lee, C.-G. (2004). A functional interaction between RHA and Ubc9, an E2-like enzyme specific for Sumo-1. *Journal of molecular biology*, 341(1):15–25.

Arts, G. J., Fornerod, M., and Mattaj, I. W. (1998). Identification of a nuclear export receptor for tRNA. *Current biology : CB*, 8(6):305–14.

Atkinson, C. (2012). Structure and Organization of Nuclear Pore Proteins. *Student Theses and Dissertations*.

Babu, J. R., Jeganathan, K. B., Baker, D. J., Wu, X., Kang-Decker, N., and van Deursen, J. M. (2003). Rae1 is an essential mitotic checkpoint regulator that cooperates with Bub3 to prevent chromosome missegregation. *The Journal of cell biology*, 160(3):341–53.

Bachi, A., Braun, I. C., Rodrigues, J. P., Panté, N., Ribbeck, K., von Kobbe, C., Kutay, U., Wilm, M., Görlich, D., Carmo-Fonseca, M., and Izaurralde, E. (2000). The C-terminal domain of TAP interacts with the nuclear pore complex and promotes export of specific CTE-bearing RNA substrates. *RNA*, 6(1):136–158.

Bader, G. D. and Hogue, C. W. (2003). An automated method for finding molecular complexes in large protein interaction networks. *BMC Bioinformatics*, 4:2.

Bai, X.-T., Gu, B.-W., Yin, T., Niu, C., Xi, X.-D., Zhang, J., Chen, Z., and Chen, S.-J. (2006). Trans-repressive effect of NUP98-PMX1 on PMX1-regulated c-FOS gene through recruitment of histone deacetylase 1 by FG repeats. *Cancer Research*, 66(9):4584–4590.

Balch, W. E., Dunphy, W. G., Braell, W. A., and Rothman, J. E. (1984). Reconstitution of the transport of protein between successive compartments of the Golgi measured by the coupled incorporation of N-acetylglucosamine. *Cell*, 39(2 Pt 1):405–16.

Ball, J. R., Dimaano, C., and Ullman, K. S. (2004). The RNA binding domain within the nucleoporin Nup153 associates preferentially with single-stranded RNA. *RNA (New York, N.Y.)*, 10(1):19–27.

Barrows, N., Campos, R., Powell, S. T., Prasanth, K., Schott-Lerner, G., Soto-Acosta, R., Galarza-Muñoz, G., McGrath, E., Urrabaz-Garza, R., Gao, J., Wu, P., Menon, R., Saade, G., Fernandez-Salas, I., Rossi, S., Vasilakis, N., Routh, A., Bradrick, S., and Garcia-Blanco, M. (2016). A Screen of FDA-Approved Drugs for Inhibitors of Zika Virus Infection. *Cell Host & Microbe*, 20(2):259–270.

Bartenschlager, R. (2013). *Current Topics in Microbiology and Immunology Hepatitis C Virus : From Molecular Virology to Antiviral Therapy*, volume 369 of *Current Topics in Microbiology and Immunology*. Springer Berlin Heidelberg, Berlin, Heidelberg.

Belote, J. M. and Lucchesi, J. C. (1980). Control of X chromosome transcription by the maleless gene in Drosophila. *Nature*, 285(5766):573–5.

Belov, G. A., Lidsky, P. V., Mikitas, O. V., Egger, D., Lukyanov, K. A., Bienz, K., and Agol, V. I. (2004). Bidirectional increase in permeability of nuclear envelope upon poliovirus infection and accompanying alterations of nuclear pores. *Journal of virology*, 78(18):10166–77.

Bennett, C. B., Lewis, L. K., Karthikeyan, G., Lobachev, K. S., Jin, Y. H., Sterling, J. F., Snipe, J. R., and Resnick, M. A. (2001). Genes required for ionizing radiation resistance in yeast. *Nature Genetics*, 29(4):426–434.

Beran, R. K. F., Serebrov, V., and Pyle, A. M. (2007). The serine protease domain of hepatitis C viral NS3 activates RNA helicase activity by promoting the binding of RNA substrate. *The Journal of biological chemistry*, 282(48):34913–20.

Bianchi, A., Crotta, S., Brazzoli, M., Foung, S. K. H., and Merola, M. (2011). Hepatitis C virus e2 protein ectodomain is essential for assembly of infectious virions. *International journal of hepatology*, 2011:968161.

Bischoff, F. R. and Görlich, D. (1997). RanBP1 is crucial for the release of RanGTP from importin beta-related nuclear transport factors. *FEBS letters*, 419(2-3):249–54.

Bischoff, F. R. and Ponstingl, H. (1991). Catalysis of guanine nucleotide exchange on Ran by the mitotic regulator RCC1. *Nature*, 354(6348):80–82.

Bleichert, F. and Baserga, S. J. (2007). The long unwinding road of RNA helicases. *Molecular cell*, 27(3):339–52.

Blevins, M. B., Smith, A. M., Phillips, E. M., and Powers, M. A. (2003). Complex Formation among the RNA Export Proteins Nup98, Rae1/Gle2, and TAP. *J. Biol. Chem.*, 278(23):20979–20988.

Blight, K. J., McKeating, J. A., and Rice, C. M. (2002). Highly Permissive Cell Lines for Subgenomic and Genomic Hepatitis C Virus RNA Replication. *Journal of Virology*, 76(24):13001–13014.

Blobel, G. (1985). Gene gating: a hypothesis. *Proceedings of the National Academy of Sciences of the United States of America*, 82(24):8527–9.

Blower, M. D., Nachury, M., Heald, R., and Weis, K. (2005). A Rae1-containing ribonucleoprotein complex is required for mitotic spindle assembly. *Cell*, 121(2):223–34.

Bogerd, H. P., Echarri, A., Ross, T. M., and Cullen, B. R. (1998). Inhibition of human immunodeficiency virus Rev and human T-cell leukemia virus Rex function, but not Mason-Pfizer monkey virus constitutive transport element activity, by a mutant human nucleoporin targeted to Crm1. *Journal of virology*, 72(11):8627–35.

Bolhy, S., Bouhlel, I., Dultz, E., Nayak, T., Zuccolo, M., Gatti, X., Vallee, R., Ellenberg, J., and Doye, V. (2011). A Nup133-dependent NPC-anchored network tethers centrosomes to the nuclear envelope in prophase. *The Journal of cell biology*, 192(5):855–71.

Bolinger, C., Sharma, A., Singh, D., Yu, L., and Boris-Lawrie, K. (2010). RNA helicase A modulates translation of HIV-1 and infectivity of progeny virions. *Nucleic Acids Res*, 38(5):1686–1696.

Bolinger, C., Yilmaz, A., Hartman, T. R., Kovacic, M. B., Fernandez, S., Ye, J., Forget, M., Green, P. L., and Boris-Lawrie, K. (2007). RNA helicase A interacts with divergent lymphotropic retroviruses and promotes translation of human T-cell leukemia virus type 1. *Nucleic acids research*, 35(8):2629–42.

Bonamassa, B., Ciccarese, F., Antonio, V. D., Contarini, A., Palù, G., and Alvisi, G. (2015). Hepatitis C virus and host cell nuclear transport machinery: a clandestine affair. *Front Microbiol*, 6:619.

Boris-Lawrie, K., Roberts, T. M., and Hull, S. (2001). Retroviral RNA elements integrate components of post-transcriptional gene expression. *Life sciences*, 69(23):2697–709.

Bourgeois, C. F., Mortreux, F., and Auboeuf, D. (2016). The multiple functions of RNA helicases as drivers and regulators of gene expression. *Nature reviews. Molecular cell biology*, 17(7):426–38.

Brass, A. L., Dykxhoorn, D. M., Benita, Y., Yan, N., Engelman, A., Xavier, R. J., Lieberman, J., and Elledge, S. J. (2008). Identification of Host Proteins Required for HIV Infection Through a Functional Genomic Screen. *Science*, 319(5865):921–926.

Bratt, E. and Ohman, M. (2003). Coordination of editing and splicing of glutamate receptor pre-mRNA. *RNA (New York, N.Y.)*, 9(3):309–318.

Brickner, D. G., Ahmed, S., Meldi, L., Thompson, A., Light, W., Young, M., Hickman, T. L., Chu, F., Fabre, E., and Brickner, J. H. (2012). Transcription factor binding to a DNA zip code controls interchromosomal clustering at the nuclear periphery. *Developmental cell*, 22(6):1234–46.

Brohawn, S. G., Leksa, N. C., Spear, E. D., Rajashankar, K. R., and Schwartz, T. U. (2008). Structural Evidence for Common Ancestry of the Nuclear Pore Complex and Vesicle Coats. *Science*, 322(5906):1369–1373.

Brown, C. R., Kennedy, C. J., Delmar, V. A., Forbes, D. J., and Silver, P. A. (2008). Global histone acetylation induces functional genomic reorganization at mammalian nuclear pore complexes. *Genes & Development*, 22(5):627–639.

Browne, S. K., Roesser, J. R., Zhu, S. Z., and Ginder, G. D. (2006). Differential IFN-gamma stimulation of HLA-A gene expression through CRM-1-dependent nuclear RNA export. *Journal of immunology (Baltimore, Md. : 1950)*, 177(12):8612–9.

Buchwalter, A. L., Liang, Y., and Hetzer, M. W. (2014). Nup50 is required for cell differentiation and exhibits transcription-dependent dynamics. *Molecular Biology of the Cell*, 25(16):2472–2484.

Bui, K. H., von Appen, A., DiGuilio, A. L., Ori, A., Sparks, L., Mackmull, M.-T., Bock, T., Hagen, W., Andrés-Pons, A., Glavy, J. S., and Beck, M. (2013). Integrated structural analysis of the human nuclear pore complex scaffold. *Cell*, 155(6):1233–43.

Burke, B. and Stewart, C. L. (2013). The nuclear lamins: flexibility in function. *Nature reviews. Molecular cell biology*, 14(1):13–24.

Burns, L. T. and Wente, S. R. (2014). From Hypothesis to Mechanism: Uncovering Nuclear Pore Complex Links to Gene Expression. *Molecular and Cellular Biology*, 34(12):2114–2120.

Burrell, C. J., Howard, C. R., Murphy, F. A., Burrell, C. J., Howard, C. R., and Murphy, F. A. (2017). Chapter 36 Flaviviruses. In *Fenner and White's Medical Virology*, pages 493–518.

Butler, D. (2016). Zika virus: Brazil's surge in small-headed babies questioned by report. *Nature*, 530(7588):13–4.

Cai, Y., Singh, B. B., Aslanukov, A., Zhao, H., and Ferreira, P. A. (2001). The docking of kinesins, KIF5B and KIF5C, to Ran-binding protein 2 (RanBP2) is mediated via a novel RanBP2 domain. *The Journal of biological chemistry*, 276(45):41594–602.

Capelson, M., Liang, Y., Schulte, R., Mair, W., Wagner, U., and Hetzer, M. W. (2010). Chromatin-bound nuclear pore components regulate gene expression in higher eukaryotes. *Cell*, 140(3):372–383.

Capitanio, J. S., Montpetit, B., and Wozniak, R. W. (2017). Human Nup98 regulates the localization and activity of DExH/D-box helicase DHX9. *eLife*, 6.

Capitanio, J. S. and Wozniak, R. W. (2012). Hematopoietic cancers and Nup98 fusions: determining common mechanisms of malignancy. *arXiv: Molecular Networks (q-bio.MN)*, pages 1–13.

Caruthers, J. M., Johnson, E. R., and McKay, D. B. (2000). Crystal structure of yeast initiation factor 4A, a DEAD-box RNA helicase. *Proceedings of the National Academy of Sciences of the United States of America*, 97(24):13080–5.

Caruthers, J. M. and McKay, D. B. (2002). Helicase structure and mechanism. *Current opinion in structural biology*, 12(1):123–33.

Castelló, A., Izquierdo, J. M., Welnowska, E., and Carrasco, L. (2009). RNA nuclear export is blocked by poliovirus 2A protease and is concomitant with nucleoporin cleavage. *J Cell Sci*, 122(Pt 20):3799–3809.

Cerutti, A., Maillard, P., Minisini, R., Vidalain, P.-O. O., Roohvand, F., Pecheur, E.-I. I., Pirisi, M., and Budkowska, A. (2011). Identification of a functional, CRM-1-dependent nuclear export signal in hepatitis C virus core protein. *PLoS One*, 6(10):e25854.

Chakraborty, P. and Grosse, F. (2010). WRN helicase unwinds Okazaki fragment-like hybrids in a reaction stimulated by the human DHX9 helicase. *Nucleic acids research*, 38(14):4722–30.

Chakraborty, P., Wang, Y., Wei, J.-H., van Deursen, J., Yu, H., Malureanu, L., Dasso, M., Forbes, D. J., Levy, D. E., Seemann, J., and Fontoura, B. M. A. (2008). Nucleoporin levels regulate cell cycle progression and phase-specific gene expression. *Developmental cell*, 15(5):657–67.

Chan, J. F.-W., Yip, C. C.-Y., Tsang, J. O.-L., Tee, K.-M., Cai, J.-P., Chik, K. K.-H., Zhu, Z., Chan, C. C.-S., Choi, G. K.-Y., Sridhar, S., Zhang, A. J., Lu, G., Chiu, K., Lo, A. C.-Y., Tsao, S.-W., Kok, K.-H., Jin, D.-Y., Chan, K.-H., and Yuen, K.-Y. (2016). Differential cell line susceptibility to the emerging Zika virus: implications for disease pathogenesis, non-vector-borne human transmission and animal reservoirs. *Emerging Microbes & Infections*, 5(8):e93.

Chatel, G. and Fahrenkrog, B. (2012). Dynamics and diverse functions of nuclear pore complex proteins. *Nucleus*, 3(2):162–171.

Chatel-Chaix, L., Melançon, P., Racine, M., Baril, M., and Lamarre, D. (2011). Y-box-binding protein 1 interacts with hepatitis C virus NS3/4A and influences the equilibrium between viral RNA replication and infectious particle production. *J Virol*, 85(21):11022–11037.

Chen, S. L. and Morgan, T. R. (2006). The natural history of hepatitis C virus (HCV) infection. *International journal of medical sciences*, 3(2):47–52.

Chen, Z. X., Wallis, K., Fell, S. M., Sobrado, V. R., Hemmer, M. C., Ramsköld, D., Hellman, U., Sandberg, R., Kenchappa, R. S., Martinson, T., Johnsen, J. I., Kogner, P., and Schlisio, S. (2014). RNA helicase A is a downstream mediator of KIF1Bβ tumor-suppressor function in neuroblastoma. *Cancer discovery*, 4(4):434–51.

Cheng, Y., Holloway, M. P., Nguyen, K., McCauley, D., Landesman, Y., Kauffman, M. G., Shacham, S., and Altura, R. A. (2014). XPO1 (CRM1) inhibition represses STAT3 activation to drive a survivin-dependent oncogenic switch in triple-negative breast cancer. *Molecular cancer therapeutics*, 13(3):675–86.

Choo, Q. L., Kuo, G., Weiner, A. J., Overby, L. R., Bradley, D. W., and Houghton, M. (1989). Isolation of a cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome. *Science (New York, N.Y.)*, 244(4902):359–62.

Christian, H., Hofele, R. V., Urlaub, H., and Ficner, R. (2014). Insights into the activation of the helicase Prp43 by biochemical studies and structural mass spectrometry. *Nucleic Acids Research*, 42(2):1162–1179.

Chu, L., Su, M. Y., Maggi, L. B., Lu, L., Mullins, C., Crosby, S., Huang, G., Chng, W. J., Vij, R., Tomasson, M. H., and Tomasson, M. H. (2012). Multiple myeloma-associated chromosomal translocation activates orphan snoRNA ACA11 to suppress oxidative stress. *The Journal of clinical investigation*, 122(8):2793–806.

Chung, H.-Y., Gu, M., Buehler, E., MacDonald, M. R., and Rice, C. M. (2014). Seed sequence-matched controls reveal limitations of small interfering RNA knockdown in functional and structural studies of hepatitis C virus NS5A-MOBKL1B interaction. *Journal of virology*, 88(19):11022–33.

Chung, K. M., Lee, J., Kim, J. E., Song, O. K., Cho, S., Lim, J., Seedorf, M., Hahm, B., and Jang, S. K. (2000). Nonstructural protein 5A of hepatitis C virus inhibits the function of karyopherin beta3. *Journal of virology*, 74(11):5233–41.

Cingolani, G., Bednenko, J., Gillespie, M. T., and Gerace, L. (2002). Molecular basis for the recognition of a nonclassical nuclear localization signal by importin beta. *Molecular cell*, 10(6):1345–53.

Cingolani, G., Petosa, C., Weis, K., and Müller, C. W. (1999). Structure of importin-beta bound to the IBB domain of importin-alpha. *Nature*, 399(6733):221–9.

Colla, E., Lee, S. D., Sheen, M. R., Woo, S. K., and Kwon, H. M. (2006). TonEBP is inhibited by RNA helicase A via interaction involving the E'F loop. *The Biochemical journal*, 393(Pt 1):411–9.

Conrad, N. K. (2008). Chapter 15. Co-immunoprecipitation techniques for assessing RNA-protein interactions in vivo. *Methods Enzymol*, 449:317–342.

Copeland, A. M., Newcomb, W. W., and Brown, J. C. (2009). Herpes simplex virus replication: roles of viral proteins and nucleoporins in capsid-nucleus attachment. *Journal of virology*, 83(4):1660–8.

Cordin, O., Banroques, J., Tanner, N. K., and Linder, P. (2006). The DEAD-box protein family of RNA helicases. *Gene*, 367:17–37.

Cordin, O., Tanner, N. K., Doère, M., Linder, P., and Banroques, J. (2004). The newly discovered Q motif of DEAD-box RNA helicases regulates RNA-binding and helicase activity. *The EMBO journal*, 23(13):2478–87.

Cortese, M., Goellner, S., Acosta, E. G., Neufeldt, C. J., Oleksiuk, O., Lampe, M., Haselmann, U., Funaya, C., Schieber, N., Ronchi, P., Schorb, M., Pruunsild, P., Schwab, Y., Chatel-Chaix, L., Ruggieri, A., and Bartenschlager, R. (2017). Ultrastructural Characterization of Zika Virus Replication Factories. *Cell Reports*, 18(9):2113–2123.

Courington, D. and Vogt, P. K. (1967). Electron microscopy of chick fibroblasts infected by defective rous sarcoma virus and its helper. *Journal of virology*, 1(2):400–14.

Cox, J. and Mann, M. (2011). Quantitative, high-resolution proteomics for data-driven systems biology. *Annual review of biochemistry*, 80:273–99.

Cronshaw, J. M., Krutchinsky, A. N., Zhang, W., Chait, B. T., and Matunis, M. J. (2002). Proteomic analysis of the mammalian nuclear pore complex. *The Journal of Cell Biology*, 158(5):915–927.

Cross, M. K. and Powers, M. A. (2011). Nup98 regulates bipolar spindle assembly through association with microtubules and opposition of MCAK. *Molecular biology of the cell*, 22(5):661–72.

D'Angelo, M., Gomez-Cavazos, J., Mei, A., Lackner, D., and Hetzer, M. (2012). A Change in Nuclear Pore Complex Composition Regulates Cell Differentiation. *Developmental Cell*, 22(2):446–458.

D'Angelo, M. A., Anderson, D. J., Richard, E., and Hetzer, M. W. (2006). Nuclear pores form de novo from both sides of the nuclear envelope. *Science (New York, N.Y.)*, 312(5772):440–3.

Davuluri, G., Gong, W., Yusuff, S., Lorent, K., Muthumani, M., Dolan, A. C., and Pack, M. (2008). Mutation of the zebrafish nucleoporin elys sensitizes tissue progenitors to replication stress. *PLoS genetics*, 4(10):e1000240.

Dawlaty, M. M., Malureanu, L., Jeganathan, K. B., Kao, E., Sustmann, C., Tahk, S., Shuai, K., Grosschedl, R., and van Deursen, J. M. (2008). Resolution of Sister Centromeres RequiresRanBP2-Mediated SUMOylation of Topoisomerase IIα. *Cell*, 133(1):103–115.

de Chassey, B., Navratil, V., Tafforeau, L., Hiet, M. S., Aublin-Gex, A., Agaugué, S., Meiffren, G., Pradezynski, F., Faria, B. F., Chantier, T., Le Breton, M., Pellet, J., Davoust, N., Mangeot, P. E., Chaboud, A., Penin, F., Jacob, Y., Vidalain, P. O., Vidal, M., André, P., Rabourdin-Combe, C., and Lotteau, V. (2008). Hepatitis C virus infection protein network. *Molecular Systems Biology*, 4:230.

Denning, D. P., Patel, S. S., Uversky, V., Fink, A. L., and Rexach, M. (2003). Disorder in the nuclear pore complex: The FG repeat regions of nucleoporins are natively unfolded. *Proceedings of the National Academy of Sciences*, 100(5):2450–2455.

Devos, D., Dokudovskaya, S., Alber, F., Williams, R., Chait, B. T., Sali, A., and Rout, M. P. (2004). Components of coated vesicles and nuclear pore complexes share a common molecular architecture. *PLoS biology*, 2(12):e380.

Di Nunzio, F., Danckaert, A., Fricke, T., Perez, P., Fernandez, J., Perret, E., Roux, P., Shorte, S., Charneau, P., Diaz-Griffero, F., and Arhel, N. J. (2012). Human nucleoporins promote HIV-1 docking at the nuclear pore, nuclear import and integration. *PLoS One*, 7(9):e46037.

Di Nunzio, F., Fricke, T., Miccio, A., Valle-Casuso, J. C., Perez, P., Souque, P., Rizzi, E., Severgnini, M., Mavilio, F., Charneau, P., and Diaz-Griffero, F. (2013). Nup153 and Nup98 bind the HIV-1 core and contribute to the early steps of HIV-1 replication. *Virology*, 440(1):8–18.

Dickmanns, A., Kehlenbach, R. H., and Fahrenkrog, B. (2015). Nuclear Pore Complexes and Nucleocytoplasmic Transport. In *International review of cell and molecular biology*, volume 320, pages 171–233.

Dingwall, C., Robbins, J., Dilworth, S. M., Roberts, B., and Richardson, W. D. (1988). The nucleoplasmin nuclear location sequence is larger and more complex than that of SV-40 large T antigen. *The Journal of cell biology*, 107(3):841–9.

D'Ortenzio, E., Matheron, S., de Lamballerie, X., Hubert, B., Piorkowski, G., Maquart, M., Descamps, D., Damond, F., Yazdanpanah, Y., and Leparc-Goffart, I. (2016). Evidence of Sexual Transmission of Zika Virus. *New England Journal of Medicine*, 374(22):2195–2198.

Doucet, C. M., Talamas, J. A., and Hetzer, M. W. (2010). Cell Cycle-Dependent Differences in Nuclear Pore Complex Assembly in Metazoa. *Cell*, 141(6):1030–1041.

Dreyfuss, G., Choi, Y. D., and Adam, S. A. (1984). Characterization of heterogeneous nuclear RNA-protein complexes in vivo with monoclonal antibodies. *Mol Cell Biol*, 4(6):1104–1114.

Duheron, V., Chatel, G., Sauder, U., Oliveri, V., and Fahrenkrog, B. (2014). Structural characterization of altered nucleoporin Nup153 expression in human cells by thin-section electron microscopy. TL - 5. *Nucleus (Austin, Tex.)*, 5 VN - re(6):601–612.

Dultz, E. and Ellenberg, J. (2010). Live imaging of single nuclear pores reveals unique assembly kinetics and mechanism in interphase. *The Journal of cell biology*, 191(1):15–22.

Dultz, E., Zanin, E., Wurzenberger, C., Braun, M., Rabut, G., Sironi, L., and Ellenberg, J. (2008). Systematic kinetic analysis of mitotic dis- and reassembly of the nuclear pore in living cells. *The Journal of Cell Biology*, 180(5):857–865.

Elton, D., Simpson-Holley, M., Archer, K., Medcalf, L., Hallam, R., McCauley, J., and Digard, P. (2001). Interaction of the Influenza Virus Nucleoprotein with the Cellular CRM1-Mediated Nuclear Export Pathway. *Journal of Virology*, 75(1):408–419.

Enninga, J., Levy, D. E., Blobel, G. G., and Fontoura, B. M. A. (2002). Role of Nucleoporin Induction in Releasing an mRNA Nuclear Export Block. *Science*, 295(5559):1523–1525.

Erkizan, H. V., Schneider, J. A., Sajwan, K., Graham, G. T., Griffin, B., Chasovskikh, S., Youbi, S. E., Kallarakal, A., Chruszcz, M., Padmanabhan, R., Casey, J. L., Üren, A., and Toretsky, J. A. (2015). RNA helicase A activity is inhibited by oncogenic transcription factor EWS-FLI1. *Nucleic Acids Res*, 43(2):1069–1080.

Fairman-Williams, M. E., Guenther, U.-P., and Jankowsky, E. (2010). SF1 and SF2 helicases: family matters. *Current opinion in structural biology*, 20(3):313–24.

Faria, A. M., Levay, A., Wang, Y., Kamphorst, A. O., Rosa, M. L., Nussenzveig, D. R., Balkan, W., Chook, Y. M., Levy, D. E., and Fontoura, B. M. (2006). The Nucleoporin Nup96 Is Required for Proper Expression of Interferon-Regulated Proteins and Functions. *Immunity*, 24(3):295–304.

Faria, P. A., Chakraborty, P., Levay, A., Barber, G. N., Ezelle, H. J., Enninga, J., Arana, C., van Deursen, J., and Fontoura, B. M. A. (2005). VSV disrupts the Rae1/mrnp41 mRNA nuclear export pathway. *Molecular cell*, 17(1):93–102.

Feldherr, C. M. and Akin, D. (1997). The location of the transport gate in the nuclear pore complex. *Journal of cell science*, pages 3065–70.

Feldherr, C. M., Kallenbach, E., and Schultz, N. (1984). Movement of a karyophilic protein through the nuclear pores of oocytes. *The Journal of cell biology*, 99(6):2216–22.

Fénéant, L., Potel, J., François, C., Sané, F., Douam, F., Belouzard, S., Calland, N., Vausselin, T., Rouillé, Y., Descamps, V., Baumert, T. F., Duverlie, G., Lavillette, D., Hober, D., Dubuisson, J., Wychowski, C., and Cocquerel, L. (2015). New Insights into the Understanding of Hepatitis C Virus Entry and Cell-to-Cell Transmission by Using the Ionophore Monensin A. *Journal of virology*, 89(16):8346–64.

Fiches, G. N., Eyre, N. S., Aloia, A. L., Van Der Hoek, K., Betz-Stablein, B., Luciani, F., Chopra, A., and Beard, M. R. (2016). HCV RNA traffic and association with NS5A in living cells. *Virology*, 493:60–74.

Fidaleo, M., Paola, E. D., and Paronetto, M. P. (2016). The RNA helicase A in malignant transformation. *Oncotarget*, 7(19):28711–23.

Field, M. C. and Dacks, J. B. (2009). First and last ancestors: reconstructing evolution of the endomembrane system with ESCRTs, vesicle coat proteins, and nuclear pore complexes. *Current Opinion in Cell Biology*, 21(1):4–13.

Finlay, D. R., Newmeyer, D. D., Price, T. M., and Forbes, D. J. (1987). Inhibition of in vitro nuclear transport by a lectin that binds to nuclear pores. *The Journal of cell biology*, 104(2):189–200.

Flather, D. and Semler, B. L. (2015). Picornaviruses and nuclear functions: targeting a cellular compartment distinct from the replication site of a positive-strand RNA virus. *Frontiers in microbiology*, 6:594.

Fontanes, V., Raychaudhuri, S., and Dasgupta, A. (2009). A cell-permeable peptide inhibits hepatitis C virus replication by sequestering IRES transacting factors. *Virology*, 394(1):82–90.

Fontoura, B. M., Faria, P. A., and Nussenzveig, D. R. (2005). Viral interactions with the nuclear transport machinery: discovering and disrupting pathways. *IUBMB Life*, 57(2):65–72.

Fornerod, M., van Deursen, J., van Baal, S., Reynolds, A., Davis, D., Murti, K. G., Fransen, J., and Grosveld, G. (1997). The human homologue of yeast CRM1 is in a dynamic subcomplex with CAN/Nup214 and a novel nuclear pore component Nup88. *The EMBO Journal*, 16(4):807–816.

Francisco-Velilla, R., Fernandez-Chamorro, J., Ramajo, J., and Martinez-Salas, E. (2016). The RNA-binding protein Gemin5 binds directly to the ribosome and regulates global translation. *Nucleic Acids Research*, 44(17):8335–8351.

Franks, T. M., Benner, C., Narvaiza, I., Marchetto, M. C. N., Young, J. M., Malik, H. S., Gage, F. H., and Hetzer, M. W. (2016). Evolution of a transcriptional regulator from a transmembrane nucleoporin. *Genes & development*, 30(10):1–17.

Fried, H. and Kutay, U. (2003). Nucleocytoplasmic transport: taking an inventory. *Cellular and molecular life sciences : CMLS*, 60(8):1659–88.

Friedemann, J., Grosse, F., and Zhang, S. (2005). Nuclear DNA helicase II (RNA helicase A) interacts with Werner syndrome helicase and stimulates its exonuclease activity. *The Journal of biological chemistry*, 280(35):31303–13.

Frieman, M., Yount, B., Heise, M., Kopecky-Bromberg, S. A., Palese, P., and Baric, R. S. (2007). Severe acute respiratory syndrome coronavirus ORF6 antagonizes STAT1 function by sequestering nuclear import factors on the rough endoplasmic reticulum/Golgi membrane. *Journal of virology*, 81(18):9812–24.

Fry, D. C., Kuby, S. A., and Mildvan, A. S. (1986). ATP-binding site of adenylate kinase: mechanistic implications of its homology with ras-encoded p21, F1-ATPase, and other nucleotide-binding proteins. *Proceedings of the National Academy of Sciences of the United States of America*, 83(4):907–11.

Fuchsová, B. and Hozák, P. (2002). The localization of nuclear DNA helicase II in different nuclear compartments is linked to transcription. *Exp Cell Res*, 279(2):260–270.

Fujii, R., Okamoto, M., Aratani, S., Oishi, T., Ohshima, T., Taira, K., Baba, M., Fukamizu, A., and Nakajima, T. (2001). A Role of RNA Helicase A in cis-Acting Transactivation Response Element-mediated Transcriptional Regulation of Human Immunodeficiency Virus Type 1. *J Biol Chem*, 276(8):5445–5451.

Fujita, H., Fujii, R., Aratani, S., Amano, T., Fukamizu, A., and Nakajima, T. (2003). Antithetic effects of MBD2a on gene regulation. *Molecular and cellular biology*, 23(8):2645–57.

Fujita, H., Ohshima, T., Oishi, T., Aratani, S., Fujii, R., Fukamizu, A., and Nakajima, T. (2005). Relevance of nuclear localization and functions of RNA helicase A. *Int J Mol Med*, 15(4):555–560.

Fullam, A. and Schröder, M. (2013). DExD/H-box RNA helicases as mediators of anti-viral innate immunity and essential host factors for viral replication. *Biochim Biophys Acta*, 1829(8):854–865.

Fuller-Pace, F. V. (2006). DExD/H box RNA helicases: multifunctional proteins with important roles in transcriptional regulation. *Nucleic Acids Res*, 34(15):4206–4215.

Fuller-Pace, F. V. (2013a). DEAD box RNA helicase functions in cancer. *RNA Biology*, 10(1):121–132.

Fuller-Pace, F. V. (2013b). The DEAD box proteins DDX5 (p68) and DDX17 (p72): multi-tasking transcriptional regulators. *Biochim Biophys Acta*, 1829(8):756–763.

Funakoshi, T., Clever, M., Watanabe, A., and Imamoto, N. (2011). Localization of Pom121 to the inner nuclear membrane is required for an early step of interphase nuclear pore complex assembly. *Molecular biology of the cell*, 22(7):1058–69.

Gall, J. G. (1967). Octagonal nuclear pores. *The Journal of cell biology*, 32(2):391–9.

Garcia-Blanco, M. A., Vasudevan, S. G., Bradrick, S. S., and Nicchitta, C. (2016). Flavivirus RNA transactions from viral entry to genome replication. *Antiviral Research*, 134:244–249.

Gastaminza, P., Cheng, G., Wieland, S., Zhong, J., Liao, W., and Chisari, F. V. (2008). Cellular determinants of hepatitis C virus assembly, maturation, degradation, and secretion. *Journal of virology*, 82(5):2120–9.

Georgiev, P., Chlamydas, S., and Akhtar, A. (2011). Drosophila dosage compensation: males are from Mars, females are from Venus. *Fly*, 5(2):147–54.

Gerace, L., Ottaviano, Y., and Kondor-Koch, C. (1982). Identification of a major polypeptide of the nuclear pore complex. *The Journal of cell biology*, 95(3):826–37.

Germain, M.-A., Chatel-Chaix, L., Gagné, B., Bonneil, r., Thibault, P., Pradezynski, F., de Chassey, B., Meyniel-Schicklin, L., Lotteau, V., Baril, M., and Lamarre, D. (2014). Elucidating novel hepatitis C virus-host interactions using combined mass spectrometry and functional genomics approaches. *Molecular & cellular proteomics : MCP*, 13(1):184–203.

Gerritsen, M. E., Williams, A. J., Neish, A. S., Moore, S., Shi, Y., and Collins, T. (1997). CREB-binding protein/p300 are transcriptional coactivators of p65. *Proceedings of the National Academy of Sciences of the United States of America*, 94(7):2927–32.

Ghildyal, R., Jordan, B., Li, D., Dagher, H., Bardin, P. G., Gern, J. E., and Jans, D. A. (2009). Rhinovirus 3C protease can localize in the nucleus and alter active and passive nucleocytoplasmic transport. *Journal of virology*, 83(14):7349–52.

Ghouzzi, V. E., Bianchi, F. T., Molineris, I., Mounce, B. C., Berto, G. E., Rak, M., Lebon, S., Aubry, L., Tocco, C., Gai, M., Chiotto, A. M., Sgrò, F., Pallavicini, G., Simon-Loriere, E., Passemard, S., Vignuzzi, M., Gressens, P., and Di Cunto, F. (2016). ZIKA virus elicits P53 activation and genotoxic stress in human neural progenitors similar to mutations involved in severe forms of genetic microcephaly and p53. *Cell Death and Disease*, 7(10):e2440.

Giaccia, A. J. and Kastan, M. B. (1998). The complexity of p53 modulation: emerging patterns from divergent signals. *Genes & development*, 12(19):2973–83.

Gibson, T. J. and Thompson, J. D. (1994). Detection of dsRNA-binding domains in RNA helicase A and Drosophila maleless: implications for monomeric RNA helicases. *Nucleic Acids Res*, 22(13):2552–2556.

Gilchrist, D., Mykytka, B., and Rexach, M. (2002). Accelerating the Rate of Disassembly of Karyopherin{middle dot}Cargo Complexes. *Journal of Biological Chemistry*, 277(20):18161–18172.

Gillespie, P. J., Khoudoli, G. A., Stewart, G., Swedlow, J. R., and Blow, J. J. (2007). ELYS/MEL-28 chromatin association coordinates nuclear pore complex assembly and replication licensing. *Current biology : CB*, 17(19):1657–62.

Glavy, J. S., Krutchinsky, A. N., Cristea, I. M., Berke, I. C., Boehmer, T., Blobel, G., and Chait, B. T. (2007). Cell-cycle-dependent phosphorylation of the nuclear pore Nup107-160 subcomplex. *Proceedings of the National Academy of Sciences of the United States of America*, 104(10):3811–6.

Gokhale, N., McIntyre, A., McFadden, M., Roder, A., Kennedy, E., Gandara, J., Hopcraft, S., Quicke, K., Vazquez, C., Willer, J., Ilkayeva, O., Law, B., Holley, C., Garcia-Blanco, M., Evans, M., Suthar, M., Bradrick, S., Mason, C., and Horner, S. (2016). N6-Methyladenosine in Flaviviridae Viral RNA Genomes Regulates Infection. *Cell Host & Microbe*, 20(5):654–665.

Goldfarb, D. S., Gariépy, J., Schoolnik, G., and Kornberg, R. D. (1986). Synthetic peptides as nuclear localization signals. *Nature*, 322(6080):641–4.

Görlich, D., Henklein, P., Laskey, R. A., and Hartmann, E. (1996a). A 41 amino acid motif in importin-alpha confers binding to importin-beta and hence transit into the nucleus. *The EMBO journal*, 15(8):1810–7.

Görlich, D., Panté, N., Kutay, U., Aebi, U., and Bischoff, F. R. (1996b). Identification of different roles for RanGDP and RanGTP in nuclear protein import. *The EMBO journal*, 15(20):5584–94.

Gough, S. M., Slape, C. I., and Aplan, P. D. (2011). NUP98 gene fusions and hematopoietic malignancies: common themes and new biologic insights. *Blood*, 118(24):6247–6257.

Greber, U. F., Senior, A., and Gerace, L. (1990). A major glycoprotein of the nuclear pore complex is a membrane-spanning polypeptide with a large lumenal domain and a small cytoplasmic tail. *The EMBO journal*, 9(5):1495–502.

Griffis, E. R., Altan, N., Lippincott-Schwartz, J., and Powers, M. A. (2002). Nup98 Is a Mobile Nucleoporin with Transcription-dependent Dynamics. *Mol. Biol. Cell*, 13(4):1282–1297.

Griffis, E. R., Craige, B., Dimaano, C., Ullman, K. S., and Powers, M. A. (2004). Distinct Functional Domains within Nucleoporins Nup153 and Nup98 Mediate Transcription-dependent Mobility. *Mol. Biol. Cell*, 15(4):1991–2002.

Griffis, E. R., Xu, S., and Powers, M. A. (2003). Nup98 Localizes to Both Nuclear and Cytoplasmic Sides of the Nuclear Pore and Binds to Two Distinct Nucleoporin Subcomplexes. *Mol. Biol. Cell*, 14(2):600–610.

Gross, C. H. and Shuman, S. (1996). Vaccinia virus RNA helicase: nucleic acid specificity in duplex unwinding. *Journal of virology*, 70(4):2615–9.

Gross, C. H. and Shuman, S. (1998). The nucleoside triphosphatase and helicase activities of vaccinia virus NPH-II are essential for virus replication. *Journal of virology*, 72(6):4729–36.

Grover, R., Candeias, M. M., Fåhraeus, R., and Das, S. (2009). p53 and little brother p53/47: linking IRES activities with protein functions. *Oncogene*, 28(30):2766–72.

Grüter, P., Tabernero, C., von Kobbe, C., Schmitt, C., Saavedra, C., Bachi, A., Wilm, M., Felber, B. K., and Izaurralde, E. (1998). TAP, the human homolog of Mex67p, mediates CTE-dependent RNA export from the nucleus. *Molecular cell*, 1(5):649–59.

Guirakhoo, F., Bolin, R. A., and Roehrig, J. T. (1992). The Murray Valley encephalitis virus prM protein confers acid resistance to virus particles and alters the expression of epitopes within the R2 domain of E glycoprotein. *Virology*, 191(2):921–931.

Gustin, K. E. and Sarnow, P. (2001). Effects of poliovirus infection on nucleocytoplasmic trafficking and nuclear pore complex composition. *The EMBO journal*, 20(1-2):240–9.

Gustin, K. E. and Sarnow, P. (2002). Inhibition of nuclear import and alteration of nuclear pore complex composition by rhinovirus. *Journal of virology*, 76(17):8787–96.

Güttler, T. and Görlich, D. (2011). Ran-dependent nuclear export mediators: a structural perspective. *The EMBO Journal*, 30(17):3457–3474.

Halaby, M.-J., Harris, B. R. E., Miskimins, W. K., Cleary, M. P., and Yang, D.-Q. (2015a). Deregulation of Internal Ribosome Entry Site-Mediated p53 Translation in Cancer Cells with Defective p53 Response to DNA Damage. *Molecular and cellular biology*, 35(23):4006–17.

Halaby, M.-J., Li, Y., Harris, B. R., Jiang, S., Miskimins, W. K., Cleary, M. P., and Yang, D.-Q. (2015b). Translational Control Protein 80 Stimulates IRES-Mediated Translation of p53 mRNA in Response to DNA Damage. *BioMed research international*, 2015:708158.

Halaby, M.-J. and Yang, D.-Q. (2007). p53 translational control: a new facet of p53 regulation and its implication for tumorigenesis and cancer therapeutics. *Gene*, 395(1-2):1–7.

Hallberg, E., Wozniak, R. W., and Blobel, G. (1993). An integral membrane protein of the pore membrane domain of the nuclear envelope contains a nucleoporin-like region. *The Journal of cell biology*, 122(3):513–21.

Hamel, R., Dejarnac, O., Wichit, S., Ekchariyawat, P., Neyret, A., Luplertlop, N., Perera-Lecoin, M., Surasombatpattana, P., Talignani, L., Thomas, F., Cao-Lormeau, V.-M., Choumet, V., Briant, L., Desprès, P., Amara, A., Yssel, H., and Missé, D. (2015). Biology of Zika Virus Infection in Human Skin Cells. *Journal of virology*, 89(17):8880–96.

Han, S. P., Tang, Y. H., and Smith, R. (2010). Functional diversity of the hnRNPs: past, present and perspectives. *The Biochemical journal*, 430(3):379–92.

Hanada, K. and Hickson, I. D. (2007). Molecular genetics of RecQ helicase disorders. *Cellular and molecular life sciences : CMLS*, 64(17):2306–22.

Hardwick, S. W. and Luisi, B. F. (2013). Rarely at rest. *RNA Biology*, 10(1):56–70.

Harel, A., Chan, R. C., Lachish-Zalait, A., Zimmerman, E., Elbaum, M., and Forbes, D. J. (2003). Importin Negatively Regulates Nuclear Membrane Fusion and Nuclear Pore Complex Assembly. *Molecular Biology of the Cell*, 14(11):4387–4396.

Hartman, T. R., Qian, S., Bolinger, C., Fernandez, S., Schoenberg, D. R., and Boris-Lawrie, K. (2006). RNA helicase A is necessary for translation of selected messenger RNAs. *Nat Struct Mol Biol*, 13(6):509–516.

Hartmuth, K., Urlaub, H., Vornlocher, H.-P., Will, C. L., Gentzel, M., Wilm, M., and Lührmann, R. (2002). Protein composition of human prespliceosomes isolated by a tobramycin affinity-selection method. *Proceedings of the National Academy of Sciences of the United States of America*, 99(26):16719 –16724.

Hawryluk-Gara, L. A., Platani, M., Santarella, R., Wozniak, R. W., and Mattaj, I. W. (2008). Nup53 is required for nuclear envelope and nuclear pore complex assembly. *Mol Biol Cell*, 19(4):1753–1762.

Hayes, E. B. (2009). Zika Virus Outside Africa. *Emerging Infectious Diseases*, 15(9):1347–1350.

He, Q. S., Tang, H., Zhang, J., Truong, K., Wong-Staal, F., and Zhou, D. (2008). Comparisons of RNAi approaches for validation of human RNA helicase A as an essential factor in hepatitis C virus replication. *Journal of Virological Methods*, 154(1-2):216–219.

He, X. and Zhang, J. (2006). Why do hubs tend to be essential in protein networks? *PLoS Genet*, 2(6):e88.

Heaton, N. S. and Randall, G. (2010). Dengue Virus-Induced Autophagy Regulates Lipid Metabolism. *Cell Host & Microbe*, 8(5):422–432.

Hellemans, J., Mortier, G., De Paepe, A., Speleman, F., and Vandesompele, J. (2007). qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. *Genome Biol*, 8(2):R19.

Hendrickson, D. G., Kelley, D. R., Tenen, D., Bernstein, B., and Rinn, J. L. (2016). Widespread RNA binding by chromatin-associated proteins. *Genome biology*, 17(1):28.

Hetzer, M. W. (2010). The nuclear envelope. *Cold Spring Harbor perspectives in biology*, 2(3):a000539.

Hetzer, M. W., Walther, T. C., and Mattaj, I. W. (2005). Pushing the envelope: structure, function, and dynamics of the nuclear periphery. *Annual review of cell and developmental biology*, 21(1):347–80.

Höck, J., Weinmann, L., Ender, C., Rüdel, S., Kremmer, E., Raabe, M., Urlaub, H., and Meister, G. (2007). Proteomic and functional analysis of Argonaute-containing mRNA-protein complexes in human cells. *EMBO reports*, 8(11):1052–60.

Hodel, A. E., Hodel, M. R., Griffis, E. R., Hennig, K. A., Ratner, G. A., Xu, S., and Powers, M. A. (2002). The Three-Dimensional Structure of the Autoproteolytic, Nuclear Pore-Targeting Domain of the Human Nucleoporin Nup98. *Molecular Cell*, 10(2):347–358.

Hoekstra, E. J., Mesman, S., de Munnik, W. A., and Smidt, M. P. (2013). LMX1B is part of a transcriptional complex with PSPC1 and PSF. *PloS one*, 8(1):e53122.

Hofmann, W. A., Stojiljkovic, L., Fuchsova, B., Vargas, G. M., Mavrommatis, E., Philimonenko, V., Kysela, K., Goodrich, J. A., Lessard, J. L., Hope, T. J., Hozak, P., and de Lanerolle, P. (2004). Actin is part of pre-initiation complexes and is necessary for transcription by RNA polymerase II. *Nature cell biology*, 6(11):1094–101.

Holt, G. D., Snow, C. M., Senior, A., Haltiwanger, R. S., Gerace, L., and Hart, G. W. (1987). Nuclear pore complex glycoproteins contain cytoplasmically disposed O-linked N-acetylglucosamine. *The Journal of cell biology*, 104(5):1157–64.

Hou, C. and Corces, V. G. (2010). Nups take leave of the nuclear envelope to regulate transcription. *Cell*, 140(3):306–308.

Huang, H., Sun, F., Owen, D. M., Li, W., Chen, Y., Gale, M., and Ye, J. (2007). Hepatitis C virus production by human hepatocytes dependent on assembly and secretion of very low-density lipoproteins. *Proceedings of the National Academy of Sciences of the United States of America*, 104(14):5848–53.

Huang, W., Loganantharaj, R., Schroeder, B., Fargo, D., and Li, L. (2013). PAVIS: a tool for Peak Annotation and Visualization. *Bioinformatics (Oxford, England)*, 29(23):3097–9.

Huber, W., Carey, V. J., Gentleman, R., Anders, S., Carlson, M., Carvalho, B. S., Bravo, H. C., Davis, S., Gatto, L., Girke, T., Gottardo, R., Hahne, F., Hansen, K. D., Irizarry, R. A., Lawrence, M., Love, M. I., MacDonald, J., Obenchain, V., Oleś, A. K., Pagès, H., Reyes, A., Shannon, P., Smyth, G. K., Tenenbaum, D., Waldron, L., and Morgan, M. (2015). Orchestrating high-throughput genomic analysis with Bioconductor. *Nature Methods*, 12(2):115–121.

Hull, S. and Boris-Lawrie, K. (2002). RU5 of Mason-Pfizer monkey virus 5' long terminal repeat enhances cytoplasmic expression of human immunodeficiency virus type 1 gag-pol and nonviral reporter RNA. *Journal of virology*, 76(20):10211–8.

Hull, S. and Boris-Lawrie, K. (2003). Analysis of synergy between divergent simple retrovirus posttranscriptional control elements. *Virology*, 317(1):146–54.

Hülsmann, B. B., Labokha, A. A., and Görlich, D. (2012). The Permeability of Reconstituted Nuclear Pores Provides Direct Evidence for the Selective Phase Model. *Cell*, 150(4):738–751.

Ibarra, A., Benner, C., Tyagi, S., Cool, J., and Hetzer, M. W. (2016). Nucleoporin-mediated regulation of cell identity genes. *Genes & development*, 30(20):2253–2258.

Ibarra, A. and Hetzer, M. W. (2015). Nuclear pore proteins and the control of genome functions. *Genes & Development*, 29(4):337–349.

Ide, Y., Zhang, L., Chen, M., Inchauspe, G., Bahl, C., Sasaguri, Y., and Padmanabhan, R. (1996). Characterization of the nuclear localization signal and subcellular distribution of hepatitis C virus nonstructural protein NS5A. *Gene*, 182(1-2):203–11.

Iost, I., Dreyfus, M., and Linder, P. (1999). Ded1p, a DEAD-box protein required for translation initiation in Saccharomyces cerevisiae, is an RNA helicase. *The Journal of biological chemistry*, 274(25):17677–83.

Ishiguro, A., Ideta, M., Mikoshiba, K., Chen, D. J., and Aruga, J. (2007). ZIC2-dependent transcriptional regulation is mediated by DNA-dependent protein kinase, poly(ADP-ribose) polymerase, and RNA helicase A. *The Journal of biological chemistry*, 282(13):9983–95.

Ishii, K., Arib, G., Lin, C., Van Houwe, G., and Laemmli, U. K. (2002). Chromatin boundaries in budding yeast: the nuclear pore connection. *Cell*, 109(5):551–562.

Isken, O., Baroth, M., Grassmann, C. W., Weinlich, S., Ostareck, D. H., Ostareck-Lederer, A., and Behrens, S.-E. E. (2007). Nuclear factors are involved in hepatitis C virus RNA replication. *RNA*, 13(10):1675–1692.

Isken, O., Grassmann, C. W., Sarisky, R. T., Kann, M., Zhang, S., Grosse, F., Kao, P. N., and Behrens, S.-E. (2003). Members of the NF90/NFAR protein group are involved in the life cycle of a positive-strand RNA virus. *The EMBO journal*, 22(21):5655–65.

Isoyama, T., Kuge, S., and Nomoto, A. (2002). The core protein of hepatitis C virus is imported into the nucleus by transport receptor Kap123p but inhibits Kap121p-dependent nuclear import of yeast AP1-like transcription factor in yeast cells. *The Journal of biological chemistry*, 277(42):39634–41.

Iwamoto, M., Asakawa, H., Hiraoka, Y., and Haraguchi, T. (2010). Nucleoporin Nup98: a gatekeeper in the eukaryotic kingdoms. *Genes Cells*, 15(7):661–669.

Iyer, N. G., Ozdag, H., and Caldas, C. (2004). p300/CBP and cancer. *Oncogene*, 23(24):4225–31.

Izaurralde, E., Kutay, U., von Kobbe, C., Mattaj, I. W., and Görlich, D. (1997). The asymmetric distribution of the constituents of the Ran system is essential for transport into and out of the nucleus. *The EMBO Journal*, 16(21):6535–6547.

Jacinto, F. V., Benner, C., and Hetzer, M. W. (2015). The nucleoporin Nup153 regulates embryonic stem cell pluripotency through gene silencing. *Genes & development*, 29(12):1224–38.

Jain, A., Bacolla, A., Chakraborty, P., Grosse, F., and Vasquez, K. M. (2010). Human DHX9 Helicase Unwinds Triple-Helical DNA Structures. *Biochemistry*, 49(33):6992–6999.

Jamali, T., Jamali, Y., Mehrbod, M., and Mofrad, M. (2011). Chapter six Nuclear Pore Complex: Biochemistry and Biophysics of Nucleocytoplasmic Transport in Health and Disease. In *International Review of Cell and Molecular Biology*, volume 287, pages 233–286.

Jankowsky, E. (2011). RNA helicases at work: binding and rearranging. *Trends in Biochemical Sciences*, 36(1):19–29.

Jankowsky, E. and Fairman, M. E. (2007). RNA helicases–one fold for many functions. *Curr Opin Struct Biol*, 17(3):316–324.

Jarmoskaite, I. and Russell, R. (2014). RNA helicase proteins as chaperones and remodelers. *Annual review of biochemistry*, 83(1):697–725.

Jayappa, K. D., Ao, Z., and Yao, X. (2012). The HIV-1 passage from cytoplasm to nucleus: the process involving a complex exchange between the components of HIV-1 and cellular machinery to access nucleus and successful integration. *International journal of biochemistry and molecular biology*, 3(1):70–85.

Jeang, K.-T. T. and Yedavalli, V. (2006). Role of RNA helicases in HIV-1 replication. *Nucleic Acids Res*, 34(15):4198–4205.

Jeganathan, K. B., Baker, D. J., and van Deursen, J. M. (2006). Securin associates with APCCdh1 in prometaphase but its destruction is delayed by Rae1 and Nup98 until the metaphase/anaphase transition. *Cell Cycle (Georgetown, Tex.)*, 5(4):366–370.

Jeganathan, K. B., Malureanu, L., and van Deursen, J. M. (2005). The Rae1-Nup98 complex prevents aneuploidy by inhibiting securin degradation. *Nature*, 438(7070):1036–1039.

Jensen, K. B. and Darnell, R. B. (2008). CLIP: crosslinking and immunoprecipitation of in vivo RNA targets of RNA-binding proteins. *Methods Mol Biol*, 488:85–98.

Jiang, J. and Luo, G. (2009). Apolipoprotein E but not B is required for the formation of infectious hepatitis C virus particles. *Journal of virology*, 83(24):12680–91.

Jin, J., Jing, W., Lei, X. X., Feng, C., Peng, S., Boris-Lawrie, K., and Huang, Y. (2011). Evidence that Lin28 stimulates translation by recruiting RNA helicase A to polysomes. *Nucleic Acids Res*, 39(9):3724–3734.

Jones, C. T., Ma, L., Burgner, J. W., Groesch, T. D., Post, C. B., and Kuhn, R. J. (2003). Flavivirus capsid is a dimeric alpha-helical protein. *Journal of virology*, 77(12):7143–9.

Joseph, J. and Dasso, M. (2008). The nucleoporin Nup358 associates with and regulates interphase microtubules. *FEBS Letters*, 582(2):190–196.

Jovanovic-Talisman, T., Tetenbaum-Novatt, J., McKenney, A. S., Zilman, A., Peters, R., Rout, M. P., and Chait, B. T. (2009). Artificial nanopores that mimic the transport selectivity of the nuclear pore complex. *Nature*, 457(7232):1023–7.

Joyce, M. A. and Tyrrell, D. L. J. (2010). The cell biology of hepatitis C virus. *Microbes and infection / Institut Pasteur*, 12:263–271.

Kalderon, D., Roberts, B. L., Richardson, W. D., and Smith, A. E. (1984). A short amino acid sequence able to specify nuclear location. *Cell*, 39(3 Pt 2):499–509.

Kalverda, B. and Fornerod, M. (2010). Characterization of genome-nucleoporin interactions in Drosophila links chromatin insulators to the nuclear pore complex. *Cell Cycle*, 9(24):4812–4817.

Kalverda, B., Pickersgill, H., Shloma, V. V., and Fornerod, M. (2010). Nucleoporins directly stimulate expression of developmental and cell-cycle genes inside the nucleoplasm. *Cell*, 140(3):360–371.

Kasper, L. H., Brindle, P. K., Schnabel, C. A., Pritchard, C. E., Cleary, M. L., and van Deursen, J. M. (1999). CREB Binding Protein Interacts with Nucleoporin-Specific FG Repeats That Activate Transcription and Mediate NUP98-HOXA9 Oncogenicity. *Mol. Cell. Biol.*, 19(1):764–776.

Kato, T., Date, T., Miyamoto, M., Furusaka, A., Tokushige, K., Mizokami, M., and Wakita, T. (2003). Efficient replication of the genotype 2a hepatitis C virus subgenomic replicon. *Gastroenterology*, 125(6):1808–17.

Kawai, S. and Amano, A. (2012). BRCA1 regulates microRNA biogenesis via the DROSHA microprocessor complex. *J Cell Biol*, 197(2):201–208.

Keminer, O. and Peters, R. (1999). Permeability of Single Nuclear Pores. *Biophysical Journal*, 77(1):217–228.

Kernan, M. J., Kuroda, M. I., Kreber, R., Baker, B. S., and Ganetzky, B. (1991). napts, a mutation affecting sodium channel activity in Drosophila, is an allele of mle, a regulator of X chromosome transcription. *Cell*, 66(5):949–59.

Kessel, R. G. (1992). Annulate lamellae: a last frontier in cellular organelles. *International review of cytology*, 133:43–120.

Kim, K. S. and Boatman, E. S. (1967). Electron microscopy of monkey kidney cell cultures infected with rubella virus. *Journal of virology*, 1(1):205–14.

Kim, M. K. and Nikodem, V. M. (1999). hnRNP U inhibits carboxy-terminal domain phosphorylation by TFIIH and represses RNA polymerase II elongation. *Mol Cell Biol*, 19(10):6833–6844.

Kim, T., Pazhoor, S., Bao, M., Zhang, Z., Hanabuchi, S., Facchinetti, V., Bover, L., Plumas, J., Chaperot, L., Qin, J., and Liu, Y.-J. (2010). Aspartate-glutamate-alanine-histidine box motif (DEAH)/RNA helicase A helicases sense microbial DNA in human plasmacytoid dendritic cells. *Proceedings of the National Academy of Sciences of the United States of America*, 107(34):15181–6.

Knipe, D. M. and Howley, P. M. (2013). *Fields virology, 6th ed.* Wolters Kluwer/Lippincott Williams & Wilkins Health, Philadelphia, 6th edition.

Knockenhauer, K. E. and Schwartz, T. U. (2016). The Nuclear Pore Complex as a Flexible and Dynamic Gate. *Cell*, 164(6):1162–1171.

Koh, H. R., Xing, L., Kleiman, L., and Myong, S. (2014). Repetitive RNA unwinding by RNA helicase A facilitates RNA annealing. *Nucleic acids research*, 42(13):8556–64.

König, R., Zhou, Y., Elleder, D., Diamond, T. L., Bonamy, G. M. C., Irelan, J. T., Chiang, C.-Y. Y., Tu, B. P., De Jesus, P. D., Lilley, C. E., Seidel, S., Opaluch, A. M., Caldwell, J. S., Weitzman, M. D., Kuhen, K. L., Bandyopadhyay, S., Ideker, T., Orth, A. P., Miraglia, L. J., Bushman, F. D., Young, J. A., and Chanda, S. K. (2008). Global analysis of host-pathogen interactions that regulate early-stage HIV-1 replication. *Cell*, 135(1):49–60.

Kopecky-Bromberg, S. A., Martínez-Sobrido, L., Frieman, M., Baric, R. A., and Palese, P. (2007). Severe acute respiratory syndrome coronavirus open reading frame (ORF) 3b, ORF 6, and nucleocapsid proteins function as interferon antagonists. *Journal of virology*, 81(2):548–57.

Korf, M., Jarczak, D., Beger, C., Manns, M. P., and Krüger, M. (2005). Inhibition of hepatitis C virus translation and subgenomic replication by siRNAs directed against highly conserved HCV sequence and cellular HCV cofactors. *J Hepatol*, 43(2):225–234.

Kosako, H., Yamaguchi, N., Aranami, C., Ushiyama, M., Kose, S., Imamoto, N., Taniguchi, H., Nishida, E., and Hattori, S. (2009). Phosphoproteomics reveals new ERK MAP kinase targets and links ERK to nucleoporin-mediated nuclear transport. *Nature structural & molecular biology*, 16(10):1026–1035.

Krull, S., Dörries, J., Boysen, B., Reidenbach, S., Magnius, L., Norder, H., Thyberg, J., and Cordes, V. C. (2010). Protein Tpr is required for establishing nuclear pore-associated zones of heterochromatin exclusion. *The EMBO journal*, 29(10):1659–73.

Kudo, N., Matsumori, N., Taoka, H., Fujiwara, D., Schreiner, E. P., Wolff, B., Yoshida, M., and Horinouchi, S. (1999). Leptomycin B inactivates CRM1/exportin 1 by covalent modification at a cysteine residue in the central conserved region. *Proceedings of the National Academy of Sciences of the United States of America*, 96(16):9112–7.

Kuhn, R. J., Zhang, W., Rossmann, M. G., Pletnev, S. V., Corver, J., Lenches, E., Jones, C. T., Mukhopadhyay, S., Chipman, P. R., Strauss, E. G., Baker, T. S., and Strauss, J. H. (2002). Structure of Dengue Virus: Implications for Flavivirus Organization, Maturation, and Fusion. *Cell*, 108(5):717–725.

Kumar, A., Hou, S., Airo, A. M., Limonta, D., Mancinelli, V., Branton, W., Power, C., and Hobman, T. C. (2016). Zika virus inhibits typeI interferon production and downstream signaling. *EMBO reports*, 17(12):1766–1775.

Kuss, S. K., Mata, M. A., Zhang, L., and Fontoura, B. M. A. (2013). Nuclear imprisonment: viral strategies to arrest host mRNA nuclear export. *Viruses*, 5(7):1824–49.

Kwong, A. D., Rao, B. G., and Jeang, K.-T. (2005). Viral and cellular RNA helicases as antiviral targets. *Nature reviews. Drug discovery*, 4(10):845–53.

Labokha, A. A., Gradmann, S., Frey, S., Hülsmann, B. B., Urlaub, H., Baldus, M., and Görlich, D. (2013). Systematic analysis of barrier-forming FG hydrogels from Xenopus nuclear pore complexes. *The EMBO journal*, 32(2):204–18.

Lai, F., Drakas, R., and Nishikura, K. (1995). Mutagenic analysis of double-stranded RNA adenosine deaminase, a candidate enzyme for RNA editing of glutamate-gated ion channel transcripts. *The Journal of biological chemistry*, 270(29):17098–105.

Lange, A., Mills, R. E., Lange, C. J., Stewart, M., Devine, S. E., and Corbett, A. H. (2007). Classical nuclear localization signals: definition, function, and interaction with importin alpha. *The Journal of biological chemistry*, 282(8):5101–5.

Laurell, E., Beck, K., Krupina, K., Theerthagiri, G., Bodenmiller, B., Horvath, P., Aebersold, R., Antonin, W., and Kutay, U. (2011). Phosphorylation of Nup98 by multiple kinases is crucial for NPC disassembly during mitotic entry. *Cell*, 144(4):539–550.

Lavillette, D., Pécheur, E.-I., Donot, P., Fresquet, J., Molle, J., Corbau, R., Dreux, M., Penin, F., and Cosset, F.-L. (2007). Characterization of fusion determinants points to the involvement of three discrete regions of both E1 and E2 glycoproteins in the membrane fusion process of hepatitis C virus. *Journal of virology*, 81(16):8752–65.

Lawrence, P. and Rieder, E. (2009). Identification of RNA helicase A as a new host factor in the replication cycle of foot-and-mouth disease virus. *Journal of virology*, 83(21):11356–66.

Le Sage, V. and Mouland, A. J. (2013). Viral subversion of the nuclear pore complex. *Viruses*, 5(8):2019–2042.

Lee, C. G. and Hurwitz, J. (1992). A new RNA helicase isolated from HeLa cells that catalytically translocates in the 3' to 5' direction. *The Journal of biological chemistry*, 267(7):4398–407.

Lee, C. G. and Hurwitz, J. (1993). Human RNA helicase A is homologous to the maleless protein of Drosophila. *The Journal of biological chemistry*, 268(22):16822–30.

Lee, K., Ambrose, Z., Martin, T. D., Oztop, I., Mulky, A., Julias, J. G., Vandegraaff, N., Baumann, J. G., Wang, R., Yuen, W., Takemura, T., Shelton, K., Taniuchi, I., Li, Y., Sodroski, J., Littman, D. R., Coffin, J. M., Hughes, S. H., Unutmaz, D., Engelman, A., and KewalRamani, V. N. (2010). Flexible Use of Nuclear Import Pathways by HIV-1. *Cell Host & Microbe*, 7(3):221–233.

Lee, S. J., Matsuura, Y., Liu, S. M., and Stewart, M. (2005). Structural basis for nuclear import complex dissociation by RanGTP. *Nature*, 435(7042):693–696.

Lee, S. J., Sekimoto, T., Yamashita, E., Nagoshi, E., Nakagawa, A., Imamoto, N., Yoshimura, M., Sakai, H., Chong, K. T., Tsukihara, T., and Yoneda, Y. (2003). The Structure of Importin- Bound to SREBP-2: Nuclear Import of a Transcription Factor. *Science*, 302(5650):1571–1575.

Lee, T. and Pelletier, J. (2016). The biology of DHX9 and its potential as a therapeutic target. *Oncotarget*.

Lelek, M., Casartelli, N., Pellin, D., Rizzi, E., Souque, P., Severgnini, M., Di Serio, C., Fricke, T., Diaz-Griffero, F., Zimmer, C., Charneau, P., and Di Nunzio, F. (2015). Chromatin organization at the nuclear pore favours HIV replication. *Nature communications*, 6:6483.

Lemke, E. A. (2016). The Multiple Faces of Disordered Nucleoporins. *Journal of Molecular Biology*, 428(10):2011–2024.

Levin, A., Horn, C., Berry-Wynne, K., Aitchison, J. D., Wozniak, R. W., and Tyrrell, D. L. J. (2014a). Nuclear Localization of Hepatitis C Virus Core Protein Early in Infection. In *21st International Symposium on Hepatitis C Virus and Related Viruses*, page 129, Banff, AB, Canada.

Levin, A., Neufeldt, C. J., Pang, D., Wilson, K., Loewen-Dobler, D., Joyce, M. A., Wozniak, R. W., and Tyrrell, D. L. J. (2014b). Functional characterization of nuclear localization and export signals in hepatitis C virus proteins and their role in the membranous web. *PLoS One*, 9(12):e114629.

Li, C., Goryaynov, A., and Yang, W. (2016a). The selective permeability barrier in the nuclear pore complex. *Nucleus*, 7(5):430–446.

Li, C., Xu, D., Ye, Q., Hong, S., Jiang, Y., Liu, X., Zhang, N., Shi, L., Qin, C.-F., and Xu, Z. (2016b). Zika Virus Disrupts Neural Progenitor Development and Leads to Microcephaly in Mice.

Li, G., Poulsen, M., Fenyvuesvolgyi, C., Yashiroda, Y., Yoshida, M., Simard, J. M., Gallo, R. C., and Zhao, R. Y. (2017). Characterization of cytopathic factors through genome-wide analysis of the Zika viral proteins in fission yeast. *Proceedings of the National Academy of Sciences of the United States of America*, 114(3):E376–E385.

Li, J., Tang, H., Mullen, T.-M. M., Westberg, C., Reddy, T. R., Rose, D. W., and Wong-Staal, F. (1999). A role for RNA helicase A in post-transcriptional regulation of HIV type 1. *Proceedings of the National Academy of Sciences of the United States of America*, 96(2):709 –714.

Li, Y., Masaki, T., Shimakami, T., and Lemon, S. M. (2014). hnRNP L and NF90 interact with hepatitis C virus 5'-terminal untranslated RNA and promote efficient replication. *Journal of virology*, 88(13):7199–209.

Li, Y., Yamane, D., Masaki, T., and Lemon, S. M. (2015). The yin and yang of hepatitis C: synthesis and decay of hepatitis C virus RNA. *Nature Reviews Microbiology*, 13(9):544–558.

Li, Z. and Nagy, P. D. (2011). Diverse roles of host RNA binding proteins in RNA virus replication. *RNA Biology*, 8(2):305–315.

Liang, Q., Luo, Z., Zeng, J., Chen, W., Foo, S.-S., Lee, S.-A., Ge, J., Wang, S., Goldman, S., Zlokovic, B., Zhao, Z., and Jung, J. (2016). Zika Virus NS4A and NS4B Proteins Deregulate Akt-mTOR Signaling in Human Fetal Neural Stem Cells to Inhibit Neurogenesis and Induce Autophagy.

Liang, Y., Franks, T. M., Marchetto, M. C., Gage, F. H., and Hetzer, M. W. (2013). Dynamic Association of NUP98 with the Human Genome. *PLoS Genet*, 9(2):e1003308.

Liang, Y. and Hetzer, M. W. (2011). Functional interactions between nucleoporins and chromatin. *Current opinion in cell biology*, 23(1):65–70.

Licatalosi, D. D., Mele, A., Fak, J. J., Ule, J., Kayikci, M., Chi, S. W., Clark, T. A., Schweitzer, A. C., Blume, J. E., Wang, X., Darnell, J. C., and Darnell, R. B. (2008). HITS-CLIP yields genome-wide insights into brain alternative RNA processing. *Nature*, 456(7221):464–469.

Lichinchi, G., Zhao, B., Wu, Y., Lu, Z., Qin, Y., He, C., and Rana, T. (2016). Dynamics of Human and Viral RNA Methylation during Zika Virus Infection.

Light, W. H., Brickner, D. G., Brand, V. R., and Brickner, J. H. (2010). Interaction of a DNA zip code with the nuclear pore complex promotes H2A.Z incorporation and INO1 transcriptional memory. *Molecular cell*, 40(1):112–25.

Light, W. H., Freaney, J., Sood, V., Thompson, A., D'Urso, A., Horvath, C. M., and Brickner, J. H. (2013). A conserved role for human nup98 in altering chromatin structure and promoting epigenetic transcriptional memory. *PLoS Biol*, 11(3):e1001524.

Lim, R. Y. H., Fahrenkrog, B., Koser, J., Schwarz-Herion, K., Deng, J., and Aebi, U. (2007). Nanomechanical Basis of Selective Gating by the Nuclear Pore Complex. *Science*, 318(5850):640–643.

Lim, R. Y. H., Huang, B., and Kapinos, L. E. (2015). How to operate a nuclear pore complex by Kap-centric control. *Nucleus (Austin, Tex.)*, 6(5):366–72.

Lim, R. Y. H., Huang, N.-P., Koser, J., Deng, J., Lau, K. H. A., Schwarz-Herion, K., Fahrenkrog, B., and Aebi, U. (2006). Flexible phenylalanine-glycine nucleoporins as entropic barriers to nucleocytoplasmic transport. *Proceedings of the National Academy of Sciences*, 103(25):9512–9517.

Lin, D. H., Zimmermann, S., Stuwe, T., Stuwe, E., and Hoelz, A. (2013). Structural and functional analysis of the C-terminal domain of Nup358/RanBP2. *Journal of molecular biology*, 425(8):1318–29.

Lin, L., Li, Y., Pyo, H.-M., Lu, X., Raman, S. N. T., Liu, Q., Brown, E. G., and Zhou, Y. (2012). Identification of RNA helicase A as a cellular factor that interacts with influenza A virus NS1 protein and its role in the virus life cycle. *Journal of virology*, 86(4):1942–54.

Lin, S.-Y., Makino, K., Xia, W., Matin, A., Wen, Y., Kwong, K. Y., Bourguignon, L., and Hung, M.-C. (2001). Nuclear localization of EGF receptor and its potential new role as a transcription factor. *Nature Cell Biology*, 3(9):802–808.

Linder, P. and Jankowsky, E. (2011). From unwinding to clamping - the DEAD box RNA helicase family. *Nat Rev Mol Cell Biol*, 12(8):505–516.

Liu, H., Sadygov, R. G., and Yates, J. R. (2004). A model for random sampling and estimation of relative protein abundance in shotgun proteomics. *Anal Chem*, 76(14):4193–4201.

Liu, Z., Kenworthy, R., Green, C., and Tang, H. (2007). Molecular determinants of nucleolar translocation of RNA helicase A. *Experimental cell research*, 313(17):3743–54.

Lloyd, R. E. (2015). Nuclear proteins hijacked by mammalian cytoplasmic plus strand RNA viruses. *Virology*, 479-480:1–18.

Lohmann, V., Körner, F., Koch, J., Herian, U., Theilmann, L., and Bartenschlager, R. (1999). Replication of subgenomic hepatitis C virus RNAs in a hepatoma cell line. *Science (New York, N.Y.)*, 285(5424):110–3.

Loïodice, I., Alves, A., Rabut, G., Van Overbeek, M., Ellenberg, J., Sibarita, J.-B., and Doye, V. (2004). The entire Nup107-160 complex, including three new members, is targeted as one entity to kinetochores in mitosis. *Molecular biology of the cell*, 15(7):3333–44.

Lopez-Denman, A. J. and Mackenzie, J. M. (2017). The IMPORTance of the Nucleus during Flavivirus Replication. *Viruses*, 9(1).

Lorgeoux, R.-P. P., Guo, F., and Liang, C. (2012). From promoting to inhibiting: diverse roles of helicases in HIV-1 Replication. *Retrovirology*, 9:79.

Lu, H., Li, W., Noble, W. S., Payan, D., and Anderson, D. C. (2004). Riboproteomics of the hepatitis C virus internal ribosomal entry site. *Journal of proteome research*, 3(5):949–57.

Lu, H., Lu, N., Weng, L., Yuan, B., Liu, Y. J., and Zhang, Z. (2014). DHX15 senses double-stranded RNA in myeloid dendritic cells. *J Immunol*, 193(3):1364–1372.

Lusk, C. P., Waller, D. D., Makhnevych, T., Dienemann, A., Whiteway, M., Thomas, D. Y., and Wozniak, R. W. (2007). Nup53p is a target of two mitotic kinases, Cdk1p and Hrr25p. *Traffic*, 8(6):647–660.

Lussi, Y. C., Hügi, I., Laurell, E., Kutay, U., and Fahrenkrog, B. (2011). The nucleoporin Nup88 is interacting with nuclear lamin A. *Molecular biology of the cell*, 22(7):1080–90.

Lussignol, M., Kopp, M., Molloy, K., Vizcay-Barrena, G., Fleck, R. A., Dorner, M., Bell, K. L., Chait, B. T., Rice, C. M., and Catanese, M. T. (2016). Proteomics of HCV virions reveals an essential role for the nucleoporin Nup98 in virus morphogenesis. *Proceedings of the National Academy of Sciences of the United States of America*, 113(9):2484–9.

Macara, I. G. (2001). Transport into and out of the Nucleus. *Microbiology and Molecular Biology Reviews*, 65(4):570–594.

Macaulay, C., Meier, E., and Forbes, D. J. (1995). Differential mitotic phosphorylation of proteins of the nuclear pore complex. *The Journal of biological chemistry*, 270(1):254–62.

Mackay, D. R. and Ullman, K. S. (2011). Coordinating postmitotic nuclear pore complex assembly with abscission timing. *Nucleus (Austin, Tex.)*, 2(4):283–8.

Maere, S., Heymans, K., and Kuiper, M. (2005). BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. *Bioinformatics*, 21(16):3448–3449.

Maeshima, K., Iino, H., Hihara, S., Funakoshi, T., Watanabe, A., Nishimura, M., Nakatomi, R., Yahata, K., Imamoto, F., Hashikawa, T., Yokota, H., and Imamoto, N. (2010). Nuclear pore formation but not nuclear growth is governed by cyclin-dependent kinases (Cdks) during interphase. *Nature Structural & Molecular Biology*, 17(99):1065–1071.

Maeshima, K., Yahata, K., Sasaki, Y., Nakatomi, R., Tachibana, T., Hashikawa, T., Imamoto, F., and Imamoto, N. (2006). Cell-cycle-dependent dynamics of nuclear pores: pore-free islands and lamins. *Journal of cell science*, 119(Pt 21):4442–51.

Mahajan, R., Delphin, C., Guan, T., Gerace, L., and Melchior, F. (1997). A small ubiquitin-related polypeptide involved in targeting RanGAP1 to nuclear pore complex protein RanBP2. *Cell*, 88(1):97–107.

Maimon, T., Elad, N., Dahan, I., and Medalia, O. (2012). The Human Nuclear Pore Complex as Revealed by Cryo-Electron Tomography. *Structure*, 20(6):998–1006.

Makhnevych, T., Lusk, C. P., Anderson, A. M., Aitchison, J. D., and Wozniak, R. W. (2003). Cell cycle regulated transport controlled by alterations in the nuclear pore complex. *Cell*, 115(7):813–23.

Manojlovic, Z. and Stefanovic, B. (2012). A novel role of RNA helicase A in regulation of translation of type I collagen mRNAs. *RNA*, 18(2):321–334.

Mansfeld, J., Güttinger, S., Hawryluk-Gara, L. A., Panté, N., Mall, M., Galy, V., Haselmann, U., Mühlhäusser, P., Wozniak, R. W., Mattaj, I. W., Kutay, U., and Antonin, W. (2006). The conserved transmembrane nucleoporin NDC1 is required for nuclear pore complex assembly in vertebrate cells. *Molecular cell*, 22(1):93–103.

Marini, B., Kertesz-Farkas, A., Ali, H., Lucic, B., Lisek, K., Manganaro, L., Pongor, S., Luzzati, R., Recchia, A., Mavilio, F., Giacca, M., and Lusic, M. (2015). Nuclear architecture dictates HIV-1 integration site selection. *Nature*, 521(7551):227–231.

Marshall, J. A., Borg, J., Coulepis, A. G., and Anderson, D. A. (1996). Annulate lamellae and lytic HAV infection in vitro. *Tissue & cell*, 28(2):205–14.

Martino, L., Pennell, S., Kelly, G., Bui, T. T., Kotik-Kogan, O., Smerdon, S. J., Drake, A. F., Curry, S., and Conte, M. R. (2012). Analysis of the interaction with the hepatitis C virus mRNA reveals an alternative mode of RNA recognition by the human La protein. *Nucleic Acids Res*, 40(3):1381–1394.

Mastrangelo, E., Pezzullo, M., De Burghgraeve, T., Kaptein, S., Pastorino, B., Dallmeier, K., de Lamballerie, X., Neyts, J., Hanson, A. M., Frick, D. N., Bolognesi, M., and Milani, M. (2012). Ivermectin is a potent inhibitor of flavivirus replication specifically targeting NS3 helicase activity: new prospects for an old drug. *The Journal of antimicrobial chemotherapy*, 67(8):1884–94.

Mateo, M., Reid, S. P., Leung, L. W., Basler, C. F., and Volchkov, V. E. (2010). Ebolavirus VP24 binding to karyopherins is required for inhibition of interferon signaling. *Journal of virology*, 84(2):1169–75.

Matreyek, K. A. and Engelman, A. (2011). The requirement for nucleoporin NUP153 during human immunodeficiency virus type 1 infection is determined by the viral capsid. *Journal of virology*, 85(15):7818–27.

Matsuoka, S., Ballif, B. A., Smogorzewska, A., McDonald, E. R., Hurov, K. E., Luo, J., Bakalarski, C. E., Zhao, Z., Solimini, N., Lerenthal, Y., Shiloh, Y., Gygi, S. P., and Elledge, S. J. (2007). ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage. *Science (New York, N.Y.)*, 316(5828):1160–6.

Matsuura, Y. (2016). Mechanistic Insights from Structural Analyses of Ran-GTPase-Driven Nuclear Export of Proteins and RNAs. *Journal of Molecular Biology*, 428(10):2025–2039.

Matunis, M. J., Coutavas, E., and Blobel, G. (1996). A novel ubiquitin-like modification modulates the partitioning of the Ran-GTPase-activating protein RanGAP1 between the cytosol and the nuclear pore complex. *The Journal of cell biology*, 135(6 Pt 1):1457–70.

Mazloom, A. R., Dannenfelser, R., Clark, N. R., Grigoryan, A. V., Linder, K. M., Cardozo, T. J., Bond, J. C., Boran, A. D., Iyengar, R., Malovannaya, A., Lanz, R. B., and Ma'ayan, A. (2011). Recovering protein-protein and domain-domain interactions from aggregation of IP-MS proteomics of coregulator complexes. *PLoS Comput Biol*, 7(12):e1002319.

Melcák, I., Hoelz, A., and Blobel, G. (2007). Structure of Nup58/45 suggests flexible nuclear pore diameter by intermolecular sliding. *Science (New York, N.Y.)*, 315(5819):1729–32.

Mellacheruvu, D., Wright, Z., Couzens, A. L., Lambert, J. P., St-Denis, N. A., Li, T., Miteva, Y. V., Hauri, S., Sardiu, M. E., Low, T. Y., Halim, V. A., Bagshaw, R. D., Hubner, N. C., Al-Hakim, A., Bouchard, A., Faubert, D., Fermin, D., Dunham, W. H., Goudreault, M., Lin, Z. Y., Badillo, B. G., Pawson, T., Durocher, D., Coulombe, B., Aebersold, R., Superti-Furga, G., Colinge, J., Heck, A. J., Choi, H., Gstaiger, M., Mohammed, S., Cristea, I. M., Bennett, K. L., Washburn, M. P., Raught, B., Ewing, R. M., Gingras, A. C., and Nesvizhskii, A. I. (2013). The CRAPome: a contaminant repository for affinity purification-mass spectrometry data. *Nat Methods*, 10(8):730–736.

Mendjan, S., Taipale, M., Kind, J., Holz, H., Gebhardt, P., Schelder, M., Vermeulen, M., Buscaino, A., Duncan, K., Mueller, J., Wilm, M., Stunnenberg, H. G., Saumweber, H., and Akhtar, A. (2006). Nuclear pore components are involved in the transcriptional regulation of dosage compensation in Drosophila. *Molecular cell*, 21(6):811–23.

Mettenleiter, T. C. (2016). Breaching the BarrierThe Nuclear Envelope in Virus Infection. *Journal of Molecular Biology*, 428(10):1949–1961.

Micallef, J. M., Kaldor, J. M., and Dore, G. J. (2006). Spontaneous viral clearance following acute hepatitis C infection: a systematic review of longitudinal studies. *Journal of Viral Hepatitis*, 13(1):34–41.

Mincer, J. S. and Simon, S. M. (2011). Simulations of nuclear pore transport yield mechanistic insights and quantitative predictions. *Proceedings of the National Academy of Sciences of the United States of America*, 108(31):351–8.

Ming, G.-l., Tang, H., and Song, H. (2016). Advances in Zika Virus Research: Stem Cell Models, Challenges, and Opportunities. *Cell Stem Cell*, 19(6):690–702.

Mischo, H. E., Hemmerich, P., Grosse, F., and Zhang, S. (2005). Actinomycin D induces histone gamma-H2AX foci and complex formation of gamma-H2AX with Ku70 and nuclear DNA helicase II. *The Journal of biological chemistry*, 280(10):9586–94.

Mishra, R. K., Chakraborty, P., Arnaoutov, A., Fontoura, B. M. A., and Dasso, M. (2010). The Nup107-160 complex and gamma-TuRC regulate microtubule polymerization at kinetochores. *Nature cell biology*, 12(2):164–9.

Mitchell, J. M., Mansfeld, J., Capitanio, J., Kutay, U., and Wozniak, R. W. (2010). Pom121 links two essential subcomplexes of the nuclear pore complex core to the membrane. *The Journal of cell biology*, 191(3):505–21.

Mohd Hanafiah, K., Groeger, J., Flaxman, A. D., and Wiersma, S. T. (2013). Global epidemiology of hepatitis C virus infection: new estimates of age-specific antibody to HCV seroprevalence. *Hepatology (Baltimore, Md.)*, 57(4):1333–42.

Monette, A., Panté, N., and Mouland, A. J. (2011). HIV-1 remodels the nuclear pore complex. *The Journal of cell biology*, 193(4):619–31.

Montague, E., Janko, I., Stanberry, L., Lee, E., Choiniere, J., Anderson, N., Stewart, E., Broomall, W., Higdon, R., Kolker, N., and Kolker, E. (2015). Beyond protein expression, MOPED goes multi-omics. *Nucleic acids research*, 43(Database issue):1145–51.

Montpetit, B., Seeliger, M. A., and Weis, K. (2012). Analysis of DEAD-box proteins in mRNA export. *Methods Enzymol*, 511:239–254.

Montpetit, B., Thomsen, N. D., Helmke, K. J., Seeliger, M. A., Berger, J. M., and Weis, K. (2011). A conserved mechanism of DEAD-box ATPase activation by nucleoporins and InsP6 in mRNA export. *Nature*, 472(7342):238–242.

Moore, M. A. S., Chung, K. Y., Plasilova, M., Schuringa, J. J., Shieh, J.-H., Zhou, P., and Morrone, G. (2007). NUP98 Dysregulation in Myeloid Leukemogenesis. *Annals of the New York Academy of Sciences*, 1106(Hematopoietic Stem Cells VI):114–142.

Moore, M. J., Zhang, C., Gantman, E. C., Mele, A., Darnell, J. C., and Darnell, R. B. (2014). Mapping Argonaute and conventional RNA-binding protein interactions with RNA at single-nucleotide resolution using HITS-CLIP and CIMS analysis. *Nature Protocols*, 9(2):263–293.

Moorhead, G. B., Trinkle-Mulcahy, L., Nimick, M., De Wever, V., Campbell, D. G., Gourlay, R., Lam, Y., and Lamond, A. I. (2008). Displacement affinity chromatography of protein phosphatase one (PP1) complexes. *BMC Biochemistry*, 9(1):28.

Morchoisne-Bolhy, S., Geoffroy, M.-C., Bouhlel, I. B., Alves, A., Audugé, N., Baudin, X., Van Bortle, K., Powers, M. A., and Doye, V. (2015). Intranuclear dynamics of the Nup107-160 complex. *Molecular biology of the cell*, 26(12):2343–56.

Morra, R., Smith, E. R., Yokoyama, R., and Lucchesi, J. C. (2008). The MLE subunit of the Drosophila MSL complex uses its ATPase activity for dosage compensation and its helicase activity for targeting. *Molecular and cellular biology*, 28(3):958–66.

Morra, R., Yokoyama, R., Ling, H., and Lucchesi, J. C. (2011). Role of the ATPase/helicase maleless (MLE) in the assembly, targeting, spreading and function of the male-specific lethal (MSL) complex of Drosophila. *Epigenetics & Chromatin*, 4(1):6.

Musser, S. M. and Grünwald, D. (2016). Deciphering the Structure and Function of Nuclear Pores Using Single-Molecule Fluorescence Approaches. *Journal of molecular biology*, 428(10 Pt A):2091–119.

Musso, D., Nilles, E., and Cao-Lormeau, V.-M. (2014). Rapid spread of emerging Zika virus in the Pacific area. *Clinical Microbiology and Infection*, 20(10):O595–O596.

Musso, D., Roche, C., Robin, E., Nhan, T., Teissier, A., and Cao-Lormeau, V.-M. (2015). Potential Sexual Transmission of Zika Virus. *Emerging Infectious Diseases*, 21(2):359–361.

Myöhänen, S. and Baylin, S. B. (2001). Sequence-specific DNA binding activity of RNA helicase A to the p16INK4a promoter. *The Journal of biological chemistry*, 276(2):1634–42.

Nagai, S., Dubrana, K., Tsai-Pflugfelder, M., Davidson, M. B., Roberts, T. M., Brown, G. W., Varela, E., Hediger, F., Gasser, S. M., and Krogan, N. J. (2008). Functional targeting of DNA damage to a nuclear pore-associated SUMO-dependent ubiquitin ligase. *Science (New York, N.Y.)*, 322(5901):597–602.

Nagata, T., Tsuda, K., Kobayashi, N., Shirouzu, M., Kigawa, T., Güntert, P., Yokoyama, S., and Muto, Y. (2012). Solution structures of the double-stranded RNA-binding domains from RNA helicase A. *Proteins*, 80(6):1699–706.

Nagy, P. D. and Pogany, J. (2011). The dependence of viral RNA replication on co-opted host factors. *Nature reviews. Microbiology*, 10(2):137–49.

Nakajima, T., Uchida, C., Anderson, S. F., Lee, C. G., Hurwitz, J., Parvin, J. D., and Montminy, M. (1997). RNA helicase A mediates association of CBP with RNA polymerase II. *Cell*, 90(6):1107–1112.

Napetschnig, J., Kassube, S. A., Debler, E. W., Wong, R. W., Blobel, G., and Hoelz, A. (2009). Structural and functional analysis of the interaction between the nucleoporin Nup214 and the DEAD-box helicase Ddx19. *Proc Natl Acad Sci U S A*, 106(9):3089–3094.

Nemergut, M. E., Mizzen, C. A., Stukenberg, T., Allis, C. D., and Macara, I. G. (2001). Chromatin Docking and Exchange Activity Enhancement of RCC1 by Histones H2A and H2B. *Science*, 292(5521):1540–1543.

Neufeldt, C. J. (2014). *A role for the nuclear transport machinery in supporting positive-strand RNA virus infection and in regulating innate immune responses by*. PhD thesis.

Neufeldt, C. J., Joyce, M. A., Levin, A., Steenbergen, R. H., Pang, D., Shields, J., Tyrrell, D. L. J., and Wozniak, R. W. (2013). Hepatitis C virus-induced cytoplasmic organelles use the nuclear transport machinery to establish an environment conducive to virus replication. *PLoS pathogens*, 9(10):e1003744.

Neufeldt, C. J., Joyce, M. A., Van Buuren, N., Levin, A., Kirkegaard, K., Gale, M., Tyrrell, D. L. J., and Wozniak, R. W. (2016). The Hepatitis C Virus-Induced Membranous Web and Associated Nuclear Transport Machinery Limit Access of Pattern Recognition Receptors to Viral Replication Sites. *PLoS pathogens*, 12(2):e1005428.

Noble, K. N., Tran, E. J., Alcázar-Román, A. R., Hodge, C. A., Cole, C. N., and Wente, S. R. (2011). The Dbp5 cycle at the nuclear pore complex during mRNA export II: nucleotide cycling and mRNP remodeling by Dbp5 are controlled by Nup159 and Gle1. *Genes & development*, 25(10):1065–77.

Nofrini, V., Di Giacomo, D., and Mecucci, C. (2016). Nucleoporin genes in human diseases. *European Journal of Human Genetics*, 24(10):1388–1395.

Nowakowski, T., Pollen, A., DiLullo, E., Sandoval-Espinosa, C., Bershteyn, M., and Kriegstein, A. (2016). Expression Analysis Highlights AXL as a Candidate Zika Virus Entry Receptor in Neural Stem Cells.

Ohtsubo, M., Okazaki, H., and Nishimoto, T. (1989). The RCC1 protein, a regulator for the onset of chromosome condensation locates in the nucleus and binds to DNA. *The Journal of cell biology*, 109(4 Pt 1):1389–97.

Ojala, P. M., Sodeik, B., Ebersold, M. W., Kutay, U., and Helenius, A. (2000). Herpes simplex virus type 1 entry into host cells: reconstitution of capsid binding and uncoating at the nuclear pore complex in vitro. *Molecular and cellular biology*, 20(13):4922–31.

Oka, M., Asally, M., Yasuda, Y., Ogawa, Y., Tachibana, T., and Yoneda, Y. (2010). The mobile FG nucleoporin Nup98 is a cofactor for Crm1-dependent protein export. *Mol Biol Cell*, 21(11):1885–1896.

Okamura, M., Inose, H., and Masuda, S. (2015). RNA Export through the NPC in Eukaryotes. *Genes*, 6(1):124–149.

Palancade, B., Liu, X., Garcia-Rubio, M., Aguilera, A., Zhao, X., and Doye, V. (2007). Nucleoporins prevent DNA damage accumulation by modulating Ulp1-dependent sumoylation processes. *Molecular biology of the cell*, 18(8):2912–23.

Pan American Health Organization and World Health Organization (2017). Regional Zika Epidemiological Update (Americas) March 10, 2017.

Panaretou, B., Prodromou, C., Roe, S. M., O'Brien, R., Ladbury, J. E., Piper, P. W., and Pearl, L. H. (1998). ATP binding and hydrolysis are essential to the function of the Hsp90 molecular chaperone in vivo. *EMBO J*, 17(16):4829–4836.

Panda, D., Gold, B., Tartell, M. A., Rausch, K., Casas-Tinto, S., and Cherry, S. (2015). The Transcription Factor FoxK Participates with Nup98 To Regulate Antiviral Gene Expression. *mBio*, 6(2):02509–14.

Panda, D., Pascual-Garcia, P., Dunagin, M., Tudor, M., Hopkins, K. C., Xu, J., Gold, B., Raj, A., Capelson, M., and Cherry, S. (2014). Nup98 promotes antiviral gene expression to restrict RNA viral infection in Drosophila. *Proceedings of the National Academy of Sciences of the United States of America*, 111(37):3890–9.

Panté, N. and Kann, M. (2002). Nuclear pore complex is able to transport macromolecules with diameters of about 39 nm. *Molecular biology of the cell*, 13(2):425–34.

Paoli, M. (2001). Protein folds propelled by diversity. *Progress in biophysics and molecular biology*, 76(1-2):103–30.

Park, N., Katikaneni, P., Skern, T., and Gustin, K. E. (2008). Differential targeting of nuclear pore complex proteins in poliovirus-infected cells. *Journal of virology*, 82(4):1647–55.

Park, N., Skern, T., and Gustin, K. E. (2010). Specific Cleavage of the Nuclear Pore Complex Protein Nup62 by a Viral Protease. *Journal of Biological Chemistry*, 285(37):28796–28805.

Pascual-Garcia, P., Jeong, J., and Capelson, M. (2014). Nucleoporin Nup98 Associates with Trx/MLL and NSL Histone-Modifying Complexes and Regulates Hox Gene Expression. *Cell Rep*, 9(2):433–442.

Pasdeloup, D., Blondel, D., Isidro, A. L., and Rixon, F. J. (2009). Herpesvirus capsid association with the nuclear pore complex and viral DNA release involve the nucleoporin CAN/Nup214 and the capsid protein pUL25. *Journal of virology*, 83(13):6610–23.

Patel, R. C. and Sen, G. C. (1992). Identification of the double-stranded RNA-binding domain of the human interferon-inducible protein kinase. *The Journal of biological chemistry*, 267(11):7671–6.

Patel, S. S., Belmont, B. J., Sante, J. M., and Rexach, M. F. (2007). Natively Unfolded Nucleoporins Gate Protein Diffusion across the Nuclear Pore Complex. *Cell*, 129(1):83–96.

Patel, S. S. and Rexach, M. F. (2008). Discovering novel interactions at the nuclear pore complex using bead halo: a rapid method for detecting molecular interactions of high and low affinity at equilibrium. *Mol Cell Proteomics*, 7(1):121–131.

Paul, S., Dansithong, W., Jog, S. P., Holt, I., Mittal, S., Brook, J. D., Morris, G. E., Comai, L., and Reddy, S. (2011). Expanded CUG repeats Dysregulate RNA splicing by altering the stoichiometry of the muscleblind 1 complex. *The Journal of biological chemistry*, 286(44):38427–38.

Paulillo, S. M., Phillips, E. M., Köser, J., Sauder, U., Ullman, K. S., Powers, M. A., and Fahrenkrog, B. (2005). Nucleoporin Domain Topology is Linked to the Transport Status of the Nuclear Pore Complex. *Journal of Molecular Biology*, 351(4):784–798.

Pause, A., Kukolj, G., Bailey, M., Brault, M., Do, F., Halmos, T., Lagace, L., Maurice, R., Marquis, M., McKercher, G., Pellerin, C., Pilote, L., Thibeault, D., and Lamarre, D. (2003). An NS3 Serine Protease Inhibitor Abrogates Replication of Subgenomic Hepatitis C Virus RNA. *Journal of Biological Chemistry*, 278(22):20374–20380.

Pause, A., Méthot, N., and Sonenberg, N. (1993). The HRIGRXXR region of the DEAD box RNA helicase eukaryotic translation initiation factor 4A is required for RNA binding and ATP hydrolysis. *Molecular and cellular biology*, 13(11):6789–98.

Pause, A. and Sonenberg, N. (1992). Mutational analysis of a DEAD box RNA helicase: the mammalian translation initiation factor eIF-4A. *The EMBO journal*, 11(7):2643–54.

Payne, C., Rawe, V., Ramalho-Santos, J., Simerly, C., and Schatten, G. (2003). Preferentially localized dynein and perinuclear dynactin associate with nuclear pore complex proteins to mediate genomic union during mammalian fertilization. *Journal of cell science*, 116(Pt 23):4727–38.

Pellizzoni, L., Baccon, J., Charroux, B., and Dreyfuss, G. (2001a). The survival of motor neurons (SMN) protein interacts with the snoRNP proteins fibrillarin and GAR1. *Current biology : CB*, 11(14):1079–88.

Pellizzoni, L., Charroux, B., Rappsilber, J., Mann, M., and Dreyfuss, G. (2001b). A functional interaction between the survival motor neuron complex and RNA polymerase II. *J Cell Biol*, 152(1):75–85.

Pemberton, L. F. and Paschal, B. M. (2005). Mechanisms of receptor-mediated nuclear import and nuclear export. *Traffic (Copenhagen, Denmark)*, 6(3):187–98.

Peters, R. (2005). Translocation Through the Nuclear Pore Complex: Selectivity and Speed by Reduction-of-Dimensionality. *Traffic*, 6(5):421–427.

Pfaller, R., Smythe, C., and Newport, J. W. (1991). Assembly/disassembly of the nuclear envelope membrane: cell cycle-dependent binding of nuclear membrane vesicles to chromatin in vitro. *Cell*, 65(2):209–17.

Piccolo, L. L., Corona, D., and Onorati, M. C. (2014). Emerging Roles for hnRNPs in post-transcriptional regulation: what can we learn from flies? *Chromosoma*, 123(6):515–527.

Pichlmair, A., Kandasamy, K., Alvisi, G., Mulhern, O., Sacco, R., Habjan, M., Binder, M., Stefanovic, A., Eberle, C.-A., Goncalves, A., Bürckstümmer, T., Müller, A. C., Fauster, A., Holze, C., Lindsten, K., Goodbourn, S., Kochs, G., Weber, F., Bartenschlager, R., Bowie, A. G., Bennett, K. L., Colinge, J., and Superti-Furga, G. (2012). Viral immune modulators perturb the human molecular network by common and unique strategies. *Nature*, 487(7408):486–490.

Pietschmann, T., Lohmann, V., Rutter, G., Kurpanek, K., and Bartenschlager, R. (2001). Characterization of Cell Lines Carrying Self-Replicating Hepatitis C Virus RNAs. *Journal of Virology*, 75(3):1252–1264.

Platani, M., Santarella-Mellwig, R., Posch, M., Walczak, R., Swedlow, J. R., and Mattaj, I. W. (2009). The Nup107-160 Nucleoporin Complex Promotes Mitotic Events via Control of the Localization State of the Chromosome Passenger Complex. *Molecular Biology of the Cell*, 20(24):5260–5275.

Porter, F. W., Bochkov, Y. A., Albee, A. J., Wiese, C., and Palmenberg, A. C. (2006). A picornavirus protein interacts with Ran-GTPase and disrupts nucleocytoplasmic transport. *Proceedings of the National Academy of Sciences of the United States of America*, 103(33):12417–22.

Porter, F. W. and Palmenberg, A. C. (2009). Leader-induced phosphorylation of nucleoporins correlates with nuclear trafficking inhibition by cardioviruses. *Journal of virology*, 83(4):1941–51.

Powers, M. A., Forbes, D. J., Dahlberg, J. E., and Lund, E. (1997). The Vertebrate GLFG Nucleoporin, Nup98, Is an Essential Component of Multiple RNA Export Pathways. *J. Cell Biol.*, 136(2):241–250.

Pritchard, C. E., Fornerod, M., Kasper, L. H., and van Deursen, J. M. (1999). RAE1 is a shuttling mRNA export factor that binds to a GLEBS-like NUP98 motif at the nuclear pore complex through multiple domains. *The Journal of Cell Biology*, 145(2):237–254.

Ptak, C., Aitchison, J. D., and Wozniak, R. W. (2014). The multifunctional nuclear pore complex: a platform for controlling gene expression. *Curr Opin Cell Biol*, 28:46–53.

Ptak, C. and Wozniak, R. W. (2016). Nucleoporins and chromatin metabolism. *Current Opinion in Cell Biology*, 40:153–160.

Quan, B., Seo, H.-S., Blobel, G., and Ren, Y. (2014). Vesiculoviral matrix (M) protein occupies nucleic acid binding site at nucleoporin pair (Rae1 Nup98). *Proceedings of the National Academy of Sciences of the United States of America*, 111(25):9127–32.

Quinkert, D., Bartenschlager, R., and Lohmann, V. (2005). Quantitative analysis of the hepatitis C virus replication complex. *Journal of virology*, 79(21):13594–605.

Rabut, G., Doye, V., and Ellenberg, J. (2004). Mapping the dynamic organization of the nuclear pore complex inside single living cells. *Nat Cell Biol*, 6(11):1114–1121.

Radu, A., Moore, M. S., and Blobel, G. (1995). The peptide repeat domain of nucleoporin Nup98 functions as a docking site in transport across the nuclear pore complex. *Cell*, 81(2):215–222.

Raices, M. and D'Angelo, M. A. (2012). Nuclear pore complex composition: a new regulator of tissue-specific and developmental functions. *Nat Rev Mol Cell Biol*, 13(11):687–699.

Ramanan, V., Trehan, K., Ong, M.-L., Luna, J. M., Hoffmann, H.-H., Espiritu, C., Sheahan, T. P., Chandrasekar, H., Schwartz, R. E., Christine, K. S., Rice, C. M., van Oudenaarden, A., and Bhatia, S. N. (2016). Viral genome imaging of hepatitis C virus to probe heterogeneous viral infection and responses to antiviral therapies. *Virology*, 494:236–247.

Ranji, A. and Boris-Lawrie, K. (2010a). RNA helicases. *RNA Biology*, 7(6):775–787.

Ranji, A. and Boris-Lawrie, K. (2010b). RNA helicases: emerging roles in viral replication and the host innate response. *RNA biology*, 7(6):775–87.

Ranji, A., Shkriabai, N., Kvaratskhelia, M., Musier-Forsyth, K., and Boris-Lawrie, K. (2011). Features of double-stranded RNA-binding domains of RNA helicase A are necessary for selective recognition and translation of complex mRNAs. *J Biol Chem*, 286(7):5328–5337.

Ray, P. S., Grover, R., and Das, S. (2006). Two internal ribosome entry sites mediate the translation of p53 isoforms. *EMBO reports*, 7(4):404–10.

Reddy, T. R., Tang, H., Xu, W., and Wong-Staal, F. (2000). Sam68, RNA helicase A and Tap cooperate in the post-transcriptional regulation of human immunodeficiency virus and type D retroviral mRNA. *Oncogene*, 19(32):3570–3575.

Reenan, R. A., Hanrahan, C. J., and Ganetzky, B. (2000). The mle(napts) RNA helicase mutation in drosophila results in a splicing catastrophe of the para Na+ channel transcript in a region of RNA editing. *Neuron*, 25(1):139–149.

Reichman, T. W., Parrott, A. M., Fierro-Monti, I., Caron, D. J., Kao, P. N., Lee, C.-G., Li, H., and Mathews, M. B. (2003). Selective regulation of gene expression by nuclear factor 110, a member of the NF90 family of double-stranded RNA-binding proteins. *Journal of molecular biology*, 332(1):85–98.

Reid, C. R., Airo, A. M., and Hobman, T. C. (2015). The Virus-Host Interplay: Biogenesis of +RNA Replication Complexes. *Viruses*, 7(8):4385–413.

Reid, S. P., Leung, L. W., Hartman, A. L., Martinez, O., Shaw, M. L., Carbonnelle, C., Volchkov, V. E., Nichol, S. T., and Basler, C. F. (2006). Ebola virus VP24 binds karyopherin alpha1 and blocks STAT1 nuclear accumulation. *Journal of virology*, 80(11):5156–67.

Ren, Y., Seo, H. S., Blobel, G., and Hoelz, A. (2010). Structural and functional analysis of the interaction between the nucleoporin Nup98 and the mRNA export factor Rae1. *Proc Natl Acad Sci U S A*, 107(23):10406–10411.

Rexach, M. and Blobel, G. (1995). Protein import into nuclei: association and dissociation reactions involving transport substrate, transport factors, and nucleoporins. *Cell*, 83(5):683–92.

Rey, F. A., Heinz, F. X., Mandl, C., Kunz, C., and Harrison, S. C. (1995). The envelope glycoprotein from tick-borne encephalitis virus at 2 Å resolution. *Nature*, 375(6529):291–298.

Ribbeck, K. and Görlich, D. (2001). Kinetic analysis of translocation through nuclear pore complexes. *The EMBO journal*, 20(6):1320–30.

Rivas-Aravena, A., Ramdohr, P., Vallejos, M., Valiente-Echeverría, F., Dormoy-Raclet, V., Rodríguez, F., Pino, K., Holzmann, C., Huidobro-Toro, J. P., Gallouzi, I. E., and López-Lastra, M. (2009). The Elav-like protein HuR exerts translational control of viral internal ribosome entry sites. *Virology*, 392(2):178–185.

Robb, G. B. and Rana, T. M. (2007). RNA helicase A interacts with RISC in human cells and functions in RISC loading. *Mol Cell*, 26(4):523–537.

Romero-Brey, I., Merz, A., Chiramel, A., Lee, J.-Y., Chlanda, P., Haselman, U., Santarella-Mellwig, R., Habermann, A., Hoppe, S., Kallis, S., Walther, P., Antony, C., Krijnse-Locker, J., and Bartenschlager, R. (2012). Three-dimensional architecture and biogenesis of membrane structures associated with hepatitis C virus replication. *PLoS pathogens*, 8(12):e1003056.

Rosenbloom, K. R., Armstrong, J., Barber, G. P., Casper, J., Clawson, H., Diekhans, M., Dreszer, T. R., Fujita, P. A., Guruvadoo, L., Haeussler, M., Harte, R. A., Heitner, S., Hickey, G., Hinrichs, A. S., Hubley, R., Karolchik, D., Learned, K., Lee, B. T., Li, C. H., Miga, K. H., Nguyen, N., Paten, B., Raney, B. J., Smit, A. F. A., Speir, M. L., Zweig, A. S., Haussler, D., Kuhn, R. M., and Kent, W. J. (2015). The UCSC Genome Browser database: 2015 update. *Nucleic Acids Research*, 43(D1):D670–D681.

Rothballer, A. and Kutay, U. (2013). The diverse functional LINCs of the nuclear envelope to the cytoskeleton and chromatin. *Chromosoma*, 122(5):415–429.

Rout, M. P., Aitchison, J. D., Magnasco, M. O., and Chait, B. T. (2003). Virtual gating and nuclear transport: the hole picture. *Trends in cell biology*, 13(12):622–8.

Rout, M. P., Aitchison, J. D., Suprapto, A., Hjertaas, K., Zhao, Y., and Chait, B. T. (2000). The yeast nuclear pore complex: composition, architecture, and transport mechanism. *The Journal of cell biology*, 148(4):635–51.

Roy, B. B., Hu, J., Guo, X., Russell, R. S., Guo, F., Kleiman, L., and Liang, C. (2006). Association of RNA helicase a with human immunodeficiency virus type 1 particles. *J Biol Chem*, 281(18):12625–12635.

Ryan, C. M., Harries, J. C., Kindle, K. B., Collins, H. M., and Heery, D. M. (2006). Functional interaction of CREB binding protein (CBP) with nuclear transport proteins and modulation by HDAC inhibitors. *Cell cycle (Georgetown, Tex.)*, 5(18):2146–52.

Sadler, A. J., Latchoumanin, O., Hawkes, D., Mak, J., and Williams, B. R. G. (2009). An antiviral response directed by PKR phosphorylation of the RNA helicase A. *PLoS Pathog*, 5(2):e1000311.

Saitoh, H., Pu, R., Cavenagh, M., and Dasso, M. (1997). RanBP2 associates with Ubc9p and a modified form of RanGAP1. *Proceedings of the National Academy of Sciences of the United States of America*, 94(8):3736–41.

Sakiyama, Y., Mazur, A., Kapinos, L. E., and Lim, R. Y. (2016). Spatiotemporal dynamics of the nuclear pore complex transport barrier resolved by high-speed atomic force microscopy. *Nat Nanotechnol*.

Saldi, T., Cortazar, M. A., Sheridan, R. M., and Bentley, D. L. (2016). Coupling of RNA Polymerase II Transcription Elongation with Pre-mRNA Splicing. *Journal of Molecular Biology*, 428(12):2623–2635.

Satterly, N., Tsai, P.-L. P.-L., van Deursen, J., Nussenzveig, D. R., Wang, Y., Faria, P. A., Levay, A., Levy, D. E., and Fontoura, B. M. A. (2007). Influenza virus targets the mRNA export machinery and the nuclear pore complex. *Proceedings of the National Academy of Sciences*, 104(6):1853–1858.

Schaab, C., Geiger, T., Stoehr, G., Cox, J., and Mann, M. (2012). Analysis of high accuracy, quantitative proteomics data in the MaxQB database. *Molecular & cellular proteomics : MCP*, 11(3):M111.014068.

Schagat, T., Paguio, A., and Kopish, K. (2007). Normalizing genetic reporter assays: approaches and considerations for increasing consistency and statistical significance. *Cell Notes, Promega Coorporation, PubHub*, (17):9–12.

Schaller, T., Ocwieja, K. E., Rasaiyaah, J., Price, A. J., Brady, T. L., Roth, S. L., Hué, S., Fletcher, A. J., Lee, K., KewalRamani, V. N., Noursadeghi, M., Jenner, R. G., James, L. C., Bushman, F. D., and Towers, G. J. (2011). HIV-1 capsid-cyclophilin interactions determine nuclear import pathway, integration targeting and replication efficiency. *PLoS pathogens*, 7(12):e1002439.

Scheel, T. K. H., Gottwein, J. M., Jensen, T. B., Prentoe, J. C., Hoegh, A. M., Alter, H. J., Eugen-Olsen, J., and Bukh, J. (2008). Development of JFH1-based cell culture systems for hepatitis C virus genotype 4a and evidence for cross-genotype neutralization. *Proceedings of the National Academy of Sciences of the United States of America*, 105(3):997–1002.

Scheel, T. K. H. and Rice, C. M. (2013). Understanding the hepatitis C virus life cycle paves the way for highly effective therapies. *Nature medicine*, 19(7):837–49.

Schmitz, A., Schwarz, A., Foss, M., Zhou, L., Rabe, B., Hoellenriegel, J., Stoeber, M., Panté, N., and Kann, M. (2010). Nucleoporin 153 arrests the nuclear import of hepatitis B virus capsids in the nuclear basket. *PLoS pathogens*, 6(1):e1000741.

Schneider, C. A., Rasband, W. S., and Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nat Methods*, 9(7):671–675.

Schoggins, J. W., Wilson, S. J., Panis, M., Murphy, M. Y., Jones, C. T., Bieniasz, P., and Rice, C. M. (2011). A diverse range of gene products are effectors of the type I interferon antiviral response. *Nature*, 472(7344):481–485.

Schooley, A., Vollmer, B., and Antonin, W. (2012). Building a nuclear envelope at the end of mitosis: coordinating membrane reorganization, nuclear pore complex assembly, and chromatin de-condensation. *Chromosoma*, 121(6):539–554.

Schütz, P., Wahlberg, E., Karlberg, T., Hammarström, M., Collins, R., Flores, A., and Schüler, H. (2010). Crystal structure of human RNA helicase A (DHX9): structural basis for unselective nucleotide base binding in a DEAD-box variant protein. *J Mol Biol*, 400(4):768–782.

Schwer, B. and Meszaros, T. (2000). RNA helicase dynamics in pre-mRNA splicing. *The EMBO journal*, 19(23):6582–91.

Selvanathan, S. P., Graham, G. T., Erkizan, H. V., Dirksen, U., Natarajan, T. G., Dakic, A., Yu, S., Liu, X., Paulsen, M. T., Ljungman, M. E., Wu, C. H., Lawlor, E. R., Üren, A., and Toretsky, J. A. (2015). Oncogenic fusion protein EWS-FLI1 is a network hub that regulates alternative splicing. *Proc Natl Acad Sci U S A*, 112(11):1307–16.

Sharma, A. and Boris-Lawrie, K. (2012). Determination of host RNA helicases activity in viral replication. *Methods in enzymology*, 511:405–35.

Shen, L. (2013). GeneOverlap: Test and visualize gene overlaps R package.

Shirasaki, T., Honda, M., Mizuno, H., Shimakami, T., Okada, H., Sakai, Y., Murakami, S., Wakita, T., and Kaneko, S. (2010). La protein required for internal ribosome entry site-directed translation is a potential therapeutic target for hepatitis C virus replication. *J Infect Dis*, 202(1):75–85.

Shulla, A. and Randall, G. (2015). Spatiotemporal analysis of hepatitis C virus infection. *PLoS Pathog*, 11(3):e1004758.

Shwetha, S., Kumar, A., Mullick, R., Vasudevan, D., Mukherjee, N., and Das, S. (2015). HuR Displaces Polypyrimidine Tract Binding Protein To Facilitate La Binding to the 3' Untranslated Region and Enhances Hepatitis C Virus Replication. *J Virol*, 89(22):11356–11371.

Singer, S., Zhao, R., Barsotti, A. M., Ouwehand, A., Fazollahi, M., Coutavas, E., Breuhahn, K., Neumann, O., Longerich, T., Pusterla, T., Powers, M. A., Giles, K. M., Leedman, P. J., Hess, J., Grunwald, D., Bussemaker, H. J., Singer, R. H., Schirmacher, P., and Prives, C. (2012). Nuclear pore component Nup98 is a potential tumor suppressor and regulates posttranscriptional expression of select p53 target genes. *Mol Cell*, 48(5):799–810.

Singleton, M. R., Dillingham, M. S., and Wigley, D. B. (2007). Structure and Mechanism of Helicases and Nucleic Acid Translocases. *Annual Review of Biochemistry*, 76(1):23–50.

Singleton, M. R., Sawaya, M. R., Ellenberger, T., and Wigley, D. B. (2000). Crystal structure of T7 gene 4 ring helicase indicates a mechanism for sequential hydrolysis of nucleotides. *Cell*, 101(6):589–600.

Sir, D., Kuo, C.-f., Tian, Y., Liu, H. M., Huang, E. J., Jung, J. U., Machida, K., and Ou, J.-h. J. (2012). Replication of hepatitis C virus RNA on autophagosomal membranes. *The Journal of biological chemistry*, 287(22):18036–43.

Sloan, K. E., Gleizes, P.-E. E., and Bohnsack, M. T. (2015). Nucleocytoplasmic Transport of RNAs and RNA-Protein Complexes. *J Mol Biol*, 428(10 Pt A):2040–59.

Smith, W. A., Schurter, B. T., Wong-Staal, F., and David, M. (2004). Arginine methylation of RNA helicase a determines its subcellular localization. *J Biol Chem*, 279(22):22795–22798.

Smoot, M. E., Ono, K., Ruscheinski, J., Wang, P. L., and Ideker, T. (2011). Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics*, 27(3):431–432.

Söderqvist, H. and Hallberg, E. (1994). The large C-terminal region of the integral pore membrane protein, POM121, is facing the nuclear pore complex. *European journal of cell biology*, 64(1):186–91.

Solmaz, S. R., Blobel, G., and Melcák, I. (2013). Ring cycle for dilating and constricting the nuclear pore. *Proceedings of the National Academy of Sciences of the United States of America*, 110(15):5858–63.

Solmaz, S. R., Chauhan, R., Blobel, G., and Melčák, I. (2011). Molecular architecture of the transport channel of the nuclear pore complex. *Cell*, 147(3):590–602.

Soniat, M. and Chook, Y. (2015). Nuclear localization signals for four distinct karyopherin-β nuclear import systems. *Biochemical Journal*, 468(3):353–362.

Spångberg, K., Wiklund, L., and Schwartz, S. (2000). HuR, a protein implicated in oncogene and growth factor mRNA decay, binds to the 3' ends of hepatitis C virus RNA of both polarities. *Virology*, 274(2):378–390.

Splinter, D., Tanenbaum, M. E., Lindqvist, A., Jaarsma, D., Flotho, A., Yu, K. L., Grigoriev, I., Engelsma, D., Haasdijk, E. D., Keijzer, N., Demmers, J., Fornerod, M., Melchior, F., Hoogenraad, C. C., Medema, R. H., and Akhmanova, A. (2010). Bicaudal D2, dynein, and kinesin-1 associate with nuclear pore complexes and regulate centrosome and nuclear positioning during mitotic entry. *PLoS biology*, 8(4):e1000350.

Stadler, K., Allison, S. L., Schalich, J., and Heinz, F. X. (1997). Proteolytic activation of tick-borne encephalitis virus by furin. *Journal of virology*, 71(11):8475–81.

Stavru, F., Hülsmann, B. B., Spang, A., Hartmann, E., Cordes, V. C., and Görlich, D. (2006). NDC1: a crucial membrane-integral nucleoporin of metazoan nuclear pore complexes. *The Journal of Cell Biology*, 173(4):509–519.

Stephens, C. and Harlow, E. (1987). Differential splicing yields novel adenovirus 5 E1A mRNAs that encode 30 kd and 35 kd proteins. *The EMBO journal*, 6(7):2027–35.

Stevens, S. W. (2010). Chapter 4 The Biology of DEAH/RHA Proteins and Their Mechanism of Action. In *RNA Helicases*, pages 99–120. The Royal Society of Chemistry.

Stewart, C. L., Roux, K. J., and Burke, B. (2007). Blurring the boundary: the nuclear envelope extends its reach. *Science*, 318(5855):1408–1412.

Strunze, S., Engelke, M. F., Wang, I.-H., Puntener, D., Boucke, K., Schleich, S., Way, M., Schoenenberger, P., Burckhardt, C. J., and Greber, U. F. (2011). Kinesin-1-mediated capsid disassembly and disruption of the nuclear pore complex promote virus infection. *Cell host & microbe*, 10(3):210–23.

Strunze, S., Trotman, L. C., Boucke, K., and Greber, U. F. (2005). Nuclear targeting of adenovirus type 2 requires CRM1-mediated nuclear export. *Molecular biology of the cell*, 16(6):2999–3009.

Sun, Q., Carrasco, Y. P., Hu, Y., Guo, X., Mirzaei, H., Macmillan, J., and Chook, Y. M. (2013). Nuclear export inhibition through covalent conjugation and hydrolysis of Leptomycin B by CRM1. *Proc Natl Acad Sci U S A*, 110(4):1303–1308.

Sun, Y. and Guo, H. C. (2008). Structural constraints on autoprocessing of the human nucleoporin Nup98. *Protein Sci*, 17(3):494–505.

Suzuki, R., Sakamoto, S., Tsutsumi, T., Rikimaru, A., Tanaka, K., Shimoike, T., Moriishi, K., Iwasaki, T., Mizumoto, K., Matsuura, Y., Miyamura, T., and Suzuki, T. (2005). Molecular Determinants for Subcellular Localization of Hepatitis C Virus Core Protein. *Journal of Virology*, 79(2):1271–1281.

Szklarczyk, D., Franceschini, A., Kuhn, M., Simonovic, M., Roth, A., Minguez, P., Doerks, T., Stark, M., Muller, J., Bork, P., Jensen, L. J., and von Mering, C. (2011). The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored. *Nucleic Acids Res*, 39(Database issue):561–8.

Tabata, T., Petitt, M., Puerta-Guardo, H., Michlmayr, D., Wang, C., Fang-Hoover, J., Harris, E., and Pereira, L. (2016). Zika Virus Targets Different Primary Human Placental Cells, Suggesting Two Routes for Vertical Transmission. *Cell Host & Microbe*, 20(2):155–166.

Talamas, J. A. and Hetzer, M. W. (2011). POM121 and Sun1 play a role in early steps of interphase NPC assembly. *The Journal of cell biology*, 194(1):27–37.

Tan-Wong, S. M., Wijayatilake, H. D., and Proudfoot, N. J. (2009). Gene loops function to maintain transcriptional memory through interaction with the nuclear pore complex. *Genes Dev*, 23(22):2610–2624.

Tanaka, N., Aronova, A., and Schwer, B. (2007). Ntr1 activates the Prp43 helicase to trigger release of lariat-intron from the spliceosome. *Genes & Development*, 21(18):2312–2325.

Tang, H., Gaietta, G. M., Fischer, W. H., Ellisman, M. H., and Wong-Staal, F. (1997). A cellular cofactor for the constitutive transport element of type D retrovirus. *Science (New York, N.Y.)*, 276(5317):1412–5.

Tang, H., Hammack, C., Ogden, S., Wen, Z., Qian, X., Li, Y., Yao, B., Shin, J., Zhang, F., Lee, E., Christian, K., Didier, R., Jin, P., Song, H., and Ming, G.-l. (2016). Zika Virus Infects Human Cortical Neural Progenitors and Attenuates Their Growth.

Tang, H., McDonald, D., Middlesworth, T., Hope, T. J., and Wong-Staal, F. (1999). The carboxyl terminus of RNA helicase A contains a bidirectional nuclear transport domain. *Mol Cell Biol*, 19(5):3540–3550.

Tang, W., You, W., Shi, F., Qi, T., Wang, L., Djouder, Z., Liu, W., and Zeng, X. (2009). RNA helicase A acts as a bridging factor linking nuclear beta-actin with RNA polymerase II. *The Biochemical journal*, 420(3):421–8.

Tanner, N. K., Cordin, O., Banroques, J., Doère, M., and Linder, P. (2003). The Q motif: a newly identified motif in DEAD box helicases may regulate ATP binding and hydrolysis. *Molecular cell*, 11(1):127–38.

Tay, M., Fraser, J., Chan, W., Moreland, N., Rathore, A., Wang, C., Vasudevan, S., and Jans, D. (2013). Nuclear localization of dengue virus (DENV) 14 nonstructural protein 5; protection against all 4 DENV serotypes by the inhibitor Ivermectin. *Antiviral Research*, 99(3):301–306.

Team, R. C. (2016). R: A Language and Environment for Statistical Computing.

Terns, M. P. and Terns, R. M. (2001). Macromolecular complexes: SMN–the master assembler. *Current biology : CB*, 11(21):862–4.

Tetenbaum-Novatt, J., Hough, L. E., Mironska, R., McKenney, A. S., and Rout, M. P. (2012). Nucleocytoplasmic transport: a role for nonspecific competition in karyopherin-nucleoporin interactions. *Molecular & cellular proteomics : MCP*, 11(5):31–46.

Tetsuka, T., Uranishi, H., Sanda, T., Asamitsu, K., Yang, J.-P., Wong-Staal, F., and Okamoto, T. (2004). RNA helicase A interacts with nuclear factor κB p65 and functions as a transcriptional coactivator. *European Journal of Biochemistry*, 271(18):3741–3751.

Theerthagiri, G., Eisenhardt, N., Schwarz, H., and Antonin, W. (2010). The nucleoporin Nup188 controls passage of membrane proteins across the nuclear pore complex. *The Journal of cell biology*, 189(7):1129–42.

Thomssen, R., Bonk, S., Propfe, C., Heermann, K. H., Köchel, H. G., and Uy, A. (1992). Association of hepatitis C virus in human sera with beta-lipoprotein. *Medical microbiology and immunology*, 181(5):293–300.

Timney, B. L., Tetenbaum-Novatt, J., Agate, D. S., Williams, R., Zhang, W., Chait, B. T., and Rout, M. P. (2006). Simple kinetic relationships and nonspecific competition govern nuclear import rates in vivo. *The Journal of Cell Biology*, 175(4):579–593.

Trotman, L. C., Mosberger, N., Fornerod, M., Stidwill, R. P., and Greber, U. F. (2001). Import of adenovirus DNA involves the nuclear pore complex receptor CAN/Nup214 and histone H1. *Nature cell biology*, 3(12):1092–100.

Tsoulfas, G., Goulis, I., Giakoustidis, D., Akriviadis, E., Agorastou, P., Imvrios, G., and Papanikolaou, V. (2009). Hepatitis C and liver transplantation. *Hippokratia*, 13(4):211–5.

Uchil, P. D., Kumar, A. V. A., and Satchidanandam, V. (2006). Nuclear localization of flavivirus RNA synthesis in infected cells. *Journal of virology*, 80(11):5451–64.

Uhlén, M., Fagerberg, L., Hallström, B. M., Lindskog, C., Oksvold, P., Mardinoglu, A., Sivertsson, ., Kampf, C., Sjöstedt, E., Asplund, A., Olsson, I., Edlund, K., Lundberg, E., Navani, S., Szigyarto, C. A., Odeberg, J., Djureinovic, D., Takanen, J. O., Hober, S., Alm, T., Edqvist, P. H., Berling, H., Tegel, H., Mulder, J., Rockberg, J., Nilsson, P., Schwenk, J. M., Hamsten, M., von Feilitzen, K., Forsberg, M., Persson, L., Johansson, F., Zwahlen, M., von Heijne, G., Nielsen, J., and Pontén, F. (2015). Proteomics. Tissue-based map of the human proteome. *Science*, 347(6220):1260419.

Uhlen, M., Oksvold, P., Fagerberg, L., Lundberg, E., Jonasson, K., Forsberg, M., Zwahlen, M., Kampf, C., Wester, K., Hober, S., Wernerus, H., Björling, L., and Ponten, F. (2010). Towards a knowledge-based Human Protein Atlas. *Nature Biotechnology*, 28(12):1248–1250.

Ule, J., Jensen, K., Mele, A., and Darnell, R. B. (2005). CLIP: a method for identifying protein-RNA interaction sites in living cells. *Methods (San Diego, Calif.)*, 37(4):376–86.

Untergasser, A., Nijveen, H., Rao, X., Bisseling, T., Geurts, R., and Leunissen, J. A. (2007). Primer3Plus, an enhanced web interface to Primer3. *Nucleic Acids Res*, 35(Web Server issue):71–4.

Upadhyay, A., Dixit, U., Manvar, D., Chaturvedi, N., and Pandey, V. N. (2013). Affinity capture and identification of host cell factors associated with hepatitis C virus (+) strand subgenomic RNA. *Molecular & cellular proteomics : MCP*, 12(6):1539–52.

Vagnarelli, P., Ribeiro, S., Sennels, L., Sanchez-Pulido, L., deLimaAlves, F., Verheyen, T., Kelly, D., Ponting, C., Rappsilber, J., and Earnshaw, W. (2011). RepoMan Coordinates Chromosomal Reorganization with Nuclear Envelope Reassembly during Mitotic Exit. *Developmental Cell*, 21(2):328–342.

Vahedi, G., Kanno, Y., Furumoto, Y., Jiang, K., Parker, S. C. J., Erdos, M. R., Davis, S. R., Roychoudhuri, R., Restifo, N. P., Gadina, M., Tang, Z., Ruan, Y., Collins, F. S., Sartorelli, V., and O'Shea, J. J. (2015). Super-enhancers delineate disease-associated regulatory nodes in T cells. *Nature*, 520(7548):558–62.

Valente, S. T. and Goff, S. P. (2006). Inhibition of HIV-1 Gene Expression by a Fragment of hnRNP U.

van der Lee, R., Feng, Q., Langereis, M. A., Ter Horst, R., Szklarczyk, R., Netea, M. G., Andeweg, A. C., van Kuppeveld, F. J., and Huynen, M. A. (2015). Integrative Genomics-Based Discovery of Novel Regulators of the Innate Antiviral Response. *PLoS Comput Biol*, 11(10):e1004553.

van Steensel, B., Delrow, J., and Henikoff, S. (2001). Chromatin profiling using targeted DNA adenine methyltransferase. *Nature Genetics*, 27(3):304–308.

Vaquerizas, J. M., Suyama, R., Kind, J., Miura, K., Luscombe, N. M., and Akhtar, A. (2010). Nuclear Pore Proteins Nup153 and Megator Define Transcriptionally Active Regions in the Drosophila Genome. *PLoS Genetics*, 6(2):e1000846.

Vasu, S., Shah, S., Orjalo, A., Park, M., Fischer, W. H., and Forbes, D. J. (2001). Novel vertebrate nucleoporins Nup133 and Nup160 play a role in mRNA export. *The Journal of Cell Biology*, 155(3):339–354.

Vo, N. and Goodman, R. H. (2001). CREB-binding Protein and p300 in Transcriptional Regulation. *Journal of Biological Chemistry*, 276(17):13505–13508.

Vogel, M. J., Peric-Hupkes, D., and van Steensel, B. (2007). Detection of in vivo protein-DNA interactions using DamID in mammalian cells. *Nature protocols*, 2(6):1467–78.

von Appen, A. and Beck, M. (2016). Structure Determination of the Nuclear Pore Complex with Three-Dimensional Cryo electron Microscopy. *Journal of molecular biology*, 428(10 Pt A):2001–10.

von Kobbe C, van Deursen JM, Rodrigues, J. P., Sitterlin, D., Bachi, A., Wu, X., Wilm, M., Carmo-Fonseca, M., and Izaurralde, E. (2000). Vesicular stomatitis virus matrix protein inhibits host cell gene expression by targeting the nucleoporin Nup98. *Molecular cell*, 6(5):1243–52.

von Mering, C., Huynen, M., Jaeggi, D., Schmidt, S., Bork, P., and Snel, B. (2003). STRING: a database of predicted functional associations between proteins. *Nucleic Acids Res*, 31(1):258–261.

von Moeller, H., Basquin, C., and Conti, E. (2009). The mRNA export protein DBP5 binds RNA and the cytoplasmic nucleoporin NUP214 in a mutually exclusive manner. *Nature structural & molecular biology*, 16(3):247–54.

Wakita, T., Pietschmann, T., Kato, T., Date, T., Miyamoto, M., Zhao, Z., Murthy, K., Habermann, A., Kräusslich, H.-G., Mizokami, M., Bartenschlager, R., and Liang, T. J. (2005). Production of infectious hepatitis C virus in tissue culture from a cloned viral genome. *Nature medicine*, 11(7):791–6.

Walker, J. E., Saraste, M., Runswick, M. J., and Gay, N. J. (1982). Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *The EMBO journal*, 1(8):945–51.

Walther, T. C., Askjaer, P., Gentzel, M., Habermann, A., Griffiths, G., Wilm, M., Mattaj, I. W., and Hetzer, M. (2003). RanGTP mediates nuclear pore complex assembly. *Nature*, 424(6949):689–94.

Wang, J.-J., Liao, C.-L., Chiou, Y.-W., Chiou, C.-T., Huang, Y.-L., and Chen, L.-K. (1997). Ultrastructure and Localization of E Proteins in Cultured Neuron Cells Infected with Japanese Encephalitis Virus. *Virology*, 238(1):30–39.

Wang, M., Weiss, M., Simonovic, M., Haertinger, G., Schrimpf, S. P., Hengartner, M. O., and von Mering, C. (2012). PaxDb, a database of protein abundance averages across all three domains of life. *Molecular & cellular proteomics : MCP*, 11(8):492–500.

Wang, P., Zhu, S., Yang, L., Cui, S., Pan, W., Jackson, R., Zheng, Y., Rongvaux, A., Sun, Q., Yang, G., Gao, S., Lin, R., You, F., Flavell, R., and Fikrig, E. (2015). Nlrp6 regulates intestinal antiviral innate immunity. *Science*.

Wang, X., Babu, J. R., Harden, J. M., Jablonski, S. A., Gazi, M. H., Lingle, W. L., de Groen, P. C., Yen, T. J., and van Deursen, J. M. (2001). The mitotic checkpoint protein hBUB3 and the mRNA export factor hRAE1 interact with GLE2p-binding sequence (GLEBS)-containing proteins. *The Journal of biological chemistry*, 276(28):26559–67.

Watanabe, K., Takizawa, N., Katoh, M., Hoshida, K., Kobayashi, N., and Nagata, K. (2001). Inhibition of nuclear export of ribonucleoprotein complexes of influenza virus by leptomycin B. *Virus research*, 77(1):31–42.

Watters, K. and Palmenberg, A. C. (2011). Differential processing of nuclear pore complex proteins by rhinovirus 2A proteases from different species and serotypes. *Journal of virology*, 85(20):10874–83.

Webster, B. M., Colombi, P., Jäger, J., and Lusk, C. P. (2014). Surveillance of nuclear pore complex assembly by ESCRT-III/Vps4. *Cell*, 159(2):388–401.

Weidensdorfer, D., Stöhr, N., Baude, A., Lederer, M., Köhn, M., Schierhorn, A., Buchmeier, S., Wahle, E., and Hüttelmaier, S. (2009). Control of c-myc mRNA stability by IGF2BP1-associated cytoplasmic RNPs. *RNA*, 15(1):104–115.

Weirich, C. S., Erzberger, J. P., Berger, J. M., and Weis, K. (2004). The N-terminal domain of Nup159 forms a beta-propeller that functions in mRNA export by tethering the helicase Dbp5 to the nuclear pore. *Molecular cell*, 16(5):749–60.

Wells, M., Salick, M., Wiskow, O., Ho, D., Worringer, K., Ihry, R., Kommineni, S., Bilican, B., Klim, J., Hill, E., Kane, L., Ye, C., Kaykas, A., and Eggan, K. (2016). Genetic Ablation of AXL Does Not Protect Human Neural Progenitor Cells and Cerebral Organoids from Zika Virus Infection.

Welsch, S., Miller, S., Romero-Brey, I., Merz, A., Bleck, C. K., Walther, P., Fuller, S. D., Antony, C., Krijnse-Locker, J., and Bartenschlager, R. (2009). Composition and Three-Dimensional Architecture of the Dengue Virus Replication and Assembly Sites. *Cell Host & Microbe*, 5(4):365–375.

Wen, W., Meinkoth, J. L., Tsien, R. Y., and Taylor, S. S. (1995). Identification of a signal for rapid export of proteins from the nucleus. *Cell*, 82(3):463–73.

Wente, S. R. and Rout, M. P. (2010). The nuclear pore complex and nuclear transport. *Cold Spring Harb Perspect Biol*, 2(10):a000562.

Wilhelm, M., Schlegl, J., Hahne, H., Gholami, A. M., Lieberenz, M., Savitski, M. M., Ziegler, E., Butzmann, L., Gessulat, S., Marx, H., Mathieson, T., Lemeer, S., Schnatbaum, K., Reimer, U., Wenschuh, H., Mollenhauer, M., Slotta-Huspenina, J., Boese, J.-H., Bantscheff, M., Gerstmair, A., Faerber, F., and Kuster, B. (2014). Mass-spectrometry-based draft of the human proteome. *Nature*, 509(7502):582–7.

Wolff, B., Sanglier, J. J., and Wang, Y. (1997). Leptomycin B is an inhibitor of nuclear export: inhibition of nucleo-cytoplasmic translocation of the human immunodeficiency virus type 1 (HIV-1) Rev protein and Rev-dependent mRNA. *Chem Biol*, 4(2):139–147.

Wolfien, M., Rimmbach, C., Schmitz, U., Jung, J. J., Krebs, S., Steinhoff, G., David, R., and Wolkenhauer, O. (2016). TRAPLINE: a standardized and automated pipeline for RNA sequencing data analysis, evaluation and annotation. *BMC Bioinformatics*, 17(1):21.

Wong, R. W., Blobel, G., and Coutavas, E. (2006). Rae1 interaction with NuMA is required for bipolar spindle formation. *Proceedings of the National Academy of Sciences of the United States of America*, 103(52):19783–7.

Wozniak, R., Burke, B., and Doye, V. (2010). Nuclear transport and the mitotic apparatus: an evolving relationship. *Cellular and molecular life sciences : CMLS*, 67(13):2215–2230.

Wozniak, R. W., Bartnik, E., and Blobel, G. (1989). Primary structure analysis of an integral membrane glycoprotein of the nuclear pore. *The Journal of cell biology*, 108(6):2083–92.

Wozniak, R. W. and Blobel, G. (1992). The single transmembrane segment of gp210 is sufficient for sorting to the pore membrane domain of the nuclear envelope. *The Journal of cell biology*, 119(6):1441–9.

Wozniak, R. W. and Goldfarb, D. S. (2008). Cyclin-like oscillations in levels of the nucleoporin Nup96 control G1/S progression. *Dev Cell*, 15(5):643–644.

Wu, A. C. and Rifkin, S. A. (2015). Aro: a machine learning approach to identifying single molecules and estimating classification error in fluorescence microscopy images. *BMC Bioinformatics*, 16:102.

Xiao, R., Tang, P., Yang, B., Huang, J., Zhou, Y., Shao, C., Li, H., Sun, H., Zhang, Y., Fu, X.-D., Ameyar-Zazoua, M., Souidi, M., Fritsch, L., Robin, P., Thomas, A., Hamiche, A., Percipalle, P., Ait-Si-Ali, S., Harel-Bellan, A., Andrade, L., Chan, E., Raska, I., Peebles, C., Roos, G., Tan, E., Brosi, R., Gröning, K., Behrens, S., Lührmann, R., Krämer, A., Brosi, R., Hauri, H., Krämer, A., Cartegni, L., Krainer, A., Cartegni, L., Hastings, M., Calarco, J., Stanchina, E. d., Krainer, A., Cioce, M., Lamond, A., Clemson, C., Hutchinson, J., Sara, S., Ensminger, A., Fox, A., Chess, A., Lawrence, J., Mata, M. d. l., Alonso, C., Kadener, S., Fededa, J., Blaustein, M., Pelisch, F., Cramer, P., Bentley, D., Kornblihtt, A., Eggert, M., Michel, J., Schneider, S., Bornfleth, H., Baniahmad, A., Fackelmayer, F., Schmidt, S., Renkawitz, R., Fackelmayer, F., Richter, A., Fackelmayer, F., Dahm, K., Renz, A., Ramsperger, U., Richter, A., Fu, D., Collins, K., Gavrilov, D., Shi, X., Das, K., Gilliam, T., Wang, C., Gupta, A., Drazba, J., Banerjee, A., Han, J., Ding, J., Byeon, C., Kim, J., Hertel, K., Jeong, S., Fu, X., Hasegawa, Y., Brockdorff, N., Kawano, S., Tsutui, K., Tsutui, K., Nakagawa, S., Hastings, M., Allemand, E., Duelli, D., Myers, M., Krainer, A., Helbig, R., Fackelmayer, F., Hofmann, Y., Wirth, B., Hsieh-Li, H., Chang, J., Jong, Y., Wu, M., Wang, N., Tsai, C., Li, H., Jády, B., Bertrand, E., Kiss, T., Jodelka, F., Ebert, A., Duelli, D., Hastings, M., Kashima, T., Manley, J., Kashima, T., Rao, N., Manley, J., Kiledjian, M., Dreyfuss, G., Kim, M., Nikodem, V., Kittler, R., Surendranath, V., Heninger, A., Slabicki, M., Theis, M., Putz, G., Franke, K., Caldarelli, A., Grabner, H., Kozak, K., Al., E., Kornblihtt, A., Krämer, A., Grüter, P., Gröning, K., Kastner, B., Kukalev, A., Nord, Y., Palmberg, C., Bergman, T., Percipalle, P., Kuo, H., Nasim, F., Grabowski, P., Licatalosi, D., Mele, A., Fak, J., Ule, J., Kayikci, M., Chi, S., Clark, T., Schweitzer, A., Blume, J., Wang, X., Al., E., Lin, S., Coutinho-Mansfield, G., Wang, D., Pandit, S., Fu, X., Liu, Q., Dreyfuss, G., Lorson, C., Hahnen, E., Androphy, E., Wirth, B., Martens, J., Verlaan, M., Kalkhoven, E., Dorsman, J., Zantema, A., Araújo, M. M. d., Bonnal, S., Hastings, M., Krainer, A., Valcárcel, J., Miyajima, H., Miyaso, H., Okumura, M., Kurisu, J., Imaizumi, K., Nesic, D., Tanackovic, G., Krämer, A., Obrdlik, A., Kukalev, A., Louvet, E., Farrants, A., Caputo, L., Percipalle, P., Pedrotti, S., Bielli, P., Paronetto, M., Ciccosanti, F., Fimia, G., Stamm, S., Manley, J., Sette, C., Pellizzoni, L., Yong, J., Dreyfuss, G., Pullirsch, D., Härtel, R., Kishimoto, H., Leeb, M., Steiner, G., Wutz, A., Rappsilber, J., Ryder, U., Lamond, A., Mann, M., Reed, R., Maniatis, T., Romig, H., Fackelmayer, F., Renz, A., Ramsperger, U., Richter, A., Roshon, M., Ruley, H., Saltzman, A., Pan, Q., Blencowe, B., Singh, N., Singh, N., Androphy, E., Singh, R., Sleeman, J., Ajuh, P., Lamond, A., Tanackovic, G., Krämer, A., Tripathi, V., Ellis, J., Shen, Z., Song, D., Pan, Q., Watt, A., Freier, S., Bennett, C., Sharma, A., Bubulya, P., Al., E., Tsai, M., Manor, O., Wan, Y., Mosammaparast, N., Wang, J., Lan, F., Shi, Y., Segal, E., Chang, H., Valente, S., Goff, S., Wang, Z., Rolish, M., Yeo, G., Tung, V., Mawson, M., Burge, C., Wang, Z., Kayikci, M., Briese, M., Zarnack, K., Luscombe, N., Rot, G., Zupan, B., Curk, T., Ule, J., Will, C., Lührmann, R., Will, C., Urlaub, H., Achsel, T., Gentzel, M., Wilm, M., Lührmann, R., Xue, Y., Zhou, Y., Wu, T., Zhu, T., Ji, X., Kwon, Y., Zhang, C., Yeo, G., Black, D., Sun, H., Al., E., Yeo, G., Burge, C., Yeo, G., Coufal, N., Liang, T., Peng, G., Fu, X., Gage, F., Yik, J., Chen, R., Nishimura, R., Jennings, J., Link, A., Zhou, Q., Yugami, M., Kabe, Y., Yamaguchi, Y., Wada, T., Handa, H., Zhang, Z., Lotti, F., Dittmar, K., Younis, I., Wan, L., Kasim, M., Dreyfuss, G., Zhou, Z., Licklider, L., Gygi, S., and Reed, R. (2012). Nuclear matrix factor hnRNP U/SAF-A exerts a global control of alternative splicing by regulating U2 snRNP maturation. *Molecular cell*, 45(5):656–68.

Xing, L., Liang, C., and Kleiman, L. (2011). Coordinate roles of Gag and RNA heli-

case A in promoting the annealing of formula to HIV-1 RNA. *J Virol*, 85(4):1847–1860.

Xing, L., Niu, M., and Kleiman, L. (2014). Role of the OB-fold of RNA helicase A in the synthesis of HIV-1 RNA. *Biochim Biophys Acta*, 1839(11):1069–1078.

Xu, D., Marquis, K., Pei, J., Fu, S.-C., Cağatay, T., Grishin, N. V., and Chook, Y. M. (2015). LocNES: a computational tool for locating classical NESs in CRM1 cargo proteins. *Bioinformatics (Oxford, England)*, 31(9):1357–65.

Xylourgidis, N., Roth, P., Sabri, N., Tsarouhas, V., and Samakovlis, C. (2006). The nucleoporin Nup214 sequesters CRM1 at the nuclear rim and modulates NFB activation in Drosophila. *Journal of Cell Science*, 119(21):4409–4419.

Yamada, J., Phillips, J. L., Patel, S., Goldfien, G., Calestagne-Morelli, A., Huang, H., Reza, R., Acheson, J., Krishnan, V. V., Newsam, S., Gopinathan, A., Lau, E. Y., Colvin, M. E., Uversky, V. N., and Rexach, M. F. (2010). A bimodal distribution of two distinct categories of intrinsically disordered structures with separate functions in FG nucleoporins. *Molecular & cellular proteomics : MCP*, 9(10):2205–24.

Yang, D.-Q., Halaby, M.-J., and Zhang, Y. (2006). The identification of an internal ribosomal entry site in the 5'-untranslated region of p53 mRNA provides a novel mechanism for the regulation of its translation following DNA damage. *Oncogene*, 25(33):4613–9.

Yao, B., Christian, K. M., He, C., Jin, P., Ming, G.-l., and Song, H. (2016). Epigenetic mechanisms in neurogenesis. *Nature Reviews Neuroscience*, 17(9):537–549.

Yarbrough, M. L., Mata, M. A., Sakthivel, R., and Fontoura, B. M. A. (2014). Viral subversion of nucleocytoplasmic trafficking. *Traffic*, 15(2):127–140.

Yates, B., Braschi, B., Gray, K. A., Seal, R. L., Tweedie, S., and Bruford, E. A. (2016). Genenames.org: the HGNC and VGNC resources in 2017. *Nucleic Acids Research*, page gkw1033.

Yi, M. (2010). Hepatitis C virus: propagation, quantification, and storage. *Current protocols in microbiology*, Chapter 15:Unit 15D.1.

Yokoyama, H., Koch, B., Walczak, R., Ciray-Duygu, F., González-Sánchez, J. C., Devos, D. P., Mattaj, I. W., and Gruss, O. J. (2014). The nucleoporin MEL-28 promotes RanGTP-dependent γ-tubulin recruitment and microtubule nucleation in mitotic spindle formation. *Nature communications*, 5:3270.

York, A. and Fodor, E. (2013). Biogenesis, assembly, and export of viral messenger ribonucleoproteins in the influenza A virus infected cell. *RNA Biol*, 10(8):1274–1282.

Yu, G., Wang, L.-G., and He, Q.-Y. (2015). ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics (Oxford, England)*, 31(14):2382–3.

Yue, Y., Liu, J., and He, C. (2015). RNA N6-methyladenosine methylation in post-transcriptional gene expression regulation. *Genes & development*, 29(13):1343–55.

Yugami, M., Kabe, Y., Yamaguchi, Y., Wada, T., and Handa, H. (2007). hnRNP-U enhances the expression of specific genes by stabilizing mRNA. *FEBS Lett*, 581(1):1–7.

Zahn, R., Osmanović, D., Ehret, S., Araya Callis, C., Frey, S., Stewart, M., You, C., Görlich, D., Hoogenboom, B. W., and Richter, R. P. (2016). A physical model describing the interaction of nuclear transport receptors with FG nucleoporin domain assemblies. *Elife*, 5.

Zammarchi, L., Tappe, D., Fortuna, C., Remoli, M. E., Günther, S., Venturi, G., Bartoloni, A., and Schmidt-Chanasit, J. (2015). Zika virus infection in a traveller returning to Europe from Brazil, March 2015. *Euro surveillance : bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin*, 20(23).

Zanluca, C., Melo, V. C. A. d., Mosimann, A. L. P., Santos, G. I. V. D., Santos, C. N. D. D., and Luz, K. (2015). First report of autochthonous transmission of Zika virus in Brazil. *Memorias do Instituto Oswaldo Cruz*, 110(4):569–72.

Zeitler, B. and Weis, K. (2004). The FG-repeat asymmetry of the nuclear pore complex is dispensable for bulk nucleocytoplasmic transport in vivo. *The Journal of cell biology*, 167(4):583–90.

Zemp, I. and Kutay, U. (2007). Nuclear export and cytoplasmic maturation of ribosomal subunits. *FEBS Letters*, 581(15):2783–2793.

Zhang, F., Hammack, C., Ogden, S. C., Cheng, Y., Lee, E. M., Wen, Z., Qian, X., Nguyen, H. N., Li, Y., Yao, B., Xu, M., Xu, T., Chen, L., Wang, Z., Feng, H., Huang, W.-K., Yoon, K.-j., Shan, C., Huang, L., Qin, Z., Christian, K. M., Shi, P.-Y., Xu, M., Xia, M., Zheng, W., Wu, H., Song, H., Tang, H., Ming, G.-L., Jin, P., M., M., B., E., D.L., H., S.A., R., D., M., B.D., F., P., B., R.W., D., A., B., T.J., C., G., K., D., S., V.A., K., A., E., E., P., R., H., G., C., J., M., H., T., P.P., G., X., Q., J., D., Z., W., T.A., J., C.H., C., K.J., Y., C., T., D., K., C., T., R.C., G., H., W., W., H., J., W., J., M., H.C., C., N., O., A., S., B.A., W.-H., L.T., V., J., L., A.M., P., A.D., H., P.G., K., N., P., E., S., K., M., R., D., W., O., J.S., Y., M.R., Y., and R., B. (2016). Molecular signatures associated with ZIKV exposure in human cortical neural progenitors. *Nucleic Acids Research*, 44(18):8610–8620.

Zhang, N., Kuznetsov, S. G., Sharan, S. K., Li, K., Rao, P. H., and Pati, D. (2008). A handcuff model for the cohesin complex. *The Journal of Cell Biology*, 183(6):1019–1031.

Zhang, S. and Grosse, F. (1994). Nuclear DNA helicase II unwinds both DNA and RNA. *Biochemistry*, 33(13):3906–3912.

Zhang, S. and Grosse, F. (1997). Domain structure of human nuclear DNA helicase II (RNA helicase A). *J Biol Chem*, 272(17):11487–11494.

Zhang, S., Herrmann, C., and Grosse, F. (1999a). Nucleolar localization of murine nuclear DNA helicase II (RNA helicase A). *Journal of cell science*, pages 2693–703.

Zhang, S., Herrmann, C., and Grosse, F. (1999b). Pre-mRNA and mRNA binding of human nuclear DNA helicase II (RNA helicase A). *Journal of cell science*, pages 1055–64.

Zhang, S., Maacke, H., and Grosse, F. (1995). Molecular cloning of the gene encoding nuclear DNA helicase II. A bovine homologue of human RNA helicase A and Drosophila Mle protein. *The Journal of biological chemistry*, 270(27):16422–7.

Zhang, S. S. and Grosse, F. (1991). Purification and characterization of two DNA helicases from calf thymus nuclei. *J Biol Chem*, 266(30):20483–20490.

Zhang, X., Odom, D. T., Koo, S.-H., Conkright, M. D., Canettieri, G., Best, J., Chen, H., Jenner, R., Herbolsheimer, E., Jacobsen, E., Kadam, S., Ecker, J. R., Emerson, B., Hogenesch, J. B., Unterman, T., Young, R. A., and Montminy, M. (2005). Genome-wide analysis of cAMP-response element binding protein occupancy, phosphorylation, and target gene activation in human tissues. *Proceedings of the National Academy of Sciences*, 102(12):4459–4464.

Zhang, Z., Yuan, B., Lu, N., Facchinetti, V., and Liu, Y.-J. (2011). DHX9 pairs with IPS-1 to sense double-stranded RNA in myeloid dendritic cells. *Journal of immunology (Baltimore, Md. : 1950)*, 187(9):4501–8.

Zhao, Y. (2016). CellVision3D.

Zhao, Y., Schreiner, S. M., Koo, P. K., Colombi, P., King, M. C., and Mochrie, S. G. J. (2016). Improved Determination of Subnuclear Position Enabled by Three-Dimensional Membrane Reconstruction. *Biophysical journal*, 111(1):19–24.

Zheng, Z. M. (2004). Regulation of alternative RNA splicing by exon definition and exon sequences in viral and mammalian gene expression. *J Biomed Sci*, 11(3):278–294.

Zheng, Z.-M. M. (2010). Viral oncogenes, noncoding RNAs, and RNA splicing in human tumor viruses. *Int J Biol Sci*, 6(7):730–755.

Zhou, H., Xu, M., Huang, Q., Gates, A. T., Zhang, X. D., Castle, J. C., Stec, E., Ferrer, M., Strulovici, B., Hazuda, D. J., and Espeseth, A. S. (2008). Genome-scale RNAi screen for host factors required for HIV replication. *Cell host & microbe*, 4(5):495–504.

Zhou, K., Choe, K.-T., Zaidi, Z., Wang, Q., Mathews, M. B., and Lee, C.-G. (2003). RNA helicase A interacts with dsDNA and topoisomerase IIalpha. *Nucleic acids research*, 31(9):2253–60.

Zhou, Y., Hong, W., and Lu, L. (2013). Imaging beads-retained prey assay for rapid and quantitative protein-protein interaction. *PLoS One*, 8(3):e59727.

Zhu, Z., Chan, J. F.-W., Tee, K.-M., Choi, G. K.-Y., Lau, S. K.-P., Woo, P. C.-Y., Tse, H., and Yuen, K.-Y. (2016). Comparative genomic analysis of pre-epidemic and epidemic Zika virus strains for virological factors potentially associated with the rapidly expanding epidemic. *Emerging microbes & infections*, 5(3):e22.

Zilman, A., Di Talia, S., Chait, B. T., Rout, M. P., and Magnasco, M. O. (2007). Efficiency, selectivity, and robustness of nucleocytoplasmic transport. *PLoS computational biology*, 3(7):e125.

Zilman, A., Di Talia, S., Jovanovic-Talisman, T., Chait, B. T., Rout, M. P., and Magnasco, M. O. (2010). Enhancement of Transport Selectivity through Nano-Channels by Non-Specific Competition. *PLoS Computational Biology*, 6(6):e1000804.

Zuccolo, M., Alves, A., Galy, V., Bolhy, S., Formstecher, E., Racine, V., Sibarita, J.-B., Fukagawa, T., Shiekhattar, R., Yen, T., and Doye, V. (2007). The human Nup107160 nuclear pore subcomplex contributes to proper kinetochore functions. *The EMBO Journal*, 26(7):1853–1864.

APPENDIX A

Appendix: Hematopoietic cancers and Nup98 fusions: determining common mechanisms of malignancy

This appendix is a reproduction of the following publication:

Capitanio, J. S. and Wozniak, R. W. (2012). Hematopoietic cancers and Nup98 fusions: determining common mechanisms of malignancy. *arXiv: Molecular Networks (q-bio.MN)*, pages 1–13.

Hematopoietic cancers and Nup98 fusions: determining common mechanisms of malignancy

Juliana S. Capitanio and Richard W. Wozniak
Department of Cell Biology, Faculty of Medicine and Dentistry, University of Alberta

Abstract: Chromosomal aberrations are very frequent in leukemias and several recurring mutations capable of malignant transformation have been described. These mutations usually occur in hematopoietic stem cells (HSC), transforming them into leukemia stem cells. NUP98 gene translocations are an example of such chromosomal aberrations; these translocations produce a fusion protein containing the N-terminal portion of Nup98 and the C-terminal of a fusion partner. Over 75% of Nup98 fusions can interact with chromatin, and lead to changes in gene expression. Therefore, I hypothesize that nup98 fusions act as rogue transcriptional regulators in the cell.
Collecting previously published gene expression data (microarray) from HSCs expressing Nup98 fusions, we can generate data to corroborate this hypothesis. Several different fusions affect the expression of similar genes; these are involved in a few biological processes in the cell: embryonic development, immune system formation and chromatin organization. Deregulated genes also present similar transcription factor binding sites in their regulatory regions. These putative regulatory transcription factors are highly interconnected through protein-protein interactions and transcriptional regulation among themselves, and they have important roles in cell cycle regulation, embryonic development, hematopoiesis, apoptosis and chromatin modification.

Background:

Leukemia is a type of hematologic malignancy that originates in the bone marrow and leads to accumulation of immature hematopoietic cells with clonal origin. These leukemic cells can out-compete normal blood cells, replacing them in the bone marrow and spreading to extramedullary sites, and thus interfering with the normal function of the hematological tissue. The Canadian Cancer Society estimates 5000 new cases of leukemia for 2011, with approximately 2500 deaths. (1)
Leukemia is a heterogeneous disease that can be subdivided according to the cell lineage affected (myeloid or lymphoid) and the degree of differentiation of leukemic cells. More recent classifications by the World Health Organization also incorporate pathologic and genetic markers, achieving more biological significance.(2)

Chromosomal aberrations and leukemogenesis: delineating causes and effects.

Our knowledge of leukemogenesis is greatly influenced by the discovery of recurring chromosomal aberrations and/or gene mutations capable of malignant transformation of cells. (3) The target cell for these mutations is not always known, but increasing evidence indicates that leukemias originate in hematopoietic stem cells (HSCs) that are transformed into leukemia stem cells (LSC) by these chromosomal aberrations. LSCs and HSCs share two important characteristics, self-renewal and differentiation of new hematopoietic tissue. The clonogenic nature of leukemias is similar to that observed in normal hematopoiesis, and only a small specific subset of leukemic cells, LSCs, are capable of indefinite proliferation, as seen only for HSCs.(4)
If we consider HSCs as the cells of origin for most leukemias, we can better comprehend the high incidence of chromosomal aberrations present in this type of cancer. In a normal organism HSCs are usually quiescent and cycle very slowly; this is a protective mechanism to minimize DNA replication errors and the generation of toxic metabolic subproducts. However, quiescence of HSCs also presents a disadvantage. Once DNA damage does occur, specifically DNA double strand breaks (DSB), these cells must repair the lesions through the non-homologous end joining (NHEJ) pathway. NHEJ is an error-prone, mutagenic pathway that often causes chromosomal aberrations.(5) Given the self-renewal property of HSCs, mutations can be transmitted and accumulated, giving rise to LSCs with a high frequency of chromosomal aberrations; in fact, most recurring translocations in leukemia display a NHEJ repair signature.(6)
Most chromosomal translocations in leukemias occur in chromosomal regions where the DNA is more susceptible to double strand breaks. These susceptibility regions can be characterized by several chromatin structural elements, including topo II DNA cleavage sites, DNase I hypersensitive sites, scaffold/matrix attachment regions (S/MAR) and retrotransposon regions (LINE and SINE).(6) Chromosomal position within the nucleus also influences the frequency of translocations. It is likely that loci localized in close proximity inside the nucleus will more frequently translocate in cases of DSBs.(7) Finally, for these translocations to lead to transformation, they must generate fusion proteins that promote an advantage to the cell.
Thousands of chromosomal translocations have been detected in leukemia cells.(8) Most of these fall within two categories: type I mutations promote increased proliferation or survival, and type II mutations impair differentiation or enhance self-renewal. At least one mutation of each type seems to be necessary for leukemogenesis.(9) Leukemic chromosomal translocations generate chimeric fusion proteins, and many of these fusion proteins have similar characteristics: they localize to the nucleus, affect transcriptional regulation, contain a DNA binding domain and cause epigenetic modifications. (10)
In fact, epigenetic changes are a common occurrence in most acute leukemias.(3) If we consider that during normal hematopoiesis a complex program of epigenetic modifications takes place, it becomes clearer how altering epigenetic modifications can affect cell differentiation and self-renewal leading to leukemogenesis.(11) HOX genes are a prime example of epigenetic regulation. The pattern of HOX expres-

sion in cells is epigenetically regulated and inherited; each cell in the hematopoietic differentiation continuum displays a specific pattern of HOX genes expressed. More primitive hematopoietic cells have higher levels of HOX expression, as cells differentiate and lose their proliferative capabilities, HOX expression decreases until it becomes absent in completely differentiated cells. Overexpression of numerous HOX genes can induce leukemogenesis, in several cases changes in their expression levels are a result of histone modifications in the 5' HOXA gene.(12)

Nucleoporin genes and cancer.

To date, the literature has revealed five nucleoporins involved in carcinogenesis, Tpr, Nup88, Nup98, Nup214(13) and Nup358.(14) Nup88 expression is up-regulated in several cancers, specially carcinomas. Increase in its protein levels is thought to deregulate NF-kB nuclear transport maintaining it constantly activated.(15) The remaining 4 nucleoporins are involved in carcinogenic gene fusions.

Tpr gene fusions with Met and NTrk1 have been described in gastric cancers and papillary thyroid carcinomas, respectively. In both cases the N-terminal coiled-coil domain of Tpr is juxtaposed with the tyrosine kinase domain of the fusion partner. This leads to dimerization independent of ligand, and to constitutive activation of kinase activity, causing deregulated signaling that leads to carcinogenesis.(13) A similar carcinogenic mechanism is seen in Nup358 fusions. In inflammatory myofibroblastic tumors, the N-terminal leucine zipper of Nup358 is fused to the tyrosine kinase domain of ALK, also leading to deregulated kinase activation.(16)

Gene fusion of NUP214 and NUP98 play a role in leukemogenesis. NUP214 gene fusions with ABL also promote constitutive kinase activation in T-ALL. The kinase domain of ABL is fused to the N-terminal coiled-coil motif of Nup214. Interaction between this motif and nup88 allows localization of the fusion to the nuclear pore complex (NPC), bringing the kinase domains to sufficient proximity for constitutive activation.(17) In very rare cases Nup214 also translocates with SET and DEK in T-ALL and AML, respectively. In these cases almost the full length SET or DEK protein is fused to the C-terminal FG repeat of Nup214.(13) SET-NUP214 can interact with HOXA gene promoters leading to their expression. HOXA expression only occurs in the earliest T-cell precursor, so the fusion blocks T cell differentiation.(18) DEK-NUP214 fusions seem to increase overall protein translation specifically in myeloid cells, possibly facilitating carcinogenesis.(19) SET and DEK are two histone interacting proteins that perform opposing roles in the regulation of access to chromatin. SET promotes and DEK restricts accessibility to chromatin by the transcriptional machinery.(20) It is possible that NUP214 translocation with these genes affects the balance between the two chromatin modifiers.

The NUP98 gene is fused to a wide range of partner genes, resulting in several hematopoietic disorders, especially acute myeloid leukemia (AML).

NUP98 translocations in leukemias.

All NUP98 translocations described thus far generate a chimeric fusion protein that retains the N-terminal portion of Nup98 and the C-terminal of the fusion partner. Most chromosomal breaks take place between exons 11 and 13 of the NUP98 gene.(21) Interestingly, an enrichment of DNAse I hypersensitivity sites and a strong prediction of S/MAR are present in this region, corroborating its increased susceptibility to translocations, as seen in figure 1 below.

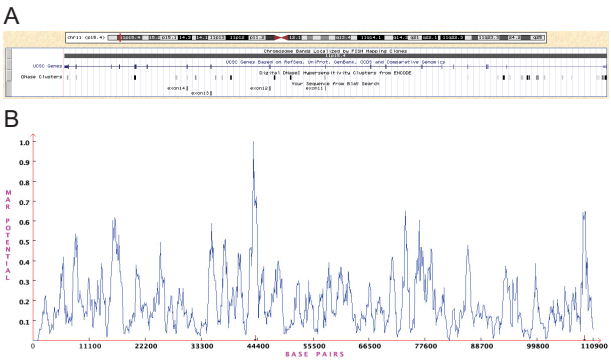


Figure 1: NUP98 gene chromatin structure. A – UCSC genome browser(52) tracks displaying NUP98 gene structure (most commonly translocated region, exons 11 to 14, marked) and DNAse I hypersensitivity sites. B – S/MAR enrichment on NUP98 gene, defined by MARFinder.(53)

The N-terminal portion of Nup98 (conserved in fusions) contains FG/GLFG repeats flanking a coiled-coil Rae1 interaction site (Fig. 2A). Even though a third of all nucleoporins contain FG repeats, Nup98 is the only GLFG repeat containing nucleoporin in humans.(22) At nuclear pore complexes (NPC), Nup98 interacts with transport molecules, mediating traffic through the NPC. The N-terminal of Nup98 interacts with XPO1, facilitating export of specific proteins from the nucleus,(23) and with TAP and Rae1, promoting mRNA export to the cytoplasm.(24)

Nup98 can also be found away from the NPC, dispersed through the nucleoplasm and in intranuclear structures called GLFG bodies (Fig. 2B).(25) In embryonic Drosophila cells, the intranuclear pool of Nup98 interacts with transcriptionally active genes and changes to the level of Nup98 present can modulate their expression, especially in developmental genes.(26) In human cells, the N-terminal GLFG repeats of nup98 have been shown to interact with histone acetyl transferases and histone deacetylases.(27, 28)

At least 27 different genes have been found translocated with NUP98 in leukemic patients (Fig. 2C). Most of these gene fusions lead to myeloid malignancies (AML, CML, MDS); however, six fusions have so far been identified in T-ALL patients. Nup98 fusions are rare (approximately 2% AML cases) however, they usually indicate a poor prognosis. Over half of Nup98 fusions are detected in patients under 20 years of age and only 25% occur in patients with therapy related malignancies.(21) The karyotype of patients with Nup98 fusions is usually simple, with no more than 3 chromosomal aberrations,(29) indicating a strong transformation potential for the fusions and arguing against an increase in genetic instability.

Nup98 fusion partners can be divided into homeodomain (HD) and non-HD containing proteins. Fusions with HD containing proteins always maintain the N-terminal GLFG domain of Nup98 fused in frame to the C-terminal HD of the partner gene.(30) NUP98-HOXA9 was the first fusion detected in an AML patient(31, 32) and it is currently the best-characterized Nup98 translocation. All non-HD containing partner genes encode putative coiled-coil motifs,(33) a domain usually involved in mediating protein-protein interactions. Chromatin recognition domains, such as plant homeodomain zinc fingers (PHD), are also recurrent in non-HD Nup98 fusion partners.(34)

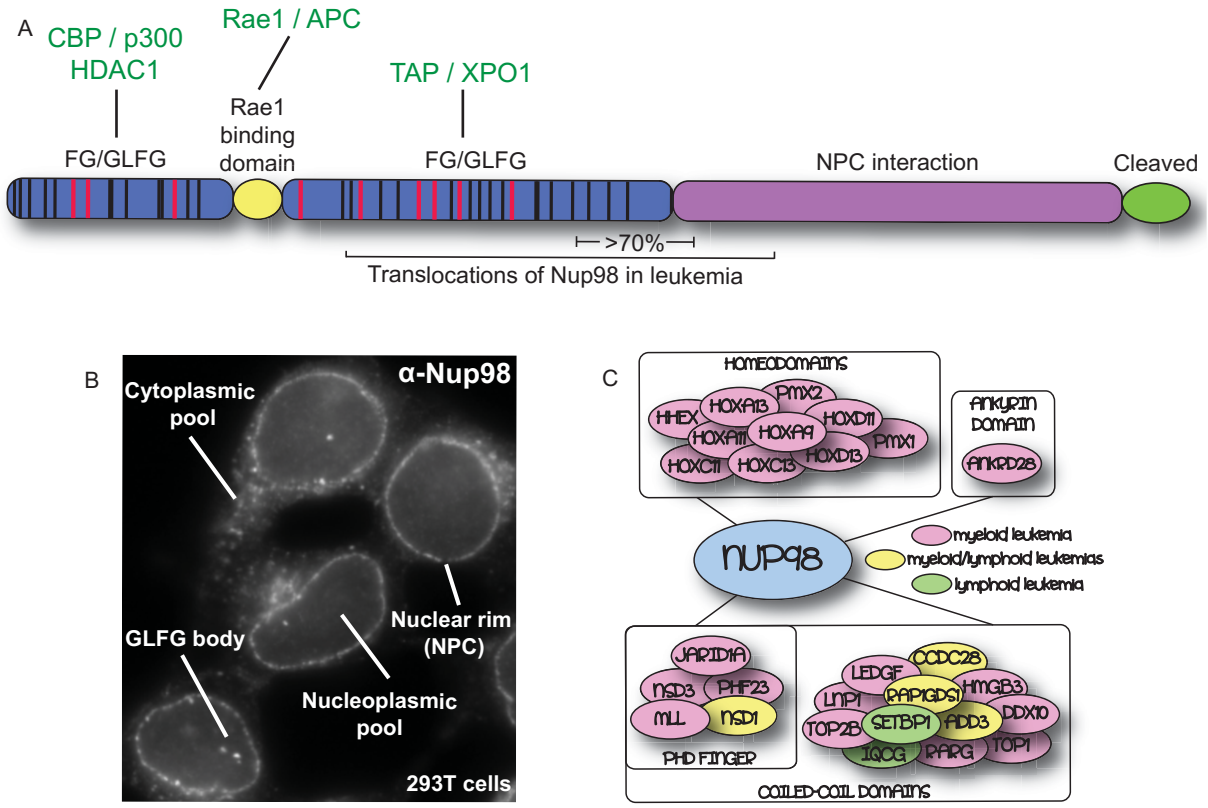


Figure 2: Nup98 and Nup98 fusions. A – Nup98 protein domains and known interactions, B – Nup98 localization in 293T cells visualized by immunofluorescence microscopy and C – Diagram of know Nup98 gene fusions and their characteristics.

So far, few common denominators have been identified when it come to the mechanism by which Nup98 fusions may lead to leukemogenesis. Characterizing their effects in altered gene expression indicates a few common targets: fusions of Nup98 with NSD1, KDM5A, PHF23, HOXA9, HOXD13, PRRX1, HHEX and DDX10 seem to increase the expression of HOXA cluster genes; NSD1, HOXA9 and DDX10 fusions also up-regulate the Hox co-factor Meis1; and HOXA9 or HOXD13 translocations increase the expression of interferon responsive genes.(21) A putative mechanism for how these fusions may alter gene expression has only been described for PHD domain containing translocations (NSD1, KDM5A and PHF23). These fusions seem to bind HOXA gene promoters (through the PHD finger) and recruit histone acetylases CBP/p300 (via GLFG domain) that modify chromatin into a transcriptionally active state.(34, 35) Nup98 fusions with HD containing proteins are assumed to act directly as transcription factors; they can in some cases collaborate with Meis1, a Hox co-factor that increases specificity and binding to target DNAs.(36)

Preliminary analysis:

Commonalities among NUP98 translocation partner genes.

Performing bioinformatics analysis of Nup98 fusion partner genes, a few recurrent themes are uncovered (Tab. I). Investigating the interaction profile of these partner proteins we notice that chromatin and/or DNA binding is a characteristic of 75% of them. The partner genes usually display direct DNA interaction or recognition of histone post-translational modifications, with 50% of all fusion partners working as transcriptional regulators. Separating Nup98 fusions leading to myeloid malignancies from those causing T-ALL, an even clearer picture appears: over 95% of myeloid related fusions can interact with DNA/chromatin, with over 2/3 of them acting as transcriptional regulators. T-ALL related fusions show no transcriptional regulators among partner genes and only one of them can interact with DNA/chromatin, likely pointing towards distinct molecular mechanisms for myeloid and lymphoid causing Nup98 fusions.

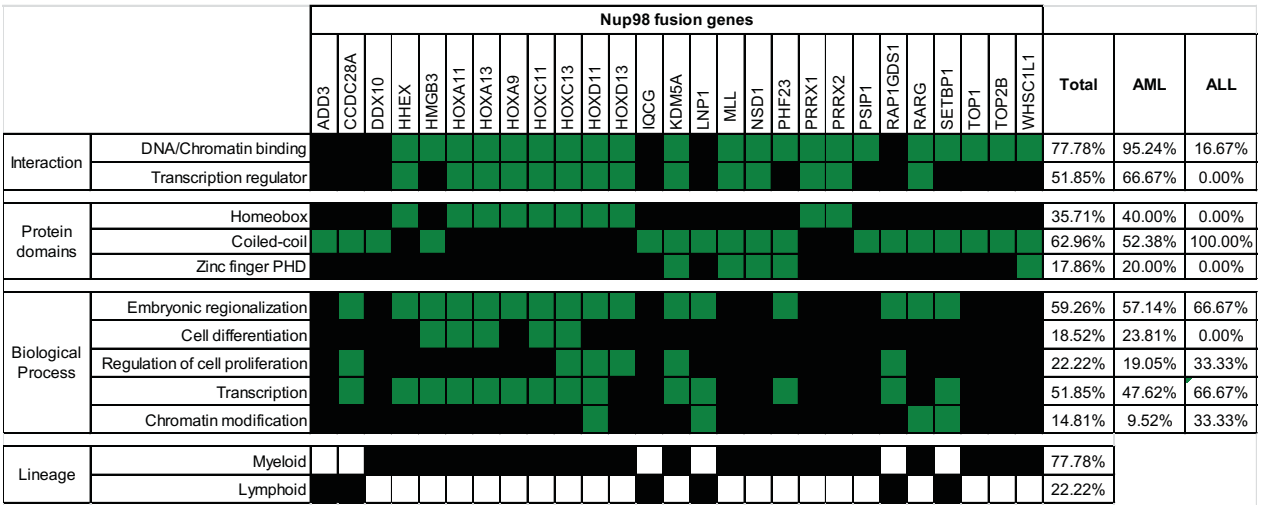
Evaluating biological processes affected by Nup98 fusion partner genes, we determine that almost 60% of them participate in embryonic regionalization and development, with over half being involved in transcription. Other biological processes over-represented among the fusion partners are regulation of cell proliferation, cell differentiation and chromatin modification (Tab. I). Interestingly, in Drosophila cells, Nup98 itself seems to regulate the transcription of developmental and cell cycle genes.(26)

The data above indicate that Nup98 fusions might function as rogue transcriptional regulators, especially in myeloid malignancies. It is possible that these fusions can affect gene expression acting directly as transcription factors (TF), altering histone modifications or deregulating other TFs.

Hypothesis:

The goal of this project is to define common mechanisms by which these different Nup98 fusions lead to malignancy. Based on the background and on the preliminary analysis provided above we hypothesize that Nup98 fusions, especially those leading to AML, might function as rogue tran-

Table I: Characteristics of Nup98 fusion partner genes.



scriptional regulators, and that their deregulated target genes might impair cell differentiation and increase self-renewal, setting the stage for malignant transformation and acute myeloid leukemia.

In this project, we propose to study the changes in gene expression caused by Nup98 translocations in bone marrow cells. Using data integration of previously published microarray experiments we will compare the effects of different Nup98 fusion proteins in the gene expression profile of bone marrow cells, leading to the discovery of specific pathways responsible for the disease phenotype. These relevant pathways can indicate key drug targets for this malignancy, and drug responses can be modeled in the existing networks, aiding in the development of new therapies for this disease.

Materials and Methods:

In order to further explore the possible role of Nup98 fusions as rogue transcriptional regulators, I collected microarray experiments of bone marrow cells transformed with different fusions for analysis. In order for this multi-experiment analysis to present biological significance, I only used results from experiments performed in similar conditions, ending up with 4 directly comparable sets: NUP98-HHEX, NUP98-HOXA9, NUP98-HOXA10, NUP98-HOXD13(37, 38).

All fusions were transduced into adult mice bone marrow cells using the retroviral vector MSCV-IRES-GFP (the empty vector was used as control), cells were FACS sorted before mRNA purification, target preparation and hybridization to Affymetrix Mouse Genome 430A Arrays. The raw (.CEL) file of each experiment was RMA (robust multichip average) normalized(39) and cross-study normalization was achieved using ComBat,(40) an empirical bayes method. The ANOVA statistical test ($p<0.05$) was used in AltAnalyze(39) to identify genes differentially expressed in control vs. Nup98 fusion samples, producing a list of genes whose expression was similarly affected by all Nup98 fusion proteins.

Functional annotation analysis of all gene lists was performed through the Database for Annotation, Visualization and Integrated Discovery (DAVID)(41) v6.7. Distant regulatory elements of co-regulated genes, such as transcription factor binding sites and CpG islands, were identified using DiRE.(42) Finally, functional protein association networks were inferred using STRING(43) version 9.0 and were imported into Cytoscape(44) for formatting.

Results and Discussion:

The expression of several genes is similarly altered in the presence of different Nup98 fusions.

Analyzing the genes whose expression was affected by all Nup98 fusions similarly, we see enrichment of a few biological processes: embryonic development, immune system formation and chromatin organization (Fig. 3A and Tab. S1). Evaluating only those genes whose expression was increased in the presence of all NUP98 translocations, we see enrichment for regulation of transcription, cell proliferation and immune system development (Fig. 3B). On the other hand, genes with decreased expression in the presence of Nup98 fusions are overrepresented for embryonic development, RNA processing and chromatin modification (Fig. 3C). These changes in the expression profile of genes involved in chromatin organization and modification can be correlated to the know epigenetic deregulation occurring in leukemic cells, especially as has been described for cells containing NUP98 translocations.(21) An increase in the expression of cell proliferation genes can also explain the expansion in the number of these LSCs. They abandon quiescence, as seen in HSCs, and actively proliferate in a deregulated manner, contributing to malignancy.

Genes with altered expression in the presence of NUP98 fusions are regulated by similar transcription factors.

Mapping the regulatory regions and transcription factor binding sites (TFBS) present in the deregulated genes, we can see that up and down-regulated genes present several regulatory regions in common, and 38 transcription factor binding site are enriched in both sets of genes (Tab. 2).

These transcription factors play relevant roles in cell cycle regulation, embryonic development, hematopoiesis, apoptosis and chromatin modifications (Fig. 4). They form a highly interconnected network,(43) indicating protein-protein interactions and transcriptional regulation among themselves (Fig. 5). Networks of TFs regulating genes with increased expression or genes with decreased expression are less interconnected than that of TFs that are present in both; however,

Table 2: Transcription factors with TFBS enriched in genes with deregulated expression in the presence of nup98 fusions.

Common TFs	Symbol	Gene ID	Occurence		Importance	
			up reg genes	down reg genes	up reg genes	down reg genes
CEBP	cebpa	12606	8.82%	4.66%	0.01544	0.0459
	cebpb	12608				
CMAF	maf	17132	4.41%	1.43%	0.0024	0.00896
CRX	crx	12951	2.94%	1.08%	0.00037	0.00329
DEC	Bhlhe40	20893	7.35%	2.87%	0.07077	0.00621
E2A	tcf3	21423	8.82%	7.17%	0.02206	0.00341
E2F1DP1	e2f1	13555	5.88%	2.15%	0.06655	0.00324
	tfdp1	21781				
E2F1DP1RB	e2f1	13555	5.88%	2.87%	0.03318	0.01098
	tfdp1	21781				
	rb1	19645				
E2F4DP1	e2f4	104394	5.88%	2.87%	0.00147	0.00229
	tfdp1	21781				
FOXJ2	foxj2	60611	7.35%	1.79%	0.13281	0.01915
FOXO4	foxo4	54601	7.35%	1.08%	0.00184	0.00184
GC	CpG islands		16.18%	19.35%	0.03808	0.13548
GCM	gcm2	107889	7.35%	4.30%	0.06147	0.03602
	gcm1	14531				
GRE	glucocorticoid res. elem.		1.47%	1.08%	0.0136	0.00197
HFH1	foxq1	15220	2.94%	0.36%	0.01454	0.00205
HIF1	hif1a	15251	11.76%	9.68%	0.02061	0.04876
HNF1	hnf1a	21405	8.82%	2.51%	0.03529	0.01356
	hnf1b	21410				
IK1	ikzf1	22778	4.41%	2.15%	0.02509	0.00403
LXR	Nr1h2	22260	7.35%	2.15%	0.07146	0.01984
	nr1h3	22259				
MYB	myb	17863	7.35%	2.51%	0.02941	0.00044
MYC	myc	17869	1.47%	2.87%	0.01418	0.02151
MYOGENIN	myog	17928	7.35%	6.09%	0.05285	0.00515
NERF	elf2	69257	8.82%	5.02%	0.09816	0.01944
NFKAPPAB50	nfkb1	18033	7.35%	2.51%	0.16176	0.00659
NFKB	rela	19697	8.82%	6.81%	0.02105	0.06719
NFY	nfia	18027	13.24%	4.66%	0.03174	0.00453
	nfib	18028				
	nfic	18029				
	nfix	18032				
P53	trp53	22059	8.82%	1.43%	0.14366	0.00409
PAX	pax1	18503	5.88%	2.51%	0.04559	0.01687
PAX9	pax9	18511	4.41%	2.87%	0.01665	0.01093
RSRFC4	mef2a	17258	2.94%	0.72%	0.00126	0.00905
SREBP	srebf1	20787	5.88%	2.87%	0.03676	0.0145
	srebf2	20788				
STAT	stat1	20846	7.35%	2.51%	0.0432	0.01443
TAL1	tal1	21349	5.88%	2.15%	0.06324	0.00225
TCF4	tcf4	21413	2.94%	2.15%	0.04403	0.01667
TEL2	telo2	71718	1.47%	2.15%	0.00729	0.01726
TST1	pou3f1	18991	1.47%	1.08%	0.01556	0.0121
USF2	usf2	22282	1.47%	2.87%	0.00023	0.02222
WT1	wt1	22431	14.71%	6.45%	0.03808	0.03055
ZF5	zfp161	22666	27.94%	23.30%	0.08507	0.13499
OCT4	pou5f1	18999	2.94%	1.08%	0.01838	0.00753

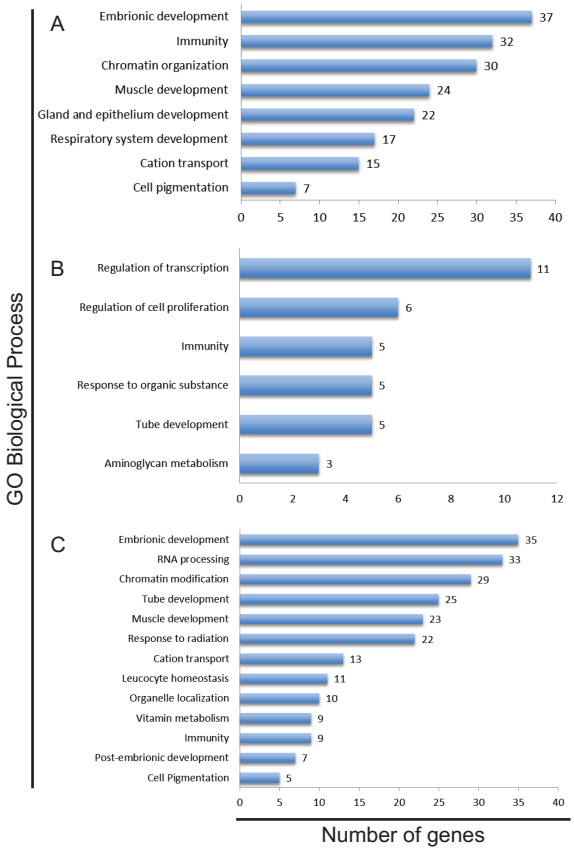


Figure 3: GO annotation biological process enrichment A) all genes differentially expressed; B) up-regulated genes; C) down-regulated genes in the presence of Nup98 fusion proteins.

each network contains at least one module of a few highly interconnected TFs (Fig. S1 and S2).

Some of these transcription factors are known to also be deregulated in other types of leukemia.

More in-depth analysis of the role of a few of the TFs putatively regulating genes with altered expression in cells bearing NUP98 translocations can uncover possible novel mechanisms by which these fusions may lead to leukemogenesis. Wilms Tumor 1 (WT1) is one of the most enriched TFBS in up and down regulated genes. WT1 is a TF with expression restricted to hematopoietic progenitor cells in the bone marrow with a role in their self-renewal. Mutations in WT1 indicate a worse prognosis in acute leukemias and can be found in approximately 10% of AML cases.(45) This transcription factor can work as both a tumor suppressor and an oncogene, and it can enhance or repress transcription of its target genes (such as MYC and BCL-2) depending on cellular conditions. (46) WT1 and N-terminal Nup98 (present in fusions) both interact with CBP, providing an interesting putative mechanism on how Nup98 fusions might be affecting this TF and its targets without affecting its expression level.

As another example, CEBPA appears as an important regulatory TF in genes deregulated by Nup98 fusions. Interestingly its own expression is reduced nearly 2 fold in Nup98 fusion expressing cells. Decreased CEBPA expression in BM cells decreases differentiation and increases proliferation of myeloid progenitors leading to leukemia. Another leukemic fusion, AML1-ETO, has also been shown to down-regulate CEBPA expression, and CEBPA mutations that abrogate its function or generate dominant negatives have also been described as leukemogenic.(47)

TFBS for MYC are also overrepresented in this deregulated gene set, albeit to a lesser extent. MYC expression increases 40% in the presence of Nup98 fusions, similarly to what is observed with several other leukemic chromosomal aberrations (AML1-ETO, PML-RARA, PLZF-RARA, FLT3-ITD) shown to induce c-myc activation. Overexpression of c-myc alone in BM cells can quickly induce fatal AML.(48)

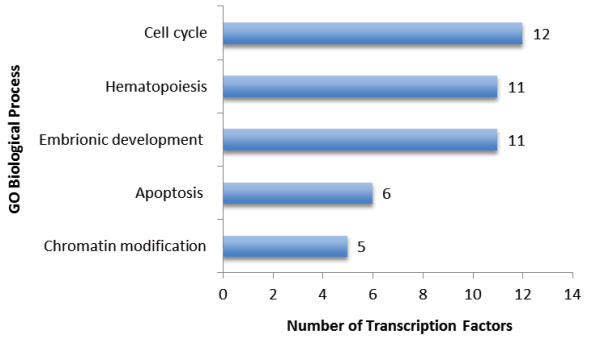


Figure 4: GO annotation biological process enrichment for all TFs with TFBS enriched in genes differentially expressed in the presence of Nup98 fusion proteins.

The identified transcription factor network is highly similar to the network of transcription factors regulating growth arrest and differentiation in human myeloid cells.

Given the above results, we can hypothesize that these NUP98 translocations are deregulating transcription factors that control differentiation and self-renewal in primitive hematopoietic cells. The network of transcription factors regulating growth arrest and differentiation in a human myeloid cell line has already been described.(49) Superposing the microarray results obtained above into this previously published network (Fig. 6), we notice that over 80% of the TFs thought to regulate growth arrest and differentiation have decreased expression in cells containing Nup98 fusions. Additionally, 70% of the TFs represented in this network have enriched TFBS in genes with deregulated expression upon NUP98 translocations. This reinforces the idea that Nup98 fusions can deregulate key TFs in myeloid cells, leading to a cascade of changes in their gene expression profile that ultimately disrupts differentiation and proliferation, promoting leukemogenesis.

Enrichment of CpG islands in the promoter region of genes with deregulated expression further indicates epigenetic deregulation in the presence of Nup98 fusions.

A final interesting observation is the enrichment of CpG islands in the promoters of genes with deregulated expression in Nup98 fusions, indicating an important role for epigenetic changes in their altered expression. Epigenetic regulation of gene expression is a hallmark of hematopoiesis, and several leukemic translocations have been shown to alter transcription by altering epigenetic markers in their target genes.(3) HOX genes are an example of this epigenetic regulation and they are highly enriched for CpG islands in humans.(50) Several Nup98 fusions promote HOX genes up-regulation,(21) some of them have been shown to alter histone post transla-

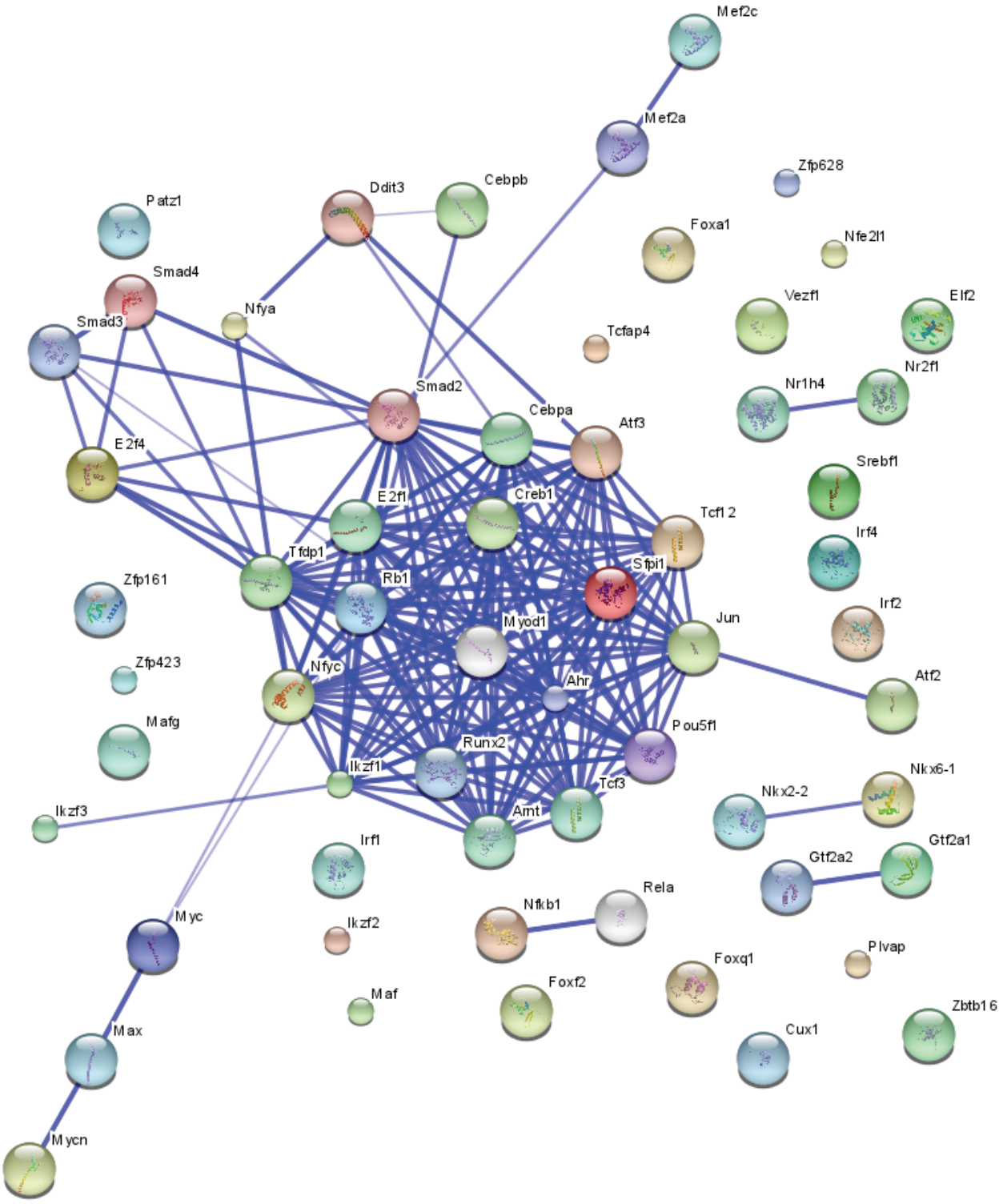


Figure 5: Protein-protein interaction network of TFs with TFBS enriched in genes differentially expressed in Nup98 translocations.

tional modifications in the HOX locus,(34, 35) and it is possible that Nup98 fusions may also lead to epigenetic deregulation affecting DNA methylation or histone modifications in the CpG islands identified in deregulated genes.

As well as DNA methylation, CpG islands in gene promoters can be silenced by polycomb group proteins.(51) In mouse hematopoietic progenitors NUP98-KDM5A fusions bind the promoters of HOXA6-A10, this stops polycomb complex proteins from silencing them, maintains H3K4me3 and acetylated histones in their promoter, and leads to increased expression of these genes.(34) Similarly, NUP98-NSD1 binds the promoters of HOXA7 and HOXA9, leading to histone

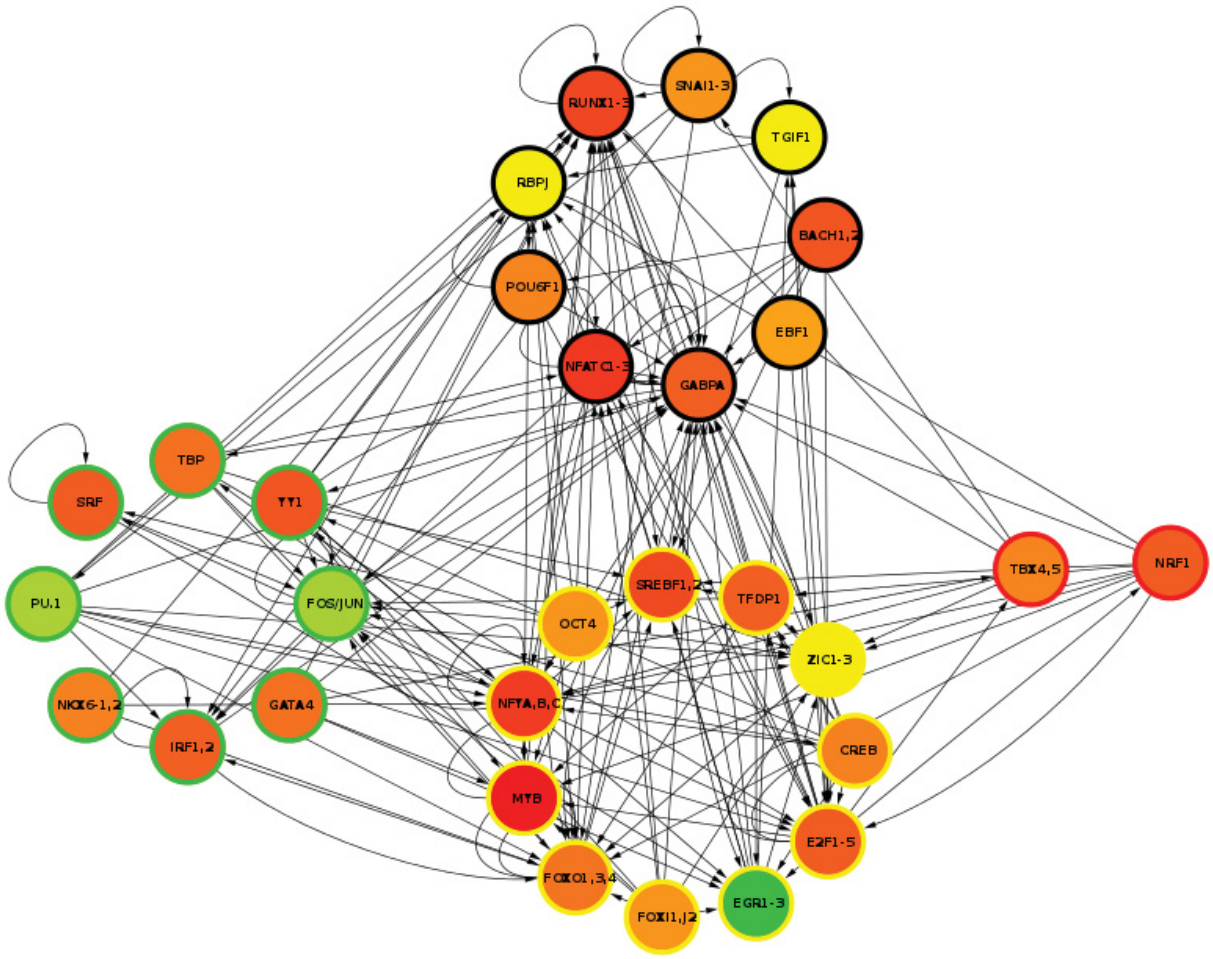


Figure 6: Transcriptional network of growth arrest and differentiation in a human myeloid leukemia cell line.(49) Superposition of results from Nup98 fusion microarrays into the previously published network of growth arrest and differentiation. Node color indicates expression level of TF (red – decreased, green – increased) and node border indicates enrichment of TFBS in groups of genes with increased (green) or decreased (red) expression (yellow border - TFBS present in up and down regulated genes, black – TFBS absent in deregulated genes).

acetylation (p300 recruitment by Nup98's N-terminal) and H3K36me3 (NSD1 C-terminal function) that increase gene expression.(35) It's possible that other Nup98 fusion proteins may have similar mechanisms of epigenetic deregulation.

Concluding remarks:

The project described here aims at improving our understanding of how NUP98 translocations lead to leukemia. Based on previous literature and the results presented we can see that most Nup98 fusion partners provide a DNA/chromatin interacting interface to these fusion proteins. This leads to the deregulation of sets of genes that increase the proliferation of these cells, as well as to the deregulation of genes that alter epigenetics. The genes with altered expression in the presence of Nup98 fusions are regulated by similar transcription factors, and these form a highly interconnected network. Part of this regulatory transcription factor network is itself down-regulated and very similar to the network of transcription factors regulating growth arrest and differentiation in human myeloid cells. It's likely therefore, that these Nup98 fusions act as rogue transcriptional regulators, affecting mainly cell differentiation and self-renewal.

References:

1. Canadian Cancer Society., Public Health Agency of Canada., Statistics Canada. Canadian cancer statistics 2011 featuring colorectal cancer. Toronto, Ont.: Canadian Cancer Society; 2011. Available from: http://site.ebrary.com/lib/abhealth/Doc?id=10471043Available from: http://site.ebrary.com/lib/macewan/Doc?id=10471043Available from: http://site.ebrary.com/lib/ualberta/Doc?id=10471043.
2. Vardiman JW. The World Health Organization (WHO) classification of tumors of the hematopoietic and lymphoid tissues: an overview with emphasis on the myeloid neoplasms. Chem Biol Interact. 2010 Mar;184(1-2):16-20. PubMed PMID: 19857474. eng.
3. Chen J, Odenike O, Rowley JD. Leukaemogenesis: more than mutant genes. Nat Rev Cancer. 2010 Jan;10(1):23-36. PubMed PMID: 20029422. Pubmed Central PMCID: PMC2972637. eng.
4. Passegué E, Jamieson CH, Ailles LE, Weissman IL. Normal and leukemic hematopoiesis: are leukemias a stem cell disorder or a reacquisition of stem cell characteristics? Proc Natl Acad Sci U S A. 2003 Sep;100 Suppl 1:11842-9. PubMed PMID: 14504387. Pubmed Central PMCID: PMC304096. eng.
5. Mohrin M, Bourke E, Alexander D, Warr MR, Barry-Holson K, Le Beau MM, et al. Hematopoietic stem cell quiescence promotes error-prone DNA repair and mutagenesis. Cell Stem Cell.

2010 Aug;7(2):174-85. PubMed PMID: 20619762. Pubmed Central PMCID: PMC2924905. eng.
6. Zhang Y, Rowley JD. Chromatin structural elements and chromosomal translocations in leukemia. DNA Repair (Amst). 2006 Sep;5(9-10):1282-97. PubMed PMID: 16893685. eng.
7. Zhang Y, Gostissa M, Hildebrand DG, Becker MS, Boboila C, Chiarle R, et al. The role of mechanistic factors in promoting chromosomal translocations found in lymphoid and other cancers. Adv Immunol. 2010;106:93-133. PubMed PMID: 20728025. Pubmed Central PMCID: PMC3073861. eng.
8. Mitelman Database of Chromosome Aberrations and Gene Fusions in Cancer http://cgap.nci.nih.gov.login.ezproxy.library.ualberta.ca/Chromosomes/Mitelman"2011 [cited 2011 December 18th, 2011].
9. Bachas C, Schuurhuis GJ, Hollink IH, Kwidama ZJ, Goemans BF, Zwaan CM, et al. High-frequency type I/II mutational shifts between diagnosis and relapse are associated with outcome in pediatric AML: implications for personalized medicine. Blood. 2010 Oct;116(15):2752-8. PubMed PMID: 20592250. eng.
10. Scandura JM, Boccuni P, Cammenga J, Nimer SD. Transcription factor fusions in acute leukemia: variations on a theme. Oncogene. 2002 May;21(21):3422-44. PubMed PMID: 12032780. eng.
11. Cedar H, Bergman Y. Epigenetics of haematopoietic cell development. Nat Rev Immunol. 2011 Jul;11(7):478-88. PubMed PMID: 21660052. eng.
12. He H, Hua X, Yan J. Epigenetic regulations in hematopoietic Hox code. Oncogene. 2011 Jan;30(4):379-88. PubMed PMID: 20972460. eng.
13. Xu S, Powers M. Nuclear pore proteins and cancer. Semin Cell Dev Biol. 2009 Jul;20(5):620-30. PubMed PMID: 19577736. Pubmed Central PMCID: PMC2706781. eng.
14. Chen ST, Lee JC. An inflammatory myofibroblastic tumor in liver with ALK and RANBP2 gene rearrangement: combination of distinct morphologic, immunohistochemical, and genetic features. Hum Pathol. 2008 Dec;39(12):1854-8. PubMed PMID: 18701132. eng.
15. Köhler A, Hurt E. Gene regulation by nucleoporins and links to cancer. Mol Cell. 2010 Apr;38(1):6-15. PubMed PMID: 20385085. eng.
16. Ma Z, Hill DA, Collins MH, Morris SW, Sumegi J, Zhou M, et al. Fusion of ALK to the Ran-binding protein 2 (RANBP2) gene in inflammatory myofibroblastic tumor. Genes Chromosomes Cancer. 2003 May;37(1):98-105. PubMed PMID: 12661011. eng.
17. De Keersmaecker K, Rocnik JL, Bernad R, Lee BH, Leeman D, Gielen O, et al. Kinase activation and transformation by NUP214-ABL1 is dependent on the context of the nuclear pore. Mol Cell. 2008 Jul;31(1):134-42. PubMed PMID: 18614052. eng.
18. Van Vlierberghe P, van Grotel M, Tchinda J, Lee C, Beverloo HB, van der Spek PJ, et al. The recurrent SET-NUP214 fusion as a new HOXA activation mechanism in pediatric T-cell acute lymphoblastic leukemia. Blood. 2008 May;111(9):4668-80. PubMed PMID: 18299449. Pubmed Central PMCID: PMC2343598. eng.
19. Ageberg M, Drott K, Olofsson T, Gullberg U, Lindmark A. Identification of a novel and myeloid specific role of the leukemia-associated fusion protein DEK-NUP214 leading to increased protein synthesis. Genes Chromosomes Cancer. 2008 Apr;47(4):276-87. PubMed PMID: 18181180. eng.
20. Gamble MJ, Fisher RP. SET and PARP1 remove DEK from chromatin to permit access by the transcription machinery. Nat Struct Mol Biol. 2007 Jun;14(6):548-55. PubMed PMID: 17529993. eng.
21. Gough SM, Slape CI, Aplan PD. NUP98 gene fusions and hematopoietic malignancies: common themes and new biologic insights. Blood. 2011 Dec;118(24):6247-57. PubMed PMID: 21948299. eng.
22. Iwamoto M, Asakawa H, Hiraoka Y, Haraguchi T. Nucleoporin Nup98: a gatekeeper in the eukaryotic kingdoms. Genes Cells. 2010 Jun;15(7):661-9. PubMed PMID: 20545767. eng.
23. Oka M, Asally M, Yasuda Y, Ogawa Y, Tachibana T, Yoneda Y. The mobile FG nucleoporin Nup98 is a cofactor for Crm1-dependent protein export. Mol Biol Cell. 2010 Jun;21(11):1885-96. PubMed PMID: 20375145. Pubmed Central PMCID: PMC2877646. eng.
24. Blevins M, Smith A, Phillips E, Powers M. Complex formation among the RNA export proteins Nup98, Rae1/Gle2, and TAP. J Biol Chem. 2003 Jun;278(23):20979-88. PubMed PMID: 12637516. eng.
25. Griffis E, Altan N, Lippincott-Schwartz J, Powers M. Nup98 is a mobile nucleoporin with transcription-dependent dynamics. Mol Biol Cell. 2002 Apr;13(4):1282-97. PubMed PMID: 11950939. Pubmed Central PMCID: PMC102269. eng.
26. Kalverda B, Pickersgill H, Shloma V, Fornerod M. Nucleoporins directly stimulate expression of developmental and cell-cycle genes inside the nucleoplasm. Cell. 2010 Feb;140(3):360-71. PubMed PMID: 20144760. eng.
27. Kasper L, Brindle P, Schnabel C, Pritchard C, Cleary M, van Deursen J. CREB binding protein interacts with nucleoporin-specific FG repeats that activate transcription and mediate NUP98-HOXA9 oncogenicity. Mol Cell Biol. 1999 Jan;19(1):764-76. PubMed PMID: 9858599. Pubmed Central PMCID: PMC83933. eng.
28. Bai X, Gu B, Yin T, Niu C, Xi X, Zhang J, et al. Trans-repressive effect of NUP98-PMX1 on PMX1-regulated c-FOS gene through recruitment of histone deacetylase 1 by FG repeats. Cancer Res. 2006 May;66(9):4584-90. PubMed PMID: 16651408. eng.
29. Romana SP, Radford-Weiss I, Ben Abdelali R, Schluth C, Petit A, Dastugue N, et al. NUP98 rearrangements in hematopoietic malignancies: a study of the Groupe Francophone de Cytogénétique Hématologique. Leukemia. 2006 Apr;20(4):696-706. PubMed PMID: 16467868. eng.
30. Moore MA, Chung KY, Plasilova M, Schuringa JJ, Shieh JH, Zhou P, et al. NUP98 dysregulation in myeloid leukemogenesis. Ann N Y Acad Sci. 2007 Jun;1106:114-42. PubMed PMID: 17442773. eng.
31. Nakamura T, Largaespada D, Lee M, Johnson L, Ohyashiki K, Toyama K, et al. Fusion of the nucleoporin gene NUP98 to HOXA9 by the chromosome translocation t(7;11)(p15;p15) in human myeloid leukaemia. Nat Genet. 1996 Feb;12(2):154-8. PubMed PMID: 8563753. eng.
32. Borrow J, Shearman AM, Stanton VP, Becher R, Collins T, Williams AJ, et al. The t(7;11)(p15;p15) translocation in acute myeloid leukaemia fuses the genes for nucleoporin NUP98 and class I homeoprotein HOXA9. Nat Genet. 1996 Feb;12(2):159-67. PubMed PMID: 8563754. eng.
33. Hussey D, Dobrovic A. Recurrent coiled-coil motifs in NUP98 fusion partners provide a clue to leukemogenesis. Blood. 2002 Feb;99(3):1097-8. PubMed PMID: 11822362. eng.
34. Wang G, Song J, Wang Z, Dormann H, Casadio F, Li H, et al. Haematopoietic malignancies caused by dysregulation of a chromatin-binding PHD finger. Nature. 2009 Jun;459(7248):847-51. PubMed PMID: 19430464. Pubmed Central PMCID: PMC2697266. eng.
35. Wang GG, Cai L, Pasillas MP, Kamps MP. NUP98-NSD1 links H3K36 methylation to Hox-A gene activation and leukaemogenesis. Nat Cell Biol. 2007 Jul;9(7):804-12. PubMed PMID: 17589499. eng.
36. Yung E, Sekulovic S, Argiropoulos B, Lai CK, Leung M, Berg T, et al. Delineating domains and functions of NUP98 contributing to the leukemogenic activity of NUP98-HOX fusions. Leuk Res. 2011 Apr;35(4):545-50. PubMed PMID: 21130494. eng.
37. Palmqvist L, Pineault N, Wasslavik C, Humphries RK. Candidate genes for expansion and transformation of hematopoietic stem cells by NUP98-HOX fusion genes. PLoS One. 2007;2(8):e768. PubMed PMID: 17712416. Pubmed Central PMCID: PMC1942085. eng.
38. Jankovic D, Gorello P, Liu T, Ehret S, La Starza R, Desjobert C, et al. Leukemogenic mechanisms and targets of a NUP98/HHEX fusion in acute myeloid leukemia. Blood. 2008 Jun;111(12):5672-82. PubMed PMID: 18388181. eng.
39. Emig D, Salomonis N, Baumbach J, Lengauer T, Conklin BR, Albrecht M. AltAnalyze and DomainGraph: analyzing and visualizing exon expression data. Nucleic Acids Res. 2010 Jul;38(Web Server issue):W755-62. PubMed PMID: 20513647. Pubmed Central PMCID: PMC2896198. eng.
40. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in

microarray expression data using empirical Bayes methods. Biostatistics. 2007 Jan;8(1):118-27. PubMed PMID: 16632515. eng.
41. Huang dW, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res. 2009 Jan;37(1):1-13. PubMed PMID: 19033363. Pubmed Central PMCID: PMC2615629. eng.
42. Gotea V, Ovcharenko I. DiRE: identifying distant regulatory elements of co-expressed genes. Nucleic Acids Res. 2008 Jul;36(Web Server issue):W133-9. PubMed PMID: 18487623. Pubmed Central PMCID: PMC2447744. eng.
43. Szklarczyk D, Franceschini A, Kuhn M, Simonovic M, Roth A, Minguez P, et al. The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored. Nucleic Acids Res. 2011 Jan;39(Database issue):D561-8. PubMed PMID: 21045058. Pubmed Central PMCID: PMC3013807. eng.
44. Smoot ME, Ono K, Ruscheinski J, Wang PL, Ideker T. Cytoscape 2.8: new features for data integration and network visualization. Bioinformatics. 2011 Feb;27(3):431-2. PubMed PMID: 21149340. Pubmed Central PMCID: PMC3031041. eng.
45. Owen C, Fitzgibbon J, Paschka P. The clinical relevance of Wilms Tumour 1 (WT1) gene mutations in acute leukaemia. Hematol Oncol. 2010 Mar;28(1):13-9. PubMed PMID: 20013787. eng.
46. Yang L, Han Y, Suarez Saiz F, Saurez Saiz F, Minden MD. A tumor suppressor and oncogene: the WT1 story. Leukemia. 2007 May;21(5):868-76. PubMed PMID: 17361230. eng.
47. Pabst T, Mueller BU. Complexity of CEBPA dysregulation in human acute myeloid leukemia. Clin Cancer Res. 2009 Sep;15(17):5303-7. PubMed PMID: 19706798. eng.
48. Luo H, Li Q, O'Neal J, Kreisel F, Le Beau MM, Tomasson MH. c-Myc rapidly induces acute myeloid leukemia in mice without evidence of lymphoma-associated antiapoptotic mutations. Blood. 2005 Oct;106(7):2452-61. PubMed PMID: 15972450. eng.
49. Suzuki H, Forrest AR, van Nimwegen E, Daub CO, Balwierz PJ, Irvine KM, et al. The transcriptional network that controls growth arrest and differentiation in a human myeloid leukemia cell line. Nat Genet. 2009 May;41(5):553-62. PubMed PMID: 19377474. eng.
50. Branciamore S, Chen ZX, Riggs AD, Rodin SN. CpG island clusters and pro-epigenetic selection for CpGs in protein-coding exons of HOX and other transcription factors. Proc Natl Acad Sci U S A. 2010 Aug;107(35):15485-90. PubMed PMID: 20716685. Pubmed Central PMCID: PMC2932574. eng.
51. Deaton AM, Bird A. CpG islands and the regulation of transcription. Genes Dev. 2011 May;25(10):1010-22. PubMed PMID: 21576262. Pubmed Central PMCID: PMC3093116. eng.
52. Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, et al. The human genome browser at UCSC. Genome Res. 2002 Jun;12(6):996-1006. PubMed PMID: 12045153. Pubmed Central PMCID: PMC186604. eng.
53. Singh GB, Kramer JA, Krawetz SA. Mathematical model to predict regions of chromatin attachment to the nuclear matrix. Nucleic Acids Res. 1997 Apr;25(7):1419-25. PubMed PMID: 9060438. Pubmed Central PMCID: PMC146595. eng.
54. Hubner NC, Bird AW, Cox J, Splettstoesser B, Bandilla P, Poser I, et al. Quantitative proteomics combined with BAC TransgeneOmics reveals in vivo protein interactions. J Cell Biol. 2010 May;189(4):739-54. PubMed PMID: 20479470. Pubmed Central PMCID: PMC2872919. eng.
55. Gehlenborg N, O'Donoghue SI, Baliga NS, Goesmann A, Hibbs MA, Kitano H, et al. Visualization of omics data for systems biology. Nat Methods. 2010 Mar;7(3 Suppl):S56-68. PubMed PMID: 20195258. eng.
56. Rhee HS, Pugh BF. Comprehensive Genome-wide Protein-DNA Interactions Detected at Single-Nucleotide Resolution. Cell. 2011 Dec;147(6):1408-19. PubMed PMID: 22153082. Pubmed Central PMCID: PMC3243364. eng.
57. Hawkins RD, Hon GC, Ren B. Next-generation genomics: an integrative approach. Nat Rev Genet. 2010 Jul;11(7):476-86. PubMed PMID: 20531367. eng.
58. Martin JA, Wang Z. Next-generation transcriptome assembly. Nat Rev Genet. 2011 Oct;12(10):671-82. PubMed PMID: 21897427. eng.
59. Cheng C, Yan KK, Hwang W, Qian J, Bhardwaj N, Rozowsky J, et al. Construction and analysis of an integrated regulatory network derived from high-throughput sequencing data. PLoS Comput Biol. 2011 Nov;7(11):e1002190. PubMed PMID: 22125477. Pubmed Central PMCID: PMC3219617. eng.

APPENDIX B

Appendix: Fluorescent *in situ* hybridization probes

Table B.1: FISH probes conjugated to Quasar 570 fluor mapping to hepatitis C virus (AB047639.1) positive strand RNA between nucleotides 701 to 2999.

Probes sequence
gttagggtgtcgatgacttt
ctgttgcataattaaccccg
aaaaggggaaaccgggtagg
aacagggccagcaagaagat
atgtagctgctactggtatt
tgtcattggagcagtcattg
cacaaccatatcgatgtgcg
cgagacgatgaacacctggg
tgcacaaaccagtggtactg
ggtagatggagcaattgcat
gaccagttcatcatcatgtc
cgatgtctatgatgacctcg
tgcatagagaagtaggccaa
aggatgacaatgaccttcgc
caatcacgttggtggaacgt
aactgccgttggtgttaatg
tcaaggcagtacggttgatg
gtgttcaaggagtcattgca
agaacaaggccgcgagaaag
gacgagttaaagcggttggt
attatcctcgtactgtaggg
tcatatcctctggattggtg
gggtgaaacagtacactggg
gtgctgttcagtaggaagac
tcttggtgaaaccagtggag
aatccgtagggcacaacaag
gcatcaggatgcttcctaaa
cccagaaccacacttaatat
agagtctgtaagggtagtgg
attgactgtgcaggggtaat
ccctacatacattcttatct
cacgagtgaagttgcatgcg
tggtagagtgcaacagagga
taagtctgagtaggtgcagg
gaagaccagttgacaaagcg
acgatgttctggtgaaggtg
agaggccatacatgtattgt
aataagagtaccacccactc
ccaacaagatgagcatccac
tgcaagacgaccaacttctc
aaatataggaggccatggca
caagctgccacgaagaagat

Continued on next page

Table B.1 – FISH probes set 3

Probes sequence
caataggtggtcaaggggac
catgagcagtaggcagaagg
aggtgcgtcataggcataag
tatcaacaaacccacgccta
gtgagtgtgaagagggtgat
cgggcagaatatagtgacgg

Table B.2: FISH probes conjugated to Quasar 670 fluor mapping to hepatitis C virus (AB047639.1) positive strand RNA between nucleotides 3734 to 9415.

Probes sequence
gtgaccagatataggtcgac
caaggtcgaaatgggtctcg
caacgtcgagtgtctcaacg
gtgctgttgtcactgaaagt
ccatatgtggagtacgtgat
cttgatgcaggataggggaa
gtcgatcacggagtcaaagt
tggttatagtgaaggtgggg
tcgttcaccagtggaaacat
tcgtagcactcacaaagcac
gctttgtttgggagaggaag
attggtaatagggcccaaac
cgatgatggaaacgcatcca
caaaagcctcatacaggacc
gagaaggatggtggtactgg
gatcttgaatgcgacgaggg
agcaaaggcaataagcctgt
gttatccaattgtggagtct
gcagcttggggaacaattta
aattgataggaaaggtcccc
tcagtcctgttacataggag
acagcataggaattaagccc
tgaagctctacctgatcaga
tcggagcaagtagaccaaga
tatgacatggagcagcacac
ttgattggcaacttttcctc
ttatggtatcgcaacagcga
gctctttgatgttgtacagt
gtcctgtcaaaagttacctt
cttggatcttgcagaatggg
acgcagaacacctcattttt

Continued on next page

Table B.2 – FISH probes set 4

Probes sequence
tcagggtaaacgatgaggcg
aatgtcatagagggccattt
cccatgctttcaagagatac
aaagtctctcagtcagcgag
cacatagcatgtgatggtgt
ctggctttctgagatgacta
caggatgttattagctccag
caacgccacagacacatttg
gccatgaattgataggggag
tatggttggagcatactgga
gaagtgtgtcattaggacca
ttatggctggaaggtccaaa
ggtgagagtatgtgtgcata
ccaactggataagtccagta
acgctgtgaaaaatgtcgcc
gaagagtaatgagcggggtc
acccctacgaaaagtaggag

Table B.3: FISH probes conjugated to Quasar 570 fluor mapping to hepatitis C virus (AB047639.1) negative strand RNA between nucleotides 432 to 3035.

Probes sequence
ttaccaaatggcttttggcg
tgggccgtcactatattctg
gattcaggagtgggtaccac
tgatattgatcaccctcttc
acctgtgcacggacagatag
tcctatattttgccatcttc
cattggagaagttggtcgtc
tgtggatgctcatcttgttg
tactcttattcctgctctta
aatacatgtatggcctctca
tccaccttcaccagaacatc
ctgcacctactcagacttac
ttgcactctaccacggaatg
ttggaggacagggacaggag
catgcaacttcactcgtggg
aatgtatgtaggggggttg
gcacagtcaattttaccatc
cactacccttacagactctg
tattaagtgtggttctgggc
ttgttttaggaagcatcctg

Continued on next page

Table B.3 – FISH probes set 5

Probes sequence
tggtttcaccaagacttgtg
gtcttcctactgaacagcac
acctacacatggggagagaa
cagaggatatgaggccgtac
accctacagtacgaggataa
aacatcgaggctttccggat
ttctacaccaaccgctttaa
caatgactccttgaacaccg
ttggcacatcaaccgtactg
cagcagaacattcagctcat
acgttccaccaacgtgattg
tcattgtcatccttctgctg
ctacttctctatgcagggag
gaggtcatcatagacatcgt
atgatcctggcgtacgtgat
tggcatgggacatgatgatg
agaatgcaattgctccatct
gacgcacatcgatatggttg
caaacatggctgtgcggcag
tgcgagagagtggggaatac
aatgacagcatcacttggca
ataccagtagcagctacatg
cttttctatcttcttgctgg
aggacggggttaattatgca
cgcaacgtgggtaaagtcat
caaggcctggggaaaaccag
gtgtgcgcacgacaaggaaa
ttggcggagtatacttgttg

Table B.4: FISH probes conjugated to Quasar 570 fluor mapping to Zika virus (KF993678.1) positive strand RNA between nucleotides 1515 to 3483.

Probes sequence
ctagaaccacgacagtttgc
agtttatccattttcaggcg
ggtacacagggagtatgaca
ccgggatcttagtgaatgtg
ctgagctggaaccttgcaag
cagtgattacagggttagcg
gcatcatcttagagttctca
cccaaatggtggatcaagtt
ccgactcctatgacaatgta

Continued on next page

Table B.4 – FISH probes set 6

Probes sequence
ctctcacagtggcttcaaat
caagactgccattctcttgg
caactgatccaaagtcccag
cttgcccaatgagttgagag
atgatttgaaagctgctcca
tgtgagaaccaggacattcc
agatccattctttgtattca
taaggccaagcacataaggg
cagagacggctgtggataag
ggcttcaacgtcgttataga
caatctacgaggggagtcag
cttgaaacagaggagatccc
cttctactgatctccacatg
tcttccaggattgcgttgag
aacgaccgtcagttgaactc
atggggtttttacagatcc
gaagtacgattttccccaag
tatttgtctttgctgctctg
tgtcaccatccacgacaaag
ttccatgctctatgtttgag
ccatgatcctccacaagaaa
acactagtgtgaaatacccc
tagcctagatcactgtgtac
tcgttcttctcactctcaat
gttttcatctcgatcaggtg
cacaatgtgtgggactttgg
cactttcttctattccatct
ccagctaaagacttgggtat
ctggtgttgtgatggctgag
tttcatttgggtcctgtaac
accgaatttcaagctcttca
tggttgatctcagagatggt
gaacgacagtgggggcattg
cataccaacagccatcttta
atgtgatcagttgatcctgc
cactccaagggagaagtgat
tgcaccatgagcagaatcac
ctgccattgatgtgcttatg
ctcattgaaaatcctcccag

Table B.5: FISH probes conjugated to Quasar 670 fluor mapping to Zika virus (KF993678.1) positive strand RNA between nucleotides 6542 to 8483.

Probes sequence
aattctggctggctcaattt
ttctctggctcaggtatgag
gttgtcctgaggagatcttt
gtaatcaagcccagaagacc
caaccatccgagttcattgg
ttagatggcttaggtcactc
caggtcaatgtccattgaga
gtcaaagcagcgtagatagc
gacggctggggtaatgaaag
gtatgaagtggtcactgcat
tcgccattaaggagtagttg
tttgcccataccaaacaaca
caaagtcccatgcgtagaat
gtagcaacctatcattagca
tagggtcaggggtgttaatt
cacgagcaaaatgatggcca
cgacagggttcttcatgatg
tcaatgtcagtcaccactat
ttgggggtcaattgtcattg
acggctactgctatgagtag
gaagttgcagctgtgatcag
tccagtacttgttcggagag
acacagtgaagtggctgtag
attagagaagctccagccaa
cagcgtttcttgttactgtg
tgtaggagtagaactccagg
tgccacatccaagatcaatg
cttgaactttgcggatggtg
ccttttgtgtatcctttcac
ccactcttaagacggactat
tcagccgccatatgaaagac
cctatatcacacagcaacgt
tccacttcaggactagatga
accatggagaggactctgag
ggtctttttttcaagccaatc
gggcacaacacttttataca
tttccatcatagtgctggtg
atctcatgtgtagagttgcg
cgtggtggacacacttttta
attcacatcctcctcatatt
cttcatgttgggagcttcag
tctcaatgcggttaccaatg

Continued on next page

Table B.5 – FISH probes set 7

Probes sequence
gtcaaagaaccacgtttccg
atgtcctatatgggtggttc
ctcatagcttccatggtaag
tattagagaggacgctgacc
ttttgacaggagcctgacaa
tcatggctattcctgtgact

APPENDIX C

Appendix: Analysis macros, scripts, and functions

This appendix is also available online.
Repositories can be found at: http://github.com/jucapitanio

C.1 ImageJ macro

C.1.1 ImageJ macro for FISH assay analysis

For 3 channel images:

```
/*
We start by creating a function that can perform all the steps
we need to create the necessary files.
This function will get a .dv file, and output a mask for the
entire cell (based on dapi + smFISH staining). Saved as .tif.
The channels are also saved separately in .tif files.
*/

function cell_images(input, output, filename) {
  /*
  Below we will use the Bio-Formats importer plugin to open the images.
  It works best if you create one variable with the location of file to
  open (newfile) and another with the options for the importer (options).
  */
  newfile = input + filename;
  options = "open=newfile autoscale color_mode=Default split_channels
  view=[Standard ImageJ] stack_order=XYZCT";
  run("Bio-Formats Importer",options);

  /*
  For each channel we save a .tif file separately.
  Then we create a Z projection of intensity sum to use for the mask
  and close the original stack file.
  */
  selectWindow(filename + " - C=0");
  saveAs("Tiff", output + "cy5_" + filename + ".tif");
  run("Z Project...", "projection=[Sum Slices]");
  selectWindow("cy5_" + filename + ".tif");
  close();
  selectWindow(filename + " - C=1");
  saveAs("Tiff", output + "cy3_" + filename + ".tif");
  run("Z Project...", "projection=[Sum Slices]");
  selectWindow("cy3_" + filename + ".tif");
  close();
  selectWindow(filename + " - C=2");
  saveAs("Tiff", output + "dapi_" + filename + ".tif");
  run("Z Project...", "projection=[Sum Slices]");
  selectWindow("dapi_" + filename + ".tif");
  close();

  /*
  We join all the projections of different channels into one stack,
  blur them and create masks.
  The masks are again Z-projected as max intensity to create a single
  mask encompassing the entire cell.
  This mask is also save as a .tif file.
  */

  run("Images to Stack", "name=Stack title=[]");
  run("Gaussian Blur...", "sigma=5 stack");
  setOption("BlackBackground", true);
  run("Convert to Mask", "method=Default background=Default calculate black");
  run("Z Project...", "projection=[Max Intensity]");
  selectWindow("Stack");
  close();
  selectWindow("MAX_Stack");
  run("Convert to Mask");
  saveAs("Tiff", output + "Mask_" + filename + ".tif");
```

```
  close();

}

/* Below are the folder were your .dv files are (input)
and the folder where you'd like to save the files (output).*/
input = getDirectory("Input directory");
output = getDirectory("Output directory");

/* This will iterate the function created above in
all the files present in input.*/
setBatchMode(true);
list = getFileList(input);
for (i = 0; i < list.length; i++)
        cell_images(input, output, list[i]);
setBatchMode(false);
```

For 2 channel images:

```
/*
We start by creating a function that can perform all the steps we need to
create the necessary files.
This function will get a .dv file, and output a mask for the entire cell
(based on dapi + smFISH staining). Saved as .tif.
The channels are also saved separately in .tif files.
*/

function cell_images(input, output, filename) {
  /*
  Below we will use the Bio-Formats importer plugin to open the images.
  It works best if you create one variable with the location of file to open
  (newfile) and another with the options for the importer (options).
  */
  newfile = input + filename;
  options = "open=newfile autoscale color_mode=Default split_channels
  view=[Standard ImageJ] stack_order=XYZCT";
  run("Bio-Formats Importer",options);

  /*
  For each channel we save a .tif file separately.
  Then we create a Z projection of intensity sum to use for the mask and close
  the original stack file.
  */
  selectWindow(filename + " - C=0");
  saveAs("Tiff", output + "cy5_" + filename + ".tif");
  run("Z Project...", "projection=[Sum Slices]");
  selectWindow("cy5_" + filename + ".tif");
  close();

  selectWindow(filename + " - C=1");
  saveAs("Tiff", output + "dapi_" + filename + ".tif");
  run("Z Project...", "projection=[Sum Slices]");
  selectWindow("dapi_" + filename + ".tif");
  close();

  /*
  We join all the projections of different channels into one stack, blur them
  and create masks.
  The masks are again Z-projected as max intensity to create a single mask
  encompassing the entire cell.
  This mask is also save as a .tif file.
  */

  run("Images to Stack", "name=Stack title=[]");
  run("Gaussian Blur...", "sigma=5 stack");
  setOption("BlackBackground", true);
  run("Convert to Mask", "method=Default background=Default calculate black");
```

```
  run("Z Project...", "projection=[Max Intensity]");
  selectWindow("Stack");
  close();
  selectWindow("MAX_Stack");
  run("Convert to Mask");
  saveAs("Tiff", output + "Mask_" + filename + ".tif");
  close();

}

// Below are the folder were your .dv files are (input) and the folder where
//you'd like to save the files (output).
input = getDirectory("Input directory");
output = getDirectory("Output directory");

// This will iterate the function above in all the files present in input.
setBatchMode(true);
list = getFileList(input);
for (i = 0; i < list.length; i++)
        cell_images(input, output, list[i]);
setBatchMode(false);
```

C.1.2 ImageJ macro for bead halo assay analysis

```
function beadhalo(input, output, filename) {
        open(input + filename);
        selectWindow(filename + "  Ch0");
                close();
                selectWindow(filename + "  Ch1");
                run("Brightness/Contrast...");
                resetMinAndMax();
                setMinAndMax(0, 2000);
                call("ij.ImagePlus.setDefault16bitRange", 0);
                run("Subtract Background...", "rolling=50");
                setAutoThreshold("Default dark");
                //run("Threshold...");
                resetThreshold();
                setThreshold(150, 2000);
                run("Convert to Mask");
                run("Close");
                run("Create Selection");
                run("Add to Manager");
                roiManager("Add");
                roiManager("Select", 0);
                roiManager("Split");
                roiManager("Select", 0);
                roiManager("Delete");
                selectWindow(filename + "  Ch1");
                close();
                open(input + filename);
                selectWindow(filename + "  Ch0");
                close();
                selectWindow(filename + "  Ch1");
                resetMinAndMax();
                setMinAndMax(0, 2000);
                call("ij.ImagePlus.setDefault16bitRange", 0);
                run("Subtract Background...", "rolling=50");
                roiManager("Deselect");
                roiManager("Show all");
                roiManager("Measure");
                saveAs("Results", output + filename + ".xml");
                close();
                roiManager("Delete");
                run("Clear Results");
```

```
}

input = getDirectory("Input directory");
output = getDirectory("Output directory");

setBatchMode(true);
list = getFileList(input);
for (i = 0; i < list.length; i++)
        beadhalo(input, output, list[i]);
setBatchMode(false);
```

C.1.3 ImageJ macro for focus forming units quantification

```
/* Start with a Operetta images folder containing subfolders for each well.
Remember to use the matlab script first to organize all
the wells in the correct directories*/

inputdir = getDirectory("Input directory with well folders");

Dialog.create("File type");
Dialog.addString("File suffix: ", "-1.tif", 5);
Dialog.show();
suffix = Dialog.getString();

processdirectories(inputdir)

function processdirectories(inputdir) {
        list = getFileList(inputdir);
        for (i = 0; i < list.length; i++) {
                if(File.isDirectory(inputdir + list[i]))
                        input = inputdir + list[i];
                        output = inputdir + list[i];
                        processFolder(input);

        }
}

function processFolder(input) {
        list = getFileList(input);
        for (i = 0; i < list.length; i++) {
                if(endsWith(list[i], suffix))
                        processFile(input, output, list[i]);
        }
}

function processFile(input, output, file) {

        newfile = input + file;
        options = "open=newfile autoscale color_mode=Default group_files
        split_channels view=Hyperstack stack_order=XYCZT swap_dimensions
        stitch_tiles dimensions axis_1_number_of_images=13
        axis_1_axis_first_image=1 axis_1_axis_increment=1 contains=[] name=file
        z_1=13 c_1=2 t_1=1";

        run("Bio-Formats Importer",options);

        imageslist = getList("image.titles");
        selectWindow(imageslist[0]);
        close();
        selectWindow(imageslist[1]);
        resetMinAndMax();
        selectWindow(imageslist[1]);
        run("Add Slice");
        run("Add Slice");
        run("Make Substack...", "  slices=2,4,5,8,7,6,9,1,10,13,12,11,3,14,15");
        selectWindow(imageslist[1]);
```

```
close();

selectWindow("Substack (2,4,5,8,7, ... 15)");
run("Make Montage...", "columns=3 rows=5 scale=0.5 last=15");
selectWindow("Substack (2,4,5,8,7, ... 15)");
close();
selectWindow("Montage");
run("8-bit");
saveAs("Tiff", output + file + "montage.tif");
selectWindow(file + "montage.tif");
rename("Montage");
resetMinAndMax();
selectWindow("Montage");
run("Subtract Background...", "rolling=50");
selectWindow("Montage");
run("Out [-]");
selectWindow("Montage");
run("Despeckle");
selectWindow("Montage");
run("Gaussian Blur...", "sigma=5");
selectWindow("Montage");
setAutoThreshold("IJ_IsoData dark");
run("Convert to Mask");
selectWindow("Montage");
run("Find Connected Regions", "allow_diagonal display_one_image
display_results regions_for_values_over=100 minimum_number_of_points=1000
stop_after=-1");
selectWindow("Montage");
saveAs("Tiff", output + file + "montage-mask.tif");
close();
selectWindow("All connected regions");
run("RGB Color");
saveAs("Tiff", output + file + "connected-regions.tif");
close();
resultsfile = output + file + "Results.xls";
selectWindow("Results");
//saveAs("Results", output + file + "Results.xls");
run("Text...", "save=[resultsfile]");
}
```

C.2 MATLAB Scripts

C.2.1 MATLAB scripts for FISH assay analysis

Analysis script 1, for 3 channel images:

```
%% Create all necessary folders to run the analysis and add them to path.
% Go into the images directory where you want to store your analysis and
% then run the section below to create all folders.

AnalysisDate = 'datehere';
mkdir(AnalysisDate);
cd (AnalysisDate);
mkdir('ImagesOriginal');
mkdir('rootCell');
cd('rootCell');
mkdir('ImageData');
mkdir('SegmentationMasks');
mkdir('AnalysisJu');
mkdir('cell masks');
mkdir('Cell plot images');
mkdir('SpotsData');
cd('ImageData');
mkdir('dapi');
```

```
mkdir('cy3');
mkdir('cy5');
cd ..\..\..\;

addpath(genpath(pwd));

clear AnalysisDate;

%% Rename the files according to the script requirements
raw_files = dir(fullfile(strcat(pwd,'\ImagesOriginal')));
raw_files = raw_files(3:end);
rename_files = cell(length(raw_files),2);
for i = 1:length(raw_files)
    rename_files(i,1) = cellstr(raw_files(i).name);
    rename_files(i,2) = cellstr(strcat('Pos', num2str(i)));
end;
rename_files;

%% Double check the correct name structure was printed in the command window
% before actually changing the names.
% The name conversion is saved in the renaming_table.mat file.
save('renaming_table.mat','rename_files');
cd('ImagesOriginal');

for i = 1:length(rename_files)
    movefile(char(rename_files(i,1)),char(rename_files(i,2)));
end;

clear raw_files rename_files;

%% Start with ImageData folder with subfolder for each channel containing
images
% for analysis.
% Prepare the images in imageJ using the macro for vRNA FISH.
% Create the position identifier Mask files with the function below. You
% must be in the rootCell directory and the mask files in the SegmentationMasks
% folder.
rootfolder = pwd;
numimg = size(dir(strcat(rootfolder, '\cell masks')),1) - 2;
date = '20160712';

parfor i = 1:numimg
    createSegmenttrans(strcat('Pos',num2str(i)));
end;

createSegImages('tif');

doEvalFISHStacksForAll
%% To create this training set I used the images Pos10, 20, 30, 40, 49, 59.
%Repeating the command below.
% This includes all the different experimental conditions (drugs).
createSpotTrainingSet('cy3_Pos10','ZikaposCy3')
%% To create this training set I used the images Pos10, 20, 30, 40, 49, 59.
% Repeating the command below.
% This includes all the different experimental conditions (drugs).
createSpotTrainingSet('cy5_Pos12','ZikaposCy5')
%%
load trainingSet_cy3_ZikaposCy3.mat
trainingSet=trainRFClassifier(trainingSet);
load trainingSet_cy5_ZikaposCy5.mat
trainingSet=trainRFClassifier(trainingSet);
load trainingSet_cy3_ZikaposCy3.mat
classifySpotsOnDirectory(1,trainingSet,'cy3')
load trainingSet_cy5_ZikaposCy5.mat
classifySpotsOnDirectory(1,trainingSet,'cy5')
%% run the function reviewFISHClassification('dye_PosX') in the 2012a
%version of MATLAB.
% Use an image different from the one used to create the training set and
```

```
%repeat the process for all channels.
% I used images Pos1, 11, 21, 31, 41 and 50 to cover all experimental
% conditions (drugs).
% Don't forget to change line 168 of this function with the full path to
% the Aro_parameters file.
% Don't forget to add the new points to the training set. After each round
% of correction save and repeat the following commands:
load trainingSet_cy3_ZikaposCy3.mat
trainingSet=trainRFClassifier(trainingSet);
classifySpotsOnDirectory(1,trainingSet,'cy3')
%%
load trainingSet_cy5_ZikaposCy5.mat
trainingSet=trainRFClassifier(trainingSet);
classifySpotsOnDirectory(1,trainingSet,'cy5')
% You'll have to do it manually and in the 2015 version of MATLAB. After
% retraining and detecting the spots repeat the reviewFISHClassification
% for new images (choose a different experimental condition). Repeat at
% least 3 or 4 times until you don't see much error in the classifications.

%% You can collect the data on all spots using the following function.
% This is not that useful, but it's good to check classification error.
% It all ends saved in the Plots folder under AnalysisJu.
rootfolder = pwd;

cy3Spotstats = strcat(rootfolder, '\AnalysisJu\SpotStats\cy3');
cy5Spotstats = strcat(rootfolder, '\AnalysisJu\SpotStats\cy5');

cd(cy3Spotstats);
spotStatsDataAligning([date 'cy3'], 0);

cd(cy5Spotstats);
spotStatsDataAligning([date 'cy5'], 0);

cd(rootfolder);

movefile(strcat(rootfolder, '\AnalysisJu\SpotStats\cy3\wormData_', date,
'cy3.mat'), strcat(rootfolder, '\AnalysisJu\Plots\wormData_', date, 'cy3.mat'));
movefile(strcat(rootfolder, '\AnalysisJu\SpotStats\cy5\wormData_', date,
'cy5.mat'), strcat(rootfolder, '\AnalysisJu\Plots\wormData_', date, 'cy5.mat'));
movefile(strcat(rootfolder, '\AnalysisJu\SpotStats\cy3\ErrorPercentagePlot_',
date, 'cy3.fig'), strcat(rootfolder, '\AnalysisJu\Plots\ErrorPercentagePlot_',
date, 'cy3.fig'));
movefile(strcat(rootfolder, '\AnalysisJu\SpotStats\cy5\ErrorPercentagePlot_',
date, 'cy5.fig'), strcat(rootfolder, '\AnalysisJu\Plots\ErrorPercentagePlot_',
date, 'cy5.fig'));

%% FROM HERE ON, USE THE NUMBER 2 SCRIPT to process everything.
```

Analysis script 1, for 2 channel images:

```
%% Create all necessary folders to run the analysis and add them to path.
% Go into the images directory where you want to store your analysis and
% then run the section below to create all folders.

AnalysisDate = 'add date here';
%AnalysisDate = '20150716';
mkdir(AnalysisDate);
cd (AnalysisDate);
mkdir('ImagesOriginal');
mkdir('rootCell');
cd('rootCell');
mkdir('ImageData');
mkdir('SegmentationMasks');
mkdir('AnalysisJu');
mkdir('cell masks');
mkdir('Cell plot images');
```

```
mkdir('SpotsData');
cd('ImageData');
mkdir('dapi');
mkdir('cy3');
cd ..\..\;

clear AnalysisDate;
%% Start with ImageData folder with subfolder for each channel containing
% images for analysis.
% Create the position identifier Mask files with the function below. You
% must be in the root directory and the mask files in the SegmentationMasks
% folder.

rootfolder = pwd;
numimg = length(dir(strcat(rootfolder, '\cell masks'))) - 2;
date = '20150925';

parfor i = 1:numimg
    createSegmenttrans(strcat('Pos',num2str(i)));
end;
%%
createSegImages('tif');
%%
doEvalFISHStacksForAll
%% To create this training set I used the images Pos10, 20, 30, 40, 49, 59.
% Repeating the command below.
% This includes all the different experimental conditions (drugs).
createSpotTrainingSet('cy3_Pos11','negvRNA')

%%
load trainingSet_cy3_negvRNA.mat
trainingSet=trainRFClassifier(trainingSet);

%%
load trainingSet_cy3_negvRNA.mat
classifySpotsOnDirectory(1,trainingSet,'cy3')

%% run the function reviewFISHClassification('dye_PosX') in the 2012a version
% of MATLAB.
% Use an image different from the one used to create the training set and
% repeat the process for all channels.
% I used images Pos1, 11, 21, 31, 41 and 50 to cover all experimental
% conditions (drugs).
% Don't forget to change line 168 of this function with the full path to
% the Aro_parameters file.
% Don't forget to add the new points to the training set. After each round
% of correction save and repeat the following commands:
load trainingSet_cy3_negvRNA.mat
trainingSet=trainRFClassifier(trainingSet);
classifySpotsOnDirectory(1,trainingSet,'cy3')
% You'll have to do it manually and in the 2015 version of MATLAB. After
% retraining and detecting the spots repeat the reviewFISHClassification
% for new images (choose a different experimental condition). Repeat at
% least 3 or 4 times until you don't see much error in the classifications.

%% You can collect the data on all spots using the following function.
% This is not that useful, but it's good to check classification error.
% It all ends saved in the Plots folder under AnalysisJu.

cy3Spotstats = strcat(rootfolder, '\AnalysisJu\SpotStats\cy3');

cd(cy3Spotstats);
spotStatsDataAligning([date 'cy3'], 0);

cd(rootfolder);

movefile(strcat(rootfolder, '\AnalysisJu\SpotStats\cy3\wormData_', date,
'cy3.mat'), strcat(rootfolder, '\AnalysisJu\Plots\wormData_', date, 'cy3.mat'));
```

```
movefile(strcat(rootfolder, '\AnalysisJu\SpotStats\cy3\ErrorPercentagePlot_',
date, 'cy3.fig'), strcat(rootfolder, '\AnalysisJu\Plots\ErrorPercentagePlot_',
date, 'cy3.fig'));

%% FROM HERE ON, USE THE NUMBER 2 SCRIPT to process everything.
```

Analysis script 2, for 3 channel images:

```
%% Add all necessary folder paths to functions and data:
%% Looping all the functions I made to analyze all images and save data

%parpool;
rootfolder = pwd;
numimg = size(dir(strcat(rootfolder, '\cell masks')),1) - 2;

cy3_spotStats_files = dir(fullfile(strcat(rootfolder,
'\AnalysisJu\SpotStats\cy3'), 'cy3_Pos*_spotStats.mat'));
cy5_spotStats_files = dir(fullfile(strcat(rootfolder,
'\AnalysisJu\SpotStats\cy5'), 'cy5_Pos*_spotStats.mat'));
dapisegstack = dir(fullfile(strcat(rootfolder,'\AnalysisJu\SegStacks\dapi'),
 'dapi_Pos*_SegStacks.mat'));
cellsegmask = dir(fullfile(strcat(rootfolder,'\SegmentationMasks'),
 'segmenttrans_Pos*.mat'));

cy3countstt = struct('nuclear',{}, 'cyto',{}, 'total',{}, 'per100nuc',{},
 'Pos', {});
cy5countstt = struct('nuclear',{}, 'cyto',{}, 'total',{}, 'per100nuc',{},
 'Pos', {});
cy3midcountstt = struct('nuclear',{}, 'cyto',{}, 'total',{}, 'per100nuc',{},
 'Pos', {});
cy5midcountstt = struct('nuclear',{}, 'cyto',{}, 'total',{}, 'per100nuc',{},
 'Pos', {});

cellarea = struct('cellareaguess',{}, 'Pos', {});


% I cannot do clear or save inside a parfor loop, so I had to make it a for
% loop again... See if therre is another way to use parallel computing, or
% if you have parfor loops inside all functions possible.

for i = 1:numimg

    segmenttrans_maskfile = cellsegmask(i).name;
    cell_area_guess = guestimatecellarea(segmenttrans_maskfile,
        segmenttrans_maskfile(14:18)); % more than 99 images, change to (14:19).
    cellarea = [cellarea, cell_area_guess];
    clear cell_area_guess segmenttrans_maskfile;

    cy3_spotStats_file = cy3_spotStats_files(i).name;
    cy5_spotStats_file = cy5_spotStats_files(i).name;
    [ locs3 ] = goodspots( cy3_spotStats_file );
    [ locs5 ] = goodspots( cy5_spotStats_file );
    [ colocalizedcy3, colocalizedcy5 ] = colocalized( locs3, locs5, 10 );
    clear locs3 locs5;

    if not(isempty(colocalizedcy3)) && not(isempty(colocalizedcy5))

        dapisegstackfile = dapisegstack(i).name;
        [ dapiiso, Vnorm, stackmid ] = DAPIisosurface2( dapisegstackfile );
        clear dapisegstackfile;

        [ coloccy3dapi ] = Spot2NEdist( dapiiso, colocalizedcy3 );
        [ coloccy5dapi ] = Spot2NEdist( dapiiso, colocalizedcy5 );
        clear colocalizedcy3 colocalizedcy5;

        [ cy3counts ] = countsum( coloccy3dapi, cy3_spotStats_file(5:10) );
```

```
        [ cy5counts ] = countsum( coloccy5dapi, cy5_spotStats_file(5:10) );

        cy3countstt = [cy3countstt, cy3counts];
        cy5countstt = [cy5countstt, cy5counts];
        clear cy3counts cy5counts;

        [ cy3midcoloc ] = stacksubset(coloccy3dapi, stackmid - 5, stackmid + 5);
        [ cy5midcoloc ] = stacksubset(coloccy5dapi, stackmid - 5, stackmid + 5);

        [ cy3countsmid ] = countsum(cy3midcoloc, cy3_spotStats_file(5:10) );
        [ cy5countsmid ] = countsum(cy5midcoloc, cy5_spotStats_file(5:10) );

        cy3midcountstt = [cy3midcountstt, cy3countsmid];
        cy5midcountstt = [cy5midcountstt, cy5countsmid];
        clear cy3countsmid cy5countsmid;

        figspanel( dapiiso, Vnorm, coloccy3dapi, coloccy5dapi, cy3midcoloc,
                cy5midcoloc, cy3_spotStats_file);

        clear cy5_spotStats_file cy3midcoloc cy5midcoloc stackmid;
        save(strcat(rootfolder,'\SpotsData\SpotsIsosurf',
                cy3_spotStats_file(5:10),'.mat'));
        clear  cy3_spotStats_file coloccy3dapi coloccy5dapi Vnorm dapiiso;

    end;
end;
figure('Visible', 'on', 'name','toClose');
close 'toClose';
clear cy3_spotStats_files cy5_spotStats_files dapisegstack cellsegmask
rootfolder numimg;
save('AnalysisSummary.mat');

%% If you want you can also export the files to csv.

struct2csv(cy3countstt, 'cy3countstt.csv')
struct2csv(cy3midcountstt, 'cy3midcountstt.csv')
struct2csv(cy5midcountstt, 'cy5midcountstt.csv')
struct2csv(cy5countstt, 'cy5countstt.csv')
struct2csv(cellarea, 'cellareaguess.csv')
```

Analysis script 2, for 2 channel images:

```
%% Add all necessary folder paths to functions and data:
%% Looping all the functions I made to analyze all images and save data

%parpool;
rootfolder = pwd;
numimg = size(dir(strcat(rootfolder, '\cell masks')),1) - 2;

cy3_spotStats_files = dir(fullfile(strcat(rootfolder,
'\AnalysisJu\SpotStats\cy3'), 'cy3_Pos*_spotStats.mat'));
dapisegstack = dir(fullfile(strcat(rootfolder,'\AnalysisJu\SegStacks\dapi'),
 'dapi_Pos*_SegStacks.mat'));
cellsegmask = dir(fullfile(strcat(rootfolder,'\SegmentationMasks'),
 'segmenttrans_Pos*.mat'));

cy3countstt = struct('nuclear',{}, 'cyto',{}, 'total',{}, 'per100nuc',{},
 'Pos', {});
cy3midcountstt = struct('nuclear',{}, 'cyto',{}, 'total',{}, 'per100nuc',{},
 'Pos', {});

cellarea = struct('cellareaguess',{}, 'Pos', {});

% I cannot do clear or save inside a parfor loop, so I had to make it a for
% loop again... See if therre is another way to use parallel computing, or
% if you have parfor loops inside all functions possible.
```

```
for i = 1:numimg

    segmenttrans_maskfile = cellsegmask(i).name;
    cell_area_guess = guestimatecellarea(segmenttrans_maskfile,
        segmenttrans_maskfile(14:18)); % if more than 99 images, change to (14:19).
    cellarea = [cellarea, cell_area_guess];
    clear cell_area_guess segmenttrans_maskfile;

    cy3_spotStats_file = cy3_spotStats_files(i).name;
    [ locs3 ] = goodspots( cy3_spotStats_file );


    if not(isempty(locs3))

        dapisegstackfile = dapisegstack(i).name;
        [ dapiiso, Vnorm, stackmid ] = DAPIisosurface2( dapisegstackfile );
        clear dapisegstackfile;

        [ cy3dapi ] = Spot2NEdist( dapiiso, locs3 );
        clear locs3;

        dist3 = struct('Pos', cy3_spotStats_file(5:10), 'Distance',
                coloccy3dapi(:,4));
        spotNEdistCy3 = [spotNEdistCy3, dist3];
        clear dist3

        [ cy3counts ] = countsum( cy3dapi, cy3_spotStats_file(5:10) );
        cy3countstt = [cy3countstt, cy3counts];
        clear cy3counts;

        [ cy3mid ] = stacksubset( cy3dapi, stackmid - 5, stackmid + 5 );

        [ cy3countsmid ] = countsum( cy3mid, cy3_spotStats_file(5:10) );

        cy3midcountstt = [cy3midcountstt, cy3countsmid];

        clear cy3countsmid;

        figspanel1dye( dapiiso, Vnorm, cy3dapi, cy3mid, cy3_spotStats_file);

        clear cy3mid stackmid;
        save(strcat(rootfolder,'\SpotsData\SpotsIsosurf',
                cy3_spotStats_file(5:10),'.mat'));
        clear  cy3_spotStats_file cy3dapi Vnorm dapiiso;

    end;
end;
figure('Visible', 'on', 'name','toClose');
close 'toClose';
clear cy3_spotStats_files dapisegstack cellsegmask rootfolder numimg;
save('AnalysisSummary.mat');

%% If you want you can also export the files to csv.

struct2csv(cy3countstt, 'cy3countstt.csv')
struct2csv(cy3midcountstt, 'cy3midcountstt.csv')
struct2csv(cellarea, 'cellareaguess.csv')
```

C.2.2 MATLAB scripts for focus forming units quantification

```
%Import sourcefilename and WellName columns from the
%ImageIndex.ColumbusIDX.txt file for the plate to be analyzed as a table
%with 1st row as column names.
```

```
Imagetable = unique(ImageIndex, 'rows');
newdirecs = unique(Imagetable.WellName);


for i = 1:length(newdirecs)
    mkdir(newdirecs{i,1})
end;

for i = 1:length(Imagetable.sourcefilename)
    movefile(char(Imagetable.sourcefilename(i)),char(Imagetable.WellName(i)))
end;
```

C.2.3 MATLAB functions

1.
```
function [ dapifillgaus, Vnorms, midstack ] = DAPIisosurface2( dapisegstackfile )
%DAPIisosurface Create an isosurface from the DAPI greyscale image.
% I had to include a step resize the image up and down or I'd run out of
% memory to create the isosurface. This looses a bit of definition though.
% Oh crap! Now I think this increases the size of the extra pixels in x y
% by 10 and not 5? I fixed it after reincreasing the image size.

dapi = open(dapisegstackfile);
dapi = dapi.segStacks{1, 1};
dapi = imresize(dapi,0.5);
dapi = PreProcessImages(dapi);

stacksize = size(dapi,3);
level = graythresh(dapi) * 0.01;

dapimaskfill = [];
parfor i = 1:stacksize
    dapimaskfill(:,:,i) = segmentNuclei(dapi(:,:,i),level);
end;
dapimaskfill = imresize(dapimaskfill,2);
dapimaskfill = dapimaskfill(1+5:end-5,1+5:end-5,:);

dapimaskfill=gaussianfilter3(dapimaskfill,1.5);

isoVal = isovaluetest(dapimaskfill) * 0.5;
dapifillgaus = isosurface(dapimaskfill,isoVal);

Vnorms = isonormals(dapimaskfill,dapifillgaus.vertices);

dapifillgaus.vertices(:,3) = (dapifillgaus.vertices(:,3)* 2.213736) - 2.213726;

midstack = round(stacksize/2);
end
```

2.
```
function [ dapifillgaus, Vnorms, midstack ] = DAPIisosurface(dapisegstackfile)
%DAPIisosurface Create an isosurface from the DAPI greyscale image.

dapi = open(dapisegstackfile);
dapiimg = dapi.segStacks{1, 1};

dapiPP = PreProcessImages(dapiimg);

stacksize = size(dapiPP,3);
level = graythresh(dapiPP);

dapimaskfill = [];
parfor i = 1:stacksize
    dapimaskfill(:,:,i) = segmentNuclei(dapiPP(:,:,i),level);
end;

imgiso=gaussianfilter3(dapimaskfill,1.5);
```

```
dapifillgaus = isosurface(imgiso);

Vnorms = isonormals(imgiso,dapifillgaus.vertices);

dapifillgaus.vertices(:,3) = (dapifillgaus.vertices(:,3)* 2.213736) - 2.213726;

midstack = round(stacksize/2);
end
```

3.
```
function [ th ] = ImgTh( img, percent )
% Find the threshould above which the double image has percent pixes above
% Input:
%       img
%       percentage
dimg=double(img(:));
[counts,bins]=hist(dimg/max(dimg(:)),0:0.01:1);
cumcounts=cumsum(counts);
countfilter=cumcounts>(percent*sum(counts));
th_index=[countfilter(2:end)-countfilter(1:end-1),0];
th=bins(th_index==1)*max(dimg(:));
if isempty(th)
    th=0;
    warning('all image save value');
end

end
```

4.
```
function [ imagebig ] = PreProcessImages( image )
% preprocess the raw image
% Taken from the 3D membrane reconstruction library

wimg3 = image;
wimg3=wimg3-ImgTh(wimg3,0.8); %usually 0.8 changing to see effect
wimg3(wimg3<0)=0;
extsize=5;
wimg3_ext=zeros(size(wimg3)+2*extsize);
wimg3_ext(1+extsize:end-extsize,1+extsize:end-extsize,1+extsize:end-extsize)
=wimg3;
wimg3=wimg3_ext;
wimg3=bpass3(wimg3,1,50,1.84478);
wimg3=wimg3/max(wimg3(:));
imagebig=wimg3;


end
```

5.
```
function [ spotdapi ] = Spot2NEdist( dapiiso, spots )
%Spot2NEdist Calculate the distance between smFISH spots and the NE
%delimited by a DAPI isosurface.

spotdapi = spots;
distances = point2trimesh(dapiiso, 'QueryPoints', spots, 'Algorithm',
'parallel');
spotdapi(:,4) = distances;

end
```

6.
```
function [ gconv ] = bpass3(img,lnoise,lobject,zxr )
%bpass for 3d

img2=zeros(size(img)+2*lobject);
img2(1+lobject:end-lobject,1+lobject:end-lobject,1+lobject:end-lobject)=img;
img=img2;
clear img2;
```

```
normalize = @(x) x/sum(x(:));

if lnoise == 0
  gaussian_kernel1 = 1;
  gaussian_kernel2 = 1;
else
%   [X,Y,Z]=meshgrid(-ceil(5*lnoise):ceil(5*lnoise),-ceil(5*lnoise):ceil(
5*lnoise),...
%       -ceil(5*lnoise/zxr):ceil(5*lnoise/zxr));
  gaussian_kernel1 = normalize(exp(-((-ceil(5*lnoise):ceil(5*lnoise))/(2*lnoise
  )).^2))';
  gaussian_kernel2 = normalize(exp(-((-ceil(5*lnoise/zxr):ceil(5*lnoise/zxr))/(
  2*lnoise)).^2))';
end

if lobject
% boxcar_kernel= normalize(ones(length(-ceil(lobject):ceil(lobject)),...
%     length(-ceil(lobject):ceil(lobject)),length(-ceil(lobject/zxr):ceil(
lobject/zxr))));
    boxcar_kernel1=normalize(ones(length(-ceil(lobject):ceil(lobject)),1));
    boxcar_kernel2=normalize(ones(length(-ceil(lobject/zxr):ceil(lobject/zxr))
    ,1));
end

gconv=convn(img,gaussian_kernel1,'same');
gconv=convn(permute(gconv,[2 1 3]),gaussian_kernel1,'same');
gconv=convn(permute(gconv,[3 1 2]),gaussian_kernel2,'same');
gconv=permute(gconv,[3 2 1]);

if lobject
bconv=convn(img,boxcar_kernel1,'same');
% bconv=convn(gconv,boxcar_kernel1,'same');
bconv=convn(permute(bconv,[2 1 3]),boxcar_kernel1,'same');
bconv=convn(permute(bconv,[3 1 2]),boxcar_kernel2,'same');
bconv=permute(bconv,[3 2 1]);

gconv=gconv-bconv;
end


gconv(gconv<0)=0;
gconv=gconv(1+lobject:end-lobject,1+lobject:end-lobject,1+lobject:end-lobject);

end
```

7. ```
function [coloccy3dapi, coloccy5dapi] = cellLocspot(colocalizedcy3,
colocalizedcy5, dapimasktiff_filepath)
%cellLocspot Check if spot localizes to nuclei or cytoplasm
%   For this you need the files from the nuclei masks folder, created with
% the 'nuclei mask with Z' imageJ macro. They are 3D masks of the dapi
% nuclei stain. You also need the colocalized spots for cy3 and cy5, they
% are the outputs from function colocalized.
% If a spot is outside the nuclei the of collumn 4 in the output is 0 if
% it's inside the nuclei the value will be 255 (the value starts at 10,
% so if you find that in col 4 something is wrong).

coloccy3dapi = colocalizedcy3;
coloccy5dapi = colocalizedcy5;
coloccy3dapi(:,3) = coloccy3dapi(:,3) / 0.24 / 9.2239;
coloccy5dapi(:,3) = coloccy5dapi(:,3) / 0.24 / 9.2239;
stackmask = tiffread2(dapimasktiff_filepath);

coloccy3dapi(:,4) = ones(1,size(coloccy3dapi,1)) * 10;
coloccy5dapi(:,4) = ones(1,size(coloccy5dapi,1)) * 10;

value = 10;
for j = 1:size(coloccy3dapi,1)
    value = stackmask(uint8(coloccy3dapi(j,3))).data(coloccy3dapi(j,1),
```
```

```
    coloccy3dapi(j,2));
    coloccy3dapi(j,4) = value;
end;

value = 10;
for j = 1:size(coloccy5dapi,1)
    value = stackmask(uint8(coloccy5dapi(j,3))).data(coloccy5dapi(j,1),
    coloccy5dapi(j,2));
    coloccy5dapi(j,4) = value;
end;

coloccy3dapi(:,3) = coloccy3dapi(:,3) * 0.24 * 9.2239;
coloccy5dapi(:,3) = coloccy5dapi(:,3) * 0.24 * 9.2239;


end
```

8.
```
function [ coloccy3dapi, coloccy5dapi ] = cellLocspot2( colocalizedcy3, colocalizedcy5, dapimasktif
%cellLocspot Check if spot localizes to nuclei or cytoplasm
%   For this you need the files from the nuclei masks folder, created with
% the 'nuclei mask with Z' imageJ macro. They are 3D masks of the dapi
% nuclei stain. You also need the colocalized spots for cy3 and cy5, they
% are the outputs from function colocalized.
% If a spot is outside the nuclei the of collumn 4 in the output is 0 if
% it's inside the nuclei the value will be 255 (the value starts at 10,
% so if you find that in col 4 something is wrong).

coloccy3dapi = colocalizedcy3;
coloccy5dapi = colocalizedcy5;
coloccy3dapi(:,3) = coloccy3dapi(:,3) / 0.24 / 9.2239;
coloccy5dapi(:,3) = coloccy5dapi(:,3) / 0.24 / 9.2239;
stackmask = tiffread2(dapimasktiff_filepath);

dapi = open(dapisegstackfile);
dapiimg = dapi.segStacks{1, 1};

coloccy3dapi(:,4) = ones(1,size(coloccy3dapi,1)) * 10;
coloccy5dapi(:,4) = ones(1,size(coloccy5dapi,1)) * 10;
coloccy3dapi(:,5) = ones(1,size(coloccy3dapi,1)) * 10;
coloccy5dapi(:,5) = ones(1,size(coloccy5dapi,1)) * 10;

value = 10;
for j = 1:size(coloccy3dapi,1)
    value = stackmask(uint8(coloccy3dapi(j,3))).data(coloccy3dapi(j,1),
     coloccy3dapi(j,2));
    coloccy3dapi(j,4) = value;
end;

value = 10;
for j = 1:size(coloccy5dapi,1)
    value = stackmask(uint8(coloccy5dapi(j,3))).data(coloccy5dapi(j,1),
    coloccy5dapi(j,2));
    coloccy5dapi(j,4) = value;
end;

value = 10;
for j = 1:size(coloccy3dapi,1)
    value = dapiimg(coloccy3dapi(j,1), coloccy3dapi(j,2), uint8(coloccy3dapi(
    j,3)));
    coloccy3dapi(j,5) = value;
end;

value = 10;
for j = 1:size(coloccy5dapi,1)
    value = dapiimg(coloccy5dapi(j,1), coloccy5dapi(j,2), uint8(coloccy5dapi(
    j,3)));
    coloccy5dapi(j,5) = value;
end;
```

```
coloccy3dapi(:,3) = coloccy3dapi(:,3) * 0.24 * 9.2239;
coloccy5dapi(:,3) = coloccy5dapi(:,3) * 0.24 * 9.2239;


end
```

9.
```
function [ colocalizedcy3, colocalizedcy5 ] = colocalized( locs3, locs5, distmax )
%colocalized Finding cy3 and cy5 points withing given distance of oposite
%dye points.
%   I'll use [idx, dist] = rangesearch(X,Y,r), which returns the distances
%   between each row of Y and the rows of X that are r or less distant.
%   For this the Z scale must be corrected to pixel distance not stack
%   number. locs3 and locs5 come from the goodspots function and the
%   distmax is the maximun distance in pixels between detected spots to
%   consider them colocalized.

idx =  rangesearch(locs3, locs5, distmax);
idx = idx(~cellfun('isempty',idx)); %remove empty cells
idx = [idx{:}]; % merge all values into single row
idx = unique(idx); %remove duplicate values
colocalizedcy3 = locs3(idx, :);

idx =  rangesearch(locs5, locs3, distmax);
idx = idx(~cellfun('isempty',idx)); %remove empty cells
idx = [idx{:}]; % merge all values into single row
idx = unique(idx); %remove duplicate values
colocalizedcy5 = locs5(idx, :);


end
```

10.
```
function [ counts ] = countsum( colocdapi, Pos )
%countsummary This will collect the number of spots in each cell location
%and and create a summary in a structure for cy3 and cy5. The input files
%come from the cellLocspot function. The Pos input refers to the name of
%the file evaluated eg. Pos_59

nuclear = size(colocdapi(colocdapi(:,4) > 0,:),1);

cyto = size(colocdapi(colocdapi(:,4) < 0,:),1);

total = nuclear + cyto;

per100nuc = nuclear / total * 100;

counts = struct('nuclear',nuclear, 'cyto',cyto, 'total',total,
'per100nuc',per100nuc, 'Pos', Pos);

end
```

11.
```
function [ cy3counts, cy5counts ] = countsummary( coloccy3dapi, coloccy5dapi, Pos )
%countsummary This will collect the number of spots in each cell location
%and and create a summary in a structure for cy3 and cy5. The input files
%come from the cellLocspot function. The Pos input refers to the name of
%the file evaluated eg. Pos_59

nuclearcy3 = length(coloccy3dapi(coloccy3dapi(:,4)==255,1));
nuclearcy5 = length(coloccy5dapi(coloccy5dapi(:,4)==255,1));

cytocy3 = length(coloccy3dapi(coloccy3dapi(:,4)==0,1));
cytocy5 = length(coloccy5dapi(coloccy5dapi(:,4)==0,1));

totalcy3 = nuclearcy3 + cytocy3;
totalcy5 = nuclearcy5 + cytocy5;
```

```
cy3per100nuc = nuclearcy3 / totalcy3 * 100;
cy5per100nuc = nuclearcy5 / totalcy5 * 100;

cy3counts = struct('nuclearcy3',nuclearcy3, 'cytocy3',cytocy3,
'totalcy3',totalcy3, 'cy3per100nuc',cy3per100nuc, 'Pos', Pos);
cy5counts = struct('nuclearcy5',nuclearcy5, 'cytocy5',cytocy5,
'totalcy5',totalcy5, 'cy5per100nuc',cy5per100nuc, 'Pos', Pos);

end
```

12.
```
function createfigure(X1, Y1, Z1, S1, C1)
%CREATEFIGURE(X1, Y1, Z1, S1, C1)
%  X1:  scatter3 x
%  Y1:  scatter3 y
%  Z1:  scatter3 z
%  S1:  scatter3 s
%  C1:  scatter3 c

%  Auto-generated by MATLAB on 27-Aug-2015 21:10:22

% Create figure
figure1 = figure;
colormap('prism');

% Create axes
axes1 = axes('Parent',figure1,'Color',[0 0 0],'CLim',[3 5]);
%% Uncomment the following line to preserve the X-limits of the axes
%xlim(axes1,[0 1000]);
%% Uncomment the following line to preserve the Y-limits of the axes
%ylim(axes1,[0 1000]);
view(axes1,[90 90]);
hold(axes1,'on');

% Create scatter3
scatter3(X1,Y1,Z1,S1,C1,'MarkerFaceColor','flat','MarkerEdgeColor','none');
```

13.
```
function createfigure1(X1, Y1, X2, Y2, C3, C4)
%CREATEFIGURE1(X1, Y1, X2, Y2, C2, C3, C4)
%  X1:  Xpos cy3
%  Y1:  Ypos cy3
%  S1:  set to 49 (7x7point squared dots)
%  C1:  set to green
%  X2:  Xpos cy5
%  Y2:  Ypos cy5
%  C2:  set to red
%  S2:  set to 49 (7x7point squared dots)
%  C3:  Col indicating nuc or cyto (4th) for cy3
%  C4:  Col indicating nuc or cyto (4th) for cy5

%  Auto-generated by MATLAB on 01-Sep-2015 14:23:11

% Create figure
figure1 = figure;

% Create subplot
subplot1 = subplot(1,2,1,'Parent',figure1);
%% Uncomment the following line to preserve the X-limits of the axes
xlim(subplot1,[0 1000]);
ylim(subplot1,[0 1000]);
hold(subplot1,'on');

% Create scatter
scatter(X1,Y1,49,'green','DisplayName','cy3','Parent',subplot1,'Marker','.');

% Create scatter
scatter(X2,Y2,49,'red','DisplayName','cy5','Parent',subplot1,'Marker','.');
```

```
% Create legend
legend(subplot1,'show');

% Create subplot
subplot2 = subplot(1,2,2,'Parent',figure1);
%% Uncomment the following line to preserve the X-limits of the axes
xlim(subplot2,[0 1000]);
%% Uncomment the following line to preserve the Y-limits of the axes
ylim(subplot2,[0 1000]);
hold(subplot2,'on');

% Create scatter
scatter(X1,Y1,49,C3,'Parent',subplot2,'Marker','.');

% Create scatter
scatter(X2,Y2,49,C4,'DisplayName','data2','Parent',subplot2,'Marker','.');
```

14.
```
function createfigure2(X1, Y1, Z1, S1, C1, Vertices1, VertexNormals1, Faces1)
%CREATEFIGURE2(X1, Y1, Z1, S1, C1, VERTICES1, VERTEXNORMALS1, FACES1)
%  X1:  scatter3 x
%  Y1:  scatter3 y
%  Z1:  scatter3 z
%  S1:  scatter3 s
%  C1:  scatter3 c
%  VERTICES1:  patch vertices
%  VERTEXNORMALS1:  patch vertexnormals
%  FACES1:  patch faces

% USE THIS TO CREATE A FIG WITH NUCLEI MASK AND SPOTS

%  Auto-generated by MATLAB on 01-Dec-2015 20:02:42

% Create figure
figure1 = figure;

% Create axes
axes1 = axes('Parent',figure1,'Color',[0 0 0],...
    'ZColor',[0.501960813999176 0.501960813999176 0.501960813999176],...
    'YColor',[0.501960813999176 0.501960813999176 0.501960813999176],...
    'XColor',[0.501960813999176 0.501960813999176 0.501960813999176],...
    'DataAspectRatio',[1 1 0.45]);
view(axes1,[-70.5 40]);
box(axes1,'on');
hold(axes1,'on');

% Create scatter3
scatter3(X1,Y1,Z1,S1,C1,'MarkerEdgeColor',[1 0 0],'Marker','.');

% Create patch
patch('Parent',axes1,'FaceLighting','gouraud','FaceAlpha',0.7,...
    'Vertices',Vertices1,...
    'VertexNormals',VertexNormals1,...
    'Faces',Faces1,...
    'FaceColor',[0 0 1],...
    'EdgeColor','none');

% Create light
light('Parent',axes1,...
    'Position',[244.817067307074 -733.002902326085 1235.67614387691],...
    'Style','local');
```

15.
```
function createfigure3(X1, Y1, Z1, Vertices1, VertexNormals1, Faces1, X2, Y2,
 Z2)
%CREATEFIGURE4(X1, Y1, Z1, S1, C1, VERTICES1, VERTEXNORMALS1, FACES1, X2, Y2,
 Z2, S2, C2)
%  X1:  scatter3 x
%  Y1:  scatter3 y
```

```
% Z1: scatter3 z
% S1: scatter3 s
% C1: scatter3 c
% VERTICES1: patch vertices
% VERTEXNORMALS1: patch vertexnormals
% FACES1: patch faces
% X2: scatter3 x
% Y2: scatter3 y
% Z2: scatter3 z
% S2: scatter3 s
% C2: scatter3 c

% Auto-generated by MATLAB on 03-Dec-2015 17:42:24

% Create figure
figure1 = figure;

% Create axes
axes1 = axes('Parent',figure1,'Color',[0 0 0],...
    'ZColor',[0.501960813999176 0.501960813999176 0.501960813999176],...
    'YColor',[0.501960813999176 0.501960813999176 0.501960813999176],...
    'XColor',[0.501960813999176 0.501960813999176 0.501960813999176],...
    'DataAspectRatio',[1 1 1]);
view(axes1,[-70.5 40]);
box(axes1,'on');
hold(axes1,'on');

% Create scatter3
scatter3(X1,Y1,Z1,50,'MarkerEdgeColor',[1 0 0],'Marker','.');

% Create patch
patch('Parent',axes1,'FaceLighting','gouraud','FaceAlpha',0.5,...
    'Vertices',Vertices1,...
    'VertexNormals',VertexNormals1,...
    'Faces',Faces1,...
    'FaceColor',[0 0 1],...
    'EdgeColor','none');

% Create light
light('Parent',axes1,...
    'Position',[244.817067307074 -733.002902326085 1235.67614387691],...
    'Style','local');

% Create scatter3
scatter3(X2,Y2,Z2,50,'MarkerEdgeColor',[1 1 0],'Marker','.');
```

16.
```
function [  ] = figspanel( dapiiso, Vnorm, coloccy3dapi, coloccy5dapi,
cy3midcoloc, cy5midcoloc, cy3_spotStats_file )
%figspanel Create figure with 4 panels showing spot localization and
%hitogram distribution of distances from spots to NE isosurface.

rootfolder = pwd;
fig = figure('Visible', 'off');

% Graph 1, all spots plotted with DAPI surface. Cy3 yellow, Cy5 red.
subplot(2,2,1)
title('All posvRNA spots - Cy3 yellow, Cy5 red')
scatter3(coloccy3dapi(:,1),coloccy3dapi(:,2),coloccy3dapi(:,3)
,'MarkerEdgeColor','yellow','Marker','.');
hold on;
whitebg('black');
scatter3(coloccy5dapi(:,1),coloccy5dapi(:,2),coloccy5dapi(:,3)
,'MarkerEdgeColor','red','Marker','.');
hold on;
pd = patch(dapiiso);
Vnorm;
pd.FaceColor = 'blue';
pd.EdgeColor = 'none';
```

```
daspect([1,1,1]) % Corrigir isso no make fig.
view(3); axis tight
camlight
lighting gouraud
pd.FaceAlpha = 0.5;

% Graph 2 all spots with DAPI surface. Cytoplasmic red, Nuclear yellow.
subplot(2,2,2)
title('All posvRNA spots - Nuclear yellow, Cytoplasmic red')
scatter3(coloccy3dapi(coloccy3dapi(:,4) > 0,1),coloccy3dapi(coloccy3dapi(:,4) >
 0,2),coloccy3dapi(coloccy3dapi(:,4) > 0,3)
 ,'MarkerEdgeColor','yellow','Marker','.');
hold on;
whitebg('black');
scatter3(coloccy5dapi(coloccy5dapi(:,4) > 0,1),coloccy5dapi(coloccy5dapi(:,4) >
 0,2),coloccy5dapi(coloccy5dapi(:,4) > 0,3)
 ,'MarkerEdgeColor','yellow','Marker','.');
hold on;
scatter3(coloccy5dapi(coloccy5dapi(:,4) < 0,1),coloccy5dapi(coloccy5dapi(:,4) <
 0,2),coloccy5dapi(coloccy5dapi(:,4) < 0,3),'MarkerEdgeColor','red','Marker','.
 ');
hold on;
scatter3(coloccy3dapi(coloccy3dapi(:,4) < 0,1),coloccy3dapi(coloccy3dapi(:,4) <
 0,2),coloccy3dapi(coloccy3dapi(:,4) < 0,3),'MarkerEdgeColor','red','Marker','.
 ');
hold on;
pd = patch(dapiiso);
Vnorm;
pd.FaceColor = 'blue';
pd.EdgeColor = 'none';
daspect([1,1,1]) % Corrigir isso no make fig.
view(3); axis tight
camlight
lighting gouraud
pd.FaceAlpha = 0.5;

% Graph 3 mid spots plotted with all of DAPI surface. Cytoplasmic red, Nuclear
 yellow.
subplot(2,2,3)
title('Mid 10 stacks posvRNA spots - Nuclear yellow, Cytoplasmic red')
scatter3(cy3midcoloc(cy3midcoloc(:,4) > 0,1),cy3midcoloc(cy3midcoloc(:,4) > 0,2
),cy3midcoloc(cy3midcoloc(:,4) > 0,3),'MarkerEdgeColor','yellow','Marker','.');
hold on;
whitebg('black');
scatter3(cy5midcoloc(cy5midcoloc(:,4) > 0,1),cy5midcoloc(cy5midcoloc(:,4) > 0,2
),cy5midcoloc(cy5midcoloc(:,4) > 0,3),'MarkerEdgeColor','yellow','Marker','.');
hold on;
scatter3(cy5midcoloc(cy5midcoloc(:,4) < 0,1),cy5midcoloc(cy5midcoloc(:,4) < 0,2
),cy5midcoloc(cy5midcoloc(:,4) < 0,3),'MarkerEdgeColor','red','Marker','.');
hold on;
scatter3(cy3midcoloc(cy3midcoloc(:,4) < 0,1),cy3midcoloc(cy3midcoloc(:,4) < 0,2
),cy3midcoloc(cy3midcoloc(:,4) < 0,3),'MarkerEdgeColor','red','Marker','.');
hold on;
pd = patch(dapiiso);
Vnorm;
pd.FaceColor = 'blue';
pd.EdgeColor = 'none';
daspect([1,1,1]) % Corrigir isso no make fig.
view(3); axis tight
camlight
lighting gouraud
pd.FaceAlpha = 0.5;

% Graph 4 histogram of distances from spots to NE, all spots
subplot(2,2,4)
title('All posvRNA spots distance to NE - Cy3 yellow, Cy5 red')
hA = histfit(coloccy5dapi(:,4),30,'kernel');
hA(1).EdgeColor = 'red';
```

```
hA(1).FaceColor = 'red';
hA(2).Color = 'red';
hA(1).FaceAlpha = 0.25;
hold on;
hB = histfit(coloccy3dapi(:,4),30,'kernel');
hB(1).EdgeColor = 'yellow';
hB(1).FaceColor = 'yellow';
hB(2).Color = 'yellow';
hB(1).FaceAlpha = 0.25;

set(fig, 'Visible', 'on');
savefig(fig, strcat(rootfolder,'\Cell plot images\',cy3_spotStats_file(5:9),'B.
fig'));
set(fig, 'Visible', 'off');

end
```

17.
```
function [  ] = figspanel1dye( dapiiso, Vnorm, cy3dapi, cy3mid,
cy3_spotStats_file )
%figspanel Create figure with 4 panels showing spot localization and
%hitogram distribution of distances from spots to NE isosurface.

rootfolder = pwd;
fig = figure('Visible', 'off');

% Graph 1 all spots with DAPI surface. Cytoplasmic red, Nuclear yellow.
subplot(1,3,1)
title('All negvRNA spots - Nuclear yellow, Cytoplasmic red')
scatter3(cy3dapi(cy3dapi(:,4) > 0,1),cy3dapi(cy3dapi(:,4) > 0,2),cy3dapi(
cy3dapi(:,4) > 0,3),'MarkerEdgeColor','yellow','Marker','.');
hold on;
whitebg('black');
hold on;
scatter3(cy3dapi(cy3dapi(:,4) < 0,1),cy3dapi(cy3dapi(:,4) < 0,2),cy3dapi(
cy3dapi(:,4) < 0,3),'MarkerEdgeColor','red','Marker','.');
hold on;
pd = patch(dapiiso);
Vnorm;
pd.FaceColor = 'blue';
pd.EdgeColor = 'none';
daspect([1,1,1]) % Corrigir isso no make fig.
view(3); axis tight
camlight
lighting gouraud
pd.FaceAlpha = 0.5;

% Graph 2 mid spots plotted with all of DAPI surface. Cytoplasmic red, Nuclear
yellow.
subplot(1,3,2)
title('Mid 10 stacks negvRNA spots - Nuclear yellow, Cytoplasmic red')
scatter3(cy3mid(cy3mid(:,4) > 0,1),cy3mid(cy3mid(:,4) > 0,2),cy3mid(cy3mid(:,4)
 > 0,3),'MarkerEdgeColor','yellow','Marker','.');
hold on;
whitebg('black');
hold on;
scatter3(cy3mid(cy3mid(:,4) < 0,1),cy3mid(cy3mid(:,4) < 0,2),cy3mid(cy3mid(:,4)
 < 0,3),'MarkerEdgeColor','red','Marker','.');
hold on;
pd = patch(dapiiso);
Vnorm;
pd.FaceColor = 'blue';
pd.EdgeColor = 'none';
daspect([1,1,1]) % Corrigir isso no make fig.
view(3); axis tight
camlight
lighting gouraud
pd.FaceAlpha = 0.5;
```

```
% Graph 3 histogram of distances from spots to NE, all spots
subplot(1,3,3)
title('All negvRNA spots distance to NE')
hB = histfit(cy3dapi(:,4),30,'kernel');
hB(1).EdgeColor = 'yellow';
hB(1).FaceColor = 'yellow';
hB(2).Color = 'yellow';
hB(1).FaceAlpha = 0.25;

set(fig, 'Visible', 'on');
savefig(fig, strcat(rootfolder,'\Cell plot images\',cy3_spotStats_file(5:9),'B.
fig'));
set(fig, 'Visible', 'off');

end
```

18.
```
function [ ] = figspanelnosave( dapiiso, Vnorm, coloccy3dapi, coloccy5dapi,
cy3midcoloc, cy5midcoloc, cy3_spotStats_file )
%figspanel Create figure with 4 panels showing spot localization and
%hitogram distribution of distances from spots to NE isosurface.

rootfolder = pwd;
fig = figure('Visible', 'off');

% Graph 1, all spots plotted with DAPI surface. Cy3 yellow, Cy5 red.
subplot(2,2,1)
title('All posvRNA spots - Cy3 yellow, Cy5 red')
scatter3(coloccy3dapi(:,1),coloccy3dapi(:,2),coloccy3dapi(:,3)
,'MarkerEdgeColor','yellow','Marker','.');
hold on;
whitebg('black');
scatter3(coloccy5dapi(:,1),coloccy5dapi(:,2),coloccy5dapi(:,3)
,'MarkerEdgeColor','red','Marker','.');
hold on;
pd = patch(dapiiso);
Vnorm;
pd.FaceColor = 'blue';
pd.EdgeColor = 'none';
daspect([1,1,1]) % Corrigir isso no make fig.
view(3); axis tight
camlight
lighting gouraud
pd.FaceAlpha = 0.5;

% Graph 2 all spots with DAPI surface. Cytoplasmic red, Nuclear yellow.
subplot(2,2,2)
title('All posvRNA spots - Nuclear yellow, Cytoplasmic red')
scatter3(coloccy3dapi(coloccy3dapi(:,4) > -10,1),coloccy3dapi(coloccy3dapi(:,4)
 > -10,2),coloccy3dapi(coloccy3dapi(:,4) > -10,3)
 ,'MarkerEdgeColor','yellow','Marker','.');
hold on;
whitebg('black');
scatter3(coloccy5dapi(coloccy5dapi(:,4) > -10,1),coloccy5dapi(coloccy5dapi(:,4)
 > -10,2),coloccy5dapi(coloccy5dapi(:,4) > -10,3)
 ,'MarkerEdgeColor','yellow','Marker','.');
hold on;
scatter3(coloccy5dapi(coloccy5dapi(:,4) < -10,1),coloccy5dapi(coloccy5dapi(:,4)
 < -10,2),coloccy5dapi(coloccy5dapi(:,4) < -10,3)
 ,'MarkerEdgeColor','red','Marker','.');
hold on;
scatter3(coloccy3dapi(coloccy3dapi(:,4) < -10,1),coloccy3dapi(coloccy3dapi(:,4)
 < -10,2),coloccy3dapi(coloccy3dapi(:,4) < -10,3)
 ,'MarkerEdgeColor','red','Marker','.');
hold on;
pd = patch(dapiiso);
Vnorm;
pd.FaceColor = 'blue';
pd.EdgeColor = 'none';
```

```
daspect([1,1,1]) % Corrigir isso no make fig.
view(3); axis tight
camlight
lighting gouraud
pd.FaceAlpha = 0.5;

% Graph 3 mid spots plotted with all of DAPI surface. Cytoplasmic red, Nuclear
 yellow.
subplot(2,2,3)
title('Mid 10 stacks posvRNA spots - Nuclear yellow, Cytoplasmic red')
scatter3(cy3midcoloc(cy3midcoloc(:,4) > 0,1),cy3midcoloc(cy3midcoloc(:,4) > 0,2
),cy3midcoloc(cy3midcoloc(:,4) > 0,3),'MarkerEdgeColor','yellow','Marker','.');
hold on;
whitebg('black');
scatter3(cy5midcoloc(cy5midcoloc(:,4) > 0,1),cy5midcoloc(cy5midcoloc(:,4) > 0,2
),cy5midcoloc(cy5midcoloc(:,4) > 0,3),'MarkerEdgeColor','yellow','Marker','.');
hold on;
scatter3(cy5midcoloc(cy5midcoloc(:,4) < 0,1),cy5midcoloc(cy5midcoloc(:,4) < 0,2
),cy5midcoloc(cy5midcoloc(:,4) < 0,3),'MarkerEdgeColor','red','Marker','.');
hold on;
scatter3(cy3midcoloc(cy3midcoloc(:,4) < 0,1),cy3midcoloc(cy3midcoloc(:,4) < 0,2
),cy3midcoloc(cy3midcoloc(:,4) < 0,3),'MarkerEdgeColor','red','Marker','.');
hold on;
pd = patch(dapiiso);
Vnorm;
pd.FaceColor = 'blue';
pd.EdgeColor = 'none';
daspect([1,1,1]) % Corrigir isso no make fig.
view(3); axis tight
camlight
lighting gouraud
pd.FaceAlpha = 0.5;

% Graph 4 histogram of distances from spots to NE, all spots
subplot(2,2,4)
title('All posvRNA spots distance to NE - Cy3 yellow, Cy5 red')
hA = histfit(coloccy5dapi(:,4),30,'kernel');
hA(1).EdgeColor = 'red';
hA(1).FaceColor = 'red';
hA(2).Color = 'red';
hA(1).FaceAlpha = 0.25;
hold on;
hB = histfit(coloccy3dapi(:,4),30,'kernel');
hB(1).EdgeColor = 'yellow';
hB(1).FaceColor = 'yellow';
hB(2).Color = 'yellow';
hB(1).FaceAlpha = 0.25;

end
```

19.
```
function [ gconv ] = gaussianfilter3( img,lnoise)
%apply a 3d gaussian filter
% 3/18/2015


normalize = @(x) x/sum(x(:));

if lnoise == 0
  gaussian_kernel1 = 1;
%   gaussian_kernel2 = 1;
else
%   [X,Y,Z]=meshgrid(-ceil(5*lnoise):ceil(5*lnoise),-ceil(5*lnoise):ceil(
5*lnoise),...
%       -ceil(5*lnoise/zxr):ceil(5*lnoise/zxr));
  gaussian_kernel1 = normalize(exp(-((-ceil(5*lnoise):ceil(5*lnoise))/(2*lnoise
  )).^2))';
%   gaussian_kernel2 = normalize(exp(-((-ceil(5*lnoise/zxr):ceil(5*lnoise/zxr))
/(2*lnoise)).^2))';
```

```
end
gaussian_kernel2=gaussian_kernel1;

gconv=convn(img,gaussian_kernel1,'same');
gconv=convn(permute(gconv,[2 1 3]),gaussian_kernel1,'same');
gconv=convn(permute(gconv,[3 1 2]),gaussian_kernel2,'same'); %
gconv=permute(gconv,[3 2 1]);

end
```

20.
```
function [ locsf ] = goodspots( spotStats_file )

%goodspots Collecting the location of only selected spots to plot.
%   I collect the X Y Z coordenates for only the spots classified as true.
%   I need to move the points by 5 in all directions to mach the dapi
%   isosurface I'll create later. I also need to reverse the X,Y
%   coordinates to match the dapi isosurface.
%   I have to transform the Z axis from stack number, to micron to pixels so
%   the values match the X and Y. I know that for these images the Z stacks
%   are 0.24um apart and that 1um is 9.2239 pixels. I'm also setting the Z
%   axis to 0 instead of 1.

load(spotStats_file)
locs=spotStats{1}.locAndClass(spotStats{1}.locAndClass(:,4)==1,1:3);
locs = locs + 5; % adjust dims to compare to dapi isosurface later
locsf = [];
locsf(:,1) = locs(:,2); %Reverse x,y coords to match dapi.
locsf(:,2) = locs(:,1);
locsf(:,3) = (locs(:,3) * 2.213736) - 2.213726;

end
```

21.
```
function [ cell_area_guess ] = guestimatecellarea( segmenttrans_maskfile, Pos )
%guestimatecellarea Try to estimate the area of the image covered by cells
%   This will be so we can normalize the number of vRNA to the actual area
%   with cells in the picture, since there is no way to estimate the number
%   of cells present.
load(segmenttrans_maskfile);
cellareaguess = sum(sum(currpolys{1,1}));
cell_area_guess = struct('cellareaguess',cellareaguess, 'Pos', Pos);
end
```

22.
```
function val = isovaluetest(data)
%ISOVALUE  Isovalue calculator.
%   VAL = ISOVALUE(V) calculates an isovalue from data V using hist
%   function.  Utility function used by ISOSURFACE and ISOCAPS.

%   Copyright 1984-2012 The MathWorks, Inc.

% All I did wa rename this and save it in the folder so I can actually see
% the isovalue being used by the isosurface function.

% only use about 10000 samples
r = 1;
len = length(data(:));
if len > 20000
  r = floor(len/10000);
end

[n, x] = hist(data(1:r:end),100);

% remove large first max value
pos = find(n==max(n));
pos = pos(1);
q = max(n(1:2));
if pos<=2 && q/(sum(n)/length(n)) > 10
```

```
  n = n(3:end);
  x = x(3:end);
end

% get value of middle bar of non-small values
pos = find(n<max(n)/50);
if length(pos) < 90
  x(pos) = [];
end
val = x(floor(length(x)/2));
```

23.
```
function [imgfillmask] = segmentNuclei(im, level)
%segmentImage segments image using auto-generated code from imageSegmenter App
%  [BW,MASKEDIMAGE] = segmentImage(IM) segments image IM using auto-generated
%  code from the imageSegmenter App. The final segmentation is returned in
%  BW and a masked image is returned in MASKEDIMAGE.

% Auto-generated by imageSegmenter app on 01-Dec-2015
%----------------------------------------------------


% Initialize segmentation with given threshold
 %mask = im>0.20585;

% Initialize segmentation with Otsu's threshold
% Uncomment below to let auto threshold work.
% This is a bad idea for the stacks since the ones on the top that should
% be weak greys will get blown up and selected. If you want this calculate
% the threshold for the entire stack first, outside this function and then
% use that here as an argument above
%level = graythresh(im);
mask = im2bw(im,level);
mask = imfill(mask, 'holes');% I added this since activecontour does not
accept masks with holes. If it breaks the code, remove it.
% Evolve segmentation
% I m changing the contraction bias to zero here to see what happens. If
% bad revert to 0.5 as before.
BW = activecontour(im, mask, 100, 'Chan-Vese','SmoothFactor',1.
5,'ContractionBias',0);

% Suppress components connected to image border
BW = imclearborder(BW);

% Fill holes
BW = imfill(BW, 'holes');

% Filter components by area, uncomment if desired.
% The nuclei in a mid stack is between 8000 and 14000 or so
% Good limits may be [2500 30000]
BW = bwareafilt(BW, [1000 30000]); % I'm changing this to see if I eliminate
% background small dirt. Initial filter [50 30000].

% Form masked image from input image and segmented image.
maskedImage = im;
maskedImage(~BW) = 0;

% Fill holes in greyscale dapi
imgfillmask = imfill(maskedImage, 'holes');
end
```

24.
```
function [ spotsubset ] = stacksubset( spotpointcoord, start, stop )
%stacksubset subset point coordnate files only for given Z stacks.
start = ((start + 5) * 2.213736) - 2.213726;
stop = ((stop + 5) * 2.213736) - 2.213726;
spotsubset=spotpointcoord(spotpointcoord(:,3)>start & spotpointcoord(:,3)<
stop,:);
end
```

C.3 R Scripts for genomics data analysis

Comparative analysis of available next generation sequencing datasets

Genome-wide Nup98 interaction with chromatin was assessed through available DamID-seq data, by comparing enriched DNA sequences from Dam-Nup98 or Dam-Nup98dCTD expressing HeLa-C cells to those of Dam-GFP expressing cells (GSE83692). Data analysis was performed as described in the corresponding dataset and its publication. Transcriptome-wide interaction of Nup98 with mRNA molecules was determined from available sequencing data for Nup98 RNA immunoprecipitations from K562 cells (GSE67963) (G Hendrickson et al., 2016). DHX9 interaction with RNA was determined from sequencing data for DHX9 RNA immunoprecipitation from TC32 cells, kindly provided by Dr. Hayriye Verda Erkizan and Professor Jeffrey Alan Toretsky (Georgetown University) (Erkizan et al., 2015). Data analysis was performed as described in the corresponding datasets and their indicated publication, transcripts were considered as interacting with target proteins if showing a fold enrichment above 1.5 and adjusted p-value < 0.05. Transcriptome-wide changes in transcript or splicing isoform abundance were determined from RNA-sequencing data for HepG2 or IMR90 cells upon Nup98 depletion (GSE83551)(Franks et al., 2016). Transcriptome changes in NB1 cells upon DHX9 depletion were determined from available RNA-sequencing data (GSE44585) (Chen et al., 2014). Transcriptome sequencing data was analyzed as previously described (Wolfien et al., 2016), using Galaxy (Afgan et al., 2016), R (R Core Team, 2016) and Bioconductor (Huber et al., 2015). An adjusted p-value $<$ 0.05 was used to identify transcripts/isoforms whose abundance was significantly altered upon target protein depletion. All datasets were aligned to human reference sequence GRCh37/hg19 and annotated with corresponding UCSC genes and Ensembl genes (Huang, Loganantharaj, Schroeder, Fargo, & Li, 2013; Rosenbloom et al., 2015; Yates et al., 2016; Yu, Wang, & He, 2015). Statistically significant overlap between gene sets were calculated using the Fisher's exact test based on the hypergeometric distribution through the R package GeneOverlap (Shen, 2013).

Load all available datasets

Datasets are stored as RData files containing lists of gene symbols.

```
## [[1]]
## [1] "Allgenes"
##
## [[2]]
## [1] "CRE.genes"
##
## [[3]]
## [1] "Dam.all"
##
## [[4]]
## [1] "RIP.all"
##
## [[5]]
## [1] "Splice.all"
##
## [[6]]
## [1] "Transcript"
##
```

```
## [[7]]
## [1] "AllKD2"
```

Comparing Nup98 and DHX9 RNA-IP interacting mRNAs:

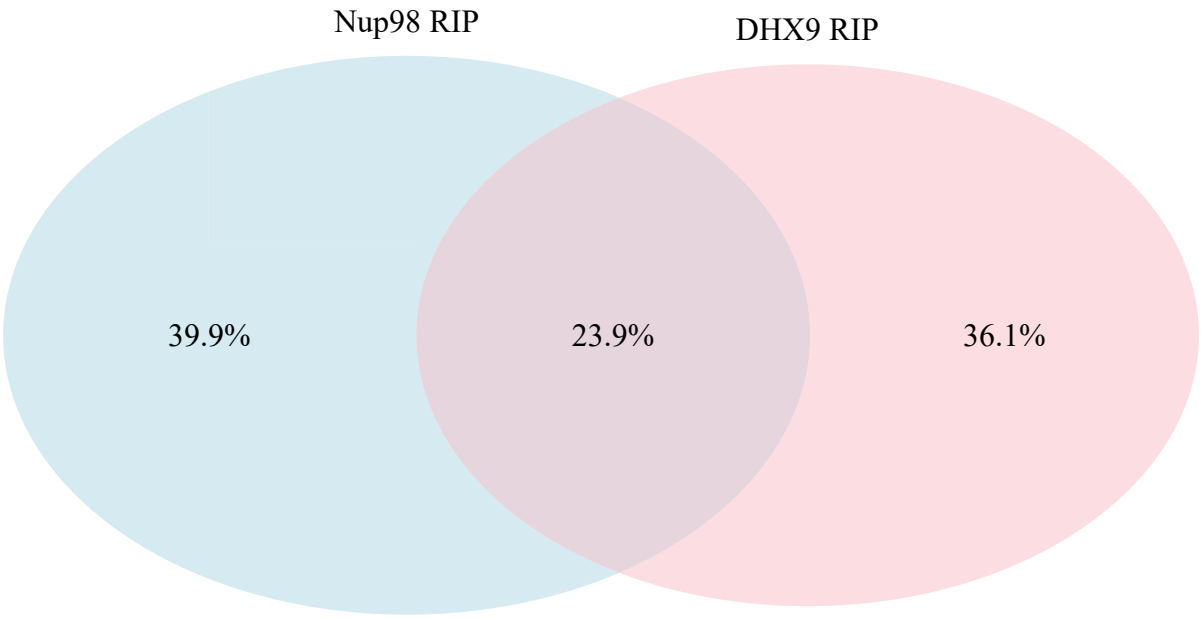
```
gs.RNASeq <- 48321 #Number of all possible genes from ENSG

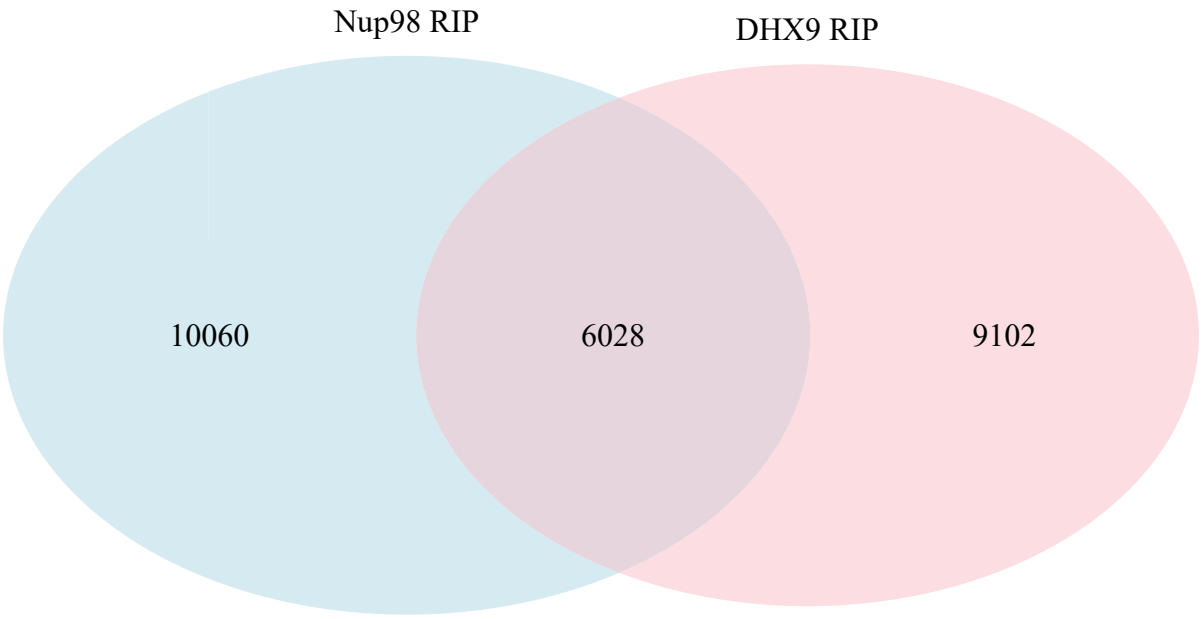
go.obj <- newGeneOverlap(RIP.all$Nup98RIP.K562,RIP.all$DHX9RIP.T32,genome.size=gs.RNASeq)
go.obj <- testGeneOverlap(go.obj)
print(go.obj)
```

```
## Detailed information about this GeneOverlap object:
## listA size=16088, e.g. STPG1 SLC7A2 CCL18
## listB size=15130, e.g. OR4G11P SAMD11 NOC2L
## Intersection size=6028, e.g. PDK4 COPZ2 PRR5
## Union size=25190, e.g. STPG1 SLC7A2 CCL18
## Genome size=48321
## # Contingency Table:
##       notA   inA
## notB 23131 10060
## inB   9102  6028
## Overlapping p-value=2.5e-93
## Odds ratio=1.5
## Overlap tested using Fisher's exact test (alternative=greater)
## Jaccard Index=0.2
```

```
## Loading required package: grid
```

```
## Loading required package: futile.logger
```


Nup98 RIP
DHX9 RIP
39.9%
23.9%
36.1%



Comparing changes in transcript splicing isoform upon DHX9 KD in NB1 cells, Nup98 KD in HepG2 cells and Nup98 KD in IMR90 cells

Table of p-values:

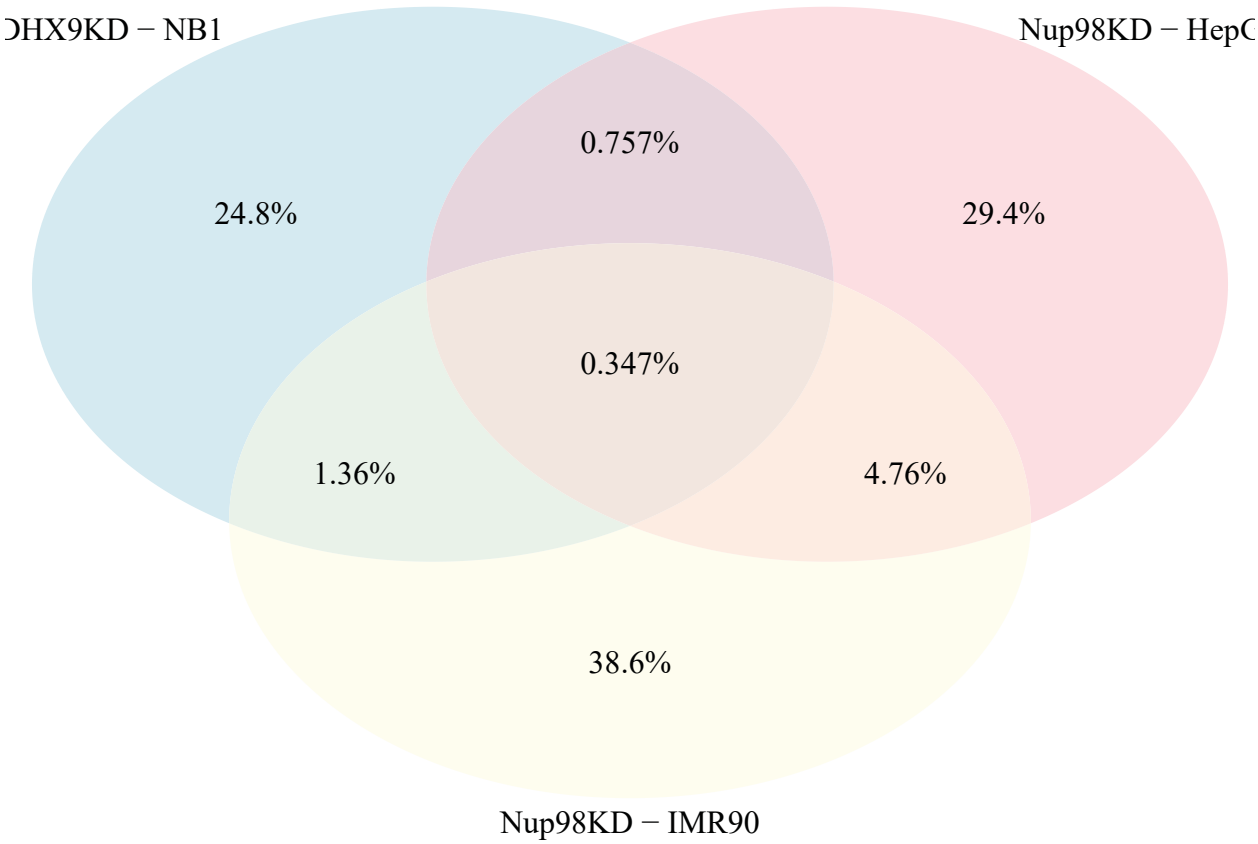
```
gs.RNASeq <- 23144 #ENSG spliced genes
names(Splice.all) <- c("Nup98KD.HepG2","Nup98KD.IMR90","DHX9KD.NB1")
gom.self <- newGOM(Splice.all, genome.size=gs.RNASeq)
getMatrix(gom.self, name="pval")
```

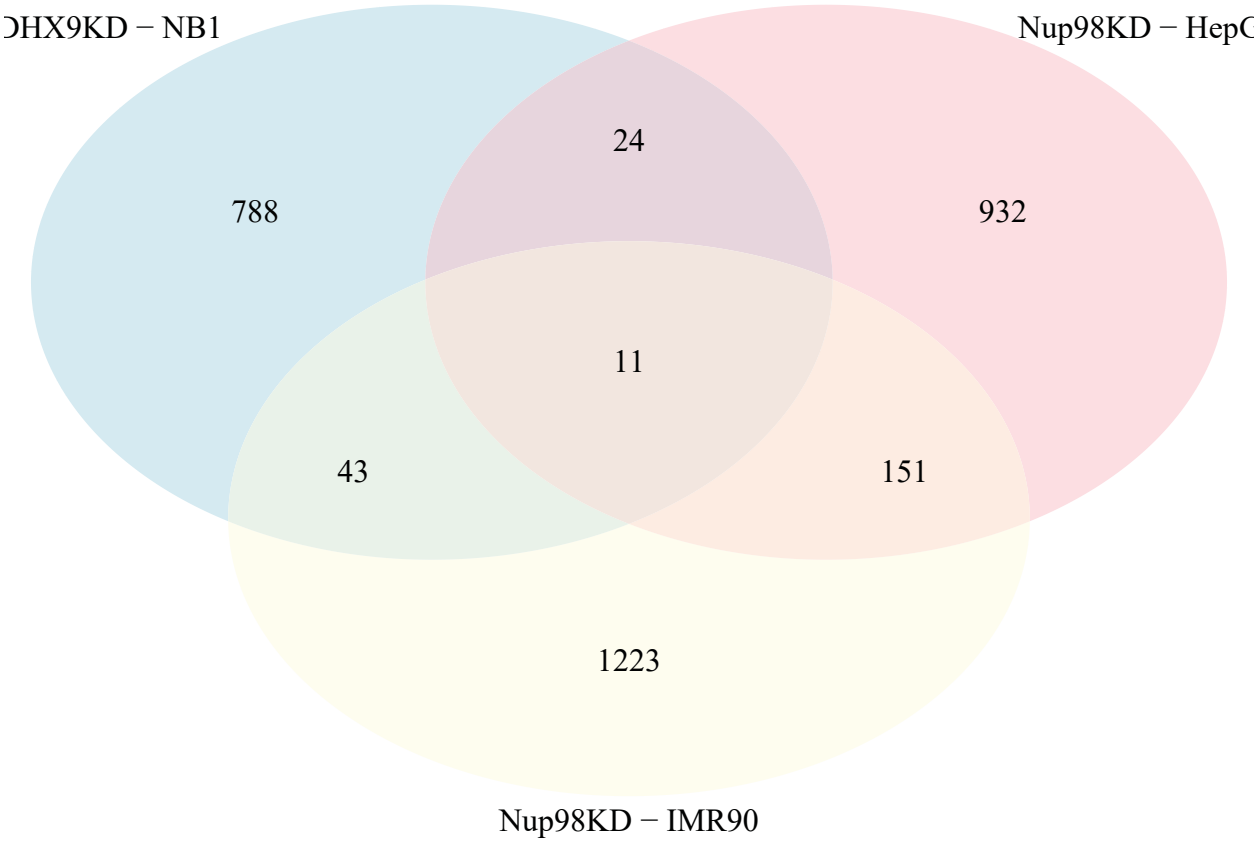
```
##                Nup98KD.IMR90   DHX9KD.NB1
## Nup98KD.HepG2 9.728468e-102 3.984383e-27
## Nup98KD.IMR90  1.000000e+00 4.956290e-24
```

Number of genes in common:

```
getMatrix(gom.self, name="intersection")
```

```
##                Nup98KD.IMR90 DHX9KD.NB1
## Nup98KD.HepG2            277        119
## Nup98KD.IMR90              0        132
```


DHX9KD − NB1
Nup98KD − HepC
0.757%
24.8%
29.4%
0.347%
1.36%
4.76%
38.6%
Nup98KD − IMR90



Compare changes in transcript level upon DHX9 KD in NB1 cells, Nup98 KD in HepG2 cells and Nup98 KD in IMR90 cells

All statistically significant genes (up or down regulated), table of p-values:

```
gs.RNASeq <- 29432 #Total number of genes detected in these transcriptomes

gom.self <- newGOM(Transcript$all, genome.size=gs.RNASeq)
getMatrix(gom.self, name="pval")
```

```
##                Nup98KD.IMR90   DHX9KD.NB1
## Nup98KD.HepG2   2.72786e-27 7.593352e-20
## Nup98KD.IMR90   1.00000e+00 2.450127e-21
```

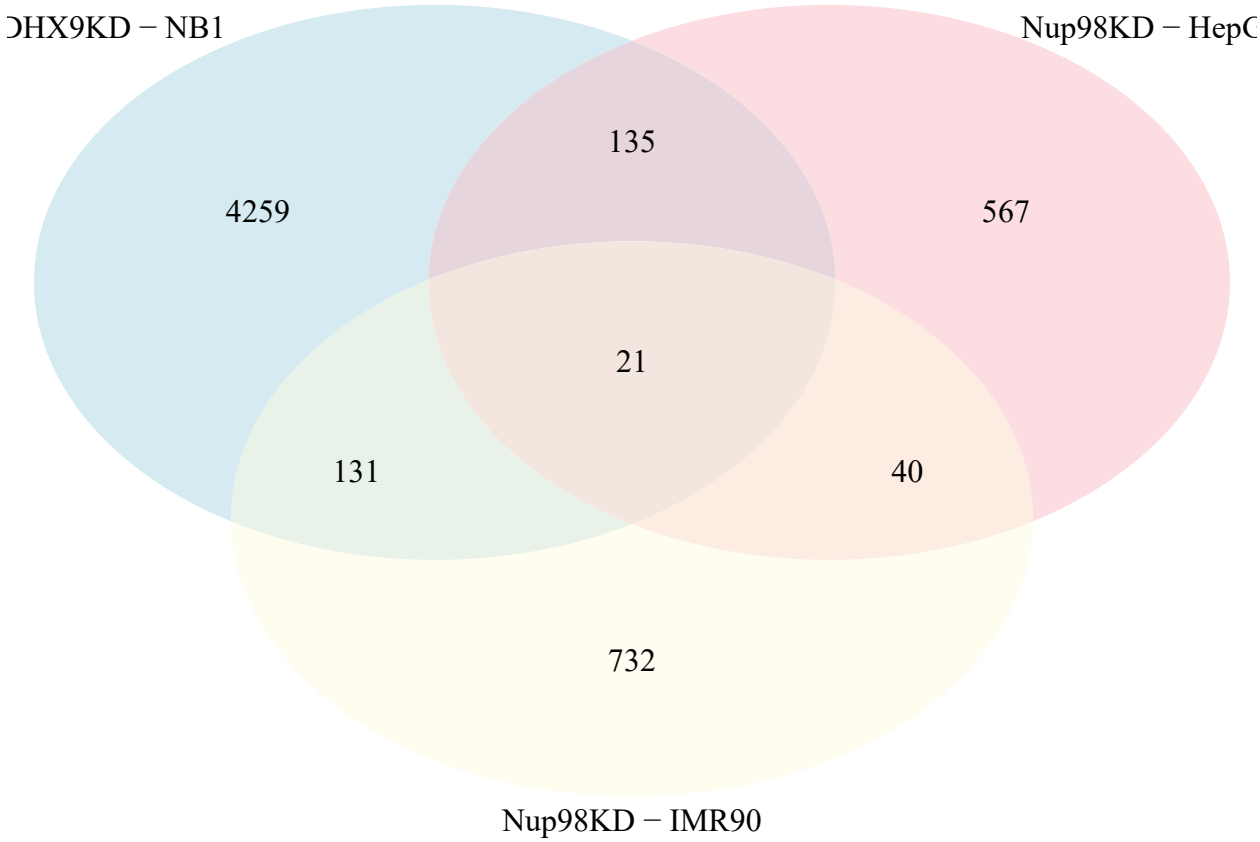
Number of genes in common:

```
getMatrix(gom.self, name="intersection")
```

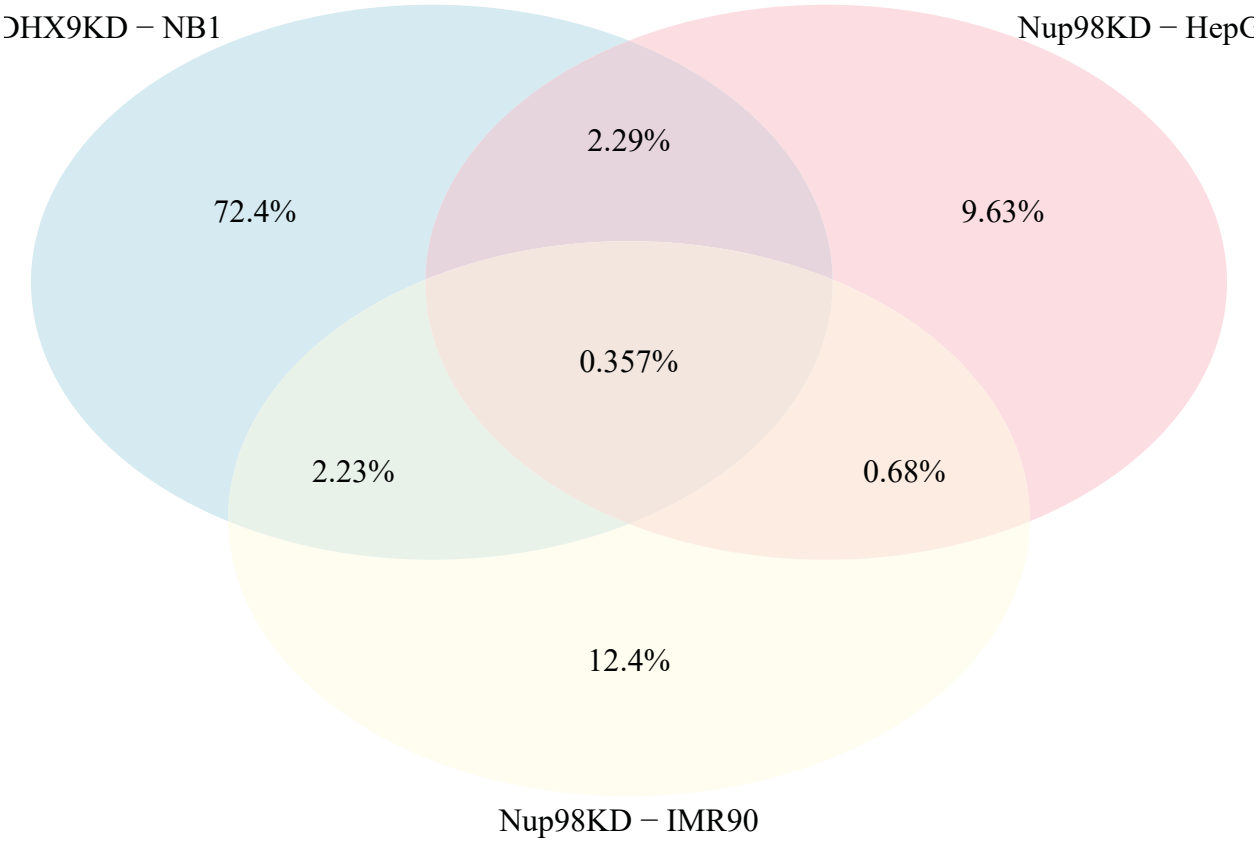
```
##                Nup98KD.IMR90 DHX9KD.NB1
## Nup98KD.HepG2             61        156
## Nup98KD.IMR90              0        152
```

```
grid.newpage()
#draw.triple.venn(area1, area2, area3, n12, n23, n13, n123,
draw.triple.venn(4546, 763, 924, 156, 61, 152, 21, category =
c("DHX9KD - NB1","Nup98KD - HepG2", "Nup98KD - IMR90"), lty =
```

```
rep("blank", 3),  fill = c("light blue", "pink", "light yellow"),
alpha = rep(0.5, 3), cat.pos = c(-40,40,180), cat.dist =
c(0.05, 0.05, 0.025))
```




```
grid.newpage()
#draw.triple.venn(area1, area2, area3, n12, n23, n13, n123,
draw.triple.venn(4546, 763, 924, 156, 61, 152, 21, category =
c("DHX9KD - NB1","Nup98KD - HepG2", "Nup98KD - IMR90"), lty =
rep("blank", 3),  fill = c("light blue", "pink", "light yellow"),
alpha = rep(0.5, 3), cat.pos = c(-40,40,180), cat.dist =
c(0.05, 0.05, 0.025),print.mode="percent")
```



Are genes with altered splicing upon Nup98 or DHX9 KD enriched in Nup98 or DHX9 RNA-IPs?

```
gs.RNASeq <- 48321 #Number of all possible genes from ENSG
gom.obj <- newGOM(Splice.all, RIP.all, genome.size=gs.RNASeq)
getMatrix(gom.obj, name="pval")
```

```
##                Nup98RIP.K562 DHX9RIP.T32         both
## Nup98KD.HepG2 3.487218e-242  0.56362514 4.035794e-36
## Nup98KD.IMR90 3.467813e-317  0.01866402 6.034974e-67
## DHX9KD.NB1    1.534022e-164  0.26566206 5.226360e-28
```

Number of genes in common:

```
getMatrix(gom.obj, name="intersection")
```

```
##                Nup98RIP.K562 DHX9RIP.T32 both
## Nup98KD.HepG2           889         339  287
## Nup98KD.IMR90          1146         474  414
## DHX9KD.NB1              668         275  223
```

Are genes with altered expression upon Nup98 or DHX9 KD enriched in Nup98 or DHX9 RNA-IPs?

```
gom.obj <- newGOM(Transcript$all, RIP.all, genome.size=gs.RNASeq)
getMatrix(gom.obj, name="pval")
```

```
##               Nup98RIP.K562  DHX9RIP.T32          both
## Nup98KD.HepG2  2.631757e-67 0.0089745245  8.901021e-15
## Nup98KD.IMR90  7.673003e-44 0.0006578878  6.382602e-16
## DHX9KD.NB1     0.000000e+00 0.0075508145 6.099347e-109
```

Number of genes in common:

```
getMatrix(gom.obj, name="intersection")
```

```
##               Nup98RIP.K562 DHX9RIP.T32 both
## Nup98KD.HepG2           403         213  142
## Nup98KD.IMR90           343         209  138
## DHX9KD.NB1             2602        1229  936
```

Are genes with altered expression also enriched for CRE regulatory elements?

Table of p-values:

```
gom.obj <- newGOM(Transcript$all, CRE.genes, genome.size=gs.RNASeq)
getMatrix(gom.obj, name="pval")
```

```
##                         CRE
## Nup98KD.HepG2  5.080665e-37
## Nup98KD.IMR90  7.192518e-49
## DHX9KD.NB1    3.594919e-127
```

Number of genes in common:

```
getMatrix(gom.obj, name="intersection")
```

```
##                CRE
## Nup98KD.HepG2  357
## Nup98KD.IMR90  361
## DHX9KD.NB1    1975
```

Are genes with altered splicing also enriched for CRE regulatory elements?

Table of p-values:

```
gom.obj <- newGOM(Splice.all, CRE.genes, genome.size=gs.RNASeq)
getMatrix(gom.obj, name="pval")
```

```
##                         CRE
## Nup98KD.HepG2 2.401469e-29
## Nup98KD.IMR90 1.509787e-31
```

```
## DHX9KD.NB1    4.693071e-30
```

Number of genes in common:

```
getMatrix(gom.obj, name="intersection")
```

```
##                CRE
## Nup98KD.HepG2 557
## Nup98KD.IMR90 695
## DHX9KD.NB1    456
```

Compare Dam-Nup98 interacting genes with CRE element containing genes

Table of p-values for statically significant gene set overlaps:

```
gom.obj <- newGOM(CRE.genes, Dam.all, gs.RNASeq)
getMatrix(gom.obj, name="pval")
```

```
##   Nup98FL.CRE Nup98dCTD.CRE
##  5.533903e-64  1.023096e-52
```

Number of genes in common:

```
getMatrix(gom.obj, name="intersection")
```

```
##   Nup98FL.CRE Nup98dCTD.CRE
##          1411           884
```

Compare Nup98DamID genes with CRE elements to genes with altered expression upon Nup98 or DHX9 KD.

Table of p-values:

```
gom.obj <- newGOM(N98Dam.CRE, Transcript$all, gs.RNASeq)
getMatrix(gom.obj, name="pval")
```

```
##                 Nup98KD.HepG2 Nup98KD.IMR90    DHX9KD.NB1
## DamNup98FLCRE    2.229494e-24  8.831719e-38 1.021099e-102
## DamNup98dCTDCRE  3.211930e-14  2.694110e-20  5.277103e-77
```

Number of genes in common:

```
getMatrix(gom.obj, name="intersection")
```

```
##                 Nup98KD.HepG2 Nup98KD.IMR90 DHX9KD.NB1
## DamNup98FLCRE              72            87        370
## DamNup98dCTDCRE            43            50        251
```

Compare Nup98DamID genes with CRE elements to genes with altered splicing upon Nup98 or DHX9 KD.

Table of p-values:

```
gom.obj <- newGOM(N98Dam.CRE, Splice.all, gs.RNASeq)
getMatrix(gom.obj, name="pval")
```

```
##                 Nup98KD.HepG2 Nup98KD.IMR90   DHX9KD.NB1
## DamNup98FLCRE    4.757813e-32  5.605825e-39 1.326171e-24
## DamNup98dCTDCRE  3.492706e-17  6.310816e-30 4.408208e-20
```

Number of genes in common:

```
getMatrix(gom.obj, name="intersection")
```

```
##                 Nup98KD.HepG2 Nup98KD.IMR90 DHX9KD.NB1
## DamNup98FLCRE             114           143         88
## DamNup98dCTDCRE            66            98         62
```

Compare Nup98DamID genes with CRE elements to mRNAs bound to Nup98 or DHX9 in RNA-IPs

Table of p-values:

```
gom.obj <- newGOM(N98Dam.CRE, RIP.all, gs.RNASeq)
getMatrix(gom.obj, name="pval")
```

```
##                 Nup98RIP.K562  DHX9RIP.T32         both
## DamNup98FLCRE   4.209607e-205 2.354996e-05 1.048003e-52
## DamNup98dCTDCRE 3.498157e-160 8.500419e-06 1.136685e-51
```

Number of genes in common:

```
getMatrix(gom.obj, name="intersection")
```

```
##                 Nup98RIP.K562 DHX9RIP.T32 both
## DamNup98FLCRE            1023         513  386
## DamNup98dCTDCRE           682         337  280
```

Compare CRE containing genes, bound by Dam-Nup98, bound to RNA-IP of Nup98 or DHX9 to genes with altered expression upon Nup98 or DHX9 KD

Table of p-values

```
gom.obj <- newGOM(N98Dam.CRE.RIP, Transcript$all, gs.RNASeq)
getMatrix(gom.obj, name="pval")
```

```
##                    Nup98KD.HepG2 Nup98KD.IMR90   DHX9KD.NB1
## CRE.DamFL.N98RIP    5.985168e-22  1.463091e-26 6.839806e-89
## CRE.DamdCT.N98RIP   8.556358e-12  5.473753e-17 1.404156e-67
## CRE.DamFL.D9RIP     5.481425e-07  1.428984e-14 1.398558e-28
## CRE.DamdCT.D9RIP    1.213211e-06  1.248059e-11 1.744279e-25
## CRE.DamFL.bothRIP   1.919544e-08  3.125601e-11 2.234810e-28
## CRE.DamdCT.bothRIP  2.450644e-06  1.218374e-10 1.971842e-23
```

Number of genes in common:

```
getMatrix(gom.obj, name="intersection")
```

```
##                    Nup98KD.HepG2 Nup98KD.IMR90 DHX9KD.NB1
## CRE.DamFL.N98RIP              58            62        290
## CRE.DamdCT.N98RIP             34            40        205
## CRE.DamFL.D9RIP               22            32        120
## CRE.DamdCT.D9RIP              17            23         89
```

```
## CRE.DamFL.bothRIP              21             24         101
## CRE.DamdCT.bothRIP             15             20          77
```

```
getNestedList(gom.obj, name="intersection")
```

```
## $Nup98KD.HepG2
## $Nup98KD.HepG2$CRE.DamFL.N98RIP
##  [1] "HMGCR"     "CDK6"      "NR6A1"     "RHOBTB1"   "HSD17B12"
##  [6] "MAP1B"     "TPM1"      "FHL2"      "IFRD1"     "CPS1"
## [11] "GDA"       "CLDN11"    "MBNL2"     "LHFPL2"    "WDR3"
## [16] "DNAJB4"    "SLC2A2"    "LDLR"      "HTR1D"     "KITLG"
## [21] "PRRG1"     "AOX1"      "GCNT2"     "CTH"       "FYN"
## [26] "LCP1"      "ABLIM1"    "MDM2"      "PRPS2"     "PDHX"
## [31] "SLC4A7"    "SLC25A12"  "TNFRSF10B" "CTSS"      "STK17B"
## [36] "SLC16A5"   "MAP4K4"    "PCYT1B"    "BCAT1"     "YAP1"
## [41] "TACC1"     "TXNIP"     "TOPBP1"    "SARDH"     "RRAS2"
## [46] "MYO10"     "MAT2B"     "CPA4"      "RNF125"    "PLEKHB2"
## [51] "ASPM"      "ARHGEF3"   "CNNM1"     "ELOVL6"    "CPEB4"
## [56] "CYGB"      "SLC39A11"  "NAV2"
##
## $Nup98KD.HepG2$CRE.DamdCT.N98RIP
##  [1] "NR6A1"     "PIM1"      "SDC4"      "FHL2"      "IFRD1"
##  [6] "CPS1"      "FDPS"      "GDA"       "TCEA1"     "DNAJB4"
## [11] "LDLR"      "HTR1D"     "KITLG"     "PRRG1"     "EMP2"
## [16] "GCNT2"     "AKR1B1"    "CPT1A"     "FYN"       "MDM2"
## [21] "WEE1"      "TNFRSF10B" "STK17B"    "SLC16A5"   "AKAP12"
## [26] "YAP1"      "MYO10"     "MAT2B"     "GIPC2"     "PLEKHB2"
## [31] "ASPM"      "CPVL"      "SLC39A11"  "NAV2"
##
## $Nup98KD.HepG2$CRE.DamFL.D9RIP
##  [1] "HSD17B12"  "TPM1"      "GDA"       "CLDN11"    "WDR3"
##  [6] "DNAJB4"    "LIPC"      "HTR1D"     "PRRG1"     "AOX1"
## [11] "GCNT2"     "TNFRSF10B" "PCYT1B"    "BCAT1"     "SARDH"
## [16] "MYO10"     "RNF125"    "PLEKHB2"   "ASPM"      "CPEB4"
## [21] "CYGB"      "NAV2"
##
## $Nup98KD.HepG2$CRE.DamdCT.D9RIP
##  [1] "PIM1"      "GDA"       "DNAJB4"    "HECA"      "HTR1D"
##  [6] "PRRG1"     "SLC6A2"    "EMP2"      "GCNT2"     "WEE1"
## [11] "TNFRSF10B" "MYO10"     "GIPC2"     "PLEKHB2"   "ASPM"
## [16] "CPVL"      "NAV2"
##
## $Nup98KD.HepG2$CRE.DamFL.bothRIP
##  [1] "HSD17B12"  "TPM1"      "GDA"       "CLDN11"    "WDR3"
##  [6] "DNAJB4"    "HTR1D"     "PRRG1"     "AOX1"      "GCNT2"
## [11] "TNFRSF10B" "PCYT1B"    "BCAT1"     "SARDH"     "MYO10"
## [16] "RNF125"    "PLEKHB2"   "ASPM"      "CPEB4"     "CYGB"
## [21] "NAV2"
##
## $Nup98KD.HepG2$CRE.DamdCT.bothRIP
##  [1] "PIM1"      "GDA"       "DNAJB4"    "HTR1D"     "PRRG1"
##  [6] "EMP2"      "GCNT2"     "WEE1"      "TNFRSF10B" "MYO10"
## [11] "GIPC2"     "PLEKHB2"   "ASPM"      "CPVL"      "NAV2"
##
##
```

```
## $Nup98KD.IMR90
## $Nup98KD.IMR90$CRE.DamFL.N98RIP
##  [1] "ADAM10"    "KLF5"      "NEO1"      "NUP98"     "NCOA2"
##  [6] "MAFF"      "SSX2IP"    "OSBPL1A"   "ATF3"      "PCDH7"
## [11] "SHB"       "SLC7A2"    "STAT3"     "ETS2"      "LHFPL2"
## [16] "PPP1R12B"  "MDGA1"     "PRDM1"     "ACADM"     "CYP1B1"
## [21] "LAMA3"     "LAMA2"     "LDLR"      "CD44"      "GRIA4"
## [26] "IL1R1"     "PDE3B"     "PLAT"      "ADM"       "ETV4"
## [31] "FMOD"      "FYN"       "ITPR1"     "KCNMA1"    "LAMA4"
## [36] "MDM2"      "SLC3A2"    "MYC"       "PAPPA"     "PAWR"
## [41] "TNFSF4"    "SLC4A7"    "PPAP2A"    "TNFRSF10B" "BTRC"
## [46] "CTSS"      "SLC1A2"    "PDE1A"     "APBA2"     "MRVI1"
## [51] "NNMT"      "TNFAIP2"   "NEDD9"     "CHL1"      "LPHN2"
## [56] "NOX4"      "CHST11"    "MAN1C1"    "NAV1"      "SLC12A8"
## [61] "CYBRD1"    "CYGB"
##
## $Nup98KD.IMR90$CRE.DamdCT.N98RIP
##  [1] "KLF5"      "NUP98"     "ITGA3"     "OSBPL1A"   "ATF3"
##  [6] "NFKB2"     "STAT3"     "WISP1"     "EPHA2"     "ETS2"
## [11] "ELN"       "LDLR"      "IL6R"      "BDKRB1"    "ADM"
## [16] "ACTA2"     "AKR1B1"    "CPT1A"     "ETV4"      "FYN"
## [21] "LAMA4"     "MDM2"      "SLC3A2"    "MYC"       "PAPPA"
## [26] "PAWR"      "SVIL"      "TNFRSF10B" "BTRC"      "SLC1A2"
## [31] "PDE1A"     "TNFSF15"   "MT2A"      "NNMT"      "TNFAIP2"
## [36] "SLC25A13"  "RAB30"     "CARD10"    "SLC40A1"   "SLC12A8"
##
## $Nup98KD.IMR90$CRE.DamFL.D9RIP
##  [1] "KLF5"      "NUP98"     "OSBPL1A"   "SHB"       "STAT3"
##  [6] "ETS2"      "SMOC1"     "MDGA1"     "PRDM1"     "ACADM"
## [11] "LAMA3"     "PTEN"      "TYRP1"     "CD44"      "GRIA4"
## [16] "IL1R1"     "ADM"       "COL15A1"   "ITPR1"     "SLC3A2"
## [21] "PAPPA"     "PAWR"      "TNFRSF10B" "SLC1A2"    "PTGES"
## [26] "APBA2"     "TNFAIP2"   "TOX"       "COL5A3"    "CHST11"
## [31] "BMP5"      "CYGB"
##
## $Nup98KD.IMR90$CRE.DamdCT.D9RIP
##  [1] "KLF5"      "NUP98"     "ITGA3"     "OSBPL1A"   "STAT3"
##  [6] "ETS2"      "PTEN"      "ELN"       "IL6R"      "BDKRB1"
## [11] "ADM"       "ACTA2"     "SLC3A2"    "PAPPA"     "PAWR"
## [16] "SVIL"      "TNFRSF10B" "SLC1A2"    "PTGES"     "TNFAIP2"
## [21] "RAB30"     "CARD10"    "BMP5"
##
## $Nup98KD.IMR90$CRE.DamFL.bothRIP
##  [1] "KLF5"      "NUP98"     "OSBPL1A"   "SHB"       "STAT3"
##  [6] "ETS2"      "MDGA1"     "PRDM1"     "ACADM"     "LAMA3"
## [11] "CD44"      "GRIA4"     "IL1R1"     "ADM"       "ITPR1"
## [16] "SLC3A2"    "PAPPA"     "PAWR"      "TNFRSF10B" "SLC1A2"
## [21] "APBA2"     "TNFAIP2"   "CHST11"    "CYGB"
##
## $Nup98KD.IMR90$CRE.DamdCT.bothRIP
##  [1] "KLF5"      "NUP98"     "ITGA3"     "OSBPL1A"   "STAT3"
##  [6] "ETS2"      "ELN"       "IL6R"      "BDKRB1"    "ADM"
## [11] "ACTA2"     "SLC3A2"    "PAPPA"     "PAWR"      "SVIL"
## [16] "TNFRSF10B" "SLC1A2"    "TNFAIP2"   "RAB30"     "CARD10"
```

```
##
##
## $DHX9KD.NB1
## $DHX9KD.NB1$CRE.DamFL.N98RIP
##   [1] "NF2"       "HLCS"      "HMGCR"     "CDK6"      "EPB41L2"
##   [6] "ESRRG"     "GMDS"      "GNS"       "MEIS2"     "NEO1"
##  [11] "PDE7A"     "PDE8A"     "RAD1"      "ZYX"       "MAP7"
##  [16] "TRIP12"    "CREBBP"    "NDUFS2"    "PCYT1A"    "IDH3A"
##  [21] "ACTR1A"    "COG5"      "USP3"      "GNB5"      "USP25"
##  [26] "SACM1L"    "SSX2IP"    "RHOBTB1"   "WSB1"      "HSD17B12"
##  [31] "TBC1D7"    "BCAS3"     "POLR1B"    "ATP10D"    "PREX1"
##  [36] "OSBPL8"    "EGLN3"     "NXN"       "ASPSCR1"   "ADAMTS17"
##  [41] "TTC8"      "TTL"       "GALC"      "NR3C1"     "NPC1"
##  [46] "LTBP1"     "CYP51A1"   "FHL2"      "NDST1"     "IFRD1"
##  [51] "ARNT"      "ATP6V1A"   "CDK2"      "FTH1"      "SLC7A2"
##  [56] "TMPO"      "VCL"       "ADAM23"    "B2M"       "EPB41"
##  [61] "PRKCE"     "TOB1"      "LHFPL2"    "SPRY2"     "NAB1"
##  [66] "MRC2"      "ELL"       "PACSIN2"   "ELL2"      "NLGN1"
##  [71] "PHF15"     "TRIM37"    "SBDS"      "IRAK4"     "DDX24"
##  [76] "TCF7L2"    "PPP1R12B"  "MSI2"      "ADSL"      "BCKDHB"
##  [81] "FBN1"      "GALT"      "ITGB3"     "GPD2"      "PCCB"
##  [86] "IL15"      "ACACA"     "DCK"       "GRIA4"     "HRH1"
##  [91] "IGF1R"     "NQO1"      "P4HA1"     "PDE3A"     "PLAT"
##  [96] "PLOD2"     "RYR2"      "ADM"       "ANXA6"     "BMP1"
## [101] "CASP4"     "CPD"       "DOCK1"     "FHL1"      "GTF2I"
## [106] "ALCAM"     "CAPNS1"    "CBFB"      "CNN3"      "COL4A2"
## [111] "CTSB"      "EPS15"     "ETV4"      "FHIT"      "FKBP3"
## [116] "IARS"      "ITGB5"     "ITPR2"     "LAMB1"     "LAMC1"
## [121] "LIFR"      "MAP2"      "SLC3A2"    "MPV17"     "PNN"
## [126] "POLA2"     "MAP2K6"    "PRPS2"     "PTPN4"     "PTPRR"
## [131] "PXN"       "PYGL"      "RAP2B"     "RDX"       "ROBO1"
## [136] "SDC2"      "SORL1"     "TBCE"      "TCF12"     "TIAM1"
## [141] "TIMP2"     "UVRAG"     "ZNF202"    "PDHX"      "PIP5K1A"
## [146] "BCAR3"     "NCK2"      "CDC42BPA"  "RGS20"     "SLC25A12"
## [151] "VAMP4"     "RIOK3"     "TNFRSF10B" "SYNJ2"     "ADCY3"
## [156] "GNG11"     "STK17B"    "TRIP11"    "BCL9"      "DFNA5"
## [161] "EPS8"      "EZH2"      "GALNT2"    "NFATC3"    "PEX14"
## [166] "PKP2"      "LATS1"     "ARHGEF2"   "MAP4K4"    "PCYT1B"
## [171] "ROCK2"     "LITAF"     "ACTN4"     "NRCAM"     "ABL1"
## [176] "ABL2"      "BCAT1"     "LIMK2"     "LOXL1"     "NFIA"
## [181] "NFIB"      "NFIC"      "SLC1A5"    "SYT1"      "GTF2IRD1"
## [186] "ACTR2"     "HIPK3"     "CTDSPL"    "ABCC4"     "STAG1"
## [191] "SEMA3A"    "YAP1"      "VAV3"      "DDR2"      "PPP1R2"
## [196] "TACC1"     "TYRO3"     "XRCC1"     "BAIAP2"    "SLC25A17"
## [201] "UNC13B"    "TRIM16"    "ZNF217"    "AKAP13"    "SSFA2"
## [206] "RAB31"     "SOX5"      "ZNF33A"    "WWP2"      "CDC42EP1"
## [211] "TRIO"      "DDX42"     "CRB1"      "AK5"       "CASP8AP2"
## [216] "DDAH1"     "FBXL5"     "LPHN2"     "RPS6KC1"   "SLCO3A1"
## [221] "SNX12"     "PCOLCE2"   "BAZ1A"     "BAZ2B"     "SOCS5"
## [226] "TMOD3"     "CYFIP1"    "PHF14"     "LAPTM4A"   "TBC1D5"
## [231] "BTBD3"     "POGZ"      "TRIM2"     "CLASP1"    "POLR1A"
## [236] "AUTS2"     "KCTD3"     "LEF1"      "CENPF"     "ANKFY1"
## [241] "GALNT10"   "CHRNA9"    "PLEKHB2"   "PACS1"     "FANCL"
## [246] "KIRREL"    "LRRN3"     "CENPJ"     "ERBB2IP"   "CLN8"
```

```
## [251] "ARHGEF3"   "TUFT1"     "EIF2B3"    "PHTF2"     "PELI1"
## [256] "KIAA1324"  "GPHN"      "PITPNM2"   "SH3RF1"    "NMT1"
## [261] "CFL2"      "TEAD1"     "CSNK1G1"   "EGLN1"     "EPB41L4A"
## [266] "DNAJC1"    "DCLRE1C"   "HIVEP3"    "ZDHHC14"   "CPEB4"
## [271] "DUSP16"    "SLC38A1"   "GABARAPL1" "SH3KBP1"   "C1QTNF6"
## [276] "KCTD10"    "CHD6"      "CAMKK1"    "GPR124"    "ZDHHC12"
## [281] "TRIM5"     "ARHGAP18"  "ADAMTSL1"  "S100A16"   "NEK7"
## [286] "STXBP5"    "TTBK2"     "JAZF1"     "NCOA7"     "ADCY5"
##
## $DHX9KD.NB1$CRE.DamdCT.N98RIP
##   [1] "HLCS"      "CHEK1"     "EPB41L2"   "ESRRG"     "CREM"
##   [6] "MPP3"      "GNS"       "PDE8A"     "PIM1"      "PTPRJ"
##  [11] "PTPRS"     "RAD1"      "SLC7A5"    "MAP7"      "PRKAB2"
##  [16] "ACTR1A"    "RASSF1"    "USP25"     "HABP4"     "MDN1"
##  [21] "PHYHIP"    "CHSY1"     "CHD5"      "WSB1"      "GADD45B"
##  [26] "KLF13"     "RAB14"     "WBSCR22"   "POLR1B"    "EML4"
##  [31] "PREX1"     "NXN"       "ASPSCR1"   "SP1"       "TTC8"
##  [36] "NR3C1"     "LTBP1"     "CYP51A1"   "PAM"       "FHL2"
##  [41] "IFRD1"     "ATP6V1A"   "ID1"       "MAP3K11"   "NFKB2"
##  [46] "UAP1"      "TMPO"      "VCL"       "FABP3"     "EPB41"
##  [51] "RAB5C"     "ATP5G1"    "ATP5G2"    "TAF12"     "SPRY2"
##  [56] "CD164"     "TCEA1"     "PACSIN2"   "PHF15"     "ACTR6"
##  [61] "TCF7L2"    "MSI2"      "ADSL"      "FBN1"      "NAGLU"
##  [66] "PCCB"      "SLC11A2"   "ALDH3B1"   "DCK"       "HRH1"
##  [71] "IGF1R"     "NQO1"      "PDE3A"     "TKT"       "ADM"
##  [76] "CASP8"     "DNASE2"    "DOCK1"     "EMP2"      "FHL1"
##  [81] "GTF2I"     "AARS"      "ALCAM"     "CBFB"      "CNN3"
##  [86] "ATF2"      "CTNNA1"    "ETV4"      "INPP1"     "ITGB5"
##  [91] "LAMC1"     "LOXL2"     "SLC3A2"    "MPV17"     "NPAS2"
##  [96] "PRDX1"     "POLA2"     "POLR2G"    "PTPN4"     "PTPRR"
## [101] "PXN"       "RANGAP1"   "ROBO1"     "SHC1"      "SMARCA4"
## [106] "TBCE"      "TCF12"     "TIMP2"     "TJP1"      "UVRAG"
## [111] "WEE1"      "PIP5K1A"   "BCAR3"     "PIK3R3"    "CTSF"
## [116] "TNFRSF10B" "SYNJ2"     "ADCY3"     "STK17B"    "BCL9"
## [121] "DFNA5"     "GALNT2"    "PSPH"      "SNRPA"     "LATS1"
## [126] "HS6ST1"    "ROCK2"     "LITAF"     "ABCG1"     "ACTN4"
## [131] "MARK2"     "ABCC1"     "ABL1"      "ABL2"      "CAPN1"
## [136] "PODXL"     "PPP1R3C"   "HSPG2"     "NFIA"      "NFIB"
## [141] "NFIC"      "SLC1A5"    "GTF2IRD1"  "DPP3"      "HIPK3"
## [146] "FARP1"     "RBM6"      "ABCC4"     "STAG1"     "MT2A"
## [151] "TBX1"      "YAP1"      "VAV3"      "DDR2"      "PPP1R2"
## [156] "LYPLA1"    "BAIAP2"    "TRIM16"    "NCOA3"     "GMEB1"
## [161] "TFEB"      "CRB1"      "PGLS"      "AK5"       "CASP8AP2"
## [166] "ESPL1"     "KPNA6"     "RPS6KC1"   "SOCS5"     "HIBCH"
## [171] "TAF5L"     "ARFGAP3"   "ERO1L"     "PHF14"     "TBC1D5"
## [176] "AUTS2"     "KCTD3"     "LEF1"      "ANKFY1"    "PLEKHB2"
## [181] "PACS1"     "EFHC1"     "CLN8"      "CPVL"      "SH3RF1"
## [186] "NMT1"      "POLD4"     "DTNB"      "LHPP"      "ZDHHC14"
## [191] "DUSP16"    "SLC38A1"   "MRPL9"     "C1QTNF6"   "CAMKK1"
## [196] "LACTB"     "ABCC10"    "SOCS7"     "NEK7"      "STXBP5"
## [201] "LACE1"     "GJC1"      "JAZF1"     "NEK8"      "OXR1"
##
## $DHX9KD.NB1$CRE.DamFL.D9RIP
##   [1] "ESRRG"     "PDE8A"     "ZYX"       "PCYT1A"    "IDH3A"
```

```
##   [6] "GNB5"      "HSD17B12"  "CRIM1"     "BCAS3"     "PREX1"
##  [11] "EGLN3"     "ADAMTS17"  "NR3C1"     "CYP51A1"   "NDST1"
##  [16] "TOP1"      "B2M"       "GNG7"      "SPRY2"     "ELL"
##  [21] "UBL3"      "PACSIN2"   "ELL2"      "TCF7L2"    "FBN1"
##  [26] "GALT"      "GPD2"      "IL15"      "ANXA1"     "GRIA4"
##  [31] "HRH1"      "IGF1R"     "NQO1"      "PDE3A"     "ADM"
##  [36] "ANXA6"     "CASP4"     "DOCK1"     "ENO1"      "GTF2I"
##  [41] "COL4A2"    "FKBP3"     "IARS"      "KCNH1"     "LIFR"
##  [46] "SLC3A2"    "NTRK3"     "PTPRR"     "PYGL"      "SDC2"
##  [51] "TBCE"      "TIAM1"     "TNR"       "PIP5K1A"   "BCAR3"
##  [56] "NCK2"      "CDC42BPA"  "TNFRSF10B" "EIF2S2"    "ADCY3"
##  [61] "GNG11"     "NEURL"     "EPS8"      "GALNT2"    "NFATC3"
##  [66] "PCYT1B"    "LITAF"     "PTGES"     "ACTN4"     "ABL2"
##  [71] "BCAT1"     "LIMK2"     "LOXL1"     "NFIA"      "NFIB"
##  [76] "SYT1"      "BAIAP2"    "SLC25A17"  "TRIM16"    "PDE10A"
##  [81] "AKAP13"    "WWP2"      "CDC42EP1"  "TRIO"      "DDX42"
##  [86] "CRB1"      "FBXL5"     "SLCO3A1"   "BAZ1A"     "TMOD3"
##  [91] "SLK"       "POLR1A"    "ANKFY1"    "PLEKHB2"   "KIRREL"
##  [96] "LRRN3"     "CLN8"      "TUFT1"     "EIF2B3"    "KIAA1324"
## [101] "GPHN"      "NMT1"      "CSNK1G1"   "EPB41L4A"  "HIVEP3"
## [106] "ZDHHC14"   "CPEB4"     "C1QTNF6"   "PHF6"      "ABLIM2"
## [111] "GPR124"    "ZDHHC12"   "TRIM5"     "ADAMTSL1"  "NEK7"
## [116] "RAB3C"     "SEPT10"    "TTBK2"     "THAP5"     "ADCY5"
##
## $DHX9KD.NB1$CRE.DamdCT.D9RIP
##  [1] "CHEK1"     "ESRRG"     "PDE8A"     "PIM1"      "RNASEH2A"
##  [6] "HABP4"     "NPTXR"     "PHYHIP"    "GADD45B"   "CRIM1"
## [11] "PREX1"     "NR3C1"     "CYP51A1"   "ID1"       "RAB5C"
## [16] "GNG7"      "ATP5G1"    "SPRY2"     "PACSIN2"   "HECA"
## [21] "TCF7L2"    "FBN1"      "NAGLU"     "ANXA1"     "HRH1"
## [26] "IGF1R"     "NQO1"      "PDE3A"     "SLC6A2"    "TKT"
## [31] "ADM"       "DOCK1"     "EMP2"      "GTF2I"     "CTNNA1"
## [36] "INPP1"     "LOXL2"     "SLC3A2"    "NPAS2"     "PRDX1"
## [41] "POLR2G"    "PTPRR"     "RANGAP1"   "TBCE"      "TJP1"
## [46] "WEE1"      "PIP5K1A"   "BCAR3"     "PIK3R3"    "TNFRSF10B"
## [51] "ADCY3"     "GALNT2"    "LITAF"     "PTGES"     "ABCG1"
## [56] "ACTN4"     "ABL2"      "PPP1R3C"   "HSPG2"     "LAMP1"
## [61] "NFIA"      "NFIB"      "DPP3"      "RBM6"      "TBX1"
## [66] "BAIAP2"    "TRIM16"    "NCOA3"     "EHD1"      "TFEB"
## [71] "CRB1"      "PGLS"      "HS2ST1"    "KPNA6"     "HIBCH"
## [76] "ARFGAP3"   "ANKFY1"    "PLEKHB2"   "CLN8"      "CPVL"
## [81] "NMT1"      "ZDHHC14"   "MRPL9"     "C1QTNF6"   "ABLIM2"
## [86] "LACTB"     "ABCC10"    "NEK7"      "NEK8"
##
## $DHX9KD.NB1$CRE.DamFL.bothRIP
##   [1] "ESRRG"     "PDE8A"     "ZYX"       "PCYT1A"    "IDH3A"
##   [6] "GNB5"      "HSD17B12"  "BCAS3"     "PREX1"     "EGLN3"
##  [11] "ADAMTS17"  "NR3C1"     "CYP51A1"   "NDST1"     "B2M"
##  [16] "SPRY2"     "ELL"       "PACSIN2"   "ELL2"      "TCF7L2"
##  [21] "FBN1"      "GALT"      "GPD2"      "IL15"      "GRIA4"
##  [26] "HRH1"      "IGF1R"     "NQO1"      "PDE3A"     "ADM"
##  [31] "ANXA6"     "CASP4"     "DOCK1"     "GTF2I"     "COL4A2"
##  [36] "FKBP3"     "IARS"      "LIFR"      "SLC3A2"    "PTPRR"
##  [41] "PYGL"      "SDC2"      "TBCE"      "TIAM1"     "PIP5K1A"
```

```
##  [46] "BCAR3"     "NCK2"      "CDC42BPA"  "TNFRSF10B" "ADCY3"
##  [51] "GNG11"     "EPS8"      "GALNT2"    "NFATC3"    "PCYT1B"
##  [56] "LITAF"     "ACTN4"     "ABL2"      "BCAT1"     "LIMK2"
##  [61] "LOXL1"     "NFIA"      "NFIB"      "SYT1"      "BAIAP2"
##  [66] "SLC25A17"  "TRIM16"    "AKAP13"    "WWP2"      "CDC42EP1"
##  [71] "TRIO"      "DDX42"     "CRB1"      "FBXL5"     "SLCO3A1"
##  [76] "BAZ1A"     "TMOD3"     "POLR1A"    "ANKFY1"    "PLEKHB2"
##  [81] "KIRREL"    "LRRN3"     "CLN8"      "TUFT1"     "EIF2B3"
##  [86] "KIAA1324"  "GPHN"      "NMT1"      "CSNK1G1"   "EPB41L4A"
##  [91] "HIVEP3"    "ZDHHC14"   "CPEB4"     "C1QTNF6"   "GPR124"
##  [96] "ZDHHC12"   "TRIM5"     "ADAMTSL1"  "NEK7"      "TTBK2"
## [101] "ADCY5"
##
## $DHX9KD.NB1$CRE.DamdCT.bothRIP
##  [1] "CHEK1"     "ESRRG"     "PDE8A"     "PIM1"      "HABP4"
##  [6] "PHYHIP"    "GADD45B"   "PREX1"     "NR3C1"     "CYP51A1"
## [11] "ID1"       "RAB5C"     "ATP5G1"    "SPRY2"     "PACSIN2"
## [16] "TCF7L2"    "FBN1"      "NAGLU"     "HRH1"      "IGF1R"
## [21] "NQO1"      "PDE3A"     "TKT"       "ADM"       "DOCK1"
## [26] "EMP2"      "GTF2I"     "CTNNA1"    "INPP1"     "LOXL2"
## [31] "SLC3A2"    "NPAS2"     "PRDX1"     "POLR2G"    "PTPRR"
## [36] "RANGAP1"   "TBCE"      "TJP1"      "WEE1"      "PIP5K1A"
## [41] "BCAR3"     "PIK3R3"    "TNFRSF10B" "ADCY3"     "GALNT2"
## [46] "LITAF"     "ABCG1"     "ACTN4"     "ABL2"      "PPP1R3C"
## [51] "HSPG2"     "NFIA"      "NFIB"      "DPP3"      "RBM6"
## [56] "TBX1"      "BAIAP2"    "TRIM16"    "NCOA3"     "TFEB"
## [61] "CRB1"      "PGLS"      "KPNA6"     "HIBCH"     "ARFGAP3"
## [66] "ANKFY1"    "PLEKHB2"   "CLN8"      "CPVL"      "NMT1"
## [71] "ZDHHC14"   "MRPL9"     "C1QTNF6"   "LACTB"     "ABCC10"
## [76] "NEK7"      "NEK8"
```

Compare CRE containing genes, bound by Dam-Nup98, bound to RNA-IP of Nup98 or DHX9 to genes with altered splicing upon Nup98 or DHX9 KD

Table of p-values

```
gom.obj <- newGOM(N98Dam.CRE.RIP, Splice.all, gs.RNASeq)
getMatrix(gom.obj, name="pval")
```

```
##                    Nup98KD.HepG2 Nup98KD.IMR90   DHX9KD.NB1
## CRE.DamFL.N98RIP    3.568999e-34  1.330982e-39 1.359331e-22
## CRE.DamdCT.N98RIP   8.727141e-18  1.145908e-29 7.103040e-18
## CRE.DamFL.D9RIP     6.249886e-11  1.731011e-07 6.247124e-11
## CRE.DamdCT.D9RIP    1.797240e-08  1.505827e-04 7.597436e-07
## CRE.DamFL.bothRIP   2.097131e-11  1.327236e-07 4.032745e-10
## CRE.DamdCT.bothRIP  7.034966e-09  2.640526e-05 7.120887e-07
```

Number of genes in common:

```
getMatrix(gom.obj, name="intersection")
```

```
##                    Nup98KD.HepG2 Nup98KD.IMR90 DHX9KD.NB1
## CRE.DamFL.N98RIP              99           121         71
## CRE.DamdCT.N98RIP             58            85         51
## CRE.DamFL.D9RIP               39            38         34
```

```
## CRE.DamdCT.D9RIP              27            23         21
## CRE.DamFL.bothRIP             34            32         28
## CRE.DamdCT.bothRIP            25            22         19
```

```
getNestedList(gom.obj, name="intersection")
```

```
## $Nup98KD.HepG2
## $Nup98KD.HepG2$CRE.DamFL.N98RIP
##  [1] "HLCS"     "THRB"     "MEIS2"    "NFX1"     "NUP155"   "NOLC1"
##  [7] "MYO6"     "CLIC4"    "GLS"      "KIFAP3"   "ANGPTL4"  "BCAS3"
## [13] "GDAP2"    "OSBPL8"   "OSBPL9"   "ZNF341"   "TPM3"     "UBR1"
## [19] "DMD"      "PMS2"     "CDK2"     "CPS1"     "SLC7A2"   "HIVEP2"
## [25] "MYO9A"    "IRAK4"    "SFMBT1"   "PKD2"     "LTA4H"    "ANXA6"
## [31] "ATR"      "BMP1"     "CCNT2"    "FLNB"     "CTH"      "ITPR1"
## [37] "KRT8"     "NDUFC1"   "POLA2"    "MAP2K5"   "SPTAN1"   "TBP"
## [43] "TSSC1"    "UVRAG"    "ZNF202"   "PIP5K1A"  "PPFIBP1"  "PRC1"
## [49] "ELF3"     "ARHGEF2"  "RASAL2"   "MAP3K8"   "ETS1"     "LMO7"
## [55] "SDCBP"    "HIPK3"    "YAP1"     "TFPI"     "FRS2"     "AKAP13"
## [61] "ZNF33A"   "TOPBP1"   "TRIO"     "CIT"      "RNF24"    "CHORDC1"
## [67] "SCMH1"    "SIRT4"    "CNOT4"    "STAU2"    "MELK"     "HELZ"
## [73] "KIAA0922" "PUM2"     "OPA1"     "PCF11"    "GALNT10"  "USP47"
## [79] "SLC30A6"  "CENPJ"    "ERBB2IP"  "TUFT1"    "NDRG4"    "NMT1"
## [85] "MTMR3"    "PTBP2"    "DHX35"    "C14orf93" "TEAD1"    "ZNF148"
## [91] "KIF13A"   "DCLRE1C"  "BICC1"    "TDRD3"    "KCTD10"   "TRIM5"
## [97] "BTBD9"    "APBB2"    "NCOA7"
##
## $Nup98KD.HepG2$CRE.DamdCT.N98RIP
##  [1] "HLCS"     "THRB"     "MARK3"    "NUP155"   "BIN1"     "NOLC1"
##  [7] "HDAC5"    "NCOR2"    "MAP4K5"   "MDN1"     "EML4"     "SP1"
## [13] "TPM3"     "DMD"      "CPS1"     "RACGAP1"  "BBX"      "PSEN2"
## [19] "LTA4H"    "CASP8"    "FLNB"     "KRT8"     "POLA2"    "MAP2K5"
## [25] "UGCG"     "UVRAG"    "PIP5K1A"  "PLA2G6"   "PRC1"     "RASAL2"
## [31] "ABCC1"    "MAP3K8"   "HIPK3"    "YAP1"     "TFPI"     "GMEB1"
## [37] "TLK2"     "ZNF92"    "CHEK2"    "CHORDC1"  "STAU2"    "RAB30"
## [43] "ARHGEF11" "HELZ"     "DHX30"    "KIAA0922" "USP47"    "TBC1D2"
## [49] "ABCB9"    "AGTRAP"   "NMT1"     "MTMR3"    "KIF13A"   "ABCC10"
## [55] "BTBD9"    "NEK8"     "OXR1"     "GPR133"
##
## $Nup98KD.HepG2$CRE.DamFL.D9RIP
##  [1] "THRB"     "DDX10"    "NOLC1"    "KIFAP3"   "ANGPTL4"  "BCAS3"
##  [7] "OSBPL9"   "TPM3"     "DMD"      "HIVEP2"   "MYO9A"    "LTA4H"
## [13] "ANXA6"    "CCNT2"    "FLNB"     "ITPR1"    "NDUFC1"   "TSSC1"
## [19] "PIP5K1A"  "PRC1"     "CDKN3"    "FRS2"     "AKAP13"   "TRIO"
## [25] "CNOT4"    "STAU2"    "HELZ"     "TUFT1"    "SMURF1"   "NMT1"
## [31] "MTMR3"    "C14orf93" "KIF13A"   "TDRD3"    "PHF6"     "TRIM5"
## [37] "JMY"      "BTBD9"    "APBB2"
##
## $Nup98KD.HepG2$CRE.DamdCT.D9RIP
##  [1] "THRB"     "BIN1"     "NOLC1"    "HDAC5"    "MAP4K5"   "TPM3"     "DMD"
##  [8] "RACGAP1"  "LTA4H"    "FLNB"     "UGCG"     "PIP5K1A"  "PRC1"     "TLK2"
## [15] "STAU2"    "RAB30"    "HELZ"     "DHX30"    "NMT1"     "MTMR3"    "KIF13A"
## [22] "ABCC10"   "JMY"      "BTBD9"    "GPR97"    "NEK8"     "GPR133"
##
## $Nup98KD.HepG2$CRE.DamFL.bothRIP
##  [1] "THRB"     "NOLC1"    "KIFAP3"   "ANGPTL4"  "BCAS3"    "OSBPL9"
```

```
##  [7] "TPM3"     "DMD"      "HIVEP2"   "MYO9A"    "LTA4H"    "ANXA6"
## [13] "CCNT2"    "FLNB"     "ITPR1"    "NDUFC1"   "TSSC1"    "PIP5K1A"
## [19] "PRC1"     "FRS2"     "AKAP13"   "TRIO"     "CNOT4"    "STAU2"
## [25] "HELZ"     "TUFT1"    "NMT1"     "MTMR3"    "C14orf93" "KIF13A"
## [31] "TDRD3"    "TRIM5"    "BTBD9"    "APBB2"
##
## $Nup98KD.HepG2$CRE.DamdCT.bothRIP
##  [1] "THRB"    "BIN1"    "NOLC1"   "HDAC5"   "MAP4K5"  "TPM3"    "DMD"
##  [8] "RACGAP1" "LTA4H"   "FLNB"    "UGCG"    "PIP5K1A" "PRC1"    "TLK2"
## [15] "STAU2"   "RAB30"   "HELZ"    "DHX30"   "NMT1"    "MTMR3"   "KIF13A"
## [22] "ABCC10"  "BTBD9"   "NEK8"    "GPR133"
##
##
## $Nup98KD.IMR90
## $Nup98KD.IMR90$CRE.DamFL.N98RIP
##   [1] "MLH1"      "NF2"       "HLCS"      "GOLGA4"    "MEIS2"
##   [6] "RAD1"      "SAFB"      "SRP54"     "SMS"       "NOLC1"
##  [11] "METTL1"    "TIMM44"    "FBXL2"     "TLK1"      "USP25"
##  [16] "SSX2IP"    "OSBPL8"    "ASPSCR1"   "STK11IP"   "TTC8"
##  [21] "TPM3"      "MITF"      "PHKB"      "CACNA2D1"  "CDK7"
##  [26] "STAT3"     "ADAM23"    "RUNX2"     "TOB1"      "TACC2"
##  [31] "KIAA0391"  "PHF15"     "TRIM37"    "IRAK4"     "PARD3"
##  [36] "PC"        "TCF7L2"    "CPEB2"     "HADHA"     "ALPL"
##  [41] "LDLR"      "ATR"       "AUH"       "CLTCL1"    "GSPT1"
##  [46] "IL7R"      "ITPR1"     "MYC"       "PHKA1"     "PIK3C2B"
##  [51] "MAP2K5"    "PRPS2"     "ROBO1"     "TGFBR3"    "ZNF202"
##  [56] "EED"       "TNFRSF10B" "HERC1"     "BTRC"      "BCL9"
##  [61] "EZH2"      "LATS1"     "RASAL2"    "ABL2"      "APBA2"
##  [66] "DMXL1"     "SDCBP"     "CTDSPL"    "PTPN13"    "TRIM16"
##  [71] "RASA2"     "CHL1"      "AKAP13"    "ZNF33A"    "TOPBP1"
##  [76] "KATNA1"    "UTRN"      "CIT"       "RNF13"     "CHORDC1"
##  [81] "SCMH1"     "RPS6KC1"   "CNOT4"     "NME7"      "SNX12"
##  [86] "NCOA6"     "TRPS1"     "KCNK2"     "ADAMTS6"   "STAU2"
##  [91] "SENP1"     "PHF14"     "MELK"      "HELZ"      "USP33"
##  [96] "KIAA0922"  "CLASP1"    "OPA1"      "VPS54"     "NOX4"
## [101] "CDKAL1"    "PLEKHB2"   "FANCL"     "CDK5RAP2"  "STRBP"
## [106] "ERBB2IP"   "TUFT1"     "MCCC1"     "PITPNM2"   "PTBP2"
## [111] "DHX35"     "C14orf93"  "TEAD1"     "ALG8"      "CYBRD1"
## [116] "TDRD3"     "SH3KBP1"   "CAMKK1"    "SYTL2"     "FOXP1"
## [121] "SOX6"
##
## $Nup98KD.IMR90$CRE.DamdCT.N98RIP
##  [1] "HLCS"      "GOLGA4"    "PTPRS"     "RAD1"      "ZFX"
##  [6] "NFKB1"     "BIN1"      "LRP8"      "NOLC1"     "LDHA"
## [11] "PTPRU"     "NCOR2"     "FBXL2"     "USP25"     "MDN1"
## [16] "UBAP1"     "WBSCR22"   "BCOR"      "ASPSCR1"   "TTC8"
## [21] "TPM3"      "MITF"      "CASP9"     "SPG7"      "STAT3"
## [26] "THRA"      "WISP1"     "BRAF"      "TCEA1"     "PHF15"
## [31] "PPHLN1"    "PC"        "TCF7L2"    "RFX5"      "LDLR"
## [36] "IL6R"      "MYC"       "PHKA1"     "PIK3C2B"   "MAP2K5"
## [41] "ROBO1"     "TJP1"      "PLA2G6"    "TNFRSF10B" "CFLAR"
## [46] "BTRC"      "CCBL1"     "BCL9"      "LATS1"     "RASAL2"
## [51] "MARK2"     "AKAP12"    "ABL2"      "TRIM16"    "RASA2"
## [56] "NCOA3"     "TLK2"      "KATNA1"    "TEP1"      "UTRN"
```

```
## [61] "ZNF92"     "CHEK2"     "RNF13"     "CHORDC1"   "FBXO22"
## [66] "RPS6KC1"   "NME7"      "ADAMTS6"   "STAU2"     "PHF14"
## [71] "ARHGEF11"  "HELZ"      "USP33"     "KIAA0922"  "NOSIP"
## [76] "PLEKHB2"   "CDK5RAP2"  "TBC1D2"    "MRPL1"     "AGTRAP"
## [81] "MPHOSPH9"  "CAMKK1"    "SOX6"      "GATS"      "OXR1"
##
## $Nup98KD.IMR90$CRE.DamFL.D9RIP
##  [1] "MLH1"      "SAFB"      "SRP54"     "NOLC1"     "TIMM44"
##  [6] "STK11IP"   "TPM3"      "PHKB"      "CDK7"      "STAT3"
## [11] "TACC2"     "TCF7L2"    "HADHA"     "TYRP1"     "IL7R"
## [16] "ITPR1"     "NRIP1"     "TNFRSF10B" "GLP2R"     "ABL2"
## [21] "EYA2"      "APBA2"     "PTPN13"    "TRIM16"    "PDE10A"
## [26] "AKAP13"    "RNF13"     "CNOT4"     "STAU2"     "HELZ"
## [31] "CDKAL1"    "PLEKHB2"   "STRBP"     "TUFT1"     "MCCC1"
## [36] "C14orf93"  "TDRD3"     "SEPT10"
##
## $Nup98KD.IMR90$CRE.DamdCT.D9RIP
##  [1] "NFKB1"     "BIN1"      "NOLC1"     "PTPRU"     "BCOR"
##  [6] "TPM3"      "STAT3"     "THRA"      "PPHLN1"    "TCF7L2"
## [11] "RFX5"      "IL6R"      "TJP1"      "TNFRSF10B" "WTAP"
## [16] "ABL2"      "TRIM16"    "NCOA3"     "TLK2"      "RNF13"
## [21] "STAU2"     "HELZ"      "PLEKHB2"
##
## $Nup98KD.IMR90$CRE.DamFL.bothRIP
##  [1] "MLH1"      "SAFB"      "SRP54"     "NOLC1"     "TIMM44"
##  [6] "STK11IP"   "TPM3"      "PHKB"      "CDK7"      "STAT3"
## [11] "TACC2"     "TCF7L2"    "HADHA"     "IL7R"      "ITPR1"
## [16] "TNFRSF10B" "ABL2"      "APBA2"     "PTPN13"    "TRIM16"
## [21] "AKAP13"    "RNF13"     "CNOT4"     "STAU2"     "HELZ"
## [26] "CDKAL1"    "PLEKHB2"   "STRBP"     "TUFT1"     "MCCC1"
## [31] "C14orf93"  "TDRD3"
##
## $Nup98KD.IMR90$CRE.DamdCT.bothRIP
##  [1] "NFKB1"     "BIN1"      "NOLC1"     "PTPRU"     "BCOR"
##  [6] "TPM3"      "STAT3"     "THRA"      "PPHLN1"    "TCF7L2"
## [11] "RFX5"      "IL6R"      "TJP1"      "TNFRSF10B" "ABL2"
## [16] "TRIM16"    "NCOA3"     "TLK2"      "RNF13"     "STAU2"
## [21] "HELZ"      "PLEKHB2"
##
##
## $DHX9KD.NB1
## $DHX9KD.NB1$CRE.DamFL.N98RIP
##  [1] "RRM1"     "ADAM10"   "CDK6"     "NEO1"     "TIMM44"   "USP3"
##  [7] "MAN1A2"   "ASXL1"    "ZFP64"    "OSBPL9"   "ZNF341"   "NPC1"
## [13] "ATP1B3"   "STAT3"    "TOB1"     "NAB1"     "MRPS27"   "SPAG9"
## [19] "CPEB2"    "ACADM"    "ADSL"     "AGL"      "PKD2"     "UBE3A"
## [25] "P4HA1"    "ADM"      "CBFB"     "CNN3"     "ITPR1"    "MPV17"
## [31] "MAP2K5"   "MAP2K6"   "RAP1A"    "SPTAN1"   "NSMAF"    "CDC42BPA"
## [37] "PRPF4B"   "PRC1"     "CDKL1"    "ILK"      "DGKI"     "LRRFIP1"
## [43] "APBA2"    "BCAT1"    "KIFC3"    "LIMK2"    "NFIB"     "XRCC1"
## [49] "BAIAP2"   "SDCCAG8"  "PKIG"     "PITPNC1"  "STK39"    "SLCO3A1"
## [55] "STAU2"    "PDCD4"    "DAAM1"    "PUM2"     "POLR1A"   "NUP54"
## [61] "SLC30A6"  "ASPM"     "STRBP"    "CHST11"   "SLC12A9"  "NAV1"
## [67] "MTMR3"    "EPB41L4A" "ALG8"     "CAMKK1"   "NCOA7"
```

```
##
## $DHX9KD.NB1$CRE.DamdCT.N98RIP
##  [1] "CANX"     "CREM"     "TPD52L2"  "PPFIA1"   "TFCP2"    "AP3S2"
##  [7] "ZFR"      "NASP"     "PAM"      "STAT3"    "MRPS27"   "ADSL"
## [13] "ACTB"     "ADM"      "CASP8"    "CBFB"     "CNN3"     "CSNK1E"
## [19] "IVD"      "MPV17"    "PPP2R5C"  "MAP2K5"   "PRPF4B"   "PRC1"
## [25] "CSRP1"    "CDKL1"    "TOP3A"    "ABCC1"    "KIFC3"    "NFIB"
## [31] "NR1H3"    "DPP3"     "PLTP"     "RAD21"    "BAIAP2"   "PKIG"
## [37] "PITPNC1"  "SLC25A13" "STAU2"    "TAF5L"    "PDCD4"    "ERO1L"
## [43] "DAAM1"    "NUSAP1"   "ASPM"     "ABCB9"    "SLC12A9"  "MTMR3"
## [49] "CAMKK1"   "ITPA"     "GATS"
##
## $DHX9KD.NB1$CRE.DamFL.D9RIP
##  [1] "HSPA4"    "TIMM44"   "ASXL1"    "OSBPL9"   "ATP1B3"   "STAT3"
##  [7] "ACADM"    "ADM"      "ITPR1"    "PTPRG"    "RAP1A"    "NSMAF"
## [13] "CDC42BPA" "PRC1"     "DGKI"     "APBA2"    "BCAT1"    "LIMK2"
## [19] "NFIB"     "BAIAP2"   "PKIG"     "SLCO3A1"  "STAU2"    "POLR1A"
## [25] "SLC35A5"  "ASPM"     "STRBP"    "CHST11"   "SLC12A9"  "MTMR3"
## [31] "EPB41L4A" "PHF6"     "RAB3C"    "THAP5"
##
## $DHX9KD.NB1$CRE.DamdCT.D9RIP
##  [1] "CANX"    "PPFIA1"  "RIOK1"   "NASP"    "STAT3"   "ADM"     "PPP2R5C"
##  [8] "PTPRG"   "PRC1"    "CSRP1"   "TOP3A"   "NFIB"    "DPP3"    "PLTP"
## [15] "BAIAP2"  "PKIG"    "STAU2"   "NUSAP1"  "ASPM"    "SLC12A9" "MTMR3"
##
## $DHX9KD.NB1$CRE.DamFL.bothRIP
##  [1] "TIMM44"   "ASXL1"    "OSBPL9"   "ATP1B3"   "STAT3"    "ACADM"
##  [7] "ADM"      "ITPR1"    "RAP1A"    "NSMAF"    "CDC42BPA" "PRC1"
## [13] "DGKI"     "APBA2"    "BCAT1"    "LIMK2"    "NFIB"     "BAIAP2"
## [19] "PKIG"     "SLCO3A1"  "STAU2"    "POLR1A"   "ASPM"     "STRBP"
## [25] "CHST11"   "SLC12A9"  "MTMR3"    "EPB41L4A"
##
## $DHX9KD.NB1$CRE.DamdCT.bothRIP
##  [1] "CANX"    "PPFIA1"  "NASP"    "STAT3"   "ADM"     "PPP2R5C" "PRC1"
##  [8] "CSRP1"   "TOP3A"   "NFIB"    "DPP3"    "PLTP"    "BAIAP2"  "PKIG"
## [15] "STAU2"   "NUSAP1"  "ASPM"    "SLC12A9" "MTMR3"
```

Subsetting genes with altered expression upon DHX9 KD and Nup98 KD that change in the same direction

Given a change in expression upon DHX9 KD, consider true if it changes in the same direction in either Nup98 KD dataset.

Percent of DHX9 KD upregulated genes also upregulated in either Nup98 KD:

```
length(Transcript$changesameboth$upreg)/length(Transcript$up$DHX9KD.NB1)*100
```

```
## [1] 16.32463
```

Percent of DHX9 KD downregulated genes also downregulated in either Nup98 KD:

```
length(Transcript$changesameboth$downreg)/length(Transcript$down$DHX9KD.NB1)*100
```

```
## [1] 24.85429
```

Percent of Nup98 KD upregulated genes in both cell lines:

```
length(Transcript$changesameboth$bothupN98)/length(Transcript$up$Nup98KD.HepG2)*100
```

```
## [1] 31.1943
```

Percent of Nup98 KD downregulated genes in both cell lines:

```
length(Transcript$changesameboth$bothdownN98)/length(Transcript$down$Nup98KD.IMR90)*100
```

```
## [1] 41.85464
```

Compare genes with similarly altered expression upon DHX9 or Nup98 KD to genes bound by Dam-Nup98

Table of p-values

```
gom.obj <- newGOM(Transcript$changesameboth, Dam.all, gs.RNASeq)
getMatrix(gom.obj, name="pval")
```

```
##                  Nup98FL    Nup98dCTD
## upreg       1.246803e-06 1.001008e-06
## downreg     1.251822e-03 4.783581e-01
## alldir      5.781723e-09 1.013736e-05
## bothupN98   7.696832e-06 7.049577e-02
## bothdownN98 3.839536e-01 2.942224e-02
## bothallN98  7.055005e-06 9.100411e-03
```

Number of genes in common:

```
getMatrix(gom.obj, name="intersection")
```

```
##             Nup98FL Nup98dCTD
## upreg            18        14
## downreg           9         2
## alldir           27        16
## bothupN98        12         4
## bothdownN98       1         2
## bothallN98       13         6
```

```
getNestedList(gom.obj, name="intersection")
```

```
## $Nup98FL
## $Nup98FL$upreg
##  [1] "GRK5"    "RHOBTB1" "YAP1"    "OAF"     "SRGAP1"  "STON2"   "ETV4"
##  [8] "MAP4K4"  "SERTAD2" "FHL2"    "TMEM158" "UTP3"    "PALLD"   "CDK6"
## [15] "SLC7A2"  "PLAT"    "NCS1"    "PCYT1B"
##
## $Nup98FL$downreg
## [1] "PPP1R12B" "SSX2IP"   "HSD17B12" "GRIA4"    "TXNDC16"  "CLIP4"
## [7] "FAM105A"  "ARL4A"    "TACC1"
##
## $Nup98FL$alldir
##  [1] "GRK5"     "RHOBTB1"  "YAP1"     "OAF"      "SRGAP1"   "STON2"
##  [7] "ETV4"     "MAP4K4"   "SERTAD2"  "FHL2"     "TMEM158"  "UTP3"
## [13] "PALLD"    "CDK6"     "SLC7A2"   "PLAT"     "NCS1"     "PCYT1B"
## [19] "PPP1R12B" "SSX2IP"   "HSD17B12" "GRIA4"    "TXNDC16"  "CLIP4"
## [25] "FAM105A"  "ARL4A"    "TACC1"
```

```
##
## $Nup98FL$bothupN98
##  [1] "C1orf9"    "CTSS"      "MDM2"      "PHLDA1"    "TTC7B"
##  [6] "CYGB"      "QPCT"      "SLC4A7"    "LHFPL2"    "FYN"
## [11] "TNFRSF10B" "MOSPD1"
##
## $Nup98FL$bothdownN98
## [1] "LDLR"
##
## $Nup98FL$bothallN98
##  [1] "C1orf9"    "CTSS"      "MDM2"      "PHLDA1"    "TTC7B"
##  [6] "CYGB"      "QPCT"      "SLC4A7"    "LHFPL2"    "FYN"
## [11] "TNFRSF10B" "MOSPD1"    "LDLR"
##
##
## $Nup98dCTD
## $Nup98dCTD$upreg
##  [1] "NFKB2"   "ITPRIP"  "YAP1"    "SRGAP1"  "TESC"    "SLC6A2"  "MT2A"
##  [8] "EMP2"    "ETV4"    "DOCK6"   "SERTAD2" "FHL2"    "PALLD"   "DENND2A"
##
## $Nup98dCTD$downreg
## [1] "TRNP1" "CLIP4"
##
## $Nup98dCTD$alldir
##  [1] "NFKB2"   "ITPRIP"  "YAP1"    "SRGAP1"  "TESC"    "SLC6A2"  "MT2A"
##  [8] "EMP2"    "ETV4"    "DOCK6"   "SERTAD2" "FHL2"    "PALLD"   "DENND2A"
## [15] "TRNP1"   "CLIP4"
##
## $Nup98dCTD$bothupN98
## [1] "ITPRIP"    "MDM2"      "FYN"       "TNFRSF10B"
##
## $Nup98dCTD$bothdownN98
## [1] "CPT1A" "LDLR"
##
## $Nup98dCTD$bothallN98
## [1] "ITPRIP"    "MDM2"      "FYN"       "TNFRSF10B" "CPT1A"     "LDLR"
```

Compare genes with similarly altered expression upon DHX9 or Nup98 KD to genes bound by Nup98 or DHX9 RNA-IPs

Table of p-values

```
gom.obj <- newGOM(Transcript$changesameboth, RIP.all, gs.RNASeq)
getMatrix(gom.obj, name="pval")
```

```
##             Nup98RIP.K562 DHX9RIP.T32        both
## upreg        7.888894e-14   0.2728807 0.012856748
## downreg      1.154609e-04   0.6057791 0.054196314
## alldir       5.666073e-17   0.3187557 0.002030835
## bothupN98    4.597944e-08   0.6898875 0.047524405
## bothdownN98  8.752413e-02   0.7697584 0.262902872
## bothallN98   5.144745e-09   0.6980879 0.024150298
```

Number of genes in common:

```
getMatrix(gom.obj, name="intersection")
```

```
##             Nup98RIP.K562 DHX9RIP.T32 both
## upreg                  61          29   18
## downreg                28          14   10
## alldir                 89          43   28
## bothupN98              33          13   10
## bothdownN98             5           2    2
## bothallN98             38          15   12
```

```
getNestedList(gom.obj, name="intersection")
```

```
## $Nup98RIP.K562
## $Nup98RIP.K562$upreg
##  [1] "EPHB2"    "OLFML3"   "BTG2"     "NFKB2"    "ITPRIP"   "GRK5"
##  [7] "ARHGAP12" "RHOBTB1"  "PDLIM1"   "HSPA12A"  "CD82"     "MDK"
## [13] "TMEM132A" "PCNXL3"   "YAP1"     "OAF"      "SRGAP1"   "TPCN1"
## [19] "TESC"     "DHX37"    "PROSER1"  "STON2"    "TLE3"     "TMEM159"
## [25] "MMP2"     "MT2A"     "EMP2"     "ANKRD13B" "JUP"      "ETV4"
## [31] "SLC44A2"  "GATAD2A"  "WTIP"     "FXYD5"    "TRPM4"    "DOCK6"
## [37] "POU2F2"   "RPIA"     "MAP4K4"   "ITGA6"    "IGFBP2"   "SERTAD2"
## [43] "FHL2"     "NR4A2"    "COL18A1"  "CRKL"     "GPD1L"    "CTNNB1"
## [49] "PFKFB4"   "UGDH"     "FAM198B"  "MSX2"     "SOX4"     "CDK6"
## [55] "SLC7A2"   "LY6E"     "PLAT"     "DAB2IP"   "MAOA"     "PCYT1B"
## [61] "CXorf38"
##
## $Nup98RIP.K562$downreg
##  [1] "TRNP1"    "PPP1R12B" "SSX2IP"   "DPYD"     "MORN4"    "HSD17B12"
##  [7] "GRIA4"    "MIR210HG" "MANSC1"   "GPR137C"  "TMOD2"    "PIGN"
## [13] "CLIP4"    "ST3GAL5"  "SCN3A"    "PNPLA3"   "PCNP"     "CPZ"
## [19] "SLIT2"    "PARM1"    "PIK3R1"   "SERPINB6" "GSTA4"    "ARL4A"
## [25] "SEMA3C"   "TMEM176B" "TACC1"    "TMEM64"
##
## $Nup98RIP.K562$alldir
##  [1] "EPHB2"    "OLFML3"   "BTG2"     "NFKB2"    "ITPRIP"   "GRK5"
##  [7] "ARHGAP12" "RHOBTB1"  "PDLIM1"   "HSPA12A"  "CD82"     "MDK"
## [13] "TMEM132A" "PCNXL3"   "YAP1"     "OAF"      "SRGAP1"   "TPCN1"
## [19] "TESC"     "DHX37"    "PROSER1"  "STON2"    "TLE3"     "TMEM159"
## [25] "MMP2"     "MT2A"     "EMP2"     "ANKRD13B" "JUP"      "ETV4"
## [31] "SLC44A2"  "GATAD2A"  "WTIP"     "FXYD5"    "TRPM4"    "DOCK6"
## [37] "POU2F2"   "RPIA"     "MAP4K4"   "ITGA6"    "IGFBP2"   "SERTAD2"
## [43] "FHL2"     "NR4A2"    "COL18A1"  "CRKL"     "GPD1L"    "CTNNB1"
## [49] "PFKFB4"   "UGDH"     "FAM198B"  "MSX2"     "SOX4"     "CDK6"
## [55] "SLC7A2"   "LY6E"     "PLAT"     "DAB2IP"   "MAOA"     "PCYT1B"
## [61] "CXorf38"  "TRNP1"    "PPP1R12B" "SSX2IP"   "DPYD"     "MORN4"
## [67] "HSD17B12" "GRIA4"    "MIR210HG" "MANSC1"   "GPR137C"  "TMOD2"
## [73] "PIGN"     "CLIP4"    "ST3GAL5"  "SCN3A"    "PNPLA3"   "PCNP"
## [79] "CPZ"      "SLIT2"    "PARM1"    "PIK3R1"   "SERPINB6" "GSTA4"
## [85] "ARL4A"    "SEMA3C"   "TMEM176B" "TACC1"    "TMEM64"
##
## $Nup98RIP.K562$bothupN98
##  [1] "TIMM17A"   "JUN"       "CTSS"      "ITPRIP"    "AMPD3"
##  [6] "PCNXL3"    "CCND1"     "MDM2"      "TTC7B"     "RHBDF2"
```

```
## [11] "CYGB"      "GDF15"     "PPP1R15A"  "GGT5"      "FAM43A"
## [16] "SLC4A7"    "USP53"     "FAM198B"   "F2RL1"     "COX7C"
## [21] "CCNG1"     "LHFPL2"    "GNPDA1"    "SOX4"      "TNFAIP3"
## [26] "COX7A2"    "FYN"       "PEG10"     "LONRF1"    "DLC1"
## [31] "TNFRSF10B" "RPL12"     "MOSPD1"
##
## $Nup98RIP.K562$bothdownN98
## [1] "MTHFR"  "PBXIP1" "CPT1A"  "LDLR"   "CTSA"
##
## $Nup98RIP.K562$bothallN98
##  [1] "TIMM17A"   "JUN"       "CTSS"      "ITPRIP"    "AMPD3"
##  [6] "PCNXL3"    "CCND1"     "MDM2"      "TTC7B"     "RHBDF2"
## [11] "CYGB"      "GDF15"     "PPP1R15A"  "GGT5"      "FAM43A"
## [16] "SLC4A7"    "USP53"     "FAM198B"   "F2RL1"     "COX7C"
## [21] "CCNG1"     "LHFPL2"    "GNPDA1"    "SOX4"      "TNFAIP3"
## [26] "COX7A2"    "FYN"       "PEG10"     "LONRF1"    "DLC1"
## [31] "TNFRSF10B" "RPL12"     "MOSPD1"    "MTHFR"     "PBXIP1"
## [36] "CPT1A"     "LDLR"      "CTSA"
##
##
## $DHX9RIP.T32
## $DHX9RIP.T32$upreg
##  [1] "ITPRIP"   "CD82"     "TMEM132A" "TPCN1"    "DHX37"    "TMEM159"
##  [7] "MMP2"     "SLC6A2"   "EMP2"     "ANKRD13B" "JUP"      "GATAD2A"
## [13] "IGFBP2"   "COL18A1"  "TCN2"     "TIMP3"    "CTNNB1"   "TMEM158"
## [19] "PFKFB4"   "PDZRN3"   "PALLD"    "UGDH"     "DENND2A"  "LY6E"
## [25] "C8orf31"  "SFRP1"    "NCS1"     "PCYT1B"   "LAS1L"
##
## $DHX9RIP.T32$downreg
##  [1] "DPYD"     "MORN4"    "HSD17B12" "GRIA4"    "TXNDC16"  "TMOD2"
##  [7] "ST3GAL5"  "SCN3A"    "PROS1"    "CPZ"      "DANCR"    "PARM1"
## [13] "GUCY1B3"  "SERPINB6"
##
## $DHX9RIP.T32$alldir
##  [1] "ITPRIP"   "CD82"     "TMEM132A" "TPCN1"    "DHX37"    "TMEM159"
##  [7] "MMP2"     "SLC6A2"   "EMP2"     "ANKRD13B" "JUP"      "GATAD2A"
## [13] "IGFBP2"   "COL18A1"  "TCN2"     "TIMP3"    "CTNNB1"   "TMEM158"
## [19] "PFKFB4"   "PDZRN3"   "PALLD"    "UGDH"     "DENND2A"  "LY6E"
## [25] "C8orf31"  "SFRP1"    "NCS1"     "PCYT1B"   "LAS1L"    "DPYD"
## [31] "MORN4"    "HSD17B12" "GRIA4"    "TXNDC16"  "TMOD2"    "ST3GAL5"
## [37] "SCN3A"    "PROS1"    "CPZ"      "DANCR"    "PARM1"    "GUCY1B3"
## [43] "SERPINB6"
##
## $DHX9RIP.T32$bothupN98
##  [1] "ITPRIP"    "RHBDF2"    "CYGB"      "PPP1R15A"  "ODC1"
##  [6] "TIMP3"     "FAM43A"    "EREG"      "COX7A2"    "LONRF1"
## [11] "DLC1"      "TNFRSF10B" "MOSPD1"
##
## $DHX9RIP.T32$bothdownN98
## [1] "MTHFR" "CTSA"
##
## $DHX9RIP.T32$bothallN98
##  [1] "ITPRIP"    "RHBDF2"    "CYGB"      "PPP1R15A"  "ODC1"
##  [6] "TIMP3"     "FAM43A"    "EREG"      "COX7A2"    "LONRF1"
```

```
## [11] "DLC1"      "TNFRSF10B" "MOSPD1"    "MTHFR"     "CTSA"
##
##
## $both
## $both$upreg
##  [1] "ITPRIP"   "CD82"     "TMEM132A" "TPCN1"    "DHX37"    "TMEM159"
##  [7] "MMP2"     "EMP2"     "ANKRD13B" "JUP"      "GATAD2A"  "IGFBP2"
## [13] "COL18A1"  "CTNNB1"   "PFKFB4"   "UGDH"     "LY6E"     "PCYT1B"
##
## $both$downreg
##  [1] "DPYD"     "MORN4"    "HSD17B12" "GRIA4"    "TMOD2"    "ST3GAL5"
##  [7] "SCN3A"    "CPZ"      "PARM1"    "SERPINB6"
##
## $both$alldir
##  [1] "ITPRIP"   "CD82"     "TMEM132A" "TPCN1"    "DHX37"    "TMEM159"
##  [7] "MMP2"     "EMP2"     "ANKRD13B" "JUP"      "GATAD2A"  "IGFBP2"
## [13] "COL18A1"  "CTNNB1"   "PFKFB4"   "UGDH"     "LY6E"     "PCYT1B"
## [19] "DPYD"     "MORN4"    "HSD17B12" "GRIA4"    "TMOD2"    "ST3GAL5"
## [25] "SCN3A"    "CPZ"      "PARM1"    "SERPINB6"
##
## $both$bothupN98
##  [1] "ITPRIP"    "RHBDF2"    "CYGB"      "PPP1R15A"  "FAM43A"
##  [6] "COX7A2"    "LONRF1"    "DLC1"      "TNFRSF10B" "MOSPD1"
##
## $both$bothdownN98
## [1] "MTHFR" "CTSA"
##
## $both$bothallN98
##  [1] "ITPRIP"    "RHBDF2"    "CYGB"      "PPP1R15A"  "FAM43A"
##  [6] "COX7A2"    "LONRF1"    "DLC1"      "TNFRSF10B" "MOSPD1"
## [11] "MTHFR"     "CTSA"
```

Compare genes with similarly altered expression upon DHX9 or Nup98 KD to genes with CRE element

Table of p-values

```
gom.obj <- newGOM(Transcript$changesameboth, CRE.genes, gs.RNASeq)
getMatrix(gom.obj, name="pval")
```

```
##                     CRE
## upreg       9.088663e-05
## downreg     4.971751e-03
## alldir      1.262064e-06
## bothupN98   3.168789e-06
## bothdownN98 2.436227e-02
## bothallN98  9.455544e-08
```

Number of genes in common:

```
getMatrix(gom.obj, name="intersection")
```

```
##            CRE
## upreg       46
```

```
## downreg      25
## alldir       71
## bothupN98    31
## bothdownN98   6
## bothallN98   37
```

```
getNestedList(gom.obj, name="intersection")
```

```
## $CRE
## $CRE$upreg
##  [1] "EPHB2"    "CA14"     "BTG2"     "NFKB2"    "ARHGAP12" "RHOBTB1"
##  [7] "PDLIM1"   "MDK"      "YAP1"     "TPCN1"    "DHX37"    "TLE3"
## [13] "CSPG4"    "MMP2"     "SLC6A2"   "MT2A"     "EMP2"     "JUP"
## [19] "ETV4"     "FXYD5"    "TRPM4"    "POU2F2"   "RPIA"     "MAP4K4"
## [25] "ITGA6"    "IGFBP2"   "FHL2"     "NR4A2"    "JAG1"     "COL18A1"
## [31] "CRKL"     "TCN2"     "TIMP3"    "CTNNB1"   "PFKFB4"   "UGDH"
## [37] "MSX2"     "SOX4"     "CDK6"     "SLC7A2"   "LY6E"     "SFRP1"
## [43] "PLAT"     "DAB2IP"   "MAOA"     "PCYT1B"
##
## $CRE$downreg
##  [1] "UQCRH"     "PPP1R12B"  "SSX2IP"    "DPYD"      "HSD17B12"
##  [6] "GRIA4"     "C14orf129" "TMOD2"     "PIGN"      "SCN3A"
## [11] "PCBP3"     "PCNP"      "ATP2B2"    "PROS1"     "CPZ"
## [16] "SLIT2"     "GUCY1B3"   "PIK3R1"    "HIST1H3E"  "SERPINB6"
## [21] "GSTA4"     "SEMA3C"    "TACC1"     "AGPAT2"    "RPS6KA6"
##
## $CRE$alldir
##  [1] "EPHB2"     "CA14"      "BTG2"      "NFKB2"     "ARHGAP12"
##  [6] "RHOBTB1"   "PDLIM1"    "MDK"       "YAP1"      "TPCN1"
## [11] "DHX37"     "TLE3"      "CSPG4"     "MMP2"      "SLC6A2"
## [16] "MT2A"      "EMP2"      "JUP"       "ETV4"      "FXYD5"
## [21] "TRPM4"     "POU2F2"    "RPIA"      "MAP4K4"    "ITGA6"
## [26] "IGFBP2"    "FHL2"      "NR4A2"     "JAG1"      "COL18A1"
## [31] "CRKL"      "TCN2"      "TIMP3"     "CTNNB1"    "PFKFB4"
## [36] "UGDH"      "MSX2"      "SOX4"      "CDK6"      "SLC7A2"
## [41] "LY6E"      "SFRP1"     "PLAT"      "DAB2IP"    "MAOA"
## [46] "PCYT1B"    "UQCRH"     "PPP1R12B"  "SSX2IP"    "DPYD"
## [51] "HSD17B12"  "GRIA4"     "C14orf129" "TMOD2"     "PIGN"
## [56] "SCN3A"     "PCBP3"     "PCNP"      "ATP2B2"    "PROS1"
## [61] "CPZ"       "SLIT2"     "GUCY1B3"   "PIK3R1"    "HIST1H3E"
## [66] "SERPINB6"  "GSTA4"     "SEMA3C"    "TACC1"     "AGPAT2"
## [71] "RPS6KA6"
##
## $CRE$bothupN98
##  [1] "C1orf9"    "TIMM17A"   "JUN"       "CTSS"      "TFAM"
##  [6] "AMPD3"     "CCND1"     "MDM2"      "PHLDA1"    "CYGB"
## [11] "PPP1R15A"  "IL11"      "QPCT"      "ODC1"      "TIMP3"
## [16] "SLC4A7"    "EREG"      "F2RL1"     "COX7C"     "CCNG1"
## [21] "LHFPL2"    "GNPDA1"    "SOX4"      "ITPR3"     "TNFAIP3"
## [26] "COX7A2"    "FYN"       "PEG10"     "DLC1"      "TNFRSF10B"
## [31] "RPL12"
##
## $CRE$bothdownN98
## [1] "MTHFR"     "PBXIP1"    "CPT1A"     "C14orf129" "LDLR"      "HIST1H1C"
##
```

```
## $CRE$bothallN98
##  [1] "C1orf9"    "TIMM17A"   "JUN"       "CTSS"      "TFAM"
##  [6] "AMPD3"     "CCND1"     "MDM2"      "PHLDA1"    "CYGB"
## [11] "PPP1R15A"  "IL11"      "QPCT"      "ODC1"      "TIMP3"
## [16] "SLC4A7"    "EREG"      "F2RL1"     "COX7C"     "CCNG1"
## [21] "LHFPL2"    "GNPDA1"    "SOX4"      "ITPR3"     "TNFAIP3"
## [26] "COX7A2"    "FYN"       "PEG10"     "DLC1"      "TNFRSF10B"
## [31] "RPL12"     "MTHFR"     "PBXIP1"    "CPT1A"     "C14orf129"
## [36] "LDLR"      "HIST1H1C"
```

Compare genes with similarly altered expression upon DHX9 or Nup98 KD to genes bound by Dam-Nup98 and containing CRE elements

Table of p-values

```
gom.obj <- newGOM(Transcript$changesameboth, N98Dam.CRE, gs.RNASeq)
getMatrix(gom.obj, name="pval")
```

```
##             DamNup98FLCRE DamNup98dCTDCRE
## upreg        7.125956e-04    0.0008336621
## downreg      1.072400e-02    1.0000000000
## alldir       2.416446e-05    0.0094183398
## bothupN98    5.491088e-07    0.0491168723
## bothdownN98  2.110880e-01    0.0087001448
## bothallN98   2.548260e-07    0.0025877944
```

Number of genes in common:

```
getMatrix(gom.obj, name="intersection")
```

```
##             DamNup98FLCRE DamNup98dCTDCRE
## upreg                   9               7
## downreg                 5               0
## alldir                 14               7
## bothupN98              10               3
## bothdownN98             1               2
## bothallN98             11               5
```

```
getNestedList(gom.obj, name="intersection")
```

```
## $DamNup98FLCRE
## $DamNup98FLCRE$upreg
## [1] "RHOBTB1" "YAP1"    "ETV4"    "MAP4K4"  "FHL2"    "CDK6"    "SLC7A2"
## [8] "PLAT"    "PCYT1B"
##
## $DamNup98FLCRE$downreg
## [1] "PPP1R12B" "SSX2IP"   "HSD17B12" "GRIA4"    "TACC1"
##
## $DamNup98FLCRE$alldir
##  [1] "RHOBTB1"  "YAP1"     "ETV4"     "MAP4K4"   "FHL2"     "CDK6"
##  [7] "SLC7A2"   "PLAT"     "PCYT1B"   "PPP1R12B" "SSX2IP"   "HSD17B12"
## [13] "GRIA4"    "TACC1"
##
```

```
## $DamNup98FLCRE$bothupN98
##  [1] "C1orf9"    "CTSS"      "MDM2"      "PHLDA1"    "CYGB"
##  [6] "QPCT"      "SLC4A7"    "LHFPL2"    "FYN"       "TNFRSF10B"
##
## $DamNup98FLCRE$bothdownN98
## [1] "LDLR"
##
## $DamNup98FLCRE$bothallN98
##  [1] "C1orf9"    "CTSS"      "MDM2"      "PHLDA1"    "CYGB"
##  [6] "QPCT"      "SLC4A7"    "LHFPL2"    "FYN"       "TNFRSF10B"
## [11] "LDLR"
##
##
## $DamNup98dCTDCRE
## $DamNup98dCTDCRE$upreg
## [1] "NFKB2"  "YAP1"   "SLC6A2" "MT2A"   "EMP2"   "ETV4"   "FHL2"
##
## $DamNup98dCTDCRE$downreg
## character(0)
##
## $DamNup98dCTDCRE$alldir
## [1] "NFKB2"  "YAP1"   "SLC6A2" "MT2A"   "EMP2"   "ETV4"   "FHL2"
##
## $DamNup98dCTDCRE$bothupN98
## [1] "MDM2"      "FYN"       "TNFRSF10B"
##
## $DamNup98dCTDCRE$bothdownN98
## [1] "CPT1A" "LDLR"
##
## $DamNup98dCTDCRE$bothallN98
## [1] "MDM2"      "FYN"       "TNFRSF10B" "CPT1A"     "LDLR"
```

Compare genes with similarly altered expression upon DHX9 or Nup98 KD to genes bound by Dam-Nup98 with CRE elements and bound by Nup98 or DHX9 RNA-IPs

Table of p-values

```
gom.obj <- newGOM(Transcript$changesameboth, N98Dam.CRE.RIP, gs.RNASeq)
getMatrix(gom.obj, name="pval")
```

```
##             CRE.DamFL.N98RIP CRE.DamdCT.N98RIP CRE.DamFL.D9RIP
## upreg           6.681342e-05      0.0011633060      0.58796092
## downreg         2.814332e-03      1.0000000000      0.08579431
## alldir          6.198465e-07      0.0098535102      0.15564014
## bothupN98       4.202096e-05      0.0256230510      0.08262414
## bothdownN98     1.573457e-01      0.0052642439      1.00000000
## bothallN98      1.295256e-05      0.0008295656      0.10553581
##             CRE.DamdCT.D9RIP CRE.DamFL.bothRIP CRE.DamdCT.bothRIP
## upreg              0.1143891        0.48637122          0.3829222
## downreg            1.0000000        0.05231912          1.0000000
## alldir             0.2246514        0.08350064          0.5251925
```

```
## bothupN98          0.2702711        0.05030326          0.2302074
## bothdownN98        1.0000000        1.00000000          1.0000000
## bothallN98         0.3051932        0.06499076          0.2609241
```

Number of genes in common:

```
getMatrix(gom.obj, name="intersection")
```

```
##             CRE.DamFL.N98RIP CRE.DamdCT.N98RIP CRE.DamFL.D9RIP
## upreg                      9                 6               1
## downreg                    5                 0               2
## alldir                    14                 6               3
## bothupN98                  7                 3               2
## bothdownN98                1                 2               0
## bothallN98                 8                 5               2
##             CRE.DamdCT.D9RIP CRE.DamFL.bothRIP CRE.DamdCT.bothRIP
## upreg                      2                 1                  1
## downreg                    0                 2                  0
## alldir                     2                 3                  1
## bothupN98                  1                 2                  1
## bothdownN98                0                 0                  0
## bothallN98                 1                 2                  1
```

```
getNestedList(gom.obj, name="intersection")
```

```
## $CRE.DamFL.N98RIP
## $CRE.DamFL.N98RIP$upreg
## [1] "RHOBTB1" "YAP1"    "ETV4"    "MAP4K4"  "FHL2"    "CDK6"    "SLC7A2"
## [8] "PLAT"    "PCYT1B"
##
## $CRE.DamFL.N98RIP$downreg
## [1] "PPP1R12B" "SSX2IP"   "HSD17B12" "GRIA4"    "TACC1"
##
## $CRE.DamFL.N98RIP$alldir
##  [1] "RHOBTB1"  "YAP1"     "ETV4"     "MAP4K4"   "FHL2"     "CDK6"
##  [7] "SLC7A2"   "PLAT"     "PCYT1B"   "PPP1R12B" "SSX2IP"   "HSD17B12"
## [13] "GRIA4"    "TACC1"
##
## $CRE.DamFL.N98RIP$bothupN98
## [1] "CTSS"      "MDM2"      "CYGB"      "SLC4A7"    "LHFPL2"    "FYN"
## [7] "TNFRSF10B"
##
## $CRE.DamFL.N98RIP$bothdownN98
## [1] "LDLR"
##
## $CRE.DamFL.N98RIP$bothallN98
## [1] "CTSS"      "MDM2"      "CYGB"      "SLC4A7"    "LHFPL2"    "FYN"
## [7] "TNFRSF10B" "LDLR"
##
##
## $CRE.DamdCT.N98RIP
## $CRE.DamdCT.N98RIP$upreg
## [1] "NFKB2" "YAP1"  "MT2A"  "EMP2"  "ETV4"  "FHL2"
##
## $CRE.DamdCT.N98RIP$downreg
## character(0)
```

```
##
## $CRE.DamdCT.N98RIP$alldir
## [1] "NFKB2" "YAP1"  "MT2A"  "EMP2"  "ETV4"  "FHL2"
##
## $CRE.DamdCT.N98RIP$bothupN98
## [1] "MDM2"      "FYN"       "TNFRSF10B"
##
## $CRE.DamdCT.N98RIP$bothdownN98
## [1] "CPT1A" "LDLR"
##
## $CRE.DamdCT.N98RIP$bothallN98
## [1] "MDM2"      "FYN"       "TNFRSF10B" "CPT1A"     "LDLR"
##
##
## $CRE.DamFL.D9RIP
## $CRE.DamFL.D9RIP$upreg
## [1] "PCYT1B"
##
## $CRE.DamFL.D9RIP$downreg
## [1] "HSD17B12" "GRIA4"
##
## $CRE.DamFL.D9RIP$alldir
## [1] "PCYT1B"   "HSD17B12" "GRIA4"
##
## $CRE.DamFL.D9RIP$bothupN98
## [1] "CYGB"      "TNFRSF10B"
##
## $CRE.DamFL.D9RIP$bothdownN98
## character(0)
##
## $CRE.DamFL.D9RIP$bothallN98
## [1] "CYGB"      "TNFRSF10B"
##
##
## $CRE.DamdCT.D9RIP
## $CRE.DamdCT.D9RIP$upreg
## [1] "SLC6A2" "EMP2"
##
## $CRE.DamdCT.D9RIP$downreg
## character(0)
##
## $CRE.DamdCT.D9RIP$alldir
## [1] "SLC6A2" "EMP2"
##
## $CRE.DamdCT.D9RIP$bothupN98
## [1] "TNFRSF10B"
##
## $CRE.DamdCT.D9RIP$bothdownN98
## character(0)
##
## $CRE.DamdCT.D9RIP$bothallN98
## [1] "TNFRSF10B"
##
##
```

```
## $CRE.DamFL.bothRIP
## $CRE.DamFL.bothRIP$upreg
## [1] "PCYT1B"
##
## $CRE.DamFL.bothRIP$downreg
## [1] "HSD17B12" "GRIA4"
##
## $CRE.DamFL.bothRIP$alldir
## [1] "PCYT1B"   "HSD17B12" "GRIA4"
##
## $CRE.DamFL.bothRIP$bothupN98
## [1] "CYGB"      "TNFRSF10B"
##
## $CRE.DamFL.bothRIP$bothdownN98
## character(0)
##
## $CRE.DamFL.bothRIP$bothallN98
## [1] "CYGB"      "TNFRSF10B"
##
##
## $CRE.DamdCT.bothRIP
## $CRE.DamdCT.bothRIP$upreg
## [1] "EMP2"
##
## $CRE.DamdCT.bothRIP$downreg
## character(0)
##
## $CRE.DamdCT.bothRIP$alldir
## [1] "EMP2"
##
## $CRE.DamdCT.bothRIP$bothupN98
## [1] "TNFRSF10B"
##
## $CRE.DamdCT.bothRIP$bothdownN98
## character(0)
##
## $CRE.DamdCT.bothRIP$bothallN98
## [1] "TNFRSF10B"
```